IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

CONTENTS

****

http://www.met.hu/firat/ido-e.html

VOL. 102 * NO. 1 * JANUARY–MARCH 1998

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: P.O. Box 39, H-1675 Budapest, Hungary or
Gilice tér 39, H-1181 Budapest, Hungary
E-mail: gmajor@met.hu or *antal@met.hu*
Fax: (36-1) 290-7387

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary;
E-mail: gmajor@met.hu or *antal@met.hu*; *Fax: (36-1) 290-7387*

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 1, January–March 1998, pp. 1–18

Generation, detection and granulometry of nanoparticles in the air

M. B. Attoui[1], **A. Renoux**[1] and **D. Boulaud**[2]

[1]*Laboratoire de Physique des Aérosols et de Transfert des Contaminations, Université Paris XII, Avenue du Général de Gaulle, 94010 Creteil, France*

[2]*IPSN CE Saclay 91191 Gif S/Yvette, France*

(Manuscript received 16 June 1997; in final form 10 November 1997)

Abstract—Nanoaerosols are present in the atmosphere in high concentrations. They are formed from the gas phase by homogeneous or heterogeneous condensation. Natural radioactive aerosol (radon daughter elements) can also be of nanometric or even sub-nanometric size. Nanometric aerosols are also found in some industrial conditions such as: furnaces, combustion, chemical reactors for synthetic materials, nanostructures, ceramics, etc. In this paper we review the nanometric aerosol generators for polydisperse aerosols based on spraying and evaporation-condensation. We also consider electric selection for monodisperse generation (with the Differential Mobility Analyzer and the Spectomètre de Mobilité Electrique Circulaire). We examine the detection of this aerosol using the *u*ltrafine *c*ondensation *n*ucleus *c*ounter (UCNC) and its sizing with the *d*ifferential *m*obility *a*nalyzer (DMA) or diffusion battery. Finally we review another device for determining the diameter of nanometric and subnanometric aerosols: the supersonic impactor.

Key-words: nanoparticles, aerosols, granulometry.

1. Introduction

Nanometric aerosols are increasingly studied by specialists both in aerosol science and nanostructure technology. In fact, it is an area where a number of different domains overlap: the vast field of aerosol mechanics (studied by aerosol physicists for more than a quarter of a century), solid-state physics and chemistry at nanometric level. However, these aerosols have been known and studied since the last century (*Coulier*, 1875; *Aitken*, 1888). Aitken was the first to measure the concentration of condensation nuclei in air (sometimes called Aitken nuclei) in 1910. In 1951 Pollack and Murphy confirmed his results. *O'Connor et al.* (1961) demonstrated that the majority of these nuclei,

resulting from gas-particle conversion, have a diameter of less than 11 nm and are electrically neutral. The aerosol produced by algae left by the sea (low tide) is also nanometric and present in high concentration in the air, more than 10^6 cm^{-3} according to the work of *Paugam* (1978).

On the other hand, the sizes of natural radioactive aerosol particles (radon daughter products) in the air in a free state vary from a few angstroms to a few nanometers (*Malet et al.,* 1996a and b; *Hopke*, 1990). Natural radioactive aerosol conforms to the laws of aerosol physics, of course, just as non-radioactive aerosols and, in addition, can be detected using techniques for radioactivity (*Renoux*, 1996; *Kerouanton et al.*, 1996; *Mesbah et al.*, 1996; *Olawin et al.*, 1995). Gas particle conversion in sea air over continental coasts gives rise to high nanometric aerosol concentration (*Tymen et al.*, 1995). Nanometric aerosols are also found in many other conditions (combustion process, chemical reactors, electric ovens, electrical discharges, etc.).

Due to its size this aerosol is, of course, greatly influenced by the activity of the carrier gas molecules. It is therefore subject to Brownian movement. It obeys Einstein's law which expresses its average quadratic movement in terms of time and its diffusion coefficient, the latter dependent upon the size of the aerosol (*Malet et al.,* 1996b).

In this article we will pass in review the means of generating this aerosol along with techniques for measuring its size and its concentration in the air.

2. Detection of the nanometric aerosol

To detect this aerosol we use either its capacity to "grow" to a size which is relatively easy to exploit or its capacity to carry electric charges producing a measurable electric current. There are, therefore, two particle concentration measuring devices based on these two properties which can detect nanometric aerosols. The condensation nucleus counter enlarges the aerosol before detecting it optically. The electrometer measures an electric current (weak in general) induced by an air flow, the aerosols having previously been electrically charged.

2.1 The condensation nucleus counter (CNC)

The condensation nucleus counter is the device used for detecting fine, ultrafine and nanometric particles. It can "see" particles down to one nanometer in diameter. This device works according to the same principle as the first counter (*Aitken*, 1888, 1891), i.e. the particles are enlarged by condensable vapor before counting. Aitken used water vapor to form water droplets by condensation on each of the nuclei, which he then counted individually and visually.

The Pollak counter (*Nollan* and *Pollak*, 1946) was the first photoelectric counter based on the attenuation of a light ray by droplets, again, of water. These counters, along with the commercial versions which followed, use adiabatic expansion to make water vapor condense on the particles. *Bricard et al.* (1972, 1974, 1976) in France, *Sinclair* and *Hoops* (1975) in the United States propose a so-called "continuous flow" counter using butyl vapor to enlarge the aerosol to be detected. The air-borne aerosols, introduced above a tank of butanol, enter a heated horizontal tube to be saturated in vapor and then arrive in a second tube maintained at a low temperature. This tube, called a condenser, is where the butanol condenses on the aerosol particles. The droplets formed in this way, in the ten to twelve microns range (*Agarwal et al.*, 1977), are detected by an optical system. This device described and calibrated by *Agarwal* and *Sem* (1980) has been marketed by the American firm TSI (Thermo System Incorporation) under the name of CNC 3020. Its counting efficiency is only 50% for diameters of 10 nm with a sampling rate of 0.3 liter per minute. TSI has introduced two counting systems into this device (photometric and individual) which allow the measurement of concentration between 0.01 to 10^7 particles per cubic centimeter. A few years later the addition of an optical laser and a bypass loop on the air circuit allowed TSI to reduce the detectable diameter to 7 nm and increase the sampling rate to 1.4 liter per minute. The rate for enlarged aerosols remained to be 0.3 liter per minute. In fact the CNC only gives an integral measure of the aerosol. It does not make it possible to discover the initial diameter of the nucleus. *Nolan* (1972) and *Liu et al.* (1975) review the different types of CNC and their development.

Scientific research in this field of metrology (optical detection after enlargement), branched out into two main directions following the development of the CNC. In the United States it focused on the detection of finer and finer diameters. In France work was concentrated mainly on increasing the rate of sampling. Research into diameter led to devices becoming available on the market. For industrial reasons, commercial or otherwise, high yield devices are not currently available, although a patent has been filed and sold by the CEA.

2.1.1 Ultrafine CNC

The so-called "ultrafine" CNC (TSI Model 3025) has a counting efficacy of more than 50% for 3 nm aerosols. It can detect even smaller diameters but with an efficiency which is much less reliable. This device is based on the work on heterogeneous condensation in a refrigerated pipe by *Metayer et al.* (1982), *Metayer* (1982), *Liu et al.* (1982), *Wilson et al.* (1983), *Stolzenburg* and *McMurry* (1984), *Bartz et al.* (1985), *Keady et al.* (1988) and *Stolzenberg* (1988). The operating principle is practically the same (*Fig. 1*), apart from the condensation. The aerosol is injected into the centre of the condenser which is,

in this case, coaxial. Filtered air saturated with butanol vapor accompanies the aerosol through the condenser to reduce losses by diffusion through the walls. Also to reduce losses, the aerosol does not go right through the saturator. The main part of the sampling (1.4 liter per minute of air-borne aerosols) is filtered, dried and then passed into the saturator before joining the rest of the sampling (0.3 liter per minute) which has not been filtered and which has passed through a capillary tube only a few centimeters in length. In this way the losses by diffusion of the ultrafine aerosol is reduced to a minimum and finer and finer nuclei are activated. This CNC (TSI Model 3025), known as the ultrafine CNC (UCNC), is limited to concentrations of 10^5 particles per cubic centimeter to avoid coagulation.

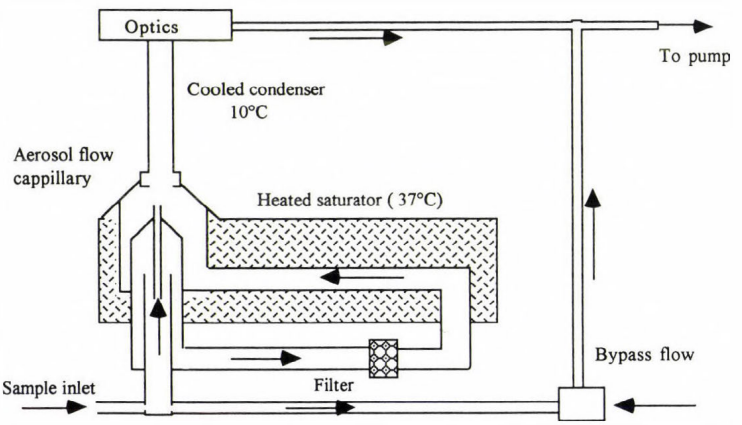


Fig. 1. Schematic cross section of the continuous-flow ultrafine CNC (UCNC).

2.1.2 High-flow CNC

French research on the CNC (CEA - Paris XII University) has investigated the problem of yield. Continuing the work of *Metayer* (1982), *Assa Achy* (1987) developed a high-flow CNC. He designed a new CNC (still based on the principle of continuous flow) capable of detecting aerosols of 10 nm with an efficacy higher than 50% and extracting 28.4 liters per minute! It should be recalled that standard TSI CNCs do not exceed 1.4 liter per minute. The CEA received a patent for this device bought by an American company which has never put it on the market. On the other hand, the possibility of distinguishing nuclei from the produced droplets has been investigated. Indeed, according to *Stolzenburg* (1988) and *Ahn* (1990) the size of the nucleus depends on the size of the final droplet. *Rebours et al.* (1996), *Rebours* (1994) continued the work

of *Metayer* (1982) and *Assa Achy* (1987) and found, using a system of high yield particle enlargement, an indisputable relation of dependence between the size of the drops and the size of the nuclei when the latter was between 4 and 20 nm, making it feasible to use a high flow nanometric granulometer. This device has an activation efficiency of 100% for nuclei of 4 nm.

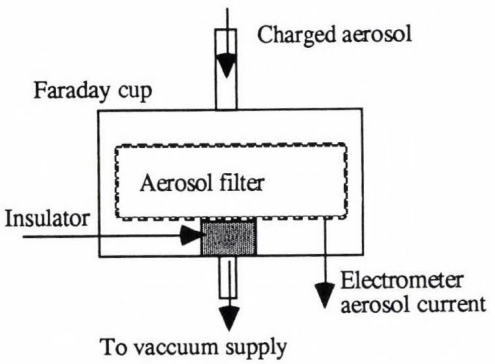


Fig. 2. Schematic cross section of the aerosol electrometer.

3.3 Electrometer

The electrometer, (*Fig. 2*) is a device composed of a Faraday cup and an ammeter capable of measuring very weak currents (in the order of 10^{-15} A). Thanks to this device the volume concentration of an aerosol can be measured if the electric charge carried by each particle is known (*Lui* and *Pui*, 1974; *Sem,* 1975). Moreover the capture or detection efficacy of the electrometer does not depend on the size of the aerosol particles (*Winklmeyer et al.,* 1991; *Wiedensohler et al.,* 1994). In this device the air containing the electrically charged particles enters the Faraday cage to cross a filter with a very high efficiency. The filter placed under a wire mesh traps the particles whilst their electrical charges are recuperated by the mesh and sent to an electrical detector (electrometer). As in case of the CNC, the electrometer gives the concentration of the aerosol but not the size of the particles. The measured current i is a function of the volume output Qv, of the aerosol entering the cup, of the number of charges n carried by each particle, of the elementary charge e and of the concentration of particles C. The current is given by the following relation:

$$i = C\, n\, e\, Q_v.$$

3. Generation of nanometric aerosols

There are several types of generators for polydisperse or quasi-monodisperse nanometric aerosols, for instance the heated wire generator (*O'Connor* and *Roddy*, 1966), exploded wire generator (*Phalen*, 1972), electric arc generator (*Boffa* and *Pfender*, 1973) and electron torch generator (*Tarroni et al.*, 1974). They all, however, have a major flaw: unstable generation over time. Only generators using homogeneous condensation and "electrospraying" can give a stable nanometric aerosol (*Fig. 3*).

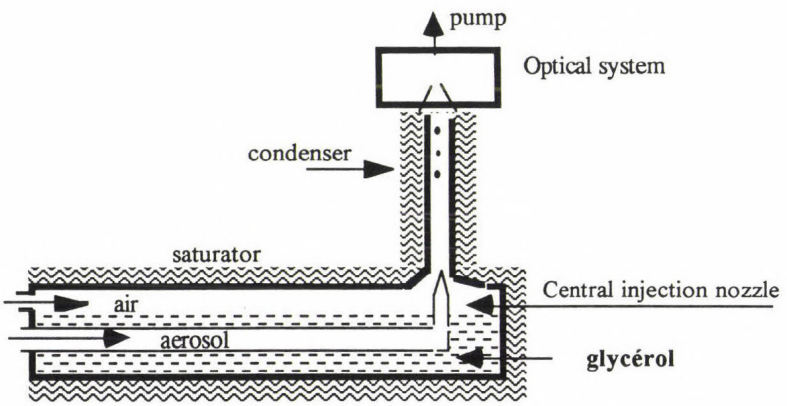


Fig. 3. Schematic cross section of the particle growth system.

3.1 Generation by evaporation condensation

Various publications deal with different examples of this type of generator. They are described by *Spurny* and *Hampl* (1965), *Spurny* and *Lodge* (1968, 1972, 1973), *Spurny et al.* (1980), *Sutugin et al.* (1970/1971) and *Dousaka et al.* (1982). They all use a furnace to vaporize a substance which, by homogeneous condensation, produces a highly concentrated polydisperse aerosol, ($\sigma_g \approx 2.5$). We will consider the devices giving the best σ_g.

To obtain high concentrations of nanometric aerosols, *Scheible* and *Porstendorfer* (1983) suggest to vaporize a substance (NaCl or Ag) in a tube furnace under a nitrogen flow. The vapor-nitrogen mixture is cooled to ambient temperature when leaving the furnace into the extension of the tube (*Fig. 4*). The aerosol obtained is not very polydisperse ($\sigma_g \approx 1.3$ to 2.0). The aerosol volume concentration exceeds 10^6 particles per cm^3. The largest diameters are obtained for the high temperatures (12 nm at 1300°C for Ag, 825°C for NaCl

and 2 nm at 1065°C for Ag, 600°C for NaCl). This generator produces an aerosol where the sizes of the particles become unstable after roughly an hour and the user has no way of influencing its concentration.

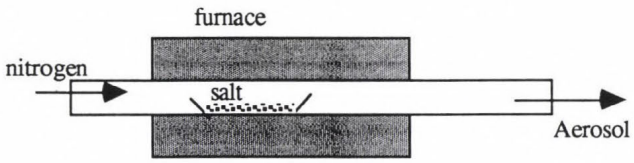


Fig. 4. Nanometric generator.

Bartz et al. (1987) suggest a slightly different nanometric generator based on the one used by *Liu* and *Levy* (1980) to produce sulfuric acid aerosols (*Fig. 5*). They pulverize an aerosol dissolved in a liquid (NaCl in water for instance) in a standard Collison atomizer before drying it and sending it into a tube furnace to be vaporized. The vapors leaving the furnace enter a mixing section through a nozzle where they are accelerated and mixed with cold, filtered, compressed air. The produced aerosol passes into a cooling section where it is brought down to the ambient temperature. The cooling and mixing sections allow improved control of the condensation of the vapors leaving the furnace. The supply of filtered compressed air prevents, by dilution, the coagulation of the aerosol. This generator thus allows an aerosol stability for more than 12 hours. The final size of the nanometric aerosol depends on how the temperature of the furnace, the rate of dilution and the concentration of the atomized solution are regulated. This makes it possible to reduce even more the σ_g of the nanometric aerosol obtained.

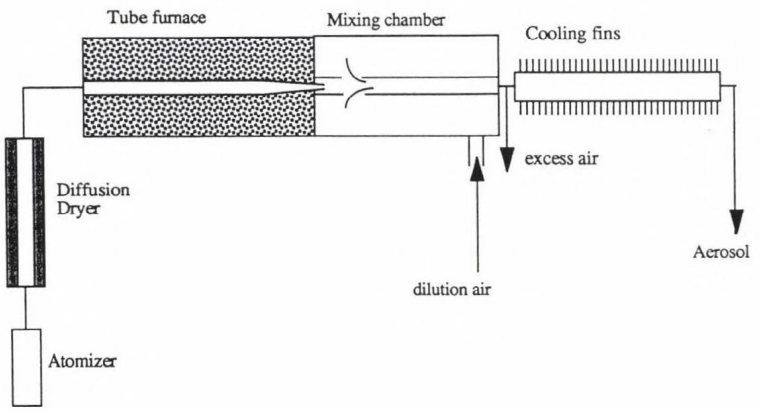


Fig. 5. Nanometric generator of Bartz.

3.2 Electrohydrodynamic pulverization (electrospraying)

Electrospraying, a phenomenon observed by *Zeleny* (1915, 1917) is based on the capacity of an electric field in the ambient air to disperse a liquid. Droplets charged by electrospraying are produced by applying a difference in potential of several kilovolts between a plate and the exit of a capillary tube containing the liquid, as we can see in *Fig. 6* below. The droplets formed are electrically charged. Their size varies, according to different conditions, from several millimetres to several tenths of microns (*Cloupeau*, 1994).

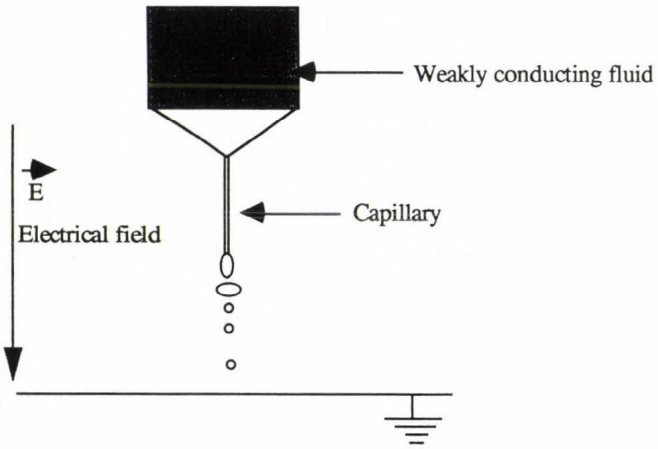


Fig. 6. Schematic cross section of electrospraying.

Chen et al. (1995) suggest a generator (*Fig. 7*) for particles from 4 nm to 1.2 μm using the electrospraying of liquid conductors. This work initiated by *Vonnegut* and *Neubauer* (1952) has produced a fine aerosol generator (*Grace* and *Mareijnissen*, 1994; *Loscertals* and *Fernandez de la Morra*, 1994). The electrospraying phenomenon is explained in detail by *Cloupeau* and *Prunet-Foch* (1989, 1990, 1994). In addition, the magazine, "Journal of Aerosol Science", devoted a special issue (Vol. 25, No. 6, 1994) to this phenomenon.

Finally, photochemical generators (*Dubstov* and *Baklanov*, 1996; *Dubstov et al.*, 1996, 1995) produce relatively monodisperse aerosols ranging in size from several angstroms to a dozen nanometers. Clusters are obtained in a reactor by exposure to radiation from a high pressure mercury lamp before enlargement in vapors supersaturated with dibutyl phthalate.

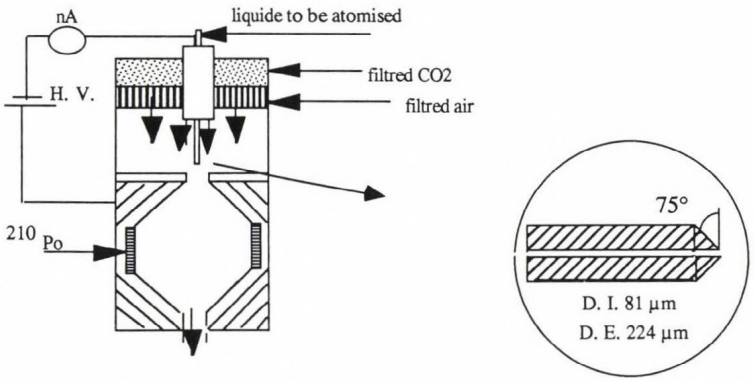


Fig. 7. Chen's generator (electrospraying).

4. Generation of monodisperse nanometric aerosols

None of the above mentioned generators produce a rigorously monodisperse aerosol. A calibrated nanometric aerosol is achieved by electrostatic classification using the electrical mobility of electrically charged aerosols placed in an electric field. There are two devices which make this possible, the DMA and the Radial DMA.

4.1 The differential mobility analyzer (DMA)

To obtain nanometric monodisperse aerosols with the generators described above the electrical properties of the aerosols are exploited in the differential mobility analyzer (*Zeleny*, 1915; *Hewit*, 1957; *Knutson* and *Whitby*, 1975; *Knutson*, 1976). This device generates particles of up to 7 nm. Size accuracy can be as much as 2% (*Liu* and *Pui*, 1974).

The DMA is composed of two coaxial cylinders between which a horizontal electric field is set up. The polydisperse aerosol, previously brought to a state of Boltzmann electrical equilibrium, is introduced along the internal walls of the external cylinder. Dried, filtered air circulates along the internal cylinder where an extractor valve is placed. A variable high voltage, between 0 and 12 kV, makes it possible to vary the electric force ($F = qE$) to which the particle penetrating into the selection area is subjected. A second force is exerted on this same particle due to the downward flow of the filtered air along the walls of the internal electrode.

The aerosol should be electrically charged beforehand and the total charge law should be known if we are to have a precise idea of the quantity of atomic charge carried by each of the particles. The Boltzmann charge law is usually

applied in this device (*Bricard*, 1977; *Liu* and *Pui*, 1974b). The global charge is nil in that case, which means that we have as many particles charged positively as negatively. As a result, within the selection area, the neutral particles pass without being diverted, the negative particles are deposited on the internal walls of the external tube, thus only the positively charged particles are trapped or selected. Consequently, there is a drop in the concentration of the selected particles in relation to that of the injected particles. This generator rarely exceeds 10^5 cm^{-3}.

The "short DMA" is used to reduce losses by diffusion in the DMA. It has been specially developed for nanometric aerosols and works on the same principle as the DMA, but is much shorter in length (11.11 cm instead of 44.44 cm) (*Winklmayr et al.*, 1991). It allows nanometric aerosols down to 3 nm to be generated (*Hummes et al.*, 1996).

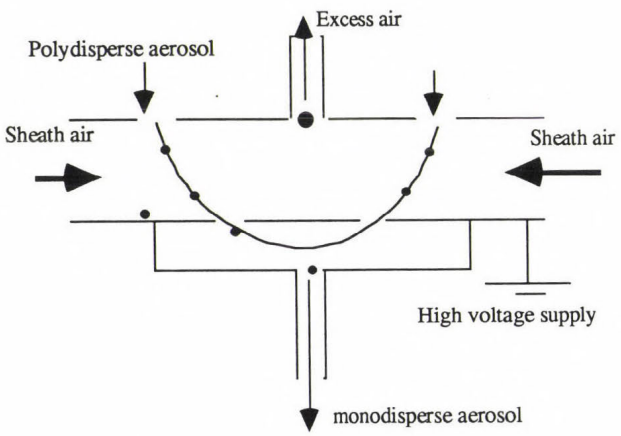


Fig. 8. Schematic cross section of a S.M.E.C.

4.2 The S.M.E.C. (Radial DMA)

This other device has been developed at CEA (*Pourprix et al.*, 1990; *Pourprix* and *Daval,* 1990) and studied by *Mesbah* (1994), *Zhang et al.* (1995), *Boulaud et al.* (1996), *Fissan et al.* (1996), *Lebronnec et al.* (1996). The electromobility spectrometer (named Radial DMA) is composed of two circular plates between which filtered air is introduced. This device with a circular geometry functions, of course, according to the same principle as the DMA (cylindrical geometry). An electric field is set up between the plates by electrifying the lower plate (*Fig. 8*). The polydisperse aerosol, injected through a valve on the upper plate, is extracted with the excess air by a valve on the same plate. The monodisperse aerosol selected is extracted by a nozzle in the lower part of the device on the

second plate. The geometry of this device, the distance between the plates and the length of the selection area all make it much easier to construct, more compact and less cumbersome than the DMA. In addition, there should be fewer losses by diffusion than with other devices of its generation.

5. Granulometry of the nanometric aerosol

After passing in the review the methods for generating and classifying nanometric aerosols, we will now consider the different methods and techniques allowing to determine the size of this aerosol. The electric, aerodynamic (or mechanical) or optical (after enlargement) properties of these aerosols are exploited for this purpose.

5.1 The DMPS/SMPS system

With the introduction of the ultrafine CNC we can now carry out a granulometry in the 5 nm to 1 μm range in an acceptable time thanks to a DMPS system (*D*ifferential *M*obility *P*article *S*ize), composed of an electro-mobility analyzer (DMA), a condensation nuclei counter, a microcomputer (PC) and a program for data inversion (*Keady et al.* 1983). *Knutson* had the idea as early as 1976. The development of the DMPS became possible with the arrival of the microprocessor and especially thanks to a program for data inversion provided by *Fissan et al.* 1983, *Kousaka et al.* (1985), *Hopel* (1978), *Haaf* (1980), *Reischl* (1981), *Alofs* and *Balakumar* (1982) and *ten Brink et al.*(1983) who solved the problems posed by the charge in a bipolar environment and multiple charges. The reaction time was from 2 to 3 min. according to the precision requested. A more powerful and more rapid software in the Windows environment forms, with the previous system (DMA-CNC-PC), what is called the SMPS (*S*canning *M*obility *P*article *S*izer). The SMPS produced by the firm TSI gives the granulometry in 60 or 30 seconds from 5 nm (*Wang* and *Flagan*, 1990). With the arrival of the short DMA we can hope to see, very soon, the detection limit of this system dropping to the 3 nm limit currently imposed by the CNC.

5.2 The diffusion battery

As pointed out in the introduction, nanometric aerosol is subject to Brownian movement. Consequently, the aerosol does not follow the lines of flow of the carrier fluid. In a pipe it tends to diffuse towards the walls and remains fixed there since the adhesion forces are very strong. Thus, there is particle loss on the walls and the aerosol concentration decreases as the air goes along the pipe. This, often undesirable effect is used in the diffusion battery to obtain infor-

mation on the size of an aerosol particle. There are several types of diffusion batteries: parallel plate batteries, cylindric tube batteries, grid batteries, etc. However, only the wire mesh battery is commercially available since it is the only one for which data processing has been developed. In fact, in this type of granulometric measuring the practically unsolvable problem is the restitution of the granulometric spectra which are numerical solutions of a standard unstable Fredholm integral equation. Two types of measures can be made with diffusion batteries: the measure of the penetration or the measure of the deposit on the internal wall of the battery in terms of the capture abscissa. Indeed, using a condensation nucleus counter, the upstream and downstream concentration (and thus the penetration) or the concentration at a given point can be determined.

The grid battery has led to a granulometric measuring device which can measure particles to 2 nm (*Dubstov et al.*, 1995, 1996; *Eremenko et al.*, 1995). Only the multiple-stage wire mesh diffusion battery developed by *Sinclair* and *Hoopes* (1975), *Sinclair et al.* (1976), is produced commercially (TSI). It has ten stages, each one with n grid, n varies according to i, the number of stages, according to the relation: $n = i(i + 1)/2$.

The theory on which this battery is based is explained by *Cheny* and *Yeh* (1980). *Sinclair* (1986) provides an excellent review of the subject. This is the battery which has been studied most over the last years (*Yeh*, 1982; *Yamada et al.*, 1988; *Cheng et al.*, 1980; *Cheng et al.*, 1988; *Schiebel*, 1984; *Wang*, 1993; *Skaptsov*, 1996). At the present time it is the only battery which gives the granulometry of the nanometric aerosol in real time when connected to an ultrafine CNC as shown in *Fig. 9*.

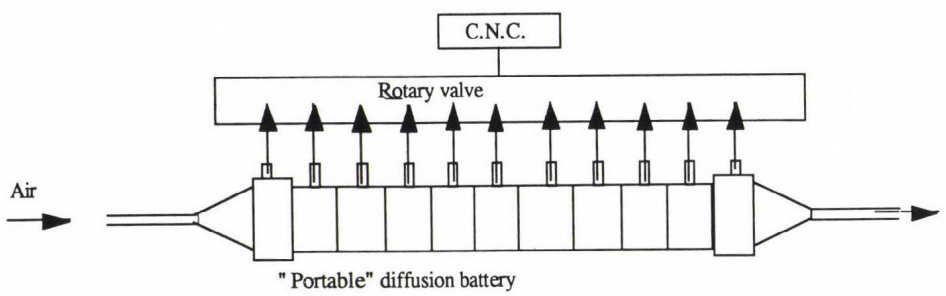


Fig. 9. Schematic diagram of the automatic diffusion battery TSI.

However specific diffusion batteries, for example for radioactive aerosols, developed in laboratories can detect aerosols down to 0.5 nm (*Ramarmurthi et al.*, 1993; *Kerouanton et al.*, 1996).

5.3 The hypersonic impactor

The hypersonic impactor (*Fernandez de la Morra*, 1990a, 1990b) is based on the principle of a shock wave separating the aerosol from its carrier gas. The shock wave is produced sequentially by sending a jet of air at atmospheric pressure through an orifice in a plane plate (infinite) located in a cavity where the pressure is only a few torrs (*Fig. 10*). The distance between the plate and the orifice determines the size of the aerosol to be separated or deposited (*Fig. 11*). The aerosol is detected thanks to an electrometer or, in the case of radon daughter products, by techniques for detecting radioactivity (*Olawoyin et al.*, 1995).

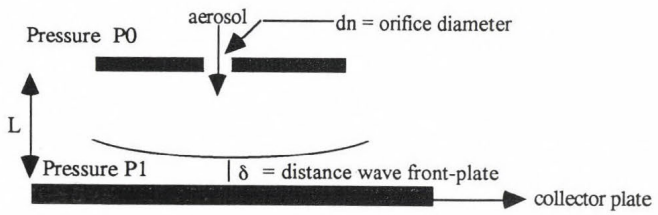


Fig. 10. Shock-wave formation.

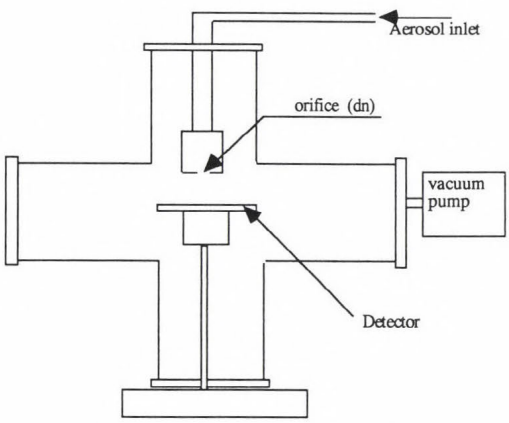


Fig. 11. Sketch of the hypersonic impactor.

For $P_o/P_1 >> 1$ and $d_n << 1$, *Fernandez de la Morra* (1990a) showed that the number of critical Stokes S_o of the flow is given by the relation

$S_o = 0.1983\ \rho_p\ d_p\ C_o^2/(d_n\ P_o)$, where ρ_p is the density of the particle, d_p is its diameter and C_o is the speed of sound in the air.

The capture efficacy of the plate and the diameter of the captured particles depends on L/d_n and P_o/P_1.

6. Conclusion

As explained above, nanometric aerosols first began to interest Aerosol Physicists more than a century ago. However, over recent years, there has been steadily increasing interest, especially in the field of metrology with more and more teams worldwide working on this type of aerosol.

Indeed, to estimate the planet's radiative balance, it is primordial to understand the effect of homogeneous nucleation on the concentration of ultrafine aerosols which act as condensation nuclei for the formation of clouds and precipitation.

For the nanostructure industry, which is developing considerably, the need for information on this aerosol is growing and this can only continue. The importance of this aerosol for the nuclear industry, in the case of an accident in a nuclear reactor, should also be remembered. As we have seen, there is still room for progress in all fields (detection, generation, granulometry, etc.)

Nanometric aerosol physics has not had its day yet — it is far from it. It is faced with new needs and new challenges from fields such as environmental protection, climate evolution, metrology, nanostructure industry, microelectronics, etc.

References

Agarwal, J.K., Sem, J. and *Pourprix, M.*, 1977: A continuous flow CNC capable of counting single particles. *Proc. of Ninth International Conference on Atmospheric Aerosols Condensation and Ice Nuclei*. Galway, Ireland.

Agarwal, J.K. and *Sem, G.J.*, 1980: Continuous flow, single-particle counting condensation nucleus counter. *J. Aerosol Sc. 11*, 343-

Ahn, K., 1988: Ultrafine aerosol measurement and particle activation and droplet growth processes in condensation nucleus counter. Thèse Université du Minnesota.

Aitken, J., 1888/89: On the number of dust particles in the atmosphere. *Proc. R. Soc. Edinburgh 16*, 135-

Aitken, J., 1890/91: On a simple pocket dust-counter. *Proc. R. Soc. Edinburgh 18*, 39-

Alofs, D.J. and *Balakumar, P.*, 1982: Inversion to obtain aerosol size distribution from measurements with a differential mobility analyser. *J. Aerosol Sci. 13*, 513-

Assa Achy, M., 1987: *Mise au point et étalonnage d'un compteur de noyaux de condensation à grand débit*. Thèse de l'Université Paris XII.

Balkanov, A.M., Dubstov, S.N., Caldow, R., Havlicek, M. and *Sem, G.*, 1995: Application of photochemical ultrafine aerosol particles generators for SMPS and DB size resolution. *J. Aerosol Sci. 26*, S751

Bartz, H., Fissan, H. and *Liu, B.Y.H.*, 1987: A new generator for ultrafine aerosols below 10 nm. *Aerosol Sci. Technol. 6,* 163-

Bartz, H., Fissan, H., Helsper, C., Kousaka, Y. Okuyama, K., Fukushima, N., Keady, P.B., Kerrigan,S., Fruin, S.A., McMurry, P.H., Pui, D.Y.H. and *Stolzenburg, M.R.*, 1985: Response characteristics for four different condensation nucleus counters to particles in the 3–50 nm diameter range. *J. Aerosol Sci. 16,* 443-

Boffa, C.V. and *Pfender, E.*, 1973: Controlled generation of monodisperse aerosols in the submicron range. *J. Aerosol Sci. 4,* 103-

Boulaud, D., Pourprix, M., Gougeon, R., Lebronec, E. and *Renoux, A.*, 1996: On the radial flow differential mobility analysers for the determination of aerosol particle density and mass. *J. Aerosol Sci. 27,* S307-

Bricard, J. Madelaine, G., Reiss, P. and *Turpin, P.V.*, 1972: Compteur de noyaux de condensation à flux continu. *C.R. Acad. Sc.* Paris t. 275, 837-

Bricard, J. Delattre, P. and *Madelaine, J.*, 1974: Counting of condensation nuclei at low pressure: its application to photolysis of gaseous impurities in the stratosphere. In *Proc. of the Third Conference on the Climatic Impact Assessment Program. Rep.* DOT-TCS-OST-74-15, US Dept. Transp. Washington, D.C.

Bricard, J. Delattre, P., Madelaine, J. and *Pourprix, M.*, 1976: Detection of the ultra fine particle particles by means of a continuous flux CNC. In *Fine particles* (ed.: *B.Y.H. Liu*). Academic Press, New York.

Bricard, J., 1977: Physique des aérosols. *Rapport CEA R 4831-2.*

Cheng, Y.G., Keating, J.A. and *Kanapilly, G.M.*, 1980: Theory and calibration of a screen-type diffusion battery. *J. Aerosol Sci. 11,* 549-

Cheng, Y.G., Yamada, Y. and *Yeah, H.C.*, 1990: Diffusion deposition on model fibrous filters with intermediate porosity. *Aerosol Sci. Technol. 12,* 286-

Chen, D. R., Pui, D.Y.H. and *Kaufman, S.L.*, 1995: Electrospraying of conducting liquids for monodisperse aerosol generation in the 4 nm to 1,8 μm. *J. Aerosol Sci. 6,* 963-

Cloupeau, M. and *Prunet-Foch, B.*, 1989: Electrostatic spraying of liquids in cone jet mode. *J. Electrostatics 22,* 135-

Cloupeau, M. and *Prunet-Foch, B.*, 1990: Electrostatic spraying of liquids: main functioning modes. *J. Electrostatics 25,* 165-

Cloupeau, M. and *Prunet-Foch, B.*, 1994: Electrohydrodynamic spraying functioning modes: a critical review. *J. Aerosol Sci. 25,* 1021-

Coulier, M., 1875: Note sur une nouvelle propriete de l'air. *J. Pharm. Chimie 22,* 165-

Dubstov, S.N., Balkanov, A.M., Makela, J.M. and *Augustin, J.*, 1996: Performance of the photochemical aerosol generator at 2-10 nm. *J. Aerosol Sci. 27,* S401-

Dubstov, S.N., Koutzenogii, K.P., Levykin, A.N. and *Skubneskaya, G.I.*, 1995: Photochemical aerosol formation of haloidbenzenes — comparaison between theory and experiment. *J. Aerosol Sci. 26,* 705-

Eremenko, S.I., Caldow, R., Baklanov, R. A.M., Havlicek, H. and *Sem, G.*, 1995: Diffusion battery particle sizing system based on the MSA data conversion algorithm: experimental examination 3. *J. Aerosol Sci. 26,* S747-

Fernandez de la Morra, J., Rao, N. and *McMurry, P.H.*, 1990a: Hypersonic impaction of ultrafine particles. *J. Aerosol Sci. 21,* 169-

Fernandez de la Morra, J., Hering, S.V., Rao, N. and *McMurry, P.H.*, 1990b: Inertial impaction of fine particles at moderate Reynolds numbers and the transonic regime with a thin plate orifice nozzle. *J. Aerosol Sci. 21,* 889-

Fissan, H., Hummes, D., Stratmann, F., Buscher, P., Neumann, S., Pui, D.Y.H. and *Chen, D.*, 1996: Experimental comparaison of four differential mobility analysers for nanometer aerosol measurements. *Aerosol Sci. Technol. 24,* 1-

Fissan, H., Helpser, C. and *Thielen, J.H.*, 1983: Determination of particle size distribution by means of an electrostatic classifier. *J. Aerosol Sci. 14,* 354-

Grace, J.M. and *Marijnissen, J.C.M.*, 1994: A review of liquid atomisation by electrical means. *J. Aerosol Sci. 25,* 1005-

Haaf, W., 1980: Construction of a new plate condenser electric mobility analyser and first results. *J. Aerosol Sci. 11*, 201-

Hewitt, G.W., 1957: The charging of small particles for electrostatic precipitation. *Trans. Am. Inst. Elect. Engrs 76,* 300-

Hopke P., 1990: A critical review of measurements of the unattached fraction of radon decay products. *U.S. Department of Energy Office of Energy Research,* No. DE-FGO2-89ER60878.

Hopel, W.A., 1978: Determination of the aerosol size distribution from the mobility distribution of the charged fraction of aerosols. *J. Aerosol Sci. 9*, 41-

Hummes, D., Neumann, S., Fissan, H., Chen, D.R., Pui, D.Y.H., Quant, F.R. and *Sem, G.J.*, 1996: Nanometer differential mobility analyser (nano DMA): experimental evaluation and performance verification. *J. Aerosol Sci. 27,* S135-

Keady, P.B, Denler, V.L., Sem, G.J, Stolzenburg, M.J. and *McMurry, P.H.*, 1988: A condensation nucleus counter designed for ultrafine particle detection above 3 nm. *12th International Conference on Atmospheric Aerosols and Nucleation,* Vienna, Austria.

Keady, P.B., Quant, F.R. and *Sem, G.J.*, 1983: Differential mobility particle sizer: a new instrument for high resolution aerosol size distribution measurement below 1μm. *TSI Quarterly 9*, 3-

Kerouanton, D., Tymen, G. and *Boulaud, D.*, 1996: Small particle diffusion penetration of an annular duct compared to other geometries. *J. Aerosol Sci. 27,* 345-

Knutson, E.O. and *Whitby, K.T.*, 1975: Accurate measurement of aerosol electric mobility moments. J. *Aerosol Sci. 6,* 443-

Knutson, E. O., 1976: Extended electric mobility method for measuring aerosol particle size and concentration. In *Fine Particles* (ed.: *B. Y. H. Liu*). Academic Press, New York.

Kousaka, V., Niida, T., Okuyama, K. and *Tanaka, H.*, 1982: Development of a mixing type condensation nucleus counter. *J. Aerosol Sci. 13,* 231-

Kousaka, V., Okuyama, K. and *Adachi, M.*, 1985: Determination of particle size distribution of ultrafine aerosols using a differential mobility analyser. *Aerosol Sci. Technol. 4,* 209-

Kousaka, V., Okuyama, K., Adachi, M. and *Mimura, T.*, 1986: Effect of Brownian diffusion on electrical classification of ultrafine aerosol particles in differential mobility analyser. *J. Chem. Engng. Jpn. 19,* 401-

Lebronec, E., Boulaud, D., Gougeon, R., M. Pourprix, M. and *Renoux, A.*, 1996: Les analyseurs differentiels de mobilité électrique, une nouvelle méthode pour la détermination de la densité et de la masse des aérosols. 1, 161 *ASFERA*, Paris.

Liu, B.Y.H. and *Pui, D.Y.H.*, 1974a: A submicron aerosol standard and the primary, absolute calibration of the condensation nuclei counter. *J. Colloid and Interface Sci. 47,* 155-

Liu, B.Y.H. and *Pui, D.Y.H.*, 1974b: Equilibrium bipolar charge distribution of aerosols. *J. Colloid and Interface Sci. 49,* 305-

Liu, B.Y.H. Pui, D.Y.H., Hogan, A.W. and *Rich, T.A.*, 1975: Calibration of the Pollak counter with monodisperse aerosols. *J. Appl. Met. 14,* 46-

Liu, B.Y.H., Pui, D.Y.H., McKenzie, R.L., Agarwal, J.K., Jaenicke, R., Pohl, F.G., Preining, O., Reichel, G., Szymansky, W. and *Wagner, P.E.*, 1982: Intercomparison of different absolute instruments for measurment of aerosol number concentration. *J Aerosol Sci. 13*, 429-

Liu, B.Y.H. and *Pui, D.Y.H.*, 1975: On the performance of the electrical aerosols analyzer. *J. Aerosol Sci. 6,* 249-264.

Liu, B.Y.H. and *Levy, J.*, 1980: Generation of submicron sulfuric acid aerosol by vaporisation and condensation. *In Generation of aerosols and Facilities for Exposure Measurements* (ed.: *K. Willeke*). Ann Arbor Science Publishing, Ann Arbor, MI.

Loscertals, I.G. and *Fernandez de la Morra, J.*, 1993: Characterisation of electrospray-generated *of Ultrafine Particles*. Delft, 28-29 May. (ed.: *J.C.M. Marijnissen* and *S. Pratssinis*). Delft University Press, The Netherland.

Malet, J., Montassier, N., Boulaud, D. and *Renoux, A.*, 1996a: Diffusion of aerosols with in flight formation in laminar tube flow. *J. Aerosol Sci. 27,* S261-

Malet, J., Montassier, N., Boulaud, D. and *Renoux, A.*, 1996b: Détermination expérimentale du coefficient de diffusion du ^{218}Po. *Actes ASFERA*, Paris.

Mesbah, B., 1994: *Le spéctromètre de mobilité électrique circulaire; performances et applications.* Thèse de l'Université Paris XII, France.

Mesbah, B., Fitzgerald, B., Hopke, P. and *Pourprix, M.*, 1996: Mobilité des particules ultrafines radioactives; nouvelle technique de mesure. *Actes de l'ASFERA*, Paris.

Metayer, Y., 1982: *Techniques de mesure de l'aérosol fin.* Thèse de l'Université Paris XII.

Metayer, Y., Perrin, M.L. and *Madelaine, G.J.*, 1982: Analysis of a continuous flow condensation nuclei counters. *J. Aerosol Sci. 13*, 170-

Nolan, P.J. and *Pollak, L.W.*, 1946: *Proc. R. Irish. Acad.*, Dublin 51 A2, 9-

Nolan, P.J., 1972: The photoelectric nucleus counter. *Sci. Proc. R. Dublin Soc. A4*, 161-

O'Connor, T.C. and *Roddy, A.F.*, 1966: *J. Rech. Atmos. 2*, 239-

O'Connor, T.C., Sharkey, W.P. and *Flanagan, V.P.*, 1961: Observations on the Aitken nuclei in Atlantic air. *Quart. J. Roy. Meteorol. Soc. 87*, 105-

Olawoyin, O.O., Raunemaa, T.M. and *Hopke, P.K.*, 1995: A system for aerodynamically sizing ultrafine radioactive particles. *Aerosol Sci. Technol. 23*, 121-

Paugam, 1978: *Etude des principaux paramètres électriques de l'air, sur un site côtier, en atmosphère peu polluée.* Thèse de l'Université de Bretagne Occidentale de Brest.

Phalen, R.F., 1972: Evaluation of an exploded-wire aerosol generator for use in inhalation studies. *J. Aerosol Sci. 3*, 395-

Pourprix, M., Daval, J. and *Berne, P.*, 1990: Device permitting the controlled deposit of fine particles on surfaces. *AAAR 90 Meeting*, Philadelphia, PA, June 18-22.

Pourprix, M. and *Daval, P.*, 1990: *Electrostatic Precipitation of Aerosols on Wafers, a Mobility Spectrometer* (eds.: *S. Masuda* and *K. Takahashi*). Pergamon Press, Vol. 2.

Rebours, A., 1994: *Etude de la détection et de la granulométrie en temps réel de l'aérosol ultrafin par un système de grossissement de particules.* Thèse de l'Université Paris XII.

Rebours, A., Boulaud, D. and *Renoux, A.*, 1996: Recent advances in nanoparticle size measurement with a particle growth system combined with an optical counter- a feasibility study. *J. Aerosol Sci. 27*, 1227-

Reischl, G.P., 1981: On the equivalence of low pressure impactor and differential mobility analyser data for submicron particle size distributions. *Proc. of 9th Annual Conference of GESELLSCHAFT Fur Aerosolforschung.* Duisburg, West Germany.

Renoux, A., 1996: Radon and natural indoor radioactive aerosol. *Időjárás 100,* 89-105.

Scheibel, H.G. and *Porstendorfer, J.*, 1983: Generation of monodisperse Ag- NaCl- aerosols with particle diameters betwen 2 and 300 nm. *J. Aerosol Sci. 14*, 113-

Sem, G.J., 1975: Design an application of an electrical size analyser for submicron aerosol particles. Analysis Instrumentation Vol. 13. *Proc. of the 21st Annual ISA Analysis instrumentation Symposium.* Instrument Society of America, 400 Stanwix St., Pittsburgh, Pa. 15222.

Skapsov, AZ.S., Baklanov, A.M., Dubtsov, S.N., Laulainen, N.S., Sem, G. and *Kaufman, S.*, 1996: Measurement of ultrafine aerosol particle size below 15 nm by a diffusion battery. *J. Aerosol Sci. 27 S,* 167-

Sinclair D. and *Hoops, G.S.*, 1975: A continuous flow condensation nucleus counter. *J. Aerosol Sci.* 6, 1-

Sinclair, D., 1986: Measurement of nanometer aerosols. *Aerosol Sci. Technol. 5*, 187-

Spurny, K. and *Hample, V.*, 1965: Preparation of radioactive labeled condensation aerosols: I. aerosols of sodium chloride, silver iodide and sulhuric acid. *Coll. Czech. Chem. Communs 30,* 507-

Spurny, K. and *Lodge, J.P.*, 1968: Aerosol filtration by means of nuclepore filters. Atmos. Environ. 2, 429-

Spurny, K. and *Lodge, J.P.*, 1972: Preparation of radioactive labeled condensation aerosols -V-. *A.J.H.A.J. 33*, 431-

Spurny, K. and *Lodge, J.P.*, 1973: Herstellugen hochdisperser modellaerosole fur storbfurschung and filter-prufung. *Staub-Reinalt. Luft. 33*, 166-

Spurny, K., Opiela, H., Weiss, G. and *Lodge, J. P.*, 1980: On the preparation of highly dispersed radioactively labeled condensation aerosols of silver and silver compounds. *Atmos. Environ. 14,* 871-

Stolzenburg, M.R. and *McMurry, P.H.*, 1984: A theoretical model for ultra-fine aerosol condensation nucleus counter. In *Aerosol* (eds.: *Liu, B.Y.H., Pui, D.Y.H.* and *Fissan, H.J.*). Elseiver New York.

Stolzenburg, M.R., 1988: *An Ultrafine Aerosol Size Distribution Measuring System.* Ph.D. Thesis University of Minnesota.

Sutugin, A.G. and *Fuchs, N.A.*, 1970: Formation of condensation aerosols under rapidly changing environmental conditions — theory and method of calculation. *J. Aerosol Sci. 1,* 287-

Sutugin, A.G., Fuchs, N.A. and *Kostev, E.I.*, 1971: Formation of condensation aerosols under rapidly changing environmental conditions - non coagulated, high dispersed aerosols. *J. Aerosol Sci. 2,* 361-

Taroni, G., Prodi,V., Melandri, L., Bompane,G.F., Zaiacomo, T. and *Formignani, M.*, 1974: Production of ultrafine monodisperse aerosols by condensation. *Proc. of the GAef-meeting*, Bad Sooden, Germany.

ten Brink, H.M., Plomp, A., Spoelstra, H. and *de Vate J.F.*, 1983: A high resolution electrical mobility aerosol spectrometer. *J. Aerosol Sci. 14,* 589-

Tymen, G., Le Bihan, O. and *Droal C.*, 1995: Evolution dynamique de la distribution en taille des particules atmosph;riques ultrafines à la station cotière de Porspoder. 20, *Actes ASFERA,* Paris.

Vonnegut, B. and *Neubauer, R.L.*, 1952: Production of monodisperse liquid particles by electrical atomisation. *J. Colloid Sci. 7,* 616-

Wang, S.C. and *Flagan, R.C.*, 1990: Scanning electrical mobility spectrometer. *Aerosol Sci. Technol. 13,* 230-

Wang, H.C., 1993: Thermal rebound of nanometer particles in a diffusion battery. *Aerosol Sci. Technol. 18,* 180-

Wiedensohler, A., Aalto, P., Covert, D., Heintzenberg, J. and *McMurry, P.H.*, 1994: Intercomparison of four methods to determine size distributions of low concentration ultafine aerosols with illustrative data from arctic. *Aerosol Sci. Technol. 21,* 95-

Wilson, J.C., Hyun, J.H. and *Blackshear E.D.*, 1983: The function and response of an improved stratospheric condensation nucleus counter. *J. Geophys. Res. 88,* 6781-

Winklmayer, W., Reischeil, A.O., Lindner, A.O. and *Berner, A.*, 1991: A new electromobility spectrometer for the measurement of aerosol size distributions in the size range from 1 to 1 000 nm. *J. Aerosol Sci. 22,* 289-

Yamada, Y., Cheng, Y.S. and *Orman, M.M.*, 1988: Evaluation of a coarse screens a diffusion cell material for ultrafine aerosols below 0,02 μm. *J. Aerosol Sci. 19,* 733-

Yeah, H.C., Cheng, Y.S. and *Orman, M.M.*, 1982: Evaluation of various types of wire screens as diffusion battery cells. *J. Colloid Interface Sci. 86,* 12-

Zhang, S.H., Akutsu, Y., Russell, L.M., Flagan, R.C. and *Seinfeld, J.H.*, 1995: Radial differential mobility analyser. *Aerosol Sci. Technol. 23,* 357-

Zeleny, J., 1915: On the conditions of instability of electrified drops, with applications to the electrical discharge from liquid points. *Proc. Camb. Phil. Soc. 18,* 71-

Zeleny, J., 1917: Instability of electrified surfaces. *Phys. Rev. 10,* 1-

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 1, January–March 1998, pp. 19–41

On the origin of model errors. Part I. Effects of the temporal discretization for Hamiltonian systems

Barbara Kádár, István Szunyogh[1]

Department of Meteorology, Eötvös Loránd University,
H-1083 Budapest, Ludovika tér 2, Hungary
E-mail: kadar@nimbus.elte.hu

and **Dezső Dévényi**

NOAA/ERL/Forecast Systems Laboratory,
325 Broadway, Boulder CO, USA; E-mail: devenyi@fsl.noaa.gov
also affiliated with
Cooperative Institute for Research in Environmental Sciences,
University of Colorado at Boulder, Boulder, CO, USA

(Manuscript received 23 September 1997; in final form 21 November 1997)

Abstract—There are two main sources of errors in numerical modeling of the atmosphere: the errors of initial conditions and that of the models themselves. The algebraic structure of equations is strongly related to their integrability, thus the problem of sensitivity to initial data cannot be handled separately from that of model formulation (i.e. of model errors). In a nonintegrable system the utilization of a numerical solution algorithm is a must. The present review paper (in two parts) deals with model errors originating from the inevitable discretization of the continuous equations in space and time. The effects of these errors are investigated in case of Hamiltonian systems in terms of phase space structure. It is pointed out in Part I that (1) even if the discretized ordinary differential equations were perfect the time integration schemes unavoidably introduce errors, (2) the time discretization errors have a strong influence on the time evolution of the probability distribution function in the phase space, thus an ensemble of numerical runs is always distorted by model errors, (3) the initial spin-up process is unavoidable, even if the true state of the system can be observed. In a Hamiltonian system the accuracy of a numerical solution can always be improved by increasing the order of the integration scheme and decreasing the length of the time step. The concepts introduced are illustrated with simple examples and with numerical experiments carried out with the 2D vorticity equation and the two-layer quasi-geostrophic model.

Key-words: model errors, integrability, Hamiltonian structure.

---

[1] Present affiliation: UCAR scientific visitor at the National Centers for Environmental Prediction, USA; E-mail: Istvan.Szunyogh@noaa.gov

1. Introduction

The quality of numerical weather prediction is limited by two factors: the limited accuracy of initial conditions and the imperfection of the models. The main attention has in recent years been focused on the effects of uncertainties in the initial conditions. Lots of effort has been devoted to reduce the analysis errors through the development of enhanced data assimilation (*Pires et al.*, 1996 and references therein) and observation techniques (*Lorenz* and *Emmanuel*, 1997; *Bishop* and *Toth*, 1996). One might have the impression that the only requirement for producing better forecasts is to reduce the initial errors as much as possible. One should not forget, however, that even if the governing laws of the atmosphere were known exactly, and the numerical models based on these knowledge were perfect, furthermore, if the initial fields were exactly known, it is very unlikely that the state of the atmosphere at any future time could be predicted within the error limit of the initial data. This is the consequence of the fact that, to our present knowledge, the atmosphere is a nonintegrable nonlinear system, and as such it shows generic chaotic behaviour.

Since the nonhydrostatic adiabatic primitive equation and its consistent simplifications possess Hamiltonian structure and most of our knowledge about the connection between chaos and integrability is related to conservative Hamiltonian systems we limit our discussions to Hamiltonian systems of equations. Though it can be argued that the atmosphere is not conservative, rather a forced dissipative system, still a number of phenomena can be described within the conservative framework. The sceptic reader can easily check in standard textbooks (e.g. *Holton*, 1992 and *Pedlosky*, 1987) that almost all rigorous results of atmospheric dynamics were achieved with conservative systems of equations. Inspection of the contraction speed of the Liouville volume validates the non-forced non-dissipative approximation for processes with characteristic time much less than the characteristic time of the contraction. Usually it is the case when physical processes are neglected in simplified atmospheric models.

As a recent example, *Camassa* and *Tin* (1996) demonstrated that the dynamics of the dissipative 5-variable Lorenz model (L5) could be completely explored assuming that it was a perturbed version of the conservative one. In other words, the careful examination of the conservative subsystem gave the key to the understanding of the full forced-dissipative system. We believe that the L5 system is not an exception and dynamics of the more complex high dimensional forced-dissipative systems cannot be captured without a solid base of knowledge on their conservative counterparts.

The authors of the present review paper would like to point out that the problem of sensitivity to initial uncertainties and that of integrability are strongly connected, and none of them should be overlooked in favour of the other. We are not dealing here with model errors resulting from the incomplete

and/or incorrect parameterizations of physical and subgrid processes. We focus on the difference between the numerical model and the system modeled in the sense of algebraic structure and integrability. Particularly, we examine the effect of discretizations that, in many cases, destroy the structure of the equations. In the followings we do not attempt to give a complete overview of numerical Hamiltonian problems, this can be found in excellent papers and books like *Mackay* (1992), *Sanz-Serna* and *Calvo* (1994), *McLachlan* and *Scovel* (1993). The main goal here was to collect and illustrate those results of this quickly developing discipline which may substantially deepen our understanding on numerous problems of atmospheric dynamics and numerical weather prediction.

Eulerian equations of atmospheric dynamics are partial differential equations (PDEs) and their numerical approximation usually involves two steps. First, a spatial discretization is performed which transforms the evolution equations into a finite set of ordinary differential equations (ODEs), then a temporal discretization scheme is applied to facilitate the numerical integration of the equations. Motivated by the contemporary results of mathematics on numerical PDEs and ODEs (*Sanz-Serna*, 1992) we follow the dynamical systems approach, i.e. our main focus is rather on the qualitative (global, geometric) than on the quantitative (local) deficiencies caused by the truncations. The first part of the paper deals with the errors introduced in the numerical temporal integration, while the second part discusses the problem of spatial truncation. The reason for reversing the logical order is that almost all results on the chaotic nature of large scale atmospheric motions are achieved by using the hypothesis that the spatially discretized equations perfectly simulates the real atmosphere. While in Part II we will demonstrate that the spatial discretization inevitably alters the qualitative, as well as the quantitative behaviour of geophysical fluid dynamical system, hereafter we will illustrate that the model solutions are inevitably imperfect even for a perfect system of ODEs and initial conditions.

At the beginning two simple examples are presented in Section 2 and then we introduce the bases of Hamiltonian formalism in Section 3. In Section 4 the effects of temporal discretizations on the algebraic structure of the discretized system is discussed and a short account of structure preserving schemes are given. We illustrate the concepts introduced in the paper with simple examples and, in Section 5, with results of numerical experiments that were carried out with a structure preserving spatial truncation of the two-dimensional vorticity equation and the two-layer quasi-geostrophic model. Section 6 summarizes our considerations.

Throughout the paper we tried to introduce every concept that is necessary for understanding the main points, but we did not intend to give a complete introduction to dynamical systems theory. For further reading and references see, for example, *Götz* (1994, 1995).

2. *Simple examples*

Pendulum. With a suitable choice of units (i.e. taking the length of the weightless rod as unit of length and $\frac{1}{g}kg$ as the unit of mass), the system of equations describing the motion of the mathematical pendulum of unit mass takes the form

$$\begin{aligned} \dot{p} &= -\sin q, \\ \dot{q} &= p, \end{aligned} \tag{1}$$

where q is the angle of the rod to the vertical and p is the angular velocity. Since the order of the equation can be reduced by one with the help of a *first integral*, the energy, this system is totally *integrable*. The possible states of the system on the two-dimensional (p,q) phase plane form smooth curves, each of which corresponds to a given energy level (*Fig. 1*). Starting the integration from an initial condition that differs from the "true state" with a small error, the erroneous trajectory we follow will remain close to the true curve except if the two trajectories are on different sides of the separatrix. In this latter case,

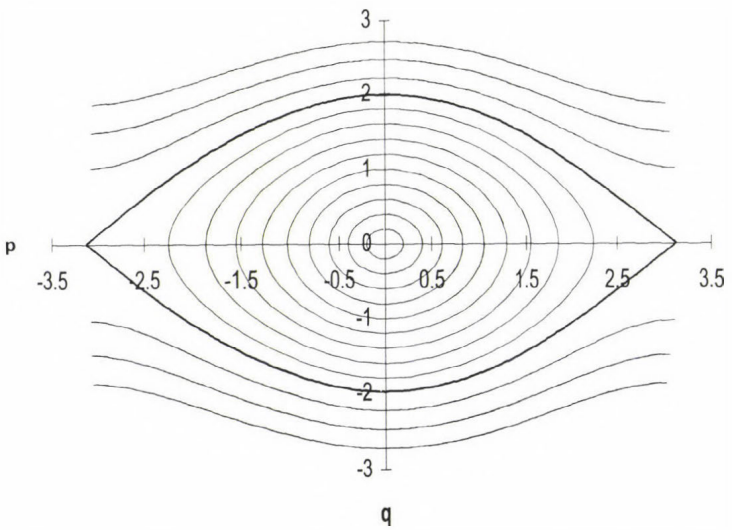


Fig. 1. Phase portrait of the pendulum. The solid lines are phase curves associated with given energy levels. Bold line indicates the separatrix.

however, if several observations of the system's trajectory are available over a given period of time, a 4D data assimilation algorithm may find iteratively an initial condition that corresponds to a qualitatively better trajectory, i.e. one that

follows the right direction of the pendulum. The highly unlikely case when either the true or the analyzed state falls on the separatrix could be handled in a similar way. *Lorenz* (1960) showed that a maximum simplification of the barotrop vorticity equation that is capable to simulate nonlinear barotrop phenomena is a vorticity triad. A vorticity triad consists of only three Fourier modes, with the sum of their wave vectors being equal to zero and the interaction of any pair of them altering the value of the third one. With a suitable transformation of variables and exploiting the invariance of enstrophy the real vorticity triad can be written in the form of Eq. (1) (see e.g. *Bokhove* and *Shepherd*, 1996).

Nondissipative 5-variable Lorenz model. The nondissipative five-component model (L5) of *Lorenz* (1986) was derived from a spectral shallow water model through extreme truncations. After further reduction with the use of an invariant quantity the equations can be written in the form (*Bokhove* and *Shepherd*, 1996)

$$\begin{aligned} \dot{q}_1 &= \epsilon p_1 , \\ \dot{p}_1 &= -\epsilon\, C \sin 2\left(q_1 + b\sqrt{\epsilon}\, q_2\right), \\ \dot{q}_2 &= p_2 , \\ \dot{p}_2 &= -q_2 - \epsilon^{3/2}\, b\, C \sin 2\left(q_1 + b\sqrt{\epsilon}\, q_2\right). \end{aligned} \tag{2}$$

As it was first observed in *Camassa* (1995), this is a nonlinearly coupled system of two integrable subsystems, a nonlinear pendulum and a linear harmonic oscillator: if $b = 0$ the last two equations give the linear differential equation of the harmonic oscillator $\ddot{q}_2 = -q_2$, while the first pair of equations reduces to $\ddot{q}_1 = -\epsilon^2 C \sin 2\, q_1$, which corresponds to the mathematical pendulum (Eq. (1)) with the length of the rod and $\frac{2\epsilon^2 C}{g}\, kg$ taken as the units of length and mass, respectively, and with q_1 measuring the half angle of the rod to the vertical. As it was mentioned previously, the pendulum is equivalent with a vorticity triad, while the solutions of the harmonic oscillator model gravity wave oscillations. The ratio of the frequencies of motions in the subsystems can be controlled with the parameter ϵ. At the limit $\epsilon \to 0$ a formal separation of timescales is present, and one can speak of p_1 and q_1 as *slow* and of p_2 and q_2 as *fast* variables. Accordingly, when $b \neq 0$ and ϵ is small Eq. (2) describes the nonlinear interactions of slow vortical and fast gravity wave-like motions. The system of Eq. (2) has the appealing feature of being the lowest order truncation of the atmospheric primitive equations. *Lynch* (1996) pointed out that structurally similar equations govern the motion of the elastic pendulum. Though the coupling terms between the pendulum and the linear oscillator are somewhat

different in the two systems the basic conclusions of the important papers *Camassa* (1995), *Camassa* and *Tin* (1996) and *Bokhove* and *Shepherd* (1996) do not depend on this detail. The reader, therefore, can always visualize the L5 model as an elastic pendulum in order to gain a better heuristic understanding of the underlying physics.

Bokhove and *Shepherd* (1996) demonstrated, that as the nonlinear coupling term b increases, huge portions of the phase space sections become chaotic. The system does not follow smooth trajectories when started from initial condition, albeit correct, in these chaotic regimes. The key process here is the increasing chaotic exchange of energy between the two subsystems: at a given instant the observer-forecaster may erroneously conclude that the pendulum will have enough energy to complete a rotation, simply because the linear oscillator charges more energy from the other subsystem than that was predicted. Once the inevitable uncertainty in the prediction of the energy exchange is larger than the error in the initial estimate of the energy, model errors start to dominate over the effects of initial uncertainties.

Remarkable theoretical results have been achieved using the geometric approach of dynamical systems theory. With these results (*Camassa*, 1995; *Camassa* and *Tin*, 1996) the vigorously debated existence/nonexistence of an integrable slowest manifold (*Lorenz*, 1980, 1986; *Leith*, 1980; *Errico*, 1984; *Lorenz* and *Krishnamurthy*, 1987) in the L5 has been proved in a local/global sense. While *Camassa* (1995), *Camassa* and *Tin* (1996) utilized the fact that the L5 model is a nonlinearly coupled system of two simple subsystems with well-known phase space geometry, *Bokhove* and *Shepherd* (1996) showed that if the formal separation of timescales is valid the model can be regarded as a slightly perturbed integrable Hamiltonian system, and the theorems of Hamiltonian perturbation theory apply. Thus, if the energy of the gravitational modes is sufficiently small at the beginning of the integration, it remains bounded, and the nonlinear system preserves (although slightly modified) its "nice" trajectories in most of the phase space. It suggests that if one can keep the energy of the linear oscillator at an infinitesimally small level at the beginning of model integration the qualitative dynamics of the pendulum remains intact, the initial and the integration errors have no catastrophic effects on the prediction. The practical realization of this idea is the normal mode initialization (*Lynch*, 1996).

The above example shows that chaos and nonintegrability occur even in very simple models. Indeed, this is generic in nonlinear dynamics. The classification of nonlinear systems as integrable or nonintegrable, however, is a problem that has no general solution yet. Some nonlinear PDEs encountered frequently in mathematical physics exhibit perplexing integrable behaviour. One illustrative example is the Korteweg-de Vries (KdV) equation. In an early experiment (in 1955) Fermi, Ulam and Pasta (FUP) were investigating the behaviour of a nonlinearly coupled chain of harmonic oscillators. They were

expecting sharing of energy between the modes (a kind of statistical equipartitioning of energy), but it turned out that the energy was cycling periodically among the modes initially generated. This astonishing integrable behaviour was explained later when the KdV equation, which is the continuum limit of the chain of oscillators used in the FUP experiment, was proved to have Hamiltonian structure and to be integrable. Its now famous, stable, nonlinear solutions are the so called *solitons*. It is interesting to mention that soliton like solutions are subjects to a non-linear superposition principle deeply related to the integrability property. For more about integrability in nonlinear dynamics see *Tabor* (1989), or other standard texts on dynamical system theory (e.g. *Ott*, 1994; *Arnold*, 1989).

The primitive equations are nonintegrable, at least, no one proved the contrary so far. From this point of view the goal of numerical modeling is to find approximate integrable models, so that in spite of the inevitable initial errors the state of the flow could be estimated over a given period of time. This is not impossible as it has been demonstrated over the history of numerical modeling and is manifested in the indisputable improvement in the skill of numerical weather prediction models.

3. Hamiltonian formulation

In the following subsections we briefly introduce the concept of Hamiltonian structure for both finite dimensional and continuous systems. The finite dimensional Hamiltonian formulation with the so called canonical variables (*generalized coordinates*, *generalized momenta*) has been widely used in classical mechanics. The pursue for canonical representation, however, restricted attention to even-dimensional systems. In the next subsections we present the more general description of finite dimensional Hamiltonian systems that led to the extension of Hamiltonian mechanics to odd-dimensional and also to infinite dimensional problems. A good introduction to the Hamiltonian methods can be found e.g. in *Olver* (1989) and in *Shepherd* (1990).

3.1 Finite dimensional Hamiltonian systems

Let $\mathbf{x}(t) = (x_1(t),\ldots,x_n(t))^T$ denote the vector of the state variables of an n-dimensional system. This system is called *Hamiltonian* if the differential equation describing its evolution in the phase space takes the form

$$\frac{d\mathbf{x}}{dt} = \mathbf{D}(\mathbf{x})\nabla H(\mathbf{x}), \tag{3}$$

where the entries of the *structure matrix* **D** satisfy the equalities

(i) skew-symmetry

$$D_{ij}(\mathbf{x}) = -D_{ji}(\mathbf{x}), \qquad i,j = 1,...,n, \tag{4}$$

(ii) Jacobi identity

$$\sum_{l=1}^{n} \left(D_{pl}\frac{\partial D_{qr}}{\partial x_l} + D_{rl}\frac{\partial D_{pq}}{\partial x_l} + D_{ql}\frac{\partial D_{rp}}{\partial x_l} \right) = 0, \qquad p,q,r = 1,...,n \tag{5}$$

and H, the so called *Hamiltonian function*, is a real valued function, usually the total energy of the system.

For the familiarization with the Hamiltonian notation, it might be instructive to see the examples of Section 1 cast in the above form.

Nonlinear pendulum. If the vector of state variables is $\mathbf{x} = (p,q)^T$, and the Hamiltonian is defined as $H = \frac{p^2}{2} - \cos q$, which is the total (kinetic plus potential) energy of the pendulum, while the skew-symmetric structure matrix is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, Eq. (3) readily gives Eq. (1). The structure matrix being a constant trivially obeys condition (5).

L5 model. To see the equivalence of Eq. (2) and Eq. (3), let $\mathbf{x} = (p_1,p_2,q_1,q_2)^T$, define the Hamiltonian as $H = \epsilon\frac{1}{2}p_1^2 - \frac{1}{2}\epsilon\, C\cos 2\left(q_1 + b\sqrt{\epsilon}\, q_2\right) + \frac{1}{2}\left(p_2^2 + q_2^2\right)$ and use the structure matrix $\begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$, where $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Note that both of the above models are examples for *canonical* Hamiltonian systems, in which $\mathbf{D} = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$ with I being the $n \times n$ identity matrix, and n is the number of degrees of freedom in the system. The q-s are called *generalized coordinates* and the p-s are the *generalized momenta*. The definition Eq. (3), however, is not restricted to this special type of structure matrices, not even to even-dimensional cases.

The reason why canonical Hamiltonian systems have got particular attention is that according to Darboux' theorem (see e.g. in *Olver*, 1989) for any finite dimensional Hamiltonian system defined on an m dimensional manifold there exist local coordinates $(p_1,\ldots,p_n,q_1,\ldots,q_n,z_1,\ldots,z_l)$, $m = 2n + l$ in which the structure matrix has the form

$$\mathbf{D} = \begin{bmatrix} 0 & -I & 0 \\ I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{6}$$

where I is the $n \times n$ identity matrix. The coordinates $z_1,\ldots,z_l$ are the so called *distinguished coordinates*, that are constants along the Hamiltonian flow and any quantity which is function of the distinguished coordinates alone is an invariant of the system. These are the so called *distinguished functions* or *Casimir invariants*.

3.2 Infinite dimensional Hamiltonian systems

The counterpart of Eq. (3) for infinite dimensional evolution equations is

$$\frac{\partial \boldsymbol{u}}{\partial t} = \mathcal{D}\,\delta\,\mathcal{H}(\mathbf{u}), \tag{7}$$

where the prognostic variable u(x(t),t) is a vector valued function of space and time, whose partial derivative with respect to time is taken, and the Hamiltonian $\mathcal{H}(\mathbf{u})$ is a functional (real valued function of a function) whose functional derivative plays the role of the gradient in Eq. (3). The skew-symmetric structure matrix of Eq. (3) is replaced by a skew-adjoint differential operator $\mathcal{D}(\mathbf{u})$ that must satisfy the continuous equivalent of the Jacobi identity (5) as well.

Let the two-dimensional vorticity equation stand here as an example. Take the two-dimensional vorticity $\zeta(\mathrm{x},t)$ as the prognostic variable, the two-dimensional Jacobian $\partial(\zeta,.)$ (with a negative sign) as $\mathcal{D}$ and $\mathcal{H}(\psi) = \frac{1}{2}\int_\Omega \nabla\psi^2 d\mathrm{x}$ as the Hamiltonian, where $\psi(\mathrm{x},t) = \Delta^{-1}\zeta(\mathrm{x},t)$ is the stream function. Upon substitution to Eq. (7) one obtains the familiar governing equation for an inviscid, nondivergent fluid flow over the model domain Ω

$$\frac{\partial \zeta}{\partial t} = \partial(\zeta,\psi). \tag{8}$$

For the Hamiltonian formulation of the shallow water equation, the baroclinic quasi-geostrophic flow over topography and that of the nonhydrostatic primitive equation see *Shepherd* (1990) and references therein.

4. *Temporal discretization*

The problem of integration in time is that in case of most integration schemes a number of important geometric properties of the flow defined by the spatially discretized system are lost under the transformation of timestepping. Such property can be, for example, energy conservation, volume preservation (*Quispel*, 1995), reversibility (*McLachlan et al.*, 1996) or, in case of Hamiltonian systems, preservation of area or symplectic structure (*Yoshida*, 1990). Lots of activity has been focused on devising temporal integration schemes (integrators) that preserve certain geometric properties of the flow, and it seems, that essentially, any properties of the ODE can be retained with an appropriate integrator. But no integration scheme can preserve "everything". As we outline below conservation of the finite dimensional Hamiltonian structure (i.e. *symplectic structure*) and exact conservation of energy are conflicting requirements of a nonintegrable system.

4.1 Preservation of structure

As it was mentioned in the previous section, some equations of fluid dynamics have *structure preserving* spatial truncation, that is the discretization in space of the continuous Hamiltonian system results in a finite dimensional Hamiltonian system of ordinary differential equations. In this subsection we briefly describe what symplectic structure means, and introduce integration schemes that preserve this property, the so called *symplectic integrators*.

For one-degree-of-freedom systems symplectic structure is just a different expression for the preservation of oriented phase plane area. Since the phase space is two-dimensional in this case, it is equivalent with phase space volume preservation which follows from Liouville's theorem. For n-degree-of-freedom (that is 2n-dimensional) systems the symplectic map is a generalization of area preservation stricter than volume preservation. In fact, if the model domain is simple enough, it is equivalent with the system being Hamiltonian. It requires that the sum of certain phase space areas should be conserved along the flow. More specifically, if an m-(not necessarily even-)dimensional system is given with the canonical coordinates $(p_1,\ldots,p_n, q_1,\ldots,q_n, z_1,\ldots,q_l)$, $m = 2n + l$, with the z_i-s being distinguished coordinates, the condition of symplectic area-preservation reads

$$\oint_{\Gamma(t=0)} \sum_{i=1}^{n} p_i \, dq_i = \oint_{\Gamma(t=T)} \sum_{i=1}^{n} p_i \, dq_i, \tag{9}$$

where Γ is a closed curve surrounding a tube of phase space trajectories and

evolving with the flow. That is here the sum of the projections of the area encircled by Γ on the (p_i, q_i) phase planes is conserved.

In order to preserve the Hamiltonian structure the integration scheme should be a discrete mapping of the phase space that also preserves the area in the above sense. Such mappings are called *canonical* or *symplectic* transformations. One of the simplest examples for symplectic integrator is the *midpoint rule*

$$\mathbf{x}^{n+1} = \mathbf{x}^n + h_{n+1}\,\mathbf{F}\left(\frac{1}{2}(\mathbf{x}^n + \mathbf{x}^{n+1}); t_n + \frac{1}{2}h_{n+1}\right), \tag{10}$$

where $\mathbf{F}$ is the right-hand side of the ODE and the superscript n denotes the approximated value of the prognostic variable $\mathbf{x}$ at the n-th time step with h being the length of a time step. Using the notation of Section 2, for a Hamiltonian system, $\mathbf{F}$ is the right-hand side of Eq. (3), i.e. $\mathbf{F} = \mathbf{D}\nabla H$. A number of Runge-Kutta methods are also symplectic, provided that their coefficients satisfy certain conditions. A detailed account of symplectic numerical methods can be found in *Sanz-Serna* and *Calvo* (1994). The interested reader can find further surveys of the progress made in the field of symplectic integrators e.g. in *Yoshida* (1993) or in *McLachlan* and *Scovel* (1993).

The difference between "normal" and symplectic methods is demonstrated on *Fig. 2*. The pendulum equation (Eq. (1)) was integrated with the explicit Euler method (Fig. 2a, d), with the implicit Euler scheme (Fig. 2b, e) and with the midpoint rule (Fig. 2c, f). The points of a circle of radius 0.3, centered at $p = 1$, $q = 0$ in the (p,q) phase plane were chosen as initial conditions. $h = \frac{\pi}{12}$ was used as time step. On Fig. 2a-c the computed points are plotted after every third time steps up to the *27*th step, while on the panels Fig. 2d-f the trajectories of the point $p = 1$, $q = 0.3$ are followed, plotting the points after each time step up to the *60*th. Although the closed curve (initially a circle) should keep its area along the phase space flow generated by Eq. (1), the Euler methods fail to preserve this property. In addition they turn the fix point $(p,q) = (0,0)$ into stable and unstable spiral points (see Fig. 2e, Fig. 2d, respectively) unknown in Hamiltonian systems. Note that the circle of initial points can be regarded as an ensemble describing the uncertainty in the position of the initial point. It is clear, that as the ensemble is propagated along the flow it is not only reflecting the varying nature of the probability density function of the initial uncertainty, but is highly influenced by model errors, as well. The main problem with these errors that they are systematic and thus cannot be simulated as a result of random processes. This fact suggests that handling of model errors in an ensemble prediction system is a highly nontrivial problem.

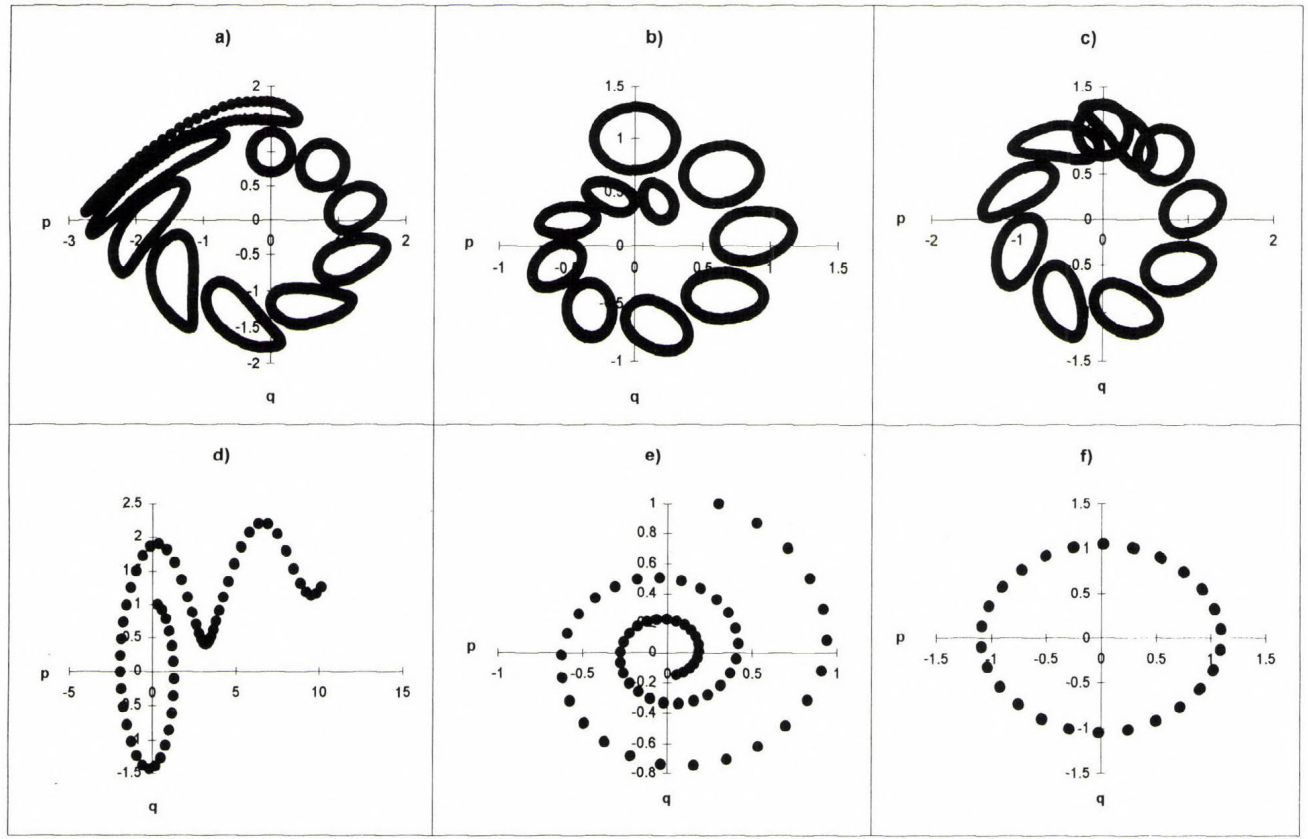


Fig. 2. The pendulum equation integrated by the explicit (*a, d*), implicit (*b, e*) Euler methods and with the midpoint rule (*c, f*). See also text in Section 4.1.

4.2 Conservation of energy

Since the invariants of motion constrain the flow into a smaller dimensional subspace of the phase space, it is assumed that the more invariants an integrator can preserve the better it is. Numerous schemes have been invented that conserve one or more invariants, e.g. energy, potential enstrophy, etc. These schemes, however, usually destroy the Hamiltonian structure of the flow. The question naturally arises, whether symplectic integrators can preserve the energy and/or other invariants. According to the theorem of *Ge* and *Marsden* (1988) *a symplectic method cannot conserve energy exactly for a general nonintegrable Hamiltonian system*. Even for an integrable system the preservation of the Hamiltonian is possible only if the time map of the scheme is identical with that of the true flow. Since, in general, the preservation of energy excludes the preservation of structure, and vice versa, a necessary compromise should be reached in case of every concrete problem at hand. In case of nonlinear wave equation (see *McLachlan*, 1994), for instance: Either one applies a symplectic scheme that conserves the structure (i.e. the shape) of the travelling waves, but then its speed will be in error, or one insists on the conservation of energy and uses an energy conserving scheme that models the speed of the wave correctly, but gives erroneous wave shape. Conservation of energy reduces the dimensionality of the problem only by one. It seems plausible, that the symplecticness, which distinguishes the class of Hamiltonian systems within that more general class of energy conserving ones, is a somewhat more distinctive characteristic of the Hamiltonian flows.

The fact that energy conservation and symplecticness are conflicting requirements of a scheme does not imply that the symplectic methods are completely wrong in terms of energy errors. In fact, it can be proved, that *for any symplectic integrator applied to an autonomous Hamiltonian problem, a modified autonomous Hamiltonian system can be found so that the numerically computed points are either exactly or "very nearly" on the trajectories of this modified system* (see e.g. *Sanz-Serna* and *Calvo*, 1994, 133 p.). Since the structure of a symplectic integrator is perfect within a negligible roundoff error this modification means a perturbation of the Hamiltonian, i.e. the model errors are encapsulated into the energy error. This perturbation leads to an initial jump in the energy of the model, but the energy error cannot grow without limit over any number of time steps. Indeed, the energy error can be reduced to the roundoff level increasing the order of the integrator and decreasing the time step. On the other hand the initial jump in the energy is an obvious model bias, the forecast will always be burdened with errors, even if we know the true governing equations and the initial conditions. The initial jump of the energy is a well known phenomenon in NWP and is generally described as a result of the equilibration process between forcing and dissipative terms. While the contribution of these terms is obvious one should keep in mind that the above

adjustment process is a must even if the source and the sink terms are all switched off.

Consider again the simple example of the pendulum. *Fig. 3* shows the "true" trajectory of the pendulum emanating from the initial point $p = 1.41615$, $q = 0.0$ (solid line) and points (small diamonds) that were computed with a symplectic Runge-Kutta scheme ($p^{n+1} = p^n - h \sin q^{n+1}$, $q^{n+1} = q^n + hp^n$). It can be shown (*Sanz-Serna* and *Calvo*, 1994) that the scheme preserves to the second order the modified Hamiltonian $H_2^h = \frac{p^2}{2} - \cos q + \frac{h}{2} p \sin q$. Since the trajectory corresponding to the approximately conserved modified Hamiltonian is also a closed curve the energy error cannot grow without limit. The facts that the energy errors remain bounded even over extremely long integrations and that it can be further reduced to arbitrarily small amount using higher order schemes and shorter time steps, are great benefits of symplectic integrators. On the other hand, the perfect conservation of structure and a good preservation of energy cannot guarantee any error bound for the individual variables of a nonintegrable chaotic system.

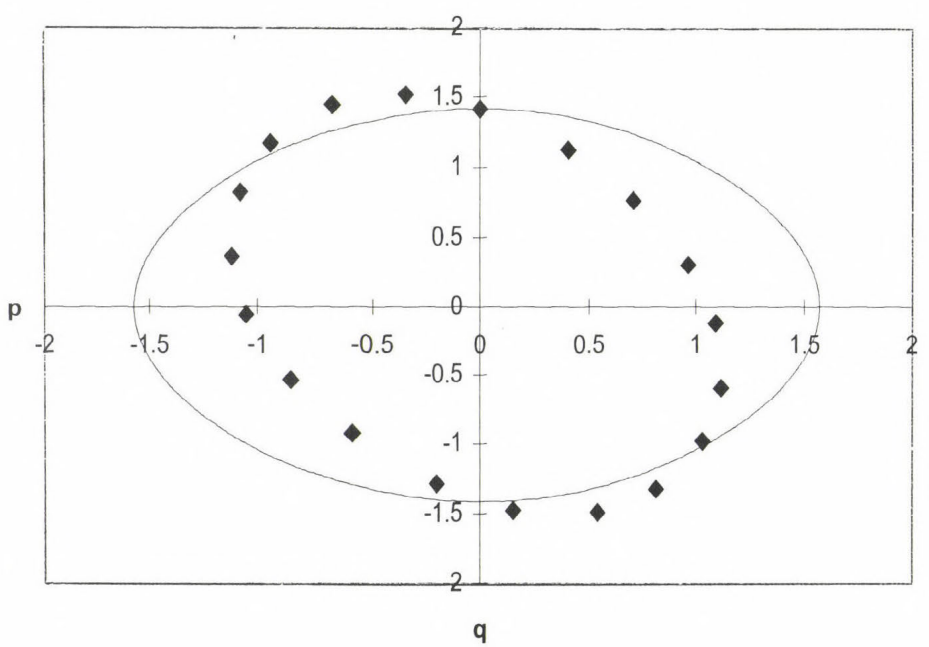


Fig. 3. The true trajectory (solid line) of the pendulum from initial point $p = 1.41615$, $q = 0.0$ and points (diamonds) computed with a symplectic Runge-Kutta method from the same initial condition.

5. *Numerical experiments*

In this section some numerical results will be presented to illustrate the above discussed features of symplectic integrators. In the experiments summarized below the structure preserving spatial truncation of the two-dimensional vorticity equation (hereafter 2DV) (*Zeitlin*, 1991) and that of the two-layer quasi-geostrophic equations (denoted by L2 QG) (*McLachlan et al.*, 1997) were used. A comparative study of the structure preserving truncation and a traditional spatial discretization applied to these equations will be detailed in the second part of this paper. For the time integration a slightly modified version of McLachlan's explicit symplectic integrator was used. The integration scheme and its modification is briefly described in the Appendix. A detailed description can be found in *McLachlan* (1993) or in *McLachlan et al.* (1997).

In the description of the experiments we refer to the resolution with TM, where M is a positive integer, and it implies that a cutoff wavenumber of M is used, that is, spectral modes associated with wave vectors $\mathbf{k} = (k_1, k_2)$, $|k_1|$, $|k_2| \leq M$ are considered in the truncated system.

Since the spatially truncated system possesses a Hamiltonian structure and the time integration scheme is symplectic, the preservation of structure is guaranteed in the integration. As a consequence, the Casimir invariants related to the structure matrix (e.g. the enstrophy) can be shown to be conserved within the error limit of number representation. On the other hand, if the system is nonintegrable, the conservation of the Hamiltonian (i.e. the energy) and of those invariants that are associated with its symmetries are inevitably burdened with errors. Thus a straightforward method to test the skill of a symplectic scheme is to examine the behavior of the energy error.

The relative energy error, $(H(t) - H(t=0))/H(t=0)$, was calculated for model runs started from randomly generated initial conditions at different resolutions. The integrations were carried out for 10^6 time steps at T1 and T2, and for 10^5 time steps at T13. In the first series of experiments first order integration scheme was used and the time step was uniformly set to 1.0. Taking the reciprocal of the average (potential) vorticity $1/\bar{\zeta}$ sec as time unit, the unit time step corresponds to $\bar{\zeta}$ nondimensional time unit. A similar time unit for the atmosphere is 10^5 sec, thus $\bar{\zeta} \times 10^5$ sec is a fair estimate of the time steps used in the different size systems. In the numerical experiments the initial vorticity field for the 2DV model were generated randomly with averages 0.09796 1/sec, 0.06091 1/sec and 0.00801 1/sec at T1, T2 and T13, respectively. With these values the unit time step corresponds to an atmospheric equivalent time step of about 160 min, 102 min and 13 min at the given resolutions. The average potential vorticity in the initial conditions generated for the L2 QG model was of the magnitude 10^{-2} 1/sec at all three resolutions, thus the unit time step in these cases agrees to an atmospheric equivalent of about 15 min.

In the 2DV runs the maximum energy error was in the magnitude of 10^{-1} at T1, 10^{-2} at T2, and 10^{-1} at T13. These significant errors are due to the fact that the time step is too large for randomly generated fields. Reduction of the time step to its hundredth resulted in relative errors two magnitudes smaller. Another way of reducing the energy errors is to apply a higher order integration scheme. Thus in the next row of experiments the second order version of McLachlan's time integrator was applied. The relative energy errors obtained with 1.0 length time steps are plotted on *Fig. 4* at T1, T2 and T13. The magnitude of the relative errors are of 10^{-2} at T1, 10^{-4} at T2 and 10^{-4} at T13. Reduction of the time steps to 0.01 results in further decrease in the errors, the magnitude of maximum relative energy errors being 10^{-7} at T1, 10^{-8} at T2 and 10^{-7} at T13.

The relative energy errors in the L2 QG runs with the first order scheme and 1.0 time steps were 10^{-3} at T1, T2 and also at T13. Using hundred times smaller time steps gave $O(10^{-5})$ relative energy errors. Second order integration in time gave relative errors in the magnitude of 10^{-6} and 10^{-5} at T1 and T2, and of 10^{-4} at T13 when the time step was set to 1.0, and of 10^{-8}, 10^{-9} and 10^{-8} at T1, T2 and T13, respectively, when the time steps were 0.01.

Using the second order scheme implied only doubling of the number of operations and was as effective as reducing the time step which required significantly more CPU time to perform the integration over the same period of time (hundred times more time steps had to be performed). The advantage of using higher and higher order schemes, however, is not a general rule since further increase of order leads to a quadratic drop in the computational efficiency, while the same factor for reducing the time step is only linear.

An important feature of the relative energy errors plotted on the panels of Fig. 4 is that in all three cases the relative error of energy oscillates around a nonzero value, but remains bounded. This is also true for the relative energy errors in the L2 QG model (not shown). This is the manifestation of the fact that the symplectic scheme conserves the energy of an approximating or perturbed Hamiltonian system. The perturbed Hamiltonian of a higher order scheme is closer to the "true" energy level than that of a lower order scheme. This is demonstrated on *Fig. 5a-b*, where the relative error of energy for the 2DV model is plotted at T2, calculated with first (Fig. 5a) and second order (Fig. 5b) schemes from the same random initial conditions. Note that the average relative perturbation of the Hamiltonian (the value around which the errors oscillate) are of opposite sign on these plots.

Fig. 5c-d show a similar pair of energy error curves. These plots were obtained with the 2DV model at T13 resolution, with first and second order scheme. The initial conditions were the discrete approximation of a rotating cone placed in the vorticity field. That is the initial field in this experiment was a discrete Fourier transform of a rotating cone of radius ρ defined in the physical space as

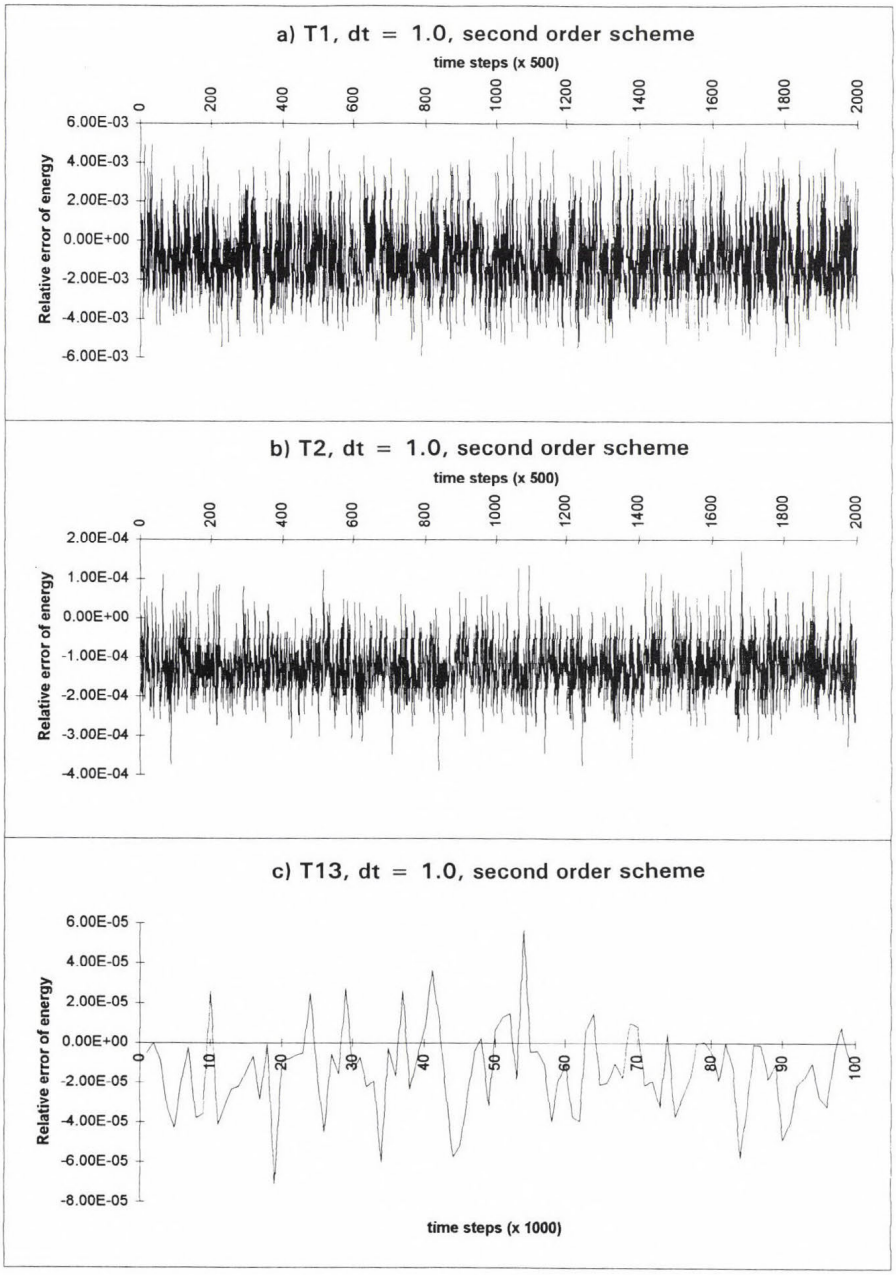


Fig. 4. Relative energy errors at T1 (*a*), T2 (*b*) and T13 (*c*) obtained by integrating the 2DV model with second order symplectic scheme and time step 1.0.

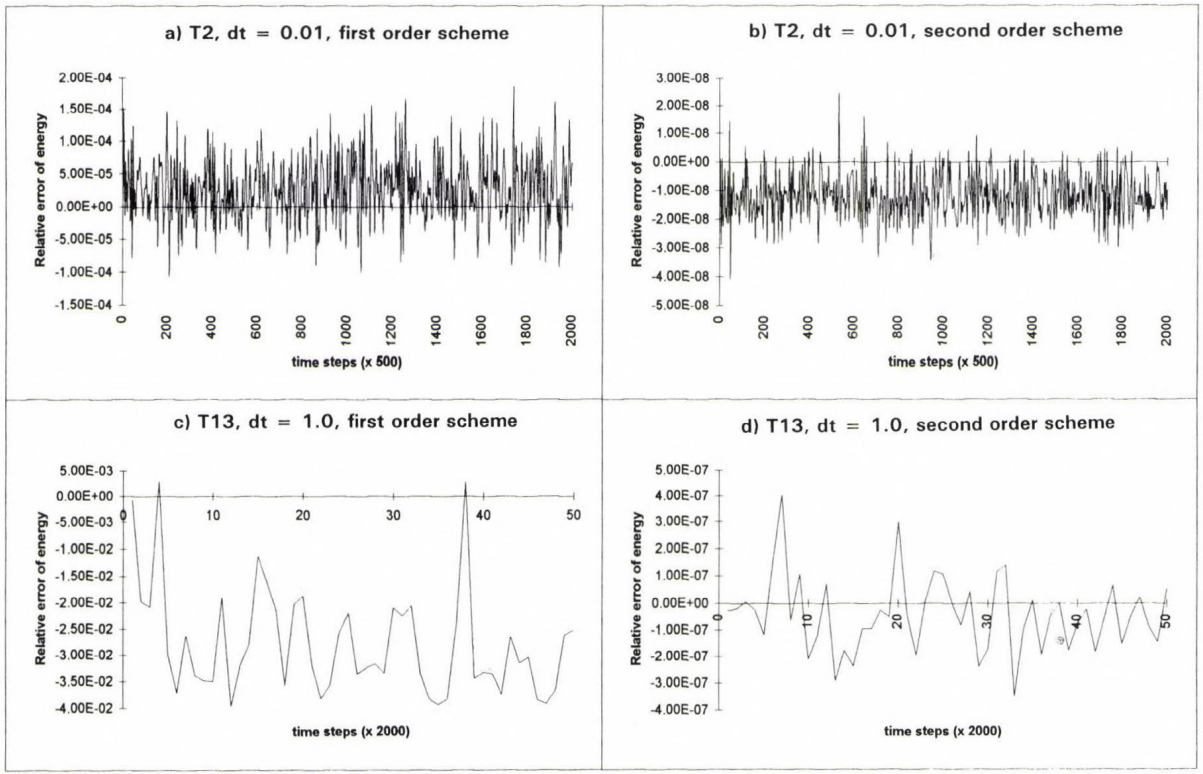


Fig. 5. Panels *a, b* show the relative energy errors obtained by integrating the 2DV model at T2 with the first (*a*) and second (*b*) order scheme started from random initial conditions. Panels *c, d* present the relative energy at T13 with first (*c*) and second (*d*) order scheme when the initial condition corresponds to the spectral transform of a cone placed in the vorticity field.

$$\zeta(\mathbf{r},t=0) = \begin{cases} 1 - \frac{|\mathbf{r}|}{\rho} & \text{for } |\mathbf{r}| \le \rho \\ 0 & \text{for } |\mathbf{r}| > \rho \end{cases}, \tag{11}$$

where $|\mathbf{r}| = ((r_1 - r_{01})^2 + (r_2 - r_{02})^2)^{\frac{1}{2}}$ and $\mathbf{r}_0 = (r_{01}, r_{02})$ is the place of the rotation axis of the cone. (Here $\mathbf{r}$ denotes vector of place in the physical domain.) This is a stationary solution of the continuous equation, thus a proper test also for the accuracy of the models (see Part II of this paper). In the numerical experiment $r_{01} = 3$, $r_{02} = 7$ and $\rho = 3$ were chosen and unit time steps were used. The relative error of energy is plotted at every 2,000th time steps up to the 100,000th time step on Fig. 5c-d. An interesting feature of numerical models can be observed on the relative energy error plots of the first order run. There are some "jumps" in the relative energy error at the beginning of the integration interval before a steady oscillation begins. Presumably, because the system "needs some time" to find that energy shell of the approximating Hamiltonian system, on which it remains afterwards (*McLachlan*, 1996). Evidently, the use of a higher order scheme resulted in the improvement of the energy conservation with orders of magnitude, and at the same time the initial "jumps" almost vanished indicating that the approximating Hamiltonian is much "closer" to that of the truncated system.

6. Summary

The "perfect model" hypothesis is a widely used presumption in atmospheric dynamics and in numerical weather prediction that makes possible to deal with the problem of initialization with the tools of dynamical systems theory. On longer terms, however, model errors tend to dominate. Their sources and effects are more complex than that of the errors in the initial conditions since they include e.g. the parameterizations of physical and subgrid processes. There are no general methods to approach these heterogeneous sources of errors. The present review article (along with an accompanying paper) focuses on the systematic errors caused only by the temporal and spatial discretizations. We examined the effect of truncations on the algebraic — Hamiltonian — structure of conservative continuous equations. This restriction of attention to conservative systems is justified by the facts that Hamiltonian formalism has been payed more and more attention recently and a number of exciting theoretical results have been obtained in this framework (see Section 1 and 2). We gave a brief introduction to the Hamiltonian formalism in Section 3.

Given a spatially discretized system of equations with Hamiltonian structure, the temporal integration plays a crucial role in determining the phase space

behaviour of the flow. The timestepping algorithm inevitably alter the structure and/or the energy level of the flow. In Section 4 the reader can get acquainted with the concept of symplecticness, a characteristic feature of finite dimensional Hamiltonian flows, and with symplectic time integrators that preserve this property. The conservation properties of these schemes are briefly reviewed. The numerical results presented in Section 5 confirm the theory that the energy conservation of symplectic integrators is remarkable, even though it is impossible to conserve the exact energy of a nonintegrable system. Despite of the attractive results in terms of energy conservation, we abstain ourselves from drawing final conclusions about the applicability of these schemes until the end of the second part of this paper.

Acknowledgements—This work was partly supported by a grant from State Scientific Research Fund (OTKA F 019268). Encouraged by the informal discussions great progress was achieved during the authors' stay at the Isaac Newton Institute for Mathematical Sciences, Cambridge, England, within the program "Mathematics of the Atmosphere and Ocean Dynamics" in the summer and autumn of 1996. The critical comments of the anonymous referees on the previous version of the manuscript are greatly appreciated.

References

Arnold, V.I., 1989: *Mathematical Methods in Classical Mechanics*. Springer-Verlag, New York, 508 pp.

Bishop, C.H. and *Toth, Z.*, 1996: Using ensembles to identify observations likely to improve forecasts. *Contributions to the 11th AMS Conference on Numerical Weather Prediction*, Aug. 19-23, Norfolk, Virginia, 72-74.

Bokhove, O. and *Shepherd, T.*, 1996: On Hamiltonian balanced dynamics and the slowest invariant manifold *J. Atmos. Sci. 53*, 276-297.

Camassa, R., 1995: On the geometry of the atmospheric slow manifold. *Physica D 84*, 357-397.

Camassa, R. and *Tin, S.*, 1996: The global geometry of the slow manifold in the Lorenz-Krishnamurthy model. *J. Atmos. Sci. 53*, 3251-3264.

Errico, R.M., 1984: The statistical equilibrium solution of a primitive-equation model. *Tellus 36A*, 42-51.

Ge, Z. and *Marsden, J.E.*, 1988: Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators. *Phys. Lett. A 133*, 134-139.

Götz, G., 1994: Application of nonlinear dynamics in atmospheric sciences. *Időjárás 98*, 65-86.

Götz, G., 1995: Predictability of nonlinear dynamical systems. *Időjárás 99*, 1-32.

Holton, J.R., 1992: *An Introduction to Dynamic Meteorology*. 3rd edition. Academic Press, San Diego, 507 pp.

Leith, C.E., 1980: Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci. 37*, 958-968.

Lorenz, E.N., 1960: Maximum simplification of dynamic equations. *Quart. J. Roy. Meteorol. Soc. 12*, 243-254.

Lorenz, E.N., 1980: Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci. 37*, 1685-1699.

Lorenz, E.N., 1986: On the existence of a slow manifold. *J. Atmos. Sci. 43*, 1547-1557.

Lorenz, E.N. and *Emmanuel, K.A.*, 1997: Optimal sites for supplementary weather observations: Simulation with a small model. Submitted to *J. Atmos. Sci.*

Lorenz, E.N. and *Krishnamurthy, V.*, 1987: On the nonexistence of a slow-manifold. *J. Atm. Sci. 44*, 2940-2950.

Lynch, P., 1996: *The Elastic Pendulum: A Simple Mechanical Model of Atmospheric Balance*. The Irish Meteorological Service Technical Note, No. 54. The Irish Meteorological Service.

Mackay, R.S., 1992: Some aspects of the dynamics and numerics of Hamiltonian systems. In *The Dynamics of Numerics and the Numerics of Dynamics* (eds.: *D.S. Broomhead* and *A. Iserles*). Clarendon Press, Oxford, pp. 137-194.

McLachlan, R., 1993: Explicit Lie-Poisson integration and the Euler equations. *Phys. Rev. Lett. 71*, 3043-3046.

McLachlan, R., 1994: Symplectic integration of Hamiltonian wave equations. *Numer. Math. 66*, 465-492.

McLachlan, R.I. and *Scovel, C.*, 1993: A survey of open problems in symplectic integration. *Proc. Conf. Integration Algorithms in Classical Mechanics* (ed.: *G. Patrick*). Under review.

McLachlan, R.I., *Szunyogh, I.* and *Zeitlin, V.*, 1997: Hamiltonian finite-dimensional models of baroclinic instability. *Phys. Lett. 229A*, 299-305.

McLachlan, R.I., *Quispel, G.R.W.* and *Turner, G.S.*, 1996: Numerical integrators that preserve symmetries and reversing symmetries. Under review.

Olver, P.J., 1989: *Applications of Lie Groups to Differential Equations*. Springer-Verlag, New York, 513 pp.

Ott, E., 1994: *Chaos in Nonlinear Dynamical Systems*. Cambridge University Press, Cambridge, 385 pp.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics*. Second edition. Springer-Verlag, New York, 710 pp.

Pires, C., *Vautard, R.* and *Talagrand, O.*, 1996: On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus 48A*, 96-121.

Quispel, G.R.W., 1995: Volume-preserving integrators. *Phys. Lett. A 206*, 26-30.

Sanz-Serna, J.M., 1992: Numerical ordinary differential equations *vs.* Dynamical systems. In *The Dynamics of Numerics and the Numerics of Dynamics* (eds.: *D.S. Broomhead* and *A. Iserles*). Clarendon Press, Oxford, pp. 81-106.

Sanz-Serna, J.M. and *Calvo, M.P.*, 1994: *Numerical Hamiltonian Problems*. Chapman & Hall, London, 207 pp.

Shepherd, T.G., 1990: Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Adv. Geophys. 32*, 287-337.

Tabor, M., 1989: *Chaos and Integrability in Nonlinear Dynamics*. John Wiley & Sons, New York, 364 pp.

Yoshida, H., 1990: Construction of higher order symplectic integrators. *Phys. Lett. A 150*, 262-268.

Yoshida, H., 1993: Recent progress in the theory and application of symplectic integrators. *Cel. Mech. Dyn. Astr. 56*, 27-43.

Zeitlin, V., 1991: Finite-mode analogs of 2D ideal hydrodynamics: Coadjoint orbits and local canonical structure. *Physica D 49*, 353-362.

Appendix

1. McLachlan's algorithm for the integration of the 2D vorticity equation

The algorithm, whose detailed description can be found in *McLachlan* (1993), is a realization of the so called *composition method*, in which the vector field defined by the right-hand side of Eq. (3) is decomposed into vector fields, that are (more) easily integrated by suitable explicit methods.

Hereafter the vector of prognostical variables will be denoted with ζ, whose components $\zeta_{\mathbf{i}}$ are the spectral coefficients of vorticity associated with the wave vector $\mathbf{i}$. The index vectors are from the finite lattice $\{\mathbf{i} \,|\, |i_1|, |i_2| \leq M\}$, where M is the cutoff wavenumber. N denotes $2 \cdot M + 1$. The algorithm described in *McLachlan* (1993) is applicable when N is prime.

The algorithm is based on the splitting of the Hamiltonian $H = \sum_{\mathbf{k}} H_{\mathbf{k}}$, where $H_{\mathbf{k}} = \frac{1}{2} \sum_{n=0}^{N-1} \frac{\zeta_{n\mathbf{k}} \zeta_{-n\mathbf{k}}}{|n\mathbf{k}|^2}$, with $\mathbf{k} \in \mathbf{K} = \{(0,1)\} \bigcup \{(1,m) \,|\, 0 \leq m < N\}$ and the indices are taken modulo N.

McLachlan's observation is that this splitting of the Hamiltonian reduces the problem of integrating the truncated spectral vorticity equation to the solution of a set of linear differential equations with constant coefficients, each corresponding to a vector field $\mathbf{D}\nabla H_{\mathbf{k}}$, with $\mathbf{k} \in \mathbf{K}$. Let denote $\exp(dt \cdot \mathbf{D}\nabla H_{\mathbf{k}})$ the resolvent of these linear differential equations. A first order approximation in the time step dt of the solution of Eq. (3), i.e. that of the equation $\dot{\zeta} = \mathbf{D}\nabla H = \sum_{\mathbf{k}} \mathbf{D}\nabla H_{\mathbf{k}}$ is $\zeta(t) = \prod_{\mathbf{k}} \exp(dt \cdot \mathbf{D}\nabla H_{\mathbf{k}}) \zeta(t=0)$ (see e.g. *McLachlan*, 1993). While

$$\zeta(t) = \exp\left(dt/2 \cdot \mathbf{D}\nabla H_{\mathbf{k}_1}\right) \ldots \exp\left(dt/2 \cdot \mathbf{D}\nabla H_{\mathbf{k}_{N+1}}\right)$$
$$\times \exp\left(dt/2 \cdot \mathbf{D}\nabla H_{\mathbf{k}_{N+1}}\right) \ldots \exp\left(dt/2 \cdot \mathbf{D}\nabla H_{\mathbf{k}_1}\right) \zeta(t=0),$$

where $\mathbf{k}_i \in \mathbf{K}$, gives a second order (in dt) approximation to the solution of Eq. (3) (*McLachlan*, 1996).

The vector field $\mathbf{D}\nabla H_{\mathbf{k}}$ defines an explicitly integrable system of equations because each component of the prognostic variable vector can be written as $\zeta_{\mathbf{j}+m\mathbf{k}}$, where $-M \leq m \leq M$ and $\mathbf{j}$ is an appropriate transition vector of the indices. Using the notation $z_m = \zeta_{\mathbf{j}+m\mathbf{k}}$ the evolution of the vector z can be given with the set of linear differential equations of $\dot{z}_m = \sum_{n=-M}^{M} a_n z_{m-n}$, the discrete Fourier transform of which gives an explicitly integrable system $d\tilde{z}/dt = \Lambda \tilde{z}$, where $\tilde{z} = \boldsymbol{F} z$, $\Lambda = diag(\boldsymbol{F}\mathbf{a})$, $\boldsymbol{F}$ being the matrix of the discrete Fourier transformation.

Thus the algorithm for integrating the 2D vorticity equation consists of the following steps:

(i) for $\mathbf{k} \in \mathbf{K}$ do,

(ii) for $\mathbf{j} = 1\text{st}, \ldots, M\text{th}$ translation of $\mathbf{k}$ do,

(iii) with $z_m = \zeta_{\mathbf{j}+m\mathbf{k}}$ set $z = \boldsymbol{F}^{-1} e^{\Lambda \cdot dt} \boldsymbol{F} z$, where $\Lambda = diag(\boldsymbol{F}\mathbf{a})$,

(iv) set $\zeta_{-(\mathbf{j}+m\mathbf{k})} = z_m^*$,

(v) end do,

(vi) end do.

2. Modifications to McLachlan's algorithm

In order to be able to carry out effective numerical integrations with the use of available FFT routines, we extended the algorithm to certain non-prime Ns. Our modifications were based on the following observations:

(i) If N is prime, the only common element in the sets $\sigma_{\mathbf{k}} = \{n\mathbf{k} \ mod N \mid 0 \leq n < N\}$ with $\mathbf{k} \in \mathbf{K}$ is the $\mathbf{0} = (0,0)$ index and $\bigcup_{\mathbf{k} \in \mathbf{K}} \sigma_{\mathbf{k}}$ covers the entire lattice.

(ii) If N is non-prime, then the $\sigma_{\mathbf{k}}$-s have more common elements — consequently they do not cover the entire lattice —, because the equation $s = nr \ mod N$ can have more than one solution for a given gridpoint (s,r).

(iii) If $N = p^l$, with p prime, the $\sigma_{\mathbf{k}}$-s cover the entire lattice if $\mathbf{k} \in \mathbf{K}'$, where $\mathbf{K}' = \mathbf{K} \bigcup \{(n,1) \mid 1 < n < N \text{ and } [n,N] > 1\}$ ([.,.] denotes the largest common factor).

It follows from the above statements that McLachlan's algorithm can be applied in $N = p^l$ cases as well, only the splitting of the Hamiltonian must be modified in the following way: $H = \sum_{\mathbf{k} \in \mathbf{K}'} H_{\mathbf{k}}$, where $H_{\mathbf{k}}(\mathbf{m}) = \dfrac{H(\mathbf{m})}{\chi(\mathbf{m})}$, where $\chi(\mathbf{m})$ gives the number of $\sigma_{\mathbf{k}}$ sets that include the index $\mathbf{m}$.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 1, January–March 1998, pp. 43–52

Interpolation of bivariate functions in connection with isoline construction problem

B. G. Vager[1] and **N. K. Serkov**[2]

[1]*Department of Calculating Mathematics,*
St.Petersburg State University of Architecture and Civil Engineering,
198005, 2nd Krasnoarmeiskaya, 4, St. Petersburg, Russia
E-mail: BVager@compmath.abu.spb.ru

[2]*State Hydrological Institute,*
199053, 2nd liniya, V.O., 23, St. Petersburg, Russia

(Manuscript received 2 January 1997; in final form 6 October 1997)

Abstract—The description and critical comparative analysis of the main methods of two-dimensional interpolation for scattered data are given. The new method of interpolation from regular rectangular grid founded on the quadratic B-splines is suggested. The results of numerical experiments are announced.

Key-words: interactive surface tensioning, interpolation, isoline, knot, kriging, scattered data, spline, smoothing.

1. Formulation of the problem

One of the actual problems in applied mathematics which is important in geophysics is the construction of isoline maps. From mathematical point of view the problem is reduced to interpolation of functions $f(x,y)$ whose values f_i are known in given scattered points P_i $(i = 1,\ldots,N)$.

The interpolant S from some given set W is usually required to be smooth, converging to f if the distance among grid points is decreasing to zero. Moreover the physical reasons involve that S should possess "locality", i.e. its value $S(x,y)$ in every point (x,y) should depend on a few number of values f_i in some neighboring grid points only.

The interpolation may be exact if

$$S(P_i) = f_i (i = 1,\ldots,N), \tag{1}$$

or with smoothing if condition (1) does not hold. The latter case usually occurs when the values f_i are the results of measurements which were done with some remarkable errors.

The well-known way to construct isoline maps is to interpolate given values to the grid points of regular rectangular grid first and then to interpolate in an arbitrary point (x,y) by means of one simple method of interpolation, by bicubic splines for example. This is a possibility to achieve a calculative efficacy.

In this paper we consider some methods of two-dimensional interpolation from scattered points in the (x,y)-plane and one new method of interpolation from a regular rectangular grid to an arbitrary point (x,y).

2. Optimal interpolation

In this method the values $f_i(i = 1,\ldots,N)$ are considered as the realization of some stochastic process f. Let stochastic process f be stationary in wide sense, i.e. a mean value $\mathbf{E}f(\mathbf{u})$ is independent from $\mathbf{u}$ and the variance

$$Var[f(\boldsymbol{u}) - f(\boldsymbol{v})] = 2g(\boldsymbol{h}) = E[(f(\boldsymbol{u}) - f(\boldsymbol{v}))^2] \tag{2}$$

depends on the difference vector $\mathbf{h} = \mathbf{u} - \mathbf{v}$ only. If the variance depends on the length $h = |\mathbf{h}|$ only, the stochastic field is called to be isotropic.

Let us formulate the problem of optimal interpolation, or the best linear unbiased estimation (*Kolmogorov*, 1941):

for $P_0 \in R^2$ we have to find weights $p_i \in R(i = 1,\ldots,N)$ minimizing

$$\sigma^2 = \boldsymbol{E}[f(P_0) - S(P_0)]^2, \tag{3}$$

where $S(P_0)$ is the linear combination of known values $f_i = f(P_i)$:

$$S(P_0) = \sum_1^N p_i f_i. \tag{4}$$

A more general problem can also be considered. One may suggest that values f_i have a random errors Δf which are not correlated and with f, having the zero mean value and the measure of error

$$\eta = \mathbf{D}(\Delta f)/\mathbf{D}f, \tag{5}$$

where $\mathbf{D}$ is the sign of dispersion; $\eta = 0$ in the errorfree case.

This method was first used by the French geologist *Krige* (1951) (as referred in the name "kriging") in statistical geophysics, and by the Russian scientist *Gandin* (1963) in climatology. The latter used an autocorrelative function (ACF) $r(x)$ instead of the variogram $g(x)$:

$$r(x) = 1 - g(x)/\mathbf{D}f. \tag{6}$$

The unknown weights p_i in Eq. (4) satisfies the system of linear equations

$$\eta p_i + \sum_{j=1}^{N} \mu_{ij} p_j = \mu_{i0}, \tag{7}$$

where $\mu_{ij} = r(x_i - x_j)$, $(i = 1,\ldots,N;\ j = 0,\ldots,N)$.

The exactness of optimal interpolation may be described by the measure of its error (*Gandin*, 1963)

$$\epsilon = S^2/\mathbf{D}f, \qquad (0 \leq \epsilon \leq 1). \tag{8}$$

The value ϵ may be written as

$$\epsilon = 1 - \sum_{1}^{N} p_i \mu_{i0}. \tag{9}$$

It can easily be shown that

$$\epsilon + \delta = 1, \tag{10}$$

where $\delta = \mathbf{D}f/\mathbf{D}S$ and $\mathbf{D}S$ is the dispersion of the random variable $S(x_0)$.

The equation (10) is equivalent to

$$\mathbf{D}f = \sigma^2 + \mathbf{D}S. \tag{11}$$

This estimation of the smoothing effect of optimal interpolation should be taken into account in the cases where individual values of the interpolant are of interest and their decreasing is not desirable (see also *Kostyukov*, 1982).

Moreover it is suggested that the ACF (or the variogram) is known i.e. there is a sufficiently large number of measurements or/and some a'priori information.

3. *The method of inverse distance raised to a power*

In this method the values of the interpolated function are weighted in some grid points P_i nearest to the point (x,y) of interpolation with the weights which are inverse to a distance raised to the power a:

$$S(x) = \sum_{i=1}^{N} p_i f_i, \tag{12}$$

$$p_i = 1/\left(r_i^a \sum_{j=1}^{N} (1/r_j^a)\right), \quad (0 \leq a \leq 2), \tag{13}$$

where

$$r_j = ||(x,y) - P_j||. \tag{14}$$

Formula (12) involves that $lim\, p_j = \delta_{ij}$ (Kronecker's delta) if $||(x,y) - P_i|| \to 0$.

The value a regulates the influence of distances: the more is a the more quickly the influence of values of the function in the grid points is decreasing as the point of interpolation is being farer and farer from the grid points.

This method can be used as smoothing. In this case the weights are choosen as

$$p_i \approx 1/(r_i^a + b), \tag{15}$$

where $b > 0$ is an empirical parameter (*Kostyukov*, 1982).

The described method has one undesirable property: isolines near the grid points are often distorted ("bull's eye" effect).

Moreover, if there are some grid points which are situated close to each other on one side of the point of interpolation and only one grid point on the contrary side with the same distance, it is evident that "the heap overcomes". This fact is not in agreement with the main physical ideas.

Next method suggested by *Babaliev* (1973) is free from the above mentioned defect.

4. *The weighted anisotropic interpolation (WAI)*

In this manner the method was named by *Kostyukov* (1982). The weights in formula (4) satisfy the system of linear equations:

$$\sum_{i=1}^{N} \pi_j r_{ij} = r_{i0} \qquad (i = 1,...,N) \tag{16}$$

where $r_{ij} = ||P_i - P_j||$, $r_{i0} = ||P_i - (x,y)||$, $p_i = \pi_i/\Sigma\pi_i$.

So, condition (1) holds, and the reciprocal arrangement of data points is taken into account. When $N = 2$ the method is equivalent to the method of inverse distance with $a = 1$ and $b = 0$ in formula (15).

5. Shepard's method

The method is based on ideas of the least squares and inverse distance methods. The interpolant $S(x,y)$ is given in any point $(x,y) \in D$ by the formula

$$S(x,y) = \sum_{k=1}^{N} W_k(x,y) S_k(x,y) / \sum_{k=1}^{N} W_k(x,y), \tag{17}$$

where the nodal functions S_k given by

$$\begin{aligned} S_k(x,y) = {} & A_{1,k}(x - x_k)^2 + A_{2,k}(x - x_k)(y - y_k) + A_{3,k}(y - y_k)^2 \\ & + A_{4,k}(x - x_k) + A_{5,k}(y - y_k) + f_k, \qquad (k = 1,...,N), \end{aligned} \tag{18}$$

are the second degree polynomials of (x,y), which interpolate the data values f_k at grid points (x_k,y_k). The coefficients $A_{1,k}$, ..., $A_{5,k}$ are obtained by a weighted least squares fit to the closest NQ $(4 < NQ < \min\{41, N\})$ data points with weights similar to W_k which are taken to be

$$W_k(x,y) = [(R_k - D_k)_+ / (R_k D_k)]^2, \tag{19}$$

where $a_+ = \max(0,a)$, $D_k(x,y) = ||(x,y) - (x_k,y_k)||$ and R_k is the radius of the influence, which varies with k and is chosen so that the NW $(0 < NW < \min\{41, N\})$ data points are within the radius. Parameters NQ and NW are taken by the user. For a sufficiently large N the recommended values are $NW = 19$ and $NQ \leq 13$.

The value $W_k(x,y)$ is not defined by formula (19) at the grid point (x_k,y_k) but $S(x,y) \to f_k$ as $(x,y) \to (x_k,y_k)$, $(k = 1,...,N)$. It can be shown that $S \in C^1(D)$. A detailed description of the method is given in the paper *Algorithm 660* (1988).

6. *Interpolation by splines*

Let's formulate the problem of seeking for a spline $S \in W_2^2(D)$ which satisfies Eq. (1) and minimizing condition

$$\int_D \left[S_{xx}^2 + 2S_{xy}^2 + S_{yy}^2\right] dx\, dy = \min_{W_2^2(D)}. \tag{20}$$

It may be shown (see *Vasilenko*, 1983, p. 49–50), that this spline is given by

$$S(x,y) = 0{,}5\sum_{i=1}^{N} \lambda_i[(x - x_i)^2 + (y - y_i)^2] \ln[(x - x_i)^2 + (y - y_i)^2] + \nu_{00} + \nu_{10}x + \nu_{01}y, \tag{21}$$

where the unknown coefficients $\lambda_1,\ldots,\lambda_N$, ν_{00}, ν_{10}, ν_{01} satisfy the system of linear equations. This system is given in the monograph (*Vasilenko*, 1983).

The system has the only solution if there are at least three data points not lying on the same straight line. The more the value of N the more the condition number of the system is.

7. *Interpolation by multiquadrics*

Hardy (1971) suggested quadratic functions, or quadrics for modeling of smoothing surfaces. These functions are given by the equation

$$f_i(x,y) = [(x - x_i)^2 + (y - y_i)^2 + b]^{1/2}, \tag{22}$$

where b is the parameter taken by the user. If $b = 0$ the Eq. (22) defines a cone, else it defines a hyperboloid with vertical axes.

The general equation of the surface is the sum

$$H(x,y) = \sum_{i=1}^{N} c_i[(x - x_i)^2 + (y - y_i)^2 + b]^{1/2}. \tag{23}$$

Coefficients c_i are obtained from the system, which arises from interpolation conditions.

The surface constructed by this way is too smooth. The results are good only if data functions are varying slowly.

To obtain the coefficients c_i it is necessary to solve the system with dense matrix. For $N > 1000$ the problem is the storage of the matrix as well as the time of calculations.

8. The Italian method

In this method (*Montefusco* and *Casciola*, 1989) the global interpolant is constructed by smooth "pasting together" of partial interpolants, given in triangles which are defined by chosen triangulation of data point set. Interpolants are constructed provided minimum of some (energetic) functional involved interactive given parameter ("tension parameter"). It gives a strong tool to users to have influence on the approximative surface behavior and allows to smooth undesired oscillations near the points with high gradients.

The algorithm consists of three consecutive steps:

(1) Triangulation of the domain D — the convex hull of the set of data points with the tops of triangulation in these points (*Lawson*, 1977);

(2) Approximation of a first partial derivatives;

(3) Construction of the global interpolant (*Nielson*, 1980).

In step (2) there is possibility for interactive influence on the interpolant by means of the tension parameter to control the form of the interpolation surface and avoid undesired oscillations. The tension parameter is recommended to vary between sufficiently wide limits to achieve a marked effect.

An example for the use of the Italian method for different values of tension parameter, when data are natural, is given below (see *Fig. 1*).

9. Numerical experiments

The objective analysis of wind, geopotential, temperature and air pollution fields was carried out by *Kostyukov* (1982) on the base of extensive natural data. The author made use of least squares polynomial approximation, inverse distance raised to a power, optimal interpolation and WAI methods. Numerical experiments using independent data showed that the best results were received when the two latter methods were applied. WAI was preferred as it does not require a supplementary knowledge of statistical structure of meteorological fields to be interpolated.

We made use of the above mentioned methods and multiquadrics, Shepard, spline functions and Italian methods also. The numerical experiments were carried out on lake depth data. The analysis of results showed that in the case of sufficiently slowly varying functions and approximately regularly placed grid ponts, all of the methods in question led to similar results which were quite good. The analogous conclusion is fair for tests when the grid points are modeled randomly and the function to be interpolated is taken as smooth

analytic function. In real situations when the function to be interpolated has high gradients the interpolant has oscillations which are uncharacteristic of the initial data. The Italian method allows to avoid this defect by a suitable choosing of the tension parameter.

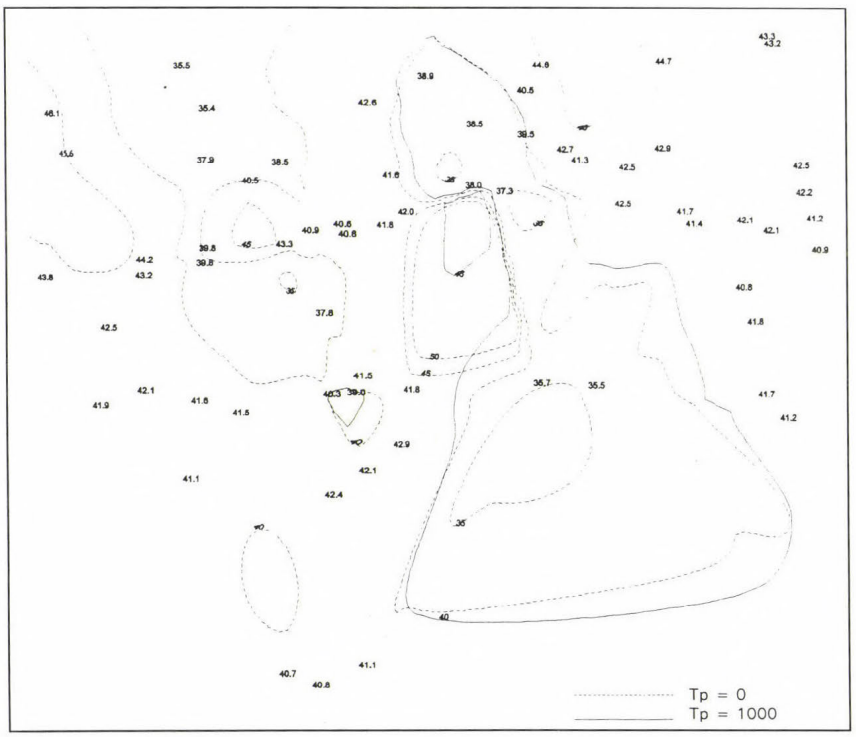


Fig. 1. Isolines of water depth field on the Kama river
(the numbers refer to the measurement points).

10. Interpolation when a grid is regular

In this case bicubic splines may be used for instance (*Vasilenko*, 1983; *Vager* and *Serkov*, 1987). The multiple solutions of the linear equation systems with high dimensional three-diagonal matrices are required to realize the above mentioned methods.

Powell (1977) formulated the interpolation problem of given function f when the data points are situated on a regular rectangular grid. Let the steps in axes x,y are d_x, d_y, respectively. We seek for bivariate function $S(x,y)$ such that

(1) $S(P_i) = f_i \qquad (i = 1,\ldots,N)$;
(2) $S \in C^1(D)$;
(3) $|S(x) - f(x)| = O(d^3)$, where $d^2 = d_x^2 + d_y^2$, if $f \in C^3(D)$;
(4) The property of locality (see the point **1** above) holds;
(5) Isolines $S(x,y) = const$ are easy to construct.

Powell (1977) suggested to construct the interpolant as a spline. His approach comes to a numerical solution of the system of the linear equations which is analogous to that arising from the finite differences solution of the Laplase equation.

We worked out a modification of the Powell's method. The value of the interpolant in each point (x,y) is given as a linear combination of the values f_i in 16 grid points nearest to the point (x,y). Coefficients of this linear combination are the quadratic splines given by **explicit** formulae, so there is no need to solve any systems. This approach allows to decrease the computing time essentially. The formula for $S(x,y)$ has the form

$$S(x,y) = \sum_{1=1}^{4} \sum_{k=1}^{4} \sigma_{i-2+1}(x)\, \sigma_{j-2+k}(y) f_{i-2+1,\, j-2+k}, \tag{24}$$

where if $x \in [x_i, x_{i+1}]$, for instance

$$\sigma_{i-1}(x) = 3\lambda^2/4 - \lambda/2\lambda, \quad \text{if} \quad 0 \le \lambda \le 1/2,$$
$$\sigma_{i-1}(x) = -\lambda^2/4 + \lambda/2\lambda - 1/4, \quad \text{if} \quad 1/2 \le \lambda \le 1,$$

where $\lambda = (x - x_i)/d_x$.

It can be showed that all conditions (1)–(5) are held.

11. Aims of further research

(1) To find the more expedient domains of using of each above mentioned interpolation method on the base of extensive numerical experiments. To set in accordance the various literatural recommendations which are often contradictional. For instance, *Kravchenko* (1984) informs that the using of splines is effective when $N < 150$. The authors of the well-known package of applied programs SURFER recommend $N > 1000$ on the contrary.
(2) To work out an effective method for approximation of bivariate functions in the case of the domain D is not one-connective (it means that there are a number of "holes" within D i.e. a lake with islands). The known methods usually work as there is no difference between "holes" and domains with a sparse arrangement of data points.

References

Algorithm 660, 1988: Collected algorithms from ACM. *ACM Trans. on Mathematical Software 14*, No. 2, 149.

Babaliev, A.M., 1973: About one method of functions of many independent variables (in Russian). *Computer Graphics and Its Application*. Novosibirsk, 118-121.

Gandin, L.S., 1963: *Objective Analysis of Meteorological Fields* (in Russian). Gidrometeoizdat, Leningrad.

Hardy, R., 1971: Multiquadric equation of topography and other irregular surfaces. *J. Geophisical Res. 78*, 1905-1915.

Kolmogorov, A.N., 1941: Interpolation and extrapolation of stationary random sequences (in Russian). *Izv. AN SSSR, Ser. Mathem. 5*, No. 1, 3-14.

Kostyukov, V.V., 1982: *Objective analysis and adjusting of meteorological fields* (in Russian). MO Gidrometeoizdata, Moscow.

Kravchenko, Yu.A., 1984: Methods of modeling of topography surfaces (in Russian). *Review information CNIIGAiK*, Moscow.

Krige, D.G., 1951: A statistical approach to some basic mine problems on the Witwatersrend. *J. Cem. Metall Min. Soc. S. Afr. 52*, 119-139.

Lawson, C.L., 1977: Software for C1 surface interpolation. In *"Mathematical Software III"* (ed.: *J.R. Rice*). Academic Press, Orlando, Fla., pp. 161-194.

Montefusco, L.B. and *Casciola, G.*, 1989: Algorithm 677. C^1 Surface interpolation. *ACM Trans. Math. Soft. 15*, No. 4, 365-377.

Nielson, G.M., 1980: Minimum norm interpolation in triangles. *SIAM J. Numer. Anal. 17*, No. 1, 44-62.

Powell, M.J.D, 1977: Numerical methods for fitting functions of two variables. *State Art Numer. Anal. Proc. Conf.*, New York, 1976, 563-604.

Vasilenko, V.A., 1983: *Spline Functions: Theory, Algorithms, Programs* (in Russian). Nauka, Novosibirsk.

Vager, B.G. and *Serkov, N.K.*, 1987: *Spline Methods in Applied Meteorology and Hydrology* (in Russian). Gidrometeoizdat, Leningrad.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 1, January–March 1998, pp. 53–66

Estimation of solar radiation components over Qena/A.R. Egypt at cloudless sky conditions (model verification)

Sayed M. El-Shazly

Department of Physics, Faculty of Science,
South Valley University, Qena/A.R. Egypt; E-mail: svalleyu@frcu.eun.eg

(Manuscript received 27 January 1997; final form 16 June 1997)

Abstract—In this paper three simple, widely applicable models were used to estimate the various components of solar radiation under clear sky conditions at Qena /Egypt. The first model is an empirical one and estimates the global solar radiation as a function of the solar elevation h as: $G = A \sin(h) + B$. The second and third models were developed by *Kasten* (1983) and *Davies* and *McKay* (1982) representing semi empirical and physical models, respectively. The results produced by the models have been compared with measured hourly and daily values. Model performance was assessed from the mean bias error (MBE), the mean absolute error (MAE) and the root mean square error (RMSE). The comparison shows that the global and direct solar radiation are more accurately estimated than the diffuse radiation. With respect to global radiation (G) all the three models perform well, but the first and second ones estimate better. The mean bias error in estimating hourly values of G is smaller using the first model (= zero in all months), while for daily values the Kasten model is better.

With respect to direct solar radiation (I), Kasten model provided better daily and hourly estimates than the Davies and McKay one. In general, the superiority of an appropriate statistical approach against a physical one is obvious. The monthly variation of the model performance indicates somewhat general consistencies between the months, which are characterized by high aerosol vibration and that show high error indices.

Key-words: solar radiation components estimation, statistical models, physical models, aerosol effect.

1. Introduction

As long as precise measurements of solar radiation are absent, one still needs models to estimate various components of it, which are necessary for solar energy utilization and studying of atmospheric conditions. This need is more significant in the developing countries, where these measurements are scarce,

their spatial density is inadequate. The meteorological and engineering literature is replete with such procedures. There are mainly two approaches. The first approach is represented by several empirical and semi empirical models, these are commonly called statistical models. They were developed to satisfy local needs and the users in other locations have to verify or revise the numerical values of the constants and coefficients before applying them. However the semi empirical models have some claim to generality. The second approach is represented by the models based on radiative transfer calculations that simulate the physical processes in the atmosphere, e.g. the attenuation due to the absorption and scattering of water vapor, ozone, Rayleigh and Mie scattering by gases and aerosols. These models are usually referred to as physical models. They require a lot of information about the constituents of the atmosphere, but they do not need solar radiation data and may be applied anywhere.

The aim of this paper is the analysis of the accuracy of two statistical and one physical model in estimating the radiative fluxes on a horizontal surface in Qena/Egypt at cloudless days. A selection was made from models and model forms which may be suitable for general application.

2. *Selected models*

In this paper we use some models, which are considered relatively simple and well suited to the available meteorological information.

The *first model* (*Kasten* and *Czeplak*, 1980; *Holstlag* and *Van Ulden*, 1983; *Van Ulden* and *Holstlag*, 1985; *Sahsamanoglou*, 1991), also the simplest one, estimates the global solar radiation (G) as a function of the solar elevation h, according to the following known relation:

$$G = \mathrm{A}\ \sin(h) + \mathrm{B}, \tag{1}$$

where A and B are coefficients estimated on the basis of hourly global solar radiation using least square method.

The *second model* (*Kasten*, 1983) estimates G in the case of cloudless sky as:

$$G = G_0\, \mathrm{C}\, \exp(-\mathrm{D}\, T_L\, m), \tag{2}$$

where G_0 is the extraterrestrial radiation, T_L is the Linke turbidity factor, m is the relative optical air mass (see Eq. (12)) and C and D have the values C = 0.84 and D = 0.027, based on data over West Germany. T_L is calculated as:

$$T_L = (1 + [(\tau_0 + \tau_w + \tau_a)/\tau_r)], \tag{3}$$

in which τ_0, τ_w, τ_a and τ_r are the spectrally integrated optical depths for ozone

absorption, water vapor absorption, aerosol attenuation and Rayleigh scattering, respectively. Kasten did not include the calculation of direct beam radiation in his model. This modification was made using :

$$I = I_0 \exp(-T_L \tau_r m). \tag{4}$$

The diffuse component is given as:

$$D = I \sin h - G. \tag{5}$$

The *third model* (*Davies* and *McKay*, 1982) calculates the direct (I) and diffuse (D) radiation, both on horizontal surface, as:

$$I = G_0 (T_0 T_r - a_w) T_A, \tag{6}$$

$$D = G_0 [(1 - T_r)/2 + (T_0 T_r - a_w)(1 - T_A) \omega g], \tag{7}$$

where T_0 is the transmissivity after absorption by ozone, T_r is the transmissivity after Rayleigh scattering, a_w is the absorptivity of water vapor, T_A is the transmissivity after extinction by aerosol, ω is the spectrally-averaged single scattering albedo for aerosol and g is the ratio of forward to total scattering by aerosol. Global radiation is expressed as the sum of I and D. The following formulas (*Iqbal*, 1983) have been used for the parameterization of the different models:

$$G_0 = I_0 \cos Z \tag{8}$$

in which I_0 is the corrected value of the solar constant and Z is the solar zenith angle given as:

$$I_0 = I_{sc} [1 + 0.033 \cos(360\, d_n/365)] \tag{9}$$

and

$$Z = \cos^{-1} (\sin \phi \sin \delta + \cos \phi \cos \delta \cos H), \tag{10}$$

where φ is the latitude of the station, δ and H are the declination and hour angles of the sun, respectively.

$$T_0 = 1 - \left[0.1611\, x_1\, (1.0 + 139.48\, x_1)^{-0.3035}\right.$$

$$\left. - 0.002715\, x_1\, (1 + 0.044\, x_1 + 0.0003\, x_1^2)^{-1}\right], \quad (11)$$

in which $x_1 = m\, U_0$, where U_0 is the ozone layer thickness in cm (NTP) taken from *Robinson* (1966) and m is the relative optical air mass calculated using the *Kasten* formula (1966):

$$m = \left[\cos Z + 0.15/(93.885 - Z)^{1.253}\right]^{-1}. \quad (12)$$

$T_r = \exp(-\tau_r\, m)$, in which τ_r is calculated according to *Louche et al.* (1986) as:

$$\tau_r = (6.5567 + 1.7513\, m - 0.1202\, m^2 + 0.0065\, m^3 - 0.00013\, m^4)^{-1}. \quad (13)$$

$$a_w = 2.4959\, x_2 \left[(1 + 79.034\, x_2)^{0.6828} + 6.385\, x_2\right]^{-1}, \quad (14)$$

in which $x_2 = wm$, where w is the perceptible water thickness in cm, calculated with the aid of the *Leckner* (1978) formula:

$$w = 0.493\, (\varphi_r/T) \exp (26.23 - 5416/T). \quad (15)$$

φ_r is the relative humidity in fraction of one and T is the ambient temperature in Kelvin, taken from the Meteorological Department of A.R. Egypt.

$T_A = \exp(-\tau_a\, m)$, in which τ_a is calculated using the following formula (*Freund*, 1983):

$$\tau_a = (-1/m) \ln\left[(I/I_0)/(T_0\, T_r - a_w)\right]. \quad (16)$$

$g = 0.93 - 0.21 \ln (m)$. A constant value of 0.75 was used for ω. τ_0 and τ_w in the Kasten model are calculated with the aid of the above given equations of T_0 and T_w as:

$$\tau_{0,w} = (-1/m) \ln T_{0,w}. \quad (17)$$

3. *General climate of Qena/Egypt*

Qena is located in the south part of Egypt at latitude 26°10'N, longitude 32°43'E and elevation 78 m above sea level. Climatically, Qena lies within the sub-tropical region characterized by warm sunny days and rather cool nights in winter (Dec–Feb) and hot, very dry summers (Jun–Aug). In spring (March-

May) the khamasin depressions are the main features which are associated with high and medium clouds. They can be vigorous and cause severe sand storms raised by strongly southern winds. The study region has almost calm weather, low cloudiness (80% of the days of the year are cloudless) and nearly no precipitation. Analysis of trend values of the most important weather elements, which were taken from the Meteorological Authority of A.R. Egypt, shows the following:

(i) Average temperature ranges from 14.5°C in January to 34°C in July.

(ii) Average relative humidity varies from 21% in May and June to 48% in December.

(iii) Significant percentage of winds is calm ($\cong$52%). The prevailing winds are W, NW, SW, and N, with percentages of occurrence of 15.9%, 11.8%, 11.7% and 4.5%, respectively. The majority of winds range from 2 to 3.1 m s^{-1} and the least occurrence of speed intervals ranges from 8.8 to 10.8 m s^{-1}.

4. Experimental data

Either for empirical and semi empirical models or for quality control of the physical ones, measured values of solar irradiance and other parameters have been used. The hourly values of global (G_h) and diffuse (D_h) solar radiation, which are used in this study were recorded through a program for measuring the solar radiation components, supervised by the author, over Qena/Egypt. A precision pyranometer (Kipp and Zonen Model CM 6B) is used to measure G, another similar one, fitted with a shadow band (radius 620 mm and width 60 mm), constructed by the author following Kipp and Zonen rules, has been used to measure D. Both pyranometers were conjuncted with a two channels solar integrator (Kipp and Zonen Model CC12) to record the values of hourly G_h and D_h in kJ/m^2 as well as daily values of G_d and D_d in MJ/m^2 for a period of 2 1/2 years (Feb 1992–Sep 1994). More details about the experimental techniques and the specification and exposure of the instruments are given in a previous contribution (*El-Shazly*, 1994). From these data, those under clear sky conditions have been selected for this study. The criterion for clear sky condition is the cloudiness to be less than one okta. From G and D data, the corresponding values of I were calculated. The measurements were performed on the roof of a building of the Literature Faculty, South Valley University, Qena/Egypt. There were no evidence of errors in the measured radiation records except for a few cases when D_h exceeds G_h near sunrise and sunset. In these instances the observations were rejected.

5. *Results and discussion*

5.1 Estimation of the coefficients A, B, C, *and* D *in Eqs. (1) and (2)*

The values of A, B, C, and D were estimated on the basis of hourly values of *G, D* and *h* with the aid of the least square method. Because the quantities of water vapor and aerosols in the atmosphere vary from month to month, so different pairs of values of A and B were estimated for each month. *Table 1* summarize the monthly values of A and B in Qena/Egypt. Also, its average values in some other areas are given in *Table 2* for comparison. With respect to C and D in the Kasten formula, they were found to be 0.801 and 0.011, respectively. These values are in a very good agreement with the values estimated by Kasten and in other locations in Egypt as given in *Table 3*. This table shows that the values of C and D are nearly the same for both Egypt and Germany. This result indicates that the considering of the turbidity factor in the empirical relations to estimate the solar radiation components under cloudless sky conditions, may make these realtions quasi independent of the climatological conditions of the stations and have some claim to generality.

Table 1. Monthly values of coefficients A and B (W/m^2) in Qena/Egypt

Month	A	B	R^2
January	1135.1	-69.5	0.920
February	1183.7	-84.0	0.900
March	1087.7	-46.4	0.940
April	1152.7	-97.1	0.990
May	1175.9	-132.9	0.980
June	1162.7	-128.2	0.950
July	1143.9	-112.9	0.960
August	1125.7	-105.9	0.950
September	1076.2	-48.6	0.930
October	1172.6	-106.0	0.920
November	1131.6	-78.0	0.940
December	1108.5	-62.2	0.990

Table 2. Average values of coefficients A and B (W/m^2) at different areas (*Sahsamanoglou* and *Makrogiannis*, 1991)

Area	A	B
Boston (42°13'N/71°07'W)	1098	-65
N. Atlantic (52°30'N/20°W)	1100	-50
Harrogate (54°00'N/1°30'W)	990	-30
Hamburg (53°38'N/9°50'E)	910	-30
De Bilt (52°06'N/5°11'E)	1041	-69
Qena (26°10'N/32°43'E)	1137	-88

Table 3. Values of coefficients C and D in the Kasten formula at different locations

Location	C	D
Hamburg (Germany)	0.840	0.027
Barani (A.R. Egypt)	0.933	0.034
Matruh (A.R. Egypt)	0.879	0.027
El Arich (A.R. Egypt)	0.887	0.032
Tahrir (A.R. Egypt)	0.835	0.027
Cairo (A.R. Egypt)	0.749	0.017
Aswan (A.R. Egypt)	0.864	0.030
Kharga (A.R. Egypt)	0.800	0.018
Qena (A.R. Egypt)	0.801	0.011

5.2 Model performance

Model performance was assessed primarily from the mean bias error (MBE), which characterizes systematic errors, and from the root mean square error (RMSE), characterizing nonsystematic errors. The mean absolute bias error (MAE) was also computed since the MBE may cancel significant positive and negative biases. These indices were calculated for each month for both hourly (kJ/m^2) and daily (MJ/m^2) totals of G, D and I:

$$RMSE = \left[\Sigma (Y_{cal} - Y_{\exp})^2 / N\right]^{0.5}, \tag{18}$$

$$MBE = \left[\Sigma (Y_{cal} - Y_{\exp}) / N\right], \tag{19}$$

$$MAE = \left[\Sigma \left|(Y_{cal} - X_{\exp})\right| / N\right]. \tag{20}$$

Figs. 1a, b and *c* summarize the performance statistics for each flux and model for daily and hourly irradiation in different months of the year. From these figures the following deductions can be made:

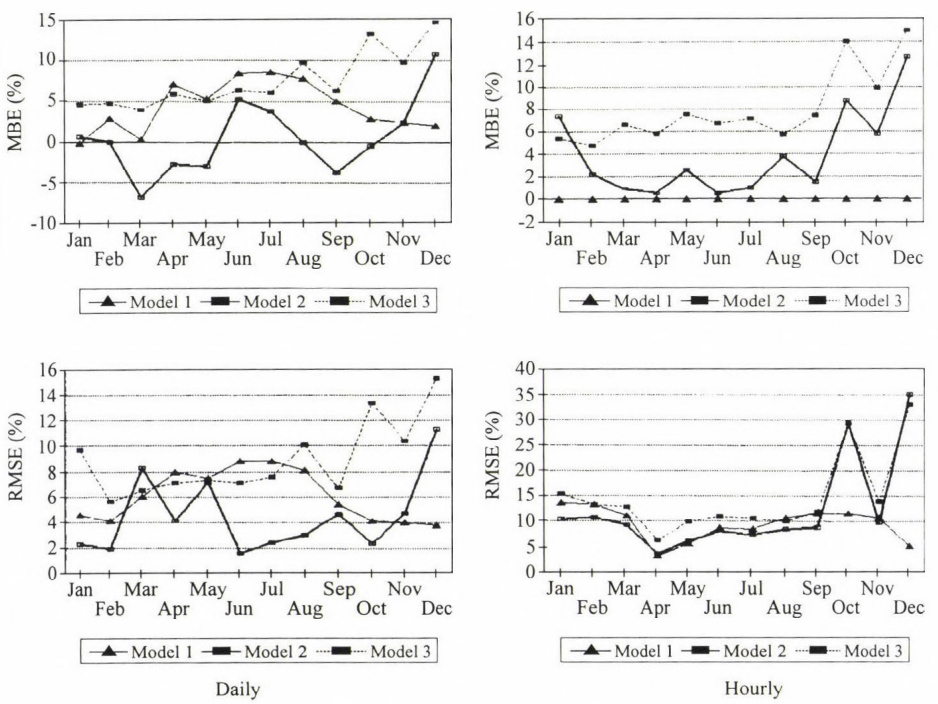
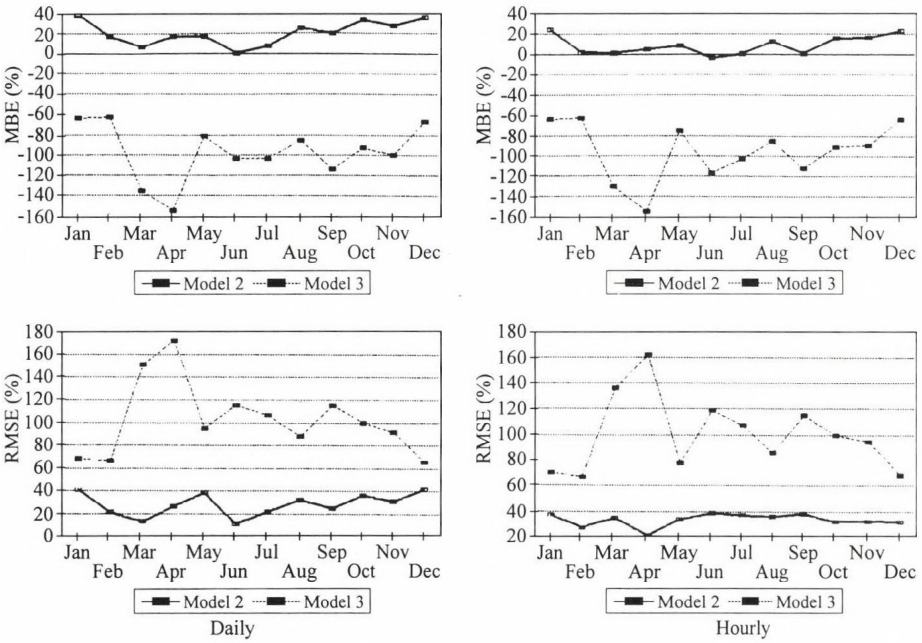
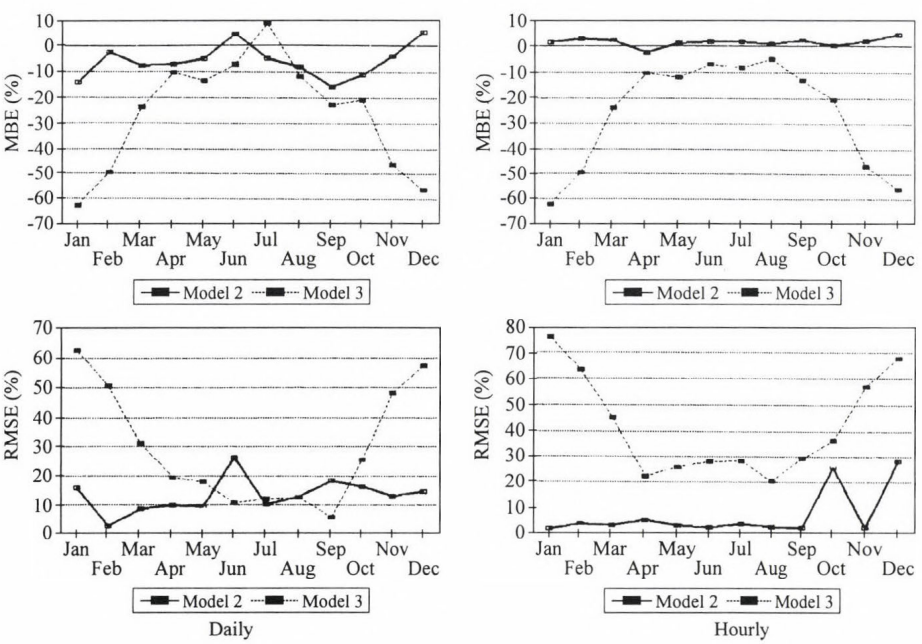


Fig. 1. Monthly variation of error indices in estimating (*a*) global solar radiation; (*b*) diffuse solar radiation; (*c*) direct solar radiation in Qena.

(b)


MBE (%)
RMSE (%)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Model 2
Model 3
Daily
Hourly

(c)


MBE (%)
RMSE (%)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Model 2
Model 3
Daily
Hourly

5.2.1 Global radiation

The three models performed similarly and estimated G with percentage of MBE and RMSE up to 14.7% and 15.3%, respectively, for daily values and 14.9% and 35.2% for hourly ones. However models 1 and 2 estimated the daily and hourly values with the least percentage of errors. The percentage of MBE in daily estimates lie in the ranges (-0.18–8.5), (-6.8–10.7) and (3.96–14.7), while they were (-0.03–0.0), (0.51–12.6) and (4.8–14.9) in hourly estimates, using models 1, 2 and 3, respectively. The corresponding percentages of RMSE were (3.8–8.8), (1.6–11.3) and (5.6–15.3) for G_d and (3.2–13.6), (3.7–35.2) and (6.3–33.2) for hourly values (G_h). As shown in Fig. 1, MBE using the first model is really zero for hourly values, which is a remarkably good result, making this model in favor to the other two models in estimating hourly values of G.

5.2.2 Diffuse and direct solar radiation

With respect to D and I, both daily and hourly, the Kasten model provided better estimates than the Davies-McKay model. However neither of the models performed well for estimating D. As it has been shown in a previous work (*El-Shazly*, 1994), models of Liu and Jordan type realize the best estimation of this radiation component in our region. This conclusion agrees very well with that found by other authors at other locations (*Davies et al.*, 1989).

For I, the Kasten model estimated I with percentages of RMSE in the range of 2.53–18.5 for daily values and 1.47–28.16 for hourly ones (with RMSE of $I_h < 4\%$ for most of the months of the year). RMSE in model 2 reaches 62.5% with most of the values in the range of 10.0–31% for daily estimates and 76% with most of the values in the range of 22.4–29.3 for hourly values. The corresponding MBE were between -15.4 and 5.4 using model 2 and between -62.5 and 8.86 using model 3 for daily estimates, while for hourly estimates it were between -2.4 and 4.56 and from -61.8 to 4.88 using the Kasten and Davies and McKay models, respectively.

For D, the percentage of RMSE for daily and hourly values, respectively, were (12.9–41.8) and (20.5–39) for the Kasten model and (65.9–171) and (67.3–162) for the Davies-McKay model. The MBE percentages were (-0.0–39.34) and (-154.5 – -62.5) and (-3.3–24.1) and (-154.4 – -62.94) for daily and hourly estimates using models 2 and 3, respectively.

As it has been shown above, the accuracy of the statistical approach is better than that of the physical one. This may be explained in view of the inadequacy of the necessary input parameters in the physical model (dust, water vapor and ozone) to estimate the various quantities related to the interaction of solar radiation with the atmosphere. Especial interest must be given to the aerosol effect, which is the major source of error in model estimation for the

real atmosphere as suggested by *Omran* (1989) and confirmed in this paper as will be shown in the next section. Measurements of these parameters with sufficient accuracy for each individual location will improve the accuracy of the physical model to be comparable or better than statistical ones.

5.2.3 Monthly variation of the performance of the selected models

The monthly performance of the three models are shown in the Fig. 1a, b, and c for G, D and I, respectively. From these figures, the following may be concluded:

(i) All models overestimate G, both daily and hourly, except in some months, in which model 2 underestimates its daily values (Mar, Apr, May, Sep and Oct). Model 1 shows no biases in hourly values of G.

(ii) Model 2 overestimated D_d, D_h (except in June) and I_h (except in April), while I_d was underestimated using this model (except in December and June). However both D and I were underestimated using model 3.

Considering the nature of the atmosphere of the study region, we can say that the fluctuation of the error indices through the months of the year may be generally due to the inadequate incorporating of aerosol in the calculations in the different models. In Qena city the main pollutants in the atmosphere are the aerosol dust particles dispersed from the surrounding east and west mountains as well as man activities (*El-Shazly*, 1989). So the attenuation by aerosols in the incoming solar radiation in this region is significant and more than that caused by other constituents (Water vapor, Ozone, and Rayleigh scattering) as it is shown in *Fig. 2*. Also its content varies considerably from day to day and month to month. A general consistency was found between the months, which are characterized by obvious aerosol variation and that show high error indices. Accordingly its effect can not be safely ignored in calculations with models attempting to imitate the physical processes, which attenuate radiation and appear in the different values of solar radiation components at the same Julian day. However there is little empirical information, which can be used in models, and aerosol effects can be incorporated crudely. The different optical characteristics of the aerosol in Qena will be the matter of an empirical study in the near future owing to its very important role in the radiation energy exchanges taking place within the earth-atmosphere system.

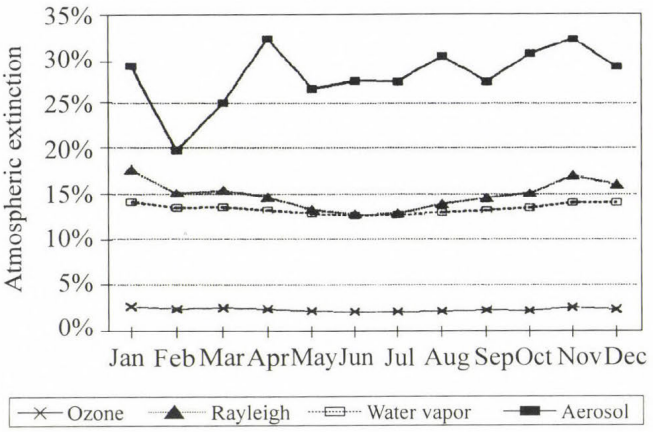


Fig. 2. Variation of the atmospheric extinction caused by different constituents in Qena/Egypt.

5.2.4 Comparison between the performances of the models at different stations

Comparing the model performances at different stations in Egypt and some other countries (*Davies*, 1989) shows the following:

(1) Model 1 provided similar high accuracy in estimating G_d in both Qena/Egypt (MBE = 4.33% and RMSE = 6.09%) and Thessaloniki/Greece (MBE = –0.01% and RMSE = 8.83%).

(2) With respect to the other two models, *Table 4* summarizes the percentage of MBE and RMSE, which were found estimating *G, D* and *I* both daily and hourly at different locations. The values of these indices show some claim to homogeneity in view of the different atmospheric conditions. The following general features may be concluded from this table:

   (a) Both models estimate *G* with smaller errors than *D* and *I* at all stations.

   (b) In Cairo/Egypt the two models perform better in estimating *I* than *D*, which is in contrast to the European stations, where *D* was estimated with slightly more accuracy than *I*.

   (c) In Qena/Egypt model 3 failed to match its performance in the other stations.

   (d) Higher values of both MBE and RMSE were found usually for hourly estimates of all solar radiation components compared to daily ones.

Table 4. Percentage of RMSE and MBE of solar radiation components at different locations

Location		MODEL 2						MODEL 3					
		G		*D*		*I*		*G*		*D*		*I*	
		MBE	RMSE	MBE	RMSE	MBE	RMSE	MBE	RMSE	MBE	RMSE	MBE	RMSE
Qena/Egypt	d	0.5	4.51	20.8	28.3	-5.8	13.2	7.5	8.99	-96.9	103	-26.25	29.5
	h	3.94	12.3	8.7	33.4	1.56	6.83	8	14.79	-95.6	100	-26.13	41.6
Matruh/Egypt	d	-0.5	6.6	-5	35.7			-7.2	10.1	-27	47.1		
El-Arich/Egypt	d	-1.1	12.3	4.7	36.1			-3.8	12.7	-5	37.5		
Cairo/Egypt	d	-0.2	9	0.5	26.5	0.36	18.8	0.32	8.1	-1.8	26	1.4	18.7
Aswan/Egypt	d	0.61	13	12.5	39			2.5	12.8	9.8	38.6		
Kharga/Egypt	d	-1.4	5.8	10	30.5			-3.9	6.4	-4.4	27.5		
De Bilt (The Netherlands)	d	3.8	19.3	-0.5	29.2	1.71	57.8	-1.5	17.7	-18.3	33.8	32	69.4
	h	3.8	39	-2.4	49.7	19.1	102	-1.5	38.7	-18.3	52.5	34.1	112.2
Hamburg (Germany)	d	5	21.8	10.8	38.6	1.8	57.4	1.2	18.8	-13.1	36.7	21.9	62.2
	h	5	35.5	9.2	56.9	1.9	80.3	1.2	33.7	-13.1	57.3	22	86
Kew/UK	d	4.7	20.5	24.9	38.8	-15	50.4	1.5	18.4	9.2	30.7	1.8	43.8
	h	4.7	34.3	24.9	52	-16	78.7	1.5	31.4	8.7	46.1	1.4	73.5

6. Conclusion

An attempt to estimate various components of solar radiation at Qena/Egypt using two main types of models, namely, statistical models (1 and 2) and a physical model (3) leads to the following conclusions:

(i) All selected models estimate the global radiation, both hourly and daily, with satisfactory accuracy, but model 1 and 2 usually better.

(ii) Model 2 estimates the diffuse and direct components better than model 3, which failed to estimate D, both hourly and daily.

(iii) Values of C and D coefficients in equation 2 agree very well with that found for other stations. This indicates the importance of introducing some meteorological parameters in the empirical formulas to improve its accuracy.

(iv) There is a general consistency between months with high variation of aerosol contents and that of large error indices.

(v) The accuracy of the statistical approach is better than that of the physical one, but they can not have general applicability.

(vi) Although model 2 produced generally the best estimates of solar radiation components in Qena, the differences between the statistical measures of errors for the best and the worst performing models may not be sufficient to be significant for solar energy or any other application.

References

Davies, J.A. and *McKay, D.C.*, 1982: Estimating solar irradiance and components. *Solar Energy 29*, 55-64.

Davies, J.A. and *McKay, D.C.* 1989: Evaluation of selected models for estimating solar radiation on horizontal surfaces. *Solar Energy 43*, 153-168.

El-Shazly, M.S., 1994: Solar radiation components at Qena/Egypt. *Időjárás 98*, 47-54.

El-Shazly, S.M., 1989: Studies of the number concentration and size distribution of the suspended dust particles in the atmosphere of Qena/Egypt. *Water, Air and Soil Pollution 45*,121-133.

Freund, J., 1983: Aerosol optical depth in the Canadian Arctic. *Atmosphere-Ocean 2*, 158-167.

Holstlag, A.A.M. and *van Ulden, A.P.*, 1983: A simple scheme for daytime estimation of the surface fluxes from routine weather data. *J. Climate Appl. Meteorol. 22*, 517-529.

Iqbal, M., 1983: *An Introduction to Solar Radiation*. Academic Press, New York.

Kasten, F., 1983: Parametrisierung der Global Strahlung durch Bedeckungsgrad und Trubungs Faktor. *Ann. Met. 20*, 49-50.

Kasten, F., 1966: A new table and approximation formula for the relative optical air mass. *Archive für Meteorol. Geophys. und Bioklim* B, 206-223.

Kasten, F. and *Czeplak, G.*, 1980: Solar and terrestrial radiation dependent on the amount and type of cloud. *Solar Energy 124*, 177-189.

Leckner, B., 1978: The spectral distribution of solar radiation at the earths surface. Elements of a model. *Solar Energy 20*, 143-150.

Louche, A., Peri, G. and *Iqbal, M.*, 1986: An analysis of Linke turbidity factor. *Solar Energy 7*, 393-396.

Omran, M.A.H., 1989: Modeling solar radiation and effects of aerosols over Egypt. *Ph.D. Thesis*, Cairo University (Egypt).

Robinson, N. (ed.), 1966: Solar Radiation. American Elsevier, New York.

Sahsamanoglou, H.S. and *Makrogiannis, T.J.*, 1991: Empirical estimation of daily sky solar radiation in Thessaloniki, Greece. *Theor. Appl. Climatol. 43*, 211-215.

Van Ulden, A.P. and *Holstlag, A.A.M.*, 1988: Estimation of atmospheric boundary layer parameters for diffusion applications. *J. Climat Appl. Meteor. 24*, 1196-1207.

BOOK REVIEW

Wilfried Schröder and *Hans-Jürgen Treder* (eds.): **The Earth and the Cosmos**. *Science Edition, IDCH IAGA/History commission DGG. 382 pages.*

This book is presented to the memory of *Professor Hans Ertel* (1907–1971). The object of this book is the introduction of Hans Ertel as scientist, person and leading manager in science.

This book uses many sources, e.g. correspondence with lot of scholars, his working life, some papers from his wide field of scientific work.

Reading this book we learn that Prof. Ertel made many important discoveries in theoretical meteorology, theoretical geophysics and geophysical hydrodynamics.

First of all Ertel's potential vorticity theorem has become more and more important in meteorology, geophysics, astrophysics, physics and geophysical hydrodynamics.

During the early 1930-ies Ertel addressed the general problem of the thermodynamics and friction of the atmosphere and ocean. His most important contributions were to show the importance of the variation of the Coriolis force with latitude and the theory of long waves in dynamical and synoptic meteorology. He studied the problems of interactions between the stratosphere and troposphere. He emphasized the importance of the advective-dynamics theory of air pressure variations and their periodicities.

In 1939 he published his famous book on theoretical meteorology.

After the Second World War *Rossby* invited Ertel to Stockholm and *Hilding Kueler* to Uppsala. Then and in the time of Rossby's visit in Berlin (1948) it was clear that Ertel's earlier studies had implied the existence of Rossby waves.

On the basis of "Ertel's vortex invariants" derived from the vorticity equations a symmetrical deformation tensor could be constructed which made it possible to make linear transformation of the vorticity components in their total-time derivatives.

One of his very important works: "A relationship between kinematical parameters of horizontal fields of flow in the atmosphere" was published in IDŐJÁRÁS (Vol. 74 (1970), 98-102).

Professor Ertel was a member of the Editorial Board of IDŐJÁRÁS. He was in correspondence with many Hungarian meteorologists, e.g. with *László Aujeszky, Frigyes Dési, György Kozma, Ferenc Rákóczi* and *Alfréd Zách*.

Hans Ertel was a member of the Deutsche Akademie der Wissenschaften since 1949, in 1952 he was chosen as one of its vice-presidents. He was at the

same time Director of the Institute of Meteorology and Geophysics in Berlin, Professor for Geophysics and Theoretical Mechanics at the Humboldt University and Director of the Institute of Physical Hydrography of the Academy of Sciences.

He helped to organize the commemoration of the work of Alexander Humboldt and Max Planck. He worked actively in the program of the Geophysical Year and in the work of the International Hydrological Decade. He was one of the leader scientists of the Carpathian Conference.

Unfortunately, Ertel published all his papers in German or Spanish, in journals which were not very well-known internationally, with the consequence that Ertel's work has not received the recognition it merits. Knowing this fact, we have to agree with *W. Schröder*, who told that Ertel's work "has been overlooked in the English scientific literature. Consequently many 'new' ideas in English books and journals are 'old', because they have been described already by H. Ertel in the years between 1930–1970".

The book consists copies of original articles, xeroxes of handwritings and letters.

We are in agreement with the subtitle of the book, it is really: "The Legacy of Hans Ertel".

Ferenc Rákóczi

ATMOSPHERIC ENVIRONMENT

an international journal

To promote the distribution of Atmospheric Environment *Időjárás* publishes regularly the contents of this important journal. For further information the interested reader is asked to contact *Prof. P. Brimblecombe*, School for Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, U.K.; E-mail: atmos_env@uea.ac.uk

Volume 32 Number 1 1998

Conference on the Benefits of the Urban Forest (sponsored by the) Sacramento Municipal Utility District, Sacramento, CA; 7 March 1995

Volume 32 Number 2 1998

NOTES TO CONTRIBUTORS OF *IDŐJÁRÁS*

The purpose of the journal is to publish papers in any field of meteorology and atmosphere related scientific areas. These may be

- reports on new results of scientific investigations,
- critical review articles summarizing current state of art of a certain topic,
- shorter contributions dealing with a particular question.

Each issue contains "News" and "Book review" sections.

Authors may be of any nationality, but the official language of the journal is English. Papers will be reviewed by unidentified referees.

Manuscripts should be sent to
Editor-in-Chief of IDŐJÁRÁS
P.O. Box 39
H-1675 Budapest, Hungary

in three copies including all illustrations. One set of illustrations has to be of camera ready quality, the other two might be lower quality.

Title part of the paper should contain the concise title, the name(s) of the author(s), the affiliation(s) including postal and E-mail address(es). In case of multiple authors, the cover letter should indicate the corresponding author.

Abstract should follow the title, it contains the purpose, the data and methods as well as the basic conclusion.

Key-words are necessary to help to classify the topic.

The text has to be typed in double spacing with wide margins. Word-processor printing is preferred. The use of SI units are expected. The negative exponent is preferred to solidus. Figures and tables should be consecutively numbered and referred to in the text.

Mathematical formulas are expected to be as simple as possible and numbered in parentheses at the right margin. Non-Latin letters and hand-written symbols should be indicated and explained by making marginal notes in pencil.

Tables should be marked by Arabic numbers and printed in separate sheets together with their captions. Avoid too lengthy or complicated tables.

Figures should be drawn or printed in black and white, without legends, on separate sheets. The legends of figures should be printed as separate list. Good quality laser printings are preferred as master copies.

References: The text citation should contain the name(s) of the author(s) in Italic letter and the year of publication. In case of one author: *Miller* (1989), or if the name of the author cannot be fitted into the text: *(Miller,* 1989); in the case of two authors: *Gamov* and *Cleveland* (1973); if there are more than two authors: *Smith et al.* (1990). When referring to several papers published in the same year by the same author, the year of publication should be followed by letters a,b etc. At the end of the paper the list of references should be arranged alphabetically. For an article: the name(s) of author(s) in Italics, year, title of article, name of journal, volume number (the latter two in Italics) and pages. E.g. *Nathan, K.K.,* 1986: A note on the relationship between photosynthetically active radiation and cloud amount. *Időjárás 90,* 10-13. For a book: the name(s) of author(s), year, title of the book (all in Italics except the year), publisher and place of publication. E.g. *Junge, C. E.,* 1963: *Air Chemistry and Radioactivity.* Academic Press, New York and London.

The final version should be submitted on diskette altogether with one hard copy. Use standard 3.5" or 5.25" DOS formatted diskettes. The preferred word-processors are WordPerfect 5.1 and MS Word 6.0.

Reprints: authors receive 30 reprints free of charge. Additional reprints may be ordered at the authors' expense when sending back the proofs to the Editorial Office.

More information: gmajor@met.hu
Information on the last issues:
http://www.met.hu/firat/ido-e.html

Published by the Hungarian Meteorological Service

Budapest, Hungary

INDEX: 26 361

HU ISSN 0324-6329

IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

CONTENTS

****

http://www.met.hu/firat/ido-e.html

VOL. 102 * NO. 2 * APRIL–JUNE 1998

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: P.O. Box 39, H-1675 Budapest, Hungary or
Gilice tér 39, H-1181 Budapest, Hungary
E-mail: gmajor@met.hu or *antal@met.hu*
Fax: (36-1) 290-7387

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary;
E-mail: gmajor@met.hu or *antal@met.hu*; *Fax: (36-1) 290-7387*

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 2, April–June 1998, pp. 71–107

On the origin of model errors. Part II. Effects of the spatial discretization for Hamiltonian systems

Barbara Kádár, István Szunyogh[1]

Department of Meteorology, Eötvös Loránd University,
H-1083 Budapest, Ludovika tér 2, Hungary
E-mail: kadar@nimbus.elte.hu

and **Dezső Dévényi**

NOAA/ERL/Forecast Systems Laboratory,
325 Broadway, Boulder CO, USA; E-mail: devenyi@fsl.noaa.gov
also affiliated with
Cooperative Institute for Research in Environmental Sciences,
University of Colorado at Boulder, Boulder, CO, USA

(Manuscript received 2 December 1997; in final form 12 March 1998)

Abstract—The effects of model errors stemming from the temporal and spatial discretizations of the continuous model equations are investigated (in two parts) for Hamiltonian systems. In this second part of the paper the effects of spatial truncation are discussed. Since the "article relabelling" symmetry is only implicit in the Eulerian description, the spatial truncation of the Eulerian equations inevitably distorts the individual conservation properties of the flow. Theoretical considerations and numerical experiments confirm that the wave mode interactions misrepresented in the finite dimensional truncation are responsible for the destruction of structure and the breaking of conservation laws. Although there is no way to circumvent this problem, there are two approaches to reduce the effects of the falsely represented interactions: (1) Increasing the resolution and initializing the fields so that the energy of high wavenumber modes is small; (2) Use of structure preserving schemes that can conserve the algebraic structure and a finite number of invariants. To illustrate these methods the results of numerical experiments carried out with a traditional (aliased) and a structure preserving truncation of the two-dimensional vorticity equation and of the 2-layer quasi-geostrophic model are presented.

Key-words: Hamiltonian formalism, invariants of motion, structure preserving spatial truncation.

---

[1] Present affiliation: UCAR scientific visitor at the National Centers for Environmental Prediction, USA; E-mail: Istvan.Szunyogh@noaa.gov

1. Introduction

In the first part of this paper (*Kádár et al.*, 1998, hereafter Part I) we discussed the effects of temporal discretizations. We pointed out, that even if the spatially discretized equations were perfect, the time integration inevitably introduces model errors whose effects in a chaotic system lead to the complete loss of the forecast skill within a finite time interval. This part concentrates on the effect of spatial discretizations and as in Part I we start considering conservative fluid dynamical processes.

Atmospheric (fluid) dynamics is based on the classical mechanical model of interacting atmospheric (fluid) particles. The number of these particles is infinite and their volume goes to zero in a mathematical analytical sense. The infinite dimensional system of equations describing the motion of the particles is called the Lagrangian form of the atmospheric (fluid) dynamical equations. These equations, however, are very rarely used for either the solution of theoretical problems or for more practice oriented forecasting purposes.

Our main interest is usually in predicting the state of the atmosphere at given geographical locations. This goal can be achieved by solving the Eulerian form of the dynamical equations, where we do not distinguish between the individual fluid particles any more. One should always keep in mind, however, that the fundamental physical model even in the case of the Eulerian equations is the one utilizing infinitesimal fluid particles. The ultimate test of any numerical solution strategy is to examine to what extent it can preserve the fundamental properties associated with the system of particles. The main difficulty with investigating the effects of spatial discretization of the Eulerian equations stems from the fact that it discretizes the space instead of the fluid and we cannot speak about fluid particles any more. This problem is more than a simple technical difficulty. Examination of the rigid body equations provided guidance for exploring the relations between the Eulerian and the Lagrangian characteristics of fluid motions (see *Arnold*, 1989, App. 2). The key step in this procedure is the identification of a metric naturally arising from the equations themselves. Once the equations are discretized in space the choice of this metric is not apparent any more (*Zeitlin* and *Pasmanter*, 1994). The discussion of this important theoretical problem is out of the scope of the present paper, but the paper by *Dowker* and *Wolski* (1992), though fails to give definite answer, gives an insight into the technical and theoretical difficulties associated with this problem. Here we follow the traditional, technically less demanding approach, i.e. we will investigate quantitative properties directly related to the original physical model which are still calculable in the discretized Eulerian systems. Conservation laws offer a unique opportunity to realize this program.

Lagrangian forms of the adiabatic primitive equations ensure the individual (particle) conservation of (potential-)vorticity-type quantities in addition to the usual mechanical invariants, such as total energy and momenta of the system.

Noether's theorem (*Olver*, 1989; *Shepherd*, 1990) connects the individual conservations to the "relabelling" symmetry: interchanging particles of the same (potential-)vorticity does not alter the motion of the system. Derivation of Eulerian equations is a reduction via these symmetries. Neither the replacement of particles along the isolines of individual invariants nor the symmetries themselves can be detected in the Eulerian description. Remarkably however, Hamiltonian mechanics provides a systematic way to find the so-called Casimir invariants associated with hidden symmetries.

Dealing with the conservative part of the equations governing fluid motions, we make use of this advantage of Hamiltonian formalism. In Section 2 we briefly summarize the invariants of motions of Hamiltonian systems. Since the effects of the different truncation methods on these conservation laws are in the forefront of our interest, in Section 3 we describe the criteria for maintaining the conservation laws in the spatial discretization of the continuous equations.

Section 4 contains the description of a traditional and a structure conserving spectral truncation of the two-dimensional vorticity equation on a double periodic plane. The continuous equation has infinite invariants of motion — reflecting the individual conservation of vorticity. We examine how the traditional discretization methods distort the algebraic structure of the original continuous equation and break most conservation laws. The structure preserving truncation does retain the Hamiltonian structure and preserve a finite number of invariants, but these are not identical with the discrete analogues of the original integrals of motion.

The quasi-geostrophic layer models, which are important tools for investigating baroclinic instability in geophysical fluids since they are capable to simulate baroclinic phenomena even at very low resolution, also possess Hamiltonian structure (*McLachlan et al.*, 1997). Since the truncation methods used for the two-dimensional vorticity equation are applicable to these models as well, the comparative study was extended to these more realistic models. Section 5 describes the basic equations and the applied truncation strategies to these models.

Fortunately, there is a remarkably efficient integration algorithm that can be used for the integration of both truncations of the vorticity equation and the quasi-geostrophic layer models (*McLachlan*, 1993). It makes feasible to carry out even extremely long numerical integrations. The role of the constraints exerted by the additional invariants conserved in the structure preserving truncation were investigated through numerical experiments. The results are presented in Section 6. Summary and conclusions are given in Section 7.

There are several statements in the main body of the paper about the structure and the conservation laws of the schemes. Since the authors tried to use the standard mathematical tools and methods of the meteorological literature in the verification of these statements, the proofs are sometimes a bit cumbersome. Since the technical details of these considerations are not necessary for understanding the main points of the paper, we placed the rigorous proofs in the Appendices.

2. *Conservation laws in Hamiltonian systems*

In Part I the basis of Hamiltonian formalism was introduced. In this section the conservation laws associated with Hamiltonian systems are summarized.

2.1 Finite dimensional systems

Once a system of ordinary differential equations is cast in Hamiltonian form, i.e. the skew-symmetric structure matrix **D** and the Hamiltonian H is determined in a way that the Jacobi identity holds for the structure functions (see Section 3.1 in Part I), the invariants of motions can be systematically identified.

The invariance of H is ensured by the skew-symmetry of **D** since

$$\frac{dH}{dt} = (\nabla H)^T \frac{d\mathbf{x}}{dt} = (\nabla H)^T \mathbf{D} \nabla H = -(\nabla H)^T \mathbf{D} \nabla H = 0. \tag{1}$$

In other words, the Hamiltonian H is invariant with respect to translation in time. Further symmetries of H lead to other invariant quantities.

There is, however, another class of invariants, the so called *Casimirs* or *distinguished functions*, which are independent of the Hamiltonian function H (*Shepherd*, 1990). These are the solutions of the homogeneous system of equations

$$\mathbf{D} \nabla C = 0. \tag{2}$$

Their invariance follows from

$$\frac{dC}{dt} = (\nabla C)^T \mathbf{D} \nabla H = -(\nabla H)^T \mathbf{D} \nabla C = 0, \tag{3}$$

where again the skew-symmetry of **D** was taken into account. The number of independent Casimirs is equal to the dimension of the kernel of **D**.

2.2 Continuous Hamiltonian systems

In continuous Hamiltonian systems there are also two classes of invariants: those that are connected to the continuous symmetries of the Hamiltonian (e.g. the Hamiltonian itself) and the *Casimirs* that are the solutions of

$$\mathcal{D} \delta \mathcal{C} = 0. \tag{4}$$

For the two-dimensional vorticity equation (see Part I, Eq. (8)), for example, the Casimirs can be written

$$\mathcal{C} = \int_{\Omega} \mathcal{J}(\zeta) d\mathbf{x}, \tag{5}$$

where $\mathcal{J}$ is an arbitrary smooth function of the vorticity ζ (*Olver*, 1989). For the system of primitive equations and its consistent simplifications (e.g. for the quasi-geostrophic layer models) the potential vorticity is an individually conserved quantity the integrals of whose functions serve as Casimir invariants. More about the conservation laws in Hamiltonian fluid dynamical systems can be found e.g. in *Shepherd* (1990).

3. Spatial discretizations of continuous Hamiltonian systems

The solution of Eulerian atmospheric governing equations requires the application of approximating numerical methods that solve a finite dimensional algebraic analogue of the original system of partial differential equations after the spatial and temporal discretization of the meteorological fields. Schemes used in the numerical simulation of atmospheric motions, however, preserve the Hamiltonian structure only in part (*Szunyogh*, 1993). Since the conservation properties of Hamiltonian systems are closely related to their algebraic structure, this fact substantially restricts the number of retained invariants of motions.

3.1 Formal conservation of the total energy

Since the conservation of total energy is solely related to the skew-symmetry of the structure matrix, the problem of constructing a total energy conserving scheme is equivalent to deriving a skew-symmetric finite dimensional analogue of the continuous differential operator $\mathcal{D}(\mathbf{u})$. The first explorers of formally energy conserving schemes (*Lorenz*, 1960; *Arakawa*, 1966) proceeded along this line, although they did not realize the underlying Hamiltonian structure.

3.2 Formal conservation of the Casimir invariants

If the numerical structure matrix **D** is skew-symmetric, the Casimir invariants of the model are determined by the kernel of **D** (Eq. (2)). Therefore, if one attempts to construct a numerical scheme preserving the finite dimensional analogue $C(\mathbf{x})$ of a Casimir invariant $\mathcal{C}(\mathbf{u})$ of the original continuous system the entries of **D** have to be set according to Eq. (2). The existence of energy and

(potential)-enstrophy preserving schemes (e.g. *Arakawa*, 1966; *Arakawa* and *Lamb*, 1980; *Salmon* and *Talley*, 1989) demonstrates that the preservation of energy and one of the Casimirs at the same time is a realizable goal. The theoretical upper limit for the number of retained Casimir invariants, however, is equal to the corank of **D**. This suggests that an optimal scheme can preserve further Casimirs.

On the other hand, *Szunyogh* (1993) observed that the conservation of the energy (skew-symmetry of **D**) is not a necessary condition for the preservation of a prechosen Casimir. Let

$$(\nabla C)^T \mathbf{D} = 0, \tag{6}$$

then

$$\frac{dC}{dt} = (\nabla C)^T \mathbf{D} \nabla H = 0. \tag{7}$$

The quantity C is a Casimir invariant in the sense that it is independent of the energy function, but it is not a Casimir in the sense of Eq. (2) if **D** is not skew-symmetric (when **D** is skew-symmetric then Eq. (7) is equivalent to Eq. (2)). The best known examples of this type of schemes are Arakawa's enstrophy conserving schemes for the two-dimensional vorticity equation (*Arakawa*, 1966) and Sadourny's potential-enstrophy conserving scheme for the shallow-water equations (*Sadourny*, 1975).

4. Truncation strategies to the two-dimensional vorticity equation

Although the two-dimensional vorticity equation is of interest in atmospheric dynamics only when the beta-effect is negligible, i.e. when the significant scales of motions are short compared to a stationary Rossby wavelength, its capability to describe the nonlinear interactions of different scale motions makes it an extremely useful environment for testing different truncation methods. The Jacobian governs the advection of the individual invariant in the flow and the errors in the approximation of this term are related to the breaking of individual conservation laws.

In this section we describe the Hamiltonian structure of the two-dimensional vorticity equation in its spectral form and its Casimir invariants. In Subsection 2 we examine how the structure (the Jacobi identity) is destroyed by traditional truncation methods. In Subsection 3 a structure preserving truncation (the so called *sine-bracket* truncation) and its conservation properties are described.

4.1 The Hamiltonian structure of the two-dimensional vorticity equation

The two-dimensional vorticity equation on a double-periodic plane transformed with the two-dimensional Fourier transformation takes the form of an infinite system of ordinary differential equations

$$\frac{d\zeta_{\mathbf{m}}}{dt} = \sum_{\substack{n_1, n_2 = -\infty \\ \mathbf{n} \neq \mathbf{0}}}^{\infty} \frac{\mathbf{m} \times \mathbf{n}}{|\mathbf{n}|^2} \zeta_{\mathbf{m}+\mathbf{n}} \zeta_{-\mathbf{n}}. \tag{8}$$

In Eq. (8) $\zeta_{\mathbf{m}}$ denotes the spectral coefficient associated with the two-dimensional integral wave vector $\mathbf{m}$, and the skew-symmetric scalar product is $\mathbf{m} \times \mathbf{n} = m_1 n_2 - m_2 n_1$. The spectral variables satisfy the reality condition $\zeta_{\mathbf{m}} = \zeta^*_{-\mathbf{m}}$, where star denotes complex conjugate. The condition $\mathbf{n} \neq \mathbf{0}$ does not restrict generality since the circulation $\zeta_{\mathbf{0}}$ is a Casimir invariant of Eq. (8), and does not contribute to the time change of any other variable because $\mathbf{m} \times \mathbf{0} = 0$. The Hamiltonian takes the form

$$H = \frac{1}{2} \sum_{\substack{n_1, n_2 = -\infty \\ \mathbf{n} \neq \mathbf{0}}}^{\infty} \frac{\zeta_{\mathbf{n}} \zeta_{-\mathbf{n}}}{|\mathbf{n}|^2}, \tag{9}$$

while the entries of the structure matrix are

$$\mathbf{D}_{\mathbf{mn}} = \mathbf{m} \times \mathbf{n} \zeta_{\mathbf{m}+\mathbf{n}}. \tag{10}$$

The skew-symmetry of the structure matrix follows from $\mathbf{m} \times \mathbf{n} = -\mathbf{n} \times \mathbf{m}$. The Jacobi identity (see Eq. (5) in Part I) can also be shown to hold.

When periodic boundary conditions are applied, the Casimir invariants, $\mathcal{C}^N = \frac{1}{\Omega} \int_\Omega \zeta^N d\mathbf{x}$ can be expressed with the spectral coefficients as follows

$$\mathcal{C}^N = \frac{1}{\Omega} \int_\Omega \zeta^N d\mathbf{x} = \sum_{\sum_{i=1}^{N} \mathbf{n}_i = \mathbf{0}}^{N} \zeta_{\mathbf{n}_1} \cdots \zeta_{\mathbf{n}_N}. \tag{11}$$

In order to point out later the deficiencies of the truncated systems in conserving the Casimir invariants, it is worth verifying directly that Eq. (2) is valid, that is

$$\left[\mathbf{D}\frac{\partial \mathcal{C}^N}{\partial \zeta}\right]_{\mathbf{k}} = 0 \tag{12}$$

for each $\mathbf{k} \in I$. Using Eq. (10) and Eq. (11) the left-hand side of Eq. (12) reads

$$\sum_{\mathbf{i}\in I} \mathbf{D}_{\mathbf{k}\mathbf{i}} \frac{\partial \mathcal{C}^N}{\partial \zeta_{\mathbf{i}}} = N \sum_{\mathbf{i}\in I} \sum_{\substack{N-1 \\ \sum_{j=1} \mathbf{m}_j = -\mathbf{i}}} (\mathbf{k} \times (\mathbf{k}+\mathbf{i}))\, \zeta_{\mathbf{k}+\mathbf{i}}\, \zeta_{\mathbf{m}_1} \cdots \zeta_{\mathbf{m}_{N-1}} . \tag{13}$$

Note that this double sum contains every N-fold product of $\zeta_{\mathbf{n}_i}$-s in which the sum of the indices equals $\mathbf{k}$. Let us examine one particular product in this summation containing the variables $\zeta_{\mathbf{e}_1} \cdots \zeta_{\mathbf{e}_N}$, where $\sum_{i=1}^{N} \mathbf{e}_i = \mathbf{k}$. These $\zeta_{\mathbf{e}_i}$-s can be put in $N!$ different orders, and in $(N-1)!$ different orders with a particular $\zeta_{\mathbf{e}_i}$ in the first position defining the coefficient of this particular product: $\mathbf{k} \times \mathbf{e}_i$. Thus the summation above can be written

$$\sum_{\{\mathbf{e}_i\} \in P_N^{\mathbf{k}}} \zeta_{\mathbf{e}_1} \cdots \zeta_{\mathbf{e}_N} \left[(N-1)! \sum_{i=1}^{N} \mathbf{k} \times \mathbf{e}_i \right], \tag{14}$$

where $P_N^{\mathbf{k}}$ contains all subsets $\{\mathbf{m}_1, \ldots, \mathbf{m}_N\}$ of indices that obey $\sum_{i=1}^{N} \mathbf{m}_i = \mathbf{k}$ (different permutations do not add new terms to the above sum). Since

$$\sum_{i=1}^{N} \mathbf{k} \times \mathbf{e}_i = \mathbf{k} \times \left[\sum_{i=1}^{N} \mathbf{e}_i \right] = \mathbf{k} \times \mathbf{k} = 0, \tag{15}$$

Eq. (13) indeed equals 0 for all $\mathbf{k} \in I$. Note that each spectral coefficient appears in an infinite number of conservation laws, which provide infinite constraints for the individual spectral coefficients over the time evolution of the flow generated by Eq. (8).

In order to be able to integrate spectral equations the wave modes taken into account should form a finite set, that is, a finite number of equations and spectral coefficients should be considered. The components of the vector of dynamical variables are the spectral coefficients $\zeta_{\mathbf{m}}$ with $\mathbf{m}$ taking its value from the lattice $-T \le m_1, m_2 \le T$. The main difficulty is the suitable modifi-

cation of the structure functions. This modification is unavoidable because there are always index pairs **m** and **n** in the truncated system for which **m** + **n** is out of the lattice. To our knowledge there is no superior truncation strategy, the different schemes correspond to different "best fit" criteria (*Machenhauer*, 1991). The general form of the modified structure functions is

$$\mathbf{D}_{\mathbf{mn}} = C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})}(\mathbf{m},\mathbf{n})\,\zeta_{F(\mathbf{m},\mathbf{n})}, \tag{16}$$

where the function $F(\mathbf{m},\mathbf{n})$ takes its value from the set of the retained indices in the truncated system.

4.2 Traditional truncation strategies

Traditional truncation strategies use $F(\mathbf{m},\mathbf{n}) = \mathbf{m} + \mathbf{n}$ and $C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})} = \mathbf{m} \times \mathbf{n}$ if $|m_1 + n_1| \le T$, $|m_2 + n_2| \le T$ and attempt to find optimal definitions for the functions $F(\mathbf{m},\mathbf{n})$ and $C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})}$ if $|m_1 + n_1| > T$ or $|m_2 + n_2| > T$.

There are two basic approaches to truncate the system of equations at a wavenumber T.

4.2.1 Aliasing-free truncation

The strategy which minimizes the mean square residue, the difference between the two sides of the truncated equation, is the aliasing-free truncation (*Machenhauer*, 1991). The structure function $\mathbf{D}_{\mathbf{mn}}$ is set to zero if $|m_1 + n_1| > T$ or $|m_2 + n_2| > T$, otherwise it is defined by Eq. (10). The fact that the residue can only be minimal, but not zero, reflects the inevitable approximating character of the scheme.

4.2.2 Aliased truncation

The aliased truncation is a "best fit" strategy in the sense that all interactions between the retained wave modes are taken into consideration and the conservation of quadratic invariants is ensured. In this case $F(\mathbf{m},\mathbf{n}) = (\mathbf{m} + \mathbf{n})\, mod M$, where $M = 2T + 1$, and $\mathbf{a}\, mod M = \mathbf{b} = (b_1, b_2)$ with b_i being the least in absolute value integer congruent with a_i modulo M (i = 1,2). There are different possible choices for $C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})}$, a straightforward one would be $C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})} = [(\mathbf{m} + \mathbf{n}) mod M] \times \mathbf{n}$. A scheme with this choice, however, does not conserve the energy, because the corresponding structure matrix is not

skew-symmetric. This problem can be solved by setting $C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})} = (\mathbf{m} \times \mathbf{n}) mod M$. This approach also preserves the enstrophy and hereafter the notion aliased truncation corresponds to this scheme. Note that $\mathbf{D}_{\mathbf{mn}}$ is different from its original untruncated form, only if $|m_1 + n_1| > T$ or $|m_2 + n_2| > T$.

4.2.3 Structure of the traditional schemes

Both techniques have the limitation misinterpreting interactions between modes $\zeta_{\mathbf{m}}$ and $\zeta_{\mathbf{n}}$ if $|m_1 + n_1| > T$ or $|m_2 + n_2| > T$. In the aliasing-free case these interactions are neglected, while in the second case the aliased interactions affect different modes than in the untruncated equations. In the numerical practice the first approach dominates, thanks to the unambiguous results of the numerical experiments of *Orszag* (1971) and to the fact that it is a least square approximation. *Lorenz* (1960) pointed out that the aliasing-free scheme ensures the conservation of quadratic invariants, provided that all possible interactions between a triad of retained wavenumbers are taken into account, even if additional interactions are disregarded. The aliasing free truncated system in general, however, is not Hamiltonian. (For proof see Appendix A.)

Roughly speaking, the aliasing-free truncation violates the algebraic structure because it neglects numerous interactions between the components of the state vector. It suggests that the aliased truncation, which is free of the above error, may preserve the structure. Unfortunately, it is not the case. Although it also preserves the quadratic invariants, the aliased truncation for $T \geq 2$ does not satisfy the Jacobi identity, i.e. it is not Hamiltonian[2]. The maximal truncation $T = 1$ seems to be an exceptional case in the sense that it does satisfy the Jacobi identity.

A finite-difference scheme is an aliased truncated spectral scheme, though the function $C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})}(\mathbf{m},\mathbf{n})$ for it may have a substantially different form from what was used in this section and the structure function $\mathbf{D}_{\mathbf{mn}}$ for $|m_1 + n_1| \leq T$ and $|m_2 + n_2| \leq T$ may also be modified. Spectral transforms of energy and enstrophy conserving finite difference schemes have the property that the function $C_{\mathbf{mn}}^{F(\mathbf{m},\mathbf{n})}(\mathbf{m},\mathbf{n})$ satisfies the conditions for the preservation of quadratic invariants (for an example see the spectral form of the nine-point Arakawa Jacobian in *Bennett* and *Middleton* (1983).

---

[2] To our knowledge the fact that spectral truncation destroys the Jacobi identity was first reported by *Morrison* (1981) in a paper that seems to be forgotten. He also proposed a structure preserving truncation strategy based on the use of Clebsch potentials, but this method doubles the dimension of the space and the practical applicability of this approach has never been tested.

4.2.4 Conservation laws of traditional schemes

Since the traditional schemes conserve the quadratic invariants, the main problem is whether they can preserve the truncated analogues of the Casimir invariants $\mathcal{C}^N$ (Eq. 11) for $N > 2$. The following statement can be made about the conservation laws of traditional truncation methods: *The aliasing-free traditional truncation does not preserve the finite analogues of the Casimirs* $\mathcal{C}^N$ *for* $N > 2$ *at any resolution. (For proof see Appendix B.)* It is also true for the aliased truncation introduced above (the proof is not detailed in this paper).

It is apparent from the proofs of these statements that for the violation of the Jacobi identity and the conservation laws the misrepresented interactions are responsible. Most of the misrepresented interactions are related to wavenumbers close to the truncation limit. Therefore, if the absolute value of the high-wavenumber spectral coefficients were small enough over the whole period of integration, the structure and the preservation of higher order Casimirs may be improved.

Take the aliasing-free truncation as an example. If all spectral coefficients associated with wavenumbers larger than T/2 are kept somehow zero then for $|\mathbf{k}| \leq T/2$ Eq. (12) is valid because the $\zeta_{\mathbf{e}_i}$-s of a nonzero N-fold product occur in all possible permutations in Eq. (13). On the other hand the components in the gradient of the Hamiltonian that are related to $|\mathbf{k}| > T/2$ are also zero. This guarantees the preservation of the Casimirs according to Eq. (7). Notice, however, that these invariants would not be Casimirs in the sense of Eq. (2), because their conservation were not independent of the particular form of the Hamiltonian. Unfortunately, however, the ideal fields with zero large wave-number coefficients cannot be maintained.

Since Liouville's theorem and ergodicity apply for these schemes, the time average of the spectral energy distribution can be estimated with the canonical equilibrium phase-space average as it was given by *Kraichnan* (1967). If the ratio of the average wavenumber $k^* = \sqrt{C^2/H}$ and the truncation wavenumber T is small enough, the energy spectrum has a maximum at large scales and declines towards the small scales proportionally to the reciprocal of the wavenumber (*Fox* and *Orszag*, 1973). Since the number of wave vectors in a wavenumber band increases linearly with increasing wavenumber, a significant amount of energy can remain on the small scale part of the spectrum in the canonical distribution. On the other hand, if the initial field is such that the energy is only on given scales and the cutoff wavenumber is large enough compared to these scales then it takes a while for the energy to flow to higher wavenumbers, and until this happens the large scale coefficients and their interaction will dominate. Thus for a flow dominated by a given scale the conservation of the structure and higher order Casimirs can be improved for a

finite-time interval by increasing the resolution. In this respect the advantage of using higher resolution numerical weather prediction models is not the enhanced representation of small scales but the improved control of the synoptic scale motions through better conservation of the invariants.

In numerical practice the favorable steep energy spectrum is also maintained by the introduction of scale-selective viscosity terms. These terms were originally applied to reduce the effects of "blocking errors" related to the nonrepresented interactions in the aliasing-free models (*Machenhauer*, 1991) and to eliminate the aliasing from the aliased schemes (*Orszag*, 1971). From our point of view, however, they can also help to improve the conservation of the structure in the inertial range. On the other hand, these viscosity terms constantly decrease the enstrophy indicating the destruction of the Casimir invariants. At the same time dissipation has only a slight effect on the conservation of energy, which leads to the decrease of k^*. The energy spectrum becomes steeper and steeper, thus the scale-selective dissipation can cause an increase at the large scale part of the energy spectrum. This is the "tail wagging the dog" effect described by e.g. *Frederiksen et al.* (1996). Long term integration of relevant atmospheric models have to account for dissipative processes anyway, but one must recall that the role of the viscosity and diffusion terms is in one part to simulate the realistic dissipation, in another to control the truncation errors. Consequently, traditional truncations can accurately simulate the conservative Hamiltonian nature of the inertial wavenumber range between the forced and dissipative scales, if the cut-off wavenumber is large enough and the initial data are properly processed.

4.3 Structure preserving truncation strategies

4.3.1 The sine-bracket equation

So far only one truncation of the vorticity equation has been shown to preserve the Hamiltonian structure of the equations, and there are strong indications that this may be a unique one (*Zeitlin*, 1991; *Dowker* and *Wolski*, 1992; *McLachlan*, 1993). This sine-bracket truncation first proposed by *Zeitlin* (1991) is named after the underlying algebraic structure that was first introduced by *Fairlie et al.* (1989) and *Fairlie* and *Zachos* (1989). The governing equation is a slight modification of the aliased truncation of Eq. (8),

$$\frac{d\zeta_{\mathbf{m}}}{dt} = \sum_{\substack{n_1,n_2=-T \\ \mathbf{n}\neq\mathbf{0}}}^{T} \frac{1}{\epsilon} \frac{\sin(\epsilon\,\mathbf{m}\times\mathbf{n})}{|\mathbf{n}|^2}\, \zeta_{\mathbf{m}+\mathbf{n} mod M}\, \zeta_{-\mathbf{n}}, \tag{17}$$

where $\epsilon = 2\pi/M$ and $M = 2T+1$. The entries of the structure matrix are

$$\mathbf{D}_{\mathbf{mn}} = \frac{1}{\epsilon} \sin(\epsilon\, \mathbf{m} \times \mathbf{n})\, \zeta_{\mathbf{m}+\mathbf{n} mod M}\,. \tag{18}$$

Due to the feature of the sine function the structure matrix is obviously skew-symmetric and the Jacobi identity also holds. Hence Eq. (17) is a Hamiltonian system.

Note that in the limit $\epsilon \to 0$ (as $T \to \infty$) the term in Eq. (17) associated with given $\mathbf{m}$ and $\mathbf{n}$ wave vectors will be identical with the corresponding term of Eq. (8). The value of $\epsilon \mathbf{m} \times \mathbf{n}$ in general, however, will not tend to zero with increasing resolution. The smoothness of the vorticity field, i.e. the fact that the spectral coefficients tend to zero as the wave vectors grow to infinity, would ensure the convergence of Eq. (17) to Eq. (8), (*Miller et al.*, 1992; *Hoppe*, 1989), but this ideal case cannot be maintained in an inviscid fluid (see previous section).

4.3.2 Invariants of the sine-bracket equation

The conservation of energy in the system Eq. (17) is a straightforward consequence of the skew-symmetry of the structure matrix. The most important new feature of the sine-bracket equation in contrast to the traditional truncations is that it preserves the Jacobi condition and possesses $2T$ independent Casimir invariants. Their general form is

$$C_s^N = \sum_{I^N} \zeta_{\mathbf{i}_1} \cdots \zeta_{\mathbf{i}_N} \cos\left(\epsilon\, 2A(\mathbf{i}_1,\ldots,\mathbf{i}_N)\right), \tag{19}$$

$$I^N = \{(\mathbf{i}_1,\ldots,\mathbf{i}_N),\ \sum_{j=1}^{N} \mathbf{i}_j = \mathbf{0} mod M\}, \tag{20}$$

where $2 \le N \le 2T$ and $A(\mathbf{i}_1,\ldots,\mathbf{i}_N)$ denotes the area spanned by the index vectors

$$A(\mathbf{i}_1,\ldots,\mathbf{i}_N) = \frac{1}{2}\left(\mathbf{i}_2 \times \mathbf{i}_1 + \mathbf{i}_3 \times (\mathbf{i}_1 + \mathbf{i}_2) + \ldots + \mathbf{i}_N \times (\mathbf{i}_1 + \ldots + \mathbf{i}_{N-1})\right). \tag{21}$$

(For proof see Appendix C.) Note that C_s^2 is the enstrophy.

Again, the cosine factor would tend to unit for a given index set if $T \to \infty$, but C_s^N will be identical with the Casimirs $\mathcal{C}^N$ of Eq. (8) for $2 < N \le 2T$ in the limit $T \to 0$ only if the vorticity field were smooth.

The impressive properties of structure conservation and of having an increasing number of invariants with increasing resolution suggest that the time

evolution of the sine-bracket system is controlled by an increasing number of natural constraints that should be important for the long-term model integrations. On the other hand it eliminates the structure-destroying effects of aliasings at the price of distorting the well represented interactions. Moreover, the invariants at low resolution substantially differ from the invariants of the original system Eq. (8), and in contrast to the quadratically growing number of dynamical variables the number of extra invariants increases only linearly with resolution. As the resolution increases Eq. (17) gives better and better approximation of Eq. (8) for the large scale part of the spectrum, but with suitably processed initial data it is also true for the traditional schemes. The theoretical considerations and numerical experiments of *Hattori* (1993) showed that the extra invariants of the sine-bracket truncation have only slight impact on the statistical mechanics of the finite-mode vorticity equation.

5. Truncation strategies to quasi-geostrophic layer models

It was shown in *McLachlan et al.* (1997) that the quasi-geostrophic layer models also have Hamiltonian structure and an infinite number of invariants. Albeit simple, these are important models of the atmosphere, because they are capable to simulate baroclinic instabilities even in case of only two layers and low resolution. Fortunately, similar truncation methods can be applied as to the two-dimensional vorticity equation, moreover, a slightly modified version of the same time integrator can be used for the numerical integrations. We briefly describe the model equations and invariants of the N-layer model, a more detailed description can be found in *Pedlosky* (1987) and *McLachlan et al.* (1997).

The N-layer quasi-geostrophic model consists of N linearly coupled layers with different but constant densities. The velocity field and the pressure field are connected through the geostrophic relation and the stream function is identified with the pressure. The continuous equations governing this system are

$$\frac{\partial q^{(i)}}{\partial t} = \partial(q^{(i)},\psi^{(i)}),\ i = 1,...,N, \tag{22}$$

where $\partial(f,g) = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}$, and the potential vorticity $q^{(i)}$ and the stream function $\psi^{(i)}$ of the i-th layer ($i = 1,...,N$) are connected through

$$q^{(i)} = -D^{(i)}\{\nabla^2\psi^{(i)} - F^{(i)}\{(1-\delta_{iN})(\psi^{(i)} - \psi^{(i+1)}) + (1-\delta_{i1})(\psi^{(i)} - \psi^{(i-1)})\}\}. \tag{23}$$

Here $F^{(i)}=f^2L^2/D^{(i)}g(\Delta\rho/\rho_0)$, where f is the Coriolis parameter, g is the acceleration due to gravity, L is a characteristic horizontal scale, ρ_0 and $\Delta\rho$ are a characteristic value of the density and its variation between neighboring layers, respectively, and $D^{(i)}$ is the relative thickness of the i-th layer in the absence of motion $\left(\sum_{i=1}^{N} D^{(i)}=1\right)$. Note that assuming equal density jumps between neighboring layers $F^{(i)}D^{(i)}=F^{(j)}D^{(j)}$ for all $i,j\in\{1,\ldots,N\}$.

Transformed via two-dimensional Fourier transformation with periodic boundary conditions, the equivalent ordinary differential equations are

$$\dot{q}_{\mathbf{m}}^{(i)}=\sum_{\mathbf{n}}(\mathbf{m}\times\mathbf{n})\,q_{\mathbf{m}+\mathbf{n}}^{(i)}\,\psi_{-\mathbf{n}}^{(i)}, \tag{24}$$

where $\mathbf{m}\times\mathbf{n}=m_1n_2-m_2n_1$ and $q_{\mathbf{m}}^{(i)}$, $\psi_{\mathbf{m}}^{(i)}$ are the spectral coefficients of the i-th layer potential vorticity and stream function, respectively, associated with the wave vector $\mathbf{m}$.

The spectral transforms of the potential vorticity and the stream function are related by

$$\left(\psi_{\mathbf{n}}^{(1)},\ldots,\psi_{\mathbf{n}}^{(N)}\right)^T=\mathbf{B}_{\mathbf{n}}\left(q_{\mathbf{n}}^{(1)},\ldots,q_{\mathbf{n}}^{(N)}\right)^T\equiv\mathbf{B}_{\mathbf{n}}\mathbf{q}_{\mathbf{n}} \tag{25}$$

with $\mathbf{B}_{\mathbf{n}}$ being the inverse of the following symmetric tridiagonal matrix

$$\begin{bmatrix} D^{(1)}(n^2+F^{(1)}) & -\hat{F} & & & & & \\ -\hat{F} & D^{(2)}(n^2+2F^{(2)}) & -\hat{F} & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -\hat{F} & D^{(N-1)}(n^2+2F^{(N-1)}) & -\hat{F} & \\ & & & & -\hat{F} & D^{(N)}(n^2+F^{(N)}) & \end{bmatrix}, \tag{26}$$

where n^2 stands for $|\boldsymbol{n}|^2$ and $\hat{F}$ for $D^{(i)}F^{(i)}$. Since the Hamiltonian of this system is

$$H = \frac{1}{2} \sum_{\mathbf{n}} \mathbf{q}_{\mathbf{n}}^{T} \mathbf{B}_{\mathbf{n}} \mathbf{q}_{-\mathbf{n}}, \tag{27}$$

and $\mathbf{B}_{\mathbf{n}}$ is also symmetric,

$$\psi_{-\mathbf{n}}^{(i)} = \left(\nabla_{\mathbf{q}_{\mathbf{n}}} H\right)^{(i)}. \tag{28}$$

Thus Eq. (24) defines a Hamiltonian system for each layer with a corresponding structure matrix

$$\mathbf{D}_{\mathbf{mn}}^{(i)} = (\mathbf{m} \times \mathbf{n})\, \mathbf{q}_{\mathbf{m}+\mathbf{n}}^{(i)}. \tag{29}$$

In this model the potential vorticity of each layer is an individually conserved quantity, thus the integrals of any functions of the potential vorticity give the integrals of motion. The Casimir invariants of the system with spectral coefficients read

$$Q^{(i)k} = \sum_{I^k} q_{\mathbf{i}_1}^{(i)} \cdots q_{\mathbf{i}_k}^{(i)}, \quad (i = 1,\ldots,N), \tag{30}$$

$$I^k = \{(\mathbf{i}_1,\ldots,\mathbf{i}_k) \,|\, \sum_{j=1}^{k} \mathbf{i}_j = \mathbf{0}\}. \tag{31}$$

As it is described in *McLachlan et al.* (1997) for the two-layer system, there exists a similar structure-preserving truncation to the N-layer model as to the two-dimensional vorticity equation:

$$\dot{q}_{\mathbf{m}}^{(i)} = \sum_{\mathbf{n} \in \mathbf{I}} (1/\epsilon) \sin(\epsilon\, \mathbf{m} \times \mathbf{n})\, q_{\mathbf{m}+\mathbf{n}}^{(i)} (\mathbf{B}_{-\mathbf{n}} \mathbf{q}_{-\mathbf{n}})^{(i)}, \tag{32}$$

where $\epsilon = 2\pi/M$, and the index set is defined by

$$\mathbf{I} = \{\mathbf{i} \in \mathbf{Z}^2 \,|\, -T \le i_1, i_2 \le T,\ \mathbf{i} \neq \mathbf{0},\ T = (M-1)/2\}. \tag{33}$$

The finite dimensional system obtained with the above truncation is Hamiltonian, and possesses $N \cdot 2T$ integrals of motion, but again the Casimirs of this system $Q_s^{(i)k}$ for $k > 2$ converge to the Casimirs of the original system in the limit $T \to \infty$ only if the potential vorticity fields are smooth. The Casimirs are

$$Q_s^{(i)k} = \sum_{\mathbf{I}^k} q_{\mathbf{i}_1}^{(i)} \cdots q_{\mathbf{i}_k}^{(i)} \cos[\epsilon\,(2A(\mathbf{i}_1,\ldots,\mathbf{i}_k))], \quad (i = 1,\ldots,N), \tag{34}$$

$$\mathbf{I}^k = \{(\mathbf{i}_1,\ldots,\mathbf{i}_k) \,|\, \sum_{j=1}^{k} \mathbf{i}_j = \mathbf{0} \, mod M\}, \tag{35}$$

where $2 \leq k \leq 2T$ and $A(\mathbf{i}_1,\ldots,\mathbf{i}_k)$ is given in Eq. (21).

In order to be able to extend our comparative study to the hierarchy of quasi-geostrophic layer models, we consider also a truncation that is the counterpart of the traditional truncation applied to the two-dimensional vorticity equation: retaining the same set of coefficients as in Eq. (32), the index $(\mathbf{m}+\mathbf{n}) mod M$ is used instead of $\mathbf{m}+\mathbf{n}$ and the coefficients of the nonlinear terms are modified to $(\mathbf{m}\times\mathbf{n}) mod M$ if $|m_1+n_1|>T$ or $|m_2+n_2|>T$. This truncation destroys the Hamiltonian structure, but ensures the conservation of energy and potential enstrophy.

6. Numerical experiments

6.1 Algorithm for integration

Numerical experiments were carried out in order to compare the time-depending characteristics of the structure preserving and of the traditionally truncated systems described in Section 4 and 5. One can argue that a fair comparison between the traditional and the structure preserving schemes should involve the aliasing-free traditional scheme, which is assumed to be superior to the aliased truncation of this study. The superiority of the aliasing-free schemes in the numerical practice, however, is not obvious since (1) the finite-difference models, though inevitably burdened by aliasing, proved to be competitive with the aliasing-free spectral schemes, moreover (2) there are numerical evidences that the aliased spectral schemes may perform even better than the aliasing-free methods (*Chen*, 1997).

A slightly modified version of McLachlan's explicit scheme (*McLachlan*, 1993) was used in our experiments for the integration of the two-dimensional vorticity equation (hereafter denoted by 2DV). It will be indicated in the descriptions of the experiments whether the first or the second order version of the time integrator was used. *McLachlan et al.* (1997) describes, that with some modification the integration algorithm used for the integration of the two-dimensional vorticity equation can be used for the structure preserving truncation of the quasi-geostrophic equation. We carried out experiments with the two-layer quasi-geostrophic model (hereafter denoted by L2 QG). For details of the

integration scheme see Appendix of Part I and *McLachlan et al.* (1997). The main advantages of this algorithm are its efficiency and the fact that it can be applied to the traditionally truncated aliased systems as well. This makes possible a fair comparison between the traditional and the structure preserving schemes even in extremely long term integrations.

6.2 Verification of the conservation properties

6.2.1 Conservation of the energy

Energy cannot be an exact invariant of the discrete Hamiltonian system even if the time stepping algorithm is symplectic. There is, however, a modified Hamiltonian in this case which is conserved and this ensures the boundedness of the energy error, in addition, the perturbation in the value of the Hamiltonian can be reduced to arbitrary small value with higher and higher order schemes (see Part I). Although the traditional spatial truncation conserves formally the energy and the enstrophy, the boundedness of energy errors over the time integration is not guaranteed. Possibly the fewer constraints exerted by the fewer formal invariants allow greater errors of energy. To investigate whether this conjecture is true, we calculated the relative energy error, $(H(t)-H(0))/H(0)$, and compared the results obtained with the differently truncated systems.

The models were integrated from randomly generated initial fields for 10^6 time steps at resolutions T1 and T2, and for 10^5 time steps at T13. The time step was uniformly set to 1.0 and the second order time integration scheme was used. Care should be taken when the results at the given resolutions are compared, because the uniform time step has different meaning for the different size systems. If the reciprocal of the average (potential-)vorticity $1/\bar{\zeta}$ is taken as a characteristic time unit, then the unit time step corresponds to $\bar{\zeta}$ nondimensional time units. Since a similarly defined characteristic time unit for the atmosphere is 10^5 sec, an atmospheric equivalent time step for the unit time step in the different size systems can be obtained by taking $\bar{\zeta} \times 10^5$ sec. The initial average vorticity was 0.09796 1/sec, 0.06091 1/sec and 0.00801 1/sec at T1, T2 and T13, respectively, in the 2DV model, while the average potential vorticity in the L2 QG model runs were of the magnitude 10^{-2} 1/sec at all three resolutions. With these initial fields, an equivalent atmospheric time step of 160 min., 102 min. and 13 min. can be attributed to the discretized systems at T1, T2 and T13 in case of the 2DV model and a time step of about 15 min. to the L2 QG model at all three resolutions.

Fig. 1 shows the relative energy errors for the structure preserving and the traditional truncation at T2 and T13. The result for the structure-preserving truncation of the model at T1 can be seen in Fig. 4a of Part I. At T1 the magnitudes of the relative errors obtained with the traditional truncation are the same (not shown), which can be the consequence of the fact that at T1 even the

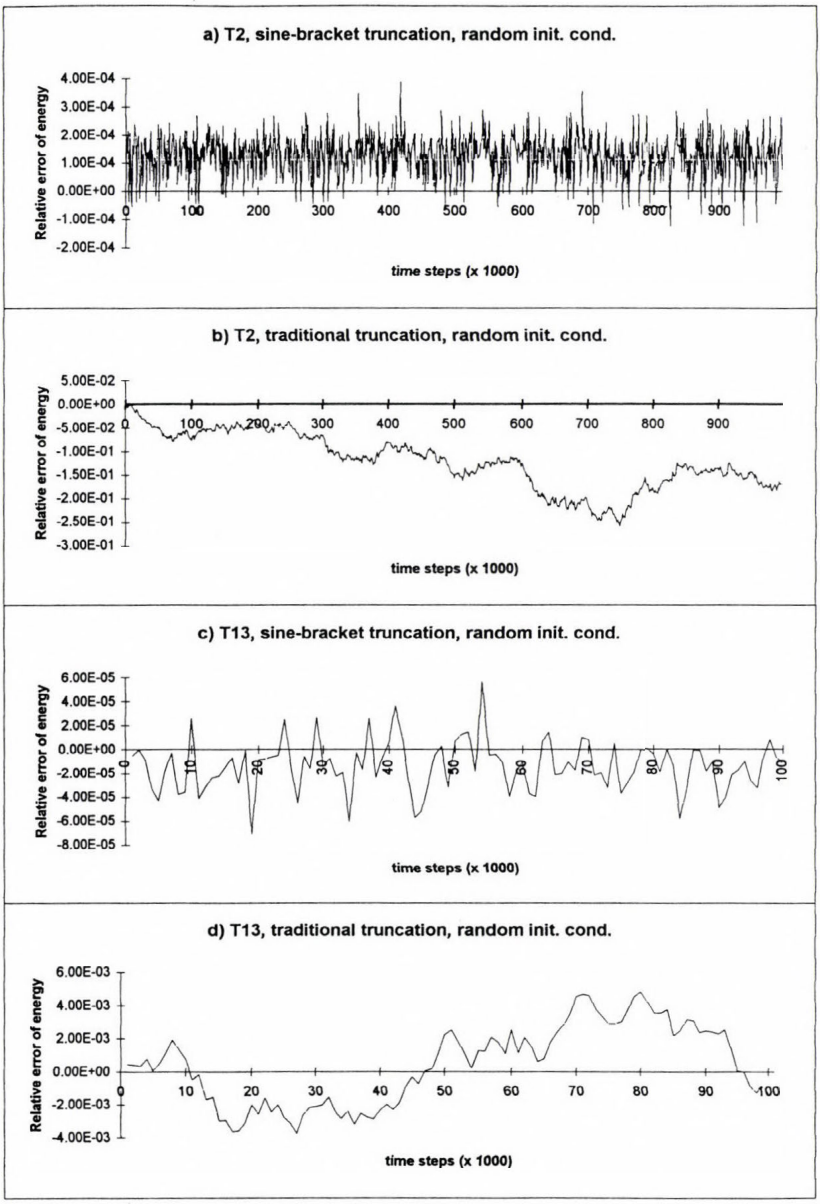


Fig. 1. Relative error of energy over a 10^6 time-step integration of the 2DV model at T2 started from randomly generated initial condition integrated with (*a*) the structure-preserving scheme, (*b*) the traditional scheme; relative error of energy over a 10^5 time-step integration of the 2DV model at T13 started from randomly generated initial condition integrated with (*c*) the structure-preserving scheme, (*d*) the traditional scheme.

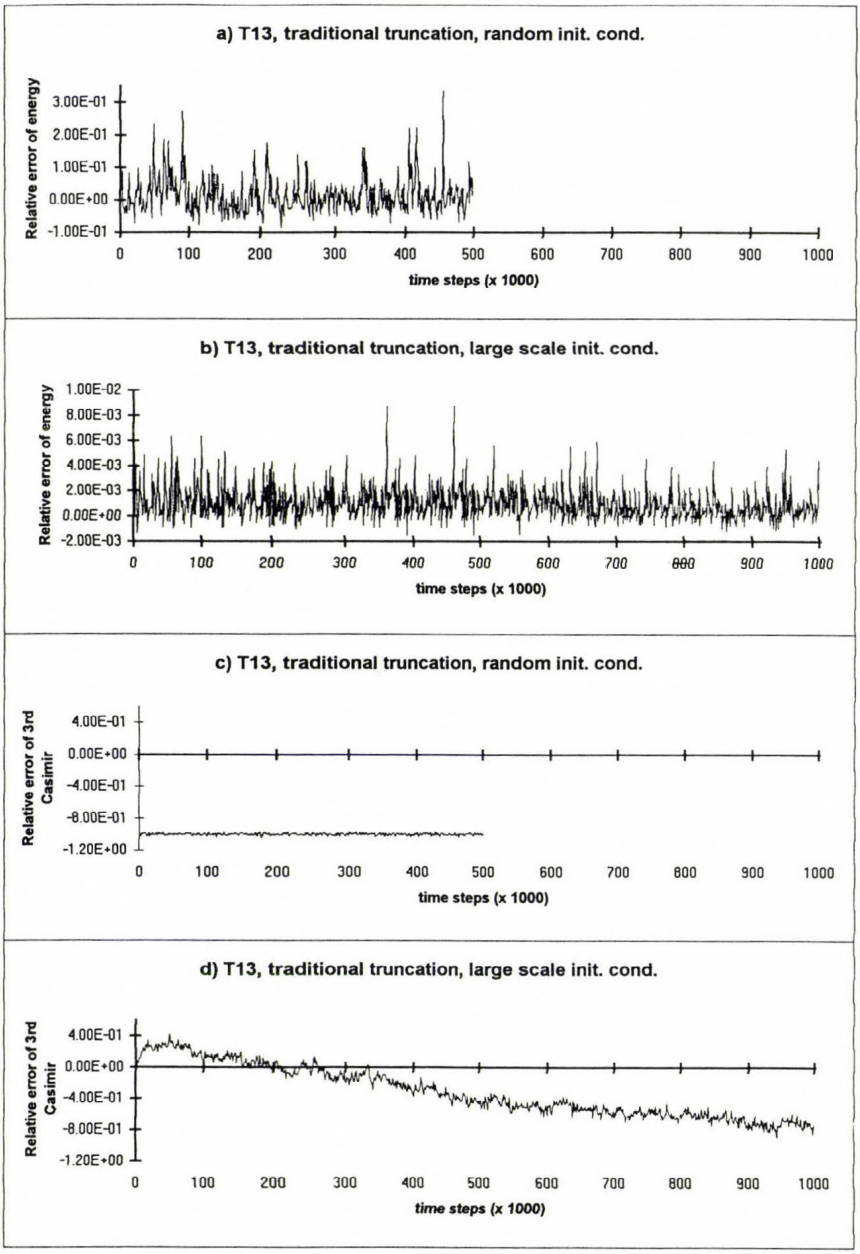


Fig. 2. Relative error of the energy and of the 3-rd Casimir (C^3) over a 5×10^5 time-step integration of the 2DV model at T13 integrated with the traditional scheme started from (*a*), (*c*) randomly generated initial condition, (*b*), (*d*) large scale ($|\boldsymbol{k}| \leq 3$) initial condition.

aliased truncation is Hamiltonian. Actually, the distinction "structure preserving" vs. "traditional" are not very useful in case of the maximum truncation, since it does not reflect the algebraic structure of the schemes. At T1 the relative errors oscillate around a nonzero value but do not exhibit any increasing tendency. At T2 there are profound differences between the performances of the schemes (see Fig. 1a, b). The error bound is three orders of magnitude larger and the relative error seems to have an increasing trend for the traditional scheme. The behavior of the structure conserving model is very similar to that of the T1 case. The difference is similar at T3, and T4 (not shown) but is somewhat less pronounced at T13 (Fig. 1c, d), especially with other initial conditions (see below the rotating cone experiment). At T13 the relative errors are oscillating around a nonzero value even in the traditionally truncated system. The relative energy curves obtained with the L2 QG model are not shown, since similar statements can be made about them as for the 2DV model runs.

It is not surprising that the truncation affects the results less at higher resolutions: a linearly growing number of constraints cannot control the evolution of a system with quadratically growing number of degrees of freedom. On the other hand, the constraints of statistical physics on the energy spectra can act only at higher resolution. Because of the validity of Liouville's theorem these constraints are present in both systems. As it was indicated in Section 4.2.4 until the ratio of the average and the cut-off wavenumber is small and the energy spectrum is steep enough toward small scales, an improvement in the preservation of structure in the traditional schemes can be achieved. This idea was tested by confining the energy in the initial condition to the large ($|\mathbf{k}| \leq 3$) scales at T13 (*Fig. 2a, b*). As expected, the energy error in the traditional system is two orders of magnitude smaller than for random initial conditions. It suggests that the effect of misrepresented interactions are less pronounced. On the other hand, similar improvement can be seen for the sine-bracket truncation due to the reduction of the role of the ill-represented interactions between the high wavenumbers. One must not forget, however, that the reduced errors of the time integration scheme for smoother fields can also contribute to the improvement of the results.

6.2.2 Conservation of the Casimir invariants

The conservation of Casimir invariants can also be characterized by their relative errors. The most important Casimir, the enstrophy (C^2) is not only a formal invariant, but it is a very effective numerical constraint for both truncation strategies. The error of the enstrophy over a time step is approximately equal to the expectation of the roundoff error for both models.

The higher order invariants, C_s^N-s of the sine-bracket model were computed from $N=3$ up to $N=2T$ for the different resolutions. The results of T2 are shown in *Fig. 3*. Due to the fact that these quantities are conserved within the roundoff error over a time step, they also impose strong constraints on the model variables. Note that the high accuracy of the Casimirs was found independent of the model resolution and of the order of the invariants. The special role of Casimir invariants in long-term numerical integration is based on this fact: the conservation of the Casimir invariants is independent of the systematic errors that occur in the model variables. The Casimirs remain independent integration constraints for any time. That is why the purely enstrophy preserving schemes have a better performance in the effective conservation of energy than those ensuring the formal invariance of only energy (see. e.g. *Sadourny*, 1975).

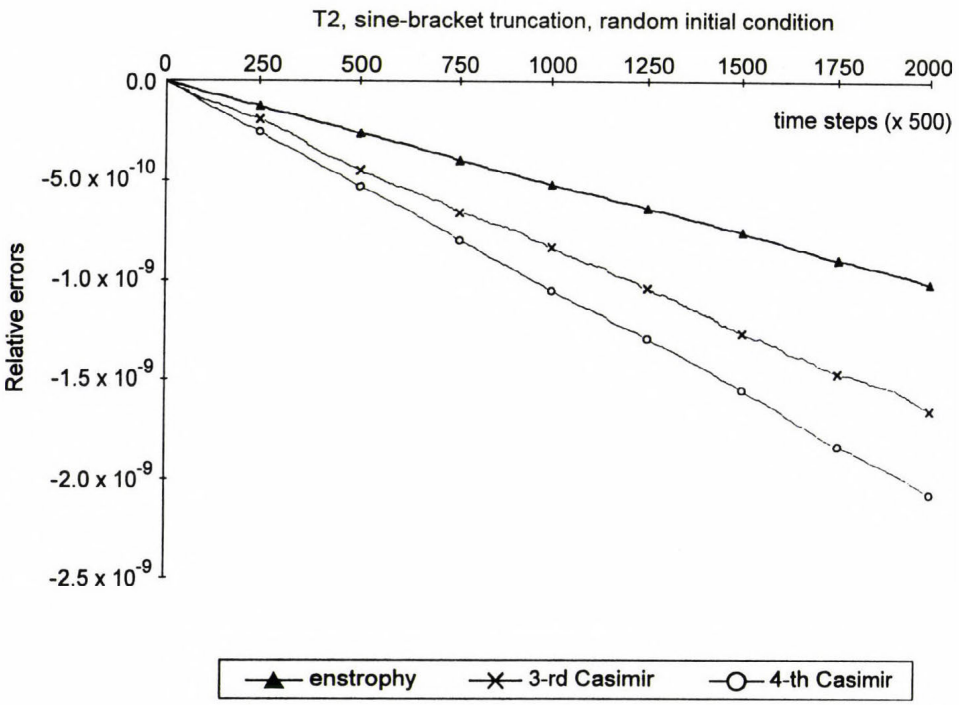


Fig. 3. Relative error of the first three Casimir invariants in the sine-bracket truncation of the 2DV model at T2. The initial condition was randomly generated, the integration was carried out for 10^6 time steps.

The smaller relative energy error for large scale initial fields can be attributed in parts to the smoothness of the vorticity field, thus it does not necessarily prove better structure conservation. A more appropriate indicator of the improvement in the preservation of structure is the relative error in the higher order Casimirs, which are not formal invariants in the traditional truncated system. *Fig. 2c, d* present the relative error of the finite analogue of Casimir $\mathcal{C}^3$ versus time for different initial conditions at T13. Started from random initial conditions the relative errors oscillate around a 100 percent right from the beginning. When the large scale random initial conditions are applied the maximum relative error is only about 40 percent during the first part of the integration. This indicates that there is indeed an improvement, though not substantial, in the preservation of structure in the traditional scheme for suitable initial conditions.

Another way of verifying conservation properties of a truncated system is to investigate the behavior of the discrete analogues of the continuous Casimir invariants in the physical space. As in *Hattori* (1993), we transformed back the spectral coefficients of vorticity to the physical space, then calculated the flatness (F) and skewness (S) parameters that are combinations of discrete analogues of continuous invariants (of $\mathcal{C}^4$, $\mathcal{C}^3$ and $\mathcal{C}^2$). Their definitions are the followings:

$$F = M^2 \frac{\sum_{\mathbf{k}\in\Omega} \zeta^4(\mathbf{k})}{\left[\sum_{\mathbf{k}\in\Omega} \zeta^2(\mathbf{k})\right]^2}, \qquad S = M \frac{\sum_{\mathbf{k}\in\Omega} \zeta^3(\mathbf{k})}{\left[\sum_{\mathbf{k}\in\Omega} \zeta^2(\mathbf{k})\right]^{3/2}}, \tag{36}$$

where $\zeta(\mathbf{k})$ denotes the vorticity at a grid point indexed by $\mathbf{k}$ of the lattice $\Omega = \{\mathbf{k} \mid -T \le k_1, k_2 \le T,\ T = (M-1)/2\}$. These quantities are frequently used to characterize vorticity distributions in a statistical sense. In our approach they can be viewed as discrete approximations to quantities that are individually conserved in the continuous case. (Note that the denominator is the power of enstrophy, which is conserved even in the truncated systems to a high degree of precision). We calculated the values of these parameters for different initial conditions in the two differently truncated systems. The results for the 2DV model at T13 are summarized in *Fig. 4*.

One apparent common feature on the panels is that the individual conservation laws are violated in all cases. Although the structure preserving schemes conserves their Casimirs accurately, these invariant quantities differ from the discrete analogues of the Casimirs of the original system, and their preservation does not guarantee the conservation of the latters.

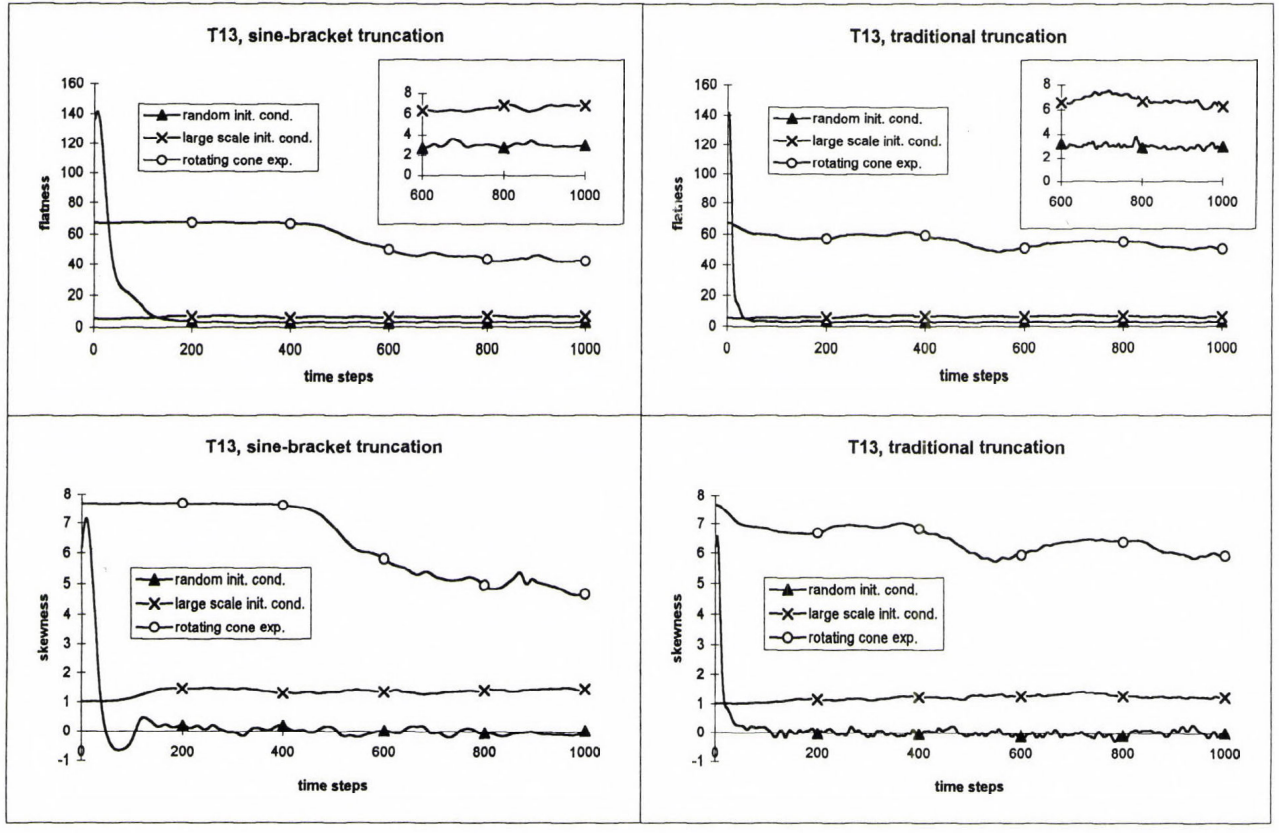


Fig. 4. Flatness and skewness (Eq. (36)) parameters in the 2DV model truncated with the traditional and the sine-bracket truncation at T13 integrated from different initial conditions (indicated on the curves) for 10^3 time steps.

When the initial fields are randomly generated the value of the parameters tend to those characterizing normal distribution ($F=3.0$ and $S=0.0$) in both truncated systems. The evolution of vorticity distribution, however, seems to be highly dependent on the initial condition — at least on the time scale represented by 10^3 time steps. In case of large scale initial field or in the rotating cone experiment (described below), the investigated statistical parameters are far from that of a normal distribution. There are differences in details, but the behavior of the two schemes in case of large scale initial condition is qualitatively similar. In both system the distribution is broader than the normal ($F \geq 3.0$) and less symmetric ($S \geq 0.0$, i.e skewed to the right).

In the third experiment the initial field was the discrete approximation of the vorticity field defined as

$$\zeta(\mathbf{r},t=0) = \begin{cases} 1 - \dfrac{|\mathbf{r}|}{r_0} & \text{for } |\mathbf{r}| \leq r_0 \\ 0 & \text{for } |\mathbf{r}| > r_0 \end{cases}, \tag{37}$$

where r_{01}, r_{02} are fixed numbers ($r_{01}=3$ and $r_{02}=7$ were used in the calculations at T13) and $\mathbf{r}=(r_1,r_2)$, $|\mathbf{r}| = \left((r_1-r_{01})^2+(r_2-r_{02})^2\right)^{\frac{1}{2}}$ and $r_0=(r_{01}^2+r_{02}^2)^{\frac{1}{2}}$. Visually, this field is a cone rotating around an axis at $\mathbf{r_0}=(r_{01},r_{02})$. The corresponding stream function field is

$$\psi(\mathbf{r},0) = \begin{cases} \dfrac{1}{4}|\mathbf{r}|^2 - \dfrac{1}{9r_0}|\mathbf{r}|^3 & \text{for } |\mathbf{r}| \leq r_0 \\ 0 & \text{for } |\mathbf{r}| > r_0 \end{cases}. \tag{38}$$

The latter can be written in the form

$$\psi(\mathbf{r},0) = \sum_{i=0}^{3} a_i \zeta^i(\mathbf{r},0) = \sum_{i=0}^{3} \frac{a_i}{i+1} \nabla_\zeta \mathcal{C}^{i+1}, \tag{39}$$

that is, this initial condition defines a steady state solution of the continuous vorticity equation which is constrained by the first four Casimir invariants (see Eq. (3)).

With this initial condition both statistical parameters have much larger values. An interesting fact, that the sine-bracket truncation succeeds in

preserving the initial value of the parameters up to about 400 time steps with this initial condition.

The cone defined in Eq. (37) is a steady solution of the continuous equation, and the initial fields are inevitably burdened with truncation errors. Unfortunately, the stability of this steady solution cannot be determined with the criteria given by Arnold for the nonlinear (Liapunov-)stability of the flow (see *Arnold*, 1989, 335 pp). Keeping in mind that the eventual breaking down of the initial vorticity pattern can be the result of the instabilities inherent in the initial conditions, we use this initial field for comparing the performance of the different schemes in terms of the effective conservation of the first four invariants and the accuracy of the individual variables. It can be expected, that if the conservation of the first four invariants in the two schemes differ substantially, the computed states in the two systems will also diverge. Since the results are affected by high truncation errors in the numerical integration, the comparative experiments were carried out at different high resolutions with the first order scheme. The cone was defined in the physical space and the spectral coefficients were computed via Fourier interpolation at a spectral truncation T13, T40, and T62. The model was integrated for 2×10^5 time steps at T13. The time evolution of the fields is shown in *Fig. 5a, c* and *Fig. 5b, d* for the structure preserving and the traditional scheme, respectively. On the base of the results neither of the schemes can be judged superior to the other. After some initial jumping the energy of both systems oscillate around an energy level about 3 percent lower than the initial level (not shown). The magnitude of energy fluctuations in the traditional model is roughly half of that in the structure conserving one. The cone completely disappears by the end of the integration. The T40 and T62 versions of both models were integrated for 5×10^3 time steps. In these experiments a cone of the same size was situated at the same part of the field as in the T13 case. *Fig. 5e-l* show the resulting vorticity and vorticity error fields at the final time. On the one hand the improvement of the results with increasing resolution is obvious, on the other the difference between the two models is hardly noticeable. The error fields in the sine-bracket model are a bit smoother and have slightly smaller extrema.

7. Summary and discussion

In this paper the effect of spatial truncation on the Eulerian atmospheric model equations was examined. We pointed out that traditional strategies inevitably violate the Jacobi identity, thus destroy the underlying Hamiltonian structure of the untruncated equations. At the same time these truncation methods necessarily ruin the preservation of nonquadratic Casimir invariants. As it was shown in Section 4 both the Jacobi identity and the conservation of Casimirs fail because of the misinterpreted interactions between wavemodes whose structure

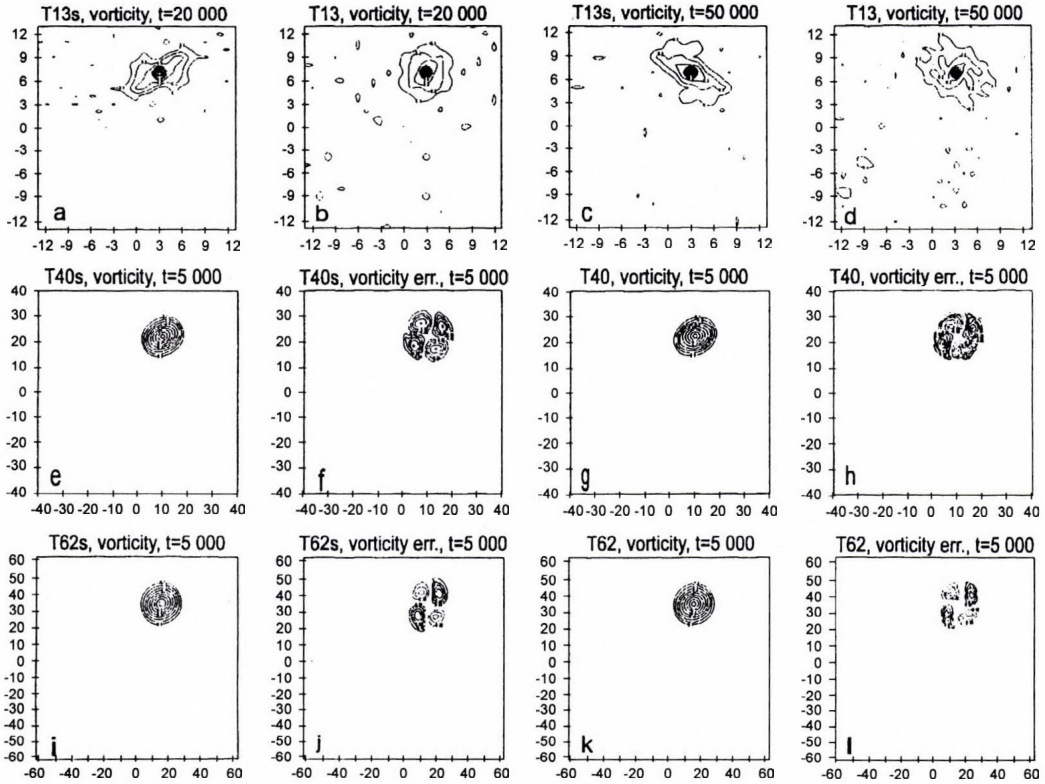


Fig. 5. Vorticity in the physical space in the rotating cone (Eq. 36) experiment (*a–d*) at T13, (*e–f*) at T40 and (*i–l*) at T62. Panels (*a*) and (*b*) show the vorticity field after $2\text{x}10^4$ time steps, while (*c*) and (*d*) correspond to the $5\text{x}10^4$-th time step in the structure-preserving and the traditional truncation, respectively. Panels (*e*), (*f*) and (*g*), (*h*) present the vorticity field and its error after $5\text{x}10^3$ time steps in the structure preserving and the traditionally truncated system, respectively, at T40. (*i*), (*j*) and (*k*), (*l*) is the same at T62.

function is distorted by the truncation. The number of misrepresented interactions is the largest for wavemodes that have a component close to the upper cutoff wavenumber. Consequently, an improvement in the conservation of the Hamiltonian structure can be achieved by reducing the role of the falsely represented interactions. This requires that the energy of the higher-wavenumber modes should be kept at a relatively low level. At this point an analogy between the higher-wavenumber modes and free gravity wave modes of the primitive equation models can be drawn. In both cases the problem can be handled by processing appropriately the initial data and continuously dissipating the energy of the problematic modes. On the other hand, the analogue of the filtered models is the structure preserving sine-bracket equation. The price for the preservation of structure and conservation of more invariants, however, must be paid: the sine-bracket equation is not aliasing-free, the high wavenumber interactions are distorted, and the higher order invariants ($N > 2$) converge to the Casimirs of the untruncated system only in the limit $T \to \infty$ and if the vorticity field is smooth. Operational and research models of today follow the first strategy using higher and higher cutoff wavenumbers and scale selective viscosity and diffusion terms. The small scale structure of such models, however, should be handled very carefully because the interactions associated with them are not constrained by higher order vorticity-type invariants. This fact might contribute to the highly unstable evolution of perturbations initially dominated by small scale structures (*Hartmann et al.*, 1995). The problem whether structure preserving schemes offer a better handling of these deficiencies can be investigated only on the base of extensive numerical experiments.

Another important aspect for the operational practice is the computational efficiency of the schemes. Executing all operations in the spectral space is supposed to have two main disadvantages: the high cost of evaluating the nonlinear terms and the difficulty to calculate some of the physical parameterization (see e.g. *Haltiner* and *Williams*, 1980). Although McLachlan's algorithm keeps the whole computational procedure in the spectral space, it reduces, though does not eliminate, the first problem to solving an equivalent diagonal system of linear equations. An additional advantage of this algorithm is that it can be even more efficient in a parallel computational environment. While it cannot be foreseen whether similar efficient algorithms can be developed for more sophisticated systems and how the problem of physical parameterization can be handled for these schemes, this method demonstrates that the utilization of hidden symmetries in the equations may result in a substantial improvement of the computational efficiency.

For theory oriented studies focusing on the phase space geometry, however, the algebraic structure preserving schemes are rather appealing. For instance, the Hamiltonian structure of the five-component Lorenz' model proved to have a crucial role in the precise definition of the slowest invariant manifold (*Bokhove* and *Shepherd*, 1996). This study, nonetheless, made use of the fact

that the maximum truncation traditional scheme, a real vorticity-triad, has Hamiltonian structure. As we stated in the main body of this paper for $T \geq 2$ there is no traditional truncated scheme, which retains the algebraic structure.

Finally, let us mention that while this study focused on the truncation of the Eulerian atmospheric equations there is another approach, the truncation of the Lagrangian equations themselves. There seems to be three different ways to realize this program: (1) application of the sine-bracket truncation to the Lagrangian equations (*Rouhi* and *Abarbanel*, 1993), (2) discretization of the action principle (*Holm et al.*, 1985) and (3) geometric models based on the semi-geostrophic approach (*Purser*, 1997 and references therein). Notice, however, that while the first two approaches have the appealing feature of preserving the structure, they destroy the particle relabelling symmetries and the related conservation laws. The most promising approach retaining the particle relabelling symmetries and the individual conservation laws is based on the semi-geostrophic equations. Opponents of this approach argue that (1) in contrast to the geostrophic equations the semi-geostrophic equations cannot be derived from the primitive equations by a formal expansion with respect to a small parameter as the Rossby number and (2) the semi-geostrophic equations are singular at the equator. Neither is an obstacle to applying this approach to a series of dynamical meteorological problems for which the Lagrangian approach is more relevant (e.g. for simulating the development of a frontal zone).

Acknowledgements—This work was partly supported by a grant from State Scientific Research Fund (OTKA F 019268). Encouraged by the informal discussions great progress was achieved during the authors' stay at the Isaac Newton Institute for Mathematical Sciences, Cambridge, England, within the program "Mathematics of the Atmosphere and Ocean Dynamics" in the summer and autumn of 1996.

References

Arakawa, A. and *Lamb, V.R.*, 1980: A Potential Enstrophy and Energy Conserving Scheme for the Shallow Water Equations. *Mon. Wea. Rev. 109*, 18-36.

Arakawa, A., 1966: Computational design for long-term numerical integration of equations of fluid of motions: Two-dimensional incompressible flow. Part I. *J. Comput. Phys. 1*, 119-143.

Arnold, V.I., 1989: *Mathematical Methods in Classical Mechanics*. Springer-Verlag, New-York, 508 pp.

Bennett, A.F. and *Middleton, J.F.*, 1983: Statistical mechanics of a finite-difference approximation to the barotropic vorticity equation. *Quart. J. Roy. Meteorol. Soc. 109*, 795-808.

Bokhove, O. and *Shepherd, T.*, 1996: On Hamiltonian balanced dynamics and the slowest invariant manifold. *J. Atmos. Sci. 53*, 276-297.

Chen, X.S., 1997: Use of the aliased spectral model in numerical weather prediction. *Mon. Wea. Rev. 125*, 2998-3007.

Dowker, J.S. and *Wolski, A.*, 1992: Finite model of two-dimensional ideal hydrodynamics. *Phys. Rev. A 46*, 6417-6430.

Fairlie, D.B., *Fletcher, P.* and *Zachos, C.K.*, 1989: Trigonometric structure constants for new infinite-dimensional algebras. *Phys. Lett. B. 218*, 203-206.

Fairlie, D.B. and *Zachos, C.K.*, 1989: Infinite-dimensional algebras, sine-brackets, and SU(∞), *Phys. Lett. B 224*, 101-107.

Fox, D.G. and *Orszag, S.A.*, 1973: Inviscid dynamics of two-dimensional turbulence, *Phys. Fluid. 16*, 169-171.

Frederiksen, J.S., *Dix, M.R.* and *Kepert, S.M.*, 1996: Systematic energy errors and the tendency toward canonical equilibrium in atmospheric circulation models. *J. Atmos. Sci. 53*, 887-904.

Haltiner, G.J. and *Williams, R.T.*, 1980: *Numerical Prediction and Dynamic Meteorology*. J. Wiley, New York, 477 pp.

Hartmann, D.L., *Buizza, R.* and *Palmer, T.N.*, 1995: Singular vectors: The effect of spatial scale on linear growth of disturbances. *J. Atmos. Sci. 52*, 3885-3894.

Hattori, Y., 1993: Effect of phase dependent invariants and ergodicity in finite-mode systems closely related with two-dimensional ideal flow. *J. Phys. Soc. Japan 62*, 2293-2305.

Holm, D.D., *Kupershmidt, B.A.* and *Levermore, C.D.*, 1985: Hamiltonian differencing of fluid dynamics. *Adv. Appl. Math. 6*, 52-84.

Hoppe, J., 1989: Diffeomorphism groups, quantization, and SU(∞). *Int. Journ. Mod. Phys. A 19*, 5235.

Kádár, B., *Szunyogh, I.* and *Dévényi, D.*, 1998: On the origin of model errors. Part I: Effects of the temporal discretization for Hamiltonian systems. *Időjárás 102*, 19-41.

Kraichnan, R.H., 1967: Inertial ranges in two-dimensional turbulence. *Phys. Fluid. 10*, 1417-1423.

Lorenz, E.N., 1960: Maximum simplification of dynamic equations. *Quart. J. Roy. Meteor. Soc. 12*, 243-254.

Machenhauer, B., 1991: Spectral methods. *Proceedings of ECMWF seminars on "Numerical methods in atmospheric models"*. Sept 9-13, Reading, UK. *Vol. 1*, 3-85.

McLachlan, R., 1993: Explicit Lie-Poisson integration and the Euler equations. *Phys. Rev. Lett. 71*, 3043-3046.

McLachlan, R.I., *Szunyogh, I.* and *Zeitlin, V.*, 1997: Hamiltonian finite-dimensional models of baroclinic instability. *Phys. Lett. 229A*, 299-305.

Miller J., *P. B. Weichman* and *M. C. Cross*, 1992: Statistical mechanics, Euler's equation, and Jupiter's Red Spot. *Phys. Rev. A*, *45*, 2328-2359.

Morrison, P.J., 1981: Hamiltonian field description of two-dimensional vortex fluids and guiding center plasmas. *PPPL-1783*, Princeton University, Princeton, New Jersey.

Olver, P.J., 1989: *Applications of Lie Groups to Differential Equations*. Springer-Verlag, 513 pp.

Orszag, S.A., 1971: On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. *J. Atmos. Sci. 28*, 1074.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics*. Second edition, Springer, New York, 710 pp.

Purser, R.J., 1997: Legendre-transformable semi-geostrophic theories. *J. Atmos. Sci.* (under review).

Rouhi, A. and *Abarbanel, H.D.I.*, 1993: Symmetric truncations of the shallow-water equations. *Phys. Rev. E*, *48*, 3643-3655.

Sadourny, R., 1975: The dynamics of finite-difference models of the shallow-water equations. *J. Atmos. Sci. 32*, 680-689.

Salmon, R. and *Talley, L.D.*, 1989: Generalization of Arakawa's Jacobian. *J. Comput. Phys. 83*, 247-259.

Shepherd, T.G., 1990: Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Adv. Geophys. 32*, 287-337.

Szunyogh, I., 1993: Finite-dimensional quasi-Hamiltonian structure in simple model equations. *Meteorol. Atmos. Phys. 52*, 49-57.

Zeitlin, V., 1991: Finite-mode analogs of 2D ideal hydrodynamics: Coadjoint orbits and local canonical structure. *Physica D 49*, 353-362.

Zeitlin, V. and *Pasmanter, R.A.*, 1994: On the differential geometry approach to geophysical flows. *Phys. Lett. A*, *189*, 59-63.

Appendix A

Proofs of statements in Section 4.2. The fact that an aliasing-free truncated system in general is not Hamiltonian follows from the lemma below:

Lemma: *Let a subsystem of the aliasing-free truncated system consist of the variables corresponding to the wave vectors* $\mathbf{i}, \mathbf{j}, \mathbf{k}, -\mathbf{i}, -\mathbf{j}, -\mathbf{k}$ *satisfying the conditions* $\mathbf{i}+\mathbf{j}=\mathbf{k}$ *and* $\mathbf{i}\times\mathbf{j}\neq 0$. *This subsystem, hereafter referred to as a vorticity-triad, possesses the following properties:*

(*i*) *It is Hamiltonian if and only if the spectral coefficients are real.*

(*ii*) *It conserves the energy and the enstrophy.*

Proof: The structure matrix for the vorticity-triad subsystem is

$$
\begin{array}{c}
\\
\mathbf{i} \\
\mathbf{j} \\
\mathbf{k} \\
-\mathbf{i} \\
-\mathbf{j} \\
-\mathbf{k}
\end{array}
\begin{array}{c}
\begin{array}{cccccc} \mathbf{i} & \mathbf{j} & \mathbf{k} & -\mathbf{i} & -\mathbf{j} & -\mathbf{k} \end{array} \\
\left[\begin{array}{cccccc}
0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{k}} & 0^* & 0 & 0^* & (\mathbf{j}\times\mathbf{i})\zeta_{-\mathbf{j}} \\
(\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{k}} & 0 & 0^* & 0^* & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{-\mathbf{i}} \\
0^* & 0^* & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{j}} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{i}} & 0 \\
0 & 0^* & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{j}} & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{-\mathbf{k}} & 0^* \\
0^* & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{i}} & (\mathbf{j}\times\mathbf{i})\zeta_{-\mathbf{k}} & 0 & 0^* \\
(\mathbf{i}\times\mathbf{j})\zeta_{-\mathbf{j}} & (\mathbf{j}\times\mathbf{i})\zeta_{-\mathbf{i}} & 0 & 0^* & 0^* & 0
\end{array}\right]
\end{array}, \tag{A.1}
$$

where star denotes matrix entries which are not zero in the original system, but contain dynamical variables that are not in the wave-triad subsystem. The structure matrix of the subsystem is clearly skew-symmetric, but the structure functions do not satisfy the Jacobi identity. This can be easily verified by substituting **i**, **j**, **–i** or **i**, **j**, **–j** for p, q, r in the Jacobi identity. Until this point, however, we did not make use of the reality condition, an important symmetry condition. Using of the reality condition the six-variable complex system can be replaced by an equivalent six-variable real system. The structure matrix for this real system is

$$
\begin{array}{c} \\ \mathbf{i}^r \\ \mathbf{j}^r \\ \mathbf{k}^r \\ \mathbf{i}^i \\ \mathbf{j}^i \\ \mathbf{k}^i \end{array}
\begin{array}{c}
\begin{array}{cccccc} \mathbf{i}^r & \mathbf{j}^r & \mathbf{k}^r & \mathbf{i}^i & \mathbf{j}^i & \mathbf{k}^i \end{array} \\
\begin{bmatrix}
0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{k}^r} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{j}^r} & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{k}^i} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{j}^i} \\
(\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{k}^r} & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{i}^r} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{k}^i} & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{i}^i} \\
(\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{j}^r} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{i}^r} & 0 & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{j}^i} & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{i}^i} & 0 \\
0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{k}^i} & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{j}^i} & 0 & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{k}^r} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{j}^r} \\
(\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{k}^i} & 0 & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{i}^i} & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{k}^r} & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{i}^r} \\
(\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{j}^i} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{i}^i} & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{j}^r} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{i}^r} & 0
\end{bmatrix}
\end{array}, \tag{A.2}
$$

where the real variables $\zeta_{\mathbf{n}^i}$ and $\zeta_{\mathbf{n}^r}$ are defined as $\zeta_{\mathbf{n}} = \zeta_{\mathbf{n}^i} + i\zeta_{\mathbf{n}^r}$ (i is the imaginary unit).

Unless all imaginary components are zero the Jacobi identity cannot be satisfied. This follows from Eq. (5) of Part I upon substitution of e.g. $\mathbf{i}^r$, $\mathbf{j}^r$, $\mathbf{i}^i$ for p, q, r:

$$
\begin{aligned}
\sum_{x=i,r}\sum_{\mathbf{l}^x}\left\{D_{\mathbf{i}^r\mathbf{l}^x}\frac{\partial D_{\mathbf{j}^r\mathbf{i}^i}}{\partial\zeta_{\mathbf{l}^x}} + D_{\mathbf{i}^i\mathbf{l}^x}\frac{\partial D_{\mathbf{i}^r\mathbf{j}^r}}{\partial\zeta_{\mathbf{l}^x}} + D_{\mathbf{j}^r\mathbf{l}^x}\frac{\partial D_{\mathbf{i}^i\mathbf{i}^r}}{\partial\zeta_{\mathbf{l}^x}}\right\} \\
= \left[(\mathbf{j}\times\mathbf{i})^2\zeta_{\mathbf{j}^i} + (\mathbf{i}\times\mathbf{j})^2\zeta_{\mathbf{j}^i}\right] \\
= 2(\mathbf{j}\times\mathbf{i})^2\zeta_{\mathbf{j}^i} = 0, \quad \text{only if } \zeta_{\mathbf{j}^i} = 0.
\end{aligned} \tag{A.3}
$$

Similarly, replacing p, q and r by the index triads $\mathbf{j}^r$, $\mathbf{i}^r$, $\mathbf{j}^i$ and $\mathbf{i}^r$, $\mathbf{k}^r$, $\mathbf{i}^i$ would require the vanishing of $\zeta_{\mathbf{j}^i}$ and $\zeta_{\mathbf{k}^i}$, respectively.

On the other hand for purely real components the structure matrix reduces to

$$
\begin{array}{c} \\ \mathbf{i}^r \\ \mathbf{j}^r \\ \mathbf{k}^r \end{array}
\begin{array}{c}
\begin{array}{ccc} \mathbf{i}^r & \mathbf{j}^r & \mathbf{k}^r \end{array} \\
\begin{bmatrix}
0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{k}^r} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{j}^r} \\
(\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{k}^r} & 0 & (\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{i}^r} \\
(\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{j}^r} & (\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{i}^r} & 0
\end{bmatrix}
\end{array}, \tag{A.4}
$$

and the structure functions can be easily shown to satisfy the Jacobi condition. Thus property (*i*) is proved. Note that an analogue statement for a system consisting of purely imaginary components is not valid because the structure functions describing interactions between the imaginary parts contain real variables.

Conservation of energy follows from the skew-symmetry of the structure matrix, while the conservation of enstrophy is the consequence of

$$\sum_{x=r,i}\sum_{\mathbf{k}^x} D_{\mathbf{l}\mathbf{k}^x}\zeta_{\mathbf{k}^x} = 0 \quad \text{for} \quad l \in \{\mathbf{i}^r,\mathbf{j}^r,\mathbf{k}^r,\mathbf{i}^i,\mathbf{j}^i,\mathbf{k}^i\}, \tag{A.5}$$

which is equivalent to Eq. (2) applied to the enstrophy $C^2 = \sum_{x=r,i}\sum_{\mathbf{k}^x} \zeta^2_{\mathbf{k}^x}$.

Feature *(i)* of the above lemma has the following important consequence:

Corollary: *An aliasing-free truncation inevitably destroys the Hamiltonian structure.*

Proof: An aliasing-free truncated system can be decomposed into a collection of wave triads. The skew-symmetry of the structure matrices of the wave-triad subsystems guarantees the skew-symmetry of the structure matrix of the complete truncated system. The truncated system, however, is not Hamiltonian, since there are wave triads, for which the "stared" coefficients of A.1 still contain wave vectors out of the lattice and must be replaced by zero. An example for such wave triad is $\mathbf{i} = (T-1, T)$, $\mathbf{j} = (1, -1)$ and $\mathbf{k} = (T, T-1)$. If $T \geq 2$, the spectral coefficients associated with the wave vectors $\mathbf{i}+\mathbf{k}$, $\mathbf{i}-\mathbf{j}$, $\mathbf{j}+\mathbf{k}$, $\mathbf{j}-\mathbf{i}$, $\mathbf{k}+\mathbf{i}$, $\mathbf{k}+\mathbf{j}$, $-\mathbf{i}+\mathbf{j}$, $-\mathbf{i}-\mathbf{k}$, $-\mathbf{j}+\mathbf{i}$, $-\mathbf{j}-\mathbf{k}$, $-\mathbf{k}-\mathbf{i}$ and $-\mathbf{k}-\mathbf{j}$ are definitely out of the truncation. At $T = 1$ the wave vectors $\mathbf{i}+\mathbf{k}$, $\mathbf{k}+\mathbf{i}$, $-\mathbf{i}-\mathbf{k}$, $-\mathbf{k}-\mathbf{i}$ are within the lattice, thus the structure matrix for this subsystem contains additional terms compared to A.1. The following entries of the structure matrix differs from that of a single wave-triad (A.2):

$$\begin{aligned}
&D_{\mathbf{i}^r\mathbf{k}^r} = (\mathbf{j}\times\mathbf{i})(\zeta_{\mathbf{j}^r} - \zeta_{(\mathbf{i}+\mathbf{k})^r}), \quad && D_{\mathbf{i}^r\mathbf{k}^i} = (\mathbf{j}\times\mathbf{i})(\zeta_{\mathbf{j}^i} - \zeta_{(\mathbf{i}+\mathbf{k})^i}),\\
&D_{\mathbf{k}^r\mathbf{i}^r} = (\mathbf{i}\times\mathbf{j})(\zeta_{\mathbf{j}^r} - \zeta_{(\mathbf{i}+\mathbf{k})^r}), \quad && D_{\mathbf{k}^r\mathbf{i}^i} = (\mathbf{j}\times\mathbf{i})(\zeta_{\mathbf{j}^i} + \zeta_{(\mathbf{i}+\mathbf{k})^i}),\\
&D_{\mathbf{i}^i\mathbf{k}^r} = (\mathbf{i}\times\mathbf{j})(\zeta_{\mathbf{j}^i} + \zeta_{(\mathbf{i}+\mathbf{k})^i}), \quad && D_{\mathbf{i}^i\mathbf{k}^i} = (\mathbf{j}\times\mathbf{i})(\zeta_{\mathbf{j}^r} + \zeta_{(\mathbf{i}+\mathbf{k})^r}),\\
&D_{\mathbf{k}^i\mathbf{i}^r} = (\mathbf{i}\times\mathbf{j})(\zeta_{\mathbf{j}^i} - \zeta_{(\mathbf{i}+\mathbf{k})^i}), \quad && D_{\mathbf{k}^i\mathbf{i}^i} = (\mathbf{i}\times\mathbf{j})(\zeta_{\mathbf{j}^r} + \zeta_{(\mathbf{i}+\mathbf{k})^r}).
\end{aligned} \tag{A.6}$$

Again it is possible to find index vectors for which the Jacobi identity fails. Let $p = \mathbf{i}^r$, $q = \mathbf{j}^r$ and $r = \mathbf{i}^i$ then

$$\begin{aligned}
&\sum_{x=i,r}\sum_{\mathbf{l}^x}\left\{D_{\mathbf{i}^r\mathbf{l}^x}\frac{\partial D_{\mathbf{j}^r\mathbf{i}^i}}{\partial\zeta_{\mathbf{l}^x}} + D_{\mathbf{i}^i\mathbf{l}^x}\frac{\partial D_{\mathbf{i}^r\mathbf{j}^r}}{\partial\zeta_{\mathbf{l}^x}} + D_{\mathbf{j}^r\mathbf{l}^x}\frac{\partial D_{\mathbf{i}^i\mathbf{i}^r}}{\partial\zeta_{\mathbf{l}^x}}\right\}\\
&= (\mathbf{j}\times\mathbf{i})\left[(\mathbf{j}\times\mathbf{i})\zeta_{\mathbf{j}^i} + (\mathbf{i}\times\mathbf{j})\zeta_{(\mathbf{i}+\mathbf{k})^i}\right] + (\mathbf{i}\times\mathbf{j})\left[(\mathbf{i}\times\mathbf{j})\zeta_{\mathbf{j}^i} + (\mathbf{i}\times\mathbf{j})\zeta_{(\mathbf{i}+\mathbf{k})^i}\right]\\
&= 2(\mathbf{j}\times\mathbf{i})^2\zeta_{\mathbf{j}^i} \neq 0.
\end{aligned} \tag{A.7}$$

Appendix B

We show that

$$C^N = \sum_{\sum_{i=1}^{N} \mathbf{n}_i = \mathbf{0}} \zeta_{\mathbf{n}_1} \cdots \zeta_{\mathbf{n}_N}. \tag{B.1}$$

is not conserved in the aliasing-free truncation if $N > 2$.
If C^N were a Casimir invariant, Eq. (2) would imply that

$$\left[\mathbf{D} \frac{\partial C^N}{\partial \zeta} \right]_{\mathbf{k}} = 0 \tag{B.2}$$

should be valid for each $\mathbf{k} \in I$.

Using the definitions of the structure functions (Eq. 10) and Eq. (B.1) the left-hand side of Eq. (B.2) reads

$$\sum_{\mathbf{i} \in I} \mathbf{D}_{\mathbf{ki}} \frac{\partial C^N}{\partial \zeta_{\mathbf{i}}} = N \sum_{\substack{\mathbf{i} \in I \\ \mathbf{i}+\mathbf{k} \in I}} \sum_{\substack{N-1 \\ \sum_{j=1} \mathbf{m}_j = -\mathbf{i}}} (\mathbf{k} \times (\mathbf{i}+\mathbf{k})) \zeta_{\mathbf{i}+\mathbf{k}} \zeta_{\mathbf{m}_1} \cdots \zeta_{\mathbf{m}_{N-1}}. \tag{B.3}$$

This double sum contains N-fold products of $\zeta_{\mathbf{e}_i}$-s that obey $\sum_{i=1}^{N} \mathbf{e}_i = \mathbf{k}$ and $\mathbf{e}_i - \mathbf{k} \in I$. If the latter condition holds for all indices in the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ then the subsum of the N-fold products containing these $\zeta_{\mathbf{e}_i}$-s in all possible permutations is zero, as it was shown in the main body of the paper (see Eq. (14)). But in case of most index sets the second condition is not valid for all $\mathbf{e}_i$ in the index set, and the summation takes the form

$$\sum_{\{\mathbf{e}_i\} \in P_N^{\mathbf{k}}} \zeta_{\mathbf{e}_1} \cdots \zeta_{\mathbf{e}_N} R^{\mathbf{k}}(\mathbf{e}_1, \ldots, \mathbf{e}_N), \tag{B.4}$$

where $R^{\mathbf{k}}(\mathbf{e}_1, \ldots, \mathbf{e}_N) = \sum_{\substack{i=1 \\ \mathbf{e}_i - \mathbf{k} \in I}}^{N} \mathbf{k} \times \mathbf{e}_i (N-1)!$. We prove that this sum cannot be zero for all vorticity fields.

Suppose that Eq. (B.4) is zero for all possible spectral coefficients. Let L denote the number of $\{\mathbf{e}_i\}$ sets in $P_N^{\mathbf{k}}$ and let us specify L different vorticity fields. Let p_i be different prime integers greater than N for $i=1,...,(M^2-1)$. Introducing new indices $r: I \to \{1,...,(M^2-1)/2\}$ that satisfy $r(\mathbf{e}) = r(-\mathbf{e})$ for each $\mathbf{e} \in I$, let $\zeta_{\mathbf{e}_i}^{(l)} = \exp(il\nu_{r(\mathbf{e}_i)})$ and $\zeta_{-\mathbf{e}_i}^{(l)} = \exp(-il\nu_{r(\mathbf{e}_i)})$, where $\nu_{r(\mathbf{e}_i)} = 2\pi/p_{r(\mathbf{e}_i)}$ and $l \in \{0,1,...,L-1\}$. This choice guarantees that the N-fold products $\zeta_{\mathbf{e}_1}^{(1)} \cdots \zeta_{\mathbf{e}_N}^{(1)}$, are all distinct. Let $a_j = \zeta_{\mathbf{e}_{j_1}}^{(1)} \cdots \zeta_{\mathbf{e}_{j_N}}^{(1)}$. The condition that Eq. (B.4) is zero for the vorticity fields just defined yields a homogeneous system of linear equations for $R^{\mathbf{k}}(\mathbf{e}_1,...,\mathbf{e}_N)$-s with constant coefficients:

$$\begin{bmatrix} 1 & 1 & ... & 1 \\ a_1 & a_2 & ... & a_L \\ a_1^2 & a_2^2 & ... & a_L^2 \\ ... & ... & ... & ... \\ a_1^{L-1} & a_2^{L-1} & ... & a_L^{L-1} \end{bmatrix}. \tag{B.5}$$

Since the a_i-s are distinct the coefficients form a nonsingular Vandermonde determinant and the system of equation has only trivial solution, that is all $R^{\mathbf{k}}(\mathbf{e}_1,...,\mathbf{e}_N)$-s should be zero. On the other hand it is easy to construct examples of $\{\mathbf{e}_i\}$ for nonzero $R^{\mathbf{k}}(\mathbf{e}_1,...,\mathbf{e}_N)$. Let $\mathbf{k} = (k_1,k_2)$ and suppose that $k_2 \geq 0$. Take the index set $\mathbf{e}_1 = (0, k_2 - T - 1)$, $\mathbf{e}_2 = (a,T)$ and $\mathbf{e}_3 = (k_1 - a, 1)$ where $|a|, |k_1 - a| \leq T$ and $\mathbf{e}_i = 0$ for $i = 4,...,N$. Since the first index $\mathbf{e}_1 - \mathbf{k}$ is out of the truncation, $R^{\mathbf{k}}(\mathbf{e}_1,...,\mathbf{e}_N) = \mathbf{k} \times (\mathbf{e}_2 + \mathbf{e}_3) = (T + k_2 + 1)k_1$ which is zero only if $k_1 = 0$. In the latter case, however, take $\mathbf{e}_1 = (0, k_2 - T - 1)$, $\mathbf{e}_2 = (0,T)$ and $\mathbf{e}_3 = (-1,1)$ and $R^{\mathbf{k}}$ will not vanish if $\mathbf{k} \neq \mathbf{0}$. The counter-example given shows that Eq. (B.4) cannot be zero for all vorticity fields.

Appendix C

Invariants of the sine-bracket truncation

As it is stated in a number of papers (e.g. *Zeitlin*, 1991; *Dowker* and *Wolski*, 1992) the truncated system Eq. (17) can be written in the equivalent "bracket form" of

$$\dot{\Omega} = \frac{i}{2\epsilon}[\Omega, \psi], \tag{C.1}$$

where i is the imaginary unit, ϵ is as defined in the main text, $\Omega(t) = \sum_{\mathbf{n}} \zeta_{\mathbf{n}}(t)\mathbf{J}_{\mathbf{n}}$, $\psi(t) = \sum_{\mathbf{n}} \frac{\zeta_{\mathbf{n}}(t)}{|\mathbf{n}|^2}\mathbf{J}_{\mathbf{n}}$ and $[.,.]$ denotes matrix commutation $[\Omega, \psi] = \Omega\psi - \psi\Omega$. The $\mathbf{J}_{\mathbf{n}}$-s are special $M \times M$ matrices (they are generators of the vector space of $M \times M$ Hermitian matrices) with the properties

$$\begin{aligned} \mathbf{J}_{\mathbf{n}} &= \mathbf{J}_{\mathbf{n}}^{*}, \\ \mathbf{J}_{\mathbf{n}}\mathbf{J}_{\mathbf{m}} &= exp(i\epsilon\,(\mathbf{m} \times \mathbf{n}))J_{\mathbf{n}+\mathbf{m} \bmod M}, \\ tr\mathbf{J}_{\mathbf{n}} &= \begin{cases} M & \text{for } \mathbf{n} = \mathbf{0} \\ 0 & \text{for } \mathbf{n} \neq \mathbf{0} \end{cases} \end{aligned} \tag{C.2}$$

Because of the multiplication rule Ω^N is also a linear combination of the $\mathbf{J}_{\mathbf{n}}$-s

$$\Omega^N = \sum_{\mathbf{i}} x_{\mathbf{i}}\mathbf{J}_{\mathbf{i}}. \tag{C.3}$$

Let $y_{\mathbf{j}} = \frac{\zeta_{\mathbf{j}}}{|\mathbf{j}|^2}$. From Eq. (C.1) and the properties of the $\mathbf{J}_{\mathbf{n}}$-s follows

$$\begin{aligned} \frac{\partial}{\partial t} tr\Omega^N &= tr\left([\Omega^N, \psi]\right) = tr\sum_{\mathbf{i}}\sum_{\mathbf{j}} x_{\mathbf{i}}y_{\mathbf{j}}[\mathbf{J}_{\mathbf{i}}, \mathbf{J}_{\mathbf{j}}] \\ &= -2i\sum_{\mathbf{i}}\sum_{\mathbf{j}} x_{\mathbf{i}}y_{\mathbf{j}} \sin\left(\epsilon\,(\mathbf{i} \times \mathbf{j})\right) tr\mathbf{J}_{\mathbf{i}+\mathbf{j} \bmod M} \\ &= -2i\,M\sum_{\mathbf{i}} x_{\mathbf{i}}y_{-\mathbf{i}} \sin\left(\epsilon\,(\mathbf{i} \times -\mathbf{i})\right) = 0. \end{aligned} \tag{C.4}$$

That is, the quantities $C_s^N = \frac{1}{M} tr\Omega^N$ are all conserved, though only $M-1$ of them are functionally independent (see e.g. *Zeitlin*, 1991). Due to the features of $\mathbf{J_n}$ matrices these invariants can be written in the following form

$$\begin{aligned} C_s^N &= \frac{1}{M} tr\left(\sum_{\mathbf{i}} \zeta_{\mathbf{i}} \mathbf{J}_{\mathbf{i}}\right)^N \\ &= \frac{1}{M} \sum_{\mathbf{i}_1,\ldots,\mathbf{i}_N} \zeta_{\mathbf{i}_1} \cdots \zeta_{\mathbf{i}_N} tr\left(\mathbf{J}_{\mathbf{i}_1} \cdots \mathbf{J}_{\mathbf{i}_N}\right) \\ &= \sum_{\sum_{j=1}^{N} \mathbf{i}_j = \mathbf{0} \, mod M} \zeta_{\mathbf{i}_1} \cdots \zeta_{\mathbf{i}_N} \exp\left(\epsilon 2A(\mathbf{i}_1,\ldots,\mathbf{i}_N)\right) \\ &= \sum_{\sum_{j=1}^{N} \mathbf{i}_j = \mathbf{0} \, mod M} \zeta_{\mathbf{i}_1} \cdots \zeta_{\mathbf{i}_N} \cos\left(\epsilon 2A(\mathbf{i}_1,\ldots,\mathbf{i}_N)\right). \end{aligned} \tag{C.5}$$

In the last equality the following symmetry of the area function was considered

$$A(\mathbf{i}_1,\ldots,\mathbf{i}_N) = -A(\mathbf{i}_N,\ldots,\mathbf{i}_1). \tag{C.6}$$

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 2, April–June 1998, pp. 109–127

Land-surface hydrology parameterization in PROGSURF: Formulation and test with Cabauw data*

Ferenc Ács[1] and **Michael Hantel**[2]

[1]*Department of Meteorology, Eötvös Loránd University, H-1083 Budapest, Ludovika tér 2, Hungary; E-mail: acs@nimbus.elte.hu*

[2]*University of Vienna, Hohe Warte 38, A-1190 Vienna, Austria*

(Manucript received 5 May 1997; in final form 20 October 1997)

Abstract—The land-surface model PROGSURF (*Prognosis of Surface Fluxes*) is introduced and applied for analysing the water budget components at the station Cabauw in the Netherlands. PROGSURF gives the main emphasis to water budget modeling. All relevant water transport processes are treated: rainfall interception, canopy drainage, infiltration, surface runoff, soil water diffusion, root water uptake, evapotranspiration, conductance of water through roots and stems and subsurface runoff. Particular attention is paid to the parameterization of transpiration. The canopy surface resistance concept is applied by using Jarvis' empirical formula. The effect of available soil moisture is parameterized via leaf water potential. The soil moisture change is simulated by a 3-layer diffusion type model.

By running PROGSURF using the Cabauw data set we obtain the following main results: evapotranspiration and runoff is –435 and 338 mm/year, respectively; annual mean of root zone soil moisture content is 0.345 m^3/m^3. Seasonal changes of water balance components are also suitably reproduced. Evapotranspiration (and thus indirectly soil moisture content and runoff) are very sensitive to the parameterization of available soil moisture; this applies both to the annual value and the seasonal change of the sensitive quantities.

Furthermore, we compare our results with those of the PILPS (Project for Intercomparison of Land-Surface Parameterization Schemes) Phase 2(a) experiment. The comparison reveals that the differences in the specification of the relevant PROGSURF parameters can be as important as the differences in the structure of the PILPS models. The results of this study may thus be useful for optimizing the transpiration schemes used in land surface models.

Key-words: surface water budget, transpiration, available soil moisture, complexity versus simplicity.

---

* The paper was presented at the 22nd General Assembly of EGS, Vienna, 21–24 April, 1997.

1. Introduction

Many aspects of land-surface hydrology parameterization are still not well understood. For example, the heterogeneity effect in up- and downscaling, the mechanism of transpiration, and the parameterization of the interaction between transpiration and other surface water budget components need to be improved. On the other hand, there is a permanent background for all modeling efforts: the aspect of *complexity versus simplicity*.

In the present study some of these aspects are considered. The surface water budget components are analysed by the aid of the land-surface model PROGSURF (**Prog**nosis of **Sur**face **F**luxes), developed jointly at University of Vienna and Eötvös Loránd University, Budapest. Most ingredients of PROGSURF are standard. The model gives emphasis to the water budget comprising a total of 4 layers (vegetation layer plus three soil layers). Transpiration is parameterized using the canopy surface resistance concept in form of *Jarvis'* (1976) multiplicative formula. This implies specification of two governing relative conductances: one describing the *atmospheric demand* (relative stomatal conductance F_{ad}); and one describing the *available soil moisture* (relative stomatal conductance F_{ma}). The effect of atmospheric forcing (radiation, air humidity and temperature) upon stomatal functioning is condensed into F_{ad}. The available soil moisture (hereafter soil moisture availability) for evapotranspiration is represented by F_{ma}. Two possible formulations of F_{ma} are investigated; both represent the relation among transpiration, soil moisture content and runoff in a different manner. F_{ma} can be parameterized either via soil moisture content (*Noilhan* and *Planton*, 1989) or via leaf water potential (*Choudhury*, 1983; *Sellers* and *Dorman*, 1987; *Lynn* and *Carlson*, 1990; *Ács*, 1994). When parameterized with soil moisture content, there is no atmospheric demand effect represented in F_{ma}. When parameterized via leaf water potential, the effect of the atmosphere is also implicitly contained in F_{ma} (*De Ridder* and *Schayes*, 1997), in addition to being represented explicitly by F_{ad}.

This study has a threefold aim. First, we document PROGSURF's land-surface hydrology parameterization, including the parameterizations just mentioned. Second, for the Cabauw data we show that PROGSURF reproduces the annual mean as well as the annual course of the water budget components. Third, we demonstrate that transpiration (and thus soil moisture content and runoff too) is largely determined by the parameterization of F_{ma}. The numerical experiments are performed in off-line mode using Cabauw data set with different modes of PROGSURF.

The paper is organized as follows: the land-surface hydrology parameterization is described in the next section, the experimental site and the measurements are discussed in Section 3, the validation results are presented in Section 4 and the sensitivity experiments are discussed in Section 5. The conclusions are given in Section 6.

2. *The model*

The land-surface model PROGSURF is mostly based on previous work of Ács (*Ács et al.*, 1991; *Ács*, 1994; *Ács* and *Hantel*, 1997, 1998). Variables and geometry of the model are schematically presented on *Fig. 1*. The core of the scheme is a 2-layer temperature prediction of the vegetation-ground system based on the force-restore method (*Noilhan* and *Planton*, 1989), plus a 3-layer diffusion type soil moisture prediction. Due to the overlap of the vegetation-ground system and the soil system the total number of layers is 4. The vegetation-ground system contains a vegetation layer and a soil surface layer. The index v refers to the vegetation; we shall use the words *vegetation* and *canopy* synonymously. Evapotranspiration is parameterized using surface resistance concepts. The canopy surface resistance is evaluated using Jarvis' proposal for estimating soil moisture availability parameter via leaf water potential. The aerodynamic resistance is separately calculated above bare and vegetated soil surfaces using Monin-Obukhov's similarity theory taking into account the atmospheric stability.

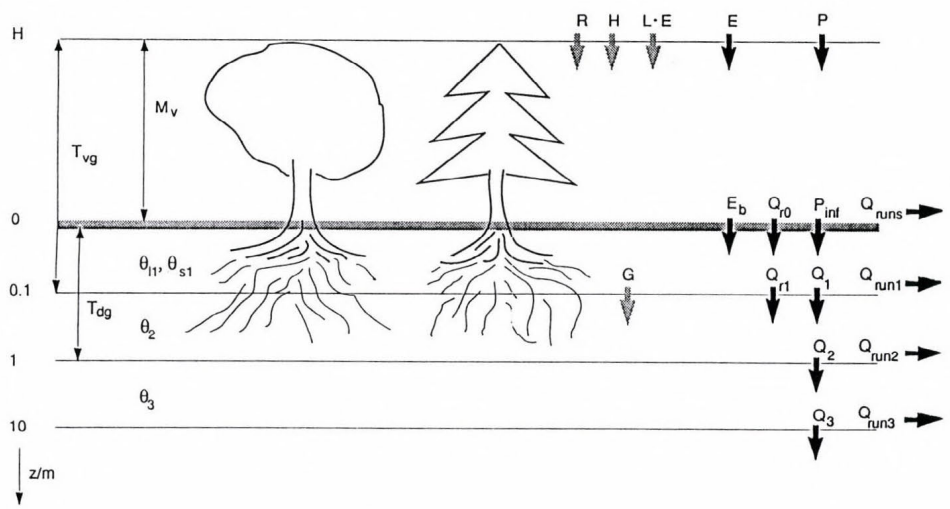


Fig. 1. Schematic diagram of prognostic variables, energy and water fluxes and layers represented by PROGSURF. (T_{vg} – temperature of vegetation-ground layer; T_{dg} – temperature of deep-ground layer; M_v – intercepted water stored in vegetation layer; θ_{l1}, θ_{s1} – liquid and solid water content in the 1st soil layer, respectively; θ_2, θ_3 – soil moisture content in the 2nd and 3rd layer; R – net radiation; H – sensible heat flux; $L \cdot E$ – latent heat flux; G – soil heat flux; E – moisture flux; E_b – bare soil evaporation; P – precipitation; Q_{r0}, Q_{r1} – root water flux across the surface and the bottom of the 1st layer; P_{inf}, Q_1, Q_2, Q_3 – soil water flux across the surface (infiltration) and the bottom of the 1st, 2nd and 3rd layer; Q_{runs}, Q_{run1}, Q_{run2}, Q_{run3} – runoff from the surface and horizontal discharge from the 1st, 2nd and 3rd soil layer. Energy fluxes are shaded, water fluxes are black and all fluxes are positive in positive z-direction, represented by downward arrows.)

PROGSURF is able to treat the case of frozen soil while the representation of snow is presently missing. Sub-gridscale variations of surface characteristics are not considered. We follow the convention to count all vertical fluxes to be positive if directed downwards. This implies, e.g., that the sign of evaporation and transpiration is practically always negative. In the following more attention is paid to the land-surface hydrology parameterization.

2.1 Soil moisture prediction

Diffusion-type moisture prediction is applied in the soil layers. The prognostic equations for the three layers (Fig. 1) are:

$$\rho_w \cdot D_1 \cdot \frac{\partial \theta_{l1}}{\partial t} = P_{\text{inf}} - Q_1 - Q_{r0} - Q_{r1} + E_b - Q_{run1} - S_p, \tag{1}$$

$$\rho_w \cdot D_1 \cdot \frac{\partial \theta_{s1}}{\partial t} = S_p, \tag{2}$$

$$\rho_w \cdot D_2 \cdot \frac{\partial \theta_2}{\partial t} = Q_1 - Q_2 + Q_{r1} - Q_{run2}, \tag{3}$$

$$\rho_w \cdot D_3 \cdot \frac{\partial \theta_3}{\partial t} = Q_2 - Q_3 - Q_{runs}, \tag{4}$$

where ρ_w is water density; D_i is the thickness of the ith soil layer; θ_{l1}, θ_{s1} is the liquid and solid moisture content in the 1st soil layer, respectively; θ_2, θ_3 is the soil moisture content in the 2nd and 3rd soil layer, respectively; P_{inf} is the water infiltrated into the soil (kg m^{-2} s^{-1}); E_b is the evaporation from bare soil surface (kg m^{-2} s^{-1}); Q_{r0} and Q_{r1} is the root water flux across surface and across the bottom of the 1st soil layer (kg m^{-2} s^{-1}); Q_1 and Q_2 is the water diffusion between the adjacent layers (kg m^{-2} s^{-1}); Q_3 is the gravitational drainage (kg m^{-2} s^{-1}) and Q_{runi} represents the lateral drainage in the ith layer (kg m^{-2} s^{-1}). Soil water diffusion terms are positive downwards and negative upwards. The gravitational and lateral drainage terms are positive representing the outflow from the system. S_p is the source/sink term (kg m^{-2} s^{-1}) representing soil water freezing or melting. It can be expressed as $S_p = -(R + H + L \cdot E - G)/L_f$, where R is the net radiation, H is the sensible heat flux, $L \cdot E$ is the latent heat flux, G is the soil heat flux and L_f is the latent heat of fusion. $S_p > 0$ when freezing and $S_p < 0$ when melting.

2.2 Water fluxes

The governing water transport processes in the soil-vegetation-atmosphere system are the precipitation, the rainfall interception, the canopy drainage, the distribution of the effective precipitation into the infiltration and surface runoff, the vertical water movement in the soil, the root water fluxes, the evapotranspiration, the conductance of water through roots and stems and the subsurface runoff. In the following the parameterizations of these processes are briefly presented.

2.2.1 Effective precipitation, infiltration and surface runoff

The effective precipitation rate on the soil surface (P_0) is defined by

$$P_0 = P - P_v + D_v, \tag{5}$$

where P is the rainfall intensity (kg m^{-2} s^{-1}), P_v is its interception by vegetation and D_v is the water drainage from the canopy. Both latter terms are parameterized after *Sellers et al.* (1986). P_0 either infiltrates into the soil (P_{inf}) or runs off on the surface (Q_{runs}) depending upon surface characteristics. This partition mechanism is parameterized as

$$P_{\text{inf}} = \begin{cases} \min(P_0, K_{S1}) & \text{if } W_1 < 1 \\ 0 & \text{if } W_1 = 1, \end{cases} \tag{6}$$

$$Q_{runs} = \begin{cases} P_0 & \text{if } W_1 = 1 \\ P_0 - K_{S1} & \text{if } W_1 < 1 \text{ and } P_0 > K_{S1} \\ 0 & \text{if } W_1 < 1 \text{ and } P_0 \leq K_{S1}, \end{cases} \tag{7}$$

where $W_1 = \theta_1/\theta_{S1}$; θ_{S1} is the saturated soil moisture content (m^3/m^3) and K_{S1} is the saturated hydraulic conductivity (kg m^{-2} s^{-1}) in the 1st soil layer. The parameterization does not take into account the orographical effects.

2.2.2 Evaporation and soil water transfer

Evaporation from bare soil is estimated by the aerodynamic formula evaluating soil surface resistance after *Sun* (1982). The water flow between the adjacent soil layers is estimated by

$$Q_i = K_{eff,i} \cdot \left[2 \cdot \frac{\Psi_i - \Psi_{i+1}}{D_i + D_{i+1}} + 1 \right], \tag{8}$$

$$K_{eff,i} = \frac{D_i \cdot K_i + D_{i+1} \cdot K_{i+1}}{D_i + D_{i+1}}, \tag{9}$$

where $K_{eff,i}$ is an effective hydraulic conductivity between the ith and $(i+1)$th ($i = 1, 2$) soil layers (kg m^{-2} s^{-1}), K_i is the hydraulic conductivity of the ith soil layer (kg m^{-2} s^{-1}) and Ψ_i is the soil water potential of the ith soil layer (m). Both latter terms are parameterized after *Clapp* and *Hornberger* (1978).

2.2.3 Water transfer in roots, stems and leaves

The root water flux (Q_{r0}) across the surface (kg m^{-2} s^{-1}) is assumed to flow in the stems of plants. The continuity of water flow in soil–vegetation–atmosphere system requires to equate Q_{r0} with the transpiration (E_{vd}):

$$E_{vd} = \frac{\rho c_p}{L\gamma} \cdot \frac{e_S(T_{vg}) - e_r}{r_v + r_{av}}, \tag{10}$$

$$Q_{r0} = E_{vd}, \tag{11}$$

where ρ is the air density (kg/m^3), c_p is the specific heat of air at constant pressure (J kg^{-1} K^{-1}), γ is the psychrometric constant (hPa/K), $e_S(T_{vg})$ is the saturation vapor pressure at T_{vg} (hPa), e_r is the vapor pressure at reference level (hPa), r_v is the canopy resistance (s/m) and r_{av} is the aerodynamic resistance above vegetation (s/m).

The root water flux (Q_{r1}) across the bottom of the 1st layer is estimated as a prespecified percentage of the flux across the ground surface (*Chen et al.*, 1997):

$$Q_{r1} = 0.3 \cdot Q_{r0}. \tag{12}$$

Q_{r0} is parameterized using leaf water potential (Ψ_v) according to *van der Hornert* (1948):

$$Q_{r0} = \rho_w \cdot \frac{\Psi_v - \Psi_R + z_T}{r_R + r_P}, \tag{13}$$

where Ψ_R is the soil moisture potential in the root zone (m), z_T is the canopy source-sink height (m), r_R is the soil resistance in the root zone (s) and r_P is the plant resistance imposed by the plant vascular system (s). r_P depends on vegetation type; in our tests we used PILPS specified value of 2.5×10^8 s.

2.2.4. *Subsurface runoff*

Subsurface runoff is the sum of lateral and gravitational drainage. Lateral drainage is estimated for each soil layer ($i = 1, 2, 3$) using the expressions

$$Q_{runi} = \begin{cases} Q_{rfi} & \text{if } WFC_i \leq 0 \\ Q_{rfi} + WF_i & \text{if } WFC_i > 0, \end{cases} \tag{14}$$

where

$$Q_{rfi} = Q_i - \min(Q_i, K_{Si}), \tag{15}$$

$$WF_1 = P_{\text{inf}} - Q_1 + E_b + Q_{r0} - Q_{r1} - Q_{rf1}, \tag{16}$$

$$WF_2 = Q_1 - Q_2 + Q_{r1} - Q_{rf2}, \tag{17}$$

$$WF_3 = Q_2 - Q_3 - Q_{rf3}, \tag{18}$$

$$WFC_i = \frac{\rho_w D_i}{\Delta t} \cdot (\theta_i - \theta_{fi}) + WF_i, \tag{19}$$

where Δt is the time step, θ_i is the actual soil moisture content and θ_{fi} is the soil moisture content at field capacity in the ith soil layer; WF_i represents the net water flux in the ith soil layer and K_{Si} is the saturated hydraulic conductivity in the ith soil layer.

The gravitational drainage rate from the bottom (Q_3) is calculated by

$$Q_3 = \begin{cases} K_{S3} \cdot W_3^{2B_3+3} & \text{if } \theta_3 \geq \theta_{f3} \\ 0 & \text{if } \theta < \theta_{f3} \end{cases}, \tag{20}$$

where W_3 is defined equivalently to W_1 (see Eq. (7)).

2.3 Resistances

To estimate evapotranspiration, it is necessary to determine aerodynamic and surface resistances. The aerodynamic resistance is evaluated using Monin-Obukhov's similarity theory taking into account the atmospheric stability. It is split into laminar and turbulent terms distinguishing transports between momentum and heat/moisture. It is separately calculated above bare and vegetated soil surfaces. Here the attention is paid only to the canopy surface resistance formulation.

2.3.1 Canopy resistance

Canopy resistance is parameterized following *Jarvis*' method (1976):

$$r_v = \frac{r_{st\,\min} \cdot F_{ad}}{LAI \cdot GLF \cdot F_{ma}}, \tag{21}$$

where r_{stmin} is the minimum stomatal resistance at optimal environmental conditions, *LAI* is the leaf area index (m^2/m^2) and *GLF* is the green leaf fraction (it expresses the fraction of green leaves ranging from 0 to 1; for parameter values see Table A2 in *Chen et al.*, 1997). F_{ad} and F_{ma} expresses the atmospheric demand and the soil moisture availability effect upon stomatal activity, respectively.

F_{ad} is split into three effects in a multiplicative manner:

$$F_{ad} = \frac{F_{vr}}{F_{ah} \cdot F_{at}}, \tag{22}$$

where F_{vr}, F_{ah} and F_{at} expresses the influence of absorbed visible radiation, air humidity and temperature, respectively. F_{vr} is parameterized after *Gates* (1981) while F_{ah} and F_{at} after *Jarvis* (1976) and *Dickinson* (1984).

F_{ma} is parameterized via leaf water potential as:

$$F_{ma} = \frac{\Psi_v - \Psi_{cr}}{\Psi_{SR} - \Psi_{cr}}, \tag{23}$$

where Ψ_v and Ψ_{cr} is the actual and the critical leaf water potential (m), respectively. At $\Psi_v = \Psi_{cr}$ the stomata are completely closed. Ψ_{cr} depends on vegetation type; for grass it is –230 m. Ψ_{SR} is the saturated soil water potential in the root zone (m); it is estimated as the weighted mean of Ψ_{S1} and Ψ_{S2} (for their values see Table A1 in *Chen et al.*, 1997) as follows:

$$\Psi_{SR} = \frac{D_1 \Psi_{S1} + D_2 \Psi_{S2}}{D_1 + D_2}. \tag{24}$$

To apply Eq. (23), it is necessary to determine Ψ_v. Combining Eqs. (10), (11), (13), (21) and Eq. (23) and after some rearranging, it is possible to get a quadratic equation for Ψ_v:

$$a \cdot \Psi_v^2 + b \cdot \Psi_v + c = 0, \tag{25}$$

where

$$a = r_{av}, \tag{26}$$

$$b = AD \cdot (r_R + r_P) - (\Psi_R - z_T) \cdot r_{av} + A \cdot (\Psi_{SR} - \Psi_{cr}) - \Psi_{cr} \cdot r_{av}, \tag{27}$$

$$c' = -AD \cdot (r_R + r_P) \cdot \Psi_{cr} + (\Psi_R - z_T) \cdot \Psi_{cr} \cdot r_{av}, \tag{28}$$

$$c = c' - A \cdot (\Psi_R - z_T) \cdot (\Psi_{SR} - \Psi_{cr}), \tag{29}$$

and

$$AD = \frac{\rho c_p}{\gamma L \rho_w} \cdot veg \cdot (1 - wif) \cdot [e_S(T_{vg}) - e_r], \tag{30}$$

$$A = \frac{r_{stmin} \cdot F_{ad}}{LAI \cdot GLF}. \tag{31}$$

The soil resistance in the root zone (r_R) is expressed as

$$r_R = \frac{\alpha_R \cdot \rho_w}{K_R \cdot D_R}, \tag{32}$$

where K_R is the hydraulic conductivity of soil in the root zone (kg m^{-2} s^{-1}), D_R is the root zone depth taken as 1 m. α_R is a vegetation specific parameter for the root zone (m^2), given by

$$\alpha_R = \frac{V_R - 3 - 2 \cdot \log\left[\frac{V_R}{1 - V_R}\right]}{8 \cdot \pi \cdot R_{\mathrm{det}}}, \tag{33}$$

where V_R is the volume of root per unit volume of soil in the root zone (m^3/m^3). Term V_R is estimated from the total root density in the root zone (R_{det}) and from the average root cross section (rcs). R_{det} is calculated after *Gerwitz* and *Page* (1974) when the root density in the soil surface layer (R_{des}) is known. For grass $R_{des} = 5500$ m/m^3 and $rcs = 3.84 \times 10^{-7}$ m^2. K_R and ψ_R are calculated as the weighted mean of their components analogously to Eq. (24).

The ψ_v value is obtained by solving Eq. (25),

$$\Psi_v = \frac{-b + [b^2 - 4 \cdot a \cdot c]^{1/2}}{2a}. \tag{34}$$

The other solution gives unrealistic results.

3. *The Cabauw data set*

There are three reasons for choosing data measured at Cabauw, the Netherlands in 1987:

(a) at Cabauw experimental station the measurements assured high quality values for the atmospheric forcing data, soil-vegetation parameters and the latent and sensible heat fluxes,

(b) the data series is long enough; it includes one full year providing a basis for testing model performance in terms of seasonal variations and

(c) the data set is chosen for the PILPS Phase 2(a) experiment which enables PROGSURF's comparison with other land-surface schemes.

Cabauw site is located in the Netherlands (51°58'N; 4°56'E); plant cover consists mainly of short grass with narrow ditches. The climate in the area is characterized as temperate maritime type with prevailing westerly winds. The soil consists of two layers: a 1-m-deep layer of silty clay on the top of a 10-m-deep layer of peat saturated with water. The most important soil hydraulic characteristics are presented in *Table 1*.

The Cabauw data, provided by the Royal Netherlands Meteorological Institute, are obtained in the scope of a semi-operational measurement program with a high level of quality control. The atmospheric forcing data (downward shortwave and longwave radiation, precipitation, surface atmospheric pressure,

near surface air humidity, temperature and wind speed) refer to the height of 20 m (reference height). For evaluation of scheme outputs, energy-flux data including sensible and latent heat fluxes, net radiation, upward longwave radiation and soil heat flux are available. The sensible and latent heat fluxes were not measured directly. The former was derived from temperature and wind profile data using modified flux-profile relationship while the latter is estimated as a residual of the surface energy balance (*Beljaars*, 1982). When any of the measurements for the energy-flux variables were not available, data were derived from a parameterization scheme which has been calibrated locally and tested extensively against observed data (*Holtslag* and *Van Ulden*, 1983). The Cabauw data used in the PILPS Phase 2(a) experiment refer and cover the entire year of 1987. Both the forcing data and the energy fluxes are available on a 30-min interval. The details are described in *Beljaars* and *Bosveld* (1997).

Table 1. Soil hydraulic characteristics at Cabauw site

Soil porosity	0.468
Soil moisture content at field capacity	0.370
Soil moisture content at wilting point	0.214
Saturated hydraulic conductivity (m/s)	3.4341×10^{-6}
Saturated soil water potential (m)	-0.045
'B' of *Clapp* and *Hornberger* (1978)	10.39

In the numerical experiments the PROGSURF is always initialized, as all schemes in PILPS 2(a) experiment, by saturating all liquid water stores and setting all temperatures to 279 K. The land-surface parameters are specified as designed in PILPS 2(a) comparison experiments (see Appendix in *Chen et al.* 1997).

4. Results of validation

The land-surface hydrology parameterization in PROGSURF is validated comparing (a) annual values of simulated and observed evapotranspiration and runoff and (b) annual course of simulated and observed evapotranspiration.

The annual partition of the amount of precipitation into evapotranspiration and runoff is presented on *Fig. 2*. In this case evapotranspiration *is practically*

equal to transpiration because the vegetation coverage is full. The results refer to the equilibrium year. Equilibrium was defined as being the first occasion that the January mean values of the surface temperature, latent and sensible heat fluxes and root-zone soil moisture did not change by more than 0.01 K, 0.1 W/m^2 and 0.1 mm, respectively, from year N to year N+1; the equilibrium time (spinup time) was then N years. PROGSURF's spinup time is 2 years.

The evapotranspiration and the runoff estimated by PROGSURF are -435 and 338 mm/year, respectively. The evapotranspiration and the runoff obtained by measurements and output analysis of the ECMWF (*European Centre for Medium-Range Weather Forecasts*) land-surface model (for details see *Viterbo* and *Beljaars*, 1995 and *Beljaars* and *Bosveld*, 1997) are about -525 and 250 mm/year, respectively.

The annual course of evapotranspiration simulated by PROGSURF (on the figure noted as Psi-PROGSURF) is presented on *Fig. 3*. The seasonal change of evapotranspiration is strong. In winter the values are small (around and under -10 mm/month); in summer they change between -70 and -100 mm/month. The absolute value of the maximum is 99 mm/month and appears in July. The root-mean-square error for simulated and observed values is 9.5 mm/month; it means that the agreement is suitable.

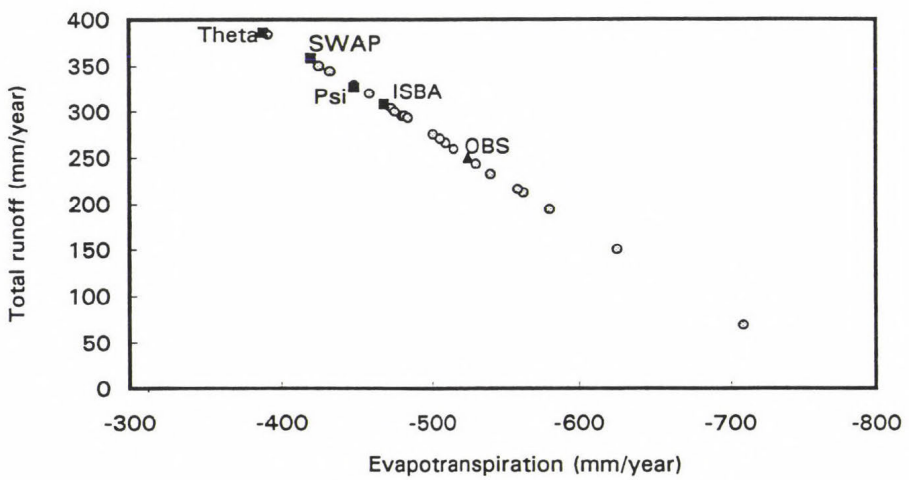


Fig. 2. Distribution of annual runoff and evapotranspiration at Cabauw experimental station in 1987. The estimations are obtained by the different versions of PROGSURF and by the models participating in PILPS Phase 2(a) experiment (for details see Fig. 10 in *Chen et al.*, 1997). *Psi* is a PROGSURF run with soil moisture availability parameter F_{ma} represented via leaf water potential; *Theta* is a PROGSURF run with F_{ma} represented via soil moisture contents; *ISBA* is a run with ISBA model; *SWAP* means a run with SWAP model; and *OBS* are the observed values.

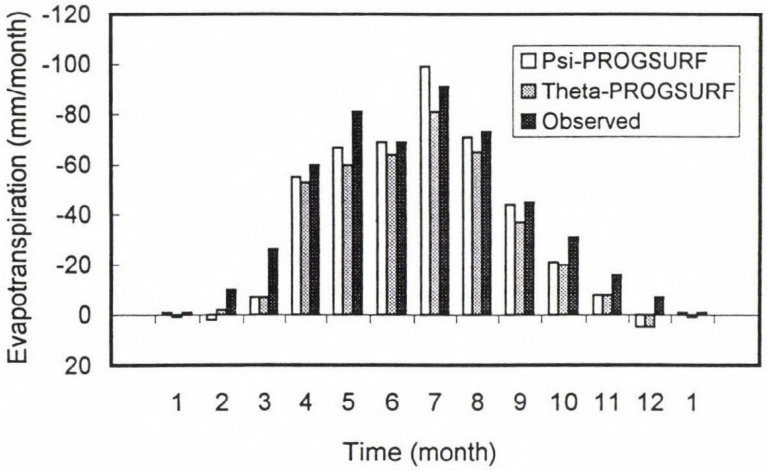


Fig. 3. Annual course of evapotranspiration simulated at Cabauw in 1987 by different versions of PROGSURF.

5. Sensitivity tests

In this chapter the sensitivity of the model performance to soil moisture availability (F_{ma}) formulation is tested. In PROGSURF, F_{ma} is estimated via leaf water potential. Both the soil moisture and the atmospheric demand effects (radiation, air temperature, humidity and aerodynamic resistances) are represented in this relatively complex formulation of F_{ma}.

Nevertheless, F_{ma} can also be parameterized in a very simple way via soil moisture contents

$$F_{ma} = \begin{cases} 1 & \text{for} \quad \theta_f \leq \theta \\ \dfrac{\theta - \theta_w}{\theta_f - \theta_w} & \text{for} \quad \theta_w < \theta < \theta_f, \\ 0 & \text{for} \quad \theta \leq \theta_w, \end{cases} \tag{35}$$

where θ, θ_f and θ_w is the actual, the field capacity and the wilting point soil moisture content in the root zone (m^3/m^3), respectively. The atmospheric demand effect is not represented in this formulation of F_{ma}.

A comparative analysis of two different versions of PROGSURF has been made to study the annual and seasonal characteristics of water budget

components from the aspect of *complexity versus simplicity*. In the Psi-PROGSURF version F_{ma} parameterized by Eq. (23), while Eq. (35) represents the parameterization of F_{ma} in the Theta-PROGSURF version.

5.1 Annual characteristics

The spinup time of Psi and Theta is 2 and 3 years, respectively. The distribution of precipitation between evapotranspiration and total runoff is presented on Fig. 2. The output obtained by Psi is in much better agreement with the observation than that estimated by Theta. The runoff produced by Theta is about 384 mm/year. In spite of this, Psi-PROGSURF's runoff is 338 mm/year which is closer to the observed value of 250 mm/year. Similarly, evapotranspiration obtained by Psi (–435 mm/year) is much closer to the observation (–522 mm/year) than the one obtained by Theta (–389 mm/year).

5.2 Seasonal changes of the water balance components

The monthly water budget is:

$$P_j + E_j + Q_{runj} = \Delta\Theta_j, \tag{36}$$

where j is the month index. P, E and Q_{run} are the monthly sum of precipitation, evapotranspiration and total runoff, respectively. Total runoff is the sum of the surface runoff Q_{runs} , subsurface lateral flows Q_{run1}, Q_{run2}, Q_{run3} and gravitational drainage Q_3. $\Delta\Theta$ is the change in total soil water storage (defined as $[\theta_1+\theta_2+\theta_3] \times 10$ m) from the beginning to the end of the month.

The annual course of observed precipitation is presented on *Fig. 4*. The annual course of E and Q_{run} obtained by the PROGSURF modes are presented on *Figs. 3* and *5*. The differences in evapotranspiration calculated by the modes are most pronounced in summer with a maximum in July. Theta-PROGSURF tends to underestimate the absolut value of the evapotranspiration as compared to the observation. The root-mean-square errors for Psi- and Theta-PROGSURF are 9.5 and 11.5 mm/month, respectively. Total runoff is mostly originated from the gravitational drainage. Q_{runs}, Q_{run1}, Q_{run2} and Q_{run3} are negligible, since their values are about 10 mm/year (not presented here). The annual amplitudes of E and Q_{run} simulated by Psi are greater than the ones simulated by Theta. During most of the winter (months 1–4) Q_{run} is between 35 and 45 mm/month, about the same for both runs. In summer Theta seems to overestimate the runoff as compared to Psi. Therefore the standard deviation of Q_{run} obtained by Psi (3.1 mm/month) is somewhat greater than the one obtained by Theta (2.3 mm/month).

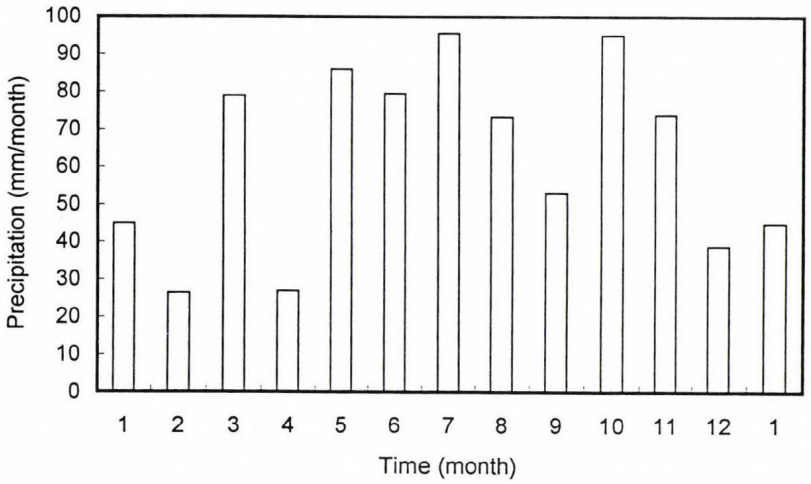


Fig. 4. Annual course of precipitation observed at Cabauw in 1987.

The annual course of soil water change in total soil depth (0–10 m) is presented in *Fig. 6*. Soil water change in total soil depth is significant in April and October. In April the soil water content decreases; in October the situation is reversed. The changes obtained by Theta are smaller than those of Psi. In summer the change is quite variable for both versions of PROGSURF. The standard deviation of $\Delta\Theta$ obtained by Psi and Theta is 3.9 and 4.0 mm/month, respectively.

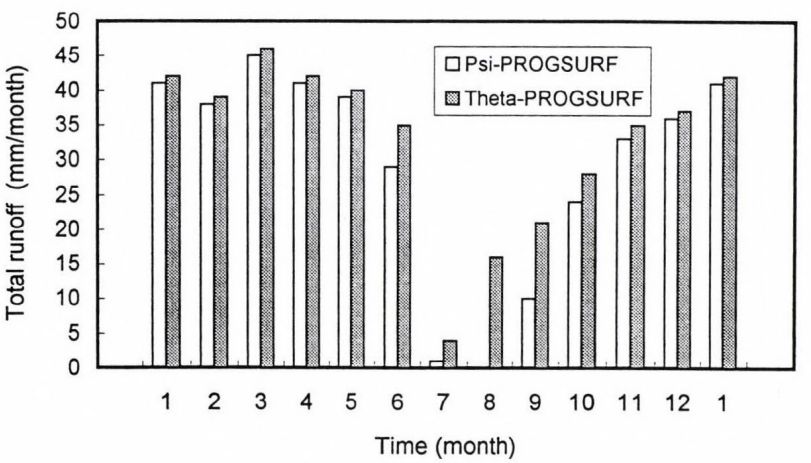


Fig. 5. Annual courses of total runoff.

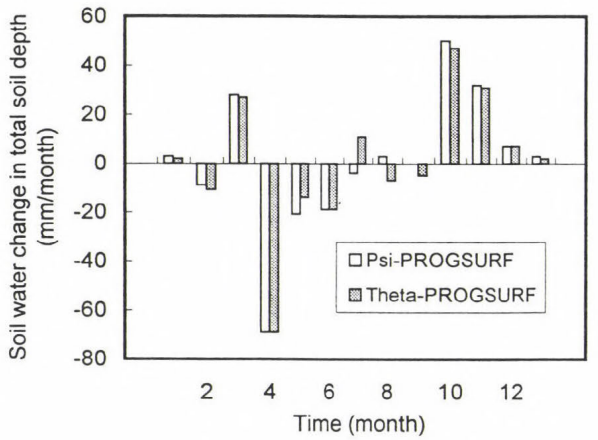


Fig. 6. Annual courses of soil water change in total soil depth of 10 m.

The annual course of soil water in the root-zone ($\theta_1 + \theta_2$, layers 0–1 m, see Fig. 1) is presented on *Fig. 7.* The annual course obtained by Psi-PROGSURF is between about 280 and 390 mm. The corresponding annual mean value is 345 mm. This is in accord with the other PILPS estimates (see Fig. 15 in *Chen et al.*, 1997). In summer, Theta overestimates soil water in the root-zone with respect to Psi. The standard deviation of $(\theta_1+\theta_2) \times 1$ m obtained by Psi and Theta is 12.2 and 11.6 mm, respectively.

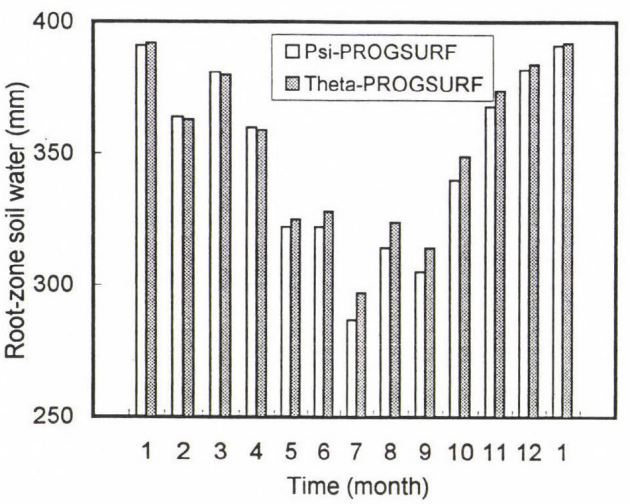


Fig. 7. Annual courses of root-zone soil water.

6. Conclusions

The land-surface hydrology parameterization in the model PROGSURF has been presented. PROGSURF contains 2 vegetation-soil layers and 3 soil layers; these systems overlap each other and represent a 4-layer model. Modeling of the transpiration is the core of the vegetation component. It is implemented using the canopy surface resistance concept with Jarvis' multiplicative formula. The moisture availability is parameterized via leaf water potential Ψ_v. Diffusion type soil moisture prediction is the core of the water transport module. Surface runoff is treated using both Horton and Dunne mechanisms (*Entekhabi* and *Eagleson*, 1989). Gravitational drainage is simply represented through soil water conductivity as parameterized by *Clapp–Hornberger*'s (1978) empirical formula.

PROGSURF has been run using the 1987 data from Cabauw, the Netherland, taking exactly the same specifications that have been applied in the PILPS campaign (*Chen et al.*, 1997). PROGSURF satisfactorily reproduces both the observed annual mean values and the seasonal changes of water fluxes and soil moisture content in the root zone. For example, the annual evapotranspiration and runoff is –435 and 338 mm, respectively. The corresponding observed values of E and Q_{run} are –522 and 250 mm.

The sensitivity of water budget components to different formulations of F_{ma} is also studied. The standard formulation of F_{ma} via leaf water potential (Psi-PROFSURF) is compared to F_{ma} represented by soil moisture contents (Theta-PROGSURF). When F_{ma} is parameterized with soil moisture contents, there is no atmospheric demand effect. On the other hand, when F_{ma} is parameterized via leaf water potential, the effect of the atmosphere is implicitly contained in F_{ma} (*De Ridder* and *Schayes*, 1997), in addition to being represented explicitly by F_{ad}. Our test results give pronounced differences for the water budget components. The deviation in evapotranspiration estimated by Psi- and Theta-PROGSURF is 58 mm/year. This value is as great as the deviations between some PILPS models of completely different structure (see Table 1 in *Chen et al.*, 1997). For example the difference in evapotranspiration between ISBA (Interaction Soil Biosphere Atmosphere) and SWAP (Soil Water-Atmosphere-Plant) is about 50 mm/year. These results suggest that the differences in the choice of the relevant PROGSURF parameters can be as important as the differences in the structure of a large class of land surface models like the ones studied in PILPS.

Further investigations are needed to understand the governing mechanism, which we have tried to represent in PROGSURF still better. Similar tests should also be performed for other climatic regions. It is hoped that further intercomparison campaigns presently in preparation may be useful for optimizing the transpiration subroutines seeking the balance between physics and complexity (*Henderson-Sellers et al.*, 1993).

Acknowledgements—The authors wish to thank the *Österreichische Akademie der Wissenschaften* within the National Austrian Committee for the IGBP and the *Hungarian Ministry for Culture and Education* under grant FKFP No. 0168/1997 for the financial support of this study. We thank *Dr. L. Haimberger* and *Dr. M. Dorninger* for their helpful discussions and help in data transfer processing. The technical help provided by *Mr. Z. Barcza* and by *Mrs. B. Berger* has been very constructive.

References

Ács, F., Mihailovic, D.T. and *Rajkovic, B.*, 1991: A coupled soil moisture and surface temperature prediction model. *J. Appl. Meteor. 30,* 812-822.

Ács, F., 1994: A coupled soil-vegetation scheme: description, parameters, validation, and sensitivity studies. *J. Appl. Meteor. 33*, 268-284.

Ács, F. and *Hantel, M.*, 1997: Transpiration – the relevant atmosphere-surface interaction process. *Book of Abstracts (Annales Geophysicae, Supplement II to Volume 15, Part II, Hydrology, Oceans, Atmosphere & Nonlinear Geophysics) of the 22nd General Assembly of the European Geophysical Society*, C 351.

Ács, F. and *Hantel, M.*, 1998: The land-surface flux model PROGSURF. Accepted in a special PILPS issue of *Global and Planetary Change*.

Beljaars, A.C.M., 1982: The derivation of the fluxes from profiles in perturbed areas. *Bound.-Layer Meteor. 24*, 35-55.

Beljaars, A.C.M. and *Bosveld, F.C.*, 1997: Cabauw data for the validation of land surface parameterization schemes. *J. Climate 10*, 1172-1194.

Chen, T.H., Henderson-Sellers, A., Milly, P.C.D., Pitman, A.J., Beljaars, A.C.M., Polcher, J., Abramopoulos, F., Boone, A., Chang, S., Chen, F., Dai, Y., Desborough, C.E., Dickinson, R.E., Dumenil, L., Ek, M., Garratt, J.R., Gedney, N., Gusev, Y.M., Kim, J., Koster, R., Kowalczyk, E.A., Laval, K., Lean, J., Lettenmaier, D., Liang, X., Mahfouf, J.-F., Mengelkamp, H.-T., Mitchell, K., Nasonova, O.N., Noilhan, J., Robock, A., Rosenzweig, C., Schaake, J., Schlosser, C.A., Schulz, J.-P., Shao, Y., Shmakin, A.B., Verseghy, D.L., Wetzel, P., Wood, E.F., Xue, Y., Yang, Z.-L. and *Zeng, Q.*, 1997: Cabauw experimental results from the project for intercomparison of land-surface parameterization schemes. *J. Climate 10,* 1194-1216.

Choudhury, B., 1983: Simulating the effects of weather variables and soil water potential on a corn canopy temperature. *Agric. Meteorol. 29*, 169-182.

Clapp, R.B. and *Hornberger, G.M.*, 1978: Empirical equations for some soil hydraulic properties. *Water Resour. Res. 14*, 601-604.

De Ridder, K. and *Schayes, G.*, 1997: The IAGL land surface model. *J. Appl. Meteor. 36*, 167-183.

Dickinson, R.E., 1984: Modelling evapotranspiration for three dimensional global climate models. In *Climate Processes and Climate Sensitivity* (eds.: *J.E. Hanson* and *T. Takahashi*). *Geophys. Monogr., Amer. Geophys. Union 29*, 58-72.

Entekhabi, D. and *Eagleson, P.S.*, 1989: Land surface hydrology parameterization for atmospheric general circulation models including subgrid scale spatial variability. *J. Climate 2*, 816-831.

Gates, M.D., 1981: *Biophysical Ecology*. Springer Verlag, 339 pp.

Gerwitz, A. and *Page, E.R.*, 1974: An empirical mathematical model to describe plant root system. *J. Appl. Ecol. 11*, 773-782.

Henderson-Sellers, A., Yang, Z.-L. and *Dickinson, R.E.*, 1993: The project for intercomparison of land-surface parameterization schemes. *Bull. Amer. Meteor. Soc. 74*, 1335-1349.

Holtslag, A.A.M. and *Van Ulden, A.P.*, 1983: A simple scheme for daytime estimates of the surface fluxes from routine weather data. *J. Clim. Appl. Meteor. 22*, 517-529.

Jarvis, P.G., 1976: The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philos. Trans. Roy. Soc. London*, Ser.B., 273, 593-610.

Lynn, B.H. and *Carlson, T.N.*, 1990: A stomatal resistance model illustrating plant vs. external control of transpiration. *Agric. and Forest Meteorol. 52*, 5-43.

Noilhan, J. and Planton, S., 1989: A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev. 117*, 536-550.

Sellers, P.J. and *Dorman, J.L.*, 1987: Testing the simple biosphere model (SiB) using point micrometeorological and biophysical data. *J. of Clim. Appl. Meteorol. 26*, 622-651.

Sellers, P.J., Mintz, Y., Sud, Y.C. and *Dalcher, A.*, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atm. Sci. 43(7)*, 505-531.

Sun, S.F., 1982: Moisture and heat transport in a soil layer forced by atmospheric conditions. *M.S. thesis, Dept. of Civil Engineering, University of Connecticut*, 72 pp.

van der Hornert, T.H., 1948: Water transport as a catenary process. *Discuss. Faraday Soc. 3*, 146-153.

Viterbo, P. and *Beljaars, A.C.M.*, 1995: An improved land-surface parameterization scheme in the ECMWF model and its validation. *J. Climate 8*, 2716-2748.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 2, April–June 1998, pp. 129–131

Short Contribution

Effect of potassium chloride (KCl) on plant water status of semi-dry rice at different time of flooding

N. S. Venkataraman, N. Kempuchetty and **S. Mohandass**

Tamil Nadu Rice Research Institute,
Aduthurai 612 101, India

(Manuscript received 26 January 1998)

Abstract—Field experiment was conducted during the wet season of 1994–95 to study the influence of time of submergence and foliar spray of potassium chloride on plant water status and grain yield of semi-dry rice. The results indicated that the foliar spray of 1% potassium chloride (KCl) on the 30th day after sowing (30 DAS) followed by flooding delayed up to 45 and 60 DAS improved the plant water status by means of reduced leaf temperature and transpiration rate.

Key-words: semi-dry rice, submergence, potassium chloride, plant water status, transpiration, leaf temperature, yield.

Semi-dry rice (dry seeded bunded rice) refers to rice which is sown on dry seedbed as an upland crop taking advantage of monsoon rain. At the fourth or fifth leaf stage, when the seasonal rains intensify or sufficient water supply is available from the canal, the field is submerged like wet land rice (*Lubigan and Moody,* 1982). The major barriers of improving the productivity of rice in this system are moisture stress at different growth stages, percolation loss and inefficient use of plant nutrients.

A field experiment was conducted at Tamil Nadu Rice Research Institute, Aduthurai, India during the wet season of August 1994 - January 1995 to study the effect of potassium chloride on plant water status of semi-dry rice (variety CR 1009) at different time of flooding (viz. 30, 45 and 60 DAS). The rice was sown on clay loam soil with pH 7.2 and organic carbon 0.38% while available N, P and K content were 220, 13.8 and 256.2 kg/ha, respectively. Foliar spray of KCl was given on the 30 DAS as drought mitigation. Measurements of transpiration rate, leaf temperature and relative leaf water content were taken

at tenth day after the foliar spray of KCl. Standard errors for mean difference (SE_D) and critical difference (CD) were worked out at 5% probability level (*Gomez* and *Gomez,* 1984).

The results are shown in *Tables 1* and *2*. Adequate supply of potassium during water stress condition made the guard cell turgid around the stomata and led to regulation of transpiration (*Breg,* 1972; *Duraisamy* and *Rosenberg*, 1977). This might be due to increased K ion content, which facilitates water uptake by increasing the osmotic potential of root cells and xylem sap (*Steineck* and *Header*, 1978) under water stress. Potassium has an effect on the maintenance of cell turgidity and helps to reduce the leaf temperature by transpirational cooling of leaves.

Table 1. Effect of time of submergence on semi-dry rice

Submergence	Transpiration (mmol H_2O m^{-2} sec^{-1})	Leaf temperature (°C)	Relative leaf water content (%)	Water use efficiency (kg ha^{-1} mm^{-1})	Dry matter production (kg ha^{-1})	Grain yield (kg ha^{-1})
30 DAS	12.53	26.43	85.16	2.74	10261	4894
45 DAS	11.36	28.78	76.41	2.84	10252	4788
60 DAS	11.10	29.39	75.60	2.67	9418	4255
SE_D	0.26	0.23	0.39	0.03	66	103
CD	0.92	0.63	1.09	0.08	184	286

Table 2. Combined effect of the time of submergence and of KCl on semi-dry rice

	KCl (1%)					
Submergence	Transpiration (mmol H_2O m^{-2} sec^{-1})	Leaf temperature (°C)	Relative leaf water content (%)	Water use efficiency (kg ha^{-1} mm^{-1})	Dry matter production (kg ha^{-1})	Grain yield (kg ha^{-1})
30 DAS	12.11	26.72	88.38	3.25	11337	5830
45 DAS	9.15	26.58	80.65	3.28	11331	5682
60 DAS	9.71	26.92	78.14	3.08	11190	5027
SE_D	0.37	0.49	0.60	0.05	113	153
CD	0.89	1.06	1.40	0.12	256	NS

In agreement with *Ali* (1985) we can say that under water stress condition, KCl spray given on 30 DAS improved the water status of plants by means of reducing the leaf temperature and the transpiration rate for a shorter period. The subsequent flooding on 45 DAS improved the uptake of nutrients and the water use efficiency favouring increased growth and yield in semi-dry direct sown rice.

References

Ali, M., 1985: Effect of presowing seed treatment, foliar nutrition and planting pattern on productivity and water use in chick pea under rainfed condition. *Legume Res. 8*, 7-11.

Breg, H., 1972: The influence of potassium on the transpiration rate and stomatal opening in *Triticum aestivum* and *Pisum sativum*. *Pl. Physiol. 26*, 250-257.

Duraisamy, P.C. and *Rosenberg, N.J.*, 1979: Reflectant induced modification of soybean canopy radiation balance. Preliminary tests with a Kaolinite reflectant. *Agron. J. 86*, 224-228.

Gomez, K.A. and *Gomez, A.A.*, 1984: *Statistical Procedure for Agricultural Research*. John Wiley and Sons, New York.

Lubigan, R.T. and *Moody, K.*, 1982: Herbicide combination for weed control in dry seeded wet land rice in Philippines. In *Proc. Conf. Plant Prot. in Tropics*. Kualalumpur, Malaysia, pp. 511-518.

Steineck, O. and *Header, H.E.*, 1978: The effect of potassium on growth and yield components of plants. In *Potassium Research Review Trends*. International Potash Institute, Switzerland, pp. 165-185.

BOOK REVIEW

M. Yoshino, M. Domrös, A. Douguédroit, J. Paszynski and *L. Nkemdirim* (eds.): **Climates and Societies — A Climatological Perspective**. Kluwer Academic Publishers, Dodrecht, Boston, London, 1997. 11 + 408 pages, 127 figures, 42 tables and several photos.

In the 1970s a comprehensive scientific activity began related to the impact of climate on human life and the effect of humans on climate. Series of international cooperations were organized, like World Climate Programme (WCP), International Geosphere-Biosphere Programme (IGBP). Intergovernmental Panel on Climate Change (IPCC), Working Group on Tropical Climatology and Human Settlements, International Association on Meteorology and Physics (IAMAP), etc. These wide international cooperations have stored up a great deal of new scientific informations and knowledges, partly published in different periodicals and journals.

This book is written by some of the world's leading climatologists and environmental scientists. It addresses many of issues raised in the debate on global change, providing new insights into climate and its integration into space and time organizations of societies. The volume contains three main parts:

(1) Climatic changes and variability.
(2) Climate on regional scale including climate related problems in major climatic regions ranging from tropics through the temperate zone to the Poles.
(3) Human-climate relationships on local scale.

The introduction was written by *Masatoshi Yoshino*. The first greater part consists of 6 chapters written by *A. Douguédroit; J. Zhang, Y. Yasuda* and *M. Yoshino; Fr. Serre-Bachet; C. D. Schönwiese; H. Le Treut; A. Douguédroit, J.-P. Marchand, M.-F. de Saintignon* and *A. Vidal*.

The introduction, entitled "Human activities and environmental change: a climatologist's wiew", presents data on emissions of greenhouse gases in several countries of Asia, Australia and the Philippines. The author emphasizes the importance of environmental education and outlines the changes in ecosystems, agricultural products, sea level.

Chapter 2.1 (Climate variability and societies) presents two approaches of climate: a physical and a social approach. Among others the reader encounters problems of telecommunication and ENSO events in this chapter. Chapter 2.2 (Climate change in post glacial period in monsoon Asia) is remarkable as far as the warm/cold or wet/dry periods in the last 18000-year period are

considered. Climatic information sources and climate reconstruction methods are the subjects of Chapter 2.3 (Approaches to climatic variations during the historical era). The next chapter (2.4) deals with some statistical aspects of observed regional and global climate change within the instrumental period including autocorrelation spectral analysis, maximum entropy spectral analysis and so forth.

Chapter 2.5 (Climate of the future: on evaluation of the current uncertainties) gives a good review on achievements of the IPCC researches. Chapter 2.6 (Impacts on the climate variability on human activities) is a comprehensive summary of the topic.

The next great part, entitled "Regional Scales Climates" consists of Chapters 3.1–3.7 and deals with topics like society-climate systems in tropical Africa, climate and life in the Caribbean Basin, climate and societies in Southeast Asia, climatic and pathological rhythms in a humid tropical area (the case of the Philippines), possible impact of the climatic change and variability on agriculture in South America, climate and agriculture in China and the climates of the polar regions.

The last part, "Local scale climates" including Chapters 4.1–4.5, gives reviews on problems of local climate and man, namely on interactions of man and climate in urban world, the influence of urbanisation on the local climate and the influence of urban climate on man, climates of tropical and subtropical societies, air pollution (a local problem becomes a global problem), agricultural landuse and local climate.

The authors of these chapters are as follows: *O. Ojo; L. C. Nkemdirom; K.U. Sirinanda; J. Pérard* and *J.-P. Besancenot; F.- Santibañez; A.-L. Jiang; T. Niedzwiedz; J. Paszynski; F. Wilmers; E. Jauregui; H. Wanner; M. Yoshino*.

Besides the great variety of subjects, this book provides numerous sources of up-to-date knowledge on past, present and future climates, their problems and impacts.

G. Koppány

NEWS

BSRN 5th Science and Review Workshop

The BSRN (Baseline Surface Radiation Network) is one of the subprogrammes of WCRP (World Climate Research Programme). Its purpose is to establish 20–30 surface radiometric stations to provide most reliable radiation data for

- deriving of climatic trends in radiation parameters;
- validating of surface radiation products derived from satellite measurements.

The BSRN started in 1988 in Würzburg (Germany). The 5th Workshop was held in Budapest from 18 to 22 May, 1988. The objective of this workshop was to bring BSRN scientists, station managers, data users and experts in areas related to BSRN activities together to review

- the status of the implementation of the network;
- the latest developments in instrumentation and operational procedures, data management and quality control;

to discuss some of the scientific progresses achieved as a result of availability of the BSRN data archive; and to consider future needs and plans for the BSRN.

The meeting was attended by 43 participants, the largest number in the history of BSRN workshops. The participants presented nearly 20 station status reports and more than 20 technical and scientific reports.

The longest discussion was paid to the "BSRN Operations Manual (Version 1.0)" compiled by *Dr. Bruce McArthur* (Canada). This manual will be available at the WCRP Office's web site. It contains sections on sampling frequency and accuracy requirements for BSRN stations, the siting of the stations, the installation of radiation instruments, solar tracking devices, data acquisition, station maintenance, radiation instrument calibration, radiation data reduction and quality assurance procedures as well as a variety of ancillary information in the annexes.

The BSRN Working Group approved a 4-page document entitled: "BSRN data release policy". It was compiled by a subgroup that had been nominated during the 1996 meeting in Boulder (USA). The basic idea is to defend the interests of the station scientists who produce the data.

Other items addressed during the meeting:

- aerosol optical depth,
- geometry of pyrheliometers and diffusometers,

- all-weather cavity pyrheliometers,
- UV-B measurements, liaison to GAW program,
- pyranometer offsets,
- photosynthetically active radiation,
- complete literature search for existing published papers on the performance and characteristics of thermopile based radiometers used for solar and terrestrial irradiance measurements,
- albedo and surface properties,
- longwave irradiance reference,
- longwave data reduction,
- shortwave leak in pyrgeometer domes,
- cloud base height,
- pyrgeometer characterization, calibration and classification.

A comprehensive report on this BSRN meeting will be compiled by *Roger Newson* and *Hans Teunissen* (WCRP Office) and it will be published by the WMO in the second half of 1998.

The next workshop is planned for the year of 2000.

Major György

ATMOSPHERIC ENVIRONMENT

an international journal

To promote the distribution of Atmospheric Environment *Időjárás* publishes regularly the contents of this important journal. For further information the interested reader is asked to contact *Prof. P. Brimblecombe*, School for Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, U.K.; E-mail: atmos_env@uea.ac.uk

Volume 32 Number 3 1998

International Conference on Atmospheric Ammonia: Emission, Deposition and Environmental Impacts

Biosphere-atmosphere exchange

D.S. Lee, C. Halliwell, J.A. Garland, G.J. Dollard and *R.D. Kingdon:* Exchange of ammonia at the sea-surface – a preliminary study, 134-439.

G.P. Wyers and *J.W. Erisman:* Ammonia exchange over coniferous forest, 441-451.

D. Fowler, C.R. Flechard, M.A. Sutton and *R.L. Storeton-West:* Long-term measurements of the land-atmosphere exchange of ammonia over moorland, 453-459.

Canopy exchange processes and physiological control of ammonia fluxes

J.K. Schjoerring, S. Husted and *M. Mattsson:* Physiological parameters controlling plant-atmosphere ammonia exchange, 491-498.

A. Neftel, A. Blatteeer, A. Gut, D. Högger, F. Meixner, C. Ammann and *F.J. Nathaus:* NH_3 soil and soil surface gas measurements in a triticale wheat field, 499-505.

J.K. Schjoerring, S. Husted and *M.M. Poulsen:* Soil-plant-atmosphere ammonia exchange associated with *Calluna vulgaris* and *Deschampsia flexuosa*, 507-512.

Impacts of ammonia and ammonium deposition

L. van der Eerden, W. De Vries and *H. van Dobben:* Effects of ammonia deposition on forests in The Netherlands, 525-532.

J. Pearson and *A. Soares:* Physiological responses of plant leaves to atmospheric ammonia and ammonium, 533-538.

Z.H. Yin, W. Kaiser, U. Heber and *J.A. Raven:* Effect of gaseous ammonia on intracellular pH values in leaves of C_3- and C_4-plants, 539-544.

Critical loads and ammonia control policies

K.R. Bull and *M.A. Sutton:* Critical loads and the relevence of ammonia to an effects-based nitrogen protocol, 565-572.

D.A. Cowell and *H.M. Apsimon:* Cost-effective strategies for the abatement of ammonia emissions from European agriculture, 573-580.

L.J.A. Lekkerkerk: Implications of Dutch ammonia policy on the livestock sector, 581-587.

Volume 32 Number 4 1998

G. Drakou, C. Zerefos, I. Ziomas and *M. Voyatzaki:* Measurements and numerical simulations of indoor O_3 and NO_x in two different cases, 595-610.

R.J. Brown and *R.W. Bilger:* Experiments on a reacting plume–1. Convertional concentration statistics, 611-628.

R.J. Brown and *R.W. Bilger:* Experiments on a reacting plume–2. Conditional concentration statistics, 629-646.

Alexey Ryaboshapko, Laura Gallardo, Erik Kjellström, Sergey Gromov, Sergey Paramonov, Olga Afinogenova and *Henning Rodhe:* Balances of oxidized sulfur and nitrogen over the former Soviet Union territory, 647-658.

Juão F.P. Gomes: Assessment of opacimeter calibration on kraft pulp mills, 659-664.

Robert L. Lee and *Erik Näslund:* Lagrangian stochastic particle model simulations of turbulent dispersion around buildings, 665-672.

L. Cheng, K.M. McDonald, R.P. Angle and *H.S. Sandhu:* Forest fire enhanced photochemical air pollution. A case study, 673-681.

Volume 32 Number 5 1998

Atmospheric Transport, Chemistry and Deposition of Mercury

Volume 32 Number 6 1998

NOTES TO CONTRIBUTORS OF *IDŐJÁRÁS*

The purpose of the journal is to publish papers in any field of meteorology and atmosphere related scientific areas. These may be

- reports on new results of scientific investigations,
- critical review articles summarizing current state of art of a certain topic,
- shorter contributions dealing with a particular question.

Each issue contains "News" and "Book review" sections.

Authors may be of any nationality, but the official language of the journal is English. Papers will be reviewed by unidentified referees.

Manuscripts should be sent to
Editor-in-Chief of IDŐJÁRÁS
P.O. Box 39
H-1675 Budapest, Hungary

in three copies including all illustrations. One set of illustrations has to be of camera ready quality, the other two might be lower quality.

Title part of the paper should contain the concise title, the name(s) of the author(s), the affiliation(s) including postal and E-mail address(es). In case of multiple authors, the cover letter should indicate the corresponding author.

Abstract should follow the title, it contains the purpose, the data and methods as well as the basic conclusion.

Key-words are necessary to help to classify the topic.

The text has to be typed in double spacing with wide margins. Word-processor printing is preferred. The use of SI units are expected. The negative exponent is preferred to solidus. Figures and tables should be consecutively numbered and referred to in the text.

Mathematical formulas are expected to be as simple as possible and numbered in parentheses at the right margin. Non-Latin letters and hand-written symbols should be indicated and explained by making marginal notes in pencil.

Tables should be marked by Arabic numbers and printed in separate sheets together with their captions. Avoid too lengthy or complicated tables.

Figures should be drawn or printed in black and white, without legends, on separate sheets. The legends of figures should be printed as separate list. Good quality laser printings are preferred as master copies.

References: The text citation should contain the name(s) of the author(s) in Italic letter and the year of publication. In case of one author: *Miller* (1989), or if the name of the author cannot be fitted into the text: *(Miller,* 1989); in the case of two authors: *Gamov* and *Cleveland* (1973); if there are more than two authors: *Smith et al.* (1990). When referring to several papers published in the same year by the same author, the year of publication should be followed by letters a,b etc. At the end of the paper the list of references should be arranged alphabetically. For an article: the name(s) of author(s) in Italics, year, title of article, name of journal, volume number (the latter two in Italics) and pages. E.g. *Nathan, K.K.,* 1986: A note on the relationship between photosynthetically active radiation and cloud amount. *Időjárás 90*, 10-13. For a book: the name(s) of author(s), year, title of the book (all in Italics except the year), publisher and place of publication. E.g. *Junge, C. E.,* 1963: *Air Chemistry and Radioactivity*. Academic Press, New York and London.

The final version should be submitted on diskette altogether with one hard copy. Use standard 3.5" or 5.25" DOS formatted diskettes. The preferred word-processors are WordPerfect 5.1 and MS Word 6.0.

Reprints: authors receive 30 reprints free of charge. Additional reprints may be ordered at the authors' expense when sending back the proofs to the Editorial Office.

More information: gmajor@met.hu
Information on the last issues:
http://www.met.hu/firat/ido-e.html

Published by the Hungarian Meteorological Service

Budapest, Hungary

INDEX: 26 361

HU ISSN 0324-6329

IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

CONTENTS

****

http://www.met.hu/firat/ido-e.html

VOL. 102 * NO. 3 * JULY–SEPTEMBER 1998

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: P.O. Box 39, H-1675 Budapest, Hungary or
Gilice tér 39, H-1181 Budapest, Hungary
E-mail: gmajor@met.hu or *antal@met.hu*
Fax: (36-1) 290-7387

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary;
E-mail: gmajor@met.hu or *antal@met.hu*; *Fax: (36-1) 290-7387*

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 3, July–September 1998, pp. 141–147

Temporal variation of the atmospheric sulfur budget over Hungary during 1980–1996

László Bozó

Hungarian Meteorological Service,
H-1675 Budapest P.O. Box 39, Hungary
E-mail: bozo@met.hu

(Manuscript received 9 March 1998; in final form 18 May 1998)

Abstract—Annual atmospheric sulfur budgets over Hungary for the period of 1980–1996 are investigated in the paper using simple mass balance equations. It was found that the sulfate budget is almost balanced over the country, however, regarding sulfur-dioxid, Hungary is still a net atmospheric exporter in spite of the fact that Hungarian SO_2 emission decreased significantly during the past years. Data for the atmospheric sulfur concentration and deposition gained from the Hungarian regional background air pollution monitoring network are also presented for the period investigated.

Key-words: SO_2-emission, sulfur deposition, atmospheric budget, long-range transport.

1. Introduction

Sulfur components play significant role in acid rain formation. In addition, fine size range sulfate particles can reduce the effect of potential global warming — at least on regional scale — through the direct reflection of incoming solar radiation back to the space. Emission of SO_2 has changed significantly in Europe during the last 15 years which should also be resulted in the temporal variation of the atmospheric sulfur budget of Hungary. On the basis of the method used, the horizontal transport of sulfur can be estimated so conclusions can be drawn on the role and its temporal variation of Hungary in the European atmospheric sulfur budget.

The calculations were made for a box over Hungary up to the tropopause without any exchange with the stratosphere. It is assumed that there is no accumulation of sulfur in the box over a year's period which means that the gains and losses of the box are equal on an annual average.

2. *Equations used for the calculations*

The mass balance equations for SO_2 and sulfate can be given as follows:

$$E + A_{im} = D_{dry} + D_{wet} + T + A_{ex}, \tag{1}$$

$$T + A^*_{im} = D^*_{dry} + D^*_{wet} + A^*_{ex}, \tag{2}$$

where

E: SO_2 emission from the surface,
A_{im} and A^*_{im}: import terms for SO_2 and sulfate,
A_{ex} and A^*_{ex}: export terms for SO_2 and sulfate,
D_{dry} and D^*_{dry}: dry deposition of SO_2 and sulfate,
D_{wet} and D^*_{wet}: wet deposition of SO_2 and sulfate,
T: mass of SO_2 transformed into sulfate over the country.

Let

$$AD = A_{im} - A_{ex}, \tag{3}$$

$$AD^* = A^*_{im} - A^*_{ex}. \tag{4}$$

These terms are calculated using the mass balance equations (Eqs. (1) and (2)):

$$AD = D_{dry} + D_{wet} + T - E, \tag{5}$$

$$AD^* = D^*_{dry} + D^*_{wet} - T. \tag{6}$$

3. *Emission of sulfur compounds in Hungary*

The natural source strength of sulfur species has been estimated on the basis of data compilation by *Várhelyi* and *Gravenhorst* (1981). These authors estimated the average continental emission to be 0.01–0.03 g S m^{-2} a^{-1}. Thus, the corresponding value for Hungary is in the order of 10^{-3} Tg S a^{-1} (1 Tg = 10^{12} g). Anthropogenic sulfur emission data — which is represented in the form of SO_2 release — are based on the official national reports of the Ministry of Environment, Hungary. The annual emission values are shown in *Fig. 1*. It can be seen that the release of sulfur dioxid has been decreasing since the early 80's. Its value was 0.82 Tg S a^{-1} in 1980 and 0.38 Tg S a^{-1} in 1996. The decrease can be explained by the reduced industrial activity of the country and — from the 90's —

by the replacement of out-of-date technologies by less energy consumer and more environment friendly ones. Coal, which was used for home heating widely during the previous decades is now also being replaced by natural gas which causes only very low sulfur emission from the heaters.

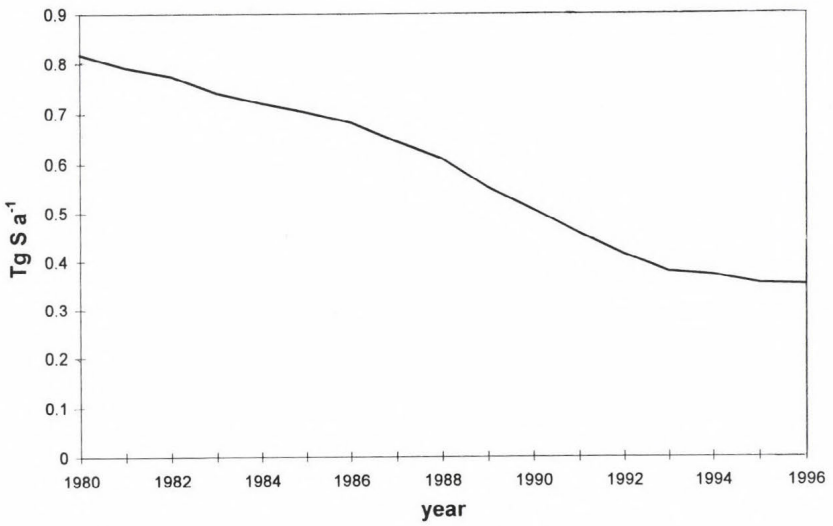


Fig. 1. Temporal variation of SO_2 emission from Hungary.

4. Concentration of SO_2 and sulfate over Hungary

Monitoring of SO_2 and sulfate under regional background conditions was started in the 70's in Hungary. The longest Hungarian measurement record is available from K-puszta. In the early 80's Farkasfa station also joined the SO_2 monitoring activity. Since 1996 there have been two additional regional background stations in operation in Hungary measuring also SO_2 and sulfate. Locations of stations are mapped in *Fig. 2.*

Temporal variation of SO_2 and sulfate concentrations under regional background conditions at K-puszta station is shown in *Fig. 3.* Due to the fact that during the period of 1992–1994 a lot of SO_2 data were missing for technical reasons, annual averages used in the atmospheric budget calculations were taken from the EMEP calculations (*Barrett et al., 1995*). One can see from the figure that SO_2 concentrations were higher in the first half of the period investigated (1980–1988) than in the second one. It is in accordance with the reduced sulfur emission from Hungary and other European countries.

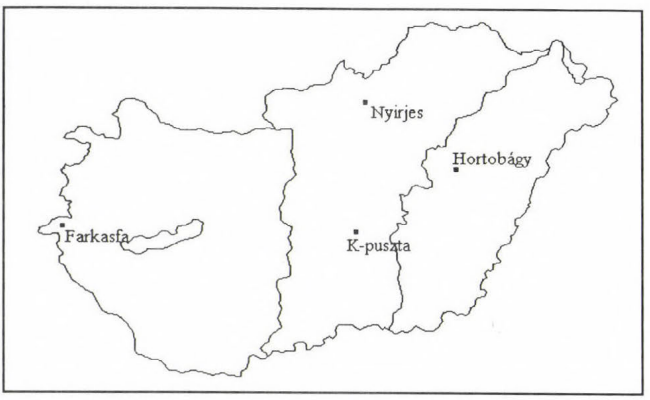


Fig. 2. Regional background monitoring stations in Hungary.

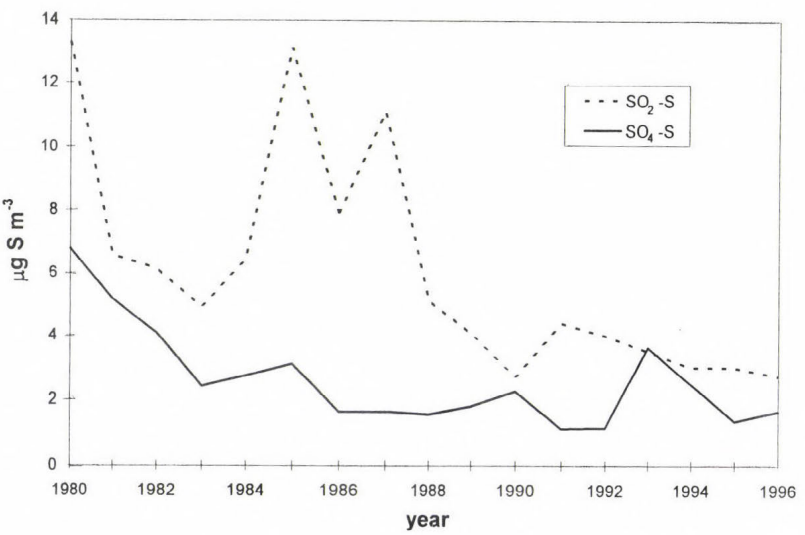


Fig. 3. Temporal variation of SO_2 and sulfate concentration.

5. Atmospheric SO_2-sulfate transformation and sulfur deposition over Hungary

The oxidation process of SO_2 in the troposphere is initiated by its reaction with free OH^- radicals producing SO_3 in two steps. Then SO_3 reacts with water vapour producing sulfuric acid. Transformation ratio can be estimated by both laboratory experiments and direct field measurements. The results show

relatively big variation (*Warneck*, 1988). Laboratory experiments indicate lower values for the intensity of transformation which can be explained by the fact that liquid phase atmospheric reactions might speed up the transformation of sulfur dioxide into sulfate. In present paper the transformation ratio of 3.6%/hour was used for the calculations measured by *Horváth* and *Bónis* (1980).

Dry deposition of atmospheric sulfur compounds can be determined by multiplying the averaged concentration near the surface and deposition velocity. This latter value was taken from the field measurements of *Horváth et al.* (1996) who carried out vertical profile measurements for SO_2 and meteorological parameters in Hungary so that to estimate the dry flux of SO_2 from the atmosphere onto the surface. Based on their measurements and calculation an annual average of 0.6 cm s^{-1} was used for atmospheric budget estimations. Dry deposition velocity for sulfate particles was taken from *Möller* (1983). Its value is 0.1 cm s^{-1}.

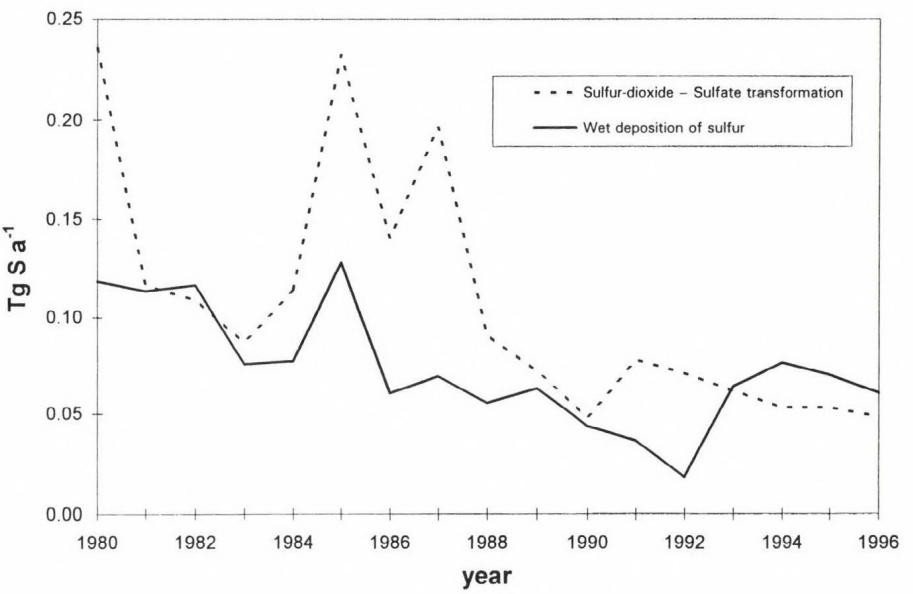


Fig. 4. Temporal variation of SO_2-sulfate transformation and wet deposition of sulfur over Hungary.

Long-term variation of sulfur wet deposition as well as SO_2-sulfate transformation is plotted on *Fig. 4*. Wet only samplers at 10 background sites of Hungary were used for collecting precipitation samples. Wet deposition of sulfur is estimated by averaging the wet deposition rates from the stations. It

can be seen from the figure that the lowest deposition rate were detected in 1992 and the highest in 1985. Since the wet deposition of sulfur dioxide and sulfate can only been measured altogether as the sulfate content of precipitation, it was needed to separate the rain-out and wash-out processes for SO_2 and sulfate. Scavenging ratios applied in the EMEP model (*Barrett et al.,* 1995) were used for our budget calculations.

6. The atmospheric sulfur budget over Hungary

The differences between horizontal atmospheric import and export terms (see Eqs. (3) and (4)) are plotted in *Fig. 5*. Sulfate budget is more balanced as compared to that of sulfur-dioxide. Its largest negative value was calculated for 1987. During the last four years it was going to be close to zero. Regarding AD, its value was strongly negative in the early 80's. During the period of 1993–1996 it is stabilized at around –0.20 Tg S a^{-1}. It means that even in 1996, more sulfur dioxide was exported horizontally in the atmosphere from Hungary than the country imported from other countries, so Hungary is still playing a net exporter role in the European sulfur budget. In *Fig. 6* it is plotted how much percent of the Hungarian sulfur emission is compensated by total sulfur deposition and transformation. Its lowest value was detected in 1990 while the highest one — over 50% — in 1985.

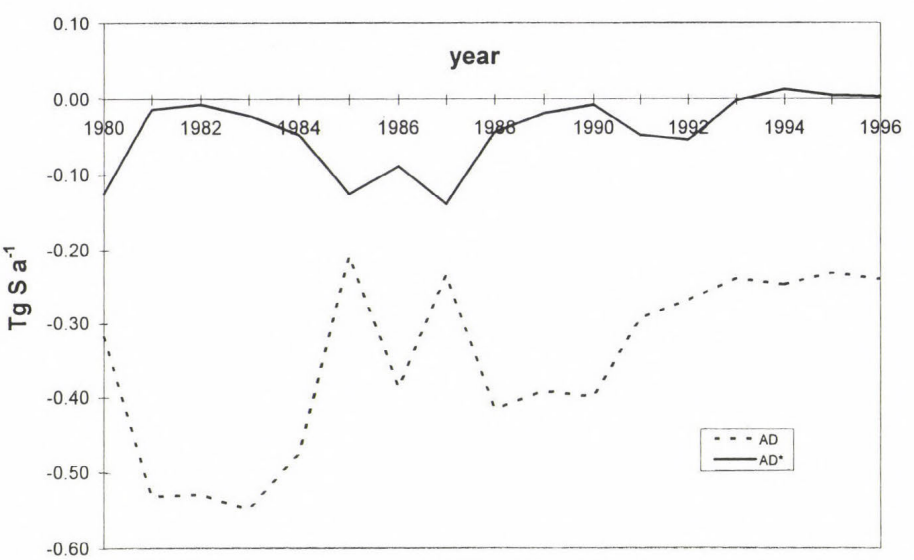


Fig. 5. Temporal variation of horizontal transport terms of sulfur over Hungary.

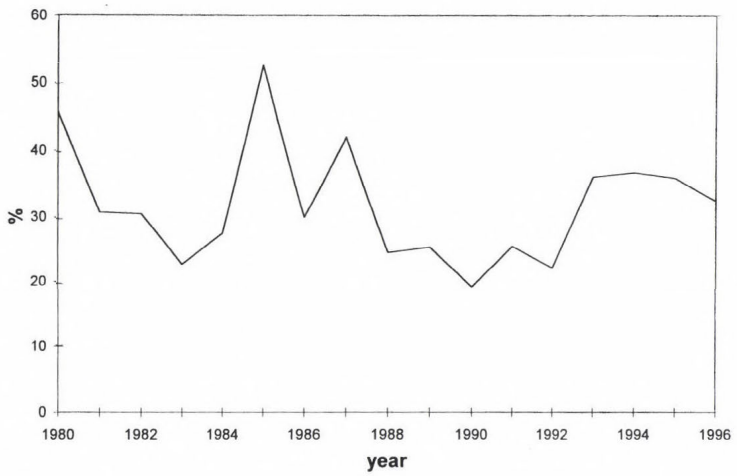


Fig. 6. Hungarian sulfur emission compensated by total deposition and transformation.

7. Conclusions

Although the sulfur dioxid emission of Hungary decreased rapidly during the last decade, the country still plays a net atmospheric exporter role for sulfur dioxide in Europe. Atmospheric sulfate budget is balanced, thus the amount of sulfate imported by the country is roughly equal to that of exported by Hungary. It is expected that sulfur budget will also vary in the future due to the continuous completion of the second Sulfur Protocol: Hungary should reduce its sulfur-dioxide emission by 50% until year of 2005 and by 60% until 2010 as related to the emission level reported in 1980. On the basis of these numbers, Hungarian SO_2 emission in 2010 must not exceed the limit of 0.327 Tg S a^{-1}.

References

Barrett, K., Seland, O., Foss A., Mylona, S., Sandnes, H., Styve, H. and *Tarrasón, L.*, 1995: European Transboundary Acidifying Air Pollution: ten years calculated fields and budgets to the end of first Sulphur Protocol. *EMEP/MSC-W, Report* 1/95.

Horváth, L. and *Bónis, K.* 1980: An attempt to estimate the rate constant of sulfur dioxide — sulfate conversion in the urban plume of Budapest. *Időjárás 84*, 190-195.

Horváth, L., Weidinger, T., Nagy, Z. and *Führer, E.*,1996: Measurement of dry deposition velocity of ozone, sulfur dioxide and nitrogen oxides above pine forest and low vegetation in different seasons by the gradient method. *Proceedings of EUROTRAC Symposium* '96. Computational Mechanics Publications, Southampton, 315-318.

Möller, D., 1983: The global sulfur cycle. *Időjárás 87*, 121-143.

Várhelyi, G. and *Gravenhorst, G.*, 1981: An attempt to estimate biogenic sulfur emission into the atmosphere. *Időjárás 85*, 126-33.

Warneck, P., 1988: *Chemistry of the Natural Atmosphere.* Academic Press, San Diego.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 3, July–September 1998, pp. 149–158

Non-parametric estimation of climate trends

István Matyasovszky

Department of Meteorology, Eötvös Loránd University,
H-1083 Budapest, Ludovika tér 2; E-mail: matya@ludens.elte.hu

(Manuscript received 6 February 1998; in final form 25 June 1998)

Abstract—One of the most important statistical characteristics of climate variables is their expected value. Under a changing climate the expected value can also change. The expected value at a given time can be simply estimated by averaging observations being in an interval around this time. A basic question is how to define the length of this interval. It may be supposed that a weighted average results in a more accurate estimate than ordinary averaging. How to choose weights in that case? As the intensity of change of climate can also change it is reasonable to use variable averaging time length. Which techniques are available for this purpose? In the paper, a summary of mathematical details is presented and three examples are shown to illustrate the importance of above questions.

Key-words: smoothing, weighted local regression, bandwidth, kernel function.

1. Introduction

Consider the observations $y_0, y_1, \ldots, y_n$ of a climate variable at $t_0 < t_1 < \ldots < t_n$ in the following form

$$y_j = f(t_j) + e_j, \quad j = 0, \ldots, n$$

where $f(t)$ is the so-called trend function representing temporal behavior of the expected value. We may generally assume that the superposing fluctuations come from a white noise process, i.e.,

$$E[e_j] = 0, \quad E[e_j^2] = d^2, \quad E[e_i e_j] = 0, \quad i \neq j$$

where E denotes the expected value operation.

There are two basic approaches to estimate trend functions. One of them includes the *parametric method* when $f(t)$ is considered as a linear combination

of known functions $g_0(t),\ldots, g_{k-1}(t)$ $(k>1)$. The task is then to estimate optimum parameters of the linear combination generally by the method of least squares. Besides the simplicity of this procedure, an important advantage is its ability to assess the statistical significance of changes in observed series. However, the functions $g_0(t),\ldots, g_{k-1}(t)$ are generally not known in practice and an appropriate choice of them is a difficult problem.

The other basic approach includes *non-parametric methods* when no assumption on the shape of $f(t)$ is needed. The trend function is then approximated locally by different polynomials at each t. This results in the following weighted averaging of observations:

$$\hat{f}(t) = \sum_{j=0}^{n} w_j(t) y_j, \tag{1}$$

which is frequently called smoothing of data. Several techniques are available to choose the weights $w_j(t)$ (*Nadaraya*, 1964; *Cleveland*, 1979; *Gasser* and *Muller*, 1984; *Eubank*, 1988), which, however, have the same asymptotic properties. Asymptotics are defined such that the support of $f(t)$ is fixed, and taking $n\to\infty$ (i.e., distance between neighboring observations becomes smaller and smaller) the properties of estimation Eq. (1) are obtained. For simplicity but without any loss of generality the support of $f(t)$ is chosen as the interval [0,1] for mathematical formulation. (Obviously, the actual support of $f(t)$ can result in [0,1] after a scale transformation of time t.)

The following section summarizes asymptotic properties of non-parametric regression techniques. Choice of weights and the averaging time is also discussed here. Section 3 presents the best estimator Eq. (1) of regression functions $f(t)$ for fixed numbers of observations. Section 4 shows three examples to demonstrate some theoretical problems in practice. Finally, a section for discussion is provided.

2. Asymptotic properties

A reasonable assumption, when estimating $f(t)$, is that observations at instants close to t have relatively large weights, while observations far from t have small weights in Eq. (1). There are several traditional versions for the concept, such as binomial, exponential, or Gaussian smoothers. However, the choice of both the actual smoother and the averaging time is quite arbitrary and these smoothers reflect only basic characteristics of time series. Therefore, there was a need for a well-developed theory in order to overcome above difficulties. To illustrate this theory we are focusing on weighted local regression (WLR) (*Cleveland*, 1979). Weighted local version of the method of least squares at t results in

$$\hat{f}(t) = \hat{a}_0,$$

where $\hat{a}_0$ is obtained by minimizing

$$\sum_{j=0}^{n} (y_j - [a_0 + a_1(t - t_j) + \ldots + a_{k-1}(t - t_j)^{k-1}])^2 K((t - t_j)/h) \tag{2}$$

with respect to a_0, $a_1,\ldots, a_{k-1}$. The kernel function K determines how to decrease the influence of a squared difference in the sum of squared differences Eq. (2) when moving off t, and the bandwidth h determines the rate of this decrease.

In order to construct an estimate having "good" properties some criteria should be satisfied for the kernel. K is defined on the interval $[-1,1]$ and is called of order k if

$$\int_{-1}^{1} K(z) z^j dz = \begin{cases} 0, & 0 < j < k \\ 1, & j = 0 \end{cases}$$

is satisfied. Then, assuming that kth derivative of f exists, under conditions $n \to \infty$, $h \to 0$, $nh \to \infty$
the bias

$$E[\hat{f}(t)] - f(t) = u_k(t),$$

the variance

$$E[(\hat{f}(t) - E[\hat{f}(t)])^2] = v_k^2(t)$$

and the mean squared error

$$MSE(t) = E[(\hat{f}(t) - f(t))^2] = u_k^2(t) + v_k^2(t) \tag{3}$$

of WLR are obtained asymptotically by

$$u_k(t) = \frac{(-1)^k}{k!} h^k B_k f^{(k)}(t), \quad v_k^2(t) = \frac{d^2}{nh} V_k f^2(t), \tag{4}$$

where

$$B_k = \int_{-1}^{1} K(z) z^k dz, \quad V_k = \int_{-1}^{1} K^2(z) dz,$$

and $f^{(k)}$ denotes kth derivative of f. Eq. (4) shows that the bias is proportional to $f^{(k)}$ and is small when the smoothing is small. The variance is proportional to f^2 and is large when the smoothing is small. A small bias, therefore, requires small bandwidths, while a small variance needs large bandwidths. A trade-off between the two requirements, i.e., an estimate of the optimal bandwidth could be defined by minimizing the asymptotic integrated mean squared error

$$IMSE = \int_0^1 (u_k^2(t) + v_k^2(t))\,dt \tag{5}$$

when the trend function and its kth derivative were known. Due to lack of knowing these functions an unbiased and consistent estimate of Eq. (5) is constructed as

$$CV(h) = \frac{1}{n+1} \sum_{i=0}^{n} (y_i - \hat{f}_i(t_i))^2 \tag{6}$$

and its minimum with respect to h gives the optimal bandwidth. Here $\hat{f}_i(t_i)$ is an estimate Eq. (1) of $f(t_i)$ but leaving out y_i.

Eqs. (3) and (4) show that the mean squared error of WLR depends on properties of f. This fact motivates the choice of locally varying bandwidths. A smaller bandwidth where f is rough ($f^{(k)}$ is, in absolute value, large) reduces bias and a larger bandwidth in the smooth regions of f ($f^{(k)}$ is, in absolute value, small) reduces variance. In general, construction of a local bandwidth estimator entails a two-step procedure. The first step produces a pilot estimator using a fixed bandwidth, and the second stage yields the local bandwidth estimator. For instance, having a global bandwidth obtained by minimizing the quantity Eq. (6) an estimation of $f(t)$ is performed, then using a finite difference scheme its kth derivative is also estimated. We mention that derivatives of trend functions can be directly estimated from observations using some generalized versions of non-parametric regressions (*Fan* and *Gijbels*, 1995). $f(t)$ and $f^{(k)}(t)$ in Eqs. (3) and (4) are then substituted by their corresponding estimates, and Eq. (3) is minimized at each t.

Note that the above technique is based on asymptotics. One can hope that usage of exact properties can yield better estimates. Exact bias and variance can be written as

$$u(t) = \int_{-1}^{1} K(z) f(t - h(t)z)\,dz - f(t),$$

$$v^2(t) = \frac{1}{nh(t)} \int_{-1}^{1} K^2(z) f(t - h(t)z)\, dz.$$

After substituting $f(t)$ by its pilot estimate, the exact mean squared error

$$MSE(t) = u^2(t) + v^2(t)$$

is minimized at each t to have local bandwidth (*Staniswalis,* 1989). A further advantage of the method is that it does not require derivatives of f. It is crucial in areas where $f^{(k)}$ equals to zero and, therefore, the estimated local bandwidth can be extremely large when using asymptotic formulae. An important question is how the local bandwidth depends on pilot global bandwidths. Theoretical considerations and simulation experiments show that for a large range of pilot bandwidths, the estimated trend functions differ slightly (*Staniswalis*, 1989).

Another important question is the choice of kernel. Evidently, it is useful to choose that kernel which delivers the smallest asymptotic variance for a fixed k. These kernels for different k can be found in *Gasser et al.* (1985). *Muller* (1984) produced kernels minimizing variance of mth ($m \geq 0$) derivative of $\hat{f}$. The only remaining question now is the order of kernels, i.e., the choice of k. Theoretical reasons and simulation work show that a small value of k is reasonable. Generally, $k = 2$ (locally linear approximation of f) is satisfactory since the choice of bandwidth is a much more serious problem than the choice of the order of the kernel (*Simonoff,* 1996).

3. Optimal smoothing

Knowing asymptotic properties of non-parametric techniques is highly useful for several reasons. However, the final question is that which method in practice is preferred, when the number of observations is given, i.e., n is fixed. After *Fan* (1993), the best local linear estimation ($k = 2$)

$$\hat{f}(t) = \sum_{j=0}^{n} \alpha_j(t) y_j \tag{7}$$

of f is the weighted local regression, and the weights can be calculated as

$$\alpha_j(t) = w_j(t) / \sum_{i=0}^{n} w_i(t),$$

where

$$w_j(t) = K((t - t_j)/h)\,(s_2 - (t - t_j)s_1), \quad j = 0,\ldots,n$$

and

$$s_l = \sum_{j=0}^{n} K((t - t_j)/h)\,(t - t_j)^l, \quad l = 1,2.$$

Furthermore, K is

$$K(z) = \frac{3}{4}(1 - z^2), \quad z\epsilon[-1,1],$$

the so-called Epanechnikov kernel (*Fan*, 1992).

The optimality of Eq. (7) is the minimax optimality over linear smoothers, which is defined as follows. Under quite weak assumptions for f (*Fan*, 1993), WLR based on the Epanechnikov kernel achieves the minimum value over all linear smoothers of the maximum of $MSE(t)$, and has a minimax efficiency at least 89.6%. In other words, the maximum of $MSE(t)$ of WLR with Epanechnikov kernel is at least 10.4% smaller than the maximum of $MSE(t)$ of any other linear smoother.

4. Examples

The first example demonstrates that non-parametric regression can be a more effective technique to estimate trends than parametric regression. Annual temperature anomalies over Northern Hemisphere from 1854 to 1993 (*Jones et al.*, 1994) are examined by fitting a common linear trend and using WLR. Evidently, linear trend basically describes the warming tendency but with considerable departures from data during several subperiods of the entire period (*Fig. 1*). The trend obtained by WLR fits much better the data. In addition to the good fit, WLR provides a smoother curve than a frequently used, say 10-year, running mean (not shown here). Another curve corresponding to WLR with local bandwidth is also presented in Fig. 1. Difference between trends estimated with global and local bandwidths is not considerable due to the small variance of data. The small variance makes it possible to use small global bandwidth (15 years) resulting in a small bias with small variance. Therefore, a somewhat larger or smaller local bandwidth at a given time t provides almost the same estimate.

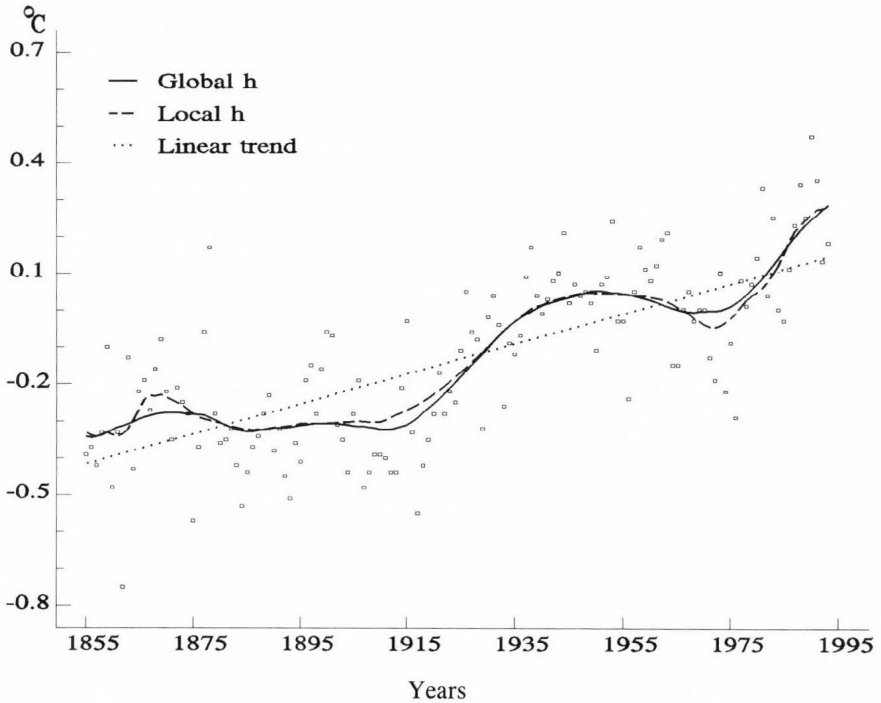


Fig. 1. Trend of annual temperature anomalies over Northern Hemisphere.

The second example represents that quantity Eq. (6) is sometimes a highly complicated function of h. How does the shape of trend functions influence the behavior of Eq. (6)? To illustrate this problem let's consider monthly SOI time series. SOI (Southern Oscillation Index) is defined as the difference between sea surface air pressures of Tahiti and Darwin, and is used to characterize ENSO phenomena. SOI data set has been obtained from NCAR (National Center for Atmospheric Research) for the period from 1868 to 1994. The question to be answered is whether a tendency in strength of ENSO exists. Global minimum of the quantity Eq. (6) has been found at 9 months. This is in a good agreement with the fact that SOI has intensive fluctuations on a 2–6-year scale, which can be reproduced by only a small bandwidth. However, a secondary minimum of Eq. (6) exists at a 36-year value. This bandwidth is related to long-term changes.

Fig. 2 shows that there is a strong decreasing tendency of SOI from the middle of the seventies, which is well reproduced by the curve estimated by WLR. To sum up, the bandwidth to be used depends on the underlying problem, and the selection between global and local minima of Eq. (6) cannot be automatically performed.

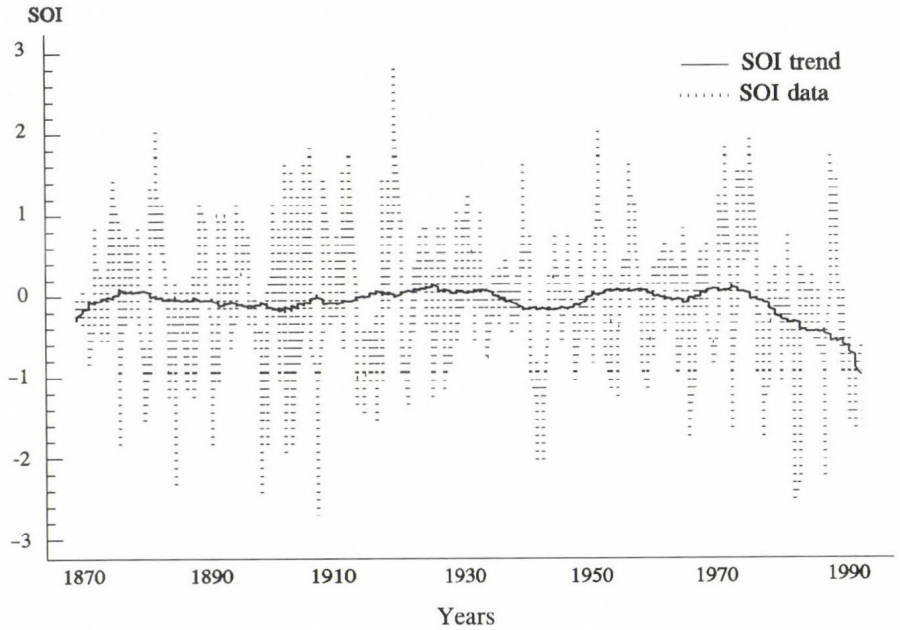


Fig. 2. Trend of SOI estimated with 36-year global bandwidth.

In the first example there was no considerable difference between curves obtained by WLR with global and local bandwidths. The following case illustrates the usefulness of local bandwidths. Central England monthly mean temperature data are available since 1659 (*Manley*, 1974). Trend of these annual mean temperatures has been estimated by WLR using global bandwidth and local bandwidth as suggested by *Staniswalis* (1989). Two important points can be drawn from *Fig. 3*. The first one is that the two curves corresponding to global and local bandwidths are very similar, and both have intensive fluctuations until the XIX century. This is probably due to the low reliability of data originating from this time. Just enough to mention the term "Central England" because location of measurements changed several times during the period of observations. In spite of deep work on data, this period cannot be considered homogeneous (*Malcher* and *Schönwiese*, 1987). The second conclusion is that the trend function estimated with global bandwidth remains highly fluctuating during second part of the whole period, while the curve obtained with local bandwidth is quite smooth. This is because first part of data produces the global bandwidth to be small (6 years), but the local bandwidth adapts quickly to homogeneous part of data and exhibits a smoothness required for "nice" trend functions.

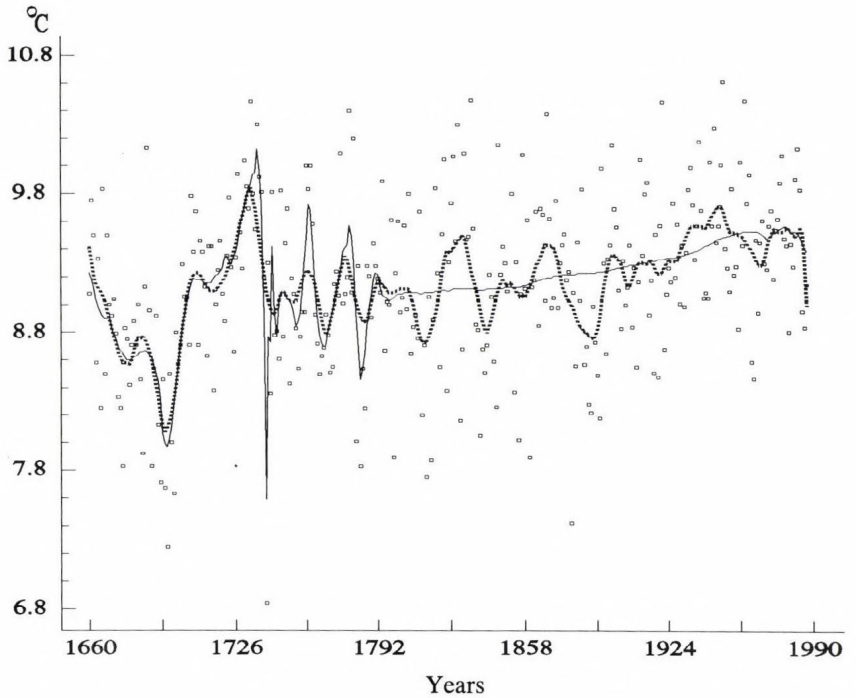


Fig. 3. Trend of Central England annual mean temperatures estimated with local bandwidth (solid line) and global bandwidth (dotted line).

5. Discussion

The paper summarized the most important mathematical details of non-parametric trend estimation, and three examples were shown to illustrate the importance of theoretical questions.

A reviewer suggested to use the expression *signal* instead of *trend* arguing that the deterministic part of climate data sets generally contain a long-term change and a short term interannual variability. This is certainly true, but every deterministic member of any data set may be inserted into a common component, called trend. Following the terminology of statistical literature, we decided to use the term *trend*, because the paper is principally a theoretical one. However, distinction between long-term change and interannual variability is really an important issue as illustrated by our second example. As the trend is a superposition of components with different time scales, the quantity Eq. (6) has various local minima, each of them corresponds to specific components of the trend function.

An important message of the third example is the usefulness of local bandwidths. Fig. 3 shows that local bandwidth can deliver a much more realistic estimate of the trend function when data set has inhomogeneities. We mention that a modified and further developed version of non-parametric trend estimation is able to detect data inhomogeneities (*Muller*, 1992). The principle of the technique is that two estimates are defined at each time t such that the first one uses data preceding t, while the second one uses only succeeding data. After some normalization, the maximum of the difference between these two estimates can be applied as a test statistic to detect change point in a time series. This test shows strong inhomogeneities in Central England temperature data set until the end of XVIII century.

References

Cleveland, W.S., 1979: Robust locally weighted regression and smoothing scatterplots. *J. Am. Statist. Ass. 74*, 829-836.

Eubank, E., 1988: *Spline Smoothing and Nonparametric Regression*. Marcel Dekker, New York.

Fan, J., 1992: Design-adaptive nonparametric regression. *J. Am. Statist. Ass. 87*, 998-1004.

Fan, J., 1993: Local linear regression smoothers and their ninimax efficiency. *Ann. Statist. 21*, 196-216.

Fan, J. and *Gijbels, I.*, 1995: Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation. *J. R. Statist. Soc. B. 57*, 371-394.

Gasser, T. and *Muller, H.G.*, 1984: Estimating regression functions and their derivatives by the kernel method. *Scand. J. Stat. 12*, 210-229.

Gasser, T., Muller, H.G. and *Mammitzch, V.*, 1985: Kernels for nonparametric curve estimation. *J. R. Statist. Soc. B. 47*, 238-252.

Jones, P.D., Wigley, T.M.L. and *Briffa, K.R.*, 1994: Global and hemispheric anomalies-land and marine instrumental records. In *Trends '93: A Compendium of Data on Global Change*. ORNL/CDIAC-65 (eds.: *T.A. Boden, D.P. Kaiser* and *F.W. Stoss*). Carbon Dioxid Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn., U.S.A.

Malcher, J. and *Schönwiese, Ch.-D.*, 1987: Homogeneity, spatial correlation and spectral variance analysis of long European and North American air temperature records. *Theor. Appl. Climatol. 38*, 157-166.

Manley, G., 1974: Central England temperatures: monthly means 1659 to 1973. *Quart. J. Roy. Meteorol. Soc. 100*, 389-405.

Muller, H.G., 1984: Smooth optimum kernel estimates of densities, regression curves and modes. *The Ann. Statist. 12*, 766-774.

Muller, H.G., 1992: Change-points in Nonparametric Regression Analysis. *The Annals of Statistics 20*, 737-761.

Nadaraya, E.A., 1964: On estimating regression. *Theory Probab. Applic. 15*, 134-137.

Simonoff, J.S., 1996: *Smoothing Methods in Statistics*. Springer-Verlag, New York.

Staniswalis, J.G., 1989: Local bandwidth selection for kernel estimators. *J. Am. Statist. Ass. 84*, 284-288.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 3, July–September 1998, pp. 159-165

Two statistical methods of long-time prognoses

László István Orbán

H-8200 Veszprém, Török Ignác u. 12/a, Hungary

(Manuscript received 21 January 1998; in final form 18 May 1998)

Abstract—In this paper the expected mean values of meteorological elements at Cluj-Napoca (Kolozsvár; Romania) are estimated by two statistical methods. The first method is based on the application of the harmonic analysis. The study is performed for 90-year time series of mean values of temperature and precipitation. The second method is correlation analysis carried out between mean temperatures and precipitation amounts of given month and the preceding months for the same 90-year data set. The methods gave acceptable prognoses for the monthly mean temperatures, but have not been proved to be applicable for precipitation forecasting yet.

Key-words: statistical forecast, harmonics, correlation analysis, long data series.

1. Introduction

Long-range forecasting frequently applies statistical methods because of the complexity of the hydrodynamical differential equations and the insufficient density of the meteorological observing network, mostly over the oceans. In this study, first of all, the estimation of the monthly mean values of temperature at Cluj-Napoca are attempted by two statistical methods. The first approach is the application of the harmonic analysis in which long time series (annual, monthly values) of meteorological elements are decomposed into harmonic components. The dominant harmonics with the greatest amplitudes are selected, and the phase angles of those are defined from the registered data of the previous period. The prognosis is based on the extrapolation of the obtained phase angles. The dominant harmonics are probability representatives of some physical process. (E.g. the harmonic with a period of 11.2 years expresses the period of sun activity.)

An application of this method to forecast the annual precipitation amount at Cluj-Napoca has already been described in journal Időjárás (*Orbán* and *Pap*, 1968). The prognosis used precipitation data of the three preceding years by

defining the phase angles of the three harmonics. That time, due to the lack of computer, a nomogram technique was presented to give an approximate solution of the trigonometric equation.

The second approach is the correlation analysis. In this case the forecast is made by using the correlation coefficients between the values of meteorological characteristics for a given term (e.g. year, season, month) and those for preceding terms. The regression coefficients are determined with the method of minimum squares.

For developing the methods, the computer assessment was provided by the Agronomic Institute of Cluj-Napoca.

2.1 Harmonic analysis

In 1992 we analysed the monthly values of the temperature and the precipitation for the previous 90 years (1080 months) on the basis of the data gained at Cluj-Napoca with harmonic analysis. *Fig. 1* shows frequencies and periods of amplitudes of the dominant harmonics (*Bozac,* 1993). (The values marked with asterisks represent the drawn values of the near harmonics.) It is worth observing that the harmonics corresponding to the sun activity (n = 8; T = 135 months) have an insignificant amplitude in the analysis of the monthly values (amplitude of the precipitation is not even indicated for its small value).

First, it was attempted to find a computer assisted method for the forecasting of the monthly mean temperature on the basis of the harmonics with the biggest amplitude (these are periods of 12; 1.1; 1.08; 1.005 months on Fig. 1). Then the mean temperature of the respective month can be approached with the equation:

$$y(t) = \sum_{k=1}^{4} x_k \sin(\theta t n_k + i_k) + t_m. \tag{1}$$

Here (t) denotes the respective month, x_k, n_k and i_k are the amplitude, the frequency and the phase angle of the corresponding harmonic respectively, while t_m is the annual mean temperature (t_m = 8.4°C for Cluj-Napoca). The total period (T = 1080) of the taken interval determines the θ value:

$$\theta = 2\pi/T \approx 0.005817764 \text{ rad/month}.$$

Considering the mean temperature of four consecutive months $y(0)$, $y(1)$, $y(2)$ and $y(3)$, by applying Eq. (1), we get the following equation system:

$$y(0) = \sum_{k=1}^{4} x_k \sin(i_k) + t_m, \tag{2}$$

$$y(1) = \sum_{k=1}^{4} x_k \sin(\theta n_k + i_k) + t_m, \tag{3}$$

$$y(2) = \sum_{k=1}^{4} x_k \sin(2\theta n_k + i_k) + t_m, \tag{4}$$

$$y(3) = \sum_{k=1}^{4} x_k \sin(3\theta\, n_k + i_k) + t_m. \tag{5}$$

This equation system contains twelve unknown factors. The i_k phase angle takes its values from the interval of $[0, 2\pi]$, while the x_k amplitudes and n_k frequencies differ only a little bit from the basic values of the harmonic analysis. We can find an exact or approximate solution of Eqs. (2) to (5). The expected mean temperature of the next month is given by

$$y(4) = \sum_{k=1}^{4} x_k \sin(4\theta n_k + i_k) + t_m. \tag{6}$$

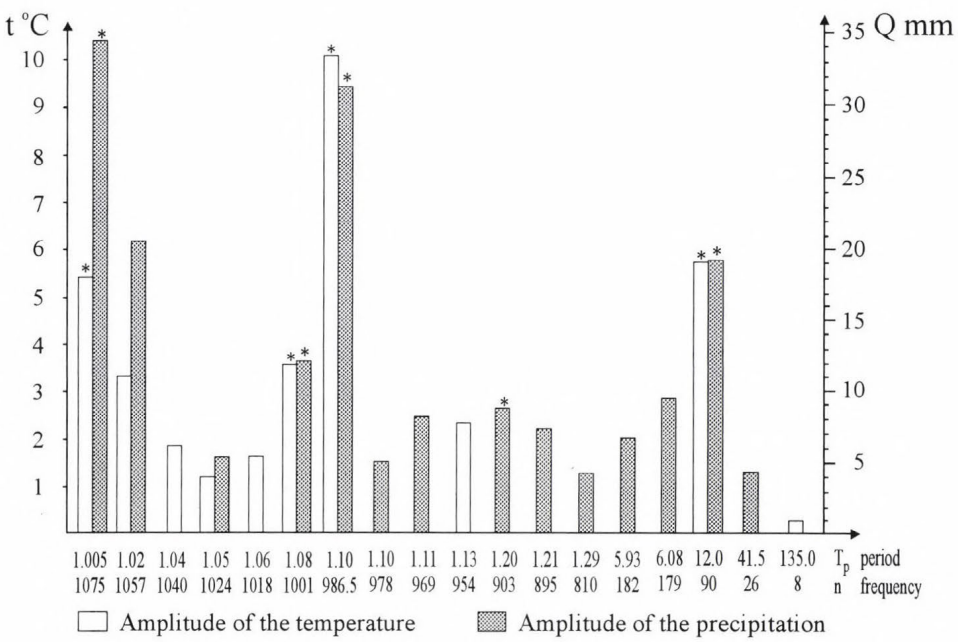


Fig. 1. The amplitudes of the dominant harmonics.

In the computer procedure we had to introduce the following expressions:

$$A_0 = \sum_{k=1}^{4} x_k \sin(i_k) + t_m - y(0), \tag{7}$$

$$A_1 = \sum_{k=1}^{4} x_k \sin(\theta n_k + i_k) + t_m - y(1), \tag{8}$$

$$A_2 = \sum_{k=1}^{4} x_k \sin(2\theta n_k + i_k) + t_m - y(2), \tag{9}$$

$$A_3 = \sum_{k=1}^{4} x_k \sin(3\theta n_k + i_k) + t_m - y(3), \tag{10}$$

$$E = |A_0| + |A_1| + |A_2| + |A_3|. \tag{11}$$

The computer calculates the possible values step by step using the developed software, and we have the best approximate solution when the value of E is minimal. When $E = 0$, the obtained factors satisfy the equation system Eqs. (2) to (5), because the sum of the absolute values can only be zero if each element equals to zero ($A_0 = A_1 = A_2 = A_3 = 0$), and then the computational equation system (Eqs. (7) to (10)) is equivalent to Eqs. (2) to (5).

However, it should be mentioned that the extent of the steps greatly influences the number of trials. If m represents the trial number for the unknown factor, then m^{12} calculations must be done because a given value has to be combined with each other value. Therefore, suitable precision by less calculation can only be achieved by arranging the operations into cycles. In cadre of a certain cycle the m value must be determined in accordance with the speed of the computer. The values obtained in the preceding cycle will provide the basic values for the succeeding cycle. In the following cycle narrower intervals and smaller steps are to be used, retaining the m value for the steps.

On the basis of this conception we made a prognosis for each month of the year of 1903 by using the mean temperatures of the preceding four months in Cluj-Napoca. We chose the year 1903 because the December of 1902 had been extremely cold (–7.1 °C mean temperature) and wanted to see the effect of such extreme values in the application of our method. *Table 1* shows the differences between the realised and the calculated monthly mean temperatures. It is noteworthy that the differences do not exceed 1 °C in the case of eight months, but there is such a month (April) when the deviation reaches 7 °C.

This example illustrates the efficiency of the above procedure well. Improvement of the procedure can be achieved in different ways. Using a faster

computer results bigger m steps and thus a more precise solution. Increasing the number of the selected harmonics and months would probably lead to better results but the number of operations would increase too. The accuracy of the prognosis may be increased by taking into account the correlation coefficients between the respective months.

Table 1. Computer prognosis for monthly mean temperatures of 1903, °C

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Predicted	-3.5	-0.7	6.0	14.5	15.1	17.0	16.5	15.7	15.4	9.2	4.4	1.0
Realised	-4.8	-0.6	6.0	7.5	14.1	17.3	19.2	17.8	14.8	9.8	4.6	0.3
Difference	-1.3	0.1	0.0	-7.0	-1.0	0.3	2.7	2.1	-0.6	0.6	0.2	-0.7

In the case of the temperature and other continual elements, not only the mean value but also the scattering of the values may be of interest. This characteristic can be determined by carrying out the harmonic analysis of the dispersion of the given element, and then we can adopt the above method for dispersion prognosis.

The application of the above method for the forecasting of the monthly precipitation amount is under way. The tests have shown that at least five harmonics should be selected. Acceptable prognosis of the precipitation can only be obtained if we select five months of the preceding period with two-month turns. For example, when predicting the precipitation in December we have to put in the software the precipitation amounts of the previous October, August, June, April and February. The possible correlation between the months can be taken into account by substituting for the E value of Eq. (11) with the relationship:

$$E = c_1|A_1| + c_2|A_2| + c_3|A_3| + c_4|A_4| + c_5|A_5|. \tag{12}$$

Here the coefficients c_1, c_2, c_3, c_4 and c_5 are proportional to the suitable correlation coefficients.

To characterise the distribution of the monthly precipitation we have to forecast the number of days with precipitation in a month and then to apply the above method for this quantity.

2.2. *Correlation analysis*

The second statistical forecasting method is based on the correlation analysis. On the computer of the Agronomic Institute of Cluj-Napoca we calculated the

correlation coefficients between the given month and the preceding months of the previous year for both the monthly mean temperature and the monthly precipitation. The calculations were made with the data gained at Cluj-Napoca during the years of 1901 and 1991. *Table 2* contains the correlation coefficients between the monthly temperatures. Since much lower correlation coefficients occurred in the study of the precipitation, they are not included in our recent paper.

Table 2. The correlation coefficients of the monthly mean temperatures based on the data gained in (Cluj-Napoca)

	The preceding year											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
												*
J	-0.0 ***	-0.03	0.15	0.01	-0.19	0.03	0.04	-0.15	0.07	0.06	-0.09	0.23
F	0.35	-0.05 **	0.08	0.07	-0.01	0.0	0.08	0.04	0.17	-0.05	0.03	0.12
M	0.13	0.28	0.06	-0.12	-0.03	-0.19	-0.05	-0.15 *	-0.09	0.07	-0.11	0.20
A	0.15	0.19	0.15	0.06	0.12	-0.09	0.09 **	0.21	0.07	-0.10	0.17	0.07
M	-0.02 *	0.07	-0.03	0.11	0.07 *	0.02	0.27	0.04	0.01	-0.15	0.03 *	-0.08 *
J	-0.21	0.02	0.12	0.14	0.26	0.03 **	0.13 **	0.19 ***	-0.06 *	-0.12	0.24	0.22
J	-0.07	-0.01	-0.04	0.20 *	0.17 *	0.29	0.32 ***	0.43 *	0.21	0.07	0.11	0.12
A	-0.03 **	-0.04 *	-0.07	0.25	0.22	0.18 *	0.45 *	0.24 **	0.18	-0.11	0.15	0.04
S	-0.30	-0.22	0.0 *	0.19	0.17	0.25	0.25	0.29	0.0	-0.03	-0.01	-0.17
O	-0.08	-0.13	-0.23	-0.11	-0.11	-0.11	-0.12	-0.05	0.11	0.01	0.09	0.03 *
N	-0.15	0.10	0.07	0.12	-0.01	-0.09	0.01	0.02	0.03	0.13	0.15	-0.23
D	-0.08	-0.10	0.09	-0.05	0.06	0.07	-0.03	-0.09	0.0	0.01	0.17	0.0
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
	The same year											

* Asterisk are above the marked values

The values marked with asterisks indicate significant correlation with the threshold of reliability 5% (*), 1% (**) and 0.1% (***).

Using these correlations we have defined the following forecasting formula:

$$y_i = e_i + \sum_{k=1}^{4} (a_{ik} x_k^2 + b_{ik} x_k). \tag{13}$$

Here y_i is the expected monthly mean temperature, while x_k indicates the registered mean temperatures of the preceding months for which the correlation coefficient are relatively high and significant. The e_i, a_{ik} and b_{ik} values were determined with the method of minimum squares on the basis of the data of the registered 1080 months. Four months proved to be satisfactory for an acceptable prognosis. The study of this method is going to be published by the Agronomic Institute of Cluj-Napoca.

We can state that this simple statistical forecasting method gives good prognosis for monthly mean temperatures, but is not applicable for monthly precipitation prediction. For the improvement of the method it is advisable to use data gained at different meteorological stations which are located at various distances around the cardinal points and to correlate one element to other elements.

It has to be admitted that the above two methods can only be carried out with the help of powerful computers, but the quantity of the calculations is still much fewer than that of the methods based on hydrodynamical equations.

References

Orbán, I.L. and *Pap, G.*, 1968: An attempt for forecasting annual precipitation amounts by using harmonic analysis (in Hungarian). *Időjárás 72*, 363-369.

Bozac, R., Orban, V., Parv, M. and *Pap, G.*, 1993: Determining the periodicity in weather trend by using harmonic analysis (in Rumanian). *Buletin USACN,* A-H, 47/2.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 3, July–September 1998, pp. 167–187

Risk and hazard assessment for accidental chlorine release using dispersion modeling

Diganta Bhusan Das, Sandeep Kumar Tetarwal and **Raj Sharma**[1]

Department of Chemical Engineering, Malaviya Regional Engineering College, Jaipur-302017, Rajasthan, India; E-mail: rajsh@recjai.ernet.in

(Manuscript received 22 August 1997; in final form 22 January 1998)

Abstract—Dispersion modeling for risk and hazard assessment for disaster management planning has been a rapidly advancing area. In the present study, a heavy gas model was used to simulate the dispersion of negatively buoyant and highly toxic chlorine to illustrate the use of dispersion modeling for risk and hazard studies in Matsya Industrial Area in the Alwar district of Rajasthan, India, which handles about 400 tons of chlorine every day. The consequences of a possible accidental release of 2 to 100 tons of chlorine were examined. Conservative estimates (e.g. in case of low wind velocity and Pasquill stability class "A") of ground level concentrations (GLC) as a function of distance and time (e.g. concentration of 1.39×10^{-11} g/m^3 at 600 m and after a time interval of 300 sec in case of 2 tons release) were obtained in the present work to be used for risk assessment studies. Hazardous zones were also identified by calculating the safe downwind distances and critical time intervals based on different threshold concentrations of chlorine (e.g. 3.5 km and 54.5 minutes for 2 tons release when critical concentration is STEL (1 ppm)). Effects of topography were incorporated by simulating for different ratios of friction velocity to wind velocity ranging from 0.2 to 0.9, and it was observed that for increasing ratios the ground level concentrations decreased significantly. Polar isopleths based on actual average wind velocities in the area were generated to examine the simultaneous effect of wind direction along with velocity on air quality. The probabilities of percentage of deaths and injuries in a given population were calculated by probit analysis; while area of lethal dosage was determined in order to estimate percentages of population affected in the area based on population density.

It was demonstrated that dispersion modeling is an effective tool for risk assessment studies based on which disaster management plans can be evolved. This would help in either minimising the number of accidents or the losses to human health and life in the event of a catastrophic accident in the area.

Key-words: risk and hazard assessment, chlorine, accidental release, Matsya Industrial Area in Alwar (India), toxic load, concentration profile, safe distance, dose-effect relationship, polar isopleth, probit analysis, area of lethal dosage.

[1] Corresponding author

1. Introduction

Even though the chemical plants are perhaps the safest of all manufacturing facilities (*Crowl* and *Louvar*, 1990), public concern about safety is a result of the potential of this industry to cause significant damage to human health and life. The methyl isocyanate (MIC) disaster of Bhopal (India) in 1984 is a case in point where thousands lost their lives and tens of thousands were affected. Accidents do and continue to occur in spite of the best safety measures, however with disaster management planning the magnitude of the effects of such accidents on human life can be minimised. To develop a disaster management plan, it is first necessary to identify the hazardous zone in case of an accidental release of toxic chemicals. Dispersion modeling for identifying these hazardous zones has been a rapidly advancing area and is also applied for emergency planning, regulatory decision making and risk assessment studies for industrial operations. If properly applied, dispersion models should also improve the quality of design and layout studies (*Bennet,* 1982; *Schreurs* and *Mewis*, 1987).

Kolluru (1991) and *Auger* (1995) have discussed the different components of risk and hazard assessment and the formation of disaster management plan for chemical process industries. This work focuses on the use of dispersion modeling as a route to disaster management planning in case of an accidental release of chlorine. As a case study, an attempt has been made to estimate or quantify the hazards owing to an accidental release of chlorine in the Matsya Industrial Area (M.I.A.), Alwar (India), where large (about 400 tons) quantities of chlorine are handled every day. Chlorine is very toxic and, being heavier than air, disperses at the ground level making it extremely hazardous to human life and health. This work should serve as an aid for evolving a disaster management plan in the event of an accidental chlorine release in M.I.A.

The Matsya Industrial Area in Alwar district of Rajasthan, India was selected as the area of study. The district of Alwar is situated in the North-East of Rajasthan between the 27.4° and 28.4° North latitudes and the 76.7° and 77.13° East longitudes at a distance of 140 km in South-West direction of Delhi and of 160 km in North-East direction of Jaipur. Sariska National Park and Wild Life Sanctuary is about 45 km from the industrial area. M.I.A. has a very large chlorine plant (capacity of 90 tons per day (TPD)) and some 35 smaller units manufacturing chlorinated paraffin wax (CPW) with a chlorine consumption of 5 to 10 TPD. On an average about 400 TPD of chlorine is handled in this area. The ambient concentrations are 6.62 $\mu g/m^3$ for SO_2 and 7.87 $\mu g/m^3$ for NO_x (*RPCB*, 1995) indicating a low level of air pollution in the area, however no information on chlorine in ambient air is available. The four seasons in the district are Summer (March–June), Monsoon (July–September), post-Monsoon (September–November) and Winter (December–February). The sky is mostly cloudy in Monsoon, whereas thin cloud cover may be found

sometimes in Summer and post-Monsoon seasons. The sky is generally clear (no clouds) in winter. The maximum temperature recorded so far is 48°C and the lowest value has touched the freezing point. The average temperature is 26°C, the average rainfall is 590 mm in the area. Maximum values of relative humidity occur during the Monsoon season (75%–90%), for the rest of the year the air is dry. For most of the year wind speed in Alwar is less than 2 m/s from North-West direction.

2. Background information

2.1 Available models

It was recognised as early as 1970 that attempting to describe dispersion of heavy gas by adopting the Gaussian model suitable for neutrally or positively buoyant clouds was inherently inadequate. Consequently many new models were proposed using the so-called "Top Hat" or "Slab" and "K-theory" or "Eddy diffusivity" approaches. "Top Hat" models assume that mass transfer occurs by entrainment across the density interface of a cloud with an assumed shape (e.g. cylindrical) and that internal mixing is fast enough for the concentration within the cloud to be uniform, whereas K-theory models numerically integrate suitably simplified equation of mass, momentum and energy conservation in two or three dimensional form (*Blackmore et al.,* 1982). One of the first "Top Hat" approaches to modeling was that of *van Ulden* (1974). Many "Slab" models were proposed soon after (*Germeles* and *Drake*, 1975; *Kaiser* and *Walker*, 1978; *Picknett*, 1981; *Fryer* and *Kaiser*, 1979 (DENZ); *Colenbrander*, 1980; *Cox and Carpenter*, 1980; *Eidsvik*, 1980; *Webber* and *Brighton*, 1984; *Witlox*, 1994a,b,c (HEGADAS-V)). Models that use the K-theory representation for heavy gas dispersion have also been discussed by many authors (e.g. *Schnatz* and *Flothmann*, 1980 (TRANSLOC); *Tauscher* (see: *Blackmore et al.*, 1982); *Su* and *Patnak* (see: *Blackmore et al.*, 1982)). *Blackmore et al.* (1982) reviewed the strengths and weaknesses, mechanistic features, the applicability to differing types and geometries of release, the ease of availability to users and the degree to which calculated results can be compared with field data. However very few work use Indian data. *Singh* (1990a) developed a heavy gas model for Indian meteorological conditions. The model has already been applied for oleum leakage (55 tons) at Sriram Fertilisers and Chemicals Ltd. in New Delhi, India, for Bhopal gas tragedy (MIC release of 40 tons) in 1984 and for the 2 tons chlorine spill at Chembur, Bombay, India, in 1985 with good results (*Singh*, 1990b). The model incorporates all the basic characteristics of heavy gas dispersion except that the local fluctuations of concentration are not predicted and dry deposition of pollutants as well as possible flashing of two-phase flow are not taken into

account. The most attractive feature of this model is the ease of computational effort while providing quick and reasonable estimates of hazards which can then be used for emergency preparedness. This model was therefore adopted in the present study.

2.2 Atmospheric stability

The concept of atmospheric stability is very important for the evaluation of assimilative and supporting capacity of the atmosphere and is used extensively in dispersion modeling. Pasquill stability criteria (*Pasquill*, 1974), as adopted by the Bureau of Indian Standards, were used for identifying the atmospheric stability class prevailing in M.I.A. Surface wind speed, intensity of solar radiation and night time sky cover are the prime factors defining the atmospheric stability classes.

2.3 Meteorology

Transportation and dilution of pollutants in the atmosphere depend strongly on the prevailing meteorological conditions (wind characteristics and strength, atmospheric stability), the topography of the location, release mechanisms, the height at which the pollutants are released and various reactions occurring in the atmosphere. As suggested by *Jindal* and *Agarwal* (1995) stability class "A" (the Pasquill stability class "A" is an extremely unstable situation when the surface wind speed is less than 2 m/s and there is strong solar radiation from the clear sky) was adopted in the present study. Wind roses were not available for M.I.A., therefore wind roses prepared by the Rajasthan State Pollution Control Board (RPCB), Jaipur, for the district of Alwar were used to represent the wind distribution in the area. As evident from the wind rose (*Fig. 1*), most of the year the wind blows from the South-East direction. During winter, which gives the worst case of stability in Alwar (*Jindal* and *Agarwal*, 1995), the maximum and minimum wind velocities as calculated from the wind roses were 2.91 m/s due West and 1.21 m/s due South. *Jindal* and *Agarwal* (1995) however suggested a wind velocity of 1.0 m/s for predicting the "worst scenarios" in the area. It was argued that if the wind velocity is high, even though the pollutants would disperse more and they would cover bigger area, people would be exposed to the toxic cloud for a shorter duration; whereas in case of low wind velocity, the cloud would move slowly and cover less area, but the time of exposure would be high. Therefore if the toxicity is a direct function of exposure time, a particular dose on the same area would be lethal when the wind velocity is low. Hence a wind velocity of 1.0 m/s as a worst case scenario was used in the simulation. In the present study chlorine release was assumed to take place because of catastrophic failure of a pressurised tank.

The height of release, which is also important for describing the dispersion phenomenon, was taken to be at ground level.

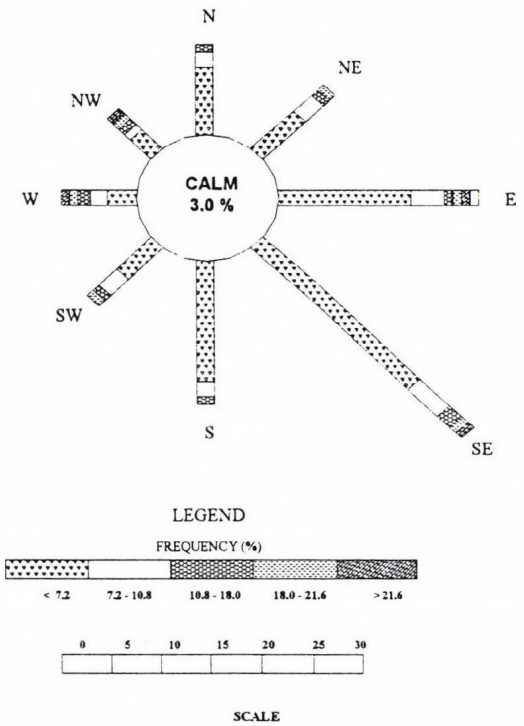


Fig. 1. Wind rose diagram of Alwar district in Winter (*RPCB,* 1995).

3. Basic model/equations used

3.1 Heavy gas model

Generally, the procedure for estimating the consequence of an accident has three components (*Mohan*, 1987):

(i) the mathematical description of the process which forms a cloud; this is necessary to provide the input parameters for the dispersion calculation for any particular event,

(ii) the mathematical description of the dispersion process which provides the concentration of a gas as a function of space and time, and

(iii) the description, in quantitative terms, of how a possible accident injures the population and damages the environment as a result of exposure to the concentration-time history derived from (ii).

In the present work the heavy gas model proposed by *Singh* (1990a) was adopted for estimating the consequences of a possible chlorine accident in M.I.A., Alwar. This model takes into account the first two components above by incorporating all the basic characteristics of heavy gas dispersion except that it does not predict the local mean fluctuations of concentration, ignores deposition and chemistry, and does not consider the possible two phase flashing flow of pollutants. The model was adopted because for all practical purposes it is possible to ignore variations in local concentrations and the presence of liquid gas in the toxic cloud. *Table 1* presents the model equations used in this study.

3.2 Simulator architecture

A computer simulator based on the equations in Table 1 was developed for predicting the concentration of chlorine as a function of distance and time under different meteorological conditions and of the released amounts. *Singh* (1990a) did not provide any value for the air entrainment co-efficient or the ratio of friction velocity to wind velocity. Therefore values suggested by *Eidsvik* (1980) and *Taylor* (1994) were used in the simulator. While simulating for risk estimation, the thermal effects of cloud heating due to temperature difference between the ground and air borne vapor were not considered because heating is important only when the cloud is considerably cold, e.g. in case of an accidental release of pressurised ammonia, therefore the equations describing the cloud heating have not been included in Table 1.

3.3 Probit analysis and area of lethal dosage

A large toxic release may give rise to the following effects on human life: (a) lethal injury (death), (b) non-lethal injury and (c) irritation depending on the breathing rate, functional capacity (e.g. oxygen uptake, muscle strength in case of paralysing agents), exposure to a given dose of the chemical and time of exposure or concentration-time profile (dosage), and most importantly individual susceptibility. In order to estimate the effects of the toxic release, it is therefore necessary to know the relationships between the concentration profile and the degree of injury. The degree of variation in dose-response can be presented in the form of a Gaussian response curve known in the terminology of toxicology and risk assessment as probit curve (*Taylor*, 1994). This dosage-response relation has been expressed by *Eisenberg* (see: *Crowl* and *Louvar*, 1990) as a probit equation. The probit equations for chlorine deaths and injuries are presented in *Table 2*.

Table 1. Model equations used in the simulation (*Singh*, 1990a)

Phenomena	Eq. No.	Parameter	Equation	Constant	Assumption	Remarks
Gravity slumping	1	Velocity at the edge of the cloud	$dr/dt = C\sqrt{\{(\rho_g - \rho_a)\, gh/\rho_g\}}$	C = 1.3* with entrainment and turbulence; C = 1.0** without entrainment and turbulence	The toxic cloud has cylindrical shape	C = 1.3 has been used in the present work
	2	Volume of puff	$V = \Pi r^2 h$		Radius and height are equal at time = 0 s	Volume and density are constant for no entrainment of air
Entrainment of air	3	Entrainment at the edges and top of the cloud	$dM_a/dt = \rho_a(\Pi r^2)U_e + 2\rho_a(\Pi rh)\alpha^* \times (dr/dt)$		Amount of air entrained initially is twenty times that of the mass of toxic material	
	4	Entrainment velocity	$U_e = \alpha' U_l R_i^{-1}$	$\alpha' = 0.5$		U_e is proportional to the difference between the local wind velocity and the velocity at the top surface, which also indicates presence of mechanical turbulence
	5	Richardson Number	$R_i = (gl_s/U_l^2)\Delta\rho/\rho_g$			It shows that entrainment velocity is dependent on the Richardson Number
	6	Entrainment at the top surface	$dM_a/dt = \rho_a\,(\Pi r^2)\,U_e$			Derived from Eq. 3. Entrainment at the edges is ignored because it is not important during gravity slumping
Conc. within the puff	7	Density of mixture at constant temperature	$\rho = (M_a+M_g)/((M_a/\rho_a) + (M_g/\rho_g))$			
	8	Instantaneous concentration	$C(x,y,z,t) = M_g G(x,y,z,t)/(\sqrt{2})\Pi^{3/2}\sigma_y^2\sigma_z$		Toxic mass has a Gaussian distribution	(x,y,z,t) gives the position of the puff in Cartesian coordinate at time *t*
	9	Function G	$G(x,y,z,t) = \exp[-\{y^2 + (x - x(t))^2\}]/2\sigma^2_y - z^2/2\sigma^2_z]$			
	10	Standard deviation in lateral direction	$\sigma_y = r/2.14$			
Post transition period	11	Standard deviation in vertical direction	$\sigma_z = h/2.14$			
	12	Radius of the cloud				
	13	Height of the cloud	$r = r_T + 2U_{fr}(t - t_T)$	$\alpha'' = 0.4$	$U_{fr} / U = 0.25$	
	14	Position of the puff centre	$h = h_T + \alpha'' U_{fr}(t - t_T)$ $x(t) = \int_0^t U(t')dt$			When the model is in heavy gas dispersion mode, the cloud point is based on its front velocity. When ambient air turbulence takes over, the cloud point is given by the ambient air

* Ref.: *Eidsvik*, 1980 ** Ref.: *Singh*, 1990a

The area covered by lethal concentration may be used to determine the percentages of population affected in a given area. The area of lethal dosage is a function of wind speed and stability condition and is calculated by the equation given in Table 2.

Table 2. Equations for probit analysis and area of lethal dosage (*Crowl* and *Louvar*, 1990)

Parameter	Eq. No.	Equation	Remarks
Probit equation for deaths	15	$Y = 1.69 \ln \Sigma C^{2.75} t - 17.1$	The relationship applies only to healthy adults and susceptible individuals such as infants, old people & people with advanced pulmonary/cardiovascular disease.
Probit equation for non-lethal injuries	16	$Y = 2.9 \ln C - 2.40$	Non-lethal injury is taken to mean hospitalisation with or without lasting impairment of health.
Area of lethal dosage	17	$A = K' U^{-n'}$	$K' = 1.72$ for Pasquill stability categories A-C. $n' = 1.06$

4. Results and discussion

Simulated results are sensitive to variables affecting atmospheric dispersion according to certain physical laws. As discussed earlier, in the present work a worst case scenario study has been conducted for M.I.A., Alwar, India, with implications of an accidental release of 2 to 100 tons of chlorine during the winter season.

4.1 Toxic load

Toxic load is essentially the quotient of the quantity of the released pollutant and the immediately dangerous to life and health (IDLH) value, and is a quantitative assessment of the potential risk associated with the use of that toxic chemical (*Singh*, 1990b). *Fig. 2* shows the toxic loads of chlorine for different accidental releases. Increasing values of toxic load indicate that the potential risk associated with a larger amount of release is higher. It necessitates the plants to use the best safety measures to lower the associated risks. An effective way of decreasing the risk would be to reduce the storing capacity of the plants.

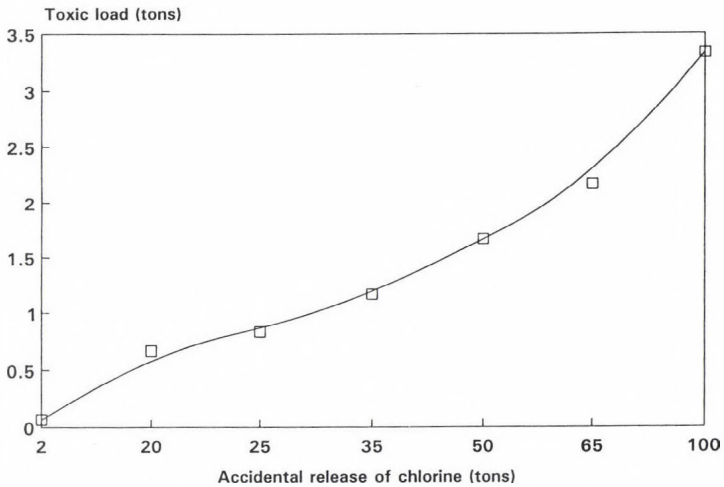


Fig. 2. Toxic load for different chlorine releases.

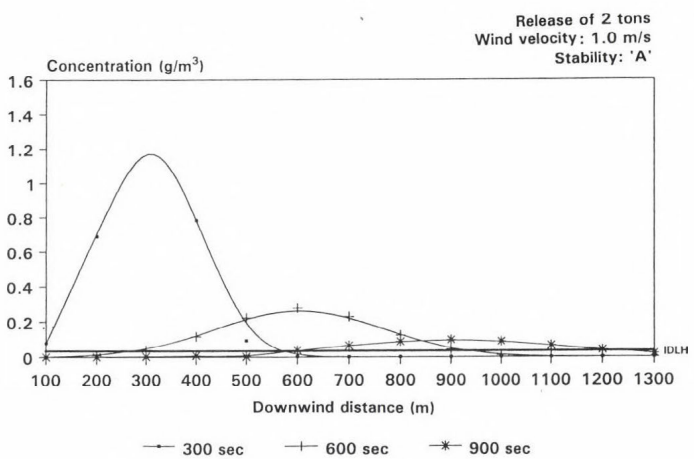


Fig. 3. Concentration profile at different time intervals.

4.2 Concentration profile

In general, at a given time the concentration of chlorine increases with the increase in downwind distance until reaching a maxima and then decreases. *Fig. 3* shows a typical concentration profile for a 2-ton release of chlorine at 300, 600 and 900 second time intervals as a function of downwind distance (100 m

to 1300 m). The puff center is the point of the maximum concentration. Hence the peaks of the curves which are the points of highest concentration, represent the points near the puff center. At points away from the puff center concentration is lower as shown in Fig. 3. The points of maximum concentration are 300 m at 300 seconds, 600 m at 600 seconds and 800 m at 900 seconds. The concentration levels are also seen to be above the IDLH value between 100 m and 550 m at time interval of 300 seconds, between 300 m and 900 m for the 600 second interval and between 600 m and 1200 m for the 900 second interval. If the wind dispersed the toxic cloud towards North-West direction, Alwar city and the industrial township would be the most affected area, whilst the effect would be the least in case of a South-Easterly wind direction. Variation of concentration at a particular downwind distance with time is a matter of importance. *Table 3* presents a typical trend of the variation of concentration with time at distance of 900 m from the point of release. Concentration is found below the lethal dose (IDLH) at 300 seconds and 600 seconds and above the IDLH value at 900 seconds.

Table 3. Variation of concentration with time at a fixed distance (ref. Fig. 6)

	Time interval, sec		
	300	600	900
Downwind distance, m	900	900	900
Concentration, g/m^3	1.39×10^{-11}	4.29×10^{-2}	9.59×10^{-2}

4.3 Safe downwind distance

For the purpose of hazard assessment, the main objective of dispersion calculation here is to determine the potential for damage at various points in the vicinity of an accidental release. At some distance from the release the concentration experienced would be below that likely to constitute a significant hazard; but within that range the proper specification and identification of hazards are a matter of great importance. Therefore safe downwind distances for different threshold concentrations (see Appendix for definitions) and different amounts of chlorine release were calculated. The simulated results are presented in *Table 4* and a typical trend is shown in *Fig. 4*. These results would find use while identifying the hazardous zones in case of an accidental chlorine release. For example the safe IDLH distance for 50 tons release is 4.9 km. The effect is estimated to last for a time period of 57 minutes.

Table 4. Safe downwind distances based on different threshold concentration of chlorine

Based on STEL concentration (1 ppm)							
Amount released (tons)	2	20	25	35	50	75	100
Safe distance (km)	3.5	7.9	8.6	11.1	12.9	14.6	-
Time interval (minute)	54.5	118.3	127.5	142.8	161.0	184.7	-
Based on NJATC (14 ppm)							
Amount released (tons)	2	20	25	35	50	75	100
Safe distance (km)	1.6	3.8	4.2	4.8	5.5	6.6	7.5
Time interval (minute)	22.8	50	53.8	60.5	68.3	78.3	86.5
Based on ERPG-3 concentration (20 ppm)							
Amount released (tons)	2	20	25	35	50	75	100
Safe distance (km)	1.41	3.5	3.8	4.4	5.1	6.1	7.0
Time interval (minute)	20.3	44.5	48.0	53.8	60.8	70	77.2
Based on LC-50 (10 min exposure) concentration							
Amount released (tons)	2	20	25	35	50	75	100
Safe distance (km)	0.7	1.8	2.0	2.4	2.9	3.6	4.1
Time interval (minute)	7.7	17.3	18.3	21.2	24.2	27.8	30.8
Based on IDLH concentration (30 ppm)							
Amount released (tons)	2	20	25	35	50	75	100
Safe distance (km)	1.32	3.3	3.6	4.2	4.9	5.8	6.5
Time interval (minute)	18.8	41.3	43	50	56.5	65	71.8

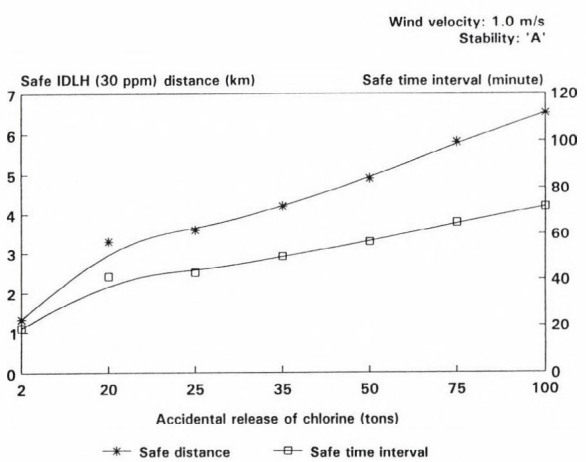


Fig. 4. Safe IDLH distance and time interval for different chlorine release.

4.4 Dose-effect relationships

To determine the chronic (long term) health effect of any specific chlorine dose on human beings in case of an accident so that immediate remedial action may be taken up, or to ascertain the extent of medical help that might be needed in the event of an emergency, it would be necessary to estimate the exposure of the population and to identify a safe range. *Table 5* presents the results simulated for a few "critical" concentrations. The estimated safe downwind distances and time intervals for different amounts of chlorine release when the concentration of interest is 3 ppm are presented in *Fig. 5*. A dose of 3 ppm for a 30-minute exposure time can be tolerated without any subjective feeling of malaise. If 2 tons of chlorine are released, a concentration of this intensity would be experienced up to a downwind distance of 2.5 km and for a period of 13.5 minutes; whereas, in case of 100 tons release, the critical distance and time interval would be 11 km and approximately 52 minutes respectively.

4.5 Effect of topography

Hills, buildings and other obstacles may cause additional turbulence. Typical effects of additional turbulence on concentration profiles for the releases of 20, 50, 65 and 100 tons of chlorine at a distance of 1.0 km and a time interval of 600 seconds are presented in *Fig. 6*. It is apparent that with the increase of the ratio of friction velocity (U_{fr}) to wind velocity (U), the concentration decreases significantly. This trend is observed because the additional turbulence produces a well-mixed plume. Since the concentration is a function of degree of mixing, a well-mixed plume gives lower concentration. Nevertheless except for 20 tons release at the U_{fr}/U ratio of 0.7 or greater, the pollutant levels would be high and dangerous.

4.6 Polar isopleths

Polar isopleths for examining the simultaneous effect of wind direction and velocity on air quality were prepared. Two types of polar isopleths are generally reported (*Bower* and *Sullivan*, 1981): wind isopleths and pollution isopleths, which are generalisations of wind and pollution roses, respectively. *Table 6* presents the average wind speeds for eight directions during the Winter season in Alwar district. *Fig. 7* shows the wind isopleth compiled from the wind data where average wind velocities are plotted against wind directions. Maximum and minimum wind speeds were found to be 2.91 m/s and 1.21 m/s in West and South directions, respectively. Pollution isopleths were prepared on the basis of actual wind velocities (Table 6) to give the dispersion of the pollutants. A pollution isopleth is presented in *Fig. 8*. The isopleth shows how chlorine would disperse with time to give a constant concentration of 90 ppm in case of 50 tons release.

Table 5. Dose-effect relationships for different chlorine release

Concentration: 3 ppm. Effect: can be tolerated without any feeling of malaise. Ex. time: 30 minute					
Amount released (ton)	2	20	35	50	100
Time interval (minute)	13.5	29.7	36	46.3	51.8
Downwind distance (km)	2.5	5.9	7.3	9.2	11
Concentration: 5 ppm. Effect: mild irritation to upper respiratory tract. Ex. time: 30 minute					
Amount released (ton)	2	20	35	50	100
Time interval (minute)	11.5	25.3	30.8	34.8	44.5
Downwind distance (km)	2.1	5.1	6.3	7.3	9.5
Concentration: 15 ppm. Effect: severe coughing, running of nose. Ex. time: 30 minute					
Amount released (ton)	2	20	35	50	100
Time interval (minute)	8	18.2	22.2	25	32
Downwind distance (km)	1.6	3.6	4.7	5.6	7.6
Concentration: 30 ppm. Effect: nausea, vomiting, over pressure feeling, shortness of breath, fits of coughing. Ex. time 30 minute					
Amount released (ton)	2	20	35	50	100
Time interval (minute)	6.5	14.8	18	20.5	26.3
Downwind distance (km)	1.3	3.53	4.1	4.7	6.4
Concentration: 40 ppm. Effect: toxic tracheobronchitis. Ex. time: 30 minute					
Amount released (ton)	2	20	35	50	100
Time interval (minute)	6	13.5	16.7	18.8	24.3
Downwind distance (km)	1.2	3.0	3.75	4.5	6.1

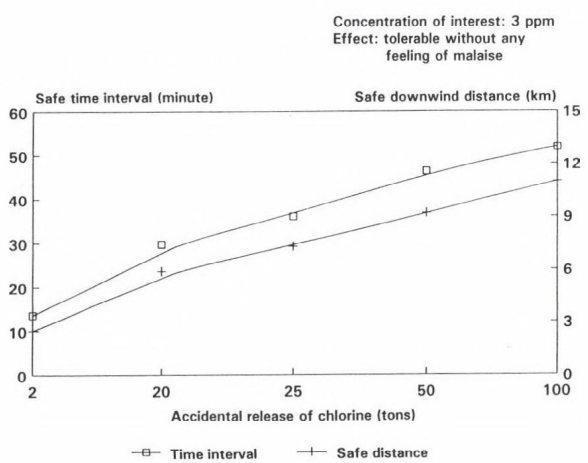


Fig. 5. Safe distance and time interval for different chlorine releases.

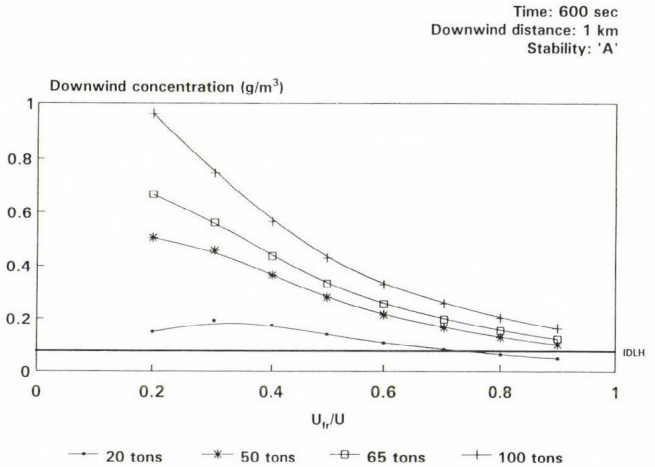


Fig. 6. Effect of topography.

Table 6. Wind directions and average velocities in winter at Alwar district of Rajasthan

Direction	North	N-East	East	S-East	South	S-West	N-West	West
Velocity, m/s	1.46	1.91	1.82	2.02	1.21	2.08	2.75	2.91

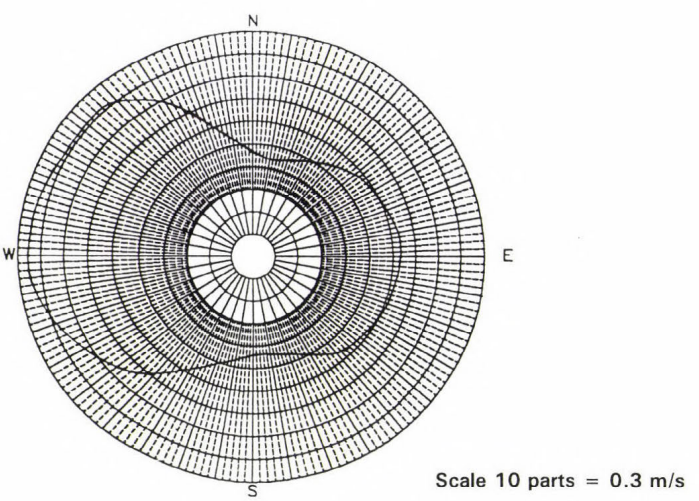


Fig. 7. Wind isopleth for Alwar district, India.

Variation of concentration around the source at a particular time interval of release is also a matter of importance. In order to illustrate this variation *Fig. 9* presents different concentrations of chlorine at 600 seconds for 50 tons release.

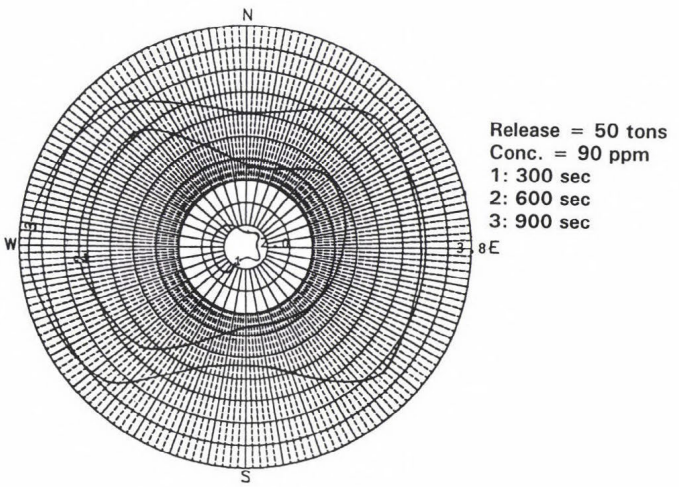


Fig. 8. Pollution isopleth for a concentration of 90 ppm at varying time interval in case of 50 tons release.

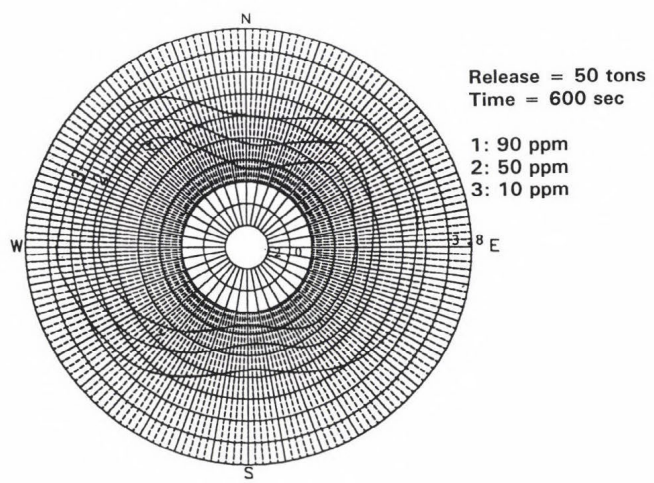


Fig. 9. Pollution isopleth for varying concentration at 600 seconds in case of 50 tons release.

4.7 Probit analysis

Probit analysis for lethality of chlorine has been applied to healthy adults including susceptible individuals such as infants, old people and people with advanced pulmonary/cardiovascular disease. The response of the subjects to a given dose is assumed to vary in a Gaussian manner (*Taylor*, 1994). These variations are presented by two cumulative distribution plots (probit curves), which indicate the probabilities of percentage of deaths (*Fig. 10*) or injuries (*Fig. 11*) of a given population for a specific dose or exposure time. Fig. 10 shows that the time of exposure causing 100% deaths for given concentration decreases with increasing concentrations and it is in the range of 4.3 hours for 30 ppm to 3.2 minutes for 150 ppm; similarly for given time intervals the concentrations for 100% death ranged from 85.0 ppm for 15 minutes to 51.0 ppm for 60 minutes. For non lethal injuries with or without lasting impairment of health, a concentration of 37.5 ppm will affect 100% of the population (*Fig. 11*).

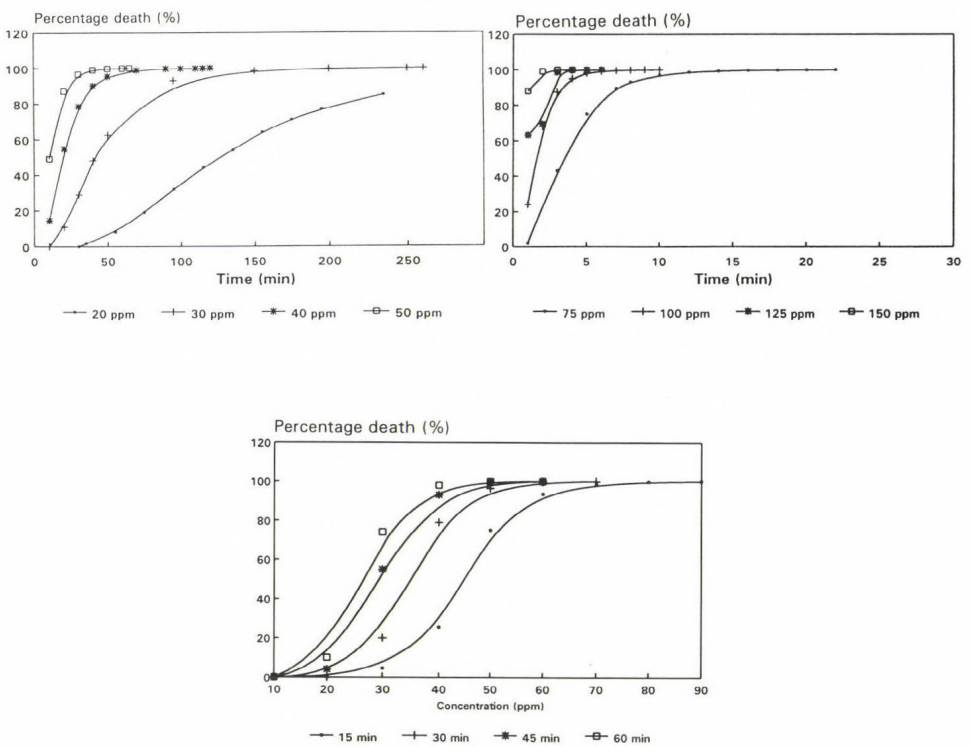


Fig. 10. Probit analysis for chlorine death.

4.8 Area of lethal dosage

Simulated results showed that the area of lethal dosage decreased with the increasing wind velocity (*Fig. 12*). This may be attributed to the fact that as the atmospheric conditions become more turbulent the chlorine disperses over greater area, hence the area of lethal dosage is smaller. For the wind velocity of 1.0 m/s used in the simulator the area of lethal dosage is 1.72 m^2. No information was available on the population density either in the vicinity of M.I.A. or in the industrial township. However, for a population density of 100 persons km^{-2}, the predicted number of people that would be most affected in case of an accident in the area is 172.

5. Summary

The paper has illustrated how a heavy gas model can be utilised for risk assessment in cases where combined effects of gravity slumping and air entrainment should be taken into consideration. The simple analytical form of the model allowed quick estimates of the consequences when the magnitudes of any or all variables like wind velocity, stability, source strength, downwind distance vary. Since the dispersion models assume the current and the future weather and release conditions to be representatives of past measurements, even though there can be a wide range of cases giving many probable outcomes, the worst meteorological conditions (e.g. stability class "A") were considered.

The source term conditions studied for predicting downwind concentrations and other measurements include instantaneous release of chlorine in the range of 2 to 100 tons. Concentration profiles as functions of time and downwind distance were obtained which can be used for various risk assessment studies. One of the estimates that can be obtained from this study is the safe downwind distance (e.g. the safe IDLH distance of 1.32 km in case of 2-ton release). The safe downwind distances can be used for identification of hazardous zone. Dose-effect relationships based on the threshold concentrations of chlorine, developed by various organisations, were established for proper specification of the extent of hazards involved. Topography or surface roughness which has a significant effect on the dispersion process, was considered in the present study for generating concentration profiles. Concentration was found to decrease with the increasing surface roughness at a particular distance. The polar isopleths, which combine the simultaneous effects of direction and velocity, would be very useful to ascertain the exact spreading of the pollutants.

The work presented here lays the foundation of a disaster management plan and demonstrates the use of dispersion modeling as an important tool for risk assessment studies. A future exercise would evolve a disaster management plan based on this work, which would demonstrate the utility of such studies in

preventing or minimising accident related damage. This would be useful either in preventing accidents in the M.I.A., Alwar, or for minimising the number of deaths in case of an accidental release of chlorine to the atmosphere.

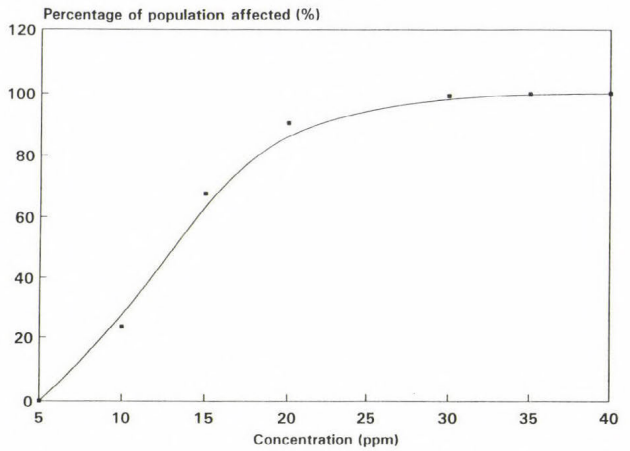


Fig. 11. Probit analysis for chlorine injury.

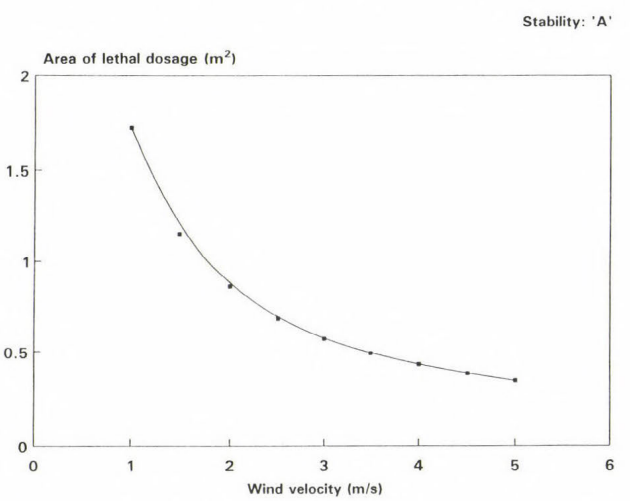


Fig. 12. Area of lethal dosage.

Acknowledgements—*Dr. A. Bhargava*, Sr. Environmental Engineer, Rajasthan State Pollution Control Board, Jaipur is acknowledged for his valuable support and help. Comments of *Dr. A. B. Gupta*, Professor, Civil Engineering Department, MREC, Jaipur during the preparation of the manuscript are also gratefully acknowledged.

References

Auger, J.E., 1995: Build Proper PSM Program from the Ground Up. *Chemical Engineering Progress*, 47-53, January.

Bennet, G.F., 1982: *Hazardous Materials Spill Hand Book*. Mc Graw Hill, New York.

Blackmore, D.R., Hermon, M.N. and *Wood Word, J.L.*, 1982: Heavy gas dispersion models. In *Dense Gas Dispersion* (ed.: *R.E. Britter* and *R.F. Griffiths)*. Chemical Engineering Monographs, *16*, Elsevier Scientific Publishing Company, New York.

Bower, J.S. and *Sullivan, E.J.*, 1981: Polar isopleth diagrams: A new way of presenting wind data and pollution data. *Atmospheric Environment 15*, 537-540.

Colenbrander, G.W., 1980: A mathematical model for the transient behaviour of dense vapour clouds. *3rd. Int. Symp. on Loss Prevention and Safety Promotion in the Process Industries*. Basel, Switzerland.

Cox, R.A. and *Carpenter, R.J.*, 1980: Further developments of a dense vapour cloud dispersion model for hazard analysis. In *Heavy Gas and Risk Assessment* (ed.: *S. Hartwig*). D. Reidel Publishing Co., Holland.

Crowl, D.A. and *Louvar, J.F.*, 1990: *Chemical Process Safety, Fundamentals with Applications*. Prentice Hall Int. Series, New Jersey.

Eidsvik, K.J., 1980: A model for heavy gas dispersion in the atmosphere. *Atmospheric Environment 14*, 769-777.

England, W.G., 1978: *Atmospheric Dispersion of Liquefied Natural Gas Vapour Cloud using SIGMET, a three dimensional time dependent Hydrodynamic computer model*. Heat Transfer and Fluid Mechanics Institute, Washington State University, Pullman, WA, June.

Fryer, L.S. and *Kaiser, G.D.*, 1979: *DENZ — A computer program for the calculation of dense toxic and explosive gases in the atmosphere*. UKAEA—SRD/SRO, R152.

Germeles, A.E. and *Drake, E.M.*, 1975: Gravity Spreading and Atmospheric Dispersion of LNG vapour clouds. *4th Int. Symp. on Transport of Hazardous Cargoes by Sea and Inland Water Ways*. Jacksonville, Florida.

Jindal, N. and *Agarwal, V.*, 1995: Estimation of hazards due to sudden and accidental release of chlorine in M.I.A., Alwar. *Senior Project Report*, Department of Chemical Engineering, M.R.E.C., Jaipur.

Kaiser, G.D. and *Walker, B.C.*, 1978: Release of anhydrous ammonia from pressurised containers — The importance of Denser-then-Air Mixtures. *Atmospheric Environment 12*, 2289-2300.

Kolluru, R.V., 1991: Understand the basics of risk assessment. *Chemical Engineering Progress*, 61-67, March.

Mohan, M., 1987: A physical basis of dense gas dispersion. *Workshop on Risk Assessment Techniques and Management*, IIT Delhi, Delhi, Oct. 26-30.

Pasquill, F., 1974: *Atmospheric Diffusion*. 2nd. Edition, Ellis Horwood, Chichester.

Picknett, R.G., 1981: Diffusion of dense gas puffs released in the atmosphere at ground level. *Atmospheric Environment 15*, 509-525.

RPCB (Rajasthan State Pollution Control Board), 1995: *Annual Report*, pp. 22 and pp. 24.

Schnatz, G. and *Flothman, D.*, 1980: A "K" model and its modification for the dispersion of heavy gases. In *Heavy Gas and Risk Assessment* (ed.: *S. Hartwig)*. D. Reidel Publishing Co., Dordrecht, Holland.

Schreurs, P. and *Mewis, J.*, 1987: Development of a transport phenomena model for accidental releases of heavy gases in an industrial environment. *Atmospheric Environment 21*, 765-776.

Singh, M.P., 1990a: Overview of atmospheric diffusion modelling. *National Seminar on Environmental Pollution Control and Management*, Jaipur, February 18-19.

Singh, M.P., 1990b: Vulnerability analysis for airborne release of extremely hazardous substances. *Atmospheric Environment 24a*, 769-781.

Taylor, J.R., 1994: *Risk Analysis for Process Plants, Pipe Lines and Transport*. E and F.N. BPON, 1st Edition.

van Ulden, A.P., 1974: On the spreading of a heavy gas near the ground. In *Proc. of 1st Int. Symp. on Loss Prevention and Safety Promotion in the Process Industries* (ed.: *C.H. Buschmann).* Elsevier Scientific Publishing Company, Amsterdam.

Webber, P.M. and *Brighton, P.W.M.*, 1984: A mathematical model of spreading, vaporising liquid pool, 3rd Battelle Inst./ Univ. of Wuppertal. *Symp. on Heavy Gas and Risk Assessment*, Bonn.

Witlox, H.W.M., 1994a: The HEGADAS model for ground-level heavy gas dispersion-I. Steady-sate model. *Atmospheric Environment 28,* 2917-2932.

Witlox, H.W.M., 1994b: The HEGADAS model for ground-level heavy gas dispersion-II. Time dependent model. *Atmospheric Environment 28*, 2933-2946.

Witlox, H.W.M., 1994c: Interfacing dispersion models in the HGSYSTEM hazard-assessment package. *Atmospheric Environment 28,* 2947-2962.

APPENDIX

Definition of critical concentration of chlorine (*Mallet*, 1993):

(a) STEL (1 ppm):	The occupational short term exposure limit. It is a 15-minute time weighted average concentration that should not be exceeded at any time during the work day.
(b) IDLH (30 ppm):	The Immediately Dangerous to Life and Health concentration. It represents the maximum concentration from which, in the event of respiratory failure, one could escape in 30 minutes without a respirator and without experiencing any escape impairing (e.g. severe eye irritation) or irreversible health effect.
(c) NJATC (14 ppm):	New Jersey Acute Toxicity Concentration — established by the state of New Jersey and used for developing its risk management program regulations.
(d) LC-50:	The concentration believed to be lethal to 50% of humans exposed for the stated period of exposure time.
(e) ERPG-3 (20 ppm):	Emergency Response Planning Guidelines-3. It is the maximum air borne concentration below which it is believed that nearly all individuals would be exposed for up to one hour without experiencing or developing any life threatening effect.

List of symbols

Symbol	*Definition*	*Used in*
A	Area of lethal dosage, m^2	Eq. 17
C	Downwind concentration, g/m^3	Eq. 8
c	Gravity slumping constant, dimensionless	Eq. 1, 2
dr/dt	Rate of change of cloud radius or, cloud front velocity	Eq. 1, 3
g	Acceleration due to gravity, m/s^2	Eq. 1
G	A function	Eq. 8, 9
h	Cloud height, m	Eq. 1, 3, 11
h_T	Height of the cloud at transition time, m	Eq. 13
K′	Constant	Eq. 17
l_s	Turbulence length scale, m	Eq. 5
M_g	Mass of toxic chemical released, ton	Eq. 7
M_a	Mass of air entrained, ton	Eq. 7
n′	Turbulence index, dimensionless	Eq. 17
r	Cloud radius, m	Eq. 2, 3, 6, 10
r_T	Radius of the cloud at the transition time, m	Eq. 12
R_i	Richardson Number	Eq. 4, 5
t	Time, sec	Eq. 12, 13
U	Wind velocity, m/s	Eq. 14
U_e	Entrainment velocity, m/s	Eq. 2, 4, 6
U_{fr}	Friction velocity, m/s	Eq. 12, 13
U_l	Turbulence velocity, m/s	Eq. 4
V	Volume of the cloud, m^3	Eq. 2
x	Downwind direction, m	Eq. 9
x(t)	Puff centre, m	Eq. 9, 14
y	Lateral direction, m	Eq. 9
Y	Probit	Eq. 15, 16
z	Vertical direction, m	Eq. 9
(x,y,z,t)	Position of the puff centre in Cartesian coordinates	Eq. 7

Greek Letters

α'	Co-efficient for top entrainment, dimensionless	Eq. 4
α''	Co-efficient for vertical entrainment, dimensionless	Eq. 13
α^*	Co-efficient for entrainment at the edges, dimensionless	Eq. 3
ρ_a	Density of air, kg/m^3	Eq. 1, 3, 6
ρ_g	Density of toxic cloud, kg/m^3	Eq. 1, 5
σ	Deviations, m	Eq. 8

Subscript

a	Air	Eq. 1, 3, 6, 7
g	Toxic gas	Eq. 1, 5, 7
T	Transition time, sec	Eq. 12, 13
y	Lateral direction, m	Eq. 7, 9
z	Vertical direction, m	Eq. 8, 9

BOOK REVIEWS

J.H. Seinfeld and *S.N. Pandis*: **Atmospheric Chemistry and Physics — From Air Pollution to Climate Change**. John Wiley and Sons, Inc., New York, Chichester, Weinheim, Brisbane, Singapore, Toronto, 1998. XXVII + 1326 pages, tremendous amount of figures, tables, references, appendices and problems to be solved.

This volume is the thickest book on the atmosphere, the reviewer has ever seen. It deals mostly with the chemistry of the atmosphere, with separate chapters on aerosol and cloud physics and air pollution meteorology. As the authors state in the preface: "The object of this book is to provide a rigorous, comprehensive treatment of the chemistry of the atmosphere, including the formation, growth, dynamics and properties of aerosols, the meteorology of air pollution, the transport, diffusion and removal of species in the atmosphere, the formation and chemistry of clouds, the interaction of atmospheric chemistry and climate, the radiative and climatic effects of gases and particles and the formulation of mathematical chemical/transport models of the atmosphere". This ambitious aim is obtained in 24 chapters. After reviewing the book one can say that, except some cases, the individual chapters are not too long and sophisticated. The length of the volume is due to the great number of chapters discussing the large subject mentioned.

If we took only Chapters 1–7, 13, and 19–24 we would receive a book on atmospheric chemistry as this branch of atmospheric science is generally imagined. The first chapter is an introduction into atmospheric characteristics, the second deals with atmospheric composition, cycles and lifetimes, the third one discusses the bases of photochemistry and reaction kinetics, while the aim of Chapters 4–7 is to give a summary of our present knowledge of stratospheric chemistry, tropospheric chemistry, atmospheric aqueous phase chemistry and aerosol properties, respectively. It is interesting that aerosols consisting of organic materials are not treated together with other particles, their properties are presented in a separate chapter (Chapter 13). It seems that the authors considered the organic particles so important that they devoted a separate chapter to their discussion (this is understandable since they are from California where organic substances were widely studied owing to their importance in photochemical smog). However, this discussion is rather separated from that of other aerosol particles. In Chapters 19–20 dry and wet depositions are presented, while in Chapters 21 and 22 our ideas on the interaction of atmospheric chemistry and climate as well as the radiative effects of atmospheric aerosols are summarized. Finally the authors' goal in the last two

chapters is to describe atmospheric chemical and statistical models. The last chapter on statistical models is of particular interest since this subject is rarely occurs in atmospheric chemistry books prepared by European writers.

Chapters 8–12 are rather unique in books discussing the atmospheric physics and chemistry. Chapter 8 are devoted to the dynamics of single aerosol particles. This means that transport processes of single particles, their interaction with the fluid (drag force) and fluid molecules (Brownian motion) as well as their motion caused by external forces (e.g. gravitational settling) are presented. The next chapter ("Thermodynamics of Aerosols") contains the principles of partitioning of atmospheric species between the vapor and particulate phases, while Chapter 10 is a rather detailed discussion of nucleation processes. Then, the mass transfer rates between condensed and gas phases are discussed (Chapter 11). This part of the book on "aerosol physics" is closed by a chapter on the dynamics of aerosol populations including such important subjects as changes in particle size distribution by condensation and evaporation and by coagulation.

Except Chapter 15 on cloud physics, the remaining chapters are composed of material what we generally call air pollution meteorology. Thus, the title of Chapter 14 is "Meteorology of Air Pollution", while Chapter 16 contains the basic principles of micrometeorology. Further, in Chapter 17 and 18 the authors present the bases of atmospheric diffusion and analytical solutions of atmospheric diffusion equations, respectively. In this last chapter they describe, among other things, the Gaussian plume equation which is discussed in a detailed way.

It goes without saying that one can raise questions concerning the structure of this book. It is not questionable, however, that the authors present an excellent material, in agreement with the subtitle of the volume, from air pollution to climate change. This means that, except atmospheric radioactivity, the reader finds practically everything in this book which is important to understand atmospheric environment of our planet. Each chapter is followed by references consisting of a lot of literature published mostly by American research workers. Further, the chapters are closed by problems of different levels "to enable the reader to evaluate his or her understanding of the material" (cited from the authors' preface). In the preface the authors also state that "The book is intended to serve as textbook for a course in atmospheric science that might vary in length from one quarter or semester to a full academic year". However, one fills that the material composing this book seems to be too large even for a course of one year. On the other hand, it is self-evident that parts of this volume can be use in different courses on environmental engineering and science as well as on meteorology and chemistry.

Thus, the reviewer concludes that the present book is a magnificent successor of the previous book of one of the authors (J.H. Seinfeld) entitled *Atmospheric Chemistry and Physics of Air Pollution* published by the same

editor in 1986. The former book is completed successfully by material emerged during the last decade. Consequently, the volume can be proposed to university professors and graduate students for their courses as well as to all persons educated in physics and chemistry for making an acquaintance with basic principles of atmospheric science.

E. Mészáros

Wilfried Schröder: **Noctilucent Clouds (Theoretical concepts and observational implications)**. Science Edition, D-28777 Bremen, 1998. 340 pages. Price: $ 20.

This book is a compilation of the works written by the author during the last decades. Since some of them have been published in German, it seems necessary to collect them into an English book, to be available for the readers who look only for English literature.

The articles can be grouped into three main topics:

— Climatology of noctilucent clouds (NLC);

— Mesosphere and NLC;

— History of observations made in Germany.

Noctilucent clouds have been observed since the eighties of the last century. They could be seen and photographed when the Sun is below the horizon by 5–15 degrees. Their average altitude is 83 km, that is they give information about such heights that could only be investigated by meteorological rockets. Their observation zone is the belt between 50 and 70 degrees latitude on both Hemispheres. The NLC are summer phenomena.

Their origin is not understood fully. Also, their relation to the wind speed and temperature is not known perfectly.

A small and enthusiastic group of scientists runs the observations and uses them to study the physics of the mesosphere. Those, who are interested in this topic, please, read this interesting book.

G. Major

NEWS

Congratulations to Professor Ernő Mészáros for the Széchenyi-prize

In March of 1998 the President of the Hungarian Republic awarded Széchenyi-prize (the highest prize for scientist in Hungary) to Professor E. Mészáros, former editor of Időjárás, in recognition of his excellent work in developing of the atmospheric chemistry as new area of atmospheric sciences in Hungary.

Ernő Mészáros began its scientific carrier in 1957 at the Hungarian Meteorological Service. That time his interest was in the cloud physics. Later on he investigated the physical and chemical properties of atmospheric aerosol particles. The prize recognized his activity in developing of the atmospheric chemistry as new area of atmospheric sciences in Hungary.

He has published several papers in high-impact scientific journals. He has always taken part in international scientific activities.

On the occasion of his 60th birthday, the Időjárás published his curriculum vitae (Vol. 100, No. 1–3, 1996).

G. Major

Symposium of the International Society for Photogrammetry and Remote Sensing (ISPRS) in Budapest

The Commission VII of ISPRS held its 1998 year Symposium at the Hungarian Academy of Sciences, 1–4 of September. The basic area of the Commission is resource and environmental monitoring, therefore the meteorologists are involved in the activity. The scientists of the Hungarian Meteorological Service presented a poster and computer demonstration of the activity of the Satellite Research Laboratory.

G. Major

ATMOSPHERIC ENVIRONMENT

an international journal

To promote the distribution of Atmospheric Environment *Időjárás* publishes regularly the contents of this important journal. For further information the interested reader is asked to contact *Prof. P. Brimblecombe*, School for Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, U.K.; E-mail: atmos_env@uea.ac.uk

Volume 32 Number 7 1998

The Transport of Air Pollutants Over Complex Terrain (TRACT) September, 1992

Volume 32 Number 8 1998

Atmospheric Sciences and Applications to Air Quality (ASAAQ)

New Directions

Volume 32 Number 9 1998

Volume 32 Number 10 1998

Volume 32 Number 11 1998

Volume 32 Number 12 1998

The Mediterranean Campaign of Photochemical Tracers — Transport and Chemical Evolution (MEDCAPHOT-TRACE), Athens, Greece 1994–1995

NOTES TO CONTRIBUTORS OF *IDŐJÁRÁS*

The purpose of the journal is to publish papers in any field of meteorology and atmosphere related scientific areas. These may be

- reports on new results of scientific investigations,
- critical review articles summarizing current state of art of a certain topic,
- shorter contributions dealing with a particular question.

Each issue contains "News" and "Book review" sections.

Authors may be of any nationality, but the official language of the journal is English. Papers will be reviewed by unidentified referees.

Manuscripts should be sent to
Editor-in-Chief of IDŐJÁRÁS
P.O. Box 39
H-1675 Budapest, Hungary

in three copies including all illustrations. One set of illustrations has to be of camera ready quality, the other two might be lower quality.

Title part of the paper should contain the concise title, the name(s) of the author(s), the affiliation(s) including postal and E-mail address(es). In case of multiple authors, the cover letter should indicate the corresponding author.

Abstract should follow the title, it contains the purpose, the data and methods as well as the basic conclusion.

Key-words are necessary to help to classify the topic.

The text has to be typed in double spacing with wide margins. Word-processor printing is preferred. The use of SI units are expected. The negative exponent is preferred to solidus. Figures and tables should be consecutively numbered and referred to in the text.

Mathematical formulas are expected to be as simple as possible and numbered in parentheses at the right margin. Non-Latin letters and hand-written symbols should be indicated and explained by making marginal notes in pencil.

Tables should be marked by Arabic numbers and printed in separate sheets together with their captions. Avoid too lengthy or complicated tables.

Figures should be drawn or printed in black and white, without legends, on separate sheets. The legends of figures should be printed as separate list. Good quality laser printings are preferred as master copies.

References: The text citation should contain the name(s) of the author(s) in Italic letter and the year of publication. In case of one author: *Miller* (1989), or if the name of the author cannot be fitted into the text: *(Miller,* 1989); in the case of two authors: *Gamov* and *Cleveland* (1973); if there are more than two authors: *Smith et al.* (1990). When referring to several papers published in the same year by the same author, the year of publication should be followed by letters a,b etc. At the end of the paper the list of references should be arranged alphabetically. For an article: the name(s) of author(s) in Italics, year, title of article, name of journal, volume number (the latter two in Italics) and pages. E.g. *Nathan, K.K.,* 1986: A note on the relationship between photosynthetically active radiation and cloud amount. *Időjárás 90*, 10-13. For a book: the name(s) of author(s), year, title of the book (all in Italics except the year), publisher and place of publication. E.g. *Junge, C. E.,* 1963: *Air Chemistry and Radioactivity*. Academic Press, New York and London.

The final version should be submitted on diskette altogether with one hard copy. Use standard 3.5" or 5.25" DOS formatted diskettes. The preferred word-processors are WordPerfect 5.1 and MS Word 6.0.

Reprints: authors receive 30 reprints free of charge. Additional reprints may be ordered at the authors' expense when sending back the proofs to the Editorial Office.

More information: gmajor@met.hu
Information on the last issues:
http://www.met.hu/firat/ido-e.html

Published by the Hungarian Meteorological Service

Budapest, Hungary

INDEX: 26 361

HU ISSN 0324-6329

IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

CONTENTS

****

http://www.met.hu/firat/ido-e.html

VOL. 102 * NO. 4 * OCTOBER–DECEMBER 1998

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: P.O. Box 39, H-1675 Budapest, Hungary or
Gilice tér 39, H-1181 Budapest, Hungary
E-mail: gmajor@met.hu or *antal@met.hu*
Fax: (36-1) 290-7387

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary;
E-mail: gmajor@met.hu or *antal@met.hu*; *Fax: (36-1) 290-7387*

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 4, October–December 1998, pp. 201–218

Statistical estimate of the vertical ozone structure

D. Spänkuch and **E. Schulz**

Meteorological Observatory Potsdam, German Weather Service
Postfach 60 05 52, D-14405 Potsdam, Germany
E-mail: spaenkuch@mop.dwd.d400.de

(Manuscript received 1 June 1998; in final form 28 September 1998)

Abstract—A multiple linear regression approach is presented to estimate layer ozone amount $\Delta\Omega$ from predictors accessible from model outputs of meteorological analysis and prediction additionally to the total ozone amount that has regularly been monitored from satellites for several years. Total column ozone is the most important predictor, in general, of the approach developed and checked for northern midlatidudes. The second most important predictor is the thermal state of atmospheric layers. The accuracy of the statistical estimate is assessed to be about 3 to 10 per cent in maximum depending on atmospheric layer and season. Comparison of the estimate with more than one hundred ozone sonde data at Hohenpeissenberg, Germany, throughout one year showed promising results between 300 and 50 hPa when outliers of about 8 to 10 per cent of cases are omitted. The outliers were mainly found for the layer between 200 to 100 hPa and are obviously related to filament-like structures in ozone there. The most satisfying results were obtained for the tropopause region 300–200 hPa with a reduction of variance of about 0.7 to 0.8. This altitude range is the most difficult one for monitoring from satellites.

Key-words: stratospheric ozone, forecasting, multiple linear regression.

1. Introduction

Atmospheric ozone has been a permanent issue for many years because of its harmful effects to the environment. Understanding the atmospheric ozone cycle and its change needs inter alia a detailed knowledge of the ozone distribution in space and time. Our current knowledge on the atmospheric ozone distribution, its variability and long-term variations and trends relies on data sources of different characteristics and quality (e.g., *WMO*, 1995). Satellites have monitored globally total column ozone Ω and the vertical ozone distribution of the upper and middle stratosphere since the late seventies. Information on the ozone profile from satellites are gained from backscattered radiation in the UV

between 252 and 306 nm by the Solar Backscatter Ultraviolet Spectrometer (SBUV) and solar occultation absorption measurements around 600 nm by the Stratospheric Aerosol and Gas Experiment (SAGE I and II). Limb emission measurements at 9.6 μm are anticipated with the Michelson Interferometric Passive Atmosphere Sounder (MIPAS) on ENVISAT. Occultation and limb measurements exhibit in general high vertical resolution down to the tropopause. The vertical resolution of SBUV data is about 8 km above 25 km altitude and increases to 15 km below that level. Intercomparisons between SAGE II and SBUV results made in the interval 1984 to 1990 agreed to better than 5% in general (*WMO*, 1995).

The first information on the vertical ozone structure, although of limited resolution, was drawn from Umkehr observations (*Götz*, 1931). Ozonesondes having high vertical resolution up to the bursting level of balloons at about 30 km altitude have been launched since the midsixties. Ozone lidars of high vertical resolution per se have been in operation since the late eighties. Stratospheric ozone research needs lidars of sufficient power, however, particularly for the upper stratosphere. All ground-based sounding systems are regionally representative. However, ground-based quasicontinuous monitoring is not feasible because of meteorological (cloudiness for lidars) as well as cost-effective reasons (both profiling systems).

Summarizing the present data situation we state a global data set of sufficient quality since 1979 for the upper stratosphere which will be updated by corresponding satellite launches in the future. Below about 20 km reliable information of sufficient quality can be presently gained only by satellite measurements with tangential observation geometry (SAGE at present) and by ozonesondes and lidars. For the tropopause region which is of particular importance to the role of ozone depletion in radiative forcing of the Earth's climate system (e.g., *Lacis et al.*, 1990; *Ramaswamy et al.*, 1992) and below, the interpretation of occultation and limb measurements is limited in the presence of clouds. Information on ozone variations at these heights are mainly from ozonesondes and lidars. Thus there is a deficit of ozone data in the troposphere and lowermost stratosphere particularly because ozonesonde and lidar measurements are only sporadically available and only regionally representative. Substantial improvement by new satellite systems is not to be expected in the near future.

However, there are some founded reasons for improvement of the data basis situation by using information of other meteorological parameters. Statistical approaches have been used with success to predict total column ozone (e.g., *Burrows et al.*, 1995; *Spänkuch* and *Schulz*, 1995, 1997a; *Long et al.*, 1996) and to check the quality of total column ozone measurements (*Spänkuch* and *Schulz*, 1997b). They are also capable for ascertaining the scope of the climatological vertical ozone structure. The physical background for significant statistical relations between the vertical ozone structure and other meteorological

parameters is given mainly by the dominance of dynamical processes up to the middle stratosphere in building up the vertical ozone structure.The approach of statistical profiling has widely been used in numerous numerical weather prediction schemes in particular to infer the vertical humidity structure. In ozone research, analogous efforts were made to evaluate the quality of Umkehr observations (*de Luisi* and *Mateer*, 1971) and to improve retrievals of satellite measurements (e.g., *Sellers* and *Yarger*, 1969; *Timofeyev et al.*, 1974; *Feister* and *Spänkuch*, 1977), but these approaches have obviously not been applied despite of partly promising results in particular by the last authors.

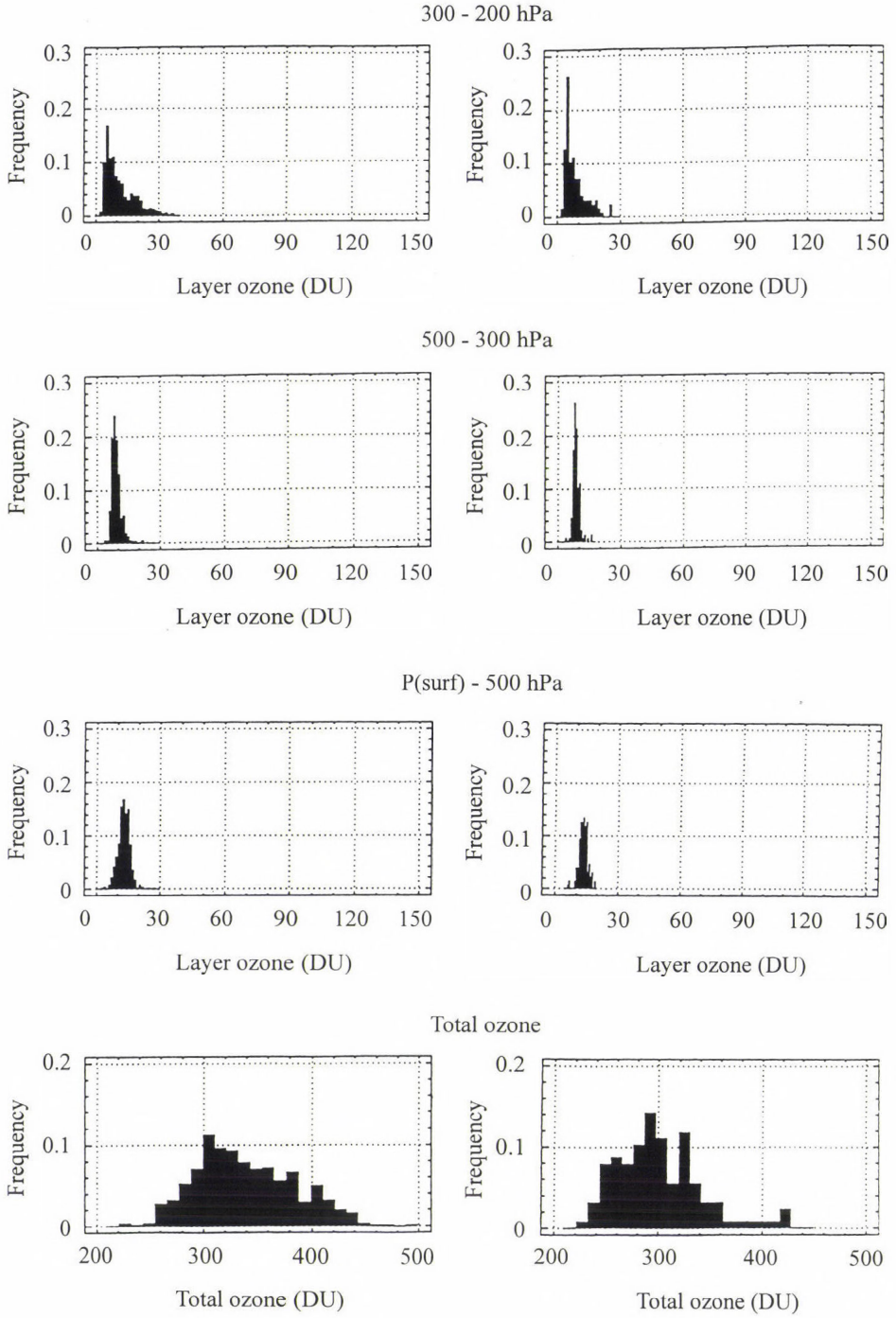
Any possible improvement of the data basis situation by statistical approaches does not mean, however, a possible replacement of measurements by statistical estimates as they reflect normal situations per se. Statistical approaches, however, have another benefit, too. Outliers are either the results of incorrect measurements or data handling or, that is much more important, they indicate the impact of some not yet recognized or, due to changing environmental conditions, newly acting processes. The history of the detection of the antarctic ozone hole is a good example in this respect (e.g., *Fabian*, 1989; *Roan*, 1989).

This paper describes the statistical approach applied to assess the vertical ozone structure in Central Europe and estimates its quality by theoretical considerations as well as by intercomparison with more than 100 ozone sonde data at Hohenpeissenberg, Germany, throughout one year.

2. Short review of the climatology of the vertical ozone structure

Climatology is the bench mark of each statistical approach. Thus some climatological knowledge is needed for a sound assessment of such approaches. This climatological background is given in *Fig. 1* as frequency distribution of layer ozone $\Delta\Omega$ at Hohenpeissenberg for the atmospheric layers indicated in this figure. The frequency distribution is given for normal years in the left panels and for disturbed years after volcanic eruptions (1983, 1992 and 1993) in the right panels. The scale of $\Delta\Omega$ is the same for all layers for easy intercomparison. The frequency distributions are given for the whole year; shorter time intervals (seasons, months) as references result in a narrowing of the distribution and a shift of the characteristic distribution parameters according to the seasonal course (e.g., *Dütsch*, 1978). Additionally, the lowest panels of Fig. 1 give the frequency distribution of total ozone Ω.

Total ozone and layer ozone from 300 to 50 hPa are lognormally or gamma distributed as the result of rare cases of high layer ozone. The distribution is broadest between 200 and 50 hPa with a standard deviation of about 15 to 20 DU and narrowest in the lower and middle troposphere with a standard of about 2 DU, respectively. Normal and disturbed years differ in the location of the


300 - 200 hPa
Frequency
0.3
0.2
0.1
0
0 30 60 90 120 150
Layer ozone (DU)
500 - 300 hPa
P(surf) - 500 hPa
Total ozone
0.2
200 300 400 500
Total ozone (DU)

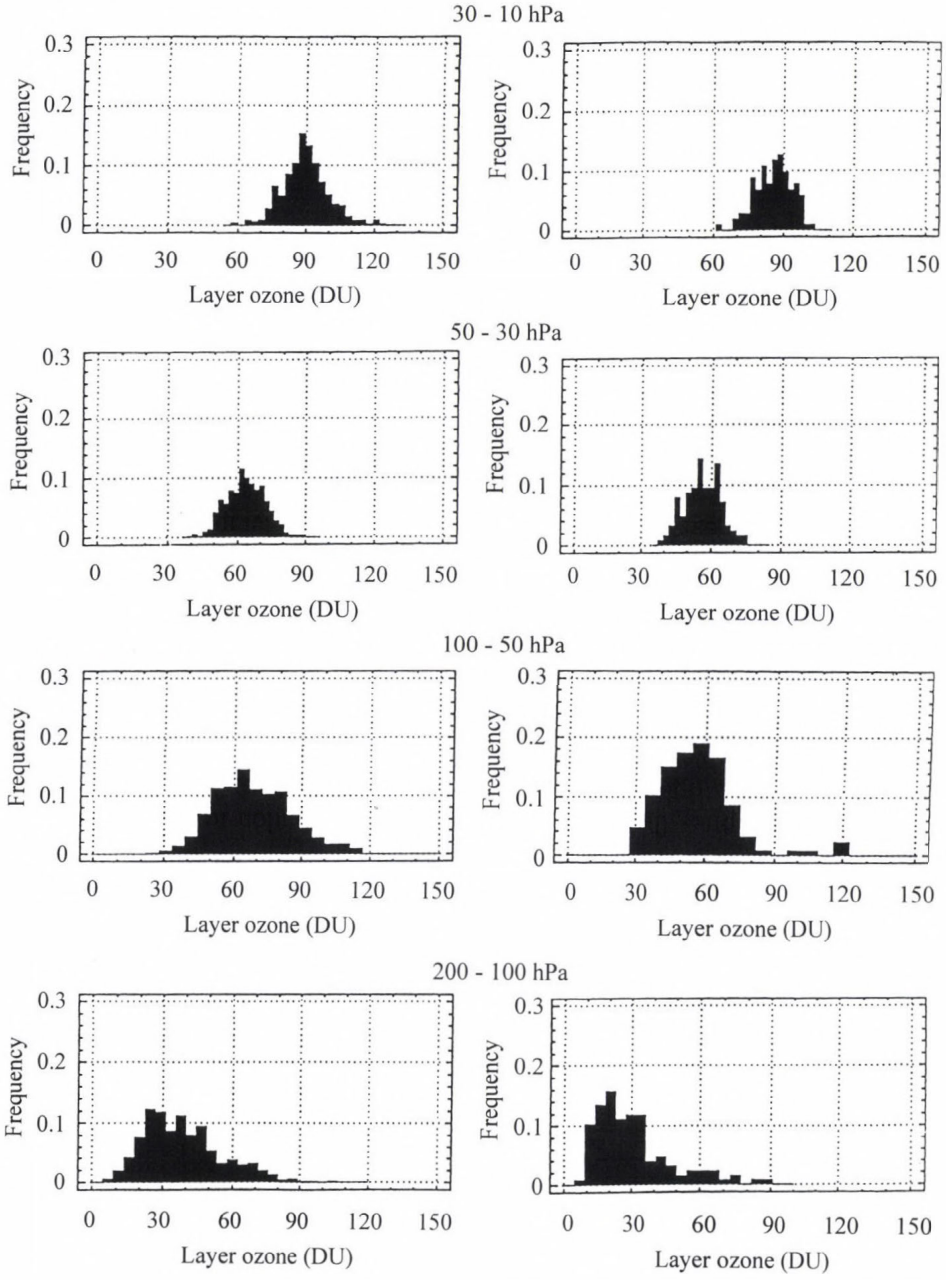


Fig. 1. Frequency distribution of total column ozone (Ω) and layer ozone amount (ΔΩ) at Hohenpeissenberg 1975 to 1994 for undisturbed years (left) and for years after volcanic eruptions (right).

distribution, but not in its spread. *Krzyścin* (1994) has already stated a shift of the Ω frequency distribution as a whole throughout the years without any change of its shape. Fig. 1 shows this finding for $\Delta\Omega$, too. The most reduction in layer ozone after volcanic eruptions, roughly 20% on average, is observed in the layer 200 to 50 hPa. The corresponding median values are 67.2 DU and 36.7 DU for the layers 100 to 50 hPa and 200 to 100 hPa for undisturbed years and 55 DU and 26 DU, respectively, after volcanic eruptions.

Two possible consequences for further steps could arise, namely modified statistical relations for disturbed and undisturbed years, and difficulties in assessing high $\Delta\Omega$ values due to the skewed distribution function of the predictands. The latter issue will be discussed in some detail later on.

3. Statistical approach

The statistical approach applied is the approach of multiple linear regression. An essential condition for the choice of potential predictors was the easy availability of these predictors in data banks confining their range to basic meteorological parameters such as geopotential, layer thickness, temperature, wind direction and velocity, and so on. Ozone data of previous days and thus persistence were not used. Thus some loss in the quality of the results is anticipated. The training sample consisted of ozone sonde and aerological data at Lindenberg (52°13'N, 14°07'E) and Hohenpeissenberg (47°48'N, 11°01'E) representing Central European conditions from the years 1975 to 1994. Multiple linear regression is a very reduced version of the general statistical approach outlined by *Pokrovsky* and *Timofeyev* (1972) where the solution is a combination of information on spectral radiance measurements and a priori meteorological knowledge by means of autocorrelation and crosscorrelation matrixes. This method was used as a pure statistical approach, i.e. without any measurement information, by *Timofeyev et al.* (1974) and later by *Feister* and *Spänkuch* (1977) who obtained encouraging results in ozone profiling when additionally to the temperature and temperature-ozone correlation the total ozone amount Ω was used as input with a maximum gain between about 300 and 50 hPa.

4. Statistical relations and reduction of variance

The quality of statistical relations is characterized by the reduction of variance achieved

$$RV = 1 - \left(\frac{\sigma_{re}}{\sigma_{cl}}\right)^2, \tag{1}$$

where σ_{re} and σ_{cl} mean the standard deviation between measurement and statistical estimate, and the climatological standard deviation, respectively. An exact physical relation results in $RV = 1$ ($\sigma_{re} = 0$). $RV = 0$ ($\sigma_{re} = \sigma_{cl}$) means that the statistical method does not give any more information against climatology. The stronger the statistical relation is the stronger, the higher the RV is.

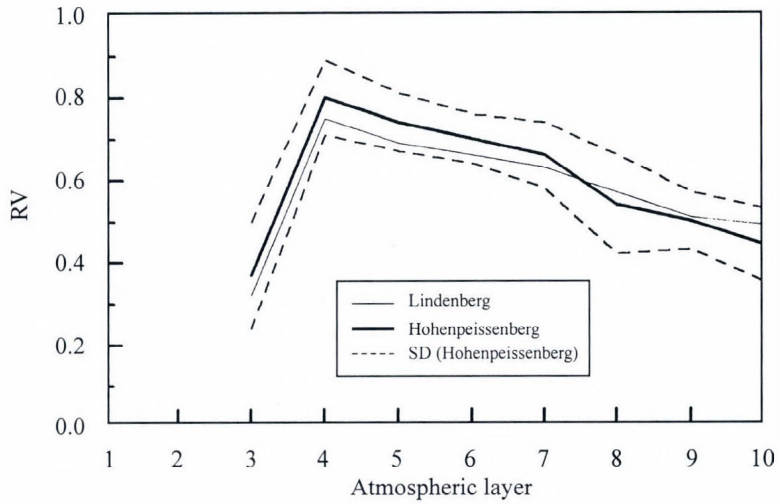


Fig. 2. Reduction of variance (*RV*) by statistical estimate of layer ozone.

Fig. 2 shows the yearly *RV* means at Lindenberg and Hohenpeissenberg, the latter one with $\pm$ the standard deviation of the monthly means. The altitude of the atmospheric layers increases with increasing number from p_s–700, 700–500, 500–300, 300–200, 200–150, 150–100, 100–70, 70–50, 50–30, and 30–20 hPa, respectively. The maximum values of *RV* are attained when all predictors of some significance even if marginal (e.g., $< 2\%$) are taken into account. The highest *RV* with about 0.8 is found for the tropopause region (layer 4: 300–200 hPa). There is a slow, steady decrease of *RV* within the stratosphere up to around 0.5 for layer 10 (30–20 hPa). Amounts larger than and around 0.7 are observed for the altitude range 300 to 100 hPa (layers 4 to 6) where monitoring from satellites is difficult and the height resolution of SBUV is limited. These results confirm previous investigations (e.g., *Spänkuch* and *Döhler*, 1975; *Fortuin* and *Kelder*, 1996) that stated maximum linear correlation between local ozone and temperature in the tropopause region and the lower stratosphere. *RV* is less than 0.3 in the troposphere. However, the climatological variance (σ_{cl}) is so small there (see Fig. 1) that the mean uncertainty of the

statistical estimate is only of the order of ± 1 Dobson units (DU) despite this low *RV*. The yearly course of *RV* at Hohenpeissenberg is given in *Fig. 3* for the atmospheric layers 300–200 hPa and 100–50 hPa, respectively. The highest *RV* is observed in general in winter and the smallest one in summer and fall.

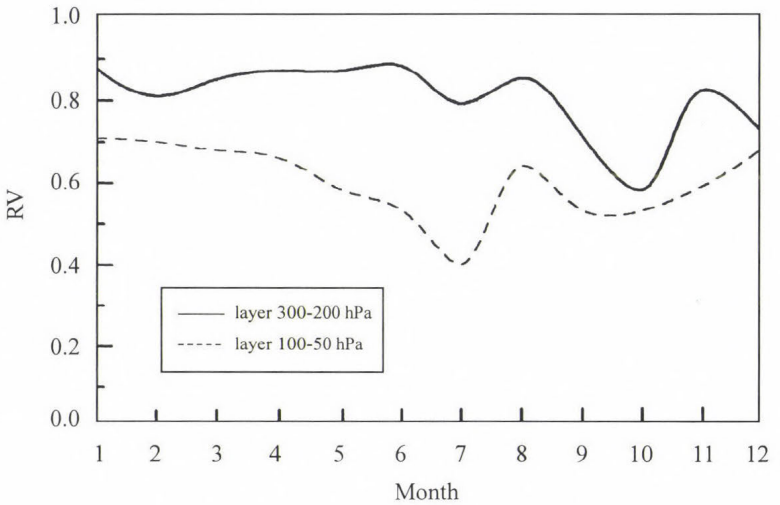


Fig. 3. Yearly course of *RV* for two atmospheric layers at Hohenpeissenberg.

Table 1 compiles the predictors and their contribution to *RV* for the four seasons for the multiple linear regression approach applied. The explanation of the predictors is given in the footnote of Table 1. The number of atmospheric layers was reduced, compared to the ones of Fig. 2, to safe space. Possible nonlinear dependence was taken into account to some extent by offering predictors with nonlinear dependence (e.g., logarithm or power) on basic meteorological parameters. There are six classes of predictors: column ozone Ω, temperature T_i at pressure levels i, layer thickness $\Delta\phi_i$, geopotential ϕ_i, normalized temperature gradient of layers $\nabla\check{T}_i$ and the saturation vapor pressure t_i or $\ln t_i$ as nonlinear parameters for the ith layer mean temperature. Although there is some redundancy by using layer thickness $\Delta\phi_i$ and saturation vapor pressure t_i as predictors simultaneously because both parameters are a function of the layer mean temperature t_m, the nonlinear dependence of t_i from t_m adds some useful information which is not appropriately taken into account by the linear dependence of $\Delta\phi$ from t_m. Table 1 contains two *RV* values, namely $RV_{max,\,total}$ when all predictors of significant even of marginal contributions were used, and RV_{total} when only predictors were used with a contribution to *RV* of

Table 1. Contribution of predictors in percentage to the reduction of variance of layer ozone for the four seasons

Layer	No	3				4				5				6				7			
	p (hPa)	300-200				200-100				100-50				50-30				30-10			
Predictor	Season	Sp	Su	Au	Wi	Sp	Su	Au	Wi	Sp	Su	Au	Wi	Sp	Su	Au	Wi	Sp	Su	Au	Wi
Ω		45	46	35	33	60	67	45	63	57	49	47	68	24	39	27		20	14	27	
ln Ω																	47				25
T_{500}							5													22	
T_{300}			7																		
T_{100}								11		13											
T_{50}						2															
T_{30}														6					9		
T_{10}																					4
$\Delta\phi_1$				14																	
$\Delta\phi_2$		8		17	13														5		
$\Delta\phi_3$			30	5	32																
$\Delta\phi_4$											4										
$\Delta\phi_5$															5	9	5				
$\Delta\phi_6$								17							4		5				
ϕ_4									6			13									
ϕ_6														5							
ln ϕ_7																					20
$\nabla\check{T}_5$							7														
$\nabla\check{T}_7$																		9		8	
t_3		30																			
t_5																					7
t_7																		23			
ln t_1											4										
RV, total		83	83	71	78	62	79	73	69	70	57	60	68	35	48	36	57	52	28	57	56
RV, max, total		86	87	73	82	72	83	77	79	75	69	64	75	48	57	41	63	56	32	64	60
Sum of predictors		3	3	4	3	2(7)	3	3	2	2	3	2	1	3	3	2	3	3	3	3	4

Explanation of the predictors: Ω – total column ozone in Dobson units, T_i – temperature (K) at pressure level i (hPa), ϕ_i and $\Delta\phi_i$ – geopotential and thickness in gpm of layer i, layers 1 and 2 are from surface pressure p_s to 500 hPa and from 500 to 300 hPa, respectively; $\nabla\check{T} = g_n \cdot (T_u - T_l) / T_l$ with T_u, T_l the temperatures at the upper and lower boundary, resp., of the layer given by the index, g_n – acceleration of gravity, t_i = saturation vapor pressure of layer i calculated from layer mean temperature.

at least 4 to 5 percent. The last line of Table 1 lists the number of predictors for RV_{total}. Three to four predictors are sufficient in general. RV_{total} can increase by about 0.1, e.g., for the layer 200 to 100 hPa in winter and spring, when 6 to 7 predictors instead of only 2 were considered.

The largest contribution to RV is from the total column ozone Ω or ln Ω, respectively, with a maximum reduction of about 2/3 for the layer 200–100 hPa (exception fall) and for the layer 100–50 hPa in winter, confirming the results of *Feister* and *Spänkuch* (1977). The RV_{total} is for the layer 200–100 hPa of the same order of magnitude in fall, too, due to the layer thickness of the sixth layer ($\Delta\phi_6$) and temperature at 100 hPa, T_{100}, as predictors.

Fig. 4 demonstrates the high correlation between Ω and layer ozone amount $\Delta\Omega$ for one single layer, the layer 200–150 hPa with a linear (upper panel) and a logarithmic regression approach (lower panel), the latter one is slightly more appropriate for large $\Delta\Omega$. The corresponding correlation coefficients are 0.80 and 0.78, respectively. The correlation was calculated with the Hohenpeissenberg data from 1975 to 1994. The single correlation coefficients between Ω and $\Delta\Omega$ are between 0.7 and 0.8 for 300 to 30 hPa, about 0.4 for 30 to 10 and 500 to 300 hPa and insignificant in the lower troposphere.

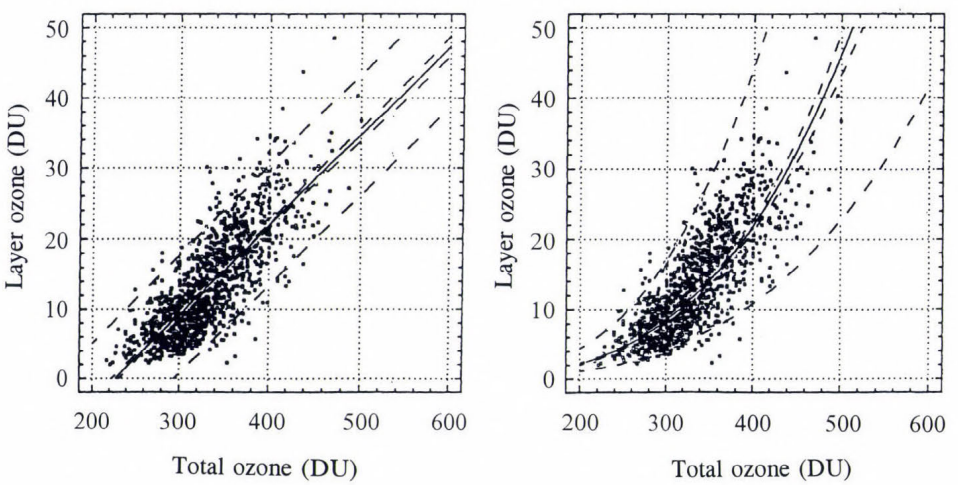


Fig. 4. The relation between layer ozone between 200 to 150 hPa and total ozone at Hohenpeissenberg.

Layer thicknesses ($\Delta\phi$) of the lower layers contribute significantly to RV in the tropopause region (300 to 200 hPa), except for spring. Their total contribution is about 0.4 and of the same order of magnitude as the contribution of Ω.

In spring, the saturation vapor pressure of layer 3, t_3, of the tropopause layer contributes 30% to *RV*. This order of magnitude is only reached for the layer 30 to 10 hPa in spring by the analogous parameter t_7. For this layer *RV* is about 0.5 and even only 0.3 in summer with half of the contribution from Ω. The high contribution of predictors characterizing the thermal state of the tropopause region is mainly due to quasi-columnar motion of air along isentropic surfaces in the lower stratosphere which explaines more than half of the variance of Ω (*Salby* and *Callaghan*, 1993).

The effect of the individual predictors is illustrated by a single example, and that for the tropopause region 300 to 200 hPa in May over Hohenpeissenberg. The maximum RV considering 6 predictors is 0.87 (see Table 1). Using as single predictors only the most effective ones, Ω and T_{200}, respectively, results in *RV* amounts of 0.55 and 0.69, respectively (*Fig. 5* upper panels). Obviously, T_{200} is a more appropriate predictor than Ω in this case. Fig. 5 shows that both single regression equations estimate ΔΩ up to 30 DU in maximum only, in contrast to the observation with 44 DU in maximum. There is an increase of *RV* to 0.75 using both predictors simultaneously (Fig. 5, lower left panel), however, the deficit remains at high ΔΩ. A sufficient estimate of ΔΩ for high ΔΩ is only reached when further predictors are added (Fig. 5, lower right panel). Note that in case of more predictors T_{200} is replaced by T_{300} (see Table 1).

5. Impact of predictor uncertainty on the quality of the statistical estimate

Another feature of interest besides the individual predictors' contribution to *RV* is the sensitivity of the quality of the statistical estimate to the quality of the predictors. This information is easily derived from the magnitude of the regression coefficients and compiled in *Table 2*. The numbers of Table 2 are related to a change in the estimate of ΔΩ in the corresponding atmospheric layer by 1 DU caused by a change of the specific predictor by the amount given. The figures of Table 2 are related to DU and K for Ω, ln Ω and T_i, respectively, and in per cent of the natural variation of the rest of the predictors. The estimate of ΔΩ is most sensitive to the temperature if this predictor is used. A change, e.g., of the ΔΩ estimate of the layer 100 to 50 hPa in spring by +1 DU is caused by a change of T_{100} by –0.5 K. This uncertainty of about 1 K results in a 2 DU uncertainty of the ΔΩ assessment already. A similar magnitude (–0.7 K) was found for T_{300} in assessing ΔΩ in the tropopause region in summer as well as for T_{100} for the 200 to 100 hPa layer in fall. The temperature uncertainty is less crucial for the other cases with temperature as a predictor.

Uncertainties in Ω are most effective for the ΔΩ assessment in that layers with a high contribution of Ω to *RV*, namely for the layers 200 to 100 hPa and 100 to 50 hPa. A 1 DU change in ΔΩ there corresponds to 3 to 5 DU in the uncertainty of Ω. However, the rule "high contribution to *RV* of a predictor

corresponds to high sensitivity to the statistical $\Delta\Omega$ estimate" does not hold unconditionally as shown by a comparison of the corresponding amounts for the layers 300 to 200 hPa and 50 to 30 hPa, respectively.

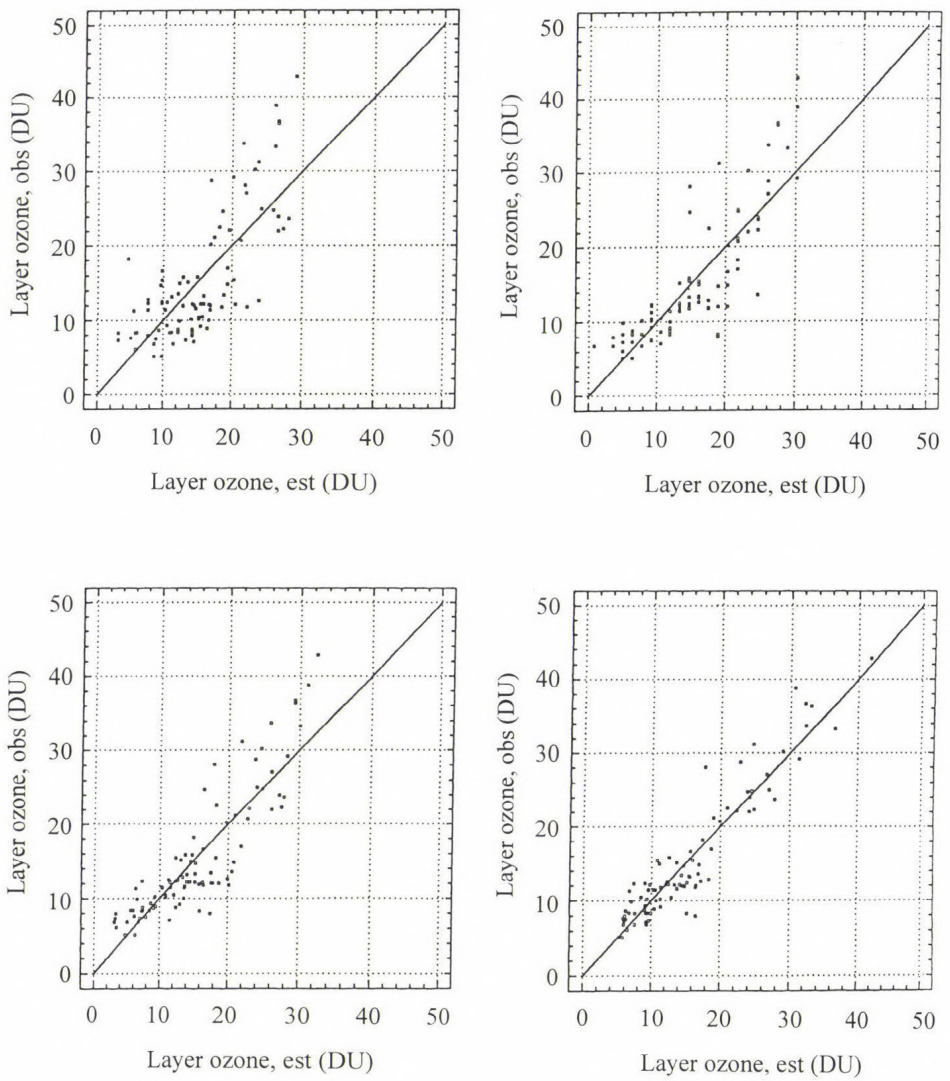


Fig. 5. Comparison of observed and estimated layer ozone in the tropopause region 300–200 hPa at Hohenpeissenberg in May with predictors Ω only (upper left panel), T_{200} only (upper right panel), Ω plus T_{200} (lower left panel) and all significant six predictors (lower right panel).

Table 2. Predictors' order of magnitude to change layer ozone estimate by 1 DU

Layer	No	3				4				5				6				7			
Predictor	p (hPa) Season	300-200 Sp	Su	Au	Wi	200-100 Sp	Su	Au	Wi	100-50 Sp	Su	Au	Wi	50-30 Sp	Su	Au	Wi	30-10 Sp	Su	Au	Wi
Ω		18	17	27	35	3	5	4	4	3	5	5	3	13	9	6		20	8	4	
ln Ω																	9				10
T_{500}							1.6													1	
T_{300}			.7																		
T_{100}								.7		.5											
T_{50}						2															
T_{30}														5					1.3		
T_{10}																					3
$\Delta\phi_1$				25																	
$\Delta\phi_2$		5		6	5														8		
$\Delta\phi_3$			4	8	6																
$\Delta\phi_4$											7										
$\Delta\phi_5$															8	9	5				
$\Delta\phi_6$								3							5		5				
ϕ_4									3			6									
ϕ_6														12							
ln ϕ_7																					7
∇T_5							8														
∇T_7																		4		3.5	
t_3		2																			
t_5																					5
t_7																		2.5			
ln t_i											80										

Explanation of the predictors see Table 1. The figures for all parameters except Ω, ln Ω and T_i are the percentage of natural variability of the corresponding parameter. For Ω and ln Ω the figures are in DU, for T_i in K.

Satellite measurements of Ω are estimated to be accurate to $\pm 2\%$ or ± 8 DU (*Fleig et al.*, 1990; *Herman et al.*, 1991), and Ω forecasting is claimed to be accurate to ± 3 to 4% or ± 10 to 12 DU (*Burrows et al.*, 1995; *Spänkuch* and *Schulz*, 1995, 1997a; *Long et al.*, 1996). The resulting uncertainty in assessing $\Delta\Omega$ for the altitude range from 200 to 50 hPa is ± 2 DU when satellite measurements will be used, and ± 3 to 4 DU in case of predicted Ω. For all other predictors a change in the estimate of $\Delta\Omega$ by 1 DU is caused by a change of the predictors by about 5 to 10 per cent of the predictors' natural variability, which is equal to the order of magnitude of the measurement and analysis error.

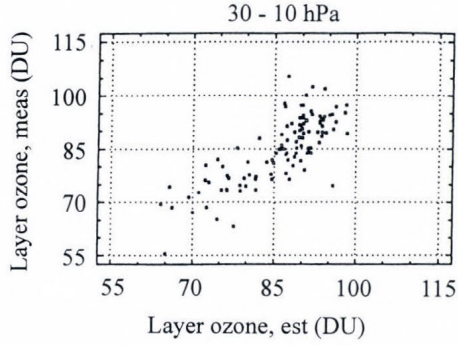
The resulting inaccuracy in the $\Delta\Omega$ estimate due to measurement and/or analysis errors is assessed to be about 3 to 4 DU when Ω is taken from satellite measurements. Consequently, the attainable maximum accuracy of $\Delta\Omega$ is about 3 to 10 per cent for the individual atmospheric layers (see Fig. 1).

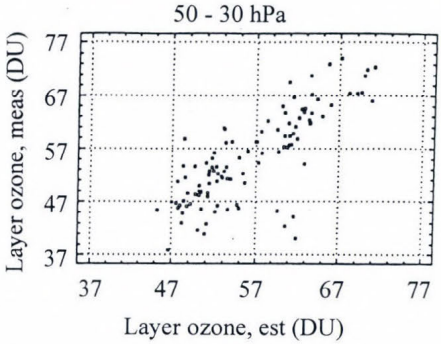
6. Results

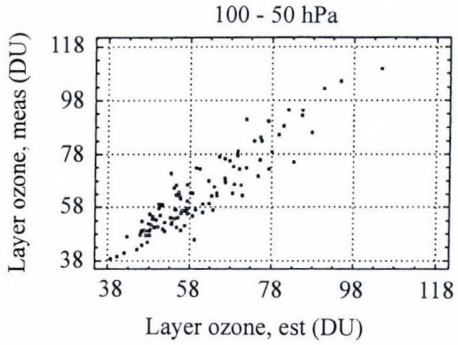
The results of layer ozone estimates at Hohenpeissenberg during the year 1996 as independent sample with more than 100 individual cases are shown in the scatter plots of *Fig. 6* and summarized in *Table 3*. Given are the results for each layer for completeness, even in case of only marginal improvement against climatology as for the tropospheric layers, and in total for the layer 300 to 50 hPa, where the most significant improvement was found. Fig. 6 reveals several interesting aspects. First, we note a negative bias of the $\Delta\Omega$ estimate ($\Delta\Omega_{est} < \Delta\Omega_{meas}$) of about 1 to 2 DU up to 50 hPa (see Table 3) and a positive bias above that level. In August 1995, the radiosonde type was changed from VIZ 1993 of VIZ Manufacturing, Philadelphia, USA, to RS 80 of Vaisala, Finland. Both radiosondes use different devices for pressure measurements resulting not only in differences in the pressure values measured (*Nash* and *Schmidlin* (1987) and *Ivanov et al.* (1991) for details), but also in differences in temperature and ozone. The resulting differences in ozone are up to 6 per cent above 26 km (20 hPa) and, due to the impact on the Dobson correction, up to about 1.5 per cent below 20 km (50 hPa) (*Steinbrecht* 1998, private communication). Thus the bias is not real and only the impact of the inhomogeneity in the Hohenpeissenberg data. The corresponding *rms* of the estimate without this virtual bias is given in brackets in Table 3.

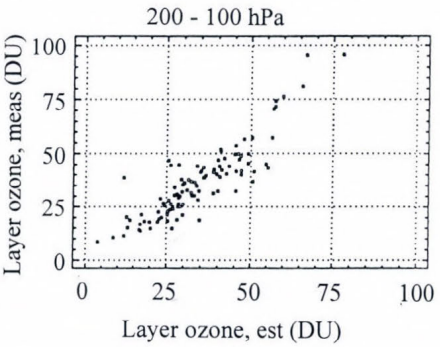
Second, there is a rate of about 8 to 10 per cent outliers with considerable underestimation in $\Delta\Omega$ for high $\Delta\Omega$ amounts in the layers below 100 hPa particularly distinct in the layer 200 to 100 hPa, with unacceptable large errors

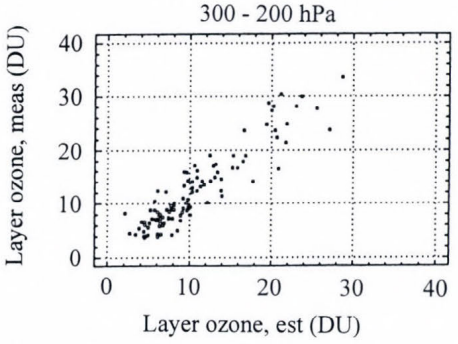
Fig. 6. Comparison of observed and estimated layer ozone at Hohenpeissenberg in 1996 (independent sample). Note the different scales of the panels.☞

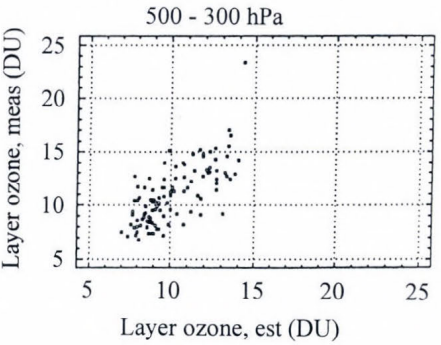

30 - 10 hPa
Layer ozone, meas (DU)
115
100
85
70
55
55
70
85
100
115
Layer ozone, est (DU)

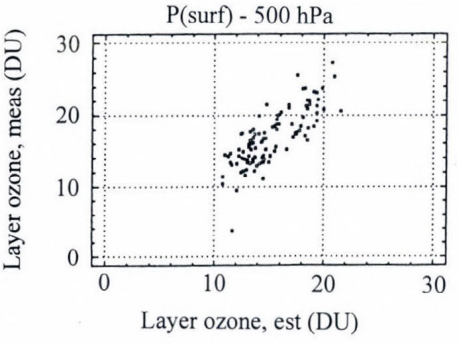

50 - 30 hPa
Layer ozone, meas (DU)
77
67
57
47
37
37
47
57
67
77
Layer ozone, est (DU)

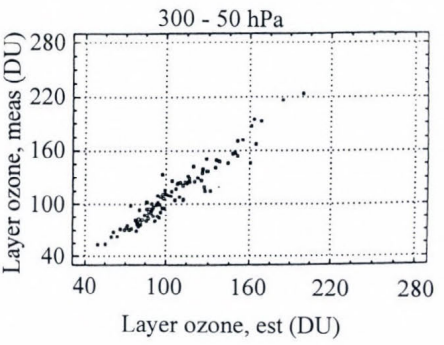

100 - 50 hPa
Layer ozone, meas (DU)
118
98
78
58
38
38
58
78
98
118
Layer ozone, est (DU)


200 - 100 hPa
Layer ozone, meas (DU)
100
75
50
25
0
0
25
50
75
100
Layer ozone, est (DU)


300 - 200 hPa
Layer ozone, meas (DU)
40
30
20
10
0
0
10
20
30
40
Layer ozone, est (DU)


500 - 300 hPa
Layer ozone, meas (DU)
25
20
15
10
5
5
10
15
20
25
Layer ozone, est (DU)


P(surf) - 500 hPa
Layer ozone, meas (DU)
30
20
10
0
0
10
20
30
Layer ozone, est (DU)


300 - 50 hPa
Layer ozone, meas (DU)
280
220
160
100
40
40
100
160
220
280
Layer ozone, est (DU)

of 15 DU and more. In all these cases filament-like structures were found. Thus the results of the approach have to be taken with care in such cases that can be characterized by a $\Delta\Omega$ estimate larger than a certain threshold to be defined by further studies. The present results suggest, e.g., about 58 DU as threshold for the layer ozone between 200 and 100 hPa, and about 160 DU for the layer 300 to 50 hPa. In the latter case the bias is about –20 DU. *Appenzeller* and *Holton* (1997) investigating the climatology of filament-like structures in the stratosphere, found three geographic areas of favourable occurrence of filaments, all linked to the edge of the polar vortex, namely Europe, Siberia and Northern Canada.

Table 3. Results of layer ozone estimate 1996

Layer (hPa)	Bias (DU)	σ_{cl}	σ_{es} (DU)	rms (DU)	r^2	Change (%)
p_s –500	–1.5	3.8	2.4	2.8 (2.2)	0.53	64
500–300	–0.8	2.8	1.9	2.1 (1.7)	0.55	63
300–200	–1.6	7.0	2.8	3.2 (2.6)	0.85	83
200–100	–1.8	16.6	7.9	8.1 (7.2)	0.77	80
100 – 50	–2.6	14.8	5.8	6.3 (5.5)	0.85	67
50 – 30	1.2	8.4	5.2	5.4 (4.7)	0.59	56
30 – 10	1.2	9.3	5.9	6.0 (5.4)	0.62	70

The most satisfying results were achieved for the layers 300 to 200 hPa and 100 to 50 hPa with the square of the linear regression coefficient of 0.85 in both cases (Table 3). Omitting the outliers which are mostly found in the $\Delta\Omega$ estimate of single layers, we state satisfying results for the lower stratosphere up to 50 hPa, as demonstrated in the corresponding panels of Fig. 6.

The statistical estimate of $\Delta\Omega$ for the troposphere in particular, and above 50 hPa is to be taken only as a kind of first guess as expected from the discussion of Chapter 3. Nevertheless, the estimate of the sign of the diurnal changes in $\Delta\Omega$ given in the last column of Table 3 is with 70 per cent for the layer 30 to 10 hPa like the one of the layer 100 to 50 hPa. Again, the highest rate with 83 per cent in the correct estimate of the sign of the diurnal change is found for the layer 300 to 200 hPa.

7. Summary and conclusions

A multiple linear regression approach is presented to estimate layer ozone amount $\Delta\Omega$ from predictors accessible from model outputs of meteorological analysis and prediction additionally to the total ozone amount that has regularly been monitored from satellites for several years. Total column ozone is the most important predictor in general of the approach developed and checked for northern midlatidudes. The second most important predictor is the thermal state of atmospheric layers. The accuracy of the statistical estimate is assessed to be about 3 to 10 per cent in maximum dependent on atmospheric layer and season. Comparison of the estimate with more than one hundred ozone sonde data at Hohenpeissenberg, Germany, throughout one year showed promising results between 300 and 50 hPa when outliers of about 8 to 10 per cent of cases are omitted. The outliers were mainly found for the layer between 200 and 100 hPa and are obviously related to filament-like structures in ozone there. The most satisfying results were obtained for the tropopause region 300–200 hPa with a reduction of variance of about 0.7 to 0.8.

Acknowledgement—The authors are pleased to acknowledge support of this research by the Bundesministerium für Forschung und Technologie (BMFT) through Grant No. 01 L0 9509 9.

References

Appenzeller, C. and *Holton, J.R.*, 1997: Tracer lamination in the stratosphere: A global climatology. *J. Geophys. Res. 102,* 13,555-13,569.

Burrows, W.R., Vallée, M., Kerr, J.B., Wilson, L.J. and *Tarasick, D.W.*, 1995: The Canadian operational procedure for forecasting total ozone and UV radiation. *Meteorol. Appl. 1*, 247-265.

Dütsch, H.U., 1978: Vertical ozone distribution on a global scale. *Pure Appl. Geophys. 116*, 511-529.

Fabian, P., 1989: *Atmosphäre und Umwelt.* Springer Verlag, Berlin.

Feister, U. and *Spänkuch, D.*, 1977: The deviation of the vertical ozone profile from statistical relations. In *Proc. Joint Symposium on Atmospheric Ozone*, Vol. I., (ed.: *K.H. Grasnick*). Dresden, 9-17 August 1976. 331-360, Berlin.

Fleig, A.J, Mc. Peters, R.D., Bhartia, P.K., Schlesinger, B.M., Cebula, R.P., Klenk, K.F., Taylor, S.L. and *Heath, D.F.*, 1990: Nimbus 7 solar backscatter ultraviolett (SBUV) ozone products user's guide. *NASA Ref. Publ.*, 19-25.

Fortuin, J.P.F. and *Kelder, H.*, 1996: Possible links between ozone and temperature profiles. *Geophys. Res. Lett. 23*, 1517-1520.

Götz, F.W.P., 1931: Zum Strahlungsklima des Spitzbergensommers. *Gerlands Beitr. Geophys. 95*, 119-154.

Herman, J.R., Hudson, R., Mc. Peters, R.D., Stolarski, R., Gu, X.-Y., Taylor, S. and *Wellenmeyer, C.*, 1991: A new self-calibration method applied to TOMS and SBUV backscattered ultraviolet data determine long-term global ozone change. *J. Geophys. Res. 96*, 7531-7545.

Ivanov, A., Kats, A., Kurnosenko, S., Nash, N. and *Zaitseva, N.*, 1991: WMO International Radiosonde Comparison. Phase III, (Dzhambul, USSR, 1989). World Meteorological Organization, Instruments and Observing Methods. *Report 40, WMO/TD-NO. 451*, Genf.

Krzyścin, J.W., 1994: Total ozone changes in the northern hemisphere mid-latitudinal belt (30–60°N) derived from Dobson Spectrophotometer measurements, 1964–1988. *J. Atmos. Terr. Phys. 56*, 1051-1056.

Lacis, A.A., Wuebbles, D.J. and *Logan, J.A.*, 1990: Radiative forcing of climate by changes in the vertical distribution of ozone. *J. Geophys. Res. 95*, 9971-9981.

Long, C.S., Miller, A.J., Lee, H.-T., Wild, J.D., Przywarty, R.C. and *Hufford, D.*, 1996: Ultraviolet index forecast issued by the National Weather Service. *Bull. Amer. Meteorol. Soc. 77*, 729-748.

de Luisi, J.J. and *Mateer, C.L.*, 1971: On the application of the optimum statistical inversion technique to the evaluation of Umkehr observations. *J. Appl. Meteorol. 10*, 328-334.

Nash, J. and *Schmidlin, F.J.*,1987: WMO International Radiosonde Comparison (U. K. 1984, USA 1985). World Meteorological Organization, Instruments and Observing Methods. *Report 30, WMO/TD-NO. 195*, Genf.

Pokrovsky, O.M. and *Timofeyev, Yu.M.*, 1972: A general statistical approach for the solution of inverse problems of atmospheric optics (in Russian). *Meteorol. Gidrol. No. 1*, 52-59.

Ramaswamy, V., Schwarzkopf, M.D. and *Shine, K.P.*, 1992: Radiative forcing of climate from halocarbon-induced global stratospheric ozone loss. *Nature 355*, 810-812.

Roan, S.L., 1989: *Ozone Crisis: The 15-year Evolution of a Sudden Global Emergency*. Wiley, New York.

Salby, M.L. and *Callaghan, P.F.*, 1993: Fluctuations of total ozone and their relationship to stratospheric air motions. *J. Geophys. Res. 98*, 2715-2727.

Seliers, W.D. and *Yarger, D.N.*, 1969: The statistical prediction of the vertical ozone distribution. *J. Appl. Meteorol. 8*, 357-361.

Spänkuch, D. and *Döhler, W.*, 1975: Statistische Characteristika der Vertikalprofile von Temperatur und Ozon und ihre Korrelation über Berlin. *Geodät. und Geophys. Veröff., Reihe II, Heft 19*.

Spänkuch, D. and *Schulz, E.*, 1995: Diagnosing and forecasting total column ozone by statistical relations. *J. Geophys. Res. 100*, 18,873-18,885.

Spänkuch, D. and *Schulz, E.*, 1997a: On short-term total-column-ozone-forecast errors. *Atmos. Environment 31*, 117-120.

Spänkuch, D. and *Schulz, E.*, 1997b: Quality checks for satellite and ground-based total ozone observations. In *Ground-level and Satellite Observations: Changes in the Mediterranean Region* (ed.: *C. Varotsos*). Springer-Verlag, Berlin-Heidelberg, 293-302.

Timofeyev, Yu.M., Kucnecov, A.D. and *Spänkuch, D.*, 1974: About the retrieval accuracy of the atmospheric ozone profile. In *Radiative Processes in the Atmosphere and at the Surface* (in Russian). Trudy Allunions-Meeting on Actinometry. Nauka, Kiev, 126-130.

WMO, 1995: *Scientific Assessment of Ozone Depletion: 1994, 1995: World Meteorological Organization Global Ozone Research and Monitoring Project-Report No. 37, Geneva.*

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 4, October–December 1998, pp. 219–237

Characterizing air pollution potential over Budapest using macrocirculation types

I. Matyasovszky and **T. Weidinger**

Department of Meteorology, Eötvös Loránd University,
H-1117 Budapest, Pázmány Péter sétány 1, Hungary; E-mail: matya@ludens.elte.hu

(Manuscript received 23 February 1998; in final form 15 June 1998)

Abstract—Time series of height of planetary boundary layer (PBL), mean virtual temperature gradient of the lower 300 m, wind speed at 925 hPa pressure level and energy of instability (in the lower 1 km in winter and lower 2 km layer in summer) were calculated using radiosonde data (1962–1992) over Budapest. These parameters principally determine the air pollution potential, i.e., the capacity of mixing of pollutants. Using daily radiosonde data (12 UTC), parameters characterizing air pollution potential are related to Hess-Brezowsky (HB) macrocirculation (MC) types in winter and summer seasons. However, the large number of HB types (30 HB types) makes it difficult to evaluate the statistical relationship, therefore, an algorithm is developed to define a family of MC types obtained by aggregating the HB types. A similarity measure of MC types is defined and used on a seasonal basis in terms of PBL parameters. The family of MC types consisting of 16, 12, 8 and 4 types are shown. For 8 types, a multivariate time series model of PBL parameters is developed and checked. The model is able to simulate statistical properties of PBL parameters under present conditions and then under different hypothetical conditions.

Key-words: air pollution potential, planetary boundary layer, macrocirculation types.

1. Introduction

The capacity of the atmosphere to disperse and dilute pollutants emitted at different scale sources depends upon prevailing meteorological conditions of planetary boundary layer (PBL), such as the height, temperature gradient (stability), wind direction and wind speed, wind shear, and energy of instability of the PBL (*Szepesi et al.*, 1977). These parameters characterize air pollution potential of a given area. The term air pollution potential has been introduced in the seventies and characterizes the ability of loading of the atmosphere by pollutants (*Wiswanadham* and *Santosh*, 1989). The statistical relationship between the macrocirculation types and air pollution potential is an important

issue in order to know how the air pollution characteristics relate to the actual weather. This problem is crucial in case of extreme pollution situations. There is a general empirical knowledge of which weather situations are problematic in the presence of pollutants in the atmosphere, but the present paper is directed at developing a stochastic model to describe this relationship quantitatively.

The ultimate purpose of our work is to develop a multivariate time series model in order to simulate statistical properties of PBL parameters under present conditions and then under different hypothetical conditions of large-scale atmospheric circulation. It is well-known that behavior of meteorological parameters strongly depend on macrocirculation. Therefore, a system of macrocirculation (MC) types is used as a relevant additional information, namely the time series model is related to MC types. The paper reports the present stage of our research.

The first step of developing the model is estimating the probability distribution of PBL parameters controlling air pollution potential. This step was performed by conditioning probability distributions on Hess-Brezowsky (HB) macrocirculation types using radiosonde data for Budapest. However, the large number of HB types (30 HB types) makes it difficult to evaluate the statistical relationship, therefore, an algorithm is developed to define a family of MC types obtained by aggregating the HB types. A similarity measure of MC types is defined and used on a seasonal basis in terms of PBL parameters. Then, a multivariate time series model is developed and is presented under 8 MC types.

First, data sets used in the analysis are presented and dependency of statistical behavior of PBL parameters on HB types is demonstrated. Next, the methodology and results of aggregating HB types resulting in a new system of MC types are reported. Then, properties of time series model reproducing observed statistics are discussed. Finally, a section for conclusions is provided.

2. Characterizing data

Archived and checked radiosonde data for Budapest are available for several years. A data set from 1962 to 1992 is used for analyzing PBL parameters. Winter (December, January, February) and summer (June, July, August) seasons are considered separately with 12 UTC observations. Four parameters have been defined to characterize PBL:

- height of PBL (H [m]),
- gradient of virtual potential temperature in the lower 300 m (γ_V [°C/100 m]),
- wind speed at the 925 hPa standard pressure level (V [m/s]),
- energy of instability (A [J/kg]) in the lower 1 km layer in winter and in the lower 2 km layer in summer in accordance with yearly course of convective PBL height.

A first overall theoretical study of PBL in Hungary was reported by *Bodolai* (1983). Here, the height of PBL is determined by significant points of virtual temperature profile. No unique definition exists for determining the height of convective PBL (*Stull*, 1988). We defined the top of the PBL at the highest significant point where the mean virtual temperature gradient between the surface and that significant point is larger than or equal to 0.95°C/100 m and the virtual temperature gradient below this significant point is larger than or equal to 0.8°C/100 m (the larger this threshold the smaller the estimated PBL height is). We used the so-called modified bubble method for dry adiabatic condition. Sensitivity analysis shows an uncertainty of 10–15 % (*Weidinger et al.*, 1996).

The energy of instability is calculated as:

$$A = \int_{z_0}^{z_1} g \frac{T_v^* - T_v}{T_v} dz, \tag{1}$$

where z_0 and z_1 denotes the lower and upper height of the layer for calculation energy of instability, g is the acceleration due to gravity, T_v and T_v^* correspond to the virtual temperature of environment and the rising air parcel from surface (z_0) by pseudo-adiabatic process.

The usefulness of these parameters is demonstrated by *Table 1*. Statistically significant relations can be identified between virtual temperature gradient and energy of instability, as well as between energy of instability and the height of PBL. In the rest of cases correlation values are quite low (but different from zero at a 5 % significance level due to large sample sizes) indicating that each variable has a large amount of information not contained in other variables.

Table 1. The mean (m), standard deviation (σ) and correlation matrix of PBL characteristics

	Winter						Summer					
	m	σ	Correlation matrix				m	σ	Correlation matrix			
H	604	380	1.00	0.30	0.03	0.53	1406	603	1.00	0.17	-0.10	0.30
γ_v	0.69	0.50	0.30	1.00	-0.09	0.76	1.15	0.27	0.17	1.00	-0.08	0.64
V	8.1	4.8	0.03	-0.09	1.00	-0.03	5.4	3.3	-0.10	-0.08	1.00	-0.14
A	-66.5	69.7	0.53	0.76	-0.03	1.00	39.9	69.1	0.30	0.64	-0.14	1.00

H – height of PBL, γ_v – gradient of potential temperature in the lower 300 m, V – wind speed at the 925 hPa standard pressure level, A – energy of instability (in the lower 1 km in winter and in the lower 2 km in summer)

Hess-Brezowsky macrocirculation types are well-known in meteorological literature; their basic properties are summarized in *Table 2* (*Hess* and *Brezowsky*, 1969; *Bartholy* and *Kaba*, 1987). In order to demonstrate the dependency of PBL parameters on HB types some results for probability distribution of parameters characterizing air pollution potential are shown next for winter and summer under two characteristic HB types. NWz (Northwest cyclonic) type usually advects oceanic air masses. NWz produces mostly positive temperature anomalies (–1 – +3°C) in winter, while causes negative anomalies (–1 – –3°C) in summer. The relative frequency of this type is 6.4% in winter and 4.9% in summer. HM (Central European high) type is characterized mostly with continental air masses resulting in negative temperature anomalies (–2.5 – –1.5°C) in winter and positive anomalies (0.5–1.0°C) in summer. Relative frequency of this type is 7.4% in winter and 6% in summer.

Table 2. Definition of HB types

Major types	Subtypes		Subtypes	
		Zonal		
West	1. West anticyclonic	Wa	2. West cyclonic	Wz
	3. Southern West	Ws	4. Angleformed West	Ww
		Half-meridional		
Southwest	5. Southwest anticyclonic	SWa	6. Southwest cyclonic	SWz
Northwest	7. Northwest anticyclonic	NWa	8. Northwest cyclonic	NWz
Central European high	9. Central European high	HM	10. Central European ridge	BM
Central European low	11. Central European low	TM		
		Meridional		
North	12. North anticyclonic	Na	13. North cyclonic	Nz
	14. North, Iceland high, anticyclonic	HNa	15. North, Iceland high, cyclonic	HNz
	16. British Islands high	HB	17. Central European trough	TrM
Northeast	18. Northeast anticyclonic	NEa	19. Northeast cyclonic	NEz
East	20. Fenno-scandinavian high anticyclonic	HFa	21. Fenno-scandinavian high cyclonic	HFz
	22. Norwegian Sea-Fenno-scandinavian high anticyclonic	HNFa	23. Norwegian Sea-Fenno-scandinavian high cyclonic	HNFz
Southeast	24. Southeast anticyclonic	SEa	25. Southeast cyclonic	SEz
South	26. South anticyclonic	Sa	27. South cyclonic	Sz
	28. British Islands low	TB	29. Western Europe trough	TrW
	30. Unclassified	Ü		

Depending on types, the mean PBL height is around 500–800 m in winter, and is 1250–1650 m in summer. The Weibull distribution fits the histograms best for both seasons and both types. The hypothesis that PBL height follows Weibull distributions in these types can be accepted at least at a 5% significance level. However, when using the entire data set the best fitting distribution is lognormal in winter (*Fig. 1*).

The gradient of potential temperature in the lower 300 m ($\gamma_v = -\Delta T_v / \Delta z$, $\Delta z = 300$ m) cannot be satisfactorily modeled by any well-known probability distribution (*Fig. 2*). In winter, the mean gradient is somewhat larger for NWz (0.74°C/100 m) than for HM (0.67°C/100 m). The instability intensifies in summer: the difference between the HB types is considerably smaller (1.2–1.05°C/100 m) than in winter (0.99–0.5°C/100 m). The HM type produces stronger instability than the cyclonic NWz type.

Flows of PBL are described by speed and direction of wind at the 925 hPa standard pressure level (*Fig. 3*). Wind roses of the HM type are similar and characterized by frequent northwestern, northern, northeastern directions in both seasons, while the relative frequency of northwestern direction exceeds 40% in NWz type. The mean wind speed in winter is between 5.6 m/s and 11.5 m/s depending on types. The summer exhibits a considerably smaller difference between mean wind speeds in each HB type. The entire set of winter wind speeds is best fitted by a Weibull distribution, while the most suitable distribution is the lognormal in summer. In winter, NWz and HM types can be described by a normal and a lognormal distribution, respectively. Both types require a Weibull distribution in summer. Statistical tests (chi-square and Kolmogorov-Smirnov tests) show that the distributions fit the data substantially better for NWz type than in the case of HM accompanied with smaller wind speeds. The wind turning between 925 hPa standard pressure level and the surface has a remarkable property (*Fig. 4*). Namely, the wind turns left (reverse from Ekman spiral) with the height with relative frequencies of 25% in winter and 50% in summer.

Energy of instability was calculated for different layers in winter and summer due to the annual cycle of the convective PBL height. The mean energy of instability is obviously negative in winter and positive in summer for each HM type. The absolute value of mean energy of instability naturally was higher in the HM anticyclonic type than in the NWz cyclonic type in both seasons (*Fig. 5*). Commonly used distribution functions cannot fit the histograms in the majority of types.

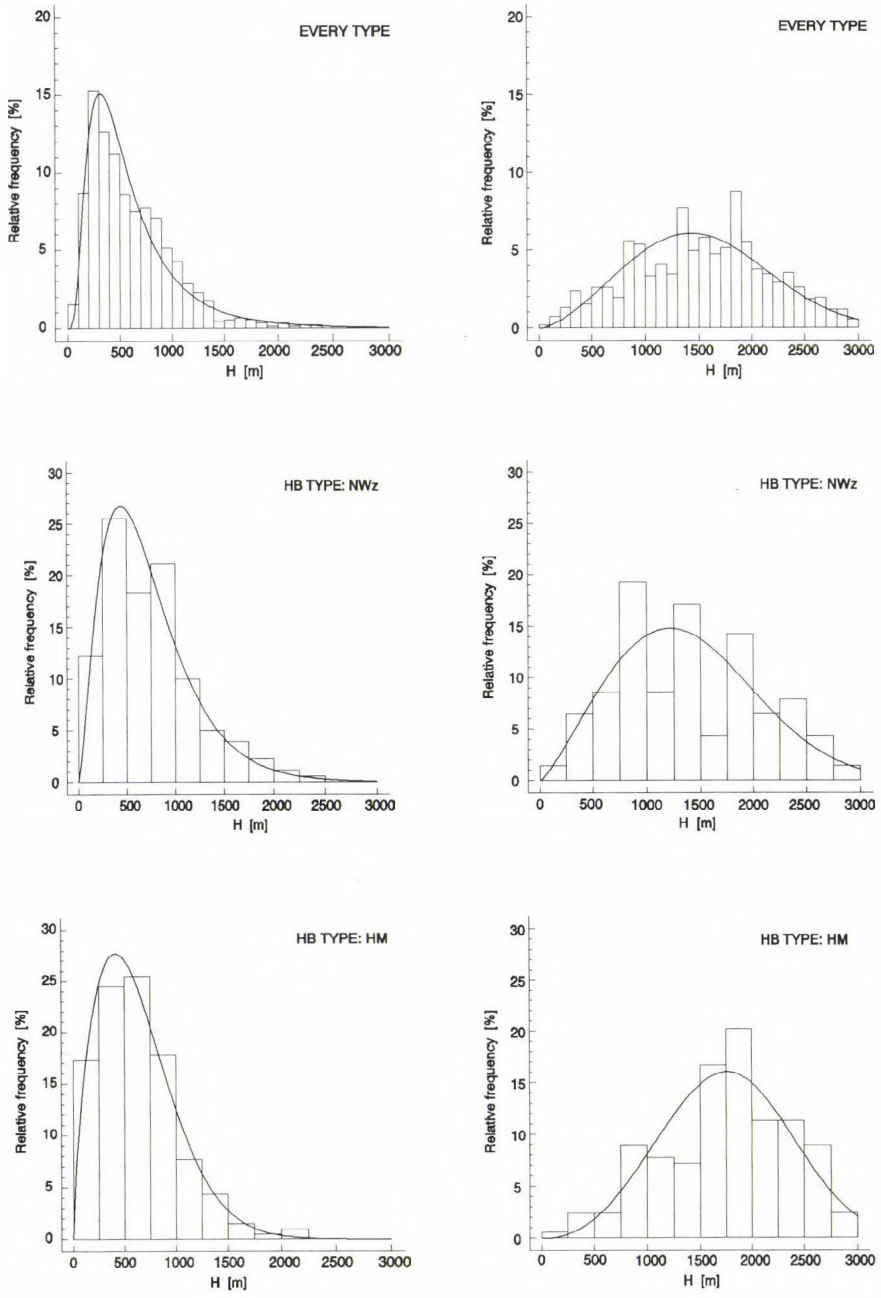


Fig. 1. Probability density of height of PBL in winter (left) and summer (right).

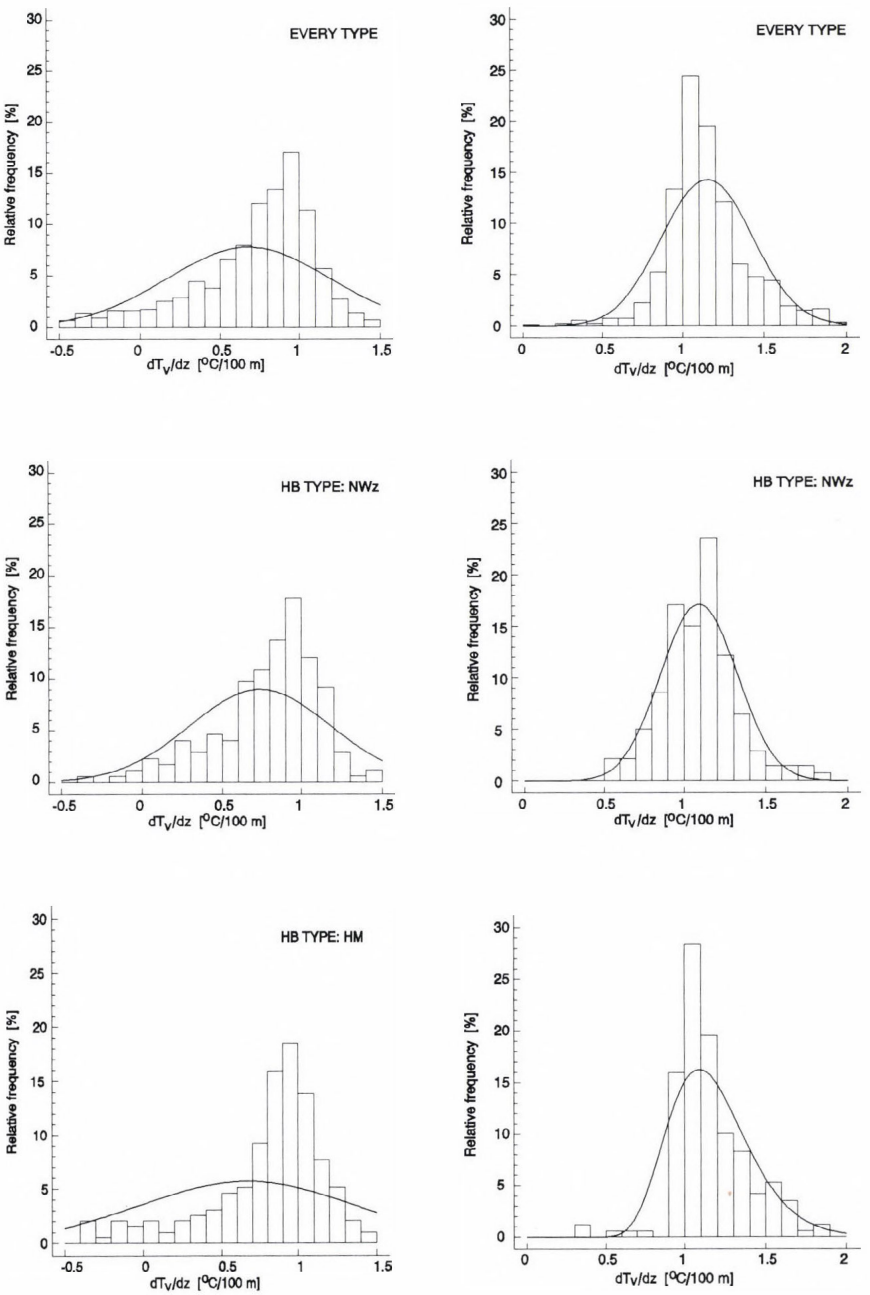


Fig. 2. Probability density of gradient of virtual temperature [°C/100 m] in the lower 300 m in winter (left) and summer (right).

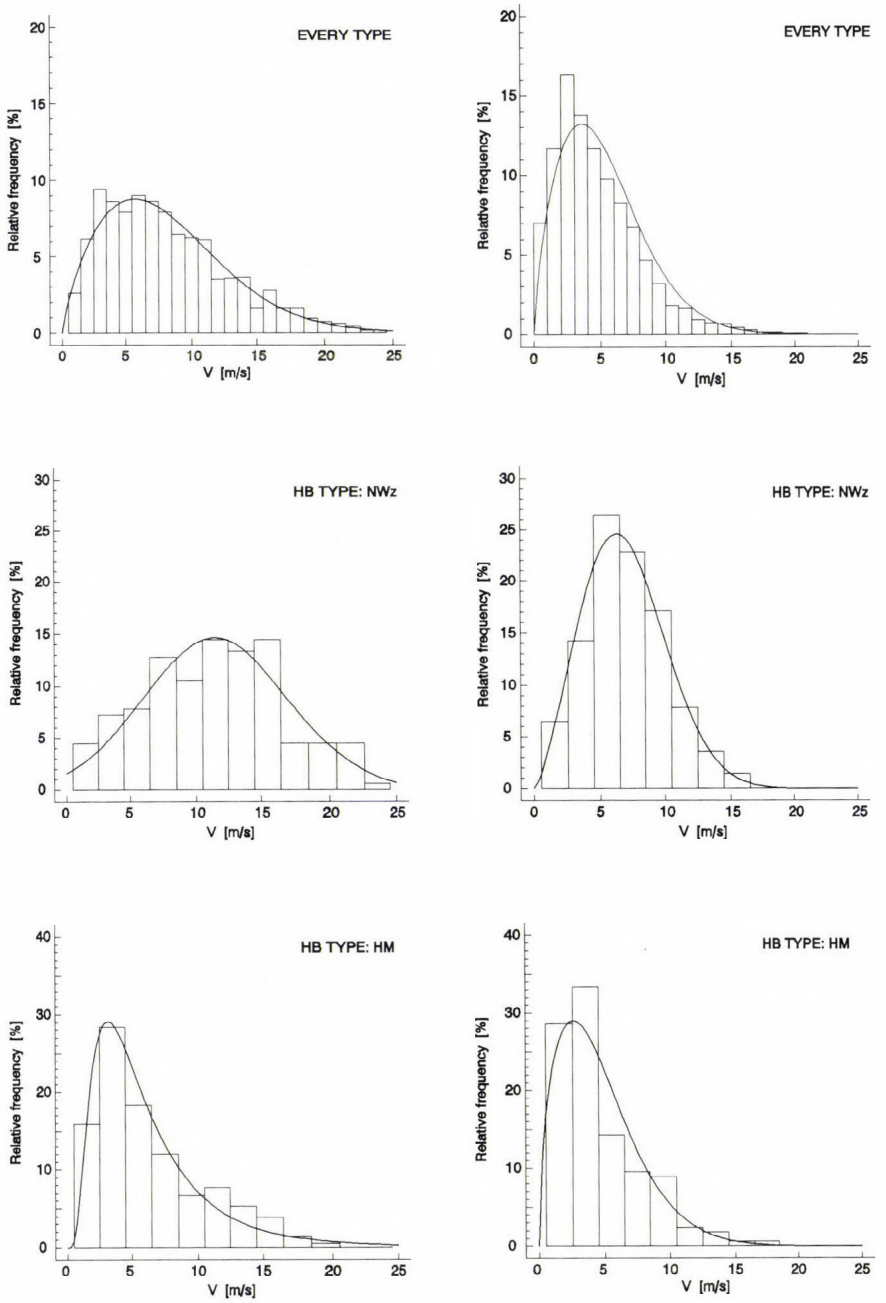


Fig. 3. Probability density of speed and direction of wind at the 925 hPa standard pressure level in winter (left) and summer (right).

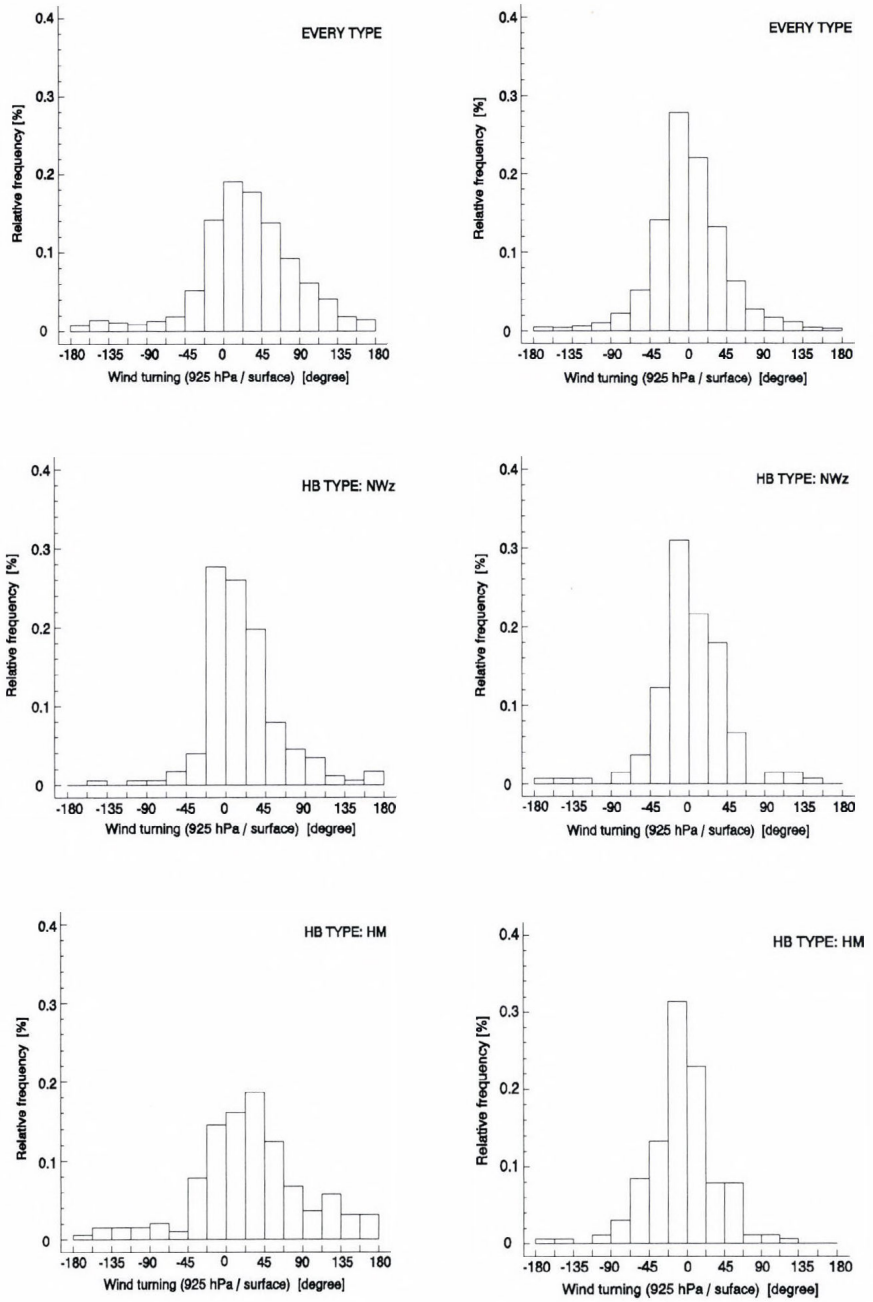


Fig. 4. Histogram of difference between wind directions (wind turning) at 925 hPa and surface levels in winter (left) and summer (right).

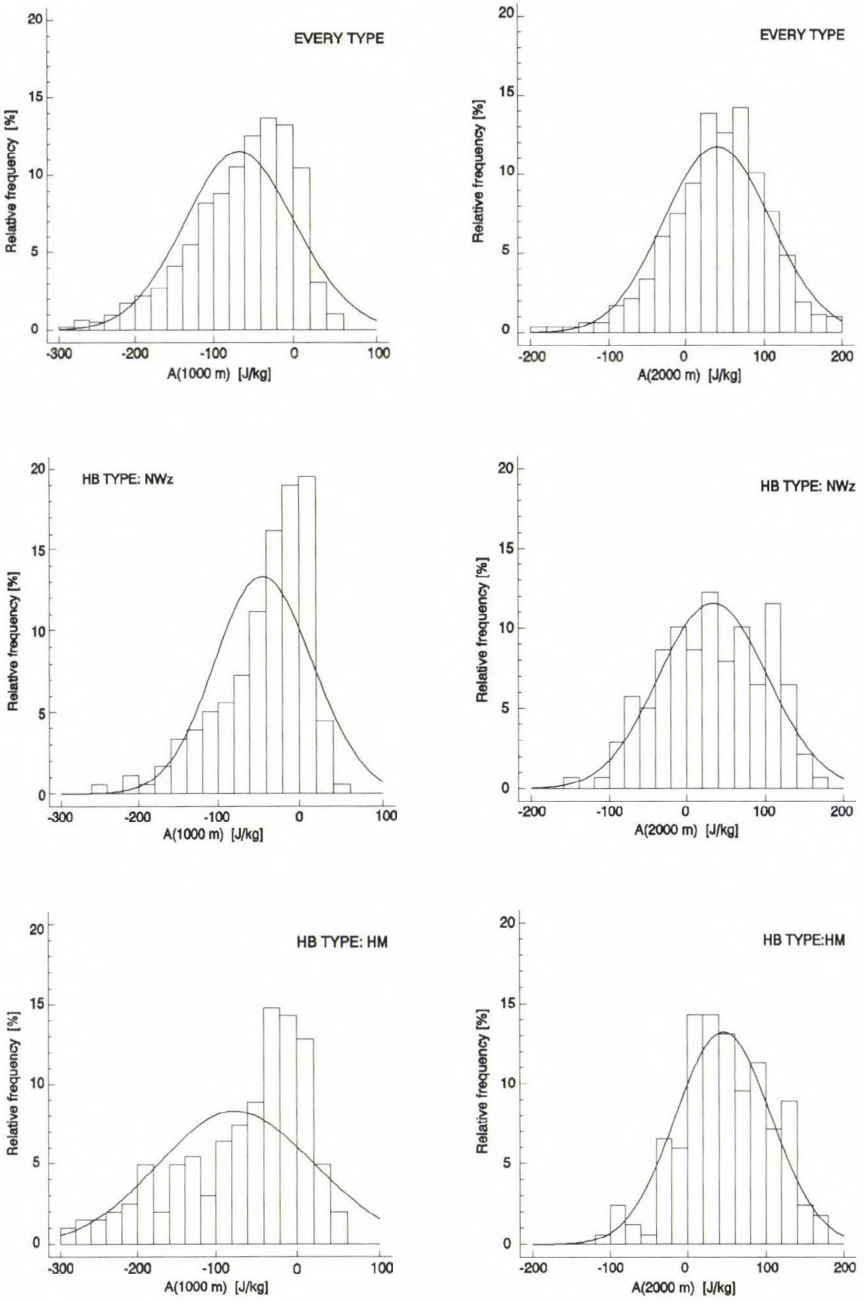


Fig. 5. Probability density of energy of instability in the lower 1km layer in winter (left) and in the lower 2 km layer in summer (right).

3. *Aggregating HB types*

3.1 Methodology

A hierarchical clustering technique has been defined in order to decrease by one the number of macrocirculation types in each step of the procedure. Let K denote the number of PBL characteristics and J denote the actual number of MC types. At the beginning of first step of the entire aggregation procedure the MC types are identical to HB types and J equals to 30. Then two MC types (HB types in the first step) are aggregated according to criteria described below and assigned to a new MC type. The rest of MC types remains unchanged and value of J decreases by one. This process is repeated until a prescribed number of MC types is achieved.

Two types are aggregated if

(a) The distance between these types is minimal.
In order to measure the distance between types i and j in terms of PBL characteristics we define the quantity

$$D_{ij} = \sum_{k=1}^{K} (m_{ik}/\sigma_{ik} - m_{jk}/\sigma_{jk})^2/(m_k/\sigma_k)^2, \qquad i,j = 1,...,J, \tag{2}$$

where subscripts i, j denote the conditioning on PBL parameters and k denotes the conditioning on MC types. m and σ correspond to mean and standard deviation, respectively. The numerator of Eq. (2) represents a Euclidian distance between means corresponding to types i and j with a normalization by their standard deviations. Denominator provides a further normalization in order to remove the differences between magnitudes of PBL parameters.

(b) Probability distributions of these types are similar.
Similarity of distributions is characterized by the chi-square distance

$$Q_{ij} = \sum_{k=1}^{K} n_i n_j \sum_{l=1}^{L} \frac{[(n_{ikl}/n_i) - (n_{jkl}/n_j)]^2}{n_{ikl} + n_{jkl}}, \qquad i,j = 1,...,J, \tag{3}$$

where n_i is the number of days in the MC type i, n_{ikl} is the number of data in the lth interval of histogram of the kth variable under MC type i, and L is the number of intervals (see e.g., *Dévényi* and *Gulyás*, 1988).

(c) Frequency of MC types is relatively uniform.

According to above considerations the aggregation procedure consists of the following steps.

(i) A few elements (5 %) of the matrix of Euclidian distances (Eq. (2)) having smallest distances is selected.

(ii) Three pairs of MC types having smallest values of the chi-square distance (Eq. (3)) are chosen from pairs of types obtained in (*i*).

(iii) Two MC types of that pair are aggregated from the three possibilities obtained in (*ii*) where the number of days in these two types is minimal.

(iv) Repeat steps (*i*)–(*iii*).

An experiment has been performed in order to examine sensitivity of aggregation procedure to changes of criteria (*i*)–(*iii*) and the result has been shown quite robust.

3.2. Results

The optimal number of MC types depends on the task in question and the sample sizes. In the actual case 8–12 MC types seem a good choice (*Fig. 6*). *Tables 3* and *4* show the main results for eight types.

Difference between PBL parameters in winter and summer is apparent. Evidently, winter generally has smaller PBL heights, smaller gradient of virtual temperature, stronger winds and negative energy of instability. A conjunctive use of the distance measure Eq. (2) and the similarity measure Eq. (3) of probability distributions of PBL parameters conditioned on different MC types seems a useful technique to reduce the number of HB types. For instance, the MC types MC1 and MC3 in summer have quite similar means of PBL parameters but with different probability distributions.

The HB types NWz and HM presented before are obviously found in different MC types in the two seasons. In winter the NWz Hess-Brezowsky type belongs to MC4, while the anticyclonic HM type is in MC2. The main characteristics of MC2 (Table 3) are similar to HM type: moderate virtual temperature gradient and wind speed, consequently high negative energy of instability. The mean PBL high is below the average. In summer, the NWz HB type forms alone an MC type (see MC5 in Table 4), and the anticyclonic HM type belongs to MC6.

In an earlier stage of the work (*Matyasovszky et al.*, 1996) the aggregation of HB types was performed without the energy of instability, and the structure of aggregated MC types was highly similar to the present case for each season and each step. For example, only 4 HB types were assigned to MC situation which differs from that of the previous case when using 8 MC types in summer. This is not surprising because the energy of instability is not independent statistically from the virtual temperature gradient and PBL height (see Table 1).

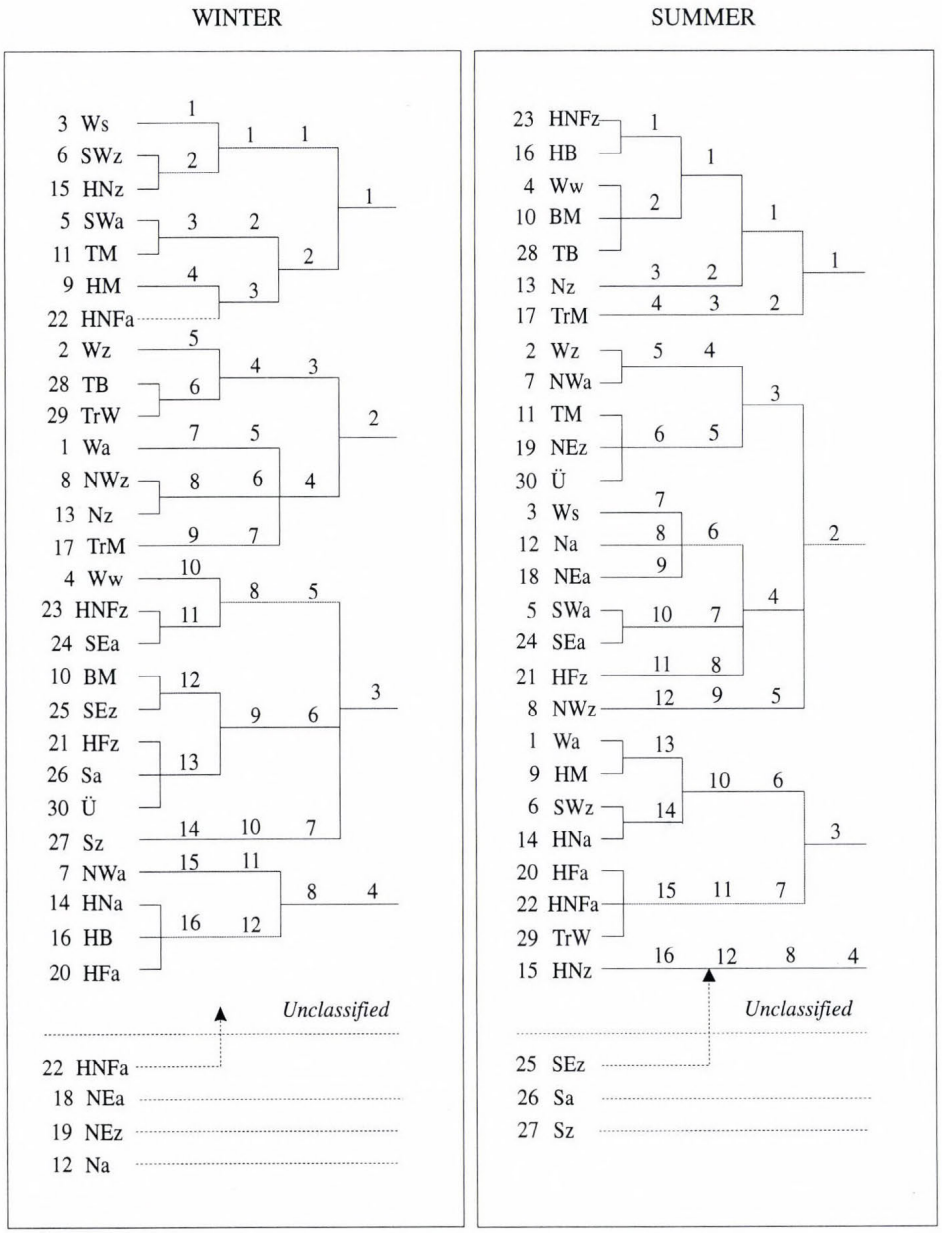


Fig. 6. The system of aggregated HB types for MC types of number of 16, 12, 8, 4. An MC type is unclassified when number of days in the type is smaller than 20.

Table 3. The distance matrix, relative frequency [%] of MC types, and mean of PBL parameters under the MC types in winter. (Numbers in bold indicate types where PBL parameters have different probability distributions at a 90% significance level.)

Type	1	2	3	4	5	6	7	8
1	0.00							
2	**0.18**	0.00						
3	**0.15**	**0.17**	0.00					
4	**0.20**	**0.25**	**0.24**	0.00				
5	**0.35**	**0.46**	**0.44**	**0.23**	0.00			
6	**0.26**	**0.33**	**0.35**	**0.15**	0.15	0.00		
7	0.39	0.49	0.43	**0.32**	**0.27**	**0.29**	0.00	
8	**0.64**	**0.71**	**0.72**	**0.49**	**0.31**	**0.40**	**0.49**	0.00
[%]	11	13	21	18	8	16	2	9
	(unclassified: 2%)							
H	504	585	577	415	559	611	464	719
γ	0.59	0.64	0.56	0.71	0.80	0.74	0.76	0.90
V	8.3	6.6	10.1	9.7	6.3	6.1	8.7	7.1
A	−84	−79	−80	−55	−62	−63	−89	−33

Table 4. The distance matrix, relative frequency [%] of MC types, and mean of PBL parameters under the MC types in summer. (Numbers in bold indicate types where PBL parameters have different probability distributions at a 90% significance level.)

Type	1	2	3	4	5	6	7	8
1	0.00							
2	**0.08**	0.00						
3	**0.17**	**0.18**	0.00					
4	**0.34**	**0.35**	**0.17**	0.00				
5	**0.23**	**0.24**	**0.13**	**0.21**	0.00			
6	**0.22**	**0.23**	**0.07**	**0.17**	**0.18**	0.00		
7	0.08	**0.12**	**0.12**	**0.30**	**0.19**	0.16	0.00	
8	**0.54**	**0.54**	**0.39**	**0.25**	0.33	**0.39**	**0.49**	0.00
[%]	23	5	22	12	5	19	11	2
	(unclassified: 1%)							
H	1361	1247	1396	1415	1296	1539	1410	1372
γ	1.15	1.09	1.15	1.13	1.09	1.18	1.15	1.09
V	5.0	6.6	6.1	5.3	7.0	4.7	4.7	5.5
A	39	11	39	44	34	46	45	39

4. Time series model of PBL characteristics

4.1 The model

To reproduce the simultaneous temporal statistical behavior of PBL parameters, a suitable model should be chosen. Autoregressive (AR) processes represent a well-developed and commonly used tool to model time series. They have been developed principally for Gaussian processes, but PBL parameters do not follow Gaussian distribution. Therefore, it is desirable to construct a transformation establishing a relationship between the distribution of the PBL parameters and the normal distribution. Time series of these hypothetical, normally distributed variables are modeled by AR processes and then the inverse of above mentioned transformation results in a time series model of PBL parameters. The temporal evolution of above mentioned hypothetical, normally distributed variables is modeled by first order autoregressive processes. The order of the model has been determined by Akaike's Information Criteria (*Akaike*, 1974).

As it was shown, the probability distribution is very different for different parameters and under different MC types, and therefore a simple analytical solution of the problem may not be expected. Therefore, one of the non-parametric techniques, the so-called Abramson method (*Abramson,* 1982) has been used to estimate probability distributions of PBL parameters. An important advantage of nonparametric methods is that they do not require assumptions on the shape of distribution functions. Details of these techniques are not discussed here, a climate oriented review of nonparametric probability density estimators can be found in *Matyasovszky* (1997).

4.2 Results

The entire data set is split into two parts and the stochastic model is developed from the first part of data. In order to validate the model, a time series of PBL parameters is simulated using the model, and several statistical characteristics of the second part of observed data and simulated data are compared. The main and standard deviation of PBL parameters are shown in *Table 5*. The stochastic model generally somewhat underestimates means in both seasons as well as standard deviations in summer. In winter, simulated standard deviations are larger than those observed except for wind speed and virtual temperature gradient. The model can reproduce not only simple characteristics like mean but probability distributions too. This is illustrated by *Fig. 7*.

In order to illustrate the accuracy of the stochastic model, another simulation experiment has been performed using the entire data set (without split sampling). In that case no visible difference can be recognized between observed and simulated probability distribution functions, which suggests that disagreements in *Table 6* or Fig. 7 are due to natural variability of PBL parameters, relative frequencies of HB types and data inhomogeneity, and not to model construction.

Table 5. Observed and simulated means and standard deviations of PBL parameters

Mean								
	Winter				Summer			
	H	γ_v	V	A	H	γ_v	V	A
Observed	628.6	0.71	7.9	-63.6	1442.1	1.16	5.4	40.6
Simulated	593.2	0.63	7.8	-71.9	1373.6	1.13	4.7	40.6
Standard deviation								
	Winter				Summer			
	H	γ_v	V	A	H	γ_v	V	A
Observed	390.3	0.47	4.7	66.8	630.6	0.29	3.3	71.1
Simulated	362.7	0.54	4.8	75.6	570.9	0.26	3.2	70.2

Table 6. Observed and simulated correlation matrices of PBL parameters with split sampling

Winter								
	Observed				Simulated			
	H	γ_v	V	A	H	γ_v	V	A
H	1.0	0.24	0.01	0.50	1.0	0.28	0.07	0.60
γ_v		1.0	-0.14	0.76		1.0	0.02	0.74
V			1.0	-0.04			1.0	0.09
A				1.0				1.0
Summer								
	Observed				Simulated			
	H	γ_v	V	A	H	γ_v	V	A
H	1.0	0.13	-0.08	0.28	1.0	0.12	-0.01	0.29
γ_v		1.0	-0.14	0.65		1.0	-0.03	0.64
V			1.0	-0.15			1.0	-0.04
A				1.0				1.0

The stochastic model reproduces the correlation matrices of PBL parameters quite well (Table 6). This is remarkable because the difference between correlations calculated from the entire data set (Table 1) and from the second half of data (Table 6) is not considerably smaller than the difference between correlations calculated from the second part of data and from simulated data.

Autocorrelations of PBL parameters decrease rapidly to zero, only one day lag values are significantly different from zero for both the observed and simulated series. For instance, one day lag autocorrelation of height of PBL was found 0.17 and 0.14 in winter and summer respectively. Corresponding simulated values are 0.17 and 0.16. The second remark is that correlation is not necessarily a best indicator of relationships between PBL parameters since the parameters are not distributed normally.

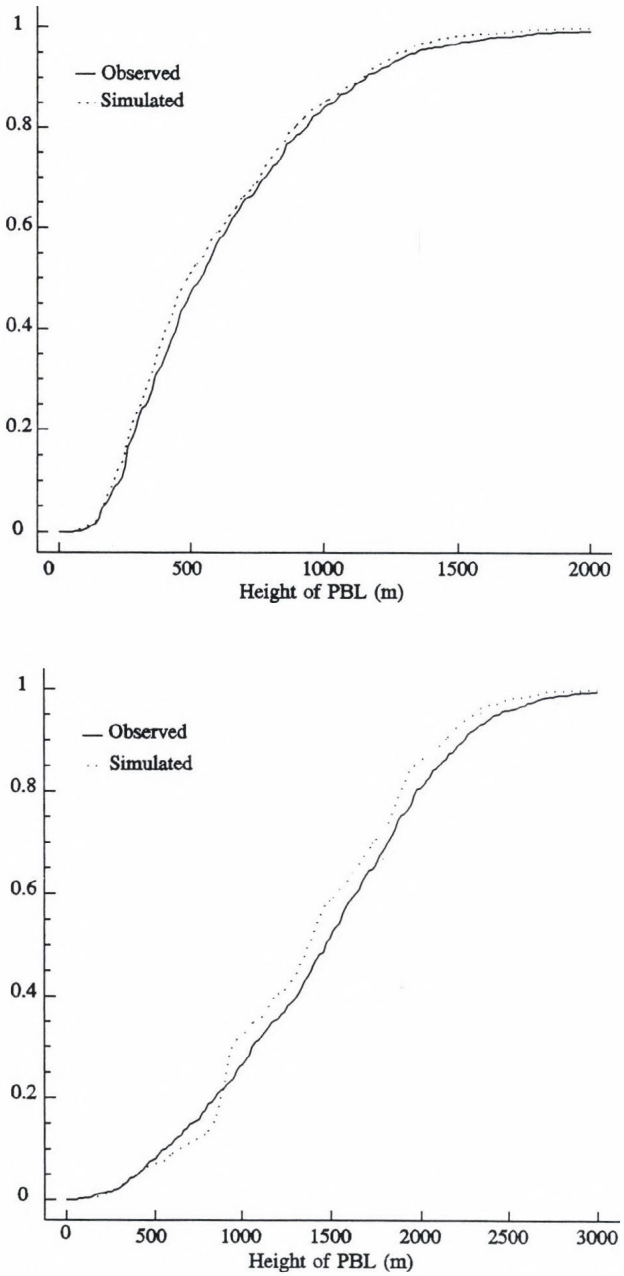


Fig. 7. Observed and simulated probability distribution functions of PBL height in winter (above) and summer (below).

5. *Conclusions*

The potential of the atmosphere to disperse and dilute pollutants depends upon various factors such as wind speed, height of PBL, vertical virtual temperature gradient, energy of instability, etc. Basic statistical characteristics of parameters controlling air pollution potential and their dependence on Hess-Brezowsky macrocirculation types have been presented in winter and summer seasons. The analysis has lead to the following conclusions:

- Statistical properties of PBL are well separated by HB types.
- The most suitable probability distribution for wind speed is frequently the lognormal distribution.
- The gradient of virtual potential temperature and energy of instability cannot be satisfactorily modeled by any well-known probability distribution.
- Calculation of the height of PBL using significant points of virtual temperature profile produces an uncertainty of 10–15 %, when the critical virtual temperature gradient is between 0.75°C/100 m and 0.85°C/100 m under the top of the PBL.
- The PBL height in different HB types can be adequately described by ordinary probability distributions.

A method has been developed in order to obtain a system of aggregated Hess-Brezowsky macrocirculation types in terms of characteristics controlling air pollution potential. Main conclusions can be summarized as follows:

- Correlations among the height of PBL, wind speed at the 925 hPa pressure level, and gradient of virtual temperature of the lower 300 m are quite low (≤ 0.3), thus, each parameter has an information essentially independent from information of the other variables. The energy of instability is statistically not independent from virtual temperature gradient of lower 300 m layer and from PBL height.
- The procedure developed for aggregating HB types is based on a conjunctive use of a distance measure of PBL parameters and a similarity measure of their probability distributions conditioned on MC types. The method provides an optimal choice for the number of days in the system of types.
- The technique resulted in statistically different conditional probability distributions of PBL parameters.

A multivariate time series model has been developed to describe simultaneous stochastic behavior of PBL parameters.

- The model reproduces satisfactorily observed statistical characteristics.
- Therefore, the model is suitable to simulate air pollution potential under hypothetical macrocirculation conditions.

It is intended to further develop and extend the stochastic model by incorporating pollutant concentrations.

Acknowledgements—Research leading to this paper has been supported by grant from Hungarian Science Foundation OTKA F-015493 and by grant from Higher Education Support Program FKFP-0193.

References

Abramson, I.S., 1982: On bandwidth estimation in kernel estimators — A square root law. *Ann. Statist. 10*, 1217-1223.

Akaike, H., 1974: A new look at the Statistical Model Identification. *IEEE Trans. Auto. Control. 19*, 716-723.

Bartholy, J. and *Kaba, M.*, 1987: Meteorological and statistical analyses and correction of the Hess-Brezowsky macrocirculation types (in Hungarian). *Országos Meteorológiai Szolgálat Kisebb Kiadványai, LVII.* Budapest, Hungary.

Bodolai, I., 1983: Main characteristics of the planetary boundary layer (in Hungarian). *Országos Meteorológiai Szolgálat Kisebb Kiadványai, LII.* Budapest, Hungary.

Dévényi, D. and *Gulyás, O.*, 1988: *Methods of Mathematical Statistics in Meteorology* (in Hungarian). Tankönyvkiadó, Budapest, Hungary.

Hess, P. and *Brezowsky, H.*, 1969: Katalog der Groswetterlagen Europas. *Ber. Deutsch. Wetterdienst*, Nr. 113, Bd. 15,2.

Matyasovszky, I., Weidinger, T. and *Németh, P.*, 1996: A system of aggregated Hess-Brezowsky macrocirculation types to characterize air pollution potential over Budapest. *Proceedings of Seventeenth International Conference on Carpathian Meteorology, Visegrád (Hungary) October 14-18, 1996*, 266-271.

Matyasovszky, I., 1997: Estimating probability density functions by kernel techniques. *Időjárás 101*, 17-31.

Stull, R.B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, Dordrecht, Boston and London.

Szepesi, D., Popovics, M., Nárai, K., Iványi, Zs. and *Mersich, I.*, 1977: Meteorological simulation of urban air pollution. Part II. Diffusion climatological bases for the simulation of the transmission (in Hungarian). *Időjárás 81*, 129-146.

Weidinger, T., Matyasovszky, I. and *Németh, P.*, 1996: A statistical analyses of PBL over Budapest using macrocirculation types. *Proceeding of International Conference on Meteorological Processes in the Boundary Layer of Atmosphere, 1996. Stara Lesná, Slovakia, October 7-11, 1996*, 188-193.

Wiswanadham, D.V. and *Santosh, K.R.*, 1989: Air pollution potential over South India. *Boundary-Layer Meteorology 48*, 299-313.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 4, October–December 1998, pp. 239–246

Some aspects of the urban heat island and relative humidity in larger Belgrade area

Miroslava Unkašević

Department of Meteorology, Faculty of Physics, University of Belgrade, Dobračina 16, 11001 Belgrade, Yugoslavia; E-mail: itosic@afrodita.rcub.bg.ac.yu

(Manuscript received 23 March 1998; in final form 24 July 1998)

Abstract—This paper presents the investigation of the urban effects on the monthly mean values of the minimum temperature and relative humidity in the area of Belgrade. The minimum urban temperature increase is explained by linking the temperatures with resident population, a parameter that reflects the size of the city itself. This examination indicates that urban air contains less relative humidity than suburban and rural ones.

Key-words: heat island, relative humidity, Belgrade area.

1. Introduction

Urbanization and industrialization have resulted in the modification of the local climate of many cities. This modification, in general, involves the alteration of the local energy and atmospheric moisture budget.

The existence of the urban heat island is usually explained as the effect of the city itself on the energy balance and the artificial heat added by combustion processes in residential and industrial buildings.

Urban-rural temperature differences are well documented, as is the heat island structure of the build-up areas as a whole (*Preston-Whyte*, 1970; *Nkemdirim*, 1976; *Goldreich*, 1985). *Voogt* and *Oke* (1997) used ground and airborne thermal infrared radiometers to estimate the surface temperature of urban areas. Many authors concluded that few studies make comparison between urban and their rural surroundings relative humidity values (*Hage*, 1975; *Kawamura*, 1985; *Brazel* and *Balling*, 1986; *Ackerman*, 1987; *Lee*, 1991).

Hage (1975) found that in winter the relative and absolute humidity were high in the city at any time because of vertical mixing and combustion sources. *Kawamura* (1985) and *Brazel* and *Balling* (1986) concluded that various

activities and landscapes associated with urbanization can be ultimate force of a kind urban "humidity island" effect. *Ackerman* (1987) suggested that urban vapour pressure and dew points were lower than the rural ones only in the forenoon and spring afternoons.

Lee (1991) concluded that in London the urban atmosphere was more humid than the rural at night in all months and during the day in winter and spring, while during summer days, the urban atmosphere was less humid.

The Belgrade basin with an estimated population of 2.0 million in 1991 covers approximately 3221 km^2 area with the fairly high mountain Avala to the south and lowlands to the north and west. The climate of Belgrade is continental in character. The Meteorological Station of the Belgrade Observatory which is situated at 132 m above mean sea level is the basic station for the analysis of the monthly mean values of minimum temperature and relative humidity.

Strong industrial zone in the Belgrade city is concentrated along coastal rivers Sava and Danube (*Fig. 1*), while commercial district and dense buildings are concentrated around the Belgrade Observatory Station. Residential areas with parks are located on highest parts of the area of Belgrade.

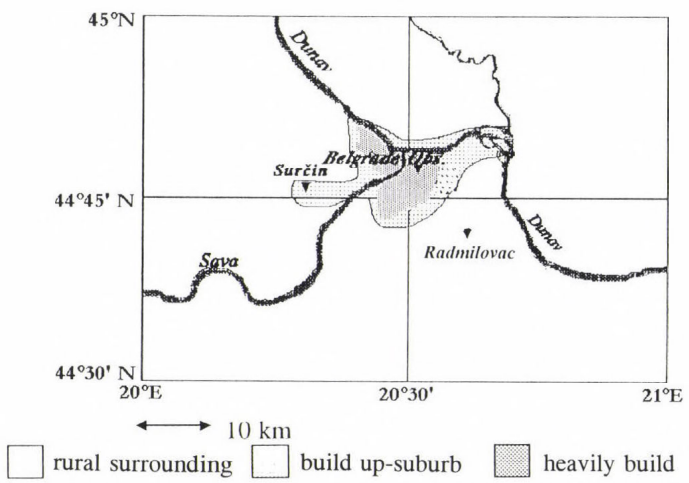


Fig. 1. Map of the Belgrade area showing the next stations:
1. Belgrade Observatory (urban), 2. Surčin Airport (suburban) and 3. Radmilovac (rural).

The daily measurements of the minimum temperature are carried out at the next location: Belgrade Observatory (urban), Surčin Airport (suburban) and Radmilovac (rural) for the period 1970–1991. The monthly mean values of the minimum temperature are determined using the daily values.

For each mean value the standard deviation has been also calculated in order to verify the data according to the ordinary criteria of the statistical distributions, e.g., the t-distribution tables with $n-1$ degree of freedom (since the standard deviation is obtained from $n \geq 30$). Besides, the Kolmogoroff-Smirnoff homogeneity test was carried out and adjustments were made to filter out inhomogeneities due to errors introduced by instruments and observed changes.

Elevations between the Belgrade Observatory (132 m), Surčin Airport (96 m) and Radmilovac (130 m) differ maximally about 36 m, so the possible effect of station elevation on the diurnal temperature range is negligible (*Kuttler et al.*, 1996).

The present report is the result of recent research of the mean magnitude of the urban heat island (1970–1991) and the humidity island (1970–1991 and 1976–1980).

2. The urban heat island

The presence of the heat island appears more strongly in the minimum temperatures than in the mean and maximum temperatures (*Landsberg*, 1970; *Oke*, 1974, 1979; *Colacino* and *Lavagnini*, 1982). It may be explained by physical mechanisms which define the next effects:

(1) Increased counter radiation due to absorption of outgoing long-wave radiation and remission by polluted urban atmosphere.
(2) Greater day-time heat storage due to the thermal properties of urban materials and its nocturnal release.
(3) Anthropogenic heat from building sides and fabrics.
(4) Decreased evaporation due to the removal of vegetation.
(5) Decreased loss of sensible heat due to the reduction of wind speed in the city.

Anthropogenic heat from building sides have a maximum in the winter season. In summer it would seem possible that effects of 1, 2 and to lesser extent 3, 4 and 5 may combine to make the city store the sensible heat by day and hence keep urban temperature higher than in the countryside (*Grimmond* and *Oke*, 1995).

It is obvious from *Fig. 2* that the mean urban minimum temperature is always higher than the suburban (Fig. 2a) and rural (Fig. 2b) ones. The maximum frequency distribution of ΔT_m (monthly mean minimum temperature difference between the urban and suburban/rural stations) are related to the temperature intervals from 1.5 to 2.5°C for the suburban (Fig. 2a) and from 2.5 to 3.5°C for the rural station (Fig. 2b).

Despite the fact that the urban heat island morphology is strongly controlled by the unique character of each city, it seems reasonable to suggest the mean magnitude of the heat island is at least 2–3°C.

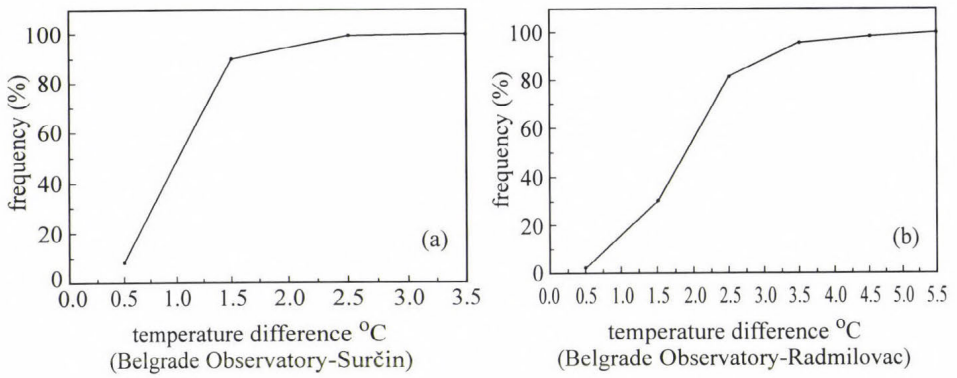


Fig. 2. Cumulative frequency distribution of the monthly mean minimum: (*a*) urban-suburban and (*b*) urban-rural temperature difference during the period 1970–1991.

The examined period (1970-1991) corresponds to the period of maximum development of the city and so it is particularly suitable for checking the validity of some formulae relating temperature variations to urban development (*Landsberg*, 1981). The minimum urban temperature increase is usually explained by linking the temperatures with resident population, a parameter that reflects the size of the city itself. For Belgrade, the statistical correlation analysis was carried out using annual data referring to the resident population (P) from 1890 until 1990, together with the corresponding minimum temperatures T_m measured at Belgrade Observatory (*Fig. 3*). The obtained result is generally similar to the result obtained by *Katsoulis* (1987), who examined long term climatic change of air temperature in Athens. Regression was carried out using the next relation:

$$T_m = 8.15 + 0.03 \log P. \tag{1}$$

This form of equation does not allow the identification of critical wind speed at which urban heat island obliterated. If winds are included, the relation between the maximum heat island (ΔT_{u-r}) and the population is given by (*Oke*, 1974):

$$\Delta T_{u-r} \approx P^{1/4}/4\,\bar{u}^{1/2} \tag{2}$$

where $\bar{u}$ is regional (non-urban) mean wind speed at a height of 10 m. Based on observations in the vicinity of Belgrade (2 million inhabitants) it appears that

this value is approximately 10 m s^{-1} in the case of "Koshava" wind (the most frequent, moderate to strong wind in greater Belgrade area).

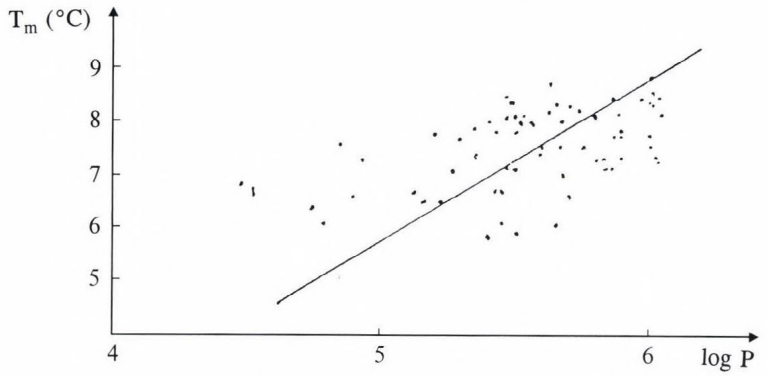


Fig. 3. Correlation between annual mean minimum temperature and log *P* (*P* is population number) for Belgrade Observatory during 1890-1990.

3. Some investigation of the relative humidity in Belgrade

In Belgrade, winter air is moist and summer air is dry, although in winter air contains less water vapour than in summer. *Table 1* presents the hourly mean (06, 13 and 20 UTC) and the monthly mean values of the relative humidity (%) measured at Belgrade Observatory during the period of 1970–1991. The monthly mean values of the relative humidity in Table 1 shows that the relative humidity decreases from winter toward summer months. Slight increase of the relative humidity recorded in June is caused by maximum precipitation which occurring in Belgrade in June.

Table 1. Averaged hourly and monthly mean values of the relative humidity (%) in Belgrade during the period 1970-1991

UTC	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
6	80	79	77	71	73	75	73	74	77	82	84	79
13	70	64	55	49	52	52	48	46	50	55	68	76
20	78	76	69	66	70	71	68	69	74	76	70	80
Mean	**76**	**73**	**67**	**63**	**65**	**66**	**63**	**63**	**67**	**72**	**78**	**78**

Table 1 also shows that although the minimum mean monthly temperature is recorded in January, the maximum of the relative humidity occurs in December.

Daily variations of the relative humidity in the summer months are about twice bigger than in the winter months. Because of these changes the air in Belgrade is frequently dry on summer days.

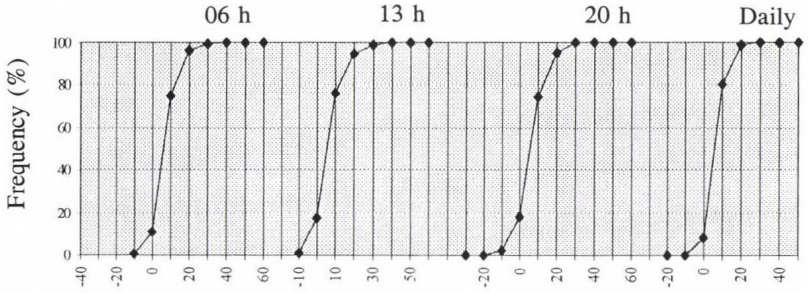
The averaged seasonal values of the relative humidity (RH) measured at Belgrade Observatory during the period 1970–1991 are:

	Winter	Spring	Summer	Autumn
Seasonal value RH (%)	76	66	64	74
Seasonal variance RH (%)	14	24	27	24

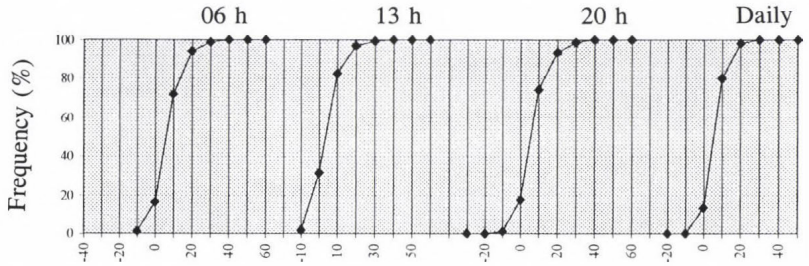
From these values of the relative humidity we can conclude that the relative humidity in autumn is 8% more than in spring, although the averaged seasonal temperature in autumn is higher than in spring. However, the averaged seasonal variances of the relative humidity are the same during spring and autumn.

The Student's t-test indicates that December, January and February have been characterized by a sharp decrease in the relative humidity levels in the period 1970–1991 according to the period 1925–1969 (*Unkašević*, 1996). This analysis suggests that such large drops in the relative humidity during the cold season must be associated with the urban heat island. This is a reflection of the strong inverse relationship of relative humidity to temperature, and thus of the relationship between ΔRH_{u-r} and ΔT_{u-r}. ΔRH_{u-r} is related to the relative humidity differences between the urban and rural sites.

In this study, suburban/rural-urban relative humidity differences at fixed hours (06, 13 and 20 UTC) during the period 1976–1980 have been used. It is obvious from *Fig. 4* that the urban air containes less relative humidity than suburban and rural air. The frequency distributions for suburban/urban-rural relative humidity differences are shifted to positive values, i.e., toward higher relative humidities in the suburban and rural areas. The reverse is true in less than 11% and 17% of the measurements at 06 UTC at the suburban and rural stations, respectively. Measurements of the relative humidity at 13 UTC indicate that this difference has an almost unchanged frequency in suburban and increased frequency in rural areas. Namely, the frequency distribution for ΔRH is shifted to negative values in less than 18% and more than 32% at 13 UTC at the suburban and rural stations, respectively. The frequency distribution at 20 UTC and on the daily basis have many of the same features as at 06 UTC.



(a) Relative humidity difference (Surčin–Belgrade Observatory)



(b) Relative humidity difference (Radmilovac–Belgrade Observatory)

Fig. 4. Cumulative frequency distribution of the hourly mean (06, 13, 20 UTC) and daily mean: (*a*) suburban-urban and (*b*) rural-urban relative humidity differences (ΔRH %) during the period 1976–1980.

4. Conclusions

Increased heating activities in the Belgrade city (more use of central heating, air conditioning, release of energy by industrial activity and traffic) substantially raised the air temperature above the city. It was shown that the monthly mean values of minimum urban temperature is always higher than the suburban and rural ones. Effect of the heat island is explained by a 2–3 °C increase of the mean minimum temperature.

Regression analysis results obtained for the correlation between the yearly mean minimum urban temperature and population number in Belgrade follow the logarithmic model.

The relative humidity was usually lower in the city than in the surrounding areas and the urban-rural relative humidity difference were strongly affected by the urban heat island.

References

Ackerman, B., 1987: Climatology of Chicago area urban-rural differences in humidity. *J. Clim. Appl. Meteor. 26*, 427-730.

Brazel, W.S. and *Balling, R.J.*, 1986: Temporal analysis of long term atmospheric moisture levels in Phoenix, Arizona. *J. Clim. Appl. Meteor. 25*, 112-117.

Colacino, M. and *Lavagnini, A.*, 1982: Evidence of the urban heat island in Rome by climatological analysis. *Arch. Meteor. Geophys. Bioklim.*, B21, 87-97.

Goldreich, Y., 1985: The structure of the groundlevel heat island in a central business district. *J. Clim. Appl. Meteor. 24*, 1237-1244.

Grimmond, C.S. and *Oke, T.R.*, 1995: Comparison of heat fluxes from summer time observations in the suburbs of four north American cities. *J. Appl. Meteor. 34*, 873-889.

Hage, K.D., 1975: Urban-rural humidity differences. *J. Appl. Meteor. 14*, 1277-1283.

Katsoulis, B., 1987: Indication of change of climate from the analysis of air temperature time series. *Climatic Change 10*, 67-79.

Kawamura, T., 1985: Urban-rural humidity differences in Japan. *Geop. Review Japan 58*, 83-94.

Kuttler, W., Barlay, A. and *Rosmann, F.*, 1996: Study of the thermal structure of a town in a narrow valley. *Atmosph Environment 30*, 365-378.

Landsberg, H.E., 1970: Climates and urban planning. Urban climates. *Techn. Paper 141, WMO*, Geneva, 364-374.

Landsberg, H.E., 1981: *The Urban Climate*. Int. Geophys. Ser. 28, Academic Press, 261 pp.

Lee, D.O., 1991: Urban-rural humidity differences in London. *Int. J. Climatol. 11*, 577-582.

Nkemdirim, L.C., 1976: Dynamics of an urban temperature field-A case study. *J. Appl. Meteor. 15*, 818-828.

Oke, T.R., 1974: Review of urban climatology. *WMO, Techn. Note 134*, WMO, Geneva, 132 pp.

Oke, T.R., 1979: Review of urban climatology. *WMO, Techn. Note 169*, WMO, Geneva, 114-119.

Preston-Whyte, R.A., 1970: A spatial model for an urban heat island. *J. Appl. Meteor. 9*, 571-573.

Unkašević, M., 1996: Analysis of atmospheric moisture in Belgrade, Yugoslavia. *Meteor. Zeit. 3*, 121-124.

Voogt, J.A. and *Oke, T.R.*, 1997: Complete urban surface temperatures. *J. Appl. Meteor. 36*, 1117-1131.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 102, No. 4, October–December 1998, pp. 247–257

Radiative cooling rate in the atmospheric boundary layer

František Smolen and **Marian Ostrožlík**

Geophysical Institute of Slovak Academy of Sciences
Dúbravská cesta 9, 842 28 Bratislava, Slovak Republic
E-mail: geofostr@savba.sk

(Manuscript received 8 September 1998)

Abstract—The long-wave radiation flux density measurements at the meteorological observatories Skalnaté Pleso (φ = 49°12'N, λ = 20°14'E, H = 1778 m a.s.l.) and Stará Lesná (φ = 49°09'N, λ = 20°17'E, H = 810 m a.s.l.) in High Tatras during the period 1991–1995 were used to study the radiative temperature changes in the atmospheric boundary layer.

Based on the changes of the long-wave radiation balance with altitude in the atmospheric layer between Stará Lesná and Skalnaté Pleso, the fluctuations of the radiative cooling rates were calculated. The effects of the thermal atmospheric stratification and the water vapor content on the radiative cooling or heating rates were determined. A special attention was paid to the daily and seasonal courses, as well as to the effect of clouds.

Key-words: radiative temperature changes, radiative cooling, radiative heating, radiative cooling rate, long-wave radiation balance, long-wave radiation flux density, thermal atmospheric stratification, atmospheric boundary layer.

1. Introduction

The atmosphere losses energy by emission of infrared radiation toward the space and toward the surface. Except for ozone, the primary long-wave radiation absorbing gases cool the atmosphere. The water vapor dominates the long-wave cooling in the troposphere. The radiative cooling rate depends in a large extent upon the cloudiness, the thermal stratification as well as on the water vapor content in the atmosphere. The concentration of other greenhouse gases in the atmosphere plays an important role, too. *Timanovskaja* and *Faraponova* (1967) have shown that the radiative cooling rate can be about 5 K h^{-1} during night-time as a consequence of the long-wave radiation flux divergence in the atmospheric ground layer, and on the other hand, during the

daylight time the radiative heating can lead to rate of 6 K h^{-1} by the long-wave radiation absorption.

To study the radiative temperature changes of the atmosphere, both the solar energy transfer and the long-wave radiation transfer should be taken into account. Comparison of the results obtained by the different calculation methods has shown that the difference between them can be up to 50 per cent in some cases (*Goody*, 1964).

The purpose of this paper is to find empirical dependence of radiative cooling on cloudiness as well as on thermal atmospheric stratification. A special attention is paid to study the effect of the low clouds on the radiative cooling rate in the boundary layer of atmosphere.

2. *Material and methods*

Long-wave radiation flux measurements, which have been used as input parameters at the study of radiative cooling rate, were carried out at the mentioned observatories during the period 1991–1995. The suitable positions of these two close sites with different altitude (the difference is about 1000 m) enabled us to study the vertical and seasonal variations of radiative cooling and heating in the boundary layer of atmosphere.

To measure the long-wave radiation fluxes, Schulze's radiation balance meters were used. They were located at 1.5 m height over the active surface. Simultaneously with the long-wave radiation flux measurements, the air temperature, air humidity and air pressure were measured, and the amount and type of clouds were observed (*Ostrožlík* and *Janičkovičová,* 1992–1995). To calculate the radiative cooling rate in the atmospheric boundary layer, the vertical changes of long-wave radiation balance $\Delta F/\Delta z$ were used.

On the basis of the long-wave radiation balance differences (ΔF) between the levels Stará Lesná and Skalnaté Pleso the radiative cooling rate $(\delta T/\delta t)_{\Delta F}$ can be calculated. To calculate this rate we used the following relationship

$$\left(\frac{\delta T}{\delta t}\right)_{\Delta F} = -\frac{1}{c_p \rho}\,\frac{\Delta F}{\Delta z}, \tag{1}$$

where T is air temperature, t is time, z is altitude, c_p is the specific heat capacity at constant pressure, ρ is the air density and $\Delta F/\Delta z \cong \text{div } F$.

3. Results

3.1 Time variability of net flux

Figs. 1 and *2* represent the daily and seasonal variability of the long-wave radiation balance at the observatories Stará Lesná and Skalnaté Pleso. Basic characteristics of the curves in these figures show good coincidence with the results of many authors (*Kondratyev et al.*, 1972; *Prokofjev* and *Ter-Markarjanc*, 1972; *Rodgers*, 1967; *Stephens et al.*, 1994; *Zajceva* and *Šljachov*, 1972).

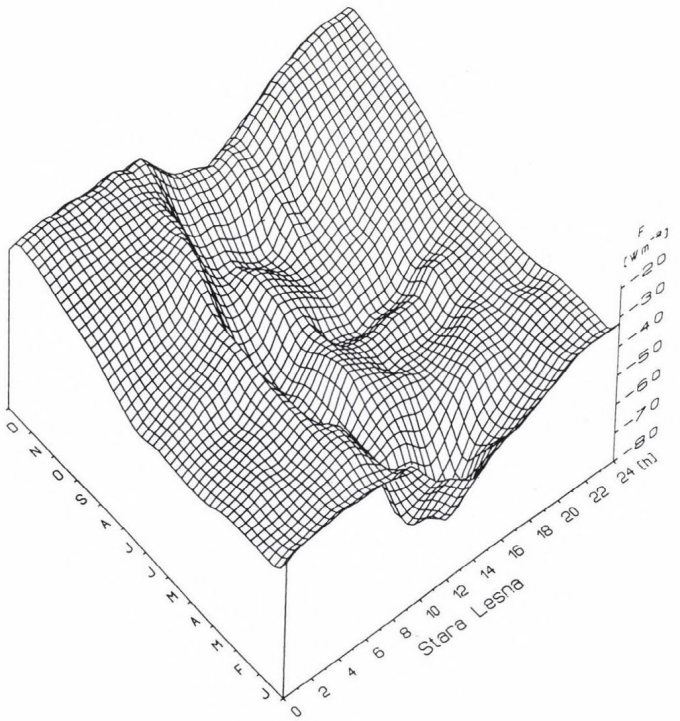


Fig. 1. Daily and seasonal variations of the long-wave radiation balance (F) in W m^{-2} at Stará Lesná during the period 1991–1995.

The net terrestrial radiation increases with altitude. Numerical results show that while the absolute value of F is 41.68 W m^{-2} at Stará Lesná, its value increases at Skalnaté Pleso up to 58.14 W m^{-2} in annual average. At clear sky (amount of clouds is less than 2/10 at both places) the corresponding values are 66.60 W m^{-2} at Stará Lesná, and 95.56 W m^{-2} at Skalnaté Pleso. The greatest

values of F occur at clear sky in March when the mean daily value of the long-wave radiation balance is 82.50 W m^{-2} at Stará Lesná and 117.23 W m^{-2} at Skalnaté Pleso. Based on these results we can state that the actual thermal energy loss owing to the long-wave radiation fluxes in the atmosphere is in average about 30 per cent higher at Skalnaté Pleso than at Stará Lesná.

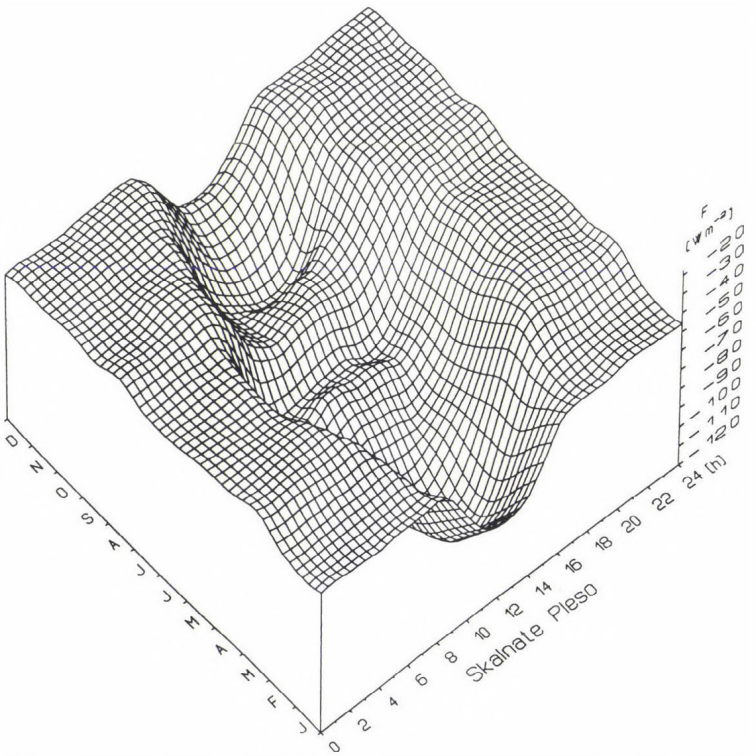


Fig. 2. Daily and seasonal variations of the long-wave radiation balance (F) in W m^{-2} at Skalnaté Pleso during the period 1991–1995.

Some anomalies in vertical changes of long-wave radiation balance during the day can be seen from February to October in evening hours after sunset. In this part of day the absolute values of F at Skalnaté Pleso are smaller than at Stará Lesná. This anomaly is caused by the fact that the downward atmospheric radiation changes slower than the radiation emitted by the active surface.

The actual thermal energy loss in the atmospheric layer between Stará Lesná and Skalnaté Pleso, expressed as a difference of the long-wave radiation balance ΔF, is illustrated on *Fig. 3*.

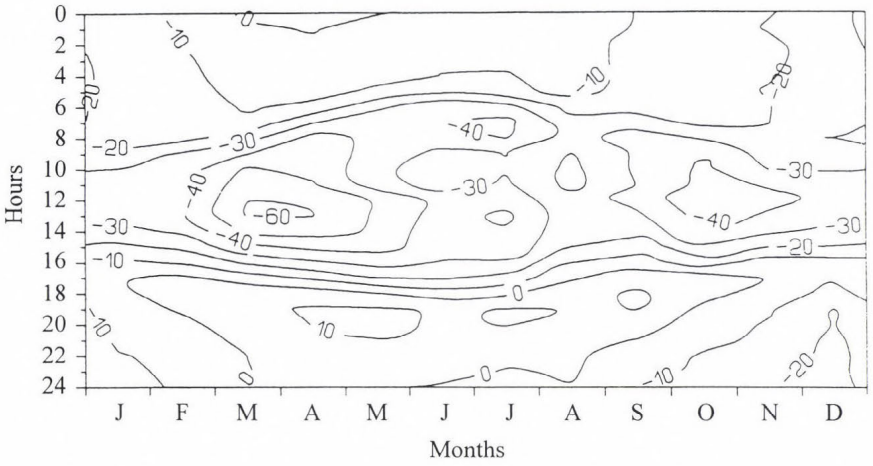


Fig. 3. Thermal energy loss in the atmospheric layer between Stará Lesná and Skalnaté Pleso as a difference of the long-wave radiation balance (ΔF) in W m^{-2} during the period 1991–1995.

3.2 Radiative temperature changes

Radiative temperature change can be expressed by the following relation: if $\Delta F/\Delta z < 0$ then $(\delta T/\delta t)_{\Delta F} > 0$, on the other hand if $\Delta F/\Delta z > 0$ then $(\delta T/\delta t)_{\Delta F} < 0$. In the first case radiative cooling, in the second one radiative heating occurs. When $\Delta F/\Delta z = 0$ then radiation equilibrium exists (*Kondratyev*, 1956).

Time variability of the radiative temperature changes $(\delta T/\delta t)_{\Delta F}$ in the atmospheric layer between Stará Lesná and Skalnaté Pleso under average cloud condition is illustrated on *Fig. 4*. The corresponding values of $(\delta T/\delta t)_{\Delta F}$ at clear sky (cloud amount is less than 2/10 at both localities) are presented on *Fig. 5*. From the comparison of the individual data we can see that the occurrence of clouds in the atmosphere causes a decrease of the radiative cooling. As average, in 43% of the days clouds occur in the layer between Stará Lesná and Skalnaté Pleso. *Stephens et al.* (1994) introduce that the clouds in relatively moist regions decrease the radiative cooling almost to the half of the clear sky values. The most intensive radiative cooling in the investigated atmospheric boundary layer can be seen in December. At average cloudiness the daily mean of $(\delta T/\delta t)_{\Delta F}$ is 0.097 K h^{-1}, at clear sky it is 0.167 K h^{-1}.

In a coincidence with results of *Prokofjev* and *Ter-Markarjanc* (1972) the most intensive radiative cooling occurs at the midday. In this part of day in March and April the mean values of $(\delta T/\delta t)_{\Delta F}$ vary between 0.237 K h^{-1} and 0.246 K h^{-1} at mean cloud conditions and in the range of 0.351 K h^{-1} and 0.583 K h^{-1} at clear sky.

In the evening hours from February to October the values of $\Delta F/\Delta z$ in the layer between Stará Lesná and Skalnaté Pleso are greater than zero. The radiative heating of the atmospheric boundary layer occurs as a consequence of long-wave radiation absorption. This effect is more expressive at mean cloud conditions than at clear sky.

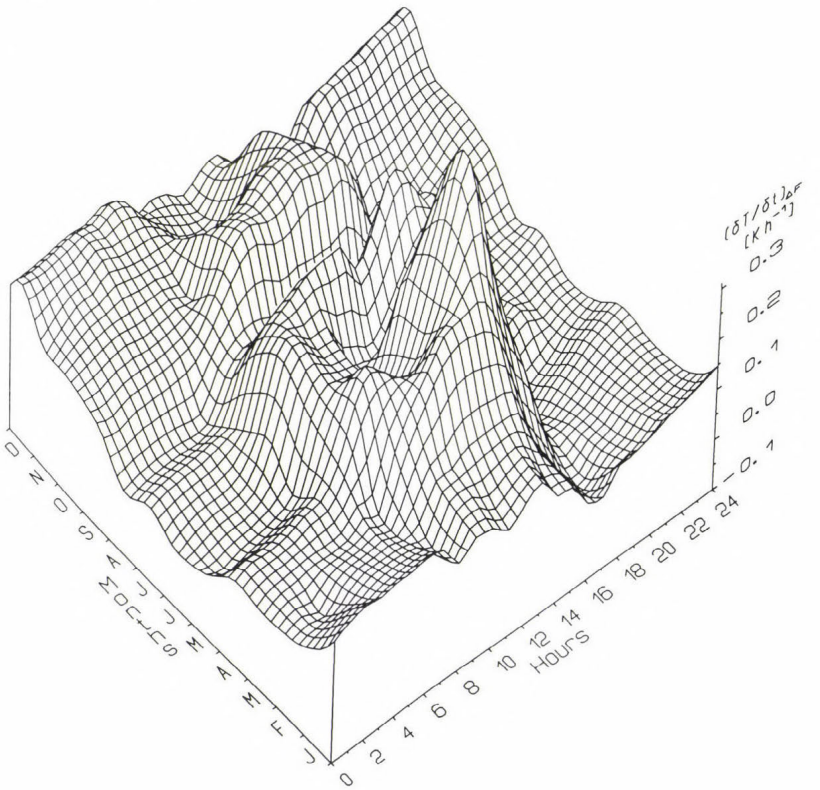


Fig. 4. Daily and seasonal variations of the radiative temperature changes $(\delta T/\delta t)_{\Delta F}$ in K h^{-1} at the average cloud conditions in the atmospheric layer between Stará Lesná and Skalnaté Pleso during the period 1991–1995.

More detailed analysis of the daily course of radiative cooling confirmed that the radiative cooling rate depends in a substantial measure on the thermal stratification of the investigated atmospheric layer. This dependence is clearly illustrated on *Fig. 6.* As we can see from Fig. 6 the values of $(\delta T/\delta t)_{\Delta F}$ rise with the increasing atmospheric stability. The greatest radiative cooling rate corresponds to strong thermal inversion. On the other hand at unstable atmospheric stratification when the vertical air temperature gradient is in the range of 1.01–1.60°C/100 m, radiative heating occurs in the atmospheric layer

between Stará Lesná and Skalnaté Pleso at both average cloud coverage and overcast sky. Radiative equilibrium occurs when the thermal gradient is in the range of 1.01–1.20°C/100 m. The greatest fluctuation of $(\delta T/\delta t)_{\Delta F}$ exists at the unstable atmospheric stratification. This fact can be caused by the convective as well as the turbulent transfer of water vapor.

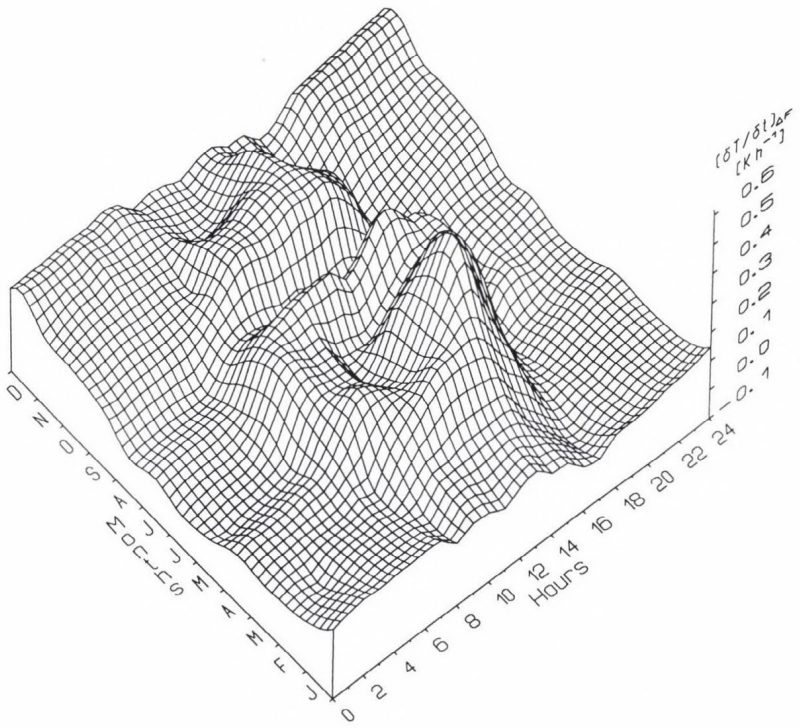


Fig. 5. Daily and seasonal variations of the radiative temperature changes $(\delta T/\delta t)_{\Delta F}$ in K h^{-1} at the clear sky in the atmospheric layer between Stará Lesná and Skalnaté Pleso during the period 1991–1995.

Functional dependence of the radiative cooling rate on the vertical change of air temperature *(dT/dz)* in the observed layer can be expressed by polynomial of degree 3

$$\left(\frac{\delta T}{\delta t}\right)_{\Delta F} = a_0 + a_1\frac{dT}{dz} + a_2\left(\frac{dT}{dz}\right)^2 + a_3\left(\frac{dT}{dz}\right)^3. \qquad (2)$$

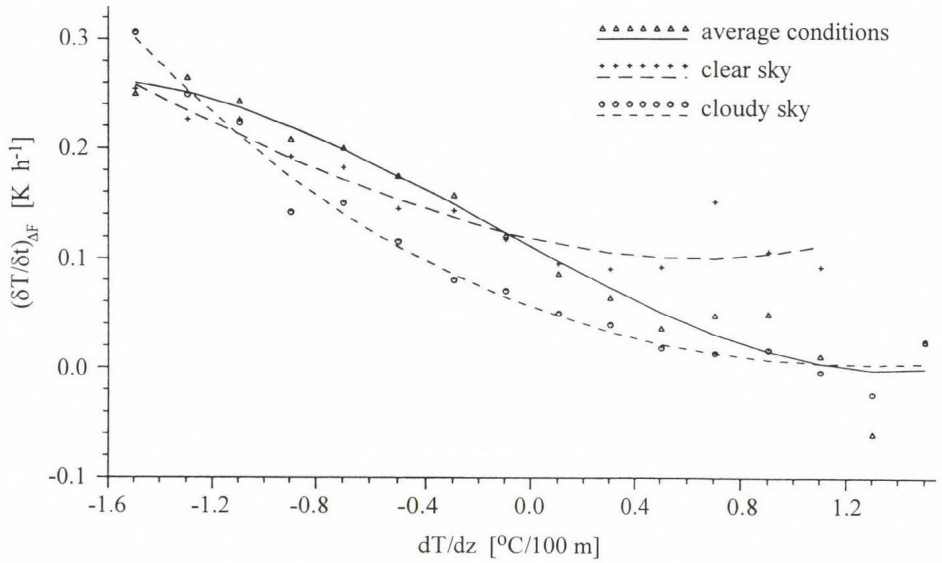


Fig. 6. Dependence of the radiative temperature changes $(\delta T/\delta t)_{\Delta F}$ in K h^{-1} on the thermal atmospheric stratification (dT/dz) in °C/100 m at the different cloud conditions in the atmospheric layer between Stará Lesná and Skalnaté Pleso during the period 1991–1995. The curves correspond to polynomial of the third degree.

The values of the regression coefficients a_0, a_1, a_2, a_3 are listed in *Table 1*.

Table 1. Coefficients a_0, a_1, a_2, a_3 of Eq. (2), and correlation coefficient (r) between experimental and calculated values

Cloud conditions	Coefficients				
	a_0	a_1	a_2	a_3	r
Clear sky	0.1179	–0.0538	0.0362	0.0064	0.9491
Overcast sky	0.0554	–0.0888	0.0433	–0.0047	0.9907
Average conditions	0.1106	–0.1287	0.0084	0.0185	0.9777

The dependence of the radiative cooling rate $(\delta T/\delta t)_{\Delta F}$ on the water vapor pressure (e) in the layer between Stará Lesná and Skalnaté Pleso is presented in graphical form on *Fig. 7*. From the course of the curves on this figure we can see that in average the radiative cooling rate decreases with the rise of the water vapor pressure. There is a substantial, more complicated situation when the sky is clear (curve 2 on Fig. 7). This dependence is characterized by two

maxima. The expressive increase of the quantity $(\delta T/\delta t)_{\Delta F}$ occurs at the water vapor pressure from 0.1 to 4.0 hPa. Later, at the further rise of the water vapor pressure till to 8 hPa, the values of $(\delta T/\delta t)_{\Delta F}$ begin suddenly decrease. Over this value we can see the slow linear rise of the radiative cooling rate and then its decrease in the atmospheric boundary layer. This dependence of $(\delta T/\delta t)_{\Delta F}$ on the water vapor pressure at the clear sky indicates the complication of this problem, which will require a more detailed analysis.

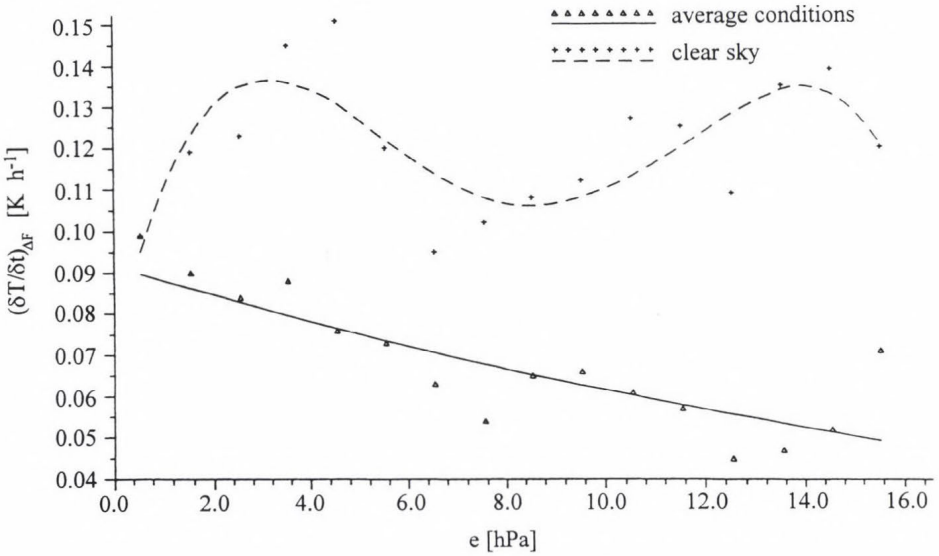


Fig. 7. Dependence of the radiative temperature changes $(\delta T/\delta t)_{\Delta F}$ in K h^{-1} on the water vapor pressure (e) in hPa at the average conditions (exponential curve) as well as at the clear sky (polynomial of degree 4) in the atmospheric layer between Stará Lesná and Skalnaté Pleso during the period 1991–1995.

Analytic expression of the radiative cooling rate dependence on the water vapor pressure at average conditions can be put down in the exponential function

$$\left(\frac{\delta T}{\delta t}\right)_{\Delta F} = 0.0916 \exp(-0.0399 e), \tag{3}$$

with correlation coefficient $r = 0.8527$ between measured and calculated values, while at the clear sky the polynomial of degree 4

$$\left(\frac{\delta T}{\delta t}\right)_{\Delta F} = a_0 + a_1 e + a_2 (e)^2 + a_3 (e)^3 + a_4 (e)^4. \quad (4)$$

The values of the regression coefficients and correlation coefficient are given in *Table 2*.

Table 2. Coefficients a_0, a_1, a_2, a_3, a_4 of Eq. (4), and correlation coefficient (r) between experimental and calculated values

Cloud conditions	Coefficients					
	a_0	a_1	a_2	a_3	a_4	r
Clear sky	-0.0706	0.0516	-0.0131	0.0012	-	0.7566

4. Conclusion

Obtained results have shown that the radiative cooling in the investigated atmospheric layer occurs as a consequence of long-wave radiation emission effect. The most intensive radiative cooling in the layer between Stará Lesná and Skalnaté Pleso was found in December. The most intensive radiative cooling occurs at noon. It was shown that from February to October there is a certain anomaly in the cooling rate values after the sunset. After the sunset the radiative heating is more expressive at mean cloud conditions than at clear sky.

Study of the radiative cooling dependence on the thermal stratification in the atmospheric boundary layer has shown non-linear dependence. The radiation cooling rate increases with the decreasing value of the thermal gradient. The highest values of radiative cooling have occurred at strong thermal inversion and low water vapor pressure. It was confirmed that in annual mean sense, clouds reduce the radiative cooling rate to level of 43% of the value that would exist if skies over measurement sites remained clear.

Acknowledgement–The authors are grateful to the Grant Agency for Science (Grant No. 98/5305/417) for partially supporting this study.

References

Goody, R.M., 1964: *Atmospheric radiation. I. Theoretical basis*. Clarendon Press, Oxford.

Kondratyev, K.Ya., Vulis, I.L., Nikoljskij, G.A. and *Štengelj, M.P.*, 1972: On calculation of the radiative temperature changes in the cloudless atmosphere (in Russian). *Trudy GGO 276*, 196-208.

Kondratyev, K.Ya., 1956: *Radiative heat exchange in the atmosphere* (in Russian). Gidrometeoizdat, Leningrad.

Ostrožlík, M. and *Janičkovičová, L.*, (eds.), 1992–1995: *Results of meteorological measurements at the observatories of the Geophysical Institute of the Slovak Academy of Sciences*. Geophysical Institute of SAS, Bratislava.

Prokofjev, M.A. and *Ter-Markarjanc, N.E.*, 1972: Measurements of radiation in free atmosphere at KENEX expedition (in Russian). *Trudy GGO, 276*, 43-61.

Rodgers, C.D., 1967: The use of emissivity in atmospheric radiation calculations. *Quart. J. Roy. Meteorol. Soc. 93*, 43-54.

Stephens, G.L., Slingo, A., Webb, M.J., Minnett, P.J., Daum, P.H., Kleinman, L., Wittmeyer, I. and *Randall, D.A.*, 1994: Observations of the earth's radiation budget in relation to atmospheric hydrology. 4. Atmospheric column radiative cooling over the world's oceans. *J. Geophys. Res. 99*, 18585-18604.

Timanovskaja, R.G. and *Faraponova, G.P.*, 1967: The radiative heat flux determination in the boundary layer of the atmosphere (in Russian). *Fyzika atmosfery i okeana 12*, 1259-1270.

Zajceva, N.A. and *Šljachov, V.I.*, 1972: On longwave radiation field transformation in the free atmosphere (in Russian). *Trudy GGO 276*, 62-70.

BOOK REVIEW

A. Renoux and *D. Boulaud*: **Les aérosols — Physique et métrologie** (The Aerosols — Physics and Metrology). Lavoisier, Technique et Documentation, Paris, London and New York, 1998. 301 pages, two parts, seven chapters, several figures and tables.

The importance of the study of the atmospheric aerosol is well known for the readers of this journal. Among other things, aerosol particles in the atmosphere influence such basic processes as cloud formation and solar radiation transfer. In this way they play an important role in the control of climate, the most essential ensemble of environmental parameters for human life. Since electric ions and radioactive isotopes are normally attached to fine aerosol particles, the aerosol also determines the electric field and radioactive properties of the air. Less known is, however, that the study of aerosol particles is of interest for different branches of our industrial activity like automobile industry, air cleaning, air conditioning or modern high technology. Further, aerosol science is applied in agricultural practice, e.g. during dispersion of pesticides, herbicides and fertilizers. Finally, the atmospheric aerosol affects our health, while artificial aerosols can be used in medical treatments by inhalation. For this reason, the basic physics and measuring methods of aerocolloidal systems are essential because of many scientific and practical purposes. In this way it is not surprising that several books have been published in this field during the last decades. However, the books available are normally in English language. Thus, the present volume is unique in such a sense that it is prepared for French speaking audience, mainly for students and doctor fellows at French universities.

The first part of the book (physical bases) consists of three chapters written by Prof. A. Renoux. The first chapter summarizes very briefly our knowledge about some basic principles of the atmospheric aerosol. It goes without saying that this is not an easy task. The main merit of this chapter is the short and coherent presentation of the radioactive properties of the atmospheric aerosol. The second chapter is devoted to the size distribution functions used in different aerosol studies as well as to practical methods of their construction. The third chapter is entitled "Physical properties of aerosols". In this important chapter the writer gives an excellent summary of the physical properties of aerosol particles including sedimentation, phoretic forces, coagulation, electric and optical properties as well as condensation and evaporation of liquid particles. The section dealing with the electric properties of aerosol particles can be highlighted in particular. The effects of electric forces are also considered by discussing the coagulation of aerosols and their adherence to a given surface.

The material of this chapter can constitute a good basis of a university course for graduate students. It can also be used by any educated person interested in the field.

The second part of the volume is the work of Prof. D. Boulaud. In three chapters he presents a concise brief survey of measuring methods used to determine the physical properties (number and mass concentration, size distribution etc.) of aerosol particles. In the fourth chapter the reader can learn very easily the basic principles of aerosol measurements including the sampling under different transport conditions. This is a magnificent introduction for the beginners in the field. Chapter 5 is entitled "Instrumental techniques". This is the main chapter of the second part. It contains such important sections as filtration, optical and electric methods as well as sampling by inertial, gravitational and centrifugal forces. Moreover, in this chapter there are two other sections, one is devoted to expansion chambers and diffusion batteries, while the second treats radioactive and piezoelectric procedures of the real time monitoring of the mass concentration. Finally, the last section gives a summary of the generation and measurements of calibration aerosols. The aim of Chapter 6 is to discuss some applications of aerosol measurements at working places and in clean electronic laboratories.

The seventh chapter of the book presents the overall synthesis of the material discussed.

After having read the volume the reviewer concludes that the authors have reached their aim. This not too voluminous book is really a good summary of the physics and metrology of aerosols. It can be used satisfactorily in university practice and it gives a good introduction to the subject. Although it is obvious that this was not in the intention of the authors, it is a pity that the chemistry of aerosols, the chemical composition and chemical identification of atmospheric particles is not discussed. For this reason it seems that an other volume is necessary on aerosol chemistry if the authors want to have a more complete material on aerosols in French. Further, it is recommended to unify in a deeper way the first and second part of the book if it will be reprinted in the future. Thus, literature is given together after the first part, while in the second part it is given separately after each section. In some cases references are not given in a uniform manner. Secondly, some definitions are repeated in the second part which are presented already previously. This is not a problem, but the symbols used in the two parts are sometimes different. Generally speaking, a more careful editing would have increased further the level of the book. The reviewer has to emphasize, however, that these formal points do not touch the technical merits of the book, which can be proposed to everybody who wants to make an acquaintance with the physics and metrology of aerosols in French language.

E. Mészáros

ATMOSPHERIC ENVIRONMENT

an international journal

To promote the distribution of Atmospheric Environment *Időjárás* publishes regularly the contents of this important journal. For further information the interested reader is asked to contact *Prof. P. Brimblecombe*, School for Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, U.K.; E-mail: atmos_env@uea.ac.uk

Volume 32 Number 13 1998

European Scientific Assessment of the Atmospheric Effects of Aircraft Emissions

Volume 32 Number 14/15 1998

Volume 32 Number 16 1998

Volume 32 Number 17 1998

Volume 32 Number 18 1998

Research on the Effects of Aircraft and Spacecraft Upon the Atmosphere

IDŐJÁRÁS

VOLUME 102 * 1998

BUDAPEST, HUNGARY

AUTHOR INDEX

TABLE OF CONTENTS

I. Papers

II. Book review

III. News

Contents of journal Atmospheric Environment, 1998

IV. SUBJECT INDEX

The asterisk denotes book reviews

A

NOTES TO CONTRIBUTORS OF *IDŐJÁRÁS*

The purpose of the journal is to publish papers in any field of meteorology and atmosphere related scientific areas. These may be

- reports on new results of scientific investigations,
- critical review articles summarizing current state of art of a certain topic,
- shorter contributions dealing with a particular question.

Each issue contains "News" and "Book review" sections.

Authors may be of any nationality, but the official language of the journal is English. Papers will be reviewed by unidentified referees.

Manuscripts should be sent to
Editor-in-Chief of IDŐJÁRÁS
P.O. Box 39
H-1675 Budapest, Hungary

in three copies including all illustrations. One set of illustrations has to be of camera ready quality, the other two might be lower quality.

Title part of the paper should contain the concise title, the name(s) of the author(s), the affiliation(s) including postal and E-mail address(es). In case of multiple authors, the cover letter should indicate the corresponding author.

Abstract should follow the title, it contains the purpose, the data and methods as well as the basic conclusion.

Key-words are necessary to help to classify the topic.

The text has to be typed in double spacing with wide margins. Word-processor printing is preferred. The use of SI units are expected. The negative exponent is preferred to solidus. Figures and tables should be consecutively numbered and referred to in the text.

Mathematical formulas are expected to be as simple as possible and numbered in parentheses at the right margin. Non-Latin letters and hand-written symbols should be indicated and explained by making marginal notes in pencil.

Tables should be marked by Arabic numbers and printed in separate sheets together with their captions. Avoid too lengthy or complicated tables.

Figures should be drawn or printed in black and white, without legends, on separate sheets. The legends of figures should be printed as separate list. Good quality laser printings are preferred as master copies.

References: The text citation should contain the name(s) of the author(s) in Italic letter and the year of publication. In case of one author: *Miller* (1989), or if the name of the author cannot be fitted into the text: *(Miller,* 1989); in the case of two authors: *Gamov* and *Cleveland* (1973); if there are more than two authors: *Smith et al.* (1990). When referring to several papers published in the same year by the same author, the year of publication should be followed by letters a,b etc. At the end of the paper the list of references should be arranged alphabetically. For an article: the name(s) of author(s) in Italics, year, title of article, name of journal, volume number (the latter two in Italics) and pages. E.g. *Nathan, K.K.,* 1986: A note on the relationship between photosynthetically active radiation and cloud amount. *Időjárás 90*, 10-13. For a book: the name(s) of author(s), year, title of the book (all in Italics except the year), publisher and place of publication. E.g. *Junge, C. E.,* 1963: *Air Chemistry and Radioactivity*. Academic Press, New York and London.

The final version should be submitted on diskette altogether with one hard copy. Use standard 3.5" or 5.25" DOS formatted diskettes. The preferred word-processors are WordPerfect 5.1 and MS Word 6.0.

Reprints: authors receive 30 reprints free of charge. Additional reprints may be ordered at the authors' expense when sending back the proofs to the Editorial Office.

More information: gmajor@met.hu
Information on the last issues:
http://www.met.hu/firat/ido-e.html

Published by the Hungarian Meteorological Service

Budapest, Hungary

INDEX: 26 361

HU ISSN 0324-6329