Acta Geodaetica et Geophysica Hungarica

VOLUME 31, NUMBERS 1–2, 1996

Akadémiai Kiadó, Budapest

AGG 31 (1–2) 1–280 (1996) HU ISSN 1217–8977

ACTA GEODAETICA ET GEOPHYSICA HUNGARICA
A Quarterly Journal of the Hungarian Academy of Sciences

Acta Geodaetica et Geophysica Hungarica (AGG) publishes original reports and reviews on geodesy and geophysics in English.

AGG is published in yearly volumes of four issues by

AKADÉMIAI KIADÓ
H-1117 Budapest, Prielle K. u. 19–35

Manuscripts and editorial correspondence should be addressed to

AGG Editorial Office
Geodetical and Geophysical Research Institute
H-9401 Sopron, P.O. Box 5
Phone: (36-99) 314-290
Fax: (36-99) 313-267
E-mail: actagg@ggki.hu

Subscription information

Orders should be addressed to

AKADÉMIAI KIADÓ
H-1519 Budapest, P.O. Box 245

Subscription price for Volume 31 (1996) in 4 issues US$ 98.00, including normal postage, airmail delivery US$ 20.00.

INSTRUCTIONS TO AUTHORS

Manuscripts should be sent to the Editorial Office (address see above). Articles not published or submitted for publication elsewhere are only accepted.

Manuscripts should be typewritten in duplicate or on disk (LATEX$*$.tex files or Word for Windows Documents or text-only ASCII files) with a written copy, or via E-mail (LATEX$*$.tex files or text-only ASCII files).

CONTENTS

Book reviews

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 3–3 (1996)

Editorial

Electronic guide to AGG

Even the best papers remain of no use if someone cannot find them. Recognizing the extraordinary possibilities offered by INTERNET the Editorial Office of Acta Geodaetica et Geophysica Hungarica disseminates information about latest volumes.

AGG information may serve as a database available on World Wide Web at

http://www.ggki.hu/agg/Acta-Geod-Geoph-Hung.html

It includes author's addresses and abstracts.

We hope that the electronic dissemination will contribute to a wider and faster distribution of results.

J Verő
editor-in-chief

Acta Geod. Geoph. Hung., Vol. 31(1-2), pp. 5-9 (1996)

ROBUST ESTIMATION USING THE CAUCHY PROBABILITY DENSITY

L BATTHA[1] and J SOMOGYI[1]

[Manuscript received February 16, 1996]

The paper deals with the robust estimation using the Cauchy-function. It gives a solution for regression problem compared with the methods of re-weighted least squares and L1-norm.

Keywords: Cauchy-function; gradient method; long-tailed distributions; non linear solution; robust estimation

1. Introduction

The detection of outliers in data sets for which the Gaussian distribution is supposed, is a very important phase of the data-processing, because outliers could lead to unacceptable results in the least-square solution. The normal distributions contaminated with outliers are termed by Huber (1964) as contaminated normal distributions, and can be ranged in the group of "long-tailed" distributions. For this reason when outliers are suspected in a data set, long-tailed probability density functions should be used to take into account uncertainties (Tarantola 1987). There are two typical long-tailed densities, the symmetric exponential function $\exp(-|x|)$, and the Cauchy-function $1/(1+x^2)$. The former has the advantage that it can be related to the L1-norm, where the location parameter is estimated by the median, making use of linear programing. The L1-norm as an effective robust method is used for the geodesy, too (e.g. Fuchs 1982, Kampmann 1986, Kampmann and Wolf 1989, Somogyi and Závoti 1989, 1992, 1993).

The Cauchy-function can be represented with a bell-shaped density curve, it has infinite variance and seems theoretically to be adequate for robust estimation (detecting suspected outliers). The use of this function leads to the solution of a non-linear system of equations which is quite new in the geodetic practice. This paper gives a short presentation of our investigation on the Cauchy-function as a robust estimator compared with the methods of reweighted least squares and L1-norm.

2. The Cauchy probability density function

The distribution with probability function

$$f(x) = (\pi\sigma)^{-1}\left[1+\{(x-\Theta)/\sigma\}^2\right]^{-1} \qquad (\sigma > 0) \tag{1}$$

is called the Cauchy distribution (Johnson and Kotz 1970).

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

The parameters Θ and σ are the so called location and scale parameters. The distribution is centered at $x = \Theta$. The median is Θ. Its variance is infinite and, for large k,

$$f(\Theta_0 \pm k\sigma) \cong f(\Theta)/k^2 \tag{2}$$

tends to zero as $1/k^2$ (Tarantola 1987). This density represents a bell-shaped curve with long-tails (see Fig. 1).

The object function $\varrho(x)$ which expresses the mathematical connection between the observations x_i and the parameter Θ, as well as the influence function $\psi(x)$ for the Cauchy distribution are:

$$\varrho(x) = \ln(1 + x^2) + C, \ \psi(x) = 2x/(1 + x^2) \tag{3}$$

which gives a non linear equations' system :

$$\sum_{i=1}^{n} \psi(x_i - \Theta) = 0 \tag{4}$$

and $\hat{\Theta}$ can be obtained by iteration.

Figure 2 shows the $\psi(x)$ curve (influence curve) of the Cauchy distribution.

3. Numerical procedure to find the minimum of the object function $f(\mathbf{x})$ for the case of linear regression

The general problem we want to solve is the following: minimize $f(\mathbf{y}) = \sum_{i=1}^{m} \varrho(y_i)$, where

$$y_i = \left(\sum_{j=1}^{n} a_{ij} x_j \right) - b_i \qquad (1 \leq i \leq m) \tag{5}$$

and $\varrho(x)$ is a nonnegative differentiable function such that

$$\varrho(x) \to +\infty \qquad (|x| \to \infty) . \tag{6}$$

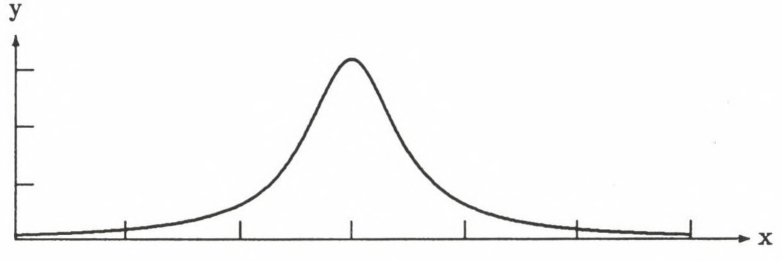


Fig. 1. Density curve of a Cauchy distribution

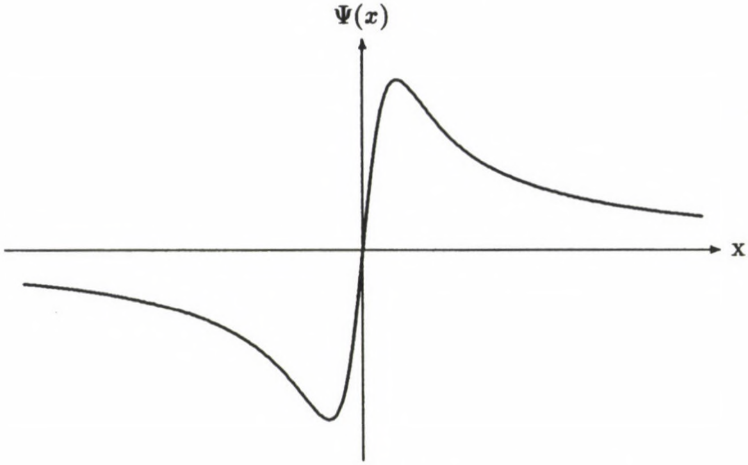


Fig. 2. The $\psi(x)$ curve of the Cauchy distribution

Our method applied to solve the minimalization problem is basically a combination of the *gradient* method (or method of *steepest descent*) and a variation of "halving the interval".

To apply the gradient method, we compute the partial derivates of f:

$$g_i = \frac{\partial f}{\partial x_j} = \sum_{i=1}^{m} \varrho'(y_i) \frac{\partial y_i}{\partial x_j} = \sum \varrho'(y_i) a_{ij} \, . \tag{7}$$

Given a current position $\mathbf{x}$, each iteration starts with computing $\mathbf{g} = \text{grad } f$. If $|\mathbf{g}| < \varepsilon_g$, where ε_g is a predetermined threshold limit, the iteration is stopped. Otherwise, we take a step in direction $-u = -\mathbf{g}/|\mathbf{g}|$ (steepest descent) to obtain $\mathbf{x}'$:

$$\mathbf{x}' = \mathbf{x} - \frac{f(\mathbf{x})}{|\mathbf{g}|} \mathbf{u} \, . \tag{8}$$

Now, we proceed with a "refinement" of the position of $\mathbf{x}'$ along the straight line $\mathbf{x}' + t(\mathbf{x} - \mathbf{x}')$. To do this, we halve the interval $0 \leq t \leq 1$, to obtain $\mathbf{x}'' = \mathbf{x}' + 1/2(\mathbf{x} - \mathbf{x}')$. If

$$|f(\mathbf{x}'') - f(\mathbf{x}')| + |f(\mathbf{x}'') - f(\mathbf{x})| < \varepsilon \, , \tag{9}$$

(where ε is the desired precision), the interval-halving is stopped. Otherwise, the interval $[\mathbf{x}\mathbf{x}']$ is replaced by $[\mathbf{x}''\mathbf{x}']$ if $f(\mathbf{x}') \leq f(\mathbf{x})$, and by $[\mathbf{x}\mathbf{x}'']$ if $f(\mathbf{x}') > f(\mathbf{x})$; and the above step is repeated. After some number of steps, condition (9) becomes true, yielding the minimum of f on the interval $[\mathbf{x}\mathbf{x}']$ at some point $y \in [\mathbf{x}\mathbf{x}']$. We now start a new iteration with x replaced by y.

In case of the Cauchy distribution, and when the observations can be expressed in linear form (as in the case of linear regression), $f(\mathbf{x})$ can be expressed in the following form:

$$f(\mathbf{x}) = \sum_{i=1}^{m} \varrho \left(b_i - \sum_{j=1}^{n} a_{ij} x_j \right) , \tag{10}$$

where $\varrho(x) = \ln(1 + x^2) + c$. From this it follows that grad $f = \mathbf{g}$, where

Table I

	Normal distribution			Long-tailed distribution		
N	Cf	Rwl	L1	Cf	Rwl	L1
1	−.046	−.046	−.025	0	−.028	−.010
2	−.034	−.034	−.016	.005	−.019	−.003
3	−.052	−.052	−.037	−.020	−.040	−.027
4	−.030	−.030	−.018	−.006	−.022	−.010
5	.012	.012	.021	.029	.017	.027
6	−.006	−.006	0	.032	−.005	.003
7	.096	.096	.099	−.902*	−.906*	−.900*
8	.098	.098	.098	.092	.092	.097
9	−.130	−.130	−.133	.857*	.861*	.863*
10	.042	.042	.036	.030	.036	.030
11	.055	.055	.045	.026	.038	.037
12	.057	.057	.045	.020	.037	.033
13	.029	.029	.014	−.015	−.005	0
14	.071	.071	.069	.019	.044	.037
15	−.047	−.047	−.068	−.106	−.077	−.087
16	.045	.045	−.021	−.021	−.011	0
17	.027	.027	0	−.047	−.010	−.023
18	−.110	−.110	−.141	−.192	−.151	−.167
19	−.058	−.058	−.092	1.852*	1.897*	1.880*
20	−.006	−.006	−.043	−.103	−.054	−.073
21	−.014	−.014	−.054	−.119	−.066	−.087
*SAD				4.484	4.410	4.394
SAD	1.065	1.066	1.059	.873	.746	.751
LSD	.077	.077	.087	.088	.054	.065
σ_y (±)	.062	.062	.062	.072	.055	.062
$\hat{a}$	.9792	.9792	1.0064	1.0414	1.0052	1.0266
$\hat{b}$	2.0004	2.0004	1.9998	1.9989	1.9997	1.9993

Given values $a = 1$, $b = 2$
*SAD = Sum of Absolute Deviations (including outliers marked with *)
SAD = Sum of Absolute Deviations (without values marked with *)
LSD = Least Square Sum of Deviations (without values marked with *)

Cf = Cauchy function
Rwl = Reweighted least-squares
L1 = L1 norm

$$g_k = -\sum_{i=1}^{m} \psi\left(b_i - \sum_{j=1}^{n} a_{ij} x_j\right) a_{ik} , \tag{11}$$

and where

$$\psi(x) = \frac{2x}{1+x^2} .$$

The initial "guess" of $\mathbf{x}$ is obtained by temporarily replacing ϱ by $\overline{\varrho}(x) = x^2$, and solving the resulting least-squares problem by a direct method (Gaussian elimination with complete pivoting).

4. Numerical example

Finally we give an example of robust solution based on the Cauchy-function described in the previous sections using a problem of linear regression.

To show the practicability of the presented method, it was generated to the y_i values of a simulated linear regression ($y_i = ax_i + b$) noise of normal distribution with a few gross errors (contaminated normal distribution). Choosing $\sigma = \pm 0.05$ units for the standard deviation of the assumed measurements, it is obtained 0.15 unit for the acceptable error.

The residuals of the regression model consisting of 21 points, statistical data, and the estimated values a and b are presented in Table I compared with the results of reweighted least-squares and L1-norm methods. The first three columns of the table present the results of the regression with normal distributed data. Outliers are marked with an asterisk.

As it appears from the table, there are no significant deviations between the different computations. The tests we have made show that the Cauchy-function can also be used for the robust estimation of linear problems. In case when the data set is used to solve a non linear problem, this method should be a very useful tool in detecting the suspected outliers.

Acknowledgement

This work was supported by OTKA (National Science Foundation) under Grant No. 014856

References

Fuchs H 1982: *Manuscripta Geodaetica*, 7, 151–207.
Huber P J 1964: *Annales Mathematical Statistics*, 35, 73–101.
Johnson N L, Kotz S 1970: Continuous univariate distributions-1. New York, John Wiley and Sons
Kampmann G 1986: *Allgemeine Vermessungsnachrichten*, 4, 139–147.
Kampmann G, Wolf H 1989: *Zeitschrift für Vermessungswesen*, 114, 493–497.
Somogyi J, Závoti J 1989: *Acta Geod. Geoph. Mont. Hung.*, 24, 237–249.
Somogyi J, Závoti J 1992: *Zeitschrift für Vermessungswesen*, 117, 336–344.
Somogyi J, Závoti J 1993: *ÖZ Vermessungswesen und Photogrammetrie*, 81, 16–24.
Tarantola A 1987: Inverse Problem Theory, Elsevier, Amsterdam

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 11–24 (1996)

A COMPARISON OF WEIGHT-FUNCTIONS IN ROBUST REGRESSION USING ITERATIVELY REWEIGHTED LEAST-SQUARES

J SOMOGYI[1] and J ZÁVOTI[1]

[Manuscript received May 20, 1996]

The paper compares different weight-functions in robust regression using iteratively reweighted least-squares. These w-functions developed from various ψ-functions are partly known and partly worked out by us. In a small Monte Carlo study different distributions were generated to the observations of the linear regression model to study the behaviour of these w-functions.

Keywords: location estimation; Monte Carlo; regression; reweighted least-squares; robustness; w-estimation; weight-function

1. Introduction

The robust statistical methods are developed with the aim of reducing or limiting the effects of the outliers and contaminations in cases when the measurement errors are assumed to follow a Gaussian distribution, or the underlying distribution differs from it. In geodesy the generally used method for data processing is the least-squares method, because it is supposed that the observation errors are normally distributed. In accordance with the geodetic traditions the most practicable robust method which could be used is the reweighted least-squares method based on weight functions. The so-called w-estimation with weight functions was proposed by Beaton and Tukey (1974) and Andrews (1974). In the geodetic literature the first such solution is the so called Danish-method (Krarup et al. 1980). Besides this several experiments have been made to solve geodetic and photogrammetric problems using iteratively reweighted LS, e.g. Veress and Youcai (1987), Faigh and Owolabi (1988), Somogyi and Kalmár (1991).

2. The concept of the W-estimation

Let the observations $y_1, y_2, \dots y_n$ be independent, and let $T_n = T_n(y_1, y_2, \dots y_n)$ be a sequence of estimates or test statistics. An M-estimate (Maximum Likelihood Estimate) defined by a minimum problem of the form

$$\sum_{i=1}^{n} \varrho\left(\frac{T_n - y_i}{\sigma}\right) \to \min\,, \tag{1}$$

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

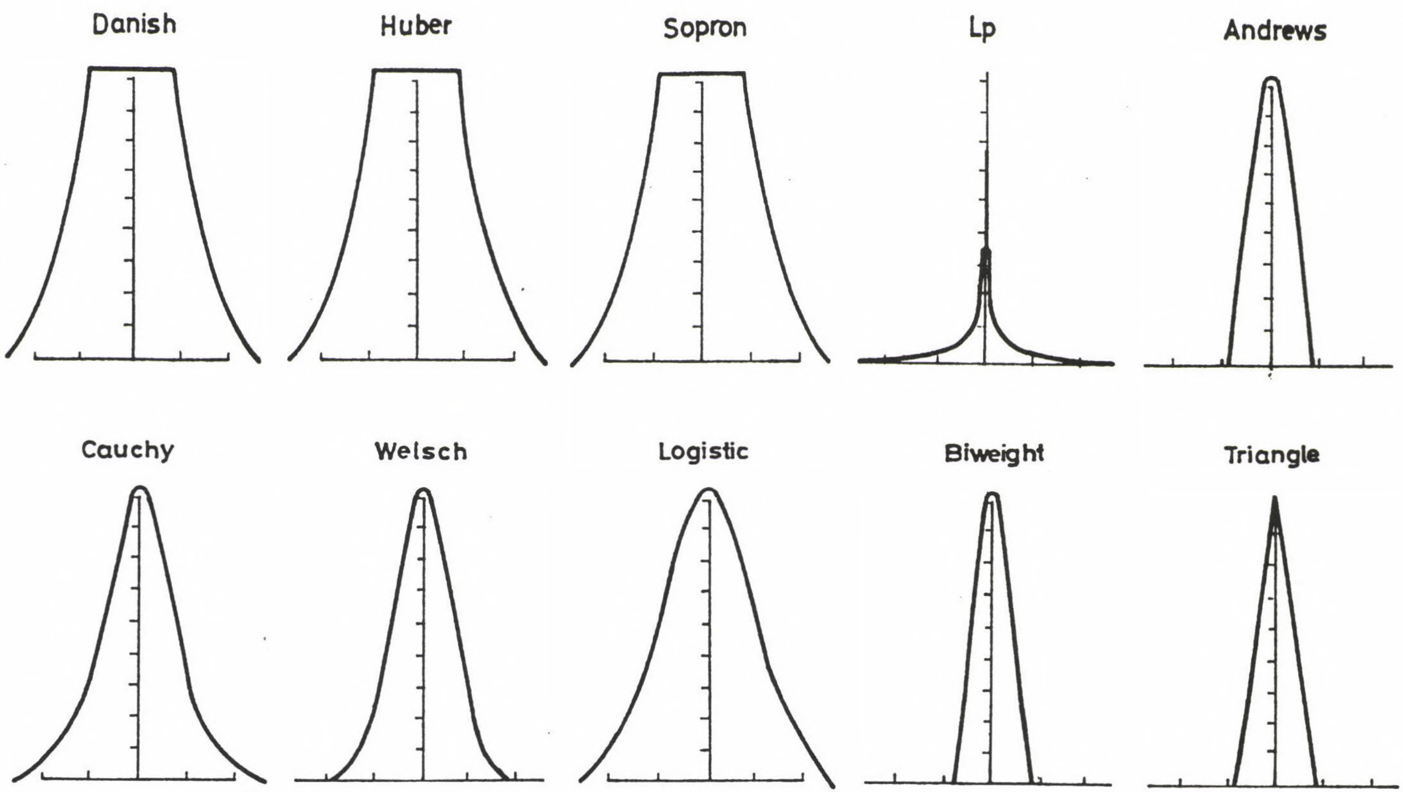


Fig. 1. Weight functions

Table I. Weight functions

Mixed normal $\psi(x)$	Unbroken $\psi(x)$	Cut off $\psi(x)$
Danish D $\quad w(x) = \begin{cases} 1 & \lvert x\rvert \le a \\ e^{-\frac{x}{a}+1} & \lvert x\rvert > a \end{cases}$	*Cauchy* Ca $\quad w(x) = \frac{1}{1+\frac{x^2}{a^2}}$	*Andrews* A $\quad w(x) = \begin{cases} \frac{\sin\frac{\pi x}{a}}{\pi x} & \lvert x\rvert \le a \\ 0 & \lvert x\rvert > a \end{cases}$
Huber H $\quad w(x) = \begin{cases} 1 & \lvert x\rvert \le a \\ \frac{a}{\lvert x\rvert} & \lvert x\rvert > a \end{cases}$	*Welsch* W $\quad w(x) = e^{-\frac{x^2}{a^2}}$	*Biweight* B $\quad w(x) = \begin{cases} \left(1-\frac{x^2}{a^2}\right)^2 & \lvert x\rvert \le a \\ 0 & \lvert x\rvert > a \end{cases}$
Sopron S $\quad w(x) = \begin{cases} 1 & \lvert x\rvert \le a \\ \frac{2}{1+\frac{x^2}{a^2}} & \lvert x\rvert > a \end{cases}$	*Logistic* Lo $\quad w(x) = a\frac{\tanh(\frac{x}{a})}{x}$	*Triangle* T $\quad w(x) = \begin{cases} 1-\frac{x}{a} & \lvert x\rvert \le a \\ 0 & \lvert x\rvert > a \end{cases}$
	Laplace Lp $\quad w(x) = \lvert x\rvert^{a-2}$	

Table II

N	Mixed normal $\psi(x)$						Unbroken $\psi(x)$							
	D		H		S		Ca		Co		Lo		Lp	
	r_i	w_i	r_i	w_i	r_i	w_i	r_i	w_i	r_i	w_i	r_i	w_i	r_i	w_i
1	–.043	1.000	–.025	1.000	–.038	1.000	–.036	.969	–.036	.967	–.022	.996	–.004	1.000
2	–.031	1.000	–.017	1.000	–.028	1.000	–.026	.984	–.026	.983	–.014	.998	.000	1.000
3	–.050	1.000	–.038	1.000	–.047	1.000	–.046	.950	–.046	.948	–.036	.989	–.026	.722
4	–.029	1.000	–.020	1.000	–.027	1.000	–.026	.984	–.026	.983	–.019	.997	–.012	1.000
5	.012	1.000	.018	1.000	.014	1.000	.014	.995	.014	.995	.019	.997	.023	.907
6	–.007	1.000	–.004	1.000	–.006	1.000	–.006	.999	–.006	.999	–.003	1.000	–.003	1.000
7	–.906	.221	–.905	.093	–.905	.029	–.906	.046	–.906	.000	–.907	.221	–.909	.001
8	.096	1.000	.093	1.000	.095	1.000	.095	.817	.094	.802	.092	.935	.086	.084
9	.867	.232	.861	.101	.866	.036	.865	.051	.864	.000	.860	.233	.850	.001
10	.038	1.000	.030	1.000	.036	1.000	.035	.971	.034	.971	.027	.994	.014	1.000
11	.049	1.000	.038	1.000	.047	1.000	.045	.952	.044	.952	.035	.990	.018	1.000
12	.050	1.000	.036	1.000	.047	1.000	.045	.952	.044	.952	.033	.931	.013	1.000
13	.022	1.000	.004	1.000	.018	1.000	.015	.994	.014	.995	.000	1.000	–.023	.886
14	.063	1.000	.043	1.000	.058	1.000	.055	.929	.054	.928	.038	.988	.011	1.000
15	–.056	1.000	–.079	1.000	–.061	1.000	–.065	.905	–.065	.899	–.084	.945	–.114	–.050
16	.035	1.000	.009	1.000	.030	1.000	.025	.984	.025	.985	.004	1.000	–.030	.545
17	.016	1.000	–.012	1.000	.010	1.000	.005	.999	.005	.999	–.019	.997	–.056	.180
18	–.112	1.000	.154	1.000	–.130	1.000	–.134	.689	–.135	.633	–.161	.828	–.202	.018
19	1.929	.000	1.894	.106	1.921	.021	1.916	.011	1.915	.000	1.887	.106	1.843	.000
20	–.020	1.000	-.058	1.000	–.028	1.000	–.034	.971	–.035	.970	–.066	.965	–.113	.051
21	–.029	1.000	–.069	1.000	–.038	1.000	–.044	.953	–.045	.951	–.078	.952	–.129	.040
$\hat{a} =$	1.0085		0.9907		0.9850		0.9938		0.9933		1.0428		1.0371	
$\hat{b} =$	1.9997		2.0001		2.0002		2.0000		2.0000		1.9995		1.9989	

$a = 1,\ b = 2$

Table II (contd.)

Cut off $\psi(x)$					
A		B		T	
r_i	w_i	r_i	w_i	r_i	w_i
−.034	.954	−.031	.954	−.032	.843
−.024	.976	−.022	.977	−.022	.889
−.044	.921	−.042	.912	−.043	.786
−0.25	.975	−.023	.973	−.024	.882
.015	.991	.016	.986	.016	.921
−.005	.999	−.005	.999	−.005	.976
−.906	.000	−.906	.000	−.906	.000
.094	.675	.093	.615	.094	.531
.864	.000	.862	.000	.863	.000
.033	.955	.031	.953	.032	.883
.043	.926	.040	.921	.042	.792
.042	.927	.039	.925	.041	.795
.012	.994	.008	.997	.010	.948
.052	.893	.047	.891	.050	.752
−.068	.818	−.074	.747	−.071	.645
.021	.982	.015	.988	.018	.907
0.001	1.000	−.006	.998	−.002	.988
−.140	.372	−.146	.215	−.143	.285
1.910	.000	1.903	.000	1.906	.000
−.040	.935	−.048	.887	−.044	.778
−.050	.898	−.059	.832	−.055	.724
0.9969		1.0013		0.9999	
1.9999		1.9998		1.9999	

where $\varrho(y,x)$ is an arbitrary loss function and σ is a robust estimation of scale.

This is equivalent to the solution for T_n of the equation:

$$\sum_{i=1}^{n} \psi\left(\frac{T_n - y_i}{\sigma}\right) = 0 , \tag{2}$$

where $\psi(x) = \frac{\partial \varrho(y;x)}{\partial x}$.

An alternative form of M-estimation is the so-called w-estimation. The w-estimators (Tukey 1970–1971) are defined as a weighted mean of the observations:

$$T_n(y_1, y_2, \dots y_n) = \frac{\sum\limits_{i=1}^{n} y_i w_i}{\sum\limits_{i=1}^{n} w_i} , \tag{3}$$

where the weights depend on the observations through $w_i = w\left(\frac{T_n - y_i}{\sigma}\right)$.

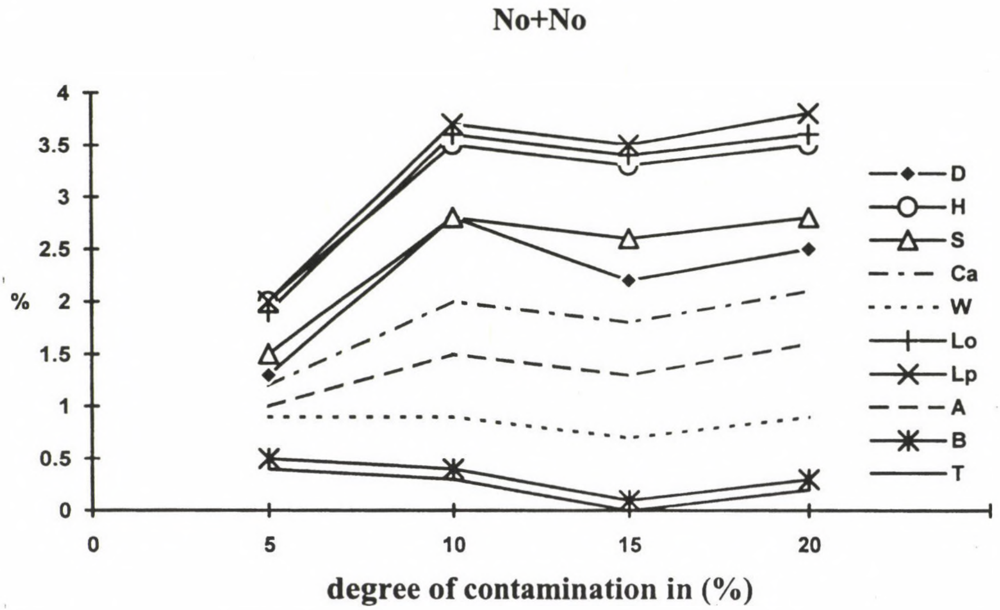


Fig. 2. Slope estimates for contaminating N (0.01 σ)

Table III. Distribution: No+No

Weight-function	Degree of contamination (%) 5	10	15	20
D	0.0463	0.0493	0.0501	0.0525
H	0.0468	0.0511	0.0546	0.0609
S	0.0464	0.0494	0.0513	0.0551
Ca	0.0463	0.0487	0.0499	0.0520
W	0.0463	0.0455	0.0463	0.0470
Lo	0.0468	0.0514	0.0554	0.0628
Lp	0.0471	0.0517	0.0561	0.0646
A	0.0463	0.0485	0.0495	0.0509
B	0.0464	0.0456	0.0463	0.0471
T	0.0464	0.0456	0.0464	0.0471

Starting from the median $T_n^{(o)}$ one computes

$$T_n^{(j+1)} = \frac{\sum_{i=1}^{n} y_i w\left(\frac{T_n^{(j)} - y_i}{\sigma}\right)}{\sum_{i=1}^{n} w\left(\frac{T_n^{(j)} - y_i}{\sigma}\right)} \tag{4}$$

Table IV. Distribution: No+Ca

Weight-function	Degree of contamination (%) 5	10	15	20
D	0.0481	0.0486	0.0484	0.0570
H	0.0487	0.0561	0.0589	0.0652
S	0.0482	0.0489	0.0490	0.0579
Ca	0.0480	0.0484	0.0482	0.0542
W	0.0479	0.0481	0.0477	0.0506
Lo	0.0486	0.0562	0.0595	0.0666
Lp	0.0497	0.0617	0.0723	0.0828
A	0.0481	0.0480	0.0476	0.0505
B	0.0480	0.0480	0.0476	0.0474
T	0.0480	0.0480	0.0476	0.0474

Table V. Distribution: No+La

Weight-function	Degree of contamination (%) 5	10	15	20
D	0.0476	0.0478	0.0481	0.0486
H	0.0483	0.0502	0.0566	0.0618
S	0.0478	0.0484	0.0493	0.0507
Ca	0.0476	0.0478	0.0480	0.0484
W	0.0475	0.0475	0.0471	0.0467
Lo	0.0482	0.0501	0.0570	0.0624
Lp	0.0485	0.0515	0.0593	0.0646
A	0.0476	0.0478	0.0470	0.0466
B	0.0476	0.0475	0.0470	0.0466
T	0.0476	0.0476	0.0471	0.0466

until the $T_n^{(j)}$ converge and gives the estimates of the unknown parameters. Therefore T_n is the so called w-estimate based on the weight function w, defined iteratively by Eq. (4).

3. The weight functions

The weight function w is defined by

$$w(x) = \begin{cases} \frac{\psi(x)}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases} \tag{5}$$

From every $\psi(x)$ function, belonging to $\varrho(x)$, the corresponding w-function can be derived. These w-functions can be classified according to the behaviour of $\psi(x)$ (soft descending, hard descending, redescending) and so put in three groups, see

No+Ca

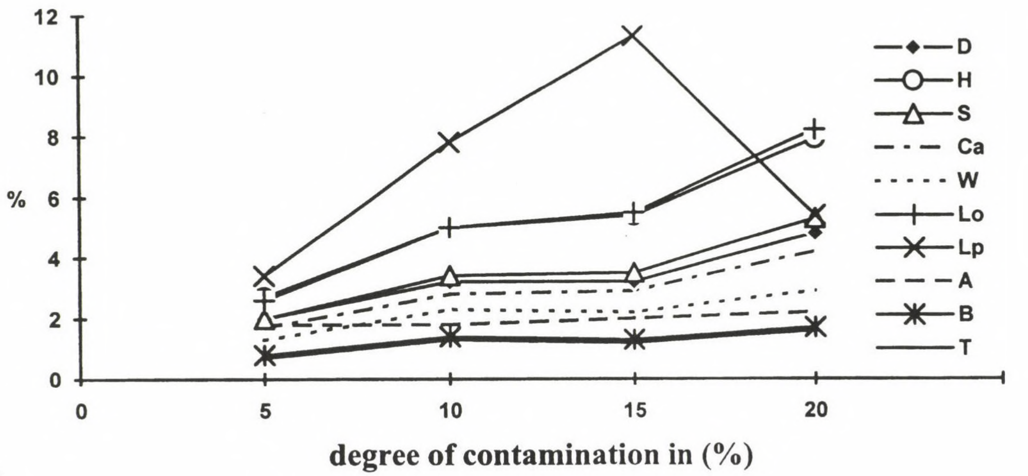


Fig. 3. Slope estimates for contaminating C (6 σ, 0)

No+La

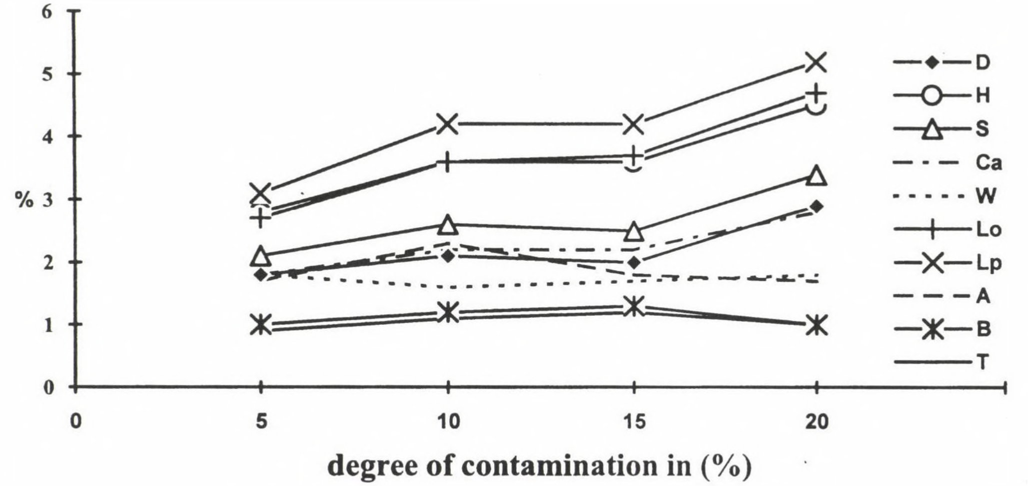


Fig. 4. Slope estimates for contaminating L (16 σ)

for instance Holland and Welsch (1977). We developed a robust program package based on w-functions, some of them are known from the literature and others have been introduced by us, Somogyi and Závoti (1993). On the basis of the behaviour of $\psi(x)$-functions we propose to classify the methods as follows: mixed normal, unbroken and cut off type. Table I shows these w-functions.

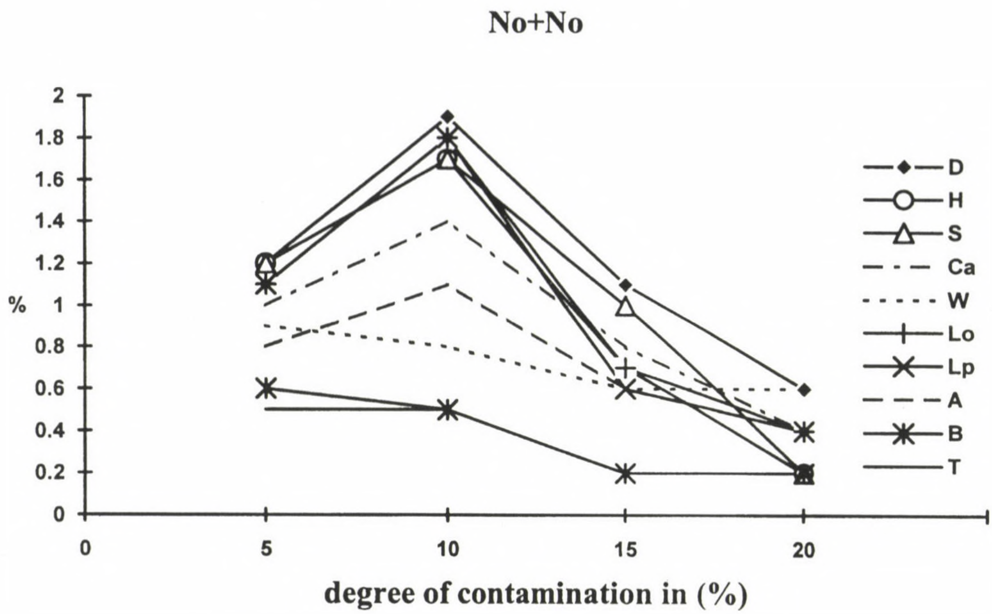


Fig. 5. Intercept estimates for contaminating N (0.10 σ)

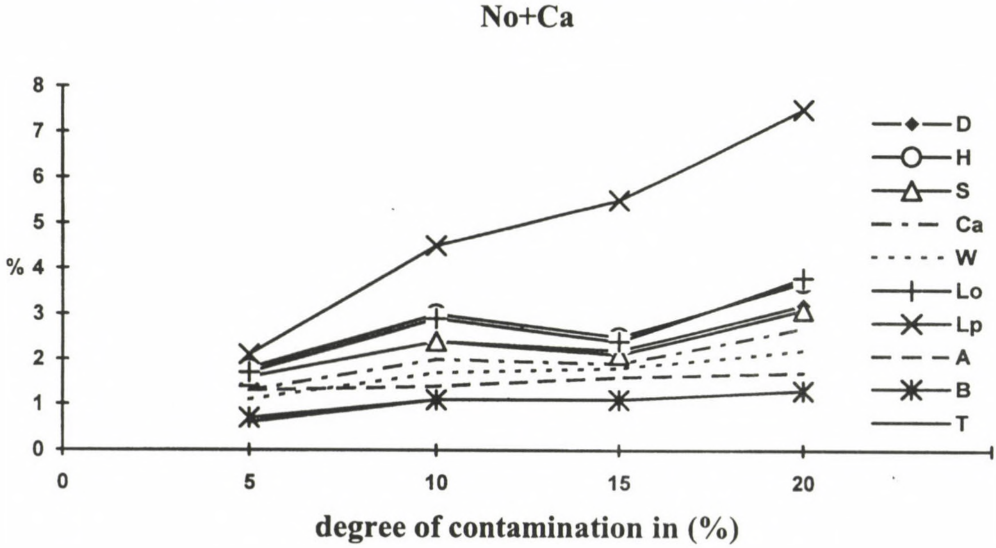


Fig. 6. Intercept estimates for contaminating C (6 σ, 0)

In order to demonstrate the trends of these w-functions, we generated to the y_i values of a simulated linear regression ($y_i = x_2 t_i + x_1$) noise of normal distribution with a few outliers. Choosing $\sigma = \pm 0.05$ units for the standard deviation of the assumed measurements, we obtained 0.15 for the acceptable error. The r_i residuals

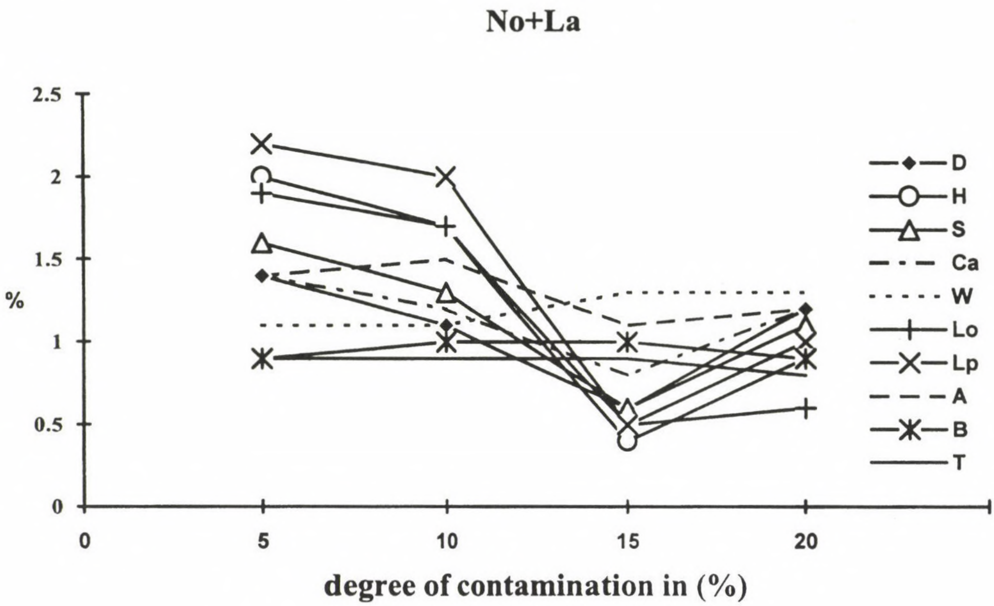


Fig. 7. Intercept estimates for contaminating L (16 σ)

and the w_i weights of the regression model consisting of 21 points are presented in Table II. From this table it can be seen, that the w_i values corresponding to the r_i residuals show a significant difference dependig on the type of the $\psi(x)$ functions. For this reason the weighted standard deviation which is used in geodesy would not give a correct information about the efficience of the estimation comparing the different weight functions. In group one weight functions D and S, in group two Ca and Lo and in group three A and B give essentially the same result. The w_i weights of functions Lp and T deviate significantly from the others, which can be explained with the form of their weight functions (see Fig. 1).

4. Linear regression model

We will consider the linear regression model

$$y = Ax + v \ , \tag{6}$$

where y and v are $n \times 1$ vector, A is a $n \times m$ matrix, x is a $m \times 1$ vector of unknown parameters.

For the robust estimate of x, we have to minimize Eq. (6), with a robust ϱ function:

$$\min \sum_{i=1}^{n} \varrho\left(\frac{y_i - a_i x}{\sigma}\right) \ , \tag{7}$$

where a_i is the i-th row of A.

From the necessary condition for a minimum we get that x satisfies

$$\sum_{i=1}^{n} a_{ij}\psi\left(\frac{y_i - a_i x}{\sigma}\right) = 0 \qquad \text{for all } j\,. \tag{8}$$

In general $\psi(x)$ is a nonlinear function and Eq. (8) provides a set of nonlinear equations the solution of which requires iterative methods.

Starting with a value $x^{(0)}$, let $x^{(j+1)}$ represent the estimate of x from the j-th iteration:

$$x^{(j+1)} = (A^{\perp}W^{(j)}A)^{-1}A^{\perp}W^{(j)}y \tag{9}$$

and

$$x^{(j+1)} = x^{(j)} + (A^{\perp}W^{(j)}A)^{-1}A^{\perp}W^{(j)}(y - Ax^{(j)})\,, \tag{10}$$

where $W^{(j)}$ denote an $n \times n$ diagonal matrix with the element

$$w_{ii}^{(j)} = w\left(\frac{y_i - a_i x^{(j)}}{s}\right)\,. \tag{11}$$

Initial estimates of parameters was given by the least squares method with unit weights. After that the iterative weighted least squares were used to determine the parameters.

The whole process can be repeated until j reaches same given value or — as an even better solution — we can use the convergence criterion:

$$\max_{k}\left\{\left|x_k^{(j+1)} - x_k^{(j)}\right| / \left|x_k^{(j)}\right|\right\} \le 0.001\,. \tag{12}$$

5. Monte Carlo study

To compare the w-functions presented in Table I, a small Monte Carlo study was performed. To the y_i values of a simulated linear regression ($y_i = x_2 t_i + x_1$), various error distributions were generated with different type of contaminations. These are as follows: No (normal), Ca (Cauchy), La (Laplace), No+No (normal + normal contaminated), No+Ca (normal + Cauchy contaminated), No+La (normal + Laplace contaminated). For the last three cases the following degrees of contaminations were used: 5,10,15,20 (in %). Choosing again the value of 0.05 for the standard deviation of the assumed measurements, to the adjustment the acceptable error is limited by 0.15.

Hundred samples of $n = 100$ were generated for each combination. This number of samples was sufficient to demonstrate the behaviour of the w-functions in the various cases.

Tables III, IV, V display the standard deviations of the location estimates from the samples. To the estimation of these values the trimmed mean was used.

In order to compare the relative performance of the w-functions, we also computed the relative efficiency given by Andrews et al. (1972):

$$\text{relative efficiency} = \frac{\text{minimum variance}}{\text{variance}}\,,$$

Table VI. Distribution: No+No

Weight-function	Degree of contamination (%) 5	10	15	20
D	1.0000	0.9229	0.9242	0.8952
H	0.9893	0.8904	0.8480	0.7718
S	0.9978	0.9210	0.9025	0.8530
Ca	1.0000	0.9343	0.9278	0.9038
W	1.0000	1.0000	1.0000	1.0000
Lo	0.9893	0.8904	0.8357	0.7484
Lp	0.9830	0.8801	0.8253	0.7276
A	1.0000	0.9381	0.9358	0.9234
B	0.9978	0.9785	1.0000	0.9979
T	0.9978	0.9785	0.9978	0.9979

Table VII. Distribution: No+Ca

Weight-function	Degree of contamination (%) 5	10	15	20
D	0.9958	0.9876	0.9835	0.8316
H	0.9836	0.8556	0.8081	0.7270
S	0.9938	0.9816	0.9714	0.8186
Ca	0.9979	0.9917	0.9876	0.8745
W	1.0000	0.9979	0.9979	0.9367
Lo	0.9856	0.8541	0.8000	0.7117
Lp	0.9638	0.7780	0.6584	0.5725
A	0.9958	1.0000	1.0000	0.9386
B	0.9979	1.0000	1.0000	1.0000
T	0.9979	1.0000	1.0000	1.0000

Table VIII. Distribution: No+La

Weight-function	Degree of contamination (%) 5	10	15	20
D	0.9979	0.9937	0.9771	0.9588
H	0.9834	0.9462	0.8304	0.7540
S	0.9937	0.9814	0.9533	0.9191
Ca	0.9979	0.9937	0.9792	0.9628
W	1.0000	1.0000	0.9979	0.9978
Lo	0.9855	0.9481	0.8246	0.7468
Lp	0.9794	0.9223	0.7926	0.7214
A	0.9979	0.9937	1.0000	1.0000
B	0.9979	1.0000	1.0000	1.0000
T	0.9979	0.9979	0.9979	1.0000

where the minimum is taken over the various weight-functions considered in this study. Thus at least one w-function (estimate) has a relative efficiency of 1, and all the other w-functions will have a value less than 1. Tables VI, VII and VIII show the relative efficiencies for the w-functions respectively.

Figures 2–4, respectively 5–7 display the deviation of the estimated slopes and intercepts coefficients (in %) from the given values.

On the basis of these tables and figures it can be said, that the estimations with the different w-functions show no great discrepancies from each other.

6. Conclusions

The purpose of this investigation was to compare 10 different w w-functions in robust regression using iteratively reweighted leastsquares, and to examine the behaviour of these w-functions under different distributions of the observations.

On the basis of results it can be seen, that the w_i values corresponding to the r_i residuals show a significant difference dependig on the type of the $\psi(x)$ functions. The influence of the different error distributions can not be determined using the various w-functions. As it appears from the tables and figures, there are no significant deviations betweeen the different computations, and the tests we have made show that each of the methods can be used for the robust adjustment of geodetic measurements.

The above statements are in accordance with the criteria of "being robust". In addition, we can note that applying a non-robust estimation may lead to very bad estimated values if the distribution of the model is not exactly normal. Stating it more explicitly, we can say that in case of contaminations of 5–20 % the mixed normal methods yield satisfactory results. We can read off from Figs 2–7 that estimations of the type "cut off" are generally superior to other methods. On the other hand, when dealing with mixed distributions, the least satisfactory results are produced by methods of the type "unbroken".

Acknowledgement

This work was supported by OTKA (National Science Foundation) under the Grants No. 014148 and 014856.

References

Andrews D F 1974: *Technometrics*, 16, 523–531.

Andrews D F, Bickel P J, Hampel F R, Huber P J, Rogers W H, Tukey J W 1972: Robust Estimates of Location: Survey and Advances. Princeton Univ. Press

Beaton A E, Tukey J W 1974: *Technometrics*, 16, 147–185.

Faigh W, Owolabi K 1988: In: Proceedings of ACSM-ASPRS Annual Convention, St.Louis, 169–180.

Holland P W, Welsch R E 1977: *Commun. Statist.-Theor. Meth.*, A6(9), 813–827.

Krarup T, Juhl J, Kubik K 1980: *International Archives of Photogrammetry*, 23(B3), 369–378.

Somogyi J, Kalmár J 1991: *Acta Geod. Geoph. Mont. Hung.*, 26, 57–68.
Somogyi J, Závoti J 1993: *Acta Geod. Geoph. Mont. Hung.*, 28, 413–420.
Tukey J W 1970–1971: Exploratory Data Analysis, Addison-Wesley, Reading, Mass.
Veress S A, Youcai H 1987: *Phot. Eng. and Remote Sensing*, 53, 167–169.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 25–35 (1996)

CONNECTION OF HEIGHT SYSTEMS OF HUNGARY AND CROATIA

S KLAK[1], L FEIL[1], N ROZIC[1]

[Manuscript received December 6, 1995]

The paper gives the outline of analyses made with reference to the measuring accuracy, datum comparison and the results of adjusting the network of Austro-Hungarian precise levelling at the territory of the Republic of Croatia, Bosnia and Herzegovina and Slovenia. The results of these analyses are very well interpolated with the problems of connecting modern height systems of Hungary and Croatia.

Keywords: Austro-Hungarian precise levelling; connecting of height systems of Croatia and Hungary; levelling network; mean sea level

1. Introduction

At the end of 1995 there was the initiative taken to connect contemporary height systems, i.e. levelling networks of geometric levelling, between the Republic Hungary and the Republic of Croatia. The initiative lead to the first discussion of expert groups from both countries which were held in mid-October 1995 in Budapest. On this occasion, the need to connect levelling networks (height systems) was fully agreed upon and principal opinions regarding the number of connection points, as well as their location, i.e. levelling routes, synchronized.

However, it is interesting, not only from the historical point of view, that Hungary and Croatia used to be not only very well connected within the former Austro-Hungarian monarchy, but also included into the same height system. The basis of this system was the precise levelling of Austro-Hungarian monarchy. Naturally, in the researches made so far in Croatia, a part of this levelling encompassing the territory of the Republic of Croatia is of a special interest, but because of its characteristic shape, also the territory of Bosnia and Herzegovina, and Slovenia. Regardless of the time span of about 90 years that have passed ever since the Austro-Hungarian precise levelling was made, a number of bench-marks has been kept until today and included by means of new measurements into the modern networks of geometric levelling. Therefore, the old and new measuring data made it possible, although originating from different epochs, to determine recent crustal movements at the territory of Croatia, Bosnia and Herzegovina and Slovenia. Apart from that, it can also be presumed, although the authors of this paper have no concrete knowledge about it, that a certain number of bench-marks comprehended by the precise levelling lines from Austro-Hungarian Monarchy and situated at the territory of

[1]Faculty of Geodesy, University of Zagreb, 10000 Zagreb, Kačićeva 26, Croatia, e-mail: Nevio.Rozic@public.srce.hr

Hungary, has retained its value and has been included into the contemporary levelling network of Hungary. Having in mind this fact, the height systems of Hungary and Croatia can be regarded as already connected, but it is necessary to carry out a new connection for the purpose of professional and scientific interpretation of the results of such works.

2. Precise levelling of Austro-Hungarian Monarchy at the territory of Croatia, Bosnia and Herzegovina and Slovenia

At the end of 19th century there were numerous fundamental geodetic works performed at the territory of former Austro-Hungarian Monarchy according to the recommendations of the international organisation operating in those days. They were supposed to be used in the process of solving a series of scientific and practical tasks. They implied the works on: triangulation, precise levelling, astronomic and gravimetric measurements, production and reproduction of maps etc. It is very significant, from today's point of view, that the results of the works were regularly and systematically published in appropriate publications. Thus, they can be used today for certain analysis of scientific or practical nature.

The precise levelling was carried out by Military Geographic Institute from Vienna according to the recommendation of the II General Conference, of government commissioners, for Central European Graduated Measurement held in 1867. The levelling was classified as precise levelling, and the works started about the year 1884 and were finished about 1906.

The authors of this paper found their special interest in the data and measuring results referring to the part of the network covering the area of Croatia, Bosnia and Herzegovina and Slovenia, Fig. 1. The consideration and processing of measurements imply only classical network bench-marks, and not special marks on stone objects, because they disappeared or were destroyed long ago. Figure 1 contains the following for the considered network: levelling lines numbers, numbers of nodal bench-marks and numbers of levelling figures in accordance with the original data of that levelling.

It can be very clearly seen, Fig 1., that the whole and parts of the levelling lines stretching over the territory of today's Hungary are included into this network. These are the entire lines: No. 258 (Nagykanizsa–Zákány), No. 259 (Zákány–Barcs) and No. 260 (Barcs–Villány), and the parts of lines No. 271 (Kranichsfeld–Nagykanizsa), No. 252 (Villány–Dalj), No. 247 (Dalj–Szeged) and No. 242 (Szeged–Timisoara).

Because of the analyses connected with the problem of the mean sea level of the Adriatic Sea which has been determined in the origin of the network (Molo Sartorio in Triest — bench-mark BV 1), the level difference between the tide gauge in Bakar (bench-mark BV) and the closest bench-mark of the network of Austro-Hungarian precise levelling has been also included into the network, although it had not been originally measured within the frame of this levelling. It should be pointed out, that the tide gauge in Bakar is the tide gauge with the longest period of measuring the sea level on the Croatian coast of the Adriatic Sea, which is much longer than the indispensable 18.6 years.

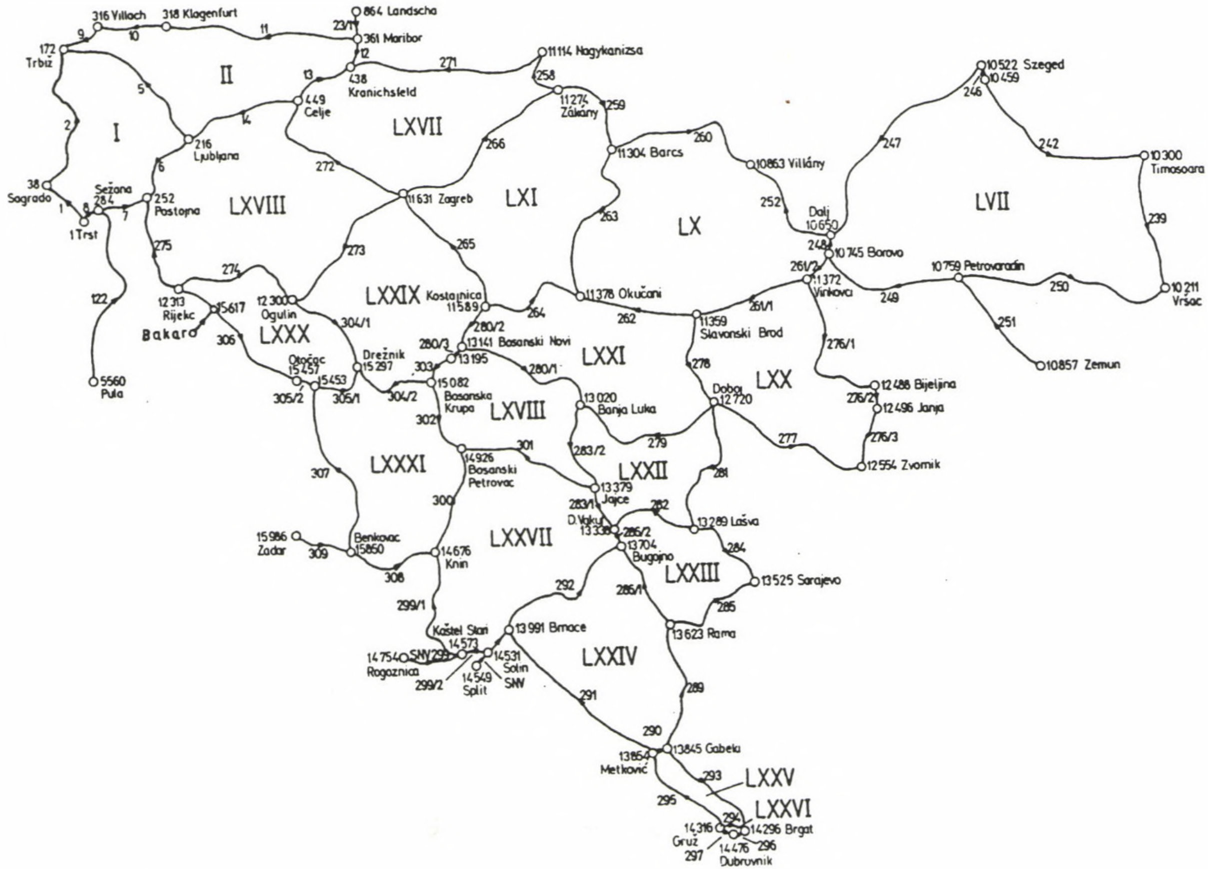


Fig. 1. Network of Austro-Hungarian precise levelling a the territory of Croatia, Bosnia and Herzegovina and Slovenia

3. Preparation and analysis of measuring data

Measuring results for the considered part of the network of Austro-Hungarian precise levelling were published in special publications "Astronomisch-geodätische Arbeiten ..." in edition of Military and Geographic Institute from Vienna, and that is:

— for the western part of the network, lines from No. 1 to No. 139, in publication (Militär-geographisches Institut 1896)

— for the northern part of the network, lines from No. 140 to No. 213, in publication (Militär-geographisches Institut 1897)

— for the south-eastern part of the network the lines from No. 214 to No. 275, in publication (Militär-geographisches Institut 1899).

The theoretical explanations of the network project, measuring method, using of equipment and similar, are contained in a special publication (Militär-geographisches Institut 1875). These publications have got the following items for every levelling line: numbers of bench-marks, distance and height differences measurement data, latitudes of bench-marks and computation corrections. By taking over these data, the files have been made on magnet media of PC computer for each line containing among other things the final results of the measured mean normal-orthometric height differences as the most important data. Table I gives the insight into one of these files for the line No. 14 (Mauthaus and der Savebrücke – Cilli).

Certain difficulties have cropped up during the defining of the files for lines from No. 276 to No. 309. These lines cover the area of Bosnia and Herzegovina and Southern Croatia (Dalmatia). For theses lines, the measuring data have namely not been attainable in such a form as for lines from No. 1 to No. 275, but the measured mean normal-orthometric height differences were computed retrogradely on the basis of tentative adjustment data and closing of levelling figures. This results from the fact that the individual lines were measured successively and then immediately individually adjusted within the frame of single levelling figures. Original data of these adjustments were published in a lot of publications: (Militär-geographisches Institut 1899, Militär geographisches Institut 1900, Militär-geographisches Institut 1901, Militär-geographisches Institut 1902, Militär-geographisches Institut 1905, Militär-geographisches Institut 1907, Militär-geographsiches Institut 1909).

The misclosures with respect to the closing of single levelling figures, inscribed levelling figure, elements for the evaluation of measuring accuracy, as well as the other indispensable data with Fig. 1 are written into the Table II.

On the basis of the data from this table, the reference probable random measurement error has been determined where:

$$u_F = \pm\frac{2}{3}\sqrt{\frac{1}{n_F}\left[\frac{\omega\omega}{F}\right]} = \pm 3.2\ \mathrm{mm}/\sqrt{\mathrm{km}}\,, \tag{1}$$

where:

n_F — number of levelling figures (including the inscribed levelling figure)

F — length of the levelling figure

ω — misclosure of levelling figure.

This accuracy can be considered as sufficiently acceptable with regard to the used instruments, equipment, working procedures and measuring methods.

If the sizes of misclosures in single levelling figures are considered more thoroughly, Table II, it can be seen that the biggest misclosure appears in LXVII levelling figure. It is interesting that this figure lays partially at the territory of Zagreb,

Table I. File for the line No. 14

Bench-mark	Geographic latitude	Length km	Height diff. m	Orth. corr. mm	Final height diff. m
BV 216	46 06 04	0.000	0.0000	0.0	0.0000
BV 476	46 06 09	0.881	3.2632	0.0	3.2632
BV 477	46 06 40	2.371	5.9028	−0.2	5.9026
BV 478	46 07 59	2.760	−4.7588	−0.6	−4.7594
BV 479	46 08 15	2.486	1.1050	−0.1	1.1049
BV 480	46 08 56	2.083	7.5083	−0.3	7.5080
BV 481	46 09 12	0.970	1.3734	−0.1	1.3733
BV 482	46 09 40	2.959	24.8850	−0.2	24.8848
BV 483	46 10 18	2.785	2.4422	−0.3	2.4419
BV 484	46 10 15	1.629	4.5329	0.0	4.5329
BV 485	46 10 05	2.917	32.5783	0.1	32.5784
BV 486	46 10 28	3.626	24.2696	−0.2	24.2694
BV 487	46 10 40	1.446	13.8935	−0.2	13.8933
BV 488	46 10 34	2.707	43.4131	0.1	43.4132
BV 489	46 11 13	2.329	61.2411	−0.5	61.2406
BV 490	46 11 20	2.352	46.2166	−0.1	46.2165
BV 491	46 12 27	3.327	−91.3124	−0.9	−91.3133
BV 493	46 13 03	2.858	−79.8990	−0.4	−79.8994
BV 494	46 13 37	2.013	−30.5210	−0.3	−30.5213
BV 495	46 14 45	2.311	−19.5116	−0.6	−19.5122
BV 496	46 14 58	1.009	−9.0578	−0.1	−9.0579
BV 497	46 14 58	1.523	−11.2908	0.0	−11.2908
BV 498	46 14 52	2.360	−12.8523	0.0	−12.8523
BV 499	46 14 48	3.031	−17.3213	0.0	−17.3213
BV 500	46 15 27	3.781	−12.0534	−0.2	−12.0536
BV 501	46 15 26	2.550	−7.2379	0.0	−7.2379
BV 502	46 15 11	3.347	−13.9920	0.1	−13.9919
BV 504	46 14 42	2.942	−9.3474	0.2	−9.3472
BV 505	46 14 28	1.135	−1.9658	0.1	−1.9657
BV 506	46 14 10	1.710	−4.8373	0.1	−4.8372
BV 449	46 13 48	2.553	−2.5675	0.1	−2.5674
Sum:		70.751	−55.9013	−4.5	−55.9058

	Year	Observer	Instrument	Lev. staff
	1874	Hoffmann	2984	C

Table II. Misclosures of levelling figures

Levelling figure	Misclosure m	Length km
I	0.0879	371.828
II	0.0862	436.124
LVII	0.0893	600.475
LX	−0.1479	455.476
LXI	0.1252	446.334
LXVII	−0.2392	410.676
LXVIII	−0.0821	516.051
LXX	0.0464	414.094
LXXI	0.0287	458.872
LXXII	0.0139	405.690
LXXIII	−0.0265	279.438
LXXIV	−0.0756	424.953
LXXV	0.0378	199.948
LXXVI	−0.0006	20.577
LXXVII	−0.1045	518.401
LXXVIII	0.0258	357.502
LXXIX	0.1849	411.274
LXXX	−0.0021	342.300
LXXXI	0.1266	505.748
External poligon	0.1742	2869.057

where there was a pretty strong earthquake in 1880. In order to determine the consequences of this earthquake, the Military and Geographic Institute from Vienna performed additional and relatively extensive levelling and triangulating works. In the concurrence of events, the measurements of precise levelling had been made in some parts of levelling lines No. 265 (Zagreb - Kostajnica), No. 266 (Zagreb - Vrbovec), No. 272 (Zagreb - Sutla) and No. 273 (Zagreb - Jaska) before the earthquake, and then within the frame of determining its possible consequences also after the earthquake. However, the differences between the two levelling, as well as the results of these examinations, indicated no significant movements caused by the earthquake (at least not for the wider area of Zagreb).

Nevertheless, the comparison with other measurements pointed out some illogicalities, because the misclosure of the LXVII levelling figure is determined by using a part of measurements carried out before the earthquake −17.79 cm, and it amounts to 23,92 cm with the part of measurement made after the earthquake, which is remarkably larger and refers to the possible blunder. This error is most probably connected with the measurements in levelling line No. 266.

4. Adjustment of network

After defining the files of single levelling lines, the work on adjustment of network by indirect observations in accordance with the principle of least squares has commenced. As it is partly to be seen from the previous chapter, the considered network, Fig. 1 has never been adjusted in that form, as separate and integer network. Namely, a larger part of this network (Slovenia, northern and western Croatia) has been adjusted within the frame of one part of the Austro-Hungarian levelling network, while single levelling figures (Bosnia and Herzegovina and South Croatia) have been successively adjusted by the method of inserting the levelling lines between two known heights.

From the files of levelling lines the data about mean normal-orthometric height differences and about the length of levelling lines have been taken over. As the given (fixed) bench-mark of the network, one has selected the bench-mark BV which is situated in the levelling line No. 1 (Triest–Sagrado). As its altitude, the value of 3.3520 m above the means sea level determined in 1875 by prof. dr. Farolfi from only one year observation (Militär-geographisches Institut 1884), has been accepted.

It should be specially pointed out, that the height difference between the tide gauge in Bakar, bench-mark BV, and the closest bench-mark of the network BV 15617 has been included into the network, although this measurement has not been made originally within the scope of the measuring the Austro-Hungarian precise levelling.

The adjustment of the network has been made by applying a special program system for adjusting the levelling networks NIVEL (Rožić 1992b), and the measuring weights have been determined reciprocally by means of levelling lines lengths. The adjusted values of heights of nodal bench-marks are given in the Table III, as well as the belonging mean errors.

On the basis of the adjustment the belonging reference probable error of measurement has been determined

$$u_\gamma = \pm\frac{2}{3}\sqrt{\frac{\mathbf{v}^t P \mathbf{v}}{n_f}} = \pm 3.3 \text{ mm}/\sqrt{\text{km}}\,, \tag{2}$$

where:

n_f — number of levelling lines

$\mathbf{v}$ — vector of measurement corrections

P —- diagonal matrix of measuring weight.

The amount of this error indicates to the complete agreement with the result of reference probable error determined on the basis of closing the levelling figures.

The relationship of the bench-mark BV heights in Bakar is very interesting. The value of the height of this bench-mark has been namely determined to the amount of 2.7803 m, and on the basis of observing the mean sea level at the tide gauge in Bakar, the height of this bench-mark was determined having the value of 2.6601 m (Bilajbegović et al. 1986). This height was determined with regard to the full period of observing the means sea level lasting from 18.6 years and it refers to the time of the 1.7.1971. (1971.5 ± 9.3 years). The comparison of these heights shows, apart from possible dynamic changes of the mean sea level in the course of time, that the mean sea level determined in Triest in 1875 is by 12.02 cm larger than the one that would be determined on the basis of full observation period. Thereby, attention should be drawn to the fact, that the mean error of the BV bench-mark height determined on the basis of adjustment amounts to ±5.35 cm, Table III.

Table III. Heights as related to Triest

Nodal bench-mark	Height m	Mean error mm	Nodal bench-mark	Height m	Mean error mm
BV	2.7803	±53.51	BV 12488	95.1209	±77.14
BV 38	31.1449	±28.73	BV 12496	107.2353	±77.36
BV 172	732.6616	±44.72	BV 12554	150.6659	±77.26
BV 216	296.1438	±42.15	BV 12720	145.1509	±68.68
BV 252	553.4157	±32.89	BV 13020	168.7783	±67.04
BV 284	361.2248	±20.48	BV 13141	123.7817	±60.96
BV 316	504.9208	±49.58	BV 13195	122.2298	±60.99
BV 318	443.3895	±54.19	BV 13379	387.7056	±67.92
BV 361	270.2014	±54.61	BV 13338	519.1964	±69.54
BV 438	272.4732	±53.65	BV 13289	360.1300	±72.10
BV 449	240.2539	±50.01	BV 13525	512.2525	±76.44
BV 5560	31.6186	±60.89	BV 13623	222.0140	±74.96
BV 10211	102.1971	±92.40	BV 13704	577.4721	±70.40
BV 864	265.9774	±60.47	BV 13845	9.1578	±80.15
BV 10300	91.1449	±92.84	BV 13854	7.7757	±80.19
BV 10459	85.9591	±88.54	BV 13991	323.3446	±74.56
BV 10522	83.6434	±88.22	BV 14316	3.5478	±86.98
BV 10650	101.0852	±72.33	BV 14476	3.6628	±87.18
BV 10745	91.4189	±71.76	BV 14296	262.6484	±86.99
BV 10759	83.0359	±83.11	BV 14531	7.7174	±74.66
BV 10857	78.2428	±92.78	BV 14549	4.5122	±75.72
BV 10863	109.3135	±72.25	BV 14573	92.2657	±74.49
BV 11114	151.5277	±61.06	BV 14676	224.0773	±69.66
BV 11274	132.2200	±60.11	BV 14754	1.4089	±80.67
BV 11304	108.4742	±64.18	BV 14926	675.7534	±65.31
BV 11359	94.8550	±67.88	BV 15082	157.5471	±60.85
BV 11372	93.3210	±70.97	BV 15297	424.0458	±58.05
BV 11378	119.9180	±64.46	BV 15453	466.0159	±58.96
BV 11589	112.5037	±60.45	BV 15457	459.4037	±58.97
BV 11631	123.6492	±55.01	BV 15850	182.0530	±70.63
BV 12300	324.3387	±55.68	BV 15986	1.1429	±76.68
BV 12313	5.0581	±48.54	BV 15617	3.8738	±52.33

It leads to the conclusion that the mean sea level in Triest (bench-mark BV 1) and Bakar (bench-mark BV) runs to about 12 cm (Triest gives higher altitudes). One should not forget the fact that this difference is determined for the mean sea levels defined in different epochs, and the period of observation lasted in Triest only for a year.

Through a special testing on the basis of measuring data of the first precise levelling originating from the former Yugoslavia, the stability of the bench-mark BV 15617, which has not changed its level position significantly in the meantime, has been checked and confirmed.

Referring to the previously determined difference of mean sea levels, the adjustment of the considered network has been done again. The bench-mark BV in Bakar has been defined as fixed (given bench-mark), for which there was the value of 2.6601 adopted as its height. The adjusted values of heights and the belonging mean errors are given in the Table IV.

In this way the network of the Austro-Hungarian precise levelling covering the area of Croatia, Bosnia and Herzegovina and Slovenia has been adjusted for the first time at once and the heights of nodal bench-marks have been determined as related to the mean sea level in Bakar. Thus, a unique height system has been defined on the basis of the measurements of Austro-Hungarian precise levelling which can usefully applied in solving a whole lot of scientific and professional problems. One

Table IV. Heights as related to Bakar

Nodal bench-mark	Height m	Mean error mm	Nodal bench-mark	Height m	Mean error mm
BV 1	3.2318	±53.51	BV 12488	95.0007	±71.06
BV 38	31.0246	±57.37	BV 12496	107.1151	±71.26
BV 172	732.5413	±56.96	BV 12554	150.5456	±70.98
BV 216	296.0236	±49.82	BV 12720	145.0306	±61.04
BV 252	553.2955	±43.75	BV 13020	168.6581	±58.47
BV 284	361.1046	±50.82	BV 13141	123.6615	±51.53
BV 316	504.8006	±59.67	BV 13195	122.1095	±51.54
BV 318	443.2692	±61.89	BV 13379	387.5854	±58.92
BV 361	270.0811	±56.98	BV 13338	519.0762	±60.76
BV 438	272.3529	±55.52	BV 13289	360.0097	±63.97
BV 449	240.1336	±51.76	BV 13525	512.1323	±68.67
BV 5560	31.4984	±76.62	BV 13623	221.8937	±66.83
BV 10211	102.0769	±87.79	BV 13704	577.3518	±1.69
BV 864	265.8571	±62.62	BV 13845	9.0375	±72.37
BV 10300	91.0246	±88.27	BV 13854	7.6555	±72.41
BV 10459	85.8389	±83.76	BV 13991	323.2244	±65.77
BV 10522	83.5232	±83.42	BV 14316	3.4276	±79.86
BV 10650	100.9649	±66.42	BV 14476	3.5426	±80.08
BV 10745	91.2987	±65.67	BV 14296	262.5282	±79.88
BV 10759	82.9157	±77.93	BV 14531	7.5972	±65.67
BV 10857	78.1226	±88.17	BV 14549	4.3920	±66.87
BV 10863	109.1933	±66.91	BV 14573	92.1454	±65.38
BV 11114	151.4074	±58.15	BV 14676	223.9570	±59.09
BV 11274	132.0997	±56.08	BV 14754	1.2886	±72.34
BV 11304	108.3539	±59.04	BV 14926	675.6331	±55.12
BV 11359	94.7348	±60.95	BV 15082	157.4269	±49.92
BV 11372	93.2008	±64.68	BV 15297	423.9256	±43.46
BV 11378	119.7977	±57.48	BV 15453	465.8957	±40.77
BV 11589	112.3835	±52.13	BV 15457	459.2834	±40.59
BV 11631	123.5290	±48.98	BV 15850	181.9328	±59.05
BV 12300	324.2185	±42.00	BV 15986	1.0227	±66.16
BV 12313	4.9379	±24.54	BV 15617	3.7536	±11.18

of the more important is by all means the problem of determining recent crustal movements, which has been solved in the meantime in one variant (Feil et al. 1992a) and Feil et al. 1992b). On the basis of adjusted heights of network nodal bench-marks, the calculations and adjustment of measurements have been made within the scope of single levelling lines. All data have served, for the purpose of easier and simpler usage, to the establishment of appropriate computer database ABAZA (Rožić 1992a) for PC by means of applying program systems DBASE III and CLIPPER. The basis conveys the following for the network: bench-mark number, number of levelling line that it is positioned in, latitude and longitude, original height of the bench-mark taken over from the corresponding publication of the Military and Geographic Institute from Vienna, adjusted height as related to the tide gauge in Bakar and the difference of these heights.

5. Connection of the network with Hungary

As it has already been stated, single lines in the marginal parts of the considered network are situated in Hungary either entirely or partly. Therefore the belonging values of adjusted heights can be stated for single bench-marks encompassed by theses lines. These values for nodal bench-marks of the network have already been given in the Table IV, and the bench-marks conveyed in the levelling line No. 258 (Nagykanizsa–Zákány), Table V, can be given as an example. For the bench-marks included into this levelling line, the descriptions of the position in words and original adjusted heights as related to Triest were taken over from the publication (Militär-geopgraphisches Institut 1899), and the adjusted altitudes as related to Bakar and differences in heights from the database ABAZA. The attention should be drawn to the fact that height differences, in this part of the network, have the same order of value, but they are not constant because of the influences of measurement errors and of the fact that the configurations of levelling networks are mutually different when originally adjusted and readjusted (previous chapter).

It can be assumed that a part of these bench-marks conveyed in these lines, have been kept until today and included into the contemporary levelling network

Table V. Data for the levelling line No. 258

Bench-mark	Position description	Height Triest m	Height Bakar m	Diff. cm
BV 11114	Nagy-Kanizsa, Stationsgebäude	151.5351	151.4074	12.77
BV 11261	Bahnwächterhaus Nr. 64	141.3649	141.2362	12.87
BV 11264	Bahnwächterhaus Nr. 62	139.0569	138.9274	12.95
BV 11267	Mura-Keresztur, Stationsgebäude	138.1750	138.0447	13.03
BV 11270	Bahnwächterhaus Nr. 2	134.8605	134.7290	13.15
BV 11273	Bahnwächterhaus Nr. 4	132.6379	132.5053	13.26
BV 11274	Zákány, Stationsgebäude	132.2328	132.0997	13.31

of Hungary within the scope of new measurements. Presuming that the retained bench-marks are stable, the relationship of height systems of Croatia and Hungary can be established in a very simple way on the basis of those old measurements. However, the relationship determined in such a way is of an approximate character, so it should be necessary to make modern and more accurate measurements.

6. Conclusion

On the basis of the previous explanations one can draw a conclusion that the network of Austro-Hungarian levelling on the territory of Croatia, Bosnia and Herzegovina and Slovenia has been for the first time adjusted as a special and entire network leaned on the mean sea level at the tide gauge Bakar. Therefore the measurements and the data of adjusting this network can be usefully applied for solving specific scientific and professional problems, as for example, the connection of levelling networks of Hungary and Croatia, at least approximately.

References

Bilajbegović A, Feil L, Klak S, Sredić S, Škeljo Lj 1986: II nivelman visoke točnosti SR: Bosne i Hercegovine, Crne Gore, Hrvatske, Slovenije i SAP Vojvodine, 1970–1973. Zbornik radova Geodetskog fakulteta, Niz D, Svezak broj 6/1–6/7.

Feil L, Klak S, Rožić N 1992a: Odredivanje visinskih pomaka na području Republike Hrvatske. Zbornik radova Geodetskog fakulteta Sveučilišta u Zagrebu u povodu 30. obljetnice samostalnog djelovanja 1962–1992., Zagreb, 17–25.

Feil L, Klak S, Roić M, Rožić N 1992b: *Österreichische Zeitschrift für Vermessungswesen und Photogrammetrie*, No. 2, 95–106.

Militär-geographisches Institut 1875: Die Astronomisch-geodätischen Arbeiten. Band VII, Wien

Militär-geographisches Institut 1884: Mitteilungen. Band IV, Wien

Militär-geographisches Institut 1896: Die Astronomisch-geodätischen Arbeiten. Band VIII, Wien

Militär-geographisches Institut 1897: Die Astronomisch-geodätischen Arbeiten. Band X, Wien

Militär-geographisches Institut 1899a: Die Astronomisch-geodätischen Arbeiten. Band XIV, Wien

Militär-geographisches Institut 1899b: Mitteilungen. Band XIX, Wien

Militär-geographisches Institut 1900: Mitteilungen. Band XX, Wien

Militär-geographisches Institut 1901: Mitteilungen. Band XXI, Wien

Militär-geographisches Institut 1902: Mitteilungen. Band XXII, Wien

Militär-geographisches Institut 1905: Mitteilungen. Band XXV, Wien

Militär-geographisches Institut 1907: Mitteilungen. Band XXVII, Wien

Militär-geographisches Institut 1909: Mitteilungen. Band XXIX, Wien

Rožić N 1992a: Kompjutorska baza podataka Austro-Ugarskog preciznog nivelmana - ABAZA. Geodetski fakultet Sveučilišta u Zagrebu, Zagreb

Ročić N 1992b: Kompjutorski program za izjednačenje nivelmanskih mreža - NIVEL. Geodetski fakultet Sveučilišta u Zagrebu, Zagreb

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 37–46 (1996)

REAL-TIME INTEGRITY MONITORING OF DUAL-FREQUENCY GPS OBSERVATIONS FROM A SINGLE RECEIVER

C D DE JONG[1]

[Manuscript received May 10, 1996]

Real-time integrity monitoring of GPS code and carrier observations is becoming increasingly important in the field of high-precision navigation and positioning. A description is given of a recursive procedure for the integrity monitoring of dual-frequency code and carrier observations from a single receiver. The procedure does not require any external information regarding station and satellite coordinates, velocities and clock behaviour and tropospheric effects and is only weakly dependent on the observation interval. It is shown that cycle slips as small as one cycle can be detected in real-time.

Keywords: GPS; hypothesis testing; integrity monitoring; Kalman filter

1. Introduction

Real-time integrity monitoring of GPS code and carrier observations is becoming increasingly important in the field of high-precision navigation and positioning. Originally, the accuracy requirements for most real-time applications were such that single frequency receivers were able to deliver satisfactory results. However, current developments show a trend towards using dual-frequency receivers for real-time applications, not only for improving positioning accuracy, but also to increase production: due to the greater redundancy, using dual frequency data allows for faster resolution of the integer cycle ambiguities of the carrier observations in real-time kinematic (RTK) applications of GPS, see e.g. (Teunissen 1993). Another trend is the increasing number of permanent GPS stations and networks for high-precision applications, such as geodynamics, maintenance of reference systems and land surveying, (Mueller 1993, de Jong et al. 1994, Gurtner 1995), which traditionally required the deployment of dual-frequency receivers, and which, due to the large amounts of data being collected and the associated long processing times, could benefit from a data validation step in the field, i.e., in real-time.

2. Testing and reliability

The real-time integrity monitoring procedure for dual-frequency GPS code and carrier observations to be discussed here can be considered independent of the actual application for which the observations are being collected (positioning, navigation, generation of differential corrections, etc.). Data processing (for the sole purpose

[1]Impalalaan 14, 5704 DB Helmond, The Netherlands

1217-8977/96/$ 5.00

of data validation) is done using a Kalman filter, see Appendix A. The integrity monitoring procedure, (Teunissen 1990a, 1990b), is also recursive and parallels the Kalman filter. The associated test statistics are based on the filter's predicted residual, which is defined as the difference between actual and predicted (using the predicted state) observation:

$$\nu_k = y_k - A_k x_{k|k-1} \,, \tag{1}$$

where ν_k is the m-vector of predicted residuals at time k, y_k is the vector of observations and $A_k x_{k|k-1}$ is the vector of predicted observations. Under the local (i.e., referring to a single epoch) null-hypothesis H_k^0 the predicted residuals are distributed as:

$$H_k^0: \quad \nu_k \sim N(0, Q_{\nu_k}) \,. \tag{2}$$

The null-hypothesis describes the case no model errors are present. The alternative hypothesis H_k^a describes the case model errors $\nabla\nu_k$ are present in the predicted residual:

$$H_k^a: \quad \nu_k \sim N(\nabla\nu_k, Q_{\nu_k}) \,. \tag{3}$$

The m-vector $\nabla\nu_k$ can be parametrised as:

$$\nabla\nu_k = C_k \nabla^k \tag{4}$$

with the m-by-b matrix ($b \leq m$) given by:

$$\begin{aligned} C_k &= (c_1\, c_2\, \dots\, c_b)_k \\ \nabla^k &= (\nabla_1\, \nabla_2\, \dots\, \nabla_b)_k^* \,. \end{aligned} \tag{5}$$

Here we will be dealing only with observation errors, which are assumed to be either outliers or slips.

The testing procedure used for the real-time validation of dual-frequency GPS code and carrier observations is known as the DIA-procedure (Teunissen 1990a, 1990b) and consists of the following steps:

1. *Detection*: An overall model test is performed to diagnose whether unspecified model errors have occurred.

2. *Identification*: After detection of model errors, identification of the potential source of these errors is required. This implies a search among the candidate hypotheses for the most likely alternative hypothesis.

3. *Adaptation*: After identification of an alternative hypothesis, adaptation of the recursive filter is needed to eliminate the presence of biases in the state vector.

In the detection step, unspecified model errors in the null hypothesis are detected using an overall model test. The LOM (Local Overall Model)-test statistic is given by:

$$T^k_{\mathrm{LOM}} = \frac{\nu_k^* Q_{\nu_k}^{-1} \nu_k}{m} \,. \tag{6}$$

If the LOM-test is rejected, the model errors have to be identified. This is done by examining the LS (Local Slippage)-test statistics, corresponding to each of the considered alternative hypotheses. For single model errors ($b = 1$, i.e., $C_k = c_k$), the LS-test statistic is given by:

$$t^k = \frac{c_k^* Q_{\nu_k}^{-1} \nu_k}{\sqrt{c_k^* Q_{\nu_k}^{-1} c_k}} \,. \tag{7}$$

The alternative hypothesis for which $|t^k|$ is at a maximum is considered to describe the most likely model error.

For the case multiple model errors are present ($b > 1$), the problem of local detection and identification of these errors can be solved recursively by examining one-dimensional alternative hypotheses. In each step, the effect of the largest LS-test statistic is removed from the LOM-statistic and all c-vectors, except the one describing the just-detected model error are orthogonalized with respect to this c-vector. This procedure is repeated until the LOM-test is accepted, (Teunissen 1990b). The result of this recursive procedure is the matrix of C_k (5).

After identification of the most likely alternative hypothesis, adaptation of the recursive filter is needed to eliminate the presence of biases in the filtered state of the system. In order to be able to adapt the filter, we first need an estimate of the identified model errors ∇^k. The best linear unbiased estimator of the b-vector ∇^k under H_k^a can be computed directly from the predicted residuals:

$$\nabla^k = (C_k^* Q_{\nu_k}^{-1} C_k)^{-1} C_k^* Q_{\nu_k}^{-1} \nu_k \tag{8}$$

with corresponding covariance matrix:

$$Q_{\nabla^k} = (C_k^* Q_{\nu_k}^{-1} C_k)^{-1} \,. \tag{9}$$

As was mentioned before, the errors ∇^k considered here consist of outliers and slips in the observations. More specifically, for the case of a slip, it will be assumed that it is a cycle slip of an integer number of cycles in the carrier observation. In that case, the error will be assumed non-stochastic, in other words, its corresponding covariance, see (9), will be zero. The adapted state vector $x_{k|k}^a$ for both types of errors is given by:

$$x_{k|k}^a = x_{k|k} - K_k C_k \nabla^k = x_{k|k} + k_k(\nu_k - C_k \nabla^k) \,, \tag{10}$$

where the state vector before adaptation $x_{k|k}$ and the gain matrix K_k are defined in Appendix A. The covariance matrix $P_{k|k}^a$ of the adapted state reads:

$$P_{k|k}^a = P_{k|k} + K_k C_k Q_{\nabla^k} C_k^* K_k^* \,. \tag{11}$$

For the expression of the covariance matrix $P_{k|k}$ of the state before adaptation, see again Appendix A.

To conclude this section, we will now briefly discuss the Minimal Detectable Bias (MDB), (Salzmann 1991). The MDB is defined as the size of the model error

∇ that can be detected with probability γ with the one dimensional test statistic (7). The quantity γ is called the power of the statistical test. For chosen reference probabilities α_0 (the probability of rejecting H_0 when it is actually true) and γ_0, the non-centrality parameter λ_0 is given by:

$$\lambda_0 = \lambda(\alpha = \alpha_0,\ b = 1,\ \gamma = \gamma_0)\ . \tag{12}$$

The expression for the MDB reads:

$$\mathrm{MDB} = \sqrt{\frac{\lambda_0}{c^*_{\nu_k} Q^{-1}_{\nu_k} c_{\nu_k}}}\ . \tag{13}$$

MDBs provide an important diagnostic tool for inferring how well particular model errors can be detected; they are referred to as the internal reliability of a system. For the particular case considered here, they may give an indication whether slips as small as one cycle in the carrier observations can be detected or not.

3. GPS observation equations

The integrity monitoring is based on the following observation equations for a carrier observation ϕ and a code observation C (both expressed in meters) at time k:

$$\begin{aligned} \phi_k &= R_k - I_k + \lambda N + \delta_\phi \\ C_k &= R_k + I_k + \delta_C \end{aligned}\ , \tag{14}$$

where:

R_k — contains all time dependent parameters, except the ionospheric effect, such as the geometric distance between receiver and satellite ϱ_k, receiver and satellite clock errors $\delta_r t_k$ and $\delta^s t_k$ and the tropospheric effect ΔT_k:

$$R_k = \varrho_k + c(\delta_r t_k + \delta^s t_k) + \Delta T_k \tag{15}$$

where c is the speed of light

I_k — ionospheric effect

N — constant carrier ambiguity

δ_Φ, δ_C — combined satellite and receiver delays (although these delays in general tend to change slowly with time, they are assumed constant here to prevent the dynamic models from becoming too complex)

λ — wavelength of one of the two GPS carriers.

4. Ionospheric modelling

In what follows it will be assumed that the ionospheric effect can be described accurately enough by its first order approximation:

$$I_k = \frac{40.3\,\mathrm{TEC}_k}{f^2} \tag{16}$$

where TEC_k is the Total Electron Content at time k and f is the frequency of one of the GPS carriers. In this case the relationship $I_{L2,k} = \kappa I_{L1,k}$ holds, where:

$$\kappa = f_1^2/f_2^2 \,. \tag{17}$$

The ionospheric effect is a function of time (and place). The dynamics of the actual effect and its time-derivatives will be modelled as a Taylor's series expansion:

$$I_k^{(i)} = \sum_{j=i}^{n_I} \frac{\Delta t_{k,k-1}^{j-i}}{(j-i)!} I_{k-1}^{(i)} \qquad i = 0, \ldots, n_I \tag{18}$$

with corresponding system noise, characterised by its spectral density $q^{(n_I+1)}$, which has dimension m^2/s^{2n_I+1}.

5. Real-time integrity monitoring procedure

Dual-frequency GPS receivers provide at least four observables (two code and two carrier measurements) from one satellite at one epoch. There will be enough redundancy to perform the integrity monitoring on a single channel basis. Using Eq. (14) and Eqs (16)–(18), the measurement model for one satellite at time k reads:

$$\begin{pmatrix} \phi_1 \\ C_1 \\ \phi_2 \\ C_2 \end{pmatrix}_k = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & -\kappa & 0 & 1 \\ 1 & \kappa & 0 & 0 \end{pmatrix} \begin{pmatrix} R_k \\ I_k \\ N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \delta_{\phi_1} \\ \delta_{C_1} \\ \delta_{\phi_2} \\ \delta_{C_2} \end{pmatrix} + \begin{pmatrix} 0 & \ldots & 0 \\ 0 & \ldots & 0 \\ 0 & \ldots & 0 \\ 0 & \ldots & 0 \end{pmatrix} \begin{pmatrix} I^{(1)} \\ \vdots \\ I^{(n_I)} \end{pmatrix}_k . \tag{19}$$

This model is not of full rank: it is not possible to separately estimate all parameters on the right-hand side, not even when more than one epoch of data is available. To make it regular, the following parameter transformation is applied:

$$\begin{pmatrix} S \\ I_1 \\ I_2 \\ I_3 \end{pmatrix}_k = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -\frac{\lambda_1}{2} & 0 \\ 0 & 1 & 0 & -\frac{\lambda_2}{\kappa+1} \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} R_k \\ I_k \\ N_1 \\ N_2 \end{pmatrix} + \begin{pmatrix} \delta_{C_1} \\ -\frac{\delta_{\phi_1} - \delta_{C_1}}{2} \\ -\frac{\delta_{\phi_2} - \delta_{C_1}}{\gamma+1} \\ \frac{\delta_{C_2} - \delta_{C_1}}{2} \end{pmatrix} \tag{20}$$

which results in the measurement model:

$$\begin{pmatrix} \phi_1 \\ C_1 \\ \phi_2 \\ C_2 \end{pmatrix}_k = \begin{pmatrix} 1 & -2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & -(\kappa+1) & 0 \\ 1 & 0 & 0 & \kappa-1 \end{pmatrix} \begin{pmatrix} S \\ I_1 \\ I_2 \\ I_3 \end{pmatrix}_k + \\ + \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} I^{(1)} \\ \vdots \\ I^{(n_I)} \end{pmatrix}_k . \tag{21}$$

Due to the changing geometry between satellite and receiver and the behaviour of the satellite and receiver clocks, the parameter S_k in (21) changes rapidly and irregularly with time and is therefore hard to predict for intervals longer than a few seconds using e.g. a low order polynomial. Therefore, this parameter will be eliminated from (21) by choosing one of the observations as reference observation and subtracting it from the remaining ones. This choice is arbitrary, but since C_1 has the simplest observation equation, it is preferred over the others. The corresponding transformation matrix T is given by:

$$T = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{22}$$

and the resulting measurement model by:

$$\begin{pmatrix} \phi_1 - C_1 \\ \phi_2 - C_1 \\ C_2 - C_1 \end{pmatrix}_k = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -(\kappa+1) & 0 \\ 0 & 0 & \kappa-1 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix}_k + \\ + \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} I^{(1)} \\ \vdots \\ I^{(n_I)} \end{pmatrix}_k . \tag{23}$$

As can be seen from (23) the state vector for each satellite consists of the following parameters:

— Three time-dependent biased ionospheric effects.

— n_I time-dependent derivatives of the ionospheric effect, see Eq. (18).

Since S_k is not present in Eq. (23) and since the ionospheric parameters change relatively slowly with time, the resulting model is rather insensitive to the observation interval. Also due to the elimination of S_k, no external information regarding station and satellite coordinates, velocities and clock behaviour and tropospheric effects is required.

Table I. Dual frequency code and carrier MDBs as function of observation interval and spectral density of the ionospheric model. For all examples the standard deviations of the L1 and L2 observations were 0.25, 0.35 m (code) and 0.0019, 0.0024 m (carrier). Observations were assumed uncorrelated. For the ionosphere, a linear model $(n_I + 1 = 2)$ was used. Testing parameters α_0 and β_0 were chosen as 0.001 and 0.80, resulting in a non-centrality parameter $\lambda_0 = 17.07$

Observation interval (s)	1		30	
Spectral density (m^2/s^3)	10^{-9}	10^{-8}	10^{-9}	10^{-8}
Code/carrier MDBs (m)				
L1	1.0/0.01	1.0/0.01	1.0/0.03	1.0/0.05
L2	1.4/0.01	1.4/0.01	1.4/0.03	1.4/0.05

The input for the dual-frequency integrity monitoring consists of the standard deviations of the observations, the order of the polynomial n_I to model the ionospheric effect and the corresponding spectral density. In Table I the MDBs of code and carrier observations for some typical values of these input parameters and observation intervals are given. From this table it can be concluded that the MDBs of the code observations are roughly equal to four times their standard deviations. For the MDBs of the carrier observations there seems to be a dependence on the chosen value for the spectral density and the observation interval. However, the MDBs are all much smaller than one cycle.

6. Hypothesis testing

If hypothesis testing is based on the assumption that the observations contain errors which are not related to the error in other observations, then the number of alternative hypotheses to be tested for each epoch would be equal to the number of observations. The c-vectors, see (5), could then simply consist of canonical unit-vectors. For the integrity monitoring procedure using the transformed model (23), the situation is more complicated: if the c-vectors would consist of unit-vectors, then only errors in the transformed observations could be identified. A better procedure would be to base the c-vectors on the original, untransformed model (21) and transform them accordingly using (22).

In the actual implementation of the integrity monitoring procedure, the number of alternative hypotheses to examine is larger than the original number of observations. Due to the transformation (22), it appeared to be difficult to identify particular combinations of cycle slips on L1 and L2, such as those for which the ratio is (almost) equal to the ratio of the two GPS frequencies ($77/60 \approx 9/7$), using the simple c-vectors, described above. However, if one expects this situation to occur, then one can also formulate the proper alternative hypothesis by a single c-vector, which, from a statistical point of view, is stronger than the two which correspond to simultaneous, independent errors on L1 and L2. In this way, it is relatively easy to identify these combinations of cycle slips.

Table II. Applied cycle slips, identified hypotheses and estimated errors. Observation interval is 30 seconds. Statistical parameters are the same as for Table I. For combined hypotheses, i.e., hypotheses describing related L1 and L2 errors, the estimated errors refer to the L1 errors. The L2 errors follow by applying the described relationship

Added slip (cycles)		Identified hypotheses	Estimated errors (cycles)		
L1	L2		L1	L2	L1*
1	0	Slip in L1	1.002	–	–
0	1	Slip in L2	–	0.999	–
1	1	Slip in L1 = Slip in L2	–	–	0.999
9	7	Slip in L1 = $9/7 \cdot$ Slip in L2	–	–	9.065
154	120	Slip in L1 = $f1/f2 \cdot$ Slip in L2	–	–	154.065
154	119	Slip in L1, Slip in L2	154.063	119.049	–

*Estimated errors, referring to related L1 and L2 errors, are expressed in L1 cycles

7. Some numerical results

The integrity monitoring procedure, as discussed in the previous sections, has been implemented in a software package for use at the reference stations of the experimental Active GPS Reference System in The Netherlands, (de Jong et al. 1994). For testing purposes, data has been collected since 1993. To show that the integrity monitoring actually works, a data set was selected with an observation interval of 30 seconds. Cycle slips were added to either L1, L2 or both. Next, the data was processed by the integrity monitoring software. Table II gives an overview of the added errors, the identified alternative hypotheses and the estimated errors. As can be seen from this table, all errors could be identified. The estimated errors were close to the integer cycle slips; these could be obtained by simply rounding to the nearest integer.

8. Conclusions and suggestions for future work

A procedure has been developed for the real-time integrity monitoring of dual frequency GPS code and carrier observations from a single receiver. The procedure is independent of the number of tracked satellites. It uses the recursive DIA procedure, (Teunissen 1990a, 1990b), to detect, identify and adapt for errors in the observations, which are assumed to be either outliers in code and carrier, or integer cycle slips in the carrier data.

The procedure does not require any external information regarding station and satellite coordinates, velocities and clock behaviour and tropospheric effects and is only weakly dependent on the observation interval, making it suitable for almost any kind application for which dual-frequency GPS data is used.

It was shown that cycle slips as small as one cycle can be identified; also "difficult" combinations of cycle slips, i.e., combinations of cycle slips in L1 and L2 for which the ratio is equal or close to the ratio of the L1 and L2 frequencies, can be identified.

An interesting extension of the integrity monitoring procedure and subject of future research could be in the field of relative positioning. Although the models developed in Section 5 apply to a single receiver, they can also be used for double difference code and carrier data. The delays δ_ϕ and δ_C will be eliminated, irrespective of the baseline length, whereas for short baselines the ionospheric effects will be greatly reduced or even eliminated. As a result the parameters $I_{1,k}$ and $I_{2,k}$ in (23) reduce to the double difference integer cycle ambiguities. These ambiguities can be determined in real-time on a single-channel basis, i.e., independent of the number of tracked satellites. This is not only interesting for routine RTK applications, but also for baselines with specific tropospheric conditions, due to e.g. large height differences or land-water transitions, for which it may sometimes be difficult to determine the integer ambiguities in the actual processing stage of the data, i.e., when determining these parameters together with the baseline components.

Acknowledgements

The work described in this contribution originated from an idea of Prof. Dr. P J G Teunissen, while the author was working as a research associate at Delft University of Technology. Current research activities by the author are carried out as a PhD student at the Technical University of Budapest, under supervision of Prof. Dr. J Ádám. Their support and enthusiasm is gratefully acknowledged.

Appendix A

Kalman filter equations

System dynamics:

$$x_{k+1} = \Phi_{k+1,k} x_k + d_k \tag{A.1}$$

where:
x_k — n-dimensional state vector at time k
$\Phi_{k+1,k}$ — n-by-n transition matrix
d_k — Process noise, associated covariance matrix Q_k.

Measurement model:

$$y_k = A_k x_k + e_k \tag{A.2}$$

where:
y_k — m-dimensional observation vector
A_k — m-by-n design matrix
e_k — observation noise, associated covariance matrix R_k.

Predicted state and associated covariance matrix:

$$\begin{aligned} x_{k|k-1} &= \Phi_{k,k-1} x_{k-1|k-1} \\ P_{k|k-1} &= \Phi_{k,k-1} P_{k-1|k-1} \Phi^*_{k,k-1} + Q_k \,. \end{aligned} \tag{A.3}$$

Filtered state and associated covariance matrix:

$$\begin{aligned} x_{k|k} &= x_{k|k-1} + K_k \nu_k \\ P_{k|k} &= (I - K_k A_k) P_{k|k-1} \end{aligned} \tag{A.4}$$

where:
ν_k — vector of predicted residuals:

$$\nu_k = y_k - A_k x_{k|k-1} \tag{A.5}$$

with covariance matrix Q_{ν_k}:

$$Q_{\nu_k} = R_k + A_k P_{k|k-1} A_k^* \tag{A.6}$$

K_k — Kalman gain matrix:

$$K_k = P_{k|k-1} A_k^* Q_{\nu_k}^{-1} \; . \tag{A.7}$$

References

de Jong C D, Lemmens R L G, van der Marel H 1994: *GPS Nieuwsbrief*, 9, 1–18.

Gurtner W 1995: Guidelines for a permanent EUREF GPS network. Presented at EUREF Symposium, Helsinki, Finland

Mueller I I 1993: In: Permanent Satellite Tracking Networks for Geodesy and Geodynamics, G L Mader ed., Springer-Verlag, Berlin

Salzmann M A 1991: *Bulletin Géodésique*, 65, 109–115.

Teunissen P J G 1990a: *IEEE Aerospace and Electronic Systems Magazine*, 5, 35–41.

Teunissen P J G 1990b: An integrity and quality control procedure for use in multi sensor integration. Proceedings ION GPS-90, Colorado Springs, 513–522.

Teunissen P J G 1993: Least-squares estimation of the integer GPS ambiguities. Invited lecture, General Meeting of the International Association of Geodesy, Beijing, China

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 47–66 (1996)

TRUNCATION EFFECTS IN USING SPHERICAL HARMONIC EXPANSIONS FOR FORWARD LOCAL GRAVITY FIELD MODELLING

G Papp[1] and Z T Wang[2]

[Manuscript received February 7, 1996]

The spherical harmonic expansion of Newton's potential is an evident way to compute the potential of density irregularities like surface topography or crustal thickness defined in a spherical coordinate system globally. Since the relatively low resolution of recent global density models does not make the detailed analysis of the capabilities and features of the method possible, it was tested in the situation of local density models. Both fictive and "real" models were used in the investigations.

The fictive models formed a series of single spherical caps having different parameters (capsize, thickness), the analytically computed exact potential of which was compared to different degree and order spherical harmonic expansions of their potential. It turned out that for a mass model as small as 3°in capsize its gravitational potential can be modelled up to the accuracy level of 99 % with a maximum degree of 360.

The real model, i.e., the model of the Neogene-Quaternary sediments in the Pannonian Basin, Hungary, was used to investigate the resolution capabilities of the spherical harmonic expansion in the situation of complicated and detailed local density models. The resolution was investigated in both the space and the frequency domain. For this purpose the potential of the real model computed in flat-earth approximation was the reference data set. The spectral relation between spherical harmonic expansion and the flat-earth computation was also determined empirically. It shows that the truncation of the series expansion implied by a finite number of the degree of expansion is not equivalent to a usual digital low-pass filtering. An effort was made to explain the deviations which result in ± 0.2 m s.d. of the geoid undulations at $l_{\max} = 360$ in the situation of $0.1° \times 0.1°$ tangential resolution of the density model.

Keywords: density distribution models; forward gravity modelling; power spectral density functions; prism integration; spherical harmonic expansion; transfer functions

1. Introduction

The gravitational potential field of a massive body is uniquely determined by its mass distribution according to Newton's law of gravitation. The prediction problem of the potential field from a volume density model is known as the forward problem in physical geodesy. Different approaches have been developed to solve it in different situations.

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

[2]Finnish Geodetic Institute, Geodeetinrinne 2, FIN–02430 Masala, Finland

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

On local areas with planar approximation, rectangular prism integration was used to predict geoid undulations (Nagy 1988, Papp and Kalmár 1995) and FFT technique was adopted for the same purpose and for the computation of terrain corrections (e.g. Forsberg 1985, Sideris 1994, Vermeer 1992) as well.

In global studies with spherical approximation, spherical harmonic expansions of the gravitational potential of the global topographic and isostatic masses were used extensively (Rummel et al. 1988, Engels et al. 1995).

In the recent work the precision, resolution and spectral properties of the latter approach were investigated by two numerical experiments in local situations.

The first one was based on a very simple, fictive and synthetic mass model (spherical cap) to test the effect of truncation on the precision of the expansion. The analytical expression of the potential of a spherical cap was used to obtain exact reference values for this experiment.

The second one was based on a real model of the Neogene-Quaternary sediments in the Pannonian Basin (Hungary) to investigate the resolution capabilities and other properties of the expansion. For this purpose the potential values obtained from the "flat-earth approximation version" of the mass model by using the rigorous potential formula of a rectangular prism (Nagy 1980) were applied as reference values. The idea behind this experiment was to check how the truncation of the series expansion relates to the usual low-pass digital filtering commonly used in e.g. (geo)physical interpretation of the gravity field.

2. Spherical harmonic expansion of the gravitational potential of disturbing masses under the surface of the Earth

As shown in Fig. 1, the Earth was assumed to be a sphere with radius R. A known disturbing mass body Ω was located totally beneath the surface of the spherical Earth.

The gravitational potential of the body Ω at point $P(r, \Theta, \lambda)$ can be expressed by the Newtonian integral

$$V_p = \int\limits_{\Omega} G \frac{dm}{l} = G \int\limits_{\Omega} \varrho(r', \Theta', \lambda') \frac{r'^2 \sin \Theta' dr' d\Theta' d\lambda'}{\sqrt{r^2 + r'^2 - 2rr' \cos \Psi}} \tag{1}$$

where r, Θ, λ are spherical coordinates; l is the spatial distance between computation point $P(r, \Theta, \lambda)$ and the point $Q(r', \Theta', \lambda')$ inside the body; Ψ is the angle between the radius vectors r and r' of P and Q respectively; $\varrho(r', \Theta', \lambda')$ is the density function; G is the gravitational constant and dm is the mass element.

The body Ω was divided into subbodies so that every subbody Ω_i forming a spherical shell with radius between r_1 and r_2 from the centre of the Earth, had a homogenous density ϱ_i. Then Eq. (1) becomes

$$V_p = \sum_i V_i = \sum_i G\varrho_i \int\limits_{\Omega_i} \frac{r'^2 \sin \Theta' dr' d\Theta' d\lambda'}{\sqrt{r^2 + r'^2 - 2rr' \cos \Psi}} \tag{2}$$

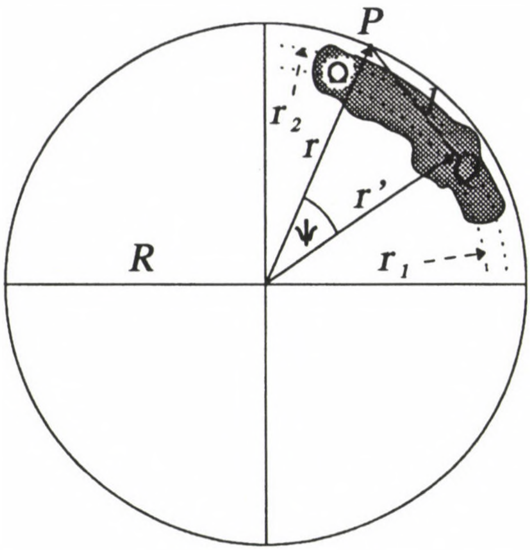


Fig. 1. Sketch figure for the determination problem of the gravitational potential in spherical coordinates

with $\cup\Omega_i = \Omega$. Of course potential V has to be piecewise integrable, i.e. the mass distribution inside the body Ω is regular.

Outside of Ω_i the gravitational potential V is a harmonic function and can be expanded into spherical harmonic series (Engels et al. 1995),

$$\begin{aligned} V_i &= G\varrho \int_{\lambda_1}^{\lambda_2} d\lambda' \int_{\Theta_1}^{\Theta_2} d\Theta' \sin\Theta' \int_{r_1}^{r_2} dr' \frac{r'^2}{\sqrt{r^2 + r'^2 - 2rr'\cos\Psi}} \\ &= G\varrho \int_{\lambda_1}^{\lambda_2} d\lambda' \int_{\Theta_1}^{\Theta_2} d\Theta' \sin\Theta' \int_{r_1}^{r_2} dr' \frac{r'^2}{r} \sum_{l=0}^{\infty} \left(\frac{r'}{r}\right)^l P_l(\cos\Psi) \\ &= G\varrho \int_{\lambda_1}^{\lambda_2} d\lambda' \int_{\Theta_1}^{\Theta_2} d\Theta' \sin\Theta' \int_{r_1}^{r_2} dr' \frac{r'^2}{r} \sum_{l=0}^{\infty} \frac{1}{2l+1} \left(\frac{r'}{r}\right)^l \cdot \\ &\quad \cdot \sum_{m=-l}^{l} \overline{Y}_{l,m}(\Theta,\lambda)\overline{Y}_{l,m}(\Theta',\lambda') \end{aligned} \tag{3}$$

for $r \geq r'$, where $\overline{Y}_{l,m}(\Theta,\lambda) = \begin{cases} \overline{P}_l^m(\cos\Theta)\cos m\lambda \ldots m \geq 0 \\ \overline{P}_l^{|m|}(\cos\Theta)\sin|m|\lambda \ldots m < 0 \end{cases}$ with $\overline{P}_l^m(\cos\Theta)$ being the fully normalized associated Legendre function. For clarity, the index i on all quantities in the right hand side of Eq. (3) was omitted. Furthermore, by performing the integral with respect to r' from r_1 to r_2 one obtains,

$$\begin{aligned} V_i &= G\varrho \int_{\lambda_1}^{\lambda_2} d\lambda' \int_{\Theta_1}^{\Theta_2} d\Theta' \sin\Theta' \sum_{l=0}^{\infty} \frac{1}{(2l+1)(l+3)} \frac{r_2^{l+3} - r_1^{l+3}}{r^{l+1}} \cdot \\ &\cdot \sum_{m=-l}^{l} \overline{Y}_{l,m}(\Theta,\lambda)\overline{Y}_{l,m}(\Theta',\lambda') . \end{aligned} \tag{4}$$

In practice Eq. (4) is not efficient because the computation of the r^{l+3} term becomes laboursome when l is large. The difficulty can be overcome by introducing depth D positive downward from the surface of the spherical Earth. Representing r by depth D,

$$\begin{aligned} r_1 &= R - D_2 \\ r_2 &= R - D_1 \end{aligned} \tag{5}$$

with D_1, D_2 representing the depth of the upper and lower boundary of the i-th subbody respectively.

Then using binomial expansion, one obtains (6),

$$\begin{aligned} r_2^{l+3} - r_1^{l+3} &= R^{l+3}\left[\left(1-\frac{D_1}{R}\right)^{l+3} - \left(1-\frac{D_2}{R}\right)^{l+3}\right] = \\ &= (l+3)R^{l+3}\frac{D_2-D_1}{R}\left\{1 - \frac{l+2}{2}\frac{D_2+D_1}{R} + \right. \\ &\left. + \frac{(l+2)(l+1)}{2\cdot 3}\frac{D_2^2+D_2D_1+D_1^2}{R^2} + \ldots\right\} . \end{aligned} \tag{6}$$

Substituting (6) into (4)

$$\begin{aligned} V_i &= G\varrho \sum_{l=0}^{\infty} \frac{R^{l+3}}{r^{l+1}} \frac{D_2-D_1}{R} \cdot \\ &\cdot \left\{1 - \frac{l+2}{2}\frac{D_2-D_1}{R} + \frac{(l+2)}{2}\frac{(l+1)}{3}\frac{D_2^2+D_2D_1+D_1^2}{R^2} + \ldots\right\} \cdot \\ &\sum_{m=-l}^{l} \overline{Y}_{l,m}(\Theta,\lambda) \int_{\lambda_1}^{\lambda_2} d\lambda' \left\{ \begin{array}{c} \cos m\lambda' \\ \sin|m|\lambda' \end{array} \right\} \int_{\Theta_1}^{\Theta_2} d\Theta' \sin\Theta' \overline{P}_l^{|m|}(\cos\Theta') \end{aligned} \tag{7}$$

the final formula suitable for computation was obtained. Equation (7) is more efficient than Eq. (4) due to the fact that in most situations the computation point P is assumed to be located on the surface of the Earth, i.e. $r = R$, thus the ratio R^{l+3}/r^{l+1} may be reduced to R^2. Furthermore, because $D \ll R$ in modelling of the Earth crust, the binomial expansion Eq. (6) up to degree 3 is sufficient.

Although theoretically Eq. (7) gives a rigorous solution, in practice, however, an approximation is unavoidable because Eq. (7) can not be expanded up to infinity. For global investigation of the geoid undulation of topographic and isostatic masses,

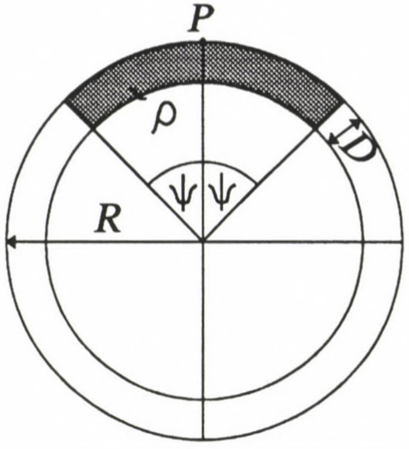


Fig. 2. Spherical cap with thickness D, capsize Ψ and density ϱ

maximum degree and order $l_{\max} = 30$ was found to be sufficient (Engels et al. 1995). In the following sections, it will be shown that a much higher value of $l_{\max}$ is required in local gravity field modelling if acceptable accuracy is desired. Two experiments have been carried out in order to investigate the properties of the expansion in local situations. The evaluation of Eq. (7) was based on a program code issued by Middel (1992).

3. Potential of a spherical cap with homogeneous density and constant thickness

In Fig. 2, a synthetic spherical cap with thickness D was assumed being located beneath the surface of the Earth (i.e. $D_1 = 0$ in Eq. (7)). The curvature radius of the cap is the same as that of the earth, its capsize is Ψ, and its density is ϱ.

The expression of the potential field for a spherical cap at an arbitrary point in the space is not known, but the potential on the central axis on its top (at point P) can be analytically expressed by Eq. (8) (see Appendix A).

$$\begin{aligned} V_p = 2\pi G\varrho\frac{1}{r}\Bigg\{&\left[-\frac{1}{2}rr'^2+\frac{1}{3}r'^3\right]_{R-D}^{R}+\left[\frac{1}{3}\sqrt{(r^2+r'^2-2rr'\cos\Psi)^3}\right]_{R-D}^{R}+\\ &+\frac{r\cos\Psi}{2}\left[(r'-r\cos\Psi)\sqrt{r^2+r'^2-2rr'\cos\Psi}\right]_{R-D}^{R}+\\ &+\frac{r^3\sin^2\Psi\cos\Psi}{2}\left[\ln(r'-r\cos\Psi+\sqrt{r^2+r'^2-2rr'\cos\Psi})\right]_{R-D}^{R}\Bigg\}, \end{aligned} \tag{8}$$

where r is the radial distance of point P; r' is the variable of integration which should be substituted by $R-D$ and R accordingly.

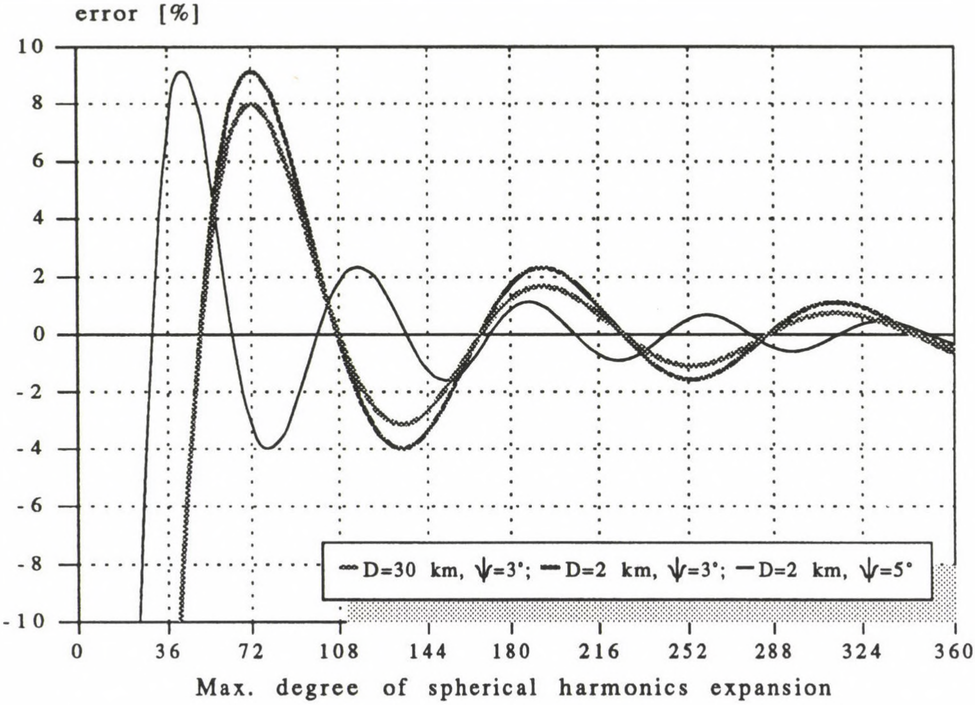


Fig. 3. Truncation errors of spherical harmonic expansions with different degree and order, in function of thickness D and capsize Ψ

The error of the gravitational potential computed from the spherical harmonic expansion was derived by subtracting the results obtained by Eq. (8) from the results obtained by Eq. (7). The percentage errors are shown in Fig. 3 and Fig. 4. From Fig. 3, where the capsize and thickness of the spherical cap varied, one can see that to reach certain accuracy level, the degree and order required depend on the capsize in such a way that the smaller is the cap size, the higher is the degree and order required. The thickness of the cap has small effect on the accuracy of the computation. From Fig. 4, where the maximum degree and order of the expansion varied, one may conclude that at point P more than 99 % of the gravitational potential of a homogeneous spherical cap can be approached with an expansion up to degree and order of 360 in the situation of a capsize as small as three degrees.

4. Geoid undulation contribution of the sediments in the Pannonian Basin, Hungary

Earlier investigation (Papp and Kalmár 1995) has shown that in the wavelength range below 300 km, the geoid undulations generated by the Neogene-Quaternary sediments dominate among the contributions of density irregularities in the lithosphere of the Pannonian Basin.

The peak-to-peak amplitude of the geoid anomalies reaches 1.5 m in the area

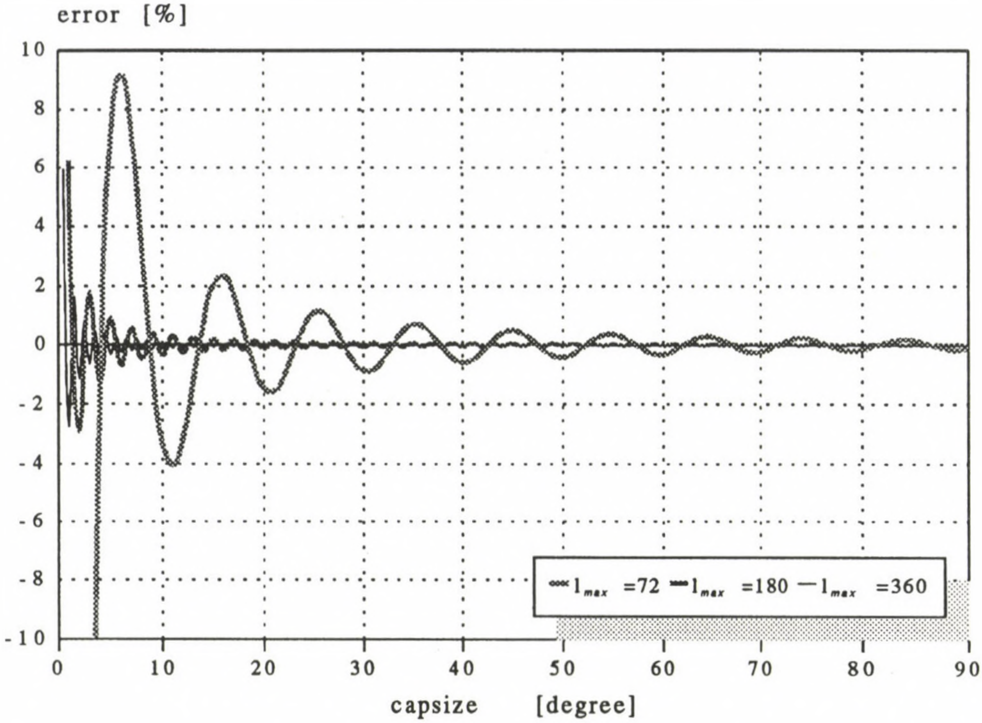


Fig. 4. Approximation errors of the gravitational potential of synthetic spherical caps with different capsize in function of different degree and order of spherical harmonic expansions

between $\varphi = 47°–48°$ and $\lambda = 17°–19°$. A very detailed 3D finite element density model (Kalmár et al. 1995) with a resolution 3–5 km in X-Y directions and some hundred meters in Z direction has been developed for the interpretation of the local features of the geoid. In the present study, however, the original model was transformed to a spherical system based on the map projection used by Cadastral Survey of Hungary, and then it was re-gridded on a sparser grid by using the utilities of the GMT software package (Wessel and Smith 1991). The model generated by this way was obviously suitable only for the spherical computations, therefore, it was transformed back to the Cartesian coordinate system of the map projection. In this process, an attempt was made to generate an "as similar as possible" model for the rectangular prism integrations. Eventually, two equivalent models were obtained (see Fig. 5). The spherical model consists of 1724 elements (spherical prisms with $0.1° \times 0.1°$ bases), whereas the Cartesian model consists of 1700 elements (rectangular prisms with 7×11 km bases, approximately corresponding to $0.1° \times 0.1°$ at the mean latitude of 47.5° of the study area). In order to simplify the computation, a homogeneous density of -460 kg/m^3, which is the mean density contrast between the sediments and the upper crust in the original 3D model was supplied.

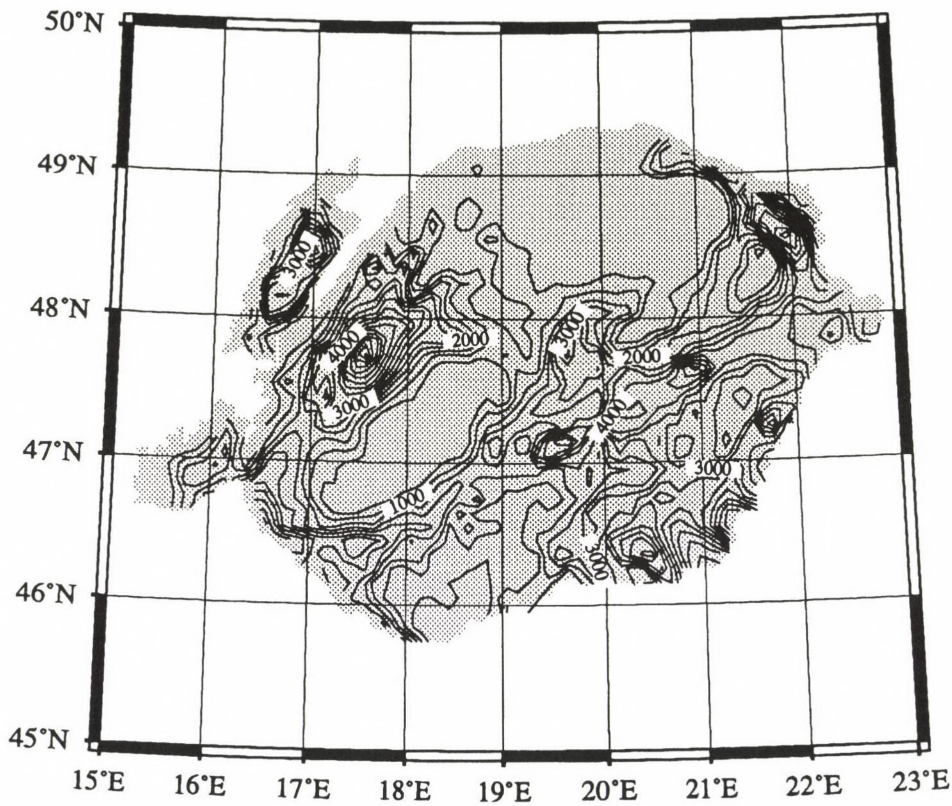


Fig. 5. Depth of the pre-Tertiary basement (thickness of the Neogene-Quaternary sediments) in the Pannonian Basin. Contour interval is 500 m. Shaded area represents the extension of available data used for density model generation

4.1 Computation of the potential in the Cartesian coordinate system

For a right rectangular prism having homogeneous density ϱ, the gravitational potential V can be computed based on Eq. (9) (Nagy 1980).

$$V = G\varrho\Bigg[xy\ln(z+r) + yz\ln(x+r) + zx\ln(y+r) - \\ -\frac{1}{2}\left(x^2\mathrm{arctg}\frac{yz}{xr} + y^2\mathrm{arctg}\frac{xz}{yr} + z^2\mathrm{arctg}\frac{xy}{zr}\right)\Bigg] \tag{9}$$

where r is the spatial distance between the computation point P and the corners of the prism; and x, y, z are coordinates of the corners related to the computation point. Eq. (9) was evaluated by the algorithm given by Nagy (1988).

Table I. Statistical parameters of the density model and of the geoid undulations obtained from forward potential modelling

	Data sets	Minimum (m)	Maximum (m)	Mean (m)	STD (m)	RMS (m)
1	thickness of sediments	0	8088.4	1742.7	±1494.0	±2295.0
2	results from integration	−8.68	−1.91	−4.40	±1.70	±4.72
3	results from series exp.	−8.41	−1.81	−4.35	±1.69	±4.66
4	difference (2–3)	−0.78	0.69	−0.06	±0.20	±0.21

4.2 Comparison of the results

The potential generated by the disturbing mass of the sediments was interpreted as disturbing potential, and therefore geoid undulations were obtained by Bruns, formula for the comparisons.

4.21 Comparison in the space domain

In order to make the direct point-by-point comparison possible, a regular spherical grid with $0.2° \times 0.2°$ spacing was defined for the computation of potential using spherical harmonic expansion. It was transformed to the plane for the rectangular prism integration. In this way, Eq. (9) and Eq. (7) were evaluated on points identical in the sense of map projection.

The geoid undulations from the spherical harmonic expansion ($l_{\max} = 360$) and from the prism integration are shown in Fig. 6 and Fig. 7, respectively. The differences between the two data sets, which can be interpreted as the effect of truncation of Eq. (7), are plotted in Fig. 8. The statistics of these data sets are listed in Table I together with the statistics of the mass model.

It can be seen from Table I that the truncation error is in decimeter level for $l_{\max} = 360$. It is in good accordance with the roughness of the density model (s.d. $= \pm 1494$ meters).

It is worthwhile to notice the strong correlation between the two sets of geoidal undulations with the mass model by comparing Fig. 6 and Fig. 7 with Fig. 5. The correlation factors are 0.60 and 0.66 for the series expansion and prism integration, respectively. The truncation error is also correlated with the mass model by comparing Fig. 8 with Fig. 5 although the correlation factor is smaller (0.19). This correlation reveals the smoothing process involved in the spherical harmonic expansion computation. Ideally, Fig. 8 should show random features and the correlation factor should be zero. The frequency domain analysis made the effects of smoothing more clear.

4.22 Comparison in the frequency domain — spectral interpretation of the results

As mentioned in the beginning of this section, computations have been carried out both on the sphere and on the plane. This situation is quite common especially

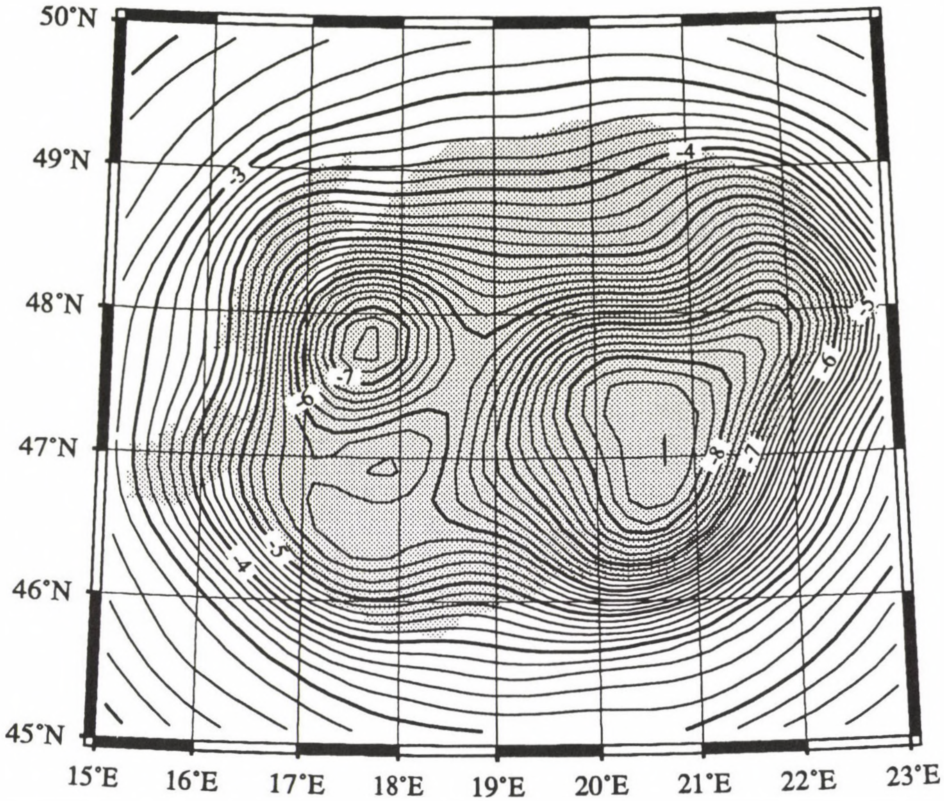


Fig. 6. Geoid undulation contribution of the sediments computed from spherical harmonic expansion approximation of the gravitational potential. Contour interval is 0.2 m. Shaded area represents the extension of the density model

in physical geodesy where global data are usually represented on the sphere/ellipsoid by spherical harmonic whereas local data are treated mostly on the plane of some kind of map projection. A typical example is the remove-restore technique used in the computation of terrain correction and geoid determination (Vermeer 1992, Vermeer and Forsberg 1992). Therefore the spectral properties of the spherical harmonics in planar approximation were studied, and an effort was made to investigate the spectral relations between the undulations obtained from the spherical harmonic expansion and the prism integration.

It was assumed that, in one hand, the undulations obtained from the prism integration contained all the spectral information that could be generated by the mass model and on the other hand, the spherical harmonic expansion up to certain degree and order worked as a low-pass filter with an approximate cut-off wavelength of two times of its highest resolution, i.e. $\lambda_{\text{cut-off}} \approx 2 \times 180°/l_{\max}$. Therefore it was supposed that a linear system existed by which the two kind of computations were related.

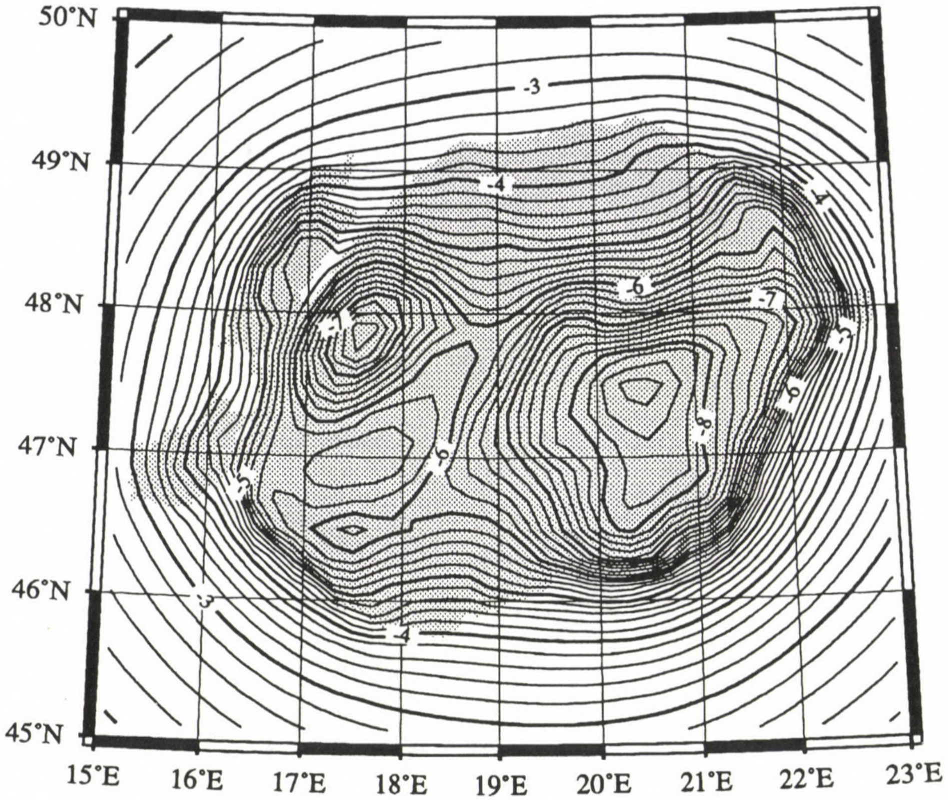


Fig. 7. Geoid undulation contribution of the sediments computed from rectangular prism integration. Contour interval is 0.2 m. Shaded area represents the extension of the density model

The frequency domain representation of the system function is the so called transfer function. In the frequency domain the relation may be written as Eq. (10).

$$R(f) = T(f)S(f) \tag{10}$$

where $S(f)$ is the spectrum of the signal $s(t)$; $T(f)$ is the transfer function and $R(f)$ is the response of the system to $S(f)$. Consequently, if $S(f)$ and $R(f)$ are known, the transfer function $T(f)$ can be determined by Eq. (11), in a somewhat similar way to deconvolution.

$$T(f) = R(f)/S(f) \; . \tag{11}$$

Applying Eq. (11) to the 2D complex spectra of the two sets of undulations, where the one obtained from the prism integration takes the place of $S(f)$ whereas the one obtained from the spherical harmonic expansion replaces $R(f)$, the transfer function was obtained. For this purpose the geoid undulations were calculated on a regular planar grid of 10×10 km including 128×128 points and then FFT combined with 100 % cosine tapering was applied to derive 2D Fourier spectra. After complex

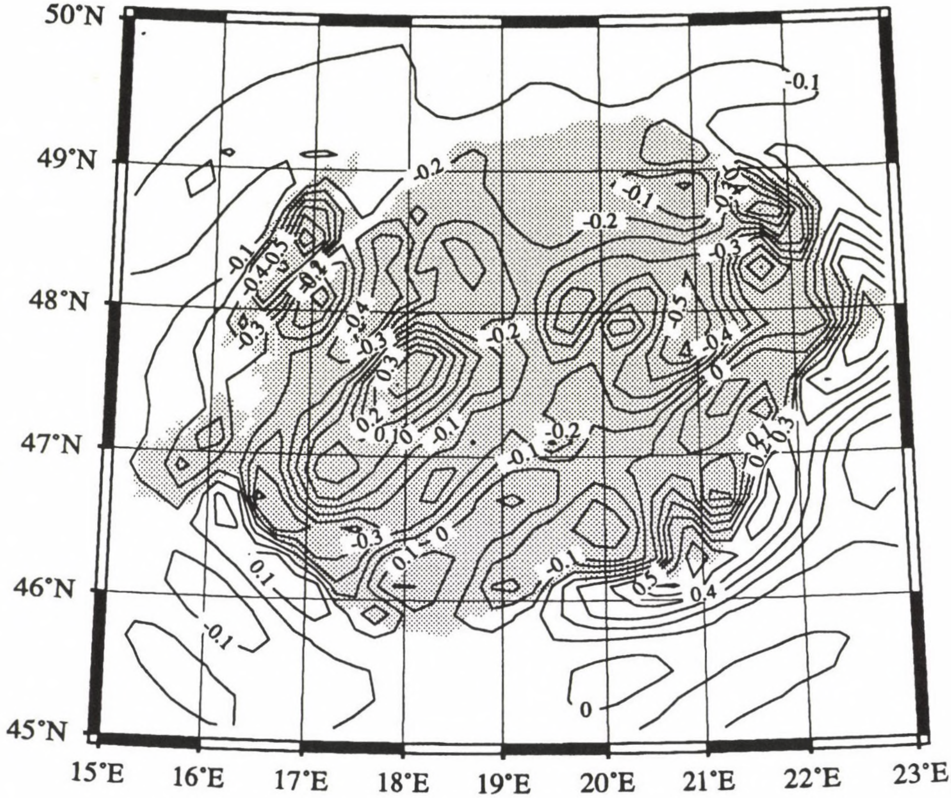


Fig. 8. Undulation differences between prism integration and spherical harmonic expansion solutions. Contour interval is 0.1 m. Shaded area represents the extension of the density model

division implied by Eq. (11) the 2D spectrum $T(f)$ was transformed to radial (1-D) average spectrum. The normalized radial transfer functions for $l_{\max} = 72$, 120, 180 and 360 are shown in Figs 9a, 9b, 9c and 9d, respectively. The basic features of low-pass filters can be seen clearly in the range of wavelengths > 60 km. Therefore the corresponding digital filters were attempted to be figured out by computing the normalized power spectrum of their impulse response functions which were obtained by using GRDFFT utility of GMT package. Between $\lambda_{\text{cut-off}}$ and λ_{pass} parameters cosine tapering was used to reproduce the smooth transition band of the transfer functions. The results are shown on the same figures for comparison. From the parameters of the digital filters (Table II), one can see that the transition bandwidth of the spherical harmonic "filter" inversely depends on the maximum degree and order $l_{\max}$ of the expansion. That is, it correlates well with the resolution of the expansion. The extraordinary width of the transition bands of $l_{\max} = 72$ and 120 is a reflection of the fact that the resolution is so low that the corresponding cut-off wavelength almost reaches the horizontal extension of the disturbing masses. However, in any situation the transition band is somewhat too wide compared to

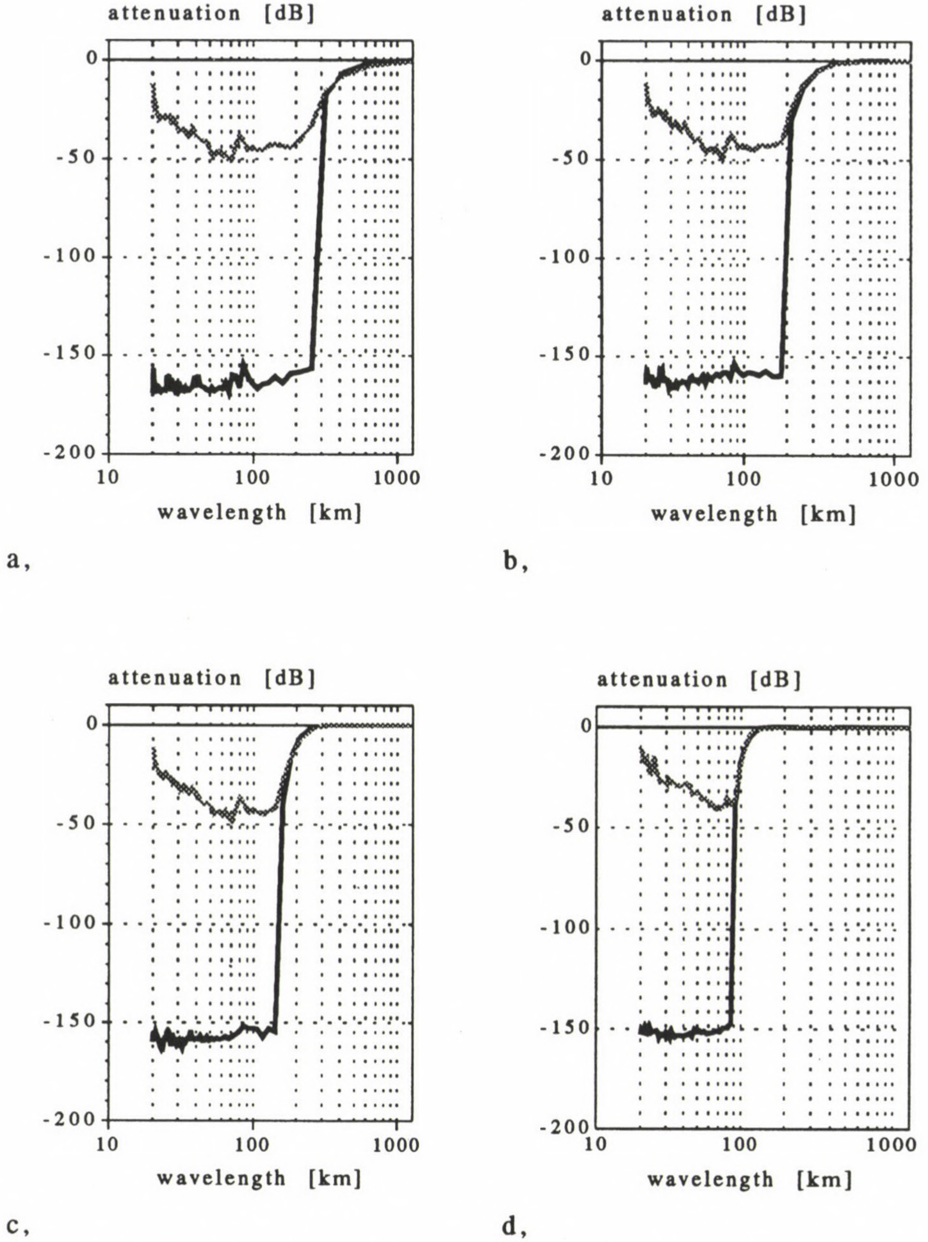


Fig. 9. Normalized radial transfer functions of different degree spherical harmonic expansions of the gravitational potential (light gray lines) and of the corresponding digital low-pass filters (black lines). a) $l_{max} = 72$, b) $l_{max} = 120$, c) $l_{max} = 180$, d) $l_{max} = 360$

Table II. Transition bands of the spherical harmonic filters and their corresponding digital filters

l_{max}	Spherical harmonic exp. $\lambda_{cut-off}$	Digital filter $\lambda_{pass} - \lambda_{cut-off}$	Transition bandwith	10 % of $\lambda_{cut-off}$
72	550 km	1000–270 km	730 km	55 km
120	330 km	600–200 km	400 km	33 km
180	220 km	300–160 km	140 km	22 km
360	110 km	150– 90 km	60 km	11 km

the common digital filters where, as a rule of thumb, the transition bandwidth is usually determined as 10 % of the cut-off wavelength. Table II lists the transition bandwidths and the corresponding values determined by the "10 % rule".

In Fig. 10 the transfer function of a low-pass filter with $\lambda_{cut-off} = 190$ km and $\lambda_{pass} = 210$ km is shown. It demonstrates that the "10 % rule" gives very good characteristics for sharp cut-off and high attenuation around 200 km wavelength which is not observable in the spherical harmonic filtering.

Another interesting phenomenon can be observed in Figs 9a–9d. The attenuation gradually decreases (in absolute sense!) below the wavelength of 50–60 km, which is not a characteristic of a common filter. In order to find the reason of this

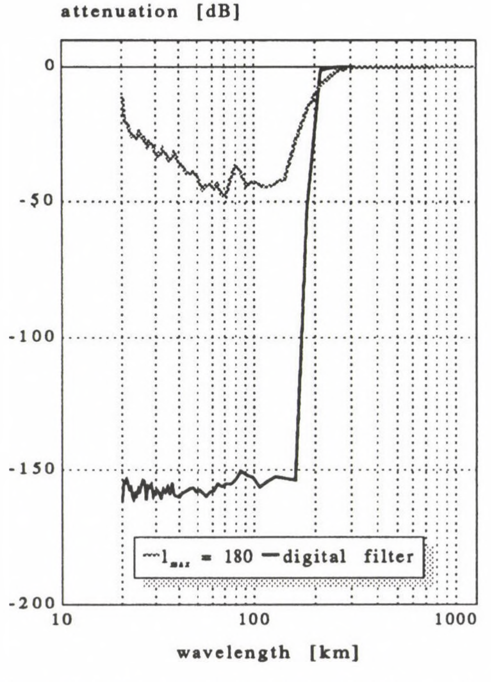


Fig. 10. Comparison of the transfer characteristics of spherical harmonic expansion with l_{max} = 180 and of a digital low-pass filter with λ_{pass} = 210 km and $\lambda_{cut-off}$ = 190 km

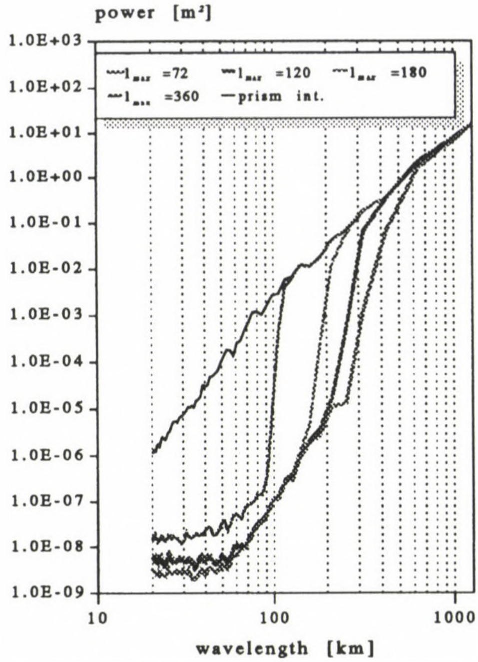


Fig. 11. Radial PSD-s of geoid undulations computed by different degree spherical harmonic expansions and rectangular prism integration

phenomenon, the radial power spectral densities (PSD-s) of all the geoid undulations obtained by forward modelling were calculated (see Fig. 11). It can be seen that the unusual decreasing of the attenuation of the transfer functions is caused by the "flat tail" starting at 50–60 km wavelength of the spectra of the undulations computed by the spherical harmonic expansion. First it was interpreted as systematic noise caused by the process of spectral analysis. However, when a digital low-pass filter with $\lambda_{\text{cut-off}} = 40$ km and $\lambda_{\text{pass}} = 60$ km was used on the spherical data set with $l_{\max} = 120$, the decrease of attenuation disappeared (Fig. 12), which implied that the noise existed *in the data*. It is assumed that this noise is always produced by the modelling computations and it indicates the highest resolution capability which can be attained from a given density model by the method applied since the same flat tail phenomenon was also observed in the PSD of the undulations obtained from prism integration carried out on a 1×1 km grid (Fig. 13). The possibility of a numerical accuracy problem was excluded because the flat tail of the spectrum did not change when the last (5^{th}, 4^{th} and 3^{rd} decimals of meter unit) digits of the undulation values were consequtively truncated and the PSD was reprocessed after each truncation.

Recalling that the range of wavelengths below $\lambda_{\text{cut-off}} \cong R\frac{2\pi}{l_{\max}}$ is the "null" space of the spherical harmonics therefore any spectral power with wavelength less than $\lambda_{\text{cut-off}}$ is irrelevant. However, from Fig. 12, one can see that there is still considerable power in the range between 50 km and the corresponding theoretical

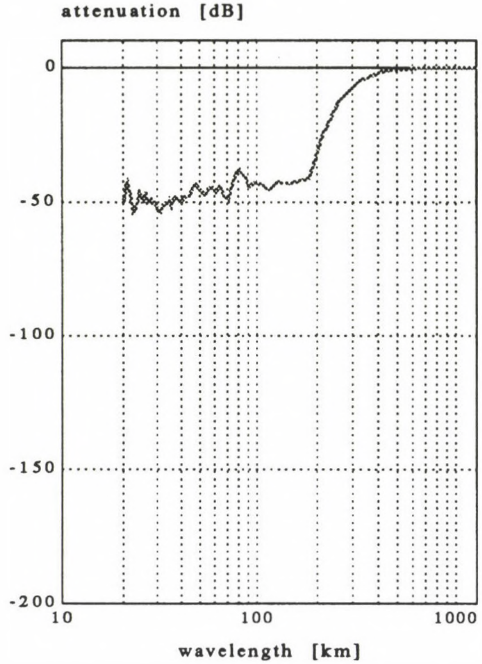


Fig. 12. Radial transfer function of $l_{max} = 120$ spherical harmonic expansion after digital low-pass filtering of the undulations obtained by it

cut-off wavelengths which gives signal contribution with amplitude about mm to cm, whereas because of the wide transition band the power above the theoretical $\lambda_{cut-off}$ is tendentiously smaller in its neighbourhood than it should be. In other words; in one hand the power is "too early" attenuated by the spherical harmonic low-pass filter and on the other hand the attenuation below $\lambda_{cut-off}$ is not sufficient. It supports the idea of "inverse aliasing" (Hipkin 1995) — a portion of the long wavelength power aliases to the shorter, theoretically forbidden wavelengths — which may be the result of the non-exact one to one correspondence between spherical and Fourier harmonics (Vermeer personal commu.). The tendency of the signal level to increase in the wavelength range below $\lambda_{cut-off}$ when l_{max} increases may also support this idea.

5. Conclusions

The forward gravitational problem can be solved by means of spherical harmonic expansions for both global and local situations. The accuracy of the expansion with a fixed l_{max} depends on the horizontal extension of the mass body and the smoothness of its bounding surface. Using spherical harmonics up to degree and order 360, the gravitational potential of masses larger than 6 degrees in size with a simple and smooth bounding surface can be modelled accurately. When it is used in situations where the mass model has rough boundaries, i.e. when it contains strong

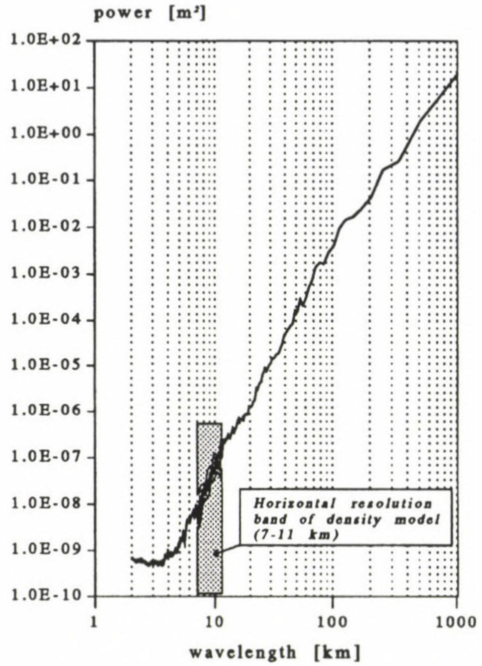


Fig. 13. Radial PSD of geoid undulations obtained by prism integration on a 1×1 km grid at 1024×1024 grid nodes

short wavelength features the truncation implied by $l_{\max}$ may cause significant signal loss at dm level.

The truncation of spherical harmonic results in low-pass filtering, but with wider transition bands than the ordinary digital filters. It indicates that the spectral characteristics of the spherical harmonics in forward local gravity modelling are different from that of the results obtained from rectangular prism integration. It means that, in case of combined application of the spherical harmonic expansion with the rectangular prism integration, attention must be paid to the different spectral characteristics resulting from the transformation from sphere to plane. For example, there is a possibility to compute the gravitational field of a global model of the topographic and isostatic masses, and then combine it with the potential generated by the local model of the disturbing masses in order to obtain a detailed solution. In this situation proper digital low-pass and high-pass filters should be used to separate global and local contributions, respectively, before their combination.

Because the efficiency becomes very low when one tries to expand in a "brute force" way the spherical harmonics up to high degrees and orders (≥ 360), the possibility of using it in local situation is limited at the moment. It has been tested that in DEC ALPHA Unix system, the expansion up to degree and order of 720 is possible.

Acknowledgements

The authors are grateful to Prof. Juhani Kakkuri, director of Finnish Geodetic Institute who provided excellent possibility for the common work. His encouragement and as well as the useful discussions with Prof. Martin Vermeer and Dr. Ruizhi Chen are appreciated in a great extent.

The investigations outlined in the paper were partly supported by Hungarian National Research Fund Project F 014284 and a research scholarship provided by Centre for International Mobility, Helsinki.

Appendix

The derivation of Eq. (8)

For a spherical surface layer with capsize Ψ and surface density μ, located at radial distance r', its potential at point P (Fig. A1) with radius r can be written as

$$V = G\mu \int_\sigma \frac{d\sigma}{l} = G\mu \int_\sigma \frac{r'^2 \sin\psi d\psi d\lambda}{l} \tag{A.1}$$

where

$$l^2 = r^2 + r'^2 - 2rr'\cos\psi \, . \tag{A.2}$$

By integrating over λ from 0 to 2π, Eq. (A.1) becomes

$$V = 2\pi G\mu \int_0^\Psi d\psi \frac{r'^2 \sin\psi}{l} \, . \tag{A.3}$$

From Eq. (A.2) by differentiation, one obtaines

$$l dl = rr' \sin\psi d\psi$$

i.e.

$$r' \sin\psi d\psi = \frac{l}{r} dl \, . \tag{A.4}$$

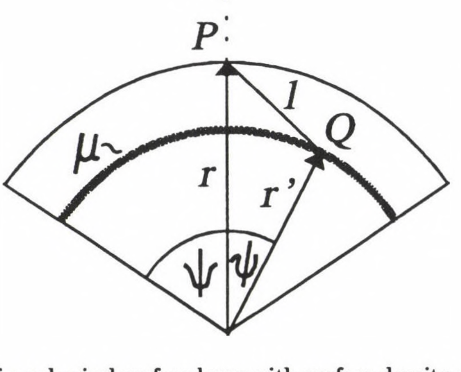


Fig. A1. Sketch figure of a spherical surface layer with surface density μ and capsize Ψ

Substituting Eq. (A.4) into Eq. (A.3), the potential of the surface layer at point P was obtained:

$$V = 2\pi G\mu \int_{l_1}^{l_2} dl \frac{r'}{r} = 2\pi G\mu \frac{r'}{r}(l_2 - l_1) =$$

$$= 2\pi G\mu \frac{r'}{r}\left[\sqrt{(r^2 + r'^2 - 2rr'\cos\psi)} - (r - r')\right] . \tag{A.5}$$

Now the potential of the spherical cap can be found by considering the surface density μ as a limit of volume density ϱ when dr' tends to be zero, that is,

$$p\mu = \varrho dr' . \tag{A.6}$$

Substituting Eq. (A.6) into Eq. (A.5) and integrating over r' from $R - D$ to R, where R is the upper limit of the integration, with $r' \leq R \leq r$, and D is the thickness of the cap, one obtains,

$$\begin{aligned} V_p &= 2\pi G\varrho \int_{R-D}^{R} dr' \tfrac{r'}{r}\left(\sqrt{r^2 + r'^2 - 2rr'\cos\psi} - r + r'\right) = \\ &= 2\pi G\varrho \tfrac{1}{r}\left\{\left[-\frac{1}{2}rr'^2 + \frac{1}{3}r'^3\right]_{R-D}^{R} + \frac{1}{3}\left[\sqrt{(r^2 + r'^2 - 2rr'\cos\psi)^3}\right]_{R-D}^{R}\right\} \\ &\quad +2\pi G\varrho \tfrac{1}{r}\frac{r\cos\psi}{2}\left[(r' - r\cos\psi)\sqrt{r^2 + r'^2 - 2rr'\cos\psi}\right]_{R-D}^{R} \\ &\quad +2\pi G\varrho \tfrac{1}{r}\frac{r\cos\Psi}{2}r^2\sin^2\psi \cdot \\ &\quad \cdot \ln\left[(r' - r\cos\psi) + \sqrt{r^2 + r'^2 - 2rr'\cos\psi}\right]_{R-D}^{R} . \end{aligned} \tag{A.7}$$

The potential of a sphere at point P can be found by Eq. (A.7) as a special case. Let $\psi = \pi$, $R - D = 0$, then

$$\begin{aligned} V_p &= 2\pi G\varrho \tfrac{1}{r}\left\{\left[-\frac{1}{2}rR^2 + \frac{1}{3}R^3\right] + \frac{1}{3}(R + r)^3 - \tfrac{r}{2}(R + r)^2 - \frac{r^3}{3} + \frac{r^3}{2}\right\} = \\ &= 2\pi G\varrho \tfrac{1}{r}\left(\frac{2}{3}R^3\right) = \tfrac{4}{3}\pi R^3 \varrho G \tfrac{1}{r} = G\frac{M}{r} \end{aligned} \tag{A.8}$$

with M the total mass of the sphere.

References

Forsberg R 1984: A study of terrain reduction, density anomalies and geophysical inversion methods in gravity field modelling. Report 355, Dept. of Geodetic Science and Surveying, Ohio State University, Columbus, Ohio

Forsberg R 1985: *Bull. Geod.*, 59, 342–360.

Engels J, Grafarend E W, Sorcik P 1995: The gravitational field of topographic-isostatic masses and the hypothesis of mass condensation. Part I and II. Technical Reports, Department of Geodesy, University of Stuttgart

Hipkin R G 1995: In: Gravity and Geoid, H Sünkel and I Marson eds, IAG Symp. 113, Springer, 529–538.

Kalmár J, Papp G, Szabó T 1995: *Computers and Geosciences*, 21, 245–257.

Middel B 1992: Computations of the gravitational potential of topographic-isostatic masses. Technical Report, No. 17, Institute of Geodesy, University of Stuttgart

Nagy D 1980: *Acta Geod. Geoph. Mont. Hung.*, 15, 17–26.

Nagy D 1988: *Acta Geod. Geoph. Mont. Hung.*, 23, 449–459.

Papp G, Kalmár J 1995: Toward the physical interpretation of the geoid in the Pannonian basin using 3-D model of the lithosphere. IGeS Bulletin, International Geoid Service at D.I.I.A.R., Politecnico di Milano (in press)

Rummel R, Rapp R, Sünkel H, Tscherning C C 1988: Comparisons of global topographic/isostatic models to the earth's observed gravity field. Report 388, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio

Sideris M G 1994: In: Lecture notes of the International School for the Determination and Use of the Geoid, International Geoid Service at D.I.I.A.R., Politecnico di Milano, 165–224.

Vermeer M 1992: In: Geodesy and Geophysics, J Kakkuri ed., Lecture Notes for NKG-Autumn School, Publications of the Finnish Geodetic Institute. No. 115, 183–200.

Vermeer M, Forsberg R 1992: *Manuscripta Geodaetica*, 17, 215–226.

Wessel P, Smith W H F 1991: *EOS Trans. AGU*, 72, 441, 445–446.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 67–78 (1996)

GROUND TILT INDUCED BY PUMPING — PRELIMINARY RESULTS FROM THE NAGYCENK TEST SITE, HUNGARY

H-J KÜMPEL[1], P VARGA[2], K LEHMANN[1], GY MENTES[2]

[Manuscript received January 22, 1996]

In-situ tilt measurements were carried out at Nagycenk near Sopron, Hungary to study ground deformation induced by ground water flow. A deformation signal, although being rather small, is expected from theoretical considerations: pore pressure gradients leading to ground water flow also deform the rock (or soil) matrix because of resistance and friction in the pore space. At Nagycenk, the pore pressure gradients are produced by pumping from a well 78 m deep. Tilt signals recorded in two shallow boreholes at a distance of about 7 m from the well are clearly related to pumping cycles. Assuming poroelasticity as the adequate rheology to describe the dominating physical process, the findings can be used to estimate petrohydraulic in-situ parameters not attainable through other methods. However, simple analytical expressions for spherical pressure disturbances in the homogeneous full-space produce tilt signals with opposite sign than observed here. This might be due to the vicinity of the stress-free surface. Tilt signals from two other locations at the test site that will be obtained soon will be used to verify this hypothesis.

Keywords: borehole; ground deformation; ground tilt; pore fluid; pore pressure; tiltmeter

Introduction

The understanding of coupled hydrological processes has been recognized to have increasing importance for many pragmatic applications among earth scientists and engineers over the past few years (Torgersen 1994). Parameter estimates are crucial for solving hydrogeologic problems but are often insufficiently known, e.g. because of the scale problem and as the nature of mechanical coupling is not yet completely understood. Petrohydraulic parameters which macro-scopically describe the ease of pore fluid flow through the rock (or soil) matrix and the mechanical response of the matrix to changing stress and/or pore pressure are of particular interest. Various laboratory tests exist to determine these parameters from rock specimens. Yet, since they have been removed from their original embeddings and are of limited size, samples disclose the rock's in-situ behaviour only partly.

A common method to evaluate some of the petrohydraulic parameters under in-situ conditions are pumping or slug tests, or their inverse procedures, injection and bail tests. They usually reveal values for the hydraulic conductivity — or the transmissivity — and a rough estimate of the storativity or storage compressibility.

[1]Geological Institute, Section Applied Geophysics, University of Bonn

[2]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

In recent years, tilt measurements have occasionally been applied to provide further insight into the hydraulic in-situ behaviour of saturated formations (Kümpel, 1989, Weise 1992, Maier and Mayr 1995). Usually, a pore pressure anomaly is stimulated and two tilt components of the resulting deformation field are observed, so that full coupling of mechano-hydraulic effects can be investigated. Since the tilt signal is small, sensors have to be of μrad resolution or better.

This paper deals with pump induced deformation in the small town of Nagycenk, 15 km east of Sopron, Hungary. The water is pumped for public supply from aquifers at a depth between 54 and 74 m. Tilt signals were recorded in two shallow boreholes at a distance of roughly 7 m and with different orientations with respect to the well. Here, we present these data and briefly outline the theoretical background for signal interpretation in the full-space case. The corresponding analytical solution predicts, however, the tilt signals to have opposite sign than that observed here. Two possible reasons will be discussed. A more complete paper is planned after tilt signals from two other locations have been recorded.

Data

Pumped wells are suitable objects for studying in-situ matrix deformation induced by pore pressure gradients, because one has control over the most influencing factors: the geometrical configuration, the energy applied to the ground through pumping, and, in general, the local geology. The well at Nagycenk was drilled in 1991. Details relevant for this investigation are listed in Table I.

Four tiltmeter boreholes were built in the vicinity of the Nagycenk well. A sketch is shown in Fig. 1. More data are listed in Table II. The annular spaces outside the PVC casings were filled with concrete from lower to top, sealing also the open bottom parts of the tubes. The open upper sections of the casings are closed by prefabricated concrete parts used in canalization works. Boreholes No. 1 and 2 end about 1.5 m below the groundwater table of the uppermost, unconfined aquifer, boreholes No. 3 and 4 end in the unsaturated zone above it. The lower sections of the former two boreholes close to a rather compact gravel formation, those of the latter two reside in a less compact sandy loam layer.

We have two tiltmeters for the investigations at Nagycenk, Table III gives technical details. A tiltmeter is installed pouring a layer of sand 0.1 m thick inside the PVC casing, tampering the sand, lowering the tiltmeter to the bottom, and fixing it in nearly vertical position by adding and tampering more sand in the free annular space. A strong mechanical coupling to the undisturbed formation is essential for obtaining meaningful ground tilt signals. The azimuth of the instrument installed is determined from the orientation of the tiltmeter's handle at its top, aligned parallel to the y-direction. This can be done to an accuracy of a few degrees. The signal and power cable is connected to a switch box at the surface that allows high or low gain settings and weak or strong low-pass filtering. The instrument can be operated in single ended and differential mode. We operated the tiltmeters in high gain, differential mode and weak damping. The switch box is connected to a power supply (batteries) and a suitable recorder.

Table I. Details of the well at Nagycenk, Hungary. Only the second and third aquifers from top are used for water supply

Parameter	Quantity	Remarks
Depth	78.5 m	cased well
Diameter	200 mm	inner casing, PVC, down to bottom
	352 mm	outer casing, steel, down to 53 m
Screen	0.7 mm	width of aquifer filter slits
	56.2 – 60.7 m	depth of upper filter section (2nd aquifer)
	65.7 – 74.0 m	depth of lower filter section (3rd aquifer)
Pump	0.009 – 0.011 m^3/s	power, active on demand, type Grundfoss SP 35-11,
	($\leq$ 40 m^3/s)	automatically activated 2 to 3 times per day
Geology	0.0 – 3.6 m	loam, loam with sand
	3.6 – 6.0 m	mixed size gravel
	6.0 – 26.0 m	loam with sand, sandy loam, layered
	26.0 – 32.4 m	median size gravel, with lime (1st aquifer)
	32.4 – 54.8 m	loam with sand, sandy loam, layered
	54.8 – 61.8 m	coarse gravel, quartz, with lime (2nd aquifer)
	61.8 – 64.4 m	sandy loam, loam
	64.4 – 74.2 m	median size gravel, with lime (3rd aquifer)
	74.2 – 100.0 m	loam with sand, loam, sandy loam, layered
		(open hole refilled up to 78 m)
Hydraulic head	3.3 m	mean value below surface
Productivity	17.5 m at 540 l/min	level from surface
	22.2 m at 720 l/min	
	26.9 m at 900 l/min	

Table II. Tiltmeter boreholes at Nagycenk Test site

Borehole	No. 1	No. 2	No. 3	No. 4
Installation depth (m)	4.3	4.3	2.8	3.8
Azimuth to well (N to E)	111	212	354	6
Radial distance r to well (m)	7.7	6.9	22.9	64
Depth z to centre of upper well sreen (m)	54.2	54.2	56.3	55.3
Distance R to centre of upper well sc. (m)	54.7	54.6	60.8	84.6
Diameter of PVC casing (mm)	300	300	150	150
Shortest distance to concrete path (m)	2.0	12.2	16.2	16
Azimuth to nearest section of concrete path (N to E)	101	98	101	101
Month when completed	07/94	07/94	03/95	10/95

So far, we have tilt data from boreholes No. 1 and 2, obtained with instruments No. 1652 and 1655. Figure 2 displays the true geometric configuration for borehole No. 1. Initially, from August 1994, the temperature at downhole depths was recorded together with the tilt signals. Its value was 13° ± 1°C and showed slight seasonal changes. Due to limited recorder capacity and for better correlation with more influential effects, later since July 1995, we recorded the switch times of the

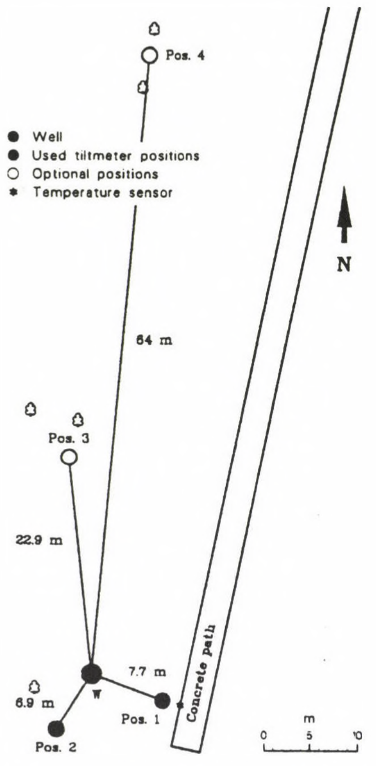


Fig. 1. Plan view of test site Nagycenk, Hungary. Topography is negligible over depicted area and in surroundings. Except for the concrete path, the surface is covered by grass. See Tables I and II for details on the pumped well and the tiltmeter boreholes at positions Nos. 1 to 4, and note the place for temperature measurements inside the concrete

pump and temperature variations in the concrete path instead (Fig. 1), the latter with a sensor fixed inside a 8 mm in diameter hole in the concrete, six centimeters below the surface. Figures 3 and 4 show typical data sets from a week in August and in October 1995. There is a clear correlation between pumping cycles and tilt variations in the X- component of position No.1 (X1) and the Y-component of position No. 2 (Y2).

It is also evident from Fig. 3 that a correlation exists between temperature variations in the concrete and tilt changes in X1, Y1, and, less pronounced, in X2. On sunny summer days, the concrete attains its highest temperature during the afternoon, whereas on cloudy days, temperature variations are moderate throughout the day. Obviously, the tilt variations reflect temperature induced ground deformation emanating from the concrete. This confirms the sensitivity of the tiltmeters and a good coupling to the ground. Such effects are typical at shallow depths, but we will not make further use of this here.

Table III. Technical details of tiltmeters

Manufacturer	Applied Geomechanics Inc., Santa Cruz, California
Model	Borehole tiltmeter, Type 722
Type of sensor	Fluid level, with electronic pick up of gas bubble position
No. of components	Two, at nominal 90 azimuth separation, named x- and y-component
Sensitivity (nominal)	0.1 μrad in High Gain Mode, 1 μrad in Low Gain Mode
Calibrated range	800 μrad in High Gain Mode, 2000 μrad in Low Gain Mode
Length of tiltmeter body	0.85 m
Diameter of tiltmeter body	54 mm
Power supply	−12 V, 0 V, +12 V
Signal output	±8 V analog
In-built temperature senor	$\leq$ 0.1 °C resolution

Pump induced tilt signals

Ground tilts due to pumping have been observed several times before (Kümpel 1988, 1989, Weise 1992, Maier and Mayr 1995). The tilt signal, that follows a characteristic, fin like shape described in detail by Kümpel (1989), is clearly visible in Figs. 3 and 4.

The theory of linear poroelasticity appears to be adequate to explain the physical process of hydraulically induced deformation. Fundamental contributions originate from Biot (1941) and Rice and Cleary (1976). Poroelasticity is a rheology very similar to thermoelasticity, where pore pressure takes the role of temperature. However, unlike in thermoelastic problems, the full coupling between deformation and pore pressure can not be neglected. The approach is macroscopic, that is pores and pore structures are not resolved in the region considered. Complete saturation with pore fluid is essential for linearity in poroelasticity. The behaviour of a porous medium differs for drained and undrained conditions, i.e. whether fluid flow through parts of the medium's surface is allowed or not. Four parameters describe the elastic stress-/pressure strain behaviour of the medium's matrix, a fifth parameter reflects the ease of fluid flow (Kümpel 1991, Wang 1993).

In the distant field of a pore pressure disturbance in a homogeneous medium, the poroelastic deformation from a spherical pressure source can be obtained through analytical solutions. Numerical techniques like the finite element method have been used to compute the deformation field for more complex situations (Kümpel 1989). The solutions also allow insight into the deflection of a borehole (Fig. 5), when we assume that the borehole completely follows the deformation field in its vicinty. By comparison with the tilt signal observed in the borehole, values for rock matrix and fluid flow parameters can be established so that theory and field data agree. In such a way, in-situ poroelastic parameters not accessible by other methods are obtained.

One may choose the shear modulus G, the Skempton ratio B (denoting the change of pore pressure per unit change of confining pressure for undrained conditions), the Poisson ratios v and v_u for drained and undrained conditions as the parameters of the rock matrix, and the intrinsic permeability k as the hydrologic

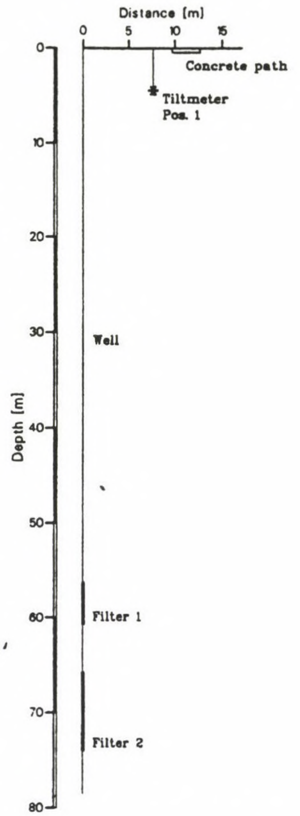


Fig. 2. True scale geometric configuration of tiltmeter position No.1, pumped well and concrete path. Groundwater is drawn from aquifers through filters 1 and 2

parameter. Note that $0 \leq v \leq v_u \leq 0.5$. Analytical solutions for the homogeneous full-space case reveal that in the surroundings of a well, which is initially in hydrologic equilibrium and in which pumping with constant rate Q starts at time $t = 0$, the shear strain φ is

$$\varphi = cK\vec{B}Q/D \tag{1}$$

(Kümpel 1989). Herein, c is a dimensionless function of the radial distance r from the axis of the well, the vertical distance z from the centre of the well screen, time t since the onset of continuous pumping, and the hydraulic diffusivity D, namely

$$c(r, z, t; D) = (3J + E^-)/(16\pi) \tag{2}$$

with
$E- = 1 - \mathrm{erf}(\Re)$
$J = [\mathrm{erf}(\Re)/\Re - 2\exp(-\Re^2)\sqrt{\pi}]/[2\Re]$
$\Re = R/\sqrt{\Re 4Dt}$
erf(.) is the error-function, and $R = \sqrt{r^2 + z^2}$ is the distance between the tilt sensor

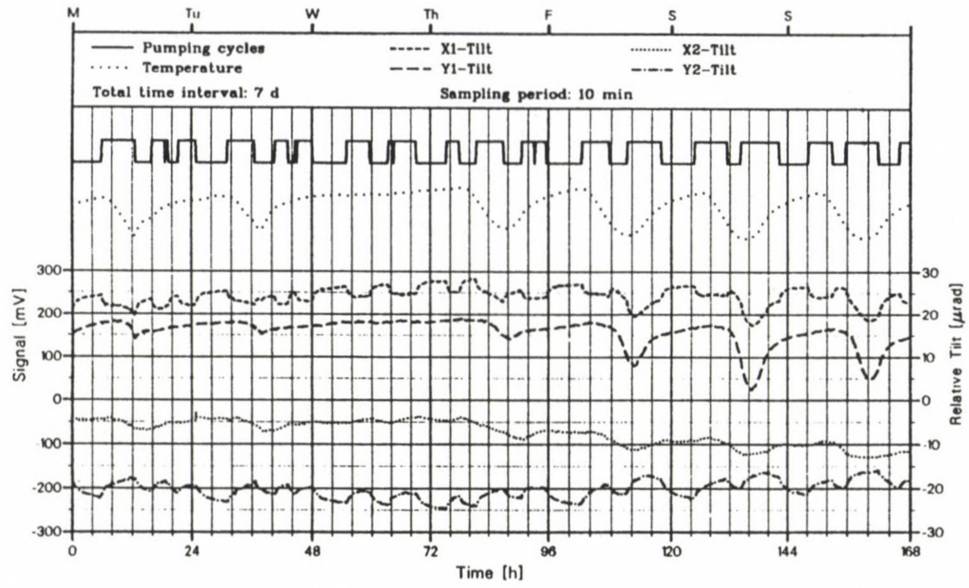


Fig. 3. Data from week 32 in 1995. Time is local time, 1 hour ahead of UTC. Synchronized, ten minutes sampling interval for all signals. Pumping signal high means pump is active. Temperature (unscaled) is that of concrete path and is inverted for better optical correlation with tilt effect, i.e. low signal corresponds to higher temperature and vice versa. X-tilt is roughly in EW, Y-tilt in NS. Positive tilt signal means relative shift of tiltmeter top towards E (for X) or N (for Y)

and the well's screen (filter). c increases steadily with t, approaching asymptotically its *maximum* value 0.017... . For the assessment of rock parameters from the maximum shear deformation at a given distance R, c can be replaced by 0.017 if $\Re \leq 0.3$ or $t \geq 2.78R^2/D$ (Kümpel 1989). It is convenient to use the estimated t_{90}-time of the tilt signal, defined as the time when 90% of the pump induced tilt amplitude is built up. One may then use the relation $t_{90} = R^2/4D$ together with $c \sim 0.017$ where the error is less than 10%. In fact, the final tilt amplitude will hardly be obtained with much higher acccuracy, because it is approached asymptotically and because of interference with non-pump-induced drift signals. The other parameters in Eq.(1) are
$K = 2rz/R^3$, a geometric coefficient, and
$\vec{B} = B/3 \cdot (1+v_u)/(1-v_u)$, obviously, $B/3 \leq \vec{B} \leq B$.

The hydraulic diffusivity D can be expressed as

$$D = 2\frac{(1-\nu)(1-\nu_u)}{(\nu_u-\nu)}KG\vec{B}^2 \tag{3}$$

(Rice and Cleary 1976), where $k = k/\eta$ is the Darcy conductivity with η denoting the dynamic viscosity of the pore fluid. The tilt signal in the surroundings of the well is a superposition of the shear strain φ as given in Eq.(1) and the rotation ψ of the tiltmeter's environs. To a first approximation one may assume that φ is the

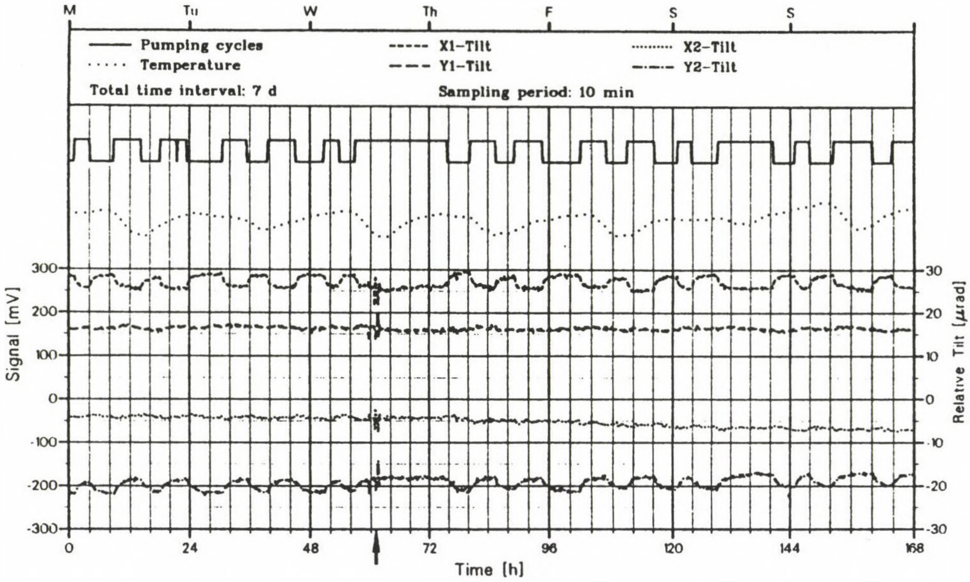


Fig. 4. Data from week 42 in 1995. Display mode as in Fig. 3. Tilt spikes around hour 61, at noon on October 18th, originate from an earthquake at Ryuku islands, 27.92 N, 130.34 E at 10.50 UTC, Mw= 6.9 (data from USGS)

observed tilt signal. The sense of the tilt signal for this case is reflected in Fig. 5. Rotation of ground regions become, however, more important close to free surfaces where shear deformation is reduced. The relevance of contributions from rotation will deserve further attention in a separate paper.

If the rock facies of the test area is known, some of the above parameters can be assessed fairly well, others can not. In our case, the permeability can be roughly estimated from the draw down tests made in the Nagycenk well for productivity control (Table I). Accordingly, using a formula of Logan for confined aquifers (cited in Krusemann and de Ridder 1973), k is around 6 Darcy ($\sim 6 \cdot 10^{-12}$ m^2).

The adopted global parameters for the test site are listed in Table IV. η is the viscosity of pure water for temperatures from 8 to 12°C, using a formula of Hardington and Cottingen (see Weast 1974). B takes always values between 0 and 1, but likely falls in the range 0.2 to 0.8 for saturated sediments. The estimated ranges for v and v_u can be argued from the fact that for increased confining pressure water can easily leave unconsolidated sediments under drained conditions, generally leading to some significant compaction of the matrix, whereas in undrained conditions the low compressibility of the pore fluid impedes a similar volume reduction.

Taking 2.5 hours as the t_{90}-time for tiltmeter positions No.1 and 2 (see Figs 3, 4), one would obtain roughly 0.1 m^2/s for the hydraulic diffusivity D. A value of 0.4 for the Skempton ratio B and order of 500 MPa for the shear modulus G would, for the full space solution, agree with pump induced tilt amplitudes of 3 μrad, as recorded here (see complete parameter set in Table IV). Similar values of B and G

Table IV. Pump rate Q and plausible rock/fluid parameters at Nagycenk test site assuming linear poroelasticity and full-space situation. The ranges of the deduced parameter values are conservative. They were calculated adopting combinations which yield the lowest and highest possible values, respectively. This explains the large range obtained for G

Pump rate	Range	Adopted value	Remarks
Pump rate Q	$30 - 34\ \mathrm{m^3}/h$	$9 \cdot 10^{-3}\ \mathrm{m^3/s}$	from observation
Fluid viscosity η	$12 - 14 \cdot 10^{-3}$ Pa s	$13 \cdot 10^{-3}$ Pa s	for pure water from 8 to 12 °C
Intrins. permeability k	$3 - 10 \cdot 10^{-12}\ \mathrm{m^2}$	$6 \cdot 10^{-12}\ \mathrm{m^2}$	from draw down tests
Hydraul. diffusivity D	$0.080 - 0.085\ \mathrm{m^2/s}$	$0.083\ \mathrm{m^2/s}$	$D \sim R^2/4t_{90}$
Skempton ratio B	0.29 – 0.39	0.34	Eq. 1
Poisson ratio v, drained	0.05 – 0.15	0.10	estimated
Poisson ratio v_u, undr.	0.40 – 0.48	0.45	estimated
$\vec{B} = B/3(1+v_u)/(1-v_u)$	0.31 – 0.50	0.39	
Shear modulus G	0.14 – 2.6 GPa	0.46 GPa	Eq. 3

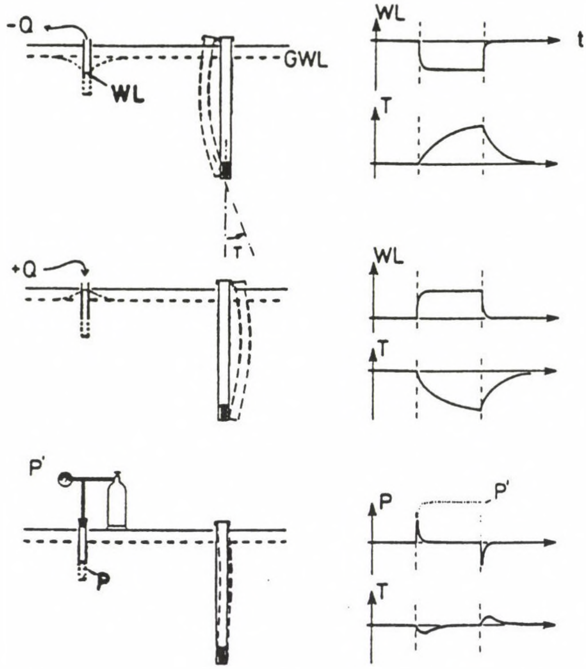


Fig. 5. Schematics of borehole deformation in the vicinity of a well with groundwater pumped (top), injected (center), or displaced through pressurized air (bottom). GWL=Groundwater level, WL=Well level, Q=Injection rate, P=Pressure at well screen, P'=Gas pressure in sealed well, T=Tilt amplitude, t=time (from Kümpel, 1988). Note that at Nagycenk, the well screens are at (much) greater depth than the tiltmeters

were found to explain the amplitudes of pump induced tilt signals in unconsolidated sediments observed elsewhere (Kümpel 1989). At Nagycenk, however, the *sign* of the tilt signals is at both positions 1 and 2 opposite than predicted by the full-space solution. Consequently, estimates of parameter values based on the adopted assumptions turn out to be invalid here, or, at least one of them must be a flaw.

For instance the axis of the pumped well could, grossly deviate from the vertical, so that the filter locations projected to the surface fall east of tiltmeter position 1 and south of position 2. Since such deviation would have to be at least 11°, an unusually high value, and nothing similar is known from the drilling record, we can rule out this option. A more likely reason for the unexpected tilt sign is that due to the shallow installation depths the tiltmeters "see" a cyclic subsidence of the surface whenever the pump is activated. In this case, the tiltmeters' tops move closer to the well axis than their lower ends, following an inward rotation of the near surface layers towards the well. This effect then dominates over the shear strain that is solely effective in the homogeneous full-space situation, when — due to higher pore pressure gradients — the tiltmeters' lower ends are attracted more strongly by the pore pressure deficits around the well filters than their tops.

Deeper insight and more accurate estimates of parameter values are expected from tilt records that will be obtained at positions No. 3 and 4. Through the clearly different geometric configurations we trace a radial profile of the tilt amplitudes at shallow depths. In addition, a more accurate estimate of the size of the induced tilt signals will be received from stacking, and deviations from strictly radial orientations of the maximum amplitudes will be analysed as hints for anisotropy in pore fluid flow.

Conclusions

We carried out tilt measurements to improve the understanding of coupled mechano-hydrogeologic processes, and to demonstrate the capabilities of tiltmeters in monitoring the deformation field around pumped wells. Like Maier and Mayr (1995) we have used tiltmeters of 'only' 0.1μrad resolution but analysed signals in the vicinity of a much deeper well. The observed, pump induced tilt signals are cleary above the noise level, but although their amplitudes are in agreement with realistic parameter values of the poroelastic sediments at the test site, they are of opposite sign than expected for the full-space situation. Possibly, this is caused by rotational effects dominating over shear deformation at the probed, shallow depths.

The results are preliminary, since data from two more tiltmeter locations will soon be available for verification and more accurate analyses. In particular, these data are expected to show how the tilt signal changes with increasing radial distance for the quasi half-space situation. We have already shown that tilt measurements can be used to study ground deformation induced by pore pressure gradients — for fundamental research aspects, for obtaining values of petrohydraulic in-situ parameters, and for specific hydrological applications where neglect of rock matrix strain is not justified.

Acknowledgements

This research was supported by grant of the German-Hungarian cooperation (project numbers in Germany WTZ X237.2 and in Hungary OMFB 76) through the International Office of Karlsruhe Research Center and the OMBF, Budapest, and by the Sonderforschungsbereich 350, German Research Foundation (DFG).

References

Biot M A 1941: *J. Appl. Phys.*, 12, 155–164.

Kruseman G P, de Ridder N A 1973: Untersuchung und Anwendung von Pumpversuchsdaten. Vlg. R Müller, Köln-Braunsfeld

Kümpel H-J 1988: In: Kurzberichte aus der Bauforschung, Fraunhofer Ges., Ber. Nr. 39, 145–146.

Kümpel H-J 1989: Verformungen in der Umgebung von Brunnen. Habil. Thesis, Faculty of Mathematics and Natural Sciences, University of Kiel

Kümpel H-J 1991: *Geophys. J. Int.*, 105, 783–799.

Maier R, Mayr S 1995: In: DGG-Mitteilungen, Sonderband, 3. DGG-Seminar Umweltgeophysik, Münster (in press)

Rice J R, Cleary ? 1976: *Rev. Geophys. Space Phys.*, 14, 227–241.

Torgersen Th 1994: *EOS Transactions*, 75, 73–78.
Wang H 1993: *PAGEOPH*, 141, 269-286.
Weast R C ed. 1974: Handbook of chemistry and physics. 55th Ed., CRC-Press Inc., Cleveland, Ohio
Weise A 1992: Neigungsmessungen in der Geodynamik — Ergebnisse von der 3-Komponenten-Station Metsähovi. Dissertation, University of Clausthal-Zellerfeld

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 79–89 (1996)

SOME CALIBRATION OF THE APPLIED GEOMECHANICS INC. BOREHOLE TILTMETER MODEL 722

GY MENTES[1], K LEHMANN[2], P VARGA[1], H-J KÜMPEL[2]

[Manuscript received June 11, 1996]

For several years we use Applied Geomechanics Inc. (AGI) borehole tiltmeters, Model 722, for measurements of ground deformation induced by ground water flow. These instruments record two components of tilt with electrolytic sensors of 0.1 μrad resolution, with their performance depending on parameters like scale factor, linearity, temperature behaviour and coupling to the ground. Different laboratory methods were applied for recalibration of the borehole tiltmeters. Calibration coefficients of four instruments were obtained using laser interferometers and mercury crapaudines and compared to those given by AGI. For full use of the resolution it appears necessary to check the calibration parameters in the suggested way and to correct the sensor accordingly.

Keywords: borehole tiltmeter; crapaudine calibration; ground deformation; laser interferometer

1. Introduction

Applied Geomechanics Inc. (AGI), Santa Cruz, California, offers the instrument series 700 which is widely applied to monitor deformation effects in a lot of applications, mostly for geotechnical but also for geophysical use. The Model 722 borehole tiltmeter we use for studying ground deformation induced by pore pressure gradients in the vicinity of pumped wells (Kümpel et al., 1996), is part of that series. This instrument is a dual-axis, low noise analog output tiltmeter with a nominal resolution of 0.1 μrad. It operates at low power consumption, has a high reliability under rugged field conditions and can be installed in shallow boreholes down to 8 m depth.

In various field studies carried out over several years, we have gained some experience with this tiltmeter type. As part of a bilateral Scientific and Technical Cooperation between Germany and Hungary, we started to check the instruments' accuracy and characteristics. This paper deals with the recalibration of the AGI borehole tiltmeter Model 722, carried out at our institutes at Sopron and Bonn. At both places, a Hewlett-Packard laser interferometer system was used for precise angle measurements using an absolute calibration method. In addition, a relative calibration method using crapaudines was applied at Sopron. We report here on the results of both methods and how they are combined to establish calibration curves for the tiltmeters. It will be shown that the recalibration can result in more accurate measurements.

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary
[2]Geological Institute, Section Applied Geophysics, University of Bonn

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

Table I. Main specifications of the AGI Model 722 borehole tiltmeter

Tilt sensing range:	± 800 μrad (HIGH GAIN), ± 2000 μrad (LOW GAIN).
Tilt resolution:	0.1 μrad (HIGH GAIN), 1 μrad (LOW GAIN).
Input voltage range:	+11 to +15 VDC, –11 to –15 VDC.
Output voltage range:	± 8 V both in HIGH and LOW GAIN (single-ended output) or ± 16 V (differential output).
Tilt scale factor:	about 0.1 μrad/mV (HIGH GAIN), about 1 μrad/mV (LOW GAIN). Both scale factors are valid in single-ended mode. The exact values differ slightly from tiltmeter to tiltmeter and are given in the calibration certificate by the manufacturer.
Filter time constant:	0.05 s (OFF position), 7.5 s (ON position).
Temperature scale factor:	0.1°C/mV.
Accuracy of temperature measurement:	± 0.75°C.
Range:	–40°C to +100 °C.

2. Technical data of the borehole tiltmeters

The Model 722 borehole tiltmeter consists of a cylindrical stainless steel body (85 cm long, 5.4 cm in diameter) containing two orthogonal electrolytic precision tilt sensors, amplifiers and signal filters for each of them and a temperature sensor. An external switch box, connected to the tiltmeter via a submersible steel-reinforced cable, hosts a panel for the tiltmeter's gain and filter switches as well as the power input and signal output connectors.

The signal amplifiers of the instrument have two or three settings: HIGH, MID –optional–, LOW GAIN. The gain factor between the settings is 10. The filters are 2-pole Butterworth low-pass devices with 40 dB/decade roll-off and a time constant settable to customer specifications. The tiltmeter can be used with FILTERED and UNFILTERED output. Furthermore, the sensor signal can be taken in asymmetrical (single-ended) or symmetrical (differential) mode, the latter optionally. Table I lists the most important specifications.

3. Installation

The tiltmeter Model 722 can be installed in a shallow borehole to monitor changes in tilt of a nearly vertical ground element over time. For a certain stability against environmental and meteorological effects on the one hand, and for handling reasons during installation and possibly readjustment of the tiltmeter's

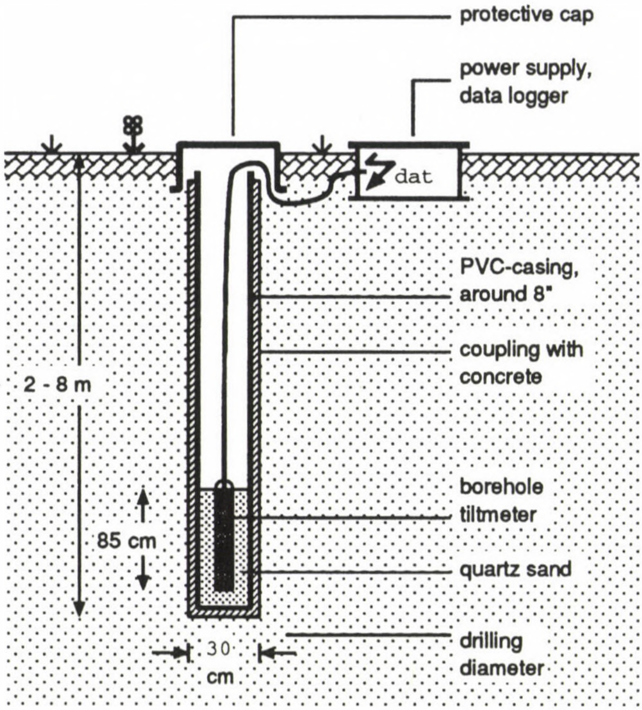


Fig. 1. Example for installation of the Model 722 borehole tiltmeter

nearly vertical alignment on the other (in case the signal is drifting out of the measuring range), it is advisable to place the instrument at depths between 2 and 8 m. The borehole has to be drilled with a diameter of about 30 cm, its casing will usually be a PVC pipe that is coupled to the surrounding formation by concrete (Fig. 1). The pipe's diameter should measure around 8″ to allow sufficient room for manipulating and levelling the tiltmeter with bars during installation. Inside the pipe the instrument is fixed through tampered quartz sand in order to obtain a strong coupling to the ground. In case of a shallow ground water table, the bottom end of the pipe should be sealed by a cap or slightly expanding concrete to prevent fluctuations of the water table inside the pipe that could affect the stiffness of the coupling.

The signal is transferred via cable to the switch box which can be placed in a trunk at the surface together with the power supply and a data logger.

4. Purpose of recalibration

Recalibration of the AGI Model 722 tiltmeters appears to be useful because of several reasons. These are, in order of relevance:

a) According to our experience, the Model 722 tiltmeters are stable instruments having a low drift. Since for our applications we are interested in measuring small deformations (Kümpel et al. 1996, Rebscher 1996), only a narrow part of the instruments' measuring range is used. A higher accuracy might in fact be achieved in a smaller range if we knew the transfer functions of the tiltmeters. The manufacturer gives only linear scale factors. These are based on linear regressions through calibrations at minimum 10 points of the measuring range.

b) For long-term measurements it is necessary to know the temporal behaviour of the instruments. Therefore, an internal calibration unit is included in some high precision tiltmeters (e.g. the Askania borehole tiltmeter; Graf 1964). For AGI tiltmeters, only recalibrations can inform about the long-term stability of the instrument's parameters (scale factor, linearity, sensitivity to temperature variations etc.).

c) The instruments were calibrated by the manufacturer at room temperature in the range 20 to 23 °C. We use them in 3 to 5 m-deep boreholes where the temperature is only 10 to 12 °C. Although AGI gives a linear relation which can generally be used to compensate for temperature effects on the scale factors and the zero points, the calibration certificate has no details about the temperature behaviour of the built-in sensors so that only approximate corrections are possible.

d) The tiltmeters are set in sand during installation. One aim of our investigations was also to learn whether this type of coupling has some influence on the accuracy of tilt measurements in boreholes, in other words to check the performance of the coupling between the instrument and the casing.

5. Calibration methods

We used an absolute and a relative calibration method for testing the tiltmeters. The absolute calibration was done with a laser interferometer system in angle reference mode. Mercury crapaudines, which were formerly applied to instruments resolving earth tides, were used for relative calibration.

5.1 Calibration by laser interferometer

This kind of calibration was made at both of our laboratories at Sopron and Bonn (Fig. 2). The tiltmeter to be calibrated is installed in an iron tube welded to a rigid bottom plate that is standing on three footscrews. The bottom plate is a rectangular triangle with the Sopron and a circular disc with the Bonn test equipment. Balls placed in ball sockets at the footscrews' tips, respectively ball-like, guarantee nearly a point support and low frictional movements of the screws. The tiltmeter can be fixed in the tube by sand or by screws. In the latter case, two sets of three screws, arranged at angular separations of 120 degrees, support the tiltmeter at the upper and lower ends in the tube. At Sopron, both tilt components can be tested at the same time, if the tilt axes are installed parallel to the perpendicular sides of the bottom plate, because the footscrews are placed at the ends of the

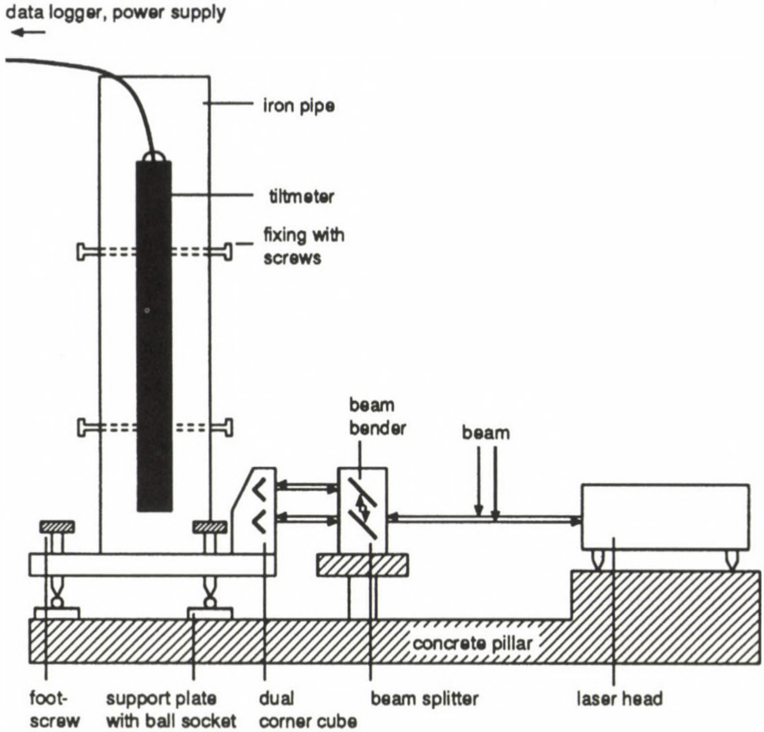


Fig. 2. Absolute calibration method using a HP laser measurement system

plate's hypotenuse. At Bonn, the footscrews form an equilateral triangle, and each tilt direction has to be tested individually. When the iron tube has to be arranged symmetrically to the laser beam axis, a well defined tilt signal can be applied by turning a footscrew on this axis. For measuring the tilt amplitude, a Hewlett-Packard (HP) laser interferometer system was used in angular measuring mode.

At Sopron, the measurements were carried out with a HP5528A Laser Measurement System (HP–5518A Laser Head, HP–55281 Angular Optics Kit), and at Bonn with a HP–5526A System (HP–5500C Laser Head, HP–5505A Laser Display, HP–10565B Remote Interferometer, HP–10558A Beam Bender, HP–10559A Reflector Mount); both have an angular resolution of 0.1 arc sec or 0.5 μrad, respectively. By comparing the tilt signal displayed through the interferometer system with the tiltmeters' voltage output for a series of tilt steps, the sensors' transfer functions can be obtained.

5.2 Calibration by crapaudines

The calibration method using crapaudines was formerly applied for testing highly sensitive horizontal pendula (Mentes 1986). A crapaudine is a cylindrical, stainless steel pressure box filled with mercury. The upper wall of the box is thinner than the others so it functions as a membrane. Increasing the pressure in the box de-

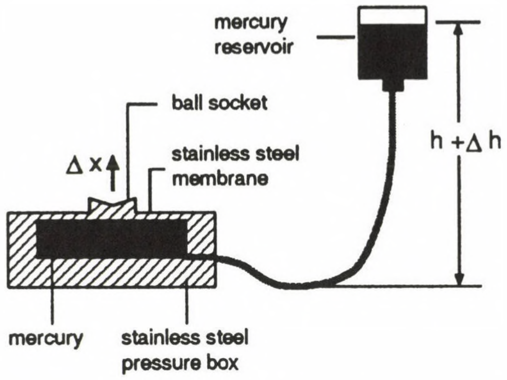


Fig. 3. Functional principle of a crapaudine. All parts are cylindrical

forms only the membrane, deformation of the other walls is negligible. The box is connected via a plastic tube to a mercury reservoir. If the height of the reservoir is increased by Δh, the membrane of the pressure box moves upwards by a distance Δx (Fig. 3): In good approximation, one has $\Delta x = k\,\Delta h$, where the proportionality factor k can be determined by means of a laser interferometer and is called calibration factor of the crapaudine. Since for our crapaudines k is rather small, Δh has to be large ($k = 0.00422$ μm/cm Hg for the crapaudine SG–5, $k = 0.00489$ μm/cm Hg for SG–6).

Calibration of the borehole tiltmeters was carried out as shown in Fig. 4. Two crapaudines were placed under footscrews of the tube's bottom plate. In several steps within the measuring range (30–40 points), the mercury reservoir was lifted to a maximum height difference Δh of 253 cm, and the output signals of the tiltmeter were recorded. The inclinations applied to the tube can be calculated from the crapaudines' scale factors, the distance between the fixed and the lifted footscrews, and Δh, yielding tilt amplitudes of 3.19 μrad in x and 3.81 μrad in y. By this method, scale factors could be determined continuously over the measuring range. An average scale factor valid for the entire range (as given by the manufacturer) can be calculated as an average of local, step-wise scale factors. The accuracy of this method is less than that of the interferometric one because of the limited accuracies of the crapaudines. The crapaudine method is nevertheless a way to check the linearity of the transfer function directly. It represents a supplementary method to gather the tiltmeters' characteristics with a different technique, i.e. with independent uncertainties. Although we made no use of it, this method is also suitable for obtaining the tiltmeters' frequency response by lifting the mercury reservoir periodically with different frequencies.

6. Results of the recalibration

In June 1995, the tiltmeter serial no. 1655 was calibrated at Sopron. The measurements were made in the institute's cellar floor test laboratory at a temperature $T = 21°$C using a HP laser interferometer. The tiltmeter was fixed in the tube

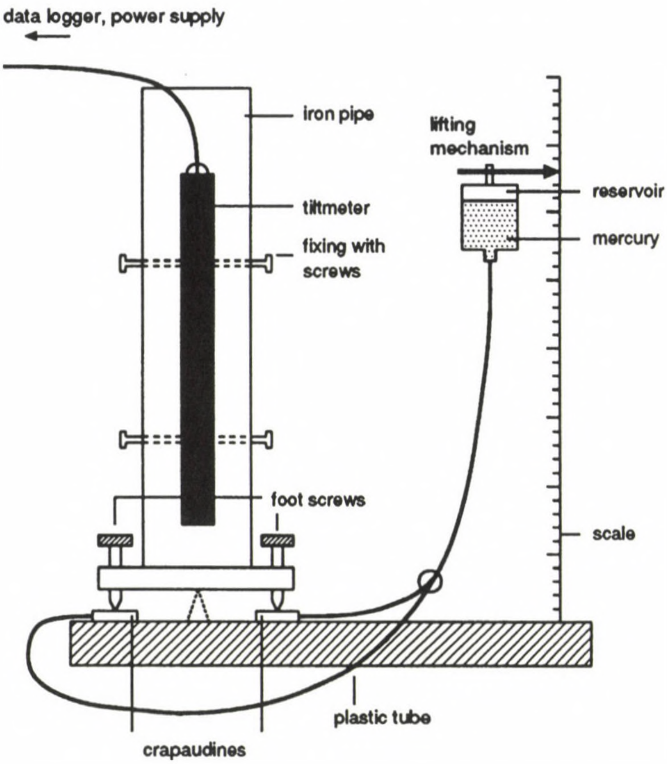


Fig. 4. Relative calibration method using crapaudines

by screws. Then, laser and crapaudine calibrations of the same instrument were simultaneously carried out in a second laboratory at $T = 12°C$. Here, the tiltmeter was fixed either with screws or with sand. These investigations allow a comparison of the tiltmeter's characteristics under the influences of temperature, calibration method and way of coupling (Table II). Fig. 5 shows a synopsis of the results.

AGI gives a linear relation between the scale factor and the temperature, described by a temperature coefficient K_S and a zero shift SK_Z. For the used tilt sensors, K_S should be about +0.05%/°C, SK_Z approximately 1–2 μrad/°C. From the scale factors obtained at 12 and 21°C, the temperature coefficients for these sensors can be calculated as $K_{Sx} = 0.5338\%/°C$ and $K_{Sy} = -0.5348\%/°C$. These coefficients are nearly equal in both directions, probably due to the symmetrical construction of the sensors. The reason for different signs must be investigated in further studies. The scale factors for both calibration methods at the same boundary conditions, temperature and coupling, confirm each other within their uncertainties. In this test run, the way how the tiltmeter is fixed to the iron tube is insignificant only for one tilt component (x). For the second component (y), the scale factors differ by about 0.00074 μrad/mV. It is possible that the estimated uncertainties, gathered from earlier calibrations at Bonn, do not represent the conditions in this

Table II. Example of the recalibration results (tiltmeter, serial no. 1655)

Calibration method	Crapaudine	Laser interferometer			AGI
Range		$\approx \pm 200 \mu rad$			$\pm 800 \mu rad$
Fixing	screws	screws	sand	screws	
Temperature [°C]	12	12	12	21	22
X-Component					
Scale factor [μrad/mV]	0.09644	0.09634	0.09665	0.10097	0.09987
Linearity error [%]	2.7	1.2	1	0.55	3
Correlation factor r^2	0.9988	0.9999	0.9998	0.9999	0.9999
Y-Component					
Scale factor [μrad/mV]	0.09681	0.09759	0.09685	0.09712	0.09997
Linearity error [%]	2.2	1	1.1	0.83	3
Correlation factor r^2	0.9996	0.9999	0.9999	0.9999	0.9997

run. A systematic effect is not obvious, because this should show up in both tilt directions. So far, we have no reason to believe that the coupling conditions at field installations with quartz sand cannot be reproduced by screws in the calibration test equipment.

Three tiltmeters, serial nos. 317, 367, 368, were calibrated in August 1995 in the cellar laboratory of the Geodetical Institute at Bonn applying the absolute method. The room temperature was similar to that of the AGI calibration so that we can compare our results to those given in the instruments' certificates without a temperature correction. In several, repeated test runs, we obtained reproducible values enabling us to describe the tiltmeters' transfer functions for each tilt sensor and each gain (see Fig. 6, as an example). These functions appear to render the bubble sensors' geometries rather than the amplifiers' behaviour. They can be used to calculate calibration parameters for individual measuring ranges in order to correct the tilt data. Within the range of ± 200 μrad, which is the one we use for some of our applications, the scale factor given by AGI had to be changed by order up to 10%. Nonlinearity effects resulted in a maximum deviation of up to 5% of the signal amplitude.

In April 1996, we examined seven more tiltmeters of Model 722 with the absolute method in the same laboratory (serial nos 2520–2525 and 2581). In that way we have data of instruments built at different periods between 1991 and 1996. The yieldings will be presented in a later paper.

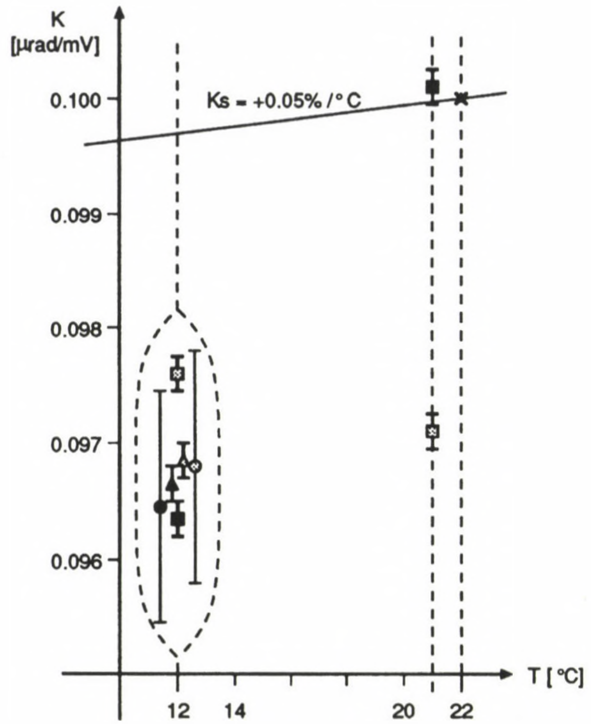


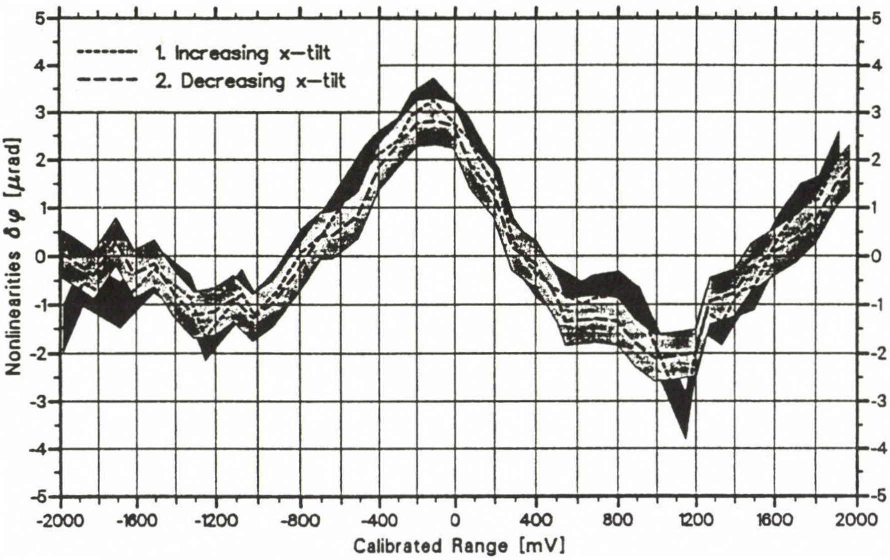
Fig. 5. Comparison of *1.)* scale factors K obtained by laser interferometer calibration, *a)* fixing with screws (squares), *b)* fixing with sand (triangles), and crapaudine calibration (circles) with that of AGI (cross), for the $x-$ (dark) and the y–components (light symbols), respectively, and *2.)* the temperature coefficients K_S and the observed temperature scaling behaviour for tiltmeter, serial no. 1655. Note that error bars represent average errors

7. Conclusions

In order to increase the accuracy of the tilt signals, two independent methods to recalibrate the Applied Geomechanics Inc. borehole tiltmeter, Model 722, were applied. An absolute and a relative method give reproducible calibration coefficients. The obtained characteristics of the instruments for each tilt sensor were evaluated and compared with those given by the manufacturer's calibration certificate. The conclusions of our investigations are:

— For larger tilt amplitudes and phenomenological research, a (corrected) linear scale factor is sufficient to describe the deformation signals that are monitored by the tiltmeter. Within small measuring ranges and for quantitative investigations, knowledge of the nonlinear characteristics appears necessary. Application of correcting terms may result in changes of signal amplitudes by up to 5%.

— Because the manufacturer's calibration temperature does not correspond to that in most field installations, it is necessary to know the temperature be-



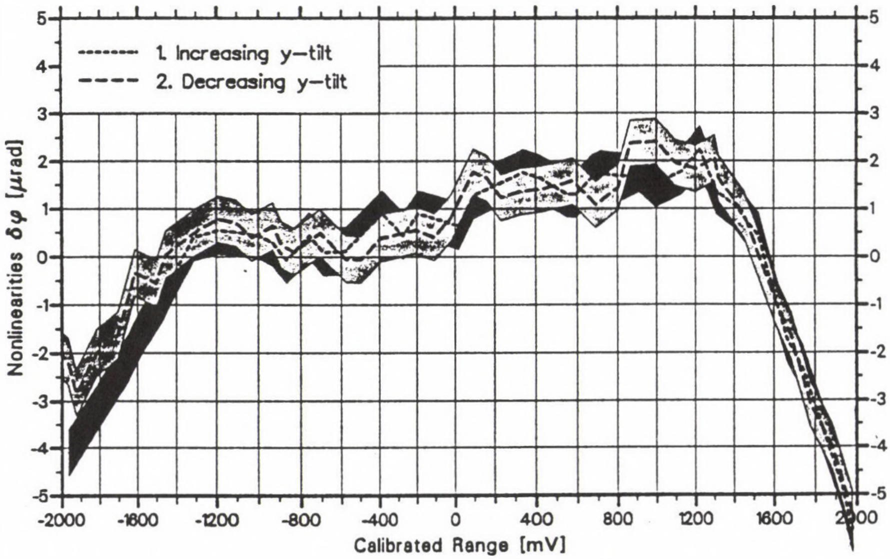


Fig. 6. Nonlinearities for tiltmeter, serial no. 317, a) x–component, b) y–component. With linear regression analysis, scale factors $K_x = 0.08881$ μrad/mV and $K_y = 0.09504$ μrad/mV were obtained within a measuring range ± 2 V ($\approx \pm 200 \mu$rad). This figure shows the deviations from the regression result. For this sensor, AGI gives $K_x = 0.099938$ μrad/mV and $K_y = 0.1001549$ μrad/mV within the range $\pm 800 \mu$rad. The two curves show opposite measurement directions, namely for increasing and decreasing tilt amplitudes. One reading was taken every 100 μrad, on average. The shaded areas mark the uncertainties for each run due to the limited resolution of the laser interferometer

haviour of the sensors. Our results do not confirm the given informations, further investigations are needed.

— The way of fixing the borehole tiltmeters to the iron tube does not significantly influence the transfer function of the instruments. For the calibration, fixing by screws can be applied to simulate coupling at field conditions.

So far, we have recalibrated eleven instruments of the discussed tiltmeter type in order to examine their transfer functions. Details of the latest tests will be published elsewhere.

Acknowledgements

This research was supported by grant WTZ X237.2 of the Scientific and Technological Cooperation between Germany and Hungary through the International Office of Forschungszentrum Karlsruhe and the OMFB, Budapest (Hungarian project no. 76), and by the Sonderforschungsbereich 350, German Research Foundation (DFG). We thank Dr.-Ing. W Schauerte, Geodetical Institute of the University of Bonn, for his help, and making available to us the geodetical laboratory and the laser measurement system. D Rebscher, Bonn, participated in early calibration tests at Sopron and T Molnár, Sopron, cooperated in the calibration tests at Sopron.

References

Graf A 1964: *Sér. Géophys.*, 69, 249–254.

Kümpel H-J 1989: Verformungen in der Umgebung von Brunnen. Habil. Thesis, Faculty of Mathematics and Natural Sciences, University of Kiel, 198p.

Kümpel H-J, Varga P, Lehmann K, Mentes Gy, 1996: *Acta Geod. Geoph. Hung.*, 31 (in press)

Mentes Gy 1986: *Acta Geod. Geoph. Mont. Hung.*, 21, 55–69.

Rebscher D 1996: Nachweis von in-situ Bodenverformungen gravitativen und vegetativen Ursprungs mittels Neigungsmessungen in Bonn–Ippendorf. Ph. D. Thesis, Faculty of Mathematics and Nature Sciences, University of Bonn, 177p.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 91–105 (1996)

EARTHQUAKES AND ASSESSMENT OF SEISMIC RISK

A MESKÓ[1]

[Manuscript received May 13, 1996]

Earthquake hazard is perceived today differently than even a decade ago. It is mainly due to some "unexpected" major earthquakes. Seismic hazard maps of many European countries are being reevaluated and updated. Seismic risk is no longer based on simply the maximum effects recorded in a few hundred years of historical records. Review and reassessment of the hazard has become a vital scientific task in Hungary.

Basic seismological quantities such as intensity, magnitude and peak accelerations and the growing awareness of seismic risk together with its foundations are briefly explained, and the probabilistic seismic risk analysis is discussed with an emphasis on the uncertainties of input data and their consequences on the reliability of estimates.

Keywords: assessment of seismic risk; earthquake hazard; intensity scale; magnitude scale; peak acceleration; seismic source

Introduction

"The vulnerability of Europe to the earthquakes is critically growing ... many populated areas become highly vulnerable to weak earthquakes which did not cause damage in the past", stated the Council of Europe in 1991. Seismic hazard maps of many European countries are being reevaluated and updated since it has been realized that seismic risk can not be simply based on maximum effects recorded in a few hundred years of historical record. Review and reassessment of hazard of several vulnerable large-scale constructions as well as reliable assessment of hazard of several potential sites of such structures to be built has become, therefore, a vital scientific task in Hungary. Previous efforts for evaluating those risks should not be neither underrated nor discarded but should be reconsidered in the view of new experiences. A large scale safety review of the Paks NPP started in 1991 and several improvements are being implemented which will enhance its capability to withstand earthquakes.

In due course of long lasting debate between earth scientists and construction engineers even basic seismological quantities such as intensity, magnitude and peak accelerations were many times misunderstood and misinterpreted. Therefore they are briefly explained as they have been defined, have evolved and are now understood by the international seismological community. The growing awareness of seismic risk together with its foundations is also briefly summarized.

In the last part of the paper the probabilistic seismic risk analysis is discussed with an emphasis on the uncertainties of input data and their consequences on the reliability of estimates. The possible large scatter of estimates is illustrated with an example.

[1]Department of Geophysics, Eötvös Loránd University, H–1083 Budapest, Ludovika tér 2, Hungary

Intensity of earthquakes

Seismologists began to use intensity as an applicable yardstick of the size of an earthquake in the 19th century. Intensity is estimated by means of the degree of damage to buildings and other constructions, the amounts of disturbances to the surface of the ground and, further away from the epicentre, the feelings of the observers.

The assessment of earthquake intensity on a descriptive scale does not depend on measuring the ground motion with instruments but rather on actual observations of the effects. Though the descriptive scale can not be very accurate it continues to be important because the long historical records from seismically active countries are based on such descriptions, and in remote, sparsely populated areas there are simply no seismographs to measure ground motion. Historical sources cover the past millennium with varying degrees of completeness and reliability while instrumental data correspond to less than one century. Under the most favorable circumstances some countries, e.g. China, Japan, countries around the Mediterranean can call upon more than a thousand years of documented historical seismicity, together with more spotty information from ancient times.

When a study of the intensity of an earthquake is made nowadays, questionnaires are circulated to a wide group of inhabitants of the affected region. In addition to the responses to these questionnaires seismologists prepare lists of all kinds of damages. Then estimated intensity values are plotted on maps and isoseismal contours are drawn.

The isoseismal map provides crude but valuable information on the distribution of the ground shaking. The rate at which shaking effects diminishes with distance can be used to estimate focal depth and the coefficient of attenuation.

The first intensity scale was developed by Rossi and Forel in the 1880's. The scale has values ranging from I to X. In 1902 the Italian seismologist and volcanologist Mercalli constructed a more refined new scale with 12 values, which served as the basis for the Mercalli-Cancani-Sieberg (MSC) scale. A somewhat modified and carefully defined version of that scale was worked out by Medvedev, Sponhauer and Kárník in 1964. This scale, called the MSK scale has also 12 values, often called degrees and denoted by °. The latest version, the "European Macroseismic Scale 1992" or MSK-92 has been worked out by the European Seismological Commission (Grünthal 1993). The new arrangement adapted for the scale is:

a) Effects on humans

b) Effects on objects and on nature
(excluding damage to buildings, effects on ground and ground failure)

c) Damage to building.

The new European Macroseismic scale (1992) defines building types and vulnerability classes as well as damage grades and quantifies "few", "many" and "most" more accurately than the previous MSK scale. There are six vulnerability classes, denoted by A, B, C, D, E and F, and the most likely vulnerability classes as well as probable and exceptional ranges are given for several types of structures. The most vulnerable types are rubble stone and fieldstone masonry buildings (class A) and the less vulnerable buildings are the reinforced concrete constructions with high level of antiseismic design (class F and probable class E). In the classification of damage the new scale uses two large groups and detailed descriptions are given separately for masonry buildings and to buildings of reinforced concrete. The names of the 5 grades are:

Grade 1: negligible to slight damage (no structural damage)

Grade 2: Moderate damage (slight structural damage, moderate non-structural damage)

Grade 3: Substantial to heavy damage (moderate structural damage, heavy non-structural damage)

Grade 4: Very heavy damage

Grade 5: Destruction (very heavy structural damage).

In the group of reinforced concrete buildings the description of grade 2 is as follows: "hair-line cracks in columns and beams, mortar falls from the joints of suspended wall panels, cracks in partition walls, fall of pieces of brittle claddings and plaster", and grade 3 reads as "cracks in columns with detachment of pieces of concrete, cracks in beams".

"Few" is understood as 0 % to 10 %, "many" is between 20 % and 50 % and "most" corresponds to the range of 60 % to 100 %. The range 10 % to 20 % might be an exceptional extension of "few" or "many", and the range 50 % to 60 % might be considered either "many" or "most".

The intensity degrees are:

I.	Not felt	VII.	Damaging
II.	Scarcely felt	VIII.	Heavily damaging
III.	Weak	IX.	Destructive
IV.	Largely observed	X.	Very destructive
V.	Strong	XI.	Devastating
VI.	Slightly damaging	XII.	Completely devastating.

As an example the definition of the intensity 9° or IX (Destructive) earthquake is quoted from the MSK-92 description.

"a) General panic. People may be forcibly thrown to the ground.

b) Many monuments and columns fall or are twisted. Waves are seen on soft ground.

c) Many buildings of vulnerability class C suffer damage of grade 3. Many buildings of class B and a few of class C suffer damage of grade 4.

Many buildings of class A and a few of class B suffer damage of grade 5.

Many buildings of class D suffer damage of grade 2, a few suffer grade 3. A few buildings of class E suffer damage of grade 2."

It should be pointed out that class D corresponds to reinforced concrete buildings with minimum level of antiseismic design and class E corresponds to reinforced concrete buildings with moderate level of antiseismic design. It involves, that only a few (normally up to 10 %) class E buildings suffer moderate damage in cases of a 9° intensity earthquakes.

An inventory of historical earthquakes, even completed with estimated magnitudes, may be instructive as to potential sources but is of little help in solving engineering problems which requires physical parameters of ground motion. This motion reflects not only the source and wave propagation characteristics but also the effects of the uppermost soil layers at the site. The accuracy of intensity estimates, obviously depends on the number and reliability of the collected data. For historical earthquakes the accuracy of the estimates depends on the availability of descriptions. Clearly, the descriptions become more and more scarce and less reliable as we go back in time. The accuracy of estimates for earthquakes which occurred more than 200 years ago can hardly be better than 1.0°.

Magnitude of earthquakes

An earthquake magnitude was originally defined by Richter (1935) as the logarithm of maximum amplitude measured in microns on the record of a standard torsion seismograph with a pendulum period of 0.8 sec, magnification of 2800 and damping factor 0.8, located at a distance of 100 km from the epicentre. A calibration curve was constructed to reduce the amplitude, observed at an arbitrary epicentral distance to that expected at 100 km. This magnitude scale is now referred to as local magnitude M_L.

A magnitude scale based on teleseismic surface waves was described by Gutenberg and Richter (1936) and developed more extensively by Gutenberg (1945). For shallow earthquakes at distances between 15° and 130° he proposed the formula

$$M_S = \log A + 1.656 \log D + 1.818 ,$$

where A is the horizontal component of the maximum ground displacement (in microns) due to surface waves with periods of 20 sec, and D is the distance. Many formulas for M have been suggested other than the one by Gutenberg with slightly different parameters. The IASPEI (International Association for Seismology and Physics of the Earth's Interior) officially adopted the following formula:

$$M_S = \log (A/T)_{\mathrm{max}} + 1.66 \log D + 3.3 ,$$

where $(A/T)_{\mathrm{max}}$ is the maximum of all A/T i.e. amplitude/period values of the wave groups on a record. For $T = 20$ sec the IASPEI formula becomes nearly identical to Gutenberg's original proposal.

Another magnitude scale is based on the amplitude of teleseismic body waves, defined by the formula

$$m_b = \log (A/T)_{\mathrm{max}} + Q ,$$

where Q is a function of epicentral distance and focal depth for eliminating the path effect from observed amplitudes. $(A/T)_{\mathrm{max}}$ is again the maximum in the wave group of either P, PP, or SH. Separate tables and charts of Q have been constructed for each phase. The first calibration curve has been empirically determined by Gutenberg and Richter (1956) and several times improved.

The convenience of describing the "size" or "strength" of an earthquake by one number (the magnitude) proved to be applicable. Unfortunately, during the attempts to create an universal magnitude several competing schemes were introduced and the resulting confusion led to lack of confidence in the quantitative aspects of magnitude.

The best practice of seismological observatories in recent years has been to concentrate on two magnitude scales. For shallow earthquakes with well-recorded surface waves M_S is used. M_S, called surface-wave magnitude, gives roughly a continuation of the local Richter magnitude and is best suited to characterize the size of damaging earthquakes. The body-wave magnitude, on the other hand, denoted by m_b has the great advantage that the registration of any earthquake, deep or shallow, local or far away, always starts with a P onset. The two generalized magnitude scales have now been developed so that there is reasonable consistency for large and moderately large global earthquakes. There are, of course, difficulties due to focal mechanism, path effects and geological structures beneath the seismological stations. Treatment of these variations continues to be an important seismological task. As a whole, magnitude determinations have uncertainties of at least 0.1 and sometimes 0.2 units.

The body wave magnitude and surface wave magnitude are different. The approximate relation between m_b and M_S is

$$m_b = 2.5 + 0.63 M_S \ .$$

Magnitude has also been related empirically to maximum or epicentral intensity. Although the correspondence is a rough one, it is nevertheless useful for estimating magnitudes of historical earthquakes. The first such formula is due to Gutenberg and Richter (1943) and reads as

$$M = aI + b + c\log(H) \ ,$$

where M is the (local or surface-wave) magnitude, I is the maximum intensity, H is the focal depth, and a, b and c are numerical constants. The constants can be slightly different for different geographical areas. The method of determination usually includes collecting all those earthquakes for which both magnitudes and maximum intensities as well as focal depths are available and determine the parameters by least mean square fitting. It is clear that the magnitude — intensity relation can not be treated as a law of nature or an exact relation between two well-defined and exactly measurable physical quantities. On the contrary, the linear dependence of magnitude on maximum intensity should be considered a useful approximation.

Hungarian seismologists (e.g. Zsíros et al. 1989, Zsíros 1990) computed magnitudes of historical earthquakes from the maximum intensity and (estimated) focal depth by the simple empirical formula

$$M = 0.6 I_0 + 1.8 \log H - 1.3 \ .$$

The parameters (i.e. $a = 0.6$, $c = 1.8$ and $b = -1.3$), have been determined from a few recent earthquakes. The scatter of points in a $M - I_0$ diagram is quite large, illustrating that maximum intensity and magnitude are not deterministically related. Roughly speaking, magnitude measures the energy and intensity measures surface effects (related to amplitude, velocity and acceleration of ground motion as a function of frequency). The relation between magnitude and intensity as well as the error margin of the intensity of historical earthquakes involves that magnitudes of historical earthquakes can not be accurately determined. As a whole, the accuracy of the derived magnitude can hardly be better than ± 0.5.

Magnitude is sometimes roughly estimated from the length of surface fault rupture length L, a procedure clearly applicable to shallow earthquakes only. Bolt (1978) gave the approximate formula based on worldwide data

$$M_S = 6.03 + 0.76 \log L$$

(L is measured in kilometers). The Japanese estimate (Matsuda 1975) reads as

$$\log L = 0.6M - 2.9\ .$$

Acceleration of ground shaking

Acceleration of ground shaking does have much bearing on the forces affecting a structure. All structures must withstand the pull of gravity. Therefore, even when no special earthquake code is applied, they will usually withstand acceleration in a vertical direction during earthquake shaking. In contrast, experience has shown that horizontal motions are critical and the back-and-forth motion of the ground can cause considerable damage. Acceleration of the shaking is usually expressed as a fraction of gravity (g). Although peak value of horizontal acceleration is important, damage to structures may be occurring throughout the entire period of strong ground shaking. The second parameter of importance in acceleration records is the duration of strong shaking, often called the bracketed duration. This is the duration of shaking above a certain threshold acceleration value, commonly taken $0.05g$ and is defined as the time between the first and the last peaks of motion that exceed this threshold value.

Acceleration is measured by strong motion seismographs often called accelerographs designed to operate near the source of the earthquake in such a way that it would not go off-scale even during the strongest shaking. Accelerographs have now provided literally thousands of records of seismic shaking, both away from and within buildings. Measurements of such accelerograms indicate that the acceleration in the shaking of firm ground in most moderately large earthquakes, at places a few tens of kilometers from the source, lies in the range of $0.05g$ to $0.35g$. For a few events it has been found that some peaks of high-frequency waves may reach unexpectedly high accelerations. (E.g. the Bear Valley, California earthquake of 4th September, 1972, $M = 4.5$, the Ancona, Italy earthquake of 21st June, 1972, $M = 4.7$ produced horizontal accelerations over $0.6g$.) Generally, the peak vertical acceleration is less than the peak horizontal acceleration, the average ratio is of about 50 percent.

Seismic ground motion on soil sites can be substantially larger than on rock sites. There are many documented cases in which the soil, because it is less stiff than the underlying rock, has amplified bedrock earthquake ground motions. An extreme example of this was the Mexico City earthquake of 19th September 1985. Amplification factors of 2 even at frequencies around the fundamental period of the ground motion are not exceptional.

When measured accelerograms are not available at the site, expected accelerograms could be computed by using an appropriate input accelerogram, chosen from a strong motion database and the transfer properties of the geological column under the site. The input accelerogram should match the magnitude, distance and focal mechanism characteristic to the seismogenic areas around the site. The geological column is usually approximated by horizontal layers. The required data is: velocities of longitudinal as well as of transversal waves and quality factors representing attenuation for both types of waves for all layers. The angle of incidence, depending on distance and geological structure is an additional variable. Computations, however, by no means make measurements superfluous.

The changing awareness of earthquake hazard

Earthquake hazard is perceived today quite differently than even one or two decades ago. It is mainly due to some "unexpected" major earthquakes.

In 1968 soviet scientists at the Institute of Physics of the Earth in Moscow produced improved seismic zonation map for the whole Soviet Union. On the map southern Armenia appeared as intensity 7 to 8 (on the MSK scale). To be more specific for the town Spitak the map stated intensity 8. Earth decided differently and on the 7th December, 1988 produced a devastating earthquake with a maximum intensity of 10.

There were some other surprises. For the town of Gazli, built on a large gas reservoir 90 km north-west of Bukhara, in the desert of Uzbekhistan, the map indicated a maximum intensity of 5 to 6, which was not enough to prevent the destruction of the town in 1976 by a magnitude 7 earthquake with a maximum intensity exceeding 9. Architects continued to believe the official map, produced by the prestigious institute and buildings were reconstructed according to the original estimate of hazard. In 1984 the town was once again demolished by another intensity 9+ earthquake.

Such unexpected earthquakes by no means happened only in the (former) Soviet Union. Canada has also received some unpleasant surprises. In 1985 two earthquakes, magnitude 6.6 and 6.9 occurred within a few month in the region of North-West Territories, mapped as having an upper bound magnitude of only 6.0. Far-away from any human settlements it startled mostly seismologists in Ottawa.

The scientists involved in the preparation of the seismic zoning made the fundamental mistake of assuming that a few hundred years of their recorded history had produced a sufficient sample of earthquakes to define the highest effects that would be produced in the future. The implicit logic involved is that it assumes the tectonic time-scale that controls the repetition of earthquakes is shorter than that of human recorded history. It is only in case of the fastest moving plate boundaries in Japan or Chile that the recurrence interval is shorter than the historical record. Even on the San Andreas fault two centuries of observation is not enough to establish recurrence of major earthquakes. The Northridge earthquake of 17th January 1994 occurred along a buried thrust fault, previously unknown. Moving to the east, away from the San Andreas fault the crust is deforming less rapidly and the recurrence interval lengthen, perhaps to a few thousand years, along faults of the Basin and Range. Still further to the east, the recurrence intervals of major earthquakes along faults are thought to have, and in some cases have been demonstrated to exceed 100 000 years.

The reason that almost every year large earthquakes cause astonishment and are labeled as surprises is because they are compared with a historical record that is grossly insufficient to display the complete picture of seismicity. A number of European countries defined their seismic hazard solely of maximum effects recorded in a few hundred years of history, that may be less than 1 percent of true recurrence interval.

In this context it might be useful to remember some recent earthquakes. In 1988 an earthquake of magnitude 6.9 struck Armenia leaving at least 25 000 dead. According to the seismic zoning map by the "scientifically approved" experts it could never have been possible. As a comparison, a year later the San Francisco earthquake of magnitude 7.1 caused only 63 fatalities.

More recently an earthquake (1995) reduced Neftegorsk's buildings to instant rubble leaving two thirds of the town's 3000 inhabitants buried in the debris. Quoting President Boris Yeltsin "Neftegorsk has been wiped off the face of the earth". Though the maximum credible earthquake at the area of the G/N project is smaller than the Neftegorsk's earthquake by 1.0 unit (between magnitudes 6.0 and 6.5) it is symptomatic that six seismic measuring stations on Sakhalin had been closed as an economy measure.

As far as seismic risk is concerned, level of uncertainty is inherently high. Anyone could realize the level of uncertainty involved in earthquake prediction. Japan, from obvious reasons, spent over 1 billion USD over the last 30 years on scientific earthquake monitoring and prediction without predicting a single major quake. E.g. the Kobe earthquake of 7th January 1995, magnitude 7.2 struck unexpectedly. The situation is very similar in California. The Northridge earthquake (17th January 1994, magnitude 6.8) was generated by a blind thrust fault unknown to researchers. But awareness and research is rewarding. Due to updated building codes and their enforcement, supported by continuous warning of seismologists, and the understanding and awareness of the general public, the latter earthquake claimed 61 lives while the earthquake of identical magnitude (6.8) in Northwest Armenia in 1988 caused at least 25 000 fatalities

The Council of Europe, realizing that earthquake risk is neither a prediction nor a hunch but a very real potential danger, has called the attention of European governments, at the occasion of the international conference on earthquake prediction (Strasbourg, 15–18 October, 1991) to the following:

"The vulnerability of Europe — its people, ecology and economy — to the earth-

quake is critically growing, due to the growth of population, destabilization of the ground and infrastructures in large cities and proliferation of constructions which, if damaged, will trigger a catastrophe. For these reasons, many populated areas become highly vulnerable to weak earthquakes which did not cause damage in the past.

Significant reduction of the damage from the earthquake can be achieved by a comprehensive set of safety measures, permanent and activated at different stages of earthquake prediction: long-term (tens of years) and short-term (weeks or less).

Earthquake predictions feasible at present are very limited in accuracy and reliability. Nevertheless, they may allow to prevent significant part of the damage, if proper safety measures are undertaken."

For many countries, it may be convenient to allow seismic hazard to languish at the state of the art of the 1970s because to update seismic hazard map would only throw into doubt and create public alarm over engineering decisions made at the time. However, earthquakes will occur in sensitive areas above the naive impositions of "maximum earthquake effects". When they do the public and media response becomes both dramatic and irrational. It is up to geoscientists, schooled in a time-scale that extends out from the historical, to call attention to the real risk and be brave enough to declare that engineering decisions are flawed and irresponsible.

Assessment of seismic risk

Seismic hazard analysis is understood as the quantitative estimation of the hazard of ground shaking due to earthquakes at a site. Recent methods of analysis incorporate geological, geophysical, seismological and earthquake engineering information's in a systematic way. Hazard analysis provides valuable information for decision making on mitigating the earthquake threat.

Comprehensive routine monitoring system of microearthquakes is sometimes established because a log-linear extrapolation of the frequency (logarithmic scale) — magnitude (linear scale) distribution of smaller earthquakes provides the only way to estimate the frequency of larger earthquakes in regions of relatively low or moderate seismicity. That kind of recurrence pattern of earthquakes has been well documented on a worldwide basis and it is extensively used for assessment of long term seismic risk (e.g. Cornell 1968, ICOLD Manual, 1987) for major engineering projects.

There are two basic approaches of assessment of seismic hazard: the probabilistic and the deterministic approach. The probabilistic assessment provides a basis for representing natural variability and allow the incorporation of uncertainties arising from incomplete knowledge. Results usually are given as probabilities of not exceeding some particular level of ground motion at a site during a time period of interest. The basic input data sets are the distribution of earthquakes in a wider neighbourhood of the site and the ground motion attenuation with distance. The deterministic analysis identifies seismic sources and estimates the Controlling Earthquake (CE) or the Maximum Credible Earthquake (MCE) for each source. For various areas of the Pannonian Basin the MCE is usually assumed to be between 6.5 and 7.0. Magnitudes and distances of the MCEs and elastic parameters of layers below the site allow the computation of site specific ground motion due to these earthquakes. Usually acceleration of ground shaking as a function of frequency or bracketed acceleration is determined. In order to apply the deterministic earthquake hazard assessment input accelerograms are used which could rightly be considered as representative accelerogram of the Maximum Credible Earthquake. The detailed geological model in the simplest case of horizontal layering should include the following parameters: thickness' of the layers, velocities of longitudinal and of transversal waves in the consecutive layers and, at last, quality factors for both types of waves in the consecutive layers. The difficulties of the deterministic approach will be discussed elsewhere.

Following the international definition (e.g. McGuire and Arabasz 1990) we understood the seismic risk analysis (SRA) as the estimation of future possible earthquake effects, specifying a time period for which the estimation is made, a quantification of the adverse effect, and the probability of the adverse effect during that time period. The probabilistic seismic hazard analysis (PSHA) is considered a vital part of any SRA and it is understood as the quantitative estimation of the hazard due to earthquake ground motion at a site, considering all possible earthquakes in an area, estimating the associated shaking at the site and calculating the probabilities of those occurrences for specified time periods.

The fundamental elements of a PHSA are the descriptions of the earthquake environment: the seismic sources, the magnitude distribution and the ground motion as a function of distance and magnitude, and a probability analysis summarizing all those input data into an integrated picture of the seismic hazard. Each set of inputs requires some hypothesis or interpretation and both the lack of data or competing interpretations introduce uncertainties. In the following discussion each element is described together with the difficulties of its unambiguous determination. The international and Hungarian practice is compared and the possible consequences of uncertainties are illustrated by some examples.

Sources of earthquakes

The first element of a PSHA is to identify where earthquakes will occur. Along tectonic plate boundaries active crustal faults or large-scale seismogenic features are usually specified as sources of earthquakes. In intra-plate tectonic setting, however, observed seismicity may not be readily associated with identifiable faults. Each seismic source is specified by some map and depth geometry that includes the hypocenters of potential earthquakes. The depiction of seismic source zones inevitably involves practical considerations relating to scale and the modeling of an imperfectly understood, complex earth. Translating faults from a geological map to a source-zone map requires interpretation and judgment relating to inferred faults, continuity of faulting etc. The subsurface geometry of faulting must be specified in addition to the fault trace. Although more complex assumptions are possible, it is common practice to assume that a fault is equally likely to rupture anywhere along its length or depth. Scale considerations may leąd to grouping small individual faults or close faults into areal or volumetric sources.

Seismic source zones are drawn to delineate discrete domains where earthquakes originate with similar character. In this regard there must be consistency between the definitions of the seismic sources and the magnitude distributions derived to described them. Various assumptions are possible for seismic source zones, but the most common are that each source describes a domain within which earthquakes

i. are equally likely in space,

ii. conform to the same magnitude distribution,

iii. have the same maximum magnitude, and

iv. are independent from event to event.

If the prediction of ground motion depends on the style of faulting or any other property of the source, this property must be specified for each source.

At several places in the Pannonian Basin association between faults and observed seismicity is problematical and weak therefore it is difficult to specify source zones of earthquakes. Source zones which are being regularly used by the Hungarian Seismological Observatory can be found in Zsíros (1990). Quadrangle shaped areas are chosen to make computation easier and depth distribution is not considered. All points within source zones are considered equally probable epicenters of earthquakes. Though this subdivision and the related assumptions might be appropriate as a first approximation, further improvements are necessary.

Magnitude distribution

For each seismic sources a probability distribution of earthquake sizes are required. Earthquake magnitude is regularly used as the standard measure of size. In Hungary, however, seismologists prefer intensity since magnitude determination for historical earthquakes seems more difficult. The advantage of dealing with intensities involves that the frequency of a given intensity at the sites can easily be derived from the intensity distribution of the seismogenic zone and the intensity attenuation relation. The uncertainty of determination becomes higher if the number of events decreases. The frequency of larger quakes has a large margin of error.

Figure 1 illustrates a magnitude distribution which is used as an estimate of a source area with dimensions 20 km × 100 km. The time history includes 110 years and the confidence intervals of the straight line fitted to the observed frequency-magnitude data are also shown. The upper-bound earthquake is assumed to be of magnitude $m_{\max} = 7.0$. Note, that the recurrence time interval of a magnitude 6 earthquake is about 500 years and that of the upper-bound earthquake is about 3000 years.

Effects of the attenuation

The frequency of intensities at the sites is the sum of contributions from each element of the source area. Due to the attenuation the contribution of an element is shifted towards smaller intensities, depending upon its distance from the site. There are two points to be mentioned now.

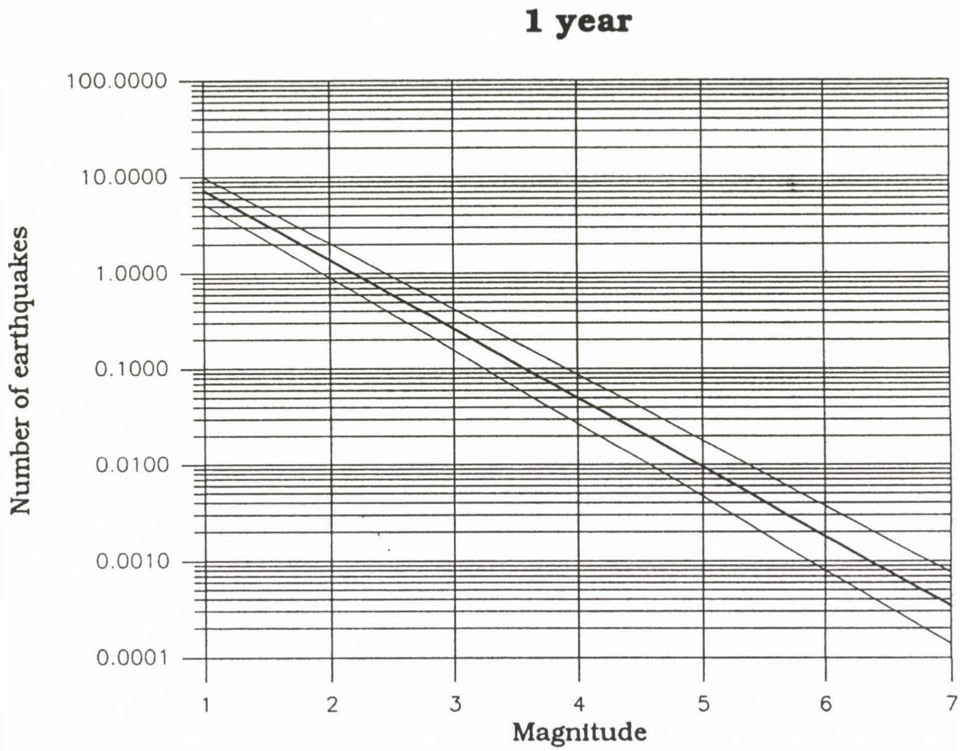


Fig. 1

The first is the upper limit used for the distribution. If one assumes that intensities higher than a given limit, e.g. $I_{max} = 9°$ will never occur in the source area, merely because they did not occur in the investigated short time period, one severely cuts the upper end of the distribution at the site. Defining a source zone at a given distance and limiting unrealistically the high intensity part of the distribution results in unrealistically low seismic risk. It seems to be a better practice estimating first the recurrence interval and determine the upper limit by applying this time limit to the intensity-frequency relation.

The second remark concerns the role of the distribution of focal depths. Attenuation depends on the focal depth as well as on the distance, and the Kövesligethy formula clearly expresses this dependence. If one uses a single attenuation curve which includes only the distance then shallow and relatively deep quakes give the same contribution. It is in contrast to observations. Some experimenting with the Kövesligethy formula and an assumed depth distribution shows that the intensity-frequency relation can change. It should also be mentioned that taking into account the depth distribution does not necessarily increases the frequency of higher intensities.

Computation of probabilities for given time intervals

Neglecting site effects, i.e. the modification of intensity due to near-surface geological conditions, the intensity of an earthquake generated in the source area can be determined from the magnitude and epicentral distance. Monte-Carlo modeling can be used to compute an intensity-frequency distribution $f(I_M)$ at the site corresponding to the assumed magnitude-frequency distribution and the assumed areal distribution of earthquakes. The probability that I_M is exceeded can also be determined for any time interval. Figure 2 illustrates intensity-probability relations, obtained from ten separate Monte-Carlo modeling for the time span of 100 years. The magnitude distribution of Fig. 1 and a uniform areal distribution have been applied while the effect of depth has been neglacted There is a large scatter. One can state, however, with some confidence that the probability of an earthquake occurring in the source area and causing an intensity 6.5 or higher at the site is about 0.5.

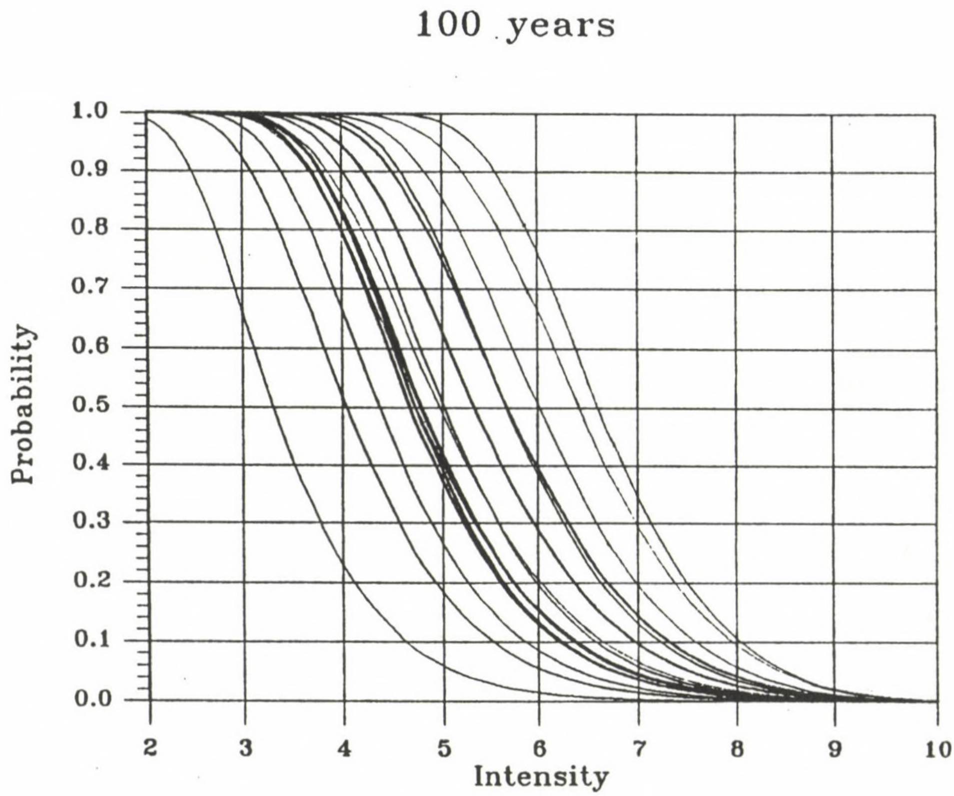


Fig. 2

References

Aki K, Richards P G 1980: Quantitative Seismology. W H Freeman, San Fransisco

Bath M 1973: Introduction to seismology, Birkhauser Verlag, Basel

Bolt B A 1976: Nuclear explosions and earthquakes. W H Freeman, San Fransisco

Bolt B A 1978: Earthquake (a primer). W H Freeman, San Fransisco

Bondár I 1994: Effect of local geological conditions on the accelerations expected in the GNBS Report, Institute of Geodesy and Geophysics, Hung. Ac.of Sci., Sopron, Vol. 4, part II.

Chi-Yu King 1992: In: Encyclopedia of Earth System Science, W E Nierenberg (ed. in chief), Vol. 2. Academic Press, 111–117.

Coppersmith K J, Youngs R R 1986: Capturing uncertainty in probabilistic seismic hazard assessments within intraplate tectonic environments. Proc. 3rd. U.S. Nat. Conf. Earthq. Eng., 1, 301–312.

Cornell A C 1968: *Bull. Seism. Soc. Am.*, 58, 1583–1606.

Council of Europe 1991: Earthquake prediction: state-of-the-art. Conclusions of the Conference, Strasbourg, France

Esteva L 19..: In: Seismic risk and engineering decision, Lomnitz, C Rosenblueth, eds, Elsevier, Amsterdam

Gutenberg E B 1945: *Bull. Seism. Soc. Am.*, 35, 3–12.

Gutenberg B, Richter C F 1936: *Science*, 83, 183–185.

Gutenberg B, Richter C F 1956: *Bull. Seism. Soc. Am.*, 46, 105–145.

I A E E 1984: Earthquake Resistant Regulations: A World List. International Association for Earthquake Engineering.

ICOLD (International Commission on Large Dams) 1987: Selecting seismic parameters for large dams. Bulletin 72, Paris

Kárník V 1971: Seismicity of the European Area (Part I). Czechoslovak Academy of Sciences, Praha

Kárník V, Schenková Z, Bune V I 1978: Scheme of Earthquake Provinces, Praha

Kárník V, Procházková D, Schenk V, Schenková Z, Brouček I 1988: *Studia Geoph. et Geod.*, 32, 44–150.

Kövesligethy R 1906: *Gerl. Beitr. Geophys.*, Band VIII, Leipzig

Matsuda T 1977: *Japan Journ. Phys. Earth*, 25, 251–260.

Matsuda T, Kinugasa Y 1991: *Episodes*, 14.

McGuire R K, Arabasz J W 1990: In: Geotechnical and Environmental Geophysics, S H Ward ed., SEG Tulsa, Oklahoma, 333–353.

Medvedev S V 1962: Engineering seismology, Moscow

Meskó A 1994: In: Natural disasters in Hungary, National Report for United Nations, Cs Nemes ed.

Müller G 1985: *Journal of Geophysics*, 58, 153–174.

Nuttli O W 1981: On the problem of maximum magnitude of earthquakes. Proc. Conf. XIII, Evaluation of Regional Seismic Hazards and Risk, USGS Open File Report 81–437, 876–885.

Polo de C M, Clark D G, Slemmons D B, Ramelli A R 1991: *Journal of Structural Geology*, 123–136.

Réthly A 1952: Earthquakes of the Carpathian Basin between 455 and 1918 (in Hungarian). Akadémiai Kiadó, Budapest

Richter C F 1958: Elementary seismology. San Francisco

Runcorn S K 1967: International Dictionary of Geophysics, Pergamon Press

Stuart W D, Aki K eds 1988: Intermediate-Term Earthquake Prediction. Birkhauser Verlag, Basel

World Data Centre A for Solid Earth Geophysics 1979: Manual of Seismological Observatory Practice Report SE-20, U.S. Dept. Commerce, P L Willmore ed., IGS, Edinburgh

Zsíros T 1990: *Magyar Geofizika*, 31, 124–132 (in Hungarian).

Zsíros T, Mónus P, Tóth L 1989: Hungarian Earthquake Catalogue (456-1986). Hungarian Academy of Sciences, Budapest

Acta Geod. Geoph. Hung., Vol. 31(1-2), pp. 107-114 (1996)

SEISMIC HAZARD ASSESSMENT FOR CSALLÓKÖZ REGION

T ZSÍROS[1]

[Manuscript received November 10, 1994]

Csallóköz region — the studied area — is a border region between Hungary and Slovakia along the river Danube, where a controversial hydroelectric power plant has been built. In this paper the seismic hazard assessment of the region is presented using the traditional probabilistic approach (Cornell 1968). The maximum expected earthquake intensities with 70 % and 90 % probabilities of not being exceeded in 100 years vary from 6.6 to 6.85 and from 7.1 to 7.7, respectively in the region. For 1000 year time span the maximum expected intensities at 70 % and 90 % probabilities of not being exceeded will be between 7.6–8.2 and 8.0–8.6, respectively.

Keywords: Csallóköz region; earthquake; Hungary; intensity; probability; seismic hazard, Slovakia; source zones

Introduction

Natural hazards, including seismic hazards, represent a potential future danger, which can be expressed by the probability that an event of a certain level should not be exceeded during a period in the future. Seismic hazard can be expressed by parameters, such as: earthquake intensity, acceleration, response spectrum, duration of excitation, frequency content, etc. All these parameters, when they reach a certain level, endanger the structures, and represent a potential hazard. Therefore, it is natural to make efforts to foresee, or possibly define the occurrence of the above-mentioned parameters in the future. A simple deterministic approach is not possible, i.e. various future events, including the earthquake, cannot be forecast with certainty. However, based on statistical analysis of past earthquakes, and their observed accelerations, the probability of the occurrence of certain accelerations in the future can be estimated. For hazard estimation, an acceleration distribution function is required

$$F(a/t) = P(A < a/t) \tag{1}$$

i.e. the probability that the acceleration A, should not be exceeded in a period of t years. This means that from acceleration distribution $F(a/t)$ the acceleration which will not be exceeded under some probability and during a period of time is determined. By convolution of the seismic hazard and vulnerability of structures and components as elements of seismic risk, the seismic risk at which the structures and components will be subjected could be determined. Since the attenuation of

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-1112 Budapest, Meredek u. 18, Hungary

accelerations has not been known in the Carpathian Basin we use the intensity decline curves in the estimation process.

The methodology which is currently used for seismic hazard assessment comprises the following steps:

a) Delineation of seismic zones (source of earthquakes)

b) Determination of earthquake frequency

c) Estimation of intensity attenuation

d) Seismic hazard assessment.

The principal technique of this process was originally suggested by Cornell (1968) and further developed by others i.e. McGuire (1976), Algermissen and Perkins (1976), Bender and Perkins (1987). Csallóköz region — the studied area — is a border region between Hungary and Slovakia along the river Danube, where a controversial hydroelectric power plant with a gate system has been built by diverting of the river.

Seismic zones

During the estimation of seismic hazard, at first, the source areas — where the earthquakes are primarily expected — must be defined. Except for the case of zones with high seismic activity, the delineation of source zones is rather difficult. Since the criteria for the delineation of source zones has not been defined exactly it provides an opportunity for debate. It is only sure, that a source area has to contain earthquakes which are needed for the estimation of recurrence curves of events. In a region of relative low seismicity — like the Pannonian Basin — the points of view contradict each other:

— One aim is that the source zone should be delimited to the possibly smallest area, since each point of the source region is considered as a source of future earthquake.

— The other aim is that, however, the source area should be as large as possible, to contain more events, and this makes the estimation of earthquake frequency more reliable.

Nevertheless further argument, information offered by geology, tectonics, geophysics and geodesy influence the determination of source areas, too. It is, however, difficult to determine their weights as long as one has no quantitative information on the measure of the effect exerted by these data and parameters on the seismicity of the area. In this study the determination of source zones have been made, first of all, on the experienced seismicity pattern. Altogether 15 source areas were delineated (Fig. 1) and this seismotectonic model is nearly the same as the one used in earlier study (Zsíros 1990) for the hazard assessment of Paks Nuclear Power Plant. Only one more seismic source zone was defined in the north Zala region of western Hungary.

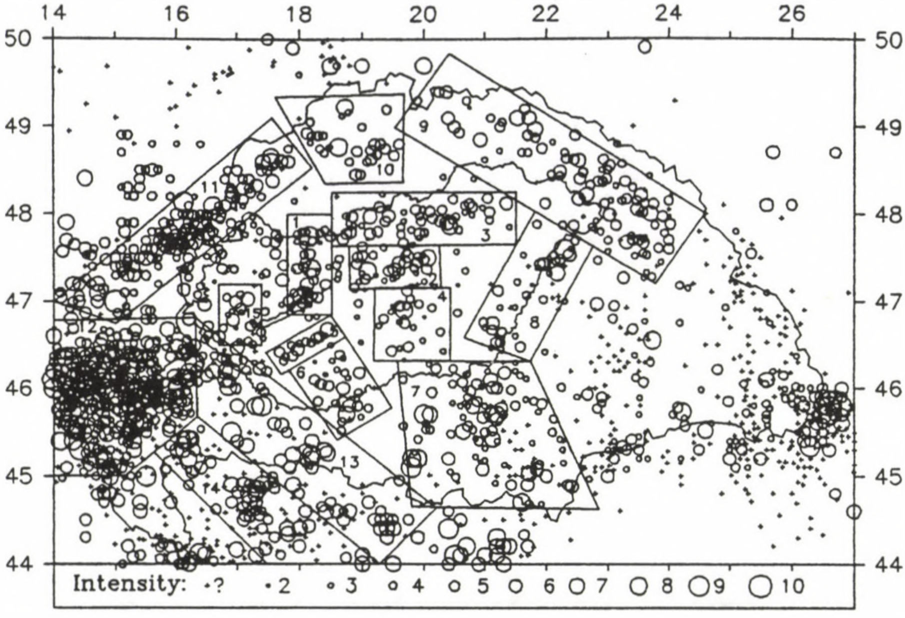


Fig. 1. Earthquake source zones used for hazard estimation with the largest epicentral events within 5 km. The present and the historic (before 1920) borders of Hungary are also plotted

For the studied area (Csallóköz region) source zone No. 1 and No. 11 are the most important ones and the contributions of the farthest zones to the seismic hazard are negligible.

The sizes and shapes of the source areas and the frequencies of the earthquakes occurring in them, respectively, can be naturally determinant for a given site studied. However, the modification of a source area means significant change in the hazard estimation only if it is connected to considerable changes in the area and/or in the earthquake frequency. In this case — due to the attenuation curves — the crucial factor is the distance between the modified source area and the site for which the hazard is estimated.

The procedure used in this study, however, allows some location uncertainty — which is chosen as 20 km for every zone — and so the hazard is calculated by using soft boundaries of the source model.

Frequency of earthquakes

Several studies (e.g. Richter 1958) proved that the frequency of earthquakes of different strength is different, i.e. there exists a linear relationship between the strength (magnitude, epicentral intensity) of the earthquakes and the logarithm of their number. With respect to the epicentral intensity I_o and the cumulative frequency N the equation:

$$\log(N) = a + b \cdot I_o \tag{2}$$

Table I. Threshold parameters used for the selection of the main shocks. R – epicentral distance (km), T – time span after the occurrence of the main shock (day), T' – time span before the occurrence of the main shock (day), I_o – epicentral intensity

I_o	R (km)	T (day)	T' (day)
2°	5	1	1
2° – 3°	5	1	1
3°	10	3	3
3° – 4°	10	3	3
4°	15	10	7
4° – 5°	15	10	7
5°	20	30	20
5° – 6°	20	30	20
6°	25	250	50
6° – 7°	25	250	50
7°	30	470	90
7° – 8°	30	470	90
8°	35	930	150
8° – 9°	35	930	150
9°	40	1400	370
9° – 10°	40	1400	370
10°	45	1830	730

is valid, where
N — number of earthquakes which exceeded I_o
a, b — constants characterizing the seismic activity of the area investigated.

For the determination of the recurrence curves (Eq. (2)) the earthquake catalogue of the Carpathian region (Zsíros 1994) was used. As the assessment method handles the occurrence of earthquakes as a Poisson process, the fore- and aftershocks must be excluded from the determination of the earthquake frequency, since they do not follow the Poisson distribution. In the selection of the main shocks a certain measure of subjectivity is unavoidably included, as it is not defined which shock must be regarded as the main one in every case. Studying the earthquake sequences of significant events in the Carpathian Basin, time and distance threshold parameters were determined (Zsíros 1993) for the selection of the main shocks. These threshold parameters of Table I were used in filtering out of the fore- and aftershocks in this study.

In the vicinity of radius R of the main earthquake with epicentral intensity I_o, all shocks with epicentral intensity $I < I_o$ are regarded as foreshocks if their occurrence time differs from the main shock's occurrence time by less than T or equal to T'. Similarly, all shocks with epicentral intensity $I \leq I_o$ are regarded as aftershocks if their occurrence time differences from the main shock's occurrence time are smaller than T or equal to T'.

In the determination of frequency recurrence curves, the crucial point is the

Table II. Time intervals belonging to the complete observation of earthquakes of different epicentral intensities

Epicentral intensity	Time span
$I_o \geq 8°$	1700–1993
$8° > I_o \geq 7°$	1800–1993
$7° > I_o \geq 6°$	1850–1993
$6° > I_o \geq 5°$	1880–1993

homogeneity and the completeness of the data set. On the basis of the investigation of the earthquakes occurred in the Carpathian Basin, the time suitable intervals belonging to different epicentral intensities are compiled in Table II.

The earthquake frequency was determined in all source zones on the basis of the criteria shown in Table II. At the application of the frequency Eq. (2) for each source area, an upper $I_{\max}$ value is determined and the occurrence of earthquakes with higher epicentral intensity is excluded. On the basis of historical seismicity of the Carpathian region $I_{\max}$ was chosen as 10^o in the 12th source zone (Friuli region), while in the other source zones as 9^o.

Intensity attenuation

An important aspect of estimating ground shaking is knowing how the ground motion produced by an earthquake is modified by transmission from source to site. For this purpose the maximum ground motion parameters used to be measured in different epicentral distances and the average attenuation curves of these parameters are constructed. Since the attenuation has not been known from instrumental ground motion data in the Pannonian Basin, the curves of earthquake intensity decline — as the only data observed — are used for the calculation of the seismic hazard.

On the basis of the isoseismal maps of about one hundred earthquakes occurred in the Carpathian Basin new intensity attenuation curves were compiled (Zsíros 1996) using the Kövesligethy model (Kövesligethy 1906). This model assumes that the energy of seismic waves declines owing to the geometrical spreading and to the absorption of the media. In this way the intensity of the shaking is the function of the epicentral intensity, the epicentral distance, the focal depth and the absorption coefficient of the media. These new intensity attenuation curves are used in the computations.

Estimation of hazard

When the sources of future earthquakes are defined (seismic zones are delineated), the activity rate (frequency) of earthquakes are known in the source areas and the intensity attenuation curves are determined, then the seismic effect on the

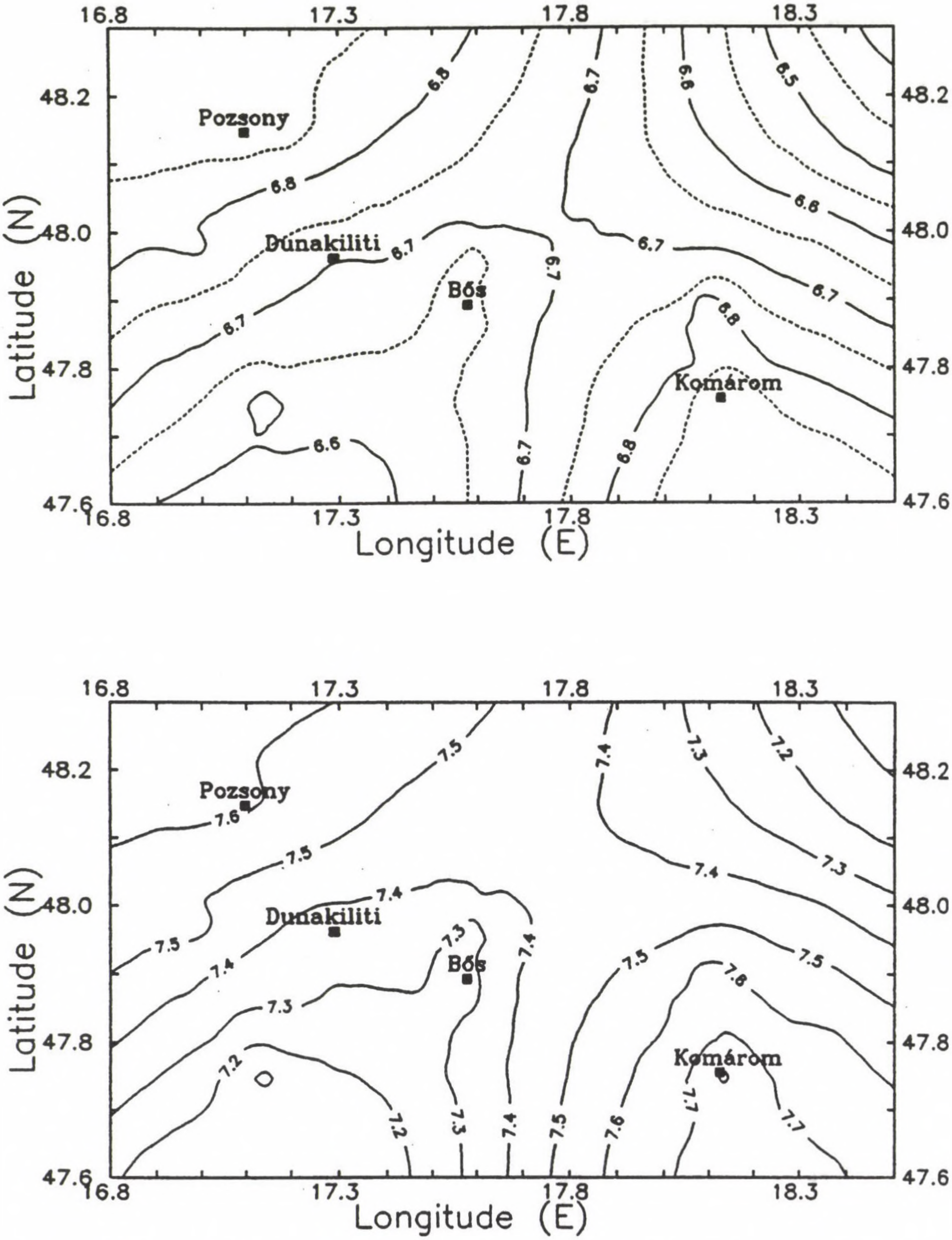


Fig. 2. Maximum earthquake intensity expected not to be exceeded a) at 70 % and b) at 90 % probability in 100 years in Csallóköz region. Slovakian names of the locations shown in the figure: Bratislava–Pozsony, Gabčikovo–Bős, Komárno–Komárom

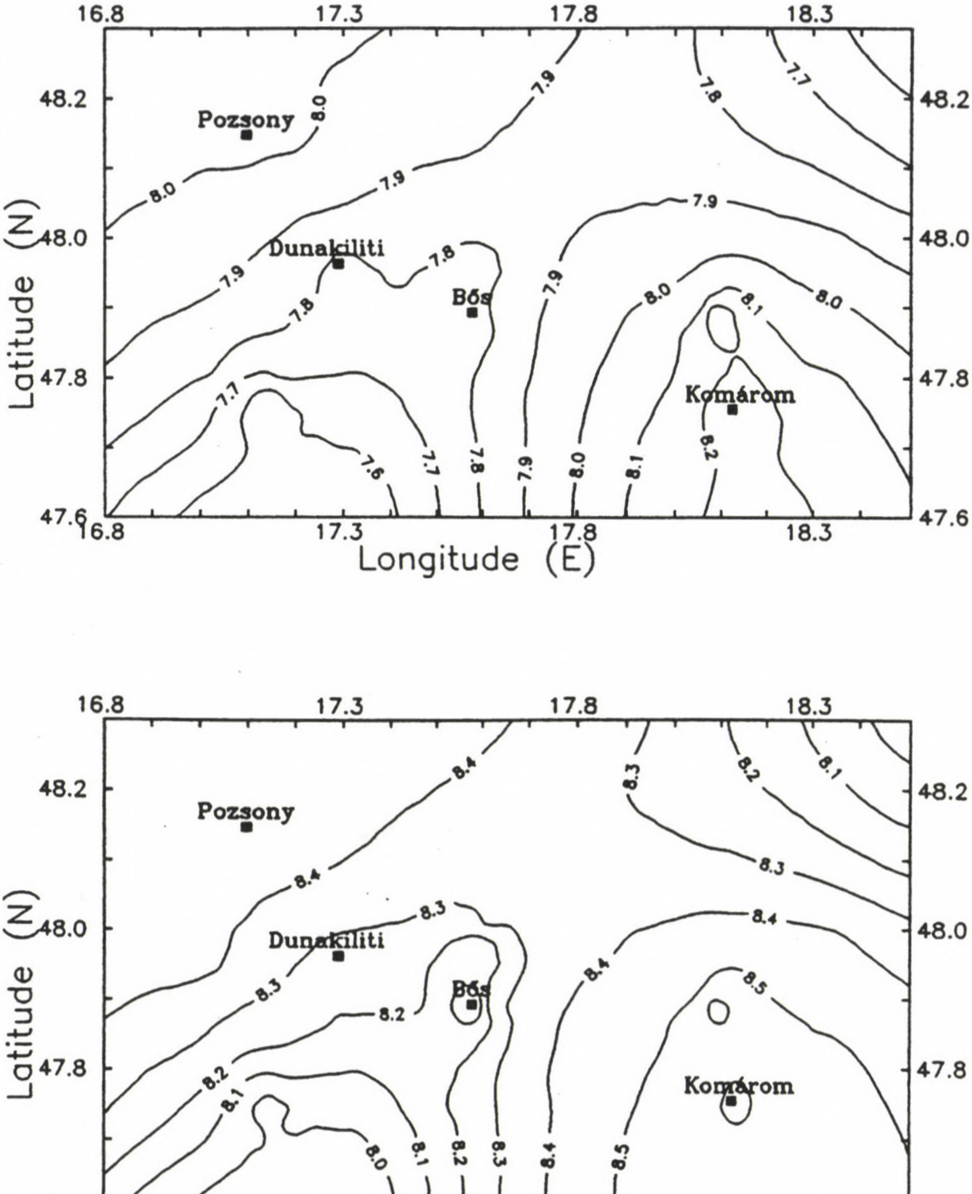


Fig. 3. Maximum earthquake intensity expected not to be exceeded a) at 70 % and b) at 90 % probability in 1000 years in Csallóköz region. Slovakian names of the locations shown in the figure: Bratislava–Pozsony, Gabčikovo–Bős, Komárno–Komárom

site in question can already be calculated. In the actual calculations SEISRISK III computer program (Bender and Perkins 1987) was used. This program — unlike former ones i.e. EQRISK (McGuire 1976) — permits earthquake rates to vary smoothly near a source boundary without being significantly different in the centre of the region from the rates assumed in the uniform-seismicity model. And it is because the locations of actual earthquakes are normally distributed.

If it is assumed that the distribution of earthquakes is Poissonian — that is earthquakes occur randomly with time — the probability of occurrence $P(I > I', t)$ can be calculated as follows

$$P(I > I', t) = 1 - \exp(-N \cdot t) \tag{3}$$

where:
N — yearly number of intensities greater than I'
t — number of years in a period of interest.

For 100 and 1000 year time intervals the seismic hazard at 70 % and 90 % probabilities is presented for Csallóköz region in Figs 2–3.

The maximum expected earthquake intensities at 70 % and 90 % probabilities of not to be exceeded in 100 years vary from 6.6 to 6.85 and from 7.1 to 7.7, respectively in the region. For 1000 year time span the maximum expected intensities at 70 % and 90 % probabilities of not to be exceeded will be between 7.6–8.2 and 8.0–8.6, respectively.

References

Algermissen S, Perkins D 1976: A probabilistic estimate of maximum acceleration in rock in the contiguous United States. U.S. Geological Survey, Open-File Report, 76–416.

Bender B, Perkins D 1987: *U.S. Geological Survey Bulletin*, 1772.

Cornell C 1968: *Bull. Seismol. Soc. Am.*, 58, 1583–1606.

Kövesligethy R 1906: *Mathematikai és Természettudományi Értesítő*, 24, 349–368 (in Hungarian).

McGuire R 1976: EQRISK Fortran computer program for seismic risk analysis. U.S. Geological Survey, Open-File Report 76–67.

Richter C 1958: Elementary seismology. W Freeman and Company

Zsíros T 1990: *Magyar Geofizika*, 31, 124–132 (in Hungarian).

Zsíros T 1993: *Magyar Geofizika*, 34, 7–20 (in Hungarian).

Zsíros T 1994: Hungarian earthquake catalog (456–1993). Seismological Observatory, Budapest (computer file)

Zsíros T 1996: Focal depth and intensity attenuation in the Carpathian region. *Acta Geod. Geoph. Hung.* (present issue)

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 115–125 (1996)

MACROSEISMIC FOCAL DEPTH AND INTENSITY ATTENUATION IN THE CARPATHIAN REGION

T Zsíros[1]

[Manuscript received October 29, 1996]

Based on the isoseismal maps of 124 earthquakes in the Carpathian region (44.5–50.0N; 15.0–27.5E) focal depths are estimated using the Kövesligethy model for the intensity attenuation. The obtained results show that 65 percent of the studied earthquakes occurred in the 12 km upper part of the crust and deeper than 60 km focal depth events are located only in the Háromszék-Vrancea zone. The most frequent depth interval is the 9–12 km depth domain and the estimation yields an average depth of 11 (±7) km and an average absorption of 0.016 (±0.027) km^{-1} in the Carpathian Basin (44.5–50.0N; 15.0–25.0E). Intensity attenuation curves of earthquakes with different epicentral intensities are compiled and the distribution of absorption versus focal depth are shown in the paper. According to the presented comparison the intensity attenuation in the Carpathian Basin is close to that experienced in South Spain.

Keywords: Carpathian region; earthquake; focal depth; isoseismal; Kövesligethy

Introduction

The focal depth of an earthquake is one of the most uncertainly determined source parameters. For historical earthquakes the only available method to determine the depth is the macroseismic way. Even today, we can only rely in some cases on the macroseismic focal depth estimations due to the lack of dense seismic network and due to the model uncertainties in depth-velocity structure of the media. Using macroseismic data, however, it is a disadvantage that both the estimation of earthquake intensity and the compilation of isoseismal map are subjective.

For the focal depth estimations, in this study, the isoseismal maps were selected from the following sources: Cvijanovic (1970), Cvijanovic et al. (1989), Drimmel and Trapp (1982), Gutdeutsch et al. (1987), Kárnik et al. (1981), Karpiv et al. (1980), Kiss (1983), Kostyuk and Rudenskaya (1982), Polonic (1980), Procházkova and Kárnik (1978), Rudenskaya et al. (1983), Shebalin et al. (1974), Szeidovitz (1984), Zátopek (1940) and Zsíros (1983a, 1983b, 1988, 1989, 1995). Every selected isoseismal map contains at least three isoseismals and the coordinates of the earthquake epicentres are between 44.5N–50.0N in latitude and 15.0E–27.5E in longitude. If the source has not contained the radii of isoseismals, the distances of the isoseismals from the epicentre were determined by the "equivalent circle" method. In this method the isoseismals are approximated by perfect circles centred at the epicentre. The epicentral distances are computed by setting them equal to the average radii of equivalent concentric circles having areas equal to that encompassed

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-1112 Budapest, Meredek u. 18, Hungary

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

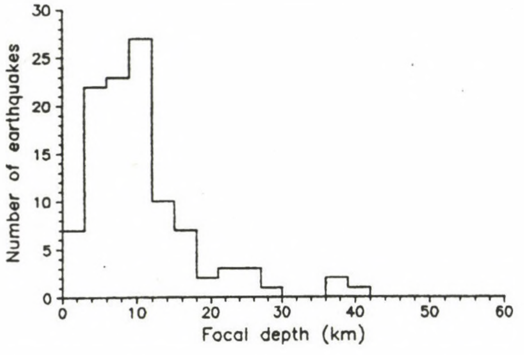


Fig. 1. Frequency distribution of focal depths in the Carpathian region (44.5–50.0N; 15.0–25.0E) based on 108 earthquakes

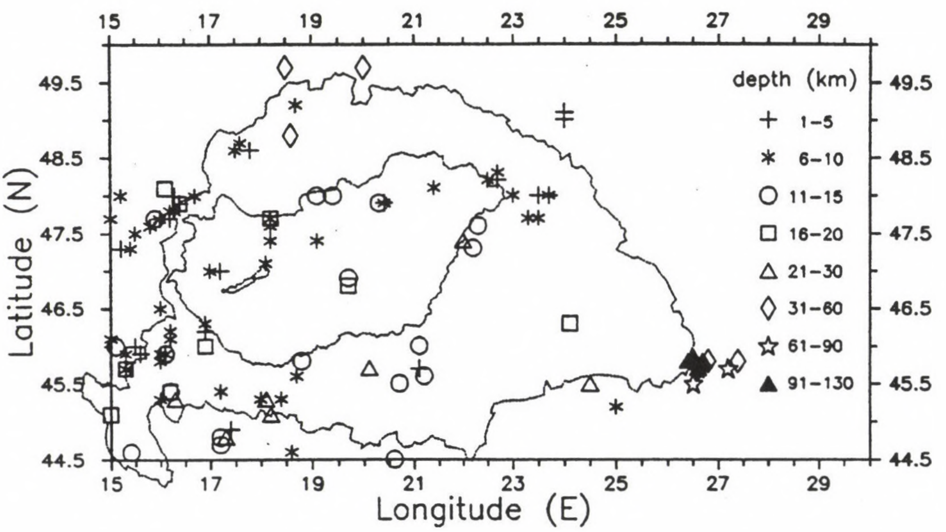


Fig. 2. Epicentre distribution of earthquakes classified into 8 different focal depth ranges with the present and the historic (before 1920) borders of Hungary

by the corresponding isoseismals. Based on the conditions mentioned above 124 earthquakes in all were collected for the focal depth estimations.

Method

The attenuation of earthquake intensity with distance is described by the Kövesligethy formula (1906)

$$I_o - I_k = 3 \cdot \log\left(\frac{D_k}{h}\right) + 3 \cdot \alpha \cdot \log(e) \cdot (D_k - h) \tag{1}$$

where

I_o	–	epicentral intensity
I_k	–	intensity at D_k hypocentral distance
D_k^2	=	$r_k^2 + h^2$
r_k	–	radius value of the isoseismal k (km)

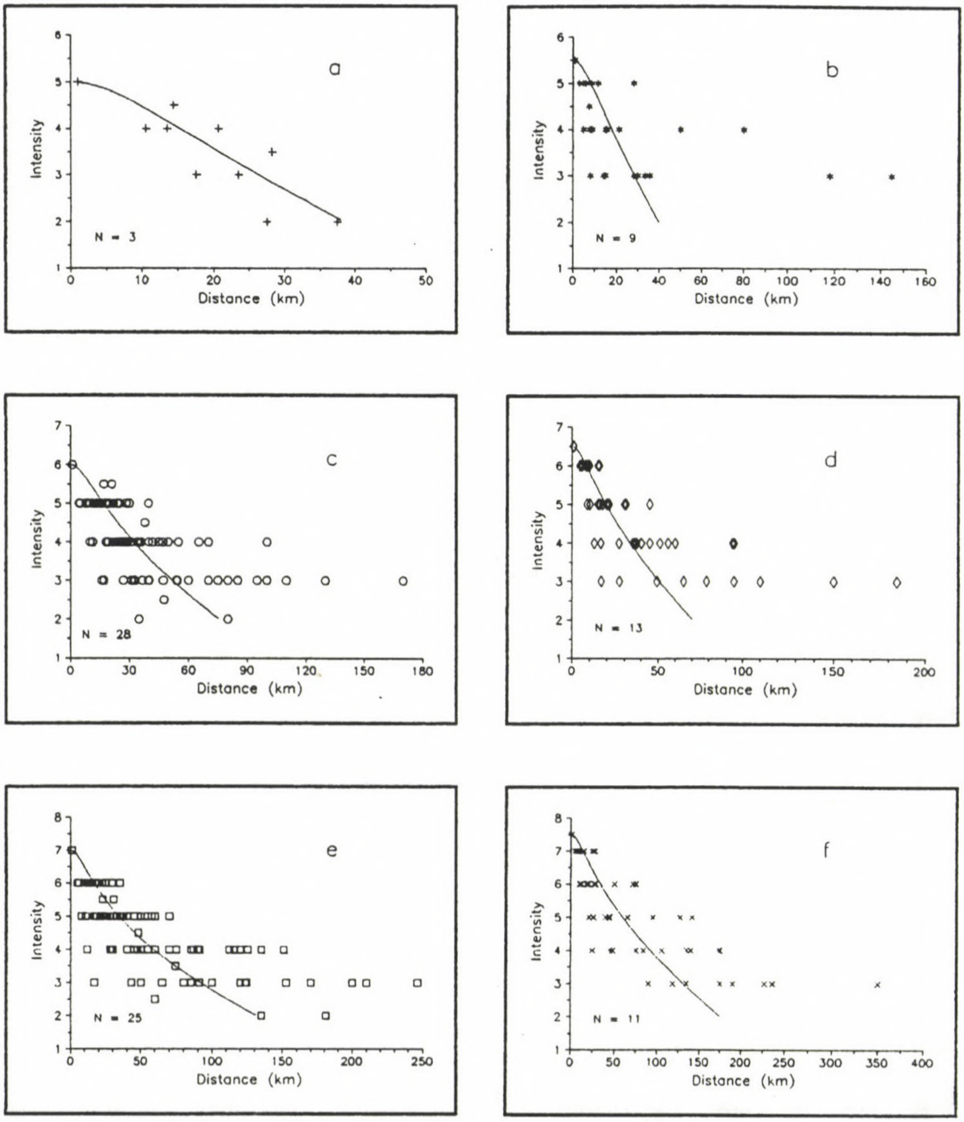


Fig. 3a. a-f: Intensity attenuation of earthquakes with different epicentral intensities in the Carpathian Basin (44.5–50.0N; 15.0–25.0)

h – focal depth (km)
α – absorption coefficient (km^{-1})
$\log(e)$ = 0.4343.

This model assumes that the energy of seismic waves declines owing to the geometrical spreading and to the absorption of the traversed media. Based on this idea the unknown h and α parameters in Eq. (1) can be estimated using a nonlinear least-square method (Bevington 1969). In the actual calculations BECS Fortran computer program (Tarcsai 1991) was used. In different seismological compilations (e.g. Macroseismische Beobachtungen, Wien) and catalogues (e.g. Ribaric 1982) published in countries of the Carpathian region usually a simplified version of the Kövesligethy method, the Blake formula (Blake 1941) is used. In the Hungarian earthquake catalog (Zsíros et al. 1988) covering the whole Carpathian Basin, the focal depth values are based either on the Kövesligethy model or on the Blake formula. However, the intensity decay of earthquakes has a better fit to the Kövesligethy model against to the Blake one (Kárnik and Algermissen 1978). Earlier, based on the Kövesligethy model, the focal depths of the central part of the Carpathian Region were already estimated (Zsíros et al. 1989) but the process used was not

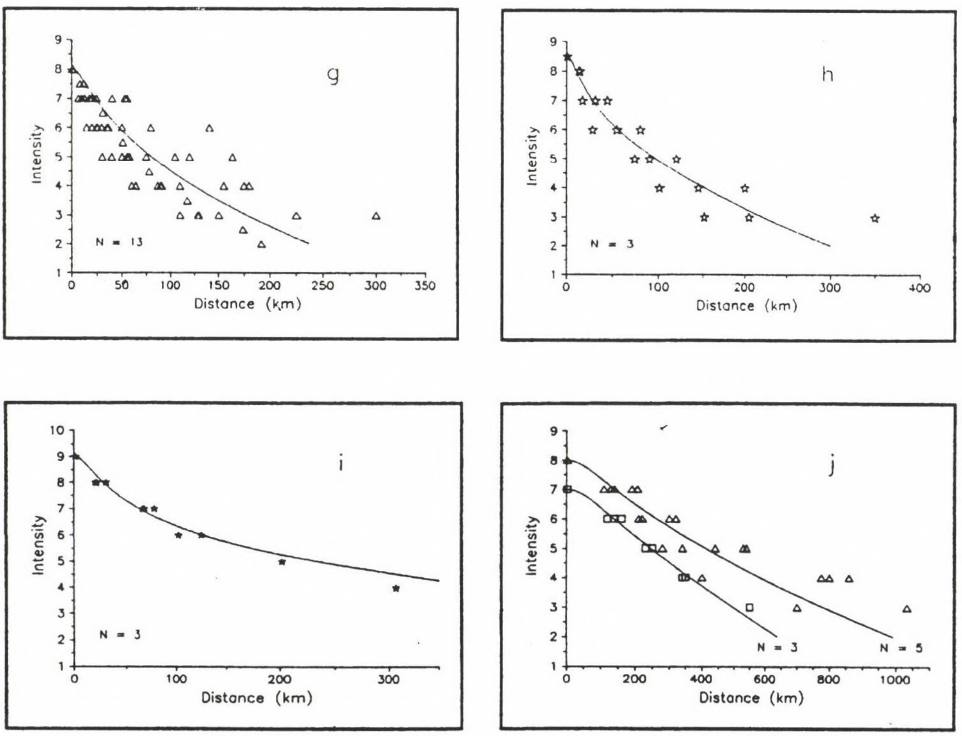


Fig. 3b. g-i: Intensity attenuation of earthquakes with different epicentral intensities in the Carpathian Basin (44.5–50.0N; 15.0–25.0), j: Intensity attenuation of intermediate focal depth earthquakes from the Háromszék-Vrancea region. N – number of events used

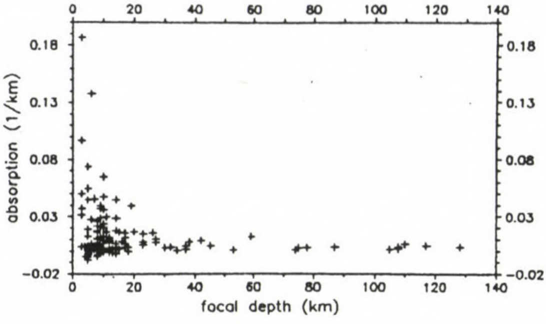


Fig. 4. Absorption coefficient (α) versus focal depth (h) estimated from 124 earthquakes in the Carpathian region (44.5–50.0N; 15.0–27.5E)

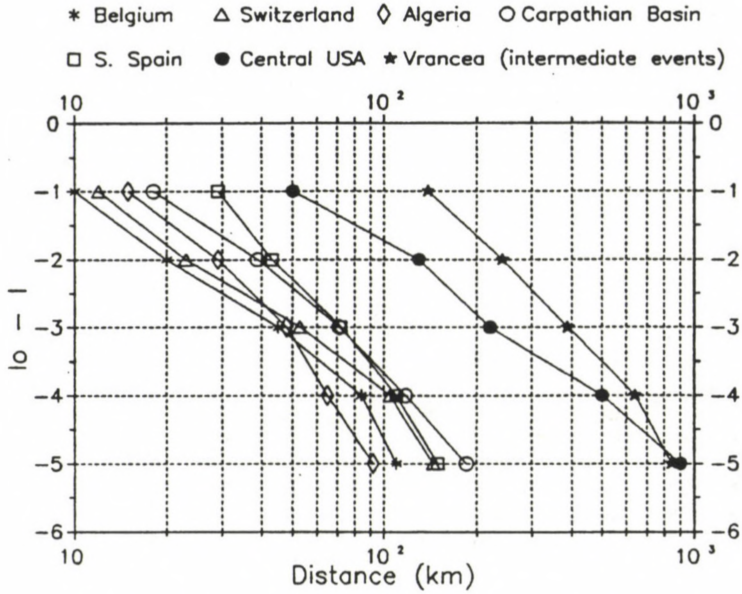


Fig. 5. Examples of intensity attenuations with distance

a non-linear one, which is needed for the solution of the non-linear equation of Kövesligethy.

Results and conclusions

The results of the depth estimations summarized in Table I show that 65 percent of the all 124 studied events occurred in the 12 km upper part of the crust and deeper than 60 km focal depth earthquakes are located only in the Háromszék-Vrancea intermediate focal zone. The frequency distribution of the depth values in

Table I. Results of macroseismic depth estimations of earthquakes occurred in the Carpathian region (44.5–50.0N; 15.0–27.5E)

Date	Time	Epicentre	I_o	h	σ_h	α	σ_α	n
1443 Jun. 05		48.8N 18.6E	8.0	37.	29.	.0044	.0076	3
1590 Sep. 15		48.1N 16.1E	9.0	18.	0.	.0034	.0001	5
1763 Jun. 28	(05:28)	47.7N 18.2E	8.5	6.	1.	.0066	.0018	5
1783 Apr. 22	(04:45)	47.7N 18.2E	8.0	16.	1.	.0083	.0005	3
1786 Feb. 27	(03)	49.7N 18.5E	7.5	38.	4.	.0080	.0008	5
1786 Dec. 03	(16)	49.7N 20.0E	7.5	42.	11.	.0088	.0024	5
1790 Apr. 06	(19:29)	45.7N 26.6E	8.0	74.	11.	.0009	.0003	4
1810 Jan. 14	(17:09)	47.4N 18.2E	8.0	8.	1.	.0176	.0018	3
1834 Oct. 15	(06)	47.6N 22.3E	9.0	13.	6.	−.0011	.0052	3
1838 Jan. 23	(18:45)	45.7N 26.6E	8.0	105.	12.	.0013	.0002	4
1858 Jan. 15	(19:15)	49.2N 18.7E	7.5	7.	0.	.0011	.0004	5
1876 Jul. 17	(12:17)	48.0N 15.2E	7.5	10.	2.	.0063	.0035	5
1880 Oct. 03	(05:18)	46.3N 24.1E	7.0	17.	2.	.0066	.0011	4
1880 Nov. 09	(06:33)	45.9N 16.1E	9.0	13.	5.	.0019	.0055	3
1884 Mar. 24	(19:59)	45.3N 18.4E	7.0	9.	2.	.0032	.0048	3
1885 Apr. 30	(23:15)	47.3N 15.4E	8.0	6.	1.	.0062	.0020	5
1888 Aug. 16	(04:25)	47.6N 18.2E	5.0	9.	1.	.0293	.0043	3
1893 Feb. 24	(17:35)	48.6N 17.8E	6.0	3.	1.	.0317	.0254	3
1893 Mar. 11	(09:25)	48.0N 23.0E	6.0	9.	1.	.0402	.0053	3
1893 Aug. 17	(15:35)	45.7N 26.6E	7.0	108.	17.	.0032	.0006	3
1894 Mar. 04	(06:35)	45.7N 26.6E	6.5	128.	75.	.0035	.0025	3
1899 Jun. 11	(00:30)	47.9N 16.4E	6.0	5.	0.	.0131	.0046	3
1901 Feb. 16	(20:06)	46.1N 15.0E	6.5	10.	1.	.0029	.0012	4
1901 Apr. 02	(16:47)	45.5N 20.7E	7.0	12.	2.	.0012	.0016	4
1901 Dec. 17	(14:12)	45.8N 16.0E	7.0	6.	0.	.0055	.0014	4
1902 Oct. 24	(17:52)	45.9N 16.0E	6.0	10.	0.	.0030	.0008	3
1902 Dec. 17	(15:20)	46.0N 15.1E	6.0	11.	8.	.0170	.0172	3
1903 Jun. 26	(04:28)	47.9N 20.4E	6.0	10.	3.	.0243	.0146	3
1903 Sep. 13	(08:02)	45.7N 26.6E	6.5	32.	7.	.0034	.0016	4
1904 Feb. 06	(02:49)	45.7N 26.6E	6.0	75.	9.	.0022	.0006	3
1905 Apr. 28	(21:04)	46.1N 16.2E	6.0	8.	1.	.0056	.0039	3
1905 May 23	(13:13)	45.9N 15.3E	7.0	10.	2.	.0477	.0109	4
1905 May 29	(11:16)	46.2N 16.2E	6.5	8.	2.	.0075	.0052	4
1905 Dec. 17	(22:16)	45.9N 16.1E	7.5	7.	0.	.0038	.0004	4
1906 Jan. 02	(04:26)	45.9N 16.1E	8.0	10.	1.	.0028	.0008	5
1906 Jan. 09	(23:05)	48.6N 17.5E	8.0	10.	2.	.0132	.0068	3
1906 Jan. 16	(02:52)	48.6N 17.5E	7.5	8.	1.	.0263	.0083	4
1906 Mar. 04	(11:38)	44.6N 15.4E	6.0	12.	2.	.0085	.0047	3
1906 Apr. 29	(09:15)	47.3N 22.2E	6.0	14.	3.	.0179	.0057	3
1908 Feb. 19	(21:11)	48.0N 16.7E	7.0	6.	0.	.0003	.0004	4
1908 May 28	(08:27)	46.9N 19.7E	7.0	12.	1.	.0119	.0015	5
1908 Oct. 06	(21:40)	45.5N 26.5E	8.0	78.	7.	.0027	.0003	4
1909 Oct. 08	(09:59)	45.4N 16.2E	8.5	12.	1.	.0073	.0006	6
1909 Oct. 10	(05:55)	45.3N 16.3E	5.5	30.	3.	.0027	.0012	3
1910 Jan. 28	(23:58)	45.4N 16.2E	7.0	15.	4.	.0018	.0023	4
1910 Jan. 29	(00:12)	45.4N 16.2E	6.0	18.	6.	−.0001	.0028	3
1910 May 11	(20:18)	47.7N 16.0E	6.5	8.	2.	.0017	.0032	3
1911 Jul. 08	(01:02)	46.9N 19.7E	8.0	12.	1.	.0098	.0009	6
1912 May 25	(18:02)	45.7N 27.2E	7.0	87.	8.	.0034	.0005	3
1912 Sep. 19	(22:15)	46.2N 16.9E	6.0	5.	0.	−.0053	.0011	3
1913 Jul. 23	(22:03)	45.7N 26.6E	5.5	59.	8.	.0124	.0023	3

Table I (contd.)

Date	Time	Epicentre	I_o	h	σ_h	α	σ_α	n
1914 Mar. 24	(09:19)	45.7N 15.3E	6.0	10.	1.	.0653	.0077	4
1916 Jan. 26	(07:38)	45.5N 24.5E	8.0	23.	6.	.0051	.0018	5
1916 Jul. 14	(20:27)	45.1N 15.0E	6.5	17.	2.	.0034	.0012	4
1917 Jan. 29	(08:23)	45.9N 15.6E	8.0	5.	0.	.0030	.0009	5
1919 Sep. 05	(20:37)	46.5N 16.0E	6.0	9.	1.	−.0017	.0015	3
1921 Jan. 15	(10:55)	45.7N 20.1E	7.0	26.	2.	.0161	.0014	4
1922 Nov. 24	(02:15)	45.6N 18.7E	7.5	10.	5.	−.0006	.0050	3
1924 Aug. 12	(16:27)	45.8N 18.8E	7.0	12.	2.	.0074	.0024	3
1925 Jan. 31	(07:05)	47.9N 20.4E	7.5	5.	0.	.0549	.0049	3
1926 May 19	(10:11)	44.5N 20.6E	6.0	12.	2.	−.0013	.0025	3
1926 Jun. 28	(20)	48.0N 23.7E	5.5	5.	0.	.0050	.0044	3
1926 Aug. 10	(01:10)	48.0N 23.7E	6.5	9.	1.	.0373	.0094	3
1926 Sep. 28	(15:42)	47.7N 16.0E	7.0	5.	0.	−.0019	.0014	4
1927 Feb. 20	(06:48)	46.3N 16.9E	6.5	9.	2.	.0125	.0083	3
1927 Jul. 25	(21:35)	47.5N 15.5E	7.0	8.	1.	−.0046	.0012	3
1928 Mar. 18	(23:50)	45.4N 17.2E	7.0	10.	1.	.0086	.0023	4
1928 Aug. 25	(21:09)	45.9N 15.6E	7.0	5.	2.	.0005	.0109	3
1929 Nov. 05	(07:50)	46.8N 19.7E	5.0	19.	4.	.0399	.0147	3
1930 Mar. 05	(23:55)	48.7N 17.6E	7.0	6.	1.	.0053	.0037	3
1930 Aug. 22	(05:49)	48.0N 19.4E	6.0	11.	1.	.0303	.0031	3
1931 Apr. 07	(01:35)	48.2N 22.5E	6.0	6.	0.	.0277	.0017	3
1934 Mar. 29	(20:07)	45.8N 26.5E	8.0	110.	22.	.0059	.0013	4
1936 Sep. 06	(04:49)	45.7N 21.1E	7.0	5.	1.	.0189	.0032	4
1937 Apr. 04	(15:40)	45.3N 18.0E	7.0	8.	1.	.0112	.0028	4
1937 Jun. 10	(01:43)	48.1N 21.4E	6.0	8.	0.	.0221	.0015	3
1938 Mar. 27	(11:16)	46.0N 16.9E	8.0	17.	1.	.0123	.0013	4
1938 Nov. 08	(03:12)	47.9N 16.4E	7.0	5.	0.	−.0082	.0006	3
1939 Mar. 23	(05:44)	47.4N 22.0E	6.0	23.	3.	.0152	.0028	3
1940 Jun. 24	(09:57)	45.8N 26.5E	5.5	34.	1.	.0005	.0003	3
1940 Oct. 22	(06:37)	45.8N 26.4E	7.0	117.	12.	.0045	.0006	3
1940 Nov. 10	(01:39)	45.8N 26.8E	9.0	53.	11.	.0009	.0003	6
1940 Dec. 17	(10:53)	45.1N 18.2E	6.0	27.	2.	.0104	.0011	3
1949 Mar. 14	(12:46)	45.7N 15.3E	6.0	17.	0.	.0160	.0007	3
1950 Aug. 31	(17:22)	44.9N 17.4E	8.0	3.	0.	.0038	.0005	5
1951 Feb. 20	(00:14)	48.0N 19.1E	6.5	15.	2.	.0166	.0025	4
1953 Sep. 13	(08:02)	47.0N 17.2E	6.5	5.	1.	.0449	.0122	4
1953 Oct. 01	(18:27)	46.0N 15.5E	7.0	3.	0.	.0040	.0039	4
1956 Jan. 12	(05:46)	47.4N 19.1E	8.0	10.	1.	.0097	.0011	6
1956 Mar. 31	(14:07)	47.0N 17.0E	6.0	10.	1.	.0363	.0078	3
1956 May 02	(11:48)	47.0N 17.0E	5.0	7.	1.	.0460	.0094	3
1956 Dec. 14	(00:12)	47.9N 20.3E	5.5	14.	3.	.0451	.0130	3
1957 Sep. 22	(14:44)	45.7N 21.1E	5.5	5.	1.	.0739	.0322	3
1959 May 27	(20:38)	45.7N 21.1E	7.5	4.	0.	.0039	.0018	5
1959 May 31	(12:16)	45.8N 27.4E	6.0	45.	1.	.0044	.0003	3
1959 Dec. 02	(18:20)	44.6N 15.4E	7.0	12.	0.	.0086	.0006	3
1960 Oct. 22	(19:18)	45.6N 21.2E	6.0	14.	1.	.0066	.0012	3
1963 Dec. 02	(06:49)	47.9N 16.4E	6.0	20.	1.	.0170	.0018	4
1964 Apr. 13	(08:30)	45.3N 18.1E	8.0	27.	2.	.0075	.0007	4

Table I (contd.)

Date	Time	Epicentre	I_o	h	σ_h	α	σ_α	n
1965 Oct. 24	(06:27)	48.2N 22.7E	7.0	3.	0.	.0967	.0138	4
1966 Nov. 11	(16:16)	45.3N 16.0E	7.0	8.	1.	.0164	.0021	4
1967 Jul. 29	(13:35)	48.0N 23.5E	5.5	3.	1.	.1869	.0921	3
1968 Dec. 03	(20:58)	44.6N 18.6E	6.5	8.	1.	.0103	.0030	4
1969 Apr. 12	(20:39)	45.2N 25.0E	6.0	10.	1.	.0004	.0021	3
1969 Oct. 26	(15:37)	44.8N 17.3E	7.5	23.	3.	.0076	.0013	5
1969 Oct. 27	(08:11)	44.8N 17.2E	8.5	14.	2.	.0034	.0007	6
1969 Dec. 31	(13:19)	44.7N 17.2E	7.0	12.	1.	.0028	.0011	4
1972 Jan. 05	(04:58)	47.8N 16.3E	6.0	8.	2.	.0074	.0073	3
1972 Apr. 16	(10:10)	47.8N 16.2E	7.5	9.	0.	–.0006	.0001	5
1972 Apr. 16	(11:05)	47.7N 16.2E	6.5	5.	0.	–.0034	.0014	3
1974 Apr. 17	(01:31)	46.0N 21.1E	6.0	14.	1.	.0288	.0046	3
1976 Jan. 14	(11:53)	49.1N 24.0E	5.5	3.	0.	.0377	.0100	3
1976 Feb. 07	(20:46)	49.0N 24.0E	6.0	3.	0.	.0504	.0048	3
1977 Mar. 04	(19:21)	45.8N 26.7E	8.0	108.	19.	.0017	.0003	5
1977 Sep. 25	(08:25)	48.3N 22.7E	6.5	6.	0.	.1379	.0066	4
1979 Mar. 08	(01:20)	47.7N 23.5E	5.5	10.	0.	.0207	.0024	3
1979 Mar. 30	(15:56)	47.7N 23.3E	6.5	9.	2.	.0218	.0067	4
1979 May 12	(21:34)	47.3N 15.2E	6.0	5.	1.	–.0006	.0028	3
1983 Apr. 14	(14:52)	47.7N 15.0E	6.5	8.	2.	–.0035	.0021	4
1985 Jun. 05	(23:54)	48.0N 16.3E	5.5	4.	1.	–.0045	.0144	4
1985 Aug. 15	(04:29)	47.1N 18.1E	7.0	10.	2.	–.0004	.0016	4

I_o – epicentral intensity
h – focal depth (km)
α – absorption coefficient (km^{-1})
s – standard deviation
n – number of isoseismals

Table II. Average focal depth and absorption coefficient of earthquakes with epicentral intensities 5° – 9° in the Carpathian Basin (44.5–50.0N; 15.0–25.0E)

I_o	h (km)	α (km^{-1})	n
5°	12±6	0.0384±.0085	3
5° – 6°	10±9	0.0406±.0607	9
6°	11±6	0.0174±.0170	28
6° – 7°	9±3	0.0223±.0377	13
7°	10±5	0.0103±.0205	26
7° – 8°	15±13	0.0109±.0164	11
8°	14±10	0.0079±.0047	12
8° – 9°	11±4	0.0058±.0021	3
9°	15±3	0.0014±.0023	3

I_o – epicentral intensity
h – focal depth
α – absorption coefficient
n – number of earthquakes

Table III. Average focal depth and absorption coefficient of earthquakes with epicentral intensities 7° and 8° in the Háromszék-Vrancea intermediate focal region

I_o	h (km)	α (km^{-1})	n
7°	104±15	0.0037±.0007	3
8°	95±17	0.0025±.0020	5

I_o – epicentral intensity
h – focal depth
α – absorption coefficient
n – number of earthquakes

Table IV. Average absorption coefficient for different focal depth intervals of earthquakes in the Carpathian region (44.5–50.0N; 15.0–27.5E)

h (km)	α (km^{-1})	n
1– 5	0.0275±.0457	22
6– 10	0.0161±.0241	46
11– 15	0.0106±.0120	21
16– 20	0.0119±.0120	9
21– 60	0.0068±.0049	16
61–130	0.0029±.0015	10

h – focal depth
α – absorption coefficient
n – number of earthquakes

the Carpathian Basin (44.5–50.0N; 15.0–25.0E) is presented in Fig. 1, where the most frequent depth interval is the 9–12 km depth domain, and the estimation yields an average depth of 11 (±7) km with an average absorption of 0.016 (±0.027) km^{-1}. The epicentre distributions of earthquakes used in this study are shown in Fig. 2, in which the focal depths are classified into 8 different depth ranges. According to this estimations the Sub-Carpathian region (47.5–48.5N; 22.5–24.0E) is the most shallow ($h \leq 10$ km) source zone and there are no earthquakes with deeper than 20 km focal depth in the central part of the Pannonian Basin (practically in the present territory of Hungary). The intensity attenuation curves of earthquakes with different epicentral intensities are presented in Fig. 3(a-j) using the average depth and absorption values (see Table II and III) in the Kövesligethy model. Though the absorption coefficient (a) is characteristic of the media traversed by the earthquake waves a dependence of the absorption from the epicentral intensity can be noticed in Table

II. The main reason of this phenomenon, however, the different distribution of the weaker ($I_o < 6°$) and the stronger ($I_o \geq 8°$) events in space. The Sub-Carpathian earthquakes are especially characterised by high intensity attenuation. A secondary factor might be the somewhat deeper focuses of the stronger earthquakes, that is the traversed media are different in the cases of weaker and stronger events. The input (isoseismal) data are also plotted in the figures to demonstrate the scattering of intensities. The distribution of the absorption coefficient (a) against the focal depth (h) in Fig. 4 shows considerable scattering in the upper 20 km crust in contrast to the deeper depth domain. It must be noted, however, that the number of data in the deeper zone is much less. For different depth domains the average absorptions are summarised in Table IV for the Carpathian region. In Fig. 5 some available observed intensity attenuations (Gupta and Nuttli 1976, Mayer-Rosa 1993) are compared with the intensity declines in the Carpathian region. According to this comparison the intensity attenuation in the Carpathian Basin is close to that experienced in South Spain.

Acknowledgments

I thank Dr. György Tarcsai (Eötvös University, Budapest) for allowing me to use his BECS program in the computations and Dr. Antal Ádám for the critical remarks on the paper. This research was partially supported by the Copernicus CT94-0238 project.

References

Bevington P R 1969: Data reduction and error analysis for the physical sciences. McGraw-Hill, New York

Blake A 1941: *Bull. Seism. Soc. Am.*, 31, 225–231.

Cvijanovic D 1970: *Geofizicka Istrazivanja*, Series C, No. 10–11, 71–102.

Cvijanovic D, Zivcic M, Krstulovic B 1989: *Acta Seismologica Iugoslavica*, 15, 57–68 (in Croatian).

Drimmel J, Trapp E 1982: Die Erdbeben Österreichs 1971–1980. Springer-Verlag, Wien

Gupta I N, Nuttli O W 1976: *Bull. Seism. Soc. Am.*, 66, 743–751.

Gutdeutsch R, Hammerl Chr, Mayer I, Vocelka K 1987: Erdbeben als historisches Ereignis. Die Rekonstruktion des Bebens von 1590 in Niederösterreich. Springer-Verlag, Wien

Kárnik V, Algermissen S T 1978: In: The assessment and mitigation of earthquake risk, UNESCO, 11–47.

Kárnik V, Procházkova D, Broucek I 1981: *Travaux Géophysiques*, 29, 155–186.

Karpiv T S, Kostyuk O P, Pronisin R S, Rudenskaya I M 1980: In: Seismological Bulletin, Kiev, 161–189 (in Russian).

Kiss Z 1983: Isoseismal maps of Hungarian earthquakes (1880–1980). Seismological Observatory, Budapest

Kostyuk O P, Rudenskaya I M 1982: In: Seismological Bulletin, Kiev, 90–123 (in Russian).

Kövesligethy R 1906: *Mathematikai és Természettudományi Értesítő*, 24, 349–368 (in Hungarian).

Mayer-Rosa D 1993: *Annali di Geofisica*, 34, 93–102.

Polonic G 1980: *Géophysique*, 24, 255–268.

Procházkova D, Kárnik V eds 1978: Atlas of isoseismal maps, Central and Eastern Europe. Geophysical Institute, Prague

Ribaric V 1982: Seismicity of Slovenia - Catalogue of earthquakes (792 A.D.-1981)

Rudenskaya I M, Sluka N T, Skarzevskij V V 1983: In: Seismological Bulletin, Kiev, 59–72 (in Russian).

Shebalin N V, Kárnik V, Hadzievski D eds 1974: UNDP/UNESCO survey of seismicity of the Balkan region, Catalogue of earthquakes, part III, Atlas of isoseismal maps. UNESCO, Skopje

Szeidovitz Gy 1984: *Földtani Kutatás*, 27, 71–80 (in Hungarian).

Tarcsai Gy 1991: BECS Fortran computer program for parameter estimation of non-linear equations. Eötvös University, Budapest

Zátopek A 1940: Earthquake Observations in Slovakia and in the Sub-Carpathian Region 1923–1939. Geophysical Institute, Praha (in Czech)

Zsíros T 1983a: *Acta Geod. Geoph. Mont. Hung.*, 18, 129–134.

Zsíros T 1983b: Hungarian Earthquake Catalogue 1978–1982. MTA Geodéziai és Geofizikai Kutató Intézet, Budapest (in Hungarian)

Zsíros T 1988: *Acta Geod. Geoph. Mont. Hung.*, 23, 389–396.

Zsíros T 1989: Macroseismic observations in Hungary (1983–1988). Seismological Observatory, Budapest

Zsíros T 1995: The Central Transylvania earthquake of 1880. Manuscript. Seismological Observatory, Budapest

Zsíros T, Mónus P, Tóth L 1988: Hungarian Earthquake Catalog (456–1986). Geodetic and Geophysical Research Institute, Budapest

Zsíros T, Mónus P, Tóth L 1989: *Gerlands Beitr. Geophysik*, 98, 146–154.

Acta Geod. Geoph. Hung., Vol. 31(1-2), pp. 127-144 (1996)

ESTIMATION OF VELOCITY PERTURBATIONS USING AMPLITUDE AND PHASE FLUCTUATIONS FOR EARTHQUAKES RECORDED AT GAURIBIDANUR SEISMIC ARRAY IN SOUTHERN INDIA

JAYANT N TRIPATHI[1] and AVADH RAM[2]

[Manuscript received November 1, 1995]

The Chernov medium approach has been applied for analysis of 24 earthquakes recorded at Gauribidanur Seismic Array (GBA) in southern India. Correlation coefficients of log-amplitude and phase fluctuations between every possible pair of seismometers are used to determine the correlation length of the medium. The Chernov theory can be well applied for the observed P-wave velocity perturbations at 0.5 Hz. Thus, the average correlation length for the Gaussian correlation function is determined to be about 18 km. The estimated velocity perturbations for GBA are around 1.4 % – 2.3 %, which is slightly greater than the corresponding estimates by previous workers for the array while the extension of the medium is around 660 km. The scattering coefficient under GBA at 0.5 Hz is $9.0 - 22.7 \times 10^{-4}$ km^{-1}. The Born approximation condition is violated for frequencies higher than 0.5 Hz due to the strong scattering coefficient for the region under GBA.

Keywords: amplitude and phase fluctuations; Chernov theory; correlation length; Gauribidanur Seismic Array; lithospheric structure; seismic scattering

1. Introduction

The observed irregular spatial variations of amplitude and travel times of seismic waves at the earth surface limit the accuracy of seismic probes, which are used for studying the Earth's interior and earthquake source mechanisms. Scattering of seismic waves is a common cause of these variations. Hence, understanding of the physical mechanism of scattering of seismic waves will provide an accurate evaluation of the capability of a seismic probe and an insight into the nature of inhomogeneities in the Earth's crust and upper mantle.

The Gauribidanur array located in southern India (Fig. 1) provides an opportunity to study the scattering mechanism for short period P-waves. Example signal variations across the array are shown in Fig. 2 for four earthquakes. The variation of signal amplitude results not only from the heterogeneities beneath the array and its seismometer sites but also from the different azimuths of the waves recorded across the array. Thus, the Chernov (1960) random media approach may be applied. The essence of this random media analysis is to provide an estimate of the root mean square (r.m.s.) variation of the velocity perturbations of the crust and

[1]Department of Earth and Planetary Sciences, University of Allahabad, Allahabad 211002, India

[2]Department of Geophysics, Banaras Hindu University, Varanasi 221005, India

upper mantle in the vicinity of the array, the extent of the inhomogeneous medium and the correlation length; the correlation length is a measure of the wavelength of the structural anomalies. Thus, the structural heterogeneities may be statistically described by Chernov Theory (1960) for the region.

This stochastic approach is desirable for obtaining the statistical properties of the heterogeneities when the medium beneath an array has multi-scale and complex heterogeneities (Nikolaev 1972, Aki 1973, Berteussen et al. 1975, 1977, Flattê and Wu 1988, Wu and Flattê 1990). Nikolaev (1972) inferred the velocity perturbations from deep seismic sounding data of a sampled region. The first statistical study of the data from the Large Aperture Seismic Array (LASA) was done by Aki (1973). Although the variance transverse correlation fluctuations of the phase and log-amplitude of the direct P-arrivals were used by the earlier workers (Aki 1973

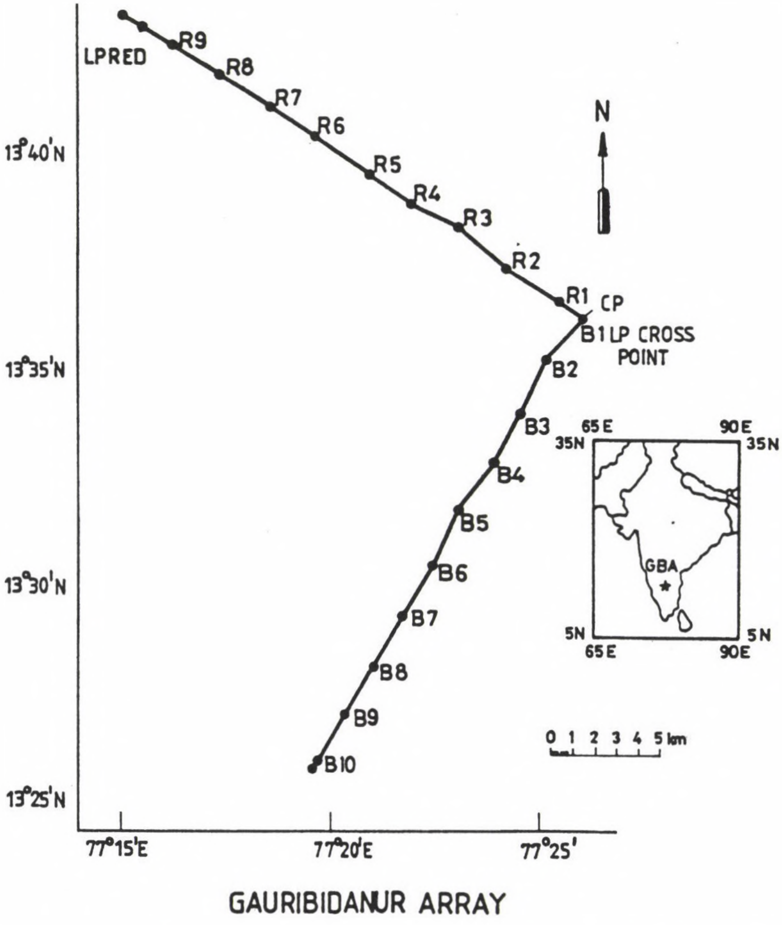


Fig. 1. Location and configuration of the Gauribidanur medium apperture seismic array (array configuration after Mowat and Burch 1974)

and Capon 1974 for LASA; Berteussen et al. 1975 for NORSAR and LASA and Berteussen et al. 1977 for GBA), the phase fluctuation information was either transformed into slowness functions (Capon 1974) or into time residuals (Berteussen et al. 1975, 1977). The stochastic model of the heterogeneous media was assumed to be a uniform isotropic random media with a Gaussian correlation function characterized by the relative velocity perturbations and average scale length a. Recently, transverse coherence functions and angular coherence functions have been used for an estimation of transmission fluctuations (Flattê and Wu 1988, Wu and Flattê 1990, Flattê and Moody 1990, Flattê et al. 1991, Flattê and Xie 1992, Tripathi and Ram 1995).

In this paper, phase fluctuation information have been directly used, instead of slowness functions or time residuals, in a Chernov random media approach for a medium-sized array (GBA) and an estimation made of the extent of the structural inhomogeneities in the vicinity of GBA. In this analysis, 24 events, having a reasonable azimuthal coverage, were considered. Significant information about the heterogeneities beneath the GBA have been obtained. Finally, these results are compared with results from similar investigations for other arrays.

2. The Gauribidanur Seismic Array (GBA)

The Gauribidanur Seismic Array (GBA), which is located in southern India, was sponsored by the U.K. Atomic Energy Authority (UKAEA) in the early sixties with the cooperation of the Bhabha Atomic Research Centre (BARC), Government of India. The array became operational in October, 1965. The array location and configuration are shown in Fig. 1. The array is L-shaped with each line (called the Blue and Red lines) containing 10 short-period (1 s) vertical-component Willmore Mk II seismometers spaced at approximately 2.5 km, resulting in an overall length of about 25 km for each line. The seismometer vaults are set in Archean rocks with unweathered gneiss lying within 2 meters of the surface over most of the region. The output of each seismometer is telemetered to a central recording laboratory where it is recorded on a separate channel of a 24-channel FM magnetic tape. A detailed description of the GBA was given by Ram and Mereu (1977) and Varghese et al. (1979).

3. Geology of the Gauribidanur Array Region

The Indian Peninsula is a triangular plateau, elevated 250–850 m (average elevation 750 m), with a vertex extending far into the Indian Ocean. The area around the Gauribidanur array is surrounded by ranges towards the east and south. The rocks beneath the array are gneissic granite of Archean age, and their general foliation trend is NNW–SSE. A thin layer of soil varying in thickness from 1.5 to 4.5 meters covers the area. Basement is composed of highly-folded crystalline schists and Archean gneisses intruded by granite and granite gneiss with wide spread formation of gneous metamorphosed rocks. Exposed basement rocks in north-eastern and southern parts of the Indian peninsula make up two-thirds of the shield area, believed to be Precambrian. The cover sediments of the platform in the north-western Paninsular regions are known as the Deccan Traps.

Table I. List of earthquakes used in the present study

Event No.	Date	Orig. Time h	m	s	Location lat	long	FD(h) km	Delta (Δ)	Az.	Mag.
1	02.09.1980	16	39	14.3	17.27°N	73.77°E	33	5.08	317.0	4.7b
2	05.04.1984	14	32	57.6	29.4 °N	81.06°E	33	16.04	11.4	4.1b
3	09.10.1981	14	31	01.6	17.82°N	85.54°E	0	8.86	60.2	4.7b
4	10.08.1981	10	58	24.4	31.10°N	77.82°E	33	17.42	1.1	4.4b
5	22.12.1980	04	36	08.0	26.7 °N	89.6 °E	33	17.28	39.1	4.4b
6	23.10.1981	23	44	44.5	29.89°N	94.93°E	0	22.9	42.0	5.1b
7	11.07.1972	05	54	44.1	36.38°N	70.73°E	206	23.47	346.3	5.4b
8	11.07.1972	20	46	11.6	24.29°N	94.70°E	48	19.46	54.4	4.6
9	08.04.1972	5	23	49.1	5.08°N	61.99°E	33	17.42	242.4	4.9
10	08.04.1972	6	42	13.3	29.67°N	89.47°E	33	19.49	32.9	4.8
11	09.04.1972	8	43	48.5	40.13°N	78.83°E	14	28.45	2.4	4.7
12	10.04.1972	22	47	36.8	35.13°N	71.61°E	49	21.58	353.7	4.9
13	10.04.1972	14	35	34.3	28.28°N	53.08°E	33	26.95	306.6	4.6
14	11.04.1972	6	0	4.6	37.37°N	62.00°E	33	27.38	332.5	4.9
15	16.04.1972	1	27	55.7	1.84°N	99.83°E	33	26.99	123.0	5.3
16	17.04.1972	2	24	49.3	33.96°N	72.88°E	45	20.69	349.2	4.8
17	17.04.1972	10	35	45.7	25.58°N	95.63°E	108	20.64	52.0	5.0
18	17.04.1972	15	12	43.5	31.94°N	59.34°E	44	24.66	320.7	4.5
19	21.04.1972	21	19	29.5	34.99°N	81.03°E	33	21.54	8.1	4.8
20	22.04.1972	13	17	57.7	17.41°N	94.31°E	33	16.69	74.7	4.8
21	27.04.1972	5	51	6.3	0.58°N	99.64°E	54	26.13	120.9	5.3
22	28.04.1972	0	52	56.8	31.26°N	84.90°E	33	18.87	20.1	5.1
23	21.08.1972	21	57	17.2	8.28°N	105.95°E	62	35.69	125.9	5.3
24	03.09.1972	16	48	28.8	35.98°N	73.42°E	36	22.57	351.5	6.3

Orig. Time = Origin Time FD = Focal Depth
Az = Azimuth Mag. = Magnitude

4. Crustal structure beneath the array

Several studies have revealed that GBA is situated on mainly homogeneous structure (Corbishley 1970, Ram and Mereu 1977). Using the local earthquake data Arora (1971) proposed a two layer crustal model in the vicinity of the array with top an upper granitic layer 16 km thick over a second layer 19 km thick, i.e. with the Moho at 35 km depth. The P-wave velocities were found to be 5.7 and 6.5 $km \cdot s^{-1}$ for these two layers above mantle of 8.0 $km \cdot s^{-1}$. In an earlier study, Arora et al. (1970) used a series of local earthquakes recorded at GBA to delineate the crustal structure of southern India. They observed that Moho-refracted phase from approximately the NE direction emerges at smaller angles compared to that for the other directions. This led them to propose non-uniform intermediate layers or a dipping Moho (about 4°) beneath GBA. However, Berteussen et al. (1977) pointed out that time residuals at GBA could be explained satisfactorily by a dipping plane coincident with the Moho (dip 6° and strike N 13°E beneath the array.

5. Seismological data

The 24 earthquakes used in this analysis are listed in Table I. The data were obtained from the GBA and from Ram (1976). The selection was made on the basis of a best possible distance and azimuth coverage. The corresponding P-wave records were also all characterized by good signal to noise ratios.

We calculated the Fourier amplitude and phase spectra for frequencies 0.4 to 1.5 Hz at an interval of 0.1 Hz and also between 3 and 5 Hz. Thus, we excluded those frequencies at which contamination by noise was expected or for which the spectral noise was prominent. Because the results at frequencies 0.1 Hz apart are not independent, we included only those frequencies which were more than 0.2 Hz apart. Our final data consisted of 192 sets for 24 earthquakes, each set consisting of 20 amplitude and 20 phase delay measurements. The obtained root-mean-square log-amplitude and phase fluctuations at different frequencies are given in Table II.

6. Statistical properties of amplitude and phase fluctuations

We calculated the variances of phase delay (ϕ) and logarithm of amplitude ($\ln A$) and the correlation between them. The rms (root mean square) phase and logarithmic amplitude fluctuations at different frequencies are listed in Table II. The rms phase fluctuations (σ_0) ranged from 0.63 to 1.03 and the log amplitude fluctuations (σ_u) ranged from 0.24 to 1.08. In Chernov theory the ratio σ_u/σ_ϕ suggests the physical mechanism of the scattering. The observed ratios were less than unity, as expected from the theory, except for three cases. All of the observed correlation coefficients (192 out of 192) were positive, as also expected by Chernov theory.

The spatial correlation of amplitude and phase delay fluctuations are important because their correlation distances are directly related to the dimension of inhomogeneity of wave medium. The spatial autocorrelation for phase and log amplitude was calculated for all possible pairings of array seismometers (Fig. 3). If all 20 seismometers of the array recorded the earthquake then the total number of array pairs will be 190. Thus, with these 190 transverse autocorrelation coefficients for log-amplitudes and phase fluctuations we can determine 190 correlation lengths and thus find the best possible result with minimal uncertainty. Statistical properties of amplitude and phase fluctuations are discussed in detail by Chernov (1960), Aki (1973), Capon (1974), Berteussen et al. (1975, 1977).

7. Chernov Theory and its application for the Single Layer Model with a Gaussian Correlation Function

Acoustic wave scattering was considered by Chernov (1960), where the wave source is a monochromatic plane wave and the wave medium is characterized by random fluctuations of wave velocity α. In the past, the space domain formulation of Chernov (1960) has been used for fluctuation analysis (Aki 1973, Capon 1974, Capon and Berteussen 1974, Berteussen et al. 1975, 1977). Explicit results were derived for the Gaussian correlation function by Chernov, assuming a uniform and isotropic random layer of thickness R with Gaussian correlation function

$$N(r) = <\mu^2> \exp(-r^2/a^2) , \tag{1}$$

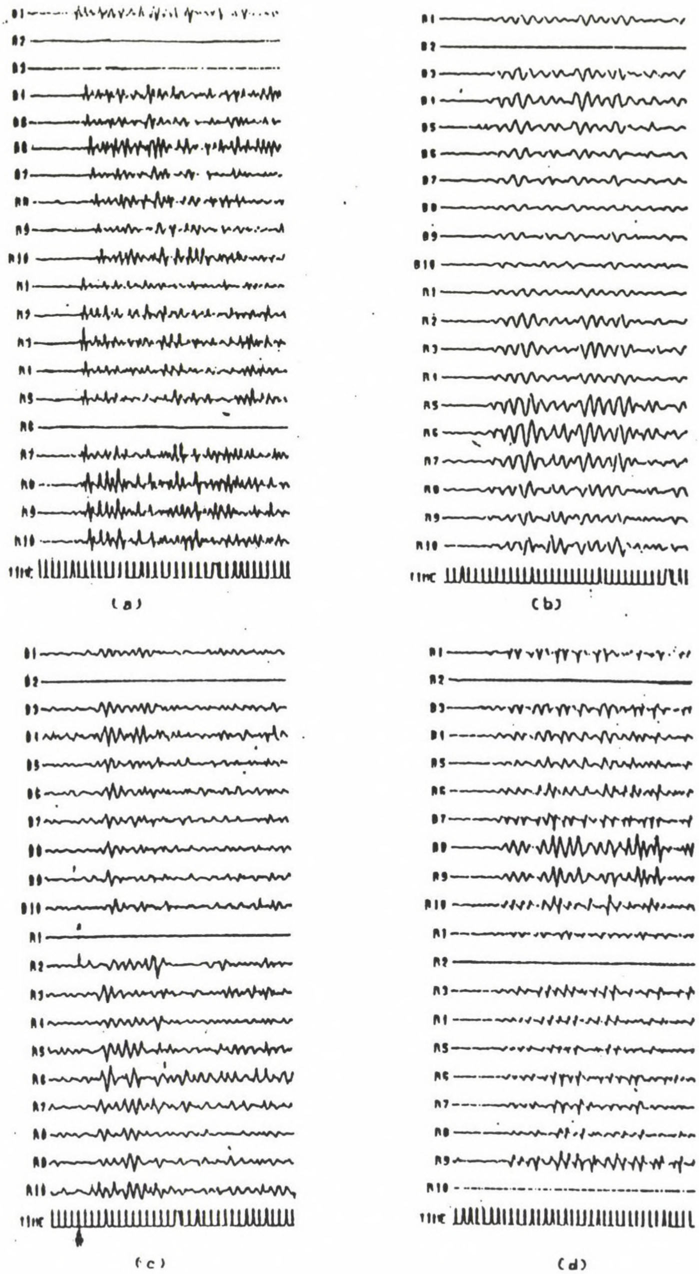


Fig. 2. Signal variations for some earthquakes recorded at GBA

Table II. Root mean square phase and log-amplitude fluctuations and correlation coefficients at different frequencies for the earthquakes shown in Table I. $D(\sigma A/\sigma\phi)$ and $D(\varrho)$ are the wave parameters obtained from $\sigma A/\sigma\phi$ and ϱ, respectively, using Eqs (6) and (5)

Event N.	Freq. Hz	σ_u	σ_ϕ	σ_u/σ_ϕ	ϱ	$D(\sigma_u/\sigma_\phi)$	$D(\varrho)$
1	.5	.50	.86	.58	.13	2.35	25.84
	.7	.46	.86	.54	.09	2.01	42.21
	.9	.45	.87	.52	.09	1.89	43.02
	1.0	.45	.88	.51	.08	1.85	45.48
	1.2	.45	.90	.50	.11	1.73	33.00
	1.5	.44	.92	.48	.10	1.62	34.05
	3.0	.41	.90	.46	.09	1.51	40.81
	5.0	.41	.91	.46	.09	1.50	40.51
2	.5	.76	.92	.83	.12	7.67	27.39
	.7	.74	.93	.79	.10	6.10	34.02
	.9	.63	.93	.68	.09	3.52	38.50
	1.0	.62	.93	.67	.07	3.37	56.84
	1.2	.64	.93	.69	.06	3.68	67.78
	1.5	.69	.91	.76	.09	5.21	40.64
	3.0	.42	.91	.46	.07	1.51	59.19
	5.0	.44	.90	.49	.13	1.70	24.47
3	.5	.44	.88	.50	.10	1.77	33.60
	.7	.43	.91	.47	.17	1.60	16.71
	.9	.44	.94	.47	.11	1.59	29.69
	1.0	.44	.87	.51	.09	1.83	38.77
	1.2	.44	.89	.50	.09	1.74	43.92
	1.5	.41	.89	.46	.09	1.53	40.12
	3.0	.43	.90	.48	.11	1.61	32.27
	5.0	.42	.94	.45	.10	1.45	35.33
4	.5	.83	.91	.91	.13	16.16	23.54
	.7	.81	.92	.87	.11	10.96	33.46
	.9	.84	.91	.92	.09	18.25	41.62
	1.0	.59	.91	.64	.10	3.01	34.71
	1.2	.87	.90	.97	.10	57.10	37.66
	1.5	.59	.91	.66	.08	3.17	51.28
	3.0	.70	.89	.79	.09	6.21	41.62
	5.0	.47	.90	.52	.10	1.89	34.93
5	.5	.82	.91	.90	.16	15.11	18.97
	.7	.69	.88	.78	.12	5.80	27.21
	.9	.67	.89	.75	.13	4.97	24.46
	1.0	.67	.89	.75	.08	4.84	52.16
	1.2	.61	.90	.68	.14	3.45	22.26
	1.5	.77	.92	.83	.09	8.06	41.28
	3.0	.44	.90	.49	.08	1.67	48.66
	5.0	.36	.93	.39	.13	1.18	24.37
6	.5	.46	.86	.49	.22	2.64	11.26
	.7	1.08	.91	1.19	.13	.00	25.93
	.9	1.07	.91	1.18	.07	.00	59.84
	1.0	1.07	.91	1.17	.09	.00	43.78
	1.2	.80	.93	.85	.08	10.12	47.79
	1.5	.74	.88	.84	.15	8.38	19.55
	3.0	.77	.93	.83	.09	7.71	42.93
	5.0	.43	.91	.48	.12	1.61	28.87

Table II (contd.)

Event N.	Freq. Hz	σ_u	σ_ϕ	σ_u/σ_ϕ	ϱ	$D(\sigma_u/\sigma_\phi)$	$D(\varrho)$
7	.5	.47	.91	.52	.12	1.86	26.60
	.7	.43	.93	.46	.08	1.54	45.78
	.9	.39	.90	.43	.13	1.38	23.44
	1.0	.44	.92	.48	.10	1.65	33.73
	1.2	.42	.86	.49	.13	1.69	23.96
	1.5	.45	.89	.50	.15	1.75	19.65
	3.0	.39	.90	.43	.12	1.38	27.62
	5.0	.42	.86	.49	.14	1.68	22.53
8	.5	.47	.93	.51	.17	1.82	16.53
	.7	.41	.88	.47	.12	1.57	29.30
	.9	.41	.88	.47	.19	1.55	13.82
	1.0	.40	.89	.45	.18	1.44	15.05
	1.2	.38	.91	.41	.14	1.28	22.47
	1.5	.44	.89	.49	.15	1.72	19.28
	3.0	.43	.93	.46	.16	1.51	18.54
	5.0	.36	.92	.39	.12	1.17	26.20
9	.5	.53	.95	.56	.22	2.17	11.41
	.7	.38	.88	.43	.17	1.36	17.00
	.9	.44	.87	.51	.19	1.84	13.65
	1.0	.54	.89	.60	.14	2.55	23.09
	1.2	.47	.89	.52	.09	1.92	44.66
	1.5	.40	.87	.47	.17	1.55	16.38
	3.0	.35	.90	.39	.15	1.16	19.97
	5.0	.33	.92	.36	.21	1.05	11.91
10	.5	.47	.90	.52	.13	1.91	23.86
	.7	.54	.91	.59	.11	2.45	29.88
	.9	.53	.91	.59	.16	2.45	17.81
	1.0	.37	.88	.42	.10	1.34	38.06
	1.2	.45	.89	.51	.16	1.80	18.53
	1.5	.39	.91	.43	.16	1.37	17.97
	3.0	.36	.94	.38	.17	1.14	17.12
	5.0	.39	.79	.49	.12	1.72	26.43
11	.5	.37	.78	.48	.17	1.61	16.68
	.7	.40	.88	.46	.10	1.52	37.95
	.9	.37	.67	.55	.09	2.14	43.29
	1.0	.34	.65	.53	.23	1.92	10.55
	1.2	.37	.92	.41	.08	1.26	49.47
	1.5	.39	.85	.46	.13	1.53	24.84
	3.0	.40	.88	.45	.17	1.49	15.97
	5.0	.34	.91	.38	.16	1.12	19.06
12	.5	.51	.88	.58	.17	2.33	16.07
	.7	.54	.89	.61	.10	2.64	34.12
	.9	.54	.90	.60	.09	2.50	40.16
	1.0	.44	.86	.51	.11	1.79	31.98
	1.2	.44	.87	.50	.07	1.76	60.05
	1.5	.46	.93	.49	.10	1.70	35.86
	3.0	.36	.89	.40	.20	1.23	13.06
	5.0	.36	.93	.39	.13	1.18	24.37

Table II (contd.)

Event N.	Freq. Hz	σ_u	σ_ϕ	σ_u/σ_ϕ	ϱ	$D(\sigma_u/\sigma_\phi)$	$D(\varrho)$
13	.5	.44	.94	.47	.21	1.57	12.19
	.7	.45	.88	.51	.07	1.82	52.98
	.9	.46	.88	.52	.12	1.91	28.49
	1.0	.46	.88	.52	.08	1.89	46.56
	1.2	.39	.85	.46	.12	1.50	26.20
	1.5	.40	.86	.47	.19	1.59	14.00
	3.0	.32	.89	.36	.20	1.06	12.87
	5.0	.39	.91	.43	.14	1.34	22.86
14	.5	.57	.91	.62	.19	2.74	14.32
	.7	.53	.90	.59	.15	2.41	21.06
	.9	.36	.73	.49	.25	1.70	8.88
	1.0	.35	.74	.47	.28	1.59	7.46
	1.2	.43	.85	.51	.19	1.82	13.55
	1.5	.49	.84	.58	.13	2.35	24.30
	3.0	.24	.90	.26	.11	.70	30.34
	5.0	.32	.90	.35	.13	1.02	25.19
15	.5	.45	.91	.50	.18	1.73	15.63
	.7	.48	.89	.54	.11	2.02	29.62
	.9	.49	.93	.52	.15	1.91	20.24
	1.0	.48	.89	.53	.11	1.98	31.43
	1.2	.47	.95	.50	.19	1.74	13.72
	1.5	.43	.84	.51	.12	1.83	27.72
	3.0	.33	.91	.36	.22	1.07	11.30
	5.0	.34	.86	.39	.18	1.19	15.56
16	.5	.49	.90	.55	.14	2.10	21.94
	.7	.51	.92	.56	.15	2.16	19.71
	.9	.41	.89	.45	.13	1.49	25.94
	1.0	.40	.89	.45	.08	1.48	48.35
	1.2	.42	.90	.47	.12	1.59	27.83
	1.5	.42	.91	.46	.06	1.51	69.01
	3.0	.39	.88	.44	.15	1.41	20.13
	5.0	.35	.92	.38	.15	1.12	20.41
17	.5	.39	.88	.44	.15	1.41	19.66
	.7	.42	.91	.46	.13	1.53	25.27
	.9	.36	.72	.50	.25	1.75	8.77
	1.2	.45	.91	.49	.20	1.69	12.59
	1.5	.39	.88	.45	.16	1.45	17.62
	3.0	.24	.92	.26	.16	.69	17.87
	5.0	.38	.87	.43	.19	1.37	14.46
18	.5	.47	.86	.55	.22	2.10	10.95
	.7	.48	.88	.54	.10	2.07	36.41
	.9	.45	.91	.49	.23	1.70	10.15
	1.0	.45	.93	.48	.14	1.64	21.22
	1.2	.43	.91	.48	.11	1.62	32.69
	1.5	.52	.88	.59	.15	2.43	19.54
	3.0	.40	.89	.46	.15	1.50	19.87
	5.0	.41	.90	.46	.12	1.50	28.12

Table II (contd.)

Event N.	Freq. Hz	σ_u	σ_ϕ	σ_u/σ_ϕ	ϱ	$D(\sigma_u/\sigma_\phi)$	$D(\varrho)$
19	.5	.57	.92	.63	.16	2.80	18.65
	.7	.41	.90	.45	.12	1.47	28.16
	.9	.42	.85	.49	.14	1.71	22.88
	1.0	.39	.76	.51	.15	1.84	21.01
	1.2	.42	.90	.47	.10	1.55	35.67
	1.5	.37	.90	.41	.10	1.28	36.33
	3.0	.35	.92	.38	.17	1.13	16.15
	5.0	.32	.91	.35	.19	1.02	13.70
20	.5	.58	.95	.61	.22	2.64	11.23
	.7	.56	1.00	.56	.08	2.20	50.39
	.9	.35	.67	.52	.20	1.88	12.77
	1.0	.34	.64	.53	.26	1.93	8.29
	1.2	.40	.88	.45	.20	1.48	13.18
	1.5	.51	.86	.59	.10	2.45	34.91
	3.0	.41	.88	.47	.11	1.56	31.97
	5.0	.30	.90	.33	.23	.95	10.62
21	.5	.57	.94	.61	.27	2.63	7.60
	.7	.55	1.00	.55	.13	2.13	24.79
	.9	.38	.72	.53	.09	1.94	43.20
	1.0	.37	.69	.54	.12	2.00	28.60
	1.2	.43	.85	.50	.18	1.78	14.95
	1.5	.50	.85	.59	.14	2.48	21.43
	3.0	.34	.95	.35	.16	1.02	18.18
	5.0	.40	.84	.48	.22	1.63	10.97
22	.5	.57	.87	.66	.10	3.24	35.57
	.7	.39	.91	.42	.11	1.33	32.43
	.9	.40	.89	.45	.06	1.47	70.29
	1.0	.50	.89	.56	.11	2.18	29.64
	1.2	.38	.88	.43	.24	1.36	9.79
	1.5	.47	.90	.53	.11	1.93	30.88
	3.0	.38	.91	.42	.16	1.33	17.49
	5.0	.32	.88	.36	.14	1.06	22.53
23	.5	.45	.90	.50	.25	1.74	8.97
	.7	.50	.91	.55	.06	2.07	65.45
	.9	.49	.90	.55	.07	2.08	53.61
	1.0	.49	.88	.55	.12	2.14	26.71
	1.2	.48	.95	.50	.28	1.75	7.44
	1.5	.42	.84	.50	.28	1.76	7.49
	3.0	.41	.93	.44	.20	1.44	13.38
	5.0	.31	.91	.34	.25	.99	8.90
24	.5	.39	.88	.44	.14	1.44	21.33
	.7	.53	.89	.60	.11	2.54	31.91
	.9	.54	.89	.60	.09	2.58	39.51
	1.0	.53	.89	.59	.12	2.49	26.65
	1.2	.46	.91	.50	.12	1.79	27.57
	1.5	.48	.87	.56	.12	2.17	28.25
	3.0	.39	.88	.44	.12	1.41	26.90
	5.0	.38	.87	.44	.25	1.42	8.71

where r is the space separation, $\mu = \delta c/C_0$, and a is the correlation length. The log-amplitude and phase fluctuations can be obtained (Chernov 1960)

$$\sigma_u^2 = <u^2> = (\sqrt{\pi}/2) < \mu^2 > k^2 aR \cdot (1 - (\tan^{-1} D)/D) \tag{2}$$

$$\sigma_\phi^2 = <\phi^2> = (\sqrt{\pi}/2) < \mu^2 > k^2 aR \cdot (1 + (\tan^{-1} D)/D) \ , \tag{3}$$

where k is the wave number, R is the layer thickness and D is the 'wave parameter' defined by

$$D = 4R/ka^2 \ . \tag{4}$$

The correlation coefficient between log amplitude and phase fluctuation is defined as

$$\begin{aligned} \varrho &= <u\phi> /(<u^2><\phi^2>)^{1/2} \ , \qquad \text{or} \\ \varrho &= (1/2) \cdot \log(1 + D^2)/(D^2 - (\tan^{-1} D)^2)^{1/2} \ . \end{aligned} \tag{5}$$

The ratio of rms log-amplitude and rms phase fluctuations is given by

$$\gamma = \frac{\sigma_u}{\sigma_\phi} = \frac{< u^2 >^{1/2}}{< \phi^2 >^{1/2}} = \frac{1 - (\tan^{-1} D)/D)^{1/2}}{1 + (\tan^{-1} D)/D)^{1/2}} \ . \tag{6}$$

We see from the Eqs (5) and (6) that the ϱ and γ are functions of D only. Thus, D can be estimated independently from these equations. From the calculated D, the layer thickness R can be obtained by using Eq. (4), if we know the correlation length a. It can be shown that when $D \gg 1$ the transverse correlation length of log amplitude and phase are approximately equal to correlation length. Thus, correlation length a can be estimated from a measurement of the transverse correlation of u and ϕ for the case of $D \gg 1$.

The transverse correlation of u and ϕ for $D \gg 1$ can be obtained (Chernov 1960) from

$$\varrho_u = \frac{\exp(-x_T^2/a^2) - \frac{1}{D} \cdot (\Pi/2 - Si(\frac{1}{D} \cdot x_T^2/a^2))}{(1 - \frac{1}{D} \tan^{-1} D)} \tag{7}$$

$$\varrho_\phi = \frac{\exp(-x_T^2/a^2) + \frac{1}{D} \cdot (\Pi/2 - Si(\frac{1}{D} \cdot x_T^2/a^2))}{(1 + \frac{1}{D} \tan^{-1} D)} \ , \tag{8}$$

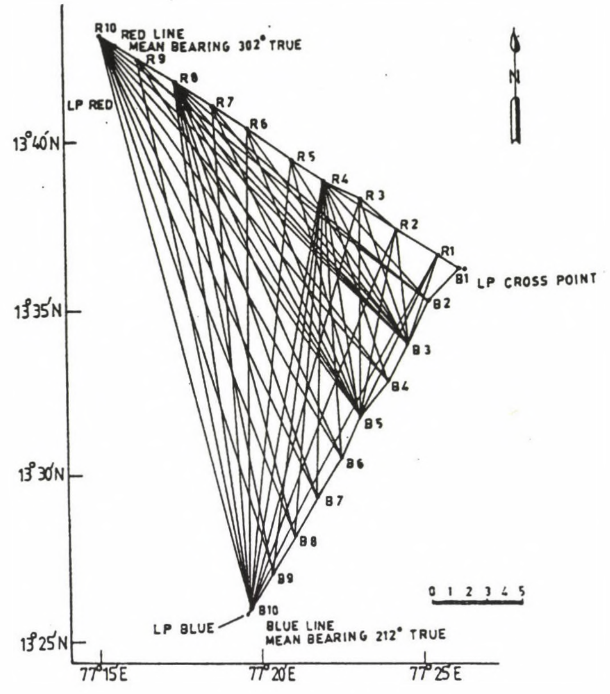
where x_T is the distance between two seismometers of a pair, Si is the function defined as

$$Si(x) = \int_0^x (\sin t)/t \ dt \tag{9}$$

and ϱ_u and ϱ_ϕ are the transverse autocorrelation coefficients for log amplitude and phase fluctuations respectively, defined as

$$\varrho_u = \frac{<u_1 u_2>}{<u^2>} \qquad \text{and} \tag{10}$$

$$\varrho_\phi = \frac{<\phi_1 \phi_2>}{<\phi^2>} \ . \tag{11}$$



GAURIBIDANUR ARRAY

Fig. 3. A schematic diagram of the Gauribidanur seismic array (GBA) showing pair combinations of some of the seismometers

The above expressions are valid only for small l/a values, which means that the sensor separations must be small. Chernov (1960) has shown that these formulae are applicable only when l/a is smaller than 3 for $D = 10$. Berteussen et al. (1977) have shown that the correlation length a for the medium beneath GBA is about 20 km. The length of each line of GBA is about 25 km and the maximum distance between two sensors (R_{10} and B_{10}) is about 32 km; the minimum distance between any two seismometers is about 1.5 km. Hence l/a varies from 0.075 to 1.6 and thus is always less than 3. In comparison, at LASA the maximum distance between sensors is about 110 km and correlation length is about 10 km. Thus $l/a = 11(\gg 3)$. Thus, application of Chernov Theory at GBA does not need modification whereas application at LASA does need modification. Berteussen et al. (1975) have generalized these equations for time residuals and Capon (1974) has modified them for slowness fluctuations. We have used these equations for phase fluctuations directly.

Thus reshuffling and adding Eqs (7) and (8) we find

$$a = \left[-x_T^2/\log\left\{\frac{1}{2}\left(\varrho_u + \varrho_\phi + \frac{1}{D}(\varrho_\phi - \varrho_u)\tan^{-1} D\right)\right\}\right]^{1/2} \tag{12}$$

and we can then calculate the correlation length a.

In deriving the above formulae, two main approximations, namely the Born and Fresnel approximations, were taken into consideration. The condition for the applicability of the Born approximation is that the scattering field is small in comparison to the primary plane wave and may be expressed quantitatively using the fractional energy lost from the primary waves by scattering.

The fractional energy lost by scattering from primary waves during propagation over distance R is given by

$$\Delta I/I = \sqrt{\Pi}\cdot < \mu^2 > k^2 a R(1 - \exp(k^2 a^2)) \tag{13}$$

for a medium characterized by the correlation function $N(r) = \exp(-r^2/a^2)$. This Born approximation is valid for small $\Delta I/I$.

The scattering coefficient for forward scattering ($ka \gg 1$) can be defined by the fractional energy loss per unit propagation distance. Thus, the scattering coefficient g for the Gaussian correlation function is given as

$$g = \Delta I/IR = (\Pi)^{1/2}\cdot < \mu^2 > k^2 a \,. \tag{14}$$

The Fresnel approximation allows replacement of a model spherical waveform by a parabolic one, thereby neglecting backscattering and requiring the scattering to be directed forward. If $ka = 2\Pi a/\lambda$ is large, most of the scattered power is concentrated within the small solid angle $\Theta \simeq 1/ka$ in the forward direction. Thus, the Fresnel approximation is valid for large ka.

The condition of large ka is a crucial one in our application of the Chernov theory to the elastic wave problem. Equations 2 to 13 were obtained for the acoustic pressure wave in a medium with velocity (or refractive index) fluctuations but with uniform density. The relative strength of the scattering source, given by the ratio of density anomaly to velocity anomaly, decreases with increasing ka. Of course, in the limit of ray theory ($ka \to \infty$), only the velocity anomaly determines the wave amplitude. Similarly, the velocity anomaly is the most important scattering source of elastic waves for large ka. Further, according to Knopoff and Hudson (1964, 1967), the P to S and the S to P conversion can be neglected if ka is large.

8. Discussion and conclusions

The wave parameter, D, can be determined independently from Eqs (5) and (6), using correlation coefficient ϱ and ratio σ_u/σ_0 respectively. The relation between σ_u/σ_0 and ϱ is shown in Figs 4a–4d for the frequencies of 0.5, 1.0, 3.0 and 5.0 Hz. The solid curve is the theoretical relationship, obtained by eliminating D from Eqs (5) and (6), and the 90 % confidence limits for the correlation coefficients were obtained by Fischer's (1954) z transformation for a sample of size 20. The corresponding confidence limits for the ratio of standard errors have been obtained using the F-distribution and were found here to be within the limits of the correlation coefficients except where the ratio was very close to unity. The majority of the data fell within the 90 % confidence limits for the frequency of 0.5 Hz (Fig. 4a). But, at frequencies of 1.0 Hz (Fig. 4b), 3.0 Hz (Fig. 4c) and 5.0 Hz (Fig. 4d), the majority of the data fell outside the 90 % limits.

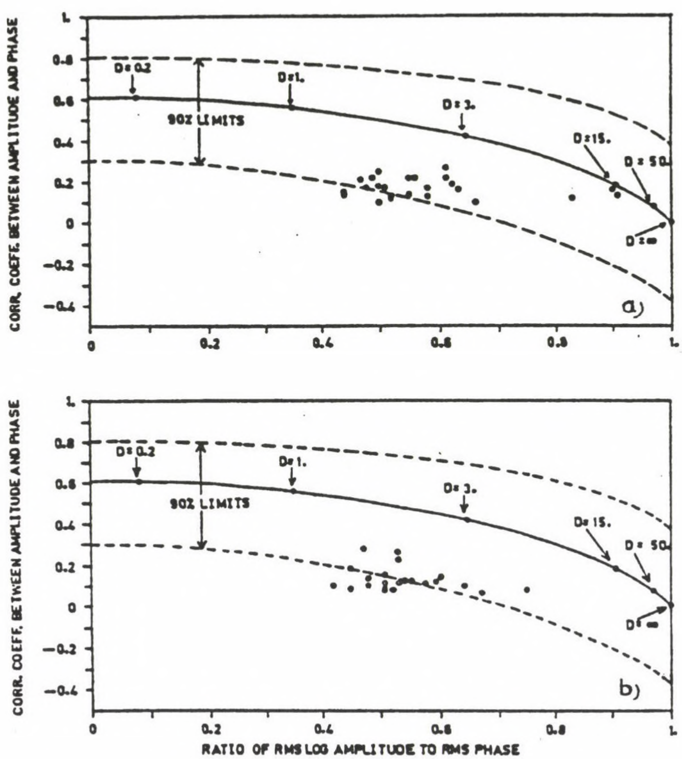


Fig. 4a. The ratio (σ_u/σ_ϕ) of rms log-amplitude fluctuations to rms phase fluctuations is plotted against correlation coefficient ϱ. The smooth curve is the theoretical relation obtained from Eqs (5) and (6) after eliminating D a) at frequency 0.5 Hz, b) at frequency 1.0 Hz

The Born approximation condition can be calculated from Eq. (13). The fractional energy loss from primary waves is $\Delta I/I \approx 0.6$ for 0.5 Hz. As $\Delta I/I$ is dependent on the square of frequency, so that at 1. Hz $\Delta I/I \approx 2.4$, the Born approximation can not be applied here for a frequency of 1.0 Hz and the Chernov formula can not be applied at 1.0 Hz and higher frequencies. However, as ka is about 6.7 at 0.5 Hz and 13.3 at 1.0 Hz, the Fresnel condition is satisfied. Aki (1973) has also concluded that Chernov formula can be applied at around 0.6 Hz but not near 1.0 Hz.

Here the data at 0.5 Hz were used to estimate D and a (Table II). The average wave parameter, D, was found to be 20.1 and the correlation length a, was found to be 18 ± 1 km. Thus, the correlation length is slightly less than the 20 km observed at GBA by Berteussen et al. (1977) but somewhat greater than the 10–16 km correlation length found for LASA and NORSAR using subsets of instruments having a comparable aperture to that at GBA. The extension of the medium, R, is found to be about 660 km at 0.5 Hz. For the average value of 0.52 for the rms log-amplitude the deduced rms variation in P-wave velocity is found to be 1.4 %. On the other hand, the rms variation in P-wave velocity is observed to be 2.3 % when

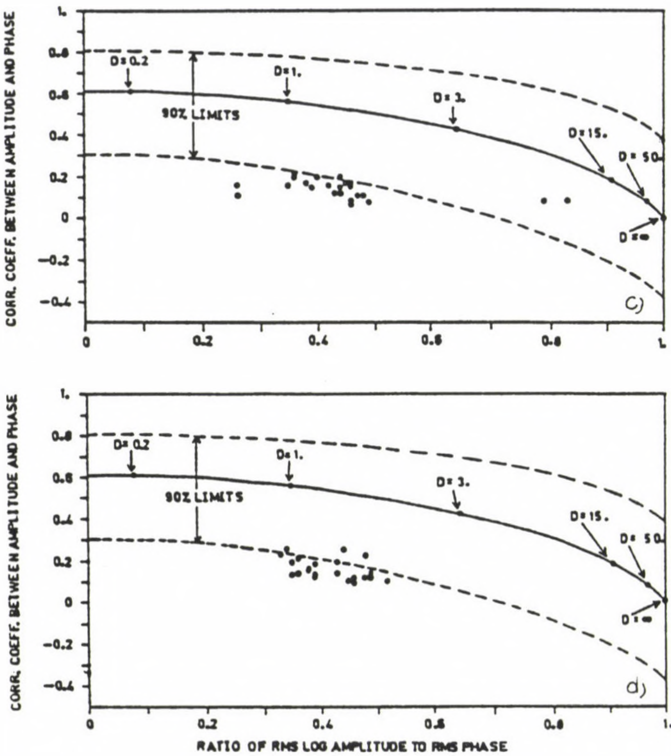


Fig. 4b. The ratio (σ_u/σ_ϕ) of rms log-amplitude fluctuations to rms phase fluctuations is plotted against correlation coefficient ϱ. The smooth curve is the theoretical relation obtained from Eqs (5) and (6) after eliminating D c) at frequency 3.0 Hz, d) at frequency 5.0 Hz

the average value of 0.898 for the rms phase is used. However, the difference is not significant. Thus, we conclude that the heterogeneity of the medium beneath GBA is characterized by a Gaussian correlation length of about 18 km with about 1.4 % – 2.3 % velocity perturbations and it extends to 660 km in depth. The scattering coefficient, g, for these values of $<\mu^2>^{1/2}$ and a is found to be $9.0 - 22.7 \times 10^{-4}$ km^{-1} at frequency 0.5 Hz, which is greater here than was found for refracted waves in the USSR (10^{-3} to 10^{-4} km^{-1} for $5 < f < 10$ Hz) (Nikolayev 1968, Nikolayev and Tregub 1970) and in the Kanto region, Japan (0.22 to 0.5 for 1.5 Hz to 24 Hz) (Aki 1980).

Berteussen et al. (1977) also noted some complex time residuals. They suggested an exceptionally homogeneous medium beneath GBA. The results of this study reveal that the crust and upper mantle structures beneath the Gauribidanur array is homogeneous but slightly less homogeneous than that proposed by Berteussen et al. (1977), but its homogeneity is less compared to similar investigations at NORSAR and at LASA. A summary of the transmission fluctuations analyses for the various single layer models is given in Table III. Here, the crustal/upper mantle scattering zone is also suggested by high semblance values for the Dharwar Cratons

Table III. Summary of transmission fluctuation analysis by Chernov theory for single layer models

		Data								
		f Hz	σ_Θ	σ_t sec	σ_u	$\sigma_{u\Theta}$	a km	Extent km	D	$\tilde{\alpha}$
Aki (1973)	LASA	0.5	0.6	0.19	0.32	0.35	10	0–60	5	4 %
Capon (1974)	LASA	0.8	0.52	0.10	0.37	0.23	12	0–136	6	1.9 %
Berteussen et al. (1975)	LASA	0.7	0.08–0.11	0.02–0.025	0.26–0.42		15	0–50		0.3–3 %
Berteussen et al. (1975)	NORSAR	0.7	0.26–1.75	0.05–0.4	0.15–0.36		15–60	0–100		0.5–2 %
Berteussen et al. (1977)	GBA			0.05–0.08			20	0–250 0–650		.15–.3 %
Present work	GBA	0.5	0.90	0.29	0.52	0.16	18	0–660	20	1.4–2.3 %

σ_Θ rms phase fluctuations
σ_t rms arrival time fluctuations
σ_u rms log amplitude fluctuations
a correlation length
D wave parameter, defined by $D = 4R^2/Ka^2$, where R is the layer thickness
$\tilde{\alpha}$ P-wave speed perturbation, defined by $\tilde{\alpha} = <(\delta\alpha/\alpha_0)^2>^{1/2}$

(Mohan and Rai 1992) and also by anomalous time delays for teleseism crossing beneath the Closepet granite and recorded at GBA (McCarthy et al. 1983). The Bababudan Group, which consists of mafic and ultra-mafic rocks of perdominantly volcanic origin, north and north-west of the array, may also have scattering effects. Thus it may be worthwhile to apply the direct inversion scheme of Aki et al. (1976, 1977) for a detailed mapping of the anomalous lithospheric structure beneath GBA.

In this study, phase fluctuations are directly used for the analysis purposes in Chernov theory instead of time residuals or slowness fluctuations for a medium aperture array. Thus, we conclude that the medium beneath GBA is less homogeneous than what due to the extremely complex geological processes that occurred during the evolution of Cratons and intrusions and the volcanic processes that took place beneath this area within the upper mantle. This study also shows that the seismic scattering tomography technique may be used for imaging broad geological features.

Acknowledgement

One of the authors (JNT) is thankful to University Grants Commission, New Delhi for financial assistance.

References

Aki K 1973: *J. Geoph. Res.*, 78, 1334–1346.

Aki K 1980: *J. Geoph. Res.*, 85, 6496–6504.

Aki K, Christoffersson A, Husebye E S 1976: *Bull. Seismol. Soc. Am.*, 66, 501–524.

Aki K, Christoffersson A, Husebye E S 1977: *J. Geoph. Res.*, 82, 277–296.

Arora S K 1971: *Bull. Seismol. Soc. Am.*, 61, 671–683.

Arora S K, Varghese T G, Krishnan C A 1970: *Nature*, 225, 261–262.

Berteussen K A 1974: NORSAR location calibration and time delay correction. NORSAR, Sci. Rep. 2–73/74. National Technical Information Services, Springfield, Va

Berteussen K A 1975: *J. Geophys.*, 41, 71–84.

Berteussen K A, Christoffersson A, Husebye E S, Dahle A 1975: *Geophys. J. R. Astr. Soc.*, 42, 403–417.

Berteussen K A, Husebye E S, Mereu R F, Ram A 1977: *Earth Planet Sci. Lett.*, 37, 326–332.

Capon J 1974: *Bull. Seismol. Soc. Am.*, 64, 235–266.

Capon J, Berteussen A 1974: *Geophys. Res. Lett.*, 1, 327–328.

Chernov L A 1960: Wave propagation in a random medium. MacGraw Hill, New York

Fisher R A 1954: Statistical methods for research workers. 12th ed. Oliver and Boyed, London

Flattê S M, Moody T N 1990: Weak-fluctuation theory via path integrals. Contribution 106 of the Institute of Tectonics of the University of California, Santa Cruz

Flattê S M, Wu R S 1988: *J. Geoph. Res.*, 93, 6601–6614.

Flattê S M, Xie XB 1992: *Geophys. Res. Lett.*, 19, 557–560.

Flattê S M, Wu R S, Shen Z K 1991: *Geophys. Res. Lett.*, 18, 1269–1272.

Knopoff L, Hudson J A 1964: *J. Acoust. Soc. Am.*, 36, 338–343.

Knopoff L, Hudson J A 1967: *J. Acoust. Soc. Am.*, 42, 18–20.

McCarthy S A, Powell C A, Rogers J J W 1983: In: Precambrian of South India, S M Naqvi, J J W Rogers (eds), Geological Society of India, Bangalore, 525–552.

Mohan G, Rai S S 1992: *Phys. Earth Planet. Int.*, 71, 36–45.

Mowat W M H, Burch R F 1974: Handbook for the stations which provide seismograms to the Blacknest seismological centre. United Kingdom, AWRE report No. 44/47/29.

Nikolayev A V 1968: Seismic properties of weakly heterogeneous media. Acad. Sci. USSR, Phys. Solid Earth, Eng. Trans. 83–87.

Nikolayev A V 1972: Seismics of heterogeneous and turbid media. Nauka, Moscow

Nikolayev A V, Tregub F S 1970: *Tectonophysics*, 10, 573–578.

Ram A 1976: The identification and interpretation of upper mantle travel time branches from slowness measurements made on data recorded at the Gauribidanur (India) and Yellowknife (Canada) seismic arrays. Ph.D. thesis, University of Western Ontario, London, Ontario, Canada

Ram, A, Mereu R F 1977: *J. R. Astron. Soc.*, 49, 87–113.

Tripathi J N, Ram A 1995: Small-scale structure of the Indian lithosphere-asthenosphere beneath Gauribidanur seismic array using the theory of wave propagation in random media (Manuscript under preparation)

Wu R S, Flattê S M 1990: *Pure and Appl. Geophys.*, 132, 175–196.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 145–159 (1996)

SHEAR WAVE VELOCITY STRUCTURE OF NORTHERN AND NORTH-EASTERN ETHIOPIA

F KEBEDE[1,4], A VUAN[3], T MAMMO[2,4], G COSTA[3], G F PANZA[4]

[Manuscript received November 27, 1995]

The non-linear inversion technique known as hedgehog is utilised to define the average crustal structure of North and North-Eastern Ethiopia. To accomplish the task a two dimensional frequency-time analysis is performed to obtain Rayleigh wave group velocity dispersion curves. Six earthquakes recorded by the broad-band digital seismograph installed at the Geophysical Observatory of Addis Ababa University are utilized. The crustal structure between the Gulf of Tadjura (western Gulf of Aden) and Addis Ababa crossing southern Afar (path I) can be approximated by a total thickness of about 22 km with average S-wave velocity in the range 2.3–3.9 km/s. The crust-mantle transition is poorly developed at greater depths and the shear wave velocity ranges from 4.0 km/s to 4.3 km/s. If the effect of the plateau part is taken into account the average total crustal thickness is found to be less than 18 km and the average S-wave velocity varies in the range 2.4–3.9 km/s. The low shear wave velocity under the Afar crust is consistent with the result of other geophysical studies. For path II, which passes through the border of the Western Ethiopian plateau, the average crustal structure is found to be approximated by a thickness of about 40 km and average S-wave velocity between 3.0 km/s and 3.9 km/s. The crust overlies a lithospheric mantle with a shear wave velocity in the range 4.1–4.4 km/s.

Keywords: anomalous mantle; crustal structure; FTAN and inversion

1. Introduction

The geological history of Ethiopia and its neighbouring countries indicates that in early Tertiary it became the site of intense magmatic and tectonic activity that has continued to the present day. From early Tertiary until the end of the Eocene period extensive vertical uplift took place with the eventual splitting of the crust into three rift systems — the Red sea, the Gulf of Aden and the Ethiopian Rift (Gass 1970) which are closely connected with (Fig. 1) in a structural and kinematic way (e.g. Mohr 1967, Mohr 1970a, 1970b, McKenzie et al. 1970, Gass 1975, Bonatti 1987, Makris and Ginzburg 1987, Bohannon et al. 1989). These three rift systems meet in the Afar triple junction which is a unique geophysical and geological province. This junction is one of the two areas in the world (the other being Iceland) where crustal spreading, dike injections and the formation of new

[1]Geophysical Observatory, Science Faculty, Addis Ababa University, POB 1176, Addis Ababa, Ethiopia
[2]Geology and Geophysics Department, Science Faculty, Addis Ababa University, POB 1176, Addis Ababa, Ethiopia
[3]Istituto di Geodesia e Geofisica, Universitá degli Studi di Trieste, Via dell'Universitá 7, 34100 Trieste, Italy
[4]International Center for Theoretical Physics, Strada Costiera 11, 34100 Trieste, Italy

lithospheric plate material can be observed by geologists (Barberi and Varet 1975, 1977, Berhe 1986). Based on a small-scale plate tectonics model, Barberi and Varet (1975, 1977) proposed within Afar the existence of spreading centres connected by transform faults. This, in turn, led to the identification of several micro plates which move relative to each other. On the other hand, Berhe (1986) argued that the Afar depression and its neighbouring regions are experiencing an extensional tectonics due to spreading and dike injection. According to this idea one may expect the presence of normal faults as the only manifestations for the mode of deformation in Afar. In general, the unresolved issue as to which tectonic model can better explain the Afar tectonics indicates the complexity of the area. Furthermore, it indicates the extent to which the crust-mantle system in the region is complex and deformed due to relative motion of plates in the region. In this respect, the Afar depression is probably the best suited area for studying the different stages of development from a true continental rift (East African rift) to a juvenile oceanic trough (Red Sea, Gulf of Aden).

To infer crustal velocities Searle (1975) used dispersion of surface waves along a path crossing the southern part of the Afar depression and concluded that the crustal thickness is about 20 km.

Using deep seismic soundings data (Berckhemer et al. 1975) and gravity data (Makris et al. 1975) the crustal structure of the Afar depression has been studied along several profiles (especially along the central and northern part of Afar). Berckhemer et al. (1975) proposed that the crust below the Ethiopian plateau has a thickness of about 38 km and overlies a mantle of normal seismic velocity, a typical shield structure. They also proposed that in Afar the depth to the top of the anomalous mantle (P-wave velocity ranging from 7.3 km/s to 7.6 km/s) varies from 16 km (northern Afar) to 26 km (southern Afar) and estimated the thickness of the anomalous mantle layer to be in the range from 15 km to 40 km. Makris et al. (1975) proposed that the crustal thickness of Afar is in the range from 14 km to 22 km and that of the Ethiopian plateau (western and eastern plateau) is in the range from 30 km to 42 km.

From travel times of some earthquakes located in the Gulf of Aden and the Red Sea and recorded at the WWSSN (World Wide Seismological Station Network) station at Addis Ababa (AAE) Searle and Gouin (1971) obtained a Pn-wave velocity of 7.95 km/s and a Sn-wave velocity of about 4.3 km/s for waves propagating through Afar. Furthermore, from the 1969 Serdo earthquakes in central Afar, Dakin et al. (1971) obtained, a P-wave velocity of 7.4 km/s at the base of the crust in western Afar.

From near earthquakes travel time observations Searle and Gouin (1971) gave an upper limit of 48 km for the crustal thickness below Addis Ababa (on the western Ethiopian plateau). The crustal spectral transfer ratio was determined for Addis Ababa from long period teleseismic P-waves by Bonjer et al. (1970) and associated with a crustal model of 38 km thickness. The upper layer is 24 km thick with average P-wave velocity of 6.0 km/s, the lower layer is 14 km thick with 6.9 km/s for P-wave and the mantle top P-wave velocity is 8.2 km/s.

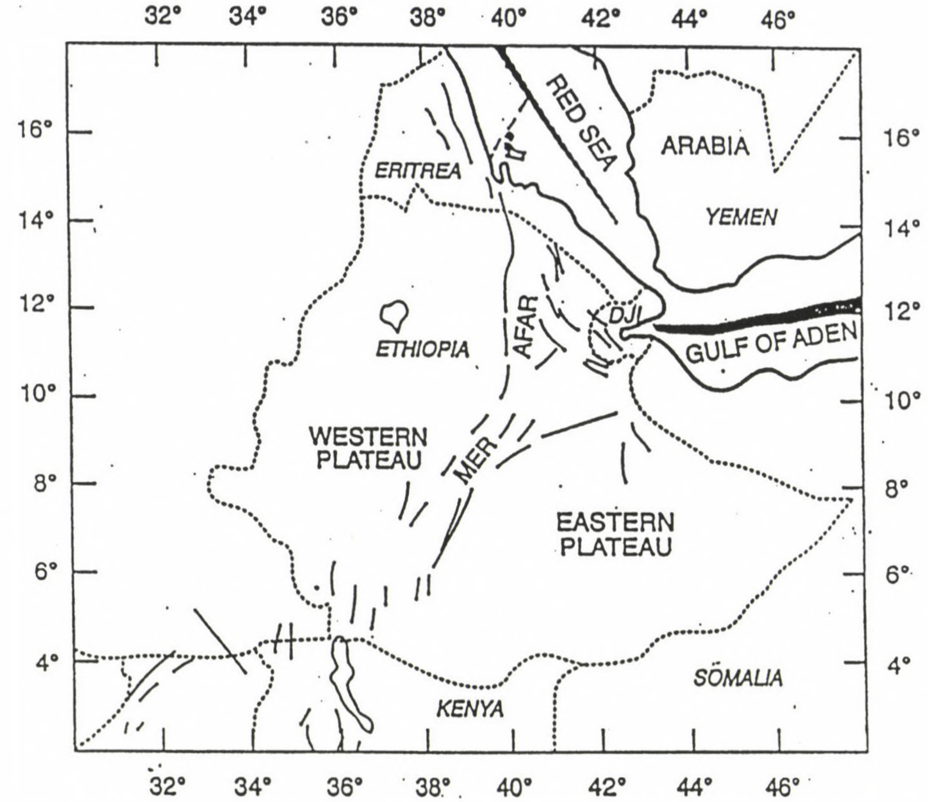


Fig. 1. Major structural elements in and around Afar depression. MER, and DJI denote the main Ethiopian rift system (part of the northern part of the East African rift system) and Djibouti, respectively. Heavy lines in the Red Sea and the Gulf of Aden show the axial troughs. Lines indicate major faults in the region. The broken line in the southern Red Sea indicates the proposed transform fault connecting northern Afar to the axial trough of the Red Sea (Barberi and Varet 1975, 1977)

2. Data source and analysis

In March 1993 the Seismological station (AAE) at the Geophysical Observatory of Addis Ababa University upgraded its recording capability by installing a broad-band three component digital seismograph system equipped with a Guralp CMG-3T seismometer and Nanometrics data acquisition system. The system sampling rate is 40 samples per second. Since the installation of the seismograph system quite a large number of local and teleseismic events have been digitally recorded at AAE. However, for most of the local earthquakes recorded by the station it was not possible to get origin time and epicentral location accurate enough for the purpose of this study. Thus, we selected those events which were large enough to be recorded by other WWSSN stations outside Ethiopia and consequently their hypocentral parameters determined. Table I gives the hypocentral parameters of

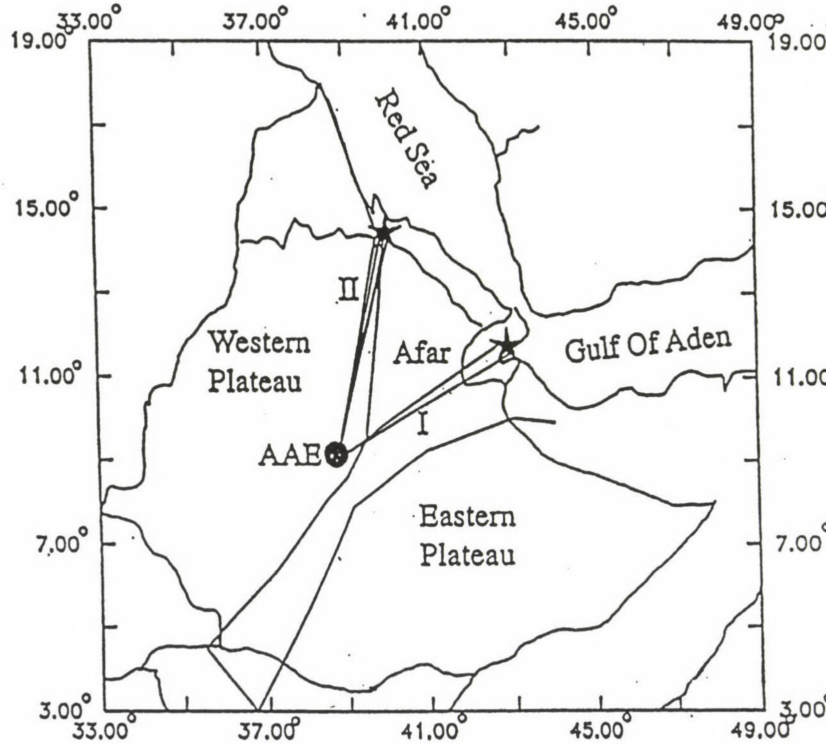


Fig. 2. Map showing major tectonic provinces in Afar and neighbouring regions. Solid star and solid circle indicate the earthquake epicenter and the AAE seismic station, respectively. Roman numbers I and II are intended to indicate propagation paths

the six earthquakes selected for this study. The parameters given in the table are taken from Earthquake Data Report (EDR) bulletin published by the United States Geological Survey. Figure 2 gives the location of epicentres of the events studied here and the paths to the AAE station, while Fig. 3 shows sample vertical component seismic records for some of the events (events 2 and 5 in Table I).

The paths between the epicentres and AAE station are divided into two groups (paths I and II, see also Fig. 2). Path I is sampled by the seismic waves generated by earthquakes occurred in the Gulf of Tadjura (western Gulf of Aden) and propagating through the southern part of the Afar depression. For this profile about 70 % of the path lies in the rift (Afar depression). Similarly, path II is followed by seismic waves originated from the northern part of Afar depression. About 90 % of the propagation path is through the western Ethiopian plateau.

In this study Rayleigh waves, recorded on vertical component broad-band digital seismograms, are utilized. Data are corrected for the instrument response and resampled with a 1s step.

The two-dimensional frequency time analysis (FTAN) (Dziewonski et al. 1969, Levshin et al. 1972) followed by phase equalization technique is employed to enhance

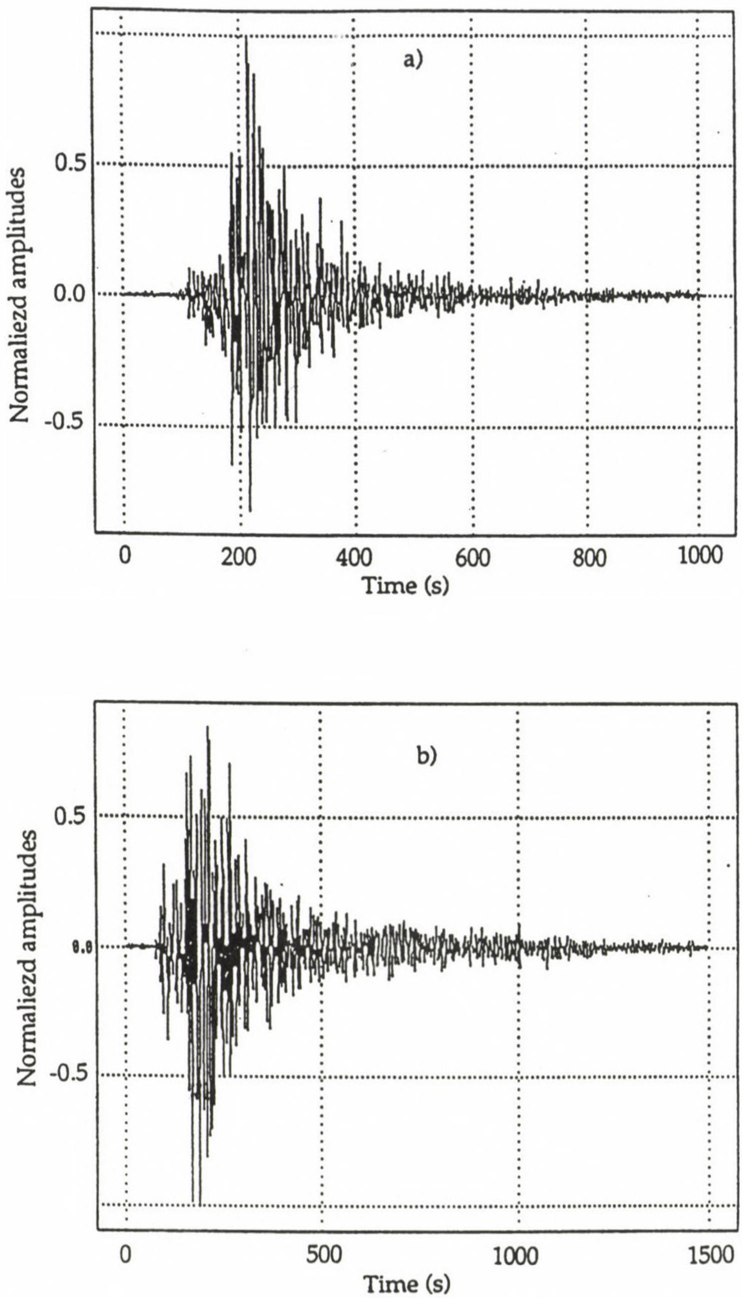


Fig. 3. Display of vertical component seismograms for some of the events studied here as recorded by the broad-band seismograph system at the AAE station: a) for event 2 in Table I; b) for event 5 in Table I

Table I. Hypocenter parameters for the six earthquakes used in this study

No.	Date	Lat. (N)	Long. (E)	mb	Ms	Dist. km
1	02/05/1993	14.470	40.093	4.7	3.7	622.7
2	02/05/1993	14.573	39.993	4.7	4.3	631.4
3	04/05/1993	14.560	39.864	4.4	–	627.2
4	05/05/1993	14.371	40.148	4.7	4.4	613.4
5	11/04/1994	11.826	42.805	5.5	4.7	540.9
6	24/04/1994	11.600	43.089	5.2	4.7	553.3

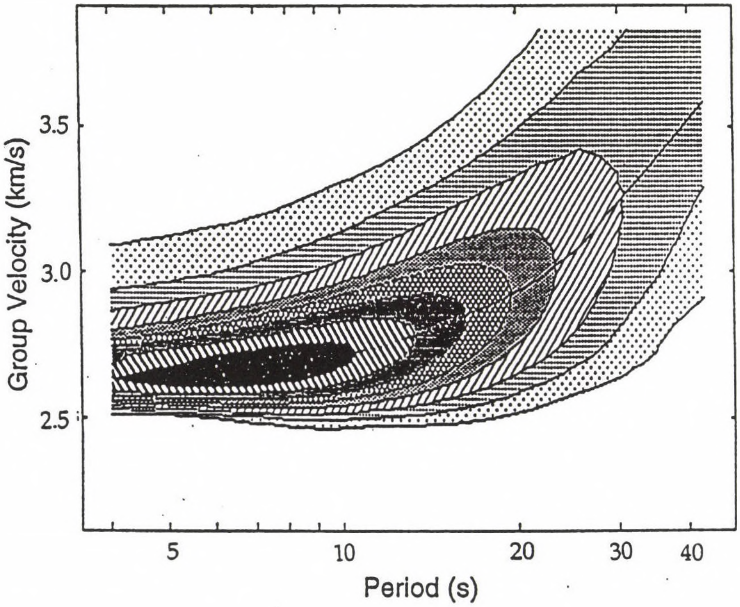


Fig. 4. Display of the two-dimensional Frequency Time Analysis (FTAN) for event 5 in Table I

the fundamental mode Rayleigh wave (see also Herrmann and Russell 1990). During the data processing, band pass filtering and windowing are utilized. A detailed description of the method is given in Levshin et al. (1992).

Figures 4 and 5 show the dispersion curves obtained by performing FTAN analysis for events number 5 (path I) and number 2 (path II), respectively.

3. Inversion for shear wave velocity structure

The dispersion curves are mainly sensitive to shear wave velocity variations and this feature can therefore be used to find the S-wave velocity distribution with depth through an inversion scheme. To accomplish the task the non-linear hedgehog

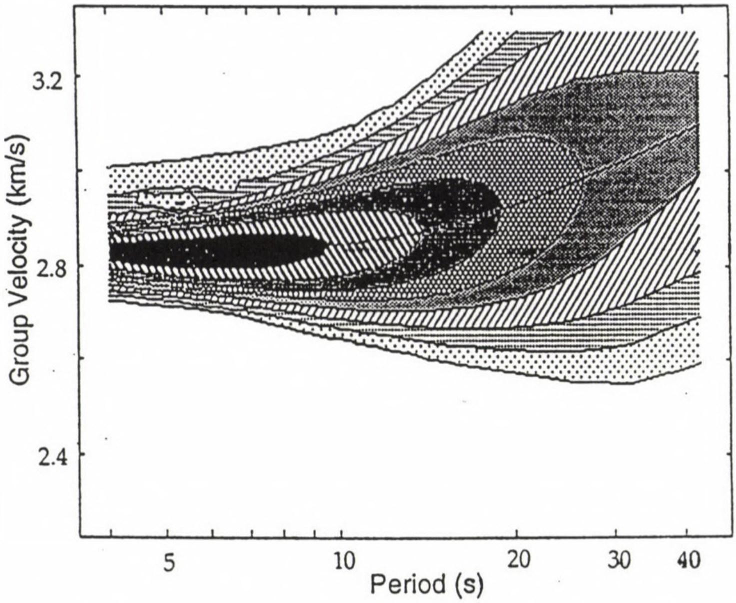


Fig. 5. Display of the two-dimensional Frequency Time Analysis for event 2 in Table I

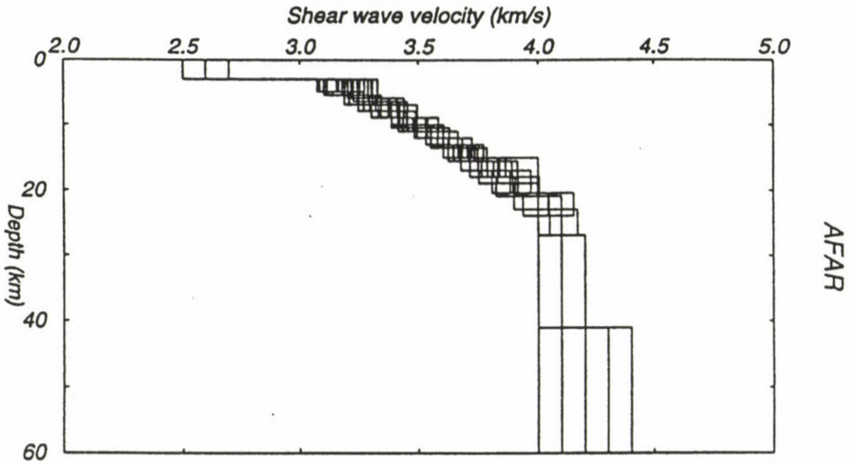


Fig. 6. Accepted shear wave velocity structures for path I (southern Afar) without correction for the effect of the plateau part

inversion procedure as developed by Kelis-Borok and Yanovskaja (1967), Valyus (1972) and Valyus et al. (1973) is utilized. A detailed description of the method, with the discussion of the appropriate parametrization of a plane layered Earth model, is given by Panza (1981).

To start the inversion process an initial model, characterised by shear and com-

Table II. Range of the parameters and their incremental steps used in the inversion, a) for the path I (AFAR), and b) for the path II (ERITREA)

a)

Parameter	Minimum	Maximum	Step
P1	2.199 (km/s)	2.701 (km/s)	0.100 (km/s)
P2	1.900 (km)	7.016 (km)	0.500 (km)
P3	0.007	0.098	0.005
P4	2.859 (km/s)	3.561 (km/s)	0.100 (km/s)
P5	3.999 (km/s)	4.501 (km/s)	0.100 (km/s)
P6	3.999 (km/s)	4.501 (km/s)	0.100 (km/s)

b)

Parameter	Minimum	Maximum	Step
P1	2.739 (km/s)	3.441 (km/s)	0.100 (km/s)
P2	1.400 (km)	6.016 (km)	0.500 (km)
P3	0.002	0.032	0.002
P4	2.759 (km/s)	3.661 (km/s)	0.100 (km/s)
P5	4.099 (km/s)	4.701 (km/s)	0.100 (km/s)
P6	2.900 (km)	21.10 (km)	3.000 (km)

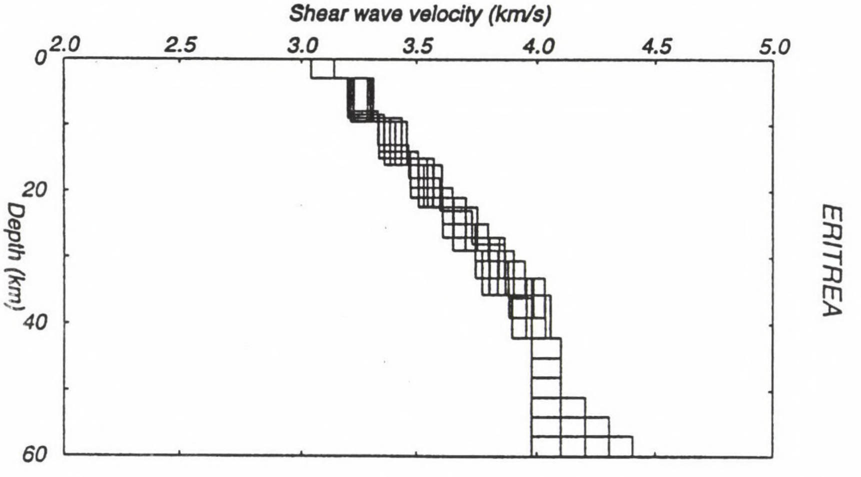


Fig. 7. Accepted shear wave velocity structures for path II as obtained in this study

pressional wave velocities and density of the different layers, must be defined. The starting model must reflect a general subsurface picture of the region under consideration. Accordingly a model proposed by Searle (1975) for path I and the AFRIC model (Gumper and Pomeroy 1970) for path II are chosen as starting models.

Density values in each layer are determined from the corresponding shear wave

Table III. Single point error in the observation determined by the variance in the group velocity measurements a) for the path I (AFAR), and b) for the path II (ERITREA)

a)

Period s	Group velocity km/s	Single point error km/s
40.0	3.390	0.100
20.0	3.139	0.095
13.3	2.945	0.080
10.0	2.809	0.075
8.00	2.655	0.070
6.66	2.579	0.055
5.71	2.529	0.040
5.00	2.489	0.030

b)

Period s	Group velocity km/s	Single point error km/s
40.0	3.159	0.100
30.0	3.090	0.080
20.0	3.000	0.060
10.0	2.904	0.050
9.00	2.897	0.040
7.00	2.889	0.030
5.00	2.886	0.020

velocities using the Nafe-Drake curve (e.g. Talwani et al. 1959). Furthermore, the compressional wave velocity needed in the input is computed by taking the Poisson's ratio equal to 0.25. All earthquakes studied here are regional earthquakes with the longest epicentral distance around 630 km (see Table I). For such an epicentral distance and the period range considered here (see Figs 4 and 5) the difference in group velocities between a flat layered and spherical earth models is negligible (see Bolt and Dorman 1961, Brune and Dorman 1963). Thus no correction for the effect of earth flattening is performed.

Once the starting model is defined the goal of the inversion is to determine the depth of the crust-mantle boundary and to retrieve the lithospheric mantle shear wave velocity. For both paths I and II, we fix the thickness of the uppermost sedimentary layer equal to 3 km while we let variable its S-wave velocity parameter (P1). Below this depth we use a linear velocity gradient modelled with 6 thin layers of equal thickness. Therefore the deeper crustal structure is described with three variable parameters (P2, P3, P4): P2 represented the thickness of the layers, P3 the velocity gradient and P4 the shear wave velocity below the sediments. Finally

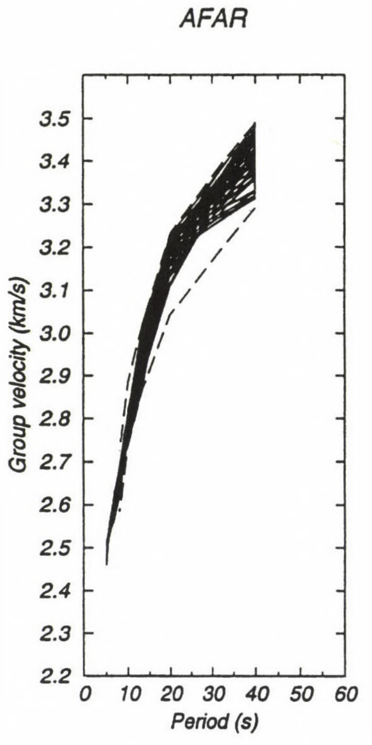


Fig. 8. Comparison of the observed dispersion curves for path I (dashed lines represent the variability range of the observations), not corrected for the effect of the plateau part, with the theoretical dispersion curves obtained from the crustal structures, solution of the inversion, shown in Fig. 6

the shear wave velocity of the upper mantle is allowed to vary up to a depth of 70 km (P5 and P6).

The inversion of the group velocities of path II is performed in a similar manner; in this case the parameter P6 represents the thickness of the layer under the crustal gradient and not its velocity. The range of the parameters and their incremental steps are shown in Table IIa for path I and in Table IIb for path II.

For each period the variance of the group velocities along similar paths is used as single point error. The models, for which the root mean square (RMS) is less than or equal to 0.055 km/s and the single point error is less than or equal to the values given in Table IIIa,b are considered acceptable.

Accepted models for the shear wave velocities are given in Figs 6 and 7 for path I (AFAR) and path II (ERITREA). The comparison of the observed dispersion curves (dashed lines, representing the range of variability of the group velocity) with the theoretical ones, computed from the solutions shown in Figs 6 and 7, is given in Figs 8 and 9 for path I (AFAR) and path II (ERITREA).

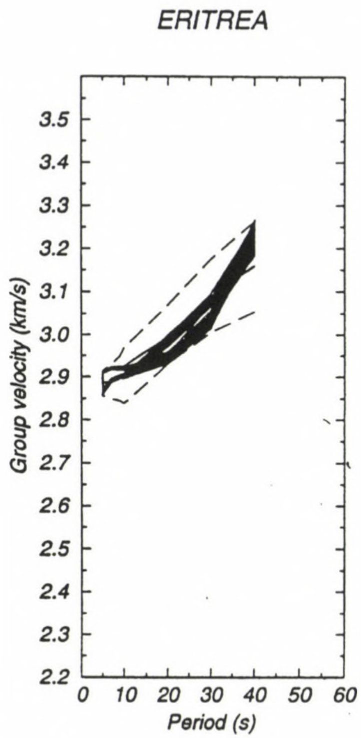


Fig. 9. Comparison of the observed dispersion curves for path II (dashed lines represent the variability range of the observations) and the corresponding theoretical dispersion curve as obtained from the crustal structures, solution of the inversion, shown in Fig. 7

4. Discussion and conclusion

Figure 6 shows the shear wave velocity structures as obtained for path I (AFAR). The crust has a total thickness ranging between 18 km to 22 km and the lithospheric mantle has a low shear wave velocity ranging from 4.0 to 4.3 km/s, even if in the inversion we have explored also normal mantle shear-wave velocity values (> 4.3 km/s).

To correct for the effect of the plateau, AFRIC model as given by Gumper and Pomeroy (1970) was employed. The observed Rayleigh wave group velocities along with the corrections needed to remove the effect of 175 km of plateau path are given in Table IV. The same parametrization applied for path I (AFAR) is used to invert the group velocities corrected for the effect of the plateau. The inverted shear wave velocity structures, AFARC, for path I, after correction for the effect of the plateau part are shown in Fig. 10. The total crust can be approximated by shear wave velocities ranging between 2.4 km/s to 3.9 km/s. The lithospheric mantle shear wave velocity varies in the range 4.0–4.3 km/s. The comparison of the observed (group velocity variability ranges in the experimental data represented by dashed lines) and theoretical dispersion curves is shown in Fig. 11. The total thickness

Table IV. Observed group velocities and corresponding correction needed to remove the effect of the western Ethiopian plateau

Group velocity km/s	Correction km/s
3.39	−0.113
3.14	+0.097
2.95	−0.031
2.81	−0.111
2.66	−0.159
2.58	−0.176
2.53	−0.185
2.49	−0.193

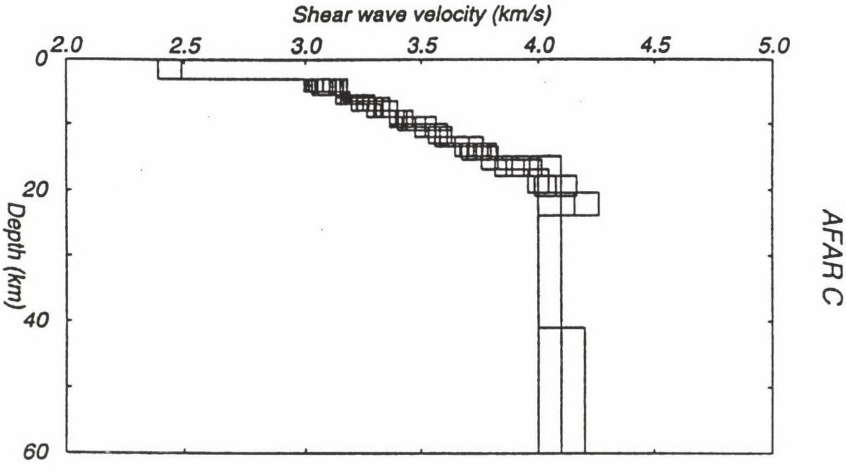


Fig. 10. Accepted shear wave velocity structures for path I as obtained in this study after correcting for the effect of the plateau part

to the top of the lithospheric mantle is less than 20 km and is smaller than the one obtained for the total path between the recording station and the epicentre (uncorrected case).

In general, our results are in quite good agreement with those of Searle (1975), both results indicating the presence of a low shear wave velocity (the presence of anomalous mantle) beneath the Afar crust.

Along several profiles (mainly in central and northern part of Afar depression) deep seismic soundings and gravity surveys have been conducted (Berckhemer et al. 1975, Makris et al. 1975, Makris and Ginzburg 1987), and they all show the presence of a low shear-wave velocity (anomalous mantle) beneath the Afar crust.

The crust for path II can be approximated by a total thickness greater than 40 km, with a shear wave velocity ranging between 3.0 km/s to 3.9 km/s. The

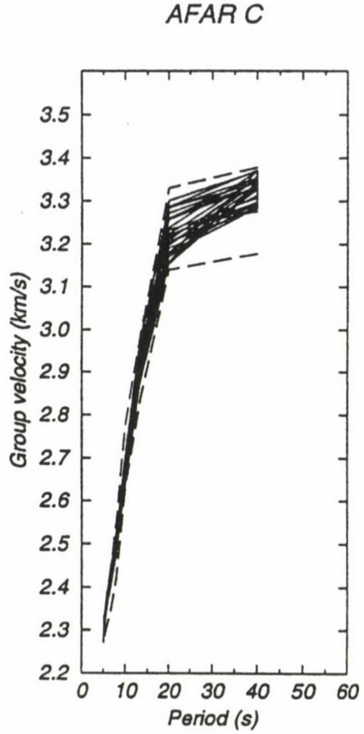


Fig. 11. Comparison of the observed dispersion curve for path I, corrected for the effect of the plateau part (dashed lines represent variability range of the observations), with the corresponding theoretical dispersion curves obtained from the crustal structures, solution of the inversion, shown in Fig. 10

lithospheric mantle shear-wave velocity varies in the range from 4.1 to 4.4 km/s lower than the typical one for stable continents and Precambrian shields (Panza 1980, Henkel et al. 1990). The crustal thickness is comparable with the 38 km proposed by Berckhemer et al. (1975) for the western Ethiopian plateau and it is thinner than the 48 km crustal thickness proposed by Searle and Gouin (1971). A layer by layer comparison of the results obtained here with the model of Berckhemer et al. (1975), shows that the significant difference between the two results is mostly limited to the top (both for thickness and shear wave velocity) crustal layers. This difference can be due either to the poor resolving power of the used data set with respect to the uppermost part of the model or to the different locations of the sampled areas.

As a general conclusion we may observe that the surface feature marking the edge of the western plateau seems to extend almost vertically at depth. This is a quite different feature with respect to the Central European rift system, where the rift boundaries seem to be not so sharp (Ahorner 1970, Panza et al. 1980, Panza 1984).

Acknowledgements

This work has been carried out at the Istituto di Geodesia e Geofisica (IGG), Universitá di Trieste, Italy. Fekadu Kebede and Tilahun Mammo thank the Swedish Agency for Research Co-operation with developing countries (SAREC) that through the International Centre for Theoretical Physics (ICTP) supported their work in Trieste. We would like to express our personal gratitude to all staff members of IGG for their help and co-operation during our stay. This research work is carried out as a contribution to the activities of the ILP Task Group II.4 (3DMET).

References

Ahorner L 1970: In: Graben problems, H Illies and St Mueller eds, Schweizerbart, Stuttgart, 155-166.

Barberi F, Varet J 1975: In: Afar Depression of Ethiopia, A Pilger and A Rösler eds, Vol. 1. Schweizerbart, Stuttgart, 174–178.

Barberi F, Varet J 1977: *Geol. Soc. Am. Bull.*, 88, 1251–1266.

Berckhemer H, Baier B, Bartelsen H, Behle A, Burkhardt H, Gebrande H, Makris J, Menzel H, Miller H, Vees R 1975: In: Afar Depression of Ethiopia, A Pilger and A Rösler eds, Vol. 1. Schweizerbart, Stuttgart, 89–107.

Berhe S M 1986: *J. Afr. Earth Sci.*, 5, 101–117.

Bohannon R G, Naeser C W, Schmidt D L, Zimmermann R A 1989: *J. Geophys. Res.*, 94, 1683–1701.

Bolt B A, Dorman J 1961: *J. Geophys. Res.*, 66, 2965–2981.

Bonatti E 1987: *Nature*, 330, 692–693.

Bonjer K P, Fuchs K, Wohlenberg J 1970: *Z. Geophys.*, 36, 287–297.

Brune J N, Dorman J 1963: *Bull. Seism. Soc. Am.*, 53, 167–210.

Dakin F M, Gouin P, Searle R C 1971: *Bull. Geophys. Obs.*, 13, 19–56.

Dziewonski A, Bloch S, Landisman M 1969: *Bull. Seism. Soc. Am.*, 59, 427–444.

Gabriels P, Snieder R, Nolet G 1987: *Geophysical Prospecting*, 35, 187–196.

Gass I G 1970: In: African magmatism and tectonics, T N Clifford and I G Gass eds, Edinburgh (Oliver and Boyd), 285–300.

Gass I G 1975: In: Afar Depression of Ethiopia, A Pilger and A Rösler eds, Vol. 1. Schweizerbart, Stuttgart, 10–18.

Gumper F, Pomeroy P W 1970: *Bull. Seism. Soc. Am.*, 60, 651–668.

Henkel H, Lee M, Lund C, Rasmussen T 1990: In: The European Geotraverse: Integrative studies, R Freeman, P Giese, St Mueller eds, European Science Foundation, Strasbourg, France, 1–47.

Hermann R, Russell D 1990: *Geophysics*, 55, 776–781.

Keilis-Borok V, Yanovskaja T 1967: *Geophys. J. R. Astr. Soc.*, 13, 223–234.

Levshin A, Pisarenko V, Pogreblnsky G 1972: *Annales Geophys.*, 28, 211–218.

Levshin A, Ratnikova L, e Berger J 1992: *Bull. Seism. Soc. Am.*, 82, 2464–2493.

Makris J, Ginzburg A 1987: *Tectonophysics*, 141, 199–214.

Makris J, Menzel H, Zimmermann J, Gouin P 1975: In: Afar Depression of Ethiopia, A Pilger and A Rösler eds, Vol. 1. Schweizerbart, Stuttgart, 135–144.

McMechan G, Yedlin M 1981: *Geophysics*, 46, 869–874.

McKenzie D P, Davies D, Molnar P 1970: *Nature*, 226, 243–248.

Mohr P A 1967: *Bull. Geophys. Obs. (Ethiopia)*, 11, 1–65.

Mohr P A 1970a: *Nature*, 228, 547–548.

Mohr P A 1970b: *J. Geophys. Res.*, 75, 7340–7352.

Panza G F 1980: In: Mechanisms of Continental Drift and Plate Tectonics. P A Davies and S K Runcorn (eds), Academic Press, 75–87.

Panza G F 1981: In: The Solution of the Inverse Problem in Geophysical Interpretation, R Cassinis ed., Plenum Press, 39–77.

Panza G F 1984: *Mem. Soc. Geol. It.*, 29, 5–13.

Panza G F, Mueller S, Calcagnile G 1980: *Pageoph.*, 118, 1209–1213.

Searle R C 1975: In: Afar Depression of Ethiopia, A Pilger and A Rösler eds, Vol. 1. Schweizerbart, Stuttgart, 113–120.

Searle R C, Gouin P 1971: *Bull. Seism. Soc. Am.*, 61, 1061–1071.

Talwani M, Sutton G H, Wortzel J L 1959: *J. Geophys. Res.*, 64, 1545–1555.

Valyus V P 1972: In: Computational Seismology, V I Keilis-Borok ed., Consult. Bureau, New York

Valyus V P, Keilis-Borok V I, Levshin A 1969: Determination of the upper-mantle velocity cross-section for Europe, Proc. Acad. Sci. USSR, 185, No. 3 (in Russian)

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 161–180 (1996)

CONSTRUCTION OF VELOCITY RELATIONS IN ABU MADI AND RUDEIS FORMATIONS IN SOME OIL FIELDS, EGYPT

ADEL A A OTHMAN[1] and ESAM A ABD EL-GAWAD[1]

[Manuscript received October 4, 1995]

Improved subsurface informations were achieved by applying a realistic velocity model. In El-Khalig and East Um El-Yusr oil fields, Rudeis Formation was penetrated by different wells. Another two wells in the Nile Delta, Hosh Isa #1 and Kafr El Dawar #1 are selected, in which Abu Madi Formation is exists and the V_{sh} values are introduced to acquire proper velocity relation. The interrelation between the velocity and the available petrophysical parameters was investigated. 15 suggested velocity models were studied and only 8 models give reasonable multiregression coefficients (C) relative to the sonic velocity. The depth is the common parameter in those 8 models. High C values are obtained with V15, which equals 0.863 in El-Khalig, and 0.457 in East Um El-Yusr. C values are raised after introduction of V_{sh} ratio in both Hosh Isa and Kafr El-Dawar wells.

Keywords: **petrophysical parameters; seismic velocity; shale content; velocity models**

Introduction

The interest to find new sources of energy, needs to improve the present technology. Exploration and production of hydrocarbons state a serious problem for search for energy. Seismic exploration has become accepted as the basic tool of the oil prospector. According to the capability of adapting to the increasing complexity of exploration requirements, reflection seismology has reached significant position. Geometric mapping of the subsurface which is considered the way to detecting of structural traps is one of the targets of seismic exploration. Connecting structural framework with a lithology, assists the investigator to locate stratigraphic traps. Estimate the quality of the deposits detects the presence of fluids and their interfaces.

The above mentioned remarkable progress is essentially due to the seismic velocities. Geophysicists began to make use of continuos velocity measurements (Schlumberger's sonic tool) in order to construct synthetic seismograms. The synthetic seismogram became a kind of bridge between geology and geophysics. Stacking velocity as one of the processing parameters showed a close approximation to the wave velocity as measured in the borehole (Cordier 1985). This velocity reconstitute accurately the depth horizons over the set of seismic investigations, both in two and three dimensions. Furthermore, velocity enables us to study the variations in facies of geological regions. Developments are still continuing and many applications which depend mainly upon seismic velocity occupied wide areas.

[1] **Al-Azhar University, Department of Geology, Post No. 11884, Nasr City, Cairo, Egypt**

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

That was a fast outline to understand the importance of the velocity parameter in modern seismology. Hence, the seismic velocity model in any proposed area for exploration, is the most important element for the prospector. Upon the velocity model depend many seismic computations and techniques. The geologic interpretation of the seismic time sections depends also on the understanding of the accurate velocity model through the seismic line. Briefly, the seismic model for exploration geophysicist is very important.

Many authors studied the wave velocity and its relation with the other petrophysical parameters in the medium of wave propagation. From many we can highlight the works of Wyllie et al. (1956, 1958), Biot (1956), Geertsma (1961), Domenico (1974, 1977), Marschall (1984). Wyllie and his coworkers related the wave velocity and the porosity of the medium of propagation. Marschall (1984) modified the so-called time average equation to another one by introducing more petrophysical parameters to the Wyllie's model to be a modified time average one. Domenico and others studied the parameters affected the longitudinal and shear wave velocities. These petrophysical parameters include the porosity, bulk density, compressibility factors, beside the water saturation in the medium. Xu and White (1995) derived a velocity model in terms of Kuster-Toksoz-Gassmann theories for clay-sand mixtures.

The main target of this study is to introduce a new velocity model relating the depth of the used parameter as a factor as well as the bulk density, the porosity, and water saturation to the velocity. Consequently we applied this idea on a hypothetical geological model to show the variation of the velocity with the combination of these parameters. Finally, the velocity models were calculated for real data measured and computed by logs recorded in drilled wells from some Egyptian oil fields.

Basic idea

Many researchers as Han et al. (1986), Klimentos and McCann (1990), Marion et al. (1992), and Mavko and Jizba (1991) observed the relation between the wave velocity in a medium and the petrophysical parameters. Velocity of propagation of seismic waves depends on a number of factors which are sometimes interconnected (Cordier 1985). One of the main factors is depth below the surface of the formation under consideration. Experimental relationships between velocity and depth of the burial have been investigated by numerous authors as Faust (1951, 1953), Wyrobeck (1959), and Acheson (1963, 1981). Approximately, the velocity of seismic waves increases with depth, but that not always true due to interference of other factors. The second factor is the lithology, in terms of mineralogical composition and texture at least. Birch (1942), Clark (1966), and Serra (1979) have published tables of velocities as function of lithology. Due to the overlap of the velocity ranges for different rocks, it is difficult to define the lithology by means of velocity only. The porosity may be the third factor. The influence of the porosity and the nature of the fluids contained in pores have been determined by many researchers. Wyllie et al. (1956, 1958) studied the relation between the velocity and the porosity of the medium, and introduced the well known relation as time average equation. His

Table I. List of symbols used for different parameters in the presented technique

C_m	=	compressibility of the matrix
C_b	=	compressibility of the rock free from fluid
C_f	=	compressibility of the fluid in the pores
C_g	=	compressibility of the gas in the pores
Φ	=	porosity of the rock
μ	=	shear modulus of the rock free from fluid
β	=	C_m/C_b
ϱ_b	=	bulk density = $\Phi\varrho_b + (1-\Phi)\varrho_m$
ϱ_m	=	density of the matrix
ϱ_f	=	density of the fluid in the pores
S_w	=	water saturation
V_W	=	wave velocity in water
V_G	=	wave velocity in gas
V_m	=	wave velocity in the matrix
V_f	=	wave velocity in the fluid
K	=	coupling coefficient between rock and fluid varying from 1.0 (no coupling) to ∞ (perfect coupling)

model may be expressed in the following form:

$$V = \frac{1}{\frac{\Phi}{V_f} + \frac{1-\Phi}{V_m}} \,. \tag{1}$$

It should be distinguished that the definition of all parameters used for equations in the present study is presented in Table I. Hence, the Wyllie's model in implicit expression could be written as:

$$V = f(\Phi) \,. \tag{2}$$

Marschall (1984) introduced a modified expression for the Wyllie's model and we can call it the modified time average model. In this model, the effects of the water saturation and the gases in porous medium was taken into consideration. The following equation expresses the Marschall's model as:

$$V = \frac{1}{\frac{\Phi S_W}{V_W} + \frac{\Phi(1-S_W)}{V_G \log\left(\frac{C_G}{C_m}\right)} + \frac{1-\Phi}{V_m}} \,. \tag{3}$$

Although, Eq. (3) may be rewritten in the coming form:

$$V = f(\Phi, S_W, C_G, C_m) \,. \tag{4}$$

Many authors have interest to relate both the longitudinal and shear velocities with the petrophysical parameters of the medium. Biot (1956) published his theory to express the velocity in a partially saturated media, relating it with different petrophysical parameters. In addition to Biot, Geertsma (1961) and Domenico

(1974, 1977) studied the effect of the porosity, bulk density, shear modulus, and the compressibility factors. Domenico (1977) introduced the following expression:

$$V_P = \sqrt{\left[\left(\frac{1}{C_b} + \frac{4}{3}\mu\right) + \frac{\frac{\Phi\varrho_b}{K\varrho_f} + (1-\beta)\left(1-\beta-\frac{2\Phi}{K}\right)}{(1-\Phi-\beta)C_m + \Phi C_r}\right]\left(\frac{1}{\varrho_b\left(1-\frac{\varrho_f\Phi}{\varrho_b K}\right)}\right)} \tag{5}$$

and

$$V_s = \sqrt{\frac{\mu}{\varrho b\left(1-\frac{\Phi\varrho f}{K\varrho b}\right)}}\ . \tag{6}$$

Hence, in a functional form Eqs (5) and (6) may be expressed as follows:

$$Vp = f(\Phi, \varrho_b, \mu, K, C_b, C_m, C_f) \tag{7}$$

and

$$VS = f(\Phi, \varrho_b, \mu, K)\ . \tag{8}$$

The first term of the Eq. (5) covers the effect of the rock matrix and the second term covers the effect of the fluid contained in the pores. Practical application of this formula assists to distinguish between two main types of rocks, consolidated and unconsolidated rocks. In the unconsolidated rocks, the nature of the contained fluids has a direct influence on the velocity of a formation. But in the consolidated rocks, the physical characteristics of the rock are much less affected by depth below the surface and the fluids contained in the reservoir has little effect on the velocity. The state of the fractures in the rock is the last factor which affect the wave velocity. Gardner et al. (1974) demonstrated experimentally the effect of the presence of microfractures on the wave velocity, and concluded that fractures lead to a reduction in the velocity of the seismic pulse through the rock.

Theoretical model

A model is a representation of a natural phenomenon or process (Koch and Link 1970). From the previous described velocity models, clearly there is a good relation between the wave velocity and the different petrophysical parameters of the medium. This encourages us to construct a velocity model applying the numerous petrophysical parameters concerning the medium in which the waves are propagated. These parameters are the depth, bulk density, porosity, and the water saturation. The available velocity models in the literature include one, two or three petrophysical parameters only. We try to introduce a new velocity model to describe the wave velocity in a medium by relating it to the depth, the porosity, the density, and the water saturation.

The velocity is estimated applying the linear multiple regression technique. The statistically estimated velocity is compared and correlated with the measured sonic velocity. Table II describes the suggested 15 velocity models which installed from the combinations of the four parameters against the velocity. The 15 models may be classified into four groups as shown in Table II. Four examples in terms of mathematical form of the suggested linear velocity models are shown in Table III.

A test was needed to clarify the capability of the suggested velocity models. Hypothetical velocity model was created and called MODEL5. Two cases of this designed model summarize the relation between the velocity and the depth. The

Table II. Designation of different forms of the suggested elastic velocity models. Correlation coefficients (C) of these suggested models computed with respect to the initial velocity (V0), for both MODEL5A and MODEL5B

No.	Velocity function	Symbol	C_{Model5A}	C_{Model5A}
15	$f(Z, \varrho_b, \Phi, S_w)$	V15	0.9694	0.8748
14	$f(Z, \varrho_b, \Phi)$	V14	0.9020	0.8746
13	$f(Z, \varrho_b, S_w)$	V13	0.9587	0.8640
12	$f(Z, \Phi, S_w)$	V12	0.9399	0.8740
11	$f(\varrho_b, \Phi, S_w)$	V11	0.5461	0.3604
10	$f(Z, \varrho_b)$	V10	0.9014	0.8637
9	$f(Z, \Phi)$	V9	0.8883	0.8739
8	$f(Z, S_w)$	V8	0.9269	0.8628
7	$f(\varrho_b, \Phi)$	V7	0.4134	0.3598
6	$f(\varrho_b, S_w)$	V6	0.4996	0.2899
5	$f(\Phi, S_w)$	V5	0.3115	0.2668
4	$f(Z)$	V4	0.8866	0.8620
3	$f(\varrho_b)$	V3	0.3974	0.2858
2	$f(\Phi)$	V2	0.1798	0.2639
1	$f(S_w)$	V1	0.1763	0.1281

Table III. Examples of velocity models expressed in functional forms. Four models, every one represents one group of the discussed 15 models

[1] $V = f(Z)$:
$\Rightarrow \text{V4} = a_0 + a_1 Z$

[2] $V = f(Z, \varrho_b)$:
$\Rightarrow \text{V10} = b_0 + b_1 Z + b_2 \varrho_b$

[3] $V = f(Z, \varrho_b, \Phi)$:
$\Rightarrow \text{V14} = c_0 + c_1 Z + c_2 \varrho_b + c_3 \Phi$

[4] $V = f(Z, \varrho_b, \Phi, S_w)$:
$\Rightarrow \text{V15} = d_0 + d_1 Z + d_2 \varrho_b + d_3 \Phi + d_4 S_W$

first case (MODEL5A) describes increasing velocity generally with depth as shown in Fig. 1a. The second one (MODEL5B) shows the velocity when it involves a low velocity layer or zone within the succession of layers as shown in Fig. 1c. As a function of the effect of the petrophysical parameters as density, porosity, water saturation besides depth on the estimated velocity by utilizing the derived velocity models, a group of synthetic seismograms were calculated and plotted side by side for the designed velocity model and shown in Figs 1b and 1d.

The correlation coefficient between the suggested velocity models and the initial (reference) one gives a reasonable results for that type of velocity models which involving the parameter Z (depth). Values of the calculated correlation coefficients (C) for the hypothetical models MODEL5A and MODEL5B are displayed in Table II. The other suggested velocity models, which includes porosity, or water saturation or bulk density in separate case or conjointly, produced low correlation coefficients compared with the measured one. It should be noted that the reference model (or the initial model,V0) in the real data means measured sonic velocity and is utilized for calibrating any calculated one. Therefore, we shall apply the models V4, V8, V9, V10, V12, V13, V14, and V15, to compute the synthetic seismograms. The amplitudes in the different computed synthetic seismograms are normalized due to V0. It is obvious that the shape and the amplitude of the seismic events are changed as the change of the suggested velocity model. In the case of increasing velocity with depth, most of the signals appear on the synthetic seismograms. But that is different for the low velocity layer example, where the amplitude of the signals is weakened after the low velocity layer.

Preparation of used data from well logs

In the present study, used parameters were prepared and calculated applying software programs as SATA and DEMY-LOG (Abd El-Gawad 1992). These programs were concluded from the broad range of log interpretation techniques. Some corrections performed for the calculated parameters. Environmental corrections were applied to density, neutron, and gamma ray logs. The determination of the effective porosity is mainly based on the combination between neutron and density logs. The water saturation was calculated through the so-called "INDONESIA" equation (Schlumberger 1972). The hydrocarbon saturation (S_h) and its subdivisions; the residual hydrocarbon saturation (S_{hr}) and the movable hydrocarbon saturation (S_{hm}) were determined according to the normal analytical relations (Schlumberger 1974):

$$\left.\begin{aligned} S_h &= 1 - S_w \\ S_{hr} &= 1 - S_{X_0} \\ S_{hm} &= S_h - S_{hr} \end{aligned}\right\} . \tag{9}$$

The average value of each of the V_{sh}, Φ_E and S_w was calculated for each zone in each well. V_{sh} and Φ_E are the shale content and effective porosity respectively.

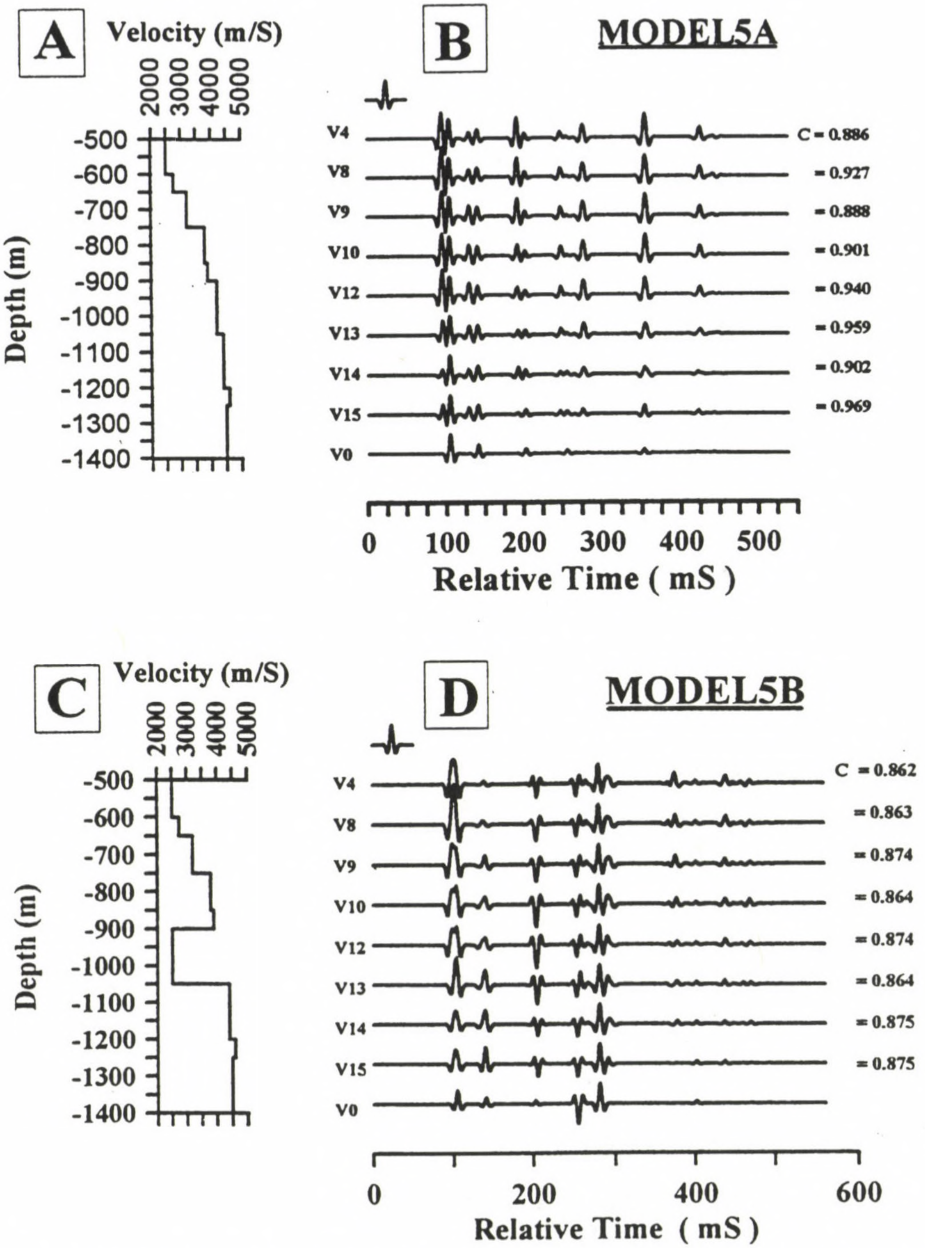


Fig. 1. A) MODEL5A is a hypothetical velocity model in which the velocity increases by increasing depth. Upper part of the figure shows special synthetic seismograms for the presented velocity models. B) MODEL5B is a hypothetical velocity model which shows a low velocity layer. The presented synthetic seismograms declare the decrease of the amplitudes of waves passed through the low velocity layer

Estimation of wave velocity models for real data

Gulf of Suez area

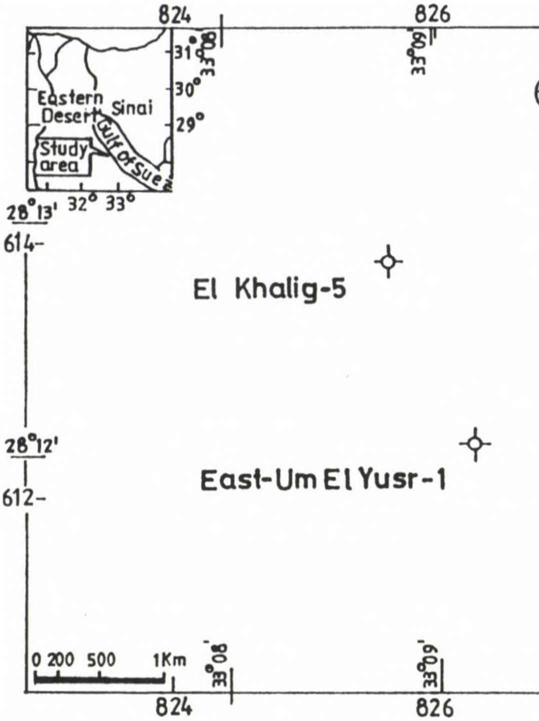
The suggested velocity models were applied for real measured data in two locations. The computed practical velocity models are derived for East Um El-Yusr #1 and El Khalig #5 in the western side of the Gulf of Suez. Location map of the studied area including the selected two wells are shown in Fig. 2a. The stratigraphic succession of the rock units in this area was configured by means of electric logs. Many investigators studied the geology of the Gulf of Suez, we note from them El-Heiny and Morsi (1992), and Harwood and Tewfik (1992). Economically, the Miocene sediments of the Gulf of Suez are very important. Most of petroleum reserves are related to these different sediments. The Lower Miocene is represented by Gharandal Group which overlies unconformably the Eocene Limestones and conformably underlies Ras Malaab Group. This group is classified into Rudeis Formation and Nukhul Formation from top to base. Rudeis Formation is the main target of the present study. Therefore, a close up details will be completed in the next item.

Rudeis Formation:

The thickness of Rudeis Formation in the present study is varying between 21 m and 708 m (East Um El Yusr #1). Based on lithology, applying the log correlation data, it can be differentiated into two informal units, upper and lower with unconformity surface (Mid-Rudeis unconformity). The upper unit consists of shale, sandstone, and limestone. The lower unit consists of calcareous shale and calcareous sandstone with a few streaks of limestone. The lower part of Rudeis Formation is absent probably due to non-deposition.

East Um El-Yusr #1

The suggested velocity models are calculated and presented in Fig. 3 and Table IVa for this well. In Fig. 3, the influence of introduction different petrophysical parameters on the wave velocity are illustrated by means of synthetic seismograms, as shown in Fig. 3b. The synthetic seismograms are displayed according to the relative time. That means at zero time starts the upper surface of the formation. If we know the preceding velocities, then we can calculate exactly the start time on the time-axis. It is clear to observe the variation in the phase, shape and amplitude of the wavelets. It should be mentioned that all plotted synthetic seismograms are normalized due to the initial measured velocity in E. Um El-Yusr #1. The aim of this illustration form to show the variations by every applied velocity model on the amplitude side by side to the initial model (Vo). Figure 3 shows some selected velocity models compared with the measured one (the solid line in the figure). V4, in which the velocity relates depth alone, is displayed in Fig. 3c. V10 which describes the influence of the bulk density in addition to the depth, as shown in Fig. 3d. The third computed velocity model is V12 to relate the velocity with the depth, the porosity and the water saturation, plotted in Fig. 3e. In case of introducing



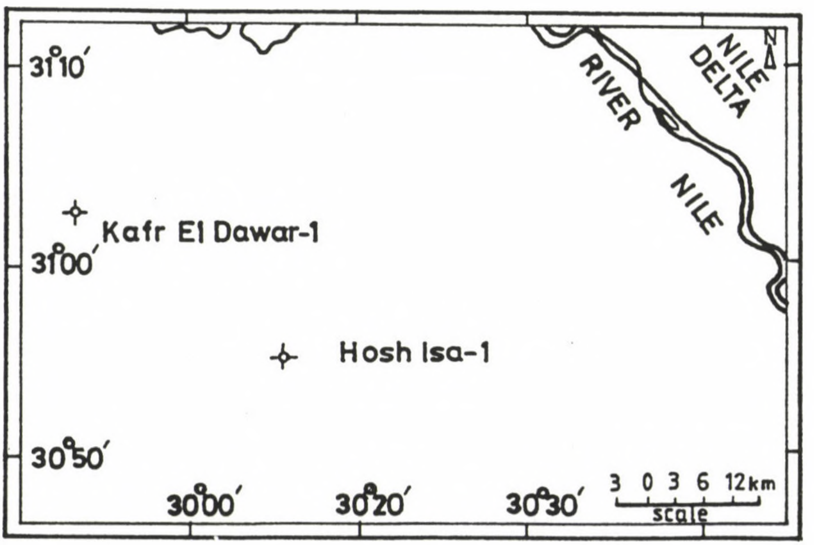


Fig. 2. A) Location map shows the sites of East Um Yusr #1 and El Khalig #5 wells, Gulf of Suez, Egypt. B): Location map pointed the positions of Kafr El Dawar #1 and Hosh Isa #1 wells. Abu Madi Formation presented in the two wells is characterized by high shale content

four parameters, as by V15, the bulk density and the depth are enclosed besides the porosity and the water saturation, Fig. 3f. For every constructed velocity model a complete statistical examinations are applied. The T- and F-tests are calculated for the available data. Analysis of variance is studied within confidence interval equal to 95 %. The T-test is applied for every parameter inserted in the computed velocity model. The tolerance between every parameter in the velocity model and the other parameters is calculated to show the interconnected relations. F-test is calculated for the computed velocity model to evaluate and determine how match the model with the normal F-distribution. Accordingly, the multiple regression coefficient, between the estimated velocity model and the measured velocity one, for every model is calculated and displayed with every plot. In Fig. 3c, V4 shows multiple regression coefficient (C) equals 0.19. The best C-value is given by V15, but the estimated velocity model shows poor matching with the derived model from the sonic log measurements. This poor matching may be due to the effect of another petrophysical parameters and are not included in the suggested model. The final results by means of the constructed velocity models for the studied well, are displayed in Table IVa. From these results, we can state, for example, if you have information only about the depth location of the Rudeis Formation in the E. Um El-Yusr #1 or the most near adjacent wells you can estimate the velocity values as function of depth. The estimated velocities matched only 19 % from the measured or the real one. The estimated velocity applying the velocity model V15 is different. Estimated velocities by means of this model (V15) matches 45.7 % from the true measured velocities derived from sonic logs.

El-Khalig #5

Through Figs 4a to 4f and Table IVb, El-Khalig #5 is displayed. In Fig. 4a, the measured velocity is illustrated in addition to the computed synthetic seismograms for the selected velocity models applying different petrophysical parameters measured or calculated in the well, as shown in Fig. 4b. The measured velocity shows decrease of the velocity with increase depth, in the upper part of the formation. In the lower part, the velocity increased with depth. This affects the appearance of the amplitudes on the synthetic seismograms specially in the lower part of the formation. Estimated velocities applying V4 model gives poor matching with the measured velocities, which the correlation coefficient equals 0.383, as shown in Fig. 4c. The estimated velocities show good similarity with the measured one for models V10, V12, and V15. For V15, estimated velocities, harmonize 86.3 % from the derived velocities from sonic logs in the present well, as displayed in Fig. 4f. Results of the constructed velocity models for El-Khalig #5 is illustrated in Table IVb. According to these results, the calculated velocities due to V4 leads to 38.3 % from the measured (actual) velocities. But applying V10 gives 86.1 % real velocities. The matching value is decreased in case of V12, which equals 73.4 % from the measured velocities. For V15, increased the correlation coefficient slightly to give 86.3 % of the real velocities. That is the best value could be obtained according to the available data in this well.

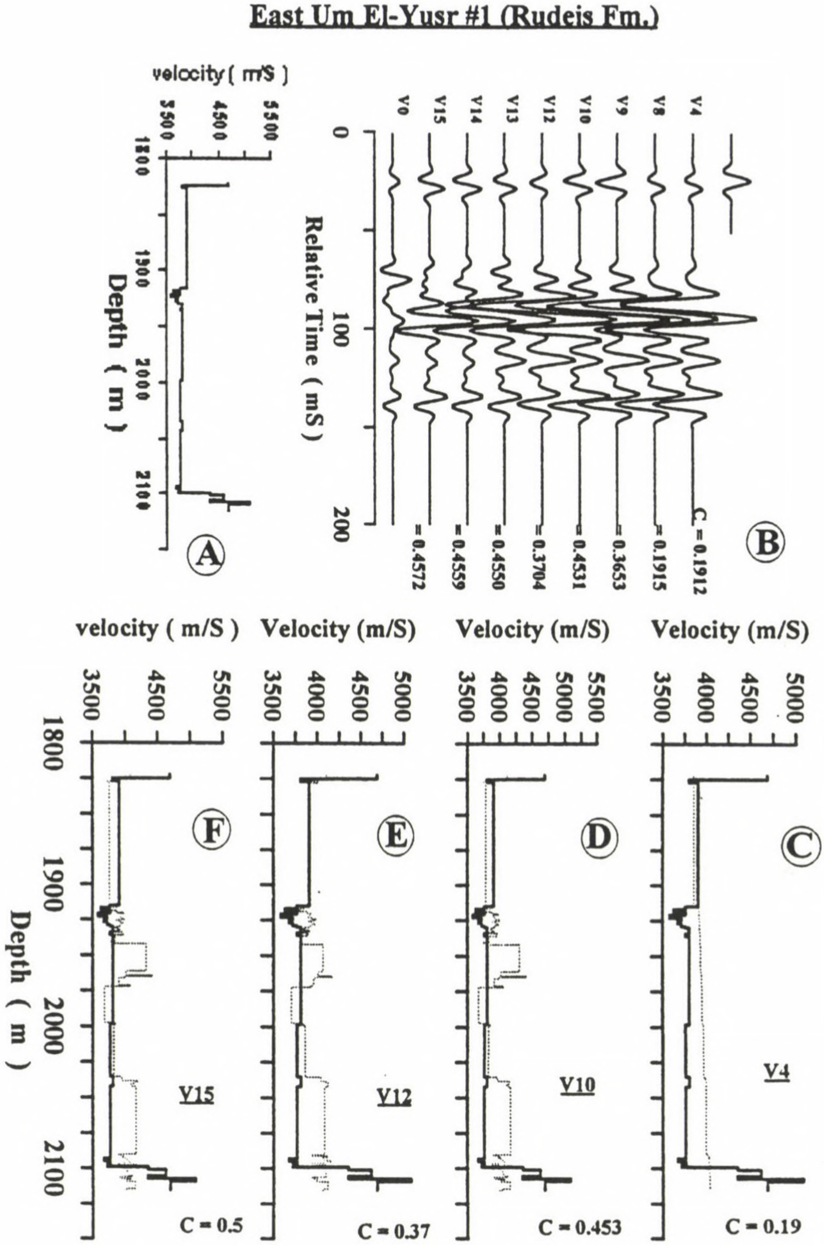


Fig. 3. A) Derived interval velocities in Rudeis Formation from sonic log in East Um Yusr #1. B) Synthetic seismograms illustrate some of the suggested velocity models. Some selected velocity models as V4 (C), V10 (D), V12 (E), and V15 (F) are displayed in dashed lines correlated with V0 (solid line). Correlation coefficient (C) between every mentioned model and V0 is shown with every one

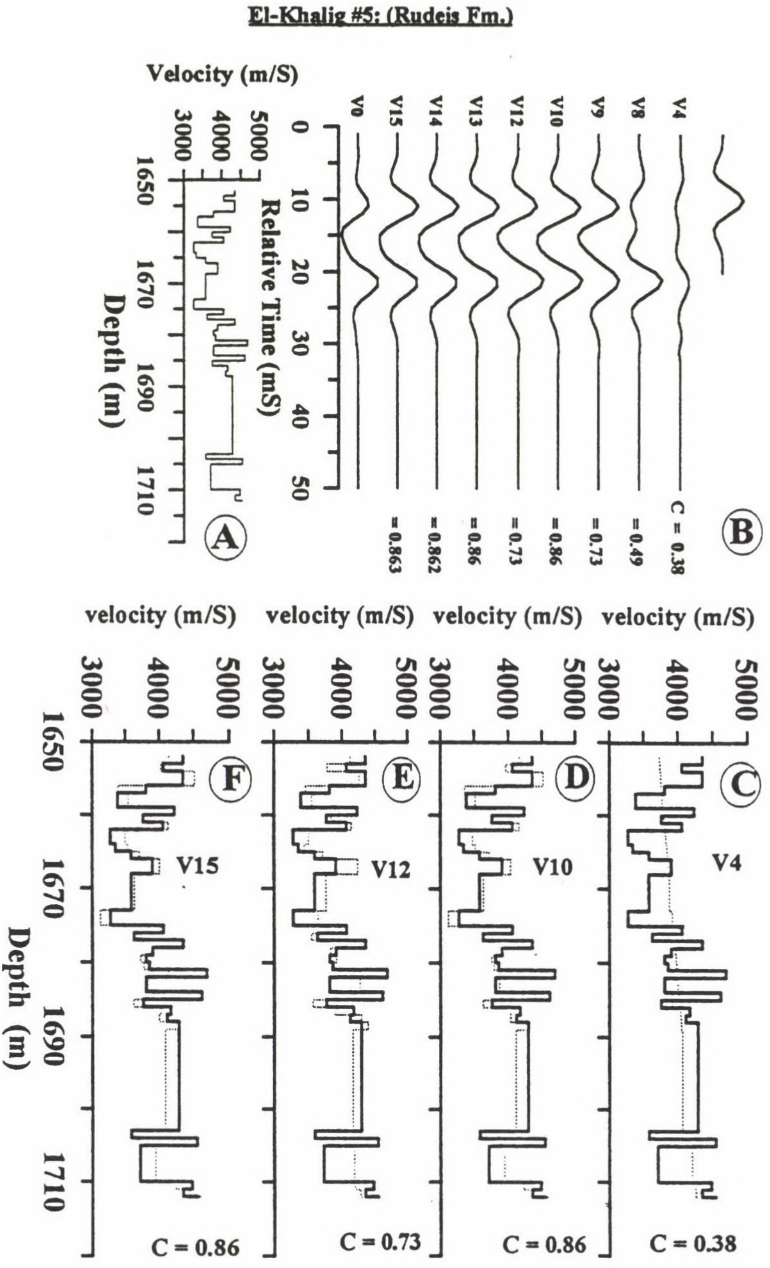


Fig. 4. A) Derived interval velocities in Rudeis Formation from sonic log in El Khalig #5. B) Synthetic seismograms for different velocity models and the computed correlation coefficient are displayed. The initial velocity model V0 (solid line) is compared with V4 (C), V10 (D), V12 (E), and V15 (F) as dashed lines

Table IV. A) Results of the constructed velocities due to different suggested models and its relationships with the proper petrophysical parameters in Rudeis Formation, East Um Yusr #1. B) Constructed velocity models with special petrophysical parameters are presented for Rudeis Formation in El Khalig #5. C) Constructed velocity models related to different petrophysical parameters are presented for Abu Madi Formation in Hosh Isa #1. D) Constructed velocity models related to different petrophysical parameters are presented for Abu Madi Formation in Kafr El Dawar #1

V15	=	$-1426.65 - 0.06Z + 2.38\varrho_b + 4.94\Phi - 219.57S_W$	A
V14	=	$-1726.25 - 0.06Z + 2.4\varrho_b + 5.59\Phi$	
V13	=	$-419.8 - 0.12Z + 2.05\varrho_b - 271.51S_W$	
V12	=	$+4764.67 - 0.09Z - 19.74\Phi - 404.67S_W$	
V10	=	$-631.9 - 0.13Z + 2.04\varrho_b$	
V9	=	$+4356.24 - 0.09Z - 19.11\Phi$	
V8	=	$+2773.47 + 0.64Z - 66.94S_W$	
V4	=	$+2714.77 + 0.63Z$	[East Um Yusr #1]
V15	=	$-5554.03 - 0.37Z + 4.27\varrho_b + 3.19\Phi - 85.37S_W$	B
V14	=	$-5785.05 - 0.35Z + 4.32\varrho_b + 5.73\Phi$	
V13	=	$-5159.45 - 0.35Z + 4.12\varrho_b - 111.37S_W$	
V12	=	$-389.37 + 3.00Z - 51.38\Phi - 255.58S_W$	
V10	=	$-5016.38 - 0.31Z + 3.99\varrho_b$	
V9	=	$-921.53 - 3.16Z - 45.46\Phi$	
V8	=	$-9207.47 + 7.56Z + 589.12S_W$	
V4	=	$-11010.05 + 8.93Z$	[El Khalig #5]
V15	=	$+1978.5 - 0.01Z + 0.19\varrho_b - 9.84\Phi + 6.71S_W$	C
V14	=	$+1776.06 + 0.04Z + 0.36\varrho_b - 5.58\Phi$	
V13	=	$+1716.29 - 0.37Z + 0.85\varrho_b + 3.74S_W$	
V12	=	$+2508.00 - 0.04Z - 10.69\Phi + 6.86S_W$	
V10	=	$+1654.94 - 0.23Z + 0.77\varrho_b$	
V9	=	$+2787.9 - 0.01Z - 7.04\Phi$	
V8	=	$+4953.01 - 0.74Z + 3.07S_W$	
V4	=	$+4646.69 - 0.59Z$	[Hosh Isa #1]
V15	=	$-464.97 + 0.5Z + 0.002\varrho_b - 5.02\Phi + 9.3S_W$	D
V14	=	$+3109.06 - 0.1Z + 0.1\varrho_b - 6.08\Phi$	
V13	=	$-2209.1 + 0.69Z + 0.25\varrho_b + 10.25S_W$	
V12	=	$-463.4 + 0.5Z - 5.03\Phi + 9.3S_W$	
V10	=	$+1405.28 + 0.06Z + 0.42\varrho_b$	
V9	=	$+3235.01 - 0.08Z - 6.42\Phi$	
V8	=	$-2302.18 + 0.81Z + 10.76S_W$	
V4	=	$+1558.93 + 0.21Z$	[Kafr El Dawar #1]

West of the Nile Delta area

In this area, two wells, in which the Abu Madi Formation exist, were studied. Abu Madi Formation is the most promising formation in the Nile Delta, since many gas discoveries were proved within its porous sands. The effect of shale content on the calculation of the velocity function is investigated. The studied wells are Kafr el Dawar #1 and Hosh Isa #1, as shown in Fig. 2b.

Abu Madi Formation:

The Formation is characterized by a series of thick layers of sands with interbedded thin shales. It represents the beginning of the Pliocene transgressive phase. This formation was deposited in deltaic environment and grades up into a shallow marine environment (Schlumberger, Egypt, 1984). The shale content increases upwards throughout the formation.

Hosh Isa #1

For Hosh Isa #1, the suggested velocity models are studied and exposed through Figs 5a to 5f and Table IVc. The variation in the phase, shape and amplitude of the wavelets is observable in the studied well. Figures 5c to 5f show some selected velocity models compared with the measured one (the solid line in the figures). V4, in which the velocity relates depth alone, is displayed in Fig. 5c. V10 which describes the influence of the bulk density in addition to the depth. The third computed velocity model is V14 to relate the velocity with the depth, the porosity and the water saturation. In case of introducing four parameters, as by V15, the bulk density and the depth are enclosed besides the porosity and the water saturation. In Fig. 5, V4 shows multiple regression coefficient (C) equals 0.23. The best C-value is given by V15, but the estimated velocity model shows moderate matching with the derived model from the sonic log measurements. This type of matching may be due to the effect of another petrophysical parameters and are not included in the suggested model. The final results by means of the constructed velocity models for the studied well, are displayed in Table IVc. From these results, we can state, for example, if you have information only about the depth location of Abu Madi Formation in the Hosh Isa #1 or the most near adjacent wells you can estimate the velocity values as function of depth. The estimated velocities using V4 matched only 23 % from the measured or the real one. The estimated velocity applying the velocity model V15 is different. Estimated velocities by means of this model (V15) matches 61.4 % from the true measured velocities derived from sonic logs.

Kafr El Dawar #1

Kafr El-Dawar #1 is displayed in Figs 6a to 6f as well as in Table IVd. In Fig. 6a, the measured velocity model is illustrated in addition to the computed synthetic seismograms for the selected velocity models applying different petrophysical parameters. The measured velocities show generally increase of the velocity with increase depth. In the upper part of the formation, a thin strip shows higher velocity than adjacent layers preceded by low velocity one. This change in velocity affects the appearance of seismic amplitudes as shown in Fig. 6b. Estimated velocities applying V4 model gives poor matching with the measured velocities, which the correlation coefficient equals 0.1, as shown in Fig. 6c. Estimated velocities show few similarity with the measured one for models V10, V14, and V15. For V15, estimated velocities, harmonize 58.7 % from the derived velocities from sonic logs in the present well. Results of the constructed velocity models is illustrated

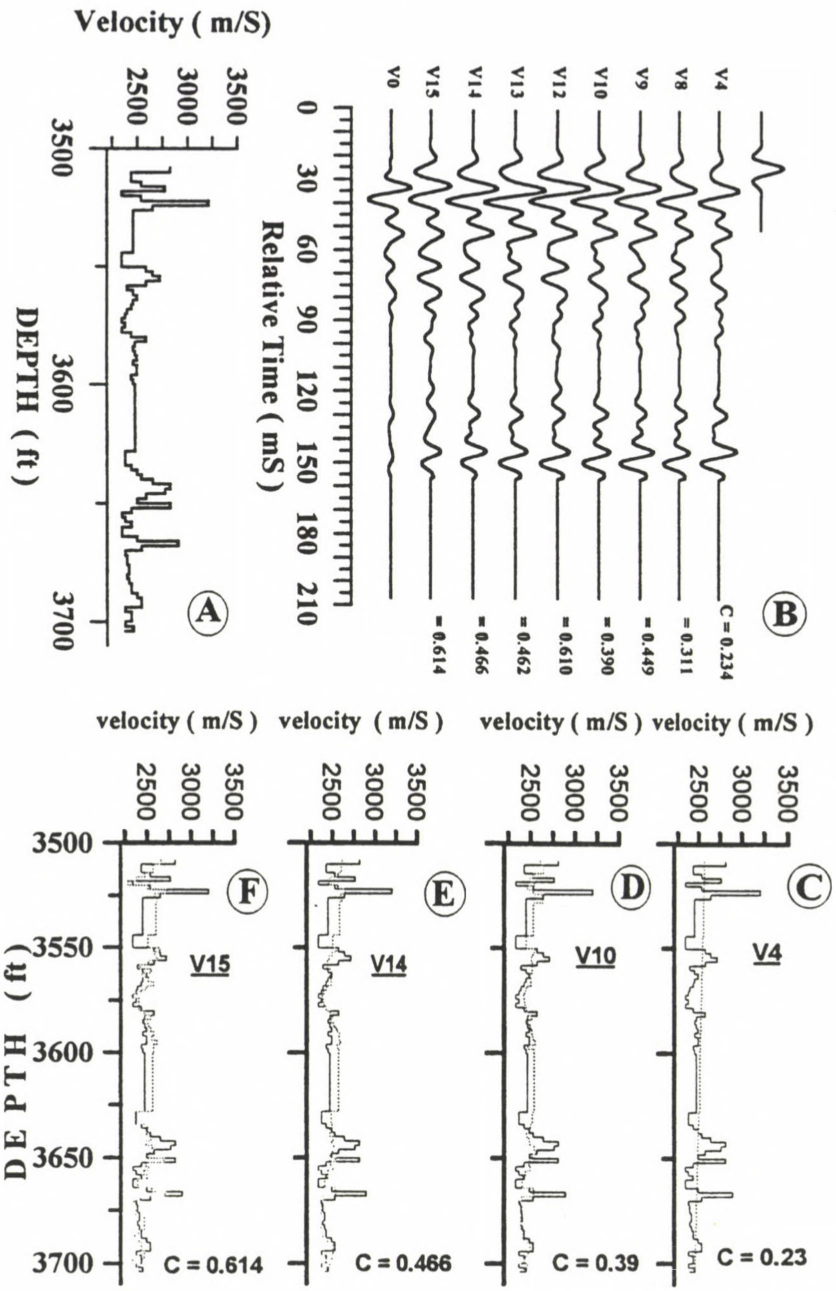


Fig. 5. A) Derived interval velocities in Abu Madi Formation from sonic log in Hosh Isa #1. B) Synthetic seismograms illustrate different velocity models correlated to the initial velocity (V0) with C-values. Initial velocity model (solid line) associated with V4 (C), V10 (D), V14 (E), and V15 (F), which expressed in dashed lines. Correlation coefficient (C) for V15 is the improved one

in Table IVd. According to these results, the calculated velocities due to V4 leads to 10 % from the measured (actual) velocities. But applying V10 gives 24 % real velocities. The matching value is increased in case of V14, which equals 42.1 % from the measured velocities. For V15, increased the correlation coefficient to give 58.7 % of the real velocities. That is the best value could be obtained according to the available data in this well.

Effect of shale content on velocity models

In studying a physical phenomenon, if we introduce all appropriate variables in an equation that will increase and raise the value of the correlation coefficient for the equation. To which extent is any parameter or variable is effective for constructing the equation. That is the role of the applied statistical techniques through F- and T-tests and others. By introducing the V_{sh} ratio specially to the data measured in the area west of the Nile Delta, according to highly shale content in rocks of Abu Madi Formation, the two wells were studied. Figures 7a and 7b describe the shale ratio (V_{sh}) in both Hosh Isa #1 and Kafr El Dawar #1. In Hosh Isa #1, as displayed in Fig. 7a, the V_{sh} value ranges from 3 % to 49 %. Introduction of V_{sh} parameter to compute the velocity model V15 increases the likelihood circumstance of the calculated velocity with the actual (sonic) one. In case of Hosh Isa #1, the new constructed velocity model (V15A) increased the correlation coefficient to 0.926 as measure of likelihood or matching between the new model and the sonic velocity.

In the same figure, the water saturation (S_w) shows decrease by increasing the V_{sh} value at the same depth. The mobile hydrocarbon saturation (S_{hm}) show similar habitat as V_{sh} by increasing the depth. The new velocity model V15A may be expressed as follows:

$$\mathrm{V15A} = 3238.71 + 0.07Z + 0.13\varrho_b - 25.57\Phi + 0.05S_W - 15.33V_{sh} \,. \tag{10}$$

Equation (10) represents a powerfully velocity model for Hosh Isa #1 which recovers 92.6 % from the actual (sonic) velocity. Results of introduction of V_{sh} values in the constructed velocity model for Kafr El Dawar #1 is considered in Fig. 7b. The correlation coefficient after applying V_{sh} ratio equals 0.96 and higher than that of velocity model without shale ratio. It should be mentioned that the shale content in this well ranges from 1 % to 55 % and it considered high, as shown in Fig. 7b. The suggested new velocity model including V_{sh} values may be expressed as:

$$\mathrm{V15A} = 2815.01 + 0.21Z - 0.01\varrho_b - 29.04\Phi + 1.25S_W - 20.88V_{sh} \,. \tag{11}$$

Applying this new velocity model for Kafr El Dawar #1 recovers 96 % from the actual velocity value.

Conclusions

The present study focused its target to construct velocity models. The constructed models utilize different petrophysical parameters are compared with measured velocity values. The applied petrophysical parameters are the porosity, bulk

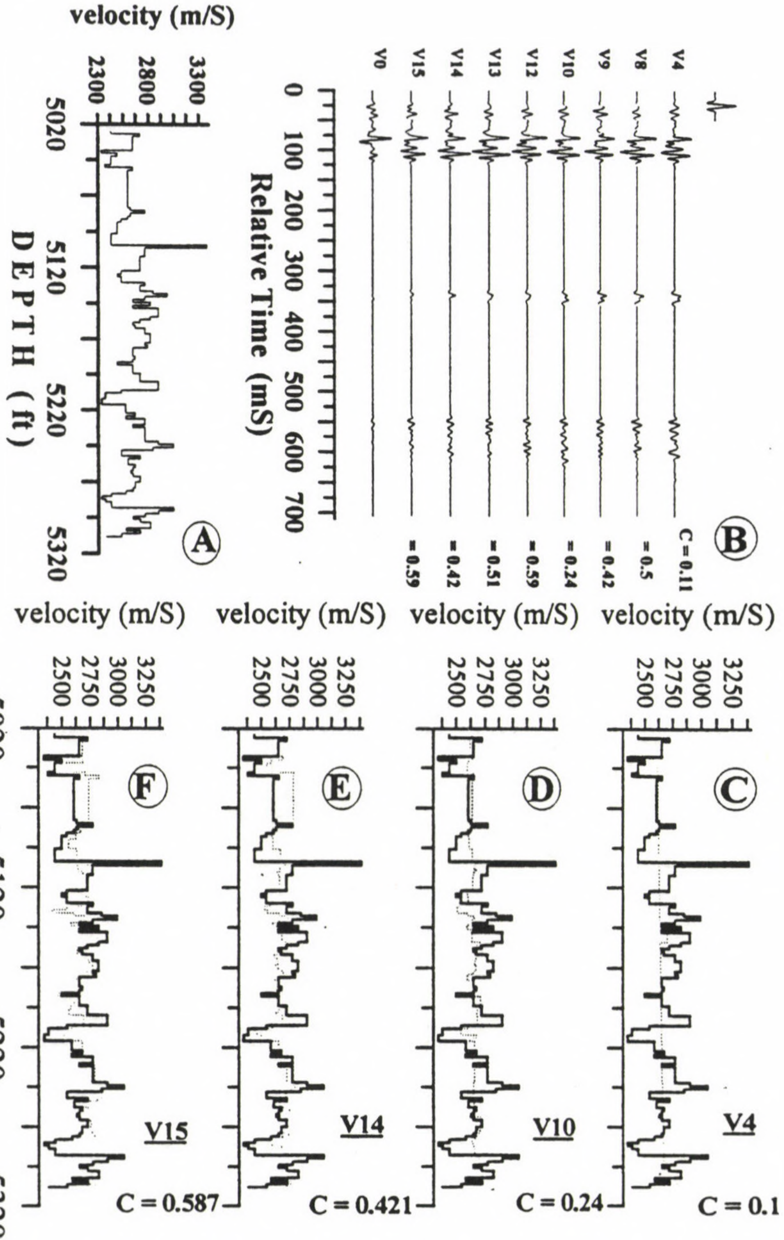


Fig. 6. A) Derived interval velocities in Abu Madi Formation from sonic log in. Kafr El Dawar #1. B) Synthetic seismograms illustrate selected velocity models correlated to the initial velocity (V0) with the computed C-values. The V0 velocity model (solid line) associated with V4 (C), V10 (D), V14 (E), and V15 (F)

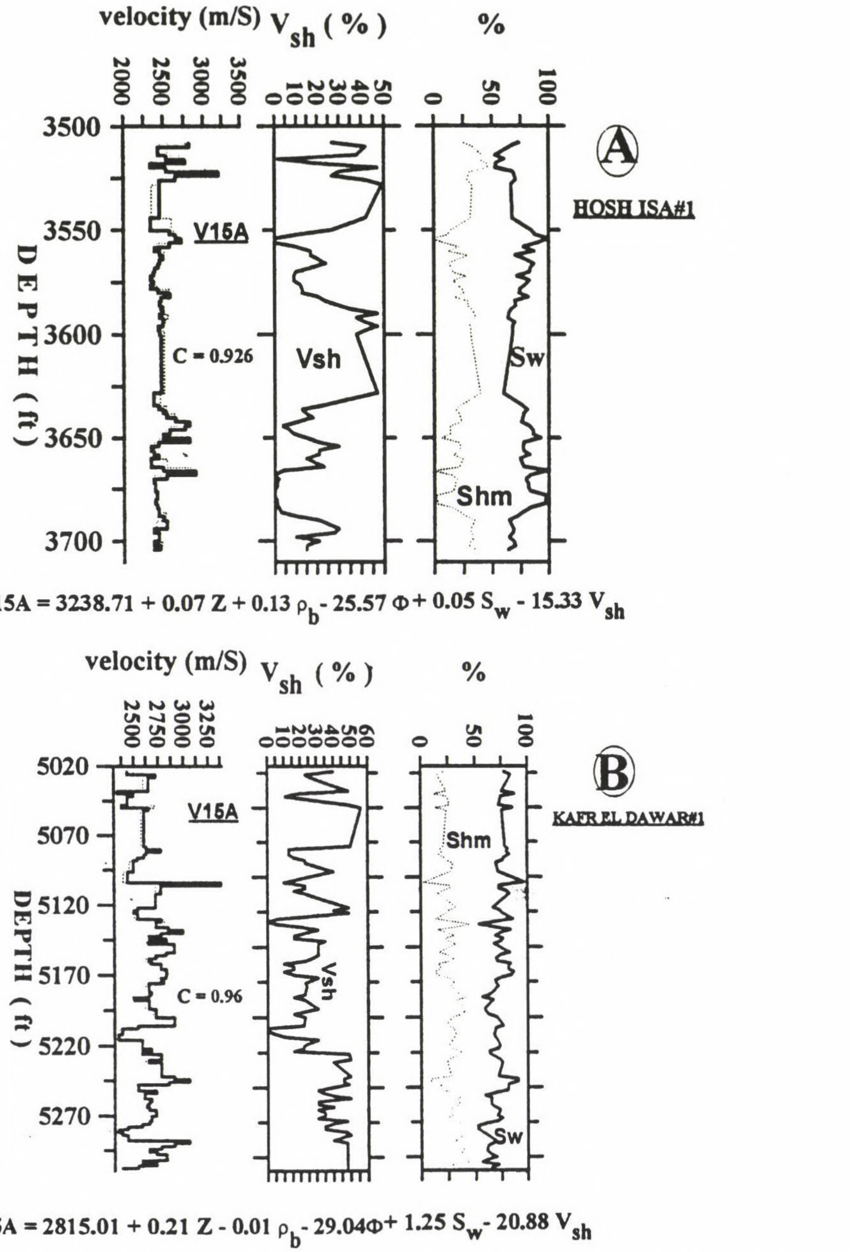


Fig. 7. A) Effect of introduction V_{sh} ratio on the velocity model V15. V15A, which the enhanced velocity model (dashed line), is associated with V0 (solid line). ϱ_b and Φ in the exposed equation expressed bulk density and porosity, respectively. V_{sh} ratio, S_{hm}, and S_w are exhibition as function with depth for Abu Madi Formation in Hosh Isa #1. B) V15A shows enhanced velocity model (dashed line), which associated with V0 (solid line). V_{sh} ratio, S_{hm}, and S_w are presented as for Abu Madi Formation in Kafr El Dawar #1

density, water saturation, in addition to the subsurface depth of the considered formation. Fifteen velocity models are suggested. Only 8 models, in which the depth as parameter is common, are chosen for investigation. V15 gives velocities in good matching with the measured one. The other models, show different degree of similarity with the real measured data. In case of V4 model, whereas the main parameter is the depth, shows poor harmonization with the derived velocities from sonic logs.

Results may be enhanced if we applied the constructed velocity models to many wells involving the considered formation. Also, the present study deals with elastic velocities neglecting the effect of the attenuation as medium character. This of course affects the recovery of the real velocity values by the estimated one.

Introduction of the shale content into the computations of the constructed velocity models shows good results. This applied to the wells of the west Nile Delta area due to the existence of the shale in higher ratios. The correlation coefficient for Hosh Isa #1 raised from 0.614 before considering of V_{sh} value to 0.926 after introduction of it. In case of Kafr El Dawar #1, the correlation coefficient increased from 0.587 without V_{sh} parameter to 0.960 after introducing it in the calculations.

References

Abd El-Gawad E A 1992: Reservoir rock study in Zeit Bay Field, Gulf of Suez, Egypt. Ph.D. Thesis, Al-Azhar University, Cairo, Egypt

Acheson C H 1963: *Geophysics*, 28, 5, 984–909.

Acheson C H 1981: *Geophysics*, 46, 5, 707–716.

Biot M A 1956: *J. Acoust. Soc. Am.*, 28, 168–191.

Biot M A 1962: *J. Acoust. Soc. Am.*, 34, 1254–1264.

Birch M 1942: Handbook of physical constants. Geological Society of America

Clark S P 1966: Handbook of physical constants. Geol. Soc. of America, memoir 97.

Cordier J-P 1985: Velocities in reflection seismology. D Reidel Pub., Dordrecht

Domenico S N 1974: *Geophysics*, 39, 759–769.

Domenico S N 1977: *Geophysics*, 42, 1339.

El-Heiny I, Morsi S 1992: Stratigraphic correlation of Neogene sediments in the Eastern Nile Delta and Gulf of Suez, Egypt. Abs. of E.G.P.C. 11-th exploration and production conf., Cairo, Egypt

Faust L Y 1951: *Geophysics*, 16, 192–206.

Faust L Y 1953: *Geophysics*, 18, 271–288.

Gardner G H F, Gardner L W, Gregory A R 1974: *Geophysics*, 39, 770–780.

Geertsma J 1961: *Soc. Petrol. Eng. J.*, 1, 234–248.

Han D-H, Nur A, Morgan D 1986: *Geophysics*, 51, 2093–2107.

Harwood C, Tewfik N 1992: The Miocene Rudeis and Kareem Formations of the Gulf of Suez, aspects of sedimentology and geohistory. Abs. of E.G.P.C. 11-th exploration and production conf., Cairo, Egypt

Klimentos T, McCann C 1990: *Geophysics*, 55, 998–1014.

Koch G S, Link R F 1970: Statistical analysis of geological data. John Wiley and Sons Inc., New York

Marion D, Nur A, Yin H, Han D 1992: *Geophysics*, 57, 554–563.

Marschall R 1984: Some aspects of deconvolution (preprint). CDC's Petroleum Seminar, Geneva

Mavko G, Jizba D 1991: *Geophysics*, 56, 1940–1949.

Schlumberger 1972: Log interpretation principles and charts. Paris

Schlumberger 1974: Applications. Paris

Schlumberger 1984: Well evaluation conference. Egypt

Serra O 1979: Delayed sonic logs. Vol. 1, SNEA (P), Pau

Wyllie M R J, Gregory A R, Gardner L W 1956: *Geophysics*, 21, 41–70.

Wyllie M R J, Gregory A R, Gardner L W 1958: *Geophysics*, 23, 459–493.

Wyrobeck S M 1959: *Geophysical Prospecting*, 7, 218–230.

Xu S, White R 1995: *Geophysical Prospecting*, 43, 91–118.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 181–190 (1996)

REGIONAL TIME AND MAGNITUDE PREDICTABLE SEISMICITY MODEL FOR NORTH-EAST INDIA AND VICINITY

D SHANKER[1] and V P SING[1]

[Manuscript received March 25, 1996]

Investigation of the time dependent seismicity in four seismogenic sources of North-East India and vicinity indicates that strong shallow earthquakes in each source exhibit short as well as intermediate term time clustering. Intervent times of strong shallow mainshocks were determined and used for each source in an attempt at long-term prediction. The following relations have been obtained.

$$\begin{aligned} \log T_t &= -0.24M_{\min} + 0.52M_p - 0.01\log m_0 - 0.38 \\ M_f &= -0.53M_{\min} + 0.78M_p + 0.79\log m_0 - 15.16 \end{aligned}$$

which relate the intervent time, T_t in years and the surface waves magnitude of the smallest mainshock considered ($M_{\min}$), the magnitude of the preceding mainshock (M_p), the magnitude of the following mainshock (M_f) and the moment rate in each source per year (m_0). On the basis of these relations and taking into account the time of occurrence and the magnitude of the last mainshock in each seismogenic source, time dependent conditional probabilities for the occurrence of the next large ($M_s \geq 6.0$) shallow mainshocks during the next 10 years as well as the magnitudes of the expected mainshocks are determined.

Keywords: earthquake model; cumulative magnitude, mainshock; prediction of earthquakes; seismic hazard; seismic moment; time and magnitude predictable model

Introduction

Most of the earthquake generation models currently used for seismic hazard evaluation assume a Poisson or other memoryless distributions (Cornell 1968). However, a temporal dependence between large earthquakes in several regions has been currently observed (Shimazaki and Nakata 1980, Sykes and Quittmeyer 1981, Wesnousky et al. 1984, Papazachos 1989, Sing et al. 1992). Two kind of time-dependent models have been proposed: the slip-predictable model, according to which the size (coseismic slip, mangitude) of a future earthquake in a seismogenic source depends on the time elapsed since the last earthquake, and the time-predictable model, according to which the time of occurrence of a future earthquake depends on the size and on the time of occurrence of the last earthquake in the seismic source. Most of the aforementioned investigations as well as other independent data (Mogi 1981, Karakaisis et al. 1991, Singh et al. 1992, Shanker and Singh 1995) favour the time predictable model.

Research work on the time dependence between strong earthquakes can be of theoretical interest because it can improve our understanding of the earthquake

[1]Department of Geophysics, Banaras Hindu University, Varanasi–221005, India

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

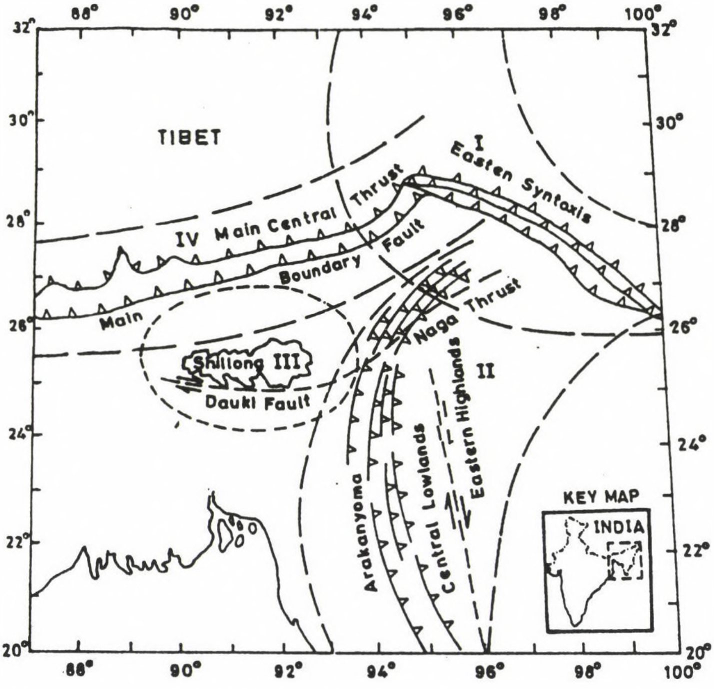


Fig. 1. Tectonic map showing division of the North-East India region and vicinity into four major zones of seismic activity (after Gupta et al. 1986)

generation process, as well as of practical importance, because it can contribute to long-term earthquake prediction and to a more accurate of seismic hazard evaluation (Kiremidjian and Anagnos 1984, Anagnos and Kiremidjian 1984).

An investigation has been carried out for last several years to identify time-dependent relations between strong earthquakes which occurred in seismogenic sources of shallow earthquakes in North-East India and its vicinity. The methodology suggested by Papadimitriou and Papazachos (1994) has been used to include a new term in the relations both for the intervent time and the magnitude of the expected main shock. This term depends on the yearly moment rate $\dot{m}_0$ in each seismogenic source.

The purpose of the present study is to test and apply this time and magnitude predictable model in the North-East India region. An attempt concerning long-term earthquake prediction in this area was made by Goswami and Sarmah (1982) on the basis of the Gumbel's extreme value theory. The most outstanding features of the North-East India region are the intensity of seismicity and the complexity of the presently active tectonics. The area includes the interaction of at least four major tectonic elements: the Eastern Syntaxis (Zone I), the Arakan Yoma and the Naga Thrust Fold Belt (Zone II), the Shillong Plateau (Zone III) and the Main Central and the Main Boundary Fault of the Himalayan Frontal Arc (Zone IV) (Fig. 1).

Method

Papazachos and Papaioannou (1993) on the basis of the intervent times of strong mainshocks in the seismogenic sources of the Aegian area, proposed relations of the following form:

$$\log T_t = bM_{\min} + cM_p + d\log \dot{m}_0 + t \quad (1)$$

$$M_f = BM_{\min} + CM_p + D\log \dot{m}_0 + m \,. \quad (2)$$

The model exposed by these relations has the advantage that all parameters (b, c, d, t, B, C, D and m) are calculated by all available data from all the seismogenic sources.

The moment rate $\dot{m}_0$, that is, the moment released per year in each seismogenic source, is a measure of the seismicity level and varies from source to source, but it can be reliably calculated if enough data are available for the source. These data concern not only the magnitudes of the few mainshocks but also all the data of strong and small shocks available for each source. The values of $\dot{m}_0$ have been determined by Just Summing the moments of the earthquakes (derived by seismic moment Purcaru and Berckhamer (1978) formula:

$$\log M_0 = 1.5M_s + 16.1$$

where M_0 is the seismic moment and M_s represent the surface wave magnitude) and dividing by the corresponding time. The accurate determination of moment rate $\dot{m}_0$ is problematic for smaller regions because in these cases its value depends on a relatively small sample of large earthquakes. On the contrary, calculation of $\dot{m}_0$ is based on large samples of data of a wide range of magnitudes and on the magnitude of the largest earthquake of the region which is usually known or can be reliably estimated. For this reason, $\dot{m}_0$ was calculated for each seismogenic region and is used in the present study.

For the determination of the parameters of Eq. (1) as well as for Eq. (2), one can use a well-known technique (Draper and Smith 1966, Weishergs 1980), which has been widely used. A computer program used on this method and written by Papazachos was used to calculate model parameters for the present study.

Seismogenic sources and data used

The division of the studied area into seismogenic sources, that is into parts of the seismic zone having individual characteristics from the seismotectonic point of view, is one of the basic requirements for the application of the methodology explained above. Following Papazachos (1989), seismotectonic criteria, seismicity level, type of faulting, and geomorphological criteria were used. The spatial distribution of intermediate ($M_s \geq 5.5$) earthquake is also taken into consideration in this separation. Thus four seismogenic sources (Fig. 2) were defined along the major tectonic features of the area. The events included in each seismogenic sources are not necessarily originated by the same fault but belong to the same tectonic

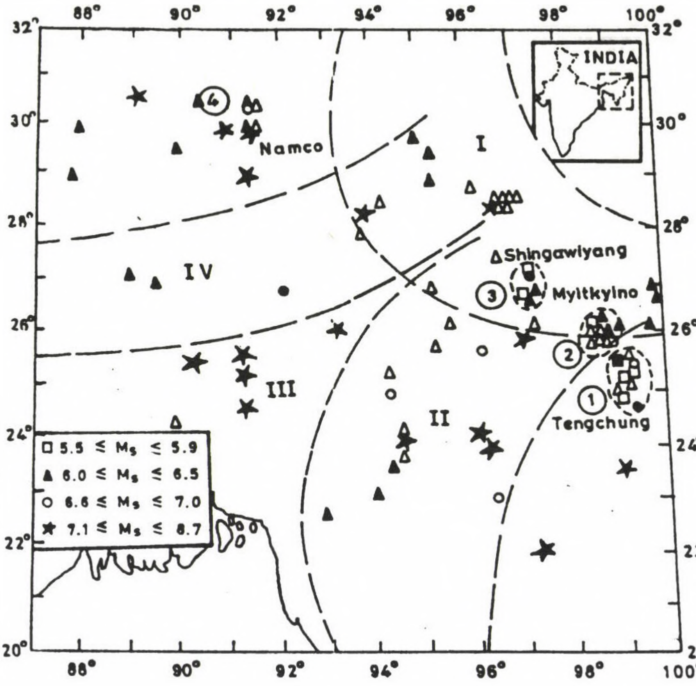


Fig. 2. Earthquake epicentral distribution of the shallow mainshocks with $M_s \geq 5.5$ for the period 1906–1984. The four seismogenic sources are demarcated by elliptical boundaries. Open triangle and circle denote foreshock/aftershocks with the magnitude 5.5–6.4, 6.5–7.4 respectively

environment. A number of faults plane solutions for the sources 1, 2 and 3 show strike slip and thrust faultings wehereas source 4 is situated in Tibetan region of extensional environments (Molnar and Tapponnier 1978).

The earthquake catalogue prepared by the National Geophysical Data Centre, Boulder, Colorado, USA was used for the period 1906 to 1984 with surface wave magnitude $M_s \geq 5.5$. The plot of epicentre is shown in Fig. 2. The clustering pattern of epicentres is seen at four places showing high level seismicity. These places were identified as seismogenic sources and demarcated by elliptical boundaries. For these zones earthquakes of $5.5 \leq M_s \leq 7.5$ were considered for statistical derivation.

Table I gives the name of the sources and the relevant information on the mainshocks in each seismogenic sources as well as moderate ($M_s \geq 5.5$) foreshocks and aftershocks. The cumulative magnitude, M of each sequence, that is the magnitude which corresponds to the total moment released by the major shocks (mainshock and large foreshocks and aftershocks) of each sequence is calculated according to the relation suggested by Hanks and Kanamori (1979). These cumulative magnitudes were used in this study instead of the magnitudes of the mainshocks.

In each seismogenic sources, after considering the first minimum magnitude $M_{\min}$, of all the mainshocks, the year of occurrence of preceding mainshock, t_p, the year of occurrence of the following mainshock, t_f, and the intervent times T, between successive mainshocks with magnitudes equal to or larger than $M_{\min}$ are

Table I. Information on the basic data used for each seismogenic sources

Source	Date D	M	Y	Epicentre $\phi°$	$\lambda°$	M_s	M
Teng Chung	28	03	1914	25.0	99.0	6.9	6.9
	25	09	1930	25.3	98.9	6.0	6.0
	16	05	1941	24.0	99.0	6.9	6.9
	11	06	1961	25.2	98.6	5.6	5.6
	29	05	1976	24.6	98.9	6.9	7.2
	29	05	1976	24.5	98.7	7.0	–
	31	05	1976	24.3	98.6	6.2	–
	02	06	1976	24.9	98.8	5.9	–
	21	07	1976	24.7	98.7	6.3	–
Myityina	16	10	1929	25.8	98.7	6.5	6.9
	28	04	1930	25.8	98.7	6.3	–
	21	09	1930	25.8	98.4	6.5	–
	04	11	1930	25.6	98.3	6.0	–
	02	12	1930	25.8	98.3	6.0	–
	18	10	1931	26.0	98.0	5.6	–
	03	01	1932	25.5	98.5	5.6	–
	11	08	1933	25.5	98.5	6.5	–
	19	01	1934	25.5	98.2	6.0	–
	31	10	1941	25.4	98.4	6.3	6.3
Shingowiyang	31	08	1906	27.0	97.0	7.0	7.0
	04	05	1955	27.0	97.0	5.8	5.8
	22	09	1962	26.5	97.0	6.0	6.0
	30	05	1976	26.6	97.0	5.6	6.2
	12	08	1976	27.7	27.1	6.1	–
Namco	14	10	1921	30.5	91.0	6.3	6.3
	03	09	1940	31.0	91.5	6.3	6.4
	04	10	1940	30.5	91.5	6.0	–
	17	11	1951	31.0	91.6	6.3	7.9
	18	11	1951	30.9	91.6	6.8	–
	18	11	1951	30.5	91.0	7.9	–
	17	08	1952	30.5	91.5	7.5	–

calculated. Then, the second minimum magnitude, $M_{\min}$ is considered, and new intervent times between successive mainshocks with magnitudes equal to or larger than the second $M_{\min}$ are calculated. This procedure is continued until the last $M_{\min}$ is considered. The results of these calculations are given in Table II.

Results

By the use of the data set listed in Table II the following relation was determined:

$$\log T_t = -0.24 M_{\min} + 0.52 M_p - 0.01 \log \dot{m}_0 - 0.38 \qquad (3)$$

with a correlation coefficient equal to 0.78 and a standard deviation equal to 0.13. Figure 3 displays a plot of

$$\log T^* = \log T + 0.24 M_{\min} - 0.52 M_p + 0.01 \log \dot{m}_0 + 0.38$$

Table II. Seismogenic Source Data used to determine the parameter of the empirical relations

$M_{\min}$	M_p	M_f	T	t_p	t_f	$\log \dot{m}_0$
			Source–1			
5.6	6.9	6.0	16.42	1914	1930	25.32
	6.0	6.9	10.64	1930	1941	
	6.9	3.6	20.06	1941	1961	
	5.6	7.2	14.97	1961	1976	
6.0	6.9	6.0	16.42	1914	1930	
	6.0	6.9	10.64	1930	1941	
	6.9	7.2	35.04	1941	1976	
6.9	6.9	6.9	27.13	1914	1941	
	6.9	7.2	35.04	1941	1976	
			Source–2			
6.3	6.9	6.3	12.04	1929	1941	25.17
	6.3	6.5	13.39	1941	1955	
6.5	6.9	6.5	25.43	1929	1955	
			Source–3			
5.8	7.0	5.8	48.70	1906	1955	24.79
	5.8	6.0	7.38	1955	1962	
	6.0	6.2	13.69	1962	1976	
6.0	7.0	6.0	56.06	1906	1962	
	6.0	6.2	13.69	1962	1976	
6.2	7.0	6.2	69.70	1906	1976	
			Source–4			
6.3	6.3	6.4	18.83	1921	1940	26.19
	6.4	7.9	11.21	1940	1951	
6.4	6.4	7.9	11.21	1940	1951	

as a function of M_p, where T, $M_{\min}$, $\log \dot{m}_0$ and M_p are the observed values. The line drawn through the data is a least squares fit. The positive correlation between the repeat time and the magnitude of the preceding mainshock suggest that the time-predictable model holds in the area under study. Figure 4 shows the frequency distribution of $\log(T/T_t)$, which is fitted by normal distribution with $\mu = 0$ and $\sigma = 0.13$.

The values of the parameters of Eq. (2) were determined by the use of all available information from Table II, and the following empirical formula was found:

$$M_f = -0.53M_{\min} + 0.78M_p + 79\log \dot{m}_0 - 15.16 \tag{4}$$

with a correlation coefficient equal to 0.53 and a standard deviation equal to 0.50.

Prediction of strong shallow mainshocks

Prediction of strong earthquakes in a long-term sense is a very common target of earthquake studies. The validity of time-predictable model expressed by the relation (3) provides a tool for such predictions, since the time of the next strong mainshock in each seismogenic sources can be estimated by this relation. It is better, however,

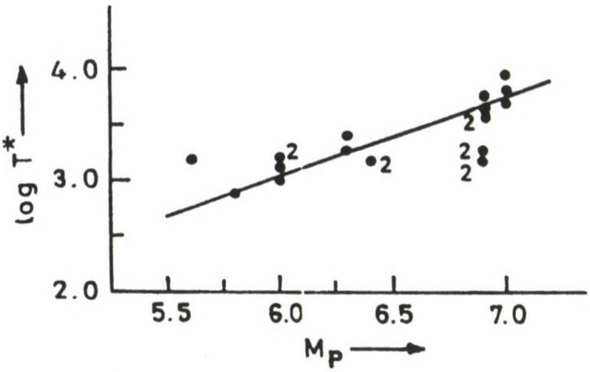


Fig. 3. Dependence of the repeat time on the magnitude of the preceding main-shock, M_p. The numeral denotes the number of coinciding points

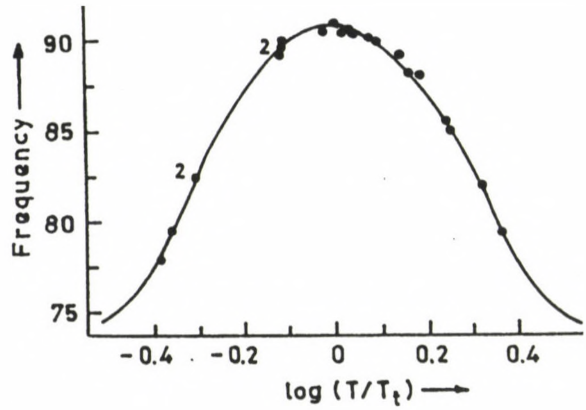


Fig. 4. The frequency distribution of the observed repeat times, T, compared to the theoretical ones, T_t

Table III. Expected magnitude M_f and the corresponding probabilities; P_{10} for the occurrence of large ($M_{\min} \geq 6.0$). Shallow main-shocks during the period 1996–2005 in North-East India

Seismogenic source	Source name	M_f	P_{10}
1	TengChung	6.7	0.02
2	Myitkyina	6.6	0.92
3	Shingowiyang	6.0	0.74
4	NamCo	8.2	0.21

to estimate the probability of occurrence of the next mainshock larger than a certain magnitude in a certain time interval, because there is a considerable uncertainly in the calculation of repeat time.

Among the probabilistic methods widely used for such purposes, the ones which take into account the time elapsed since the occurrence of the last strong shock look more appropriate. Following Papazachos and Papaioannou (1993) it was found that the ratio T/T_t follows a lognormal distribution. One can calculate the conditional probability that a mainshock with $M \geq M_{\min}$ will occur in a seismogenic source during the next Δt years (from now) if the last earthquake occurred in the source, t years ago (from now), by the relationship:

$$P(\Delta t) = \frac{F\left(\frac{L_2}{\sigma}\right) - F\left(\frac{L_1}{\sigma}\right)}{1 - F\left(\frac{L_1}{\sigma}\right)} \tag{5}$$

where

$$L_2 = \log\left(\frac{t+\Delta t}{T_t}\right)$$

$$L_1 = \log\left(\frac{t}{T_t}\right)$$

and F denotes the complementary cumulative value of the normal distribution with mean equal to zero and standard deviation σ (0.13).

Given the date and the magnitude of the last event in a seismogenic source as well as the uncertainties of the model expressed by its standard deviation ($\sigma = 0.13$), the probabilities of occurrence of the next shallow mainshocks with $M_s \geq 6.0$ during the next 10 years were computed. Table III gives information on the magnitudes, M_f, of the expected mainshocks and the corresponding probabilities, P_{10}, based on the model expressed by the relation (3) and (4). The estimated probabilities are high ($P_{10} \geq 0.50$) for sources Myitkyina and Shingowiyang. We must note that the absolute values of the probabilities are of relative importance, that is their changes from source to source are of significance. This is so because these values may change if a larger sample of data is used.

Discussion

This study attempts to quantitatively assess the probabilities of occurrence and the magnitudes of future large ($M_s \geq 6.0$) shallow mainshocks in the four seismogenic sources of North-East India. The estimation of conditional probabilities and the magnitudes of the expected events was based on the time and magnitude predictable model expressed by the relations (3) and (4).

There are several cases where this model was applied, on the basis of the methodology followed in the present study; i.e. for the Aegean area (Papazachos 1988a, 1988b, 1989, 1991, 1992, 1993, Papazachos and Papaioannou 1993), for the seismic zone of South and Central America (Papadimitriou 1993), for Japan area (Papazachos et al. 1994) and for Iran (Karakaisis 1994). The "Time and magnitude predictable model" tested in the present paper differs from the time and slip-predictable models of Shimazaki and Nakata (1980). In those original models the intervent time is proportional to the coseismic slip of the preceding or of the following mainshock, but for the model expressed by Eqs (1) and (2) these relations are different. On the other hand, those original models are applied in a single fault or simple plate boundary, while the model tested in the present paper is applied in seismogenic sources which may include several faults.

The model presented by relation (3) and (4) containing different terms have some physical meaning. The term $bM_{\min}$ and $BM_{\min}$ ($b = -0.24$, $B = -0.53$) expresses the observational fact that the larger the magnitude of an earthquake in a seismic region the larger its repeat time. The term cM_p ($c = 0.52$) indicates that the larger the magnitude of a mainshock in a seismogenic region the longer the time till the next mainshock. This indicates that the generation of an earthquake in a fault of a seismogenic region affects the stress level in other faults of the same region. The meaning of the term CM_p ($C = 0.78$) in relation (4) is that a large mainshock in a seismogenic region is followed by a small one. This can be explained by assuming that large earthquakes reduce much the stress level in all faults of the region and, since the tectonic loading is constant, the stress level in a small fault reaches soonner the maximum stress. For this reason, small mainshocks are followed by larger ones as it indicated by this same term.

The term $d \log \dot{m}_0$ and $D \log \dot{m}_0$ ($d = -0.01$ and $D = 0.79$) express the tectonic loading which is excerted on the seismogenic region and is assumed to be constant. It is expected that, for certain $M_{\min}$ and M_p, the repeat time, T to decrease (d negative) and the magnitude of the following mainshock to increase (D positive) with the increase of the tectonic loading, that is with the seismicity level of the corresponding seismogenic region. Here the value of the parameter 'd' is very small that means the moment rate will play an insignificant role in the estimation of the theoretical repeat time. In the relation (4) the parameter 'C' is positive that contradicts the principle of the model. This is due to the limited available data set used. The estimates of the seismic potential presented in this study are meant to be interpreted as longterm forecasts only.

Acknowledgements

The authors would like to express their sincere thank to University Grant Commission for their financial support. Gratitude is also extended to Prof. E Papadimitriou for her many helpful suggestions and some necessary calculations.

References

Anagnos T, Kiremidjian A S 1984: *Bull. Seismol. Soc. Am.*, 74, 2593–2611.

Draper N R, Smith H 1966: Applied regression analysis. John Wiley, New York

Goswami H C, Sarmah S K 1982: *Bull. Seismol. Soc. Am.*, 72, 999–1009.

Gupta H K, Rajendran K, Singh H N 1986: *J. Geol. Soc. of India*, 28, 345–365.

Hanks T C, Kanamori H 1979: *J. Geoph. Res.*, 84, 2348–2350.

Karakaisis G F 1994: Long term earthquake prediction in Iran based on the time and magnitude predictable model. *Phys. Earth Planet. Int.* (in press)

Karakaisis G F, Kourouzidis E M, Papazachos B C 1991: Behaviour of seismic activity during a single seismic cycle. Paper presented at International Conferences on Earthquake Prediction: State-of-the-Art, Council of Europe, Strasbourg, France

Kiremidjian A S, Anagnos T 1984: *Bull. Seismol. Soc. Am.*, 74, 739–755.

Mogi K 1981: In: Earthquake Prediction, D W Simpson, P G Richards eds, An International Review, Maurice Ewing Ser., AGU, Washington, Vol. 4, 43–52.

Molnar P, Taponnier P 1978: *J. Geoph. Res.*, 5361–5375.

National Geophysical Data Centre, Boulder, Colorado, USA, Seismicity data of North-East India, 1906–1984

Papadimitriou E E, Papazachos B C 1994: *J. Geoph. Res.*, 99, 15389–15398.

Papazachos B C 1988a: Seismic hazard and long term earthquake prediction in Greece. European School of Earthquake Sciences, Course on Earthquake Hazard Assessment, Athens, Greece, 1–10.

Papazachos B C 1988b: Long-term prediction of earthquakes in seismogenic sources of Greece. UN Sem. on Prediction of Earthquakes, Lisbon, Portugal, 1–10.

Papazachos B C 1989: *Bull. Seismol. Soc. Am.*, 79, 77–84.

Papazachos B C 1991: Long-term earthquake prediction in South Balkan region based on a time dependent seismicity model. Paper presented at 1st General Conference of the Balkan Physicists' Union, Thessaloniki

Papazachos B C 1992: *Pure Appl. Geophys.*, 138, 287–308.

Papazachos B C 1993: *Nat. Hazards*, 7, 211–218.

Papazachos B C, Papaioannou Ch A 1993: *Pure Appl. Geophys.*, 140, 593–612.

Papazachos B C, Papadimitriou E E, Karakaisis G F, Tsapanos T M 1994: *Bull. Seismol. Soc. Am.*, 84, 426–437.

Purcaru G, Berckhamer H 1978: *Tectonophysics*, 49, 189–198.

Shanker D, Singh V P 1995: *Acta Geod. Geoph. Hung.*, 30, 207–216.

Shimazaki K, Nakata T 1980: *Geophys. Res. Letter*, 7, 279–282.

Singh V P, Shanker D, Singh J 1992: *Proc. Ind. Acad. Sci., Earth Planet. Sci.*, 101, 361–368.

Sykes L R, Quittmeyer R C 1981: Repeat times of great earthquakes along simple plate boundaries in earthquake prediction. D W Simpson, P G Richards eds, An International Review, Maurice Ewing Ser., AGU, Washington, Vol. 4, 217–247.

Wesnousky S G, Scholz C H, Shimazaki K, Matsuda T 1984: *Bull. Seismol. Soc. Am.*, 74, 687–708.

Acta Geod. Geoph. Hung., Vol. 31(1-2), pp. 191-216 (1996)

REGIONAL MAGNETOTELLURIC (MT) ANISOTROPY IN THE PANNONIAN BASIN (HUNGARY)

A ÁDÁM[1]

[Manuscript received February 7, 1996]

The deep MT sounding curves measured in the Pannonian Basin inside the Carpathian arcs show anisotropy characterized by the following features:

- **The extreme MT sounding curves (Rhomax and Rhomin) strongly differ from each other from the S-interval on, i.e. in the increasing branch of the MT curves characterizing the resistive crystalline basement of the sedimentary cover. Rhomin curves give for the resistivity of the basement extremely low values.**
- **Directions of Rhomax values are near to N-S or NNW-SSE.**
- **The apparent depths to the conductive asthenosphere derived from Rhomax and Rhomin curves are very different. In the case of Rhomax curves — except a few very isotropic areas — it is $\geq$ 100 km and in case of Rhomin, as shallow as $\leq$ 50 km. Their geometric mean better approximates the depths measured by seismological method and calculated from geothermal data in the hot Pannonian Basin with a heat flow of about 100 mWm^{-2}.**
- **Extreme local basement structures like deep narrow extensional 3D basins (e.g. Békés graben) do not change this character of MT curves at long periods.**
- **A numerical model with anisotropic basement resistivity does not completely explain these features. The main source can probably be found in the E(ENE)-W(WSW) directed deep crustal fractures cutting through the whole crust and forming basin and range structure.**

Keywords: anisotropy; asthenosphere; conductivity anomaly; deep sounding; magnetotellurics; numerical modeling; tectonics

1. The main features of the deep MT sounding curves measured in the Pannonian Basin inside the Carpathian arcs

- The extreme MT sounding curves (Rhomax and Rhomin) strongly differ from each other from the S-interval on, i.e. in the increasing branch of the MT curves characterizing the resistive crystalline basement of the sedimentary cover. Rhomin curves give for the resistivity of the basement extremely low values.

- Directions of Rhomax values are near to N-S or NNW-SSE.

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

— The apparent depths to the conductive asthenosphere derived from Rhomax and Rhomin curves are very different. In the case of Rhomax curves — except a few very isotropic areas — it is ≥ 100 km and in case of Rhomin, as shallow as ≤ 50 km. The Rhomax indication is much weaker, i.e. the decreasing branch of the MTS curves is far from the asymptotes. The Rhomin curves better approximate the depths measured by seismological method and calculated from geothermal data in the hot Pannonian Basin with a heat flow of about 100 $\mathrm{mWm^{-2}}$.

Typical MT curves are shown in Fig. 1 from SE-Hungary (*HÓD*MEZŐVÁSÁRHELY) with the results of 1D inversions. (See MT sites along PGT 1 and 4 Pannon Geotraverses and Rhomax directions in Fig. 2.)

— Similar difference appears in the apparent depth of the crustal conductors calculated by Rhomax and Rhomin curves (Fig. 3) measured in the area of the Transdanubian CA. (In Fig. 4 the rose diagram of the Rhomax direction is also shown.) This statement refers here also to the apparent depth of the asthenosphere derived from these curves. In this case *only* the Rhomax curves are distorted by static shift, therefore these are of B polarization.

— Extreme local basement structures like deep narrow extensional 3D basins (e.g. Békés graben) do not change this character of MT curves at long periods.

2. Numerical models for explanation of the regional anisotropy

2.1 Effect of the anisotropic basement rocks

Following the numerical technique of O'Brien and Morrison (1967) model-computations have been carried out both for anisotropic sediments and anisotropic basement rocks (Ádám et al. 1972a). As stated, the angle difference of the increasing branches of the extreme sounding curves (S-interval) at measuring sites in the Great Hungarian Plain may be generated by the anisotropy of the metamorphic basement rocks. It was also concluded that an anisotropic medium with thickness of some tens of km-s is needed for the interpretation. It should be remarked that 1D inversion of the measured curves does not give so high anisotropy coefficients ($\lambda = \sqrt{\varrho_t/\varrho_\ell}$) as those used by O'Brien and Morrison in their calculation ($\varrho_1 = 1000\ \Omega$m, $\varrho_2 = 50\ \Omega$m). In the case of the measuring site PGT4–2 (Hód) the resistivity of the basement was inverted to 270 Ωm and 38 Ωm, respectively.

Concerning the depth to the asthenosphere essential difference appears between the calculated anisotropic and the experimental models.

In case of σ_1 and σ_2 basement conductivity the decreasing branches of the calculated anisotropic model curves, indicating the conductive asthenosphere meet each other and so give the same asthenospheric depth (Fig. 5a). Having an intermediate conductive layer in the layer sequence — this can be the asthenosphere if the ultimate conductive layer corresponds to the phase-transition of peridotite at a depth of 400 km — there is no difference between depths of this intermediate conductor derived by inversion from the two extreme model curves.

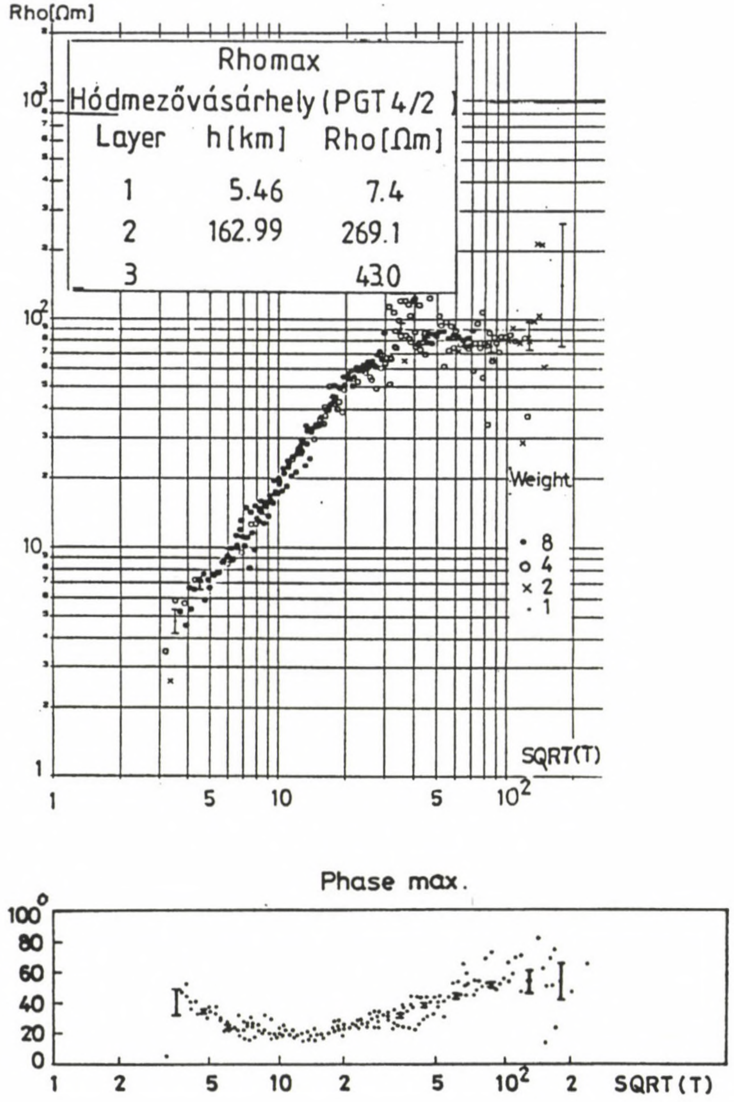


Fig. 1a. Characteristic MT sounding curves (Rhomax) from the measuring site PGT-4/2 (near *Hód*mezővásárhely) (shown in Fig. 2)

The cause of this great difference can only be found in the structural inhomogeneities as concluded by Ádám et al. (1972a) summarizing the results of the anisotropic model calculations.

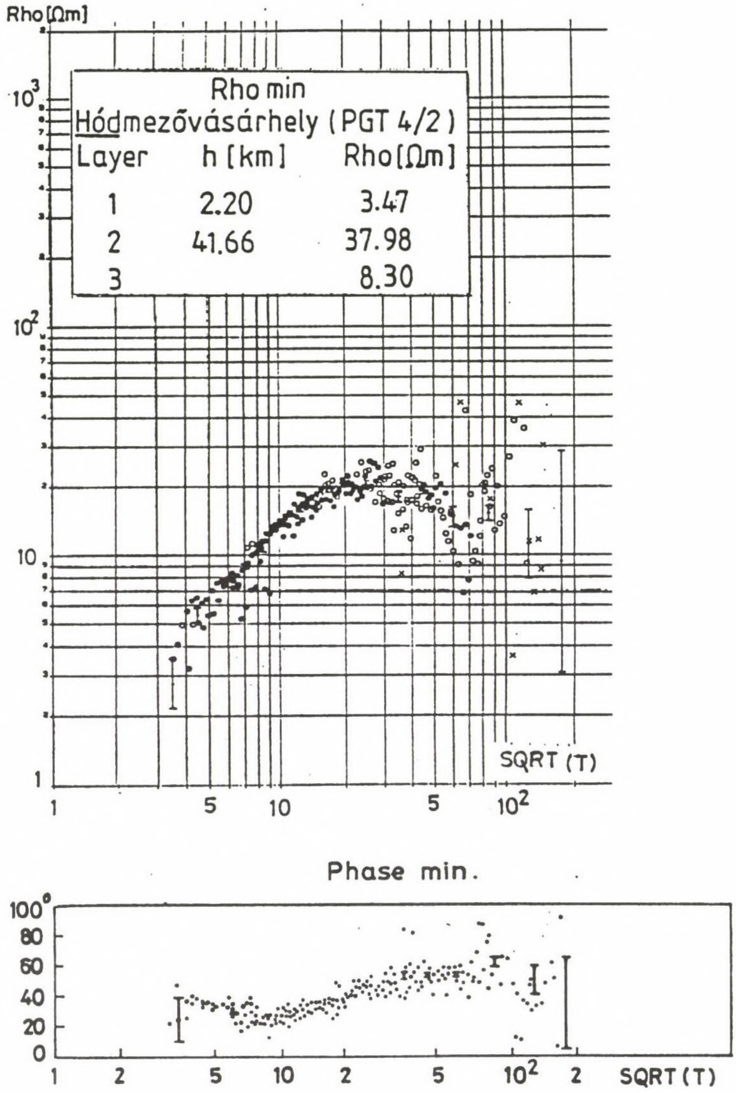


Fig. 1b. Characteristic MT sounding curves (Rhomin) from the measuring site PGT-4/2 (near *Hód*mezővásárhely) (shown in Fig. 2)

2.2 Numerical models for crustal inhomogeneities

In the following inhomogeneity models will be studied to find explanation for the difference in the apparent asthenospheric depth given by Rhomax and Rhomin curves in the Pannonian Basin. First 2D models and then their 3D modulations will be examplified.

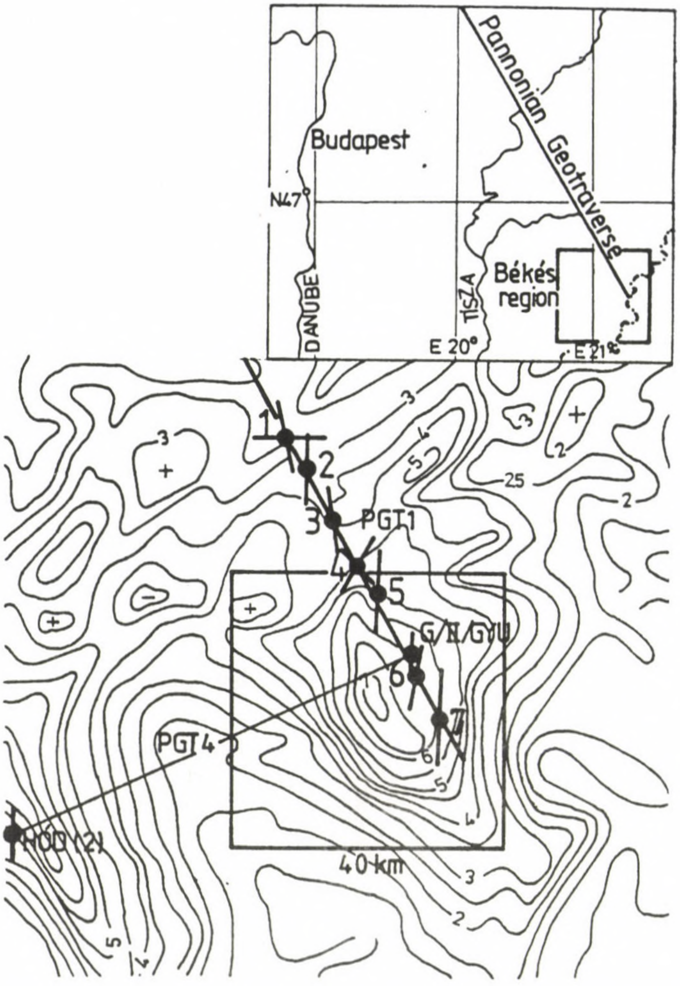


Fig. 2. Isopach map on the subbasins Békés and Makó with Rhomax directions

2.21 2D models and their effect on the layer parameters in case of so-called formal interpretation (FI)

2.211 Horst and graben models

From 2D models, the EM distortions of the basement "graben" and "horst" will be studied at first after Berdichevsky and Dmitriev (1976). The distortion of the horst model results in resistivity increase in the B-polarization. Its consequence is that the Rhomax curves supply greater depth values to the conductive half space ($\sigma_3 = \infty$) than the real one which is well determined by the E polarized Rhomin curves (Fig. 6). According to the isopach map of the Pannonian Basin the resistive basement elevations (horsts) are very frequent, nevertheless, the probability of occurrence of the opposite structures, i.e. of the grabens could be similar.

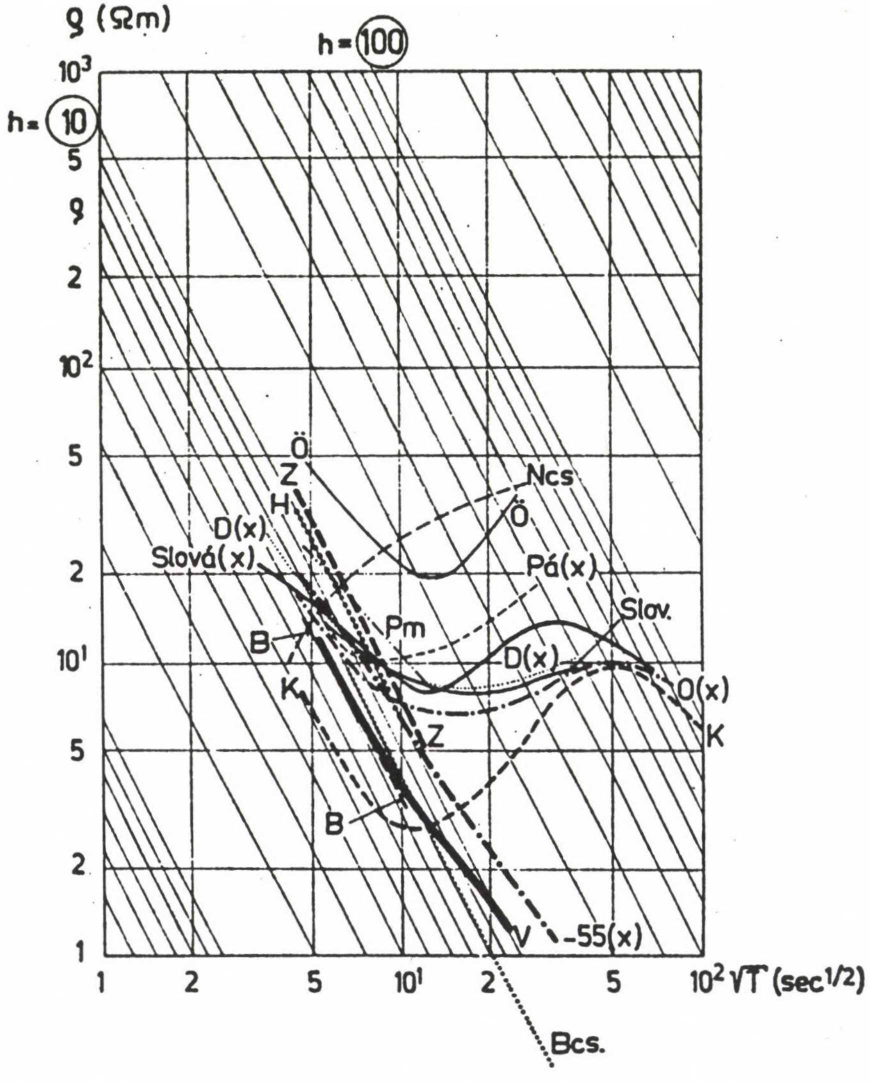


Fig. 3a. MT sounding curves (Rhomax) measured in the area of Transdanubian CA (Ádám 1981)

In case of the grabens the character of the polarization of the Rhomax and Rhomin curves turns over. The Rhomin curves of B polarization do give in this case shallower depth value for the conductive half space (:asthenosphere) than the real one expressed here by the Rhomax curves of E polarization. The apparent depth indicated by Rhomin curves depends on the relative position of the measuring sites in the graben similarly to the horst model.

If the graben has very steep walls, the distortion of the so-called "edge effect" (wall effect) appears in the B polarized Rhomin curves and they indicate false conductors at

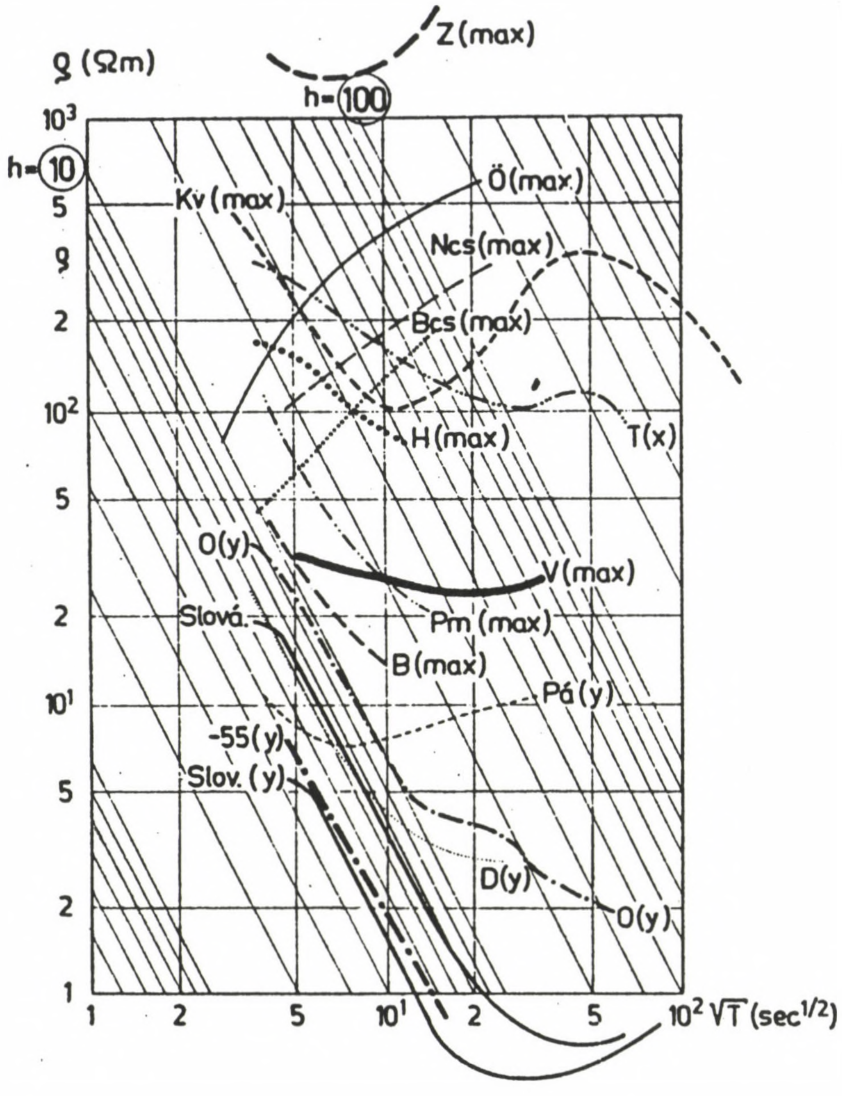


Fig. 3b. MT sounding curves (Rhomin) measured in the area of Transdanubian CA (Ádám 1981)

shallow depth. A classical example of the edge effect is that caused by Pre-Kopet-Dag downwarp (Berdichevsky and Dmitriev 1976). Although similar effects can also occur in the very deep graben of the Pannonian Basin, this type of anomaly cannot explain the general — regional — character of MT sounding curves measured here.

The effect of a sinusoidal relief in the third layer (asthenosphere) is shown after Berdichevsky and Dmitriev (1976) with E ($\varrho_{\|}$) and B ($\varrho_{\perp}$) polarized MT curves and apparent asthenospheric depths derived from them (Fig. 7). The distortion of the B polarized curves similar to that obtained in case of the horst and graben of the resistive

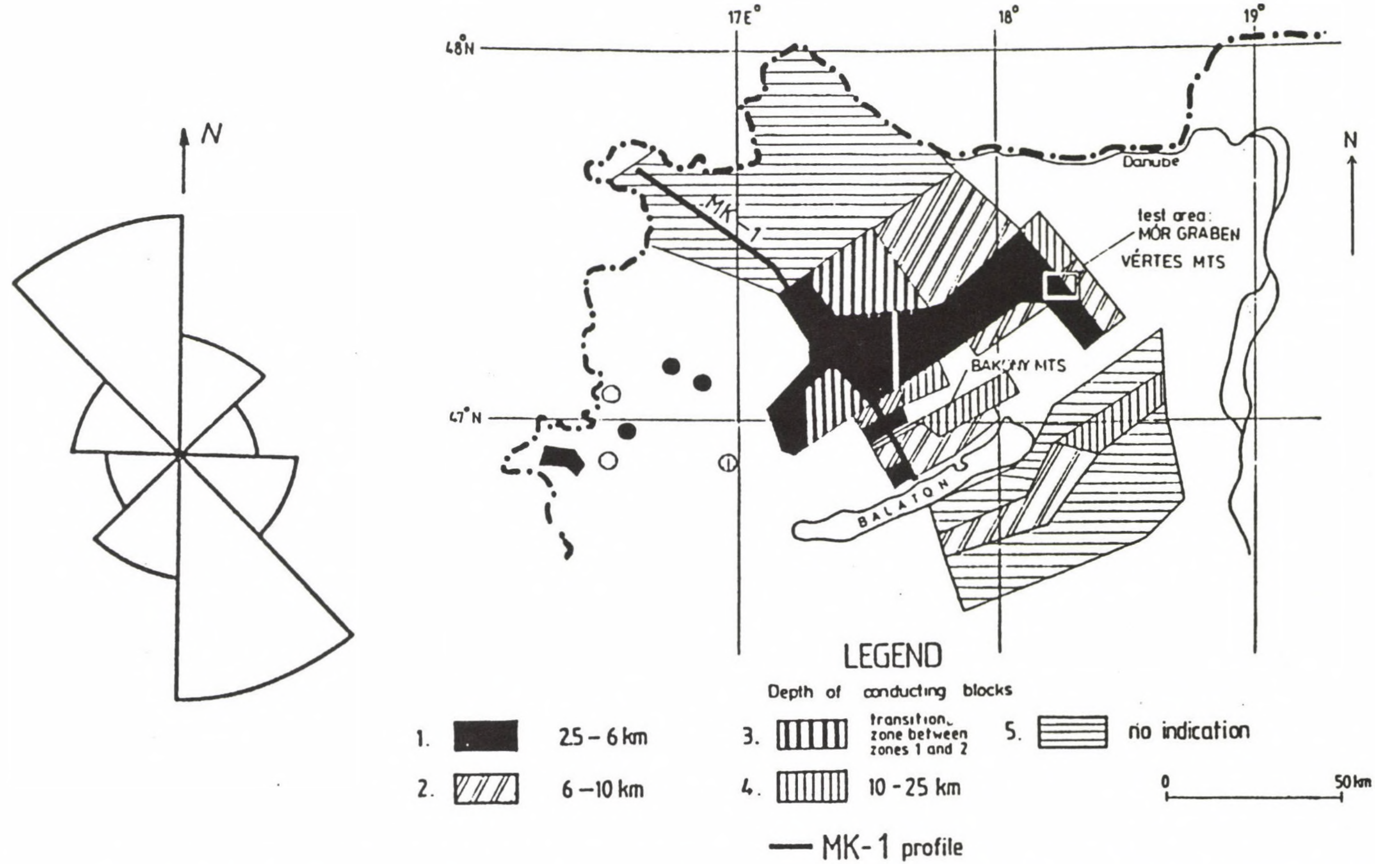


Fig. 4. Depth distribution of the conductors in the area of the Transdanubian conductivity anomaly (CA) with rose diagram of Rhomax directions

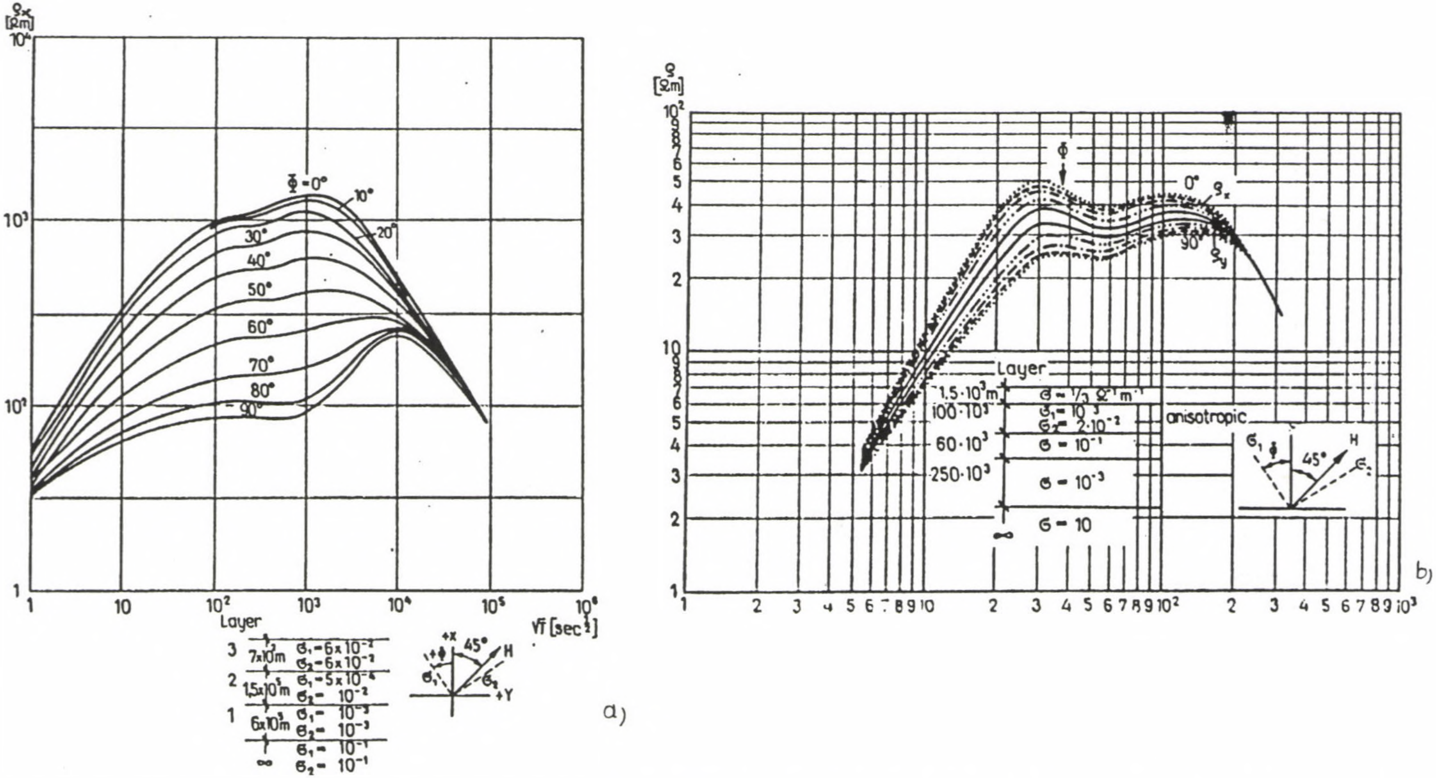


Fig. 5. Theoretical MT sounding curves of layered models with anisotropic basement rocks in a sedimentary basin without (a) and with (b) an intermediate conductor (Ádám et al. 1972a)

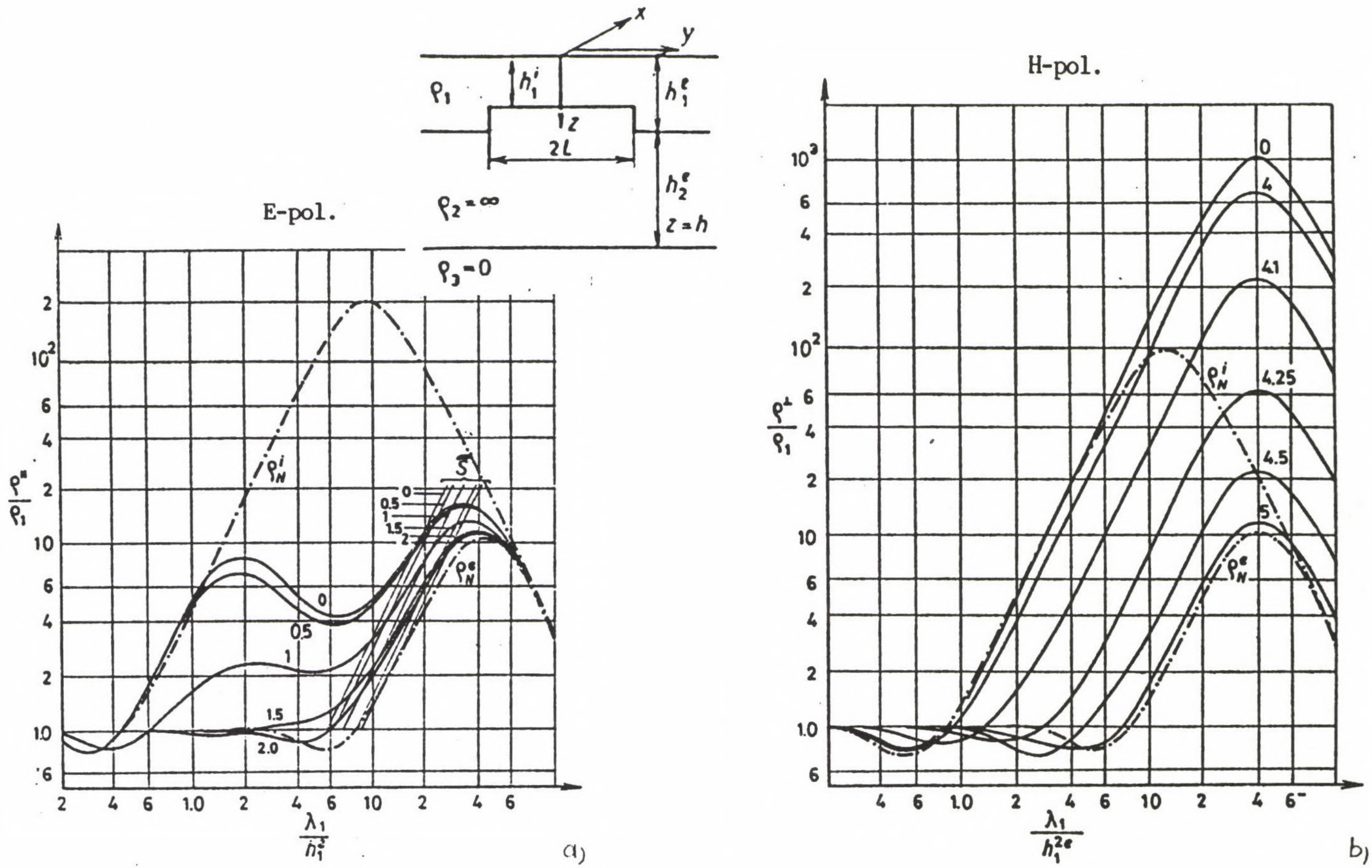


Fig. 6. E and B polarized sounding curves above horst model given by Berdichevsky and Dmitriev (1976). The parameters are according to the inlet: a) E-pol. $\varrho_2/\varrho_1 = \infty$, $\varrho_3/\varrho_1 = 0$, $h_1^i/h_1^e = 0.05$, $h_2^e/h_1^e = 20$, $L/h_1^e = 1$, curves $\varrho^{\parallel}$ and $\tilde{S}$-lines are digitized by $|y|/h_1^e$, b) B-pol. $\varrho_2/\varrho_1 = \infty$, $\varrho_3/\varrho_1 = 0$, $h_1^i/h_1^e = 0.1$, $h_2^e/h_1^e = 20$, $L/h_1^e = 4$, curves $\varrho^{\perp}$ are digitized by $|y|/h_1^e$

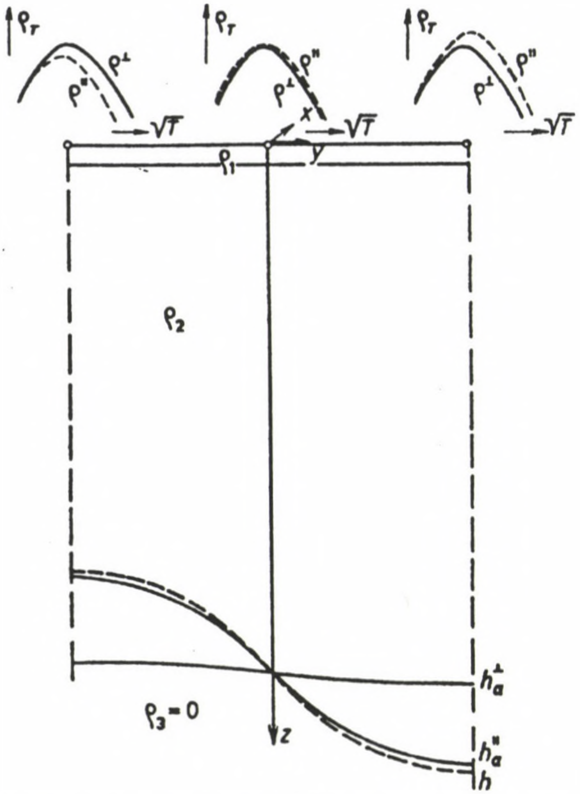


Fig. 7. Depth profiles in a model of sinusoidal relief. $h_1/h_2 = 25$; $\alpha = 0.2$; $S_1/S_2^{\min} = 100$; $L/h_1 = 400$; $L_\tau\min = 2$; h is the true depth to the conducting formation, $h_a^{\perp}$ and $h_a^{\parallel}$ is apparent depth obtained from curves $\varrho^{\perp}$ (B-pol.) and $\varrho^{\parallel}$ (E-pol.) (Berdichevsky and Dmitriev 1976)

basement! Therefore beneath a sedimentary basin with upwelling asthenosphere (see later) the two effects on B polarized curves can decrease each other.

2.212 Conducting dikes

As a next step, the distortion effect of the longitudinal and transversal fracture zones containing loose and broken rocks with electrolytes of low resistivity should be studied in the Pannonian Basin as their directions and those of the extreme MT resistivity values are strongly related to each other.

Single conductive dike

In the centre of a conductive dike the E polarized sounding curve gives the real depth to the top of the dike (Tátrallyay 1977). The apparent depth increases with distance from the dike centre (Ádám 1987). The indication of B polarized curves for the conductive dike is much weaker and therefore the depth determination is also

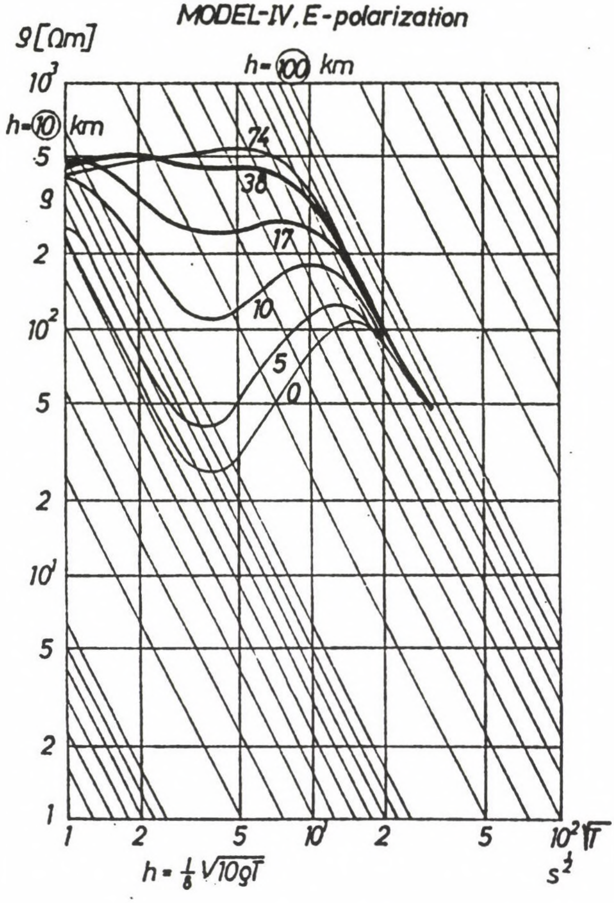


Fig. 8a. E polarized sounding curves calculated above a layer model having a conductive dike and conductive basement (asthenosphere) (Ádám 1987)

more unreliable (Fig. 8). In B polarization the conducting dike, similarly to the sedimentary basin (graben model) decreases the apparent depth to the conducting half space (4th layer in our layer model) giving the lowest apparent depth in the dike centre.

Parallel conducting dikes

Along a profile above a dike system the difference between the E and B polarized Rho and phase (φ) values do not change at a given period. The indication of the dikes by B polarized sounding curves is here also weaker and therefore the apparent depth values are more unreliable than in case of the E polarized curves (Fig. 9a). Moving away from the centre of the dike system an apparent deepening of the top of the dikes appears in case of E polarization (Fig. 9b). The B polarized curves give shallower depth to the asthenosphere than the real one (Fig. 9c).

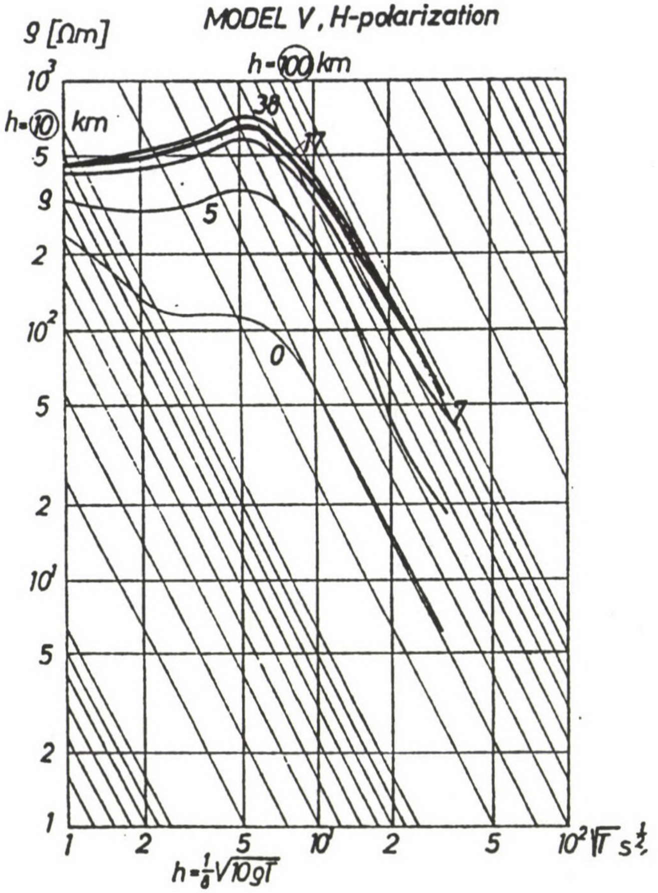


Fig. 8b. H polarized sounding curves calculated above a layer model having a conductive dike and conductive basement (asthenosphere) (Ádám 1987)

This conducting dike system in the upper mantle (in the asthenosphere) may cause similar mantle anisotropy as experienced by Mareschal et al. (1993) in the Canadian shield and can explain the differences between the apparent asthenospheric depths derived from Rhomax and Rhomin curves in particular if the B polarized curves are distorted by static shift. Mareschal et al. (1993) explained the dike system partly by graphite partly by fluid in deep fracture zones or by "fossilized stress fields" in the deep rocks.

The described dyke system successfully explains the EM distortion observed in the area of the Transdanubian *crustal* anomaly (Ádám 1992). As was shown by Ádám (1981), strong static shift due to different near-surface formations was superposed to the weak indication of the B polarized sounding curves. Figure 10 gives relation between the near-surface formations of different resistivity and the apparent depth of the conductive blocks derived from Rhomax and Rhomin curves

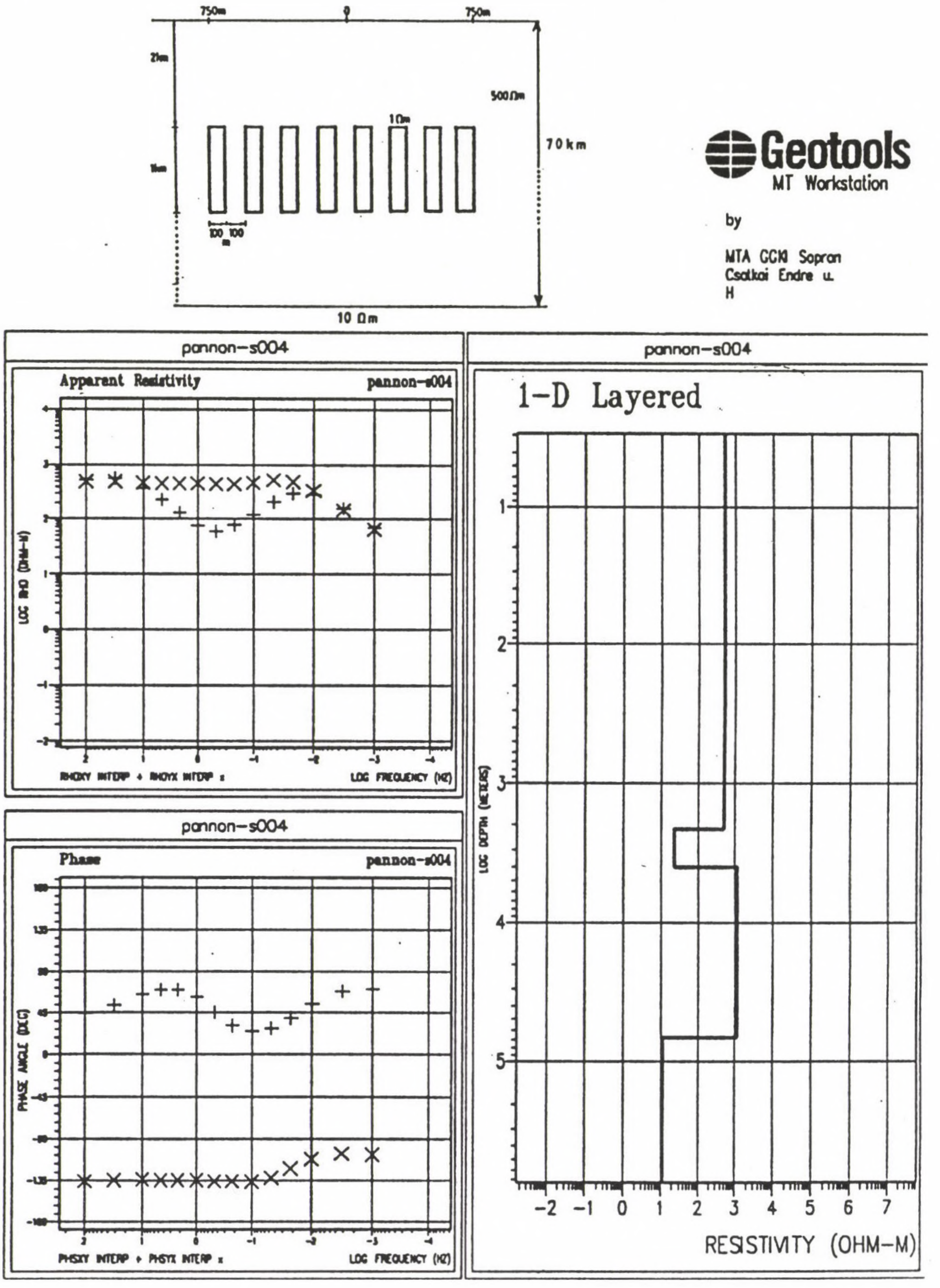


Fig. 9a. Conducting dyke system embedded in resistive matrix with conductive asthenosphere. E and B polarized sounding curves with 1D inversion of E-pol. curves

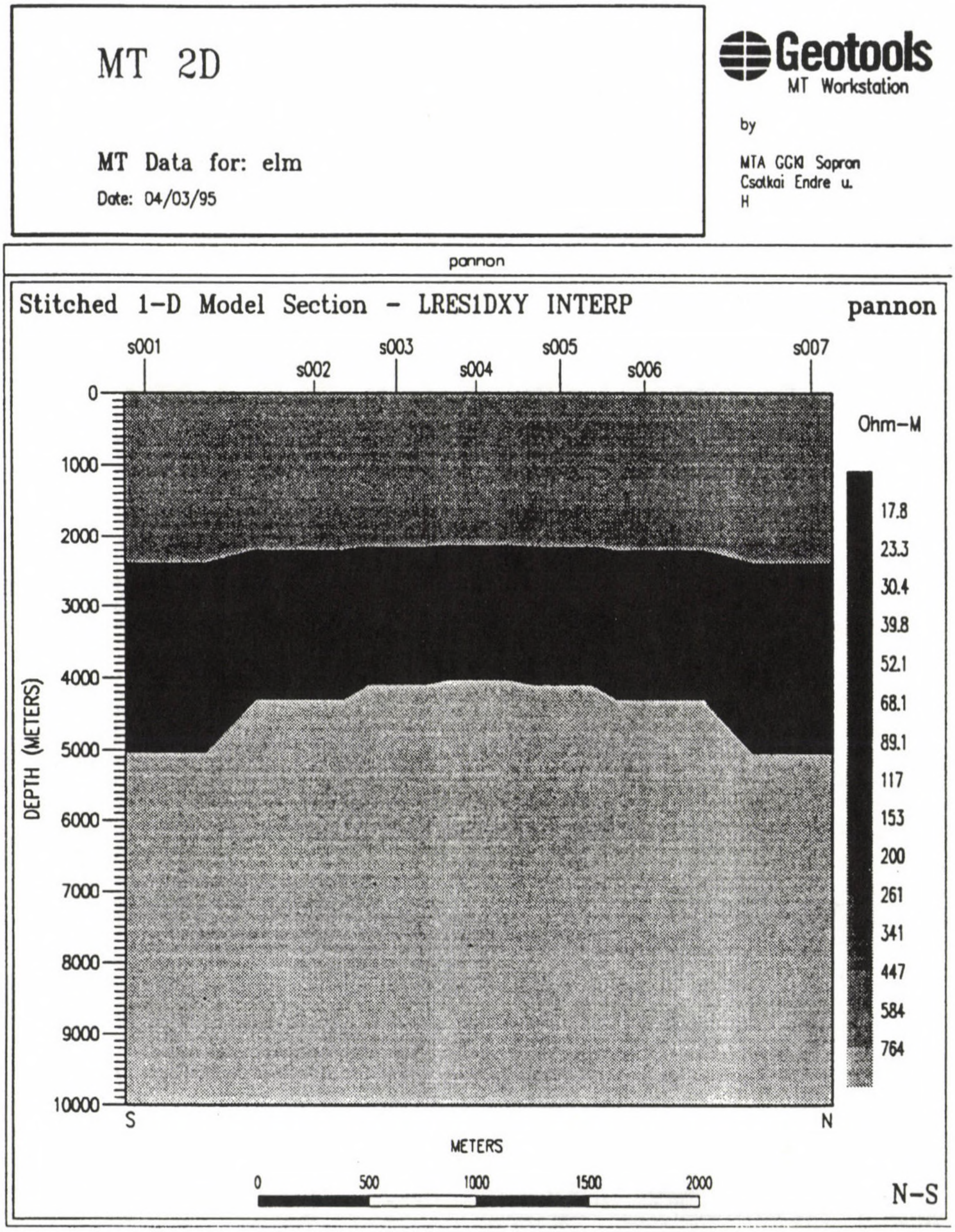


Fig. 9b. Conducting dyke system embedded in resistive matrix with conductive asthenosphere. Apparent "E-pol" structure of the dyke system

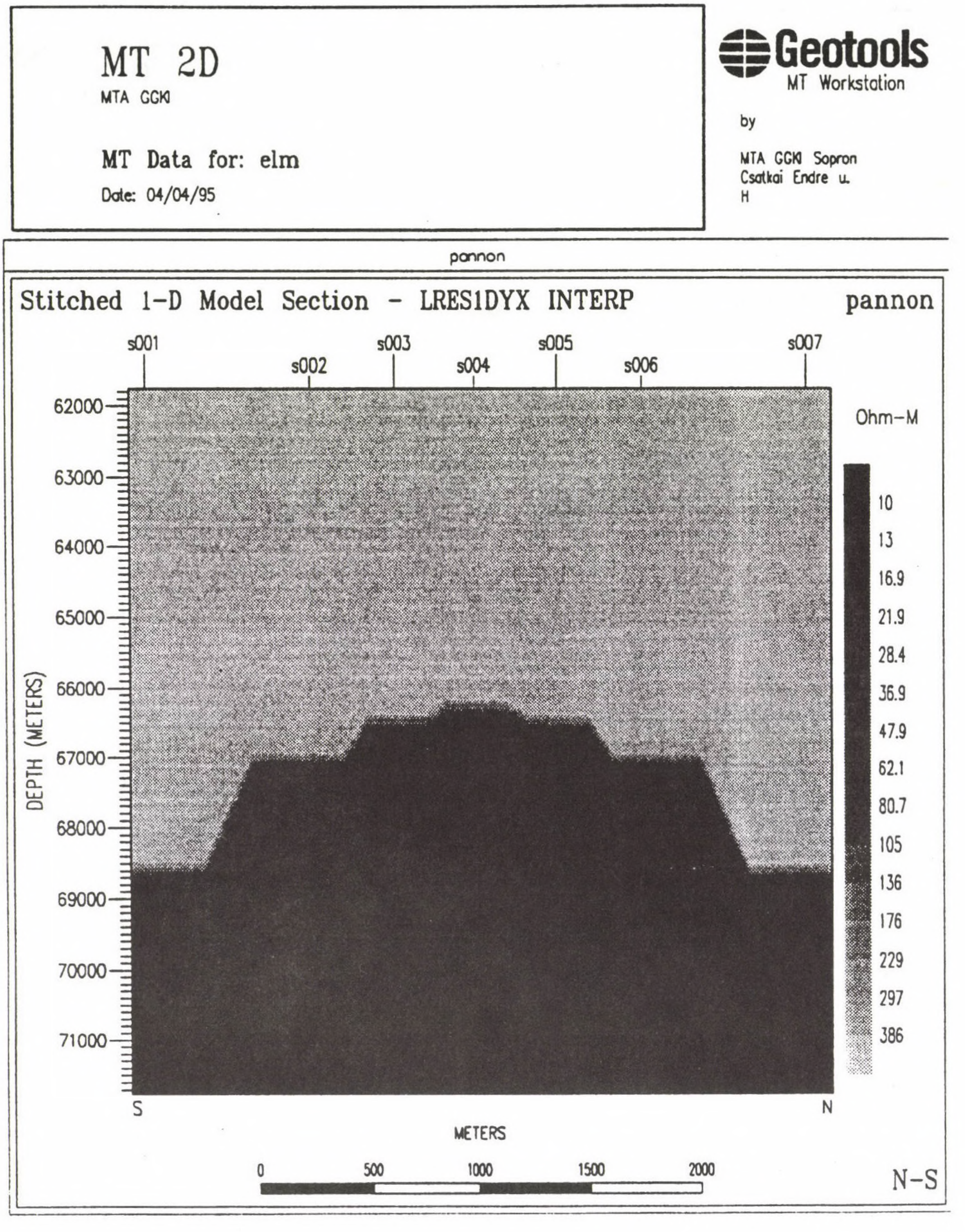


Fig. 9c. Conducting dyke system embedded in resistive matrix with conductive asthenosphere. "B-pol" indication of the conducting asthenosphere

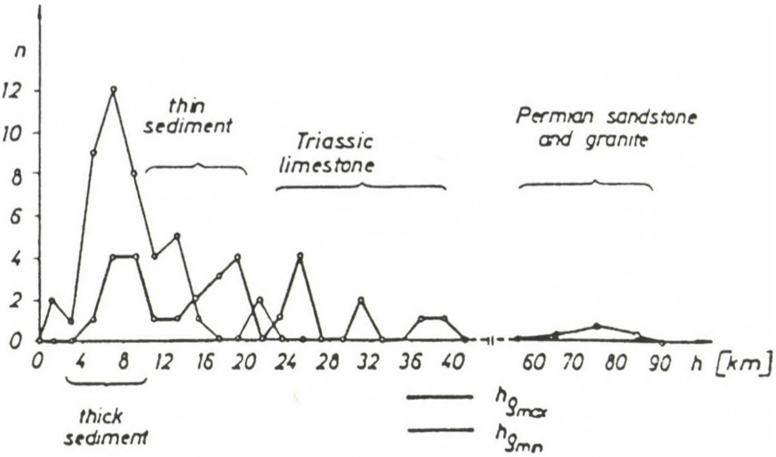


Fig. 10. Relation between the apparent depth of the crustal conductor in the area of the Transdanubian CA and near-surface geological formations (Ádám 1981)

for this area. The static shift practically appears only on the Rhomax curves as B polarized ones giving systematically greater apparent depth values. Hence, the static shift should also be taken into account when studying the regional anisotropy.

2.22 3D effects

A detailed study has been published about the EM distortions in the 3D Békés graben and their effect to the determination of the asthenospheric depth (Ádám et al. 1993). According to model calculations of Berdichevsky and Dmitriev (1976) (Fig. 11) and to our experimental ones *apparent* asthenosphere depth values in the 3D graben derived from the extreme MT curves continuously decrease towards the centre of the graben and are smaller than the real one (Fig. 12). In 1994 a similar effect has been experienced in an other subbasin, in the Makó graben (Ádám 1994).

These grabens have very steep walls as shown by the isopach map, nevertheless, the edge effect can be excluded. Namely, the edge effect should appear perpendicularly both to the eastern-western walls (see measuring sites 6 and 7 in Fig. 12) and to the similarly steep northern wall (sites 4 and 5). In this latter case just the Rhomax represents the quasi B polarization, also an edge effect does not appear here. The situation in the Makó graben is the same. There is great difference between apparent asthenospheric depths derived from Rhomax and Rhomin (MT anisotropy).

It is interesting to remark that in the area of the Békés and Makó grabens there is an almost 90° turn in the direction of Rhomax curves between $T = 16$ and $T = 100$ s from E-W to N-S direction. The latter remains constant at the greater periods (Fig. 13 for site HÓD). The Rhomax direction certainly reflects local effects below 100 s. (This local effect can also be found in the telluric ellipses of same period and in the Wiese arrow corrected for the regional effect.) A similar phenomenon appears in the centre of the Great Hungarian Plain at measuring site Túrkeve-Csodaballa.

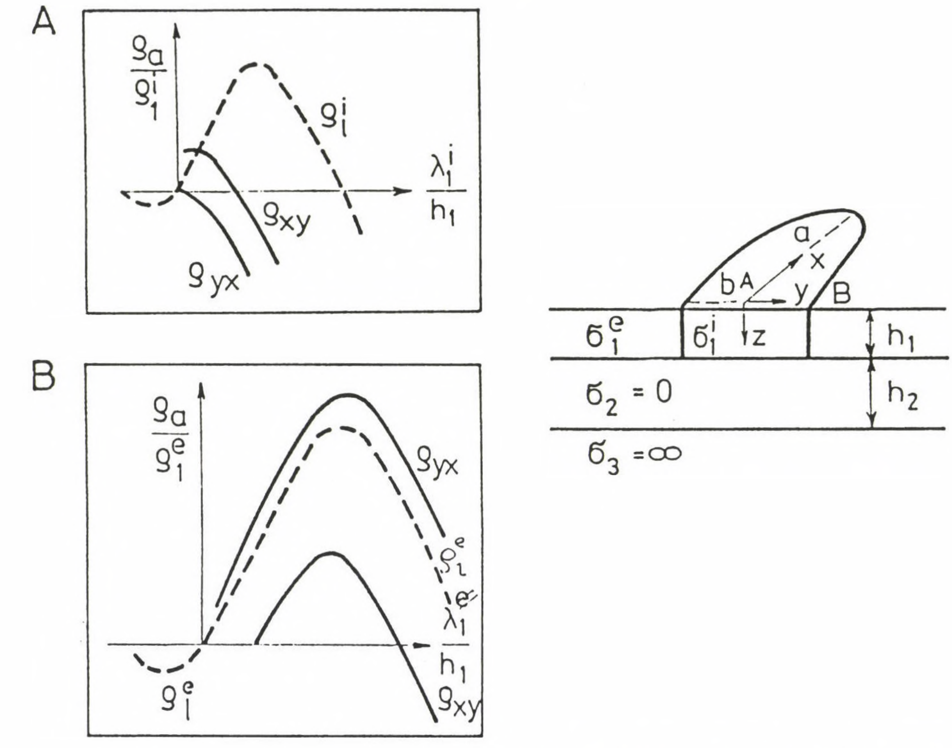


Fig. 11. Berdichevsky and Dmitriev's (1976) 3D model (conducting embedding) and its effect to the quasi E and B polarized MTS curves

The rotation of Rhomax direction represents the transition from a 2D sedimentary basin to a system of 3D subbasins (Békés and Makó grabens) having steep walls. In the 2D basin covered by thick (at least 2–3 km) sediment the great axis of the impedance ellipses (polar diagram) is directed in E-W, i.e. in the strike direction of the 2D structure. When the 3D subbasins start to deepen from the lowest homogeneous level of 2D structure at the period corresponding to this (penetration) depth, the Z_{xy} polardiagram begins to rotate into the northern dip direction of the subbasins. This direction corresponds to the quasi strike direction (quasi E polarization) of the 3D "grabens" in its centre and quasi dip direction (quasi B polarization) at the same time at their northern walls.

As shown by Fig. 13, there is a continuity in the change of the apparent asthenospheric depth between the areas inside and outside the Békés graben (i.e. its northern foreground) and the difference is too much between the two extreme depth values. This phenomenon can only be explained by a regional 2D structure of "horst character" which includes the "3D micro-structures". Therefore the surrounding of these latter ones should be involved into the interpretation.

The mentioned regional 2D structure in the foreground of Békés and Makó subbasins is boundered by a basement elevation (threshold) coming into beeing probably by the Békés tectonic line in E.NE-W.SW. If we want to explain the

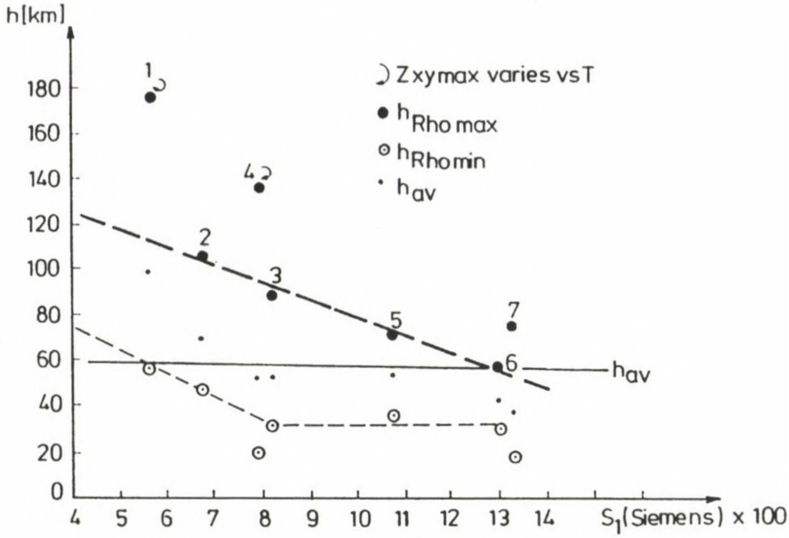


Fig. 12. The relation between the apparent depth to the asthenosphere and the conductance of the surface sediment (S_1) in the Békés graben (Ádám et al. 1993)

regional anisotropy similar effects should be generalized in connection with the almost parallel longitudinal (and transversal) tectonic line system — appearing also in Horváth's (1993) Neogene tectonic map (Fig. 14) and Kalmár et al's (1995) gray scale map of the pre-Tertiary basement (Fig. 15) — forming "basin and range" structures from the resistive basement and it seems that the "horst" effect prevails in the character of the distortion.

Returning to the case history of the 3D Békés graben, the most reliable asthenospheric depth at the rim of this subbasin can be get by the extrapolation of the Rhomin depth values towards the rim in Fig. 11 where the 3D effect can be neglegted. The depth of the asthenosphere is 60–70 km here.

The Wiese induction vectors of long periods (T $\geq$ 20 min) in the Pannonian Basin are southward directed, i.e. in the dip direction of the block structures, representing the direction of B polarization, too (Fig. 16). Nevertheless, their regional content is questionable, which may be due to the Carpathian conductivity anomaly in the Western and Eastern Carpathians.

The role of the conductive basement fractures in the MT anisotropy of the mantle, according to Mareschal et al's (1993) idea — on the basis of the parallel conductive dike model — should be further studied first of all their physical reality in the plastic mantle.

Most recently a new question emerged: how does influence the local elevation of the conducting asthenosphere (mantle plume) beneath deep extensional subbasins (e.g. along the Pannonian Geotraverse (PGT 1) the Békés graben, Jászság subsidence) the regional MT anisotropy. At present there is only a rough indication of these local upwelling of the asthenosphere in the MT image (Ádám et al. 1996) which is not enough for correct modelling. Information on the shape, the dimen-

Table I. Numerical models for the explanation of the regional MT anisotropy

Characteristic difference between Rhomax and Rhomin curves	Numerical model	Consequences
Angle difference in S-interval	Anisotropic basement rock ($\varrho \to \infty$)	It satisfactorily explains the angle difference between the increasing branches of the extreme MT curves
		It does not give explanation for the difference in the apparent depth of conductors appearing in the E- and B-pol. curves
Difference in the apparent depth to the crustal and mantle conductor (asthenosphere) given by the extreme MT curves	2D "horst" basement model	The B-pol. curves (Rhomax) give *greater apparent depths* for the asthenospheric layer than the real one indicated by E-pol. curves
	2D "graben" in the basement	The B-pol. (Rhomin) curves give *shallower apparent depths* for the asthenosphere than the real one indicated by E-pol. curves
	2D single conductive "dike" embedded in resistive matrix on conductive half space (asthenosphere)	Both the E- and B-pol. curves give apparent depth value for the top of the dike which increases with distance from the dike centre. The indication of B-pol. (Rhomax) curve is much weaker therefore more unreliable than the E-pol. curve (Rhomin). The B-pol. curves — in addition — can be strongly distorted by galvanic effects. Concerning the asthenosphere the B-pol. indication similar as in case of a graben model and gives apparently shallower depth, than the real ones
	2D parallel conductive dike system embedded in resistive matrix on conductive half space (asthenosphere)	The B-pol. curves in the centre of the system can give *greater apparent depth* — if any — to the top of the dike system than the real one. The indication of the asthenosphere is similar as in case of a 2D graben (and of a single long dike).
	3D effects: 3D sedimentary basin with conducting asthenospheric layer	Both *quasi* E- and B-pol. curves indicate the asthenosphere shallower than it is in reality
	Elevation of the conducting asthenosphere beneath narrow and deep extensional subbasins (e.g. Békés graben, Jászság subsidence)	The character of the E- and B-pol. curves could be very similar to those of crustal (near-surface) conducting dikes well approximating the measured deep sounding curves (Fig. 1) nevertheless the contrast in B-pol. effect of the relief of the resistive basement and of conductive asthenosphere has to be taken into account (Fig. 7)

Conclusion: The great difference between the apparent asthenospheric depths derived from the extreme (E- and B-pol.) MT curves can be explained by "*basement horst structure*" or "*parallel conductive dyke system*" in the asthenosphere corresponding to the local elevation of the asthenosphere.

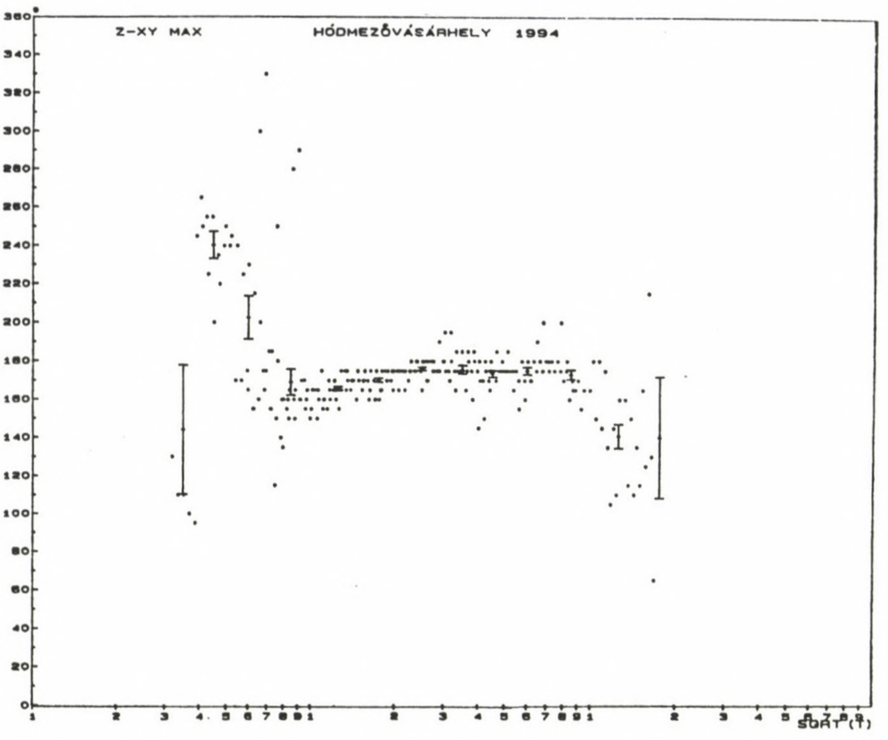


Fig. 13. The change of Rhomax direction in function of period in MT site HÓD in the Makó graben (see in Figs 1 and 2)

sion and regional distribution of these geometrical changes in the asthenosphere is needed for the modelling, nevertheless, their effect — similarly to the idea of Mareschal et al's (1993) model — cannot be excluded. The character of the E and B polarized sounding curves of a conducting dike (Ádám 1987) is very similar to the shape (indication) of the deep MT curves in the Pannonian Basin as shown in Fig. 1. As mentioned in connection with Fig. 7 the contrast of the B polarization effect of the relief of the resistive basement and of the conductive asthenosphere has to be taken into acount.

Conclusions

It is our fundamental interest to find the cause(s) of MT anisotropy to be able to explain the great difference between the deep structures derived from the extreme (Rhomax and Rhomin) sounding curves in many MT measuring sites in the Pannonian Basin. MT soundings carried out in quasi-isotrope, homogeneous area (e.g. in Nagycenk observatory, Ádám et al. 1981) hint at the greater reliability (i.e. withouth distortion) of those conductive structures which have been calculated by MT curves supposedly being of E polarization (Ádám et al. 1982). These conductive structures are in better agreement with other physical parameter of the Pannonian Basin, first of all with its high geothermal anomaly.

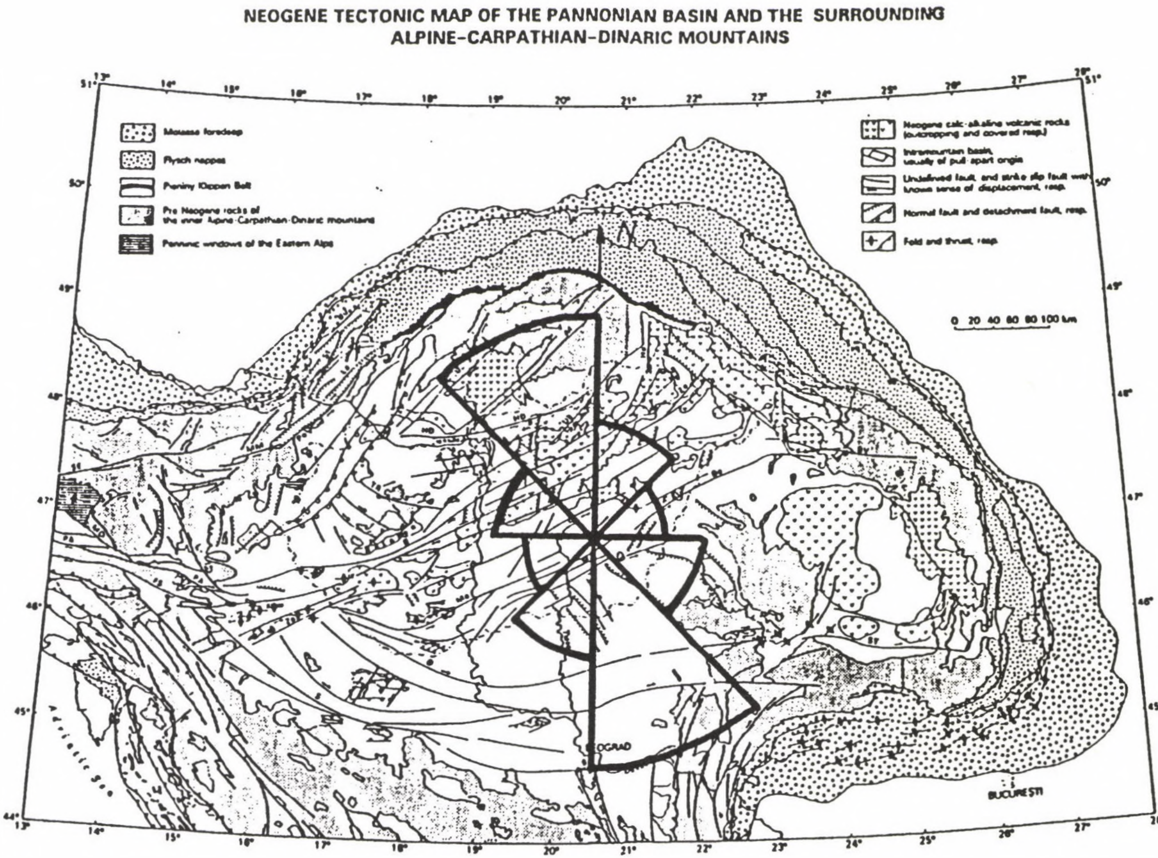


Fig. 14. Neogene tectonic map of the Pannonian basin and the surrounding Alpine-Carpathian-Dinaric Mountains (Horváth 1993) with the rose diagram of Rhomax directions in the area of the Transdanubian CA

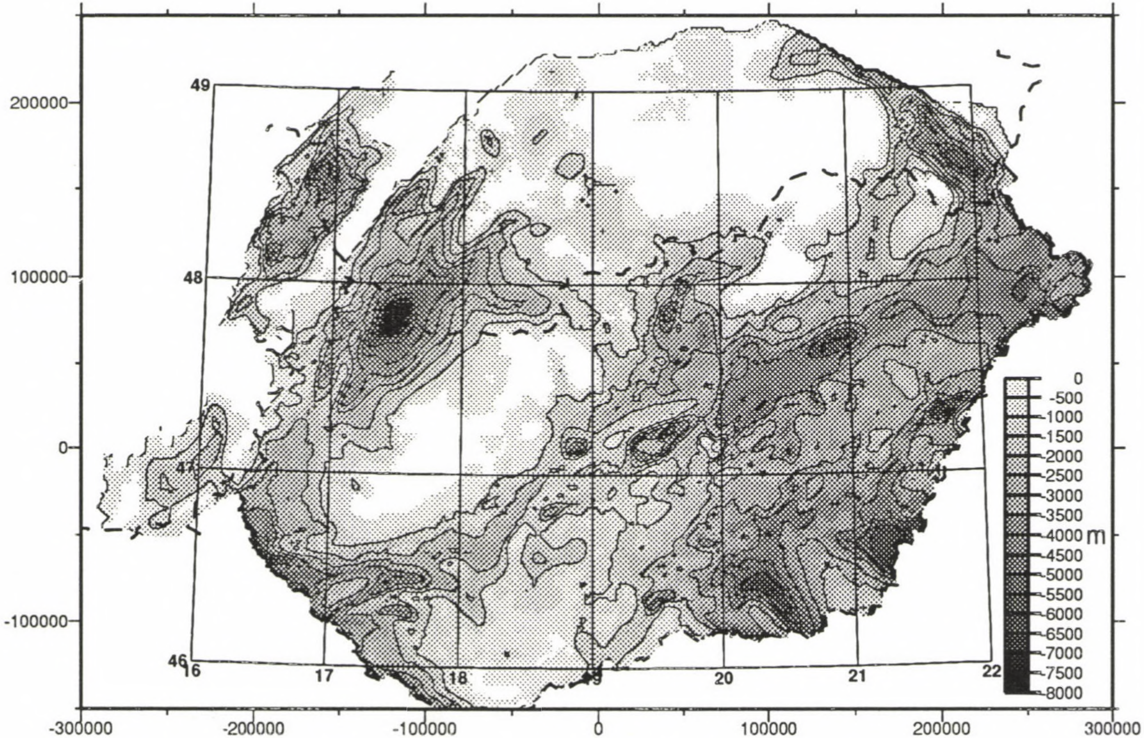


Fig. 15. Gray-scale map of pre-Tertiary basement of the Pannonian Basin (Kalmár et al. 1995)

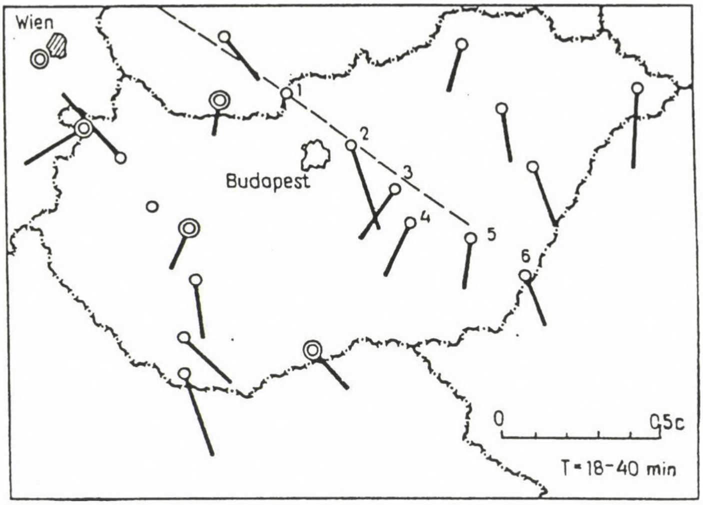


Fig. 16. Induction vectors of long periods (T $\geq$ 20 min) in the Pannonian basin (Ádám et al. 1972b)

The map of the Neogene tectonics (Horváth 1993) and of the crystalline basement structures of the Pannonian Basin (without sediment) (Fülöp and Dank, MÁFI 1987) both hint at the fracture tectonics of the basement, the mechanically strongly crushed, loose and therefore conductive linear structures having at least two main directions (longitudinal and transversal) which are closly related to the MT anisotropy. The structure of this type of fracture cutting through the whole crust is shown by Posgay and Szentgyörgyi's (1991) seismic section. In connection with this strike slip zone (in Szolnok flisch area), Ádám and Steiner (1993) studied the probability of the deepening of the crustal and mantle conductive layers here.

On the basis of the geologic-tectonic structure of the Pannonian Basin, of the studied numerical models, of the experience got from the investigation of the graphitic blocks of the Transdanubian crustal anomaly (parallel isolated conductive dikes); static shift due to different geological formations, and of the detailed study of the 3D subbasins, the following conclusions can be drawn concerning the MT regional anisotropy observed here.

It should be taken into account:

— the intrinsic anisotropy of the sediments and first of all that of the metamorphic crystalline basement rocks. This anisotropy does not distort the determination of the conductive asthenosphere;

— the almost parallel tectonic line system of the Pannonian Basin directed in E(ENE)-W(WSW) and perpendicular to them (Horváth 1993) which form "basin and range" structure in the basement and probably represents a conductive dike system, too, cutting through the crust;

— a quite new aspect could be the local upwelling of the asthenosphere beneath deep extensional subbasins boundaries of which are determined by the above mentioned tectonic lines.

The determining effect in the MT anisotropy is given by this block-structure. This structure — as has been shown in case of the 3D Békés and Makó graben — in most cases causes a "horst type 2D B-polarized distortion", i.e. the depths to the crustal and mantle conductive layers derived from B polarized Rhomax curves are much greater than the real ones and the depths are better approximated by E polarized Rhomin curves. As remarked in connection with the B polarized sounding curves measured in the area of the Transdanubian conductivity anomaly first of all these Rhomax curves are distorted by the static shift, too.

Some conclusions are summarized in Table I.

Acknowledgements

The author is indebted to the National Research Fund (OTKA) for the financial support (OTKA Project T-014882). This publication belongs to the EUROPROBE project, too.

References

Ádám A 1981: *Acta Geod. Geoph. Mont. Hung.*, 16, 97–113.

Ádám A 1987: *Gerlands Beitr. Geophysik*, 96, 17–31.

Ádám A 1992: Geological and methodical significance of the Transdanubian conductivity anomaly (in Hungarian). Inaugural lecture, Akadémiai Kiadó, Budapest, 1–40.

Ádám A 1994: Report on the change of conducting formations in the Pannonian basin by deep MTS (in Hungarian). Manuscript, GGRI, Sopron

Ádám A, Steiner T 1993: *Phys. Earth Planet. Int.*, 81, 1–8.

Ádám A, Verő J, Cz Miletits J, Holló L, Wallner Á 1981: *Acta Geod. Geoph. Mont. Hung.*, 16, 333–351.

Ádám A, Vanyan L L, Varlamov Ɖ A, Yegorov I V, Shilovski A P, Shilovski P P 1982: *Phys. Earth Planet. Int.*, 28, 251–260.

Ádám A, Verő J, Varga Gy 1972a: *Acta Geod. Geoph. Mont. Hung.*, 7, 125–136.

Ádám A, Verő J, Wallner Á, 1972b: *Acta Geod. Geoph. Mont. Hung.*, 7, 251–287.

Ádám A, Szarka L, Steiner T 1993: *J. Geomag. Geoelectr.*, 45, 761–773.

Ádám A, Prácser E, Szarka L, Varga G 1996: Mantle diapirs or EM distortions in the Pannonian Basin? (Inversion of the deep magnetotelluric (MT) sounding along the Pannonian Geotraverse.) Submitted to Geophysical Transactions

Berdichevsky M N, Dmitriev V I 1976: In: Geoelectric and Geothermal Studies, KAPG Geophysical Monograph, A Ádám ed., Akadémiai Kiadó, Budapest, 165–221.

Fülöp J, Dank V ed. 1987: Geologic Map of Hungary without Cenozoic layers. MÁFI

Horváth F 1993: *Tectonophysics*, 226, 333–357.

Kalmár J, Papp G, Szabó T 1995: *Comp. and Geosci.*, 21, 245–257.

Mareschal M, Kellett R, Kurtz R, Bailay R, Ji Schaocheng 1993: Small scale electrical heterogeneities in the upper mantle of the Canadian Shield. IAGA Bulletin No. 55, Part B. Abstracts, 137.

O'Brien D P, Morrison H F 1967: *Geophysics*, 27, 668–678.

Posgay K, Szentgyörgyi K 1991: *Magyar Geofizika*, 32, 1–15.

Tátrallyay M 1977: *Acta Geod. Geoph. Mont. Hung.*, 12, 279–285.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 217–229 (1996)

THE STRUCTURE OF THE UPPER CRUST IN THE SOUTH PORTUGUESE ZONE FROM MAGNETOTELLURIC STUDIES — PRELIMINARY RESULTS

E P de Almeida[1], F A Santos[1], L A Mendes-Victor[1]

[Manuscript received May , 1996]

Five MT soundings were carried out in the South Portuguese Zone, in order to get a first approach to the conductive structure associated with the regional geological features. The profile running roughly NNE-SSW is 40 km long and it crosses the main flysch formations. The recording frequency ranges of 180–1/125 Hz were appropriate for defining the upper crust structure and for imaging major geoelectrical features in the middle crust. The preliminary processing included impedance tensor analysis, 1D inversion and modelling of the apparent resistivity and phase responses. A 2D model of electrical conductivity structure associated with the southern part of the portuguese mainland was obtained from forward 2D numerical modelling. The resulting structures are consistent with results previously obtained from seismic refraction data and shows a deep fault (strike roughly N115E), that may represents the south limit of the pyrite belt. A less resistivity zone (50–20 Ohm m) occurs in the Upper Paleozoic crust and can be tentatively associated with the presence of blackshist facies rocks. Two conductive blocks in the upper crust (resistivity of 30–35 Ohm m) were clearly defined. Those blocks might correspond to major shear zones associated with faults. The derived resistivity structure suggests that the proposed detachment level might occur at 15 km depth.

Keywords: electrical conductivity; Iberia; magnetotellurics; modelling; Portugal

Introduction

The available geophysical information concerning deep structures of western Iberia has been obtained from seismic refraction surveys and from gravity and aeromagnetic maps. Generally a global agreement, between the regional variscian zones and the geophysical characteristics of each zone, was recognized. However, several features are still without a clear geophysical and geological interpretation (Mendes Victor et al. 1993, Miranda et al. 1989, Torres and Lisboa 1988).

According to the accepted model for the formation and evolution of the SW variscian foldbelt, the South Portuguese "terrain" and the Iberia "terrain" crustal structures would be the result of a continent-continent collision with a NE to E subduction (Ribeiro et al. 1988, Silva et al. 1990). The South Portuguese Zone (SPZ) consists of Upper Paleozoic low-grade sediments ante volcanics, the Pyrite belt, and is separated from the Lower Paleozoic metasediments and abundant granitic

[1]Departamento de Física da Universidade de Lisboa and Centro de Geofísica da Universidade de Lisboa, R. Escola Politécnica, 58, 1250 Lisboa, Portugal, e-mail: fraissa@skull.cc.fc.ul.pt

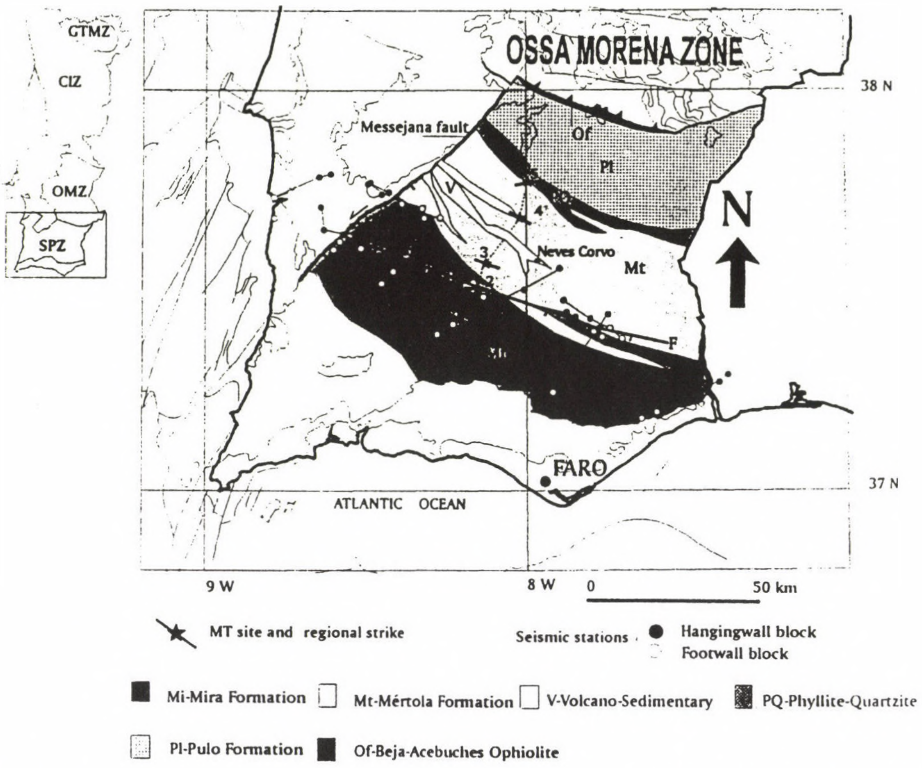


Fig. 1. Location of MT sites (numbers) and main geological setting of the region (simplified from Oliveira (1990) and Matias (1996)). The open and solid circles represent footwall and hangingwall blocks interpreted from seismic refraction profiles. GTMZ, Galicia-Trás-os-Montes subzone; CIZ, Central Iberian Zone; OMZ, Ossa-Morena Zone; SPZ, South Portuguese Zone

intrusions of the Ossa Morena Zone (OMZ) by an Hercynian suture that contains ophiolitic sequences (Beja-Acebuches ultramafics).

To improve the knowledge on the deep geology, a profile of deep magnetotelluric (MT) soundings, oriented approximately in the NE-SW direction has now being carried out by the Centro de Geofísica da Universidade de Lisboa/Faculdade de Ciências as a part of the Europrobe/Iberia project (Fig. 1). The main goals are: 1. to determine the electrical resistivity distribution in the crust, 2. to detect the faults and sole detachment levels, mainly if they are high dipping and 3. to confirm the existence of the suture (given that suture zones appear with high anomalous conductivity in other parts of the world).

In this paper the preliminary results concerning the resistivity distribution in the upper and middle crust of the South Portuguese Zone will be presented.

Magnetotelluric data

The MT profile is about 40 km long, extending from the southern limit of the SPZ to the "pyrite belt" (Fig. 1). Distances between stations were between 8 and 14 km. The MT data were acquired, using a Garchy MT system (CNRS-GeoInstruments, France), with four channnels. The recording frequency range was 180–1/125 Hz. The measurement direction of the horizontal fields were nearly N45E and N135E, in accordance with the main direction of the variscian structures. The time series have been processed, after visual inspection, by using the cascade decimation and a robust method based on Sutarno and Vozoff (1991) work. Although each measuring site was carefully selected, all the recorded data contain a considerable amount of noise in the 1–1/10 Hz band, mainly associated with power lines, mining activity and urban activity.

The MT impedance tensor at each station was examined to determine its dimensionality using the three normalized dimensional weights (D1, D2 and D3) introduced by Kao and Orr (1982). According to those authors, the dimensional weights attempt to assess the relative importance of 1D, 2D and 3D structures contributions simultaneously. The dimensional weights for the sites 1 and 3 are shown in Fig. 2a, together with skew values. In both stations the largest dimensional weight is D1. However, the reciprocal behavior of the D1 and D2 weights, in station 3, suggests that the contribution of the 2D structures increases with depth. The data exhibit low values of skew (< 0.2) and D3 weight, suggesting that the contribution of 3D conductive structures is not important. As was pointed out by Kurtz et al. (1990), there is no set acceptable upper value for skew, but generally values greater than 0.2 are taken to indicate the presence of 3D structures.

The impedance tensor, at each station, was analyzed to determine the regional direction, following the methods described by Zhang et al. (1987), and by Chakridi et al. (1992). Taking into account tectonic considerations, a regional strike nearly E-W, on average, was obtained using the former method (Fig. 2b). The values obtained from the latter method are ranging from N82E to N120E (Fig. 1).

It has been a common practice to find individual strike directions at each site and to present the apparent resistivity and phase curves in the rotate frames. If the results from different sites are referred to different coordinate axes, a simple interpretation of the data is impossible, and a regional strike must be obtained (Gamble et al. 1982). Figure 1 shows that there is a stable strike of about N115E for four sites in good agreement with the direction of the main geological structures in the SPZ. Exception is site 5 with a strike of N82E, revealing the influence of electrical structures related to the "pyrite belt". Despite this local deviation we decided to take a N115E regional strike as representative of the region.

After the regional strike had been selected, the MT data were rotated into a coordinate system with the Y axis along the strike direction, so that the electric fields in the X direction (N25E) represent the B-polarization and the orthogonal electric fields are the E-polarization.

Figures 3 and 4 show B and E polarization of the rotated curves for the five sites. The data acquired over the Baixo Alentejo flysch group (sites 1, 2 and 3), do not

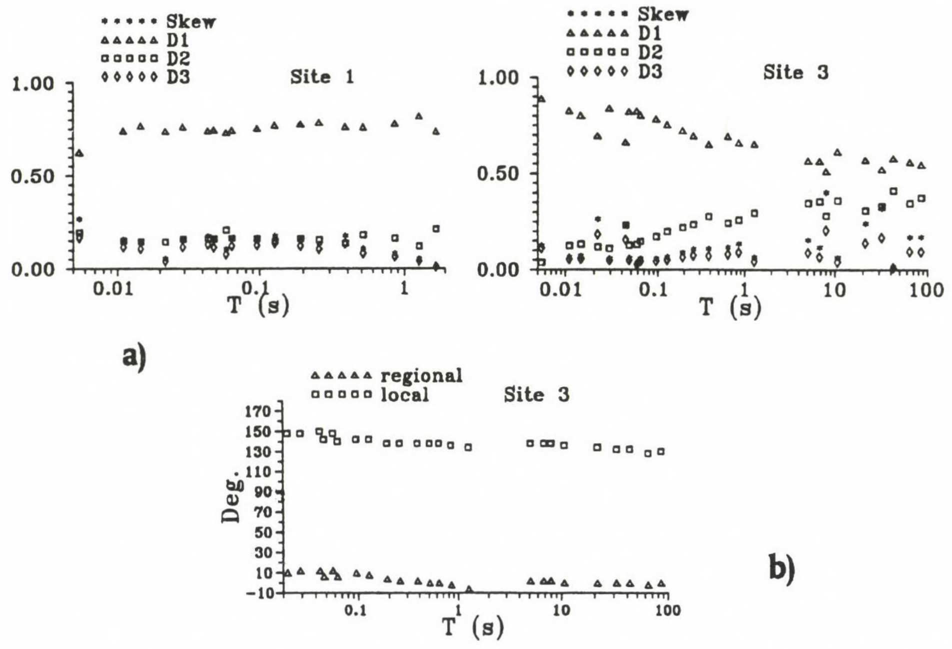


Fig. 2. The variation with period of, a) dimensional factors D1, D2, D3 and Skew for sites 1 and 3, b) regional and local strikes for site 3 based on Zhang et al. (1987) algorithm

show a strong scattering in the apparent resistivities, all the values corresponding to the shortest periods being in the same decade. The data acquired over the "pyrite belt" (Figs 4b, 4c, sites 4 and 5), show higher apparent resistivity values. No control of the possible static shift effect is possible using MT data alone (Kurtz et al. 1993), and statistically the number of stations is not significant to provide an estimate of the level of the regional apparent resistivity curve. Therefore we decided to deal with the superficial distortion by imposing constraints from the previous seismic refraction studies.

Qualitative analysis shows that the stations 1 and 2 are characterized by a similar shape, suggesting the presence of local conductors beneath the uppermost layers. At station 3, located near the south border of the Mértola formation, the apparent resistivity in the E-polarization, separates from the B-polarization in the range of 0.1 to 1 s. The divergence of both polarizations suggests the presence of lateral conductors. The apparent resistivity curves corresponding to northern sites (4 and 5), are different from those to south, revealing the influence of the volcano sedimentary complex.

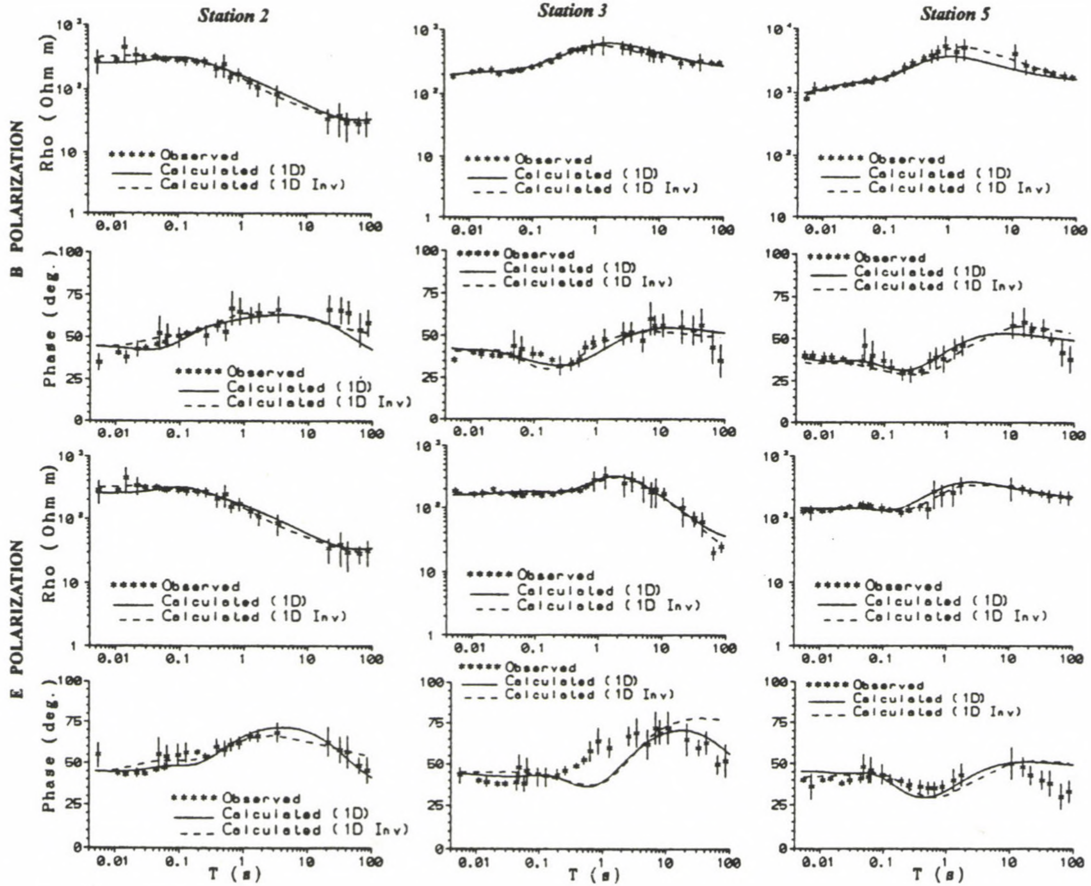


Fig. 3. MT data and 1D inversion and modelling responses for three representative sites. Note that the apparent resistivity curves at site 5 were not static-shift corrected

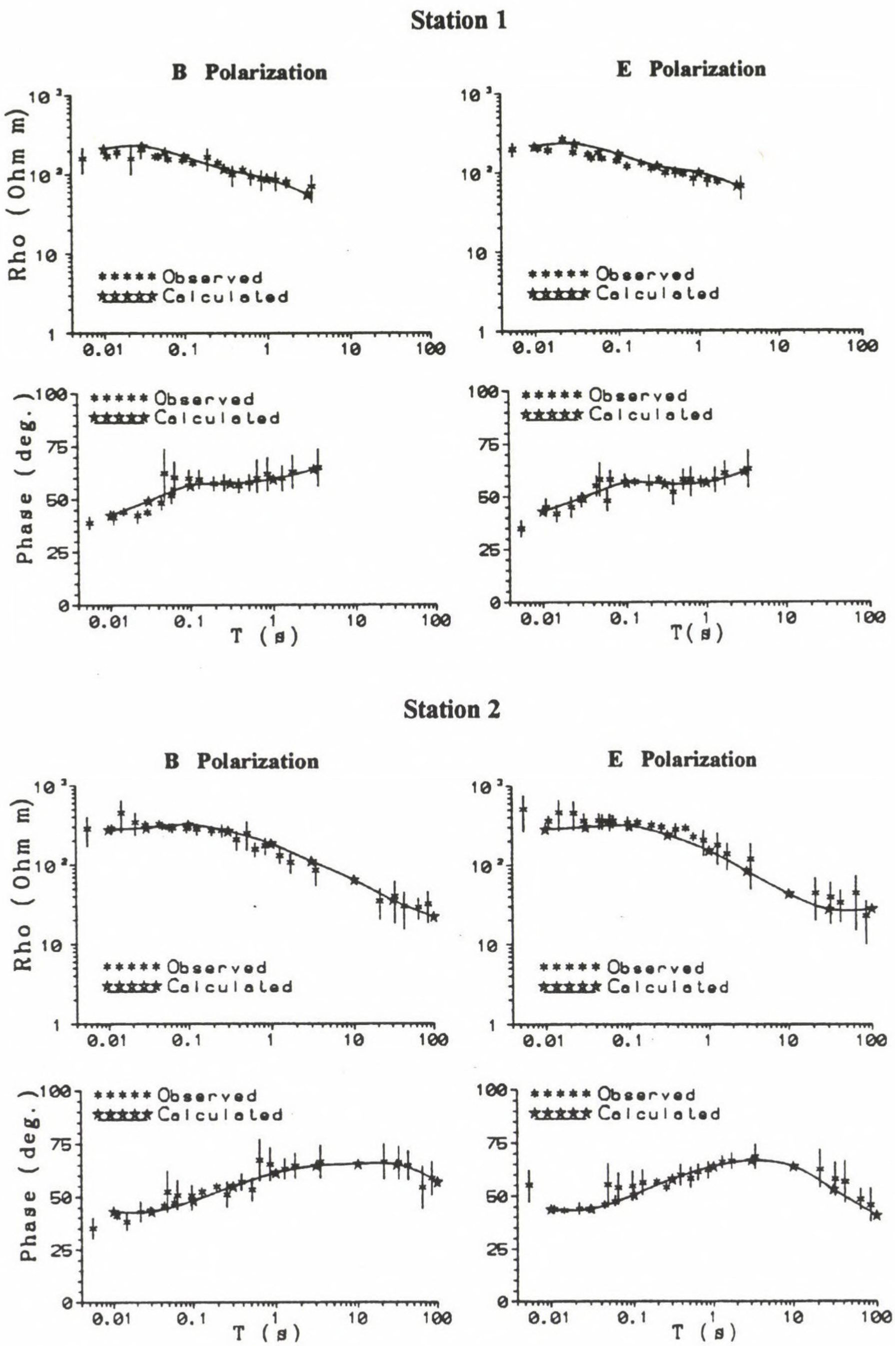


Fig. 4a. MT data and 2D model responses. Note that the apparent resistivity curves at site 5 were static-shift corrected

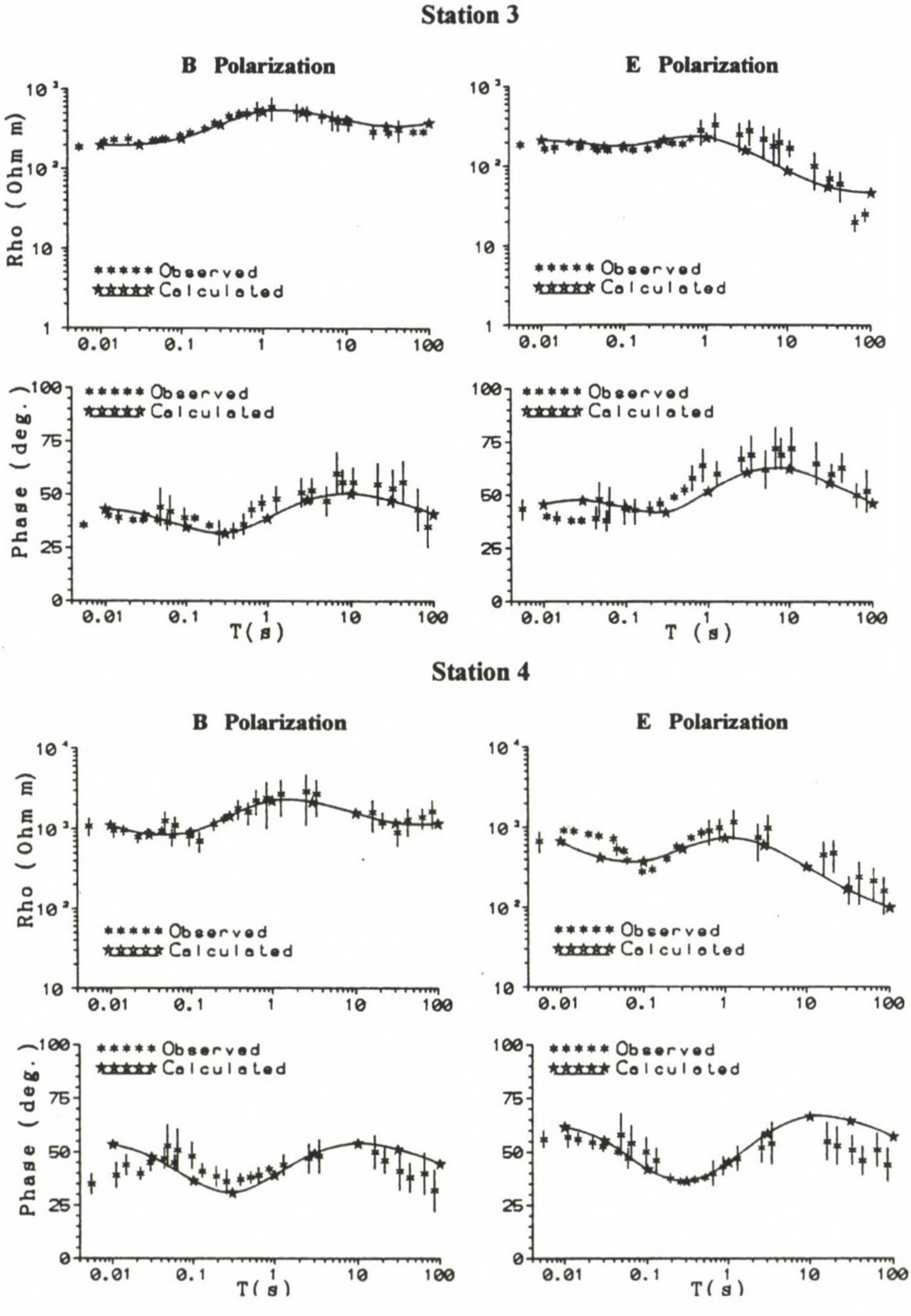

Station 3
B Polarization
E Polarization
Rho (Ohm m)
Phase (deg.)
***** Observed
Calculated
T(s)
T (s)
Station 4
B Polarization
E Polarization
Rho (Ohm m)
Phase (deg.)
***** Observed
Calculated
T(s)

Fig. 4b (contd.)

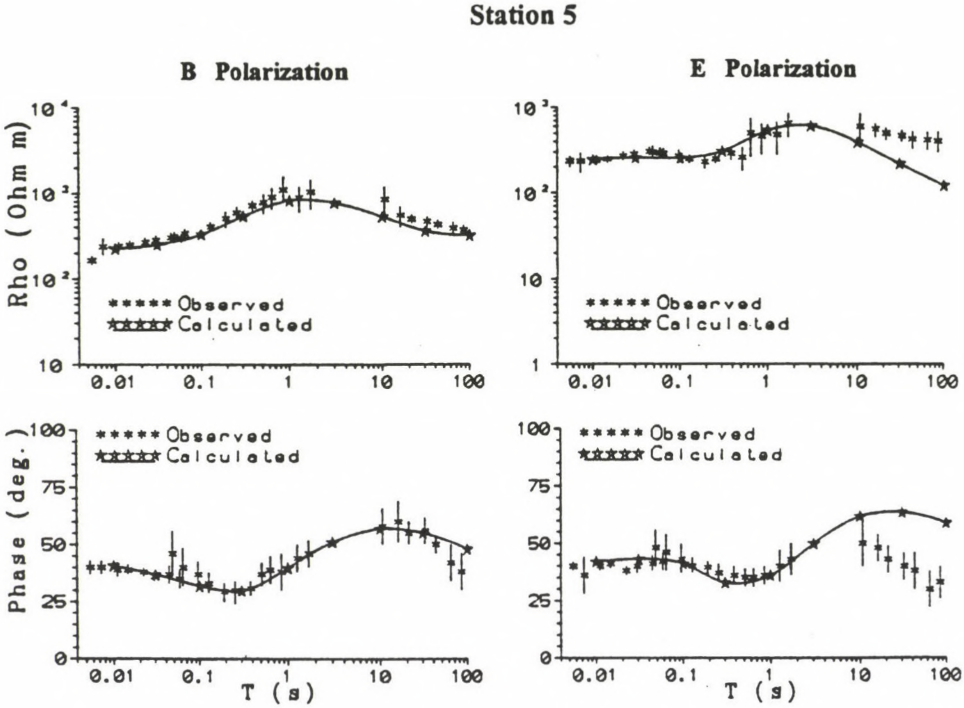


Fig. 4c (contd.)

Interpretation

1D inversion and modelling

As a preliminary step to 2D modelling the rotated data were inverted for minimum 1D earth structure using a program based on the algorithm presented by Johansen (1977). The obtained models for both polarizations are presented in Fig. 5. Figure 3 shows the data and model responses for three representative sites of the model: site 1 for the southern part, sounding 3 near the main deep fault and site 5, located on the phyllite-quartzite formation.

Taking into account seismic information related to interfaces, 1D trial-and-error modelling (Patra and Mallick 1980) was also performed to estimate the electrical conductivity distribution at each site. The results were obtained by adjusting the parameters of the model in order to reach a satisfactory fitting between the calculated and observed values of apparent resistivity and phase (Fig. 3). The parameters for each site were correlated with those of the adjacent sites in order to obtain the most consistent model. The final results are presented as layered models for both polarizations in Fig. 5b. These results were obtained from MT data in the rotated coordinate system (N25E-N115E) without static-shift corrections. The agreement between both experimental and calculated curves is generally satisfactory. Exception is the E-polarization phase at site 3, in the period range from 0.1 to 10 s, which

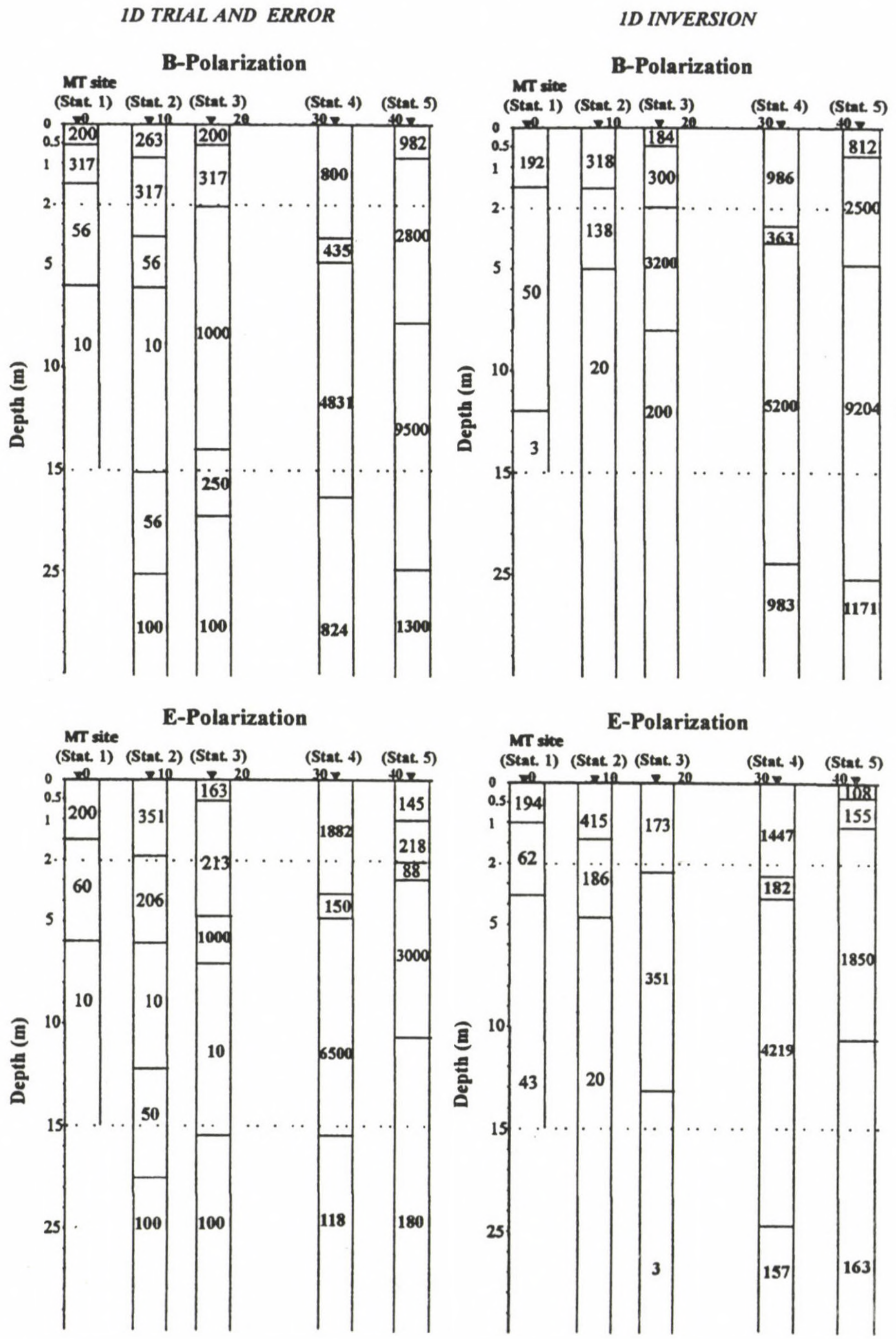


Fig. 5. 1D electrical resistivity models obtained a) from inversion for minimum layers and b) from trial-and-error modelling taking into account seismic information. Note the vertical scale changing at 2 km depth (dashed line)

is not well fit by the 1D inversion and modelling (Fig. 3). This feature seems to be associated with a 2D conductive structure, as it will be discussed further below.

Both presented composite 1D cross-sections show the main structural trends, in particular the lateral contact associated with a fault south of site 3, and the low resistivity zone beneath the Mira flysch formation (sites 1 and 2).

Considering the second layer as being homogeneous, with a resistivity of 300 Ohm m, and constraining the thickness of the flysch group to 4–5 km, the data from site 5 was static-shift corrected.

2D modelling

The 2D trial-and-error model was obtained from the apparent resistivity curves rotated in accordance with the mean regional strike (N115E), using the finite-element algorithm described by Rijo (1977). Taking into account the interpretation of the seismic refraction profiles (Matias 1996), some relevant interfaces in the crust were constrained. The 1D (B-polarization) resistivity distribution was used as a starting model to compute the 2D model. The final resistivity model with its geological interpretation is shown in Fig. 6. Detailed interpretation of the resistivity structure is not possible because very few soundings are available. However, three outstanding features are seen in the model:

1. The Mira and Mértola formations, with resistivities ranging from 200 to 300 Ohm m and thickness varying between 1.5 and 3 km, seem to be quite uniform along whole profile;
2. The Upper Paleozoic (at depths greater than 4 km) presents a strong change on resistivity. Indeed, unexpected low resistivity values (50–20 Ohm m) occur in the southern part of the studied zone;
3. Two blocks with low resistivity (30-35 Ohm m) beneath stations 3 and 5 are clearly defined.

Figure 4 shows the agreement between the observed and calculated data (both polarizations). In general the B-polarization model response, calculated at 3 points per decade on a logarithmic scale, agree fairly well with the field data. The model does not fit the E-polarization data as well as the B-polarization at stations 3 and 5. Nevertheless, a reasonable fit was obtained for E-polarization from the southern part of the profile, i. e., sites 1 and 2, located relatively far removed from the fault and with a predominant 1D structure. At sounding 5 the E-polarization model response does not fit the observed data, probably because the site has a different strike preference (N82E).

Discussion

The electrical structure of the uppermost crust is characterized by two layers with resistivity ranging from 200 to 300 Ohm m that, in according to exposed geology, correspond to the Mira formation (Namurian) and the Mértola formation

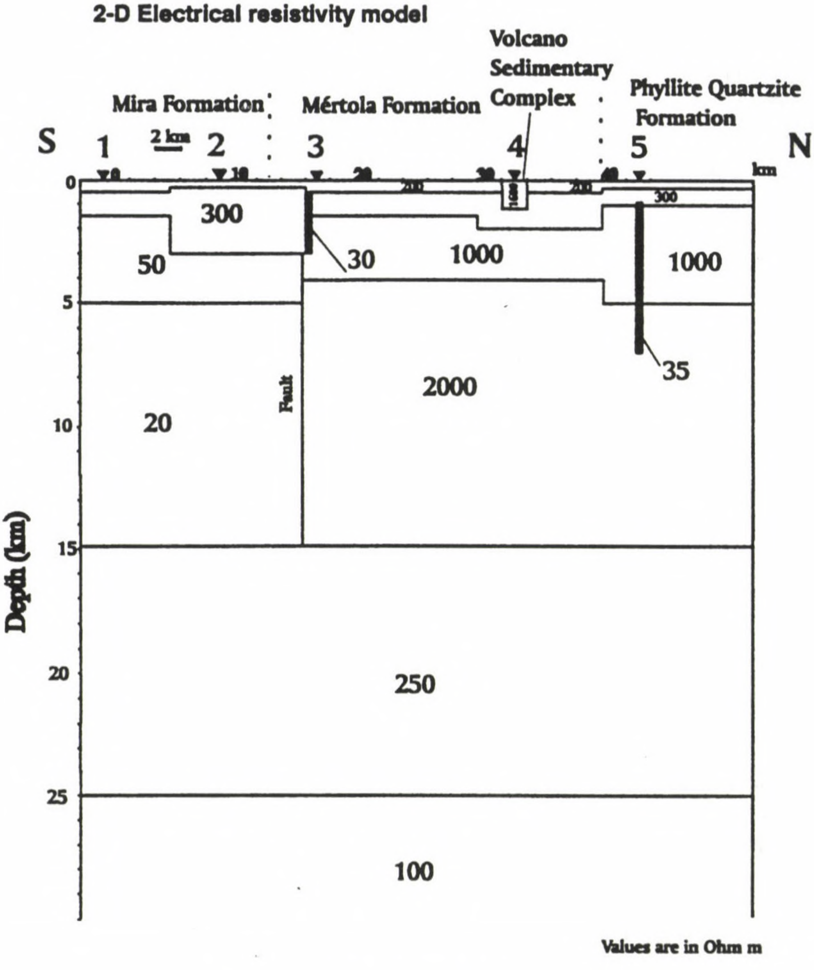


Fig. 6. 2D electrical resistivity model

(Upper Visean). The lithology is dominated by shales, siltstones and greywackes that form the Baixo Alentejo flysch group (Oliveira 1990). In the southern part of the profile those two zones are underlain by a less resistive layer (50 Ohm m). In this area the frequency range is not able to resolve the interface at a depth of 5 km, that is shown in seismic refraction profiles (Matias 1996). The zone of 20 Ohm m is required to produce the descending apparent resistivity values which occur with increasing period at site 2. However the top and penetration of this structure are not well controlled. The question concerning the origin of the good conductivity of those zones has not any answer, at present. However, the available geological information suggests that the low resistivity will probably be associated with black (graphitic) schist (Ribeiro, personal communication).

Beneath the Mértola formation, a resistive layer (1000 Ohm m) is required at

1.5 to 4 km depth, to explain the increase of the apparent resistivity at sites 3, 4 and 5. The bottom of this layer was constrained by seismic data. The spatial sampling of the MT soundings is too large to detail the deep transition between Mira and Mértola formations. However, the model shows that it can be interpreted as a near vertical fault (normal/reverse ?). The narrow conductive block extending at least to 3 km is only defined by the E-polarization data at site 3, and may be associated with fluid transport effects through the shear zone. The model shows that, in the southern part, the sedimentary layers are thicker than in the northern part. These results are in good accordance with the seismic interpretation, as can be seen in Fig. 1, where the hangingwall and footwall blocks are represented by solid and open circles respectively.

A high resistivity (1600 Ohm m) unit is modeled beneath site 4, corresponding to local features related to volcano sedimentary formation. The northern part of the model shows that the thrust fault associated with the phyllite-quartzite formation is a quite shallow tectonic structure, which is in good agreement with the known geology of the region.

The conductive block underneath site 5 is required to fit the E-polarization data and probably is associated with a fault at the north edge of the "pyrite belt". However we can not exclude the possibility that, at least the upper part of this conductive feature, might represent disseminated mineralization.

There are few longer period data. Therefore the structure at depths greater than 10 km is very poorly resolved. The interfaces at 15 and 25 km were constrained from seismic data. Some sensitivity tests performed by perturbing the depth of the 15 km interface show that it can vary from 12 to 18 km without significant data misfit. It is interesting, however, to see that the geoelectrical structure at depths greater than 15 km seems to be uniform along whole the studied region, confirming the geological hypothesis that the Lower Paleozoic was not deformed (Ribeiro et al. 1988).

According to Matias (1996), a seismic velocity anisotropy, in the layer beneath 5 km depth, is revealed by seismic data. The velocity changes with azimuth being the fast velocity axis oriented WNW-ESE. The value of anisotropy (6–7 %) in the upper crust can be attributed to foliation in a schistose formation, and an equivalent electrical anisotropy, associated with the 2000 Ohm m layer, would be expected. However, the available MT data seem not support this hypothesis. The characteristic divergences in orthogonal resistivity components seem to be quite consistent with 2D structure.

Conclusions

On the basis of our MT results, the SPZ can be divided into two zones, the deep boundary between them running in an almost N115E direction. MT and seismic data modelling suggest that this boundary is a fault with a thicker sedimentary formation to the south. Besides, the regional lateral resistivity variation in the Upper Paleozoic indicates that the fault is deep and suggests new aspects for the geodynamic evolution of the SPZ, mainly, that the detachment level, between Upper and Lower Paleozoic, will probably be deeper than supposed (Silva et al. 1990). The 2D model suggests that it may occur at 15 km depth, associated with the decreasing on the resistivity.

The MT data and interpretation have limitations and many questions remain yet unsolved. Nevertheless, new and interesting results were obtained in this preliminary analysis. Further research is needed with new and deeper MT soundings, to improve the knowledge on the electrical structure of this part of the Iberian massif.

Acknowledgements

The authors thank A Ribeiro for fruitful discussion on variscian geology. Constructive comments by an anonymous referee are greatfully acknowledged.

References

Chakridi R, Chouteau M, Mareshal M 1992: *Geophys. J. Int.*, 108, 917–929.

Gamble T D, Goubau W M, Miracky R, Clarkes J 1982: *Geophysics*, 47(6), 932–937.

Johansen H K 1977: *Geophys. Prospecting*, 25, 677–691.

Kao D, Orr D 1982: *Geophys. J. R. astr. Soc.*, 70, 323–337.

Kurtz R D, Delaurier J M, Gupta J C 1990: *J. Geophys. Research.*, 95(B7), 10292–10946.

Kurtz R D, Craven J A, Niblet E R, Stevens R A, 1993: *Geophys. J. Int.*, 113, 483–498.

Matias L M 1996: A sismologia experimental na modelao da estrutura da crusta em Portugal Continental. PhD thesis, University of Lisbon (in preparation).

Mendes Victor L A, Miranda J M, Matias L M 1993: In: Iberian Lithosphere Heterogeneity and Anisotropy (ILIHA), Monografia n. 10, J Mercua and E Carreño eds, 179–196.

Miranda J M, Galdeano A, Rossignol J C, Mendes Victor L A 1989: *E. Planet. Sc. Lett.*, 95, 161–172.

Oliveira J T 1990: In: Pre-Mesozoic Geology of Iberia, R D Dallmeyer and E Martinez Garcia eds, Springer-Verlag, Berlin, Heidelberg, 334–347.

Patra H P, Mallick K 1980: Geosounding Principles, 2: Time varying geoelectric soundings. Methods in geochemistry and geophysics, Elsevier, Amsterdam, 419 pp.

Ribeiro A, Silva J B, Dias R, Araújo A, Marques F, Merino H, Fonseca P 1988: In: Proc. of the 5th EGT Workshop. The Iberian Peninsula, E Banda and L A Mendes Victor eds, European Science Foundation, 57–64.

Rijo L 1977: Modelling of electric and electromagnetic data. Ph.D. Thesis, University of Utah

Silva J B, Oliveira J T, Ribeiro A 1990: In: Pre-Mesozoic Geology of Iberia, R D Dallmeyer and E Martinez Garcia eds, Springer-Verlag, Berlin, Heidelberg, 334–347.

Sutarno D, Vozoff K 1991: *Geophysics*, 56(12), 1999–2007.

Torres L, Lisboa M 1988: In: Proc. of the 5th EGT Workshop. The Iberian Peninsula, E Banda and L A Mendes Victor eds, European Science Foundation, 111–121.

Zhang P, Roberts R G, Pedersen L B 1987: *Geophysics*, 52, 267–278.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 231–242 (1996)

FIRST RESULTS FROM FOUR DEEP MT SOUNDINGS ALONG A WEST–EAST PROFILE OF 40 KM LONG AT 35°S LAT. AND 70°W LONG. IN THE ANDEAN CHAIN, ARGENTINA, SHOWING A VERY HIGH CONDUCTIVE CRUST

H G FOURNIER[1], M J MAMANÍ[1], M MUÑOZ[2], A N MAIDANA[1], I ROKITIANSKY[3], E BORZOTTA[1], B CASTIGLIONE[1], J VENENCIA[1]

[Manuscript received May 6, 1994]

Four MT soundings in a profile of 40 km long, namely from West to East: El Desecho, Los Molles, Alfalfalito and Puesto Mendez Viejo, show a crust having a very high conductivity:

1. On the West side, El Desecho and Los Molles, and on the East side, Puesto Mendez Viejo, having an entire column with very great conductivity from near the surface to the Moho, situated at nearly 60 km depth.
2. In the central part, Alfalfalito, having the volcano El Infiernillo, at about 7 km on the NE side, with a resistive column from near the surface until 20 km depth. Hot sulphurous water springs are in the Los Molles area and a basaltic overflow is visible near the sounding of Alfalfalito.

Induction vectors estimated for the sites indicate a current canalization running south of Los Molles valley (see Fig. 1). The field work was made at the end of 1991 and at the beginning of 1996. In the latest case, we have used the new "EMI" equipment.

Keywords: Andes Chain; Argentina; magnetotelluric; magnetovariational

Introduction

In Fig. 1 we see the four soundings' disposition made along an East-West profile of 40 km long near a zone of one active and many extinguished volcanos situated at the west extreme of the profile. The general direction of the four soundings' profile arrives at the same latitude as the Peteroa active volcano (Tormey et al. 1989). The beginning of this volcanic active zone is the Plio-Pleistocene period. The latest eruption date is 9 February 1992.

There is no seismic activity between 35° and 36° South Latitude in this region.

[1]Cricyt–Conae–Geofisica, Casilla de Correo 131, 5500 Mendoza, Argentina
[2]Departamento de Geología y Geofísica, Universidad de Chile, Casilla 2777, Santiago, Chile
[3]POB 3387/7, 252146 Kiev–146, Ukraine

I. Magnetotelluric study

The general trend of the phenomenon in the first stage of this study is given in Fig. 2a to Fig. 2d. The two tensorial resistivity and the two-phase curves are given for each sounding. Strike directions are shown in Fig. 3.

In Fig. 4 we group the four chosen tensorial curves given in Fig. 2. These curves are chosen because they are on the Rho minimum values — these directions of the Rho minimum are generally parallel to the tectonic directions. In this way, we can obtain the interpretation sections using the basic formula (Fournier et al. 1963). Table I shows the list of thicknesses and resistivities for the four soundings, following 1D modelling, used to build Fig. 5, which gives a schematic section of the layers for the four soundings. Figure 5 is logarithmic in depths.

The elements to interpret the section 0–10 km on the West extreme and 0–20 km on the East extreme of the profile into the Earth's crust are: 1. the chemical differentiation; 2. the distance from the seismic silent zone (35° – 36° LS); 3. the Peteroa active volcano at the same latitude; 4. the geochemical change of the eruptive rocks from North and South directions of the studied sites (Hickey et al. 1986, Stern 1987); 5. from 35° S, to North direction, the volcanic arc is situated on the water dividing line (NS) and to the South of this parallel, at the west side of that same line (Muñoz Bravo et al. 1989).

For the section 10–58 km depth to the west extreme and 20–58 km depth to the east extreme: 1. The maximum depth of the conductive layer (58 km) is in

Fig. 1. Position of the sounding sites, between Curicó in Chile and Malargue in Argentina. ED: El Desecho; LM: Los Molles; AL: Alfalfalito; PMV: Puesto Mendez Viejo. Four induction vectors indicate the approximate presence of a conductivity anomaly (CA) running W–NW to E–SE direction, South of Los Molles valley. *: sounding sites for the next campaign of 1996–1998

Table I. Layers composition from the four sounding results

El Desecho 6 layers			Los Molles 6 layers			Alfalfalito 6 layers			Puesto Mendez Viejo 7 layers		
Σ km	km	ohm.m	Σ km	km	ohm.m	Σ km	km	ohm.m	Σ km	km	ohm.m
									0.5	0.5	13.
2.	2	5.4	0.4	0.4	1.6	1.5	1.5	22.	3.0	2.5	5.5
4.	2.	3.9	1.	0.6	0.27	1.65	0.15	0.8	5.5	2.5	13.
6.	2.	3.0	13.	12.	2.	20.	18.	650.	18.	12.5	2.5
57.	51.	2.58	58.	45.	3.3	58.	38.	15.	58.	40.	60.
447	390	2000	443	385	2000	448	390	2000	453	395	2000
∞	∞	1	∞	∞	1	∞	∞	1	∞	∞	1

agreement with the results given by the gravimetric measurements (57 km) for the lower limit of the crust: (Fraga and Introcaso 1988), see Fig. 6. 2. The thickness change of the continental crust from 30–35 km at 37° S to 60 km at 34° S (Lommitz 1962, Lowrie and Hey 1981)). 3. The Sb index of partial melting of mantle peridotite (Onuman and López Escobar 1987).

In the region comprised by this work the volcanos are located on extensional fractures whose strikes are in NNW–SSE direction. The inactive volcano V on Fig. 1, situated at the centre of our profile, (Alfalfalito – AL) could stand on one of these fractures which is visible in South-East direction.

In general we should be very careful with the interpretation of this profile because on one hand we have only four soundings and, in the other hand, we are near the crest of the Andes.

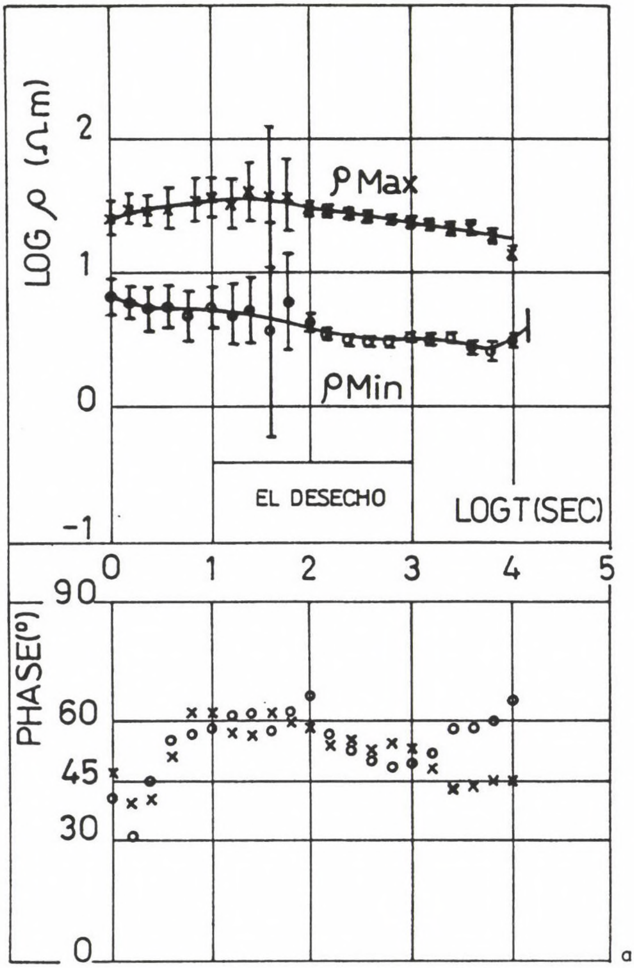


Fig. 2a. Tensorial sounding curves of the four sites: El Desecho. The two tensorial resistivities and the two phase curves are given for each sounding

II. Magnetovariational study

Using the H, D and Z magnetograms corresponding to high periods, obtained from the magnetotelluric sounding work, it was possible to make a magnetic variation analysis. Induction vectors, according to Wiese, were obtained for El Desecho: Fig. 7a, Los Molles: Fig 7b, Alfalfalito: Fig 7c and Puesto Mendez Viejo: Fig. 7d along Los Molles valley.

In order to accomplish the study, the following linear relation was used:

$$Z = aH + bD ,$$

where H, D and Z are the magnetic variation field components in x, y and z directions. a and b are complex coefficients. A manual processing was made using events almost in phase in order to obtain a and b as real numbers.

Dividing the previous expression by D or H and making its plotting, a and b were calculated for the range of periods between $20' - 150'$. Then, the induction

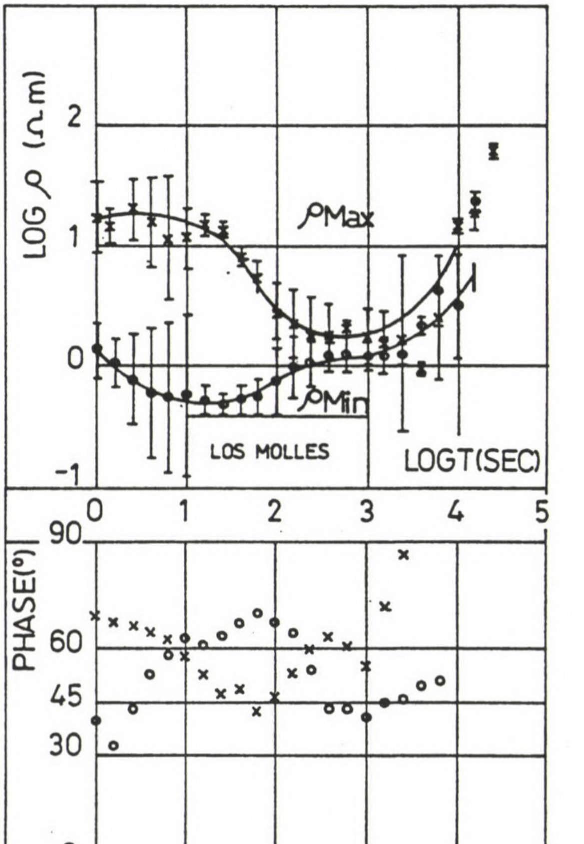


Fig. 2b. Tensorial sounding curves of the four sites: Los Molles. The two tensorial resistivities and the two phase curves are given for each sounding

vectors were obtained as:

$$c = \bar{a}\mathbf{i} + \bar{b}\mathbf{j} .$$

These vectors indicate the presence of current canalization and conductivity anomalies in the ground, which are perpendicular to induction vectors.

The results can be seen in Fig. 1. Four induction vectors indicate the presence of a conductivity anomaly (CA) running W–NW to E–SE direction, south of Los Molles valley. Only an approximated position of the CA is suggested in Fig. 1, because with the information we have now, it is impossible to know with precision the central line of this CA. Therefore, more field data will be necessary to get in the future in order to assure our results. It is interesting to see, as it is shown in Fig. 1, that the probable continuation of the CA just mentioned, in West direction, would arrive near Peteroa active volcano, situated on the border between Chile and Argentina. That detail suggests a possible genetic link between these two tectonic features.

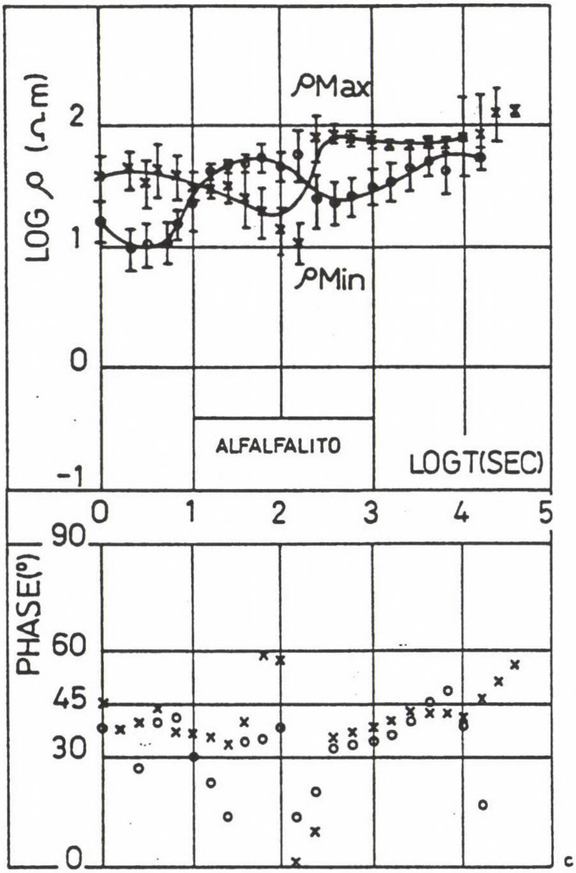


Fig. 2c. Tensorial sounding curves of the four sites: Alfalfalito. The two tensorial resistivities and the two phase curves are given for each sounding

On the other hand, it was possible to compare magnetograms corresponding to Los Molles and Alfalfalito sites with synchronous magnetograms obtained from Pilar Geophysical Observatory. This observatory is located in the central region of Argentina (Córdoba province), far from the Andean range, on an almost cratonic zone, at a distance from Los Molles of about 700 km in NE direction. The comparison was made in the range of periods between 10 min to 2 hours.

As a result, it was visible that the geomagnetic variation field measured in Los Molles and Alfalfalito locations is approximately the same as the Observatory's. This situation suggests that thermal manifestations, visible in Los Molles valley, are essentially local and superficial in character, perhaps derived from Peteroa active volcano, placed 50 km west from Los Molles site, on the border between Chile and Argentina, as mentioned before.

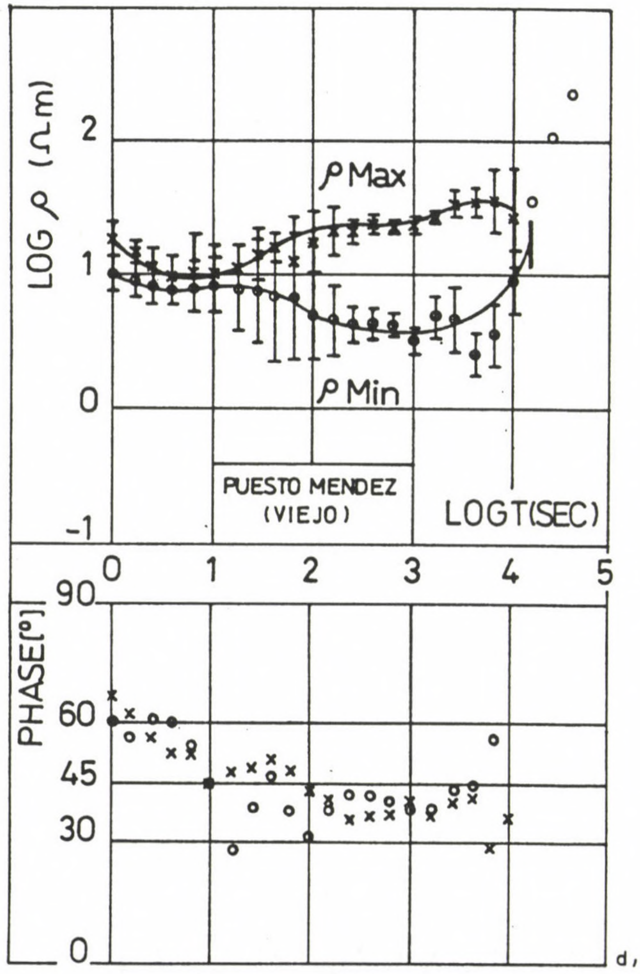


Fig. 2d. Tensorial sounding curves of the four sites: Puesto Mendez Viejo. The two tensorial resistivities and the two phase curves are given for each sounding

Conclusion

In general, we can say that the crust in the region of Alfalfalito site seems to be very distrubed in relation with the surrounding area due to, among other factors, magmatic processes whose superficial evidences we have mentioned before. For the

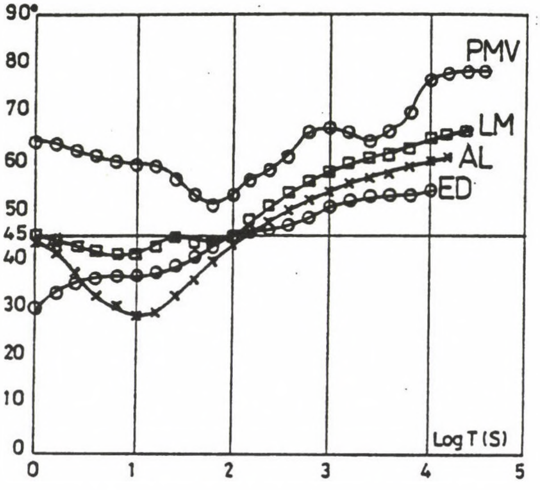


Fig. 3. Strike curves for the four sites: ED: El Desecho; LM: Los Molles; AL: Alfalfalito; PMV: Puesto Mendez Viejo. The mentioned angles are taken from magnetic North clockwise direction. These strikes correspond to given curves shown in Fig. 2

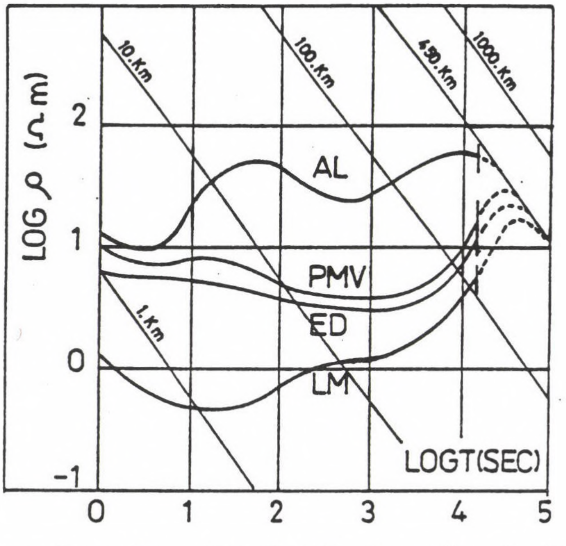


Fig. 4. Chosen curves taken as the good ones for doing the unidimensional intepretation for each sounding: ED: El Desecho; LM: Los Molles; AL: Alfalfalito; PMV: Puesto Mendez Viejo. We have chosen the curves having the value of Rho minimum. These curves are generally parallel to the Andean chain axis

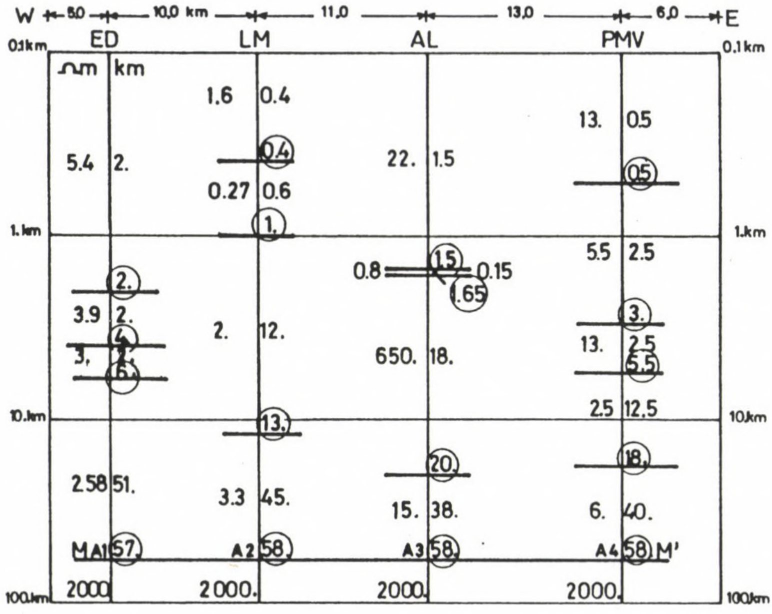


Fig. 5. Logarithmic section of the studied zone. We see the three soundings, El Desecho, Los Molles and Puesto Mendez Viejo having a very conductive nearly complete column; the conductivity increasing from 10000 S to the East side until 15000 S to the West side. It is interesting to mention that because in the extreme West of our profile there is, as we have mentioned before, one active volcano (see Fig. 1)

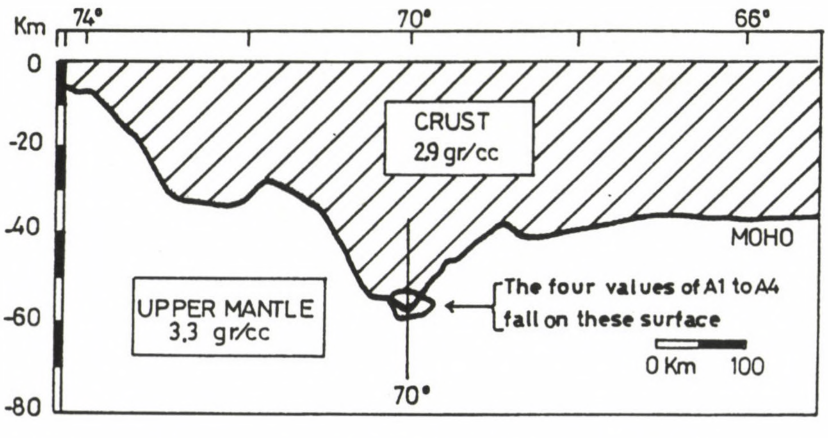


Fig. 6. Gravimetric model for the crust in the region of the parallel 35° S after Fraga and Introcaso (1988). Our results correspond well with this model. The four values of A1, A2, A3 and A4 from Fig. 5 fall on the surface S. Consequently, we may say that M to M' line is the Moho one

three other soundings, the complete sections form near the surface to the Moho are very conductive: from West to East, about 15000 S to 10000 S. The induction vectors estimated for the four sites show a current canalization south of Los Molles

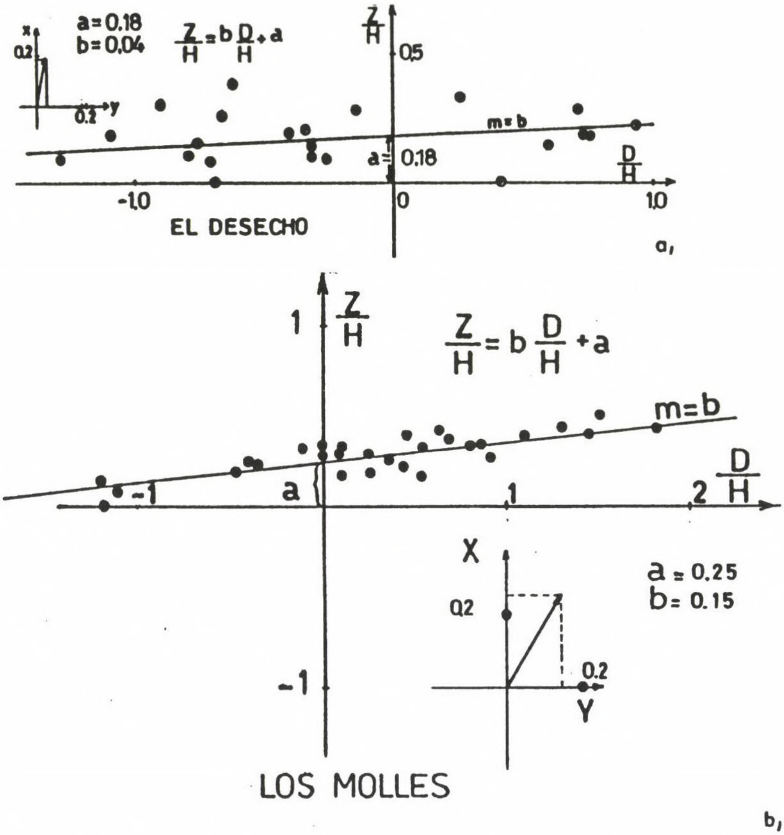


Fig. 7a. Graphical estimation of a and b coeffeicients and Wiese induction vector representation. a) for El Desecho; b) for Los Molles

valley (see Fig. 1). With the information we have now, it is impossible to know the distance between this canalization and Los Molles valley and how strong this canalization is.

These conclusions should be taken only as an approximation to the study of the geological structure of this region of the Andean range. In the future it would be necessary to undertake more field work in this area to obtain more information to be able to confirm or reject these conclusions.

Future work

We need more soundings at each extremity of the profile, see Fig. 1: two at the East of Puesto Mendez Viejo, three between El Desecho and the volcanic border and three at the West of the border with Chile. In this way we could understand much better the profile having 100 km long with 14 MT soundings in the East-West direction. We could do valuable bidimensional interpretation. We are preparing this supplementary research for the period 1996–1998.

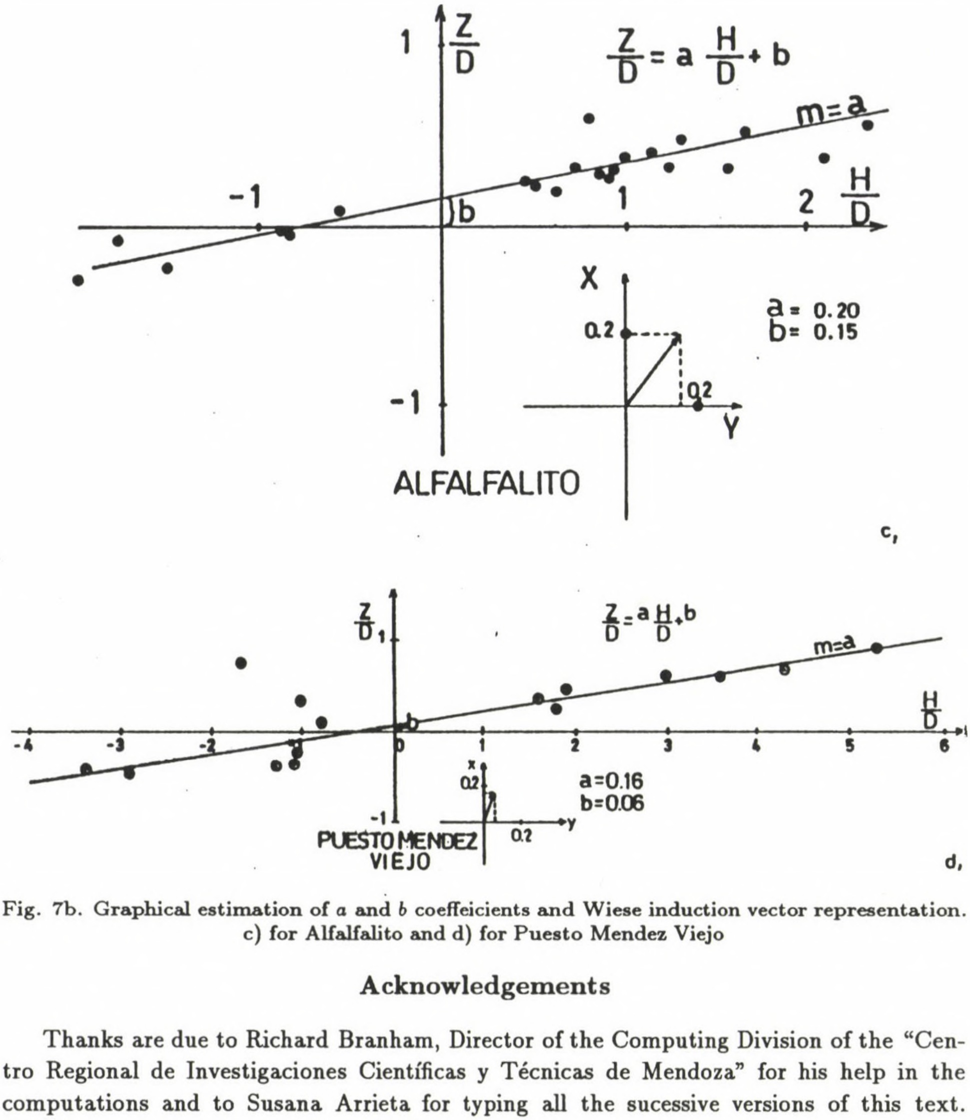


Fig. 7b. Graphical estimation of a and b coeffeicients and Wiese induction vector representation. c) for Alfalfalito and d) for Puesto Mendez Viejo

Acknowledgements

Thanks are due to Richard Branham, Director of the Computing Division of the "Centro Regional de Investigaciones Científicas y Técnicas de Mendoza" for his help in the computations and to Susana Arrieta for typing all the sucessive versions of this text. Thanks are due to Ana María Bonbled for drawing the figures.

References

Fournier H G, Ward S H, Morrison H F 1963: Magnetotelluric evidence for the low velocity layer. Space Sci. Lab., Univ. of Calif., Berkeley, Tech. Rep. 222 (89); 4 (76).

Fraga H, Introcaso A 1988: Modelo gravimétrico cortical en la latitud 35° S del sector argentino-chileno. V Congreso Geológico Chileno, Depart. de Geolog. y Geofis., Univ de Chile

Hickey R L, Frey F A, Gerlack D C, López Escobar L 1986: *J. Geoph. Res.*, 91, 5963–5987.

Lommitz C 1962: *J. Geoph. Res.*, 67, 351–363.
Lowrie A, Hey R 1981: *Geol. Soc. Am. Mem.*, 154, 741–754.
Muñoz Bravo J, Stern C R, Bermudez A, Delpino D, Dobbs M F, Frey F A 1989: *Asoc. Geol. Argent.*, 44, 270–286.
Onuman N, López Escobar L 1987: *J. Volcanol. Geotherm. Res.*, 33, 283–298.
Stern C R 1987: In: Int. Symp. Andean Volcanism, Tucumán, Argentina, Abstracts, Actas 4, 202–203.
Tormey D R, Frey F A, Escobar L L 1989: *Asoc. Geol. Argent. Rev.*, 44, 420–430.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 243–255 (1996)

PLASMASPHERIC PARAMETERS DERIVED FROM WHISTLER OBSERVATIONS AND THEIR RELATION TO IONOSPHERIC ABSORPTION AND GEOMAGNETIC PULSATIONS

F MÄRCZ[1], P BENCZE[1], J VERŐ[1]

[Manuscript received March 18, 1996]

Plasmaspheric electron densities derived from ground-based whistler observations (at Tihany, Hungary) have been analysed for showing their changes with L and geomagnetic activity. Ionospheric absorption and the f_0F2 parameter simultaneously determined at two high mid-latitude stations (Kühlungsborn and Juliusruh, Germany) are studied together with the plasmaspheric parameter during a selected interval including a geomagnetic disturbance. Results show that both the absorption and the f_0F2 parameter increase after geomagnetic storms and these events can be coupled with changes in the plasmasphere. Satellite observations during the same interval confirm these findings. Geomagnetic pulsation activity (at the Nagycenk station, Hungary) is rather attenuated if the $f_0\mathrm{F2_{max}}$ values increase to a limit of about 10 MHz, as detected earlier for winter months of the corresponding hemisphere around solar activity maximum. Additionally, a quite peculiar short-term attenuation was revealed in the springtime pulsation activity, too.

Keywords: attenuation of geomagnetic pulsations; geomagnetic activity; ionospheric absorption; ionospheric f_0F2 parameter; plasmaspheric electron density

1. Introduction

The character of the plasmasphere being the inner region of the magnetosphere is somewhat similar to that of the neighbouring ionosphere. The ionized particle populations of the two regions are quite closely coupled with each other. Consequently, ionospheric phenomena should be influenced by changes in the plasmaspheric plasma. Electron densities in the plasmasphere can be determined by direct measurements using satellite-borne equipments and they can also be studied by means of a ground-based technique i.e. by analysing the characteristics of whistlers. Whistlers are VLF waves generated by lightning flashes and they propagate along geomagnetic field lines between the hemispheres. The whistler-based method has the advantage that electron densities can more frequently be determined at an appropriately chosen site and this allows to study changes with time in a certain area of the magnetosphere.

At the Tihany Observatory (Hungary, geom. latitude: 43°N) whistler observations were quite continuously carried out in the seventies and the data set was used by the research group of the Geophysical Department of the Lorand Eötvös

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

University (Budapest) for determining some parameters of the plasmasphere. They derived the electron density for the equatorial plane, the electron density at a height of 1000 km and the tube content. As whistlers travel along rather different paths, the deduced plasmaspheric parameters belong to different L-values. Thus we have information for various parts of the plasmasphere and in some cases we can follow the changes of the mentioned parameters with time, too.

Most of the successful ground-based methods applied for the study of the ionosphere are based on the propagation of radio waves. One of these methods (the so-called A3 method) is aimed at the determination of the attenuation of radio waves obliquely reflected from the ionosphere. The determined parameter, the so-called ionospheric absorption depends on the electron density of that part of the ionosphere where the radio waves are reflected from. It is well known that the absorption deduced for the mid-latitude lower ionosphere shows a peculiar change after certain geomagnetic storms. Ionospheric absorption increases during the main phase of the storm and in addition it remains enhanced well after the main phase for several days, sometimes even for a fortnight. The enhanced absorption is due to the increased ionization in the lower ionosphere which is caused by magnetospheric particles precipitating there from the radiation belts due to wave-particle interaction (Lauter and Knuth 1967). Thus, the mentioned ionospheric event is connected with magnetospheric processes, moreover changes in parameters of the interplanetary space might also play a role in the appearance of the event as shown by Mrcz (1988, 1991, 1992). Co-ordinated ground-based and satellite measurements (Larsen et al. 1976) confirmed the relation between increased ionization in the ionosphere and particle precipitation from the plasmasphere associated with geomagnetic storms (which could finally lead to the enhancement of ionospheric absorption).

As that kind of satellite measurements which timely corresponds to a certain absorption event is not so often available, it seemed to be worth using plasmaspheric parameters derived from whistler observations in a comparative study, i.e. to analyse the changes of the plasmaspheric electron density together with those in ionospheric absorption at times of geomagnetic disturbances. In addition to the latter object, a simultaneous analysis of geomagnetic pulsation activity and a parameter of the upper ionosphere (f_0F2) will also be discussed in the present study. Namely, it was shown that a significant minimum appears in the annual variation of pulsation activity in winter of the corresponding hemisphere around the solar maximum years (Verő and Menk 1986, Zieger 1991). The minimum coincides with maxima occurring both in the equatorial particle density derived from whistler data (Verő 1965) and in the electron density of the upper ionosphere as shown by f_0F2 (Verő and Menk 1986). The latter authors found that the attenuation of pulsations is connected with a limit for f_0F2 (being about 10–11 MHz). The present study intends to test whether this limit might be confirmed on the basis of further data.

2. Results

2.1 Equatorial electron densities derived from whistler data

Electron density values determined for the equatorial plane on the basis of Tihany whistler data have been available for the interval between 1970 and 1975. Samples were drawn from this set according to an aspect which made it possible to study the changes of electron density within a rather wide L-zone and in dependence of geomagnetic activity. Actually, 163 electron density values have been selected out of the whole set and they covered the zone between $L = 1.3$ and 3.7. This sample has been divided into two groups by roughly distinguishing the high and low geomagnetic activities on the basis of the daily ΣK_p-value. The group for high geomagnetic activity ($\Sigma K_p > 25$) consisted of 76 data and the remaining 87 cases formed the group for low activity ($\Sigma K_p < 25$).

The individual electron density values are plotted versus L in Figs 1a (for $L = 1.3 - 2.5$) and 1b (for $L = 2.5 - 3.7$) where they are denoted by different symbols according to geomagnetic activity. In spite of the spread of data, the decrease of the equatorial electron density with increasing L is clearly shown in both figures. Nevertheless, a relation between the electron density and geomagnetic activity is not striking. For showing up a general change in the electron density according to L and considering also geomagnetic activity, we have grouped the electron densities step by step (using a distance of 0.1 in L) and have determined the median value

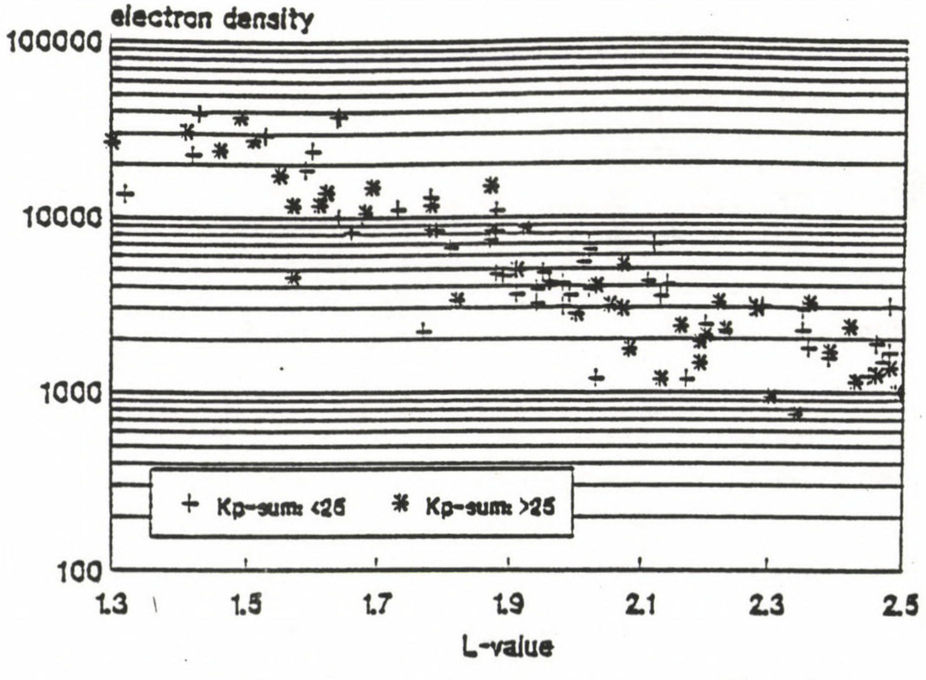


Fig. 1a. Equatorial electron densities derived from whistler observations versus low L-values for days of different geomagnetic activity

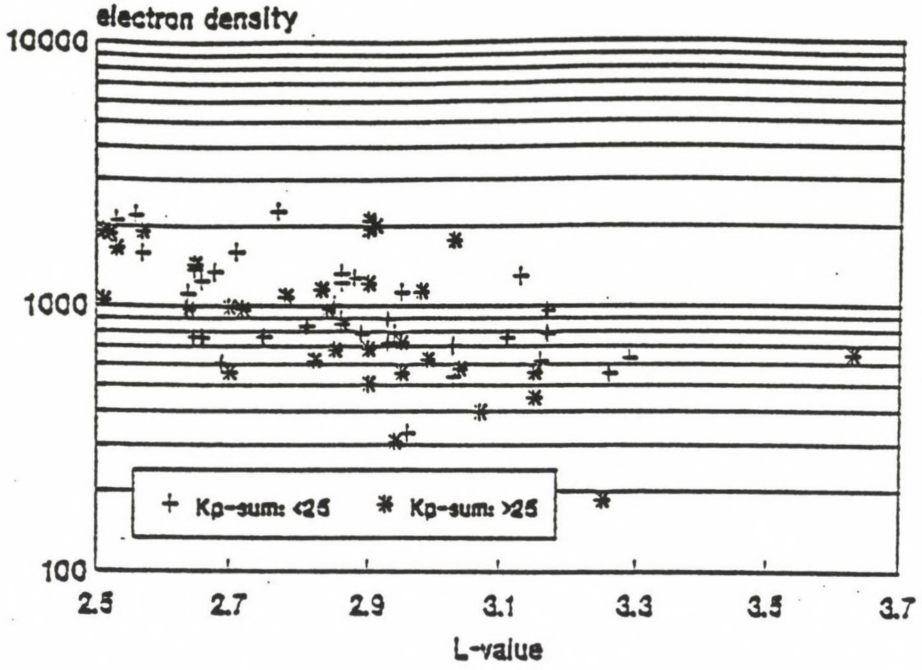


Fig. 1b. The same as in Fig. 1a, but for higher L-values

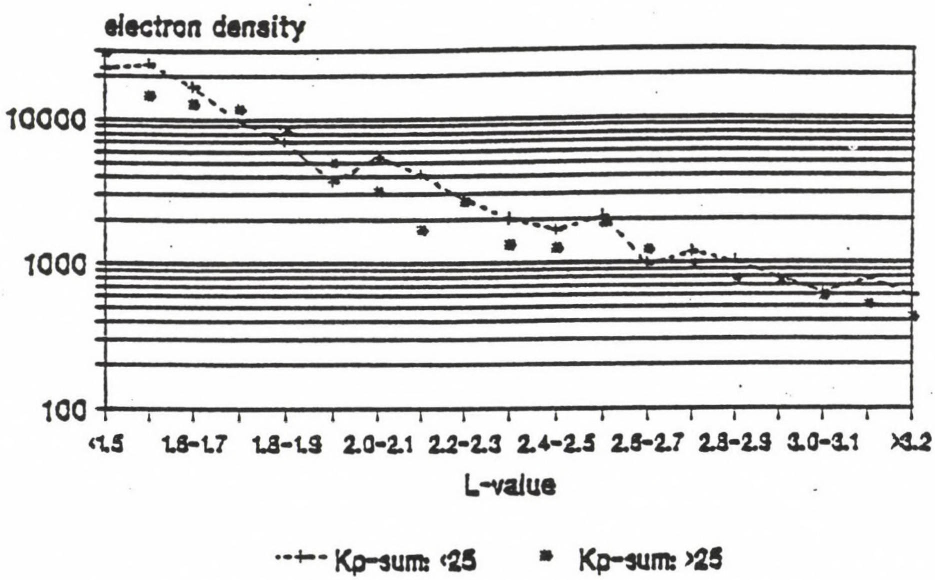


Fig. 2. Medians of electron densities versus L and depending on geomagnetic activity

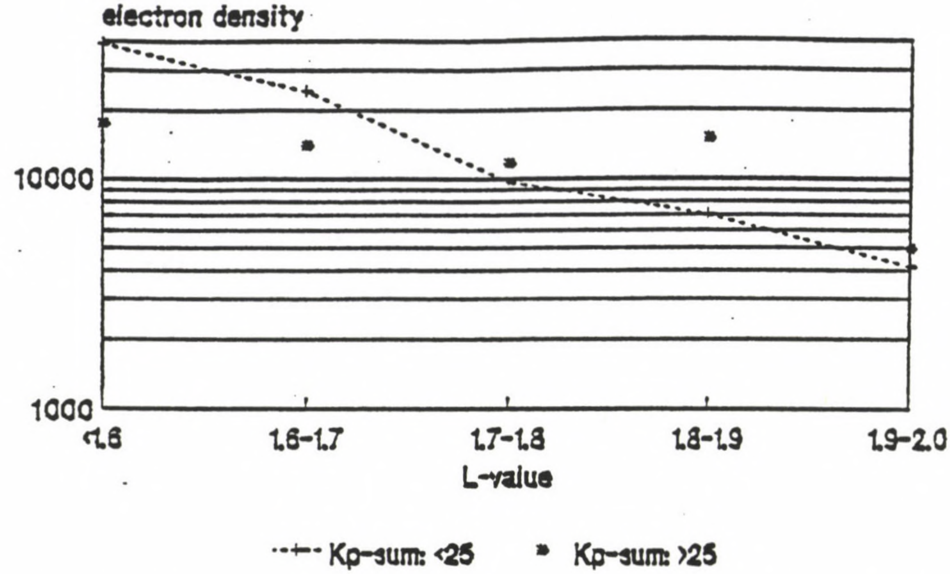


Fig. 3. The same as in Fig. 2, but determined on the basis of whistlers of high quality for a narrow L-zone

of each group both for high and low geomagnetic activities. (Electron densities derived for L-values below 1.5 and above 3.2 were quite rare therefore these cases have been united into two extreme groups.) The medians are plotted versus L in Fig. 2 where they are distinguished according to geomagnetic activity. (In this figure, the dotted line linking the values given for low activity — $\Sigma K_p < 25$ — merely helps in distinguishing the medians belonging to different activity levels.) The medians for high activity are generally less than those for low activity above $L = 2.0$, except the value in the zone $L = 2.6 - 2.7$. Nevertheless, the electron densities are higher for high geomagnetic activity as compared with those for low activity in the zone between $L = 1.7$ and 2.0. We could confirm the reality of this behaviour based on a set of selected data as shown in the following.

The colleagues at the Geophysical Department of the Lorand Eötvös University qualified the whistlers used for determining the electron densities. As the spread of the individual electron density values depends on the quality of whistler observations, we discarded values which were derived from whistlers of low quality. Thus for the zone from $L < 1.6$ to $L = 2.0$, we calculated new medians from the remainig values and these are shown in Fig. 3. At L-values below 1.7, the electron densities determined for high geomagnetic activity are less than those for low activity, however, the opposite is true for the zone between $L = 1.7$ and 2.0 i. e. the finding in Fig. 2 seems to be a real one.

2.2 Plasmaspheric and ionospheric parameters associated with geomagnetic disturbances

As described in the introductory part, ionospheric absorption may inrease after certain geomagnetic disturbances. The preceding section has shown that the equatorial electron densities in the plasmasphere and their changes with L might also depend on geomagnetic activity. Consequently, it seemed to be worth analysing both parameters associated with geomagnetic disturbances. In addition, a parameter characterizing the upper ionosphere (the daily $f_0\mathrm{F2}_{\mathrm{max}}$) should also be included in the analysis. The electron density data derived on the basis of whistler observations are sporadic, they are rarely available for consecutive days. This circumstance made more difficult the simultaneous investigation of the ionospheric parameters and plasmaspheric electron densities for a long period. Nevertheless, we could find some cases for which these parameters were at our disposal during a quite appropriate interval, even if gaps also appeared (mainly in the plasmaspheric data). One of these cases is treated below.

In the second half of December 1971, there was a rather strong geomagnetic storm as indicated by the daily ΣK_p-values in Fig. 4. Following a quiet period, the geomagnetic activity suddenly increased on December 17 ($\Sigma K_p = 42$) and it decreased to the pre-storm level after 3 days. This storm was followed by a smaller one on December 22. The simultaneous changes in ionospheric absorption are also shown in Fig. 4. They are given as departures from the corresponding monthly median expressed in percents (Lss%). The absorption in the lower ionosphere was determined at a middle latitude station (Kühlungsborn, geogr. latitude of the reflection point: $\sim 55°$N) by the A3 method (at 245 kHz) for sunset. For showing the changes in the upper ionosphere, daily values of the $f_0\mathrm{F2}_{\mathrm{max}}$ parameter determined at the Juliusruh station (geogr. latitude: 54.6°N) are used. Departures from the monthly median expressed in percents are shown in this case, too.

The equatorial electron densities derived from whistler observations at the Tihany observatory have been available for three consecutive days (on December 18, 19 and 20) during the investigated interval. As shown in Fig. 5, the electron density data cover a rather wide zone (between $L = 1.3$ and 3.1) on the day (Dec. 18) immediately after the onset day of the geomagnetic storm which is characterized by a quite high activity ($\Sigma K_p = 27$). For the two following days, several electron density data are plotted at L-values above 1.9; on these days the geomagnetic activity was lower, especially on December 20. Figure 5 shows that the plasmaspheric electron densities are usually higher for a wide range of L on the day characterized by high geomagnetic activity as compared with those derived for the two days of lower activity. The present results are peculiar since the situation was opposite for $L > 2$ in Fig. 2, where the electron densities were generally lower for high geomagnetic activity than for the lower one. Thus for certain geomagnetic storms, it can be expected that the increased electron density extends to a larger area of the plasmasphere and not only to a narrow zone immediately below $L = 2.0$ as could be suggested on the basis of Figs 2 or 3. For an additional comparison, Fig. 5 also shows the electron densities derived for two L-values on a quiet day (Dec. 9: $\Sigma K_p = 15$) well

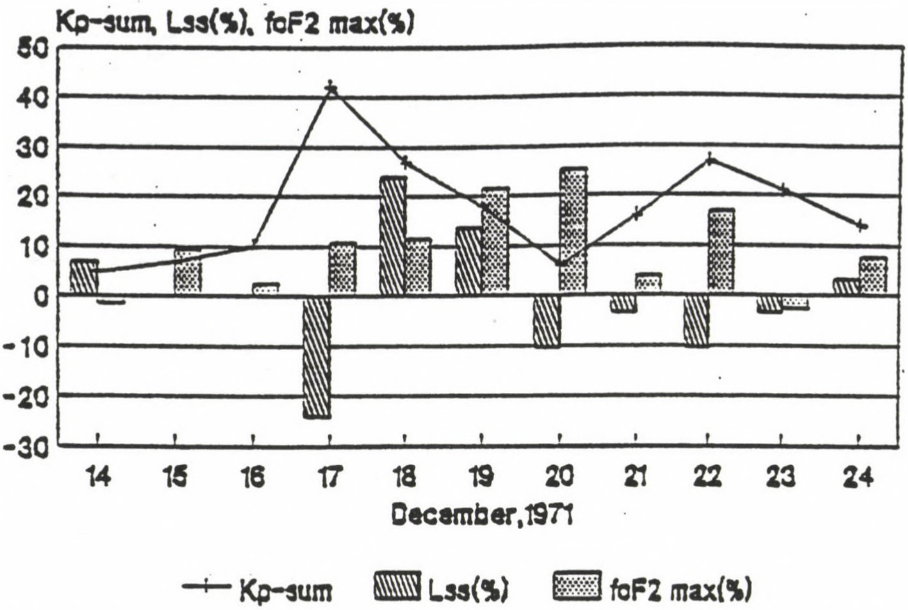


Fig. 4. Changes in geomagnetic activity (ΣK_p), ionospheric absorption (Kühlungsborn, Lss%) and the f_0F2_{max} parameter (Juliusruh) around a geomagnetic storm in December, 1971

before the onset of the geomagnetic storm. (The line linking the two values hints at the trend of the change with L.) The corresponding electron densities for the investigated high activity day (Dec. 18) are also higher than these reference values.

As shown in Fig. 4, ionospheric absorption was well below the normal on the onset day of the geomagnetic storm, however, it suddenly increased on the following day for which an electron density enhancement has been revealed within a wide area of the plasmasphere (see Fig. 5). This hints at the fact that the flux of energetic electrons in the outer radiation belt was abundant due to acceleration processes and radial diffusion, thus the conditions for wave-particle interaction were favourable. Consequently, the precipitation of particles from the plasmasphere into the ionosphere could indeed cause an enhanced ionization there which finally led to the increase of the ionospheric absorption on December 18. Absorption remains high on the next day, too, in spite of the low electron densities shown for a wide range of L in Fig. 5. On this basis it can be concluded that the intense particle precipitation did only cease on December 20 when the ionospheric absorption is somewhat below the normal. The following smaller geomagnetic disturbance is not accompanied by an enhancement of the ionospheric absorption, i.e. it can be supposed that the plasmaspheric conditions should be unfavourable for particle precipitation because of the decay of particles in the radiation belt during the former geomagnetic disturbance, or the spectrum of precipitating particles changed during and after this storm. If the latter supposition is true it would mean that particles with higher energies (e.g. electrons of > 40 keV which could penetrate deep into the lower ionosphere after the previous storm and caused the enhanced absorption)

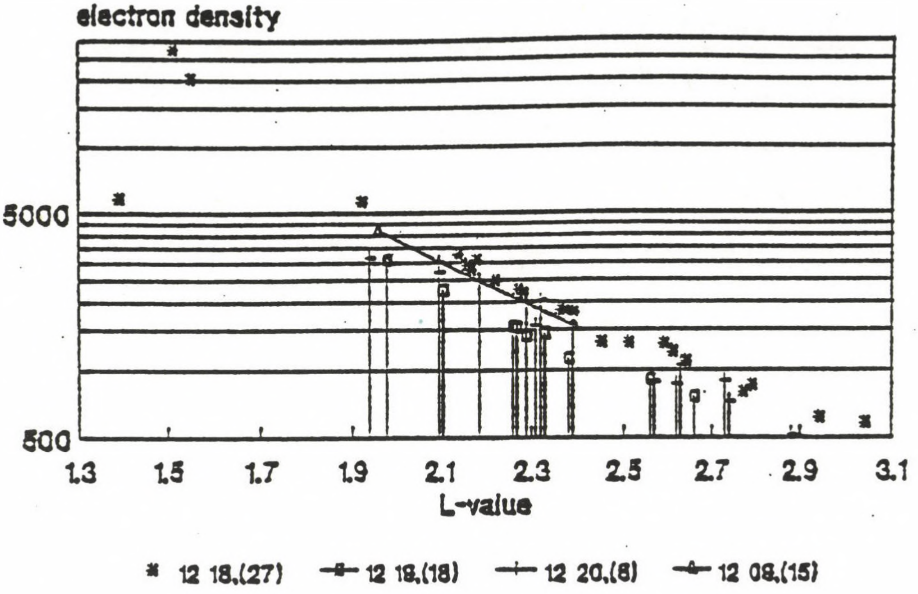


Fig. 5. Equatorial electron densities versus L for days following and preceding the geomagnetic storm in December, 1971

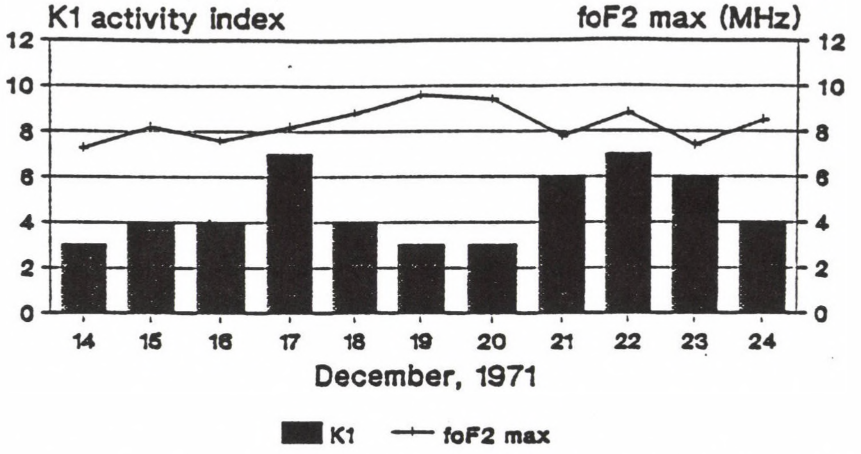


Fig. 6. Changes in pulsation activity (as shown by the K1-index determined at Nagycenk) and in the f_0F2_{max} parameter (Juliusruh) in winter 1971

are missing from the spectrum and those with lower energies were only preserved. These particles can merely reach the higher ionospheric levels and not the lower ionosphere where the attenuation of the used radio wave (245 kHz) takes place, thus no absorption enhancement can be observed. A parameter characterizing the upper ionosphere (f_0F2_{max}) is also given in Fig. 4 and its changes suggest that the previous considerations might be acceptable; namely the f_0F2_{max} is high from the

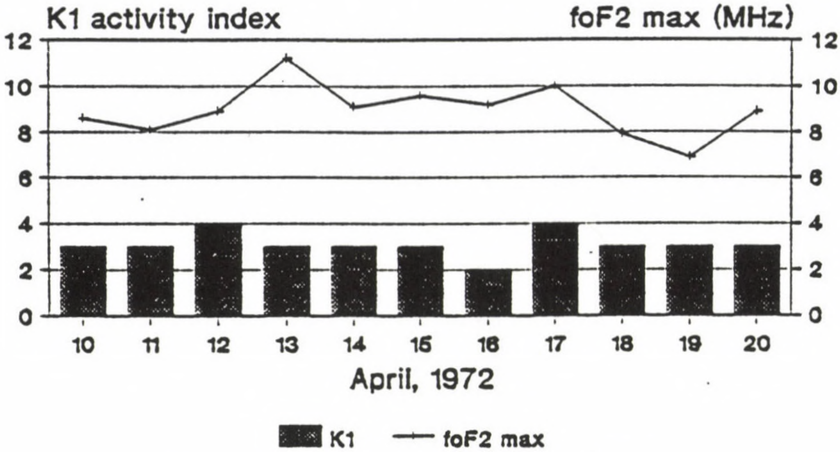


Fig. 7. The same as in Fig. 6, but for the spring 1972

beginning of the first storm and it remains enhanced till the second geomagnetic disturbance. Nevertheless, it is to be noted that the electron density in the F-region also depends on composition changes and the increase of the electron density between the two geomagnetic disturbances and during the second one might be due to thermospheric winds altering the composition by lifting the F-region.

2.3 Damping of geomagnetic pulsations associated with changes in $f_0F2_{\max}$

It was mentioned in the introduction that a minimum appears in the annual variation of pulsation activity which occurs in winter of the corresponding hemisphere around solar maximum years and this might be related to a limit for the corresponding f_0F2 value, however, the limit might only be estimated (being about 10–11 MHz, according to Verő and Menk 1986). As the event studied in the previous section occurred in a northern winter month, in December 1971 (and this year was in the descending phase of solar cycle 20, however, near to the preceding maximum), we extended our analysis for finding a relation between the used $f_0\mathrm{F2}_{\max}$ parameter and an index characterizing the activity of geomagnetic pulsations.

Geomagnetic pulsations are regularly observed at the Nagycenk Observatory (Hungary, geom. latitude: 47.2° N), and the data processing yields various indices for characterizing the activity of different kind of pulsations. We have used the K1-index characterizing the pulsation activity in the 0–2 min period range for our investigation. Figure 6 shows the values of the K1-index together with those of the $f_0\mathrm{F2}_{\max}$ parameter for the same interval (between December 14 and 24, 1971) investigated previously. The $f_0\mathrm{F2}_{\max}$ approached to the 10 MHz value on December 19 and 20. It is striking that the pulsation activity (as shown by the K1-index) drops to a minimum on the same days (there is no other similarly low value during the interval studied, except that on December 14). The ionospheric parameter decreased on the following day, simultaneously the K1-index began to increase and

Table I. f_0F2 maximum hourly medians (around noon, in MHz) at Juliusruh and averaged pulsation activity indices (K1) at Nagycenk for winter and spring months in 3 consecutive years

Month	January		April	
Year	f_0F2	K1	f_0F2	K1
1971	9.2	4.26	7.5	5.00
1972	7.8	4.19	8.2	3.77
1973	7.3	5.84	6.0	6.00

it remained at a rather high level for some days. The results of Fig. 6 yield a further confirmation for the earlier findings, namely they convincigly show for the present case that the attenuation in pulsation activity is really associated with the level of the f_0F2 parameter representing the electron density in the upper ionosphere.

An additional case is presented in Fig. 7 which occurred some months later (in April 1972) than the previous one. The same parameters ($f_0\mathrm{F2}_{\mathrm{max}}$ and K1-index) are shown for the interval between April 10 and 20. Pulsation activity is on a rather low level during the whole period. The $f_0\mathrm{F2}_{\mathrm{max}}$ surpasses the 10 MHz value on April 13 and remains near to this level till April 17. As experienced earlier, the attenuation of pulsations most likely occurs in winter months of the corresponding hemisphere around solar activity maximum. Notwithstanding, the presented low pulsation activity appeared in spring and this is somewhat peculiar. There is, however, another peculiarity which might explain this "spring damping" of pulsations: the f_0F2 parameter (around noon) was unusually high in this spring month. During the period investigated here, it was even higher than the maximum hourly median for a winter month of the same year (January, 1972) or in January of the following year (1973). For comparison, Table I shows the highest hourly medians for f_0F2 (appearing around noon) in three consecutive years (1971, 1972 and 1973) both for January and April as well as the corresponding monthly averages of the K1-indices. It is evident from Table I that the $f_0\mathrm{F2}_{\mathrm{max}}$ values shown for the investigated interval in Fig. 7 are indeed extremely high for this season, and the K1 average for April 1972 is also the lowest in comparison with those determined for April of the neighbouring years. As the low level of pulsation activity coincides with these extreme $f_0\mathrm{F2}_{\mathrm{max}}$ values, we suggest that an attenuation of pulsations can take place in an unusual (spring) season, too, if the conditions in the upper ionosphere and plasmasphere are similar to those being usual in winter of the actual hemisphere. It is to be admitted that not the high f_0F2 values themselves (which are equivalent to the high electron densities in the upper ionosphere) might directly be responsible for the attenuation, they should only be regarded as indicators of coupled magnetospheric processes leading to the low pulsation activity. Namely, it is to be noted that the electron density in the plasmasphere depends on the electron density in the F-region because of the striving of the ionosphere/plasmasphere system after equilibrium by diffusion. Additionally, a great part of the pulsations are due to L shell resonance. This means that not only their period but also their amplitude is affected by the electron density in the plasmasphere. Increasing electron density can increase the period and above a certain limit it can also reduce the amplitude (like a damped oscillator).

Equatorial electron densities derived from whistler observations are sporadic for the interval in Fig. 7. For the day with the highest f_0F2_{max} (on April 13), an electron density of about $7000/cm^3$ could, however, be determined at $L \sim 1.9$ (i.e. in the inner part of the plasmasphere) and the geomagnetic activity was moderate ($\Sigma K_p = 20$). This electron density is twice higher than that found on December 19, 1971 for almost the same L-value and geomagnetic activity, as can be seen in Fig. 5. Thus, in addition to the increased electron density in the upper ionosphere (indicated by f_0F2_{max}) we also have an indication for the highly increased plasmaspheric electron density on April 13, 1972 when simultaneously the pulsation activity is rather damped (by a factor of about two). It is not inconsistent to consider that this extreme conditions in the plasmasphere might contribute to the attenuation of geomagnetic pulsations even in spring, albeit it is more regular in winter.

3. Discussion and conclusions

The analysis of electron densities derived for the equatorial plane from selected whistler observations has proven a decrease of this parameter with increasing L for a quite wide zone of the plasmasphere, as could be expected. Nevertheless, the presented results revealed a dependence of this change on geomagnetic activity and a peculiarity of the change within the inner plasmasphere (between $L = 1.7$ and 2.0). In this narrow zone, the electron densities are higher for high geomagnetic activity than for the lower one, whilst the opposite is true in the investigated other zones of the plasmasphere. Additionally, we found certain geomagnetic storms for which this behaviour does not hold. Data for a selected case (Fig. 5) show that the electron densities can also be enhanced outside of the mentioned narrow zone outwards to about $L = 3$, after a strong geomagnetic storm. This storm was followed by an increase of ionospheric absorption, i.e. in accordance with accepted interpretations (e.g. Lauter and Knuth 1967) it is suggested that electrons did precipitate from the plasmasphere into the ionosphere due to wave-particle interaction and they led to the enhancement of ionization in the lower ionosphere causing the increase of absorption. For the present case, the appropriate abundance of electrons in the plasmasphere was detected on the basis of a ground-based method using whistler observations.

The above mentioned conditions are confirmed by the results of Larsen et al. (1976). They carried out ground-based measurements with coordinated satellite observations of quasi-trapped and precipitating electrons over Ottawa (45° N, 76° W). By means of the ground-based partial reflection measurements, electron concentration profiles for D-region altitudes (between $\sim$ 60 and 90 km) have been derived. For the period after the December 17, 1971, geomagnetic storm (which has been analysed in the present study, too), Larsen et al. (1976) reported significant fluxes of precipitating electrons of $>$ 130 keV near Ottawa and a simultaneous excess ionization in the D-region determined by ground-based technique. They proved that the D-region poststorm conditions (at middle latitudes above Ottawa) were caused by precipitating electrons from the radiation belt, and they argued that the quasi-trapped electron fluxes might be 3–4 orders of magnitude higher over Ottawa than those over other mid-latitude stations in central Europe ($\sim$ 10°E). In comparison with the intense electron precipitation over Ottawa, however, a less significant electron precipitation should be expected over stations in Germany, as stated by Larsen et al. (1976). Our results given in Figs 4 and 5 clearly confirm this expectation. As shown by the increased ionospheric absorption (Fig. 4), the ionization in the D-region over Kühlungsborn (Germany) was really enhanced, but only for two days after the onset of the geomagnetic storm of December 17, 1971, and an increased electron density was also found for a wide zone of the plasmasphere on December 18 (Fig. 5). Consequently, our results allow to conclude that plasmaspheric parameters determined on the basis of whistler observations yield a useful tool for interpreting certain plasmaspheric and ionospheric relations, especially in the case of missing satellite data.

Results given in Figs 6 and 7 hint at a rather strong attenuation of geomagnetic pulsations associated with high electron conentration in the upper ionosphere (as shown by $f_0\mathrm{F2}_{\mathrm{max}}$). In the first case (Fig. 6), the attenuation takes place in winter corresponding to earlier findings (Verő 1965, Verő and Menk 1986). In this case of strong damping, the limit to which the $f_0\mathrm{F2}_{\mathrm{max}}$ approached was around 10 MHz; this value does not contradict the limit estimated in a recent study (Verő et al. 1995) on the basis of data for another period (winter 1991), even if the latter limit was somewhat higher. Figure 7 showing another case of strong damping yields a quite surprising result as the attenuation of pulsation activity appeared in spring (April, 1972). Moreover, it lasted rather long and was accompanied by $f_0\mathrm{F2}$ values which are unusual for this season as can be seen from a comparison in Table I. Based on samples of electron densities derived from whistler data for $L \sim 1.9$, the plasmaspheric conditions also departed from the normal in the investigated period. The unusual $f_0\mathrm{F2}_{\mathrm{max}}$ values well indicate these extreme conditions in the plasmasphere, i.e. the increased particle density. As suggested previously (Verő et al. 1995), the source of damping of pulsations should be outside of the upper ionosphere; the attenuation is most likely associated with the enhanced particle density in the equatorial zone of the plasmasphere at around $L = 2$. The present results confirming this suggestion detected a new aspect: it seems that a short-term strong attenuation of pulsation activity might occasionally appear in another season, too (not exclusively in winter), if the plasmaspheric conditions are winterlike, i.e. highly favourable for damping, namely if the field line resonance is attenuated by the increased electron density in the plasmasphere.

Acknowledgements

The authors thank the colleagues of the Geophysical Department of the Lorand Eötvös University (Budapest) for the supply of data derived from whistler observations. The investigations were supported by the Hungarian State Grant OTKA No. T-3091 and 1171.

References

Larsen T R, Reagen J B, Imhof J S 1976: *J. Geoph. Res.*, 81, 2200–2212.
Lauter E A, Knuth R 1967: *J. Atmos. Terr. Phys.*, 29, 411–417.
Märcz F 1988: *Ann. Geophysicae*, 6, 195–200.
Märcz F 1991: *J. Atmos. Terr. Phys.*, 53, 319–323.
Märcz F 1992: *Planet. Space Sci.*, 40, 979–983.
Verő J 1965: *J. Geoph. Res.*, 70, 2254–2257.
Verő J, Menk F W 1986: *J. Atmos. Terr. Phys.*, 48, 231–245.
Verő J, Best I, Vellante M, Lühr H, de Lauretis M, Holló L, Märcz F, Strestik J 1995: *Ann. Geophysicae*, 13, 689–697.
Zieger B 1991: *J. Geoph. Res.*, 96, 21115–21123.

Acta Geod. Geoph. Hung., Vol. 31(1-2), pp. 257-263 (1996)

ARE SHELLS OF FIELD LINE RESONANCE OF FINITE THICKNESS OR INFINITELY THIN?

J VERŐ[1]

[Manuscript received July 19, 1996]

The problem of the thickness of shells of geomagnetic field lines, along which resonance may take place, is discussed. These shells play an important role in the excitation of geomagnetic pulsations (Pc3-4). It is found that shells have a finite thickness of about one or a few hundred km.

Keywords: field line resonance; geomagnetic pulsations; mathematical modelling

Field line resonance (FLR) is now a well established source for a significant part of the geomagnetic pulsations Pc3-4 (Chen and Hasegawa 1974, Southwood 1974). The main characteristics of pulsations with an FLR origin is that their period changes with geomagnetic latitude or L-value (Orr 1978, Cz Miletits 1980, Cz Miletits et al. 1990), while the change with time is much less (Verő 1996). The rate of the change vs. latitude was found in average to be about 10 percent for one geomagnetic degree. Nevertheless, the change is often not smooth, the period is constant for a certain range of latitudes, followed by a very rapid change (Cz Miletits 1980). That is why the idea of "shells" has emerged. Now, "shells" may have two different meanings: according to the first concept, resonances along different field lines are coupled, therefore shells pulsating with a certain (constant) period have a finite thickness (Kurchasov et al. 1987, Verő and Cz Miletits 1994), be it expressed in L-value, be it in surface distance within which the period remains constant. In the other concept shells are not coupled, the period changes smoothly with L-value or latitude, and finite thickness "shells" do not exist, each shell of a certain period is infinitely thin. As shells of field lines play a central role in the following discussions, let us consider the effect of infinitely thin "shells" on the pulsations which result from the field-line resonance. (In such a case, the "thickness" or "surface width" of the shells may be defined as a distance where the amplitude of the signals coming from the actual shell drops to a certain fracture, e.g. to 0.7 times the maximum value). The resulting signal is then the sum of signals coming from an infinite number of shells. It can be approximated at a given location by the sum of many shell resonances, each of a slightly different period and of an amplitude being determined by the distance from the actual location. We use in the following model computation the sum of an increasing number of components; in the first approximation, there is a central period, corresponding to the field line resonance of the station, the next two components — with amplitudes less by a factor of

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

2 — come from the distance corresponding to the shell width where the periods differ by 10 percent from the central period; this distance was supposed to be 100 km. These values correspond to the experimentally observed rate of change of the period at the latitude of the Nagycenk observatory (Verő and Cz Miletits 1994) for which the central period, 20 s, was also used. The next pair of components are at a twice greater distance, with periods and amplitudes computed corresponding to the previous rule, and so on until the amplitude drops below 1 percent of the initial (unit) value. In the second approximation, all distances are halved, else the process is the same. In the third step, distances are again halved and so on. The amplitude in the time moment 0 is made equal 1 in each approximation. (If the rate of the period change vs. latitude is different, the resulting dynamic spectra only change in scale, but not in form). This means that the following F_m values are to be computed:

$$F_m = \frac{F_t}{F_0} = \frac{\sum\limits_{n=-7.2^m}^{+7.2^m} e^{-\ln 2(n_0 n)^2} \cos\left[\frac{\pi}{10}(1+0.1n_0)^n\right] t}{\sum\limits_{n=-7.2^m}^{+7.2^m} e^{-\ln 2(n_0 n)^2}}$$

where F_m is the computed signal, F_t and F_0 are the values of the function at t and 0 respectively, and

$$n_0 = \frac{1}{2^m}$$

where m runs from 0 to 3 (the serial number of the approximation, the computer-time increases exponentially with the number of components), n from $-$ to $+$ (at this point the amplitude of the corresponding component drops below 0.01 times the value in the origin), and t from 0 to, say 3000 (s). It is supposed that at time $t = 0$ all components (resonances) are in-phase, that is why cos-functions appear in the formula.

Figures 1 to 4 show the waveforms and the dynamic spectra for the 3000 s long portion after the initial impulse for $m = 0$ to 3. In the case $m = 0$ (Fig. 1), a beating structure can be rather well seen similarly to "normal" beats, with some additional features due to interference from more than two components. The most important such feature is the existence of longer "super-beats", being more than 2000 s long (over hundred cycles). In Fig. 2, for $m = 1$, these long beats become shorter, about 200 s, but their length is variable, and as an additional feature, a quiet interval follows the initial impulse. At higher values of m (Fig. 3, $m = 2$), these features get more and more accentuated, especially the quiet interval after the initial impulse gets longer and longer. At $m = 3$ (Fig. 4), the quiet interval is about 1000 s long. Moreover, some kind of dispersion is also seen, the period increases within each "beat", and at the beginning of the next beat, it drops again to the initial value, below the central period, 20 s. The period changes between about 17 and 25 s. None of the features mentioned in connection with the model given here are present in observed beats.

A consequence of this mathematical model is that field line resonance is more likely to appear in shells of finite thickness, thus shells of field lines are coupled to

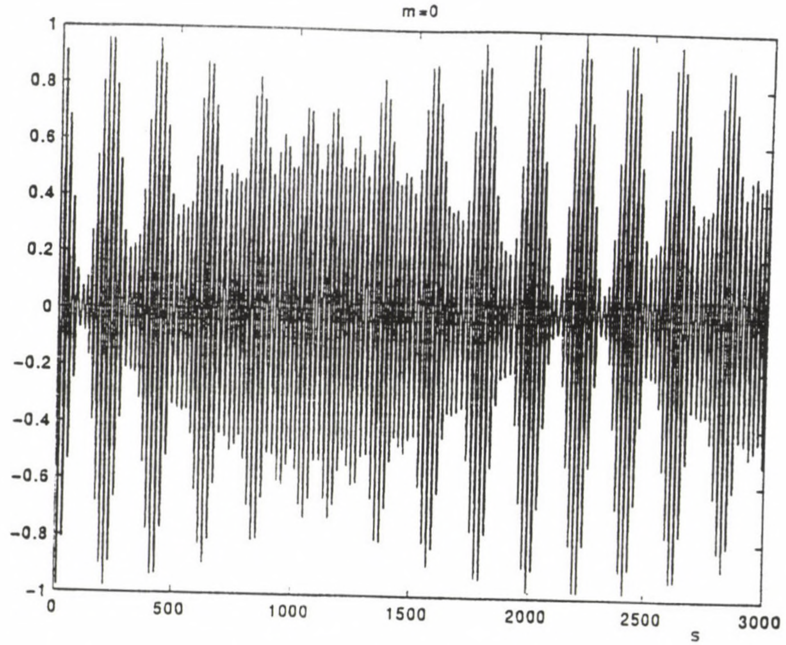


Fig. 1a. Waveform of signals resulting from Eq. (1) in the case of $m = 0$. The beating structure is clear, but "super-beats" appear, too, which are unknown in pulsation records

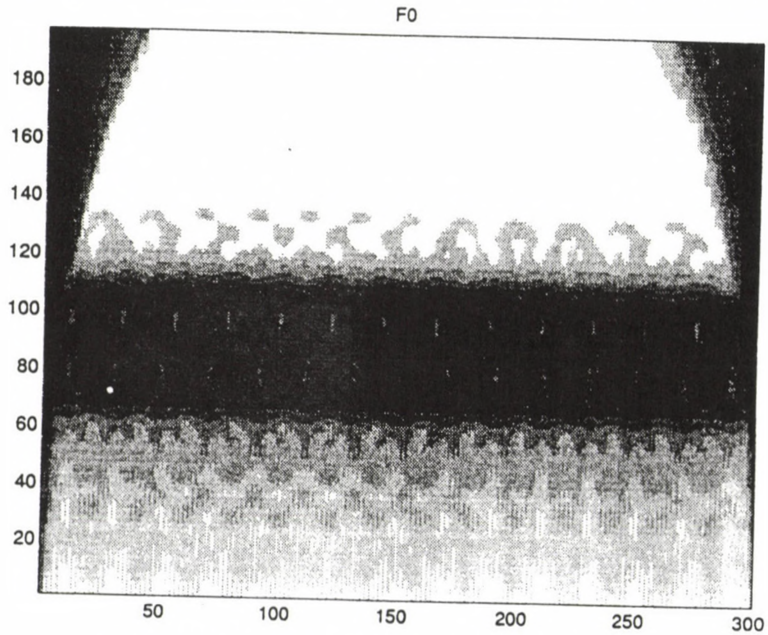


Fig. 1b. Dynamic spectrum of the waveforms in Fig. 1a

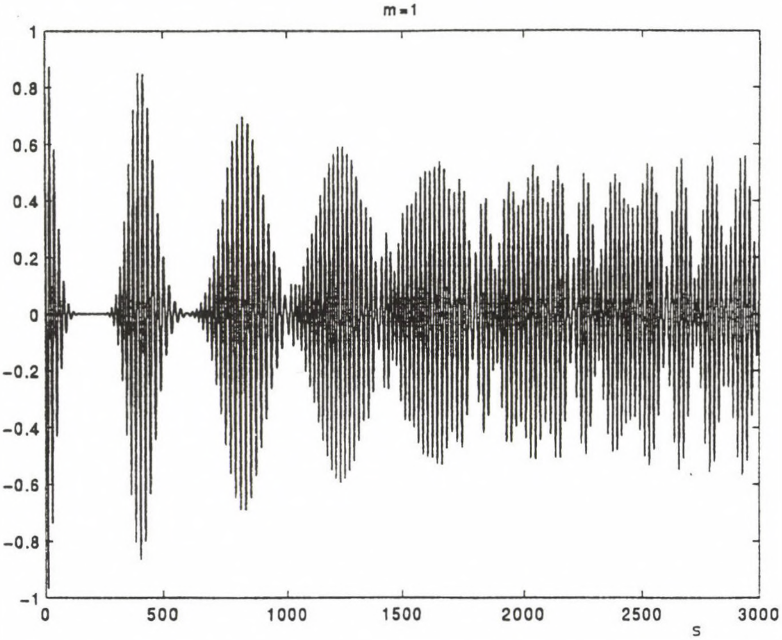


Fig. 2a. Waveform as in Fig. 1a, but for $m = 1$. With increasing m, the signal differs more and more from pulsation records. Here the beating structure is less clear than in Fig. 1a — the beats in the initial part are too long, later they get confused

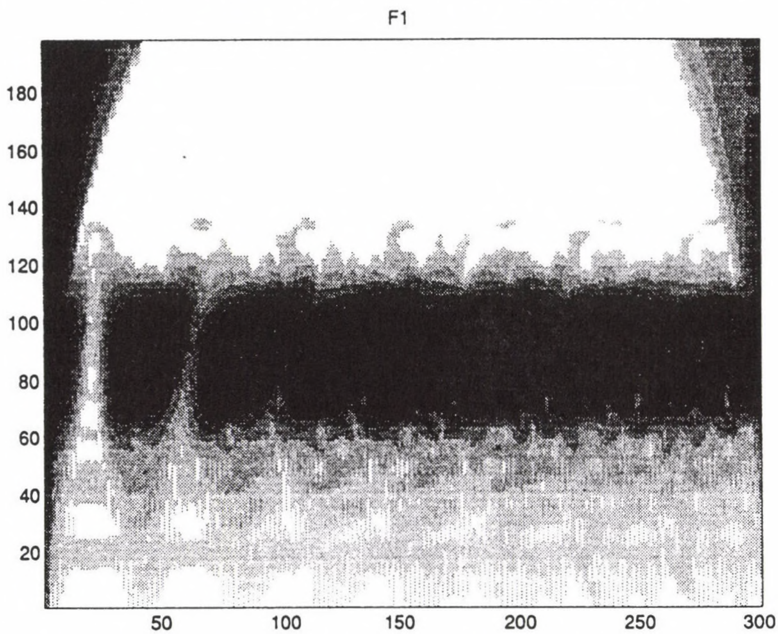


Fig. 2b. Dynamic spectrum of the waveforms in Fig. 2a. In the initial long beat, dispersion is apparent being unknown in pulsations

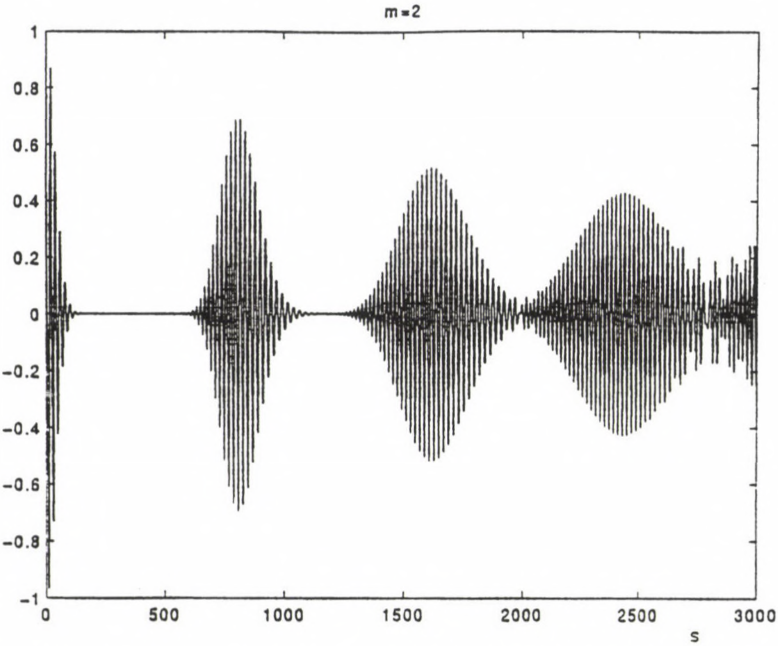


Fig. 3a. Waveform as in Fig. 1a, but for $m = 2$. The initial part (impulse) is followed by a quiet interval, then long "super-beats" follow again, at the end a very confused waveform with traces of a beating structure are seen

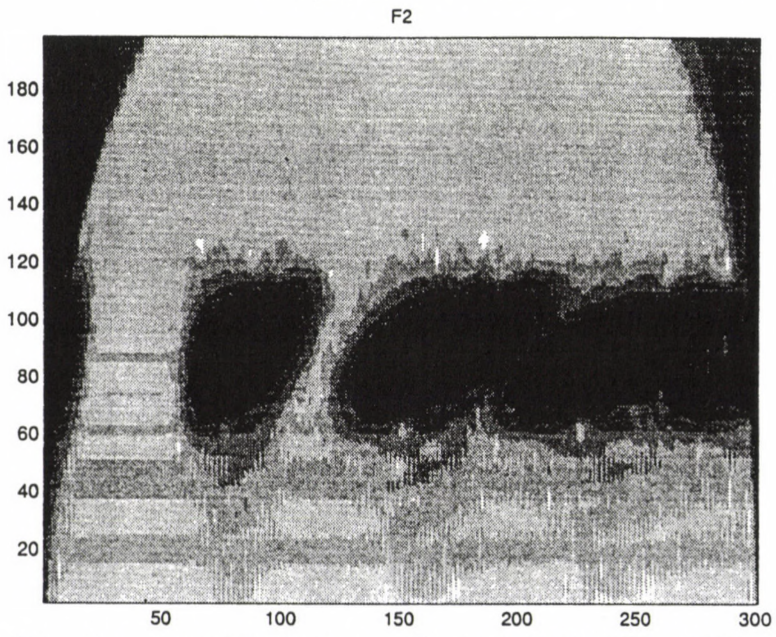


Fig. 3b. Dynamic spectrum of the waveforms in Fig. 3a. Both the initial quiet interval and the dispersion in the superbeat got more accentuated

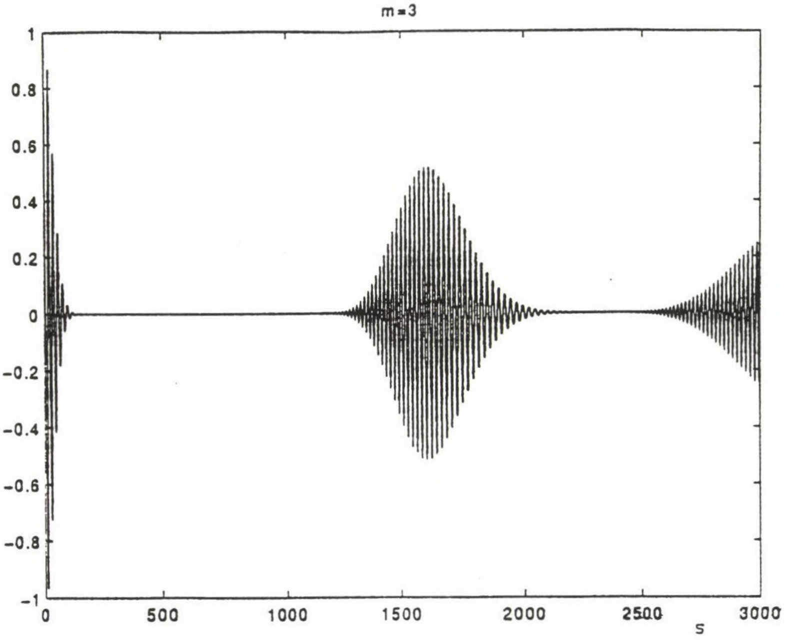


Fig. 4a. Waveform as in Fig. 1a, but for $m = 3$. The difference from pulsation records is very great here, as the initial impulse is followed by about 20 min of calm, then a nearly 10 min long "super-beat" follows, then again calm for more than 10 min

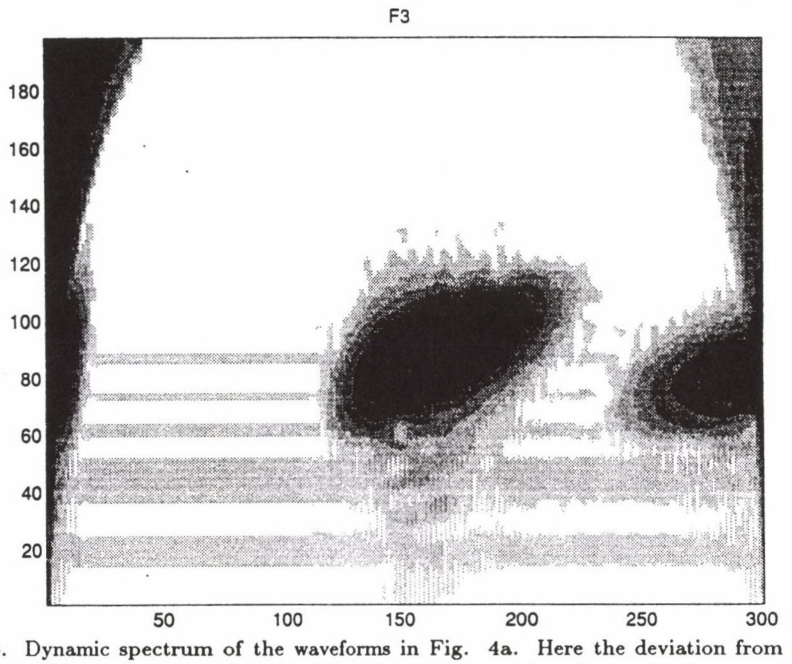


Fig. 4b. Dynamic spectrum of the waveforms in Fig. 4a. Here the deviation from dynamic spectra of pulsations are very clear, too: the dispersive superbeat is very far from anything seen in pulsations

each other. The thickness of the shell is between one hundred and a few hundred km (as found e.g. by Kurchasov et al. 1987, Verő and Cz Miletits 1994) and this is thus the thickness of the shell which resonates with a certain period, and the period changes stepwise and not smoothly with latitude or L-value. All facts gathered hitherto are clearly in favour of the stepwise change.

This research has been supported by the Hungarian State Grants Nos 1171 and 19568.

Rerefences

Chen L, Hasegawa A 1974: *J. Geophys. Res.*, 79, 1024–1032.

Cz Miletits J 1980: *J. Atm. Terr. Phys.*, 42, 563–567.

Cz Miletits J, Verő J, Szendrői J, Ivanova P, Best A, Kivinen M 1990: *Planet. Space Sci.*, 38, 85–95.

Kurchasov Yu P, Nikomarov Ya S, Pilipenko V A, Best A 1987: *Ann. Geoph.*, 5A, 147–174.

Orr D 1978: *J. Atmos. Terr. Phys.*, 35, 1–50.

Southwood D J 1974: *Planetary Space Sci.*, 22, 483–491.

Verő J 1996: *J. Geophys. Res.*, (in press)

Verő J, Cz Miletits J 1994: *J. Atm. Terr. Phys.*, 56, 433–445.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 265–270 (1996)

TO THE HISTORY OF THE BIRTH OF SOLAR-TERRESTRIAL PHYSICS

W SCHRÖDER[1]

[Manuscript received March 25, 1996]

The early history of solar-terrestrial physics is supplemented by some, mostly German contributions.

Keywords: **geomagnetism; history; solar-terrestrial physics; sunspots**

I

Cliver (1994) recently discussed the different historical roots of solar-terrestrial physics. He dealt understandably mostly with the Anglo-American influence. It is nevertheless clear that the history of solar-terrestrial physics was strongly influenced, even initiated by German scientists. Thus in the following this interesting field, representing a connection between astronomy and geophysics, is laborated in some details.

II

Hamel (1995) discussed recently the interesting controversy about the ideas on comets. Something similar is also valid for the auroras: they belong to the phenomena of heavens which had been theologically interpreted and understood for several centuries (Schröder 1984). Their interpretation, their transfer in an exact representation with the means of exact sciences succeeded only in the 18th century after the great aurora of March 1716 had prolongedly impressed people in Germany.

The role played by Johann Wilhelm Ritter (1776–1810) is to be emphasized in the scientific investigation of this phenomenon. He dealt with fireballs and magnetic thunderstorms among others, as auroras were called in those times. Causes and connections had been understood always in a telluric framework. They had been considered as phenomena completely of the atmosphere and of geomagnetism, respectively, while the latter was considered as an innate property of the Earth. Similarly to F W J Schelling (1775–1854), Ritter meant by magnetism the cross-connections and cause of the mentioned phenomena (see e.g. Ritter 1803, Hardenberg 1804, Treder 1984).

In the years 1803/04 Ritter published a paper in Gilberts Annalen with the characteristic title "On northern lights and their period, and on the connection of northern lights with magnetism, that of magnetism with fireballs, lightnings and electricity". There is no reference to an extraterrestrial source; Ritter remarked, however, at another place which has to be counted to the ancestors of the later

[1]Hechelstrasse 8, D–28777 Bremen Rönnebeck, Germany

Table I. Ritter's auroral maxima compared to the solar cycle (years)

Auroral maximum	1720/23	1739/42	1751	1760	1769/70	1779	1788	1797/98
Solar maximum	1718.2	1738.7	1750.3	1761.5	1769.5	1778.4	1788.1	1805.2

solar-terrestrial physics: he wrote on "over-mature thunderstorms" and on "immature auroras" in a letter to Karl von Hardenberg. This proves that a mere description of the phenomenon was transgressed: he asked about the causes even if the explanations were and had to remain of a speculative character (see in Ritter 1966).

Ritter investigated the connection between magnetism and "galvanism" (i.e., electric currents) which was later quantitatively proven by Ritter's friend and corresponding partner, H Oerstedt (1777–1841), then by A M Ampère (1775–1836) and by M Faraday (1797–1867) based on experiments carried out in the years 1820, 1821 and 1831 respectively. (This led to the Maxwellian theory of electromagnetism).

Ritter succeeded nevertheless in reaching an important conclusion: in 1803 he called attention to the fact that auroras ("northern lights") appeared in certain time intervals more frequently. Table I summarizes the time of maxima of auroras as given by Ritter and the solar maxima as they recently accepted.

At Ritter's time, the solar cycle had not yet been discovered therefore he could not consider this connection. Using the years of the solar maxima which became known after Ritter's time, a significant coincidence is obtained. Moreover it was evident for Ritter that auroras have a connection with magnetism; he supposed that changes in the occurrence of auroras are consequence of a general disturbance of the magnetic orderliness.

Ritter's remarks characterize — in spite of being unsatisfactory — a progressive change of the ideas.

Alexander von Humboldt discussed in Volume 4 of his "Kosmos" the connection between magnetic storms and auroras which he had systematically studied since 1803. Humboldt referred to observations at high latitudes (especially to observations made by Sabine, see later). For Humboldt, "magnetic thunderstorms" and auroras were two kinds of appearance of the same single phenomenon. He mentioned that E Halley (1650–1742) noted already in 1716 this connection owing to the big auroral event in March 1716. Besides the 1716 aurora led to a re-evaluation as an authority like Ch Wolff, the philosopher held an open lecture in Halle to soothe the anxious public. He tried to convince the audience that it was a natural phenomenon — this statement was at that time a completely new idea (Schröder 1984, 1995).

Humboldt returned to the connection between auroras and magnetic storms in the correspondence with C F Gauß and with H C Schumacher. He asked especially for references on great auroras and at observed geomagnetic storms. Humboldt stuck, however, to the notion of "magnetic thunderstorm" and did not suppose a connection Earth-Sun.

The following development is closely connected with the study of the Sun and with geomagnetic observations: there is a complex connection between them where different scientists chose different paths of approach. Alexander von Humboldt and later Arago, too, published interesting surveys of these early times (Humboldt 1858, Arago 1865). Humboldt succeeded in an observation of the aurora on December 20, 1806: The deviation from the known values in declination reached during this night 26′29″, thus the connection between geomagnetic variations and auroras was clearly established (Humboldt 1808, 1858, Biermann 1977, 1979).

Measurements of the geomagnetic field led geophysicists J von Lamont (1805–1879) and E Sabine (1788–1883) to a new conception. (Humboldt referred to this in "Kosmos", see Lamont 1867). Both found a connection between geomagnetic variations and a possible variability of the Sun. Lamont called attention to a possible period of 10 1/3 years in the magnetic declination data. Independently of Lamont, Sabine supposed a periodic change in the intensity of the magnetic field which should have some purely cosmic origin. He considered the source in the periodic changes of the Sun. Both discoveries happened in a time when Schwabe from Dessau published his solar observations. He concluded that there is a period of about 10 years in the occurrence frequency of sunspots (Humboldt 1858, Lamont 1867).

III

Lamont's, Sabine's and Schwabe's results indicated a terrestrial-cosmic connection in geomagnetic phenomena. It became evident that geomagnetic variations originate from the Sun. Was this already the birthday of solar-terrestrial physics? This problem has to be discussed in more details to be able to give an exact response to this question. In spite of the fact that connections were found and described, nevertheless, empirical data were lacking. They were only presented by the Swiss astronomers Rudolf Wolf and Hermann Fritz (Fritz 1873, 1881, 1878, Wolf 1877).

The task was to show a causal connection between terrestrial and solar events. For this purpose, long series of data and special methods of processing were needed. The decisive step in this direction was made by Rudolf Wolf (1816–1893). He presented the necessary data as follows in Table II.

From these data he deduced an average sunspot period of 11.111 ± 0.038 years. Further he remarked that years of sunspot maximum coincide with years when "northern light phenomena ...occur remarkably often" (Wolf 1877, pp. 659–670).

Wolf introduced as measure of sunspot occurrence the relative sunspot number $r = k(f + 10g)$, where f is the number of spots counted on a certain day, g is the number of sunspot groups, k is a factor depending on the person of the observer. Wolf traced sunspot cycles back far into the 18th century thus they became available till 1749. Since 1849, data have been produced without gaps.

Wolf tried continuously to improve his results and therefore he consequently continued his research. It is to be emphasized here that he published in 1857 a list of auroras (owing to his historical interest) containing 5500 data. This enabled further studies to understand better solar-terrestrial connections.

Table II. Years of sunspot maxima and minima after Wolf

Maximum (year)	Minimum (year)
1626.0 ± 1.0 (Scheiner)	1645.0 ± 1.0 (Hevelius)
1717.5 ± 1.0 (Rost)	1755.5 ± 0.5 (Zucconi)
1816.3 ± 1.0 (Stark)	1810.5 ± 1.0 (Fritsch)
1829.5 ± 1.0 (Schwabe)	1823.2 ± 0.5 (Stark)
1837.5 ± 0.5 (Schwabe)	1833.6 ± 0.5 (Schwabe)
1848.6 ± 0.5 (Schwabe)	1845.0 ± 0.5 (Schwabe)

Fortunately, Wolf's work was continued by another scientist. This person was Hermann Fritz (1830–1893), a professor of machine construction and technical drawing. He was a non-professional astronomer who interested himself with astronomy and geophysics in his spare time.

Hermann Fritz supplemented Wolf's auroral catalogue with a lot of observations. The result was the today famous "List of observed auroras", published in 1873 with support from the Austrian Academy of Sciences. This book created the empirical basis for future work in this field. Fritz started on the basis of this catalogue as well as with other solar and geomagnetic data to look exactly for possible causal connections between solar and terrestrial phenomena. He succeeded already in 1862 to prove the correlation between the frequencies of occurrence of auroras and sunspots, moreover that auroral maxima and minima occur within the eleven-year period simultaneously with those of sunspots.

A further step forward succeeded Fritz in the direction of the geographical distribution of auroras. (Muncke called already attention to the fact (see Schröder 1984, 1995) that auroras are more or less frequent at certain places and at different geographic latitudes, respectively). He coded the observations according to the sites of observation and deduced from this system a series of curves which he called isochasms (from the Greek $\chi\alpha\sigma\mu\eta$, an extraordinary or awe-inspiring phenomenon) and which connect places where auroras occur with the same frequency. Fritz's curves from 1866 and 1874 give with considerable accuracy the occurrence frequency of auroras, and from it, he concluded to the existence of an auroral zone.

The Dutch Society of Science at Haarlem posed in 1876 a competitive question. Fritz took part at this competition and obtained in 1878 the "Golden Medal". In the same year his book was published on "The connection between sunspots with the magnetic and meteorological phenomena of the Earth". In the book he dealt with southern lights, too, i.e. with auroras of the southern hemisphere. The comparison of the northern and southern hemispheres resulted in remarkable similarities.

Fritz's work on the auroras, as well as Wolf's studies of the Sun are important cornerstones in the understanding of solar-terrestrial connections. All prerequisites became available at this time both from the geomagnetic and from solar side to enable a deeper understanding of the physical causes of this connection.

Table III. Developments in solar-terrestrial physics in the 18–19th centuries

	C Wolff (1816) J W Ritter (1803) A Humboldt (1807)	E Halley (1716)	
J Lamont (1845/67) E Sabine (1852)		H Schwabe (1843)	
	A Gautier (1869)	R Wolf (1857, 1877)	H Fritz (1862, 1873) J Lovering (1867) J Loomis (1860)
		W Foerster (1872)	
	E Goldstein (1879)	O Jesse (1872)	
A Angström (1867/68)	F Zöllner (1881)	W Boller (1898)	
C Vogel (1872)	K Birkeland (1899/1912)	C Störmer (1904/12)	

Note: years indicated in the table refer to highlights of the activity in solar-terrestrial physics and year of the publication of important works, respectively

IV

The next step was due to come from the Sun's monitoring. Carrington observed on September 1, 1859 a chromospheric disturbance. Later it was found on the records of the Kew observatory that at this time all the three elements of the geomagnetic field suddenly changed. Marchand noticed additionally that the maximum magnetic disturbance occurred when spots crossed the central meridian on the Sun. It became soon evident that some connection has to exist there. Else Wilhelm Foerster (1832-1921) dealt with the same problem, too. He noted in 1872 that one had to look for the source of these disturbances in the solar environment. Especially interesting physical considerations were made at that time by Johann Karl Friedrich Zöllner (1834–1882). He supposed among others the source of geomagnetism in currents of the fluid outer core of the Earth which cause by friction electric currents (Zöllner 1881).

Zöllner's study of the aurora on October 25, 1870 is also remarkable when he as first recorded the red auroral line. The source of the light in connection with auroras he supposed to be "glowing gas particles in our atmosphere". Moreover Zöllner declared unambiguously that it must be a connection between processes on the Sun and certain events on the Earth. Zöllner emphasized very clearly this causal physical connection between Sun and Earth, a fact which had been since mostly disregarded (also by Cliver). The further development in the knowledge about the connection between Sun and Earth was due to the use of spectroscopic methods, to the results of the study of cathode- and channel-rays (Goldstein), to the experimental study of auroras (terella experiments) by Kristian Birkeland (1867–1917), further to the use of photography in the measurement of the height of auroras. At the beginning of the 20th century it became evident that auroras are phenomena of the high terrestrial atmosphere being influenced and controlled by solar activity.

The "birthday" of solar-terrestrial physics was thus in the mid-19th century. It is by no means the result of the activity of a single scientist or even of a single group of scientists, it is more the result of a many-sided approach to the problem. Contributions from the geomagnetism were as necessary as those from solar physics; moreover, exact methods of statistical treatment and processing of a big amount of data were necessary, too. It was possible on the basis of a large number of single steps to reach a re-evaluation of the phenomenon aurora and from it, to a deeper interpretation of the connections Sun-Earth.

References

Arago F 1865: Populäre Astronomie. Wigand, Leipzig

Biermann K R ed. 1977: Briefwechsel zwischen Alexander von Humboldt und Carl Friedrich Gauss. Akademie-Verlag, Berlin

Biermann K R ed. 1979: Beifwechsel zwischen Alexander von Humboldt und Heinrich Christian Schumacker. Akademie-Verlag, Berlin

Cliver E W 1994: *EOS Transactions AGU*, 75, 609.

Fritz H 1873: Verzeichnis beobachteter Polarlichter. Wien

Fritz H 1878: Die Beziehungen der Sonnenflecken zu den magnetischen und meteorologischen Erscheinungen der Erde. Haarlem

Fritz H 1881: Das Polarlicht. Leipzig

Hamel J 1995: *Die Sterne*, 71, 18.

Hardenberg K A 1804: *Magaz. f. d. neuesten Zustand d. Naturkunde*, 8, 285.

Lamont J 1867: Handbuch des Magnetismus. Leipzig

Ritter J W 1803: Einiges über Nordlichter und deren Periode, und über den Zusammenhang des Nordlichts mit dem Magnetismus, und des Magnetismus mit den Feuerkugeln, dem Blitze und der Electricität. *An. Physik*, 15.

Ritter J W 1966: Briefe eines romantischen Physikers. F Klemm, A Hermann eds, Moos, München

Schröder W 1984: Das Phänomen des Polarlichts. Wiss. Buchgesellschaft, Darmstadt

Schröder W 1995: Sunspot Cycles and Auroral Activity. Science Edition/IAGA, Bremen

Störmer C 1955: The polar aurora. Oxford

Treder H J 1984: In: Natur und geshichtlicher Prozess. v H J Sandkühler ed., Frankfurt, Suhrkamp

v Humboldt A 1808: Die vollständigeste aller bisherigen Beobachtungen über den Einfluss des Nordlichtes auf die Magnetnadel. Bilb. Ann., 29.

v Humboldt A 1858: Kosmos, Band IV. Cotta'scher Verlag, Stuttgart/Augsburg

Wolf R 1877: Geshichte der Astronomie. München

Zöllner F 1881: Wissenschaftliche Abhandlungen mit zwei populären Verträgen, Vierter Band. Leipzig

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 271–273 (1996)

A NOTE ON THE SUGGESTED VARIABILITY OF THE LENGTH OF THE SOLAR CYCLE DURING RECENT CENTURIES AND SOME CONSEQUENCES

W SCHRÖDER[1] and H-J TREDER[2]

[Manuscript received July 9, 1996]

It is shown that there exists neither physical explanation nor plausible mechanism which could initiate or support any variability in the length of the solar cycle. Furthermore, any suggestions concerning a relationship with terrestrial climate are questionable and without any supporting physical evidence.

Keywords: **auroral cycle; length of the solar cycle; Maunder mimnimum**

Recently Lassen and Friis-Christensen (1995) have published papers in which they suggested that variations in the length of the solar cycle have occurred in recent centuries. Furthermore they suggested that a possible relationship exists with terrestrial climate. From our present physical understanding these conclusions cannot be accepted.

Furthermore the authors have used a rather limited number of literature references and have overlooked many critical papers published in recent years concerning solar variability and the so-called Maunder minimum. Because of these limitations the conclusions of the authors cannot be accepted in all details.

The existence of Maunder or Spörer minima is not well defined or documented if all available auroral data are considered. Thorough research on aurora and sunspots shows no unusual change for the 16th and 17th centuries, or for the period 1450–1550 and 1710–1750. If we use all available sources (including those in handwritten form in libraries and archives) we can show that there was no unusual pattern of occurrence of auroras during the period 1450–1750 as observed from middle latitudes in central Europe. In fact, using all the available data, we found a regular pattern of occurrence of auroras during this time. Auroras are indicators of solar activity (or solar wind) and we see no unusual variability in occurrence of minima. Table I indicates the times of possible maxima and minima in solar activity (and auroras) for 1550–1715 as deduced from the auroral data.

A consideration of Table I shows little or no agreement with the suggestions published by Lassen and Friis-Christensen and in related literature. Another aspect is the length of the solar cycle. From the table it seems that the sunspot cycle may have varied in length but we need more data for the period 1450–1715 before a hypothesis can be substantiated (cf. Table I). If there was a change in solar activity or in its components, as suggested by Lassen and Friis-Christensen and others, we must find another explanation because auroras were not observed in every year of

[1]Hechelstrasse 8, D–28777 Bremen Rönnebeck, Germany
[2]Rosa-Luxemburg-Strasse 17a, D–14482 Potsdam, Germany

1217-8977/96/$ 5.00

Table I. Table of auroral and sunspots maxima/minima

Max. Epoch	Auroras	Min. Epoch	Auroras
1549 (1547*)	1548f	1554 (1554+)	1552/1553
1560 (1558+*)	1560	1565 (1567+)	1566
1571 (1571.4**)	1571/1572	1577 (1578*)	1577/1578
1582 (1582.5**)	1581/1582	1588 (1587*)	1589?
1593 (1593*)	1592/1593	1600 (1598*)	1600/1601 (??)
1604 (1604.7**)	1604	1610 (?)	1610/1611
1615 (1615.8)	?	1619 (?)	1619/1620
1626 (1626.9**)	1626	1634 (?)	1636?
1638 (1638.1**)	1639	1645 (?)	1644 (?)
1649 (1649.2**)	1648	1655 (?)	1654
1660	1661	1666 (?)	1666/1667
1671	1671/1672	1677 (?)	1678/1679
1682	1681/1682	1688 (?)	1688/1689
1693	?	1699 (?)	1699/1700
1704 (1704.7**)	1704/1705	1710 (?)	1710/1711
1715 (1715.9**)	1716	1721 (?)	1721/1722

*see Schove (1979); ** Wittmann (1978a, 1978b)
Maxima given with p: 11.116 years by Wittmann (1978a, 1978b)
Aurora Maxima determined with 5 auroras and more.

each alleged minimum (maximum). Concerning the problem of auroras and solar activity we note also the papers by Vitinsky (1978), Jiang and Xu (1986), Leftus (1993) and Landsberg (1980).

Furthermore concerning the relationship between solar activity and climate change, a topic also discussed in literature, we should note that Landsberg (1980) in his classical paper on the so-called Maunder minimum (not referred by Lassen and Friis-Christensen) made it clear that there was no significant correlation.

A further point is that a real change of solar constant is theoretically a very powerful effect. Since the energy of solar radiation results from nuclear reactions in the centre of the Sun, the solar constant should show only small fluctuations. A variation of climate over a few decades must have terrestrial origin. A possible mechanism for a variation of the solar constant in correlation with solar activity is that with increasing solar activity a small part of the heat energy becomes converted into magnetic and kinetic energy of the solar atmosphere and this energy dissipates during decreasing solar activity. The mean value of the solar constant would remain unchanged and independent of the amplitudes of the 11-year solar cycles. The point is that such a hypothetical variation of the solar constant would show minimum values during rising solar activity and maximum values during decreasing activity.

The question of the variation in solar irradiance (and the evidence for it) remains speculative, and we should continue to monitor solar variations closely (cf. Schröder 1988, Haubold and Beer 1991, Legrand et al. 1992). Following the paper by Lassen and Friis-Christensen and other publications we suggested that there should be a more critical discussion of the whole problem of the relationships between solar activity, the length of the solar cycle, the frequency of auroras, the possible correlation with climate, and the existence or non-existence of the so-called Maunder or Spörer minima. Such critical appraisals should take into account the full body of auroral data from middle latitudes, not simply those which support one side of the argument, and journals should be prepared to publish critical reviews on the subject.

An accurate and detailed knowledge of the literature (cf. Schröder 1984) and of the various historical sources (published and unpublished) is clearly needed and hopefully the paper by Lassen and Friis-Christensen and others will stimulate further interest in achieving these goals.

References

Haubold H J, Beer J 1991: In: Solar-terrestrial variability and global change, W Schröder and J-P Legrand eds, Science Edition, Bremen-Rönnebeck

Jiang Y, Xu Z 1986: *Astrophys. and Space Sci.,* 118, 159–162.

Landsberg H 1980: *Arch. Met. Geophys. Bioklim.*, B28, 181.

Leftus V 1993: *Solar Phys.*, 145, 377–388.

Legrand J P, LeGoff M, Mazaudier C, Schröder W 1992: *Com. rend. ser. gen.*, 8, 181–219.

Lassen K, Friis-Christensen E 1995: *J. Atm. Terr. Phys.*, 57, 835–845.

Schröder W 1984: Das Phänomen des Polarlichts (Aurora Borealis). Wissenschaftliche Buchgesellschaft, Darmstadt

Schröder W 1988: *Meteorol. Atmos. Phys.*, 38, 246–251.

Vitinsky Y I 1978: *Solar Phys.*, 57, 475–478.

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 275–277 (1996)

In memoriam
Prof. em. Dr. h. c. Dr. István Hazay
(1901–1995)

Mourning in their hearts, members of the Hungarian Academy of Sciences, professors and staff of the Technical University of Budapest, a wide circle of Hungarian geodesists, surveyors, and cartographs, his former students and colleagues accompanied Professor István Hazay, the oldest Hungarian geodesist, full member of the Hungarian Academy of Sciences, to his grave.

In the course of his nearly 95-year life, Professor Hazay was a dominant personality of the history of 20th-century geodesy, surveying, and cartography in Hungary.

Safe and sound, he survived two world wars and the hard times following them, with his star constantly rising from being a junior staff engineer through being the head of the Hungarian National Survey and a university professor to the full membership of the Hungarian Academy of Sciences. (His detailed biography and the list of his publications were published on the occasion of his 80th and 90th birthdays by Acta Geod. Geoph. Mont. Hung., pp. 3–4, Vol. 16/1981 and pp. 469–470, Vol. 26/1991, respectively.)

His excellent engineering and management skills, coupled with singular tolerance and humanity, elevated him to be a leader very soon. In all his management activities: as head of the National Survey, Professor and Head of Department, Dean of the Faculty of Civil Engineering, Rector of the University, as well as President of the Geodetic Commission of the Hungarian Academy of Sciences, he set a brilliant example of professional competence, humanity, and well-meaning helpfulness.

In his personality, the scientist of a wide intellectual horizon is extremely luckily amalgamated with the engineer of excellent practical common sense. He always directed his scientific research activity towards the solution of tasks raised by engineering practice.

Besides, Professor Hazay's pedagogical affinity must also be particularly emphasized. Generations of engineers enjoyed his excellent lectures, studied from his perfectly built university lecture notes, textbooks, and handbooks.

His outstading scientific activity is hallmarked by approximately 100 special papers, studies and 10 books and textbooks, a considerable proportion of which has been reprinted in several editions. His book titled "Adjusting Calculations in Surveying", perfectly built up from the didactical point of view, has been published in English as well. One of his works of major importance is the "Handbook of Geodesy and Surveying" published in Hungarian in 3 volumes, with him being the editor and partly the author of it.

Among his prominent results, the most significant ones include the elaboration of the teaching materials of a number of special subjects in geodesist, surveyor, and cartograph training as well as writing their university lecture notes, textbooks and professional books; the solution of many problems of the projections of large-scale mapping; the recognition of the principle of mechanical and static analogue adjustment and the elaboration of its procedure; the adjustment of weighted observations with different dimensions; the establishment of the so-called "method of dominant stations" for the adjustment of extent geodetic control networks, which was also applied in the modern re-establishment of the Hungarian national geodetic control net. It was this that brought him the greatest professional recognition of the state, the "Kossuth Prize".

Our Lord blessed him with incredible health, physical strength, and the full possession of all his faculties. His retirement at the age of 70 rather meant only a formal alteration in his way of employment. Having freed himself from administrative tasks, he continued his scientific work as well as his university lectures. He rendered great help to all of us by critically reviewing our papers. He did not take leave from his students as long as the age of 90.

Professor Hazay's career was accompanied by numerous social and government recognitions, the value of which is especially increased by the fact that he preserved his political independence throughout all his life. His oeuvre is hallmarked by the "Kossuth Prize"; ministerial and government orders and decorations; the "Lázár-Deák" commemorative medallion and the "Antal Fasching" medal of geodesists; the commemorative medallion and the title "doctor honoris cause" of the Technical

University of Budapest, as well as by the golden technical doctor's and the diamond engineer's degrees. In his personality, the engineer of excellent practical common sense was extremely luckily amalgamated with the accomplished teacher and the scientist of a wide intellectual horizon.

His life rendered the shining example of professional competence and humanity every time and at all places.

He reached the summit of his career in all the fields of his activity: as a practicing and later managing geodesist, as well as a professor and scientist. Yet, the deepest memory of him is left by his human greatness. He created a friendly atmosphere around himself in which his colleagues could get to know the loving husband and father worried about his family besides the scholarly professor; he always saw to it in due time that he should have proper professional and managing replacement, paving the way for his successors with tender fatherly care so that they could follow his footsteps; he rejoiced at all the successes of his students, colleagues, and successors. All his life was characterized by the unity of professional devotion and humanity.

The passing of Professor Hazay has meant the loss of an excellent Hungarian scholar of geodesy, surveying, and cartography; a school-creating professor of technical university training, more specifically, of geodesists and surveyors; a person immensely loving his family and his profession, serving an eternal example by his human properties worthy of respect and affection, a personality generally esteemed.

"Requiescat in pacem."

P Bíró

Acta Geod. Geoph. Hung., Vol. 31(1–2), pp. 279–280 (1996)

Book reviews

M PELTZ, L DINU: *Bibliografiă Geologică şi Geofizică a Republicii Socialiste România.* Vols VII and VIII (1976–1980), Ministerul Geologiei Institutul de Geologie şi Geofizica, Bucureşti, 1985, 1994 284 and 347 pages

These valuable volumes reporting on 1917 and 2008 papers respectively, consist of the following parts

- Introduction (Rumanian, French and English)
- List of abbreviations and titles of periodicals, publishing houses, scientific events and quoted institutions
- The order of compiling the bibliographic citations
- The bibliographic citations in alphabetical order of authors
- The index of authors
- The subject index
- The geographic index achieved on the locations
- The geographic index according to the geological-structural unites.

This collection is great help for all scientists who want to make researches in any topic connected to the Rumanian geology and geophysics. The authors deserve a substantial amount of credit for making this great work.

A Ádám

W SCHRÖDER, M COLACINO eds: *Interdivisional Commission on History of the IAGA and History Commission of the German Geophysical Society.* Bremen-Roennebeck, Germany, 1996, 292 pages

This book comprises selected papers presented at the IUGG General Assembly (Boulder, USA, July 1995) at the symposium of the interdivisional commission on history of the IAGA. The present volume is the 7th in the series of proceedings of the IUGG or IAGA meetings. The contents of the book include totally 15 contributions dealing with different topics from solar terrestrial physics to geomagnetic observations, from climate variations to volcanoes, from archaeoastronomy and its geophysical implications to the variability of solar activity. Most of the reviews contain detailed natural history observations and show the connections between topics of current interest and their historical background.

It opens with a paper entitled "The prime mover of volcanoes — History of a concept" by G P Gregori and W Dong. It focuses on the concept of prime mover of volcanism in the 'western science' from the birth of sicence until the first few decades of our century. It is a comprehensive study of totally 64 pages (more than one fifth of the whole volume) with marvellous illustrations. The authors present also a very detailed table and comments (28 pages) containing a chronological scheme of ideas and models. It is called the heart of the paper. Similarly the very last part of this edited volume (written by G P Gregori and L G Gregori) is also a study of outstanding interest synthesizing the different aspects of archaeoastronomy (60 pages). Owing to the thorough work of the editors these two excellent studies give a very good frame for the whole volume.

Without mentioning all contributions and authors of the collection one by one, the following highlighted examples illustrate the wide range of topics covered by the book: e.g. E Piervitaly et al. present an analysis of the droughts as seen through religious events for a 3-centuries-long period, and also an analysis of the very cold summers and their causes in the 19th centruy. P A McNoe examined the auroral and magnetic effects of high altitude nuclear explosions. L Cafarella et al. summarize the results of the historical observations of the geomagnetic field in Italy and point the importance of these measurements. The paper of M A Miah and M A Samad entitled 'History of the environment of Bangladesh' introduces the data of population growth, the land use and land cover changes in the country and the effects of erosion and change of river courses. Regarding its content, the otherwise interesting and informative report, however, hardly fits in the row of the aforementioned contributions.

This publication will no doubt be of valuable assistance to anyone working in the field of Earth's sciences or archaeology and will also be useful for future reference. According to G Gregori "an appreciation is definitely due to all contributors and to the entire scientific community who supported them". The particular value of this book is that it presents a wide range of issues and shows the richness of history of the geophysical disciplines.

I Wesztergom

Manuscripts should include the following components which should be presented in the order listed (tables and illustrations should be separated in case of manuscripts on disk, too).

1. Title, name(s) of the author(s), affiliation, dateline, abstract, keywords
2. Text, acknowledgements
3. References
4. Footnotes
5. Legends
6. Tables
7. Illustrations

1. The *affiliation* should be as concise as possible and should include the complete mailing address of the authors. The *date of receipt* will be supplied by the editors. The *abstract* should not exceed 250 words and should clearly and simply summarize the most important methods and results. 5–10 significant expressions describing the content are used as *keywords*. Authors may recommend these keywords.

2. The *text* should be in English and as short and clear as possible. In case of typewritten manuscripts, please note the following:

— avoid possible confusion between o, O (letters) and 0 (zero), l (letter) and 1 (one), ν (Greek nu) and u, v (letters), etc.

— explain ambiguous and uncommon symbols by making marginal notes in pencil

— formulas are to be numbered consecutively with the number in parentheses to the right of the formula. Text references to the equations may then be made by the number in parenthesis. The word equation in this context is to be abbreviated to Eq. and Eqs in the plural

— the International System of Units (SI) should be used.

3. *References* are accepted only in the Harvard system. Citations in the text should be as:

...(Bomford 1971)...or Bomford (1971)...

...(Brosche and Sündermann 1976)...

...(Gibbs et al. 1976b)...

The list of references should contain names and initials of all authors (the abbreviation et al. is not accepted here): for journal articles, year of publication, the title of the journal in abbreviated form, volume number, first and last page. For books or chapters in books, the title is followed by the publisher and place of publication.

All items must appear both in the text and in the list of references.

4. *Footnotes* should be typed on separate sheets.

5. *Legends* should be short and clear. The place of the tables and figures should be indicated in the text, on the margin.

6. *Tables* should be numbered serially with Roman numerals. Vertical lines are not used.

7. All the *illustrations* should contain the figure number and author's name in pencil on the reverse. The most important point with figures is clearness. Photographs and half-tone illustrations should be sharp and well contrasted. Colour photographs will be accepted, but the extra cost of reproduction in colour must be borne by the authors (approx. US$ 260 per page).

Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement. Information is sent to the first author if no other wish is expressed. Fifty offprints are supplied free of charge, for further copies please contact the Editorial Office.

Printed in Hungary
Akadémiai Nyomda, Martonvásár

Acta Geodaetica et Geophysica Hungarica

18

VOLUME 31, NUMBERS 3–4, 1996

INCLUDING PAPERS OF THE FIFTH WINTER SEMINAR SOPRON ON GEODYNAMICS (WSS '96) "GRAVITY IN TIME AND SPACE"

Akadémiai Kiadó, Budapest

AGG 31 (3–4) 281–480 (1996) HU ISSN 1217–8977

ACTA GEODAETICA ET GEOPHYSICA HUNGARICA
A Quarterly Journal of the Hungarian Academy of Sciences

Acta Geodaetica et Geophysica Hungarica (AGG) publishes original reports and reviews on geodesy and geophysics in English.

AGG is published in yearly volumes of four issues by

AKADÉMIAI KIADÓ
H-1117 Budapest, Prielle K. u. 19–35

Manuscripts and editorial correspondence should be addressed to

AGG Editorial Office
Geodetical and Geophysical Research Institute
H-9401 Sopron, P.O. Box 5
Phone: (36-99) 314-290
Fax: (36-99) 313-267
E-mail: actagg@ggki.hu
Internet: http://www.ggki.hu/agg/Acta-Geod-Hung.html

Subscription information

Orders should be addressed to

AKADÉMIAI KIADÓ
H-1519 Budapest, P.O. Box 245

Subscription price for Volume 31 (1996) in 4 issues US$ 98.00, including normal postage, airmail delivery US$ 20.00.

INSTRUCTIONS TO AUTHORS

Manuscripts should be sent to the Editorial Office (address see above). Articles not published or submitted for publication elsewhere are only accepted.

Manuscripts should be typewritten in duplicate or on disk (LATEX*.tex files or Word for Windows Documents or text-only ASCII files) with a written copy, or via E-mail (LATEX*.tex files or text-only ASCII files).

CONTENTS

Book reviews

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 281–281 (1996)

Preface

The fifth Winter Seminar Sopron on Geodynamics (WSS '96) took place in Sopron from the 20th to the 24th February 1996. It was organized by the Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences under the auspices of the International Association of Geodesy, further of the European Geophysical Society and of the Central European Initiative's Committee of Earth Sciences. The organizers acknowledge financial support from the Hungarian National Committee for Technological Development. The Seminar attracted some 50 participants from 12 countries. Fifteen lectures were presented, among them six were invited by the organizers for a deep and detailed discussion of the main topics of WSS '96 entitled "Gravity in Time and Space".

The first Winter Seminar in Sopron was organized in 1987 and it dealt with problems connected to the Earth's rotation (lecturers were K Bretterbauer, I Abonyi, H Moritz, V Dehaut, H Jochmann, P Melchior, M I Yurkina). The second took place in 1989 and problems connected with deformations of the Earth were discussed by H Moritz, E Groten, E W Grafarend, K Arnold and I Abonyi. The 1992 Seminar was concentrated on the question of the inner structure and dynamics of the Earth (invited lecturers were S M Molodensky, C Denis, V N Zharkov, A Ádám, G F Panza and further important papers were presented by M Stavinschi and S Frank). The 1994 Winter Seminar concentrated on techniques and methods used in geodynamical research. The invited lecturers included H Schuh, I Fejes, G Kirchner, B Ambrosius, P Wilson, W Zürn, Gy Mentes, F Sanso.

The 1996 Winter Seminar was aimed to discuss important problems of gravimetry (invited lecturers: E Grafarend, E Groten, J Hinderer, I Marson, G Papp, P Varga). About half of the presented papers were devoted to theoretical problems and the participants discussed the use of gravimetric data for the study of the dynamics of the Earth and its structure, temporal variations of the gravity field due to internal and external volume and surface forces and problems of isostasy. Another part of the contributions dealt with the determination of the gravimetric geoid, the use of different mathematical and physical tools (e.g. finite element method, FFT) to describe phenomena connected to the gravity field of our planet. Other lectures presented considerations about the use of different techniques to measure the gravity field, including problems of the absolute gravimetric measurements, measurements made with superconductive gravimeters, further problems of gravimetric measurements in laboratory.

— · —

We publish in the present volume a scientific contribution by Milan Burša ("The indirect tidal torque and precession — nutation dynamics") together with the papers presented at the Winter Seminar. The topic of Burša's paper is near to the subject and spirit of our Winter Seminars and therefore we are happy to include it in this volume.

L Bányai, G Papp, P Varga

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 283–292 (1996)

THE GRAVITATIONAL POTENTIAL OF A DEFORMABLE MASSIVE BODY GENERATED BY TIDAL AND LOAD POTENTIALS

E GRAFAREND[1], J ENGELS[1], P VARGA[2]

The gravitational potential of a deformable massive body is determined for the case of a deformation generating tidal potential as well as load potential especially for a symmetric non-rotating isotropic elastic model Earth (SNREI). In particular, the source term for the time-varying incremental gravitational potential is identified as the time-varying incremental mass density. Due to *mass conservation* during deformation the incremental mass density is inferred from the divergency of the product of initial mass density and the vector-valued displacement field. The variation of the gravitational field caused by the displacement field under the *Love-Shida* functions $\{h, l, h', l', h", l"\}$ relates to the *Love* functions $\{k, k', k"\}$ which are connected by the integral relation, namely the *Love-Shida* functions are not independent.

This integral relation is tested numerically for the Molodensky (1953) Earth model with respect to an initial PREM mass density model due to Dziewonski and Anderson (1981).

Keywords: Love-Shida hypothesis; Love-Shida numbers; Poisson equation; symmetric non-rotating elastic isotropic (SNREI) Earth

The decomposition of the *space-time* gravity potential $W(\mathbf{x}, t)$, also called geodetic *height*, into a stationary part $W(\mathbf{x})$ of gravity potential and a time-variant part, the "deformation potential" $\delta w(\mathbf{x}, t)$ has been proposed by Grafarend (1990) in order to emphasize the transfer from *gravitostatics* to *gravitodynamics* for a deformable massive body like the Earth. The "deformation potential" is also called the incremental gravity potential. The deformation potential $\delta w(\mathbf{x}, t)$ fulfills the field equation of Poisson type

$$\operatorname{div} \operatorname{grad} \delta w(\mathbf{x}, t) = -4\Pi G \delta \varrho(\mathbf{x}, t) .$$

In this paper lower case variables describe the undeformed state.

The continuity equation

$$\delta \varrho(\mathbf{x}, t) = -\operatorname{div} [\varrho(\mathbf{x}) \mathbf{d}(\mathbf{x}, t)]$$

is accepted if no loss of mass during the deformation is assumed. G denotes the gravitational constant; $\varrho(\mathbf{x})$ and $\mathbf{d}(\mathbf{x}, t)$ characterize the initial mass density and the displacement vector of the deformable Earth body, respectively. The special solutions of the incremental Poisson equation are represented by the Newtonian type formulas

$$\delta w(\mathbf{x}^*, t) = -G \int \frac{\operatorname{div} [\varrho(\mathbf{x}) \mathbf{d}(\mathbf{x}, t)]}{||\mathbf{x}^* - \mathbf{x}||} \, d^3\mathbf{x} , \tag{1}$$

[1]Department of Geodetic Science, Stuttgart University, Keplerstr. 11, D–7000 Stuttgart 1, Germany

[2]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary, e-mail: varga@ggki.hu

where lower case x denotes the undeformed state, $\mathbf{x}^*$ is the field point on the Earth's surface and $\mathbf{x}$ is the source point in the Earth's interior; $||\mathbf{x}^* - \mathbf{x}||^{-1}$ could be expanded into *orthonormal scalar spherical surface functions* (defined in the Appendix) i.e. associated Legendre functions

$$\frac{1}{||\mathbf{x}^* - \mathbf{x}||} = \frac{1}{\mathbf{r}^*} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{1}{2n+1} \left(\frac{r}{r^*}\right)^n Y_{nm}(\lambda^*, \phi^*) Y_{nm}(\lambda, \phi) . \tag{2}$$

The series (2) converges uniformly if $r \leq r^*$. $\{\lambda, \phi, r\}$ are the spherical coordinates.
In addition:

$$\operatorname{div}[\varrho(\mathbf{x})\mathbf{d}(\mathbf{x}, t)] = \varrho(x) \operatorname{div}(\mathbf{x}, t) + \langle \operatorname{grad} \varrho(\mathbf{x}) | \mathbf{d}(\mathbf{x}, t) \rangle \tag{3}$$

and the displacement vector field $\mathbf{d}(\mathbf{x}, t)$ can be represented in *orthonormal vector spherical surface functions* (defined in the Appendix)

$$\begin{aligned} \mathbf{d}(\mathbf{x}, t) &= \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} r^{nm}(r, t) \mathbf{R}_{nm}(\lambda, \phi) + \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} s^{nm}(r, t) \mathbf{S}_{nm}(\lambda, \phi) + \\ &+ \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} t^{nm}(r, t) \mathbf{T}_{nm}(\lambda, \phi) . \end{aligned} \tag{4}$$

The series (4) converges uniformly for properly chosen decay functions $r^{nm}(r, t)$, $s^{nm}(r, t)$ and $t^{nm}(r, t)$. r^{nm}, s^{nm} constitute the spheroidal displacement, while t^{nm} toroidal displacements. It can be derived by substituting (2)–(4) into (1)

$$\begin{aligned} \delta w(\mathbf{x}^*, t) &= -\frac{G}{r^*} \int \sum_{n_1=0}^{\infty} \sum_{m_1=-n_1}^{+n_1} \frac{1}{2n_1+1} \left(\frac{r}{r^*}\right)^{n_1} Y_{n_1 m_1}(\lambda^*, \phi^*) Y_{n_1 m_1}(\lambda, \phi) \cdot \\ &\cdot \sum_{n_2=0}^{\infty} \sum_{m_2=-n_2}^{+n_2} \left\{ \left[\frac{2}{r} r^{n_2 m_2} + \frac{d}{dr} r^{n_2 m_2} - \frac{[n_2(n_2+1)]^{1/2}}{r} s^{n_2 m_2} \right] \cdot \right. \\ &\cdot \varrho(\mathbf{x}) Y_{n_2 m_2}(\lambda, \phi) + \\ &+ \frac{1}{r \cos\phi} \frac{\partial \varrho}{\partial \lambda} \left[\frac{s^{n_2 m_2}}{[n_2(n_2+1)]^{1/2}} \frac{1}{\cos\phi} \frac{\partial}{\partial \lambda} Y_{n_2 m_2}(\lambda, \phi) + \right. \\ &+ \left. \frac{t^{n_2 m_2}}{[n_2(n_2+1)]^{1/2}} \frac{\partial}{\partial \phi} Y_{n_2 m_2}(\lambda, \phi) \right] + \\ &+ \frac{1}{r} \frac{\partial \varrho}{\partial \phi} \left[\frac{s^{n_2 m_2}}{[n_2(n_2+1)]^{1/2}} \frac{\partial}{\partial \phi} Y_{n_2 m_2}(\lambda, \phi) - \right. \\ &- \left. \frac{t^{n_2 m_2}}{[n_2(n_2+1)]^{1/2}} \frac{1}{\cos\phi} \frac{\partial}{\partial \lambda} Y_{n_2 m_2}(\lambda, \phi) \right] + \\ &+ \left. \frac{\partial \varrho}{\partial r} [r^{n_2 m_2} Y_{n_2 m_2}(\lambda, \phi)] \right\} d^3\mathbf{x} . \end{aligned} \tag{5}$$

In Eq. (5) the first term follows from the first term of (3), while the other ones come from the second part of Eq. (3). On the other hand the displacement field of the tidal potential assuming a radial mass density distribution of a symmetric non-rotating elastic isotropic (SNREI) Earth could be represented according to the Love-Shida hypothesis as

$$\begin{aligned} d_\lambda &= l(r)(\gamma r\cos\phi)^{-1}\partial V_{\rm tid}/\partial\lambda \\ d_\phi &= l(r)(\gamma r)^{-1}\partial V_{\rm tid}/\partial\phi \\ d_r &= h(r)\gamma^{-1}V_{\rm tid}\,, \end{aligned} \tag{6}$$

where l and h are the Shida and Love numbers, respectively, γ is the mean gravity acceleration and $V_{\rm tid}$ denotes the tidal potential. The Fourier-transformed tidal potential in spherical surface functions is

$$V_{\rm tid} = \sum_{n=0}^{\infty}\sum_{m=-l}^{+l} V_{nm}Y_{nm}(\lambda^*,\phi^*)\,. \tag{7}$$

Therefore (6) gives

$$\begin{aligned} d_{\lambda_{nm}} &= l(r)(\gamma r\cos\phi)^{-1}V_{nm}\partial Y_{nm}(\lambda,\phi)/\partial\lambda \\ d_{\phi_{nm}} &= l(r)(\gamma r)^{-1}V_{nm}\partial Y_{nm}(\lambda,\phi)/\partial\phi \\ d_{r_{nm}} &= h(r)\gamma^{-1}V_{nm}Y_{nm}(\lambda,\phi)\,. \end{aligned} \tag{8}$$

Note that in the expressions above the summation convention is applied over repeated indices. *Hence only spheroidal, no toroidal components* are generated by the Love-Shida hypothesis, namely

$$\begin{aligned} r^{nm}(r) &= h_n(r)\gamma^{-1}V_{nm} \\ s^{nm}(r) &= l_n(r)[n(n+1)]^{1/2}\gamma^{-1}V_{nm} \\ t^{nm}(r) &= 0\,. \end{aligned} \tag{9}$$

If a SNREI model is used, the Love and Shida numbers are only regarded as radial functions, i.e. $\varrho=\varrho(r)$, $l=l(r)$ and $h=h(r)$, while $\partial\varrho/\partial\lambda=\partial\varrho/\partial\phi=0$, $\partial l/\partial\lambda=\partial l/\partial\phi=0$ and $\partial h/\partial\lambda=\partial h/\partial\phi=0$. We are left with a special form of Eq. (5), namely for an arbitrary surface point

$$\begin{aligned} &\delta w_{nm}(\mathbf{x}^*t) = \\ &\quad -G\int\Big\{\{\rho(r)(1/r^*)[1/(2n_1+1)](r/r^*)^{n_1}Y_{n_1m_1}(\lambda^*,\phi^*)Y_{n_1m_1}(\lambda,\phi)\}\cdot \\ &\quad\cdot\Big\{\Big[2/(\gamma r)h_n(r)+(1/\gamma)\frac{\partial h_n(r)}{\partial r}-n_2(n_2+1)/(\gamma r)l_n(r)\Big]\cdot \\ &\quad\cdot V_{n_2m_2}Y_{n_2m_2}(\lambda,\phi)+[(1/\gamma)\partial\varrho(r)/\partial r h_n(r))^{n_2}V_{n_2m_2}Y_{n_2m_2}(\lambda,\phi)]\Big\}\Big\}d^3\mathbf{x}\,. \end{aligned}$$

The use of orthonormality relations of the surface spherical harmonics and with (9) the above formula give

$$\begin{aligned}\delta w_{nm}(\mathbf{x}^*,t) &= -4\Pi G\cdot\int_0^{r^*}\left\{\frac{\varrho(r)}{r^*}\left[\frac{1}{2n+1}Y_{nm}(\lambda^*\phi^*)\right]\cdot\right.\\
&\cdot\left\{\left[\frac{2}{\gamma r}h_n(r)+\frac{1}{\gamma}\frac{dh_n(r)}{dr}-\frac{n(n+1)}{\gamma r}l_n(r)\right]V_{nm}+\right.\\
&+\left.\left.\left[\frac{d\varrho(r)}{dr}\cdot\frac{1}{\gamma}h_nV_{nm}\right]\right\}\right\}\cdot\left(\frac{r}{r^*}\right)^n r^2dr=\\
&=\frac{-4\pi GY_{nm}(\lambda^*,\phi^*)V_{nm}}{\gamma r^*(2n+1)}\int_0^{r^*}\left\{\varrho(r)\left[2\frac{h_n(r)}{r}+\frac{dh_n(r)}{dr}-\right.\right.\\
&\left.\left.-\frac{n(n+1)l_n(r)}{r}\right]+h_n(r)\frac{d\varrho(r)}{dr}\right\}\left(\frac{r}{r^*}\right)^n r^2dr\,. \qquad (10)\end{aligned}$$

With some modifications Eq. (10) leads to the following expression

$$\begin{aligned}\delta w_{nm}(\mathbf{x}^*,t) &= -4\Pi G/[\gamma(2n+1)(r^*)^{n+1}]V_{nm}Y_{nm}(\lambda^*,\phi^*)\cdot\\
&\cdot\int_0^{r^*}\left\{\varrho(r)\left[\frac{dh_n(r)}{dr}r^{n+2}+2h_n(r)r^{n+1}-\right.\right.\\
&\left.\left.-n(n+1)l_n(r)r^{n+1}\right]+\frac{d\varrho(r)}{dr}h_n(r)r^{n+2}\right\}dr\,. \qquad (11)\end{aligned}$$

The partial integration of the last term on r.h.s of (11) gives

$$\int_0^{r^*}\frac{d\varrho(r)}{dr}h_n(r)r^{n+2}dr=-\int_0^{r^*}\varrho(r)\left[\frac{dh_n(r)}{dr}r^{n+2}+h_n(r)(n+2)r^{n+1}\right]dr$$

because

$$\int^{r+\varepsilon} 0\frac{d(\varrho(r)r^{n+2}h(r))}{dr}dr=0$$

for every $\varepsilon > o$. This problem is discussed in detail by Grafarend et al. (1996).

Therefore Eq. (1) can be given in simpler form:

$$\begin{aligned}\delta w_{nm}(\mathbf{x}^*,t) &= 4\pi G/\left[\gamma(2n+1)(r^*)^{n+1}\right]V_{nm}Y_{nm}(\lambda^*,\phi^*)\cdot\\
&\cdot\int_0^{r^*}\varrho(r)\left[n(n+1)l_n(r)r^{n+1}+nh_n(r)r^{n+1}\right]dr\,. \qquad (12)\end{aligned}$$

This equation describes the relation between the potential variation due to deformation (r.h.s.) and the change in gravity potential (l.h.s.) what can be given in following form on the basis of Eq. (7)

$$\delta w_{nm}(\mathbf{x}^*, t) = k_n(r^*) V_{\mathrm{tid}_{nm}} = k_n(r^*) V_{nm} Y_{nm}(\lambda^*, \phi^*), \tag{13}$$

where $k(r)$ is the Love number to describe the potential variations due to the deformations. Equations (12) and (13) lead to a relation between the Love-Shida numbers

$$\begin{aligned} k_n(r^*) &= 4\Pi G n / \left[\gamma(2n+1)(r^*)^{n+1}\right] \cdot \\ &\cdot \int_0^{r^*} -\varrho(r) r^{n+1} [(n+1) l_n(r) + h_n(r)] dr . \end{aligned} \tag{14}$$

To carry out numerical calculations — as it is usual in the case of calculation of the Love-Shida numbers — a relative system of the units was introduced in (14), where the unity of the distance is the radius of the Earth (r^*) and of the gravity the mean gravity acceleration (γ). For the density the unit is the mean density of the Earth ($\varrho_{\mathrm{mean}} = 5.517$ g$\cdot$ cm^{-3}). Evidently in this system $G = 3/4\Pi$. This way Eq. (14) can be written as

$$k_n(r^*) = \frac{3}{2n+1} \int_0^{r^*} \varrho(r) \cdot r^{n+1} [(n+1) l_n(r) + h_n(r)] dr . \tag{15}$$

In case of $n = 2$

$$k_2(r^*) = \frac{6}{5} \int_0^{r^*} \varrho(r) \cdot r^3 (3 l_2(r) + h_2(r)] dr . \tag{16}$$

For the numerical calculation of the Love-Shida numbers for the case of the PREM (Dziewonski and Anderson 1981) was used with a system of sixth order differential equations given by Molodensky (1953). The setep of integration was 10^{-3} (~ 6.4 km). The variation of the second order Love-Shida numbers together with their derivatives along the radius of the Earth are shown in Table I.

For the orders $n = 2 - 10$ the surface values are listed in Table II. Due to the fact that (14) was obtained without any consideration of the boundary conditions at the surface of the Earth, at the core-mantle boundary or at the centre of our planet it is valid for the other types of the spheroidal Love-Shida numbers, too. The numerical values of the load numbers for the orders $n = 1 - 1000$ are listed in Table II.

Results of the calculations with the use of (16) gives $k_2(r^*) = 0.29296$. In Eqs 12 and 14–16 the first term of the r.h.s. is responsible for the horizontal (H_o) and the second for the vertical (V_e) displacements. They input into the numerical value of $k_2(r^*)$ is almost equal: 0.1471 and 0.1458 respectively. In case of the load numbers the ratio H_o/V_e is significantly different from 1:

$$\begin{aligned} \left.\frac{H_o}{V_e}\right|_{n=2} &= \tfrac{-0.1106}{-0.1862} = 0.594 \\ \left.\frac{H_o}{V_e}\right|_{n=10} &= \tfrac{-0.0054}{-0.0546} = 0.099 \end{aligned}$$

Table I. Tidal Love-Shida numbers of the order $n = 2$ and their derivatives along the radius

r/r^*	l	dl/dr	h	dh/dr	$k+1$	dk/dr
0.00	0.0000	−0.0621	0.0000	−0.0360	0.0000	0.0521
0.05	0.0537	−0.5412	−0.0887	−0.0400	0.0033	0.0171
0.10	0.1942	−1.7003	−0.2533	0.2453	0.0066	0.0319
0.15	0.2833	−2.0011	−0.4837	1.0652	0.0125	0.0721
0.20	0.3707	−2.1051	−0.6315	7.2051	0.0219	0.8754
0.25	0.4137	−2.7061	−0.8735	29.7740	0.0342	1.9061
0.30	0.3836	−3.0652	−1.0210	30.0072	0.0541	3.2399
0.35	0.2832	−1.5618	−0.0870	12.6395	0.1788	2.4262
0.40	0.2099	−0.8268	0.3195	5.8140	0.2796	2.1096
0.45	0.1746	−0.4361	0.4957	2.6784	0.3712	1.9993
0.50	0.1602	−0.2210	0.5767	1.2537	0.4609	1.9789
0.55	0.1483	0.0917	0.6115	0.8937	0.5486	1.9900
0.55	0.1483	−0.0934	0.6115	0.3040	0.5486	0.7745
0.60	0.1448	−0.0463	0.6215	0.1190	0.5984	1.0487
0.65	0.1431	−0.0273	0.6250	0.0319	0.6257	1.2772
0.70	0.1418	−0.0311	0.6254	−0.0113	0.7216	1.4767
0.75	0.1396	−0.0544	0.6241	−0.0367	0.8001	1.6577
0.80	0.1358	−0.0959	0.6218	−0.0576	0.8872	1.8250
0.85	0.1293	−0.1561	0.6138	−0.0810	0.9824	1.9822
0.90	0.1193	−0.1923	0.6134	−0.1228	1.0846	2.0073
0.95	0.1074	−0.2648	0.6061	−0.1711	1.1870	2.0631
1.00	0.0839	−0.4255	0.5959	−0.2338	1.2929	2.0174

Table II. Love-Shida numbers at the surface of the Earth for $n = 2 - 10$

n	$h(r^*)$	$k(r^*)$	$l(r^*)$
2	0.5959	0.2929	0.0839
3	0.2936	0.1007	0.0160
4	0.1817	0.0514	0.0113
5	0.1354	0.0344	0.0096
6	0.1131	0.0269	0.0080
7	0.1004	0.0228	0.0066
8	0.0921	0.0203	0.0056
9	0.0860	0.0183	0.0048
10	0.0813	0.0171	0.0041

For the study of mutual the dependence of the Love-Shida and of the load numbers on the inner structure of the Earth the numerical values of $h_2(r^*)$ and $l_2(r)$ or of $h_n'(r)$ and $l_n'(r)(n = 2, 10)$ were reduced by 30 percent (what is a high disturbance) in a spherical layer of thickness 0.1 and this layer was moved from the centre to the surface with a step 0.1. Numerical results obtained for the case $n = 2$ both in case of Love-Shida and of load numbers are almost independent from the

Table III. Load Love-Shida numbers at the surface of the Earth for $n = 1 - 1000$

n	$h'(r^*)$	$k'(r^*)$	$l'(r^*)$
1	−0.2966	0	0.1123
2	−0.9841	−0.2969	0.0245
3	−1.0430	−0.1869	0.0715
4	−1.0458	−0.1245	0.0606
5	−1.0774	−0.0952	0.0481
6	−1.1339	−0.0806	0.0405
7	−1.2010	−0.0719	0.0361
8	−1.2710	−0.0661	0.0333
9	−1.3409	−0.0622	0.0314
10	−1.4104	−0.0600	0.0301
100	−2.9692	−0.0147	0.0078
1000	−5.0055	−0.0022	0.0008

Table IV. Dependence of $k_2(r^*)$ and its components on the disturbances by 30 percent of $h_2(r)$ and $l_2(r)$ at different levels. Undisturbed values: $k_2(r^*) = 0.29296$, $H_o = 0.1471$, $V_e = 0.1458$

Disturbed layer	$k_2(r^*)$	H_o Horizontal component	V_e Vertical component	Ho/V_e
0.0–0.1	0.29295	0.14714	0.14582	1.009
0.1–0.2	0.29303	0.14703	0.14600	1.007
0.2–0.3	0.29318	0.14616	0.14782	0.991
0.3–0.4	0.29192	0.14537	0.14655	0.992
0.4–0.5	0.28600	0.14336	0.14264	1.005
0.5–0.6	0.28339	0.14236	0.14103	1.010
0.6–0.7	0.28110	0.14128	0.13982	1.011
0.7–0.8	0.27621	0.13898	0.13723	1.013
0.8–0.9	0.27167	0.13718	0.13450	1.020
0.9–1.0	0.27219	0.13831	0.13388	1.033

disturbances (Tables IV and V) introduced, what means, from the adopted model of the Earth, too. The largest deviation from the undisturbed value is ~ 7 percent. In case of $n = 10$ order load (Table VI) $(k'_{10}(r^*))$ the near surface disturbance is the biggest and its deviation from the undisturbed value is ~ 23 percent.

Finally it is remarkable that there are different equations to describe relations between different classes of the spheroidal Love-Shida numbers (Molodensky 1977, Saito 1978, Varga 1983, Okubo and Saito 1983, Merriam 1985).

The relation obtained in this paper — as it was mentioned — describes relation of Love numbers of the same type and valid for SNREI models of the Earth, and

Table V. Dependence of $k_2'(r^*)$ and its components on the distrubances by 30 percent of $h_2'(r)$ and $l_2'(r)$ at different levels. Undisturbed values: $k_2'(r^*) = 0.29675$, $H_o = 0.1106$, $V_e = -0.1862$

Disturbed layer	$k_2'(r^*)$	H_o Horizontal component	V_e Vertical component	Ho/V_e
0.0–0.1	−0.29675	−0.11057	−0.18618	0.593
0.1–0.2	−0.29651	−0.11016	−0.18635	0.591
0.2–0.3	−0.29583	−0.10846	−0.18737	0.579
0.3–0.4	−0.29386	−0.10696	−0.18689	0.5723
0.4–0.5	−0.28898	−0.10594	−0.18298	0.579
0.5–0.6	−0.28697	−0.10588	−0.18110	0.585
0.6–0.7	−0.28467	−0.10532	−0.17559	0.593
0.7–0.8	−0.27998	−0.10419	−0.17559	0.593
0.8–0.9	−0.27661	−0.10515	−0.17146	0.613
0.9–1.0	−0.27831	−0.10997	−0.16840	0.6580

Table VI. Dependence of $k_{10}'(r^*)$ and its components on the distrubances by 30 percent of $h_{10}'(r)$ and $l_{10}'(r)$ at different levels. Undisturbed values: $k_{10}'(r^*) = -0.06004$, $H_o = -0.00543$, $V_e = -0.05461$

Disturbed layer	$k_{10}'(r^*)$	H_o Horizontal component	V_e Vertical component	Ho/V_e
0.0–0.1	−0.06004	−0.00543	−0.05461	0.100
0.1–0.2	−0.06004	−0.00543	−0.05461	0.100
0.2–0.3	−0.06004	−0.00543	−0.05461	0.100
0.3–0.4	−0.06004	−0.00543	−0.05461	0.100
0.4–0.5	−0.06003	−0.00539	−0.05463	0.099
0.5–0.6	−0.06003	−0.00531	−0.05464	0.097
0.6–0.7	−0.05995	−0.00532	−0.05463	0.097
0.7–0.8	−0.05950	−0.00514	−0.05436	0.096
0.8–0.9	−0.05628	−0.00396	−0.05232	0.075
0.9–1.0	−0.04650	−0.00586	−0.04064	0.144

therefore it is a relative of the Eq. (6) given by Love in 1909. It is the only relation — in contrary to the above mentioned ones — which does not follow from the sixth-order differential equation system of motion usually used to calculate the Love-Shida numbers.

Acknowledgements

Research work described in this work was supported by the Deutsche Forschungsgemeinschaft and the Hungarian Academy of Sciences in frame of the bilateral project "Geodynamische Untersuchungen mit Hilfe der generalisierten Love-Shida Hypothese".

Appendix

Scalar and vector spherical functions

1. Scalar spherical functions

$$Y_{nm}(\lambda,\phi) :=$$
$$\frac{(2n+1)^{\frac{1}{2}}}{\varepsilon_m}\left(\frac{(n-|m|)!}{(n+|m|)!}\right)^{\frac{1}{2}} P_{nm}(\sin\phi)\begin{bmatrix}\cos m\lambda\,\forall m=0,1,\ldots,n-1,n\\ \sin|m|\lambda\,\forall m=-n,-n+1,\ldots,-2,-1\end{bmatrix}$$

$$\langle Y_{n_1m_1}|Y_{n_2m_2}\rangle := \frac{1}{4\Pi}\int\limits_0^{2\Pi} d\lambda \int\limits_{-\Pi/2}^{+\Pi/2} d\phi\cos\phi Y_{n_1m_1}(\lambda,\phi)Y_{n_2m_2}(\lambda,\phi)=\delta_{n_1n_2}\delta_{m_1m_2}$$

(orthonomality)

$$\varepsilon_m := \begin{bmatrix} 1 & \forall m=0\\ 1/2^{1/2} & \forall m\neq 0\end{bmatrix}.$$

2. Vector spherical functions

$$\left.\begin{array}{lcl} R_{nm} & := & e_r Y_{nm}(\lambda,\phi)\\ S_{nm} & := & e_\lambda\frac{1}{[n(n+1)]^{\frac{1}{2}}}\frac{1}{\cos\phi}\frac{\partial}{\partial\lambda}Y_{nm}(\lambda,\phi)+e_\phi\frac{1}{[n(n+1)]^{\frac{1}{2}}}\frac{\partial}{\partial\phi}Y_{nm}(\lambda,\phi)\\ T_{nm} & := & e_\lambda\frac{1}{[n(n+1)]^{\frac{1}{2}}}\frac{\partial}{\partial\phi}Y_{nm}(\lambda,\phi)-e_\phi\frac{1}{[n(n+1)]^{\frac{1}{2}}}\frac{1}{\cos\phi}\frac{\partial}{\partial\lambda}Y_{nm}(\lambda,\phi)\end{array}\right\}$$

$$\begin{array}{lclclcl} \langle R_{nm}|R_{n'm'}\rangle & = & \langle S_{nm}|S_{n'm'}\rangle & = & \langle T_{nm}|T_{n'm'}\rangle & = & \delta_{nn'},\delta_{mm'}\\ \langle R_{nm}|S_{n'm'}\rangle & = & \langle R_{nm}|T_{n'm'}\rangle & = & \langle S_{nm}|T_{nm}\rangle & = & 0\end{array}$$

$$\left.\begin{array}{lcl} \text{rot } R_{nm} & = & \frac{1}{r}[n(n+1)]^{\frac{1}{2}}T_{nm}\\ \text{rot } S_{nm} & = & -\frac{1}{r}T_{nm}\\ \text{rot } T_{nm} & = & \frac{1}{r}S_{nm}+\frac{1}{r}[n(n+1)]^{\frac{1}{2}}R_{nm}\end{array}\right\}$$

$$\left.\begin{array}{lcl} \text{div } R_{nm} & = & \frac{2}{r}Y_{nm}(\lambda,\phi)\\ \text{div } S_{nm} & = & -\frac{1}{r}[n(n+1)]^{\frac{1}{2}}Y_{nm}(\lambda,\phi)\\ \text{div } T_{nm} & = & 0\end{array}\right\}$$

References

Dziewonski A, Anderson D 1981: *Phys. Earth Planet. Int.*, 25, 297–356.
Grafarend E 1990: The spacetime height system. Proceedings of the Workshop on Precise Positioning, Hannover

Grafarend E, Engels J, Varga P 1996: The spacetime gravitational field of a deformable body. *Journal of Geodesy* (in press)

Love A E H 1909: *Proc. R. Soc.*, 32, 73–88.

Merriam J B 1985: *J. Geoph. Res.*, 99, 7795–7802.

Molodensky M S 1953: Elastic tides, free nutations and some questions concerning the internal structure of the Earth (in Russian). Trudy Geofis. Inst. Akad. Nauk. SSSR, 19(146), 3–42.

Molodensky S M 1977: *Phys. Solid Earth*, 13, 147–149.

Okubo S, Saito M 1983: *Bull. Géod.*, 57, 167–179.

Saito M 1978: *J. Phys. Earth.*, 26, 13–16.

Varga P 1983: *Manuscripta Geodaetica*, 8, 85–91.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 293–303 (1996)

HIGH PRECISION GEOID EVALUATION FOR GERMANY — GEOID AND QUASIGEOID

E GROTEN[1]

The evaluation of high precision geoids is discussed in view of the 1995 quasigeoid computed for Germany (Belikov and Groten 1995). Error sources existing in computer-oriented modern techniques associated with digital terrain models etc. are outlined and over-optimistic accuracy estimates for single points are analyzed.

Keywords: digital terrain model; geoid; Germany; gravity field; quasigeoid

1. Introduction

In the recent past great M S Molodensky has died and there are only a few people, such as Yurkina, Ostach and others, are left of that group which founded the quasigeoid theory and its application. They also changed mostly their focus of interest. Petrovskaya is still working on that field.

There are also otherwise changes: We still wait for satellite gradiometry where global high-resolution gravity field components would be detected. Airborne gravimetry strongly gained impact with GPS being available for Eötvös correction evaluation and also positioning.

Moreover, the role of physical geodesy has changed: On the one side, in unifying Vertical Datums, we need absolute global geoids where the N_0-term, i.e. the zero degree term in the spherical harmonics series of the geoid height, N, plays a dominant role which depends primarily on GM = the geocentric gravity constant and the volume of the earth which is now much better evaluated from global satellite stations in a unified reference frame. GPS contributed substantially to that determination. Bursa (Bursa et al. 1995) proposed to replace the semimajor axis, a_1, of the terrestrial ellipsoid by the ratio W^0 over GM where W^0 is the potential at the geoid; we call the ratio R_0 which is independent of permanent tide and also those quantities which are seldom fully understood by non-experts and lead to a lot of confusion of users of global geoids such as astrometrists etc. which usually cannot clearly separate zero-tide, mean-tide, tide-free from other geoids and from each other.

The other aspect is that engineers, cadaster and other surveyors now need to convert GPS-heights into orthometric heights so that very precise relative and/or local geoids are needed. GPS-height determination becomes competitive to leveling for distances such as 10 km or so, thus subcentimeter accuracy is desired.

Numerical possibilities and large data bases for digital height models have completely changed the earlier situation. New demands, new applications and consequent computational aspects will be discussed below.

[1]Technische Hochschule Darmstadt, Institut für Physikalische Geodäsie, Petersenstrasse 13, D–64287 Darmstadt, Germany, e-mail: groten@ipgs.ipg.verm.th-darmstadt.de

2. A review of Molodensky's concept

Moritz and, before him, Arnold had basically changed the original Molodensky quasigeoid theory into a geoid-oriented determination of the potential field. Molodensky's normal height, h', i.e. the deviation of the telluroid $U(Q) = W(P)$ from the ellipsoid where $W(P)$ is the geopotential at the earth's surface, is a useful purely geometric quantity in that concept of approximating the earth's surface but has no physical meaning as the orthometric height, h, as deviation of the earth's surface from the geoid or N.N. The problem of a global Vertical Datum inherent in $W^0 = U^0$ (see Appendix, where W^0 = geoidal potential, U^0 = normal potential at the ellipsoid) is now of greatest interest in ecological, environmental and global geodynamic studies. Note that over long intervals W^0 can no longer be cosidered as a temporal constant. Molodensky attempted to approximate the physical meaning of the geoid by introducing the well-known quasigeoid without implementing in this way a level surface. With millimeter discussion now in connection with MSL (= mean sea level) studies the extent of approximation of the quasi-geoid to the geoid is no longer of interest. This is also true in view of the high accuracy of VLBI, SLR and GPS. Insofar, Moritz's and Arnold's geoidal concepts become new more important than before. We need physical not abstract geometrical quantities.

In the mid-seventies it was already discussed at Institut für Physikalische Geodäsie (IPG) that without detailed terrain models Molodensky's G_i-corrections to Stokes' formula cannot be implemented in practice. Moreover, in various studies it was clearly demonstrated that even in hill areas G_i-type corrections in the linear approximation of Molodensky's concept can hardly be evaluated. It would be needed in steep topography to have gravity measurements in the neighbourhood zone at distances of 20–50 m all over. This will never be possible in a world-wide scale. Also for satellite gradiometry it will not be feasible to apply such local corrections all over. Thus G_i-type corrections arise not only in Alpine but also in hill areas if we want to locally guarantee high precision, too. Heck's studies of non-linearities of Molodensky's concept and the investigations of others are not yet entering practical evaluation of geoid or quasigeoid heights. Thus the comparisons of different geoids are only partial comparisons.

Under the influence of modern computational possibilities it became fashionable (BGI/IGES) to go back to Fourier techniques applied by Tsuboi, Jung, Grafarend and myself (Groten 1966) into geodesy and combine them with the FFT-technique introduced in 1963 into spectral analysis. By combining such techniques with spherical harmonic analysis, which dates back to Vening-Meinesz and predecessors, we end up with the numerous new papers presently published by Canadian, Danish and other groups (Sneeuw and Bun 1966). As a computer-oriented version of a linear Molodensky concept it is certainly worthwhile to take these papers into account. More important is, however, the progress going on in DHM (digital height models) now available globally in 5×5 km and regionally in smaller blocks. We replace in this way the actual topography by stepwise topography of mean elevations. Nevertheless, if we take 40×40 m blocks, as available in Germany, we see a tremendous increase in computational work. We have found out in our new geoid 1995 for Germany that by using soon available 20 × 20 m blocks an almost incredible increase is faced in computational work which is no longer economic.

The use of DHM for gravity interpolation in areas such as the Odenwald was carefully tested by us in numerous studies (Euler et al. 1985, 1986). In principle, it is the only technique as a cross-correlation technique, taking into account also density variations in the underground, to implement economically Molodensky's approach. Superior to FFT techniques, however, appears today, with modern computers at hand, the opposite way where spherical functions are applied locally, instead of, as in case of FFT, using flat techniques globally. Already a decade ago in geoid computations in a cooperation between IPG and Freeden's group (Freeden et al. 1995) the regional application of spherical function systems were used and successfully applied. Wavelets can also be used to superimpose local fields to global satellite derived gravity and other fields. Presently two-dimensional flat wavelets are applied by us to represent local fields. Freeden in a series of papers presented regional gravity fields using two-dimensional spherical wavelets. This techniques appears to be an optimal method in evaluating different superimposed fields, such as the G_i-corrections and terrestrial field components in combination with global spherical harmonics expansions of the gravity field deduced from satellite orbit analysis or from satellite altimetry.

3. Modern determination of geoid and quasigeoid

More than ten years ago, at a Symposium organized by Professor Chen at Beijing on precise determination of the gravity field we were asked to investigate refined corrections to geoid and deflection determination. At this meeting it was also shown (Groten 1984) that refinement by applying atmospheric corrections, using higher degree normal gravity fields etc. did not substantially increase the accuracy of present conventional geoid evaluation.

As local conventional height datums can nowadays only be connected with each other with accuracy not better than 0.2 or 0.3 m, which corresponds to 0.1 mGal, the combination of spherical harmonics determined from satellite orbit analysis with terrestrial can to a certain extent remove the impact of local vertical datum errors but not completely. This is only one aspect of a number of systematic errors inherent in presently available gravity fields. We may statistically remove various perturbing influences but the Darmstadt "fiasko" (see below) illustrates that locally not all singularities may be removed even though it is not at all clear whether this is really an error in the gravimetric geoid or, as discussed below, in the GPS-data or in the transformation process in using relative geoid and terrestrial datum.

Another source of systematic and random distortion is caused by the transition from the irregularly distributed terrestrial gravity anomaly data (free air values) to a regular grid in the integration process for solving the integrals discussed in the appendix (this is related to G_i-corrections as well as to the Stokes or spherical harmonic integrals). Fort he IPG-geoid of Germany (Groten 1995a, 1995b) Belikov selected a technique described by Belikov and Groten (1995a, 1995b) which leads to a direct transition from irregularly distributed gravity to the integration. Even though also the other (G_i-type corrections, ellipsoidal etc.) corrections were transformed into computer-oriented formulas different from the original formulas it does not appear interesting to repeat here computational details.

The biases associated with datum problems were also treated in the original publication. Still the result of our geoid 1995 of Germany is a quasigeoid. This is justified as the official heights in Germany will be in the future normal heights, h', which should be transformed by our geoid into relative ellipsoidal heights as obtained by GPS observations.

4. The geoid as a reference surface

Quasigeoid and geoid coincide by definition on the ocean, at least in neglecting ocean topography. What we need today is a physical reference for environmental studies where temporal variations such as MSL (= mean sea level), geodynamic processes of the solid earth etc. are studied. Zeman and various other investigators including myself studied temporal variations of the geoid and the gravity field, mainly based on Mathers' concepts. What appears nowadays in the literature in terms of global temporal changes of the gravity field is mainly a mixture of atmospheric, oceanic and solid earth variations. So they have to be clearly distinguished. The vertical changes of the geoid height associated with vertical changes of the earth's surface are obviously between 5 and 10 percent of the latter depending on the mechanism. When I investigated recently for Antarctica and the arctic area, Spitzbergen etc., such a deformation in connection with ice movement and melting, the importance of the geoidal reference and its variations became clear to me.

But also in areas such as the Mediterranean Sea the extremely careful determination of the geoid as a reference surface for altimetric results, as those obtained from ERS-1, Topex-Poseidon ERS-2 etc., and the modeling of ocean circulation is straightforward. Presently, the lack in accuracy in absolute and relative geoids is a primary obstacle in a clear interpretation of altimetric and tide gauge data in most part of the oceans; in Baltic, North and Mediterranean Sea still an optimal situation appears as was shown in various studies carried out at IPG.

5. The Darmstadt "fiasko"

Sir Harold Jeffreys preferred smooth and concluded "the smoother the better". Least-squares collocation as an autoregressive method is tending, on the one hand, to smooth things out; on the other hand, with the usual heterogeneous distribution of boundary data we face instabilities in autoregressive techniques. The Odenwald, south of Darmstadt, where part of our test fields are located, is an old mountain chain with a lot of erosion and almost vertically penetrating layers of basalt and granite layers of strongly varying density. Even under such unfavorable conditions gravity interpolation using DHM data led to very good results in implementing Molodensky type G_i-corrections, similarly to Pellinen's concept.

Least-squares techniques, contrary to robust methods, somehow, more or less, absorbes large single deviations. In our local evaluation for Hessen for that area single discrepancies up to 10 cm were found in a geoid which is basically accurate to 1.5 cm; this is not astonishing under these circumstances; also stepwise homogeneous terrain modelling, unless smoothed, leads to significant errors.

The comparison of gravimetric geoid heights with those determined from a combination of leveling by GPS and leveling measurements is not yet problem-free as single GPS-measurements can be strongly perturbed in height. The general comparison of GPS + leveling data with our geoid, contrarily to that single comparison, gave discrepancies always below ± 5 cm which corresponds to a 3σ-limit of our geoid, σ = standard deviation.

6. The deflections of the vertical

More than by the geoid itself are the variations of it reflected by its horizontal derivatives, i.e. by the deflections of the vertical. For three selected areas these quantities are shown in Figs 1 to 6 in the Appendix. These deflections were taken from the central part in Germany, i.e. Thüringen, where the accuracy is highest as the distance is there maximum from the boundaries of Germany where the quality of gravity anomalies and terrain data decreases. The second area of particular interest is Hessen where special tectonic areas demonstrate the relation with special features in the deflection fields. In the third area, Rheinland-Pfalz, the correlation of deflections with the lower Rhine valley, an active tectonic area, and similar features is of interest. However, there, along boundaries to Belgium, Luxembourg etc. the accuracy of gravity and, consequently, of deflections is lower.

The derivatives of deflections were interpreted by Bursa (1975) in terms of indicators for seismic activities. However in view of the problems inherent in second derivatives of the geopotential along the earth's surface such interpretations are not unproblematic, in general.

7. Conclusions

The progress in geoid evaluation depends on satellite gradiometry globally and on the improvement of airborne gravimetry locally. Moreover, the aspects have shifted to Vertical Datum and absolute geoid problems. Tremendous progress has been achieved but new methods such as spherical splines, spherical wavelet techniques could substantially contribute to this problem, besides further improvement of DHM models. In areas like Germany now centimetric geoids are possible. However, the computational costs would be tremendous. It is confusing, mainly for oceanic geoid, that J_2-terms in global models, such as the OSU 91A, are given in the "tide-free" system whereas in the real world we have the "mean" geoid whose flattening, however, is the same as in case of the "zero" geoid.

Acknowledgement

These results were first presented at the Oberwolfach Mathematical Seminar in autumn 1995 together with M Belikov; they are published here for the first time.

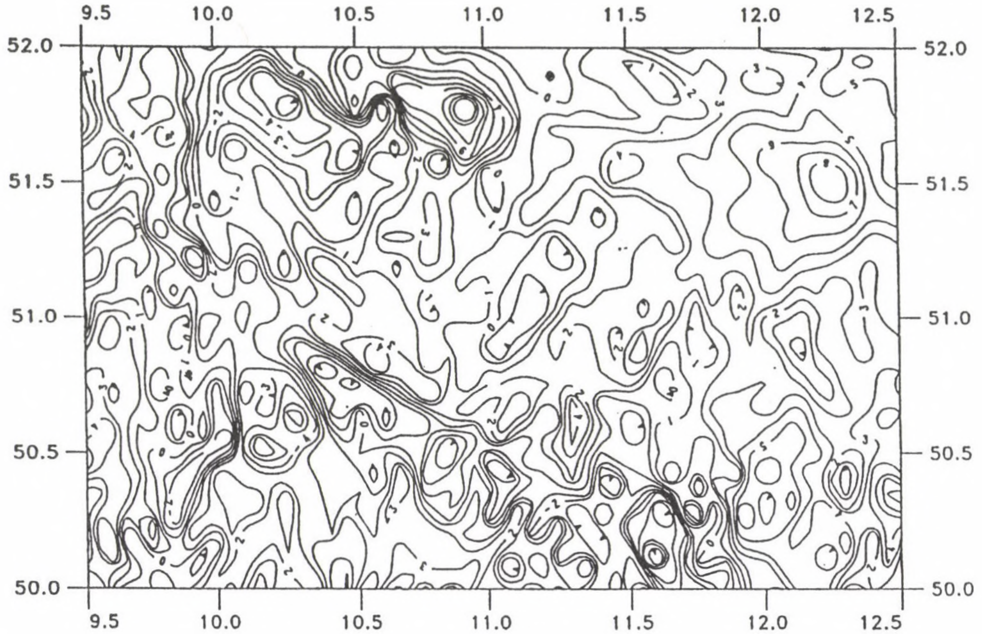


Fig. 1. Deflections of the vertical ('ETA' component) for Thüringen. Reference system: GRS 80, contour interval: 1.0 arcsec

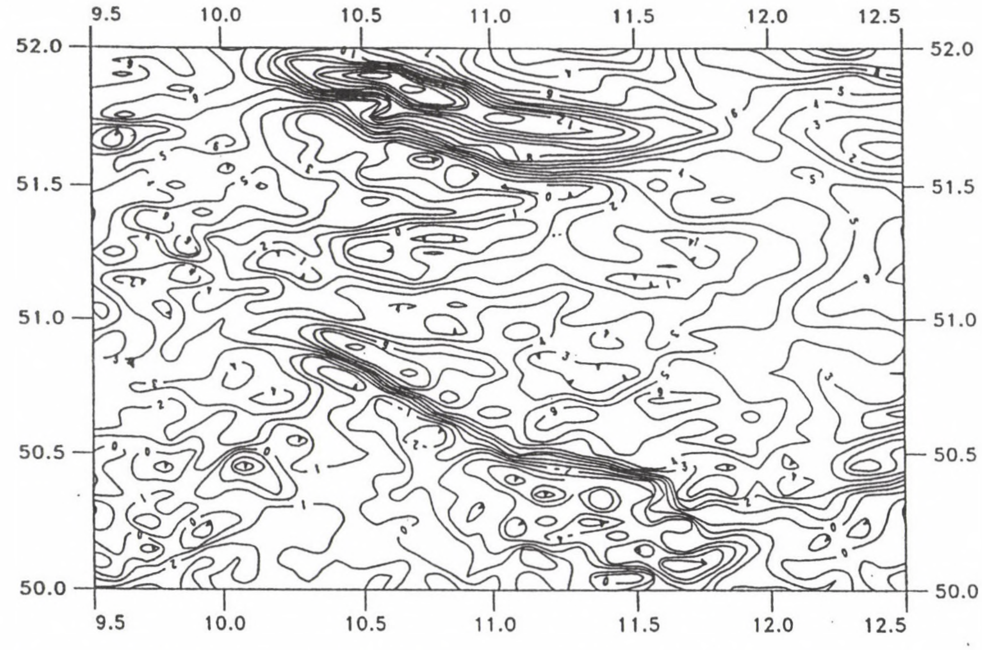


Fig. 2. Deflections of the vertical ('XI' component) for Thüringen. Reference system: GRS 80, contour interval: 1.0 arcsec

Appendix

General Aspects

$$T(P) = W(P) - U(P)$$

where $U = U(a, J_2, \omega, GM)$ and

$$N = T/\gamma$$

$$\begin{aligned} \text{if } W^0 = U^0 \text{ based on GM (actual)} &= \text{GM (ellipsoid)} \\ \text{vol. (Geoid)} &= \text{vol. (ellipsoid)} \end{aligned}$$

Here: N = geoid height, T = disturbing potential, W = gravity potential, U = normal potential, W^0 = const. = geopotential at the geoid, U^0 = normal potential at ellipsoid, γ = normal gravity, a = semimajor axis of earth's ellipsoid,

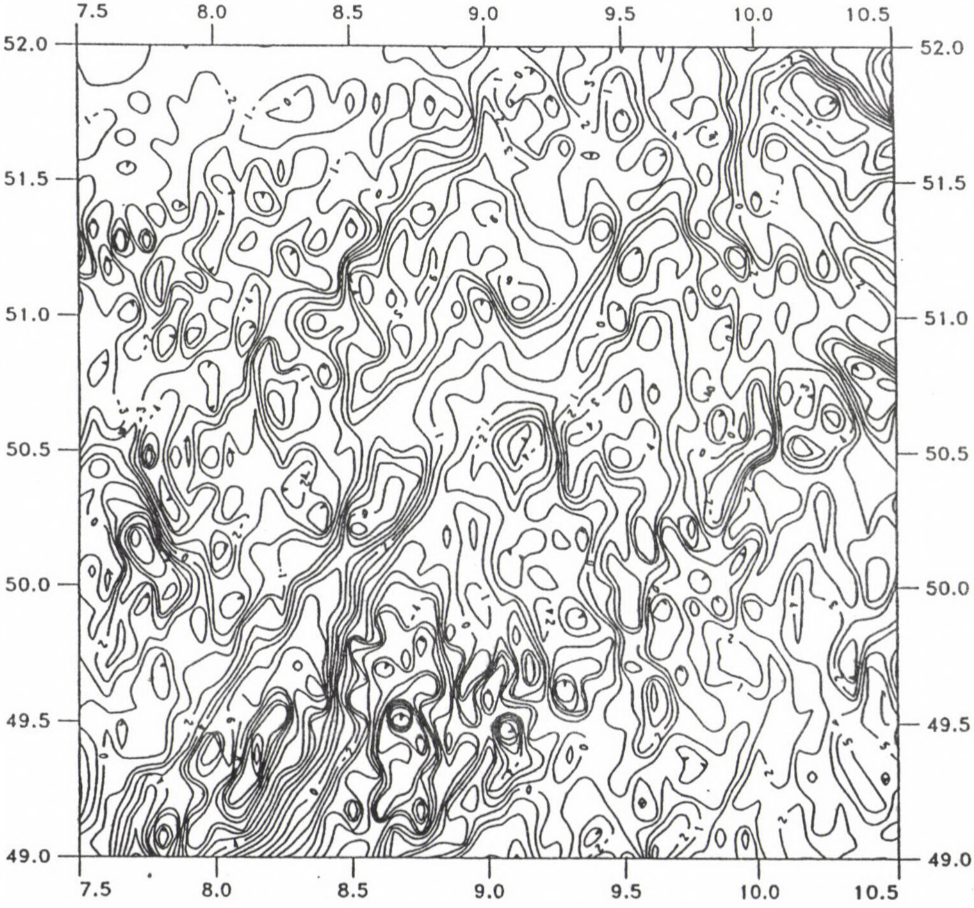


Fig. 3. Deflections of the vertical ('ETA' component) for Hessen. Reference system: GRS 80, contour interval: 1.0 arcsec

J_2 = second degree zonal harmonic of geopotential, ω = earth's rotation spin, GM = geocentric gravitation constant. BVP = boundary value problem, Δg = gravity anomaly, δg = gravity disturbance.

Stokes: A Robin-type of free BVP with boundary values at the surface of geoid, S in terms of $\Delta g(S)$.

Hotine: A Neumann-type fixed BVP with $\delta g(S)$ using GPS — or satellite and VLBI coordinates → important for the future.

Absolute Geoid: As $W^0 \neq U^0$ always in reality, an N_0 arises in N which is not contained in Stokes' theory; it is of great interest now in view of absolute geoids. In the German geoid it is of the order of 30 cm.

Tidal components: As the direct luni-solar (and planetary) attraction would lead to a non-harmonic term in T it is eliminated in tidal correction treatment of Δg and δg when now the subsequent tidal deformation potential is preserved. This

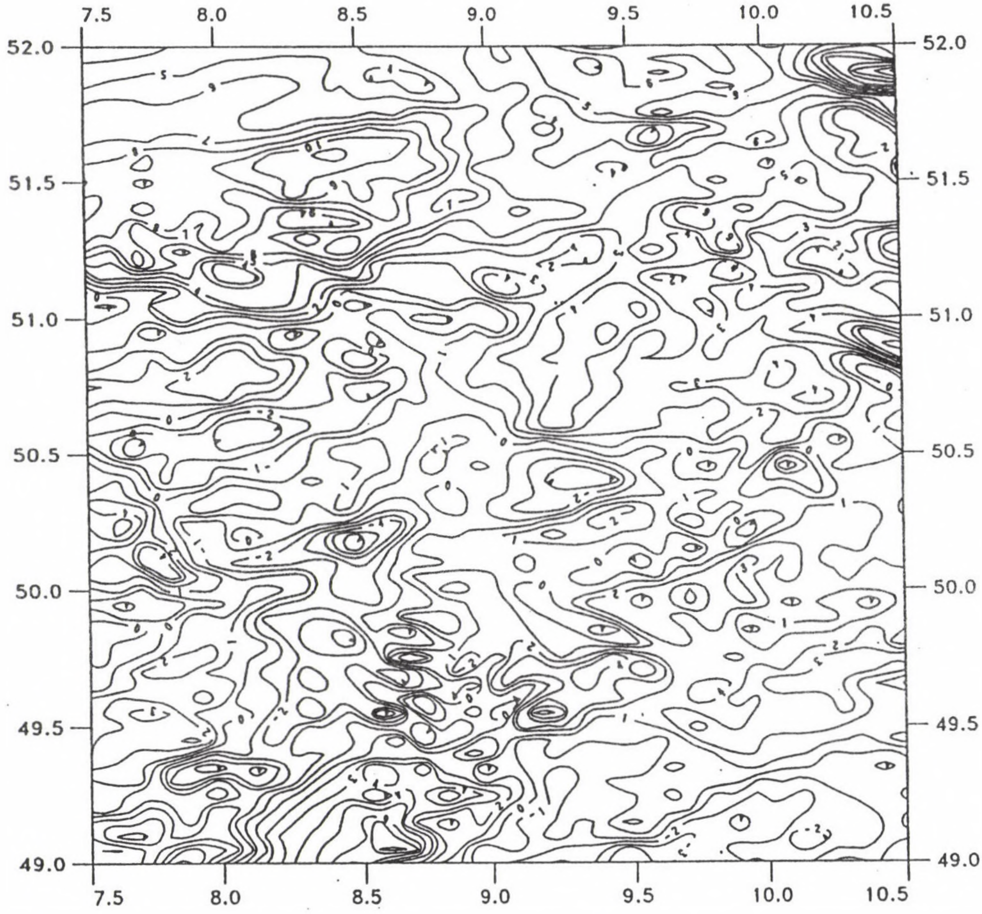


Fig. 4. Deflections of the vertical ('XI' component) for Hessen. Reference system: GRS 80, contour interval: 1.0 arcsec

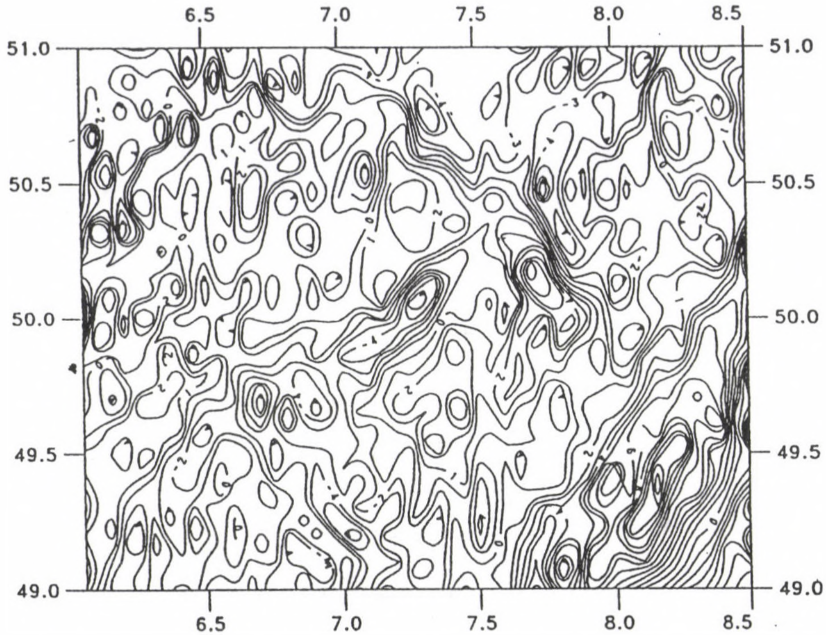


Fig. 5. Deflections of the vertical ('ETA' component) for Rheinland-Pfalz. Reference system: GRS 80, contour interval: 1.0 arcsec

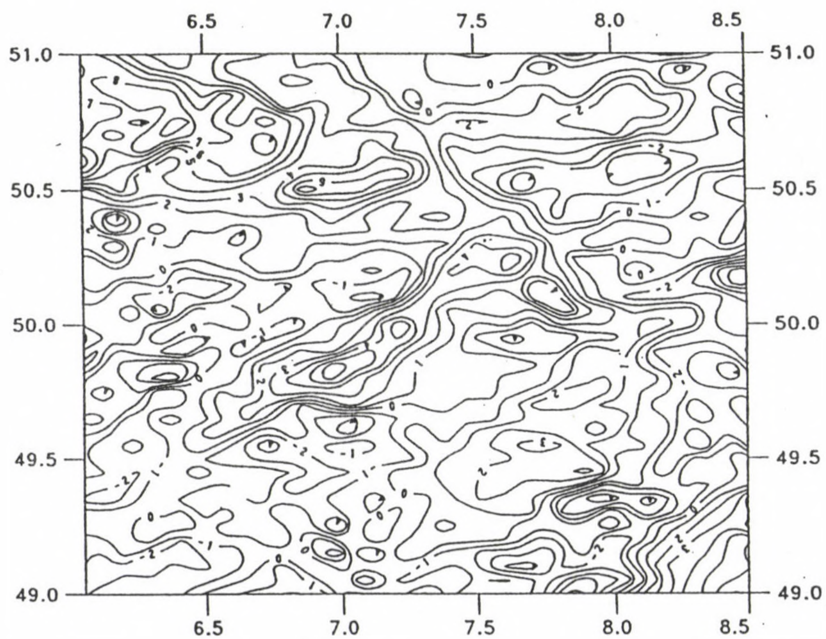


Fig. 6. Deflections of the vertical ('XI' component) for Rheinland-Pfalz. Reference system: GRS 80, contour interval: 1.0 arcsec

appears illogical and, consequently, still finds criticism and leads to confusion to non-experts who cannot clearly separate tide-free, zero-tide, mean-tide etc. geoids.

Geocentricity: The N_1 in N-expansion is well controlled due to the stability of the Stokesian problem; consequently, $N_1 \doteq 0$ can be implemented even though temporal geocentric location variations of the order of a few centimeters (ocean-, atmospheric mass shift, incl. tidal mass shift) cannot yet be completely modelled.

Computational aspects: Spherical harmonics for remote zones and integral approach for the neighbouring cap led to numerous approaches if improved convergence, based on Molodensky's original formulae, is applied. All these techniques need no further discussion.

Molodensky's concept: By converting the free into a fixed oblique BVP with respect to the telluroid $U(Q) = W(P)$ the integral equation is transformed into a Stokes type integral formula with G_i-type corrections to the kernel. By adding ellipsoidal and numerous other corrections a linear approximation may be finally useful which is, however, difficult to be implemented in practice (lack of data).

Pellinen's concept: Assuming constant crustal density the regression of free air gravity with elevation led to more or less realistic implementations of Molodensky's evaluation of geoid and deflections.

Belikov's concept: Pseudo-harmonic high frequency approximation (up to 10 km or so wavelength).

Darmstadt-geoid: for Germany, 1995: geoid at 4' intervals so that in the interior part resolution of 5 km appears possible. The relative accuracy of ± 1.5 cm decreases along the boundaries as in neighbouring countries the accuracy of the gravity field (incl. datum deviations) is different.

Modern alternative: Spherical wavelets in superimposing local to global field components.

References

Belikov M, Groten E 1995a: *AVN*, 8–9, 305–309.

Belikov M, Groten E 1995b: *AVN*, 8–9, 356–366.

Bursa M 1975: *Geofysikalni Sbornik*, 23, 431.

Bursa M, Bystrzycka B, Radej K, Vatrt V 1995: *Studia geoph. et geod.*, 39, 365–374.

Euler H, Groten E, Hausch W, Stock B 1985: Precise geoid computation at mountainous coastlines in view of vertical datum definition. Veröff. ZIPE, Akad. Wiss. Potsdam 81, Teil 1, 109–139.

Euler H, Groten E, Hausch W, Kling Th 1986: New results obtained for detailed geoid approximations. Boll. Geod. Sci. Aff., 4.

Freeden W, Schreiner M, Franke R 1995: A survey on spherical spline. Univ. Kaiserslautern

Groten E 1966: Bestimmung der Schwere an der Erdoberfläche aus Fluggravimetermessungen. DGK C94.

Groten E 1984: In: Proc. Local Gravity Field Approximation, Beijing, China, 217–274.

Groten E 1995a: Die GPS-Technik. Zukünftige Möglichkteiten und Grenzen. DVW Hessen-Thüringen, 3, 8–16.

Groten E 1995b: *AVN*, 8-9, 336–342.

Sneeuw N, Bun R 1996: *Journal of Geodesy*, 70, 224–232.

New Geoids in the World, Intern. Assoc. Geodesy, BGI-IGES, Special Issue, Bull. d'Inf. 77, IGES Bull. 4, Toulouse

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 305–327 (1996)

GRAVITY AND THE EARTH'S GLOBAL STRUCTURE AND DYNAMICS

J HINDERER[1]

This paper is devoted to the study of the Earth's changes in surface gravity in relation with global dynamics, either of deep origin or located in the atmosphere or oceans. A Love number formalism allows to easily express these changes with the help of generalised gravimetric factors. We investigate the wide spectral range which is observable with superconducting gravimeters, from the seismic frequency band to periods larger than one year. We especially pay attention to the gravity detection of core modes in the liquid core and to the Slichter mode of translation of the solid inner core. In the tidal bands, we show how accurate measurements provide a useful information on various phenomena like mantle (an-)elasticity, ocean and atmospheric loading. In addition, we show that the observation of the Free Core Nutation resonance in diurnal tides indirectly suggests an increase in the ellipticity of the core-mantle boundary with respect to its hydrostatic value. A similar resonance is theoretically predicted for the rotation of the inner core (Free Inner Core Nutation) but we show that its detection will be much more difficult to achieve than the previous one. Finally, we also present results indicating the ability of the superconducting gravimeters to observe the gravity effect due to the Earth's polar motion because of their small long-term instrumental drift.

Keywords: calibration; resonance; superconducting gravimetry; tides

1. Introduction

In this study we are primarily concerned with gravity changes which can be observed at the Earth's surface with relative gravimeters. Special emphasis will be given to superconducting gravimeters (SG) because of their large sensitivity and low instrumental long term drift which allows to explore a wide spectral range ranging from the seismic frequency band (periods less than 1 hour) to the Chandlerian band (periods above 1 year). Figure 1 from Crossley and Hinderer (1995) indicates the observable gravity spectrum from 1 sec to 435 days. The amplitudes are given here in a normalised form where a pure harmonic signal of unit amplitude leads to a delta function of unit amplitude at the frequency of the wave. The peak to the left is the noise generated by the oceans and is the dominant feature on the high frequency part of seismometers and gravimeters (see e.g. Peterson 1993). Below one hour (more precisely below 54 min which is the period of the fundamental ${}_0S_2$ spheroidal mode) we have the elastic normal mode band of the Earth which are excited by major earthquakes. For periods larger than 1 hour, we enter the so-called subseismic band (see Smylie and Rochester 1981) where are located the gravity-inertial modes of the fluid core and the solid inner core; notice that, on the contrary to the classical normal modes, the restoring force is here predominantly provided by

[1]Ecole et Observatoire de Physique du Globe,5, rue Descartes, 67084 Strasbourg, France, e-mail: jacques@pallas.u-strasbg.fr

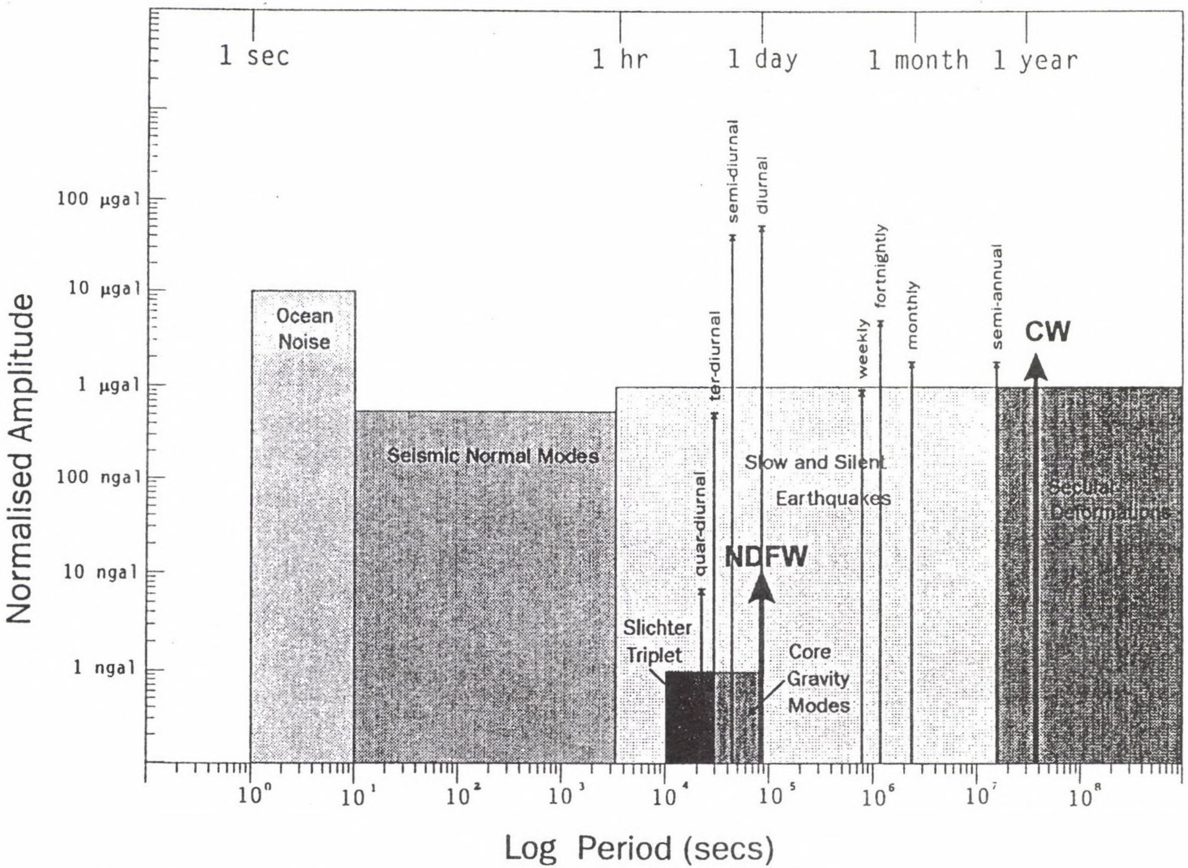


Fig. 1. Surface gravity spectrum (from Crossley and Hinderer 1995) indicating the wide spectral range observable with superconducting gravimeters

gravity contrast (Archimedian) or rotation (Coriolis) rather than elasticity. A very important mode called the Slichter mode describing the translational motion of the solid inner core belongs to this part of the spectrum. Towards lower frequencies we reach then the tidal frequency bands from 6 hours (quarter-diurnal tides) to 1 year (S_a) including the large diurnal and semi-diurnal tides. As we will see, there is a resonance effect in the diurnal part due to the presence of a free rotational mode of the fluid core which can be observed in gravimetry. At the right end of the spectrum, we have the Chandler period which corresponds to a free mode of the Earth's rotation. It has a period of about 435 days and its signature in gravity has been revealed with good quality SG records. We also have to mention here that many other phenomena can be investigated with gravity observations: atmospheric contributions (meteorological continuum, planetary scale waves), oceanic contributions (especially at tidal frequencies but also free modes), rotational contributions (subdaily to Chandler period) and finally hydrogeological effects.

2. Surface gravity effects

We review now below how gravity observations can retrieve useful information on the Earth's structure and dynamics; more details can be found elsewhere (Hinderer and Legros 1989, Hinderer et al. 1991a).

We briefly recall the basic equations used for expressing the elasto-gravitational static deformation of an Earth model with fluid parts. As usually, we suppose that the Earth is hydrostatically pre-stressed and spherically symmetric. In the solid parts, the inner core and the mantle, we have the Navier equation of motion, which is linearised to the first-order with respect to displacement, density and gravitational potential changes and where the relationship between stress and displacement is given by Hooke's law of elasticity; in addition, we also have the Poisson equation relating the gravitational potential to the density and the mass conservation law. The linearised set of equations for a spheroidal mode of deformation can be written as a first-order linear differential system:

$$\dot{y}_j(r) = c_{ij}(r) y_j \qquad i,j = 1, ...6 \tag{1}$$

where $c_{ij}(r)$ are algebraic functions of radius $\mathbf{r}$, the elastic shear modulus, compressibility, and density of the Earth model. We have then a set of six equations for the six unknowns y_i (first introduced by Alterman et al. 1959) related to the displacement u, the traction T and the mass redistribution potential ϕ,:

$$\begin{aligned} u &= \sum_n \left[y_{1n}(r) Y_n(\theta,\lambda) \frac{r}{r} + r y_{3n}(r) \nabla Y_n(\theta,\lambda) \right] \\ T &= \sum_n \left[y_{2n}(r) Y_n(\theta,\lambda) \frac{r}{r} + r y_{4n}(r) \nabla Y_n(\theta,\lambda) \right] \\ V + \phi &= \sum_n \left[y_{5n}(r) Y_n(\theta,\lambda) \right] \\ y_{6n}(r) &= \frac{d y_{5n}(r)}{dr} - 4\pi G \rho(r) y_{1n}(r) \,. \end{aligned} \tag{2}$$

In Eq. (2), $\rho(r)$ and $V(r)$ are the unperturbed density and gravitational potential, G the gravitational constant, $\mathbf{r}$ the radius vector, $Y_n(\theta,\lambda)$ a surface spherical harmonic function of degree n (colatitude θ; longitude λ):

$$Y_n(\theta,\lambda)=\sum_m Y_n^m(\theta,\lambda)=\sum_m P_n^m(\cos\theta)[\cos m\lambda,\sin m\lambda]\ \text{ for }\ 0\le m\le n$$

and $P_n^m(\cos\theta)$ are the associated Legendre polynomials of degree n and order m. In Eqs (1) and (2), the radial coefficients have the following interpretations:

$y_1(r)$	radial displacement
$y_2(r)$	normal stress
$y_3(r)$	transverse displacement
$y_4(r)$	transverse stress
$y_5(r)$	gravitational potential
$y_6(r)$	gravitational flux density.

The toroidal mode of deformation is of no interest here because it does not perturb the gravity in the case of a spherical Earth model.

In the fluid parts, the liquid outer core or a surface layer representing a global ocean or the atmosphere, we have similar equations, except that Hooke's law reduces to a state equation relating the change in the fluid pressure to the elastic compressibility and to the displacement in the initial pressure gradient (e.g. Crossley and Rochester 1980).

The fluid equations (without viscosity) are often further simplified in order to de-couple the pressure perturbation from the gravity perturbations. Two major approximations for the Earth's liquid core are in general considered: the Boussinesq approximation where the fluid is taken as incompressible and the density and gravitational disturbances are negligible, and the subseismic (below acoustic frequencies) approximation which takes into account the core compressibility, but modifies the equation of state. In the oceans and atmosphere, two classical simplifications are based on the hydrostatic or the geostrophic hypothesis (see e.g. Volland 1988).

The Love number approach provides a convenient and useful description of the response of the solid parts of the Earth to global dynamics (e.g. tides, rotation) or to dynamics which may, for reasons of simplicity, have to be calculated only in the fluid regions (e.g. core modes). With this in mind, we may express the solutions y_i in the solid parts with the help of Love numbers and 'source' terms, where the word 'source' implies the result of a dynamical calculation (e.g. oscillations in the fluid core) or a known forcing (e.g. tidal potential).

Taking into account the displacement in the initial gravitational field (terms in y_1), as well as the mass redistribution potential (terms in y_5) in addition to any direct potential (if existing), it is easy to get the surface gravity change (Hinderer and Legros 1989):

$$-\frac{a}{n}\Delta g_n(a)=\delta_n V_n+\delta_n' V_n'+\bar{\delta}_n\frac{P_n}{\rho}+\bar{\delta}_n^1\frac{P_n^1}{\rho} \tag{3}$$

with the help of generalized gravimetric factors:

$$\begin{aligned}
\delta_n &= 1+\frac{2}{n}h_n-\frac{n+1}{n}k_n \\
\delta'_n &= 1+\frac{2}{n}h'_n-\frac{n+1}{n}k'_n \\
\bar{\delta_n} &= \frac{2}{n}\bar{h}_n-\frac{n+1}{n}\bar{k}_n \\
\bar{\delta}^1_n &= \frac{2}{n}\bar{h}^1_n-\frac{n+1}{n}\bar{k}^1_n\,.
\end{aligned} \tag{4}$$

We kept here only some major types of excitation sources like volume potential V_n, surface loading potential V'_n, outer surface pressure P_n and core-mantle boundary (CMB) pressure forces P^1_n respectively. Each gravimetric factor of degree n is related to a specific source term of the same degree. If a dynamic process involves source terms of different degrees, then the surface gravity changes are obtained by summing up the different harmonic contributions.

3. Superconducting gravimeters

The superconducting gravimeters (SG) are devices using the principle of magnetic levitation of a superconducting sphere instead of the mechanical spring in the case of the classical gravimeters. The levitation force results from a current trapped in a superconducting coil. In the French instrument, the coils and sphere are in Niobium which becomes superconducting below 10 K (Kelvin); to reach such a low temperature, the sensor has to be operated in liquid helium (evaporation temperature 4.2 K). The first prototype was built in the seventies and the first results from a continuous recording period came out some years later (Warburton and Goodkind 1977, 1978). Later on, these instruments were sold by a company called GWR Instruments based in San Diego (California). The first instruments were installed in Europe in 1980, in Brussels (Belgium) and in Bad Homburg (Germany). They were followed by the Chinese and French instruments. Up to now, there are about 18 superconducting gravimeters in operation worldwide which will participate to the project called GGP (Global Geodynamics Project) (see Crossley and Hinderer 1995 for a review). This project consists in a network of worldwide distributed SG which will start on July 1, 1997.

The mean gravity force is compensated by a mean magnetic force generated by two superconducting coils and, by definition, the perfect stability of the currents in the coils (there is no ohmic dissipation) induces the stability of the magnetic force. Therefore the long term instrumental drift is small and much smaller than for the classical spring meters where the unavoidable creep of the spring was a major obstacle to achieve long term stability. In addition, there is a capacitive detection system of position changes of the superconducting sphere and a magnetic feedback force to null any change. The output of the sensor is therefore a voltage which is directly proportional to changes in gravity.

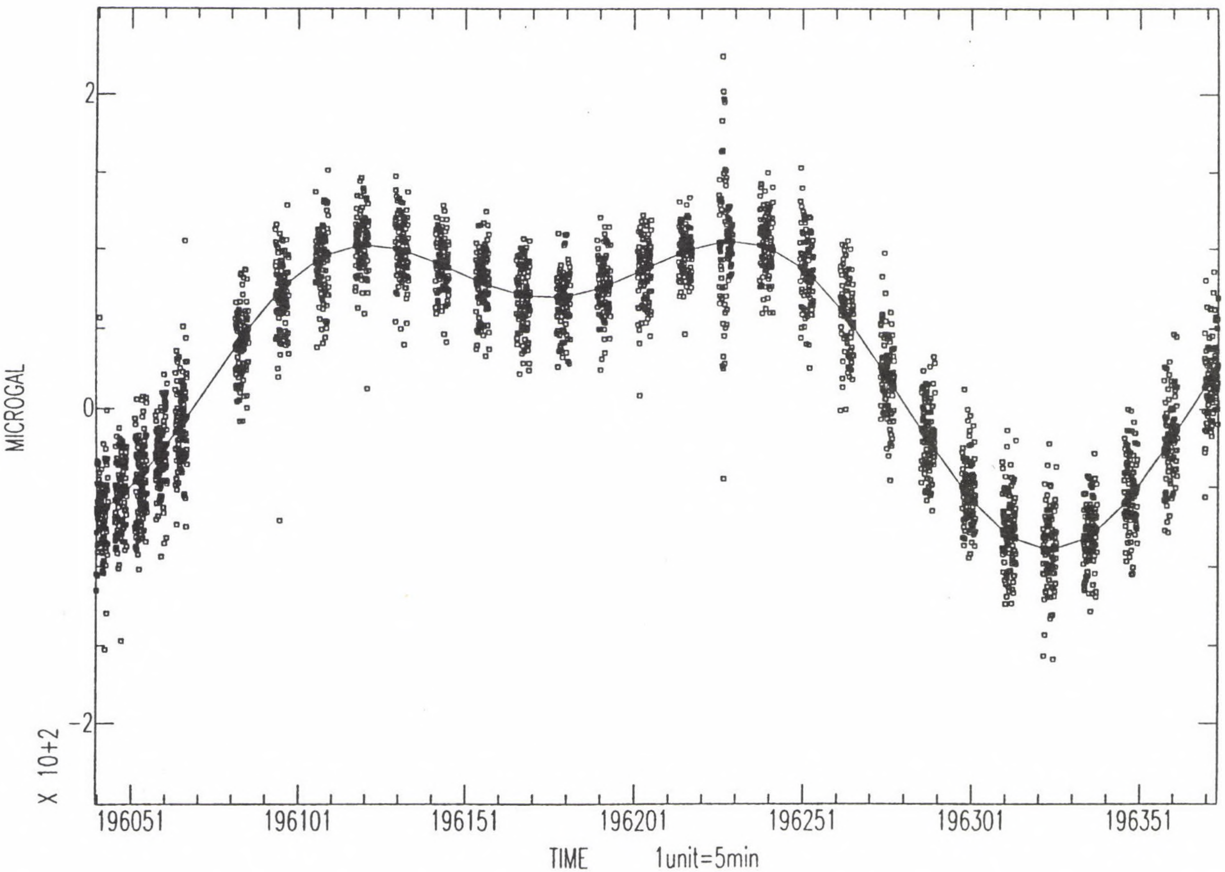


Fig. 2. Comparison at a same site (Strasbourg) between absolute gravity measurements (dots) and relative gravity changes as observed by a superconducting gravimeter (solid line)

The SG is a relative gravimeter which needs to be calibrated i.e. one has to know the conversion factor between the relative gravity change expressed in microgal (1 μGal = 10^{-8} N $\cdot$ kg^{-1}) and the feedback voltage (in volt). A precise calibration factor is needed for interpreting geophysically gravity changes; as an example, let us mention all the applications inferred from the tidal gravimetric amplitude factors (latitude dependence, ocean or atmospheric loading contributions, mantle anelasticity, heat flow correlation).

There are different methods based on the gravitational attraction of various bodies (e.g. Varga et al. 1995), inertial accelerations with an oscillating platform (e.g. Richter 1995) or elevation in the Earth's gravity field (calibration line). Another method relies on parallel registration with calibrated spring meters or with absolute gravimeters (AG) (e.g. Hinderer et al. 1991b, Bower et al. 1991). An example of such a comparative study in Strasbourg is given in Fig. 2 where the dots represent the absolute measurements with a JILAG-5 instrument roughly every 15 sec and the full line the (calibrated) gravity changes measured by the French superconducting gravimeter. The calibration uncertainty arising from an experiment over several tidal cycles is slightly better than 1 %. It also appears that the determination of the calibration factor is dependent on various parameters like the laser frequency of the AG and, more specifically, its short-term fluctuation. More recently, the conclusions from an AG/SG intercomparison study in Strasbourg using repeated measurements at different epochs has shown that the calibration value is stable in time and that its accuracy is able to reach a few ppm (Hinderer et al. 1995a).

4. Seismic band

The seismic frequency band (periods lower than 54 min which corresponds to the fundamental spheroidal eigenmode ${}_0S_2$) is usually studied with the help of broadband seismometers (BS) like the STS-1 or STS-2 (Wielandt and Streckeisen 1982), or even with LaCoste-Romberg spring meters (Zürn et al. 1991), but recent studies (Richter et al. 1995, Freybourger et al. 1996) have shown the ability of superconducting gravimeters to retrieve with equivalent quality (same signal to noise ratio) the seismic eigenmodes at least up to frequencies around 0.1 Hz. In fact, it appears that differences in the seismic mode retrieval are essentially caused by site noise; on the other hand, there is a clear advantage of the SG with respect to the BS in retrieving low frequency signals, especially in the subseismic band (periods above 54 min). Figure 3 shows the power spectral densities (PSD) from a one month record (September 1994) of the Strasbourg SG and a BS located in a mine at Echery in the Vosges (70 km away from Strasbourg). It appears clearly that the PSD are of quite comparable level in the seismic band but that the SG has a lower noise level in the subseismic band; also indicated is the New Low Noise Model (NLNM) from Peterson (1993) obtained from averaging many seismometers in many stations worldwide distributed. The noise shown by the BS increases with decreasing frequencies much more than the SG does and this is probably caused by uncorrected thermal effects (in addition to long-term creep) on the seismometer spring (Freybourger et al. 1996).

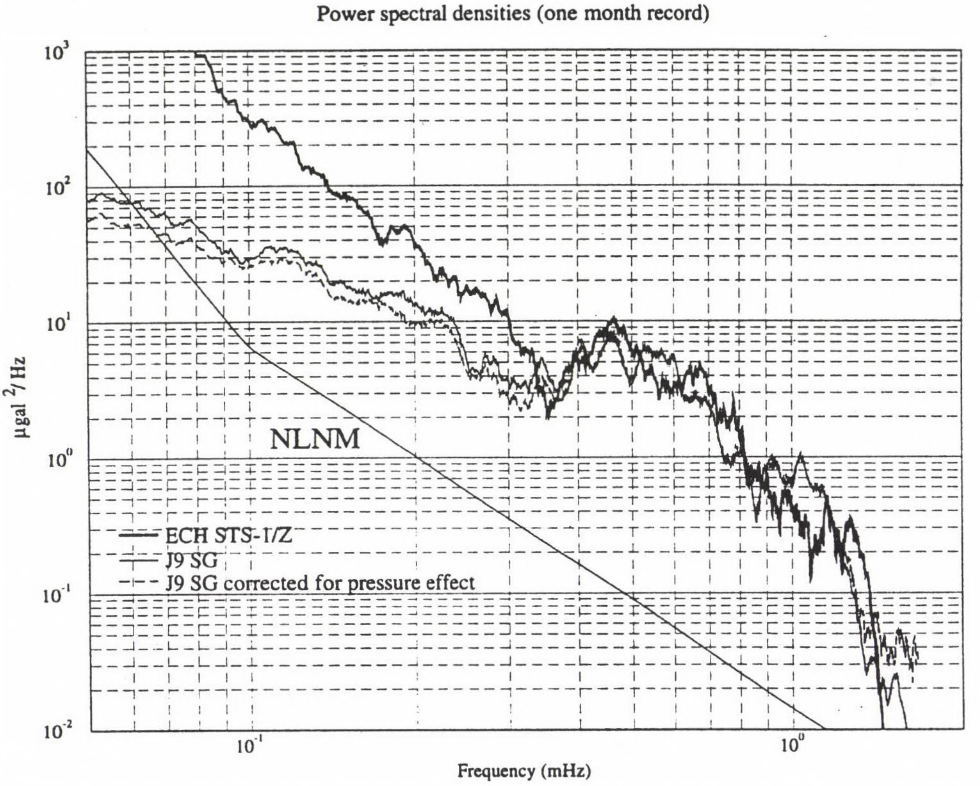


Fig. 3. Noise comparison between a superconducting gravimeter (J9 SG) and the vertical component of a broadband seismometer (ECH STS-1/Z) in the seismic and subseismic frequency bands (from Freybourger et al. 1996)

5. Subseismic band

On the contrary to seismic normal modes where elasticity provides the main restoring force, the core modes are gravity-inertial modes because of their dependence on both inertial forces and gravity restoring forces (buoyancy related to the stability of the density stratification via the Brunt-Vaissala frequency). These oscillations have a theoretical spectrum which is still controversial and debated in the literature; for more details we refer the interested reader to several recent studies on the subject (Smylie et al. 1992, Crossley and Rochester 1992, Rochester and Peng 1973, Wu and Rochester 1994). One of the major problems in the Earth deformation at long periods is the incorporation of the Coriolis term in the equations of motion which acts to couple together modes of different spatial structure; some authors even have dismissed the validity of the use of a truncated chain of terms in a spherical harmonic expansion of the core motion. Rieutord (1991) found that the inclusion of viscosity makes the problem of core oscillations well posed and tractable even when using a spherical harmonic expansion approach. Moreover, the use of

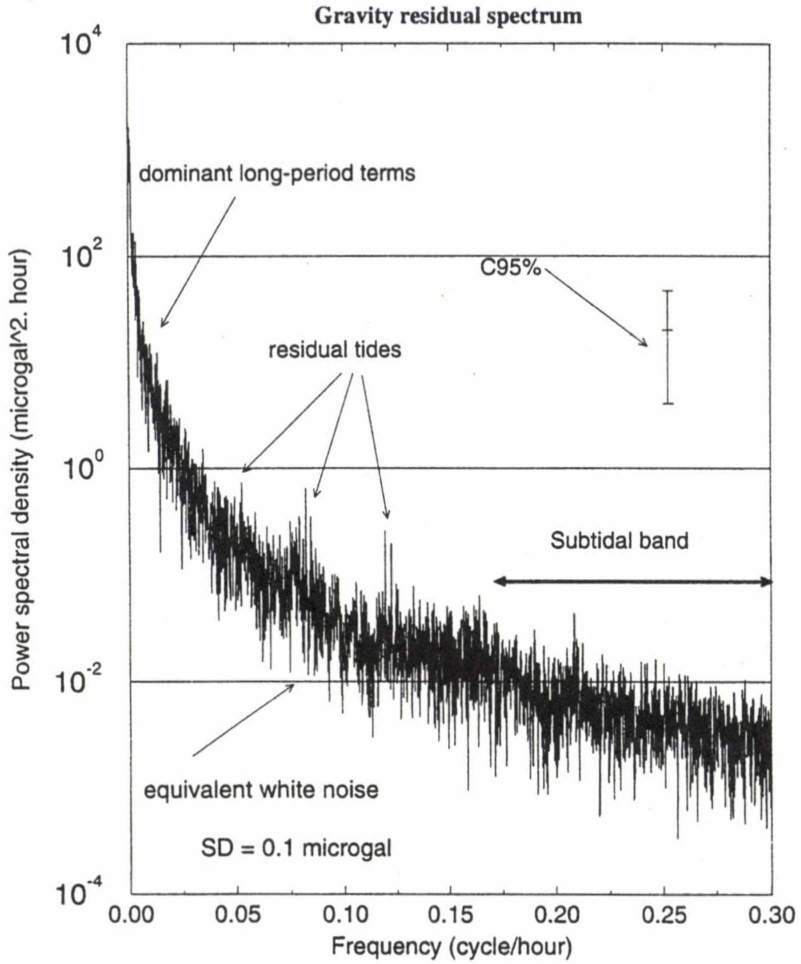


Fig. 4. Typical coloured-type gravity spectrum from a 2 year record of a superconducting gravimeter (Strasbourg SG) indicating a reduction of noise with increasing frequencies

static Love number in this problem is not trivial and we refer to Crossley et al. (1991, 1992) for more details.

Classically it is believed that earthquakes might be able to excite these modes as they do for seismic normal modes. Recent computations based on the standard dislocation source function (Crossley 1988, Crossley et al. 1991, Crossley 1993) have shown that the excitation is necessarily weak, essentially because of the high rigidity contrast existing between the mantle and the liquid core. In fact, even for the largest earthquake ever recorded (like the 1960 Chile event), most of the core modes would be excited to a surface gravity level of less than a nanoGal. Such a level is slightly below our current level of detectability, which is typically a few nanoGal in the period range of several hours as indicated by Fig. 4, even for the high quality SG records where careful corrections are made for tides, local pressure fluctuations, instrumental drift and disturbances (Hinderer and Crossley 1993, Jensen et al. 1995). For longer periods, the situation is even worse because of larger noise contributions which lead to a f^{-2} (where f is the frequency) power law spectrum as suggested by Jensen et al. (1995). The random "brown" residual gravity signal depicted by Fig. 4 can result from various causes of instrumental origin or due to environmental contributions (tectonics, hydrogeology, atmosphere). Among other possible source of excitation, tidal forcing or the motion of chemically driven disturbances within the core from the inner-core boundary (ICB) have also been suggested. It has been shown that seismic anelasticity (essentially imperfections in the fluid core bulk incompressibility) is more important in the damping process of the core modes than magnetic field dissipation in the core (Crossley and Smylie 1975) or viscous boundary layer effects. This leads then to a long persistence of excited core modes with typical decay times of about one year.

The first claim of detection of core modes was made by Melchior and Ducarme (1986) using the Brussels SG data, in which they found significant the detection of a 15 nanoGal signal at 13.9 hr following two large earthquakes. Subsequent studies by Melchior et al. (1988) or Aldridge and Lumb (1987) initially supported this claim and its possible connection to a gravity-inertial oscillation of the fluid core. Later on, most of the studies did not confirm the existence of this mode (Zürn et al. 1987, Mansinha et al. 1990, Florsch et al. 1991). It is now believed that the Brussels oscillation was due to regional combined ocean-atmosphere effects (Dehant et al. 1993b). There was an attempt to detect core modes from stacking IDA (International Digital Accelerometers) gravity data but nothing significant could be found (Cummins et al. 1991). The first stack using several SG records in Europe was initiated by Smylie et al. (1993) and has led to the claim of detection of the Slichter triplet (see next section) but no other core mode was revealed. The lack of persistent periodicities in the core-mode band was later confirmed by the analysis of a 2 year record of the French and Canadian SG data (Hinderer et al. 1994, 1995b, Jensen et al. 1995).

6. Slichter mode

A special case of core oscillations of degree one is the triplet of translational modes of the Earth's solid inner core and named in honour of Slichter (1961) who first pointed out the possibility that the inner core may oscillate within the fluid outer core. The triplet arises because of the splitting of the eigenmode caused by the effect of the Earth's rotation and ellipticity (see e.g. Dahlen and Sailor 1979). Earthquakes can in principle excite this mode (Crossley et al. 1991) but another potential candidate could be volcanic eruptions (which are less energetic than earthquakes but have a strong degree 1 contribution). It can be shown that the damping caused by the core anelasticity is weak as for the other degree core modes but an alternative mechanism involving the thermodynamics of the assumed phase transition at the ICB (Wu and Rochester 1994, Crossley et al. 1994) suggests that large heat dissipation could occur during the motion which would imply a very strong damping obviously inhibiting any detection in surface gravity.

Even if the emergence of the peaks above the mean noise level is marginal, the major factor behind the claim of Smylie (1992) was the coincidence of the peaks with Smylie's computed eigenperiods for the triplet (Smylie et al. 1994). In this last study it was also shown that no equivalent triplet in the pressure product spectrum was present therefore eliminating any contamination from the atmosphere.

On the observational side, numerous studies were devoted recently to re-examine the claim for detection of the Slichter triplet. These studies (e.g. Hinderer et al. 1995b, Jensen et al. 1995) could never univocally confirm the existence of the triplet, whatever the method of analysis or the quality of the SG data sets.

The large ($27 \cdot 10^{20}$ Nm) and deep (670 km) Bolivian earthquake of June 9, 1994 at first appeared to be an excellent event to excite the Slichter mode. Unfortunately, theoretical computations based on standard seismic models indicate a very weak surface gravity perturbation of 0.02 nanoGal which is by several orders of magnitude smaller than the typical noise level of SG in the core-mode band. The analysis of SG data after this earthquake from Strasbourg and BFO (Black Forest Observatory) in Germany led to spectra where persistent long period seismic modes could be seen but no evidence was found for the Slichter triplet or any core-constrained seismic modes (Jensen et al. 1994).

The difficulty of identifying a weak spectral peak above a high background noise level is of course well-known. A recent study (Florsch et al. 1995a) has shown that it is not surprising from a statistical point of view to find, in a typical gravity spectrum from a long SG record, many peaks above a meaningful level, such as 3 standard deviations.

It is therefore worth in the future to assemble data sets from many SG stations in order to allow stacking for common signals. This will strongly reduce the possible misidentification of spectral peaks due to random noise or artefacts generated by the numerical methods of analysis. This is exactly one of the goals of the GGP which will start in 1997 for a 6 year common observing period.

7. Tidal bands

The lunisolar tidal gravity changes, as well as those caused by fluctuations in the Earth's rotation, can be expressed with the help of an external volume potential and lead to the following gravity effect:

$$-\frac{a}{n}\Delta g_n(a) = \delta_n V_n \tag{5}$$

with the standard volume gravimetric coefficient:

$$\delta_n = 1 + \frac{2}{n}h_n - \frac{n+1}{n}k_n \, .$$

The unit coefficient represents the direct effect of the potential, the term in h_n represents an elevation change through the unperturbed gravity field and the term in k_n expresses the mass redistribution potential due to the elastic deformation.

Because the tidal potential of degree n is inversely proportional to d^{n+1}, where d is the distance between the Moon or Sun and the Earth, the major contribution

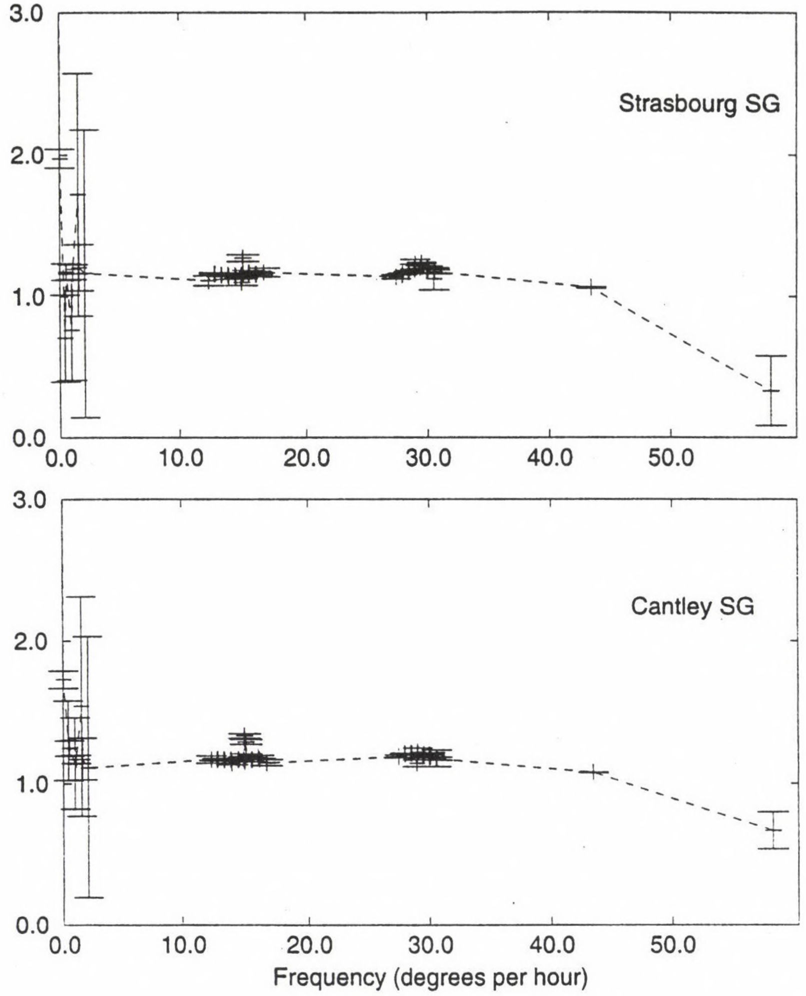


Fig. 5. Frequency dependence of the tidal gravimetric amplitude factors after a standard least squares adjustment of the French and Canadian SG data to tides and correction for instrumental drift and local atmospheric pressure changes

to the total gravity is from the degree $n = 2$, but additional terms coming from degrees 3 and 4 are also significant at the level of precision of the best gravimeters. The gravimetric factor δ_n is, in a spherical symmetrical model, independent of order m. If the Earth model includes ellipticity and rotation, there is a coupling between terms of different degrees (Wahr 1981) and a slight latitude dependence appears then in the gravimetric factor.

Why is it important to measure accurately the tidal gravimetric factors? This is because they provide an answer on how the Earth deforms elastically or anelastically at some reference frequencies ranging from the subseismic band (quarter-diurnal tides) to long-period tides (1 year or even the 18.6 year Bradley term). Even if

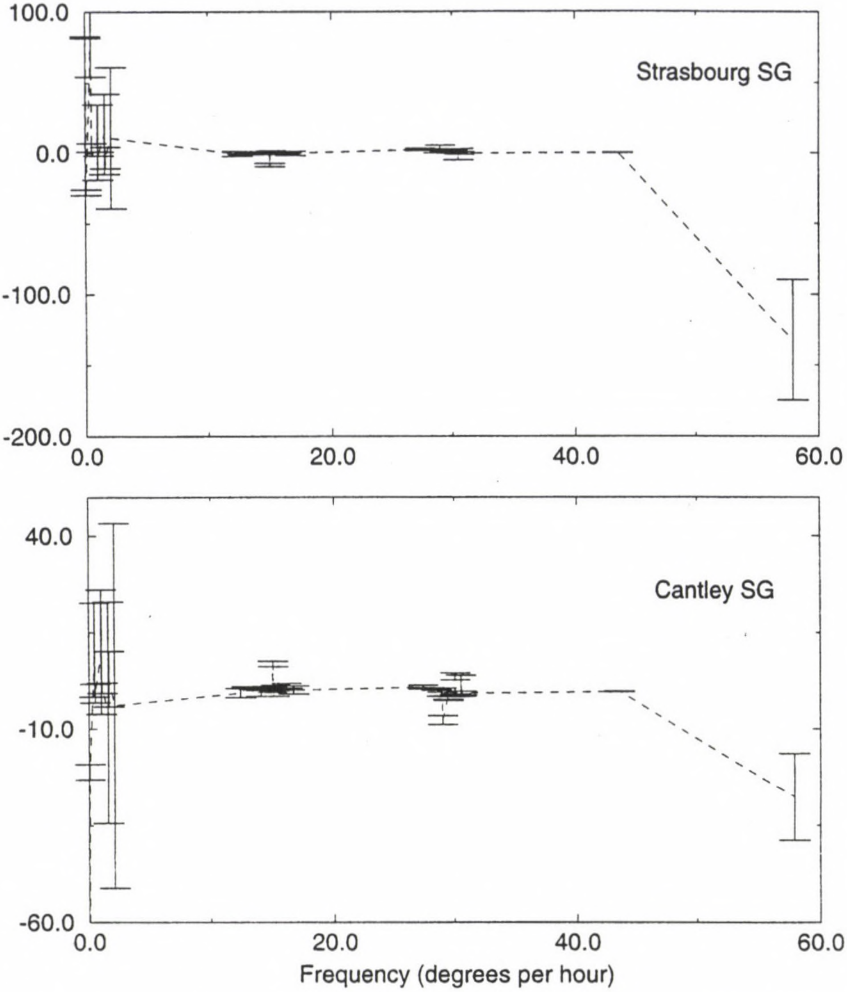


Fig. 6. Frequency dependence of the tidal gravimetric phase factors after a standard least squares adjustment of the French and Canadian SG data to tides and correction for instrumental drift and local atmospheric pressure changes

the problem is obviously mixed with ocean tidal loading corrections, it remains a very useful source of information on the Earth's rheology at intermediate time scales between rapid motions (seismology) and slow tectonic motions (post-glacial rebound). In addition to the elastic transfer function, the tidal gravimetric factors are also able to reveal large-scale processes like the flattening effect due to the Earth's shape and axial rotation (Wang 1994) or even correlations with heat flow (Rydelek et al. 1991, Melchior 1995). Figure 5 and 6 show the distribution as a function of frequency of the tidal gravimetric amplitude factors and phase factors respectively from an analysis of the French SG.

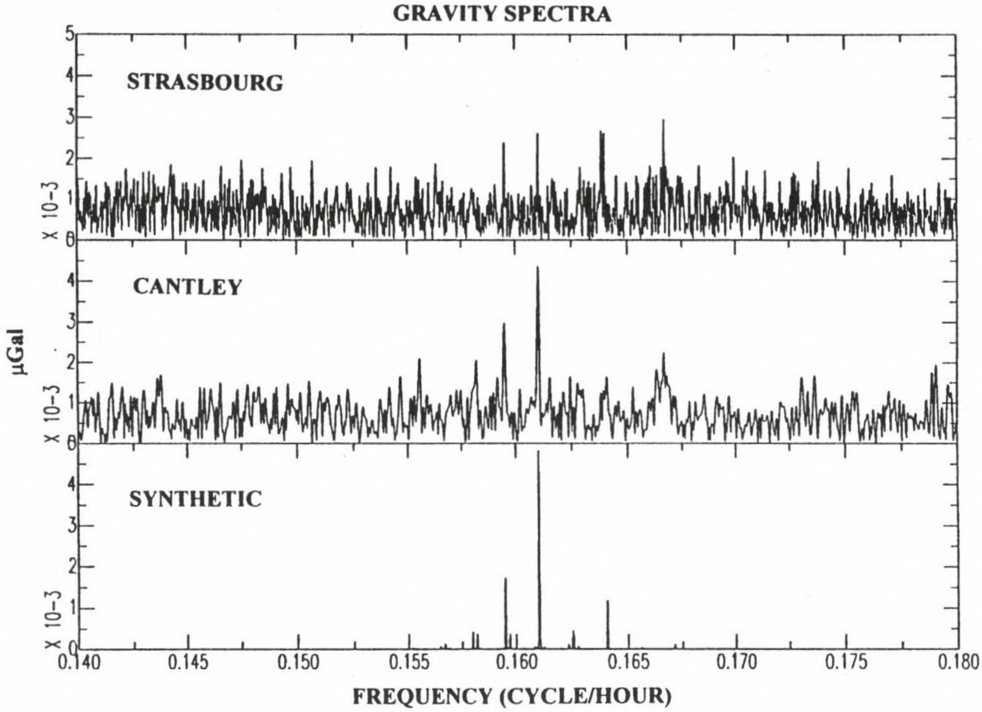


Fig. 7. Indentification of quarter-diurnal tides of a few nanoGal amplitude in superconducting gravimeter data (from Florsch et al. 1995b) indicating the high sensitivity of the cryogenic instrument

As we have reported earlier in this paper, the sensitivity needed to detect core modes in the case of a strong excitation (including the Slichter triplet) is the nanoGal. There are at least two proofs of the ability of the SG to reach such an accuracy. One was given by Merriam (1992) from the observation of gravity changes induced by small microbarometric pressure changes in the atmosphere. Another was given by Florsch et al. (1995b) who found the trace in the French and Canadian SG records of the quarter-diurnal tidal waves (degree and order 4) of some nanoGal amplitude, as illustrated by Fig. 7.

Moreover the accurate measurement of tidal gravimetric factors offers the possibility to detect resonances in the Earth's response as will be shown in the next section.

8. Resonances in gravimetry

Apart contributions from the oceans and the atmosphere, the most important deviations in the amplitude of the gravimetric factors with respect to the elastic value of a standard Earth model are caused by resonance phenomena. Two major ones are related to rotational eigenmodes of an Earth's model with a liquid core and a solid tilted inner core.

8.1 Free Core Nutation

The best known resonance process in geodesy is the diurnal resonance caused by the existence of a free rotational mode of the coupled core-mantle system called FCN (Free Core Nutation).

The main effect is due to the action of the pressure field related to the differential rotation appearing between the tilted core (Poincaré flow) and the mantle on the elliptical shape of this boundary but additional topographic torque components might also exist for a more general bumpy interface (Hide 1969, Hinderer et al. 1990). Moreover, if the liquid core is assumed to be viscous and electrically conducting, visco-magnetic friction may also act at the boundary leading to a coupling which is partly dissipative (Sasao et al. 1980). This mode can be excited by diurnal lunisolar tides and leads to a resonance in surface gravity (Hinderer et al. 1991a). A single gravimetric factor, that we denote for brevity $\bar{\delta}_2^1$, is sufficient to express the surface gravity change as a function of the CMB pressure distribution resulting from the outer core rotation. The diurnal tidal potential V_2 is able to excite this rotation and the amplitude is found by solving the Euler equations for conservation of angular momentum. We have then for the total tidal gravity change (e.g. Neuberg et al. 1987):

$$\Delta g(\sigma, a) = -\left[\delta_2 + \frac{\bar{\delta}_2^1 A}{\sigma - \sigma^{oc}}\right]\frac{2V_2}{a} \tag{6}$$

where σ^{oc} is the eigenfrequency of the FCN and A a constant term having the dimension $(\text{time})^{-1}$ and depending on several geodynamical parameters. For tidal waves of frequency σ close to σ^{oc}, there is a resonant amplification in the tidal gravimetric factors, which has been clearly observed in precise gravity records since many years (e.g. Melchior 1983). Moreover, the high quality of tidal measurements now available in particular with SG allows to set constraints on the period and damping of the NDFW (Neuberg et al. 1987, Defraigne et al. 1994, 1995) and on the rotational pressure gravimetric factor $\bar{\delta}_2^1$ (Hinderer et al. 1991c). The FCN resonance in gravimetry for tidal waves close to the eigenfrequency is shown in Fig. 8 for the French (J9) and Canadian (CSGI) SG from a 2 year common observing period (Hinderer et al. 1995c). Far away from the resonance, the resonant load is zero and we see that the resonance process enhances the waves ψ_1 and ϕ_1 while it reduces K_1 and P_1.

A summary of values for the FCN eigenperiod and quality factor retrieved from SG is given in Table I; we have also added the result from VLBI (Very Long Baseline Interferometry) as well as the theoretical values for an inelastic Earth model in the hydrostatic assumption.

The results show that there is a significant discrepancy between the observed eigenperiod and the theory. This has been attributed to a non-hydrostatic contribution to the flattening of the CMB (Neuberg et al. 1990, Wahr and De Vries 1989) and the current opinion is that this excess flattening may be due to the dynamics of the overlying mantle is now favoured by several authors (Hager et al. 1985, Forte et al. 1994). It is worth noting that the CMB deformation which is generated

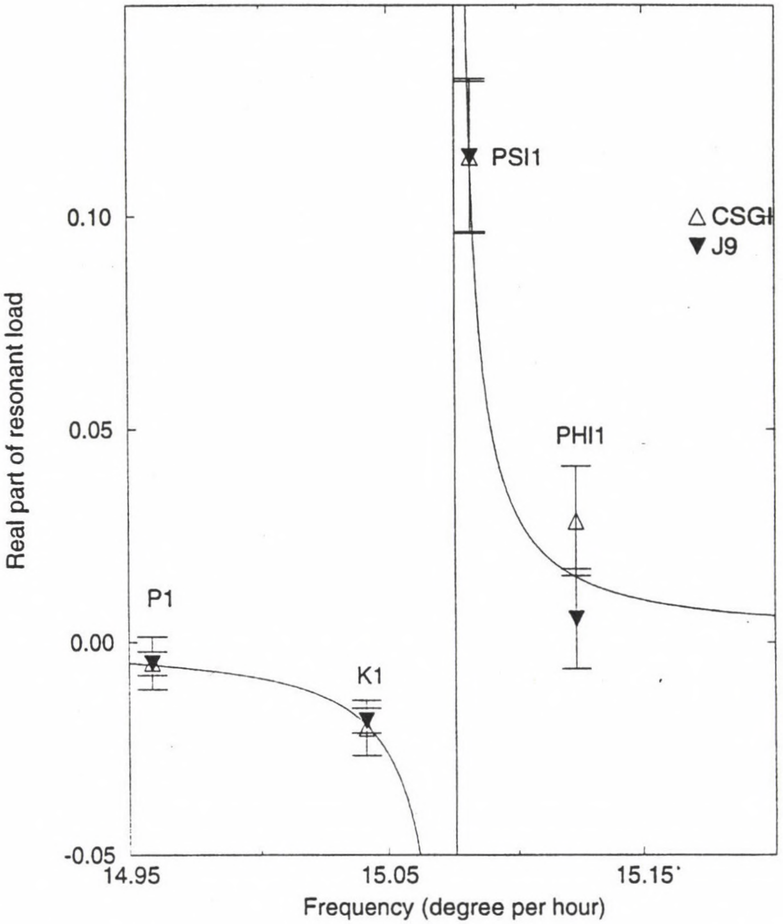


Fig. 8. Free Core Nutation (FCN) resonance in tidal gravimetry as observed by the French and Canadian superconducting gravimeters

by mantle heterogeneities reconstructed using the history of subducting slabs leads indeed to an increase in ellipticity of the right order of magnitude (Greff and Legros 1996). On the contrary, ellipticity values inferred from seismological studies (e.g. Morelli and Dziewonski 1987) do not. Let us also mention that deviations in the core flow from the uniform vorticity approximation (Poincaré flow) would change the pressure field at the CMB and therefore the pressure coupling contribution in the expression of the eigenfrequency.

If mantle anelasticity is taken into account, A in Eq. (6) becomes a complex-valued quantity (Wahr and Bergen 1986). Other contributions to the FCN eigenfrequency, from either the presence of an oceanic or atmospheric surface layer (Legros and Amalvict 1989) or from that of the solid inner core (Mathews et al. 1991a, 1991b, De Vries and Wahr 1991, Legros et al. 1993), are found to be small.

Table I. FCN retrieval in superconducting gravimetry

Author	Method	T	Q
Wahr and Bergen 1986	theory	473.8	78000
Herring et al. 1986	VLBI	434.6 ± 0.6	22000–100000
Neuberg et al. 1987	stacked gravity B + BH SG	431.0 ± 6.0	2800 ± 500
Sato et al. 1994	stacked gravity JAP SG	436.7 ±15.0	3200–∞
Defraigne et al. 1994	stacked gravity B + BH + J9 SG	425 ±6	2400–8300
Florsch et al. 1994	J9 SG	430.7 ±1.0	1700–2500
Merriam 1994	CSGI SG	430 ±6	5500–10000
Hinderer et al. 1995c	stacked gravity J9 and CSGI SG	429 ± 8	7700–∞

B: Brussels SG; BH: Bad Homburg SG; JAP: 3 Japanese SG

As previously said, the FCN parameters are determined in gravimetry from the resonant amplification of tidal waves. Another possibility would be to observe the eigenmode directly and to infer its eigenperiod and damping. The detection in VLBI of the motion in space of the freely excited FCN has been claimed recently by Herring et al. (1991) with an amplitude of less than a milliarcsec (and eigenperiod close to 430 days). The gravity change induced by the rotational potential (Hinderer and Legros 1989) associated with the corresponding wobble in a geographic frame would be of the order of 0.01 nanoGal. This level is clearly below the present detectability threshold from high-quality SG records, as discussed in the section devoted to the search for core modes. It is therefore unlikely that a direct gravimetric detection will be possible in the near future.

8.2 Free Inner Core Nutation

Besides the FCN, there is another free mode recently predicted by the theory, which is caused by the rotation of the solid inner core with respect to the outer core and mantle. This mode, called Free Inner Core Nutation (FICN), was computed independently by De Vries and Wahr (1991), Mathews et al. (1991a, 1991b) and

Dehant et al. (1993a), assuming gravitational and inertial coupling torques between the inner core and the rest of the Earth, as well as contributions to the inertia tensors due to the elastogravitational deformation. These studies found more or less comparable numerical values for the nearly diurnal prograde eigenfrequency relative to this new rotational mode. Because this mode lies in the diurnal tidal band (in the vicinity of the prograde annual wave S_1), it is interesting to investigate its resonant contribution in gravity.

Similarly to the FCN, the tidal gravimetric factor is modified to:

$$\Delta g(\sigma, a) = -\left[\delta_2 + \frac{\bar{\delta}_2^2 B}{\sigma - \sigma^{ic}}\right] \frac{2V_2}{a} \tag{7}$$

where σ^{ic} is the FICN eigenfrequency, B another constant term of dimension $(\text{time})^{-1}$ and $\bar{\delta}_2^2$ a single gravimetric factor relative to boundary conditions at the inner core-outer core interface.

The largest effect concerns the prograde annual wave S_1 and is found to be a few parts in 10^4 (De Vries and Wahr 1991). The uncertainty related to the atmospheric forcing at this frequency (e.g. Warburton and Goodkind 1978) makes any observation difficult to interpret in terms of inner core properties.

9. Earth's rotation

The direct effect of the Earth's rotation (either in wobble or in axial spin) derives from a centrifugal potential. For its spherical harmonic contribution restricted to the degree two (e.g. Hinderer et al. 1982), the resulting gravity change can be expressed with the help of a volume potential. The gravity change caused by a polar motion of components $X(t)$ and $Y(t)$, in the conventional terrestrial frame, at a station of colatitude θ and longitude λ, is:

$$\Delta g(\theta, \lambda, t) = \delta_2 \Omega^2 a \sin 2\theta [X(t) \cos \lambda - Y(t) \sin \lambda] \tag{8}$$

where Ω is the Earth's uniform axial rotation rate. This contribution is a maximum at mid-latitudes ($\theta = 45°$) and can reach about 3 μGal for a polar motion amplitude of $| X + iY | \approx 7 \cdot 10^{-7}$, as it is the case for the mean value of the Chandlerian component of 435 sidereal day period (Lambeck 1980). There is also an annual term in the polar motion which leads to a gravity fluctuation of similar amplitude. These long-period gravity changes of rotational origin have been indeed observed with superconducting gravimeters (Richter and Zürn 1988) because of the low instrumental drift of these instruments with respect to the previous spring gravimeters. Another example is given by Fig. 9 (from Richter et al. 1995) which shows the polar motion effect on gravity as observed by a SG at the Richmond station in Florida, USA. One easily sees that the long term fluctuation is close to the theoretically predicted change; there is of course also a higher frequency content in the observed gravity that might be caused by environmental effects. In attempting to use the long-period rotational gravimetric factor δ_2 for studies on the mantle anelasticity and/or ocean loading, it is necessary to obtain longer gravity records,

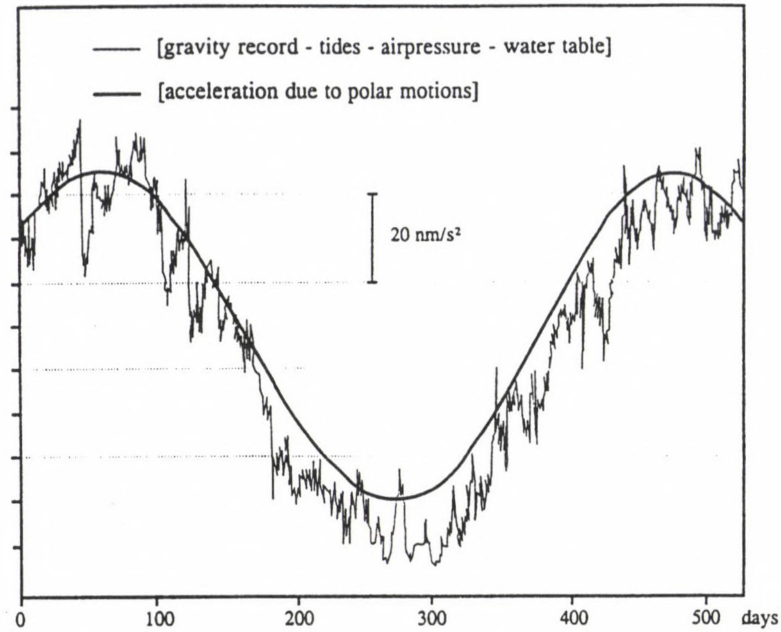


Fig. 9. Long-period gravity change induced by the Earth's rotation and its observability by a cryogenic gravimeter at Richmond, Florida, USA (from Richter et al. 1995)

in order to improve the separation between the annual components (having various origins) and the Chandlerian one.

In addition to these long-term wobble changes, there are also rapid motions with timescales between several weeks and several months (Eubanks et al. 1988) which would lead to gravity changes (with similar period) smaller than a μGal and quite difficult to identify.

The axial rotation of the Earth leads to a centrifugal volume potential that can be separated into a purely radial part and into a degree $n = 2$ and order $m = 0$ spherical harmonic part (e.g. Hinderer et al. 1982). It can be shown that the seasonal contributions, as well as the so-called 50 day period fluctuation (e.g. Langley et al. 1981), would produce gravity effects smaller than the ones caused by the polar motion but, in principle, detectable.

10. Oceanic and atmospheric loading

We have already mentionned the fact that ocean loading is responsible for anomalies in the (solid Earth) tidal gravimetric factors both in amplitude and phase. Inversely, precise tidal measurements can help in improving the knowledge of oceans tides (e.g. Francis and Melchior 1996). Moreover, it can be shown that SG tidal measurements are capable of sensing ocean tides to a few millimetres and also of exhibiting small non-linear tidal lines (Merriam 1995) in addition to the dynamic linear response of the oceans (see also Platzman et al. 1981). Notice that, in the case of ocean loading, the generally non-static character of the ocean tides (and the restriction to the oceanic domain) does not allow to use a single loading gravimetric factor to compute its gravity contribution inland.

In the context of atmospheric contributions to gravity, a loading factor δ_n' applies to any atmospheric loading process of large spatial extension excluding however those processes which generate surface pressure fluctuations not directly related to the integrated vertical density profile (of dynamic origin for instance). Of special interest is also the problem of the ocean response to atmospheric pressure changes (see e.g. Dickman 1988) and how this interfers with loading effects seen in gravity. This type of modelling has to be used to investigate the gravity effects due to the thermally driven solar tides in the atmosphere like S_1, S_2 and higher harmonics (Haurwitz and Cowley 1973, Warburton and Goodkind 1978). This approach can also be used to study the direct effect of the large-scale atmospheric loading which is responsible, for instance, for the seasonal changes in the Earth rotation by exchanging angular momentum. It can be shown that, for the annual term, the direct gravity effect from atmospheric loading can be as large as the gravity change indirectly caused by the annual polar motion (Hinderer and Legros 1991).

11. Conclusion

We have shown that high quality observations of gravity changes with superconducting gravimeters are able to provide a better knowledge of the Earth's global dynamics as well as its internal structure. In the high frequency band (seismic band), the SG have comparable performances to spring gravimeters or broadband seismometers to retrieve the elastic normal modes and the major source of noise contamination is more site-related rather than instrumental. However, for decreasing frequencies (subseismic band), SG are superior to long-period seismometers to detect the gravity-inertial core modes (probably because of thermal perturbations on the leaf-spring used in the seismometers). Because of uncertainties in the theory of these modes (eigenfrequencies, excitation, damping), there is presently no clear identification of them in SG records. We have paid special attention to the Slichter triplet related to the translation of the solid inner core and we have shown that there is presently no confirmation of an initial claim for detection from a stack of European SG. The investigation of the tidal bands provides useful information on the (an-)elastic transfer function of the Earth with the help of tidal gravimetric amplitude and phase factors. We have also shown that quarter-diurnal tides could be identified in high quality SG records and this fact clearly proves the ability of the cryogenic instruments to retrieve extremely weak amplitude signals (at the nanoGal level). We have then investigated resonance processes which perturb the tidal gravimetric factors like the one caused by the FCN (rotation of the liquid core) or the FICN (rotation of a tilted inner core). The first one is nicely observed and allows to retrieve several parameters directly related to properties of the core-mantle boundary. An example is the need for a slight increase in the flattening of the CMB from accurate observations of the FCN eigenperiod.

The second resonance (FICN) predicted by the theory is yet to be observed but this will require a lot of efforts because of perturbations of atmospheric origin on the tides in the vicinity of the eigenmode. In the long-period band which was totally hidden in standard gravity records before the use of the cryogenic instruments (because of strong drift effects), there are signals due to the polar motion (annual and Chandlerian terms), as well as oceanic and atmospheric contributions. The observation of these signatures in gravity helps to understand the rheological behaviour of the Earth at long periods (polar gravimetric factor) together with the interactions of the external layers with the solid Earth (loading effects).

A major problem which always appears in the attempts to detect weak gravity signals is to separate them from random noise spectral peaks. To do better, a network of cryogenic instruments is needed in order to promote stacking methods which will enhance the signal to noise ratio in almost all the problems. This is goal of the Global Geodynamics Project (GGP) establishing a network of about 18 SG stations located worldwide (Europe, Japan, United States, Canada, China, Antartica). The official start of the 6 year observing period of GGP will be July 1, 1997.

References

Aldridge K, Lumb L I 1987: *Nature*, 325, 421–423.

Alterman Z, Jarosch H, Pekeris C L 1959: *Proc. R. Soc. A.*, 252, 80–95.

Bower D R, Liard J, Crossley D, Bastien R 1991: In: Proc. ECGS, Workshop on 'Non-tidal gravity changes: intercomparison between absolute and superconducting gravimeters', Cahiers du Centre Européen de Géodynamique et de Séismologie, Walferdange, Luxemburg, 3, 129–142.

Crossley D J 1988: *Geophys. Monograph. Ser., Am. Geophys. Union*, 41–50.

Crossley D J 1993: *J. Geomag. Geoelectr.*, 45, 1371–1381.

Crossley D J, Hinderer J 1995: In: Proc. Second IAG Workshop on 'Non-tidal gravity changes: Intercomparison between absolute and superconducting gravimeters', Cahiers du Centre Europaen de Géodynamique et de Sismologie, Luxemburg, 11, 244–274.

Crossley D J, Rochester M 1980: *Geophys. J. Roy. astr. Soc.*, 60, 129–161.

Crossley D J, Rochester M 1992: *Geophys. J. Int.*, 108, 502–506.

Crossley D J, Smylie D E 1975: *Geophys. J.*, 42, 1011–1033.

Crossley D J, Hinderer J, Legros H 1991: *Phys. Earth Planet. Int.*, 116, 68–97.

Crossley D J, Rochester M, Peng Z 1992: *Geophys. Res. Lett.*, 19, 1679–1682.

Crossley D J, Jensen O, Rochester M, Wu W 1994: The Slichter triplet – Why have we not seen it? Poster presented at the SEDI 94 Symposium, Whistler, Canada

Cummins P, Wahr J, Agnew D, Tamura Y 1991: *Geophys. J. Int.*, 106, 189–198.

Dahlen F A, Sailor R V 1979: *Geophys. J. R. astr. Soc.*, 58, 609–623.

Defraigne P, Dehant V, Hinderer J 1994: *J. Geophys. Res.*, 99, 9203–9213.

Defraigne P, Dehant V, Hinderer J 1995: *J. Geophys. Res.*, 100, 2041–2042.

Dehant V, Hinderer J, Legros H, Lefftz M 1993a: *Phys. Earth Planet. Int.*, 76, 259–282.

Dehant V, Ducarme B, Defraigne P 1993b: *Geophys. Monograph Series, Am. Geophys. Union*, 72, 35–44.

De Vries D, Wahr J 1991: *J. Geophys. Res.*, 96, 8275–8293.

Dickman S 1988: *J. Geophys. Res.*, 93, 14941–14946.

Eubanks T M, Steppe J A, Dickey J O, Rosen R D, Salstein D A 1988: *Nature*, 334, 115–119.
Florsch N, Hinderer J, Crossley D J, Legros H, Valette B 1991: *Phys. Earth Planet. Int.*, 68, 85–96.
Florsch N, Chambat F, Hinderer J, Legros H 1994: *Geophys. J. Int.*, 116, 53–63.
Florsch N, Legros H, Hinderer J 1995a: *Phys. Earth Planet. Int.*, 90, 197–210.
Florsch N, Hinderer J, Legros H 1995b: *Bull. Int. Marées Terr.*, 122, 9189–9198.
Forte A, Woodward R, Dziewonski A 1994: *J. Geophys. Res.*, 99, 21857–21877.
Francis O, Melchior P 1996: *Geophys. Res. Lett.*, 23, 2251–2254.
Freybourger M, Hinderer J, Trampert J 1996: *Phys. Earth Planet. Int.* (in press)
Greff M, Legros M 1996: *Geophys. J. Int.*, 125, 567–576.
Hager B, Clayton R, Richards M, Comer R, Dziewonski A 1985: *Nature*, 313, 541–545.
Haurwitz B, Cowley A D 1973: *Pageoph.*, 102, 193–222.
Herring T, Gwinn C, Shapiro I 1986: *J. Geophys. Res.*, 91, 4745–4755.
Herring T, Buffett B, Mathews P, Shapiro I 1991: *J. Geophys. Res.*, 96, 8529–8273.
Hide R 1969: *Nature*, 222, 1055–1056.
Hinderer J, Crossley D 1993: *Geophys. Monograph Series, Am. Geophys. Union*, 72, 1–16.
Hinderer J, Legros H 1989: *Geophys. J.*, 97, 481–495.
Hinderer J, Legros H 1991: In: Proc. 11th Int. Symp. Earth Tides, J Kakkuri (ed.), Schweitzerbart. Verlag, Stuttgart, 425–429.
Hinderer J, Legros H, Amalvict M 1982: *Geophys. J. Roy. astr. Soc.*, 71, 121–132.
Hinderer J, Legros H, Jault D, Le Mouël J L 1990: *Phys. Earth Planet. Int.*, 59, 329–341.
Hinderer J, Legros H, Crossley D 1991a: *J. Geophys. Res.*, 96, 20257–20265.
Hinderer J, Florsch N, Mäkinen J, Legros H, Faller J E 1991b: *Geophys. J. Int.*, 106, 491–497.
Hinderer J, Zürn W, Legros H 1991c: In: Proc. 11th Int. Symp. Earth Tides, J Kakkuri ed., Schweitzerbart. Verlag, Stuttgart, 549–555.
Hinderer J, Crossley D, Xu H 1993: *Bull. Inf. Marées Terrestres*, 116, 8612–8626.
Hinderer J, Crossley D, Xu H 1994: *Geophys. J. Int.*, 116, 252–266.
Hinderer J, Florsch N, Mäkinen J 1995a: Calibration of a superconducting gravimeter from repeated absolute gravity measurements, paper presented at the XXI IUGG Gen. Assembly, Boulder, USA.
Hinderer J, Crossley D, Jensen O 1995b: *Phys. Earth Planet. Int.*, 90, 183–195.
Hinderer J, Crossley D, Xu H 1995c: The accuracy of tidal gravimetric factors and nearly diurnal free wobble resonance parameters in superconducting gravimetry, in Proc. 12th Earth Tide Symp., Beijing, 289-309.
Jensen O, Hinderer J, Crossley D 1994: *Eos*, 75, 69.
Jensen O, Hinderer J, Crossley D 1995: *Phys. Earth Planet. Int.*, 90, 169–181.
Lambeck K 1980: The Earth's variable rotation: Geophysical causes and consequences, Cambridge University Press
Langley R B, King R W, Shapiro I I, Rosen R D, Salstein D A 1981: *Nature*, 294, 24–31.
Legros H, Amalvict M 1989: In: The Earth's rotation, in Physics and evolution of the Earth's interior. Vol. 4, Gravity and Low-frequency Geodynamics, R Teisseyre ed., Elsevier, Amsterdam, 187–294.
Legros H, Hinderer J, Lefftz M, Dehant V, 1993: *Phys. Earth Planet. Int.*, 76, 283–315.
Mansinha L, Smylie D, Sutherland B 1990: *Phys. Earth Planet. Int.*, 61, 141–148.
Mathews P, Buffett B, Herring T, Shapiro I 1991a: *J. Geophys. Res.*, 96, 8219–8242.
Mathews P, Buffett B, Herring T, Shapiro I 1991b: *J. Geophys. Res.*, 96, 8243–8257.

Melchior P 1983: The tides of the planet Earth, 2nd edn., Pergamon Press, Oxford
Melchior P 1995: *Phys. Earth Planet. Int.*, 88, 223–256.
Melchior P, Ducarme B 1986: *Phys. Earth Planet. Int.*, 42, 129–134.
Melchior P, Crossley D, Dehant V, Ducarme B, 1988: *Geophys. Monograph Series, Am. Geophys. Union*, 72, 1-12.
Merriam J B 1992: *Geophys. J. Int.*, 109, 488–500.
Merriam J B 1994: *Geophys. J. Int.*, 119, 369–380.
Merriam J B 1995: *Geophys. J. Int.*, 123, 529–540.
Morelli A, Dziewonski A M 1987: *Nature*, 325, 678–683.
Neuberg J, Hinderer J, Zürn W 1987: *Geophys. J. Roy. astr. Soc.*, 91, 853–868.
Neuberg J, Hinderer J, Zürn W 1990: *Geophys. Monograph Series, AGU*, 59, 11–16.
Peterson J 1993: In: U. S. Geological Survey Report, 93–322.
Platzman, G W, Curtis G A, Hansen K S, Slater R D 1981: *J. Phys. Oceanogr.*, 12, 549–669.
Richter B 1995: In: Proc. ECGS Workshop on 'Non-tidal gravity changes: intercomparison between absolute and superconducting gravimeters', Cahiers du Centre Européen de Géodynamique et de Séismologie, Walferdange, Luxemburg, 11, 125–146.
Richter B, Zürn W 1988: In: The Earth's rotation and reference frames for Geodesy and Geodynamics, A Babcock and G Wilkins eds, Kluwer, Dordrecht, 309–315.
Richter B, Wenzel H G, Zürn W, Klopping F 1995: *Phys. Earth Planet. Inter.*, 91, 131–148.
Rieutord M 1991: *Geophys. Astrophys. Fluid Dyn.*, 59, 185–208.
Rochester M, Peng Z 1993: *Geophys. J. Int.*, 113, 575–585.
Rydelek P, Zürn W, Hinderer J 1991: *Phys. Earth Planet. Inter.*, 68, 215–229.
Sasao T, Okubo S, Saito M 1980: In: Proc. IAU Symp. 78, Nutations and the Earth's Rotation, E P Fedorov, M L Smith, P L Bender eds, Reidel, Dordrecht, 165–183.
Sato T, Tamura Y, Higashi T, Takemoto I, Nakagawa I, Morimoto N, Fukuda Y, Segawa J, Seama N 1994: *J. Geomag. Geoelectr.*, 46, 571–586.
Slichter L B 1961: *Proc. Nat. Acad. Sci., USA*, 47, 186–190.
Smylie D E 1992: *Science*, 255, 1678–1682.
Smylie D E, Rochester M 1981: *Phys. Earth Planet. Int.*, 24, 308–319.
Smylie D E, Jiang X, Brennan B J, Sato K 1992: *Geophys. J. Int.*, 108, 465–490.
Smylie D E, Hinderer J, Richter B, Ducarme B, Mansinha L 1993: *Geophys. Monograph Series, Am. Geophys. Union*, 72, 91–96.
Smylie D E, Hinderer J, Richter B, Ducarme B 1994: *Phys. Earth Planet. Inter.*, 80, 135–157.
Varga P, Hajósy A, Csapó G 1995: *Geophys. J. Int.*, 120, 745–747.
Volland H 1988: Atmospheric tidal and planetary waves. Kluwer Academic Publ., Dordrecht
Wahr J 1981: *Geophys. J. R. astr. Soc.*, 64, 677–704.
Wahr J, Bergen Z 1986: *Geophys. J. astr. Soc.*, 87, 633–668.
Wahr J, De Vries D 1989: *Geophys. J.*, 99, 511–519.
Wang R 1994: *Geophys. J. Int.*, 117, 562–565.
Warburton R J, Goodkind J M 1977: *Geophys. J.*, 52, 117–136.
Warburton R J, Goodkind J M 1978: *Geophys. J.*, 48, 281–292.
Wielandt E, Streckeisen G 1982: *Bull. Seism. Soc. Am.*, 66, 987–996.
Wu W-J, Rochester M 1994: *Phys. Earth Planet. Int.*, 87, 137–154.
Zürn W, Richter B, Rydelek P, Neuberg J 1987: *Phys. Earth Planet. Int.*, 49, 176–178.
Zürn W, Wenzel H G, Laske G 1991: *Bull. Inf. Marées Terr.*, 110, 7940–7952.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 329–337 (1996)

ABSOLUTE GRAVIMETRY IN THE STUDY OF SPACE AND TIME VARIATIONS OF THE GRAVITY FIELD

I MARSON[1]

The modern ballistic methods to determine the absolute value of the gravity acceleration have a great variety of applications ranging from the classical problem of geodetic metrology, the determination of the gravity datum and eventually of the gravity scale, to the study of time dependent gravity variations of long period due to geodynamics or global phenomena. Such projects take advantage of the high accuracy (at the μGal level), high stability (virtually drift free) and high transportability of the absolute gravity meters. The measurements are performed at given sites in a given, generally rather short (1 or 2 days), time. The choice of the sites is guided by their geological stability (for geodetic metrology purposes) and by their geodynamical environment (in case of studies of geodynamic processes). All sites are however affected by a gravity variability at short periods mostly due to changes of environmental parameters. To increase the reliability of absolute gravity measurements, efficient methods to separate the wanted signal from the environmental noise must be employed.

Keywords: absolute gravimetry; ballistic gravity meters; gravity noise; sea level change; time-dependent gravity

1. Introduction

In an experiment of absolute gravimetry one measures the acceleration due to the gravity force of the Earth by direct measurement of space and time of a free falling mass in a vacuum. The greatest part of the instruments are based on the free-fall method (the most common are JILAg and FG5); only few are based on the symmetrical rise-and-fall principle (the most known are IMGC and JAEGER-BIPM). The motion of the mass (a retroreflecting mirror) is detected by an interferometer. The time of occurrence of predetermined interference fringes is measured by a high precision time interval counter. The motion of the moving mass is subjected only to the gravity force, whereas the position of the reference mirror of the interferometer is affected by microseismic noise and vibrations induced by the dropping or throwing mechanism. The value of g is determined after averaging over a certain number of independent measurements: typically 100 for the rise-and-fall instruments and 1000–2000 for the free-fall ones. Generally, the mean square error, for both types of instruments, is of the order of 1 μGal, the total error budget of the order of 3 μGal. The acquisition rate is typically of 10–20 seconds for the free-fall instruments and 2-3 minutes for the rise-and-fall ones. A minimum set of 16 data can be acquired in 3 (free-fall) or 30 (rise-and-fall) minutes. The highest Nyquist frequency of absolute gravimetry is therefore of $5 \cdot 10^{-3}$ or $5 \cdot 10^{-4}$ Hz. Metrological

[1]DINMA, University of Trieste, Via Valerio 10, I-34127 Trieste, Italy,
e-mail: marson@univ.trieste.it

1217-8977/96/\$ 5.00

standards of time (atomic clock) and space (wavelength of a He-Ne laser stabilised on an iodine absorption cell) are used. The measurement is therefore absolute (i.e. it does not depend from another measurement performed at a different site or in a different time) and virtually drift free. Since the measurement of g results from a direct measurement of time and space, theoretically it should be independent from the particular device used for the measurement. Absolute gravimetry is therefore a rather good method, if not the only one, to study time dependent gravity variations of long period (of the order of decades).

Absolute gravimetry is also the most efficient method to solve the basic problem of geodetic metrology, i.e. the definition of the gravity datum and of the gravity scale in a given geographical area. Modern absolute gravity meters are highly transportable so that several absolute sites can be established in a rather short time (one site can be observed in less than one day). Unlike relative gravimetry, the accuracy is totally unrelated with the distance in space and gravity of the observation points and the looping procedure, so time-consuming in relative gravimetry, is totally unnecessary. The definition of the gravity datum is today a rather important problem in the definition of a world-wide unified gravity system which could allow for the definition of a world-wide geoid, important for cartography, transportation, navigation and so on. The recent availability of large data bases of gravity data, not yet referred to a unified gravity system, has re-opened this classical problem of geodesy.

2. Time-dependent gravity variations

The study of crustal deformations plays a key role in the assessment of earthquake and volcanic hazards and in the determination of mean sea level changes. These three research topics have a self-evident social impact and require long terms projects, of the order of decades. Because the geodynamical processes which might originate the variations of the gravity field are slow and because the related signals are small so that it is necessary to allow that a certain time goes by so to build up a signal bigger than the significance threshold. This means that we should initiate the observations projects now and keep them running for decades so that the data might be available when they will be necessary. Very likely we are acquiring data which will be useful to the next generation or generations. It is therefore mandatory to set up our projects in the most accurate way as possible and to take all the precautions to avoid doubts or ambiguities.

A crustal deformation process implies a variation of the position (co-ordinates and height) and a variation of the gravity field. This last because the gravity field is directly affected by the variation of the position of the measuring point (mainly of the vertical component) and because the process which is causing the crustal deformation might also induce changes of density in the earth interior (for instance pre-seismic dilatancy), dislocation or transfer of internal masses (for instance in volcanic activity). It is rather clear, therefore, that the combination of gravity and position changes allows the computation of changes of the potential and can provide important information on the dynamics of the phenomena. Moreover, gravity changes can be used to validate observed height variations. It is therefore important to measure simultaneously the position and the value of the gravity acceleration of the selected site.

Absolute gravimetry measures g by counting wavelengths and measuring time intervals: metrological standards are used in both space and time measurements. Space geodetic techniques (VLBI, SLR and GPS) use lasers or microwave sources and time intervals to measure the "relative" position of the site. The metrologies of absolute gravimetry and space geodesy are therefore similar. The co-location of absolute gravimetry and space geodesy hence defines a sort of absolute metrology with which it is possible to determine an absolute reference system suitable for long term studies of crustal deformation phenomena.

So far, the focus of the attention has been on the peculiarity of absolute gravimetry to define gravity datum and gravity scale and to study time-dependent gravity variations of long period. There is no doubt that the metrological characteristics of the measurements give to absolute gravimetry a leading role in such studies. Always, but particularly when dealing with long term gravity variations, we require high accuracy of the measuring device and high stability of the observation site. As the accuracy requirements increase, more problematic becomes the issue of the gravity variability of the observation site. Several factors can affect this variability in the medium to short period components: microseismic noise, barometric pressure, water table, solid and fluid tides, loading effects, polar motion are the more important. Some of them could be efficiently modelled or computed if the corresponding physical parameters are given (for instance once known the co-ordinates of the pole it is rather straightforward to compute the associate gravity effect). However in most cases the problem of removing the medium-short period components can be solved only by direct measurements.

In a long term project we are aiming at the detection of a signal of low amplitude (of the order of few μGal per year) and long period (of the order of year or years) affected by a noise which can have a higher amplitude than the wanted signal and periods ranging from few seconds (microseismic noise) to several months (water table).

An appropriate measurement strategy can be helpful in the attempt to increase the signal to noise ratio. To this purpose, we recall that the sampling rate of the absolute gravity meters ranges from 2–3 minutes in case of symmetrical rise and fall to 10–20 seconds in the case of the free fall ones. If we acquire data at the highest sampling rate continuously for, let us say, two days, the components of shortest periods can be detected directly from the absolute measurements. So we could safely state that the noise components with periods ranging from few minutes to half a day can be detected and removed from the data just analysing the absolute data set. If we increase the observation time to weeks or months it is possible, at least from a theoretical point of view, to detect also components of longer periods. The typical measurement strategy is however of the first kind (sampling rate of 10–20 seconds and observation period of one-two days) which limits the frequency range of the noise components which can be detected. In any case, it is obvious that the most critical components are those of longer periods.

The most efficient method to deal with these noise components requires the use of a superconducting gravity meter. As known these instruments are able to detect changes in the gravity field with a precision better than 0.1 μGal by measuring the variation of the intensity of the current flowing in a superconducting coil to keep the test mass in the reference position by means of magnetic levitation. The sampling rate of a superconducting gravity meter is higher than that of the fastest absolute gravimeter so that it has a clear advantage in the high frequency range. A superconducting gravity meter is a relative instrument which requires a calibration, and might be affected by a certain degree of drift. For these reasons a superconducting instrument is an excellent instrument to measure gravity changes with periods from seconds to year(s), but not for long term studies, at least for the very long term ones. Both calibration and drift characteristics could be assessed by frequent co-location of absolute and superconducting gravity meters. There is a mutual interest in the co-location of these two types of instruments: the absolute one can calibrate and assess the drift rate of the superconducting one, while this last can measure the noise components which are aliased by the inadequate sampling interval of the absolute gravity meter.

The ideal observation site for long term studies would be, therefore, equipped with a permanent GPS receiver, a superconducting gravity meter periodically (two to four times per year) controlled by an absolute gravity meter. The cost for installing and running such a permanent site is rather high and it is affordable only by few agencies or institutes in the world.

Because of their amplitude and period, rather critical are the gravity effects of water table and air pressure changes. The influence of water table changes on gravity measurements has been carefully studied by Delcourt-Honorez (1995 and related references). In these extensive studies, the hydrogeological perturbing effect on the local gravity at the Royal Observatory of Belgium in Brussels has been computed by measuring the water level variations at long, short term and due to Earth tidal variations. Several periods have been observed in three aquifers at different depths (-36 m, -62 m and -67 m) with different amplitudes ranging from 0.07 m to 0.18 m. It is interesting to remark that these relatively small water table variations have been detected in the gravity records acquired with a superconducting gravity meter and properly modelled. The total hydrogeological effect is the result of the sum of the gravity effect induced by the water masses in motion and of the gravity effect induced by the land surface displacement. A maximum effect of 4 μGal has been observed. Gravity effects up to 30 μGal induced by water table variations of about 8 m, with seasonal periods, have been also observed (Giorgetti et al. 1987). Long term gravity variations studies need therefore the measurement and modelling of the water table effect. This obviously requires the availability of water wells, knowledge of the physical characteristics of the aquifers and geological structures and efficient mathematical models. If these data are not available, one might make use of the regional water table data, which should be generally available. A superconducting gravity meter, or at least an earth tide gravimeter, could be installed for at least one year at the base station and the acquired gravity data could be used to compute the gravity admittance of the regional water table data.

The effect of the atmospheric pressure on gravity measurements has been deeply studied by Sun et al. (1993). Atmospheric pressure changes affect gravity through direct gravitational attraction of the air masses and through the induced elastic deformation of the Earth. The authors computed the gravity changes at the surface of a spherical, radially stratified elastic Earth, under the action of the atmospheric pressure by performing a numerical convolution between the mass loading Green functions and the local and regional barometric pressure data. Rather interesting is the geographical distribution of the barometric data which extends to more than 1000 km around the station and over 6 layers until about 12 km above the earth's surface. The conclusions reached by the authors show that the maximum peak to peak amplitude of the direct Newtonian attraction reaches 22.3 μGal at Brussels for an extension of the column load of more than 1000 km. The corresponding admittance is of -0.451 μGal/mbar which is larger than the theoretical value of -0.419 μGal/mbar. Another interesting result is that the atmospheric effects on gravity are affected by the lateral extension of the column load. In fact the admittance of the total gravity effect varies from -0.395 μGal/mbar for loads of 100 km of radius to -0.333 μGal/mbar for loads of more than 1000 km. Regional barometric pressure data over 1000 km around the station are hardly available. Fortunately about 80 percent of the total effect can be computed using only the first 100 km, which could be an achievable target. Also in this case it could be worthwhile to install a recording gravity meter for at least one year and compute the gravity admittance of the regional barometric pressure data.

3. Overview on ongoing projects

The political events of these last years have opened large gravity data sets acquired in several countries in Central-Eastern Europe which were considered as strictly confidential. This has brought up the problem of establishing a unified gravity system. More than twenty years ago a world-wide gravity reference network (IGSN71) was established. It was the result of a common adjustment of more than 20000 relative gravity observations and less than 10 absolute gravity ones. Today we are approaching this problem entirely with absolute gravimetry through a joint effort of several agencies operating absolute gravity meters and hosting countries. In a rather short time, all the European countries will have a unified gravity reference system. Also national fundamental gravity networks are being mostly, if not entirely, based on absolute sites (Austria, Germany, Italy). A third and good example of the use of absolute gravimetry for the establishment of the gravity datum concerns with the need to define it in the Antarctic continent in the framework of the several geodetic and geophysical projects which are being performed in that continent. This is clearly a case of a remote area where the difficulties and uncertainties associated with relative gravimetry put serious doubts on the usefulness of relative gravimetry for this purpose. It should be also kept in mind that several summer bases can be reached only by ship after something like 15 days of navigation. On the contrary, absolute gravimetry offers an easy and very accurate mean to solve this problem.

Among the long term projects involving gravity variations with the time, I would like to draw the attention on the assessment of the mean sea level and the study of the sea level variations (Zerbini et al. 1995). Sea level has a particular interest in geophysics because the ocean itself is one of the major components of the earth system and because sea level is the boundary between the ocean and the atmosphere. The increased attention and evidences for anthropogenic changes induced in the Earth system has focused one particular aspect of sea level: a rise in global sea level is considered as one of the more severe consequences of the predicted global warming. Hence sea level could be an important indicator of global warming, especially if an acceleration of the sea level rise can be detected. Changes in global sea level due to volumetric changes in the ocean water are thought to constitute the climate signal in the sea level, which is due to two factors: a) melting of land based ice adds water in the ocean; and b) warming of the ocean increases the volume of the water. The extraction of the climate signal is a rather delicate task. One reason is the complicated mass budget of the ocean, which is not only affected by the mass changes in the cryosphere but also by a number of other factors such as groundwater exploitation, deforestation, reservoir construction, irrigation and so on. Other major difficulties arise from the problem of determining the global sea level rise from local observations. Currently satellite altimetry data are still not sufficient to determine the present global sea level rise. But even if they were, the historical data will remain of importance for the detection of any acceleration in the sea level rise as well as the study of local effects and impacts. Most of the available sea level data come from coastal tide gauges. Tide gauges measure local sea level relative to a benchmark on land. Thus, vertical crustal movement is one of the factors affecting the local relative sea level signal. Based on both geological evidence of past sea level changes and estimates of the mass added to the ocean as a consequence of the present warming, non-tectonic trends in relative sea level are expected to be of the order of 1 mm/year. On the other hand, interannual coastal sea level variations may reach up to 100 mm over a decade. These variations can seriously bias trends estimated from short records. Warrick and Oerlemans (1990) state that accurate trends can be computed only given 15–20 years of data, while Pirazzoli (1986) sets a lower limit at 50 years. Clearly the required record length depends both on the local magnitude of the decadal sea level variability and the tolerable uncertainty in the trend estimate. For the Mediterranean, Baker et al. (1995) showed that in order to have errors less than 0.5 mm/year, records spanning at least 40 years are required. This discussion addresses correctly the issue of the need of long term projects.

From an oceanographic point of view, vertical crustal movement is a perturbation in the sea level measurements. Tectonic movements, sedimentation, groundwater or oil extraction, all may result in vertical crustal movements of regional or even local scale. Post-glacial rebound contributes to regional scale vertical movements. Furthermore, changes in the surface mass distribution in both hydrosphere and cryosphere induce a viscoelastic deformation of the Earth affecting the global geoid and consequently the sea level. In many locations the crustal component is of the same order or even in excess of the long term sea level variations. Generally, vertical crustal movements with associated changes of the geoid are considered as a major factor masking the sea level changes due to changes in the volume of the ocean water. Nowadays, space geodesy provides the means to monitor, in a global reference system, horizontal as well as vertical velocities of stations on the Earth's surface to a high degree of accuracy. SLR, VLBI and GPS measurements are being used for tide gauge benchmark fixing thus unifying the tide gauge network. Fixing the tide gauge benchmarks at subcentimeter accuracy is necessary in order to be able to separate sea level variations from vertical crustal movements and potential changes which may originate from different phenomena.

Independently from space geodetic techniques, an alternative approach to monitoring site vertical velocities is provided by the measurements of absolute gravity at tide gauge benchmark. This because absolute gravity constitute a separate check on the vertical crustal velocities and contributes to assess models of crustal deformation at inland sites.

In the SELF project (supported by the European Union) selected tide gauges in the Mediterranean Sea, in Spain, Marocco, France, Italy and Greece, have been referred to a permanent shore mark. Simultaneously GPS observations were performed to tie the permanent shore mark to selected fiducial stations of the global reference system. This last is the one provided by the SLR/GPS solution SSC(DUT)94CO1R derived by Noomen et al. (1995). Simultaneously with GPS observations, two dual-frequency ground based transportable Water Vapour Radiometers have been also used at selected sites to improve the estimation of the tropospheric path delay due to water vapour content in the atmosphere. Absolute gravity has also been measured both at tide gauge and at the reference sites. As a matter of fact, at present, by means of SLR, VLBI and GPS techniques a single unified terrestrial reference frame is attainable that, with the independent check provided by absolute gravimetry, permits vertical surface elevation changes of the order of 1–2 cm to be detected. Space geodetic techniques and absolute gravity applied simultaneously in the above mentioned "absolute metrology" provide a global vertical datum accurate to 1 cm.

The Mediterranean lies across a major plate tectonic boundary. Most of its length is occupied by a major belt of Alpine folding, locally interrupted by coastal plains or deltas in which accumulations of alluvium and compaction is likely. Many parts of the coast are subjected to seismic and volcanic activities; subduction zones are presently associated with the Hellenic and the Calabrian Arcs. In this framework it is clear that absolute gravity can add valuable information to the studies of geodynamics of the Mediterranean area. Twenty-six absolute sites have been established in the SELF project. About half of them will be re-observed after three-four years, in the framework of a second project funded by the Commission of the European Union (SELF II) to assess the actual accuracy of the observations and to estimate the gravity variability of the observation sites. The gravity variability, as well as the gravity admittance of environmental parameters, such as barometric air pressure and water table changes, will be studied in a dedicated experiment which is being performed at the VLBI site of Medicina (Italy). In this experiment, a superconducting gravimeter will record the variations of the gravity field for at least one year. Three absolute gravity meters (2 free-fall (FG5) and one rise-and-fall (IMGC)) will be periodically co-located with the superconducting gravity meter to assess its drift rate and calibration factor. The co-location will also serve as intercomparison of absolute gravity meters. Local and regional scale air pressure and water table data will be acquired as well. The aim of the experiment is to delineate a measurement and data analysis strategy able to efficiently solve the problem of the detection and modellization of the environmental noise in the gravity signal.

In a discussion about long term gravity projects, a particular mention must be made to the Global Geodynamics Project (GGP) which is aimed at the monitoring of changes in the Earth's gravity field at periods from seconds to years. The measurements will be taken over a time span of 6 years at a small number of permanent observations sites where there is superconducting gravimeter currently installed. The 6-years period has been chosen as a minimum length of data required to separate annual and 14 month Chandler wobble components in the gravity record. Absolute gravimeters will be co-located with the superconducting gravity meters at least twice a year (four times is recommended). Precise global measurements of the Earth's gravity field are essential to answer a number of important questions in geophysics like the existence and detectability of internal gravity waves in the Earth's liquid core, the gravity effect of the global atmospheric loading and mass re-distribution on the solid Earth, the changes in gravity associated with slow and silent earthquakes, tectonic motions, sea level changes and post-glacial rebound.

4. Conclusions

The study of several phenomena involving the Earth's physics require long term observations, which, in certain cases, must be carried out for decades. If the physical process involves crustal deformation, changes of the densities of the Earth's interior, transfer of mass within the Earth or from atmosphere to oceans or vice-versa, absolute gravimetry has a certain particular role. This because it is the only geodetic method able to measure the value of the observed field instead of a space or time variation of the same field. Modern absolute gravimeters are highly accurate and transportable, which makes the method reliable and affordable. However the gravity variability of the measurement site can strongly bias an absolute measurement or the trend deduced from a series of absolute measurements. To monitor and study this variability a superconducting gravimeter frequently controlled by an absolute gravity meter constitutes the most efficient tool, even if it is rather expensive. The co-location of different measuring techniques is in fact guiding the most updated projects dealing with crustal deformations and sea level variations studies. In these cases the co-location includes space geodetic techniques (SLR, VLBI and GPS), absolute gravimetry and, in selected sites, superconducting gravity meters. This co-location assures the possibility to observe changes in the potential, poses mutual constraints on the measured vertical velocities of the earth's surface, and completes the kinematics picture, deduced from space geodesy, with a component of the dynamics, deduced from the changes of the gravity field. Several studies on the assessment of natural hazards (seismic, volcanic and sea level rise) can benefit from the use of this technique.

References

Baker T F, Cerutti G, Corrado G, Kahle H G, Marson I, Muller M, Paradissis D, Pezzoli L, Plag H P, Pomrehn W, Richter B, Romagnoli C, Spencer N E, Tomasi P, Tsimplis N M, Veis G, Verrone G, Wilson P, Zerbini S 1995: In: I. Devel., EUR 15921 EN, 3232-338. Troen ed., Climate Change and Impacts, Proc. Symp., Copenhagen, 1993. Eur. Comm. Sci. Res.

Delcourt-Honorez M 1995: In: Proceedings of the Second Workshop on Non Tidal Gravity Changes. Intercomparison between Absolute and Superconducting Gravimeters. Conseil de L'Europe. Cahiers du Centre Europeen de Geodynamique et de Seismologie, Vol. 11.

Giorgetti F, Marson I, Palmieri F, 1987: Ground water table changes detected by microgravity variations: an attempt to calculate the effective porosity. *Geologia Applicata e Idrogeologia*, 22

He-Ping Sun, Ducarme B, Dehant V, 1993: Effect of the atmospheric pressure on gravity measurements. Paper presented at the General meeting of International Association of Geodesy, Beijing, 1993.

Noomen R, Springer T A, Ambrosius B A C, Herzberger K, Kuijper D C, Mets G J, Overgaauw B, Wakker K F 1995: Crustal deformations in the Mediterranean area computed from SLR and GPS observations. *J. Geodynamics*, ????

Pirazzoli P A, 1986: *J. Coastal Res.*, 81, 1–26.

Warrick R, Oerlemans J 1990: In: Climatic Change, J T Houghton et al. eds, The IPCC Scientific Assessment. Cambridge University Press, 257–282.

Zerbini S, Plag H P, Baker T, Becker M, Billiris H, Burki B, Kahle H G, Marson I, Pezzoli L, Richter B, Romagnoli C, Sztobryn M, Tomasi P, Tsimplis M, Veis G, Verrone G, 1995: *Global and Planetary Change*, 14, 1–48.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 339–358 (1996)

GRAVITY FIELD APPROXIMATION BASED ON VOLUME ELEMENT MODEL OF THE DENSITY DISTRIBUTION

G Papp[1]

The rapidly increasing number of geoscientific data and databases referring to the inner structure of the solid Earth opens new deal also in gravity field modelling. From the point of view of gravity modelling the most essential geo-parameter is the volume density because it determines the gravitational field of the Globe. Possessing information about the 3D density distribution with a sufficient "density" in global and regional/local extensions as well, a suitable systematization, characterization and discretization of it — which is the phase of model generation — makes the effective application of forward modelling possible. In this lecture notes a short review of principles and basic problems related to the aims, construction and use of such models and also to the interpretation of the results obtained by the modelling are presented. Several numerical examples are used to demonstrate the capabilities of forward gravity modelling in order to give general ideas of the reader interested.

Keywords: density distribution; forward gravity modelling; mass points; planar approximation; rectangular prisms; spherical approximation; volume elements

Introduction

The description and modelling of the Earth's external gravity field has became a standard task of physical geodesy nowdays. Many times the need for high resolution approximations arises from fields not related directly to geodesy itself therefore it increased the scientific importance of geodetic activities in a great extent. Geodesy, however, can fulfill the new requirements and challenges only if it integrates the achievements and results which were obtained by other branches of geosciences, physics, mathematics etc. In this lectures notes — as an obvious example for the necessary interdisciplinary cooperation — the highlight is put on an old approach to the gravity field approximation based on the solution of Newton's gravitational potential integral:

$$V(x,y,z) = G \iiint_B \varrho(x',y',z')\ell^{-1}(x-x',y-y',z-z')dx\,dy\,dz\,, \qquad (1)$$

where V is the gravitational potential at $P(x,y,z)$, G is the gravitational constant, ϱ is the volume density distribution function inside the bounding surface B of the masses of the Earth, ℓ is the distance between $P(x,y,z)$ and $P'(x',y',z')$ having volume density of $r(x',y,'z')$ (Fig. 1).

So far the forward method of gravity field modelling was neglected because of the lack of detailed information about the inner density or mass distribution of

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

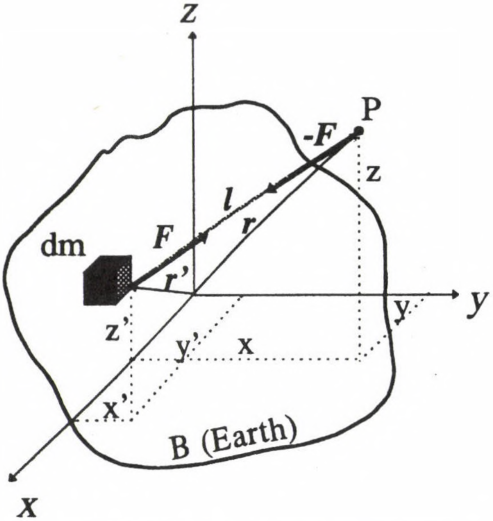


Fig. 1. Gravitational attraction of an infinitezimal dm mass element composing the body of the Earth at point P

the Earth. Without the sufficient and global knowledge of this Eq. (1) remains only a theoretical possibility and it provides approximations not competitive with the spatial resolution of other methods based on solutions of different kinds of Geodetic Boundary Value Problems. The recent maximum resolution for global potential models is 0.5 degree whereas for local cases it is some hundred metres or less (some tens of metres) if the gravity survey serves exploration purposes. This "unbelievable" broad spectrum can be obtained only by the combination of data having different information (spectral) content of the gravity field. On one hand the development of satellite techniques together with the decrease of gravimetrically unsurveyed areas provide the global and regional components whereas on the other hand the increasing availability of dense elevation models (Digital Terrain Models) of the physical surface of the Earth gives the local contribution to the high frequency end of the spectrum. This composition, however, already involves forward modelling since the so-called residual terrain correction computations applied extensively in the remove-restore technique of gravity field modelling utilize somehow Eq. (1) (Forsberg 1994):

$$\Delta g_{\mathrm{RTM}} = G\varrho \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{z=h_{\mathrm{ref}}(x,y)}^{z=h(x,y)} \frac{z - h_P}{[(x_Q - x_P)^2 + (y_Q - y_P)^2 + (z_Q - z_P)^2]^{3/2}} dx_Q\, dy_Q\, dz_Q \,, \tag{2}$$

where z is the elevation coordinate of the terrain surface represented by $h(x, y)$ DTM, h_P is the elevation of the actual computation point, ϱ is the constant density

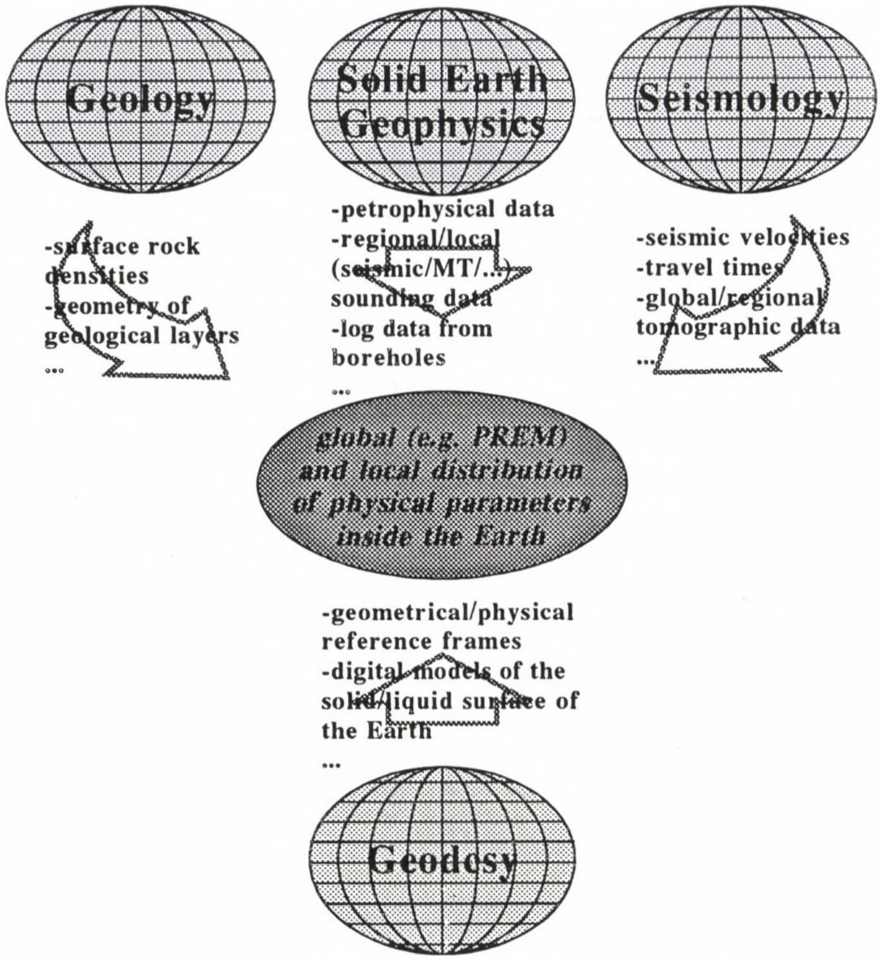


Fig. 2. Scheme of the contribution provided by different branches of geosciences

of the topographic masses and $h_{\mathrm{ref}}(x, y)$ is the DTM of the so-called mean elevation surface.

DTM-s are usually byproducts of photogrammetry, cartography etc., however, for the purposes of physical geodesy they represent the uppermost density inhomogeneities of the Earth's crust and the boundary surface for Molodensky's type BVP as well. But DTM-s are not the only one byproduct of other geoscientific branches which could be used for geodetic purposes. There is a lot of efforts done by geophysicists, geologists, seismologists etc. to explore the inner structure of the Earth and one faces a big challange if tries to collect and systematize all of this knowledge for a specific aim e.g. for forward gravity field modelling. Some problems related to the integration of information, the solution of Eq. (1) and the interpretation of results obtained by modelling will be discussed in the next sessions using as much practical examples as it is possible in the framework of this lecture notes.

Construction of volume density models

General questions

One has to face the reality. On one hand, probably, there is no hope to describe the density distribution function analytically (which implies necessary generalization and discretization of the problem (Houlding 1994)) and on the other hand at the moment data about the inner structure of the Earth in sufficient density are available only at a limited number of sites all over the world and even those — except some cases (e.g. Walach 1987, Schmidt and Götze 1996, Papp 1996) — are not integrated and arranged into a 3D form suitable for geodetic purposes. Because of the geometry of solid Earth exploration methods rather complex interpretation along profiles — called geotraverses (e.g. Posgay et al. 1996, Ádám et al. 1996) — is preferred and the extension of investigations in 3D still needs a lot of work on the well surveyed areas.

Obviously only those data can be used in gravity modelling which describes or somehow connected to density jumps and differences inside the Earth. Basically there are two kinds of data which can be used to build up a model of density distribution: geometrical and physical ones. It should be noted that for individual exploration methods the geometrical and physical parameters are usually correlated and the selection of one kind of the parameters determines the other types of parameters (e.g. seismic velocities and travel times versus depth of reflecting layers). Figure 2 shows why significant efforts needed to collect as much data as possible from different sources (different sources mean different types of exploration methods like seismic sounding, magnetotellurics, magnetic surveys, borehole drilling etc.). This is the only way to control the independence of input data for volume density model generation.

Usually two steps are neded in the process. As a first step of discretization subspaces (geological units) separated by discontinuities, faults and different lithological character must be located where the density distribution is supposed to be homogeneous and known. This is the creation of a geometrical base for the model and it involves digitization of geological maps, sounding profiles, borehole locations etc. From this raw data a 3D frame has to be formed having a structure which allows analytical solution of Eq. (1).

Second the units of the created frame must be characterised ("filled up") with volume density values which are usually density contrasts related to the reference density of the "average" crust or — in the case of topography — the real surface density values.

There are some possibilities to realize a 3D geometrical structure modelling the density or mass distribution. All of them are more or less discretized abstractions of the partly continuous real density distribution. Selecting one of them is the question of the purpose of gravity field modelling.

Basic aims and their classification in gravity field modelling

In this Lecture Notes the scientific and geodetic aspects of modelling are focused on and the exploration aspects are only touched.

Basically the scale of problems investigated distinguishes these two kinds of approaches. Geodetic research is related usually to global/regional problems whereas exploration geophysics is interested in the solution of local problems. At the moment the scale of problem is highly correlated with the attainable accuracy of modelling and therefore the accuracy of local investigations ($\leq \pm 0.1$ mGal) (Kearey and Brooks 1991) is better than that of the global/regional modelling ($\pm(1-5)$ mGal).

Concerning the starting conditions there are two possible directions (forward and inverse) of the modelling. The forward and inverse gravimetric problems represent the two sides of the same coin. In the forward problem the density distribution is known whereas the gravity related quantities (potential, gravity anomaly, derivatives of potential ...) are searched for. Therefore it is a properly-posed problem having unique solution.

In the inverse problem the casting is opposite. The density distribution is unknown whereas the external gravity related quantities are observables. Because of the change of role, the inverse problem is ill-posed and it has non-unique solution since the same gravitational field can be generated by unnumbered, absolutely different density distributions.

There are also differences concerning the primary aims of forward an inverse modelling.

The main purpose of forward modelling is to show how the geoscientific "evidences" (that is the explored density distribution) contribute to the observable picture of the gravity field whereas in inverse modelling the primary goal is not necessarily to derive realistic models of the density distribution from the gravity field but to obtain a (e.g. mass point) model which reproduces the observed gravity field with sufficient accuracy and higher speed than other methods (e.g. spherical harmonic expansion) can provide (Heikkinen 1981).

Obviously in the practice there is a strong interrelation between forward and inverse modelling and usually their subsequent use and combination provide the best results. Generally three steps are necessary to obtain it.

1. Computation of the effect of a "model of evidences".

2. Removing of this effect from the observations.

3. Interpretation of the residuals by inversion.

Conditions for volume element modelling of the density distribution

There are two basic conditions to be satisfied. First the model generated should represent a completed, overlapping-free filling of the space of available data. It means that where there are no gaps in the source data there should not be gaps in the model (continuity) and the volume elements should join to each other side by side and all around in the 3D space without overlapping (uniqueness) (Houlding 1994).

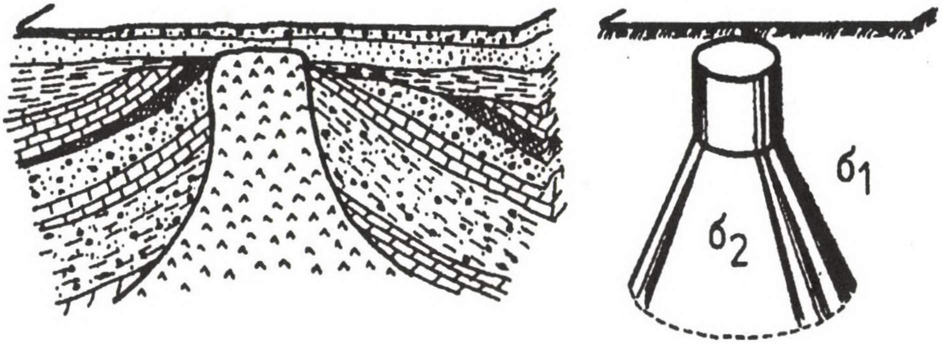


Fig. 3. Construction of a single disturbing mass (magmatic intrusion) by combining of different kinds — a cylinder and a truncated cone — of volume elements (Steiner 1982). σ_1 is the density of the surrounding e.g. sedimentary rocks whereas σ_2 is the density of the magmatic material

In exploration geophysics the spatial resolution of modelling has primary importance because of economical reasons and it can be increased only by the detailed and "as accurate as possible" mapping of the gravity field. Usually it makes no difficulties because very local and shallow geological structures (salt domes, oil traps, ore bodies, etc.) are to be explored the models of which can be built up by the combination of a small number of different types of (cylindrical, prismatic, spherical, conical, polihedral, etc.) volume elements in order to approximate the fine geometrical details in a way consistent with the resolution of the gravity field observed (Steiner 1982). But the construction of such particular and object-specific models (Fig. 3) cannot be fully automated therefore their control is very difficult. In geodetic applications because of the large number of global/regional data the automation of data manipulation cannot be neglected since it is the only way to fulfill the conditions above mentioned. That is why only homogeneous systems of volume elements with computationally efficient geometry are preferred. Regional examples (Kalmár et al. 1995) show that the fully automated generation of a geometrical frame for a density model based on rectangular prisms is possible if a dense digital model (a grid) of a subsurface density discontinuity is available. Even some kind of optimization (minimization of the number of prisms) can be obtained by appropriate algorithms. The gain of runtime in the computation of the gravitational effect of an optimized model may be 90 % in some situations whereas the loss of information (caused by the optimization/minimization process) is only 1–5 % generally relative to the information obtained from an elementary rectangular volume element representation based on the grid cells directly. Such order of decrease can be achieved only if the spatial density and the reliability of geometrical information and of the physical parameters are significantly different.

For example one may possess a detailed and accurate mapping of a subsurface interface obtained from seismics, magnetotellurics and magnetics but there is only a poor distribution of boreholes on the area investigated. Therefore rock samples give only site specific information about the volume density distribution. From sparse samples only general trends of the density distribution with low reliability can be

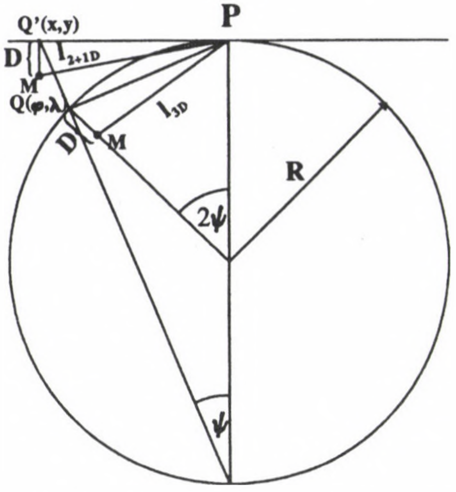


Fig. 4. Scheme of the stereographic map projection and transformation of 3D distances to 2+1D distances

recovered so it would be nonsense to use all the detailed information contained by the geometrical data. The reliability of the final results (e.g. gravity anomalies) will be, eventually, determined by the low reliability of the density values so the complexity of the geometrical structure may be adjusted (reduced) to the simplicity of the density distribution data.

An overview of mass points, rectangular prisms and other volume elements

The simplest and most abstract case of discretization is the so-called mass point approach whereas rectangular or curvilinear prisms represent a still simple but close-to-continuous approximation. A mass point possessing mass M_{point} can be considered as an infinitezimal volume element because its gravitational potential $V(r)$ is equal to the external ($r > R$) potential of a sphere of radius R and density ϱ if $M_{\text{point}} \equiv M_{\text{sphere}}$:

$$V(r) = G\frac{M_{\text{point}}}{r} = \frac{4}{3}\pi\varrho G\frac{R^3}{r}, \tag{3}$$

where G is the gravitational constant.

The simplicity of Eq. (3) makes the use of mass points very attractive, but the degree of abstraction limits its application. At the moment it seems that rather global investigations can benefit from its simplicity for representing the Earth's disturbing potential by an arbitrary set of mass points obtained by gravity field inversion. It would be difficult to interpret and connect such a model to real geological structures even (as it will be demonstrated) using mass points in order to realize known structures can lead to results distorted in some but not unusual situations.

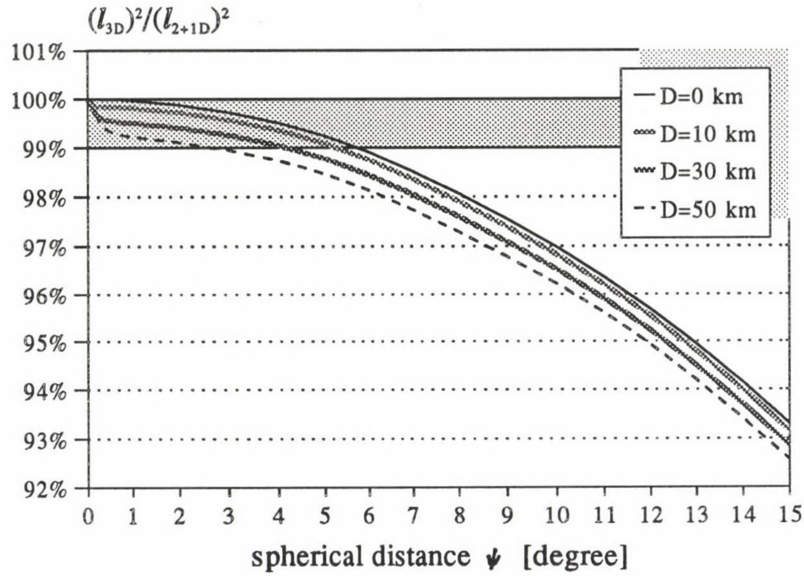


Fig. 5. Change of ratio of $(\ell_{2+1D})^{-2}/(\ell_{3D})^{-2}$ Newtonian gravitational mass attraction kernel functions in function of ψ and D (R = 6371 km)

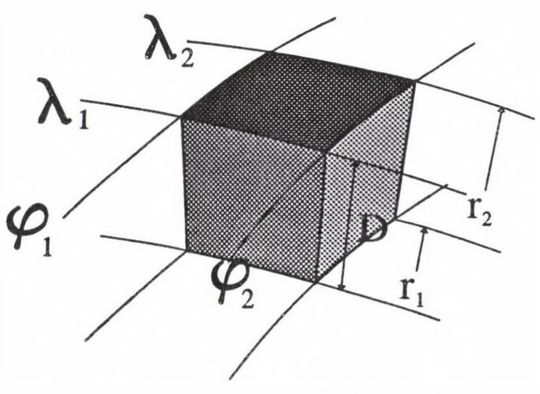


Fig. 6. Spherical volume element. φ is the geodetic latitude, λ is the geodetic longitude and D is the thickness of the element, r is the radial distance

The gravitational potential of a rectangular prism is also given in closed, analytical form (e.g. Nagy 1980):

$$V\Big|_{x_i y_j z_k}^{P} = G\varrho \Big| xy\ln(z+r) + yz\ln(x+r) + zx\ln(y+r) - $$
$$-0.5x^2\tan^{-1}\frac{yz}{xr} - 0.5y^2\tan^{-1}\frac{xz}{yr} - 0.5z^2\tan^{-1}\frac{yx}{zr}\Big|, \tag{4}$$

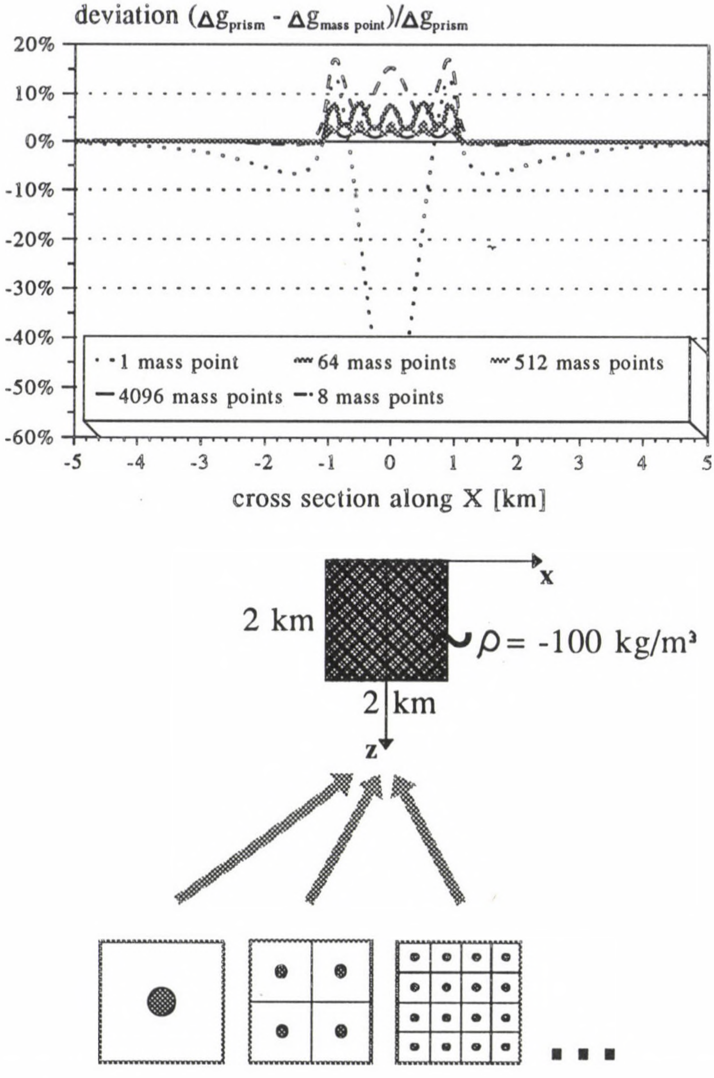


Fig. 7. Comparision of the gravitational effects generated by a single rectangular prism (2 km × 2 km × 2 km) and its mass point representations having different resolution

where $r = \sqrt{(x_P - x_i)^2 + (y_P - y_j)^2 + (x_P - x_k)^2}$ $(i, j, k = 1, 2)$ is the distance between the corners of the prism and the computational point P, ϱ is the volume density of the prism, x, y and z are coordinate differences between the coordinates of P and the coordinates of the actual corner of the prism. Since there are eight corners of a prism the computation of Eq. (4) with respect to indices i, j, k has to

Fig. 8. Horizontal arrangement of the prism system modelling the mass of Neogen-Quaternary sediments in the Pannonian Basin. Gray scale shows the depth of the bottom plane of a single prism

be repeated eight times and finally the results have to be summed up:

$$U_P = U\Big|_{x_2y_2z_2}^{P} - U\Big|_{x_2y_2z_1}^{P} - U\Big|_{x_2y_1z_2}^{P} + U\Big|_{x_2y_1z_1}^{P} - \\ -U\Big|_{x_1y_2z_2}^{P} + U\Big|_{x_1y_2z_1}^{P} + U\Big|_{x_1y_1z_2}^{P} - U\Big|_{x_1y_1z_1}^{P} . \qquad (5)$$

Comparing Eq. (3) and Eq. (4) one can see that the computation of the gravitational effect of a density model built up from prisms needs a lot of but not irrealistic runtime[2]. Characterizing the density distribution by a system of rectangular prisms very good regional geoid modelling results consistent at decimetre level ($\sigma_{\Delta N} = \pm(10 - 20)$ cm) with geoid modelling by using surface gravity data were obtained in the Pannonian Basin, Hungary (Papp 1996).

Analytical expressions are also available for other volume elements having more difficult geometry than the rectangular prism has (e.g. Talwani 1973, Meskó 1982). Many of them were derived to provide better geometrical approximation of single mass anomalies for obtaining precise information about their position, topology and extension. Probably the only one alternative of the rectangular prism in regional forward modelling is the polihedral approach combined with intensive usage of interactive computer graphics (Götze and Lahmeyer 1988). It can be very efficient if the source data referring to the structure of the Earth's interior are given in a form

[2]On an HP9000/720 workstation platform the average estimated runtime is about 3×10^{-4} s/prism/computation point. Computing the gravitational effect of a model consisting of 10000 prisms on a grid of 100×100 points takes about 2.8 hours.

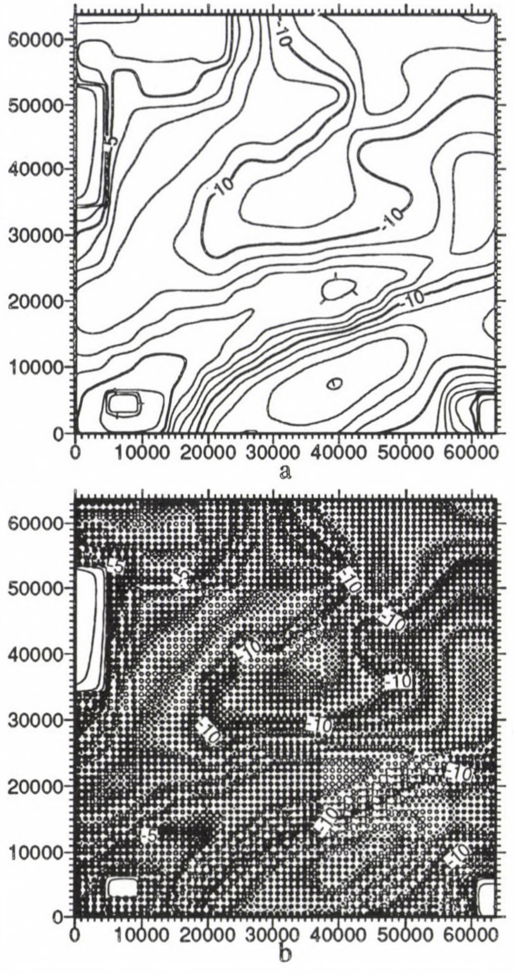


Fig. 9. Gravity anomaly contribution of the Neogen-Quaternary sediments modelled by a) a system of rectangular prisms having volume density of -100 kg/m^3 and b) its mass equivalent mass point representation. Contour interval is 1 mGal. Plane coordinates are given in metres in the map projection system used by Cadastral Survey of Hungary

of 1D/2D (e.g. seismic) profiles in a quasi-parallel position. In this situation the nodal points of the profiles represent the corners of joining polihedrons. This approach has already been applied succesfully in regional modelling of the lithospheric structure in e.g. the Central Andes (Götze et al. 1994).

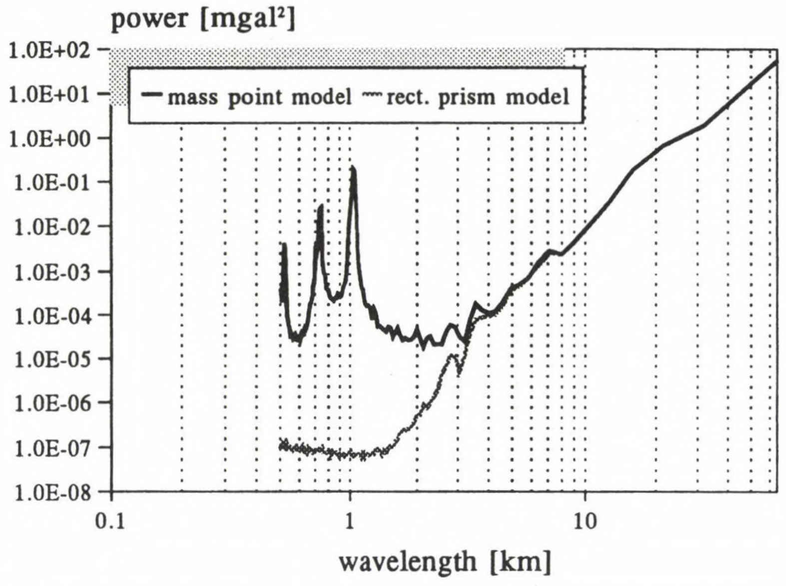


Fig. 10. Radial power spectra of gravity anomalies generated by mass equivalent mass point and rectangular prism representations of a specific model of the Neogen-Quaternary sediments in the Pannonian Basin (Kalmár et al. 1995)

Solution of the forward modelling in spherical coordinate system

It is obvious that the rectangular volume elements mentioned above can mostly be used in planar approximation. This approximation, however, valid only for a limited radius around the actual computational point because of the distortion implied by the map projection which is used to transform real 3D distances (ℓ_{3D}) determined by φ, λ, D to 2+1D distances (ℓ_{2+1D}) determined by x, y, D (see Fig. 4):

$$\ell_{3D}^2 = (R-D)^2 + R^2 - 2(R-D)R\cos 2\psi \tag{6}$$

$$\ell_{2+1D}^2 = D^2 + 4R^2 \tan^2 \psi\,, \tag{7}$$

where D is the depth of a volume element (e. g. mass point), R is the mean radius of the spherical Earth, ψ is the spherical distance from the pole of projection (explained in Fig. 4).

The magnitude of distortion of ℓ_{2+1D} relative to ℓ_{3D} is estimated from Fig. 5 which indicates that up to 500 km the effect of distortion (change of the Newtonian mass attraction kernel function < 1 %) may be neglected if crustal density anomalies are to be converted to gravity anomalies. Rigorous computation of ℓ_{3D} (from ℓ_{2+1D}) is possible by using exact equations of the projection used, but in the situation of regional modelling the curvature of the Earth can be taken into account by applying approximate correction ε:

$$\ell_{3D} = \ell_{2+1D} + \varepsilon\,. \tag{8}$$

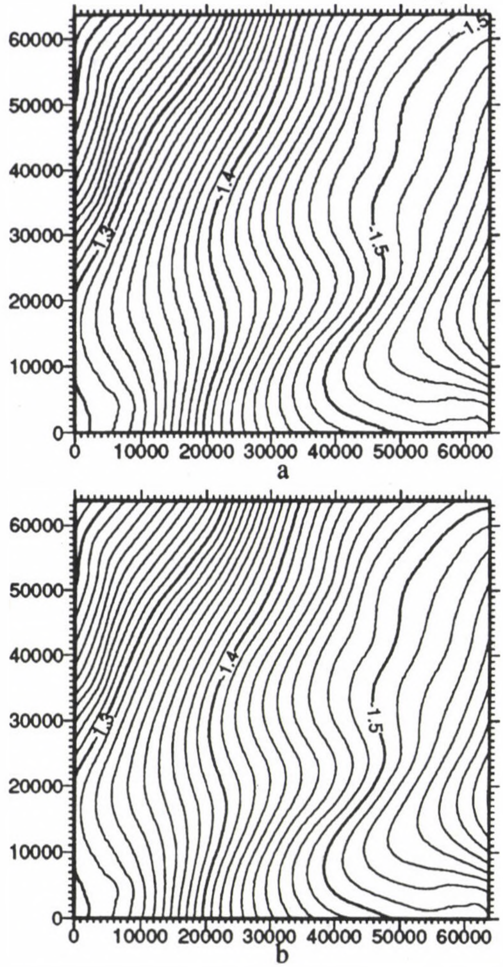


Fig. 11. Geoid undulation contribution of the Neogen-Quaternary sediments modelled by a) a system of rectangular prisms having volume density of -100 kg/m^3 and b) its mass equivalent mass point representation. Contour interval is 0.01 m. Plain coordinates are given in metres in the map projection system used by Cadastral Survey of Hungary

From Fig. 5 one can see that the ratio ℓ_{3D}/ℓ_{2+1D} may be approximated by a simple function depending on only ψ if $D \ll R$:

$$\frac{\ell_{3D}}{\ell_{2+1D}} = m(\psi, R, D) \simeq m(\psi) \tag{9}$$

$$\ell_{3D} = m(\psi)\ell_{2+1D} \,. \tag{10}$$

Substituting Eq. (10) into Eq. (8) a simple formula is obtained for computing

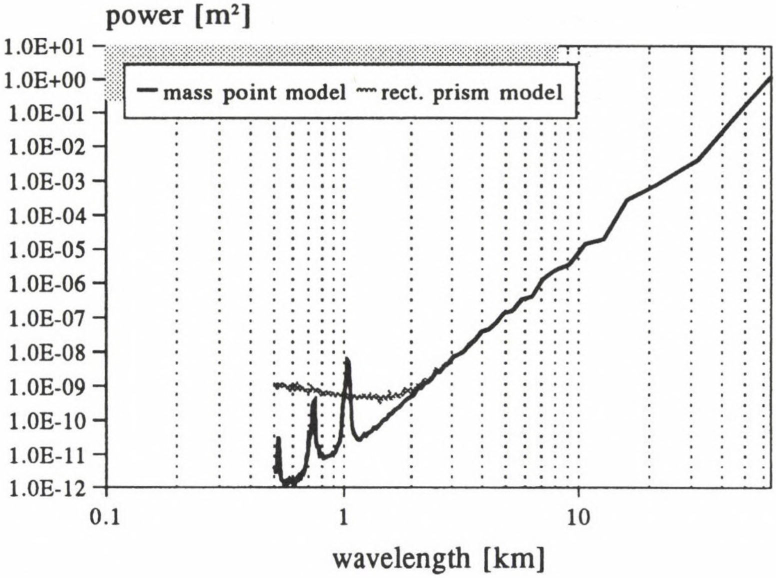


Fig. 12. Radial power spectra of geoid undulation contributions generated by mass equivalent mass point and rectangular prism representations of a specific model of the Neogen-Quaternary sediments in the Pannonian Basin (Kalmár et al. 1995)

corrections of ℓ_{2+1D} distances:

$$\varepsilon = \ell_{2+1D}(m(\psi) - 1) \tag{11}$$

from which new coordinates for e.g. the corners of the rectangular prisms can be derived.

The application of corrections can be avoided if spherical prisms defined by $\Delta\phi, \Delta\lambda$ and D, arcs of meridians and parallels and radial thickness (Fig. 6), respectively, are used to realize the geological model of the area investigated (Engels et al. 1995). For a global model this is the most obvious way to discretize and characterize the density distribution and this form allows the analytical computation of the gravitational potential utilizing spherical harmonic expansion of Newton's integral (Papp and Wang 1996):

$$\begin{aligned} V = G\varrho \sum_{l=0}^{\infty} \frac{R^{l+3}}{r^{l+1}} \frac{D_2 - D_1}{R} \cdot \\ \cdot \left\{ 1 - \frac{l+2}{2} \frac{D_2 - D_1}{R} + \frac{(l+2)}{2} \frac{(l+1)}{3} \frac{D_2^2 + D_2 D_1 + D_1^2}{R^2} + \ldots \right\} , \\ \sum_{m=-l}^{l} \overline{Y}_{l,m}(\Theta, \lambda) \int_{\lambda_1}^{\lambda_2} d\lambda' \left\{ \begin{array}{c} \cos m\lambda' \\ \sin |m|\lambda' \end{array} \right\} \int_{\Theta_1}^{\Theta_2} d\Theta' \sin\Theta' \overline{P}_l^{|m|}(\cos\Theta') , \end{aligned} \tag{12}$$

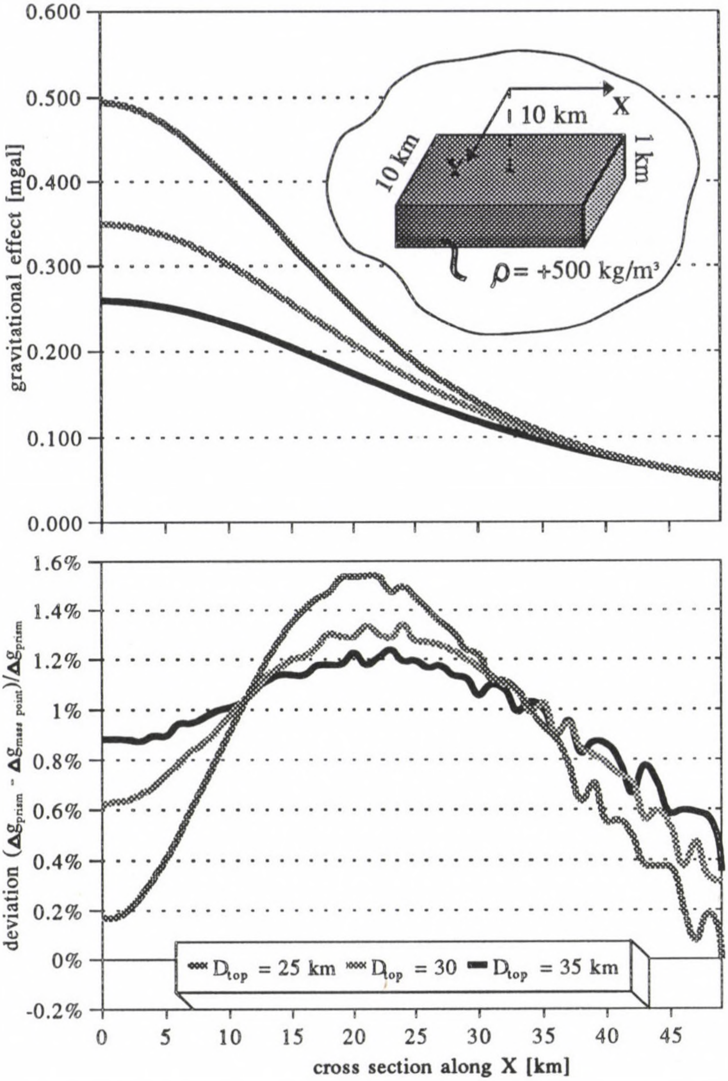


Fig. 13. Comparison of the gravitational effects generated by a single rectangular prism (10 km × 10 km × 1 km) and its single mass point representation located at the center of mass of the prism. Top: gravitational effect of the prism, bottom: relative difference curves

where $P_l^m(\cos\Theta)$ is the fully normalized associated Legendre function, D_1 and D_2 are the depth of the upper and lower boundary of the spherical prism, respectively, and $\overline{Y}_{l,m}(\Theta,\lambda)$ is the fully normalized surface spherical harmonic.

Unfortunately this natural approach results in time consuming computations which limits the use of high degree expansions. The maximum degree of expansion should be determined by the resolution of the density model itself therefore there

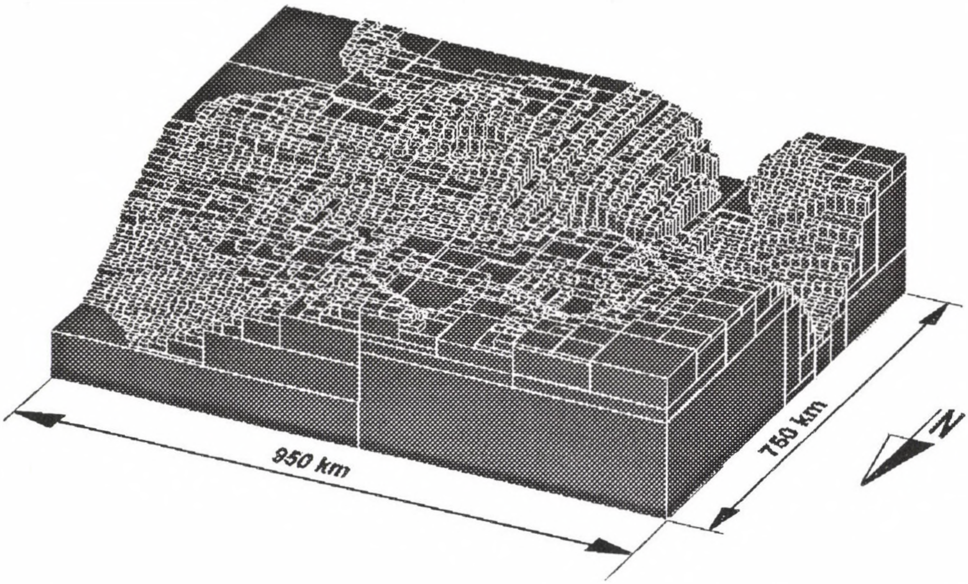


Fig. 14. 3D prism model of the upper mantle in a part of Central Europe. The vertical extension is 27 km (D_{top} = 25 km, D_{bottom} = 52 km)

is a loss of information if detailed models are available (Papp and Wang 1996). For example a near surface density model of the sediments defined by spherical prisms and having a moderate $0.1° \times 0.1°$ tangential resolution may generate a gravitational field which contains valuable spherical harmonic coefficients up to the degree of 3000–4000. The computation of such a high degree expansion, however, is limited by two factors: numerical precision and runtime.

Mass points versus rectangular prisms

The different characteristics of the gravitational fields generated by mass equivalent mass point and rectangular prism representations will be demonstrated by simple examples. There are two basic questions concerning the equivalence of this two kinds of discretization method: 1. how dense system of mass points is required to simulate the effect of a single prism and 2. how does the distance of the model from the computation point influence the convergence of these two kinds of solution.

Obviously it is suspected that in the situation of modelling of the near surface density distribution the differences are not negligable. Figure 7 shows that a rectangular prism of 2 km × 2 km × 2 km should be replaced by a big number of mass points if good agreement between the gravity fields generated by the mass point and the prism models is required. Good agreement, however, does not mean only statistically good agreement (Table I). Concerning the fine structure of the difference fields computed in the axial profile shown in Fig. 7 strong periodic features can be

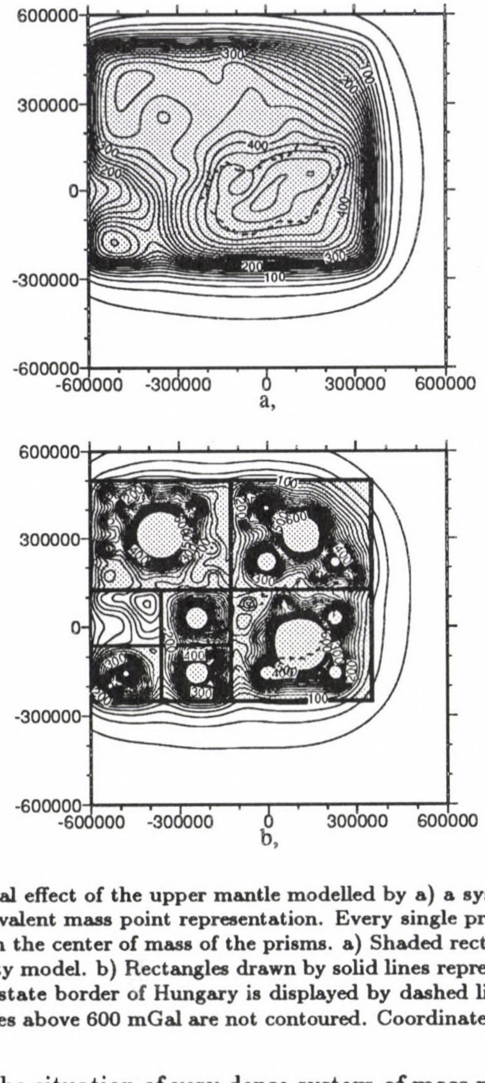


Fig. 15. Gravitational effect of the upper mantle modelled by a) a system of rectangular prisms and b) its mass equivalent mass point representation. Every single prism is replaced by a single mass point located in the center of mass of the prisms. a) Shaded rectangle shows the horizontal position of the density model. b) Rectangles drawn by solid lines represent the location of the six largest prisms. The state border of Hungary is displayed by dashed line. Contour interval is 20 mGal, anomalies above 600 mGal are not contoured. Coordinates are given in metres

observed even in the situation of very dense system of mass points (e.g. $16 \times 16 \times 16$ equally spaced mass points replacing the reference prism of 8 km^3). This artificial characteristic is resulted in by the mass point discretization itself and cannot be avoided. Mass points — no matter how densely they are spaced — generate such a high frequency signal in the synthetic gravity field. Its amplitude is a function of both mass point spacing and the distance between the model and the computation point. This effect can also be clearly seen if — instead of a single prism — the whole system of rectangular prisms of non-uniform size (Fig. 8) modelling the Neogene-Quaternary sediments of the Pannonian Basin (Kalmár et al. 1995) is replaced by a mass point model for forward gravitational field computation. In this numerical experiment every prism was represented by a 3D grid of mass points equally spaced

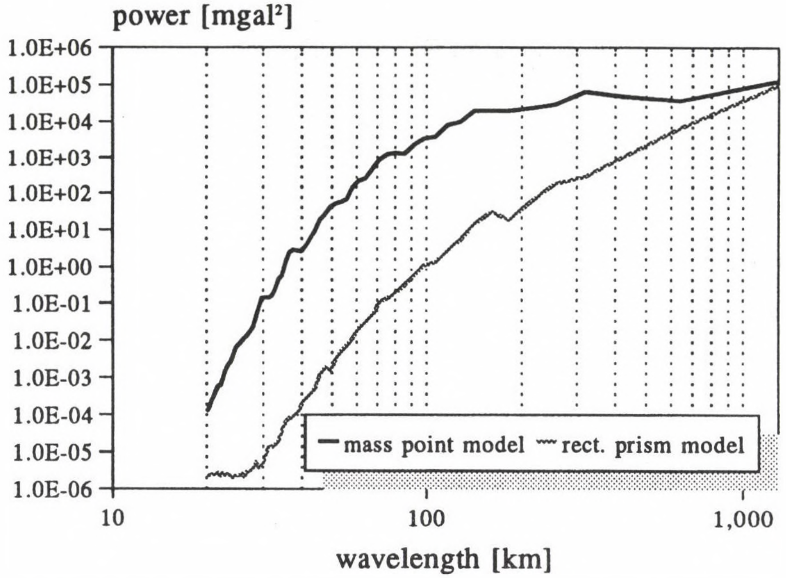


Fig. 16. Radial power spectra of gravity anomalies generated by mass equivalent mass point and rectangular prism representations of a specific model of the upper mantle below a part of Central Europe (Papp and Kalmár 1996)

by approx. 1 km and 0.5–1 km in horizontal and vertical directions, respectively. In this way more than 260,000 mass points covering an area of $\approx 110,000$ km^2 were used to substitute the density distribution of the sediments discretized by 3,303 rectangular prisms. The gravity anomaly and the geoid undulation contribution of both models are seen in Figs 9 and 11. In Fig. 9 the presence of the high frequency signal components generated by the near-surface mass points is so obvious that other components of the gravity field can hardly be recognized. The radial power spectrum of the gravity anomalies makes this phenomenon clear (Fig. 10) since strong and sharp peaks are observed in the (wavelength) range of the mass point spacing ($\lambda \leq 1$ km). Concerning the map of geoid undulations (Fig. 11) the distortion of the field generated by mass points is not obvious at all but the radial power spectrum uncovers the differences undoubtedly (Fig. 12).

But what about the modelling of density anomalies situated deeper than the sediments are?

Figure 13 shows that e.g. at the depth of the continental Mohorovičić discontinuity the difference between the gravitational effect of a single rectangular prism (having a spatial extension of 10 km $\times$ 10 km $\times$ 1 km) and its mass equivalent single mass point representation remains below 1–2 %. It indicates that below the depth of upper crustal density anomalies the mass point approach may lead to satisfactory results if some percent error is acceptable in the computations. It can also be, however, concluded from the shape of the difference curve that there is still systematic, long wavelength deviation between the gravitational effects obtained from the two

Table I. Statistics of 101 gravity anomalies computed in a profile along x (see Fig. 7) from a single prism and its different mass point representations. S.D. = Standard Deviation

Gravitational effect [mGal]	Prism	1 mass point	8 mass points	512 mass points	4096 mass points
Minimum	−3.46	−5.34	−3.11	−3.41	−3.40
Maximum	−0.04	−0.04	−0.04	−0.04	−0.04
Average	−0.90	−1.04	−0.82	−0.88	−0.88
Median	−0.26	−0.27	−0.26	−0.26	−0.26
S.D.	±1.17	±1.50	±1.05	±1.14	±1.15

approach. At the depth range considered in Fig. 13 a single prism of the specified size may be replaced by a mass point located in the centre of the prism. But if the dimensions exceed some tens of kilometres the mass point representation becomes as irrealistic as it is in the situation of near surface density anomalies. Using again a whole system of 1,820 mass points (located in the centre of mass of every single prism) instead of 1,820 rectangular prisms of different size (Fig. 14) modelling the mass of the upper mantle below a part of Central Europe (Papp and Kalmár 1996) the differences between the gravitational fields (Fig. 15) obtained from mass point and prism representations can be investigated. It is seen that locally there are significant discrepancies. At those places where large prisms — having some hundred kilometres horizontal extension in both X and Y directions — are substituted by single mass points, huge and irrealistic gravitational highs are present (empty circles without contour lines displayed). The power spectra (Fig. 16.) help to interpret the differences the main feature of which is the distinguishable larger power what is generated by the mass point model. It affects almost the whole range of spectrum but the strong high frequency components caused by the spacing of mass points cannot be identified in this situation.

Drawing the conclusions, for local/regional forward modelling mass point representation is not able to give satisfactory results. If the spatial density of mass points is high relative to the variation of the spatial distribution of geometrical and physical characteristics then in one hand the computational efforts required are unnecessarily increased and on the other hand high frequency distortions are introduced. If the number of mass points is in the same magnitude than that of the mass equivalent prisms then not only the high frequency part is affected by distortions but the whole spectrum of the gravitational field generated.

Acknowledgement

Large part of the work was supported by National Scientific Research Fund (OTKA) projects: F014284 and T0155960.

References

Ádám A, Szarka L, Prácser E, Varga G 1996: *Geophysical Transactions*, 40, 45–78.

Engels J, Grafarend E W, Sorcik P 1995: The gravitational field of topographic-isostatic masses and the hypothesis of mass condensation. Part I and II. Technical reports, Department of Geodesy, Univ. Stuttgart

Forsberg R 1994: In: Lecture Notes of the International School for the Determination and Use of the Geoid, F Sansó ed., IGeS, DIIAR, Politecnico di Milano, 101–134.

Götze H-J, Lahmeyer B 1988: *Geophysics*, 53, 1096–1108.

Götze H-J, Lahmeyer B, Schmidt S, Strunk S 1994: In: Tectonics of the Southern Central Andes. Reutter, Scheuber, Wigger eds, Springer Verlag, Heidelberg, 7–21.

Heikkinen M 1981: Solving the shape of the Earth by using digital density models. Finnish Geodetic Institute, Report 81:2

Houlding S W 1994: 3D geoscience modeling. Springer-Verlag, Berlin Heidelberg New York

Kalmár J, Papp G, Szabó T 1995: *Computers and Geosciences*, 21, 245–257.

Kearey P, Brooks M 1991: An introduction to geophysical exploration. 2nd ed., Blackwell scientific publications, Oxford

Meskó A 1984: Digital filtering. Applications in geophysical exploration for oil. Akadémiai Kiadó, Budapest

Nagy D 1980: *Acta Geod. Geoph. Mont. Hung.*, 15, 17–26.

Papp G 1996: Modelling of the gravity field in the Pannonian Basin (in Hungarian). Ph-D dissertation. Geodetic and Geophysical Res. Inst. of Hung. Acad. Sci., Sopron

Papp G, Kalmár J 1996: Toward the physical interpretation of the geoid in the Pannonian basin using 3-D model of the lithosphere. IGeS Bulletin, DIIAR-Politecnico di Milano, Milano (in press)

Papp G, Wang Z T 1996: *Acta Geod. Geoph. Hung.*, 31, 47–66.

Posgay K, Takács E, Szalay I, Bodoky T, Hegedűs E, Kántor J I, Tímár Z, Varga G, Bérczi I, Szalay Á, Nagy Z, Pápa A, Hajnal Z, Reilkoff B, Mueller S, Ansorge J, De Iaco R, Asudeh I 1996: *Geophysical Trans.*, 40, 1–44.

Schmidt S 1996: In: Proc. 7th Int. Meeting on Alpine Gravimetry. B Meurers ed., Special issue of Österreichische Beiträge zu Meteorologie und Geophysik, Wien, Heft 14, 137–144.

Steiner F 1982: Geophysics I. Gravity and magnetic methods (in Hungarian). Lecture notes. MNME Bányamérnöki Kar, Tankönyvkiadó, Budapest

Talwani M 1973: In: Methods in computational physics, Vol. 13, Geophysics, Academic Press, New York, San Francisco, London, 344–389.

Vermeer M 1995: *Manuscripta Geodetica*, 20, 362–378.

Walach G 1987: In: The gravity field of Austria, H Sünkel ed., Geodätische Arbeiten Österreichs für die Internationale Erdmessung, Neue Folge, Band IV, 3–10.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 359–370 (1996)

LABORATORY GRAVIMETRY

LABORATORY CALIBRATION OF GRAVIMETERS AND PROBLEMS RELATED TO THE GRAVITATIONAL CONSTANT

P VARGA[1]

Three severe problems of recent laboratory gravimetry are described. First of all the calibration of the gravimeters is discussed. On this basis a question of non-Newtonian gravimetry can be answered. Finally problems related to the gravitational constants G are described and a new method to determine its value is proposed.

Keywords: conversation of the angular momentum; gravimeters; gravitational earth tides; Newtonian and non-Newtonian gravity potential; torsion balances

1. Introduction

Laboratory gravimetry is concerned with experiments where the experimental conditions are under control, in contrast to observations where they are not. The experiments of this type can be done in rooms, in cavities under the Earth's surface. Instruments can be placed into carefully controlled enclosures and left undisturbed for long periods. Measurements which belong to the laboratory gravimetry are among others:

— observation of temporal variations of the gravity field

— determination of the gravitational constant

— determination of the absolute value of gravity

— calibration of gravimeters

— tidal gravimetry

— gravitational problems related to Earth rotation

— influence of external forces on the gravity field (e.g. hydrogeological effects; influence of air pressure)

— gravitational effect of the liquid core

The aim of present study is to discuss a possible way to obtain accurate and reliable calibration values for gravimeters in a laboratory.

On the basis of laboratory calibration it can be proved that the gravitational constant obtained in a laboratory has a more realistic value than that determined

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

by means of the gravity field of big geological masses. Other experimental and theoretical considerations show that there is no realistic temporal variation — in geological time scale — of the gravitational constant G. The gravitational constant was the first fundamental constant of physics. Nevertheless, its value is known with less accuracy. The torsion balance used in experiments to determine the value of G may not be the best way in which this unfavoruable situation can be altered. Possibly the solution can be the use of the special gravimeter calibration device equipped with a heavy cylindrical ring (Varga 1989, Varga et al. 1995) with superconducting gravimeters calibrated along calibration lines where gravity values were determined by means of absolute gravimeters.

2. Laboratory calibration of gravimeters

Beside the gravimeter manufacturers the first high precision laboratory absolute calibration device was prepared by Brein (1962). We can speak about three different calibration methods:

a) Parallel recording with two or more devices in this case the gravity record of a well calibrated gravimeter is compared with the ones to be calibrated. The calibration obtained in this way, however, is not an absolute determination of the instrumental scale. The accuracy obtained with this method can reach 0.1 percent. To make the calibration of this type absolute for the calibration an absolute ballistic instrument must be used. Up to the present with the use of them an accuracy better than 0.5 percent was reached (Varga and Hegymegi 1985, Hinderer et al. 1991).

b) Vertical acceleration of the gravimeter. This method was proposed by Brein (1962), realized by Valliant (1973), later by VanRuymbeke (1989) and Richter (1987). These authors used an oscillating platform and reached a relative accuracy of 0.25 percent (VanRuymbeke 1989).

c) Artifically induced gravity charges can be generated on different ways. Earlier it was common practice to calibrate gravimeters by the tilt of the instruments. In this way high quality and realistic calibration values cannot be achieved in principle because the instrument's system is in a deformed state during this operation. The other possibility of the generation of gravity changes is the vertical displacement of the gravimeters. In this way, however, only very small ($\sim 10\ \mu\text{Gal} = 100\ \text{nms}^2$) gravity changes can be generated (Bonatz 1971), the vertical gradient is very sensitive to the local mass disturbances and cannot be measured directly. Another possibility of calibration experiments of this type is: exposing the gravimeter to the influence of a big and well defined mass (see for e.g. Groten (1970) or Goodkind (1991)). In this way gravity variations (Δg) of the order of 100 μGal can be generated. This change in gravity is sufficient for an acceptable calibration accuracy, if no deformation occurs during the calibration process, and if both the geometry and the masses participating in the experiment are known with very high accuracy.

Calibration device of latter type was proposed and designed by Varga (1989) and it was installed in the Geodynamical Observatory Budapest in 1989–1990.

The principle of the equipment is simple: a suspended cylindrical ring with an inner diameter somewhat bigger than the width of the instrument to be calibrated is raised and lowered vertically and moved over the gravimeter equipped by a distant reading device and installed on a column of suitable height.

The advantages of this calibration procedure are as follows:

— the homogeneity of the generated gravity field at the extrema is very high;

— the raised and lowered ring does not load the ground around the instrument;

— the gravimeter remains stationary during the procedure what is a necessary condition for a small instrumental drift;

— the experiment is symmetrical with respect to the gravimeter and owing to technical reasons the gravity change brought about by the ring is greater than that caused by another geometrically regular body.

The disadvantages of the method are:

— this solution is immobile, it can be used only at the place where the device has been installed;

— the device in his recent form is rather sensitive to seismic noise.

All technical problems and the results of calibrations are described by Varga et al. (1991) and Varga et al. (1995).

It is worth mentioning that this device allows absolute calibration with an accuracy of (0.1–0.2) percent and it can be used to solve different problems of gravimetry:

— to study the non-linearity of the gravimeters (Götze and Meurers 1983, Becker 1984) and to detect small instrumental imperfectnesses (Meurers 1996);

— a calibration of tidal gravimeters is possible with an accuracy of 0.1 percent. This condition is important to prove the latitude dependence of earth tidal parameters and to discover mantle heterogeneities on the basis of gravimetric tidal records;

— to study the dependence of the gravitational constant G on the scale of the masses used for the determination of this fundamental constant of the nature.

The calibration device installed in the Geodynamical Observatory Budapest to generate gravity variations has a heavy stainless ring (density 8.0056 ± 0.0060 gcm^{-3}) the total mass of which is $3.2 \cdot 10^6 \pm 21$ g and generates a gravity change of ± 56 μGal. The first calibration of an LCR gravimeter was carried out by this device in August 1991 with an accuracy 0.2 percent. Calibrations carried out later on in 1993 support this error value, and the average of the runs carried out with two LCR instruments of the Eötvös Loránd Geophysical Institute of Hungary gave 0.1 percent uncertainty.

A device similar to ours to calibrate superconducting gravimeters was completed at ENEA Laboratories near Lago Brasimone (Italy) (Achilli et al. 1995). Because of the very high sensitivity of the criogenic gravimeters (of the order of 0.001 μGal) for a calibration precision approaching 0.3 percent a ring with a weight of 273 kg was only needed. The generated artifical gravity variation was 6.7 μGal.

To perform high quality calibrations numerous error forces must be investigated. These are described in detail by Varga et al. (1995). They are among others the gravimetric effect of air pressure variation (here the attraction and loading effect of the atmosphere as well as the buoyancy of the ring must be considered), the accurate position determination of the ring accuracy (in our case it was 0.1 mm and 1 arcsec), the magnetic effect (in the case of experiments carried out in Budapest the maximum magnetic variation of the ring was 14 μT what does not influences the LCR devices remarkably), the temporal variation of the digital voltmeter constant used to record the gravimeter output (1 percent variation of this value leads to a fictive gravity variation of the order of 0.1 μGal).

The gravity field exerted by the cylindrical ring can be exactly calculated along the axis of symmetry ($z = 0, y = 0$) at any point $P(\xi)$ with

$$\begin{aligned} g(\xi) &= 2\pi G\varrho\Big[\sqrt{R_2^2+(z_2-\xi)^2} - \sqrt{R_2^2+(z_1-\xi)^2} - \\ &\quad -\sqrt{R_1^2+(z_2-\xi)^2} + \sqrt{R_1^2+(z_1-\xi)^2}\Big] \end{aligned} \tag{1}$$

(here ϱ is the constant density of the annular mass, g is the gravitational constant, R_1 and R_2 are the inner and external radii, z_2 and z_1 denote the top and bottom of the calibration mass). To calculate the gravity field of the cylindrical ring outside of the axis is also of importance: the gravity effects caused by deviations from axial symmetry must be calculated. This can be done by the theorem of potential theory: if a body has an axial symmetry the gravity field caused by it along this axis determines the distribution of the gravity along any perpendicular to its direction (e.g. along x and y) expressed with the series of Legendre polinomials.

According to the experience accumulated during the operation with the calibration device developed by us it was concluded that the most serious error sources are the instrumental drift together with the earth tidal effect and the microseismic noise.

a) Removal of the tidal effect and the instrumental drift

For the Geodynamical Observatory Budapest on the basis of earth tidal observations carried out there in 1988–1992 the earth tidal parameters (amplitude ratio δ and phase difference κ) are known with high accuracy. A comparison campaign with the Askania recording gravimeter of the Technical University Prague proved that the reliability of δ factors of the Observatory for main tidal constituents is 0.1–0.2 percent (Table I).

The influence of Earth tides was removed with the use of the Cartwright-Taylor-Edden development and with the δ and κ values determined from gravity records

Table I. Gravity earth tide parameters determined for the Geodynamical Observatory Budapest (A) and results obtained with the gravimeter of the Technical University Prague (B)

Epoch	A 1988–1992		B 1988–1989	
	δ	κ	δ	κ
Diurnal Waves				
O_1	1.1515	+0.1°	1.1496	+0.1°
K_1	1.1378	−0.1°	1.1362	0.0°
Semidurnal Waves				
M_2	1.1813	−0.6°	1.1823	−0.8°
S_2	1.1764	−0.5°	1.1855	−0.5°

carried out in the Geodynamical Observatory Budapest. Of course this procedure has problems at the 10^{-1} μGal level needed for 0.1 percent calibration accuracy, because the tidal parameters are known with sufficient accuracy only for the few biggest tidal waves. The residual of the lunisolar effect can be excluded from the observation data together with the instrumental drift.

The accurate removal of the instrumental drift is one of the most important problems in the laboratory calibration procedures. For this purpose a hypothesis was used: the gravity effect of the ring on the gravimeter is the same at a given position of the ring and the measured Δg at this point is the drift. The tidal and atmospheric effects must be removed already before the determination of the drift curve. By repeated up and down movements of the ring an accurate drift curve can be obtained which allows to exclude this effect with the reliability of 0.1 μGal or even better.

b) The microseismic noise

Microseisms or seismic noise oroginate because of external influences first of all from the atmosphere and the sea. These waves always exist on seismic records but with varying intensity (Bath 1979, Bullen 1979, Bernhard 1900). The nature of the microseismic waves is not quite clear. Recent observations allow the following classification:

— short period microseisms ($T < 2$ s) depend on near disturbances and they decrease significantly with distance

— $T \sim 6$ s generated by cyclons of some hundred (up to 1000) km distance

— $T = 9 - 10$ s microseisms are produced by large low-pressure areas at greater distances

— $T = 17 - 20$ s microseisms are observed more seldom (few times per year). They can be ascribed to coastal effects.

The microseismic noise at the Geodynamical Observatory Budapest was recorded with an Askania type gravimeter. It can be concluded that at this place the microseismic activity has a seasonal amplitude variation of 5–30 μGal at periods between 5 and 10 s. What is very important for high quality laboratory gravimetric measurements: the observed microseisms undergo systematic beating with periods from 1 to 4 min. This last phenomenon is the most remarkable error source of the calibration procedure described in this paper. The corresponding gravity variations range from 1–2 μGal to 10–30 μGal. (This gravity amplitude values can be used only qualitatively because the linearity of the used Askania gravimeter was not proved at this periods.) The influence of microseisms can be significantly reduced by increasing the number of the observations. In this way the microseismic noise can be reduced to 0.2 percent even in the case of a single calibration procedure (Table II). Of course calibrations should be performed at times of low microseismic noise level.

Because of the need of very accurate determination of the extrema it is necessary to introduce adjustment calculations. This can be the least square method (the L_2 norm) in case of a Gaussian error (noise) distribution. In the observations of the gravity during the calibration procedure a number of outliers — possibly due to the long-periodic (1–5 minutes) beating of the microseisms — were detected which cannot be handled effectively with the commonly used least square method. For the adjustment calculations the robust estimation method must be used (Somogyi and Závoti 1993) possibly instead of least squares.

On the basis of our experience with the calibration device equipped with a heavy cylindrical ring it can be concluded that it is one of the most reliable methods of laboratory calibration of gravimeters. Its accuracy is at present 0.1–0.2 percent and it can be further developed by using more sensitive instruments and more convenient annular masses.

Table II. Root mean square (rms) error values in μGal relative to the expected theoretical variation of the gravity

Number of observations	rms errors
6	0.47 (0.40%)
12	0.41 (0.36%)
40	0.29 (0.25%)
50	0.25 (0.22%)
70	0.25 (0.22%)

3. Scale and time dependence of the Newtonian gravitational constant

There are several publications which report on the large-scale (10–1000 m) determination of gravity constants G_∞ with numerical values significantly differring from the results of G_0 values determined in laboratories. A theoretical basis of this deviation can be given by introduction of the so called non-Newtonian gravity potential. This phenomenon can be given analytically in the following form (Stacey et al. 1987):

$$V = -\frac{G_\infty m_1 m_2}{r}\left(1 - \alpha \cdot e^{-\beta}\right) = V_N + V_y \,. \tag{2}$$

If $\alpha = 0$, $V = V_N$ where V_N is the Newtonian potential. If $\alpha \neq 0$, V_y appears additionally which is called the Yukawa term. If $\beta \ll 1$ we get

$$G_0 = G_\infty(1+\alpha)\,. \tag{3}$$

On the basis of large scale Airy type experiments Stacey et al. (1987) got for G_∞ a value which is bigger than the G_0 value obtained in laboratories. The difference is $G_\infty - G_0 \sim 0.01$ what leads to $\alpha \sim 0.0075 \pm 0.0036$. To study this problem a special gravimeter calibration line was installed at the Geodynamical Observatory in Budapest under the surface of the Earth. This underground calibration line consists of 14 stations with a range of 100±1 μGal. Gravity differences, separation and the elevation difference between neighbouring stations are 100 μGal, 2–5 m and less than 2 cm, respectively. Because the line is horizontal it was possible to measure the differences with a computer controlled Eötvös torsion balance. The instrumental constant of the torsion balance is obtained by the measurement of the sensor masses, the length of the arm of the balance and the torsion of the wire. This means: the gravity values of the underground calibration line were determinded without the use of gravimeters. The difference of the calibration factors obtained for the same gravimeter along the calibration line (i.e. by means of gravitational effect of big geological masses) and by the cylindrical ring is of the order of 10^{-3} what means that the difference between the G_∞ obtained from large-scale Airy-type experiments and the laboratory type G_0 is at most 10^{-3}.

Another problem connected to the gravitational constant G is its temporal variation as supposed by many authors. This follows from Dirac's expanding Universe model proposed in 1937 what leads to a decreasing constant of gravitation and to the geophysical theory of the expanding Earth. On the basis of Dirac's theory Jordan concluded (1966) that the Earth radius a increases with a speed $da/dt = 0.5\ \mathrm{mm \cdot y^{-1}}$. Similar expansion value was obtained by Egyed (1957) $da/dt = 0.7\ \mathrm{mm \cdot y^{-1}}$ who supposed that originally the surface of our planet was as big as the areas of all recent continents together. The most recent and complete description of these theories can be found at Carey (1988).

The critical review of da/dt and consequently of dG/dt can be carried out on the basis of the study of the influence of earth tides on the long-term variations of the angular speed. Studies of this type are usually based on the principle of conversion of angular momentum and it is supposed that the Earth-Moon system is isolated. For the sake of simplicity it can be supposed that the Moon revolves around the Earth on a circular orbit in the plane of the terrestrial equator. The law of conversion of the angular momentum can be written as

$$\frac{d(Jw)}{dt} = \frac{1}{3}\frac{MM_m}{M+M_m}R_m^2\frac{dn_m}{dt}\,. \tag{4}$$

In Eq. (4) M, J, w is the mass, the inertia tensor and the angular speed of the Earth, respectively. M_m, R_m, n_m stands for the mass of the Moon, for the Earth-Moon distance and for the Moon's orbital speed. Kepler's law can be written as

$$n_m^2 R_m^2 = G(M+M_m)\,. \tag{5}$$

Its time derivative is

$$2n_m R_m^3 \frac{\partial n_m}{\partial t} + 3n_m^2 R_m^2 \frac{\partial R_m}{\partial t} = \frac{\partial G}{\partial t}(M+M_m) + G\frac{\partial (M+M_m)}{\partial t}\,. \tag{6}$$

In r.h.s. of Eq. (6) it can be evidently supposed that the time derivative of the gravitational constant is not a time dependent value ($\partial G/\partial t = C$) while the second term is equal to zero. Therefore

$$\frac{\partial n_m}{\partial t} = -\frac{3}{2}\frac{n_m}{R_n}\frac{\partial R_m}{\partial t} + \frac{\partial G}{\partial t}\frac{M+M_m}{2n_m R_m^3} = -\frac{3}{2}\frac{n_m}{R_m}\frac{\partial R_m}{\partial t} + C^*\,. \tag{7}$$

In r.h.s. of Eq. (7) C^* is of course a constant value. Introducing $\partial n_m/\partial t$ into the basic equation (4):

$$\frac{d(Jw)}{dt} = -\frac{1}{2}\frac{M_m M}{M_m+M}n_m R_m + \frac{\partial G}{\partial t}\frac{M+M_m}{2n_m R_m^3} = L_m + C^* \tag{8}$$

(L is the tidal torque).

From astronomical data (see e.g. Zharkov et at. 1996):

$$\frac{d(Jw)}{dt} \approx -4.1\cdot 10^{16} N\cdot m\,.$$

The total tidal torque composed by the atmospheric (L_{AT}), the earth (L_{ET}) and the oceanic (L_{OT}) tidal torques

$$L_m = L_{AT} + L_{ET} + L_{OT}\,.$$

With numerical values

$$\begin{aligned} L_{AT} &= 0.5\cdot 10^{16} N\cdot m \\ L_{ET} &= -0.5\cdot 10^{16} N\cdot m \\ L_{OT} &= -5.0\cdot 10^{16} N\cdot m. \end{aligned}$$

Consequently in Eq. (8) it is $\frac{\partial G}{\partial t} \geq 0$ what is in contradiction with Dirac's theory on the expanding Universe as well as with the theory of the expanding Earth, because an increasing gravitational constant requires a compressing Earth.

Table III. Gravity constant values published after 1969 (Chen and Cook 1993)

Authors	Year	$G \cdot 10^{-11} Nm^2 kg^{-2}$
Rose et al.	1969	6.6699±0.0014
Facy and Poinkis	1970, 1971	6.6714±0.0006
Renner	1974	6.668±0.002
Sagitov et al.	1978	6.6745±0.0008
Luther and Towler	1982	6.6726±0.0005
De Boer	1987	6.6670±0.0007
		6.6706±0.0028

Table IV. Relative errors of basic physics constants (A selection based on Brockhaus' lexicon Naturwissenschaften und Technik, 1989)

Avogadros constant	$5.2 \cdot 10^{-10}$
Boltzman's constant	$1.2 \cdot 10^{-4}$
Elementary charge	$2.8 \cdot 10^{-6}$
Faraday constant	$2.8 \cdot 10^{-6}$
Gravitational constant*	$8.5 \cdot 10^{-4}$
Mass of the neutron	$5.1 \cdot 10^{-6}$
Planck's constant	$5.5 \cdot 10^{-6}$
Rydberg's constant	$8.3 \cdot 10^{-6}$
Speed of the light	$4.0 \cdot 10^{-6}$

* The error value of the gravitational constant is the value given with the current CODATA value (Cohen and Taylor 1986)

4. The gravitational constant, its numerical value and accuracy

The value of the gravitational constant is known with less accuracy than other fundamental constants of physics. Authors of the best G value determinations claim to their experimental results an accuracy of 10^{-4}. Table III shows that the disagreement between the individual results is of the order of 10^{-3}.

Moreover it can be concluded that G is the least well known constant of fundamental physics (Table IV).

There are several explanations why G is known with such a low accuracy. First of all there is a "psychological problem": at this time there are no big research problems in the science which would urgently need a more accurate value for the gravitational constant. The second problem is connected with the weakness of gravitational attraction in scales used in laboratories. For example: a force interaction of two masses of $1g$ at the distance of 1 cm is 10^{-12} Newton while the pressure of the light of the sun is 10^{-10} Newton or the acting of forces between a proton and a neutron is 10^{-8} Newton. Additionally there is a metrological difficulty: G is defined by the fundamental quantities, time, length and mass to be determined in absolute scale. This circumstance leads of course also to experimental difficulties. The scatter of the data listed in Table III suggests that there can be a systematic error in the gravitational constant determined by different scientists. In fact the basic idea of the measurement of G is the same in all experiments. The heart of them is the torsion balance which was used in the beginning on static way and later on, after the succesful attempt of Eötvös at the very end of the last century, dynamically. It was discovered however by Kuroda in 1995 that the torsion force is dependent on the frequency with which the torsion bar is oscillating. The variability of the elastic constants is particularly significant at low frequencies of oscillating torsion balances used in laboratory experiments. According to Maddox (1995) the frequency dependence of the elastic parameters of the materials used in the torsion balances is the main source of the systematic and big differences between the different laboratory G determinations.

In spite of the recent lack of interest to the G and the considerable experimental difficulties it is important to try to increase the accuracy of the gravitational constant. It seems that one way can be in this direction the use of the laboratory calibration device developed by us. This experimental tool has a clear geometry and the used quantities (mass of the ring, the position of it etc.) are already or can be obtained with an accuracy necessary to determine G with relative error of 1 part in 10^4 (or even a few times 10^5). To reach this level in our knowledge about the value of the gravitational constant needs some development of the calibration device.

The influence of microseismic noise must be reduced significantly. As it was mentioned above the systematic beating with a period of some minutes caused by the microseisms characterised with periods between 5 and 10 seconds produces gravity variation of ~ 10 μGal. This influence must be reduced either with an appropriate antiseismic isolation or with a special feed-back system.

If the construction of new superconducting gravimeters allows an effective way to increase the gravity effect the reduction of the inner diameter of the ring used in the calibration device can be possible. If the inner diameter of the ring is reduced from 30 cm to 20 or to 15 cm, the corresponding gravity effect generated by the cylindrical ring of the mass of 3200 kg will be ± 89 or ± 118 μGal instend of ± 56 μGal.

The accurate removal of the instrumental drift (together with possible residuals of the lunisolar and meteorological effects) are of first order importance. To carry out more accurate drift determinations beside the more sophisticated measuring technics new statistical data processing methods — like the robust estimates — are also needed.

Of course to get uncertainties of 10^{-4} or even better the spring gravimeters — used earlier — must be replaced by transportable superconducting gravimeters with reduced diameter. Such "thin" instruments were demonstrated by the GWR company during the XXIst Assembly of the International Union of Geodesy and Geophysics (July 1995, Boulder, Colorado, USA). An important feature of these instruments is that their sensor has an axial symmetry.

The superconducting gravimeters must be calibrated first along the gravity lines measured with absolute gravimeters. The accuracy of these calibration lines in 10^{-5} (Atzbacher and Gerstenecker 1993). Afterwards with this gravimeters the gravity effect generated by a ring moved up and down must be measured. The gravity effect caused by the ring is known with an accuracy of $\sim 10^{-5}$. If the value of G is suitable the measured and generated-calculated gravity values must coincide. With other words by the comparison of these two gravity values the gravity constant can be obtained.

5. Conclusions

a) The methods of modern laboratory gravimetry — among them the device with a heavy cylindrical ring — allows calibration accuracy of 0.1 percent.

b) At this time there is no basis to conclude that the gravitational constant determined in a laboratory is different from those obtained with the use of gravity effect of big geological masses.

c) Similarly at present there is no ground to speak the about temporal variation of G.

d) There is a possibility to improve our knowledge on the numerical value of the gravitational constant with the use of the calibration device equipped with an annular mass.

References

Achilli V, Baldi P, Casula G, Errani M, Focari S, Guerzoni M, Palmonari F, Raguni G 1995: *Bulletin Géodesique*, 69, 73–80.

Atzbacher K, Gerstenecker C 1993: *J. Geodynamics*, 18, 1–4, 107–121.

Bath M 1979: Introduction to Seismology. Birkhäuser Verlag

Becker M 1984: *Deutsch. Geod. Komm.*, C, 294, München, 1–99.

Bernhard P 1990: *Physics Earth and Planetery Interiors*, 63, 145–150.

Bonatz M 1971: *Bull. d'Inf. Marrées Terrestres*, 67, 3066–3069.

Brein R 1962: *Bull. d'Inf. Marrèes Terrestres*, 28, 648–652.

Brockhaus Naturwissenschaften und Technik 1989: F A Brockhaus, Mannheim

Bullen K E 1979: An Introduction to the Theory of Seismology. Cambridge University Press

Carey S W 1988: Theories of the Earth and Universe Stendford University Press, Standford, California

Chen Y T, Cook A 1993: Gravitational Experiments in the Laboratory. Cambridge University Press

Cohen E H, Taylor B N 1986: Adjustment of the fundamental physical constants. Codata Bulletin, 62.

Egyed L 1957: *Geologische Rundschau*, 46, 101–121.

Goodkind J M 1991: In: Proceedings of the Workshop Non Tidal Gravity Changes. Walferdange, Luxemburg Cahiers du Centre Européen de Géodynamique et de Séismologie, 3, 81–90.

Götze H J, Meurers B 1983: *Geophysics*, 52, 136–139.

Groten E 1970: In: Symp. Int. Marées Terrestres, Strasbourg. Obs. Royal Belg. Comm. A.9. Ser. Geophys., 96, 210–212.

Hinderer J, Florsch C N, Mäkinen J, Legros H, Faller J E 1991: *Geophys. J. Int.*, 106, 491–497.

Jordan P 1966: Die Expansion der Erde. Verlag Vieweg and Sohn, Braunschweig

Maddox J 1995: *Nature*, 377, 573.

Meurers B 1996: Comparison of feedback calibration methods results from LCR D–9. *Acta Geod. Geoph. Hung.* (present issue)

Richter B 1987: Das supraleitende Gravimeter. Deutsche Geodätische Komission, Reihe C, 329.

Somogyi J, Závoti J 1993: *Acta Geod. Geoph. Mont. Hung.*, 28, 465–490.

Stacey F D, Tuck G J, Moore G I, Holding S C, Goodwin B D, Zhou R K 1987: *Rev. of Modern Physics*, 59, 157–174.

Stubbs C W 1989: *Nature*, 338, 301–302.

Valliant H D 1973: *Phil. Trans. Roy. Soc. London*, 274, 227–230.

Van Ruymbeke M 1989: *Bull. Géod.*, 63, 223–235.

Varga P 1989: *Österreichische Beiträge zur Meteorologie und Geophysik.* 2, 111–121.

Varga P, Hegymegi L 1985: *Bull. d'Inf. Marées Terrestres*, 96, 5848–5859.

Varga P, Csapó G, Becker M, Groten E 1991: Laboratory calibration of LCR type gravimeters. XXth General Assemly IUGG, Vienna

Varga P, Hajósy A, Csapó G 1995: *Geophys. J. Int.* 120, 745–752.

Zharkov N Z, Molodensky S M, Brzezinsky A, Groten E, Varga P 1996: The Earth and its Rotation. Wichmann

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 371–380 (1996)

THE DETERMINATION AND USE OF THE GEOID IN HUNGARY

A KENYERES[1]

Reflecting to challenge of the GPS technique detailed, high resolution gravimetric and GPS-gravimetric geoid solutions have been computed to the territory of Hungary. We used for these computations the uniquely dense Hungarian Gravimetric Survey, gravimetric data provided from the neighbouring countries and data from the Hungarian GPS Network. The gravimetric solutions were computed by the numerical integration of the modified (Meissl-type) Stokes-integral and the GPS-gravimetric solution was realized in a special 'remove-restore' procedure using selected points of the Hungarian GPS Network. All the steps of the computational procedure have been checked and the results were compared to the GPS-geoid. The HGR95C version was chosen as the *'official' geoid solution for the epoch 1995.* The naming suggests that the development is not finished, we are ready to continue to produce new versions if the database is significantly changes — we expect major improvement from the new DMA/GSFC 1996 geopotential model. In order to test our solutions the computation of the Hungarian geoid is planned using an FFT software package.

Beyond the pure gravimetric solutions, the so-called GPS-gravimetric geoid was also computed. This is planned to be a *tool*, used for GPS heighting purposes. The GPS-gravimetric geoid was successfully tested at the Hungarian GPS Network (OGPSH), providing height above the geoid for all non-levelled GPS points. The accuracy of these heights are expected not to be better than ±3 cm, mainly because of the GPS height component confinement. The GPS sites were measured at 1–2 hours sessions, which may cause that systematic distortions remained in the height component.

Keywords: geoid; GPS; gravimetric geoid; Hungary

1. Introduction

The geoid is a particular equipotential surface of the Earth's gravity field, which in definition — with a certain approximation — coincides with the mean sea level. The knowledge of the geoid has basic importance both in classical and modern geodesy and also in the field of geophysical interpretation. As the height reference surface of geodetic measurements, the geoid played important role e.g. in the reduction of distance measurements to a reference surface, but became an indispensable tool indeed with the advent of the satellite based positioning techniques (GPS), providing the connection between the geocentric and relative reference systems. As a result of the latest research and developments the digital, high resolution and cm-accuracy geoid solutions might be applied on a new, promising field of application: the geoid together with the GPS technique might partially replace — within given circumstances and restrictions — the height determination with levelling.

The first attempts for geoid determination in Hungary (astronomic levelling lines) can be dated back to the early fifties (Homoródi 1952), and the first complete

[1]FÖMI Satellite Geodetic Observatory, H-1373 Budapest, POB 546, Hungary, e-mail: kenyeres@sgo.fomi.hu

geoid solutions were computed by Gazsó and Taraszova (1984) in the sixties and seventies. These astronomic and later astrogravimetric solutions still had no wide practical applications because of the confinements in input data availability and of classification.

Starting in the mid-eighties and reflecting to the challenge of the GPS technique, a geoid development project was initiated at the FÖMI Satellite Geodetic Observatory, aiming at the realization of an up-to-date digital geoid for Hungary. Following intensive methodological foundation, data collection and software developmental phases the first digital gravimetric geoid solutions were completed in 1991. These geoid solutions were based on the evaluation of the truncated and modified Stokes integral, using local, regional gravimetric data and global geopotential model. The accuracy of the HGQ91 solution was estimated to 2–4 ppm based on comparisons with GPS/levelling derived geoidal heights.

During the last couple of years basic developments of the data bases took place multiplicating the available gravity measurements for Hungary and obtaining new data from the neighbouring areas. In addition the possible involvement of the GPS/levelling data into the future geoid forced not only to recompute the Hungarian geoid using the old software and the new data, but to start methodological development in order to produce geoid solution, which already includes GPS/levelling data as well.

This paper intends to introduce the computational method, the data available and the latest gravimetric and GPS-gravimetric geoid solutions produced for the territory of Hungary.

2. The data used

2.1 Gravity data

The Hungarian Gravity Survey is one of the most detailed in the world, the latest version used for the geoid determination contains more than 380000 point gravity measurements, which corresponds to 4–5 point/km^2. This database is three times larger than the set used for the first gravimetric geoid determination in 1991 (see Table I). The Hungarian gravimetric database is owned by the Lorand Eötvös Geophysical Institute and was kindly provided for the geoid determination purposes.

The interpolation of the gravity survey data was done by the carefully selected MINCURV method. This method was chosen from among 10 different softwares (Sárhidai and Kovácsvölgyi 1995). The interpolation of the free-air anomalies was supported by the 500 m $\times$ 500 m digital terrain model, compiled and sold by the Hungarian Military Service. The gravity database outside of the country is much weaker, we could collect data from Austria, Rumania, Slovakia and Ukraina on a $7.5' \times 5'$ grid, which corresponds to about a 10 km $\times$ 10 km grid. It was impossible to get any up-to-date data from the territory of the former Yugoslavia. In contrary of the contrast in the data set distribution, the whole database — a rectangle including the territory of Hungary and its 200 km surroundings — was interpolated to a uniform 1 km rectangular grid given in the Hungarian National Projection

Table I. Comparison data sets used for geoid computations in 1991 and 1995

	PROJECT 1987–91	PROJECT 1993–95
Gravity data (point)	≈ 120000	≈ 380000
Interpolation method	Franke (1978) - no tests	MINCURV - tested
Grid	800 m × 800 m	1000 m × 1000 m
DTM	not used	500 m × 500 m
Geopotential model	OSU86B	OSU91A
Method	Modified Stokes-integral	same

System (EOV), because later the computation of the geoid with an FFT software is also planned. The effect of the global gravity field, over the compiled local data base was taken into account with using the tailored version of the OSU91A model (Basic 1994). This model was fitted to the regional gravity data of Central Europe and it was expected to provide an improved reference field for the geoid determination.

2.2 GPS measurements and levelling data

Based on GPS-derived ellipsoidal heights and levelling data, accurate geocentric geoidal heights may be computed. This GPS-geoid could be used for checking existing geoid solutions or may be merged into the determination process, expected to improve the solution.

In Hungary a very dense GPS network (OGPSH) is being established, with more than 1100 sites — corresponding to about 10 km mean point distance — and about 30 % of the points will be levelled. More than 60 % of the GPS network was already measured and this work will be finished in 1997. The levelled GPS points will be extremly useful for geoid testing and improvement.

3. The computational procedure

The only way to produce a high resolution and possibly high accuracy geoid solution is to use dense gravimetric data bases, which may be supplemented with other data types (e.g. astronomic and/or GPS measurements) for collocation-type solutions. But the gravimetric geoids have a significant drawback — they require large extent of input data, possibly all over the globe. This condition is in most cases hardly fulfilled, therefore the implementation of a global geopotential model is essential. In all gravimetric methods' a model is combined with the appropriate local gravity database and digital terrain model, available in and around the area interested.

The gravimetric solutions for the territory of Hungary are quasigeoids but we use also in the following the terms of geoid and geoidal height instead of the terms of quasigeoid and height anomaly.

3.1 The gravimetric geoid

All the gravimetric methods originate from the well known Stokes-integral formula, the difference between the methods may be found in the mathematical handling of the input data. The most widely used procedures to compute geoid are the FFT and the collocation, offered by sophisticated software packages (e.g. HFTG-BVP and GRAVSOFT). All of these solutions have some advantages and drawbacks. The FFT is an efficient and quick tool, but possibly requires input data on a uniform grid, the collocation may accept irregular data distribution but requires enormous computer capacity for large matrix inversions.

The numerical integration method of the Stokes-integral also needs gridded data, but may accept the compilation of different grids and the various modifications of the Stokes-integral might be easily handled. The main drawback of this procedure is that the numerical integration is very much time consuming. We should note here that a very efficient software tool was developed for the quick computation of the numerical integration (deMin 1994), which may compete with the earlier mentioned methods.

All the gravimetric geoid solutions for Hungary have been computed using an in-house developed numerical integration software, called DIEGO. The numerical integration method was chosen because we wanted to produce geoid solutions based on the modified Stokes-integral and we have a special gravimetric database — very dense within Hungary but sparse and unevenly distributed outside of the country.

Several types of the Stokes-integral modification techniques (Wong and Gore 1969, Molodensky et al. 1962, Meissl 1971, Heck and Grüninger 1983) have been checked and tested. The essence of the modificaton techniques that the local (sufficiently dense) gravity data are used within a given truncation cap, the integration performed within this cap applying a specific modification of the Stokes-integral and the effect of the remote zone — outside the cap — is taken into account including a global geopotential model. Testing the different truncation solutions we wanted to find out, which method is the most efficient for small (less than 3 degrees) integration cap sizes. It was found that for such circumstances the Meissl-type modification gives the best results (see Fig. 1), therefore this truncation method was chosen for the Hungarian geoid determination. In the following the formalism of Meissl's truncation technique is briefly outlined, detailed description may be found at Bölcsvölgyiné Bán and Kenyeres (1990).

The original Stokes-integral after the modificaton procedure can be rewritten in the following general form:

$$\hat{N} = \frac{R}{4\pi\gamma} \iint\limits_{\sigma_o} \Delta g K(y) d\sigma + \frac{R}{2\gamma} \sum_{n=2}^{M} \overline{Q}_n(y_o) \Delta g_n + \delta N ,$$

where

$K(y)$: the modified $S(y)$ Stokes-function
σ_o : the integration cap radius
M : the degree of the geopotential model used
$Q_n(y_o)$: the truncation coefficient series
Δg_n : nth degree gravity anomaly computed from the geopotential model
δN : effects, not taken into accont (error).

In the case of Meissl-type modification the $K(y)$ kernel is very simple:

$$K(y) = S(y) - S(y_o)$$

where $S(y_o)$ is the value of the Stokes-function at the border of the truncation zone.

The Q_n truncation coefficient series for the Meissl-method can be computed in the following way:

$$Q_n^{ME}(y_o) = Q_n + \frac{S(y_o)[P_{n-1}(y_o) - y_o P_n(y_o)]}{n-1} .$$

An example for the truncation coefficients may be found in Fig. 1.

As it is clearly seen from Fig. 1 the Q_n truncation series converges fastest for the Meissl and the Heck/Gruninger methods (this latter is the combination of the Wong/Gore and Meissl), but the Meissl-method offers a much simpler solution.

The gravimetric geoid solution for the territory of Hungary was computed on a 5 km uniform grid given in the National Projection System (EOV). Several geoid versions using different truncation cap radii (0.25–2 degrees) have been computed

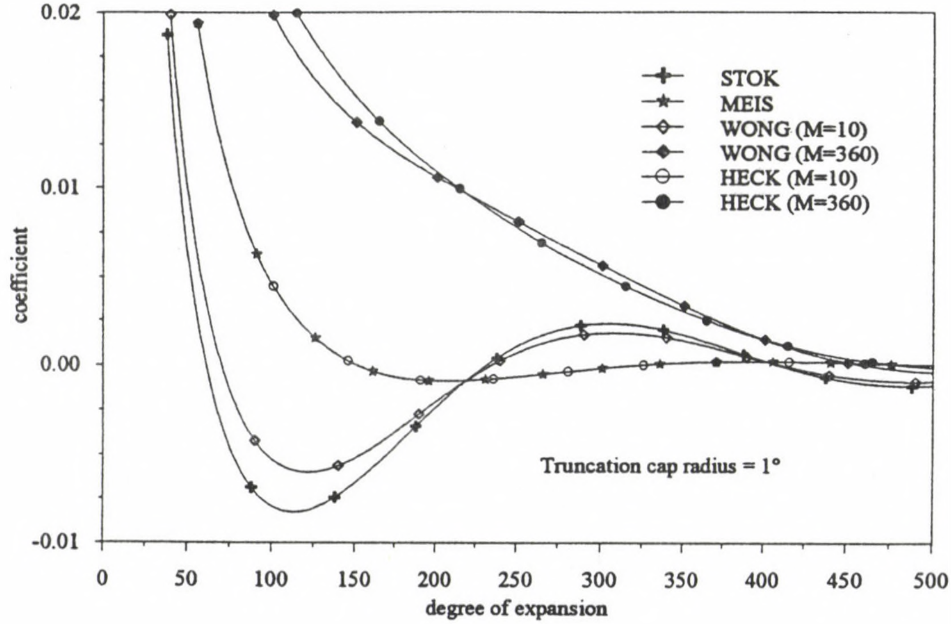


Fig. 1. Comparison of the different truncation coefficient series

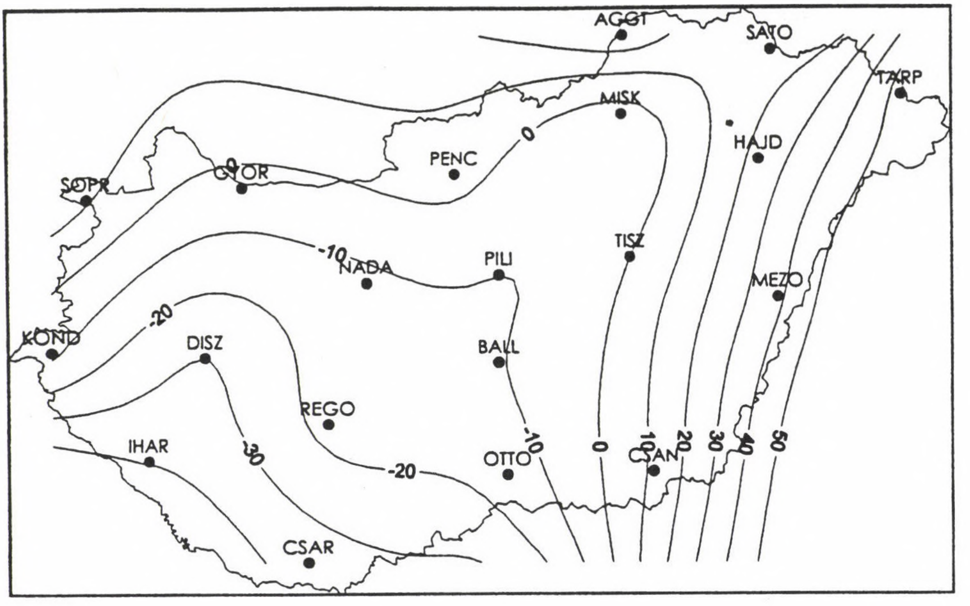


Fig. 2. The comparison of the HGR95C and the GPS-geoids. The differences are in cm unit after a substraction of a 57 cm mean

and tested. Finally the version HGR95C, with 1 degree truncation cap radius was chosen as the 'official' geoid solution for the territory of Hungary. This solution is available in digital form on a given grid format and also in the DXF format proper for GIS users (Kenyeres 1995a). In addition a coloured isoline map has been prepared with some additional information for demonstration purposes (Kenyeres 1995b). The HGR95C solution can be seen in Fig. 3.

The relative accuracy of the geoid solution is estimated to about 1–2 ppm based on comparisons with GPS+levelling data available at the moment in the zero order GPS network (see Fig. 2) and in the eastern portion of the densified OGPSH network.

3.2 The GPS-gravimetric geoid

As it is well known the gravimetric geoid solutions are suffering from long wavelength distortions arising from the incorporation of global geopotential models. This effect especially deteriorates the expected accuracy in the neighbourhood of areas, where the model is weaker due to lack of trustful input data. Unfortunately Hungary is in such bad situation — the comparison of the GPS-geoid and the gravimetric geoid has shown discrepancies of several decimetres (see Fig. 2). We may expect improvement only from the release of the new DMA/GSFC 1996 model.

For practical purposes — GPS-heighting — a geoid solution conform with the GPS-geoid data would be desirable, which is achievable with the inclusion of GPS data into the geoid determination process. This could be done on different ways,

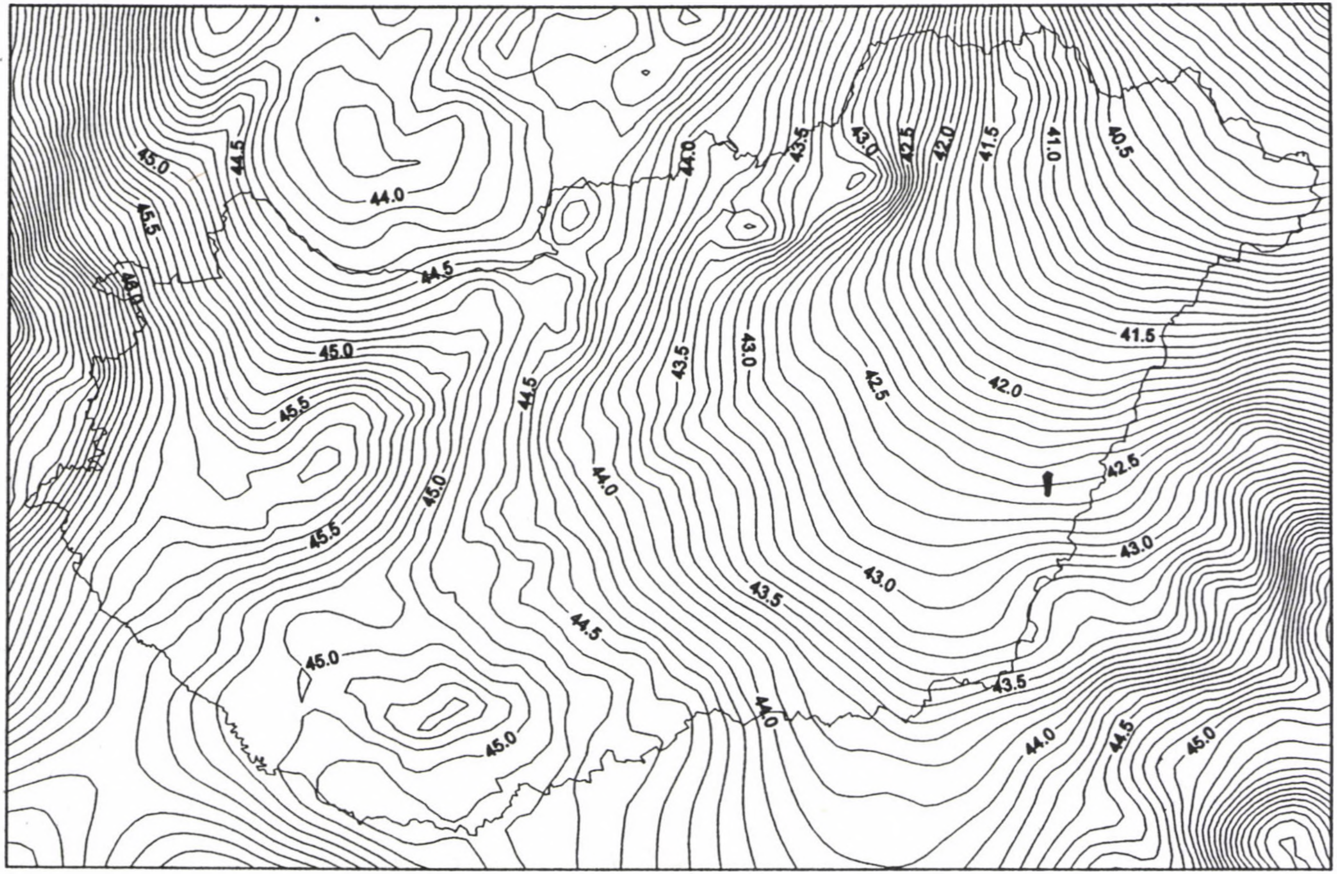


Fig. 3. The HGR95C gravimetric quasigeoid solution for the territory of Hungary

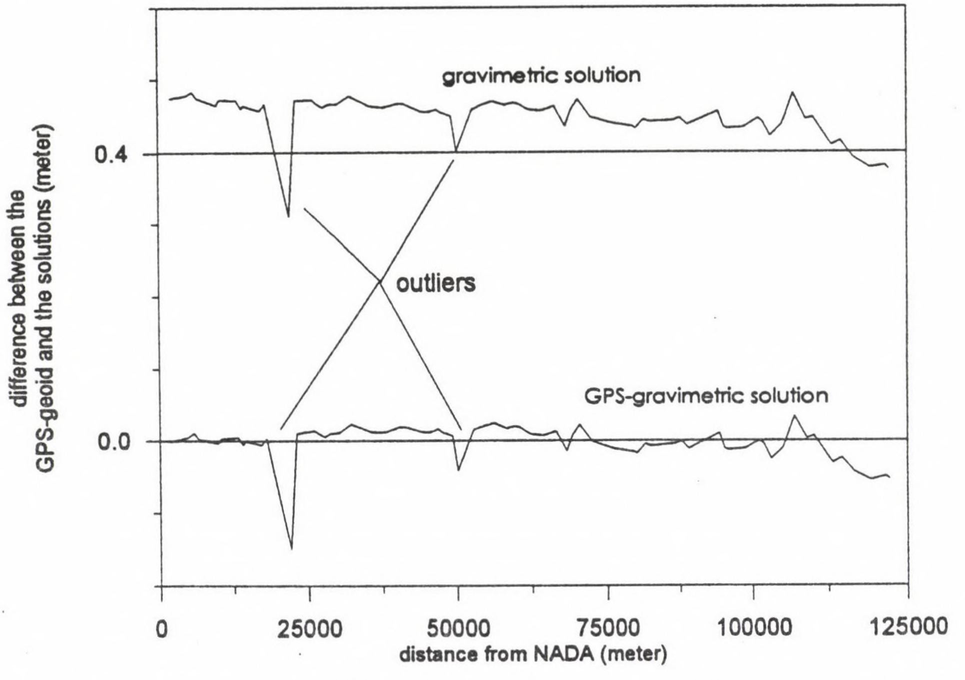


Fig. 4. The comparison of the GPS-geoid and the gravimetric and GPS-gravimetric solutions

one natural solution is the collocation. Here we shall introduce the so called GPS-gravimetric geoid, which is the result of a special 'remove-restore' procedure and modifies the gravimetric solution in a way to fully agree with the GPS-geoid. The resulting GPS-gravimetric geoid will not be any more a geoid solution in the 'classical' sense, it will be a 'tool' designed for the efficient GPS-heighting works. The ideas, how to produce this solution and why it is recommended for the practice may be found at Kenyeres (1992, 1995c).

4. The use of the geoid in Hungary — examples

4.1 Example I. Test of the geoid solutions

In order to test the quality and feasibility of the gravimetric and GPS-gravimetric solutions a GPS-campaign was performed still in 1993 along a spirit levelling line connecting 2 zero order GPS network points NADA and DISZ. In this campaign 70 benchmarks were included, the distance between the end-sites is about 122 km. At the benchmarks the GPS-geoid, the gravimetric and GPS-gravimetric geoidal heights have been computed and compared. The graphical representation of the results may be found in Fig. 4. The gravimetric solution has a shift and a tilt compared to the GPS-geoid, which was eliminated in the GPS-gravimetric solution. This latter can be *directly* compared to the GPS-geoid on the 2-3 cm level. Some outliers have been found, which are supposed to be the result of benchmark displacements.

4.2 Example II. Height determination of the densified OGPSH points without levelling

The Hungarian GPS Network (OGPSH) consists of more than 1100 points, which are mainly settled on the existing 1st to 4th order geodetic network. Most of the points has no levelled heights and the levelling of these points is not possible that time. The GPS measurements together with the introduced GPS-gravimetric geoid solution offers a unique opportunity to give height above the geoid for all the points at the 2–3 cm accuracy level. We have to involve the trustful levelled GPS points into the process of the GPS-gravimetric geoid determination, then for the remaining not-levelled network sites a simple subtraction may provide a limited accuracy height above the geoid. The first test have been performed for the eastern part of Hungary, the OGPSH network has been measured only there by middle of 1996. Figure 5 shows the test results of the fitting phase for the levelled points.

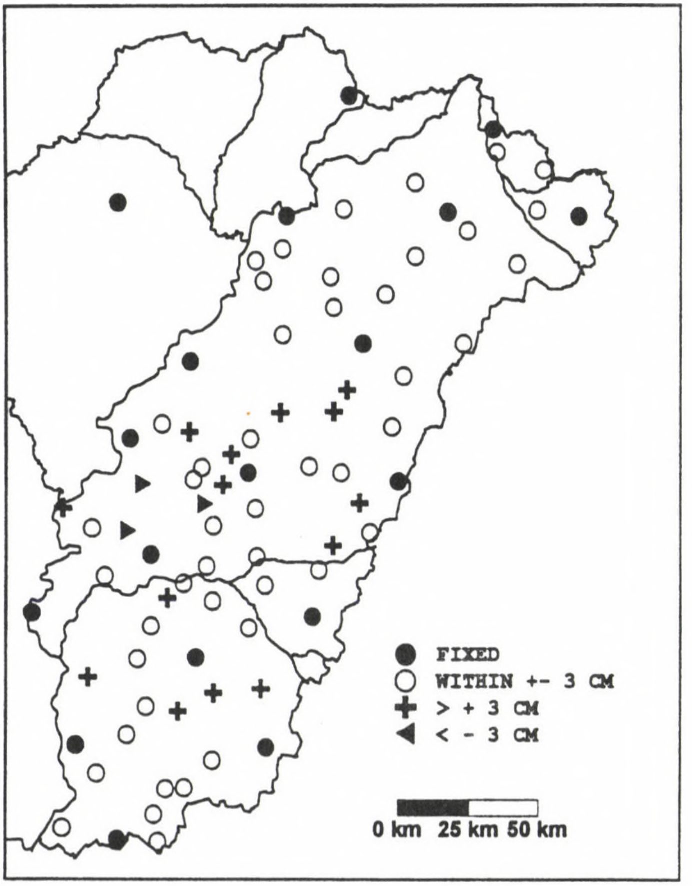


Fig. 5. The performance of the GPS-heighting in the Tiszántúl area

Keeping fixed some selected points, the residuals of the fitting are shown for the remaining levelled sites. As it is seen for most of the points the residuals are less than 3 cm and interestingly the higher residuals are concentrating on some areas. The reason is not yet found, some suspected points will be re-levelled in September, 1996.

Acknowledgements

The gravimetric geoid solution for the territory of Hungary was accomplished by the financial support of the National Committee for Technology Development — OMFB — (OMFB-FM-FÖMI 93-950003 project). The gravimetric data was provided for this purpose by the Lorand Eötvös Geophysical Institute. Special thanks to the colleagues from the neighbouring countries, who provided gravimetric data for geoid determination purposes.

References

Basic T 1994: Personal communication

Bölcsvölgyiné Bán M, Kenyeres A 1990: *ÖZfVPh*, 77, 159–178.

deMin E 1994: In: Proceedings of the IGC/ICG Joint Symposium, 1994, Graz, Austria. Springer Verlag, IAG Symposia series, No. 113, 471.

Gazsó M, Taraszova G 1984: The astrogravimetric determination of the quasigeoid in Hungary. A FÖMI tudományos kiadványa, Vol. 5.

Heck B, Grüninger W 1987: In: Proceedings of the IAG Symposium IUGG XIX. General Assembly Vancouver, Canada

Homoródi L 1952: *Földméréstani Közlemények*, 1, 1–10 and 2, 61–71.

Kenyeres A 1992: In: Proceedings of the 1st Continental Workshop on the Geoid in Europe. Prague, 1992, 482–490.

Kenyeres A 1995a: The database of the Hungarian Geoid Determination Project. Product description. Manuscript, prepared in the frame of the OMFB 93-95 0003 project

Kenyeres A 1995b: The Gravimetric Geoid Map of Hungary, Version HGR95C. Prepared in the frame of the OMFB 93–95 0003 project

Kenyeres A 1995c: In: Proceedings of the 3rd Int. Seminar on "GPS in Central Europe", Penc, Hungary, 1995. Reports on Geodesy, No. (3)16, 413–422.

Meissl P 1971: Preparation for the Numerical Evaluation of Second Order Molodensky-type Formulas. OSU/DGSS, Report No. 163, The Ohio State University, Columbus, Ohio

Molodensky M S, Eremeev V P, Yurkina M I 1962: Methods for Study of External Gravitational Field and Figure of the Earth. Israel Program for Scientific Translations, Jerusalem

Sárhidai A, Kovácsvölgyi S 1995: The interpolation of free-air anomalies to the territory of Hungary (in Hungarian). Manuscript, prepared in the frame of the OMFB 93–95 0003 project

Wong L, Gore R 1969: *Geophys. J. RAS*, 18.

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 381-388 (1996)

ESTIMATION OF THE CAVITY EFFECT USING THE F.E.M.

P KOSTECKÝ[1] and I KOHÚT[2]

The presence of the cavity changes the strain field caused by external force in solid. One way how to estimate the magnitude of this effect is computation of strain field in solid with and without cavity and their comparison. As a computational method was chosen finite element method (F.E.M.) for ability to consider the arbitrary shape of cavity and influence of inhomogeneity of material. In presented work is described the computer code developed for such purpose and some testing results are presented.

Keywords: cavity; finite element method; strain field; tidal force

Introduction

The reason for attempting a solution of the cavity problem was that measurements of deformations caused by tidal forces are burdened by error ensued from local effects. Measurements of tidal-induced strain and tilt are usually performed in underground facilities in order to minimize the influence of environmental factors. Measured values differs from the theoretical ones calculated under assumption the instrument is directly surrounded by a large volume of continuous material. Topogeological features of surroundings have a certain contribution to this deviation. The sources of deviations of such kind are the presence of the cavity, the influence of surface topology and inhomogeneities of elastic constants of rock structures.

These effects are induced by the tidal potential itself and it is impossible to separate their influence by frequency analysis. This was the motivation to solve the outlined problem by detailed modelling of strain field in the vicinity of the cavity.

Classical results for the cavity effect estimation were given by Eshelby (1957) (idealized analytical solution) and Harrison (1976) (first order F.E.M. approximation). The later had given the correction factors for the tilt amplitudes. Similar calculations were done by Brimich and Brestenský (1992) (F.E.M.) and Brimich and Hvoždara (1996) (boundary element method).

Problem definition

The reason we considered 2-dimensional (2D) plane strain model is that such a case is much more simple in computational aspects than a 3D model and it could serve as an anticipatory result for future calculations in the 3D domain. Moreover, in some cases the instruments are located in tunnels in vicinity of which the material properties do not change significantly, what could be, in some approximation,

[1]Comenius University, Fac. Math. Phys., Dept. Geophys., Mlynská dolina F-1, 842 15 Bratislava, Slovak Republic

[2]Geophysical Institute, Slovak Academy of Sciences, Dúbravská 9, 842 28 Bratislava, Slovak Republic

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

sufficiently well modelled by a 2D case. Other simplifications we have considered are linearity of the problem and its independence of time (static case), assuming the load of domain caused by tidal forces varies sufficiently slow. The reality is more complicated, mostly by the uncertainty of determination of rock structure properties in the vicinity of the measurement point. But we think it has value to solve such a simple problem, too.

The strain tensor e_{ij}^{∞} in model without cavity is coupled with this one e_{ij}^{c} for the domain with cavity by Eq. (1):

$$e_{ij}^{c} = S_{ijkl} e_{ij}^{\infty} . \tag{1}$$

For our 2D (plane strain) case this relation can be rewritten as Eq. (2):

$$\begin{aligned} e_{xx}^{c} &= S_{xx} e_{xx}^{\infty} + S_{xy} e_{yy}^{\infty} \\ e_{yy}^{c} &= S_{yx} e_{xx}^{\infty} + S_{yy} e_{yy}^{\infty} . \end{aligned} \tag{2}$$

Correction factors S_{ij} are functions of the cavity shape, e.g. for ellipsoidal cavity they are functions of the principal axes ratio. They were calculated and graphically depicted by Harrison (1976) (under the assumption that the Poisson's ratio of the material is 0.25).

For calculation of the strain field in solid we have to solve the well-known Lamé system of second-order partial differential equations of elliptic type, describing the linear elasticity problem in terms of displacements:

$$-(\lambda + \mu)\,\mathrm{grad}\,\mathrm{div}\,\mathbf{u} - \mu\,\mathbf{\Delta u} = \mathbf{F} . \tag{3}$$

λ, μ are the elastic constants of the medium (given by Young modulus and Poisson's ratio), $\mathbf{u}$ is the displacement vector, $\mathbf{F}$ represents the external (area) force field. Two types of boundary conditions have been considered:

a) $\mathbf{u} = \mathbf{0}$ at the fixed part of the boundary,

b) $\dfrac{\partial \mathbf{u}}{\partial \mathbf{n}} = 0$ at the free part of boundary (normal derivative).

The strain tensor is given by the relation:

$$e_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right) \tag{4}$$

Description of code

The classical finite-element solution of the linear elasticity problem is often based on the use of simple linear (or bilinear) shape functions. It may be sufficient in some cases, but for a more precise solution it is suggested to use of higher order approximations. For the solution in displacements to be of at least C^1 continuity, it is necessary (for 2D case) to use classical polynomial shape functions of minimum order 5, what follows from Ženíšek's theorem (1970).

Thus we aimed for developing code which is based on such a high order approximation. Babuška and Suri (1990) suggested and worked out the use of the p-method finite-element solution based on hierarchic shape functions. In our work the advantageous features of this method were utilized. Since we had no access to commercially available high order F.E.M. codes, we have to develop our own one. As it was mentioned before, it solves the 2D plane strain linear elasticity problem. It employs the hierarchic set of shape functions for triangular elements (up to the order 8). We try to build in it the possibility of the use of curved border element (isoparametric) up to fourth order of curvature. External forces can be prescribed as area forces or as nodal (single) ones. In this version of code a total fixing of attributed degrees of freedom is only possible (condition a), but in the later versions we can only use partial fixing (by putting only u_x or u_y to zero). The condition b can be, in the F.E.M. formulation, taken into account simply by equating the corresponding right-side terms to zero.

The code itself (written in the Pascal language and implemented on PC-class computers) consists of three parts:

preprocessing: reading of input data, renumbering of mesh elements for the front-width optimization, precomputation of the scalar products table, transformation of the external forces;

mainframe: computation of local stiffness matrices and their assembling, elimination of subsequent front variables, solving for final front variables and back-substitution;

postprocessing: computation displacements and stresses for all nodes, parametrically prescribed side points and general element points.

The solution values at the node and side points may be improved by the super-convergent recovery technique (introduced by Zienkiewicz (1992)). This procedure is based on the fact that solution error is strongly concentrated in the vicinity of element border lines.

Testing calculations

As the testing example, the mesh covering a rectangular domain with approximately elliptical cavity was selected (Fig. 1). The mesh with cavity consists of 70 nodes, 166 sides and 96 elements. The same mesh but without cavity, has 99 nodes, 266 sides and 168 elements. Vertical sides of the domain are taken as fixed. The direction of tidal force field was taken in this example at the angle of 20° to the vertical axis and the Poisson's ratio of the material was considered 0.3. Both meshes were processed and 2 strain fields were obtained.

For the demonstration of differences between both cases, the relative differences in the sense:

$$r_{ij} = \frac{e_{ij}^{\infty} - e_{ij}^{c}}{e_{ij}^{\infty}} \tag{5}$$

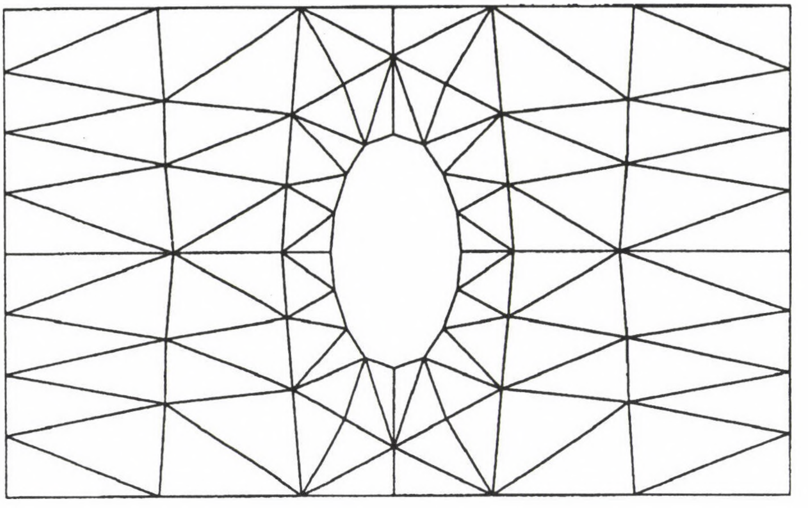


Fig. 1. Testing mesh with cavity

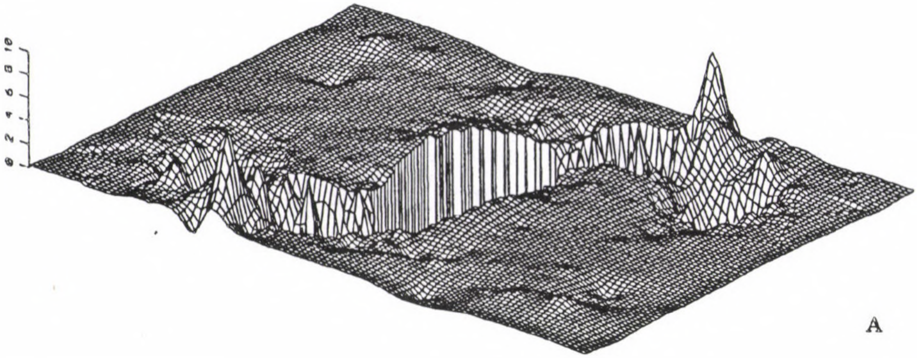
were computed, $(i,j = x,y)$ and e_{ij}^{∞}, e_{ij}^{c} stands for the appropriate component of the deformation tensor without and with cavity, respectively. The calculations were performed for the orders of the shape functions 1, 2 and 4. The results are graphically represented on the Figs 2–4.

Conclusion

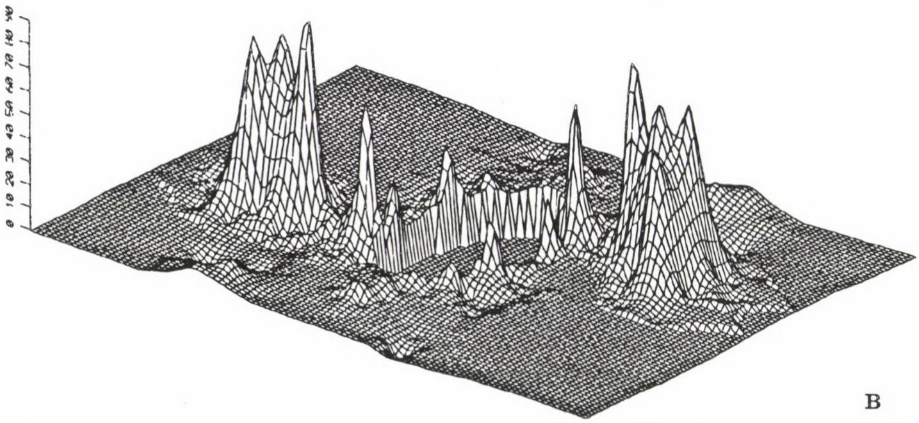
Although the example presented is far from being realistic, it shows for all components of the deformation tensor remarkable differences between first and higher order F.E.M. approximation calculations. This seems to be the justification for the necessity to use the higher order F.E.M. approximation for the cavity effect calculations, for latter to be applicable to the real tidal measurements.

The bigger and more complicated meshes are under test to evaluate the code limits and to obtain the improved correction factors for the cavity effect estimations which would be of practical significance. The code is still at the stage of refinement. We are aware that for realistic calculations, the ratio of cavity dimensions to the dimensions of whole domain under investigation must be at least 1:20 (Harrison 1976), what is not the case of the mesh presented, but it will be fulfilled in the future testing examples.

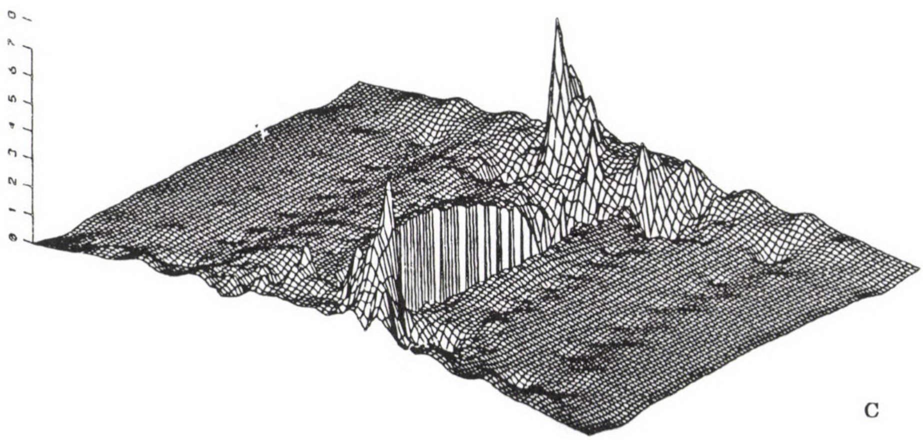
At present, a more realistic mesh — closely corresponding to the real situation of the Geophysical Institute of the Slovak Academy of Sciences Vyhne tidal station — was successfully processed. This mesh consists of 556 elements and calculations by the code described were performed up to the order 7 (more than 33000 degrees of freedom) F.E.M. approximation.



A



B



C

Fig. 2. Distribution of r_{xx} [A], r_{yy} [B], r_{xy} [C] for order 1

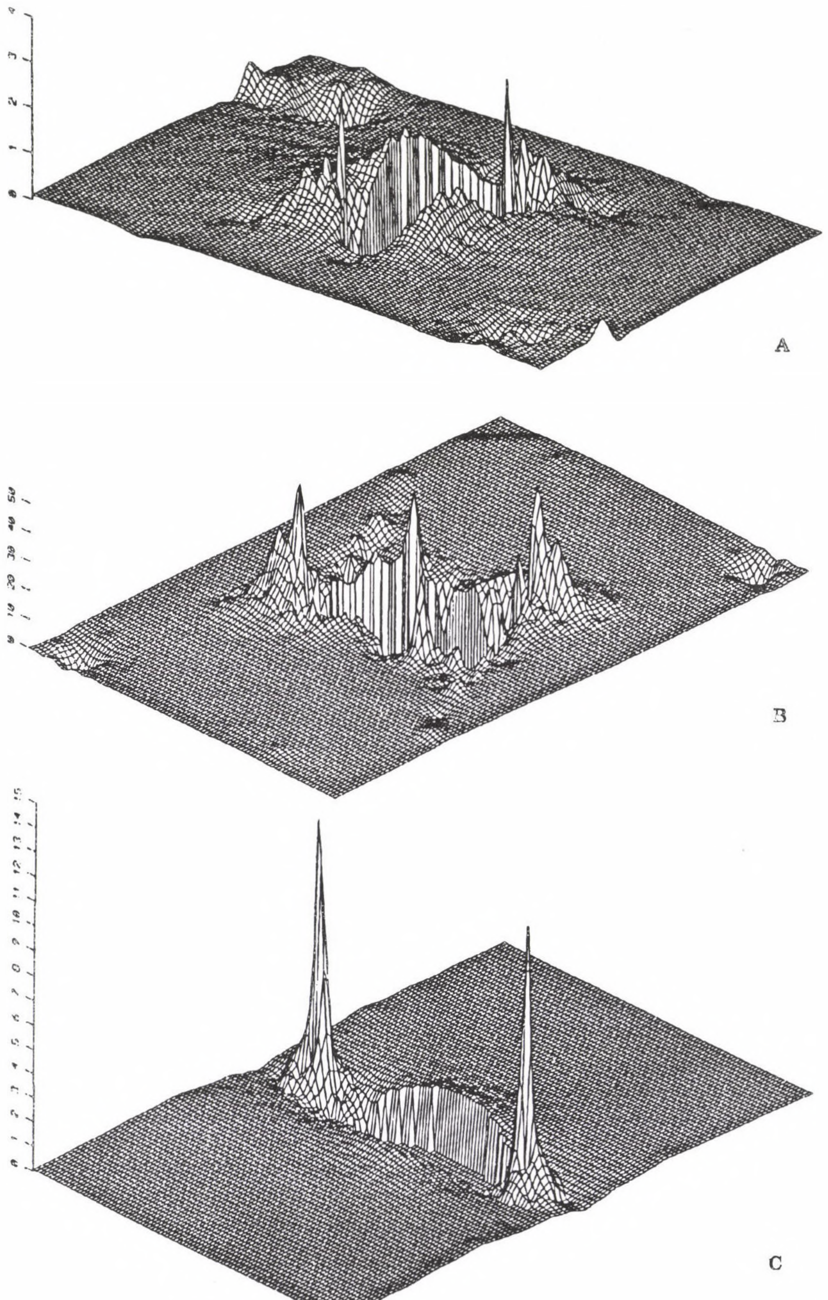


Fig. 3. Distribution of r_{xx} [A], r_{yy} [B], r_{xy} [C] for order 2

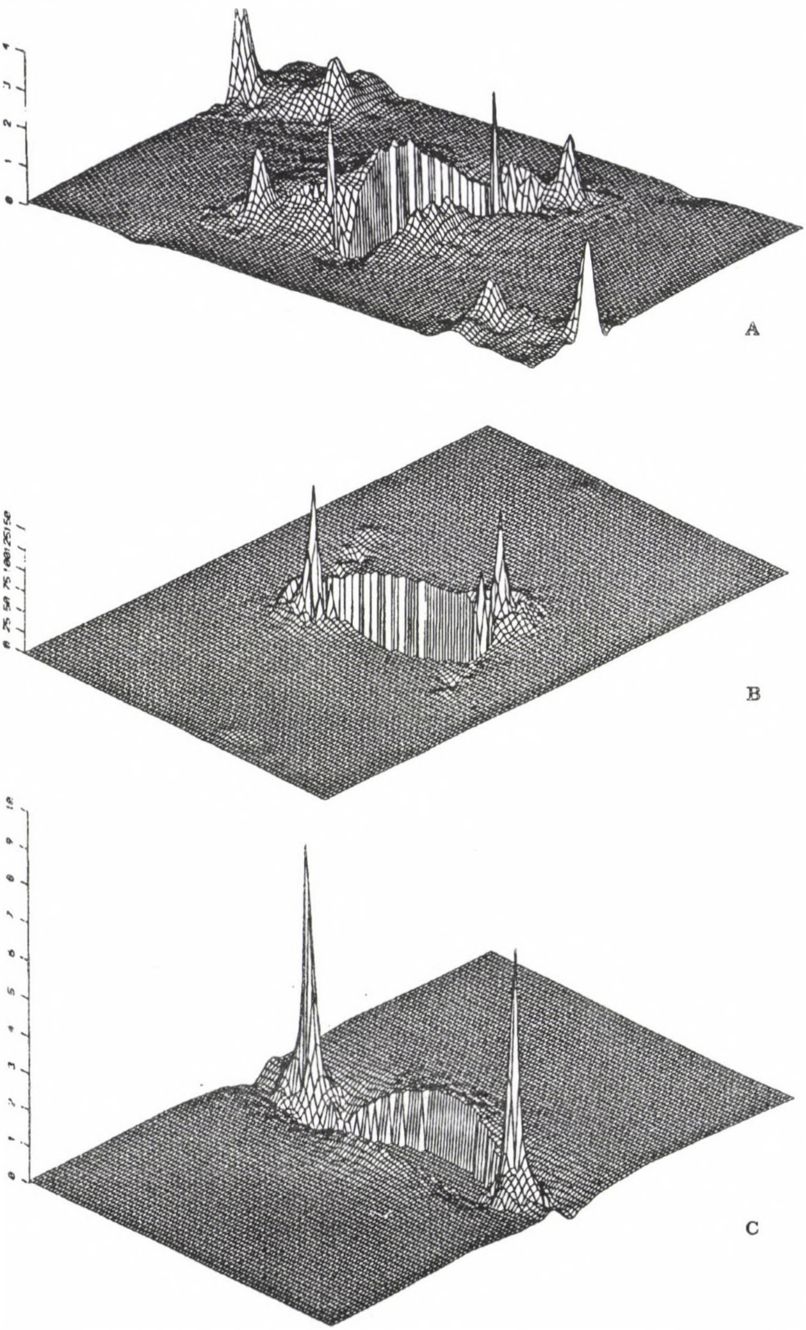


Fig. 4. Distribution of r_{xx} [A], r_{yy} [B], r_{xy} [C] for order 4

References

Babuška I, Suri M 1990: *Comp. Meth. in Appl. Mech. and Engng.*, 36, 2503.
Brimich L, Brestenský J 1992: *Contr. Geophys. Inst. S.A.S.*, 22, 92.
Brimich L, Hvoždara M 1996: *Contr. Geophys. Inst. S.A.S.*, 27 (in press)
Eshelby J D 1957: *Proc. Roy. Soc.*, Ser. A 241, 376.
Harrison J C 1976: *J. Geophys. Res.*, 81, 321.
Kohút I, Kostecký P 1995: *Acta Astron. et Geophys. Univ. Comenianae*, 17, 20.
Ženíšek A 1970: *Numer. Math.*, 15, 283.
Zienkiewicz O C, Zhu J Z 1992: *Int. J. Num. Meth. Engng.*, 33, 1331.

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 389-403 (1996)

INVESTIGATION OF THE ISOSTATIC ANOMALY OF THE EASTERN ALPS

B MEURERS[1]

The regional characteristic of the isostatic anomaly in the Eastern Alps cannot be explained by simple models with partially over- and undercompensated crust. This also holds for the geodynamical interpretation of observed recent crustal movements. The correlation with the isostatic anomaly is not convincing even when investigating the low frequency content of the field. Gravity effects of upper mantle structures as reported in the Western and Central Alps may remarkably change the regional trends also in this area. Nevertheless rough estimates of the isostatic state are possible by statistical analysis of the relation between gravity anomalies and mean elevations. The present investigation shows the Eastern Alpine area being close to isostatic equilibrium. The coincidence of high uplift rates with local negative isostatic anomalies indicates isostatic rebound effects to be possible.

Keywords: Bouguer gravity; Eastern Alps; isostatic anomaly; recent crustal movements; regression analysis

1. Introduction

Recent crustal movements are a well-known phenomenon in orogenic belts. They are mostly due to still active tectonic processes but partly also may be guided by isostatic compensation effects. The aim of this study is to investigate what information concerning the isostatic state of high mountainous areas can be gained from isostatic anomalies. This investigation is based on analyzing Bouguer gravity instead of geoidal undulations. Because of being connected with the vertical derivative of a potential, Bouguer anomalies much more reflect the density distribution of the outer parts of the earth than geoid heights. Consequently isostasy can here only be studied locally and not regarding the global characteristics.

2. The Bouguer and isostatic anomaly of the Eastern Alps

Figure 1 shows the Bouguer anomaly of the Eastern Alps based on mean Bouguer gravity values in a $3' \times 5'$ grid (Senftl 1965). This grid and additionally a detailed data set (Meurers et al. 1987) have been used to calculate the isostatic anomaly (Fig. 2, Wagini et al. 1988). The high frequency content is visible especially in the map derived from the high resolution data in the central part of the Eastern Alps (Fig. 3, Meurers 1992). The mean isostatic anomaly is +8 mGal and +4 mGal, respectively. All data refer to the geodetic reference system 1980 (Moritz 1980). Compensation according to the Airy-Heiskanen model (e.g. Heiskanen and Moritz

[1]Institute of Meteorology and Geophysics, University of Vienna, Hohe Warte 38, A-1190 Wien, Austria, e-mail: bruno.meurers@univie.ac.at

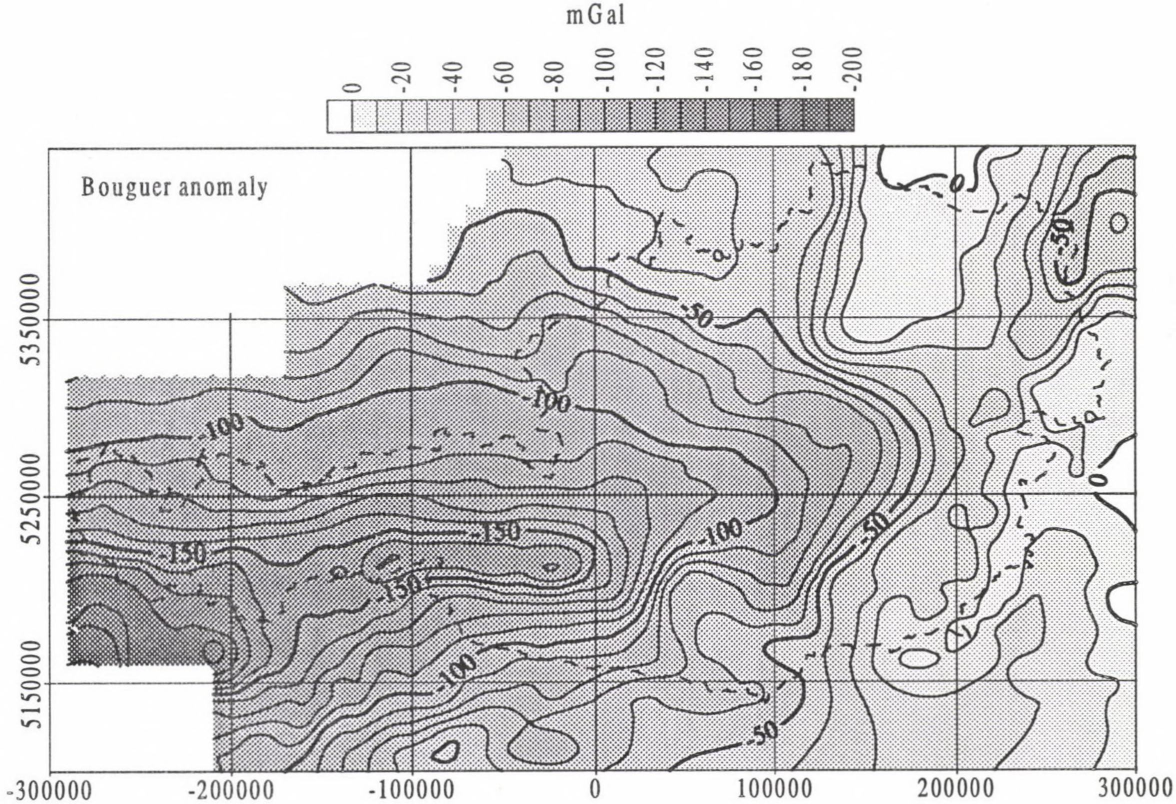


Fig. 1. Bouguer anomaly of the Eastern Alps (Senftl 1965). Co-ordinates in m

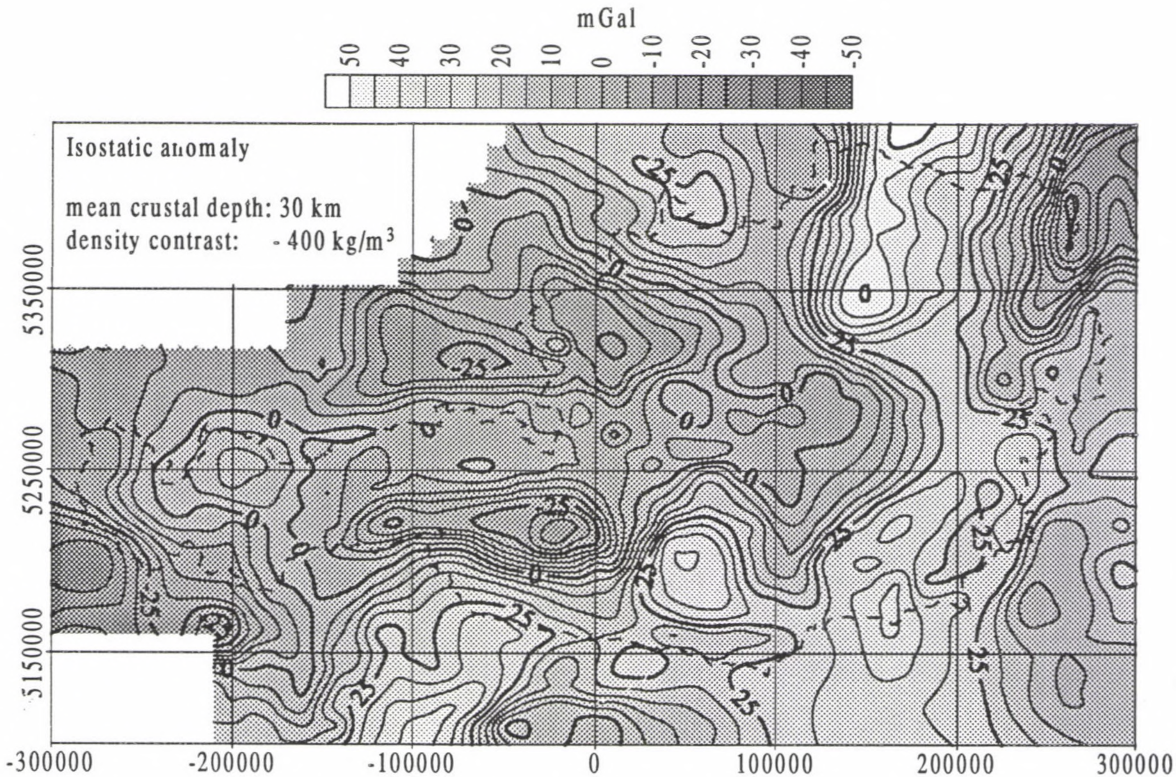


Fig. 2. Isostatic anomaly of the Eastern Alps (Wagini et al. 1988). Co-ordinates in m

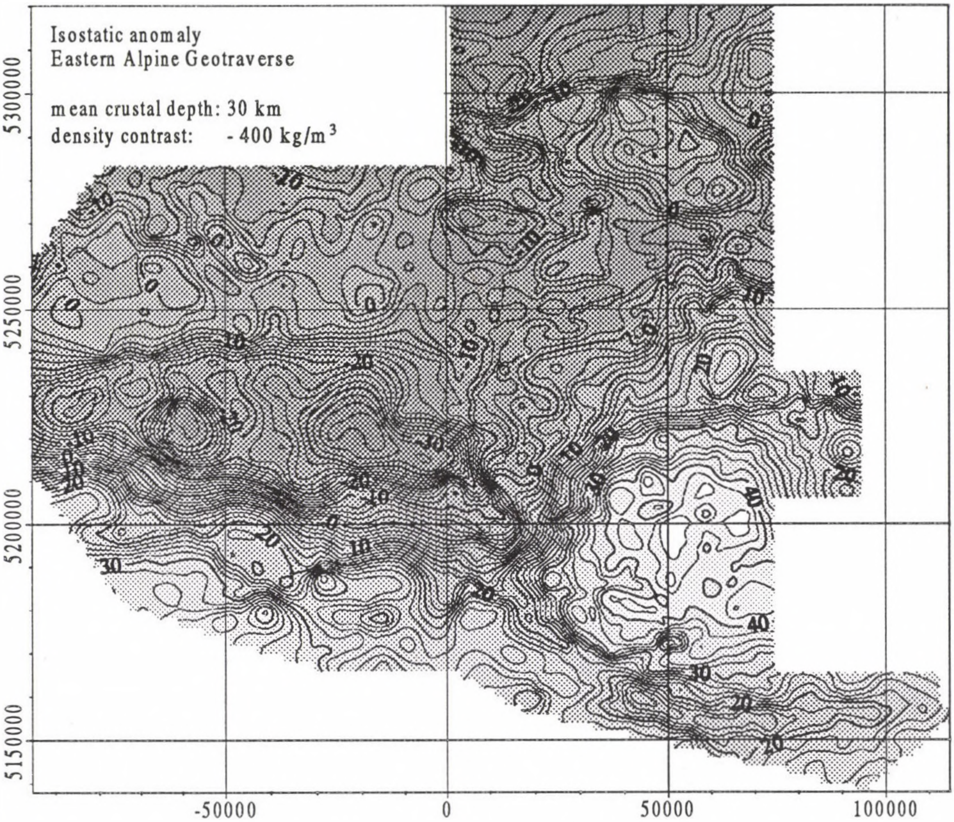


Fig. 3. Isostatic anomaly of the Eastern Alpine Traverse (Meurers 1992). Co-ordinates in m

1967) is assumed at the crust-mantle boundary. The isostatic correction can be calculated by applying Parker's (1972) Fourier transform expression:

$$F[g_1(x)] = 2\pi G\varrho e^{-|k|T}\sum_{n=1}^{\infty}\frac{(-|k|)^{n-1}}{n!}F[t_b^n(x)]\,. \tag{1}$$

G is the gravitational constant, while x and k symbolize the two-dimensional co-ordinate vectors in space and wavenumber domain respectively. $t_b(x)$ denotes the crust-mantle undulation with respect to the mean crustal thickness T at sea level. According to the Airy-Heiskanen concept it depends on the ratio between mean crustal density ϱ and the density contrast $\Delta\varrho$ at the crust-mantle boundary and the topographic heights $h(x)$:

$$t_b(x) = \frac{\varrho}{\Delta\varrho}h(x)\,. \tag{2}$$

Mean topographic heights are available by digital terrain models defined in different grid sizes (Steinhauser et al. 1984). Equation (1) corresponds to a flat

Table I. Variation of the isostatic anomaly of the Eastern Alps due to changing the density contrast and the compensation level

T km	$\Delta\varrho$ kg m^{-3}	Min mGal	Max mGal	Mean mGal
30	300	−4.7	0.5	+1.0
25	400	−0.6	9.6	−3.0

approximation which of course is not sufficient. Therefore the gravity effect of all sources outside the Hayford-zone O_2 (166.7 km) was removed by applying the method of Simpson et al. (1983). Adding the gravity both of topography and root beyond 166.7 km to the antipodes according to Eq. (2) finally results in the complete isostatic correction term. All effects of these distant masses can be taken from Kärki et al. (1961).

3. Interpretation problem

Isostatic anomalies contain generally the gravity effects of under- or overcompensation with respect to certain boundaries, e.g. the crust-mantle discontinuity or the boundary between lithosphere and asthenosphere. But the geodynamical interpretation is rather difficult. There are severe objections against such interpretation especially when the mean isostatic anomaly is concerned:

— Model parameters assumed for calculating the Bouguer and isostatic anomaly strongly influence the isostatic mean (e.g. normal gravity, density contrast between lower crust and upper mantle, crustal thickness at sea level, reference level). Table I shows the variation of the mean isostatic anomaly of the Eastern Alps depending on the definition of the Airy-Heiskanen root by different parameters (Wagini et al. 1988). Both the mean and the regional trend are strongly affected as shown by the extremes.

— Crustal density variations are able to cause local anomalies with non zero average even when the mean density variation is zero. A perfectly compensated crust does not have necessarily a non zero isostatic anomaly.

— The anomaly pattern changes dramatically when geological corrections of known upper crustal structures are taken into account. Figure 4 shows a section of the isostatic anomaly map of the Eastern Alps after removing the gravity effect of the Molasse basin. For this purpose three-dimensional modelling (Götze and Lahmeyer 1988) was applied using the 3D interpretation of the Bouguer gravity of Southern Germany (Müller 1988). In the present study the Airy-Heiskanen root has not been modified by inserting the crustal density distribution into Eq. (2), therefore Fig. 4 gives an estimate of the maximum effects. However, such modifications do not fully compensate the geological correction.

— The real density distribution is *a priori* unknown at calculating the mass corrections. Applying the mean crustal density is a suitable approximation but too simple to cope with reality. Another problem concerns the free-air correction which does not properly consider the anomalous vertical gradient and its vertical derivatives. Therefore reduction anomalies are unavoidable especially in high mountainous areas. They are dominated by high-frequency features, but also distinct regional trends can be present. In the central part of the Eastern Alps reduction anomalies up to more than 20 mGal occur (Meurers 1992). Table II compares two Airy-Heiskanen root models in the Eastern Alps and the corresponding gravity effect. The first fully agrees with the classical Airy-Heiskanen concept based on assuming constant crustal density. In the second model the root function is modified by considering topograpic density variations (Steinhauser et al. 1984, Walach 1987).

— The long-wavelength gravity effect caused by sub-crustal sources influences both the average of the isostatic anomaly and its regional pattern. Three-dimensional modelling of the Western Alps indicates high amplitude contributions from sources below the crustmantle boundary (Schwendener and Mueller 1985). This may also be true in the Eastern Alps. Therefore the interpreta-

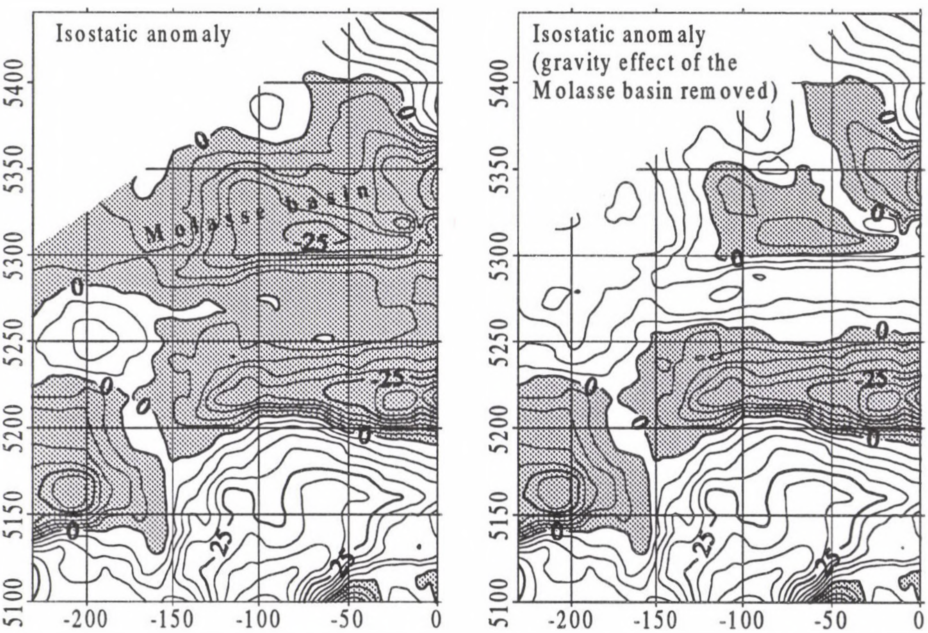


Fig. 4. Geological correction of the isostatic anomaly according to a 3D density model of the Bavarian Molasse sediments. Co-ordinates in km

Table II. Variation of the root depth and the isostatic anomaly of the Eastern Alps due to considering a topographic load with variable surface density

		Min	Max	Mean
Depth of Airy-Heiskanen root, km				
Density	$\varrho = \text{const} = 2670\ \text{kg m}^{-3}$	29.4	47.6	33.5
Model	$\varrho = \varrho(x)$	29.4	47.4	33.6
	Difference	−1.0	0.8	−0.1
Gravity effect of the Airy-Heiskanen root, mGal				
Density	$\varrho = \text{const} = 2670\ \text{kg m}^{-3}$	−164.4	−5.6	−52.0
Model	$\varrho = \varrho(x)$	−163.2	−8.6	−53.8
	Difference	−3.2	8.0	1.9

tion even of the regional field is not easy. Moreover, the correlation of recent crustal movements with isostatic anomalies is probably strongly influenced too. This problem will be discussed later.

Because of these reasons the conclusion on the isostatic state has to be drawn with much care especially when the average anomaly is close to zero like in the present case.

4. Height dependence of gravity anomalies

Therefore a further attempt was performed to analyse the state of compensation of the crust by investigating the relation between gravity anomalies and mean elevations. The advantage of such procedure is that the result does not depend on the anomaly average. A two-dimensional model was designed with the topographic heights defined within a certain interval by a Gaussian error function:

$$h(x) = \begin{cases} h_{\max} e/-(ax)^2 & -\frac{d}{2} \le x \le \frac{d}{2} \\ h_0 & |x| > \frac{d}{2} \end{cases} . \tag{3}$$

The horizontal extent of the topographic structure, the wavelength d, is controlled by the exponential factor a, the height maximum $h_{\max}$ and the constant elevation h_0 outside the interval:

$$d = \frac{2}{a}\sqrt{\ln\left(\frac{h_{\max}}{h_0}\right)} \Rightarrow h_0 = h(x)|_{x=\pm d/2} \, . \tag{4}$$

Table III. Model parameters used for investigating the relation between gravity anomaly and mean elevation

	Alps	Andes
ϱ	2670 $\mathrm{kgm^{-3}}$	2670 $\mathrm{kgm^{-3}}$
$\Delta\varrho$	−400 $\mathrm{kgm^{-3}}$	−500 $\mathrm{kgm^{-3}}$
T	30 km	40 km
$h_{\max}$	3 km	3 km
h_0	0.2 km	0.2 km

The depth of the Airy-Heiskanen root follows from Eq. (2):

$$t(x) = h(x)\frac{\varrho}{\Delta\varrho} + T\,. \tag{5}$$

Two sets of parameters (Table III) have been used characterizing the different types of continental crust in the Alps and the Central Andes (Götze et al. 1991). These assumptions are very simple, but quite realistic. Using same parameters in a nonlinear inversion of the Eastern Alps' Bouguer gravity (Granser et al. 1989) yields to Moho-depths, which generally agree well with results obtained by seismic investigations (e.g. Miller et al. 1978, Posgay et al. 1991).

Figure 5 shows the results of a linear regression analysis regarding the Bouguer gravity-height relation for different topography wavelengths of both models. The Bouguer correction term can never be obtained except for a wavelength approaching to infinity. The height dependence varies remarkably in case of over- or undercompensated topography. The dashed lines describe such situation for the Alpine model. In spite of the simple model assumptions the observed regression coefficients fit very well to the coefficients which correspond to realistic figures for the crosswise extent of Eastern Alps and Central Andes (Götze et al. 1991). This indicates a more or less compensated crust in both areas.

A similar analysis was performed concerning the isostatic anomalies. In this case the functional relation between anomaly and mean height also depends on the state of equilibrium and therefore enables us to distinguish between different compensation models. However, the high frequency content of isostatic anomalies troubles the determination of functional relations in case of real data. Nevertheless one can try to evaluate the trend in the observed data set. Some models with different compensation scenario have been calculated, two of them are shown in Figs 6 and 7. Figure 6 refers to a simple over- and under-compensated crust obtained by 10 % increase and decrease respectively of the undulation in Eq. (2). In such a case a clear trend is present in the anomaly-height relation. Figure 7 describes the gravity effect of a root shifted with respect to topography by 20 km. That means undercompensation in the South and overcompensation in the North of the mountain crest. Such a model prevents a distinct anomaly trend. Disregarding the problem of geological corrections this feature fits well to the long-wavelength component of the isostatic anomaly in the Eastern Alps (Fig. 8), which also shows

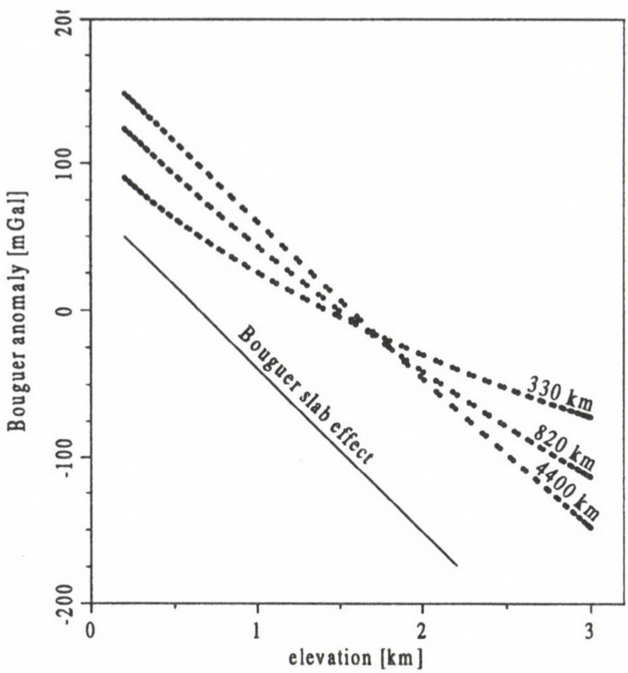


Fig. 5. Regression coefficient of the relation between Bouguer anomaly and mean elevation as function of the horizontal extent of a topographic mountain structure

the minimum in the North and the maximum in the South. The filled dots refer to the synthetic model, the open symbols show the relation between real data extracted from the low-pass filtered maps of the isostatic anomaly and the heights. Only data within the section of Fig. 4 are considered. The regression coefficient of the relation between isostatic anomalies and heights is close to zero both for the model and the real data. However, such a model does not agree with seismic refraction and reflection observations in the Swiss Alps (Kahle et al. 1980). Just contrarily maximum Moho depths are found here south of the Alpine crest corresponding to the Alpine tectonic pattern of plate collision. Three-dimensional modelling of the Bouguer gravity (Schwendener and Mueller 1990) suggests a deep structure in the upper mantle which is interpreted as subducted lithospheric slab in agreement with seismological results (Mueller 1989). Its gravity effect is rather big but completely masked by shallower structures and therefore not directly recognized in the gravity maps. Unfortunately a similar analysis cannot be performed in the Eastern Alps, as seismic results concerning the crustal structures are not sufficient to separate the gravity effects from crustal and sub-crustal sources. Besides, the ambiguity of gravity modelling has to be kept in mind. Cassinis et al. (1990) present a two-dimensional interpretation of a gravity profile in the Western Alps with no need to include sources in the upper mantle.

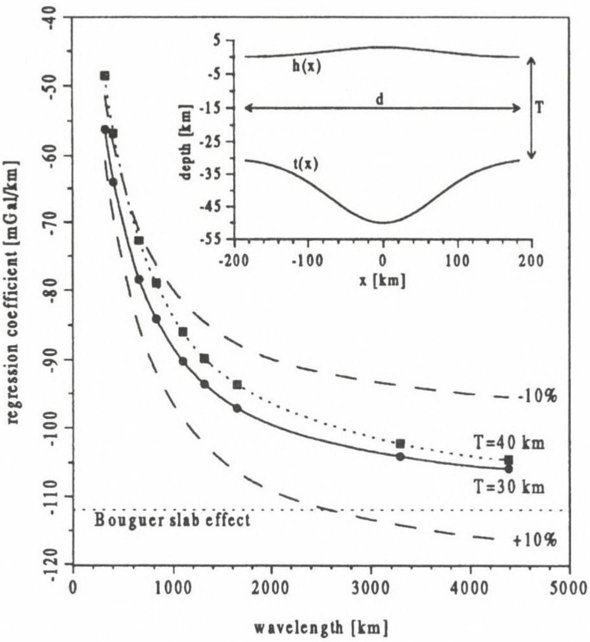


Fig. 6. Relation between isostatic anomaly and mean elevation for a simple crustal model. Topography is overcompensated (+) and undercompensated (−) by 10 % of the equilibrium root depth. The parameter corresponds to the exponential factor in Eq. (3) and characterizes the horizontal extent of the topographic mountain structure

5. Recent crustal movements and isostatic anomaly in the Eastern Alps

In the Eastern Alps recent crustal movements can be derived from repeated precise levelings (Höggerl 1989). The movement data shown in Fig. 9 (BEV 1991) refer to a station situated in the Bohemian massif which can be regarded as geologically most stable area in Austria. However, these results must be interpreted very carefully because at most locations they do not much exceed the standard deviation. Many sites are located within large valleys or inner Alpine basins where local subsidence effects may occur. Nevertheless this map can serve as indicator of regional trends in that area. Uplift prevails in the western part of Austria with maximum figures close to the crest area of the Alps, while subsidence is observed in the North and East. Crustal movement data are significantly only related to the long-wave component of the isostatic anomaly. This holds also, if the sites within the Vienna basin or the Bohemian massif are excluded which in a narrow sense do not belong to the Alpine area. However, this analysis does not gain full information since the levelling sites are unevenly distributed regarding the isostatic anomaly trend. Nevertheless, the coherence of local isostatic minima with high crustal uplift indicates that rebound effects cannot be excluded, but in this case then must occur locally. More convincing coherence has been reported from the Western Alps, where the up-

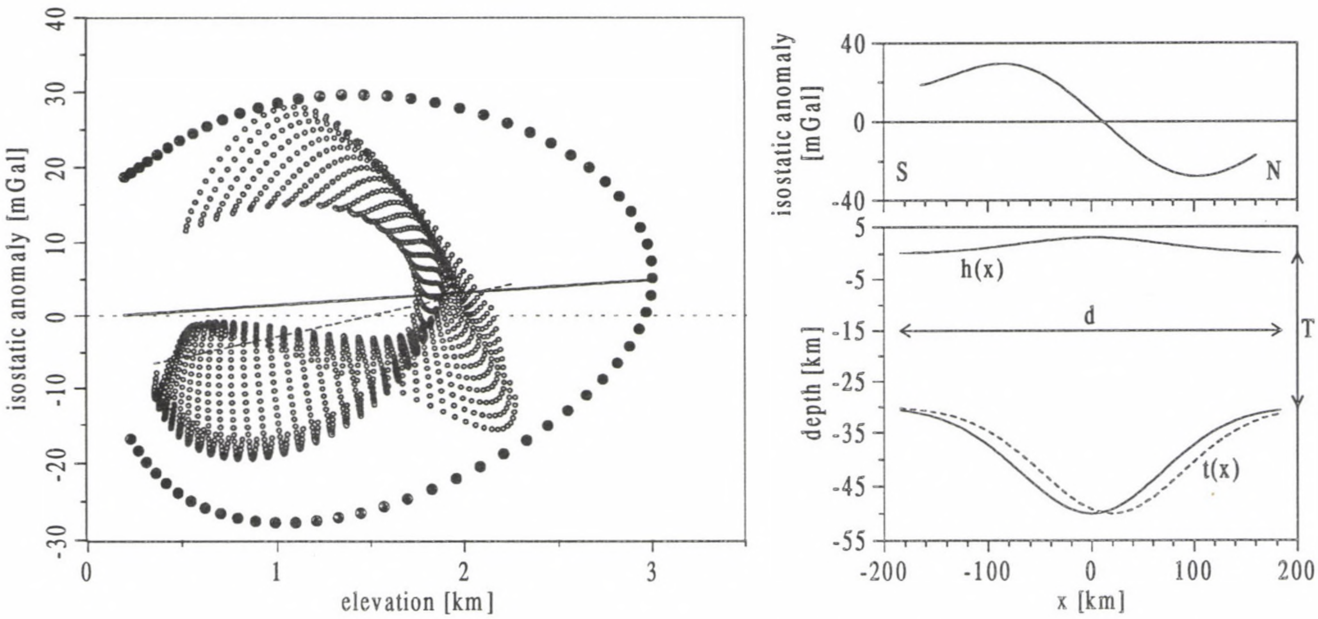


Fig. 7. Relation (full dots) between isostatic anomaly and mean elevation for a simple crustal model. Topography is undercompensated in the South and overcompensated in the North by shifting the root (dashed line) 20 km towards the North. The open circles characterize the relation between the observed long wavelength components of the isostatic anomaly and mean elevation within the section shown in Fig. 4

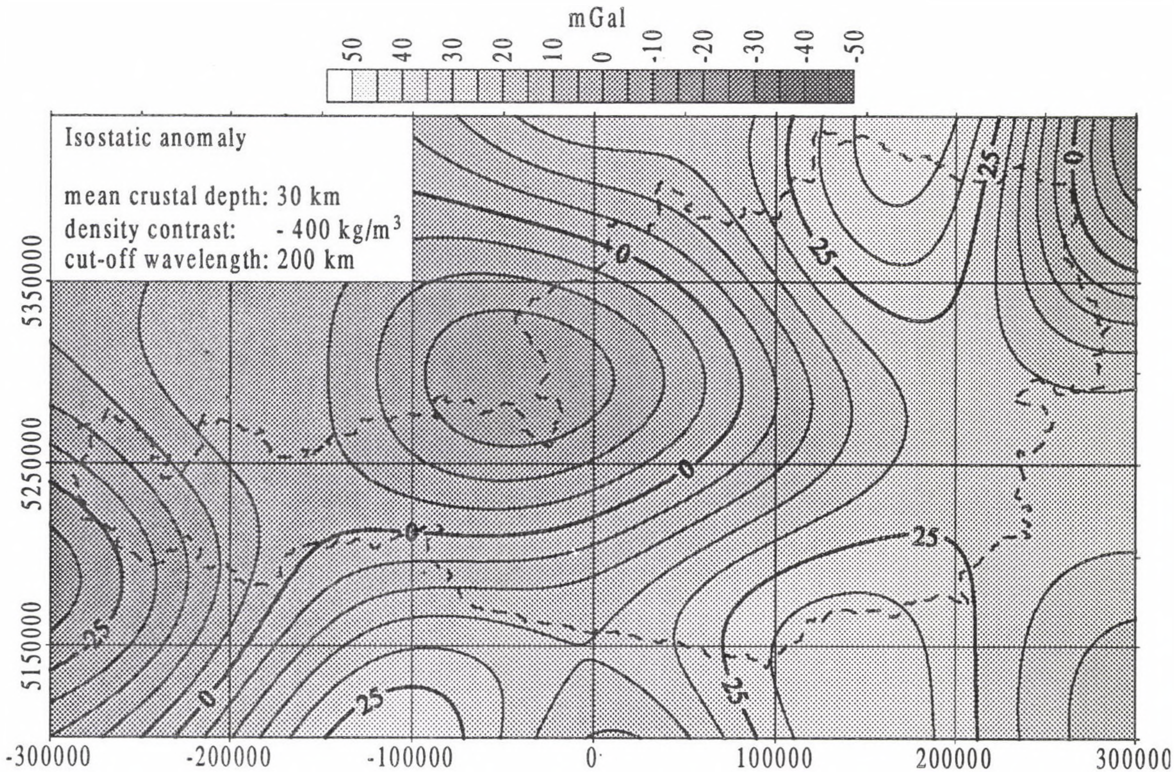


Fig. 8. Long wavelength component of the isostatic anomaly of the Eastern Alps (Fig. 2) obtained by wavelength filtering. Cut-off wavelength 200 km. Co-ordinates in m

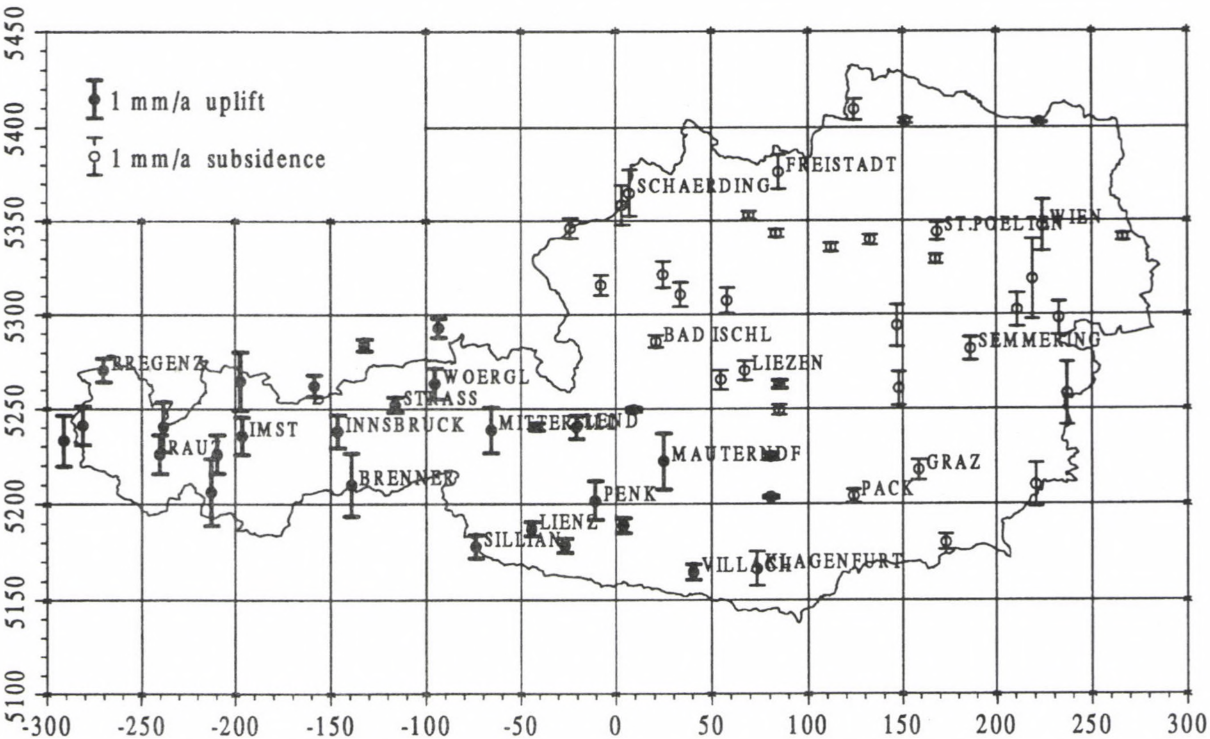


Fig. 9. Recent crustal movements of the Eastern Alps and surrounding area (BEV 1991). Co-ordinates in km

lift was explained by isostatic rebound (Kahle et al. 1980). A possible explanation for the weak correlation between isostatic anomalies and crustal movement data may be the existence of upper mantle density inhomogeneities mentioned above. They strongly influence just the long-wavelength component of gravity anomalies as shown by Schwendener and Mueller (1990). On the other hand we have to consider that recent crustal movements are also controlled by still active tectonic collision processes.

6. Conclusions

This example clearly shows that the geodynamical interpretation of isostatic anomalies is impossible without investigating in detail crustal and sub-crustal sources. The general assumptions applied for calculating the anomaly are too simple. Therefore the isostatic anomaly should be treated as residual field obtained by geophysically based high-pass filtering (e.g. Simpson et al. 1986). In the Eastern Alps the regional characteristic of the isostatic anomaly cannot be interpreted by simple two layer models as effect of a partially over- and undercompensated crust. This also holds for the explanation of the observed recent crustal movement pattern. Gravity effects of upper mantle structures ascribed to subducted lithosphere (Mueller 1989) may also change the regional trends in this area. Still active and maybe dominating tectonic processes can be another or additional explanation. Nevertheless rough estimates of the isostatic state are possible. So from the statistical analysis of the relation between gravity anomalies and mean elevations we can conclude that the Eastern Alpine area is close to isostatic equilibrium. The coincidence of high uplift rates with local negative isostatic anomalies indicates isostatic rebound effects at least in a local sense.

References

BEV 1991: Recent crustal movements in Austria. Bundesamt für Eich- und Vermessungswesen, Wien

Cassinis R, Cassano E, Capelli V 1990: *Tectonophysics*, 176, 183–192.

Götze H J, Lahmeyer B 1988: *Geophysics*, 53, 8, 1096–1108.

Götze H J, Meurers B, Schmidt S, Steinhauser P 1991: In: Andean Magmatism ans Its Tectonic Setting. R S Harmon and C W Rapela eds, Boulder, Colorado, Geological Society of America Special Paper 265, 279–290.

Granser H, Meurers B, Steinhauser P 1989: *Geoph. Prosp.*, 37, 279–292.

Heiskanen W A, Moritz H 1967: Physical geodesy. W H Freeman and Co., San Francisco

Höggerl N 1989: *Österr. Beitr. Met. Geoph.*, 2, 161–173.

Kärki P, Kivioja L, Heiskanen W A 1961: Topographic-isostatic reduction maps for the world for the Hayford zones 18–1, Airy-Heiskanen system, T=30 km. Publications of the Isostatic Institute of the International Association of Geodesy, 35.

Kahle H, Mueller St, Klingele E, Egloff R, Kissling E 1980: In: Earth rheology, isostasy and eustasy. N Mörner ed., John Wiley, New York, 377–388.

Meurers B 1992: *Österr. Beitr. Met. Geoph.*, 6, 146.

Meurers B, Ruess D, Steinhauser P 1987: In: Geodynamics of the Eastern Alps. H W Flügel and P Faupl eds, Verlag Deuticke, Wien, 334–344.

Miller H, Ansorge J, Arič K, Perrier G 1978: In: Alps, Appenines and Hellenides. H Closs, D Roeder, K Schmidt eds, E Schweizerbart'sche Verlagsbuchhdlg, Stuttgart

Moritz H 1980: *Bull. Geod.*, 54, 3, 395–405.

Mueller St 1989: In: Alpine Tectonics. M P Coward, D Dietrich D, R G Park eds, Geological Society Special Publication No. 45, 303–328.

Müller K 1988: Anwendung moderner gravimetrischer Interpretationstechniken zur Erkundung der Kruste in Süddeutschland. Dissertation, TU Clausthal

Parker R L 1972: *Geophys. J. R. Astr. Soc.*, 31, 447–455.

Posgay K, Albu I, Mayerova M, Nakladalova Z, Ibrmajer I, Blizkovsky M, Arič K, Gutdeutsch R 1991: *Geoph. Transactions*, 36, 7–13.

Schwendener H, Mueller St 1985: In: Second EGT Workshop — The Southern Segment. D A Galson, St Mueller eds, European Sience Foundation, Strasbourg, 115–120.

Schwendener H, Mueller St 1990: *Tectonophysics*, 176, 193–214.

Senftl E 1965: Schwerekarte von Österreich. 1:1000000. Bundesamt für Eich- und Vermessungswesen

Simpson R W, Jachens R C, Blakely R J 1983: A FORTRAN program for calculating the gravitational attraction of an Airy isostatic root out to 166.7 km. USGS, Open file report, 83–883.

Simpson R W, Jachens R C, Blakely R J, Saltus R W 1986: *J. Geoph. Res.*, 91, 8348–8372.

Steinhauser P, Ruess D, Zych D, Haitzmann H, Walach G 1984: In: Proc. XVIIIth General Assembly of the IUGG, IAG, Vol. 1, 322–338.

Wagini A, Steinhauser P, Meurers B 1988: Isostatic residual gravity map of Austria. USGS, Open file report, 87–402.

Walach G 1987: In: The gravity field in Austria. Geod. Arb. österr. f. d. Int. Erdmessung, Neue Folge, Band IV, 3–9, österr. Komm. f. d. Int. Erdmessung, Graz

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 405-410 (1996)

THE AUSTRIAN ABSOLUTE GRAVITY PROJECTS

D RUESS[1]

In Austria 1986 the ballistic absolute gravimeter JILAG-6 was purchased in 1986 by seven scientific institutes (ZAMG Wien, IMG – UNI Wien, GBA Wien, IG – MU Leoben, IAGP – TU Graz, ITG – TU Graz, IWF – ÖAW Graz). In a very short review, the activities on absolute gravity observations are given. Several projects in determining the absolute gravity were initiated since 1987 in cooperation with the BEV.

Keywords: absolute gravity; Austria; ballastic gravimeter

Objectives of the absolute projects

The most important purposes are

— datum of the Austrian gravity base net (ÖSGN)

— measurements in geodynamic sensitive zones

— regarding the repeatability of measurements.

In addition co-operations in bilateral and international absolute gravity projects were started.

Absolute gravity projects in Austria

Base stations

The first 4 absolute base stations of the Austrian gravity base network (ÖSGN) were established in 1980 and measured by the absolute gravity meter IMGC-Italy (Marson and Steinhauser 1981). Remeasurements with the JILAg-6 were made in 1987/1988 (Ruess and Steinhauser et al. 1989). Further 3 base stations were founded in Innsbruck, Klagenfurt, Vienna to stabilise the ÖSGN. Also the other absolute stations of the projects mentioned below are included in the adjustment of the ÖSGN (Ruess and Gold 1995).

Obergurgl / Tyrol

In 1987 the station Obergurgl in the Central Alps was established. Measurements have been repeated twice a year in spring and in autumn (Ruess 1995b). On one hand the aim was to detect seasonal gravity effects due to precipitations (rainfall in summer, snowfall in winter), on the other hand the repeatability of gravity

[1]Federal Office of Metrology and Surveying (BEV), Schiffamtsgasse 1–3, A–1025 Wien, Austria, e-mail: diethard.ruess@univie.ac.at

Table I. Absolute gravity activities in Austria

Period	Stations	Observations	Project
1980	4	4	Base stations (IMGC)
1987–1988	7	8	ÖSGN
1987–	1	16	Obergurgl / Tyrol
1988–1995	4	6	Hochkar Calibration Line (HCL) Göstling–Hochkar
1989–1990	3	4	local calibration line Leoben–Präbichl
1988	3	3	Rhine Valley
1987–1988	1	2	Mannswörth
1991–1992	6	10	Vienna Basin
1996–1998	4	0	Karawanken (periadriatic lineament)
Total	29	49	

measurements could be checked. The time series is shown in Fig. 1. There it seems to be a periodic change in the results. There is not yet a realistic explanation of the reason for such an effect. Instrumental effects like instability of the laser intensity and response to the laser comparator cannot be excluded. On the other hand, the maximum possible difference in the comperator effect is less then 15 μGal (Sasagawa et al. 1995). At the JILAg-6 instruments tests on different setups show differences below 5 μGal in the results. Further investigations in analysing the Obergurgl-curve are going on.

Vienna Basin

In Mannswörth one station was observed in 1987/1988 at the trench of Schwechat, 6 stations were established 1991/1992 in a profile Hirtenberg–Kaisereiche across the trench of the Southern Vienna Basin to control the effects of subsetting together with levelling control measurements (Ruess et al. 1993).

Hochkar Calibration Line (HCL) Göstling-Hochkar

The HCL was established in 1982 to check the La Coste and Romberg gravimeters in Austria (Meurers and Ruess 1985). The endpoints of this line were fixed in 1988 by absolute measurements. Remeasurements were realized in 1995, also the 2 intermediate points were added. Comparisons between the results of relative measurements and the absolute measurements show very good fit (Table II).

Leoben–Präbichl calibration line

3 stations of the local gravimeter calibration line were determined in 1989/1990 (Posch et al. 1989).

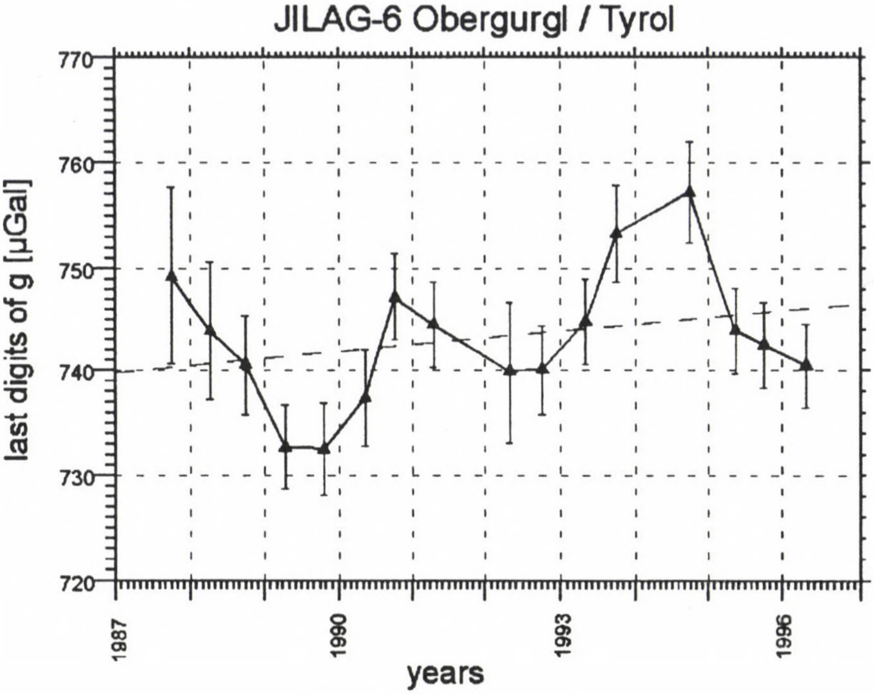


Fig. 1. Repeated measurements in the Central Alps

Table II. Differences between relative network adjustment and absolute connections at the 4 main stations of the Hochkar Calibration Line

ÖSGN Number	Location	Latitude [°]	Longitude [°]	Height m	Adjusted network 980 ...	±	Absolute gravity 980 ...	±	Δg
1–071–00	Göstling, Nepomuk	47.8084	14.9363	529	683 149	3	683 149	4	0
1–101–10	Lassing, toll-house	47.7469	14.9017	685	641 999	4	642 005	5	–6
1–101–20	Aiblboden, house	47.7352	14.9091	1113	556 266	3	556 269	5	3
1–101–30	Hochkar, Schiheim	47.7191	14.9177	1485	484 821	3	484 824	3	3

Rhine Valley

3 absolute stations were established in the Austrian part of the Rhine Valley by the Geophysical Institute at the Montanistic University in Leoben to check the transition zone between the Eastern and the Western Alps (Posch and Walach 1989).

Table III. JILAg-6 absolute gravity activities outside of Austria

Period	Stations	Observations	Country/project
1989–1994	1 (2)	3	comparisons at BIPM (Sévres, France)
1990-1994	6	6	Germany
1992–1995	4	4	Czech Republic
1993	4	4	Slovakia
1991–1995	5	6	Hungary
1994	5	5	Switzerland
1997	2	–	Slovenia
1997–1998	4	–	UNIGRACE
Total	25	28	

Karawanken

4 new absolute stations are prepared in the Karawanken in the area of the Periadriatic Lineament. First absolute measurements should start in the late autumn 1996, before the Austrian absolute gravity meter will be upgraded by some technological developments of the FG5 instruments. The aim of this project is to detect geodynamic effects in combination with levelling and GPS observations. Further absolute stations will be occupied in the surrounding area in the north (Klagenfurt) and in the south (Slovenia). Also a reference station in a stable geological zone (Altenburg, Bohemian Massif) will be measured.

International/bilateral contributions in absolute gravity

— In 1989 and 1994 the JILAg-6 attended the international comparisons of absolute gravity measurements at BIPM in Sévres/France. Checking the results of the last comparisons 1994 an offset of the data was detected by the JILAG-instruments, caused by a comparator effect (Sasagawa et al. 1995). Generally the correction due to this effect was received with 14 μGal. Indeed the contribution of the comparator was 8 μGal for the JILAg-6 instrument. After introducing this new correction, the results of the Austrian instrument fit very well.

— In Germany 2 base stations in Nordrhein-Westfalen were measured in Espelkamp and Aachen (1990). Further 3 base stations were observed in Sachsen in Landwüst, Freiberg and Collm (1994). At the base station Clausthal a comparison with the observations of the JILAG-3 instrument were made in 1991.

— In the Czech Republic in 1990 the station Pečny was remeasured (Šimon 1992). In 1995 three new stations were established in Litomerice, Benešov and Valtice.

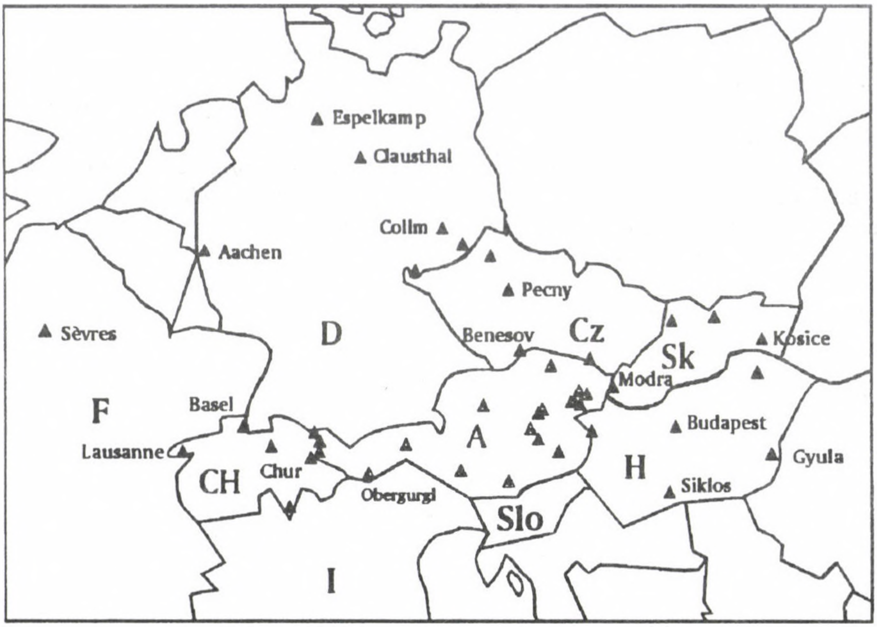


Fig. 2. JILAG-6 gravity observations in Europe

— In Slovakia 4 stations were measured in 1993 in Zilina, Ganovce, Košice and Modra.

— In Hungary the stations Budapest and Siklós were occupied in 1991, in 1993 Kőszeg and Szerencs, in 1995 Gyula and Siklós remeasured (Csapó et al. 1993).

— In 1994 in Switzerland 4 new stations were established in Lausanne, Basel, Zürich and Mt. Ceneri, the station Chur was remeasured.

— In Slovenia it is planned to observe two new stations in 1997.

— UNIGRACE (unification of gravity systems of Central and Eastern European Countries) is a project that has been derived from the former activities concerning absolute gravity measurements and sea level monitoring in the framework of CERGOP. Absolute measurements will be established in Germany, Austria, Bulgaria, Croatia, Czech Republic, Finland, Hungary, Italy, Poland, Romania,. Slovakia, Slovenia using the absolute gravimeters of Germany, Austria, Finland, Italy, Poland. Measurements are intended in 1997/1998.

References

Csapó G, Meurers B, Ruess D, Szatmári G 1993: *Geophysical Transactions*, 38, 251–259.

Marson I, Steinhauser P 1981: *EOS, Trans. Am. Geoph. Un.*, 62, 258.

Meurers B, Ruess D 1985: Errichtung einer neuen Gravimeter — Eichlinie am Hochkar. *Österr. Zeitschr. f. Verm. u. Photogr.*, Heft 3.

Posch E, Walach G 1989: *Österr. Beitr. zu Meteorol. u. Geophysik*, 2, 147–152.

Posch E, Tengg St, Walach G 1989: *Österr. Beitr. zu Meteorol. u. Geophysik*, 2, 123–124.

Ruess D 1993: *Österr. Beitr. zu Meteorol. u. Geophysik*, 8, 113–126.

Ruess D 1995a: Das Österreichische Schweregrundnetz. EVM Nr. 80, BEV Wien

Ruess D 1995b: Die Absolutschwerestation Obergurgl. Institut für Hochgebirgsforschung der Leopold-Franzens-Universitt Innsbruck, Jahresbericht 1995, 57-60.

Ruess D, Gold W 1995: The Austrian Gravity Base Net 1995 stabilised by Absolute Gravity Measurements and Connected to the European Network Adjustment. XXI. IUGG gen. meet. in Boulder, Colorado, USA, VGI 3/96.

Ruess D, Steinhauser P, Jeram G, Faller J 1989: *Beitr. zu Meteorol. u. Geophysik*, 2, 95–110.

Ruess D, Duma G, Nief M 1993: *Österr. Beitr. zu Meteorol. u. Geophysik*, 8, 127–142.

Sasagawa G S, Klopping F, Van Dam T M 1995: *Metrologia*, 32, 185–192.

Šimon Z 1992: Fourth step of the absolute gravity measurement at the Observatory Pecný. Geodetický a kartografický obzor 38/80, No. 8.

Steinhauser P 1987: Absolutschweremessungen in Österreich. Oppolzer-Gedchtnissymposium 1986, Geodt. Arbeiten Österr. f. d. Int. Erdmessung, Bd. 5.

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 411-421 (1996)

TOPOGRAPHIC-ISOSTATIC MODELS FITTING TO THE GLOBAL GRAVITY FIELD

GY TÓTH[1]

The topographic-isostatic potential of the earth's crust can be modeled very simply using average crustal parameters, a global isostatic model and a numerical dataset of mean continental and oceanic heights. In lack of the detailed data for density and crust thickness, a least squares estimation is suggested by the author to determine the global horizontal variation of crustal parameters.

These variations can be determined using a reasonable minimum principle to yield a minimum variance high frequency residual geoid. Mathematical solution of the problem is discussed using the product-sum formula of spherical harmonics.

Some calculations were also performed with these models and significant variations in crustal density up to ± 600 kgm^{-3} and crust thickness ± 7 km were detected with respect to the Airy model of uniform parameters. Also a different behaviour of the oceanic and continental crust seems to be indicated by these models.

Keywords: geopotential; lateral density variations; product-sum formula of spherical harmonics; topographic-isostatic model

1. Introduction

It is not easy to model the behaviour of the earth's crust on a global scale. The topography and the gravitational potential related to it are well known or can be modelled easily (Balmino et al. 1973). On the contrary, the exact mechanism of compensation and deep inhomogeneities inside the crust and upper mantle are much disputed or not so well known. Therefore we are in a need of crust models which are geophysically reasonable and reflect the gravitational contribution of crustal inhomogeneities rather well. There exist also global models of the earth's crustal thickness (Cadek and Martinec 1991).

The disturbing potential due to the density irregularities inside the crust is termed *topographic-isostatic potential.* This term also indicates the two principal density sources related to the crust. One needs the global models to evaluate topographic-isostatic potential. One of the most simple models which can also be interpreted physically is the Airy-Heiskanen isostatic model. Rummel et al. (1988) have developed a very efficient FFT-based technique for the computation of the gravitational effect of this model and of the Vening-Meinesz model. Also an attempt was made to fit these models globally to the earth's gravity field.

The following conclusion must be drawn from the study of such global topographic-isostatic models: Since the isostatic behaviour of the earth's crust is dependent on a number of factors these simple global models cannot be expected to reflect the full picture. At this point a useful idea may arise. Because even these simple models depend on a number of factors, e.g. crust density and thickness, let

[1]Department of Geodesy, Technical University of Budapest, H–1521 Budapest, Hungary

1217-8977/96/$ 5.00

us allow these parameters vary globally. This way, as we will see later on, one may restrict himself to those lateral variations of crust parameters which are related to the earth's topography. A spectral approach will prove to be very useful to determine these lateral parameter variations based on the knowledge of the geopotential. Available crust density and thickness data can also be integrated into these models. Some results and problems will be discussed as well.

2. Airy topographic-isostatic model and its potential

Here only a short review of the spherical harmonic analysis of the Airy topographic-isostatic potential will be given. For more details the reader should consult with (Rummel et al. 1988).

In the Airy model it is supposed that the light sialic (SiAl) crust matter of density ϱ_{cr} floats on the more heavy simatic (SiMa) material of the upper mantle of density ϱ_m. Each crust column is supposed to be in a perfect equilibrium state.

For ocean columns with depth h^* the anti-root thickness d^* and for continental columns of height h' the root thickness d' exist. By introducing the factor c_h,

$$\begin{aligned} c_h &= 1 && \text{for continents} \\ c_h &= 1 - \frac{\varrho_{cr}}{\varrho_w} && \text{for oceans}, \end{aligned}$$

one may define the *eqivalent heights* $h = c_h h'$ or $h = c_h h^*$. The eqilibrium equation is then of the following:

$$d = \frac{\varrho_{cr}}{\Delta\varrho} h = kh \,, \tag{1}$$

where $\Delta\varrho = \varrho_m - \varrho_{cr}$ is the density jump on the Mohorovičić discontinuity, k is the compensation factor.

The compensation factor k is constant for the flat earth approximation and equal to

$$k = k_0 = \frac{\varrho_{cr}}{\Delta\varrho} \,. \tag{2}$$

For a spherical earth approximation k is slightly different. It can be computed as a function of h' from the mass balance principle (see e.g. Sünkel 1986):

$$\frac{kh'}{R-D} = \left\{1 - \left(1 - \frac{D}{R}\right)^{-3} k_0 c_h \left[\left(1 + \frac{h'}{R}\right)^3 - 1\right]\right\}^{\frac{1}{3}} . \tag{3}$$

Where R denotes mean earth radius and D is the thickness of crust.

If $\delta\varrho$ denotes mass irregularities to the relativ Airy model, the topographic-isostatic potential T^{Airy} is defined by the Newton's law of gravitation:

$$T^{\text{Airy}}(P) = G \iiint\limits_{\text{crust}} \ell^{-1}(P,Q)\delta\varrho(Q)dv(Q) \tag{4}$$

where G is Newton's gravitational constant, $\ell(P,Q)$ denotes the distance of P and Q and dv is the volume element of crust, T^{Airy} = potential anomaly.

Since $T^{\text{Airy}}(P)$ is harmonic outside a sphere and its spherical harmonic expansion is convergent outside the earth, it can be developed into a spherical harmonic series,

$$T^{\text{Airy}}(P) = \frac{GM}{r(P)} \sum_{n=1}^{\infty} \left(\frac{R}{r(P)}\right)^n \sum_{m=-n}^{n} \overline{C}_{nm}^{\text{Airy}} \overline{Y}_{nm}(P) , \tag{5}$$

where $\overline{C}_{nm}^{\text{Airy}}$ $(n = 1, 2, \ldots\ m = -n, -n+1, \ldots, -1, 0, 1, \ldots, n)$ denote the fully normalized spherical harmonic coefficients of the topographic-isostatic potential, M is the total mass of earth, $r(P)$ is the geocentric distance and

$$\overline{Y}_{nm}(P) = \sqrt{2^{1-\delta_{mo}}(2n+1)\frac{(n-|m|)!}{n+|m|)!}} \overline{P}_{n|m|}(\cos\Theta_P) \cdot$$

$$\cdot \begin{Bmatrix} \cos|m|\lambda_P, & \text{if} & m \geq 0 \\ \sin|m|\lambda_P, & \text{if} & m < 0 \end{Bmatrix} \tag{6}$$

are fully normalized spherical harmonics and Θ_P, λ_P denote polar distance and longitude of P respectively and δ_{ij} is the Kronecker's delta. The $\overline{P}_{nm}$ are the unnormalized Legendre functions of the first kind defined by the equation

$$P_{nm}(t) = \frac{1}{2^n n!}(1-t^2)^{\frac{m}{2}} \frac{d^{n+m}}{dt^{n+m}}(t^2-1)^n \qquad \begin{matrix} n = 0, 1 \ldots \\ m = 0, 1, \ldots, n \\ (t = \cos\Theta) \end{matrix} . \tag{7}$$

The summation in Eq. (5) begins at $n = 1$ because there is no mass surplus or deficit in this compensation model.

If the following surface spherical harmonic coefficients of the equivalent topography h and of its square h^2 are introduced,

$$h\overline{c}_{nm} = \frac{1}{4\pi} \iint_{\sigma} \left(\frac{h}{R}\right) \overline{Y}_{nm} d\sigma , \tag{8a}$$

$$h^2\overline{c}_{nm} = \frac{1}{4\pi} \iint_{\sigma} \left(\frac{h}{R}\right)^2 \overline{Y}_{nm} d\sigma , \tag{8b}$$

then these integrals can be evaluated very efficiently and the spherical harmonic coefficients of the topographic-isostatic potential can be approximated up to the second order through the following expression (Rummel et al. 1988):

$$\overline{C}_{nm}^{\text{Airy}} = \frac{3}{2n+1} \frac{\varrho_{cr}}{\overline{\varrho}} \left\{ \left[1 - \left(\frac{R-D}{R}\right)^n\right] h\overline{c}_{nm} + \frac{n+2}{2} \cdot \right.$$

$$\left. \cdot \left[1 - \frac{\varrho_{cr}}{\Delta\varrho} \left(\frac{R-D}{R}\right)^{n-3}\right] h^2\overline{c}_{nm} \right\} , \qquad \begin{matrix} n = 0, 1, \ldots \\ m = -n, \ldots, 0, \ldots, n \end{matrix} \tag{9}$$

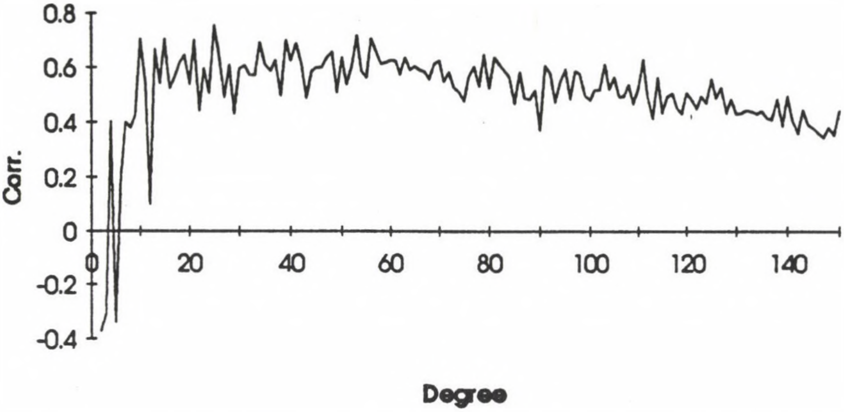


Fig. 1. Degree correlations between the Rapp (1981) geopotential and the Airy model's topographic-isostatic potential in the range of $n = 2 - 150$

where $\overline{\varrho} = \frac{3M}{4\pi R^3}$ is the mean earth density (cca. 5514 kgm^{-3}) and $\iint\limits_{\sigma} \ldots d\sigma$ denotes integration over the unit sphere σ.

From a $1° \times 1°$ dataset of mean topographic heights (64800 data for the whole earth) the $h\overline{c}_{nm}$ and $h^2\overline{c}_{nm}$ (where $h^2\overline{c}_{nm}$ denotes spherical harmonic coefficients of the h^2 function) coefficients can be computed up to degree and order 180, and the corresponding T^{Airy} and

$$N^{\text{Airy}} = \frac{T^{\text{Airy}}}{\gamma} \tag{10}$$

topographic-isostatic geoid can be determined (here γ denotes normal gravity).

To judge the fit (or similarity) of the two functions T^{Airy} and the T disturbing potential of the earth the following measure of similarity, the degree correlation coefficient is often used (Rummel et al. 1988):

$$c_n = \frac{\sum\limits_{m=-n}^{n} \overline{C}_{nm} \overline{C}_{nm}^{\text{Airy}}}{\sigma_n \sigma_n^{\text{Airy}}} \tag{11}$$

where σ_n^2 is the signal variance, $\sigma_n^2 = \sum\limits_{m=-n}^{n} \overline{C}_{nm}^2$. Also the average correlation coefficient between degrees of series expansion n_1 and n_2 is used:

$$\overline{c}(n_1, n_2) = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} c_n \,. \tag{12}$$

When the geopotential model Rapp (1981) is used for comparison, Fig. 1 shows the degree covariances for a simple Airy model of average crustal thickness D=30 km.

The fit between the two potential coefficient sets is rather poor even in the higher degree range (above degree 90) where great part of the gravity signal is to be expected to be yielded by the topographic-isostatic mass irregularities. Even a global fitting of average crust parameters, as it was done by Rummel et al. (1988) does not significantly make this correlation better. So we see that this simple Airy model has its strict limitations and it cannot be expected to reflect adequately the upper density inhomogeneities inside the crust on a global scale.

3. Topographic-isostatic models pertaining to lateral variations in the Airy model

When the crust density ϱ_{cr} and crust thickness D vary according to the lateral position of the point P, some changes can be introduced with respect to their average values:

$$\varrho_{cr}(P) = \overline{\varrho}_{cr} = \Delta\varrho_{cr}(P) \tag{13a}$$

$$D(P) = \overline{D} + \Delta D(P) . \tag{13b}$$

The following linearization is valid at the Taylor-point of the parameter vector $\overline{\delta}_0 = (x\overline{\varrho}_{cr}, \overline{D})$

$$\Delta T^{\text{Airy}} = \left.\frac{dT^{\text{Airy}}}{d\delta}\right|_{\delta_0} \Delta\delta . \tag{14}$$

From Eq. (5) it is clear that the following expression must be valid for the above potential change

$$\Delta T^{\text{Airy}}(P) = \frac{GM}{r(P)} \sum_{n=1}^{\infty} \left(\frac{R}{r(P)}\right)^n \sum_{m=-n}^{n} \left.\frac{d\overline{C}_{nm}^{\text{Airy}}}{d\delta}\right|_{\delta_0} \Delta\delta(P)\overline{Y}_{nm}(P) , \tag{15}$$

since only the coefficients $\overline{C}_{nm}$ are depending on the parameter vector $\delta = (\varrho_{cr}, D)$. If now we restrict ourselves to the linear term in Eq. (9) the following expression will be yielded for the derivative vector in Eq. (15):

$$\left.\frac{d\overline{C}_{nm}^{\text{Airy}}}{d\delta}\right|_{\delta_0} \Delta\delta = \frac{3}{2n+1}\frac{\overline{\varrho}_{cr}}{\overline{\varrho}} \left\{ \begin{matrix} 1-\left(1-\frac{\overline{D}}{R}\right)^n \\ n\frac{\overline{D}}{R}\left(1-\frac{\overline{D}}{R}\right)^{n-1} \end{matrix} \right\} \cdot$$

$$\cdot \frac{1}{4\pi} \iint_\sigma \left\{ \begin{matrix} \left(\frac{\Delta\varrho_{cr}}{\overline{\varrho}_{cr}}\right) \\ \left(\frac{\Delta D}{\overline{D}}\right) \end{matrix} \right\} \frac{h}{R}\overline{Y}_{nm} d\sigma . \tag{16}$$

If we denote the sum of these two terms by $\Delta\overline{C}_{nm}^{\text{Airy}}$, it is easy to show that the following expression is valid up to a second order approximation.

$$\Delta\overline{C}_{nm}^{\text{Airy}} = \frac{3}{2n+1}\frac{\overline{\varrho}_{cr}}{\overline{\varrho}} \left[1-\left(1-\frac{\overline{D}}{R}\right)^n\right] \cdot$$

$$\cdot \left\{ [1-M_n]\frac{1}{4\pi} \iint_\sigma \delta_+ \frac{h}{R}\overline{Y}_{nm} d\sigma + M_n \frac{1}{4\pi} \iint_\sigma \delta_- \frac{h}{R}\overline{Y}_{nm} d\sigma \right\} , \tag{17}$$

where

$$M_n = \frac{n-1}{4}\frac{\overline{D}}{R}\left(1 + \frac{n-5}{6}\frac{\overline{D}}{R}\right), \tag{18}$$

and

$$\delta_+ = \frac{\Delta\varrho_{cr}}{\overline{\varrho}_{cr}} + \frac{\Delta D}{\overline{D}}, \tag{19a}$$

$$\delta_- = \frac{\Delta\varrho_{cr}}{\overline{\varrho}_{cr}} - \frac{\Delta D}{\overline{D}}. \tag{19b}$$

The relative magnitude of the second term in Eq. (17) is 0.112 for $n = 90$ and 0.195 for $n = 150$ if $\overline{D} = 30$ km, so it is to be expected that the δ_+ term is dominant. This means that only the sum of these two functions can be determined through their product by the h/R function if we accept the Airy model's physical background.

4. An optimum criterion for topographic-isostatic crust models

The gravity potential of the earth includes the topographic-isostatic potential of the real earth's crust. This potential is included in the gravity potential in such a way that the shorter the wavelength of the gravity potential terms in the spherical harmonic expansion, the higher the contribution of the topographic-isostatic potential is to it. This is a well-known property of the Newton's kernel $\ell^{-1}(P,Q)$ in Eq. (4). So simply saying the crust becomes the most important density source of the gravity potential as the spatial frequency of T increases. This also means that the shorter the wavelength, the smaller the disturbing effect of non-crustal masses is.

When the topographic-isostatic potential is modelled, our model has to reflect the gravity potential well at shorter wavelengths. This criterion can be used to judge between such models. From this point of view, the above criterion may be used to select a best or optimal model. When the residual gravity potential is $\Delta T = T - T^{\text{model}}$ for a certain topographic-isostatic model, the optimality criterion can be expressed as

$$\sum_{n=n_1}^{n_2} \beta_n \sigma_n^2(\Delta T) = \text{minimum}. \tag{20}$$

The de-smoothing factor β_n, which expresses the amplification of higher frequency harmonics, can be determined theoretically from an ideal model of uncorrelated uniform density inhomogeneities (Völgyesi and Tóth 1992). When $D_{\max}$ denotes the maximum depth of crustal density sources, the theoretical amplification function is

$$\beta_n = 1\left(1 - \frac{D_{\max}}{R}\right)^{2n+3}. \tag{21}$$

5. Optimal linear topographic model determination

The determination of an optimal linear topographic model (OLTM) (Völgyesi and Tóth 1992) requires mathematically the determination of the parameter function δ_+ defined on the surface of the earth. In the following we will denote this function simply by δ. The spherical harmonic coefficients $\overline{C}_{nm}^{\text{model}}$ of the optimal model will then be computed from the formulae below, which are coming from the Eq. (17) if the δ_- term is neglected.

$$\overline{C}_{nm}^{\text{model}} = \overline{C}_{nm}^{\text{Airy}} + \Delta\overline{C}_{nm}^{\text{Airy}} , \tag{22}$$

where

$$\Delta\overline{C}_{nm}^{\text{Airy}} = t_n \cdot h\delta c_{nm} , \tag{23}$$

$$t_n = \frac{3}{2n+1}\frac{\overline{\varrho}_{cr}}{\overline{\varrho}}\left[1-\left(1-\frac{\overline{D}}{R}\right)^n\right][1-M_n] \tag{24}$$

and

$$h\delta c_{nm} = \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi}\frac{h(\Theta,\lambda)}{R}\cdot\delta(\Theta,\lambda)\overline{Y}_{nm}(\Theta,\lambda)\sin\Theta\, d\Theta\, d\lambda\,. \tag{25}$$

These are the surface spherical harmonic coefficients of the *product function* $(h/R)\delta$. In the following we shall see how they may be represented by the surface spherical harmonic coefficients of its component functions h/R and δ.

Let the functions h and δ be represented mathematically by the following spherical harmonic series and coefficients:

$$h(\Theta,\lambda) = R\sum_{\ell=0}^{\infty}\sum_{k=-\ell}^{\ell} hc_{\ell k}\cdot\overline{Y}_{\ell k}(\Theta,\lambda)\,, \tag{26}$$

$$\delta(\Theta,\lambda) = \sum_{i=0}^{\infty}\sum_{j=-i}^{i} oc_{ij}\cdot\overline{Y}_{ij}(\Theta,\lambda)\,, \tag{27}$$

$$hc_{\ell k} = \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi}\frac{h(\Theta,\lambda)}{R}\overline{Y}_{\ell k}(\Theta,\lambda)\sin\Theta\, d\Theta\, d\lambda\,, \tag{28}$$

$$oc_{ij} = \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi}\delta(\Theta,\lambda)\overline{Y}_{ij}(\Theta,\lambda)\sin\Theta\, d\Theta\, d\lambda\,. \tag{29}$$

The mathematical tool needed for the computation of the spherical harmonic coefficients in the Eq. (25) by the coefficients (28) and (29) is the *product-sum conversion formula of spherical harmonics*. In an abbreviated form the following equation is valid between the aforementioned coefficients:

$$h\delta c_{nm} = \sum_{i=0}^{\infty}\sum_{j=-i}^{i} a(n,m,i,j)\cdot oc_{ij}\,. \tag{30}$$

The $a(n,m,i,j)$ coefficients can be determined from the $hc_{\ell k}$ coefficients and the Clebsch-Gordan (or Gaunt-Wigner) coefficients. The details of this relationship were treated in detail in Völgyesi and Tóth (1992), here we note only the following abbreviated form of this relation:

$$a(n,m,i,j) = \sum_{\ell} N_{ij\ell nm} C(i,\ell,n;0,0,0) \cdot C(i,\ell,n;j,m-j,m) \cdot hc_{\ell,m-j}\,, \tag{31}$$

where $N_{ij\ell nm}$ denote certain constants and $C(i,\ell,n;j,k,m)$ denote the Gaunt coefficients.

When we arrange the spherical harmonic coefficients and the corresponding $a(n,m,i,j)$ coefficients according to single indices instead of the double indices n,m and i,j then the following linear system of equations will be resulted:

$$\mathbf{C}^{\mathrm{model}} = \mathbf{C}^{\mathrm{Airy}} + \mathbf{A} \cdot \mathbf{oc}\,, \tag{32}$$

where in the matrix $\mathbf{A}$ one finds the elements $t_n \cdot a(n,m;i,j)$. The optimal parameter vector $\mathbf{oc}$ may now be estimated (up to a certain maximum degree and order $i_{\mathrm{max}} = K$) to make the variance of the high frequency residual field minimum as it was prescribed by the condition (20). This is a very well known least squares parameter estimation procedure for the minimizing parameter vector.

This way the optimum parameter function δ through its harmonic coefficients will be determined and then the computation of harmonic coefficients of the optimal model is quite straightforward from the linear system (32).

6. Results and conclusions

PC-based computer programs were developed by the author in MS FORTRAN to determine such best-fitting topographic-isostatic models. In our calculations the spherical harmonic coefficients were that of the Rapp (1981) model up to the degree and order 150. This limit was chosen because of the restricted computer memory however, it is planned to extend this upper limit up to 360 on more powerful computers. Also a common dataset of $1° \times 1°$ mean topographic elevations were used to produce the harmonic coefficients of equivalent topography up to the same degree and order 150.

Two different resolution best-fitting model was computed up to degrees $K = 5$ and 10. The average crust thickness for the reference Airy model was 30 km, the average crust density was 2670 kgm^{-3} and the density jump across the Moho was chosen to be 600 kgm^{-3}. The maximum crustal density source depth in Eq. (21) was 70 km. As an illustration in Fig. 2 the geoid height differences of the simple Airy topographic-isostatic model and of the optimum parameter function model of maximum degree 10 can be seen according to our computations. The corresponding rms topographic-isostatic geoid height differences are the following:

between the Airy model and the K=5 model:	r.m.s difference: ±4.0 m
between the Airy model and the K=10 model:	r.m.s difference: ±4.7 m
between the K=5 model and the K=10 model:	r.m.s difference: ±3.8 m

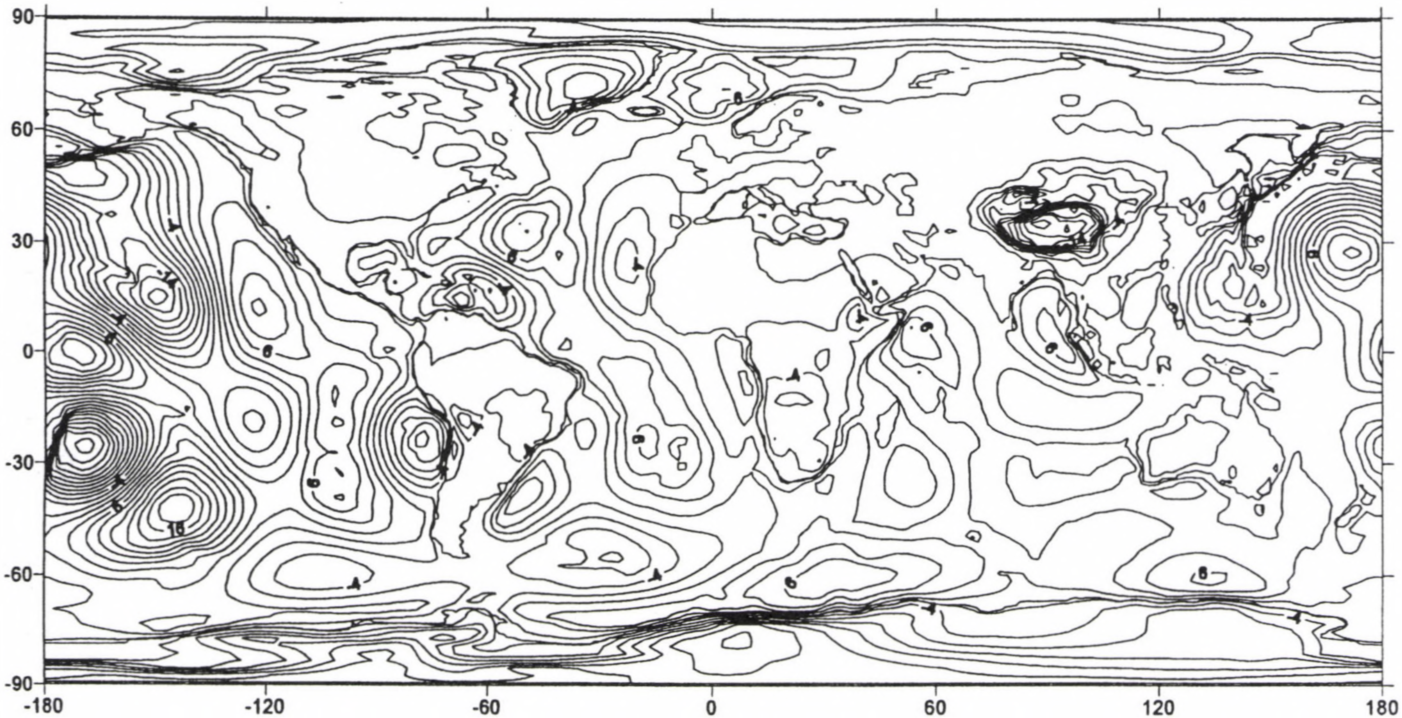


Fig. 2. Geoid height differences of the simple Airy topographic-isostatic model and degree 10 optimal model (in meters)

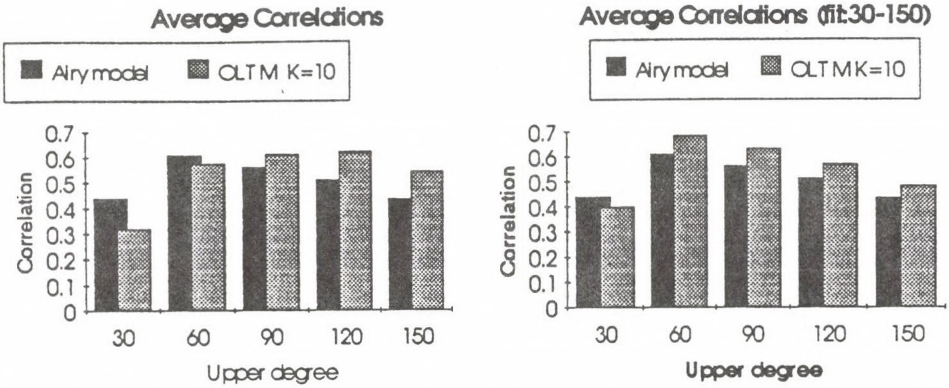


Fig. 3. Average correlations of the Airy and best-fitting models with respect to the Rapp (1981) geopotential model in the fit ranges 90–150 and 30–150

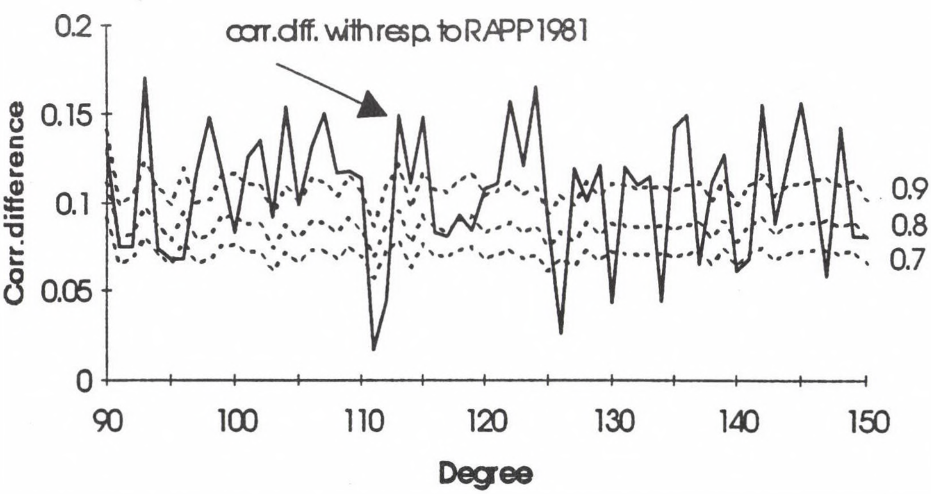


Fig. 4. Confidence levels for the difference of correlation coefficients of best-fit model of degree 10 and the Airy model if the reference geopotential model is the Rapp (1981) model

The average correlations of the degree 10 model and the Airy model with respect to the Rapp (1981) geopotential model are shown for two different fit ranges in Fig. 3. We applied a statistical test to decide whether these changes are significant or not and the confidence limits for correlation differences can be seen in Fig. 4. Finally the rms. contributions of continental and oceanic parts are visualized in Fig. 5.

Now it is necessary to point out some interesting features of these results. First, the resulting changes of crust density and thickness are reasonable, since from the parameter function (27) the numerical values are ± 600 kgm^{-3} and ± 7 km respectively. Second, it can be stated that a clear improvement of global topographic-isostatic models, compared to the simple Airy model may be achieved by allowing

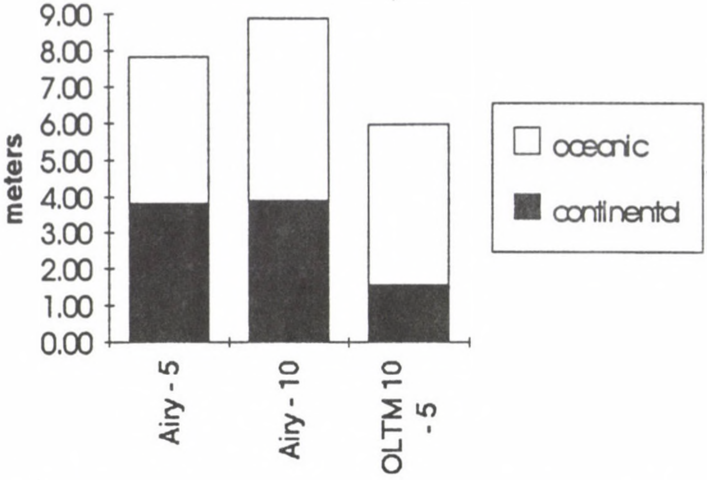


Fig. 5. Oceanic and continental r.m.s. contributions of different degree best-fitting topographic-isostatic models

horizontal variations of certain crustal parameters as crust thickness and density. The only source of information for this improvement was the geopotential itself. Third, our models show significant lateral density changes and variable crust density as a possible physical cause of topographical geoid anomalies. The significant mass deficit due to the large ice sheet of Antarctica is also clearly indicated by large negative values of the parameter function in Fig. 2. Negative values are mostly correlated with large mountain zones and ocean bottom areas. Positive values are associated with ocean trenches and old continental massifs. These results suggest the nonlinearity of compensation instead of the strict linear relation in the Airy model.

References

Balmino G K, Lambeck K, Kaula W 1973: *J. Geophys. Res.*, 78, 478–521.

Cadek O, Martinec Z 1991: *Studia geophys. et geod.*, 35, 151–165.

Rapp R H 1981: The earth's gravity field to degree and order 180 using seasat altimeter data, terrestrial gravity data, and other data. Reports of the Dept. of Ohio State Univ., Rep. No. 322.

Rummel R, Rapp R H, Sünkel H, Tscherning C C 1988: Comparisons of global topographic/isostatic models to the Earth's observed gravity field. Rep. No. 388, Dept. of Geod. Sci, The Ohio State University, Columbus, Ohio

Sünkel H 1986: In: Mathematical and numerical techniques in physical geodesy, Lecture Notes in Earth Sciences, 418–462.

Völgyesi L, Tóth Gy 1992: *Periodica Polytechnica, Ser. Civil Eng.*, 36, 207–241.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 423–431 (1996)

OBSERVATIONS OF FREE OSCILLATIONS OF THE EARTH BY SUPERCONDUCTING GRAVIMETER GWR T020

H VIRTANEN[1]

In August 1994, the superconducting gravimeter of the Finnish Geodetic Institute at the station Metsähovi started its registrations. At this site environmental noise level is low allowing the detection of weak geophysical phenomena.

Big earthquakes usually excite the free oscillations of the Earth. In this study the performance of the GWR T020 for observing this phenomenon in the seismic band up to one hour's period at a few nanoGals sensivity level is considered. The overall existence of the radial mode ${}_0S_0$ is demonstrated for the two years period of operation.

Keywords: earthquakes; free oscillations of the Earth; superconducting gravimeter

Introduction

The Finnish superconducting gravimeter GWR T020 was installed in August 1994 in the Metsähovi research station, 30 km west of Helsinki (Virtanen and Kääriäinen 1995a). A superconducting gravimeter has a low drift and a high resolution for revealing changes in gravity at the nanoGal level. In addition the instrument has a large range in the spectral domain starting from seconds up to months, so it can be used in studies from seismic bands to long periodic phenomena such as the Chandler period of polar motion. It has been proved that at the station Metsähovi the environmental noise level is low and therefore evidence of even weak effects can be expected. Seismic events appear to disturb a considerable part of the recording time. In order to estimate the significance of the disturbances, we have made some statistical studies on the number and duration of these effects. Finding that the mean standard deviation in one minute of data on a silent day is about 40 nGal, a threshold limit of 80 nGal for a disturbance of the gravity signal was adopted. This limit is somewhat arbitrary but was adopted after some practical tests. It appears that about 33 % of observing time includes discernible disturbances of seismic origin above this threshold, which is a surprisingly large rate (Virtanen and Kääriäinen 1995b).

Large earthquakes can influence recording over three days. Therefore we have studied our ability to make use of these big events, which can excite free oscillations of the Earth, employing the superconducting gravimeter as a long periodic seismometer.

The first observations of free oscillations were those of the large Kamchatka earthquake in 1952 (Benioff et al. 1954). Some few oscillations were found which

[1]Finnish Geodetic Institute, Geodeetinrinne 2 (PL 15), FIN–02431 Masala, Finland, e-mail: Heikki.Virtanen@fgi.fi

apparently were excited by the seismic event. Later on, a large number of different modes based on theoretical Earth models have been observed with seismometers and gravimeters in connection with large earthquakes. The gravity effects of these phenomena are very weak, being at the level of a few nanoGals.

Earthquake on October 4th, 1994, in Kuril Islands

For tentative study the biggest event appearing in our recording between August 11th 1994 and September 5th 1996 was chosen. This earthquake occurred on the Kuril Islands on October 4th, 1994. The time of origin was 13:23 UT, the moment magnitude (Mw) was 8.2 and the focal depth 33 km (Table I). The original one second samples of gravity signal were decimated to one minute using a symmetrical low pass digital filter of order 1225. Tides and influence of air-pressure were subtracted from these calibrated one minute samples using the programme PRETERNA (Wenzel 1995). The residual channel of one minute sampled data was used later on in all spectral analysis ranging from three minutes to one hour. Several tests were made to find the suitable length of time series and starting time after a quake. Different spectral methods and improvements for signal to noise ratio using digital filtering and precise short term linear drift corrections were investigated as well.

As a result the longer period part of the spectrum studied is presented in Fig. 1. The origin of the time series is 0 UT on October 5th, 10.5 hours after the quake. The length of the record was 48 hours containing 2880 samples. The period, given as cycles per hour (cph), ranges from 12 minutes to 60 minutes. The spectral method

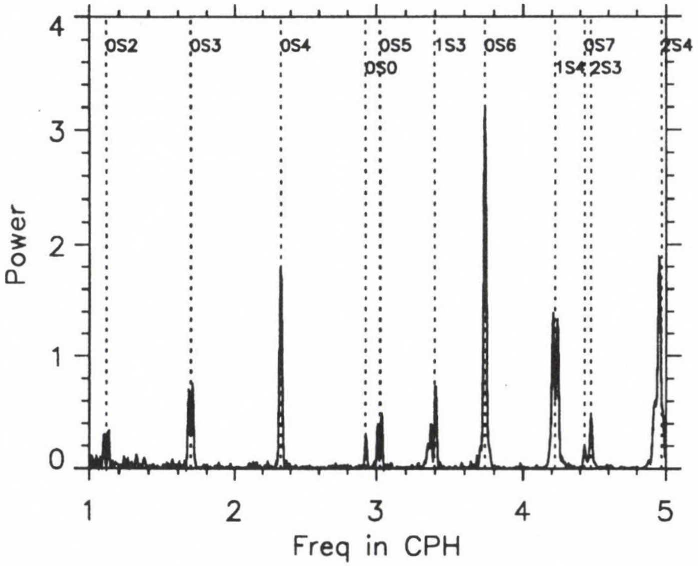


Fig. 1. The power spectrum of the long period part (Lomb-Scargle periogram) of the record during 5–6 Oct. 1994. Length of the record 48 hours

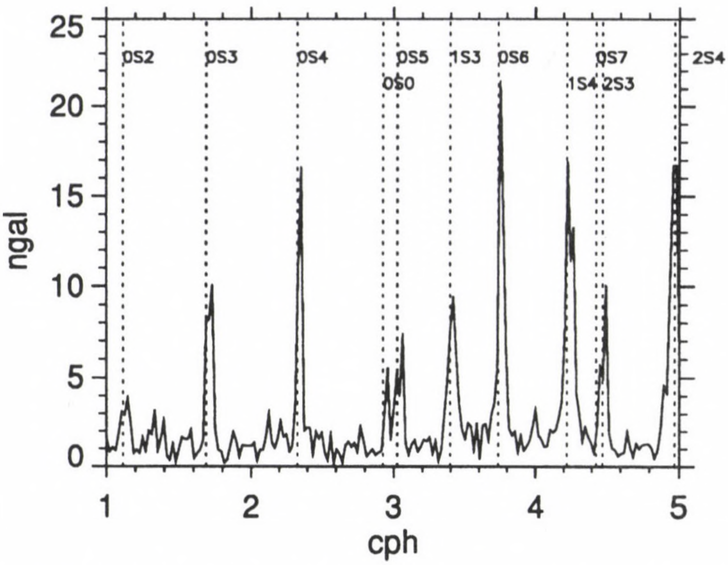


Fig. 2. The amplitude spectrum of the longer period part (FFT) of the record during 5–6 Oct. 1994. Length of the record 48 hours

used is the Lomb-Scargle periogram (Press and Teukolsky 1988), no filtering was applied but short term linear drift was removed before the analyses. Several modes rise clearly above the noise level and can be identified (Anderson and Hart 1978, Denis 1993, Teisseyre 1989) with fine structure in some lines. Figure 2 presents results from the same data when Fast Fourier Transform (FFT) method and highpass filtering were applied before computing. The order of the applied filter was 144 and the lower frequency limit one hour. The highpass filtering improves the quality of the spectra remarkably in many cases. It removes residual of tides and drift effect and gives better signal to noise ratio. The amplitude of the radial mode ${}_0S_0$ was observed to be 5 nanoGals.

The Lomb-Scargle (LS) methods gives better spectral resolution and adopts unevenly spaced data, too. The advantage of FFT methods is the allowance of real amplitudes while the LS method gives a better resolution for frequencies and thus is more adequate to find different modes. In the next presentations we have applied, however, the FFT method and used the previously mentioned highpass filtering before calculations.

Figure 3 presents the whole spectrum up to 20 cycles per hour (cph). The time serie used is the same as presented in Fig. 2. The amplitudes of higher frequencies are remarkably larger than those of lower frequencies. Some eigenperiods of numerous identified modes are labelled e.g. the toroidal mode couplings of ${}_0S_{11}$ and ${}_0T_{12}$. In Fig. 4 we have shifted the starting time of the calculated time series with 12, 24, 36 and 48 hours respectively in regard to the starting time given in Fig. 3. The length of the record is the same, 48 hours. It can be seen that in the shorter

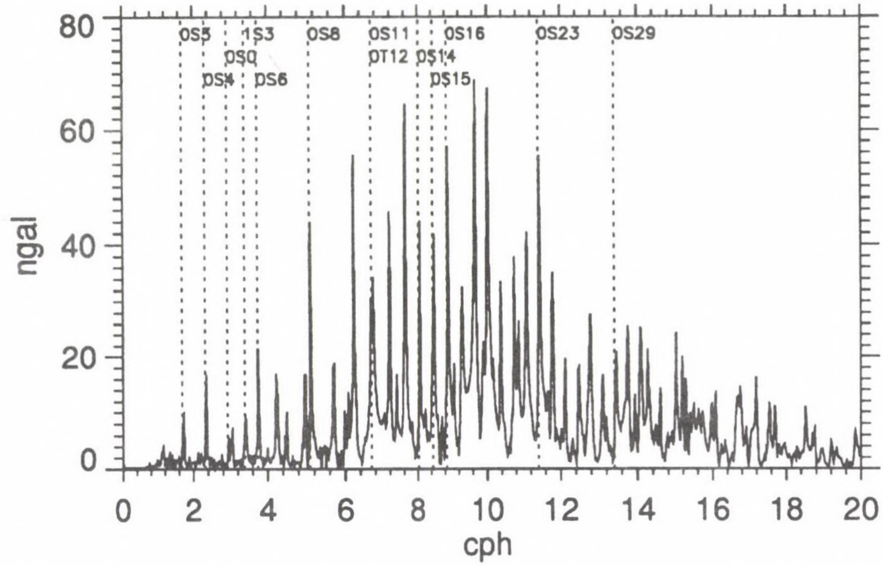


Fig. 3. The amplitude spectrum during 5–6 Oct. 1994. Length of the record 48 hours

part of the spectra originally strong peaks decay in about two days, these spheroidal modes thus having a low Q-value. As an example the amplitude attenuation of the ${}_0S_{15}$ (8.477 cph) peak after the earthquake is depicted in Fig. 5. This mode, also presented with a dashed line in Fig. 4, disappears in about 60 hours.

Figure 6 depicts the ${}_0S_0$ mode starting 6 days after the earthquake. The length of the time series used is 7200 samples or 120 hours. The amplitude of the mode is 5 nanoGal and mean noise level is about 0.5 nanoGal. All spheroidal modes except ${}_0S_0$ have faded out in two or three days. The radial "breathing" mode ${}_0S_0$ has the highest Q-factor (≈ 7000) of all free oscillation modes, the lifetime being about three months. Owing to repeating earthquakes, it theoretically should be present almost all the time but not be discernible due to the background noise. For comparison, Fig. 7 shows an 120 hours spectrum for a quiet term from April 21th to April 25th 1996. No big events have occurred since two months and the background noise only is visible showing the observational threshold value.

Summary of observations of ${}_0S_0$ mode

Results of this Kuril Islands' earthquake has proved that the Finnish superconducting gravimeter at Metsähovi station is capable to observe the free oscillations. In order to know to which extent we have recordings suitable for study of the free oscillation, we studied systematically the whole recording period of the GWR T020 from August 11th 1994 to September 5th 1996, i.e., 740 days. We used tide free, air-pressure corrected and highpass filtered one minute data as presented in the previous chapter. The length of time series was 48 hours and the starting time was shifted

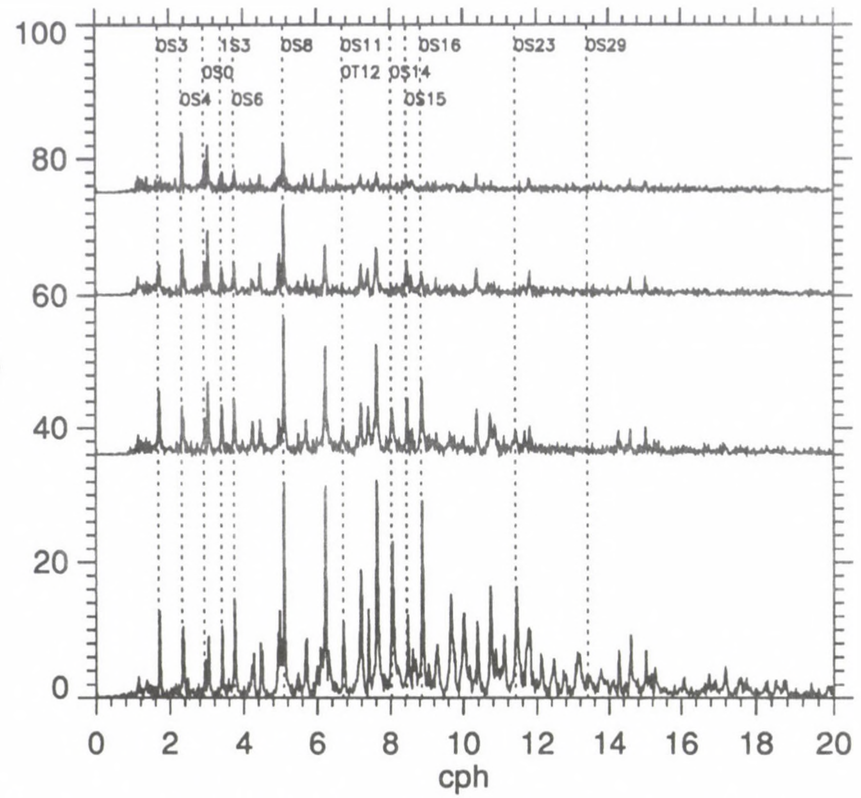


Fig. 4. Attenuation of spheroidal modes. Length of the record 48 hours. The starting time of the time series have been shifted with 12 (bottom), 24, 36 and 48 (top) hours in regard to the starting time in Fig. 3

by 12 hours between the computations. As a result, in addition to digital data we got about 1500 hardcopies of a spectrum, which were compared to catalogues of significant earthquakes (NEIC 1994–1996). During this period 35 earthquakes have occurred causing an observable spheroidal mode of free oscillations phenomena. The smallest magnitude (Mw) was 6.4. Clear excitation of radial mode ${}_0S_0$ happened in 7 cases as presented in Table I. These shallow earthquakes were, however, not the strongest of all observed events. Evidently both the triggering mechanism and the magnitude rule the excitation of ${}_0S_0$ modes.

Figure 8 gives the two dimensional summary for amplitudes of spectra for the years 1994–1996. The horizontal axis is the day of year and the vertical axis shows frequencies from 2cph to 4cph. The ${}_0S_0$ mode and some other spheroidal modes are labelled. The black color indicates amplitude exceeding one nanoGal value. Vertical grey and black lines are due to saturated spectra because of earthquakes, refilling, desensizitation of the gravimeter etc. There is also a data gap in the year 1995 from July 4 to July 15 (days 186–196).

The ${}_0S_0$ mode can be seen in the middle as a horizontal black line. In Table

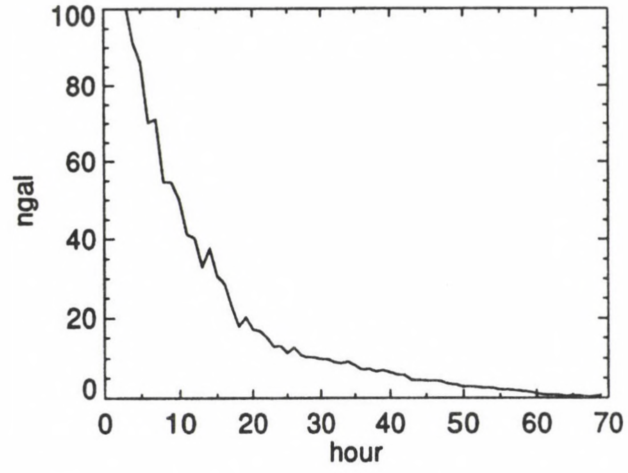


Fig. 5. The amplitude attenuation of ${}_0S_{15}$ (8.477 cph) after the earthquake on Oct. 4th, in Kuril Islands

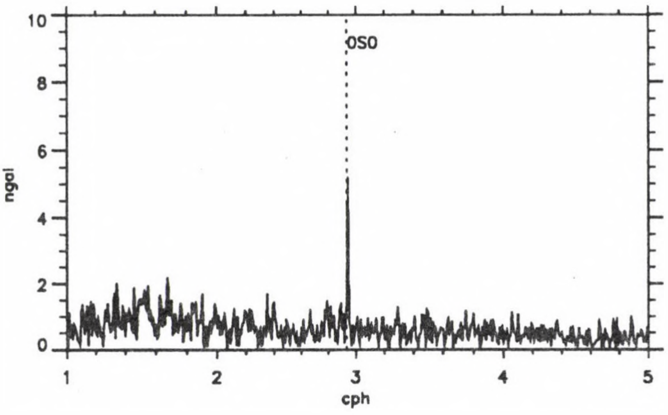


Fig. 6. The observed radial mode ${}_0S_0$ during 10–14 Oct. 1994. Length of record is 120 hours

I the observed duration of ${}_0S_0$ is given. We can say that the ${}_0S_0$ mode has been observed at the one nanoGal level at least in 25 % of the total recording time at Metsähovi.

Conclusion

The superconducting gravimeter GWR T020 at Metsähovi seems to be very suitable for free oscillation studies. The quality of spectra is good and results are comparable with other superconducting gravimeter installations (Zürn et al. 1991, Banka and Crossley 1995). The sampling rate of one minute is found to be suitable for this spectral region. The selection of length for the time series is important: in

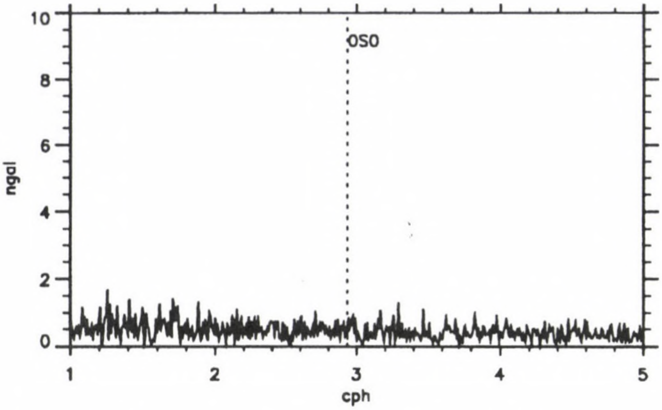


Fig. 7. The observed spectrum for a quiet term 19–24 Apr. 1996

Table I. Earthquakes excitated ${}_0S_0$ mode observed during 1994–1996

Year+Day	Date	Time UT UT	Location	Mag .Mw	Duration days
94277	Oct 4th	13:23	Kuril Island	8.2	46
95136	May 16th	20:13	Loyalty Island region	7.3	12
95211	Jul 30th	05:11	Near coast of northern Chile	7.5	43
95282	Oct 9th	15:35	Near Coast of Jalisco	7.6	7
95337	Dec 3th	18:01	Kuril Island	7.8	19
96048	Feb 17th	05:59	Irian Jaya region, Indonesia	7.9	34
96162	Jun 10th	04:03	Andreanof Island, Aleutian	7.9	13

case of too short time series, the quality of spectra will be poor and longer series also give a better signal to noise ratio. The lengths of the series are limited for practical reasons because the spectra of successive earthquakes will disturb each other. In the studies for the decay of modes the length of the spectrum should be rather short. In some cases at beginning of an earthquake the signal range of the gravimeter is fully saturated and we have to cut out this data. However, using an appropriate filter it is possible to study the spectrum at the very beginning of an earthquake, which is important in determining the low Q-values.

Results show that we have on average one free oscillation event per month concerning the short term spheroidal modes. Relative amplitudes of peaks seem to be different in several cases. Calculation of relative amplitudes, Q-factors for different modes and comparison to earthquake catalogues will be very valuable. In addition, 2–3 events per year are expected, which can excite the radial mode ${}_0S_0$. Comparison of the results with those obtained with a long periodic seismometer will be of interest in the future.

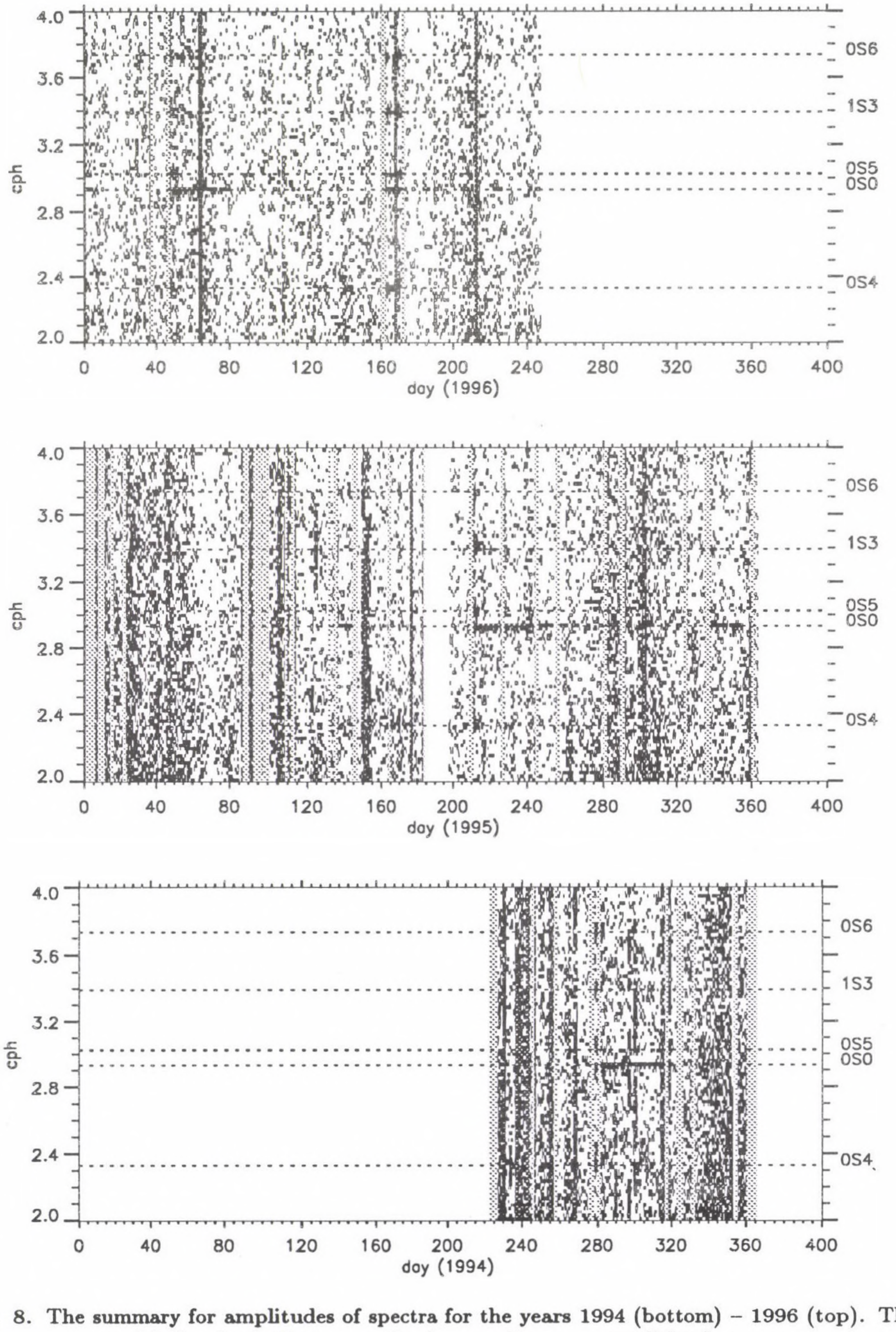


Fig. 8. The summary for amplitudes of spectra for the years 1994 (bottom) – 1996 (top). The black color indicates amplitude exceeding one nanoGal value

References

Anderson D, Hart R 1978: *J. Geoph. Res.*, 83, 5869.

Banka D, Crossley D 1995: *Bulletin d'information Marées Terrestres*, 122, 9148.

Benioff H, Gutenberg B, Richter C F 1954: *Trans. Am. Geophys. Un.*, 35, 979–987.

Denis C 1993: *Acta Geod. Geoph. Mont. Hung.*, 28, 15–131.

NEIC, National Earthquake Information Center, Golden, Colorado, USA, Significant earthquakes of the world 1994–1996. Internet: http://www.neic.cr.usgs.gov/neis/eqlists/

Press W, Teukolsky S 1988: Search algorithm for weak periodic signals in unevenly spaced data, Computers in Physics Nov./Dec.

Teisseyre R ed. 1989: In: Gravity and Low-Frequency Geodynamics. Physics and Evolution of the Earth's Interior 4, PWN-Polish Scientific Publishers, Warsaw, 362–415.

Virtanen H, Kääriäinen J 1995a: The installation of and first results from the superconducting gravimeter GWR20 at the Metsähovi station. Rep. Finn. Geod. Inst. 95, 1, Helsinki

Virtanen H, Kääriäinen J 1995b: Non-gravity effects on the superconducting gravimeter at the Metsähovi station. Poster, XXI General Assembly of the IUGG, Boulder

Wenzel H-G 1995: Earth tide data processing package ETERNA 3.20, Black Forest Observatory, Schiltach, Karlsruhe

Zürn W, Wenzel H-G, Laske G 1991: *Bulletin d'information Marées Terrestres*, 110, 7940.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 433–438 (1996)

THE INDIRECT TIDAL TORQUE AND PRECESSION-NUTATION DYNAMICS

M BURŠA[1]

[Manuscript received July 25, 1996]

The direct and indirect tidal torques exerted on the planet M by the tide forming body m are discussed from the point of view of their mutual direction. It is proved that the tidal force function enables the exact determination of both the direct and indirect tidal torques without density distribution within M. The Stokes parameters of M determining the external gravitational field of M are sufficient for the solution. The gravitational field of m is simplified to be spherically symmetrical for brevity only; the exact general solution is outlined.

Keywords: direct tidal torque; indirect tidal torque; nutation; precession

1. Introduction

There are two different torques exerted on a planet assumed to be not perfectly rigid (body and mass M, its center O) by an external body (body and mass m, its center O'): the direct one (N) giving rise to the planetary precession-nutation dynamics, and the indirect or additional one (δN) due to the transfer of masses within M. Torque N is identical with the precession torque (Melchior and Georis 1968) and it is nearly perpendicular to the planetary meridian plane $\Lambda = \Lambda_{O'}$ passing through center of mass O'. Torque δN is situated nearly in the meridional plane above and it does not contribute to the precession-nutation dynamics practically. It is just the aim of the paper to prove and discuss the fact above in detail.

2. Direct tidal torque

The direct tidal torque is defined as follows:

$$N = \int_{\tau} [\varrho \times \operatorname{grad} V_t]\, \sigma d\tau ; \tag{1}$$

τ stands for the volume of M, ϱ is the planetocentric radius-vector of volume element $d\tau$, σ the density at $d\tau$, V_t the tidal potential due to body m

$$V_t = \frac{Gm}{\Delta_{OO'}} \sum_{n=2}^{\infty} \left(\frac{\varrho}{\Delta_{OO'}}\right)^n \sum_{k=0}^{n} \frac{(2-\delta_{k0})(n-k)!}{(n+k)!} \cdot$$

$$\cdot\, P_n^{(k)}(\sin\phi) P_n^{(k)}(\sin\delta_{O'}) \cos k(\Lambda_{O'} - \Lambda) . \tag{2}$$

[1]Astronomical Institute, Acad. Sci. Czech Republic, Bočni II/1401, 141 31 Praha 4, Czech Republic, e-mail: astdss@csearn.bitnet

Notations: $\Delta_{OO'} = \overline{OO'}$ is the distance between mass centers O and O', $\delta_{O'}$ is the planetocentric declination of O', $P_n^{(k)}$ is the Legendre associated function of degree n and order k.

Because of the density distribution required, the exact determination of N by Eq. (1) could seem to be problematic. However, integral (1) can be calculated without any information about function σ (Burša 1982). Its exact solution can be reached by the Stokes parameters of the planet determining its external gravitational field. To avoid a triple integration of (1) in components, we introduced the so-called tidal force function V_{tff} defined as

$$V_{\mathrm{tff}} = \int_M V_t dM \,. \tag{3}$$

On the basis of Eq. (3), components N_j $(j = 1, 2, 3)$ of Eq. (1) can be expressed in the planetocentric coordinate system x_j, the axes of which are the axes of the planet's ellipsoid of inertia, as follows:

$$\begin{aligned} N_1 &= \left[\frac{\partial V_{\mathrm{tff}}}{\partial \psi} - \cos\vartheta \frac{\partial V_{\mathrm{tff}}}{\partial \varphi}\right] \frac{\sin\varphi}{\sin\vartheta} + \cos\varphi \frac{\partial V_{\mathrm{tff}}}{\partial \vartheta} \,, \\ N_2 &= \left[\frac{\partial V_{\mathrm{tff}}}{\partial \psi} - \cos\vartheta \frac{\partial V_{\mathrm{tff}}}{\partial \varphi}\right] \frac{\cos\varphi}{\sin\vartheta} - \sin\varphi \frac{\partial V_{\mathrm{tff}}}{\partial \vartheta} \,, \\ N_3 &= \frac{\partial V_{\mathrm{tff}}}{\partial \varphi} \,. \end{aligned} \tag{4}$$

Formula (4) is well known from the precession-nutation theory; ψ, ϑ, φ are Euler's angles (precession, nutation, proper rotation, respectively) bounding x_j-axes fixed with the planet and axes X_j of a reference system inertial in the Newtonian sense. All Eulerian angles are positive if rotation is counterclockwise.

Tidal force function (2) in a developed form reads

$$V_{\mathrm{tff}} = \frac{GMm}{\Delta_{OO'}} \sum_{n=2}^{\infty} \sum_{k=0}^{n} \left(\frac{a_o}{\Delta_{OO'}}\right)^n \left(J_n^{(k)} \cos k\Lambda_{O'} + \right.$$
$$\left. + S_n^{(k)} \sin k\Lambda_{O'}\right) P_n^{(k)}(\sin\delta_{O'}) + \Delta V_{\mathrm{tff}} \,. \tag{5}$$

Notations: a_o is scaling arbitrary length parameter, as usual, the equatorial radius of the planet associated with the Stokes parameters

$$\begin{matrix} J_n^{(k)} \\ S_n^{(k)} \end{matrix} = \frac{(2 - \delta_{kO})(n-k)!}{M a_o^n (n+k)!} \int_M \varrho^n P_n^{(k)}(\sin\phi) \begin{matrix} \cos k\Lambda \\ \sin k\Lambda \end{matrix} dM \tag{6}$$

rendering them to be dimensionless; ϕ stands for the planetocentric latitude of dM, Λ for its planetary longitude. The last term ΔV_{tff} in Eq. (5) represents the contribution to V_{tff} due to the deviations in the gravitational field of the tide forming body m from the ideal spherical symmetry.

After inserting (5) into Eq. (4), and making use of the derivatives of the Legendre associated functions with respect to the Eulerian angles (Burša 1974), components N_j of the direct tidal torque can be expressed in the final form as

$$\begin{aligned}
N_1 &= -G\tfrac{Mm}{\Delta_{OO'}}\left(\tfrac{a_e}{\Delta_{OO'}}\right)^2\Big[\tfrac{3}{2}J_2^{(o)}\sin 2\delta_{O'}\sin\Lambda_{O'}+ \\
&\quad+\tfrac{3}{2}J_2^{(1)}\cos^2\delta_{O'}\sin 2\Lambda_{O'}-3S_2^{(1)}(\sin^2\delta_{O'}-\cos^2\delta_{O'}\sin^2\Lambda_{O'})+ \\
&\quad+3J_2^{(2)}\sin 2\delta_{O'}\sin\Lambda_{O'}-3S_2^{(2)}\sin 2\delta_{O'}\cos\Lambda_{O'}\Big]\,, \\
N_2 &= G\tfrac{Mm}{\Delta_{OO'}}\left(\tfrac{a_e}{\Delta_{OO'}}\right)^2\Big[\tfrac{3}{2}J_2^{(o)}\sin 2\delta_{O'}\cos\Lambda_{O'}- \\
&\quad-3J_2^{(1)}(\sin^2\delta_{O'}-\cos^2\delta_{O'}\cos^2\Lambda_{O'})+\tfrac{3}{2}S_2^{(1)}\cos^2\delta_{O'}\sin 2\Lambda_{O'}- \\
&\quad-3J_2^{(2)}\sin 2\delta_{O'}\cos\Lambda_{O'}-3S_2^{(2)}\sin 2\delta_{O'}\sin\Lambda_{O'}\Big]\,, \\
N_3 &= G\tfrac{Mm}{\Delta_{OO'}}\left(\tfrac{a_e}{\Delta_{OO'}}\right)^2\Big[\tfrac{3}{2}J_2^{(1)}\sin 2\delta_{O'}\sin\Lambda_{O'}- \\
&\quad-\tfrac{3}{2}S_2^{(1)}\sin 2\delta_{O'}\cos\Lambda_{O'}+6J_2^{(2)}\cos^2\delta_{O'}\sin 2\Lambda_{O'}- \\
&\quad-6S_2^{(2)}\cos^2\delta_{O'}\cos 2\Lambda_{O'}\Big]\,.
\end{aligned} \tag{7}$$

We restricted ourselves to the degree $n=2$ tidal terms, however, there is no principal difficulty to derive the higher degree terms, after applying the derivatives of the corresponding Legendre associated functions with respect to the Eulerian angles referenced above.

Note that if x_j-axes coincide with the axes of the planet's inertia ellipsoid, then $J_2^{(1)}=0$, $S_2^{(1)}=0$ and $S^{(2)}=0$. That is why, only the zonal and sectorial tidal terms are responsible for the precession-nutation dynamics.

In the case of the exact rotational symmetry of the planet's gravitational field, the direction cosines of vector N are

$$\cos(N,x_1)=-\sin\Lambda_{O'},\ \cos(N,x_2)=\cos\Lambda_{O'},\ \cos(N,x_3)=0\,. \tag{8}$$

It means, if the planet's gravitational field were simplified to be rotationally symmetrical, $J_2^{(2)}=0$, the direct tidal torque would be normal to meridian plane $\Lambda=\Lambda_{O'}$. The fact above is well known from the precession-nutation dynamics. However, if $J_2^{(2)}\neq 0$, the direction of torque N derivates from the ideal one as $2J_2^{(2)}/J_2^{(o)}$. If planet Earth, the deviation angles in (N,x_1) and (N,x_2) are of the order 1.5×10^{-3} (300 arcseconds) in magnitude, 3.0×10^{-3} in (N,x_3). The variations above are the same functions of $\Lambda_{O'}$ as torques N_j themselves. The values above are maximum values.

3. Indirect tidal torque

The indirect or additional tidal torque δN exerted on the planet by the tide forming body is due to the transfer of planetary masses. In that case the Stokes parameters (6) are not constant, distortions $\delta J_n^{(k)}$ and $\delta S_n^{(k)}$ occur. They are functions of the tidal parameters and they can be derived without density distribution within M. However, the Love numbers $k_{n,k}$ and phase lag angles $\varepsilon_{n,k}$ for each harmonic term of degree n and order k should be known.

At an arbitrary point $(\overline{\varrho}, \overline{\phi}, \overline{\Lambda})$ situated at the boundary surface of the planet it should hold that

$$\frac{Gm}{\Delta_{OO'}} \sum_{n=2}^{\infty} \left(\frac{\overline{\varrho}}{\Delta_{OO'}}\right)^n k_{n,k} \sum_{k=0}^{n} \frac{(2-\delta_{k0})(n-k)!}{(n+k)!} \cdot$$

$$\cdot P_n^{(k)}(\sin\overline{\phi}) P_n^{(k)}(\sin\delta_{O'}) \cos k(\Lambda_{O'} - \varepsilon_{n,k} - \overline{\Lambda}) = \tag{9}$$

$$= \frac{GM}{\overline{\varrho}} \sum_{n=2}^{\infty} \sum_{k=0}^{n} \left(\frac{a_o}{\overline{\varrho}}\right)^n (\delta J_n^{(k)} \cos k\overline{\Lambda} + \delta S_n^{(k)} \sin k\overline{\Lambda}) P_n^{(k)}(\sin\overline{\phi}) .$$

From Eq. (9) the relations for the tidal distortions in the Stokes parameters follow, exactly, after solving the first (Dirichlet's) boundary-value problem, as

$$\begin{matrix} \delta J_n^{(k)} \\ \delta S_n^{(k)} \end{matrix} = \frac{(2-\delta_{kO})(n-k)!}{(n+k)!} k_{n,k} \frac{Gm}{GM} \left(\frac{\overline{\varrho}}{\Delta_{OO'}}\right)^3 \cdot$$

$$\cdot \left(\frac{\overline{\varrho}}{a_o}\right)^2 P_n^{(k)}(\sin\delta_{O'}) \begin{matrix} \cos k(\Lambda_{O'} - \varepsilon_{n,k}) \\ \sin k(\Lambda_{O'} - \varepsilon_{n,k}) . \end{matrix} \tag{10}$$

The corresponding contribution to the tidal force function due to the anelastic planet's response can be expressed in the form analogous to Eq. (5) as

$$\delta V_{\text{tff}} = \frac{GMm}{\Delta_{OO'}} \sum_{n=2}^{\infty} \sum_{k=0}^{n} \left(\frac{a_o}{\Delta_{OO'}}\right)^n \cdot$$

$$\cdot (\delta J_n^{(k)} \cos k\Lambda_{O'} + \delta S_n^{(k)} \sin k\Lambda_{O'}) P_n^{(k)}(\sin\delta_{O'}) . \tag{11}$$

After inserting Eq. (10) into Eq. (11), and again restricting ourselves for brevity to the tidal terms of degree $n = 2$, and putting $k_{2,0} = k_{2,1} = k_{2,2} = k$ (not to be confused with symbol k for the order of harmonics), $\varepsilon_{2,0} = \varepsilon_{2,1} = \varepsilon_{2,2} = \varepsilon$, the final form reads

$$\delta V_{\text{tff}} = k \frac{Gm^2}{\Delta_{OO'}} \left(\frac{R}{\Delta_{OO'}}\right)^5 \left\{ \left[P_2^{(0)}(\sin\delta_{O'})\right]^2 + \right.$$

$$\left. + \frac{1}{3} \left[P_2^{(1)}(\sin\delta_{O'})\right]^2 \cos\varepsilon + \frac{1}{12} \left[P_2^{(2)}(\sin\delta_{O'})\right]^2 \cos 2\varepsilon \right\} , \tag{12}$$

R stands for the planet's mean radius.

The components δN_j of the indirect tidal torque exerted on the planet by the tide forming body because of planet's anelastic response can be expressed from

Eq. (4), after inserting $V_{\rm tff} = \delta V_{\rm tff}$:

$$
\begin{aligned}
\begin{matrix}\delta N_1\\ \delta N_2\end{matrix} &= \frac{3}{2}k\sin\varepsilon\frac{Gm^2}{\Delta_{OO'}}\left(\frac{R}{\Delta_{OO'}}\right)^5\sin 2\delta_{O'}\begin{matrix}\cos\Lambda_{O'}\\ \sin\Lambda_{O'}\end{matrix}\,,\\
\delta N_3 &= -3k\sin\varepsilon\frac{Gm^2}{\Delta_{OO'}}\left(\frac{R}{\Delta_{OO'}}\right)^5\cos^2\delta_{O'}\,.
\end{aligned}
\qquad (13)
$$

Components δN_1 and δN_2 are strongly time dependent, $\Lambda_{O'}$ being equal to the negative value of the planetocentric hour angle $T_{O'}$ of $O'(\Lambda_{O'} = -T_{O'})$. However, component δN_3 is only weakly time dependent, because not containing $T_{O'}$, and the time variations in $\Delta_{OO'}$ and in $\cos^2\delta_{O'}$ are relatively small, as usual. Evidently, the indirect tidal torque δN is situated exactly in the meridian plane $\Lambda = \Lambda_{O'}$, and it is directed normal to $\overline{OO'}$ (Burša 1996). It means, the angle $(N, \delta N)$ formed by the direct and indirect tidal torques is nearly 90°. They are perpendicular exactly if the planet's gravitational field is rotationally symmetrical. That is why, δN does not influence the planet's precession-nutation dynamics practically. Neglecting ε^2,

$$\frac{\partial V_{\rm tff}}{\partial \Delta_{OO'}} = 0\,,$$

because

$$\left[P_2^{(0)}(\sin\delta_{O'})\right]^2 + \frac{1}{3}\left[P_2^{(1)}(\sin\delta_{O'})\right]^2 + \frac{1}{12}\left[P_2^{(2)}(\sin\delta_{O'})\right]^2 = F(\delta_{O'}) = 0\,.$$

Moreover,

$$\frac{\partial V_{\rm tff}}{\partial \delta_{O'}} = 0\,,$$

because

$$\frac{\partial F(\delta_{O'})}{\partial \delta_{O'}} = 0\,.$$

The contributions of the terms not depending on ε^1 to the indirect tidal torque is zero. This concerns also the long-term and/or permanent zonal tidal term. However, the last affects the coefficient in the precession constant $H = -J_2^{(0)}Ma_o^2/C$, as well as, principal moment of inertia C.

4. Conclusion

1. The direct tidal torque is fully responsible for the precession dynamics.
2. The indirect tidal torque due to the planet's anelastic response does not affect the precession torque practically. It is nearly perpendicular to the direct tidal and/or precessional torque.
3. The long-term tidal variation in the second zonal Stokes parameter affects coefficient H in the precession constant and only in this way the tidal transfer of mass within M due to m is reflected in the precession dynamics.
4. The derivatives of the tidal force function reflecting the planet's anelastic response, with respect to the planetocentric distance and declination of the tide forming body mass center, equal zero.

References

Burša M 1974: *Studia geoph. et geod.*, 18, 103–119.
Burša M 1982: *Studia geoph. et geod.*, 26, 319–322.
Burša M 1996: *Studia geoph. et geod.*, 40, 111–118.
Melchior P, Georis B 1968: *Phys. Earth Planet. Int.*, 1, 267–287.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 439–444 (1996)

ON THE EFFECT OF GREENHOUSE GASES IN THE UPPER ATMOSPHERE

P BENCZE[1] and A POOR[1]

[Manuscript received June 10, 1996]

The cooling of the upper atmosphere due to the greenhouse effect is investigated by means of ionospheric data at mid-latitudes. Comparing the results of this study to a former investigation by Bremer (1992) it has been found that on the one hand the effect is not limited to the summer months, on the other hand — in contrast with Bremer's data — it indicates a diurnal variation showing an effect in day-time. The reason of this difference is discussed.

Keywords: cooling of the upper atmosphere; greenhouse effect; ionospheric data; upper atmosphere

Introduction

As a result of the increased attention paid to the global change and the greenhouse effect by the scientific community, it has recently been suggested that the greenhouse effect might affect the upper atmosphere, too. The influence of the greenhouse effect in the upper atmosphere would be validated by the circumstance that the radiatively active greenhouse gases and especially the most effective of them CO_2 reradiate the absorbed long wave length terrestrial thermal radiation not only towards the Earth, but also in all directions; that is also upwards. As it is known, the rate of the temperature decrease is proportional to the variation of the flux of the radiation with height and inversely proportional to atmospheric density (Liou 1980). Thus, in case of the reradiation directed downwards it causes a temperature increase, since the density increases and the flux of the radiation is reduced.

However, in case of the reradiation directed upward, the density decreases and the flux of the radiation increases. On the basis of phase height measurements it has experimentally been shown by Taubenheim et al. (1989) that the pressure decreased at a height of 80 km from 1961 to 1987 by about 10 %. Almost at the same time model calculations of the possible effect of greenhouse gases in the upper atmosphere were carried out. The model calculations of Roble and Dickinson (1989) indicated a cooling in the thermosphere of about 50K, if the concentration of trace gases (CO_2, CH_4) is doubled. Using the basic theory of the ionosphere, Rishbeth (1990) investigated the effect of this cooling in the ionosphere. He has shown that the cooling and the associated composition changes would decrease the height of the maximum electron density of the E region by about 2 km and that of the peak elektron density of the F region by about 20 km during the 21st century.

[1]Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences, H-9401 Sopron, POB 5, Hungary

The first experimental investigations referring to changes in the ionosphere due to the greenhouse effect showed a decrease of the height of the F region maximum electron density ($hmF2$) of about 0.24 km/year for the period from 1957 to 1990 (Bremer 1992). Further studies found long term variations of the electron density (of the order of $+0.17$ MHz/year for ΔfoE and -0.14 MHz/year for $\Delta foF2$) in the mid-latitude ionosphere based on 32 year's data (Givishvili et al. 1995). However, these changes were attributed to a long term decrease of atomic and molecular oxygen density in the mid-latitude thermosphere. It is to be noted that Bremer (1992) obtained for the same quantities much smaller values (-0.0003 MHz/year for ΔfoE and -0.003 MHz/year for $\Delta foF2$) referring to the same latitude ($\sim 55°$N) and determined on the basis of 33 year's data.

1. Data and method

The investigations were carried out by means of data measured at the Japanese ionospheric station Wakkanai (45°24′N, 141°41′E) in the period from 1958 to 1987. At the selection of the ionospheric station, it is also necessary to take into account that not all ionospheric station publish the values of $M(3000)$ $F2$ needed for the calculation of $hmF2$. The data of an Asian ionospheric station were used to see how the greenhouse effect is validated in the upper atmosphere in this part of the northern hemisphere. For the sake of comparability, the method used by Bremer has been applied (Solé 1995); that is for the determination of the height of the maximum electron density of the $F2$ layer $hmF2$ the expression

$$hmF2 = \frac{1490}{M(3000)F2 + \Delta M} - 176\,,$$

where

$$M = \frac{F1 \cdot F4}{foF2/foE - F2} + F3\,.$$

In this formula $F1$, $F2$, $F3$ and $F4$ are coefficients depending on solar activity given by the relative sunspot number R and on geomagnetic latitude (Bilitza et al. 1979).

Furthermore, the effect of the solar and geomagnetic activities had to be removed from the ionospheric parameters. Thus, first the coefficients determining the relation between the ionospheric parameters (f) and the sunspot number R, as well as the geomagnetic activity index Ap are computed according to the equation

$$f = A + BR + CAp\,.$$

This procedure was not used in case of $foF2$ and foE, since on the one hand the correction term ΔM is of the order of 0.1 km. On the other hand in the calculation of ΔM the ratio of these two parameters appears, thus partly eliminating the effect of both the solar and the geomagnetic activities. The changes of the ionospheric parameters can be calculated using the formula

$$\Delta f = f_m - f\,,$$

where by f_m the measured value of the ionospheric parameter is denoted. Then, the long term trend of the ionospheric parameter can be computed by means of the linear regression equation

$$\Delta f = a + b(n_e - n_s) ,$$

where a and b are the regression coefficients, n_e and n_s are the last and the beginning year of the period used for the study.

The significance of the results obtained with the ionospheric parameters was also determined by the Fisher's F parameter

$$F = \frac{r^2}{1 - r^2}(N - 2) .$$

Here r is the correlation coefficient between Δf and the corresponding interval $(n_e - n_s)$, N is the number of years in the period, to which the investigation was extended.

2. Results and discussion

The results obtained by the method described above using the ionospheric parameters of the ionospheric station Wakkanai from the period 1958–1987 are plotted in Fig. 1. The long term changes are studied only in case of the ionospheric parameter $hmF2$, since the effect of cooling in the upper atmosphere can most easily be detected in the altitude of the ionospheric layers. It can be seen that the quantity Δf ($\Delta hmF2$) showing the height change of the maximum electron density in the F region — attributable to the greenhouse effect — indicates a long term decrease of $hmF2$. The daily variation of Δf shows a day-time minimum (negative

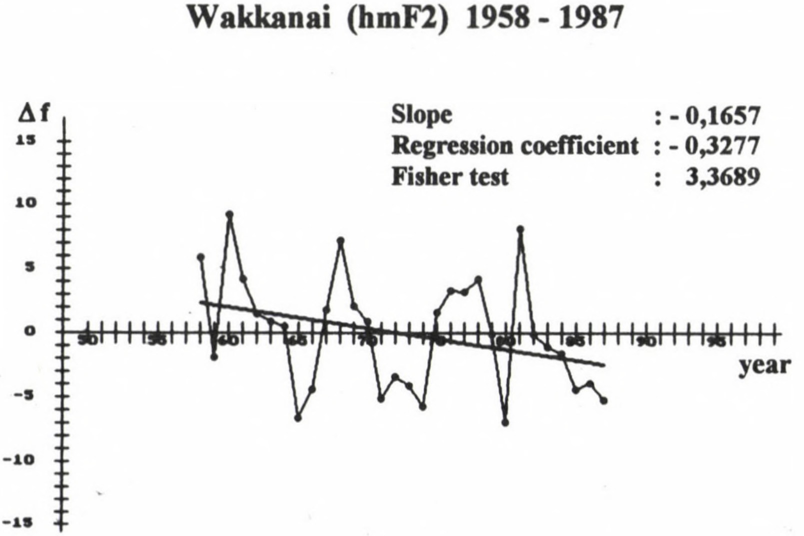
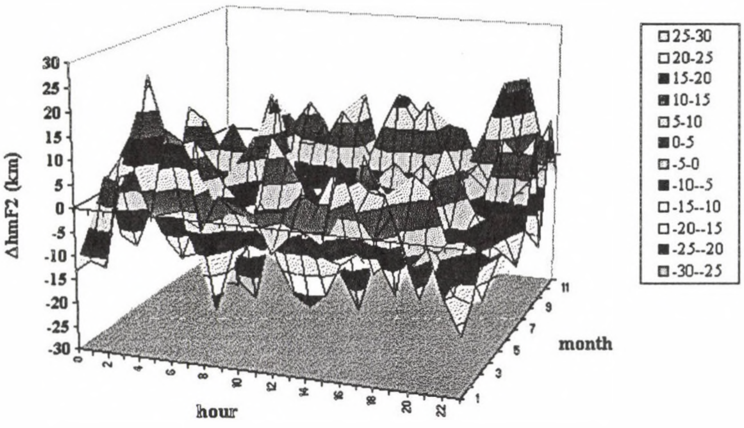


Fig. 1.

Wakkanai Δf 1987



Wakkanai Δf 1958

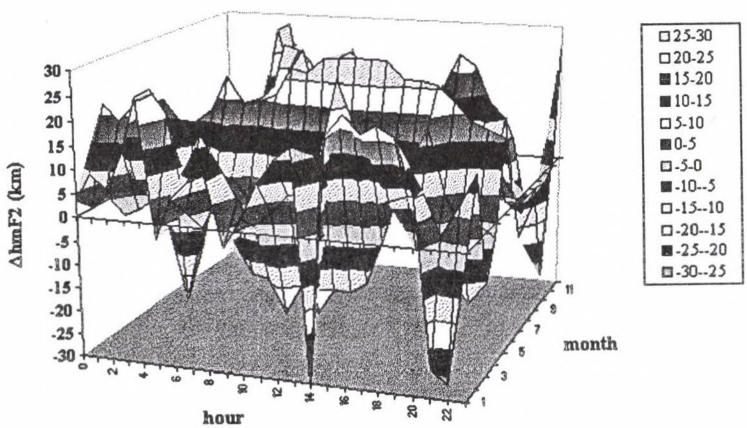


Fig. 2.

value meaning contraction), which means a maximum cooling during the day (Fig. 2). Considering the seasonal variation of Δf, no clear seasonal variation has been found, however, b — the regression coefficient — indicates negative values during almost the whole year (upper part of Fig. 3). In the lower part of Fig. 3 the results of the significance of the data obtained by the Fisher test (F) are also shown. It can be seen that the negative trend is significant with a confidence level of almost 95 % in February, with the confidence level of 95 % in July and with a confidence level of almost 99 % in October.

Considering the diurnal variation of $\Delta hmF2$, the daily variation of the trend of the decrease obtained in the height $hmF2$ of the $F2$ layer is opposite to the

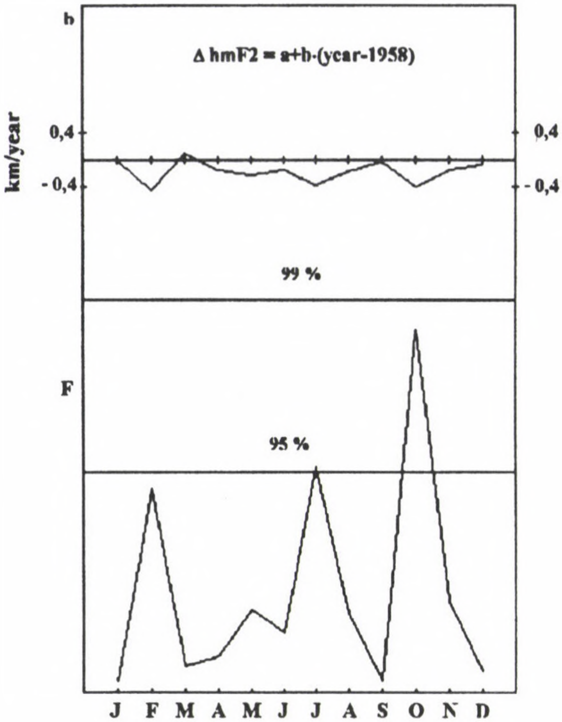


Fig. 3.

diurnal variation of the CO_2 mixing ratio. As it is known, the diurnal variation of the CO_2 mixing ratio shows maximum values in the course of the night. Thus, the maximum cooling would occur during the night when the CO_2 mixing ratio indicates maximum values, at least in the vicinity of the surface.

It is not quite clear whether this contradiction is due to the insufficient length of the data series, or it might be attributed to the vertical propagation of the daily variation (the fraction of CO_2 does essentially not change with height in the homosphere, disregarding a slight photodissociation in the lower thermosphere). In case of the daily variation, the phase shift might be explained by vertical propagation. The lifetime (residence time) of CO_2 in the atmosphere is of the order of 15 years, which is much greater than the time constants of either the vertical component of the turbulent diffusion ($\sim$ 1 day) or the time constant of the transport by the vertical wind ($<$ 1 month) in the lower thermosphere. This means that the CO_2 mixing ratio depends first of all on the dynamical conditions. However, on the one hand in the thermosphere the separation of atmospheric gases depending on their molecular weight begins in the gravity field of the Earth. Consequently, the concentration of CO_2 decreases faster with height, than O_2 or N_2 because of its greater molecular weight, approximating the decrease of Argon concentration with height. On the other hand, the meridional wind is directed in the F region from the summer hemisphere to the winter hemisphere. Thus, CO_2 can be transported in northern summer from the northern hemisphere to the southern hemisphere. In northern winter CO_2 is transported from the southern to the northern hemisphere, but it may occur that this transport is less effective because of the smaller production of CO_2 in the southern hemisphere. It can be assumed that the exchange of CO_2 between the hemispheres results in the inrease of the CO_2 mixing ratio in the northern hemisphere in winter as compared to summer. In consequence of this process, the cooling in the upper atmosphere would be greater in winter than in summer. It is to be noted that Bremer (1992) obtained a seasonal variation of $\Delta hmF2$, which is in agreement with the latter seasonal variation, and also with the seasonal variation of the CO_2 mixing ratio in the vicinity of the ground; that is, it shows maximum cooling in winter, when the CO_2 mixing ratio is maximum.

Acknowledgements

The authors should like to thank for the supply of data to Mr. German Solé, Observatori de l'Ebre, Spain.

These investigations had been supported by the Hungarian Space Office in the framework of a grant.

References

Bilitza D, Sheikh N M, Eyfrig R 1979: *Telecom. J.*, 46. 549.
Bremer J 1992: *J. atmos. terr. Phys.*, 54, 1505–1511.
Givishvili G V, Leshchenko L N, Smeleva O P, Ivanidze T G 1995: *J. atmos. terr. Phys.*, 57, 871–874.
Liou K N 1980: An Introduction to Atmospheric Radiation. Academic Press, New York
Rishbeth H 1990: *Planet. Space Sci.*, 38, 945–948.
Roble R G, Dickinson R E 1989: *Geophys. Res. Lett.*, 16, 1441–1444.
Solé G 1995: Private communication
Taubenheim J, Cossart G V, Entzian G 1989: *Handbook for MAP*, 27, 70–71.

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 445–451 (1996)

EINSTEIN AND GEOPHYSICS

W SCHRÖDER[1] and H-J TREDER[2]

[Manuscript received August 8, 1996]

The article gives an overview on Albert Einstein's activity in relation to geophysics. Various aspects of his fundamental investigations and their significance for geophysical research are discussed.

Keywords: Albert Einstein; Earth's rotation; inertial mass; theory of relativity

1. Introduction

Einstein's activity has transformed the principles of physics in atomistics, in quantum mechanics and in relativity theory and enabled us to describe processes in the cosmos and inside the body of the universe. Einstein's General Relativity Theory initiated a new chapter in astronomy, but also in geophysics and made theoretical astronomy possible: relativistic astrophysics and models of the universe based on Einstein's ideas belong nowadays without doubt to the main fields of cosmic geophysics.

In addition to these conceptual inventions and to his basic work on relativistic astrophysics and on related fields of science — which interested again and again geophysicists such as Emil Wiechert and Hans Ertel among others — Einstein dealt repeatedly with other geophysical problems and published several papers on them (see also Pais 1986)

2. The principles

During his stay in Berlin (1914 to 1932), Einstein had a close scientific and personal connection with eminent Potsdam geodesists and geophysicists, with G Helmert, A Schmidt and Julius Bartels. Among meteorologists he was especially friendly with Heinrich Ficker, with Hans Ertel's professor. That is how he got in a remarkable connection during his studies with Einstein. It is moreover known how Einstein strove to call the Göttingen geophysicist, Emil Wiechert to Potsdam. In addition, Einstein was between 1926 and 1933 a member of the editorial board of Gerlands Beiträge zur Geophysik edited during this period by the Viennese climatologist V Conrad.

[1]Hechelstrasse 8, D–28777 Bremen-Rönnebeck, Germany
[2]Rosa-Luxemburg-Strasse 17a, D-14482 Potsdam, Germany

1217-8977/96/$ 5.00 ©1996 Akadémiai Kiadó, Budapest

3. Connection with geophyisical problems

Newton did not believe in his law of gravity (deduced from Kepler's laws) for many years as the value of the gravitational acceleration seemingly did not coincide with the value of the gravitational acceleration of Moon. The Earth's radius R — as Newton knew it — was to small. The later geodetical determination of the Earth's radius by C Picard led to a correction of g so that the product of the gravity constant f with the mass of the Earth led to the same result fM from the Galilean accelaration g and from the Moon's accelaration v^2/r. Newton was willing to give the manuscript of his famous book, Principia Mathematica Philosophiae Naturalis only after this to the printer's.

Galilei's and Newton's notions on the inertia of material, i.e. the notion of mass led to C Huygens' and Newton's theoretical ideas about the geometrical structure of rotating bodies being verifiable only in "cosmic dimensions". As a sophisticated experiment arose the determination of the flattening of the Earth. According to Newton's principles the Earth's body (in a first approximation) is an ellipsoid flattened in the direction of the rotation axis. The numerical value of this flattening, the deviation from the ideal spheric form owing to the inertial force due to rotation can be *a priori* estimated following Newton's ideas.

The demonstration of the Earth's rotation is according to Newton's principles in laboratory possible, too, and the axioms of mechanics tell us that the astronomical value of the Earth's angular velocity exactly corresponds to his computations after Huygens' centrifugal forces. This postulate is as Huygens-Mach principle a problem of general relativity and cosmology (Mach-Einstein doctrine).

The astronomer Foucault has shown in 1851 that the rotation velocity of the Earth can be measured in a closed laboratory. This is else the same effect which is known in geography as von Baer's law. Einstein discussed explicitely the hydrographic consequences of the Earth's rotation in a paper with the title "On the cause of meander development of rivers and von Baer's law" (Einstein 1952, pp. 152-160).

The problem of the quantitative coincidence of astronomic, geodetic and experimental determinations of the Earth's rotation became more and more evident since Kammerlingh-Ones (1879) and owing to continuously improving geodetic and astronomic data (including those of artificial satellites). Meanwhile the classical question about the "rotation of the cosmos" (so-called Gödel-cosmos) from which the rotation of the Earth is to be subtracted, remained basically open.

The theory of the figure of the Earth, being a special problem of geodesy, but also one of general geophysics, too, was based by A C Clairaut in 1743 on Newtonian principles. He worked simultaneously on the theory of the Moon's orbit following Newton as on a problem of his time being a decisive test for the exact validity of the Newtonian principles of celestial mechanics, including the gravity law. Clairaut found initially a deviation of the actual movement of the Moon from the computed one. L Euler concluded on this basis to the necessity of a correction of the Newtonian law of gravity, namely a term of screening, corresponding to an absorption constant with the dimension $g^{-1}cm^2$.

A more exact calculation on the basis of Clairaut's theory of the Earth's figure gave a result which corresponded within the limits of measurement errors to the Newtonian law. The validity was confirmed by an argument going back to Laplace which was proven by geophysicists after 1900 with a very high precision.

The geophysical methods for the determination of the Earth's figure, namely using a reversion pendulum (F W Bessel, about 1840) and the investigation on the ratio of mass and inertia initiated by L Eötvös (around 1900), proved with increasing accuracy that the absorption constant λ must be very small. Present estimations yield

$$\lambda \leq 10^{-15} \mathrm{g}^{-1} \mathrm{cm}^{2} .$$

A most important characteristics of the Earth's rotation in the framework of the Newtonian theory of gravity was prognosticized first by Einstein. He computed in 1919 the changes in the moment of inertia of the Earth caused by the partial tides of the Moon due to half monthly and node movements of the Moon's orbit and the changes $\Delta\omega$ in the rotation velocity ω due to the change of inertia. The relative amplitude $\Delta\omega/\omega$ of 10^{-9} was, however, according to A von Brunn too small to be detected empirically at that time (1919).

Independently of Einstein computed H Jeffreys (see Moritz and Hofmann-Wollenhof 1993) the periodic changes in the length of the day $T = 1/\omega$ caused by partial solar and lunar tides. All later studies were based on Jeffreys' work. Today the periodic changes in the rotation velocity of the Earth can be measured using atomic clocks.

Einstein emphasized the basic importance of the deviation of the astronomic definition of time from that of the continuous stream of Newtonian (inertial) time.

The general theory of relativity excelled till recently only by its especially logical and mathematical elegance and by its epistemological brisance (see also Einstein 1969a, 1969b). The basic principles for it differ, however, from those of the Newtonian theory of gravity (and from possible conceptual developments of this theory) only by effects which were nearly all just outside of the limits of measurement accuracy and which could be explained in the framework of the then testing possibilities by other ways, too. Thus e.g. the formula for the perihelion movement of Mercury could be explained by a random exotic distribution of masses in the vicinity of the Sun. Similarly the rather unaccurately known value of the aberration of the light by the gravity field of the Sun could be interpreted in a post-Newtonian framework with some supplementary hypotheses.

The extraterrestrial physics has meanwhile proven that there can be no doubt about the validity of the principles of general relativity (see Einstein 1969a, Steenbeck and Treder 1984, too). Without a detailed discussion, we mention here the Mössbauer-arrangement, the atomic clock, the radioastronomy, the active space research including satellite research and electronic computation technics. They enabled us not only to answer questions such as "Newton or Einstein", but also to ask about the possible fine structures in the Einsteinian theory. It is, however, proven that these corrections may be not more than a few percent of the Einsteinian values.

Einstein's theory of relativity is therefore the basis and initial point for all kinds of future cosmological research. It is to be added that — in addition to the finest effects in celestial mechanics — the most audacious astrophysical statements proved to be correct, too. This concerns the prognosis of stellar-like objects with superheavy material in them and the existence of an immense quantity of low frequency light quanta which fill the whole cosmos as cosmic background radiation (called also 3°-Kelvin-radiation).

One of the present problems of the theory of gravity is the determination of the ratio of heavy and inertial mass:

$$\frac{M}{m} = \frac{\text{heavy mass}}{\text{inertial mass}} .$$

There are three corresponding theories: the first of them goes back to Galilei and to Newton and was chosen by Einstein as basis for his "principle of equivalence of inertia and gravity" and for his theory of general relativity. This equivalence tells us that M/m is a universal constant, i.e. the ratio of heavy and inertial mass is for all bodies and everywhere and at all times the same. In contrast Mach's principle on the relativity of inertia declares the dependence of the mass inertia from the average gravity potential Φ of the cosmos. On the other hand, Dirac and later P Jordan supposed that the gravity constant f depends on the average mass density ζ of the cosmos.

In an expanding cosmos (see Einstein 1969b) the ratio M/m is increasing according to the Mach-Einstein doctrine with time t. One has $M/m \sim t$.

According to Dirac's hypothesis, contrarily, the gravitational load $M = fm$ decreases with time and this results in

$$\frac{M}{m} = f \approx \zeta \approx t^{-1} .$$

The role of geophysics (especially of gravimetry and geodesy) in experimental gravity research is due to the relative great mass of the Earth's body with respect to all bodies being available in the laboratory corresponding to the fact that the gravity field of a body is on its surface (at constant thickness) linearly increasing with its radius R.

Questions about the possible fine corrections to Einstein's general relativistic theory of gravity conduct us again and again to problems in geophysics. Such problems are e.g. the measurement of the variation of the Galilean acceleration g of free fall depending on the one hand on time of the day and on season, on the other hand on the measurement method ("spring balance" or "free fall") (see also Treder 1980, Steenbeck and Treder 1984). The difficulty is here less the accuracy of the measurement to be reached than the estimation of other geophysical and geodetical effects which are not of purely stochastic character.

A significant characteristic of the Earth's body is its relatively high age of 4.6 billion years (practically the same age as that of the solar system). This age is comparable to the so-called age of the universe of about 15 billion years, that is why some changes in the basic physical constants or in other elementary structures of the cosmos during its lifetime (as it is demanded by some cosmological hypotheses) have to be reflected in the geologically and paleo-geophysically studied history of proportions and parameters of the Earth.

P Jordan's book "Expansion of the Earth" became important for the whole complex of these problems. Hans Ertel, a geophysicist, dealt intensively with these problems, too.

In addition to the gravity field, geophysics describes the geomagnetic field, too. W Gilbert, Galilei's contemporary laid around 1600 the foundation for the experimental study of magnetism together with that of the geomagnetic field (see also Pais 1986, Einstein 1969a, 1969b, Treder 1980). The mathematical and measurement-technical developments reached by Gauss and Weber around 1840 came into being also by a world-wide study of the geomagnetic field initiated by A von Humboldt.

The problems of geomagnetism remained closely connected with problems of physical field theories even later, following the development of the Faraday-Maxwellian electrodynamics, of the electron theory by Lorentz and of the quantum theory and relativity. Hypothetical ideas about a unified field theory of gravity and electromagnetism led to different speculations about the origin and properties of geomagnetism and connected hypothetically geomagnetism with the mass and moment of inertia of the Earth. Thus, the study of geomagnetism (and in connection with it, of geoelectricity) yielded at least estimations in the range of orders of magnitude about the gravo-electric or gravo-magnetic effects looked for since Faraday.

Einstein's general relativity informs us about the creation of the gravity field by electromagnetic fields inasmuch as according to Maxwell's tensor it represents a source of the gravity field being an inhomogeneity at the right side of Einstein's gravity equations. Weyl's, Kaluza's, Eddington's and especially Einstein's own unified relativistic field theories include as physical novelty an electromagnetic field induced by changing gravity field as a reversed effect. Schrödinger connected in 1947 this problem with the problem of cosmic gyromagnetism.

It is remarkable that exact measurements of gravity, geodesy and celestial mechanics can only be interpreted and understood together, on the basis of an increasingly accurate knowledge about the dynamics of the Earth's interior, a further study of Newton's and Einstein's gravity theory and of the general relativistic field theories. New results about the figure of the Earth and of the terrestrial fields are connected with such specialities as basic problems of the geometry of Riemann-Einstein's space-time metrics and the further development of general relativity toward a unified field theory.

4. Einstein's special geophysical problems

One of Einstein's publication from the year 1928 (Remark on periodic changes of the Moon's longitude being unexplained by Newtonian mechanics) and another from the same year (New possibilities for a unified field theory of gravity and electricity) deal with problems which were intensively studied by him. The first one includes the forecast of a period of 18.6 years of the average rotation velocity of the Earth and the problems resulting from it in connection with the deviation of the astronomic time from the continuous stream of Newtonian time.

Another publication by Einstein (Über den Äther, 1924) has a most significant footnote:

"On the basis of electrodynamical analogy it seems to be evident to suppose a connection in the form

$$d\eta = C \cdot dm \frac{\vartheta \cdot \vec{r}}{r^2}$$

where dm is a mass moving with the velocity ϑ, $\vec{r}$ and $r = |\vec{r}|$ is the distance to this mass. (This formula is, however, valid in a first approximation only for cyclic movements). The relation between solar field and terrestrial field is correctly given by this formula. The constant C has the dimension

$$\sqrt{\frac{\text{gravity constant}}{\text{velocity of light}}}\,.$$

From this, the order of magnitude of the constant C can be deduced. By substituting these values in the formula, it yields — applied for the rotating Earth — a value being correct for the geomagnetic field. This connection is remarkable, nevertheless, the coincidence may be by pure chance."

Einstein's formula for the gravo-magnetic effect coupled for a moment with the rotational impulse moment of heavy bodies is found in this paper. The coupling constant should be a universal value for which Einstein had supposed a gravity constant. High precision measurements by A Picard and E Kessler initiated by this work (1925) have shown that the coupling constant has to be smaller at least by a factor of 10^{-3}. Blackett and Schrödinger were conducted by this idea to ask about the existence of such a "gyromagnetism" and whether it may have significance for cosmic (especially for the terrestrial) magnetic field. Blackett believed in it. He has shown that the coupling factor should be less by the factor

$$\frac{\text{electron mass}}{\text{proton mas}} = \frac{1}{2}10^{-3}$$

than originally supposed by Einstein.

The basis for the supposition of a universal gyromagnetismus was his opinion that gyro-magnetism and gravo-magnetism will necessaryly follow from a unified field theory including gravity and electricity. This supposition is included in the second cited paper.

5. Outlook

Albert Einstein initiated with his multilaminated works geophysical research, too, and contributed to further studies, often of an interdisciplinary character. Other topics, as e.g. to a unified theory of the general fields led to suppositions which go over the Riemann-Einstein geometry and they may get important at high measurement accuracy for geodesy and geophysics, too (Quo Vadimus? 1990, Moritz 1993).

Everett's big project about the extraterrestrial study of the Earth's figure is a forerunner of such scientific works. Geophysics will get a significance for relativistic physics similarly to astrophysics which obtained this role some years ago.

References

Einstein A 1952: Mein Weltbild, Berlin, Ullstein, 166-170.

Einstein A 1969a: Grundzüge der Relativitätstheorie, Reprint of the 21th edition, 1956, Berlin, Akademie-Verlag

Einstein A 1969b: Über spezielle und allgemeine Relativitätstheorie, Reprint of the 5th edition, 1956, Berlin, Akademie-Verlag

Jeffreys H 1928: Possible tidal effects on accurate time keeping. Monthly Not. Roy. Astron. Soc., Geophys. Suppl. 2, No. 1

Jordan P 1960: Die Expansion der Erde. Folgerungen aus der Diracschen Gravitationshypothese. Braunschweig, Vieweg

Moritz H 1993: In: The Earth and the Universe, W Schröder ed., Sci. Edition, IAGA, Bremen

Moritz H, Hofmann-Wollenhof B 1993: Geometry, relativity and geodesy. Karlsruhe, Wichmann

Pais A 1986: Raffiniert ist der Herrgott - Albert Einstein, eine wissenschaftliche Biographie, Braunschweig, Vieweg

Quo Vadimus? Geophysics for the next generation 1990: G D Garland and J R Apel eds, IUGG Vol. 10, Washington

Steenbeck M, Treder H-J 1984: Experimentelle Schwerkraftforschung. Berlin, Akademie-Verlag

Treder H-J 1980: Fundamental principles of general relativity theories New York, Plenum

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 453–457 (1996)

HERMANN FRITZ AND THE FOUNDATION OF AURORAL RESEARCH

W SCHRÖDER[1]

[Manuscript received August 26, 1996]

Some extracts from Hermann Fritz's publications and letters relating to auroral research are presented so as to give an idea of his contribution to this subject during the 19th century.

Keywords: aurora observations; isochasm; sunspot numbers

1. Introduction

Hermann Fritz's contributions published a century ago are often cited in the international literature, as e.g. by Akasofu (1964), Akasofu et al. (1966), Elvey (1964) and Stringer and Belon (1967). It is, however, surprising that no biography is available on his life and activity. In the following it is striven to give an overview on his life and activity in auroral research based on his accessible published works and of his hitherto unpublished letters.

2. A short biography

Hermann Fritz was born on September 3, 1830 in Bingen am Rhein, Germany. Having finished his study at the Technical College Darmstadt, he was undermaster for technical drawing at the Eidgenössisches Polytechnikum (today Technical College) at Zürich, Switzerland. He was appointed titular professor there in 1872 and lectured on General Mechanics. In this function he published books and papers on the Economical Use of Combustibles (1876), a Handbook on Agricultural Machines (1880) and on Mutual Connections between Chemical and Physical Properties of Chemical Elements and Compounds (1892).

In addition to these technical papers Fritz was interested in solar-terrestrial physics, too. In this field, he was influenced by Rudolf Wolf, an astronomer being known for his merit in the research of the solar (sunspot) cycle. Wolf was also a professor at the Polytechnikum and Director of the Confederation's Astronomical Observatory in Zürich. Fritz was in a close scientific connection with him. Fritz died on August 16, 1893 due to a heart attack.

[1]Hechelstrasse 8, D–28777 Bremen-Rönnebeck, Germany

3. Auroral research before Hermann Fritz's time

The initial stage of auroral research was discussed by Hellmann (1922). There were a lot of often contradictory opinions during several decennia about the nature of aurora.

The first scientific monography was published by Mairan (1733). A connection between aurora and simultaneous irregular variations of the magnetic needle was first observed by Celsius and Hiorter in 1741. Some hundred years later (1845), Lamont noted periodic changes in the regular daily variation of the magnetic needle. Sabine added in 1851 that these changes are closely connected to the appearance of the Sun's surface.

Wolf discovered in 1852 (Wolf 1865) that years with many auroras are often identical with years of high sunspot occurrence. Wolf's result was based on data collected by himself and published in a catalogue containing more than 6300 data on auroras. This is the point where Hermann Fritz joined this work.

4. The "List of Observed Auroras"

The first task was for Wolf and Fritz to collect more observations of auroras. The result of their cooperation was the "List of Observed Auroras, compiled by Fritz" (1873). In addition to personal collection of corresponding data, an important source of this list was Lovering's catalogue (1868) which contained American observations. There are about 300 sources of the data listed on pp. 5 to 13 of this book.

Fritz distributed auroral observations into five groups according to latitude and longitude. Yearly sums are also given for these groups. At the end, some observations of auroras from the Southern hemisphere are also listed. The printing was made on cost of the German Imperial Academy of Sciences. It was the base for future research made by Fritz himself and by other auroral scientists.

5. Aurora and solar activity

Fritz's merit is to prove the close quantitative connection between occurrence of auroras and sunspot numbers. In his book on the aurora (1881, p. 196) he told that he could prove this connection at the end of the year 1862. He published this result for the first time in 1864 (Fritz 1864). In a subsequent publication (Fritz 1865) he presented the parallel course of the auroral and solar activities with much more clarity; he wrote (pp. 257 and 258):

"...that finally aurora is closely connected with, and has a parallel variation with the sunspot appearance, namely so that in time of richest occurrence of spots auroras appears most often and inversely, minima correspond to each other, too, and that maxima of sunspot numbers are less evident, while the case is just the opposite for auroras."

Wolf succeeded to make a distinct progress in 1865 (Wolf 1865) in the description of solar activity by introducing relative sunspot numbers, R calculated as:

$$R = k(10g + f)$$

where g is the number of sunspot groups, f the number of spots and k is a reduction factor depending on the telescope and on the observer. Using old observations of sunspots, Wolf could retrace the series of monthly average sunspot numbers till 1749. Fritz relied in later publications (1878, 1881, 1893) on these relative sunspot numbers. In this last publication he could make use of data collected during the First Polar Year, 1882/83, too.

There had been some controversy about the first detection of the connection between sunspots and auroral occurrence, as the discovery was often attributed to the American E Loomis (1811–1889). In his book on the aurora (1881) Fritz objected in a footnote against accepting Loomis' priority. Loomis had based on his (Fritz's) first publication and he published information from their correspondence without reference.

6. Geographic distribution of auroral occurrence

Hints at the geographic distribution of auroras are found already at Muncke (1825). Loomis (1860) gave a further, even if very rough sketch on the distribution of auroras. This sketch is restricted to an approximate indication of observation zones with 40 and 80 cases in each year.

Definitive research on the geographical distribution of auroras is due to Fritz. He introduced the notion "isochasm", being a curve which connects all the points on the Earth's surface where the occurrence frequency of auroras is expected to be the same. Fritz published in 1874 a map (Fritz 1874) which represents the distribution of auroras in the Northern hemisphere. In this publication, further in his book "The Aurora" (1881) he published tables on the occurrence frequency of auroras in different places of the Northern hemisphere.

Fritz did not extend the study of the frequency distribution of auroras on the Southern hemisphere as available data were inhomogeneous. Boller (1898) was the first who presented a comprehensive catalogue and a map on auroral occurrence of the Southern hemisphere.

7. Further studies of the aurora

Fritz's book, "The Aurora" (1881) being mentioned several times in the present paper was on continuing importance. It contains a summarizing account about the actual knowledge on auroras with a two-coloured map of isochasms for the Northern hemisphere. This book contains also a separate chapter on "The Height of the Aurora above the Earth's Surface". In a number of recent publications about geophysics it is supposed that the first reliable determinations of the height of auroras succeeded only in the years after 1930. It is, however, so that Galle (1872) tried as early as for the aurora on February 4, 1872 to determine the height of the aurora. Immediately after the publication of Fritz's monograph Jesse (1883, 1884) determined the height of some auroras and for the aurora observed on October 2, 1882, he found a height of 122.2 ± 4.5 km. He discussed also the possibility for the determination of auroral arcs and gave summarizing references to height determinations made before 1884.

8. Conclusion

Fritz's long-lasting merit in geophysics is without doubt due to his auroral research, mainly of his proof of the parallel variation in sunspot numbers and auroral frequency, further in the description of the geographical distribution of auroral occurrence using isochasms as introduced by himself. Nevertheless, Fritz dealt also with other meteorological and geophysical problems. This is shown by an extract from his letter on May 7, 1877 where he wrote:

"My special field of research is aurora, hail and electric phenomena as far as their relation to sunspots is concerned. When studying the collected data, one gets different ideas; a series of short papers was the result of such ideas; they are partly published, too. Most of them are found in the New Year's Publications of the Zürich Naturalists' Society with the title: 'From Cosmic Physics'."

The publication mentioned by Fritz in this letter was published in 1875 (Fritz 1875). From among his geo- and astrophysical publications, two books are to be mentioned on the Sun (Fritz 1885) and on periodic phenomena in meteorology and cosmology (Fritz 1889).

References

Akasofu S I 1964: *J. Atm. Terr. Phys.*, 26, 1167–1174.

Akasofu S I, Chapman S, Meinel A B 1966: In: Handbuch der Physik, S Flügge (Hrsg.), Springer, Berlin-Heidelberg-New York, 59/1, 1–158.

Boller W 1898: *Gerlands Beitr. Geophys.*, 3, 56.

Elvey C T 1964: *Planet. Space Sci.*, 12, 783–797.

Fritz H 1863: *Vierteljahr. Naturforsch. Ges. Zürich*, 8, 97–126.

Fritz H 1864: *Vierteljahr. Naturforsch. Ges. Zürich*, 9, 111–139.

Fritz H 1865: *Vierteljahr. Naturforsch. Ges. Zürich*, 10, 229–286.

Fritz H 1873: Verzeichnis beobachteter Polarlichter. Wien

Fritz H 1874: *Petermanns Geogr. Mitt.*, 20, 347–358.

Fritz H 1875: Aus der Kosmischen Physik. Neujahrsbl. Naturforsch. Ges., Zürich

Fritz H 1878: Die Beziehungen der Sonnenflecken zu den magnetischen und meteorologischen Erscheinungen der Erde (Preisschrift). Haarlem

Fritz H 1881: Das Polarlicht. Leipzig

Fritz H 1882a: *Z. Österr. Ges. Meteorol.*, 15, 73–76.

Fritz H 1882b: *Petermanns Geogr. Mitt.*, 28, 376–380.

Fritz H 1883: *Z. Österr. Ges. Meteorol.*, 18, 321–334.

Fritz H 1885: Die Sonne. Basel

Fritz H 1887: *Meteorol. Z.*, 4, 149–159.

Fritz H 1889: Die wichtigsten periodischen Erscheinungen der Meteorologie und Kosmologie. Leipzig

Fritz H 1893: *Vierteljahr. Naturfosch. Ges. Zürich*, 38, 77–107.

Galle J G 1872: *Wschr. Astron. Meteorol. Geogr. N. F.*, 15, 113–118.
Hellmann G 1922: *Veröff. Preuss. Meteorol. Inst.*, 315, 47–58.
Jesse O 1883: *Z. Österr. Ges. Meteorol.*, 18, 238–239.
Jesse O 1884: *Z. Österr. Ges. Meteorol.*, 19, 405–408.
Mairan J J 1733: Traité physique et historique de l'aurore boréale. Paris
Muncke G W 1825: Nordlicht in Gehler's Physikalischem Wörterbuch
Loomis E 1860: *Amer. J. Sci. Arts.*, 30, 89.
Lovering J 1868: On the Periodicity of the Aurora Borealis. Mem. Amer. Acad. New Ser. 10.
Stringer W J, Belon A E 1967: *J. Geoph. Res.*, 72, 4415–4421.
Wolf R 1864a: *Vierteljahr. Naturforsch. Ges. Zürich*, 9, 302–303.
Wolf R 1864b: *Vierteljahr. Naturforsch. Ges. Zürich*, 9, 111–139.
Wolf R 1865: *Vierteljahr. Naturforsch. Ges. Zürich*, 10, 349–384.

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 459-468 (1996)

CHANGES IN THE THEORY OF AURORA BOREALIS[1]

W SCHRÖDER[2]

[Manuscript received October 5, 1996]

Some development phases of the theories on aurorae during the last centuries are discussed. The aurora of March 17, 1710 marks a turning point in the history of auroral research. This aurora occurred in an important period of the intellectual development in Europe, at a time when after the long depression of the Middle Ages people were becoming most interested in natural science and phenomena relevant for geosciences. In the 19th century a connection between the sunspot cycle and auroral occurrences was realized by several people as e.g. by Fritz (1873). Further studies on the nature of auroras have been done in the 19th century by Birkeland, Wiechert and Zöllner. Resulting from their studies, a strong connection has been found between solar activity and the evolution of auroras in the upper atmosphere. Some more recent discussions on theories of magnetic storms and auroras have been published in papers by Chapman/Ferraro, Cowling/Alfvén and Alfvén. These phases of the auroral theories are briefly summarized in the present paper.

Keywords: aurora; geomagnetic storm; history of aurora; magnetohydrodynamics; solar-terrestrial physics

1. Introductory remark

Year 1998 will be the 110th birthday of Sidney Chapman and the 90th birthday of Hannes Alfvén. Both achieved significant results in cosmic physics by important works on the theory of magnetic storms and of auroras. It is noteworthy that the starting points differred in the two cases. Moreover, the discussion between Chapman/Cowling/Ferraro on the one side and Alfvén on the other is an example of the change in the understanding of auroras. As the mentioned scientists were interested in the history of science — e.g. Chapman made an detailed study about the occurrence of auroras in previous decennia — a schematic history of the auroral research seems to be a welcome contribution honouring both of them.

2. Change in auroral research in the 18th century

The shift from a geocentric to a heliocentric concept of the world meant initially no new evaluation in the classification of many natural phenomena. This concerns especially auroras. They were considered as long as till the mid-18th century as a "miraculous face", a "terrifying marvel". Their occurrence meant for people fear from war, epidemies or other existential consequences. A classification as natural phenomenon remained inacceptable even after the Copernican revolution.

[1]Dedicated to Professor Hannes Alfvén (1908–1995)

[2]Hechelstrasse 8, D-28777 Bremen-Rönnebeck, Germany

1217-8977/96/$ 5.00

A basic change concerning the interpretation of auroras occurred following the aurora on March 17, 1716 which was observed everywhere in Europe. It resulted in mass-meetings: the consequences were so far-reaching that the Halle professor Christian Wolff felt himself prompted to speak to the public with the aim to tranquilize them. This lecture by Wolff can be evaluated as the decisive turning point in the history of contemporaneous cosmic-terrestrial physics. In this lecture Wolff emphasized that this celestial phenomenon did not mean a metaphysical omen; it should be understood as a certain kind of thunderstorm which could not fully develop (Wolff 1716). He went on: *"Why should this early birth tell us more than a completely ripe fruit? From all what I told it is sufficient to say that the interpretation attributed to our meteoro could be found neither in the Script, in common sense, nor in the practical experience, but it stems only from the idea which people deduced from his own lunacy. We look therefore at it as at a curious game of nature..."* (Wolff 1716, p. 35).

It is attributed to this concept which was shared by Wagner in one of his later papers (1716) that at least a lot of people started to give a new sense to aurora and together with it, to a number of unusual celestial phenomena. Naturally, superstition was not radically eradicated, nevertheless, Wolff's and Wagner's concept can be understood in the framework of "unity of cosmic and terrestrial physics" (Treder 1973, p. 192). This new direction of thoughts showed a new approach to natural phenomena.

It is remarkable that several 18th century authors (e.g. de Mairan, Weidler and Wiedeburg) published new and detailed discussions on the essence and origin of aurora. The difficulties of the explanation lay in the lack of observations (in spite of the fact that de Mairan (1733) pioneered in collecting corresponding material, Legrand 1985) and in the fact that aurora was no distinguished object of current research. The interest in auroras remained further on connected with eminent appearances which simply "could not be overlooked". The limits of the then knowledge about terrestrial atmosphere show the following sentences by the universal scientist Lomonosov: *"When I look in the hours of rest at heaven, I am remembered not without regret that many chapters of natural science are explored in the tiniest details, but the knowledge about the air coat is in a deep obscurity"* (Lomonosov 1961, p. 263).

Hjorter's and Celsius' activity brought a progress. They succeeded in 1741 to make observations which proved the connection between magnetic storms and the occurrence of aurora.

Recently Treder (1987) emphasized the significance of geophysical observations. It is of historical interest that e.g. Triewald had already introduced such experiments concerning auroras. He let the sunlight fall through a narrow slot on a prism in a dark room and from the prism on the surface of brandy in a bottle. The coloured rays from the surface to the wall resulted in a form of "Northern lights", thus he interpreted this experiment as having certain similarity with aurora. Triewald guessed following Descartes that the aurora is a certain kind of reflection from the snow and ice stored in the polar region.

The suspicion that aurora is an electric phenomenon which should be connected to the Sun was expressed several times in the 18th century (e.g. by Winkler). Wiedeburg who was born in Jena went farthest in this discussion. He wrote: *"I think ... the material of aurora be the electric one"* (Wiedeburg 1770, p. 69).

The decive breakthrough, the expansion of the physical problem to the cosmic conditions, however, did not succeed at that time.

3. Solar-terrestrial physics and aurora in the 19th century

Cosmic research has developed significantly in the 19th century in close connection with the work of the Royal Society, with Alexander von Humboldt's activity, and with that of the Göttinger Magnetischer Verein and of other scientific enterprises. The connection of a change in the variation of the magnetic needle with the occurrence of auroras became soon a proven result. Decisive developments were to come from solar physics.

Ritter had referred to the periodicity in the occurrence of auroras, nevertheless, the most important empirical step forward was the discovery of the periodicity in the occurrence of sunspots by the Dessau pharmacist Heinrich Schwabe. It became clear that certain regularities are present in cosmic physics. The fact that behind this laid "a homogeneity of the laws of nature in the whole cosmos" (Treder 1987, p. 4) could not be yet declared. Rudolf Wolf discovered by comprehensive research an average sunspot period of 11.11 ± 0.038 years. Wolf extended his result published in 1852 by inference about the periodicity of auroras: *"...that the length of the period deduced by myself from the sunspots fits better to the magnetic variations than that of 10 1/3 years deduced directly by Lamont from the latter, but also ... just the years being rich in spots were also spectacularly rich in auroras"* (Wolf 1877, pp. 659 and 660).

Thus a bridge could be established between Earth and Sun. Statistical investigations, especially those made by Fritz further confirmed the supposition and proved a close physical connection between solar and terrestrial-atmospheric phenomena. An observation made on September 1, 1859 brought about the desired progress: Carrington observed on this day a chromospheric disturbance. Records of the observatory Kew have shown that simultaneously all the three elements of the geomagnetic field changed suddenly. In addition, Marchant found that the maximum geomagnetic disturbance occurred then when the spots crossed the central meridian, i.e. the meridian of the Sun was directed toward the Earth. The differences in the observations remained always less than 36 hours. Therefore a close (physical) connection between Sun and Earth had to be suspected.

During the discussion which followed with the aim to extend the knowledge in solar-terrestrial physics, scientists of the Berlin astronomical observatory have played a significant role. Wilhelm Foerster expressed soon the opinion that solar processes should be connected somehow to auroras. Farthest going ideas were expressed by Karl Friedrich Zöllner who had good connections with Foerster. He wrote: *"...all circumstances at the solar surface which are connected with changes of the streaming processes there, create similar changes in the state of the streaming processes in the Earth's interior and create so variations of the telluric magnetism"* (1881, p. 391).

Thus Zöllner's considerations can be embedded into H Wettstein's ideas who published significant ideas about plate tectonics much earlier than Wegener's continental theory became known. Wettstein supposed clear basic physical bonds between solar and terrestrial processes; he wrote: *"All difficulties disappear if geomagnetism is traced back to electric currents which originate from the dislocational effect of solar gravity"* (Wettstein 1880, p. 86).

Wettstein declared against any terrestrial cause of auroras and reached to the remarkable sentence that they are "the result of the effect of solar radiation" (Wettstein 1880, p.86).

Foerster's, Wettstein's and Zöllner's contributions led to new considerations; a real progress could, however, only be reached through new physical perception. The study of auroral spectra (Angström) showed the path; it became evident that the cause for the light phenomena of aurora should first be found.

The work on cathode and channel rays were carried out in the same time. Jacobus Lourens Sirks expressed as early as 1873 the supposition that electric currents around the Earth are caused by the Sun. Eugen Goldstein saw this in 1879 much clearer when he remarked that different phenomena "could be most comfortably brought in connection with electric currents which flow from a central body through the interplanetry space" (see Singer 1977 and Treder 1983). Goldstein supposed that the Sun radiates also electric rays in addition to light beams, and "that for electric communication with the Sun, the Earth itself had not to be a source of electricity and pole of the current, but discharges which have the pole on the Sun, produce radiation which propagates from the Sun into the interplanetary space"(cf. Singer 1977). Goldstein's idea about the continuous existence of streams which flow always and everywhere anticipates in a certain degree Langmuir's and Tonks' plasma conception. As Bierman told based on Tonks' communication, Langmuir and Tonks looked for an appropriate name for their discovery. Tonks told that Langmuir ran a morning into the laboratory and cried "Well, we'll call it the 'plasma'". The name plasma was foreign so for Goldstein as for Wiechert and for other physicists of that time, it is, however, admirable how their publications anticipate some modern ideas. In the year 1897 it was not settled experimentally if cathode rays are charged particles or electromagnetic waves. Emil Wiechert published his most important works in these years. He supposed that these rays are charged particles, but not molecules or atoms of gas, but much smaller elementary particles which are independent from the gas filling the cathode tube. He called them "electric atoms" and guessed that the electric charge is the same as that of ions in electrolysis.

The similarity between aurora and discharges in rarefied gases led to the supposition of an electric origin of the phenomenon; this aspect was treated by a number of authors. In addition to Goldstein, Paulsen supposed that cathode rays from the Sun should be their source. Arrhenius supposed, however, that cathode rays are to be substituted by cosmic dust with negative charge moving through the effect of the radiation pressure.

This discussion signalized a basic change in the development of auroral theory at the end of the 19th century and together with it, a new evaluation of the solar-terrestrial physics.

4. Development in observation and in theory of auroras

Kristian Birkeland participated in this discussion due to his interest in experimental proof of auroral theories. In his case one can speak about an interest in "laboratory geophysics" fully in concordance with Treder's ideas on geophysical experiments. He used as aid the terella model and could explain two important details of the aurora: namely

a) that electrically charged particles appear at the night-side of the Earth in particular spots or in bundles, and

b) the existence of the auroral zone (cf. Dessler 1983, too). A qualitative interpretation of Birkeland's results was given by Carl Störmer. He studied the trajectories of particles emitted by the Sun in the geomagnetic field. Based on this study, he found an at least partial coincidence with observations; magnetically deviated particles occurred with higher probability in a zone around the pole. The position of this zone was, however, different from that observed.

It is known and Störmer's papers prove it that the computation of the paths of charged particles in an inhomogeneous field is very difficult. Störmer's papers (e.g. 1955) show many aspects of the problem. In this connection a publication by Ertel (1933) should be mentioned which was esteemed by Störmer. In this paper he deduced a theorem which enabled him the graphic representation of electron paths in arbitrary magnetic fields. This theorem tells us that "an electron moves in an arbitrary magnetic field so that the vector of its absolute velocity remains constant in relation to a system of coordinates which rotates with the angular velocity $v = \frac{e}{m}h$ " (Ertel 1933). (Here e/m is the ratio of charge to mass of the electron, H is the magnetic field at the position of the electron.)

The discrepancies between theory and observation led Störmer and co-worker to the development of an observation network. He tried to improve knowledge about auroras by a big observational network in Norway (Brekke 1985). Both for Birkeland and Störmer the main problem was, in addition to theoretical ones, the following: too little was known about auroras, and this lack was not filled by catalogues compiled by Fritz, Tromhölt, Rubenson etc. Results on kinematics and morphology, as well as on physics of the aurora were especially lacking.

Exactly in this time photography was for the first time systematically used for geophysical purposes by Otto Jesse from the Berlin astronomical observatory. In the framework of a programme to measure the height of the noctilucent clouds, he established a chain of simultaneous stations where more than thousand single determinations of the noctilucent clouds were carried out. Jesse's same procedure was then used by Störmer in the Norwegian campaign.

By this programme which lasted several years basic knowledge was collected about the morphology and physics of auroras; Störmer strove especially to connect theoretically deduced data with observations. Even if the Birkeland-Störmer theory explained some observations (appearance on the night side of the Earth, position of the arcs, rays and draperies), some reservations remained. Lenard (1910, 1911) supposed that there are unknown radioelements in the Sun which emanate α-radiation with the obtained initial velocity. He has shown that the absorption coefficients computed for this radiation coincided with the depth of penetration observed (i.e. with the height of aurora). Another view was expressed by Vegard (1931). He suspected that α-radiation, i.e. particles with positive charge are responsible for auroras. Especially the order of magnitude of the product HR (H is the scalar magnitude of the magnetic field, R the curvature radius of the path) yielded correct values with the high m/e-ratio of α-particles at normal values of their velocity. Moreover, a number of observed details, as e.g. the sharp lower boundary of curtains could be explained without constraints. It remained, however, difficult to fit lower appearing auroral forms to the absorption conditions of α-radiation.

Störmer tried to imply β-radiation of normal velocity, too, with the supposition that the anormal position of the zone of maximum occurrence was caused by disturbing magnetic fields outside of the Earth's atmosphere. He did not exclude, however, the excitation of some auroras by α-radiation.

5. New concepts in the theory of auroras

The situation at the beginning of the 20th century was rather puzzling. It had been made even more critical that nobody less than Lord Kelvin himself told at the meeting of the Royal Astronomical Society in 1892 "that the connection between magnetic storms and sunspots is unrealistic" (Kelvin 1892, p. 302). Kelvin's point of view was not generally accepted; it is evident e.g. from Wiechert's work that he stuck to a causal connection between solar events and geomagnetic disturbances.

In addition to the Norwegian school, Wiechert was most interested in aurora; statistical studies made by Fritz and Boller are to be mentioned, too.

Sidney Chapman's scientific style was very different from Birkeland's. At the beginning of his activity, he published short notes on problems of geomagnetic variations as well as on other topics. It was in 1918 that he started to work on geomagnetic storms; in spite of his reference to Birkeland, a significant influence of Birkeland's works is not evident on Chapman. That is connected with their different attitudes: Chapman was a mathematician; he dealt always with the mathematical treatment of a physically prepared question; he strove to a mathematically unambiguous solution of the given problem. Birkeland was different, he only returned to the mathematical approximation when this seemed to him unavoidable. In a simplified form it was so that Birkeland was more interested in the details of the phenomenon, while Chapman thought about a general solution and tried to present it.

Chapman and Ferraro (who was at that time a student) presented around 1930 an interpretation of the initial phase of magnetic storms. For both of them the space around the Earth played an important role, where interplanetary plasma is compelled by the magnetic field to corotate with the Earth. Following Lindemann's ideas they started from the supposition that the cloud of particles consisted of an externally neutral plasma and contained both electrons and protons in equal quantity. This could be accepted, nevertheless, they could not give an explanation for the main phase of magnetic storms and for the accompanying auroras. Ferraro described this situation as follows: *"Our first ideas about the ring current, unlike* *those relating to the first phase, were not based on hydromagnetic concepts, and our theory of the main phase is generally considered unacceptable"* (Ferraro 1952, p. 17).

Between 1930 and 1933 Cowling worked with Chapman; he learned from him in 1940 or 1941 about Alfvén's publications; up to this time, Alfvén's work remained practically unknown for Cowling; this fact is quite remarkable. Here we see already the different mechanisms of Chapman's and Alfvén's activities. It is a decisive factor in the explanation why Alfvén's work could not reach a general early acceptance that his works remained rather unknown. And an additional point: if one considers Chapman's publications, one does not find direct discussions with other authors, even not with those who criticized violently his results. Therefore it is not surprising that Chapman cited in his letters to Cowling Alfvén's publications and asked him to check them as some points seemed to him incorrect; he thought that Cowling is the more suitable person to lead the discussion.

As a consequence, Cowling (1942) contested some results by Alfvén. Certain considerations about plasma physics seemed to be completely correct, nevertheless, Alfvén's generalization of singular cases and of their combinations was less plausible for the theory of magnetic storms. Cowling has shown that the Chapman-Ferraro theory yielded an acceptable interpretation in a certain distance from, but not in the vicinity of the Earth. On the contrary, Alfvén's results seemed to be irrealistic for all what happened in a certain distance from the Earth. It followed, however, from his theory a correct interpretation of the aurora for a source region outside of a certain distance from the Earth.

Ferraro explained in 1953 the discussion between Chapman, him and Alfvén, too: *"...Alfvén's theory is open to a number of criticisms. Cowling has pointed out that the supposed motion of the charges whose electric field would radically alter the motion and so destroy agreement with observations. There is a further energy difficulty which would seem to preclude the discharge of the charges to the polar regions in the manner envisaged by Alfvén"* (Ferraro 1952, p. 15).

The theory originally proposed by Lindemann (1919) than developed by Chapman and Ferraro from 1931 on was later discussed by Martyn (1951), Alfvén (1954) and Singer (1957). It was shown that particles are not in a thermodynamical balance with their surroundings from the very beginning. In the cross-section of the beam the velocity has not to be of Maxwellian distribution. Parker (1958) has e.g. shown that particles may originate in the lower solar corona through hydromagnetic waves. The solar magnetic field is so strongly disturbed that the field lines may be of radial direction in the lower part of the corona.

In the following years there was no discussion worth mentioning between Chapman's and Cowling's schools. This was likely due to the fact that Chapman was initially not very interested in auroral physics (Chapman and Ferraro 1930, 1968, 1969). This situation has changed after 1940 when he started to spend time in Alaska for research. On the other hand he restrained himself of direct discussions, as it was already mentioned, that was not his style. That he is why he rejected the first two drafts of Cowling's criticism about Alfvén. Only the third draft was accepted by him (Cowling 1987).

Alfvén has continuously modified his theoretical ideas during this time. Chapman became more and more interested in auroral morphology, in the collection of earlier observations and continued his investigation partly with his student Akasofu. The controversy Chapman-Alfvén was overtaken in the sixties — to quote Axford's words — both by observation and theory.

The new considerations from the field of magnetohydrodynamics, as well as the increasing knowledge about the space in the vicinity of the Earth and about the interplanetary space led to new interpretations. In this connection two aspects are of special significance: firstly Lars Vegard, the Norwegian pioneer of auroral research (1880–1963) published in this time some interesting remarks on the connection of the terrestrial and solar corona to cosmic phenomena, especially to aurora. Vegard was who has proven the unfoundedness of Wegener's "geochoronium" theory, told about beams of mixed rays which cause both aurora and magnetic variations (see Vegard 1931).

The second point is that Chapman and Alfvén have sometimes missed the point in each others' contributions as remarked already by Dessler (1983). Dessler has participated at several conferences where Chapman lectured on his theory of geomagnetic storms and where Alfvén indicated always that it made no sense in physics and that it cannot be fitted to the results described by Birkeland. When Alfvén spoke, Chapman stood up and told that Alfvén's results cannot be conciled with his and with his students' results and added: "We are presently preparing a paper that will shortly be submitted that will clarify these issues." Alfvén protested but Chapman got on and it did not come to a discussion. It would be interesting to know how would the Chapman versus Alfvén discussion ended if Birkeland could participate in it. Chapman had been well known in science when Alfvén presented his ideas, he authored a great number of publications, whereas Birkeland's life ended when Chapman's started: Alfvén started at a time when Chapman had already published significant contributions.

The Chapman/Ferraro/Cowling versus Alfvén discussion being later expanded by basic contributions of other authors (see e.g. Krummheuer 1981) on magnetohydrodynamics, as well as by a new understanding of the near-Earth and interplanetary space is to be understood as an example how new insights could be obtained. This also examplifies the role of scientists in the process of the development of disciplines (see also Rompe and Treder 1982, Treder 1983). Treder wrote concerning physics (but it is also true for earth science): *"... personages of great physicists are important for the development of physics as they propose hodogetic hypotheses and initiate ideas and programmes which are not evident, nevertheless, which exactly imply the path of development"* (Treder 1983, p. 8).

Alfvén's early vision to understand aurora as cosmic plasma which penetrates the upper atmosphere has undoubtedly stimulated subsequent discussion of the solar-terrestrial physics. The controversy of Chapman's and Alfvén's schools if one wants anyway to use the expression "school", are to be seen fully in accordance with Treder's previous thesis. Both — Chapman and Alfvén — as well as their co-workers Ferraro and Cowling contributed surely to the development of cosmic physics.

Table I. Important phases of the development of auroral theory

Year	Author	Contribution
1716	Wolff	Identification of aurora as natural phenomenon
1733	de Mairan	Monography on aurora
1741	Hjorter Celsius	Connection aurora/geomagnetic storms
1803	Ritter	Possible periodicity of auroras
1828	von Humboldt	Proposal of international geomagnetic observations
1843	Schwabe	Periodicity of sunspots
1852	Wolf	Connection of sunspots and auroral frequency
1862	Fritz	Parallel changes of solar activity and auroral frequency
1871	Zöllner	Hints on the importance of solar physics for geomagnetism and aurora
1879	Goldstein	Cathode rays from the Sun as origin of aurora
1901/1913	Birkeland	Experiments and calculations to the auroral physics
1907	Störmer	Computation of paths of charged particles in magnetic fields
1918	Chapman	First comprehensive analysis of geomagnetic storms
1918	Chapman	Theory of geomagnetic storms
1919	Lindemann	Criticism on Chapman's theory
1931/1942	Chapman Ferraro	Theory of geomagnetic storms
1939/1942	Alfvén	Theory of geomagnetic storms and aurora
1942	Cowling	Criticism on Alfvén's theory
1950	Alfvén	Cosmical electrodynamics
1951	Biermann	Conception of the solar wind
1958	Parker	Theory of solar wind
1959	Gold	Introduction of the notion magnetosphere

Acknowledgements

I am indebted to Professors Hannes Alfvén, F.R.S. and Thomas George Cowling, F.R.S. for discussions and many useful hints. S-I Akasofu, W Axford and AJ Dessler helped me with bibliographic references.

References

Akasofu S L, Chapman S 1972: Solar-terrestrial physics. Oxford

Alfvén H 1939: A theory of magnetic storms and of the aurorae. Kgl. Sv Vetenskapsak, Stockholm

Alfvén H 1948: *Phys. Soc. Cassiot Rep.*, 132, 132.

Alfvén H 1951: *Nature*, 167, 984.

Alfvén H 1983–1987: Pers. Mitt.

Biermann L 1966: Sonnenkorona und interplanetarer Raum. S.-B. Bayer. Akad. Wiss. Math.naturw. Kl. 90, 163.

Brekke A 1985: In: Historical events and people in geosciences, W Schröder ed., Bern/Frankfurt

Chapman S, Ferraro V C A 1930: *Nature*, 126, 129.

Chapman S, Ferraro V C A 1931: *J. terr. and Atmosph. Elect.*, 36, 77, 171.

Chapman S, Ferraro V C A 1932: *J. terr. and Atmosph. Elect.*, 37, 147.

Chapman S, Ferraro V C A 1968–1969: Pers. Mitt.
Cowling T G 1942: *Terr. Magn. Atmosph. Electric.*, 47, 209.
Cowling T G 1987: Pers. Mitt.
Dessler A J 1983: *Geophys. Monogr.*, 28, 22.
Ertel H 1933: *Gerlands Beitr. Geophysik*, 38, 142.
Ferraro V C A 1952: *J. Geoph. Res.*, 57, 15.
Ferraro V C A 1960: Theories of magnetic storms and aurorae. Lect. Queens Mary College, London
Ferraro V C A 1963: *Math. Gaz.*, 47, 183.
Kelvin W T 1892: *Proc. Roy. Soc. London*, A52, 302.
Krummheuer F 1981: Zur Theorie erzwungener MHD-Überlagerungen in der Magnetosphäre. Diss., Göttingen
Legrand J-P 1985: *Comp. rend., série générale*, 2, 487.
Lenard P 1910: *Heidelberger Ber.*, 17.
Lenard P 1911: *Heidelberger Ber.*, 12.
Lomonossow M W 1961: Ausgewählte Schriften in zwei Bänden. Bd. II. Naturwissenschaften, Berlin
Mairan J J 1733: Traité physique et historique de l'auroré boréale. Paris
Rompe R, Treder H-J 1982: Über die Einheit der exakten Wissenschaften. Berlin
Singer O 1977: In: Sternzeiten II. H-J Treder (Hrsg.), Berlin, 59.
Störmer C 1912: *Arch. d. sci. phys. et nat.*, 107.
Störmer C 1955: The polar aurora. Oxford
Treder H-J 1973: In: Nicolaus Copernikus 1473–1973. J Herrmann (Hrsg.), Berlin, 179.
Treder H-J 1983: Große Physiker und ihre Probleme. Studien zur Geschichte der Physik. Berlin
Treder H-J 1987: *Gerlands Beitr. Geophysik*, 96, 4.
Vegard L 1931: *Gerlands Beitr. Geophysik*, 32, 288.
Wagner R Ch 1716: Erzehlung derer zu Helmstaedt am gewichenen 17ten Martii. Helmstaedt
Wettstein H 1880: Die Strömungen des Festen, Flüssigen und Gasförmigen und ihre Bedeutung für Geologie, Astronomie, Klimatologie und Meteorologie. Zürich
Wiedeburg J E B 1770: Beobachtungen und Muthmaßungen über die Nordlichte. Jena
Wolf R 1877: Geschichte der Astronomie. München
Wolff Ch 1716: Gedanken über das ungewöhnliche Phaenomenon. Halle
Zöllner F 1881: Wissenschaftliche Abhandlungen mit zwei populären Vorträgen. Vierter Band. Leipzig

Acta Geod. Geoph. Hung., Vol. 31(3–4), pp. 469–477 (1996)

THE CHANGES OF ELECTRICAL RESISTIVITY OF COAL SAMPLES UNDER LOADING

MARCAK H[1] and TOMECKA-SUCHOŃ S[1]

[Manuscript received November 11, 1996]

The inelastic deformations occurring in rock produce a network of channels. The sides of the channels are soaked with brine or other conducting solutions and thus form an electrically conductive net. We briefly analyse here the properties of the net and propose a model for the changes of electrical resistivity of coal under stress. The model consists of two kinds of conducting links, of which one depends on the petrographic properties of rock, and the other is a result of non-elastic development of dilatancy. We show that the net built up as an effect of inelastic deformations may influence the electric properties and constitute a tool for estimating the development of micro-cracks in coal. Laboratory experiments show that the changes in the micro-crack system, such as the beginning of dilatancy, crack propagation and fusion, are associated with changes of the electrical resistivity of rocks subjected to variable loads, and may be recognized on the experimental curve plotting resistivity as a function of applied stress. Thus, the changes of electrical resistivity can be used for assessing the stresses occurring in rocks as a result of mining activity.

Keywords: coal mining; cracks; electrical resistivity; stress

1. Introduction

Since many years Poland is producing large quantities of coal which is often mined under complicated geological conditions. Mine working carries along various risks such as rockbursts, gas eruptions, flows of gas in coal in a controlled or uncontrolled manner. All these phenomena are accompanied by the developments of cracks in coal. Recently, much effort has been directed towards obtaining gas from coal. Because of the low permeability of the coal bed the gas does not flow readily. The changes of coal permeability are associated with changes of external loads. Structurally, coal may be considered as a colloid composed of micellar cores connected with relatively loose transverse bonds. The petrographic structure of coal (Remner et al. 1984, Harpalani and McPherson 1986) permits to assume a bidispersive character of coal porosity, consisting of a macropore subsystem determined by beddings and joint divisibility of the coal bed, and of a micropore subsystem (Fig. 1). Such a structure allows the formation of a large number of empty volumes whose size and surface change under the influence of external factors, such as the presence of fluids filling the channels, external pressure, and gas absorption.

The petrographic structure of coal causes the formation of hydraulic channels which can be classified into two different systems:

[1]University of Mining and Metallurgy, Institute of Geophysics, Al. Mickiewicza 30, PL–30 059 Kraków, Poland, e-mail: tomecka@geol.agh.edu.pl

— a micropore system forming channels of typically 100 Å;

— macropores, mesopores and large pores which are relatively large objects with a relatively small specific surface.

The macropores build a network of channels which may significantly alter the electrical resistivity of coal, inasmuch as the wettening of the crack surface may form electrical connections in which the conductance has a ionic character. The macropore system may change under stress. Marcak and Siemek (1995) have shown that the stress P in the elements which build up the pore-containing system may be expressed by the formula

$$P = \frac{s - \phi p}{1 - \phi} \tag{1}$$

where s is the mean stress caused by external loads, p is the intrapore pressure, and ϕ is the porosity of the coal sample. This equation shows that the stress may develop in the rock matrix either as a result of external load or by hydraulic pressure.

The relationship between permeability and stress is an important factor to consider in the exploitation of gas from coal deposits. Permeability is an essential factor for gas production. When projects of gas exploitation are being made, particular attention must therefore be paid to a possible lack of permeability, which unfortunately is rather importunate. Hydrofracturing increases the opening of fractures, and hence enhances the permeability of the coal bed, often bringing the output of gas up to an economic level. Seismoacoustic measurements have demonstrated that the gas output through coal beds is accompanied by recordable acoustic phenomena.

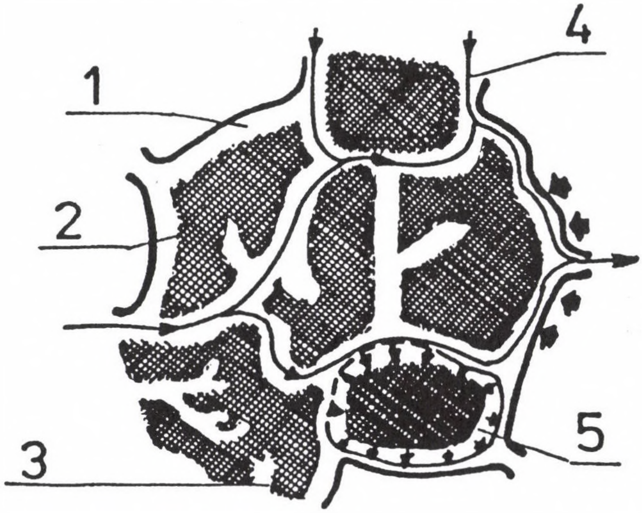


Fig. 1. Model of empty pore structure in coal, and the factors which influence the hydraulic properties of coal: 1. macropores; 2. battle channels; 3. micropores; 4. fluid pressure (widening the channels; 5. overburden rock loading (narrowing the channels) (after Nodzeński 1989)

In particular, Majewska and Marcak (1989) found a correlation between seismoacoustic activity (i.e. the number of acoustic events occurring in a given interval of time) and gas output through rock subjected to triaxial stress. Moreover, it has been proven that electrical resistivity can be employed to assess the properties of fractures.

2. Description of the development of the crack system

Experimental studies carried out by Marcak and Tomecka-Suchoń (1991) have shown that external loading causes changes in the resistivity of rock samples. Before investigating this phenomenon somewhat further, we need first to analyse the development of microcracks associated with the non-elastic strain caused by an external load. The changes occurring in the system of pores and channels in a rock medium under the influence of stress can be divided into two classes, namely the changes caused by elastic strains and the changes caused by non-elastic strains. The latter are particularly important in accounting for the changes of electric resistivity of the rock mass.

Processes of the first class consist principally in closing or expanding the existing channels and pores. They were analysed in detail by Marcak and Siemek (1995). At relatively high stress the elastic strain does not play an important role, the principal cause of change in the system of hydraulic channels being non-elastic strain. We briefly discuss here two aspects of the non-elastic strain development:

— creation and fusion of cracks, and related processes, may be described in terms of a statistical theory (Czechowski 1994) analogous to the well-known kinetic theory of gases, based on an evolution equation similar to Boltzmann's;

— non-elastic strain may also be conceived in terms of the percolation theory.

In the first approach, three processes are considered in the description of the evolution of a crack system, namely the formation of new cracks, the propagation of cracks, and the fusion of cracks. Each of the three processes starts at a characteristic threshold stress value. Thus, the process of formation of new cracks is initiated at a given compressive stress τ_{nuc}, and the already existing cracks start to propagate at a much higher stress level τ_{pro}. Finally, the fusion of cracks starts when the density of cracks in the rock medium exceeds some critical value. Under such conditions, and provided some additional conditions concerning the statistical distribution of crack lengths is fulfilled, we may write (Czechowski 1994)

$$\frac{\partial f(\vec{x},r,t)}{\partial t} + \frac{\partial[vPf(\vec{x},r,t)]}{\partial r} =$$
$$\frac{1}{2}\int_0^r f(\vec{x},r_1,t)f(\vec{x},r-r_1,t)\,\sigma v P_1 dr_1 \;-\; f(\vec{x},r,t)\int_0^\infty f(\vec{x},r_1,t)\,\sigma v P_1 dr_1 \;+\; N(r)\,. \tag{2}$$

In this integro-differential equation, the probability density function f contains as arguments the position of any crack centre $\vec{x}$, the crack length r, and the time t. P is the probability that a crack propagates, v is the velocity of propagation, $\sigma = \sigma(r, r_1)$ is the total cross-section for crack fusion, and $N(r)$ is an additional term which describes the nucleation of cracks of length r.

Czechowski (1994) pointed out that the features of the distributions fulfilling the stated assumptions imply a relative growth of long cracks with increasing stress τ. As long as $\tau < \tau_{\rm nuc}$ no new cracks appear. If $\tau \geq \tau_{\rm nuc}$ the distribution of the number of cracks versus length is exponential. However, no cracks longer than some value $r_{\rm nuc}$ do appear. Eventually, if the actual stress exceeds the critical value $\tau_{\rm pro}$, new cracks form as the result of the fusion occurring during the propagation of already existing cracks. Such a process obviously does not conserve the total number of cracks and changes the distribution function f in a way which is mathematically described by the Boltzmann-Czechowski Eq. (2). The process favours the formation of long cracks, which begin to predominate.

The latter two processes, namely propagation and fusion of cracks, constitute the principal mode of non-elastic stress relaxation and energy release during the final stage of destruction of fragile rock. It seems plausible enough to suppose that the distribution of crack length is correlated with the difference between stress already existing in the rock and the destructive stress.

The second approach is via the theory of percolation. The latter also definitely demonstrates a change of the crack system with time. Actually, in the percolation theory we investigate the evolution of s-dimensional crack clusters in time. It can be shown that at the start of the process, the number n_s of s-dimensional crack clusters is given by (c being a constant)

$$n_s \approx s^{-\mu} \exp\left[-(\rho - \rho_c)\, s^c\right] . \tag{3}$$

This happens before the crack density reaches a critical level ρ_c permitting the percolation process, *i.e.* the connection of two opposite sides of a rock volume by means of a continuous system of cracks. At this stage, the exponential predominates in the cluster distribution function.

Near the critical percolation density ρ_c the pattern of the distribution changes and becomes a power law, namely (Stauffer 1979)

$$n_s = n_0 s^{-\mu} \tag{4}$$

where n_0 represents the number of one-element clusters, and μ is a constant. The percentage of large elements in the number of newly formed elements obviously increases when ρ gets close to ρ_c. The clusters formed according to percolation theory possess a complex multidimensional structure. For this reason, conclusions about the structure of conductive elements in coal based on the considerations presented here must get additional justification. Such a justification is provided by the fractal character of the energy of seismoacoustic signals (Marcak 1994), and the decrease of the fractal dimension of the seismoacoustic energy distribution law which occurs before a large shock.

3. Mathematical models for investigating the electrical resistivity of coal

The nature of electrical conduction in coal is demonstrated clearly by the results of an experiment carried out in the department of geophysics of the University of Mining and Metallurgy, Cracow. In this experiment, a ground coal sample was reconstituted by means of resins, and its electrical resistivity was measured. It turned out that the latter was significantly higher than that of the original sample with a different petrographic structure. This result indicates that the electrical conductivity of natural coal is of a ionic character.

The resistivity of the system of connections forming the basal electric net depends on the number and quality of the conducting branches emerging from each node of the net. We shall model the network as a result of random connections, assuming that connections can only occur between neighbouring nodes. For such a network the conductivity of the whole system can be found by calculating the determinant $|\sigma_{ij}|$ of the network matrix $[\sigma_{ij}]$. If N is the number of nodes of the network, the determinant is of rank N. The elements of the matrix $[\sigma_{ij}]$ are given as follows. The element in the n^{th} position ($1 \leq n \leq N$) on the principal diagonal, i.e. σ_{nn}, represents the sum of the conductivities in all the branches emerging from the n^{th} node. Any off-diagonal element σ_{ij} ($i \neq j$) gives the value of the conductivity between nodes i and j. If the only mode of connection, resulting from the structure of the net, are connections between neighbouring nodes, the matrix $[\sigma_{ij}]$ contains non-zero values only on two parallel lines bordering the principal diagonal on each side. A zero value indicates that no connection between two particular modes exists.

During the stage of non-elastic straining cracks develop in the deforming medium in such a way as to connect various nodes in pairs, or even connect several nodes simultaneously (Fig. 2). The first kind of crack development is represented in our mathematical model by the appearance, outside the principal diagonal and the two nearest off-diagonal parallel lines, of non-zero elements of relatively high values of conductivity. In particular, if the cracks have length distributions such as those discussed in the previous section, non-zero elements occur in the matrix $[\sigma_{ij}]$ in other positions than the three parallels characteristic of nearest-neighbour connections. The distances of these elements from the main diagonal have the same distributions as the distribution of crack lengths.

The determinant $|\sigma_{ij}|$ may be calculated by expansion into minors associated with the elements of a given row or column. The appearance of new non-zero elements outside the three parallel lines necessitates that we take into account additional components during the evaluation of the determinant. Each additional crack yields an additional element, increasing the value of the net conductivity. The additional value depends on the values of the elements σ_{ij} and on the quality of the net, which is described by an appropriate subdeterminant. On the other hand, the appearance of a crack connecting several nodes changes the conductivity between the subsequent neighbouring elements of the net. The differences in conductivity may be interpreted in this case as the consequence of non-elastic changes in channels. The latter have been discussed in the previous sections and are important for the

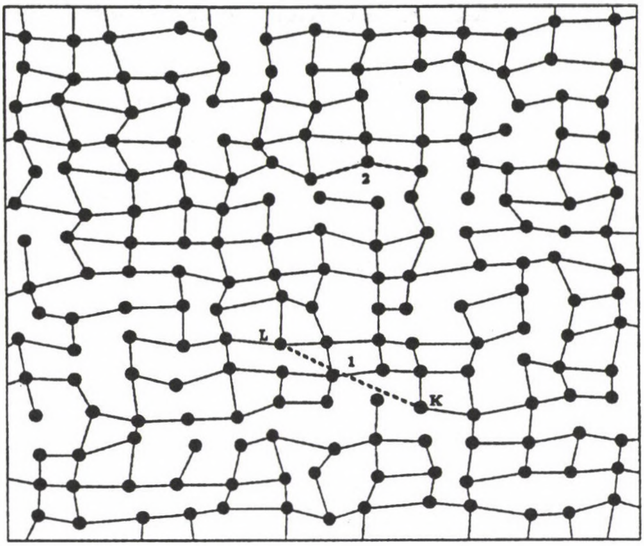


Fig. 2. Simulation of the electrical conductivity: 1. connection between nodes L and K; 2. connection between neighbouring nodes

electrical conductivity of the considered channel. Both kinds of changes obviously cause an increase of the determinantal value $|\sigma_{ij}|$. Thus, they increase both the electric conductance of the net of cracks.

4. Laboratory tests

We carried out experiments on cylindrical coal samples with a diameter of 3 cm and a length of 6 cm. Before performing the measurements, the samples were dried in a vacuum chamber and then saturated with distilled water which dissolved the mineral salts contained in the pores and cracks. Owing to this, the electrical resistivity of the sample depends on the initial density of cracks.

We assume that the conduction of electric current in rocks is of a ionic nature. According to this hypothesis, the electric conductivity of rocks is determined by the network of cracks filled with electrolyte, and the contribution of the rocky matrix to overall conductivity may be neglected.

Laboratory investigations to determine the relationship between sample resistance and applied pressure were carried out on cylindrical rock samples by means of a triaxial press. Copper electrodes were glued to opposite (upper and lower) surfaces of the sample. An indium pad was placed between the sample and the electrodes to secure a uniform stress distribution over the whole surface of the sample. The upper electrode in the sample container was isolated from the press head with a thin mica disc and a 5 millimetre ceramic (Al_2O_3) insulating disk. The lower electrode was grounded through the body of the press. The lateral cylindrical surface of the sample was isolated from the hydraulic fluid with a thin-walled tube

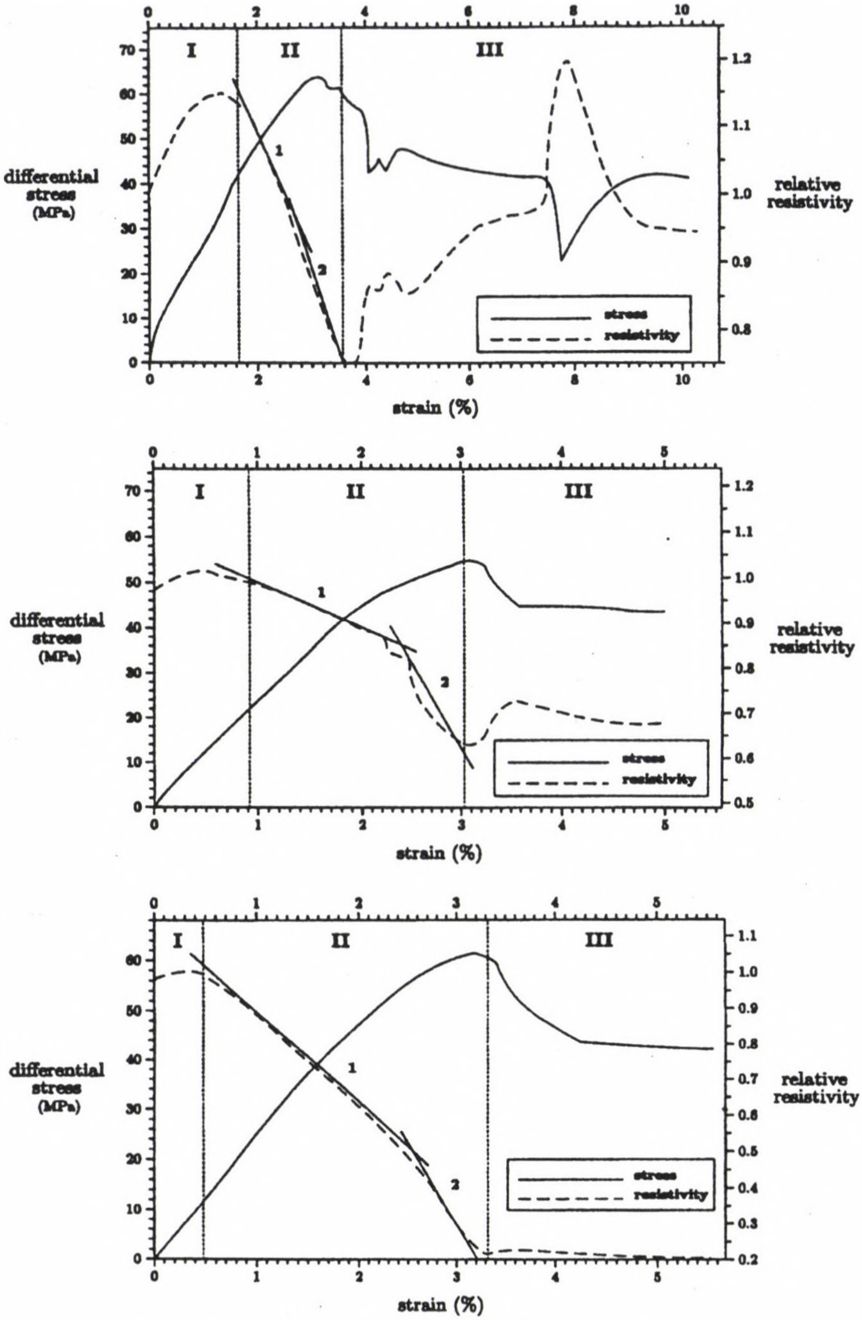


Fig. 3. Changes of resistivity of a coal sample under triaxial stress: I — crack closing stage; II — internal crack development stage; III — post-destructive stage; 1 — crack formation; 2 — crack propagation and fusion

made of a thermoshrinkable foil. Prepared in this way, the sample was placed into the container of the triaxial press. Its resistance was measured after stabilization of the confining pressure, under gradually increasing axial stress. We carried out measurements on a series of samples under a stable confining pressure of 20 MPa.

In addition to determining the linear strain and the resistance changes, we also measured the radial strain induced by known stresses by means of an electric sensor placed in such a way that its inductance changed during the straining of the sample. The change in sensor inductance affects the frequency of impulses generated by a special device, and these frequency variations are recalculated in terms of strain values using previously established calibration characteristics of the sensor. The results of our measurements, plotted in Fig. 3, demonstrate convincingly a relationship between sample resistance and the stage of the sample destruction. The beginning of dilatancy is marked on the resistance curves as an evident change in the curvature and most often by a deviation from the linearity of the strain-resistance relation. The plot obviously shows the changes of the crack distribution caused by propagation and fusion of cracks. The relative changes in electrical resistivity plotted in Fig. 3 prove that the measurements of electrical properties make it possible to identify the different stages in the evolution of a crack system.

5. Conclusions

The knowledge of non-elastic strain in coal is important for planning mining activities, especially for the assessment of the risk of rockbursts and for the exploitation of gas absorbed in coal. The evolution of cracks may be recognized by means of the statistical distribution of the crack lengths. Parameter changes of the latter can be monitored by electrical resistivity methods. In particular, the properties of the crack lengths distributions at different stages of rock destruction as well as laboratory experiments confirmed the possibility to employ electrical measurements for identifying non-elastic deformation in coal. Moreover, resistivity measurements can be used for determining changes in coal permeability caused by hydrofracturing or by an explosive source, and thus allow us to assess the degree of success of such an economic operation.

Acknowledgements

This work was funded by research project No. PB.0089/S6/94/06 (registration 9 S601 028 06) granted by the Committee for Scientific Research (Komitet Badań Naukowych).

References

de Arcangelis L, Redner S, Lonigilio A 1985: *Phys. Rev.*, B 31, 4725–4727.

Czechowski Z 1994: A kinetic model of the evolution of cracks. Publ. Inst. Geophys. Pol. Acad. Sci., A-22 (262), 47 pp.

Harpalani S, McPherson M J 1986: Mechanism of methane flow through solid coal. Proc. 3rd Mine Ventilation Symp. (cited in: Nodzeński 1989)

Majewska Z, Marcak H 1989: *Mining Science and Technology*, 9, 169–179.
Marcak H 1994: *J. Appl. Geophys.*, 31, 205–212.
Marcak H 1995: In: Proc. 2nd Int. Conf. Mechanics of Jointed and Faulted Rock, MJFR-2 Vienna (Austria), 777–782.
Marcak H, Siemek J 1995: *Archiwum Górnictwa* (in Polish), 40, 149–157.
Marcak H, Tomecka-Suchoń S 1991: A model of electrical conductivity of rock samples under triaxial stress (in Polish). Publ. Inst. Geophys. Pol. Acad. Sci., M-15 (235), 287–300.
Nodzeński A 1989: High-pressure desorption of carbon dioxide from coal in terms of the release of gas from a coal seam (in Polish). Zeszyty Naukowe AGH Chemia, Kraków
Remner D J, Ertekin I, King G R A 1984: In: Proc. SPE Eastern Regional Meeting (West Virginia) (cited in: Nodzeński 1989).
Stauffer D 1979: *Phys. Rep.*, 54, 1–74.

Acta Geod. Geoph. Hung., Vol. 31(3-4), pp. 479-480 (1996)

Book reviews

F J GRUBER: *Formelsammlung für Vermessungswesen.* 8th edition, Dümmler Verlag, Bonn (1996, VIII+149 pp, 193 figs)

This book contains definitions and calculation formulas collected from many areas of the practice field of surveying. The sixth edition of this publication has been discussed in this journal (Acta Geod. Geoph. Hung., Vol. 29, pp. 273–274) that is why the present review is shorter and the extensions since the 6th edition are mainly emphasized.

In the Chapter on general basic principles and definitions (pp. 1–7) an extension is found among others in connection with the Gauss-Krüger system of coordinates.

Chapters on basic mathematical formulas (pp. 8–30), on basic tasks of surveying (pp. 31–40), on angle measurement (pp. 41–49), on distance measurement (pp. 50–66), on point determination (pp. 67–86), on transformations in plane (pp. 87–94) remained essentially unchanged.

The Chapter on altimetric survey contains several extended sections, including those on levelling, precise levelling, adjustment of levelling, tower height measurement.

Chapters on surveying for field engineering (pp. 108–126), adjustment (pp. 127–131), statistics (pp. 132–145) remained unaltered, too.

This usefulness of this book has increased by the additions both in education and in practical surveying.

Gy Szádeczky-Kardoss

G BRANDSTÄTTER, F K BRUNNER, G SCHELLING eds: *Ingenieurvermessung 96.* Vol. 2, Dümmler Verlag, Bonn, 1996, pp. V+261, (pagination starts in each part again)

The second part of the lectures held at the XIIth International Course on Surveying Engineering in Graz, 1996 is contained in this book, covering three fields of surveying: C. Quality assurance in industry and construction (7 lectures), D. Data models and information systems (9 lectures) and E. Interdisciplinary engineering constructions (10 lectures).

From Topic C, the role of quality management in bureaus of engineers, in construction industry, in calibration of geodetic and surveying equipments are to be mentioned.

From Topic D, I found important the lectures on the current status of the development of Geographical Information Systems (GIS) and their future perspectives, on object oriented database systems and on data updating. As practical examples, two lectures dealt with the digital map of the city Graz.

In the last Topic E, problems connected with the construction of particle accelerator detectors, with the automobil industry, with tunnel construction and with surveying tasks of the construction of the EUROPIPE-line are discussed. The book ends with three lectures on the surveying tasks in connection with the construction of the Gotthard basic tunnel.

Gy Szádeczky-Kardoss

W SCHRÖDER ed.: *Catalogue of Aurorae Borealis (502 to 1735).* History Commission of the German Geophysical Society and Interdivisional Commission on History of the IAGA, 1996, pp. 44, 10 USD

This small book contains extracts from Ch Kirch's notes on aurora observations. Christfried Kirch (1694–1740) was son of a Berlin astronomer, and continued his father's astronomical observations there. He published the results in many journals, then in 1730 in a book, entitled "Observationes astronomicae selectiores in observatorio regio Berolinensi habitae...". He observed comets, meteors (fireballs), planets, sunspots, solar and lunar eclipses, and described several auroras, among them those of 1716, 1717, 1718 and 1729. His papers include a list of auroral observations, which is reproduced in this book. From the beginning till 1097, there are sporadic observations, then from 1500 on, the list is rather comprehensive, most likely he used sources no more available. There are no observations between 1657 and 1707. Both the draft and the fair copies are reproduced with supplements. This manuscript is followed by a printed description of the aurora in the night 16 to 17 November 1729 as it was observed in Berlin together with some thoughts about the northern light. The issue could be purchased at Andreas Rüdiger below the Berlin townhall. The report is dated from November 29, 1729, less than two weeks after the phenomenon, and it was printed with support from high personages. In addition to a description, he considers the frequency of appearance, the geographical distribution (in Sweden it is much more frequent than in Germany), the height (higher than clouds, and he even supposes that the height must be significantly more, as it is seen in immense areas), then he lists ancient observations, as e.g. by Grégoire de Tours, the 6th century Gallian bishop, and remarks the lack of auroras in the second half of the 17th century, then the re-appearance from 1707 on.

This book is an interesting contribution to science history and to the prehistory of aurora research.

J Verő

ACTA GEODAETICA et GEOPHYSICA
A Quarterly Journal
of the Hungarian Academy of Sciences

VOLUME 31

AKADÉMIAI KIADÓ, BUDAPEST
1996

CONTENTS

Numbers 1–2

Book reviews

Numbers 3–4

Book reviews

AUTHOR INDEX

KEYWORD INDEX

Manuscripts should include the following components which should be presented in the order listed (tables and illustrations should be separated in case of manuscripts on disk, too).

1. Title, name(s) of the author(s), affiliation, dateline, abstract, keywords
2. Text, acknowledgements
3. References
4. Footnotes
5. Legends
6. Tables
7. Illustrations

1. The *affiliation* should be as concise as possible and should include the complete mailing address of the authors. The *date of receipt* will be supplied by the editors. The *abstract* should not exceed 250 words and should clearly and simply summarize the most important methods and results. 5–10 significant expressions describing the content are used as *keywords.* Authors may recommend these keywords.

2. The *text* should be in English and as short and clear as possible. In case of typewritten manuscripts, please note the following:

— avoid possible confusion between o, O (letters) and 0 (zero), l (letter) and 1 (one), ν (Greek nu) and u, v (letters), etc.
— explain ambiguous and uncommon symbols by making marginal notes in pencil
— formulas are to be numbered consecutively with the number in parentheses to the right of the formula. Text references to the equations may then be made by the number in parenthesis. The word equation in this context is to be abbreviated to Eq. and Eqs in the plural
— the International System of Units (SI) should be used.

3. *References* are accepted only in the Harvard system. Citations in the text should be as:

...(Bomford 1971)...or Bomford (1971)...
...(Brosche and Sündermann 1976)...
...(Gibbs et al. 1976b)...

The list of references should contain names and initials of all authors (the abbreviation et al. is not accepted here): for journal articles, year of publication, the title of the journal in abbreviated form, volume number, first and last page. For books or chapters in books, the title is followed by the publisher and place of publication.

All items must appear both in the text and in the list of references.

4. *Footnotes* should be typed on separate sheets.

5. *Legends* should be short and clear. The place of the tables and figures should be indicated in the text, on the margin.

6. *Tables* should be numbered serially with Roman numerals. Vertical lines are not used.

7. All the *illustrations* should contain the figure number and author's name in pencil on the reverse. The most important point with figures is clearness. Photographs and half-tone illustrations should be sharp and well contrasted. Colour photographs will be accepted, but the extra cost of reproduction in colour must be borne by the authors (approx. US$ 260 per page).

Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement. Information is sent to the first author if no other wish is expressed. Fifty offprints are supplied free of charge, for further copies please contact the Editorial Office.

Printed in Hungary
Akadémiai Nyomda, Martonvásár