IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

Special Issue: Weather forecasting and related applications
Guest Editors: **András Horányi** and **Tamás Weidinger**

CONTENTS

****

http://www.met.hu/Journal-Idojaras.php

VOL. 114 * NO. 1–2 * JANUARY– JUNE 2010

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: Gilice tér 39, H-1182 Budapest, Hungary
P.O. Box 39, H-1675 Budapest, Hungary
E-mail: bozo.l@met.hu or antal.e@met.hu
Fax: (36-1) 346-4809

Indexed and abstracted in Science Citation Index Expanded™ and Journal Citation Reports/Science Edition
Covered in the abstract and citation database SCOPUS®

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary
E-mail: kenderesy.k@met.hu or antal.e@met.hu

Editorial

Weather forecasting and related applications

During autumn of 2008, Meteorological Scientific Days (which is a traditional event of the meteorological society) were organized at the Hungarian Academy of Sciences on the subject of weather forecasting. The organization of this scientific event was ensured by the Working Committee of Atmospheric Dynamics under the Meteorological Scientific Committee of the Hungarian Academy of Sciences. After this workshop (see the presentations at http://www.met.hu/seminars/metnap2008.php), the idea arose to devote a special issue of the quarterly journal of the Hungarian Meteorological Service to this topic. Therefore, an invitation was submitted to the participants of the Meteorological Scientific Days and to other scientists working on the field of theoretical and applied weather forecasting (including researchers from other disciplines, who are applying weather forecasting information) to propose scientific papers to this special issue. Finally, eight articles were gathered, which are covering wide-range of topics on weather forecasting and related applications: data assimilation (including ensemble data assimilation), physical parameterization, ensemble forecasting, the special case of wind prediction for wind-power stations, environmental modeling, and urban heat island issues. All these subjects give a broad, but certainly not exhaustive overview of the weather forecasting and related applications in Hungary.

Unfortunately, there is also a bitter actuality of this special issue: the sudden and unbelievable death of our colleague and friend *Dezső Dévényi* (who was also member of the Editorial Board of IDŐJÁRÁS and also contributed to the special issue with the review of one of the article), therefore, we would like to dedicate the entire special issue to the memory of *Dezső Dévényi* (see the obituary of him in the same issue), since he had outstanding and long-lasting achievements to the topics of this special issue, i.e., on weather forecasting in general and numerical weather prediction in particular.

Finally, we would like to express our thanks to the Editor-in-Chief of IDŐJÁRÁS supporting the idea of the special issue, special gratitude is going to the authors of the articles and also to the dedicated reviewers making the critical review of the articles keeping the high standards of the journal. We are also grateful to the Executive Editor of the journal taking care about all the practical issues of the recent volume.

András Horányi[1] and *Tamás Weidinger*[2]
Guest Editors

[1]Hungarian Meteorological Service
[2]Eötvös Loránd University, Budapest

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–June 2010, pp. 1–19

Diagnosis and tuning of background error statistics in a variational data assimilation system

Gergely Bölöni[1*] and **Kristian Horvath**[2]

[1]*Hungarian Meteorological Service*
P.O. Box 38, H-1525 Budapest, Hungary; E-mail: boloni.g@met.hu

[2]*Meteorological and Hydrological Service of Croatia*
Gric 3, Zagreb, HR-10000, Croatia; E-mail: horvath@cirus.dhz.hr

[*]*Corresponding author*

(Manuscript received in final form November 9, 2009)

Abstract—A very important source of information in data assimilation is the background, which is usually a short-range forecast of the numerical weather prediction (NWP) model valid at the assimilation time. The background is corrected by the atmospheric observations providing the analysis during the process of data assimilation. Errors of the background are estimated statistically and are taken into account with the aim of giving a proper weight to both the background and observations. Although, in theory, the background errors can be defined as the difference between the truth and the background, it is a great challenge to generate appropriate background error samples in practice for the computation of the background error statistics, because the true state is always unknown. A possible way to improve the background error representation is the development of the background sampling methods themselves. Another complementary approach is the "a posteriori" diagnosis or tuning of the already predefined background error statistics by a certain sampling technique. This paper presents two attempts for such a tuning within the ALADIN limited area model (LAM) and its 3d-var assimilation system used operationally at the Hungarian Meteorological Service (HMS). The first one is based on the comparison of the predefined statistics with those expected in a statistically optimal assimilation system. The second one is inspired by single observation experiments and it addresses the improvement of the multivariate statistical balance between humidity and the other analyzed variables.

Key-words: data assimilation, background errors, "a posteriori" tuning, multivariate statistical balance

1. Introduction

Data assimilation methods based on statistical optimality (*Gandin*, 1963; *Lorenc*, 1986; *Thépaut* and *Courtier*, 1991; *Bouttier* and *Courtier*, 1999) take into account the error of each available information type for weighting their

contribution to the analysis $\mathbf{x}_a$, i.e., the initial condition for the forecast. This is done by solving the so-called BLUE (best linear unbiased estimation) analysis equation:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}_b)\right). \quad (1)$$

Main available information are the atmospheric observations $\mathbf{y}$ and the background $\mathbf{x}_b$, which is usually a short-range forecast of the NWP model. The notation H stands for the so-called non-linear observation operator, which projects from the space of the model variables to that of the observed variables. The operator $\mathbf{H}$ is the linearized of H. The matrices $\mathbf{B} = \mathrm{E}\left(\boldsymbol{\varepsilon}_b \boldsymbol{\varepsilon}_b^T\right)$ and $\mathbf{R} = \mathrm{E}\left(\boldsymbol{\varepsilon}_o \boldsymbol{\varepsilon}_o^T\right)$ denote the background and observation error covariance matrices respectively, where E stands for the statistical expectation. The corresponding errors are defined as the difference from the true state of the atmosphere $\mathbf{x}_t$, i.e., $\boldsymbol{\varepsilon}_b = \mathbf{x}_b - \mathbf{x}_t$ can be written for the background errors and $\boldsymbol{\varepsilon}_o = \mathbf{y} - \mathbf{H}(\mathbf{x}_t)$ for the observation errors. In the lack of $\mathbf{x}_t$ the sampling of these errors is a real challenge in the field of data assimilation. For the sampling of background errors, many ideas have been developed, like the method of innovations (*Hollingsworth* and *Lönnberg*, 1986; *Lönnberg* and *Hollingsworth*, 1986) or the NMC (*Parrish* and *Derber*, 1992) and Ensemble (*Fisher,* 2003; *Belo Pereira* and *Berre*, 2006; *Stefanescu et al.,* 2006) methods. Also attempts have been taken to characterize LAM specific background errors with the lagged-NMC approach (*Siroká et al.*, 2003). With the application of the above sampling techniques the representation of the background errors has been improved in some aspects indeed. In this paper, however, we would like to point out, that as $\mathbf{x}_t$ is never known, it is impossible to compute perfect $\mathbf{B}$ and $\mathbf{R}$ matrices with any method, and it is always reasonable to tune certain elements of the predefined covariance matrices "a posteriori". We will concentrate on the tuning of the $\mathbf{B}$ matrix statistics primarily, which was applied to the 3d-var system of the ALADIN model (*Horányi et al.*, 1996) in our experiments. In Chapter 2 we describe a tuning experiment aiming to improve the background error variances based on statistical optimality criteria of analysis residuals. In Chapter 3 we show a tuning to improve the multivariate balance in the analysis based on the simplified analysis Eq. (1) and single observation tests.

2. Tuning of the background error statistics based on covariances of residuals

Based on the optimal estimation theory, several approaches exist for "a posteriori" validation and tuning. A technique applied earlier in ARPEGE/ALADIN was based on the assumption that the expected value of the variational cost function J at the minimum is proportional to the number of observations used, i.e.,

$E(J_{min})=p/2$ hold, where p is the number of observations used in the analysis (*Talagrand*, 1998; *Désroziers* and *Ivanov*, 2001; *Chapnik et al.*, 2004; *Sadiki* and *Fischer*, 2005; *Fischer et al.*, 2005). Another method was proposed by *Désroziers et al.* (2006) later on, which is the theoretical basis for the tuning described in this chapter.

2.1. The method of tuning

It is explained in the above-mentioned paper that in a linear analysis system, where the **B** background and the **R** observation error covariance matrices are properly estimated, the following equations hold:

$$\begin{aligned} E\left(\mathbf{d}_b^o \mathbf{d}_b^{oT}\right) &= \mathbf{HBH}^T + \mathbf{R} \\ E\left(\mathbf{d}_b^a \mathbf{d}_b^{oT}\right) &= \mathbf{HBH}^T \\ E\left(\mathbf{d}_a^o \mathbf{d}_b^{oT}\right) &= \mathbf{R} \\ E\left(\mathbf{d}_b^a \mathbf{d}_a^{oT}\right) &= \mathbf{HAH}^T \end{aligned} \tag{2a}$$

where

$$\begin{aligned} \mathbf{d}_b^o &= \mathbf{y} - \mathbf{H}(\mathbf{x}_b) \\ \mathbf{d}_b^a &= \mathbf{H}(\mathbf{x}_a) - \mathbf{H}(\mathbf{x}_b) \\ \mathbf{d}_a^o &= \mathbf{y} - \mathbf{H}(\mathbf{x}_a) \end{aligned} \tag{2b}$$

are residuals of the model and the observations provided by the assimilation system. The notation **A** in Eq. (2a) stands for the analysis error covariance matrix. The optimality of the analysis system can be diagnosed if one computes the covariances of residuals on the left-hand side and substitutes the predefined error statistics to the right-hand side of Eq. (2a). The residuals on the left-hand side can be easily obtained as a by-product of the assimilation system. The second and the third equations in Eq. (2a) can be applied for instance for the diagnosis of the background and observation error standard deviations:

$$\sigma_{bd} = \sqrt{\frac{1}{P}\sum_{i=1}^{P} d_{bi}^a d_{bi}^o} \qquad \sigma_{od} = \sqrt{\frac{1}{P}\sum_{i=1}^{P} d_{ai}^o d_{bi}^o}\,, \tag{3}$$

where d_{bi}^a, d_{oi}^a, and d_{bi}^o stand for the individual i realizations of the background and analysis departures (i=1, …, P). Then the misfit of the predefined standard deviations (σ_{bp} and σ_{op}) can be obtained as follows:

$$r_b = \frac{\sigma_{bd}}{\sigma_{bp}} \qquad r_o = \frac{\sigma_{od}}{\sigma_{op}}\,. \tag{4}$$

These misfit ratios can be used as guidance for tuning the predefined standard deviations in the assimilation system. One has to notice, that through Eq. (1), the Eq. (2b) residuals depend on σ_{bp} and σ_{op}, which means that ideally the tuning with the misfit ratios and the computation of the analysis (Eq. (1)) should be done iteratively until r_b and r_o converge. It has also been shown by *Désroziers et al.* (2006) in a simplified system, that the convergence can be reached in a few iterations.

2.2. Estimation of the misfit ratios in our analysis system

The above method was applied in the ALADIN 3d-var system used at HMS (*Bölöni*, 2006) for the estimation of the misfit ratio for the background error standard deviations. The Eq. (2b) residuals were taken from two assimilation cycles run over an autumn (October–November 2005) and a summer (June 2006) period. In these assimilation cycles all the available observations were included, which are used operationally in the ALADIN 3d-var system at HMS (*Randriamampianina*, 2006). However, for the diagnosis of the background error standard deviations, only the residuals based on direct observations (i.e., radiosondes and aircrafts) were used. There was no iteration applied in the estimation for the sake of simplicity. In the fully spectral version of the ALADIN 3d-var system, predefined background error standard deviations are uniform horizontally and variable in the vertical, so they are available as one value for each model level and analyzed meteorological variables (i.e., temperature (T), specific humidity (q), vorticity (ζ), divergence (η)) (*Berre*, 2000). On the other hand, the diagnosed standard deviations can be obtained at the observation locations and for the observed quantities (T, q, and wind u, v components). In our estimation, a vertical averaging was applied both for the predefined and diagnosed values so that they became comparable independently from the height. For the comparison of predefined (ζ,η) and diagnosed (uv) wind error values, an average wind standard deviation (σ_b(uv)) was defined that can be computed as:

$$\sigma_{bd}(uv) = \sqrt{\frac{1}{2}\left(\sigma_{bd}^2(u) + \sigma_{bd}^2(v)\right)} \tag{5}$$

from the diagnosed u and v standard deviations values, and as:

$$\sigma_{bp}(uv) = \sqrt{-\frac{1}{2}\Delta^{-1}\left(\sigma_{bp}^2(\zeta) + \sigma_{bp}^2(\eta)\right)} \tag{6}$$

from the predefined ζ and η standard deviations, where Δ denotes the Laplace operator, i.e., the second space derivative in horizontal. *Table 1* shows the misfit

ratios computed as explained above and also the predefined and diagnosed background error standard deviations both for the autumn and summer periods.

Table 1. Predefined, diagnosed background error standard deviations and misfit ratios (dimensionless) for specific humidity (kg/kg), temperature (K) and average wind speed ($m\ s^{-1}$) computed over the autumn (October 26–November 10, 2005) and the summer (June 05–20, 2006) periods

Variable	**Predefined**	**Autumn period**		**Summer period**	
		Diagnosed	**Misfit ratio**	**Diagnosed**	**Misfit ratio**
	(σ_{bp})	(σ_{bd})	(r_b)	(σ_{bd})	(r_b)
Specific humidity (q)	2.27×10^{-4}	5.34×10^{-4}	2.35	5.82×10^{-4}	2.56
Temperature (T)	0.4917	0.7071	1.43	0.8010	1.62
Wind (u,v)	1.4840	1.9878	1.33	1.9203	1.29
Average	**0.65**	**0.89**	**1.36**	**0.9**	**1.38**

Large misfit ratios for humidity suggest that the humidity background error standard deviations are the less accurate as predefined by the original **B** matrix in our assimilation system. The misfit ratios are slightly different for the two different time periods, which suggests a possible seasonal dependence of the accuracy in the predefined background error modeling.

2.3. The tuning experiments

The analysis system can be tuned by multiplying the predefined standard deviations with the misfit ratios introduced in the previous section. In our experiments such tuning was done only for the background error standard deviations in two steps with different complexity:

1. ENS1 uses a uniform misfit ratio for all the variables and vertical model levels. This means that the misfit ratio used was averaged over the vertical levels and over the variables q, T, and wind.
2. ENS2 uses a variable dependent misfit ratio still averaged over the vertical levels.

It is to be mentioned, that the predefined statistics were computed based on the Ensemble sampling method. Both the ENS1 and ENS2 tuning options were tested in real assimilation cycling experiments over the summer period. It means a 16-day assimilation cycle (64 analysis steps) using the operational observational dataset. Simulations of 48-hour production runs were started from 00 UTC after 2 days of warm up cycling. The impact of the tuning was measured through rmse and bias score (e.g., *Wilks*, 1995) computations of the above-mentioned production forecasts against surface and radiosonde observations. The reference

assimilation cycle for the tuning experiments (ENS0) was the same as ENS1 and ENS2, except that no tuning of the background error standard deviations was applied, i.e., the predefined values were used. In experiment ENS1, a uniform misfit ratio, r_b=1.3 was used as a multiplication factor to increase the predefined background error standard deviations. This value was chosen based on the average misfit ratios shown in *Table 1*. The verification results show that the ENS1 tuning had no impact on the 2-meter parameters (not shown) but did improve the humidity forecasts on 700 hPa and the wind forecasts on 250 hPa (*Fig. 1b* and *d*). The impact on geopotential and temperature in the altitude was found to be rather neutral (*Fig. 1a* and *c*).

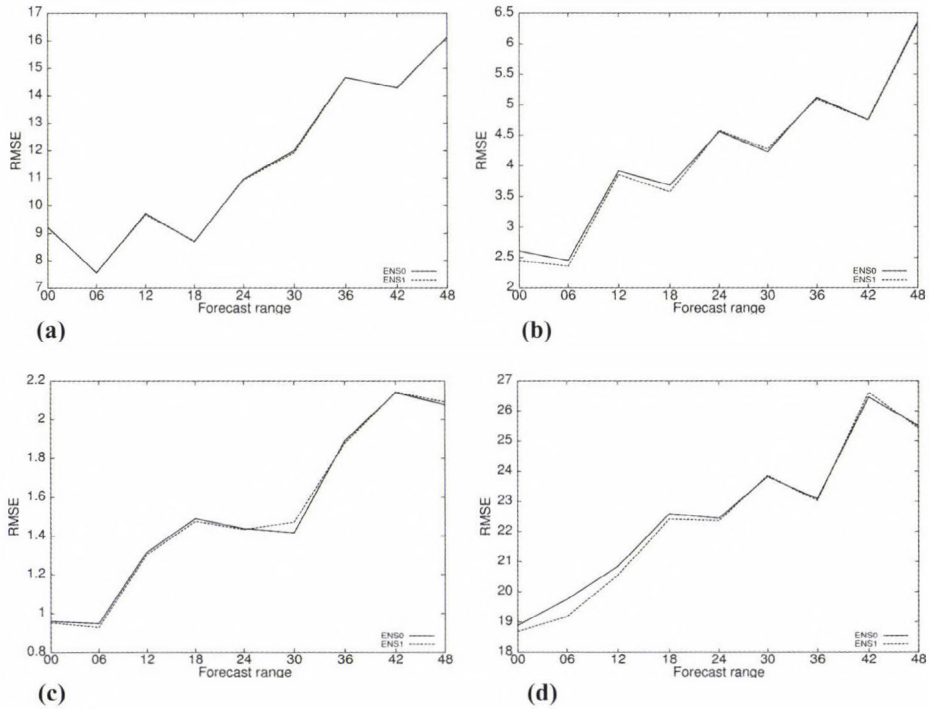


Fig. 1. Verification scores (rmse) computed for (a) geopotential at 500 hPa ($m^2\ s^{-2}$), (b) wind speed at 250 hPa ($m\ s^{-1}$), (c) temperature at 850 hPa (K), (d) relative humidity at 700 hPa (%). Dashed: experiment ENS1, solid: reference ENS0. The scores are computed over the period June 7–20, 2006.

In experiment ENS2, the misfit ratios computed from the autumn period were used. These are r_b=1.33 for the wind and mass variables (vorticity, divergence, surface pressure, and geopotential), r_b=1.43 for temperature, and r_b=2.35 for humidity, as shown in *Table 1*. The tuning had no impact on the 2-meter fields (not shown). Results for the atmospheric variables are displayed in *Fig. 2*. One

can see that ENS2 gives clearly worse results than the reference ENS0 for temperature on 850 hPa and also somewhat for relative humidity on 700 hPa. There is a positive impact of the tuning for the wind on 250 hPa.

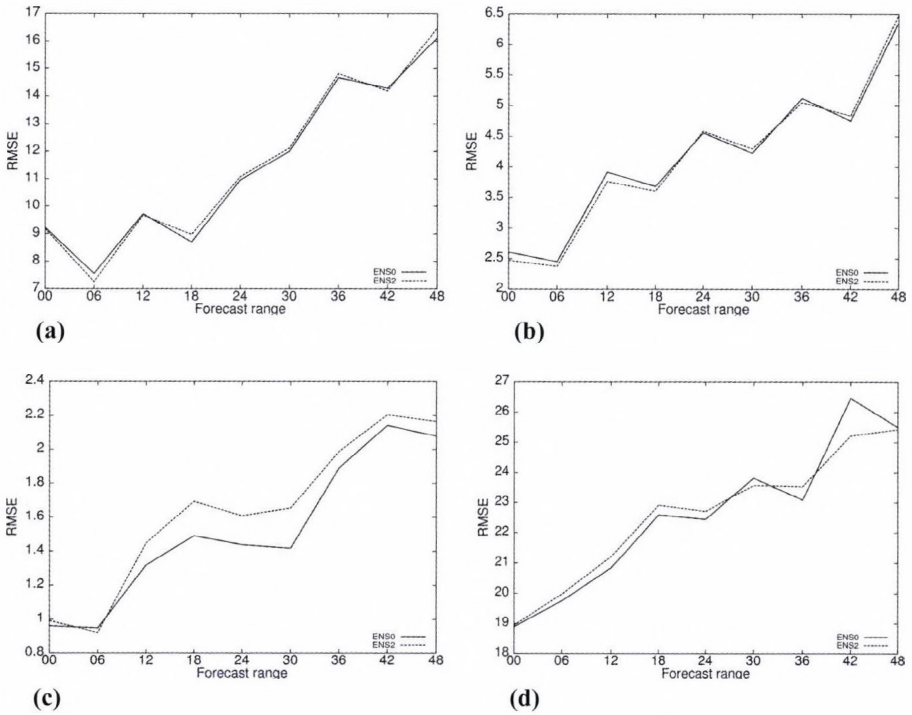


Fig. 2. Verification scores (rmse) computed for (a) geopotential at 500 hPa ($m^2\ s^{-2}$), (b) wind speed at 250 hPa ($m\ s^{-1}$), (c) temperature at 850 hPa (K), (d) relative humidity at 700 hPa (%). Dashed: experiment ENS2, solid: reference ENS0. The scores are computed over the period June 7–20, 2006.

The above results show that the uniform tuning (ENS1) provides better results than the variable dependent one (ENS2), which looks curious, because in theory Eq. (3) holds for each variable separately as well. In *Fig. 3* one can see the predefined and the tuned background error standard deviation profiles used in the experiments. It is obvious that the tuned profiles (ENS1 and ENS2) are quite similar to each other for wind. The difference is larger for temperature, and it is the largest for humidity. This suggests that the disappointing results of experiment ENS2 are mostly due to the large misfit ratios for humidity and temperature. A probable reason for the poorer result of the ENS2 experiment is that the misfit is averaged in the vertical, which can lead to an exaggerated tuning on those levels where the real misfit is drastically smaller than the average.

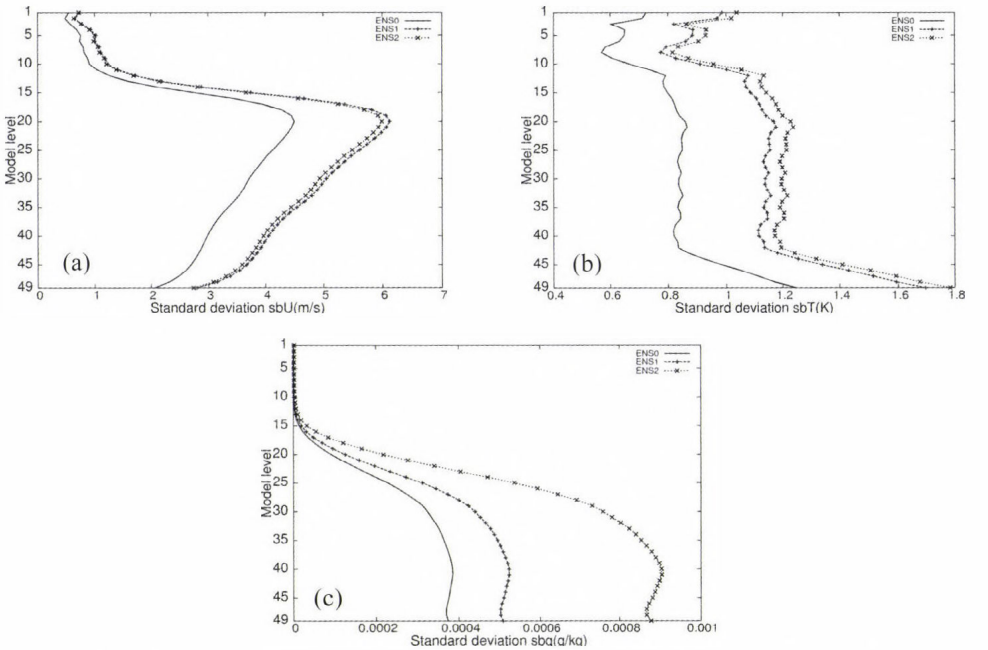


Fig. 3. Predefined (ENS0) and tuned (ENS1 and ENS2) profiles of total background error standard deviations of (a) wind, (b) temperature, (c) specific humidity.

3. Tuning of the multivariate balance

The ALADIN 3d-var is a multivariate analysis system provided by the statistical balance described by *Rabier et al.* (1998), *Courtier et al.* (1998), *Berre* (2000), and *Gustafsson et al.* (2001). In short, this statistical balance couples the analyzed variables in a meteorologically meaningful way by propagating a part of the increments of a given variable to those of another one. This insures the dynamical balance of the variables in the analysis. Experiments done at HMS showed that temperature observations have a large impact on the relative humidity analysis, which results in degradation of the relative humidity analysis and forecast verification scores in some weather situations. On the other hand, the humidity observations influence the temperature analysis in a very limited extent. This asymmetry in the multivariate balance brought us to study the analysis equation in a simplified framework and to propose a tuning of the error variances for a more symmetric balance.

3.1. Theoretical considerations

It can be shown that in a single point model (i.e., $\mathbf{H} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix), restricting to two variables only (T: temperature and Rh: relative humidity), the analysis Eq. (1) can be written as:

$$\delta T = \frac{\mathrm{cov}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right)+\sigma^2\left(\boldsymbol{\varepsilon}_{o,Rh}\right)}\Delta Rh\,, \tag{7}$$

where δ denotes the analysis increment (analysis minus background) and Δ stands for the observation increment or innovation (observation minus background). Based on *Daley* (1991) and *Hollingsworth* (1987), Eq. (7) can be decomposed as:

$$\delta Rh = \frac{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right)+\sigma^2\left(\boldsymbol{\varepsilon}_{o,Rh}\right)}\Delta Rh \tag{8a}$$

$$\delta T = \frac{\mathrm{cov}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}\delta Rh\,, \tag{8b}$$

where (8a) is the univariate or filtering step, which provides humidity analysis increment from the humidity innovation and (8b) is the multivariate propagation step, which transforms the humidity analysis increment to temperature analysis increment. Multiplying and dividing the right-hand side of (8b) with $\sigma_b\left(\boldsymbol{\varepsilon}_T\right)$ and using the definition of the correlation, Eq. (8b) can be rewritten as:

$$\delta T = \mathrm{corr}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)\frac{\sigma\left(\boldsymbol{\varepsilon}_{b,T}\right)}{\sigma\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}\delta Rh\,. \tag{8c}$$

In case of a temperature innovation, one can write the corresponding equations similarly to Eqs. (8a) and (8b) as follows:

$$\delta T = \frac{\sigma^2\left(\boldsymbol{\varepsilon}_{b,T}\right)}{\sigma^2\left(\boldsymbol{\varepsilon}_{b,T}\right)+\sigma^2\left(\boldsymbol{\varepsilon}_{o,T}\right)}\Delta T \tag{9a}$$

$$\delta Rh = \mathrm{corr}\left(\boldsymbol{\varepsilon}_{b,T}\boldsymbol{\varepsilon}_{b,Rh}\right)\frac{\sigma\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma\left(\boldsymbol{\varepsilon}_{b,T}\right)}\delta T\,. \tag{9b}$$

The asymmetry in the multivariate propagation step can be quantified as the ratio of T and Rh analysis increments induced by T and Rh innovations:

$$S = \left(\frac{\delta Rh}{\delta T}\right)_{\Delta T} \Big/ \left(\frac{\delta T}{\delta Rh}\right)_{\Delta Rh}\,. \tag{10}$$

The ratio **S** was computed within the ALADIN 3d-var system using single observations ($\Delta T = 1\,K$ and $\Delta Rh = 2\%$), taking the analysis increments right at the origin point (*Fig. 4*).

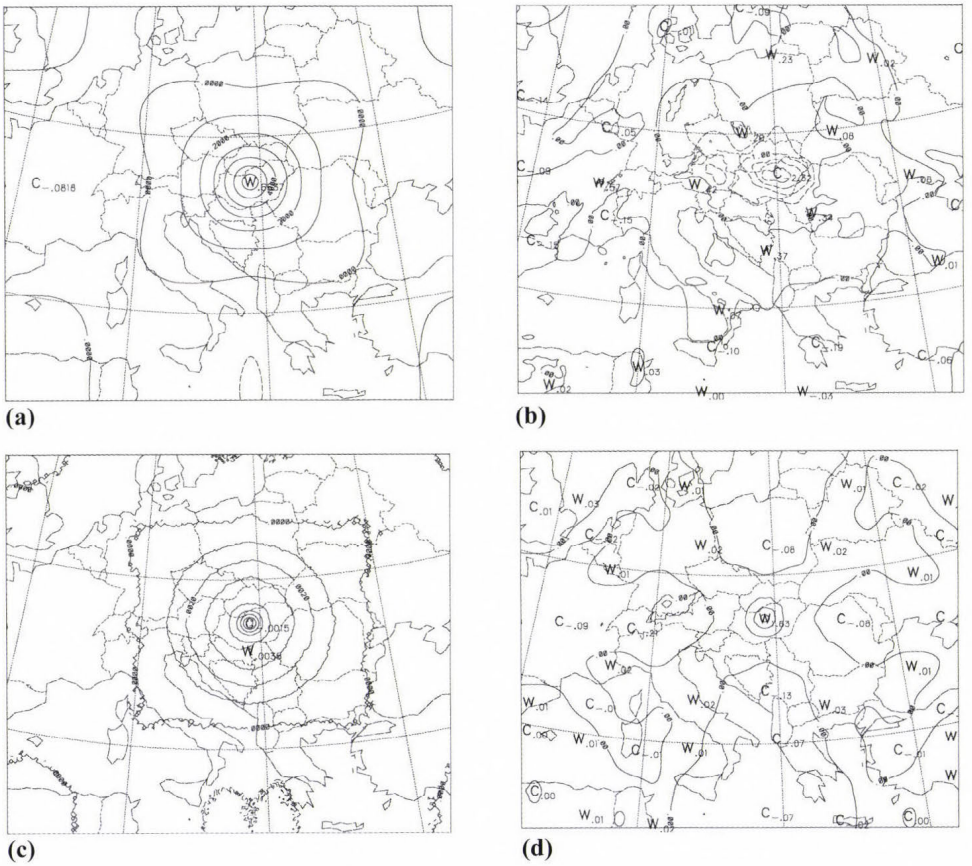


Fig. 4. Analysis increments due to single observations: (a) temperature increment induced by a temperature innovation, (b) relative humidity increment induced by a temperature innovation, (c) temperature increment induced by a relative humidity innovation, (d) relative humidity increment induced by a relative humidity innovation. The innovation values are $\Delta T = 1\,K$ and $\Delta Rh = 2\%$, observation error used are $\sigma(\varepsilon_{o,Rh}) = 0.12\%$, $\sigma(\varepsilon_{o,T}) = 1.1\,K$.

Note that the single observation experiments were run using NMC background error statistics sampled from differences of 36- and 12-hour forecasts of the operational ALADIN model over the period May–July 2004. Substituting the analysis increments shown in *Fig. 4* into Eq. (10), S≡3688 was found. It is easier to interpret this value if one writes up Eq. (10) using the Eqs. (8c) and (9b), which gives:

$$S = \left(\frac{\sigma\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma\left(\boldsymbol{\varepsilon}_{b,T}\right)} \right)^2 \cong 3688. \tag{11}$$

This shows that the symmetry of the multivariate coupling depends only on the variance ratios of the two variables. Taking the square root of the above equation one gets:

$$\sqrt{S} = \frac{\sigma\left(\boldsymbol{\varepsilon}_{b,Rh}\right)}{\sigma\left(\boldsymbol{\varepsilon}_{b,T}\right)} \cong 60 \frac{\%}{K}. \tag{12}$$

According to the above equation, a 1 K temperature error is associated to a 60% error in relative humidity and vice versa, which points out the rather asymmetric behavior of the multivariate propagation steps taking into account typical changes of temperature and humidity in the troposphere. The proposal for the tuning is thus to decrease the above ratio of standard deviations, which can be done either by increasing $\sigma(\boldsymbol{\varepsilon}_{b,T})$ or by decreasing $\sigma(\boldsymbol{\varepsilon}_{b,Rh})$. Examining the univariate steps in case of the T (*Fig. 4a*) and Rh innovations (*Fig. 4d*), one can see that the Rh analysis increment is about the 80% of the Rh innovation, while the T analysis increment is about the 60% of the T innovation. This means that the trust in the humidity background is indeed very small, which was the basis for assuming that $\sigma(\boldsymbol{\varepsilon}_{b,T})$ is well chosen and $\sigma(\boldsymbol{\varepsilon}_{b,Rh})$ is overestimated. As a consequence, it was decided to run tuning experiments by reducing the humidity background error standard deviations. An additional remark is that beside reducing humidity background error standard deviations, we considered to reduce also the humidity observation error standard deviations in order to keep Eq. (8a) univariate filtering step unchanged.

It should be clarified here, that the ALADIN 3d-var uses specific humidity (q) instead of relative humidity as analysis control variable, which also implies that the $\sigma(\boldsymbol{\varepsilon}_{b,q})$ specific humidity background error standard deviations were chosen for tuning instead of $\sigma(\boldsymbol{\varepsilon}_{b,Rh})$. The reason for using Rh in the above elaboration was that the asymmetry of the multivariate balance was observed first in relative humidity scores, and also the reader might judge relative humidity changes in the atmosphere easier than those of specific humidity.

Another feature that requires explanation here is that Rh single observation increments in *Figs. 4b* and *4d* show some obvious anisotropy in spite of the fact that the ALADIN horizontal background error structure is designed to be isotropic by origin (*Berre*, 2000). The clue for this seeming contradiction is that the anisotropy is not included by the structure functions directly but rather by the computation Rh from q. Namely, to derive and plot relative humidity

increments, first the $q \rightarrow Rh$ computation is performed both on the background and analysis fields, and then their difference is taken. As the $q \rightarrow Rh$ computation is a non-linear function of specific humidity and temperature, the resulting relative humidity increment will not keep the isotropic structure even if both specific humidity and temperature increments alone are isotropic.

3.2. The method of tuning

The extent of reduction of humidity errors was determined on the basis of an error estimation proposed by *Hollingsworth* and *Lönnberg* (1986) (HL method), which is independent from the NMC method. In short, the HL method estimates the observation and background errors based on the first equation of Eq. (2a), where the innovations are composed by the difference of radiosonde observations and model forecasts. By sorting the innovations in vertical and horizontal distance categories, one can estimate a vertical profile of observation and background error standard deviations and horizontal correlation length scales. This independent method was used as a guideline for estimating new background and observation standard deviations for humidity within our analysis system. The HL method was applied for the same period as the one used for the NMC sampling (May 2–August 2, 2004). The HL standard deviations were first computed on the standard levels of the radiosonde observations, and then they were interpolated on the model levels for the comparison with the NMC statistics.

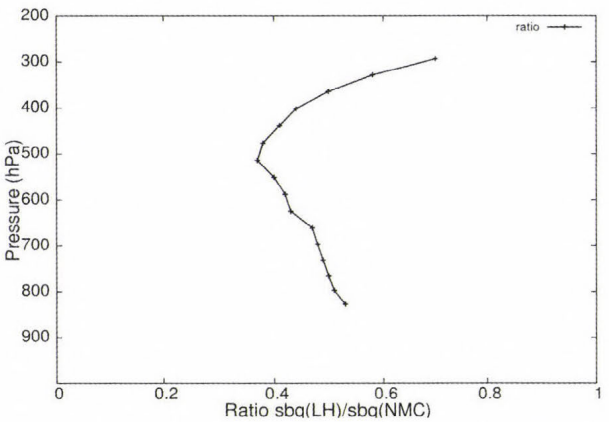


Fig. 5. Vertical profile of specific humidity background error standard deviation ratios obtained by the HL and NMC methods.

In *Fig. 5* the ratio of HL and NMC background error standard deviations are shown for specific humidity. Levels under 850 hPa were not considered for the tuning due to the small amount of radiosonde observations. Despite the strong vertical variability, the ratio is smaller than 1 for all the vertical levels,

which means that the HL estimation of $\sigma(\boldsymbol{\varepsilon}_{b,q})$ is smaller than the NMC one in the whole troposphere. This is in accordance with the notion of $\sigma(\boldsymbol{\varepsilon}_{b,q})$ being overestimated in the NMC statistics, which was also suggested by the theoretical considerations described earlier. Thus, it was decided to tune the $\sigma(\boldsymbol{\varepsilon}_{b,q})$ and $\sigma(\boldsymbol{\varepsilon}_{o,q})$ profiles used in the ALADIN 3d-var experiments according to the ratio profile shown in *Fig. 5*. Additionally, in some of the experiments $\sigma(\boldsymbol{\varepsilon}_{b,q})$ was drastically reduced (multiplied by a factor of 0.005) above 250 hPa in order to prevent to propagate high altitude humidity increments to the lower troposphere. This latter modification was based on the experiences at ECMWF (*Andersson et al.*, 1998).

3.3. The tuning experiments

The tuning experiments were run with the following three settings:

1. EX1Q: modified $\sigma(\boldsymbol{\varepsilon}_{b,q})$ between 850 and 250 hPa according to the HL/NMC ratio, reduced $\sigma(\boldsymbol{\varepsilon}_{b,q})$ above 250 hPa.
2. EX2Q: same as in EX1Q but also $\sigma(\boldsymbol{\varepsilon}_{o,q})$ is reduced in the same heights and in the same extent as $\sigma(\boldsymbol{\varepsilon}_{b,q})$.
3. EX3Q: same as EX2Q except that there was no tuning applied above 250 hPa.
4. REFQ: reference run using the predefined $\sigma(\boldsymbol{\varepsilon}_{o,q})$ and $\sigma(\boldsymbol{\varepsilon}_{b,q})$ values.

As a first step, the computation of Eq. (10) was repeated with the tuned error standard deviations based on single observation experiments again. It was proven that the tuning indeed reduced the asymmetry in the expected way as $\sqrt{S} = 35\%/\mathrm{K}$ was found with the EX1Q and $\sqrt{S} = 33\%/\mathrm{K}$ was obtained with the EX2Q settings. As a second step, complete assimilation cycles were run for a two-week period (August 29–September 10, 2004) with all the above experimental settings in order to see the impact of the tuning in a real operational-like context. The assimilation cycles were run with a 6-hour frequency, using surface, radiosonde, aircraft, and satellite (ATOVS AMSU-A and B) observations. Production forecasts (up to 48 hours) were run from the 00 UTC analyses. In general, the tuning had a smaller impact in these real cycling experiments than expected in terms of verification scores, which was surprising after the firm effect in the single observation experiments. This can be explained by the fact that the single observation experiments are much closer to the simplified system Eqs. (7)–(10) than the complex assimilation cycling experiments, where also the spatial correlations may play an important role. The largest impact was found for the experiment EX2Q, which will be referred in the further comparison against the reference run REFQ. In *Figs. 6* and *7*, bias and rmse scores for the analyses are shown for the EX2Q and REFQ runs. The tuning

resulted in a better (worse) fit to the observations concerning the temperature (humidity) analysis rmse. The impact on the wind analysis is rather mixed.

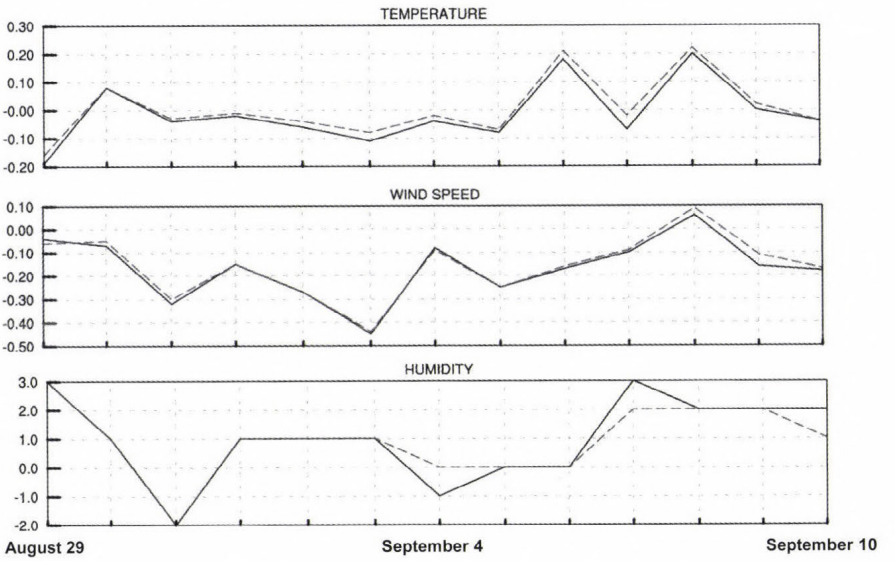


Fig. 6. Evolution of analysis bias scores for the experiments EX2Q (solid line) and the reference REFQ (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29 – September 10, 2004.

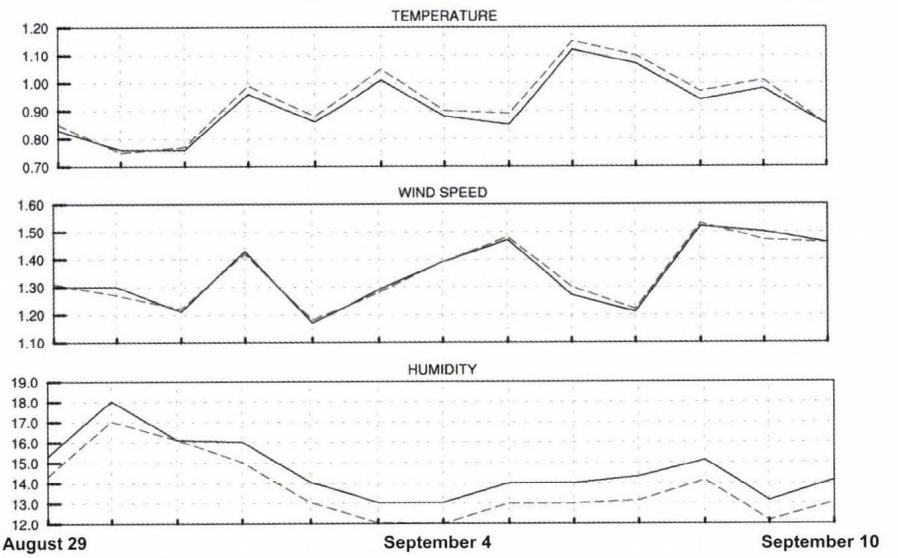


Fig. 7. Evolution of analysis rmse scores.....→

In our explanation the good fit of temperature comes indeed from the improved multivariate analysis, namely from a larger multivariate propagation of the humidity increments into those of temperature through the process described by Eq. (8b). On the other hand, the increased analysis departure of humidity may have two reasons: (i) multivariate humidity increments induced by those of temperature and wind became too small by excessively decreasing $\sigma(\boldsymbol{\varepsilon}_{b,q})$; and (ii) the univariate step of the humidity analysis Eq. (8a) was degraded by reducing $\sigma(\boldsymbol{\varepsilon}_{b,q})$, in spite of our compensation by reducing $\sigma(\boldsymbol{\varepsilon}_{o,q})$ at the same time.

Figs. 8 and *9* show that the positive impact on temperature decreases quickly with the forecast range. Nevertheless, humidity 6-hour forecast scores are improved by the performed tuning both in terms of bias and rmse. The impact on wind is kept mixed for this forecast range, while the overall effect of the tuning on longer forecast ranges is rather neutral (not shown).

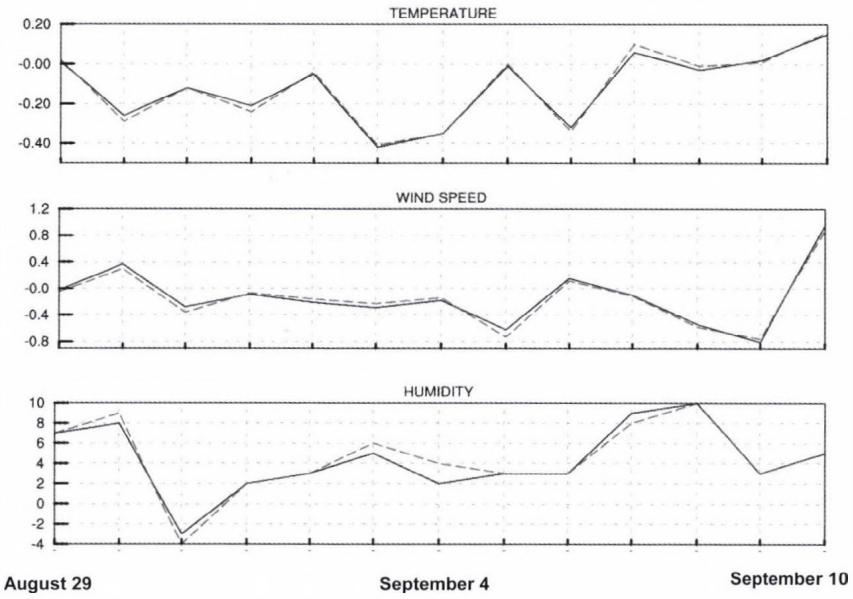


Fig. 8. Evolution of 6-hour forecast bias scores for the experiments EX2Q (solid line) and the reference (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29–September 10, 2004.

Fig. 7. Evolution of analysis rmse scores for the experiments EX2Q (solid line) and the reference REFQ (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29 – September 10, 2004.

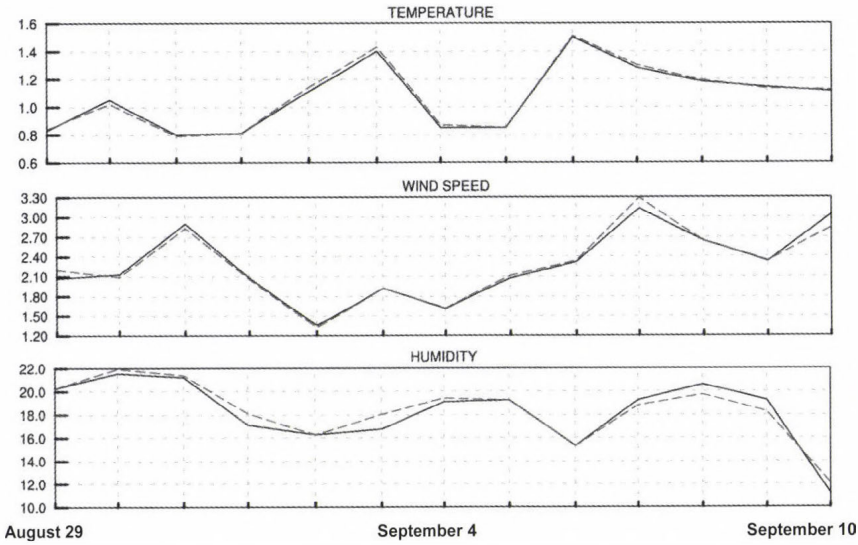


Fig. 9. Evolution of 6-hour forecast rmse scores for the experiments EX2Q (solid line) and the reference (dashed line). From top to bottom: temperature at 850 hPa (K), wind speed at 500 hPa (m s^{-1}), relative humidity at 700 hPa (%). The scores are computed over the period August 29–September 10, 2004.

4. Discussion

Two attempts to improve the Hungarian version of the ALADIN 3d-var analysis system were described. Both of them involve the tuning of the background and observation error standard deviations "a posteriori", but on a different theoretical basis.

The first tuning shown is based on the statistical optimality criteria derived by *Désroziers et al.* (2006) and aims to tune the errors for the full set of the background variables. The tuning was applied both in a variable dependent way and also as a uniform tuning for all the variables. A clear positive impact of the uniform tuning was found, while in the variable dependent case the verification showed a degradation of the forecasts. This seems to be a contradiction, as the more complex tuning had a worse impact than the simpler one. A probable explanation of this feature is that no vertical dependence of the tuning was included in any of the experiments, but a vertically averaged tuning ratio was used instead, being a considerable limitation of the approach. It is assumed that the degrading effect of the vertical averaging appears more intensively in the variable dependent tuning if the background error of the given variable shows a large vertical dependence (i.e., in the case of temperature and humidity). The uniform tuning has been introduced in the operational ALADIN 3d-var data assimilation system of HMS after the experiments.

The second tuning shown is based on the simplified BLUE analysis equations, proposing a modification only for the humidity errors. On the basis of single observation experiments, it was shown that the predefined error statistics produce an asymmetric multivariate balance between humidity and the other variables. Namely, humidity increments due to wind and temperature innovations are excessive compared to temperature and wind increments due to humidity innovations. On the basis of the simplified BLUE Eqs. (7)–(10), it was derived that the asymmetry can be decreased by reducing the background error standard deviations of humidity, which was demonstrated in a set of single observation experiments. To estimate the decrease in the humidity background error standard deviations the HL method was used (*Hollingsworth* and *Lönnberg*, 1986), which suggested a decrease of approximately half of the predefined standard deviation values on average. Full-observation experiments were performed with the modified multivariate propagation step, as well as experiments where the observation error standard deviations of humidity were also decreased to the same extent in order to keep the univariate analysis step unchanged. In real assimilation cycling experiments, the overall impact of the tuning was found to be quite moderate. This was probably caused by additional effects of spatial correlations of background errors, which were not taken into account in the simplified equations used as the basis of the tuning. However, a positive impact on 6-hour humidity forecasts was found with the application of the tuning procedure, suggesting that methodology and humidity tuning performed have a potential to improve humidity scores of numerical weather prediction models, at least for the shorter forecast ranges.

Note that the two tuning approaches described in the paper were performed independently from each other, mainly because they were motivated by independent problems. However, the two approaches are surely in a strong interaction, as both of them account for the modification of background error standard deviations. One can notice that the two tuning approaches propose to modify the humidity background error standard deviations in the opposite direction. This probably comes from the important fact that the tuning by residual covariances (Chapter 2) was based on predefined statistics sampled with the Ensemble method, while in the tuning of the multivariate balance (Chapter 3), the predefined statistics were sampled by the NMC method. Namely, background error standard deviations sampled by the Ensemble method are found to be smaller than those sampled by the NMC method (*Belo Pereira* and *Berre*, 2006; *Stefanescu et al.*, 2006), which may explain the seemingly opposing guidance on humidity tuning by the two tuning experiments. The change in the sampling method was due to better performance of the ALADIN 3d-var system of HMS using the Ensemble method, which entailed also its operational use. It is, thus, desirable to repeat the tuning of the multivariate balance within the Ensemble sampling framework in the future and to see its interaction with the tuning based on residuals covariances.

Acknowledgements—The authors would like to thank first the great help and encouragement given by *Loïk Berre* during the work related to the tuning of the multivariate balance. We also appreciate very much the advices of *Roger Randriamampianina* and *Claude Fischer* in several issues during our experimentations. The work of *Helga Tóth* is appreciated concerning the implementation of the HL diagnostic tool. We are very grateful for *Edit Adamcsek* for her useful suggestions regarding the text and formulas. This work was supported by the Hungarian National Scientific Foundation (OTKA T049579) and Ministry of Science, Technology and Sports of Republic of Croatia (MZOS project 0004001), as well as scholarship grants awarded by the Regional cooperation for Limited Area modeling in Central Europe and Hungarian Scholarship Board.

References

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljevic, D., Brankovic, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Viterbo, P., Cardinali, C., Thépaut, J-N., 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III : Experimental results. *Q J Roy Meteor Soc 124*, 1831-1860.

Belo Pereira, M. and *Berre, L.*, 2005: The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon Weather Rev 134*, 2466-2489.

Berre, L., 2000: Estimation of synoptic and meso scale forecast error covariances in a limited are amodel. *Mon Weather Rev 128*, 644-667.

Bouttier, F. and *Courtier, P.*, 1999: Data assimilation concepts and methods. *ECMWF Lecture Series (Data assimilation and use of satellite data http://www.ecmwf.int/newsevents/training/lecture_notes/pdf_files/ASSIM/Ass_cons.pdf)*, 72.

Bölöni, G., 2006: Development of a variational data assimilation system for a limited area model at the Hungarian Meteorological Service. *Időjárás 110*, 309-327.

Chapnik, B., Désroziers, G., Rabier, F., Talagrand, O., 2004: Properties and first application of an erro- statisticstuning method in variational assimilation. *Q J Roy Meteor Soc 130*, 2253-2275.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., and *Fisher, M.*, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q J Roy Meteor Soc 124*, 1783-1808.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, Cambridge, 457 pp.

Désroziers, G. and *Ivanov, S.*, 2001: Diagnosis of adaptive tuning of information error parameters in a variational assimilation. *Q J Roy Meteor Soc 127*, 1433-1452.

Désroziers, G., Berre, L., Chapnik, B. and *Poli, P.*, 2006: Diagnosis of observation, background and analysis error statistics in observation space. *Q J Roy Meteor Soc 131*, 3385-3396.

Fischer, C., Montmerle, T., Berre, L., Auger, L., Stefanescu, S., 2005: An overview of the variational assimilation in the Aladin/France numerical weather prediction system. *Tellus 131*, 3477-3492.

Fisher, M., 2003: Background error covariance modelling. *ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, 45-63.

Gandin, L.S., 1963: *Objective Analysis of Meteorological Fields.* Gidrometeorologicheskoe Izdatelstvo, Leningrad. English translation by Israeli Program for Scientific Translations, Jerusalem, 1965.

Gustaffson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B., Mogensen, K.S., and *Thorsteinnson, S.*, 2001: Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus 53A*, 425-446.

Hollingsworth, A., 1987: Objective analysis for numerical weather prediction. Short and medium-range numerical weather prediction. *Proc. of the WMO/IUGG NWP Symp., Tokyo.* Meteorological Society of Japan, 11-59.

Hollingsworth, A. and *Lönnberg, P.*, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I. The wind field. *Tellus 38A*, 111-136.

Horányi, A., Ihász, I. and *Radnóti, G.*, 1996: ARPEGE/ALADIN: A numerical weather prediction model for Central-Europe with the participation of the Hungarian Meteorological Service. *Időjárás 100*, 277-301.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Q J Roy Meteor Soc 114*, 205-240.

Lönnberg, P. and *Hollingsworth, A.*, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus 38A*, 137-161.

Parrish, D.F. and *Derber, J.C.*, 1992: The National Meteorological Centre's spectral statistical interpolation system. *Mon Wea Rev*, 1747-1763.

Rabier, F., Andersson, E., McNally, A., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A., Bouttier, F., 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Q J Roy Meteor Soc 124*, 1809-1829.

Randriamampianina, R., 2006: Impact of high resolution observations in the ALADIN/HU model. *Időjárás 110*, 329-349.

Sadiki, W. and *Fischer, C.*, 2005: A posteriori validation applied to the 3D-VAR ARPEGE and ALADIN data assimilation system. *Tellus* 57A, 21-34.

Siroká, M., Fischer, C., Cassé, V., Brozková, R. and *Geleyn, J-F.*, 2003: The definition of mesoscale selective forecast error covariances for a limited area variational analysis. *Meteor Atmos Phys* 82, 227-244.

Stefanescu, S.E., Berre, L., and *Belo Pereira, M.*, 2006: The evolution of dispersion spectra and the evluation of model differences in an ensemble estimation of error statistics for a limited area analysis. *Mon Wea Rev 134*, 3456-3478.

Talagrand, O., 1998: A posteriori evaluation and verification of analysis and assimilation algorithms. *Workshop on Diagnosis of Data Assimilation Systems*, 2-4 November 1998, ECMWF, 17-18.

Thépaut, J.-N. and *Courtier, P.*, 1991: Four-dimensional data assimilation using the adjoint of a multilevel primitive equation model. *Q J Roy Meteor Soc 117*, 1225-1254.

Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, San Diego, California, 467 pp.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–July 2010, pp. 21–37

Application of the Ensemble Transform Kalman Filter technique at the Hungarian Meteorological Service: Preliminary results

Edit Adamcsek*, Gergely Bölöni, Petra Csomós, and **András Horányi**

Hungarian Meteorological Service
P.O. Box 38, H-1525 Budapest, Hungary
E-mails: adamcsek.e@met.hu; boloni.g@met.hu; csomos.p@met.hu; horanyi.a@met.hu

**Corresponding author*

(Manuscript received in final form March 18, 2010)

Abstract—The provision of the best background error estimates for data assimilation is crucial to derive the best initial conditions for a numerical weather prediction model. Using the Ensemble Transform Kalman Filter (ETKF) method a forecast ensemble can be generated at each analysis time, which can be used to estimate the background error statistics. The error growth implied by the forecast ensemble, which is measured with the forecast differences, characterizes the reliability of the actual background. In the present paper, the theoretical background of the Ensemble Transform Kalman Filter method is presented first. The ETKF transformation technique was applied in our tests for generating analysis ensemble containing the initial conditions for the model, then experiments were run with different setups to create forecast ensembles. At this stage of the research, estimation of the background errors was left aside, and as primary validation of this technique, the ensemble generation scheme and the transformation itself were studied. According to the results it can be said that the ensembles generated with the ETKF rescaling scheme have too small spread, which should be further improved with the proper application of the inflation procedure.

Key-words: data assimilation, Kalman Filter, Ensemble Transform Kalman Filter, transformation matrix, inflation factor, background errors

1. Introduction

Numerical weather prediction (NWP) is an intensively progressing field of meteorology, where the main objective is to numerically solve the set of hydro-thermodynamics equations for the prediction of the future states of the atmosphere. The numerical forecasts can be decomposed into two major steps: the determination of the initial conditions of the model and the model integration

itself. Since a small difference in the initial atmospheric state can evolve into a large change in the forecast (*Lorenz*, 1963), it is indispensable to have an estimation of the state of the atmosphere as accurate as possible. Data assimilation is the method for creating the analysis, the initial condition of the model, by combining observed values and background information (basically numerical weather prediction forecasts). A proper statistical estimation of observational and background errors is crucial in the assimilation process since it ensures the necessary weights of the available information in the course of assimilation. At the Hungarian Meteorological Service (HMS) the ALADIN 3d-var data assimilation system is used in the operational context, and for the time being, background errors are considered to be constant in time (*Bölöni*, 2006) and they are computed from a relatively small ensemble of ALADIN forecasts valid for a period of early spring. However, background forecast errors might be of variable quality depending on the season and weather situation. Kalman Filter is an approach to take into account the time dependency of background errors in the analysis cycle by ensuring the evolution of the background error covariance matrix from one analysis time step to another. The implementation of Kalman Filter, however, is computationally expensive in large dimensional systems such as numerical weather prediction models (*Evensen*, 2007), which excludes the possibility of a real time operational application with the currently available computer technology. In order to reduce the computing costs, ensemble techniques (so-called Ensemble Kalman Filter methods) have been introduced, where the background error covariance matrix is estimated from a small size sample of background forecasts. Different approaches exist for the generation of perturbations for these forecast ensembles, there are two main categories: perturbed observation Ensemble Kalman Filter (*Houtekamer* and *Mitchell*, 1998, 2001; *Hamill* and *Snyder*, 2000) and Ensemble Square-Root Kalman Filter (*Anderson*, 2001; *Whitaker* and *Hamill*, 2002; *Tippett et al.*, 2003; *Ott et al.*, 2004; *Hamill* and *Whitaker*, 2005). The so-called Ensemble Transform Kalman Filter (ETKF), which is a type of the Ensemble Square-Root Kalman Filters, has been used for targeted observations first (*Bishop et al.*, 2001; *Majumdar et al.*, 2001, 2002). Now it is tested for generating initial conditions for ensemble prediction systems (*Wei et al.*, 2006; *Johansson et al.*, 2009). The ETKF technique can be used for creating flow-dependent background errors as well. Since it is computationally inexpensive yet effective scheme, at the Hungarian Meteorological Service this method has been chosen for implementation. In this scheme the background perturbations are converted into analysis perturbations with the help of a transformation matrix. Using these dispersions an analysis ensemble is created, which will constitute the initial conditions for those model integrations which will form the ensemble system for the background error computations.

The paper is organized as follows. In the next section the basic concepts of data assimilation in general and Ensemble Transform Kalman Filter in particular

will be summarized. The transformation matrix, which is the heart of the ETKF procedure is presented in Section 3. Experiments realized so far are described in Section 4. Section 5 provides the summary, discussion, and conclusions of the present study.

2. Data assimilation and Ensemble Transform Kalman Filter

Data assimilation systems provide initial conditions for NWP models by using the actual observations and a background forecast valid at the analysis time as primary information for the assimilation procedure. First, the notations are summarized. From now on let x_a, x_f, and y denote the analysis and background fields, and the observations, respectively. Let n be the degree of freedom of the NWP model ($n \approx 10^7$, typically the number of variables to be determined for the entire model grid), and let p be the number of observations ($p \approx 10^4 - 10^6$). Then x_f and x_a are vectors of size n, and y is a vector of size p. Let $\mathcal{H}$ denote the observation operator, which maps the state variables on the model grid to the observational points, and let H denote its Jacobian matrix at the point x_f (i.e., around the background state, $H = \frac{\partial \mathcal{H}}{\partial x}(x_f)$). The background and observation error covariance matrices P_f and P_o, respectively, are defined as follows:

$$P_f \coloneqq E\left(\varepsilon_f \cdot {\varepsilon_f}^T\right) \text{ and } P_o \coloneqq E(\varepsilon_o \cdot {\varepsilon_o}^T), \tag{1}$$

where $E(\cdot)$ denotes the expected value, $\varepsilon_f = x_t - x_f$ and $\varepsilon_o = \mathcal{H}(x_t) - y$ stand for the background and observation errors, respectively, where x_t is the unknown, true state of the atmosphere. We call the attention that $\varepsilon_f \cdot {\varepsilon_f}^T$ is the diadic product of vectors ε_f and ${\varepsilon_f}^T$ (and similarly for $\varepsilon_o \cdot {\varepsilon_o}^T$). Similarly to Eq. (1), the analysis error covariance matrix can be defined as

$$P_a \coloneqq E(\varepsilon_a \cdot {\varepsilon_a}^T), \tag{2}$$

where $\varepsilon_a = x_t - x_a$. We define the matrix

$$K = P_f H^T (P_o + H P_f H^T)^{-1} \tag{3}$$

called the gain or weight matrix. Optimal least-square data assimilation methods estimate the true state of the atmosphere by the analysis using the following estimation (*Bouttier* and *Courtier*, 1999):

$$x_a = x_f + K(y - \mathcal{H}(x_f)). \tag{4}$$

In an assimilation system based on Eq. (4), the following relation is valid between the analysis and the background error covariance matrices (*Bouttier* and *Courtier*, 1999):

$$P_a = (I - KH)P_f \,, \tag{5}$$

where I denotes the identity matrix. In practice, the P_f background error covariance matrix is often assumed to be constant in time in data assimilation systems, however, it is rather well known that background errors depend on the actual weather situation to a great extent, therefore, it is desirable to release this consideration. The main idea of the Kalman Filter methods is to update matrix P_f at each analysis step. The standard Kalman Filter method was introduced by *Rudolf Emil Kalman* (1960), who has shown that the following equation can be derived for the forecast error covariance matrix at the time level t_i:

$$P_f(t_i) = M_{i-1}^{i} P_a(t_{i-1}) {M_{i-1}^{i}}^{T} + P_M(t_i), \tag{6}$$

where M_{i-1}^{i} is the linear model forecast operator at time level t_{i-1} (acting from time t_{i-1} to t_i, which is a matrix of size $n \times n$), $P_a(t_{i-1})$ is the analysis error covariance matrix at time t_{i-1} and $P_M(t_i)$ is the covariance matrix of the linear model error at t_i. Kalman introduced his technique originally only for the case of linear model M, however, it can be generalized for non-linear model $\mathcal{M}$, as well. Then M is the linearized of $\mathcal{M}$. Updating the P_f matrix based on Eq. (6) is computationally expensive. The evolution of the error covariance matrix requires at least $2n$ model integrations, and the storage of matrices of dimension $n \times n$, which is not practically tractable for an operational model with $n \approx 10^7$. In order to overcome these difficulties ensemble techniques were introduced. Such method is the Ensemble Transform Kalman Filter (ETKF), which is an approximation to the traditional Kalman Filter. It uses an ensemble (i.e., a statistical population) to estimate the error covariance matrices P_a and P_f. The matrix P_a can be estimated for k ensemble members as follows (see *Houtekamer* and *Mitchell*, 2001):

$$P_a \approx \frac{1}{k-1} \sum_{j=1}^{k} (x_{a,j} - \bar{x}_a) \cdot (x_{a,j} - \bar{x}_a)^T. \tag{7}$$

Let us consider

$$Z_a = \frac{1}{\sqrt{k-1}} (z_{a,1}, z_{a,2}, \dots, z_{a,k}), \tag{8}$$

the matrix of size $n \times k$, where the $z_{a,j} = x_{a,j} - \bar{x}_a (j = 1, \dots, k)$ values are the analysis dispersions, i.e., the differences between the jth member $x_{a,j}$ and the

ensemble average $\bar{x}_a$. Note, that with this notation the analysis error covariance matrix P_a can be written from Eq. (7) as

$$P_a = Z_a Z_a^T, \tag{9}$$

which is the product of matrix Z_a and its transpose. (It is remarked here that the Ensemble Transform Kalman Filter is one of the so-called square-root filters, because Z_a is the mathematical square-root of P_a according to Eq. (9); see *Tippett et al.*, 2003.) Using the above notations for Z_f and Z_a, the Kalman Filter method can be derived from the formula

$$Z_f = MZ_a, \tag{10}$$

where M is the linearized of the model operator $\mathcal{M}$. This means that the background dispersions are obtained by integrating the analysis dispersions with the linearized model. In the case of ETKF, instead of making k integrations as in Eq. (10), a relationship is assumed between the dispersions of the analysis and the dispersions of the background, that is

$$Z_a = Z_f T, \tag{11}$$

where $Z_f = \frac{1}{\sqrt{k-1}}\left(z_{f,1}, z_{f,2}, \dots, z_{f,k}\right)$ contains the background dispersions $z_{f,j} = x_{f,j} - \bar{x}_f$ $(j = 1, \dots, k)$, and T denotes the (for the time being unknown) transformation matrix of size $k \times k$ describing this relationship.

The Ensemble Transform Kalman Filter method (the schematic algorithm can be seen in *Fig. 1*) provides an ensemble generation scheme so as to estimate the background error covariance matrix by the covariance matrix of the ensemble forecast perturbations. Let us assume that k $(x_{f,1}, \dots, x_{f,k})$ forecasts have been created at the initial time. Then the matrix of the dispersions Z_f can be computed. The algorithm transforms forecast perturbations into analysis perturbations by the transformation matrix T, whose derivation is shown in the next section. The matrix $Z_a = Z_f T$ contains the dispersions to be added to the analysis x_a computed from the control member (e.g., $x_{f,1}$). The new background members $x_{f,j}^{new}$ are determined by integrating the model with the new analysis ensemble members $x_{a,j}$ $(j = 1, \dots, k)$ as initial states. From their dispersions the new values of the matrix P_f can be computed, then the procedure can be restarted.

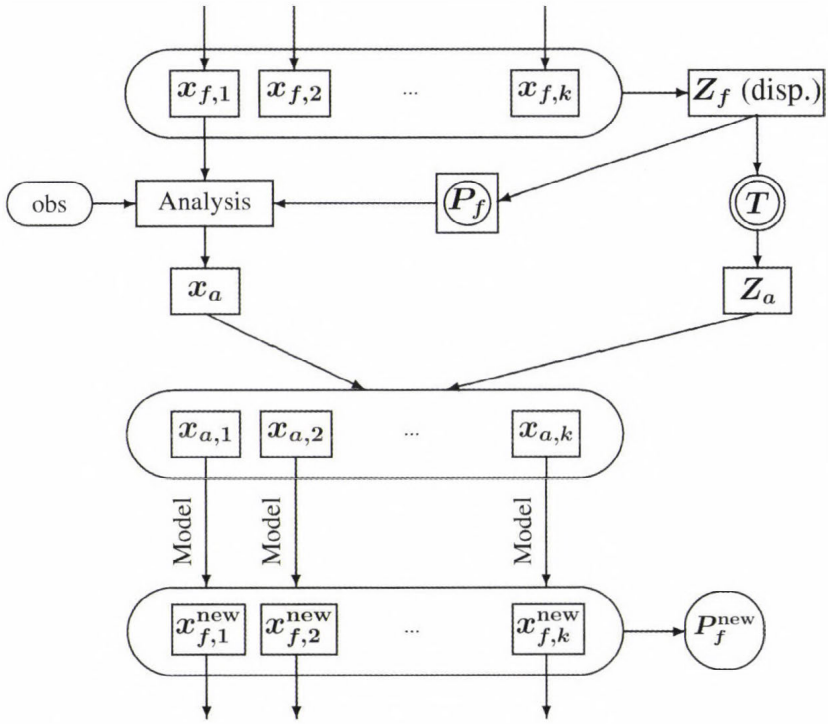


Fig. 1. The schematic algorithm of the Ensemble Transform Kalman Filter.

3. The transformation matrix

According to the Ensemble Transform Kalman Filter method the transformation matrix T is derived from the error covariance update, Eq. (5), provided that the matrix of background dispersions is the square-root of the background error covariance matrix. Hence, matrix T can be determined based on the following formulas:

$$\begin{aligned} Z_a &= Z_f T, \\ P_a &= (I - KH)P_f \ , \quad \text{where } K = P_f H^T (P_o + H P_f H^T)^{-1}, \\ P_f &= Z_f Z_f^T . \end{aligned}$$

Bishop et al. (2001) showed that according to the above assumptions, matrix T can be determined as

$$T = C(\Gamma + I)^{-1/2}, \tag{12}$$

where

$$Z_f^T H^T P_o^{-1} H Z_f = C \Gamma C^T. \tag{13}$$

Thus, matrix C contains the normalized eigenvectors of matrix $Z_f^T H^T P_o^{-1} H Z_f$, and the diagonal matrix Γ contains the corresponding eigenvalues.

The analysis perturbations $z_{a,j}$ obtained by the use of this T matrix are not centered around their mean, so the mean of the generated ensemble will not be identical to the analysis x_a. Ideally, one would like the ensemble mean to be equal to the best available estimate of the true state, that is, the analysis. Therefore, a centering method is required, which converts the analysis dispersions into centered ones. *Wang et al.* (2004) demonstrated that the method is to post-multiply the transformation matrix T by the matrix C^T to form the new, centered analysis dispersions: $Z_a = Z_f T C^T$. Hence, matrix T can be obtained as

$$T = C(\Gamma + I)^{-1/2} C^T. \tag{14}$$

However when the size of the ensemble k is significantly smaller than the rank of the true forecast error covariance, $P_a = Z_a Z_a^T$ underestimates total analysis error variance. To increase the ensemble covariance, the inflation method is used, that is, the analysis perturbations are multiplied by an estimated inflation factor Π: $Z_a = Z_f T \Pi$, which makes the transformation matrix to be computed as

$$T = C(\Gamma + I)^{-1/2} C^T \Pi. \tag{15}$$

In the following, the derivation of the estimation of the inflation factor can be read. The objective of the inflation method is to insure that the background ensemble variance P_f is consistent with the control forecast error. In order to insure this consistency, an equation is derived, which needs to hold. The following consistency diagnostic derived by *Desroziers et al.* (2005) is applied:

$$E(d \cdot d^T) = P_o + H P_f H^T, \tag{16}$$

where $d = y - \mathcal{H}(x_f)$ is the innovation vector: the difference of the observation vector y and its background counterpart $\mathcal{H}(x_f)$, that is the background forecast mapped into the observation space. Now define $\tilde{d}_i$ as the innovation vector at time level t_i, normalized by the square root of the observation error covariance matrix, that is, $\tilde{d}_i = {P_o}^{-1/2}(y_i - \mathcal{H}(x_i^f))$, where index i refers to the time level t_i. Let $\widetilde{\mathcal{H}}$ be the observation operator normalized by the square root of the observation error covariance matrix, that is $\widetilde{\mathcal{H}} = {P_o}^{-1/2}\mathcal{H}$, and let $\widetilde{H}$ be its linearized version. Now Eq. (16) can be rewritten as

$$E\left(\tilde{d}_i \cdot \tilde{d}_i^T\right) = I + \tilde{H} P_f(t_i) \tilde{H}^T, \tag{17}$$

where $P_f(t_i)$ indicates the background error covariance matrix estimated from the forecast ensemble at t_i. For the estimation of the inflation factor Π, this formula can be simplified as *Wang* and *Bishop* (2003) proposed:

$$\tilde{d}_i^T \tilde{d}_i = trace(I + \tilde{H} P_f(t_i) \tilde{H}^T). \tag{18}$$

The data of the next forecast being valid at t_{i+1} is not available at the update time t_i. Therefore, one assumes that the statistics of the next forecast will be similar to that at t_i, which may be a good approximation since the time interval between t_i and t_{i+1} is 6 hours in our experiments. Now the inflation factor is obtained by first checking if $\tilde{d}_i^T \tilde{d}_i$ is equal to $trace(I + \tilde{H} P_f(t_i) \tilde{H}^T)$, i.e., if Eq. (18) holds. If not, an inflation factor is needed: $Z_a = Z_f T \Pi_i$. Since $Z_f T \Pi_i = Z_f \Pi_i T$, the inflated ensemble covariance matrix is $\left(Z_f \Pi_i\right)^T Z_f \Pi_i = \Pi_i^2 Z_f^T Z_f = \Pi_i^2 P_f(t_i)$. To make Eq. (18) hold, the inflated covariance matrix is to be used, and the suitable inflation factor can be determined by the equation:

$$\tilde{d}_i^T \tilde{d}_i = trace\left(I + \tilde{H} \Pi_i^2 P_f(t_i) \tilde{H}^T\right). \tag{19}$$

From Eq. (19) using Eq. (13) we obtain:

$$\Pi_i = \sqrt{\frac{\tilde{d}_i^T \tilde{d}_i - p}{\sum_{j=1}^{k-1} \lambda_j}}, \tag{20}$$

where λ_j, $j = 1, \ldots, k-1$, are the diagonal elements of Γ. Using this inflation factor Eq. (19) will hold, which means that the new, inflated background ensemble variance P_f is consistent with the control forecast error.

4. Experiments

The preliminary experiments with ETKF were based on the ALADIN limited area model (LAM) and were embedded into an operational environment using real observations. A detailed description of the operational ALADIN LAM used in Hungary has been given by *Horányi et al.* (1996, 2006). The ETKF system needs an initial forecast ensemble, which was provided by the operational LAMEPS (Limited Area Model Ensemble Prediction System) used at HMS, which is the downscaling of the singular vector based global ensemble system PEARP (Prévision d'Ensemble ARPege) (*Hágel* and *Mile*, 2009). The members

of the same LAMEPS system were used as lateral boundary conditions for ETKF. As this particular ensemble includes one control forecast $x_{f,1}$ and ten perturbed members $x_{f,2}, \ldots, x_{f,11}$, the ensemble size of eleven was chosen in the experiments as well. The analysis in ETKF was provided by the ALADIN 3d-var data assimilation system used at HMS (*Bölöni*, 2006). Although the ultimate goal is to make the background error covariance matrix P_f flow-dependent, in these first experiments it was not yet updated from the ETKF ensemble, but always remained the static one used operationally. Therefore, the experiments have been limited to the examination of the transformation and perturbations so far.

4.1. Basic test of the transformation matrix

In this first experiment only one data assimilation step was performed. A detailed explanation on the practical computation of the transformation matrix and the perturbations was given by *Csomós* and *Bölöni* (2008). For the estimation of the inflation factor Eq. (20) was used. The innovation vectors were obtained from the initial forecast ensemble and observations (the same as used for the calculation of the transformation matrix). The very first estimation of the inflation factor was computed as $\Pi_1 = 10.06$. The analysis perturbations generated by ETKF were compared with those obtained by running eleven real 3d-var assimilations. The analysis perturbations in the latter case were provided in the following manner (as shown in *Fig. 2*). Eleven 3d-var assimilations were run using the same ensemble of background forecasts as in the ETKF run. In these computations the same background error covariance matrix (now denoted by B referring that it is constant in time) and the same set of observations were used, namely those used in the ETKF run. Finally, the perturbations were obtained by subtracting the analysis $x_{a,1}$ from each $x_{a,i}$ $(i = 1, \ldots, 11)$. Note that now the analysis $x_{a,1}$ is identical to the ETKF analysis x_a (see in *Figs. 2* and *3*).

The goal of this comparison was to investigate in what extent the transformation matrix approximates the real data assimilations. Although the two methods are different and the analysis perturbations are not expected to be the same, some resemblance must be shown. It was found that the spatial structure of the ETKF perturbations was very similar to those provided by the real 3d-var assimilation runs. As an example, one corresponding member of the perturbation sets is shown from the real assimilation and from the ETKF in *Fig. 4*. As a statistical comparison of the two analysis ensembles, the correlation of the corresponding members was computed, i.e., $cor\left(x_{a,i}^{a11}, x_{a,i}^{etkf}\right)$ $(i = 1, \ldots, 11)$, where the superscript corresponds to the derivation of the analysis. In *Fig. 5* the correlations are shown for some standard and physically important verification variables: 300 hPa wind components, 700 hPa relative humidity, and 850 hPa temperature. (In standard atmospheric situation the above-mentioned pressure levels are the corresponding model levels, i.e., level 15, 30, and 36, respectively.)

The analysis ensemble members are highly correlated, their correlation is above 0.974 for all variables, therefore, based on visual and then statistical comparison, it can be concluded that the transformation itself works well.

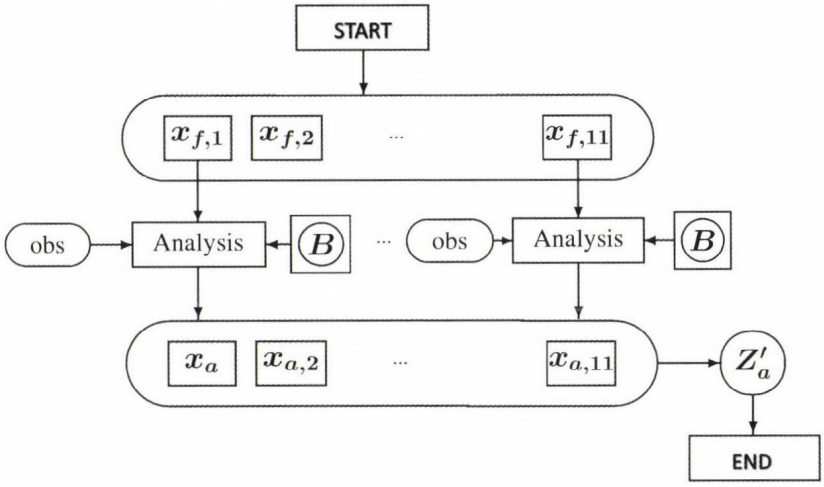


Fig. 2. The scheme of the eleven assimilations used for the test of the transformation matrix.

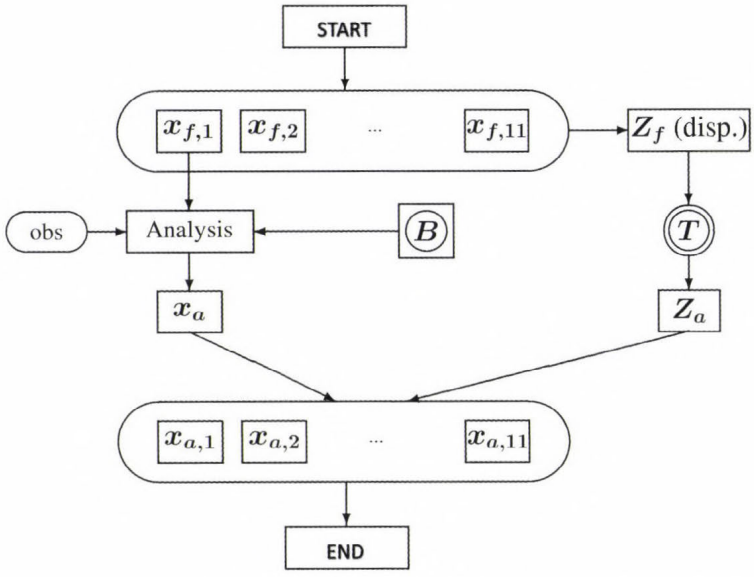


Fig. 3. The schematic algorithm of the Ensemble Transform Kalman Filter based ensemble generation scheme. The background error covariance matrix is now denoted with B referring to its static character used in the test with 11 ensemble members.

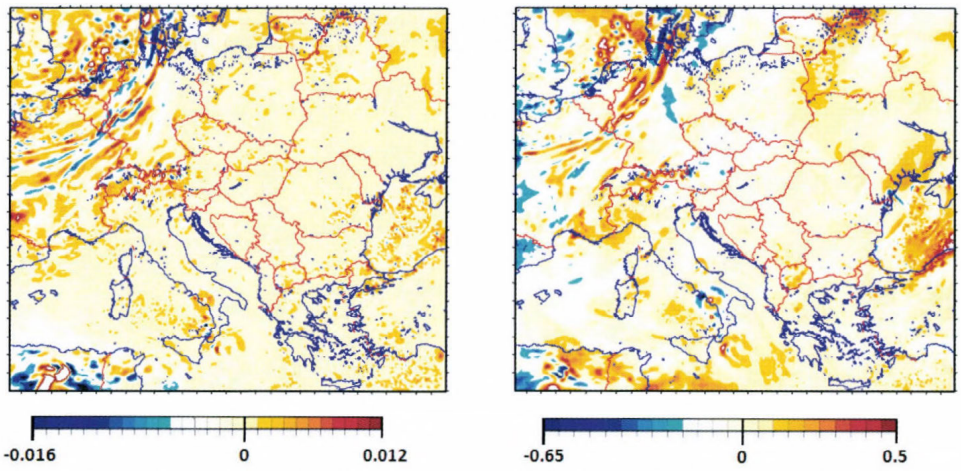


Fig. 4. The 5th analysis perturbation for temperature on the 36th model level obtained from the ETKF ensemble generation scheme (left) and obtained from the eleven assimilations (right).

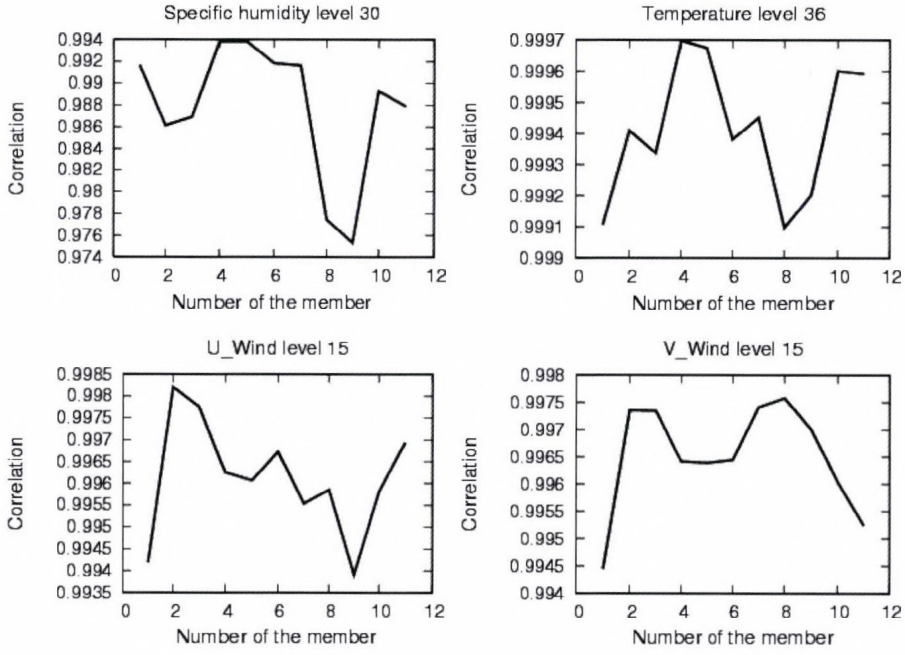


Fig. 5. Correlation of the corresponding analysis ensemble members for specific humidity at around 700 hPa (upper left panel), temperature at around 850 hPa (upper right panel), and zonal and meridional wind components at around 300 hPa (lower panels). On the x-axis the number of the ensemble members is shown, while y-axis shows the correlation value of the two analysis perturbation.

4.2. Experiments with inflation

In this experiment, the algorithm was organized in a cycle as shown in *Fig. 1*. The matrix P_f was still not updated, so the operational forecast error covariance matrix (denoted by B) was used at all analyses steps. The test was run from 00 UTC, September 3 to 18 UTC, September 9, 2008, which is a randomly chosen period, with a 6-hour cycling frequency in agreement with the operational data assimilation cycling setup at HMS. As already mentioned above, the lateral boundary conditions (LBCs) were taken from the operational LAMEPS system. Thus, a sufficient sample of analysis and background dispersions was collected for their statistical analysis.

In order to study the influence of the inflation factor on the perturbations, the following two experiments were run. In the first one, which provided a reference, there was no inflation used. In the second one the inflation factor was estimated in each analysis time as described above and built into the transformation matrix. The evolution of this parameter is shown in *Fig. 6* with solid line, it varies between 10.09 and 22.28.

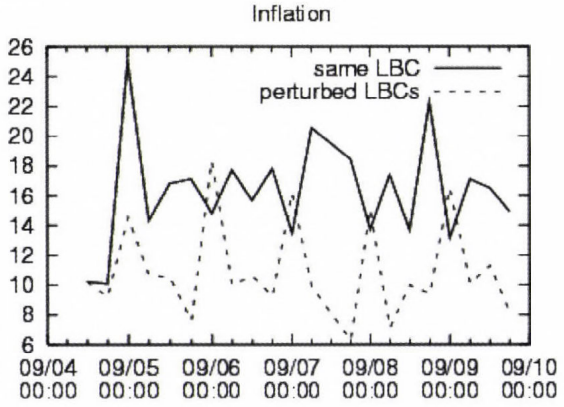


Fig. 6. Evolution of the inflation factor. The solid line represents the experiment with unperturbed LBCs and the dotted line is for experiment with perturbed LBCs.

In these tests the main intention was to focus purely on the impact of the inflation on the perturbations, therefore, the LBCs were kept constant for all the ETKF members, i.e., all of them were coupled to the same LAMEPS member. For the evaluation of forecast dispersions, the rank histograms (also known as Talagrand diagrams) were examined. For more details about the Talagrand diagram, *Hamill* (2000) is referred. These histograms were generated by verifying the forecast ensembles against ECMWF (European Center for Medium-Range Weather Forecasts) analysis. In *Fig. 7* the Talagrand diagram for the background ensembles are shown without and with inflation. It is

apparent that the extreme ranks are overpopulated, which indicates that there is a lack of variability in the ensemble. Although *Hamill* (2000) pointed out, that the U-shape of a Talagrand diagram can indicate conditional bias in the sampling (due to different weather regimes), in this specific case, where such a small sample was evaluated, we do think that the U-shape refers to under-variability. However, certainly the small sample size might prevent to have a really representative diagnostics. As some values appear in the middle ranks on the right panel in *Fig. 7*, the inflation had some minor impact on the variability, but it did not change the spread significantly. In order to gain additional variability in the background ensemble, it was decided that the LBCs are also to be perturbed.

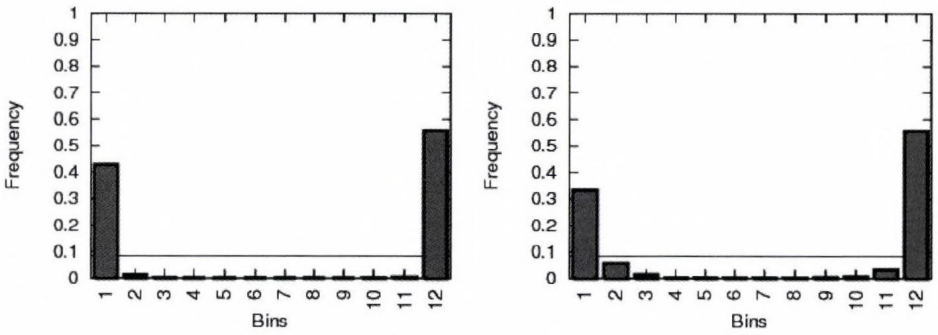


Fig. 7. Talagrand diagrams for forecast ensembles generated without inflation (left) and with inflation (right) for the period 00 UTC, September 3 to 18 UTC, September 9, 2008 for 925 hPa geopotential.

In a limited area model, such as ALADIN, the forecasts depend a lot not only on the initial, but also on the lateral boundary conditions. Thus, in this experiment, uncertainties through LBCs were also included beside the initial perturbations. This was achieved by coupling each member of the background forecast ensemble with a different member of the PEARP global system. It is to be mentioned that there is an inconsistency in this approach, since the perturbation generation methods are fundamentally different in the global and in the LAM system, e.g., the PEARP system applies the singular vector method and not ETKF. This coupling strategy was a technical constraint for we have no access to any global ETKF ensemble system. On the other hand, the effect of this inconsistency on our experiments is unknown. Both the inflated and un-inflated ETKF cycles were rerun with perturbed LBCs. The evolution of the inflation factor can be seen in *Fig. 6* with dashed line, it varies between 6.32 and 18.34. It can be clearly seen, that in this case lower inflation was estimated. The corresponding Talagrand diagrams are shown in *Fig. 8*. In both the inflated and un-inflated cases, the middle ranks are much more populated compared to the previous test. This indicates that the perturbation of the LBCs increases the

ensemble spread in a larger extent than the applied inflation itself. However, it is to be noticed that the Talagrand diagram still shows a U-shape profile, which means that the ensemble does not reflect the true state in spite of the perturbed LBC. This finding suggests that the LBCs are not the only factors in forecast uncertainties and most probably there is still room for improvement in the construction of the initial conditions in order to have a better spread of the forecast ensemble.

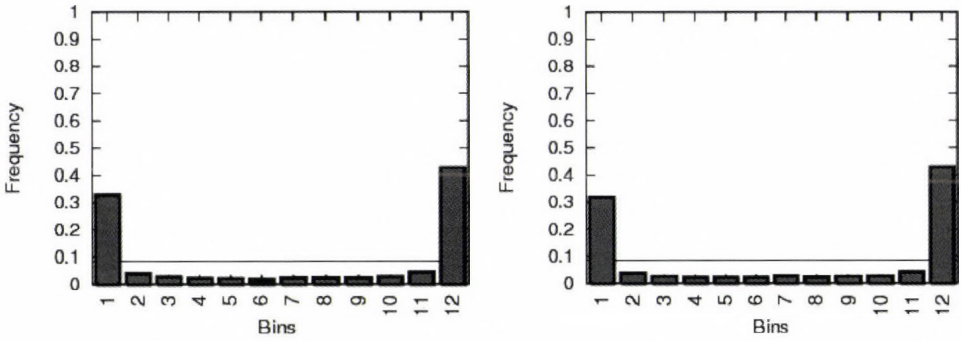


Fig. 8. Talagrand diagrams for forecast ensembles created with perturbed LBCs for the period 00 UTC, September 3 to 18 UTC, September 9, 2008 for 925hPa geopotential. The ensembles were generated without inflation (left) and with inflation (right).

5. Summary and discussion

The primary objective of the application of the Ensemble Transform Kalman Filter (ETKF) method at the Hungarian Meteorological Service is to compute flow-dependent background errors for the operational ALADIN 3d-var data assimilation system. Hence, the implementation was embedded into the operational version of the ALADIN model in Hungary. Yet, the present paper concentrates on the preliminary testing of the ETKF algorithm and performs the first, basic validations of the system. Consequently, the computation of the time-dependent background errors was left aside and the transformation matrix (as heart of the ETKF algorithm) and inflation technique were tested.

As regards the transformation matrix it was found, that the ETKF transformation technique produces analysis perturbations with a reasonable spatial structure as compared to the perturbations created by 3d-var data assimilations. The analysis ensembles generated in two different ways were shown to be similar in terms of correlation. Further tests were carried out to understand the impact of the inflation in an ETKF cycling environment (but still keeping the background errors constant in time). A 7 days period was randomly chosen and the ensemble spread of the background errors was diagnosed with the help of Talagrand diagrams. The Talagrand diagram diagnostics indicated

that the spread of the analysis and forecast ensembles was insufficient in spite of the fact that the inflation slightly improved the characteristics of the forecast ensemble. It was also proven that the perturbed lateral boundary conditions further improved the spread of the ensemble, however, it was still far from being optimal. This latter finding indicates that there is still room for improvements in the initial analyses perturbations with the application of a more optimal inflation factor.

Otherwise, the insufficient spread in the ETKF ensemble might have several sources. One reason may be the operational environment that was used in the implementation. The ETKF study made by *Wang* and *Bishop* (2003) showed good results in a simplified experimental setup assuming that the number, quality, and location of observations are similar at all analysis times, while none of these assumptions are met in an operational data assimilation system. The inflation method was derived by assuming that the statistics are not changing significantly from one analysis time to the other, so this inflation strategy needs to be modified in an operational system.

Another potential cause of the small spread of the ETKF implementation is that in the experiments the matrix P_f was not updated using the generated forecast ensemble, although it is assumed in the ETKF theory. This inconsistent setup was due to the fact that an ensemble of size eleven does not provide a sufficient sample to compute a matrix P_f. One possibility would be to increase the ensemble size, however, this would not be feasible in an operational framework with the present available computational resources. Two other options would be to accumulate the ensemble dispersions over several analysis times or to construct a so-called hybrid P_f matrix with a climatological and a flow-dependent component, where the first would come from a long-time accumulation and the latter from the actual analysis time. Both of these solutions would, however, weaken the flow-dependence of P_f. A future challenge will be to find an affordable solution that keeps the dependence of the background errors on the actual weather.

Since the primary goal for the ensemble generation is to provide a sample of forecast differences for the computation of a time-dependent P_f matrix at every assimilation step, one has to think of the proper diagnostics to be used to evaluate the quality of the ETKF ensemble. In this context, the correlation of the background perturbations and the forecast error are planned to be computed, as *Wei* and *Toth* (2003) suggested. It is believed that this will give a better insight into the characteristics of the ensemble forecasts used in the ETKF procedure.

Acknowledgements—The authors are very grateful to all the participants of the ALADIN international project, who contributed to the development of the numerical weather prediction tools (especially the data assimilation configurations of the ALADIN model) for the study. Particularly, the support of *Claude Fischer* is highly appreciated. We are also grateful to the members of the NWP team of the Hungarian Meteorological Service with special mention to *Sándor Kertész*, who laid down the first theoretical and practical basis of ETKF in Hungary. For the EPS verification

issues, the support of *Edit Hágel* and *Máté Mile* was greatly recognized and appreciated. The useful discussions with *Jelena Bojarova* are also particularly acknowledged. Many thanks for the two anonymous reviewers for their useful comments, which helped to improve the quality of the paper. The authors would like to dedicate this paper to the memory of their colleague and friend *Dezső Dévényi*, who had inevitable role in the initiation of numerical weather prediction in general and data assimilation in particular at the Hungarian Meteorological Service.

References

Anderson, J.L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon Weather Rev 129,* 2884-2903.

Bishop, C.H., Etherton, B.J. and *Majumdar, S.J.*, 2001: Adaptive sampling with the Ensemble Kalman Transform Filter. Part I: Theoretical aspects. *Mon Weather Rev 129,* 420-436.

Bouttier, F. and *Courtier, Ph.*, 1999: Data assimilation concepts and methods. *ECMWF Meteorological Training Course Lecture Series* (http://www.ecmwf.int/newsevents/training/rcourse_notes/-DATA_ASSIMILATION/ASSIM_CONCEPTS/Assim_concepts.html).

Bölöni, G., 2006: Development of a variational data assimilation system for a limited area model at the Hungarian Meteorological Service. *Időjárás 110,* 309-328.

Csomós, P., Bölöni, G., 2008: First steps towards the Ensemble Transform Kalman Filter technique at the Hungarian Meteorological Service. *HIRLAM Newsletter 54,* 9-19.

Desroziers, G., Berre, L., Chapnik, B., and *Poli, P.*, 2005: Diagnosis of observation, background and analysis error statistics in observation space. *Q J Roy Meteor Soc 131,* 3385-3396.

Evensen, G., 2007: *Data Assimilation, The Ensemble Kalman Filter.* Springer-Verlag, Berlin. pp. 37-38.

Hágel, E. and *Mile, M.*, 2009: The limited area ensemble prediction system of the Hungarian Meteorological Service. *ALADIN Newsletter 35.*

Hamill, T.M., 2000: Interpretation of rank histograms for verifying ensemble forecasts. *Mon Weather Rev 129,* 550-560.

Hamill, T.M. and *Snyder, C.*, 2000: A Hybrid Ensemble Kalman Filter-3D variational analysis scheme. *Mon Weather Rev 128*, 2905-2919.

Hamill, T.M. and *Whitaker, J.*, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Mon Weather Rev 133,* 3132-3147.

Horányi, A., Ihász, I. and *Radnóti, G.*, 1996: A numerical weather prediction model for Central-Europe with the participation of the Hungarian Meteorological Service. *Időjárás 100,* 277-301.

Horányi, A., Kertész, S., Kullmann, L. and *Radnóti, G.*, 2006: The ARPEGE/ALADIN mesoscale numerical modeling system and its application at the Hungarian Meteorological Service. *Időjárás 110,* 203-327.

Houtekamer, P.L. and *Mitchell, H.L.*, 1998: Data assimilation using an ensemble filter technique. *Mon Weather Rev 126,* 796-811.

Houtekamer, P.L. and *Mitchell, H.L.*, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon Weather Rev 129,* 123-137.

Johansson, Å., Bojarova, J., Gustafsson, N. and *Vignes, O.*, 2009: ETKF rescaling scheme for HIRLAM: application to ensemble forecasting. *HIRLAM Newsletter 55A, 55-62.*

Kalman, R.E., 1960: A new approach to linear filtering and prediction problems. *ASME Journal of Basic Engineering 82,* 35-45.

Lorenz, E.N., 1963: Deterministic nonperiodic flow. *J Atmos Sci 20,* 130-141.

Majumdar, S.J., Bishop, C.H., Etherton, B.J., Szunyogh, I. and *Toth, Z.*, 2001: Can an ensemble transform Kalman filter predict the reduction in forecast-error variance produced by targeted observations? *Q J Roy Meteor Soc 127,* 2803-2820.

Majumdar, S.J., Bishop, C.H., Etherton, B.J. and *Toth, Z.*, 2002: Adaptive sampling with the Ensemble Transform Kalman Filter. Part II: Field program implementation. *Mon Weather Rev 130,* 1356-1369.

Ott, E., Hunt, B.R., Szunyogh, I., Zimin, A.V., Kostelich, E., Corazza, M., Kalnay, E., Patil, D.J. and *Yorke, J.A.*, 2004: A local Ensemble Kalman Filter for atmospheric data assimilation. *Tellus 56A,* 415-428.

Tippett, M.K., Anderson, J.L., Bishop, C.H., Hamill, T.M. and *Whitaker, J.S.*, 2003: Ensemble square root filters. *Mon Weather Rev 131,* 1485-1490.

Wang, X., Bishop, C.H., 2003: A comparison of breeding and Ensemble Transform Kalman Filter ensemble forecast schemes. *J Atmos Sci 60,* 1140-1158.

Wang, X., Bishop, C.H., Julier, S.J., 2004: Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Mon Weather Rev 132,* 1590-1605.

Wei, M. and *Toth, Z.*, 2003: A new measure of ensemble performance: Perturbations versus Error Correlation Analysis (PECA). *Mon Weather Rev 131,* 1549-1565.

Wei, M., Toth, Z., Wobus, R., Zhu, Y., Bishop, C.H., Wang, X., 2006: Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus 58,* 28-44.

Whitaker, J. and *Hamill, T.M.*, 2002: Ensemble data assimilation without perturbed observations. *Mon Weather Rev 130,* 1913-1924.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–June 2010, pp. 39–55

Sensitivity of local convective precipitation to parameterization of the field capacity and wilting point soil moisture contents

Ferenc Ács[1*], **Ákos Horváth**[2], **Hajnalka Breuer**[1], and **Franz Rubel**[3]

[1]*Department of Meteorology, Eötvös Loránd University*
P.O. Box 32, H-1518 Budapest, Hungary
E-mails: acs@caesar.elte.hu, bhajni@nimbus.elte.hu

[2]*Hungarian Meteorological Service*
Vitorlás u. 17, H-8600 Siófok, Hungary; E-mail: horvath.a@met.hu

[3]*WG Biometeorology, Institute of Medical Physics and Biostatistics*
University of Veterinary Medicine Vienna, Austria; E-mail: franz.rubel@vu-wien.ac.at

[*]*Corresponding author*

(Manuscript received in final form December 10, 2009)

Abstract—Sensitivity of local convective precipitation to parameterization of the field capacity (θ_f) and wilting point (θ_w) soil moisture contents was analyzed. The analysis was performed using simulation results of the Penn State – NCAR MM5 Modeling System obtained on nine days in 2006 and 2007. 24-hour accumulated precipitation fields over the territory of Hungary were considered. Precipitation fields were statistically analyzed. The observed fields were obtained from rain gage data applying a kriging interpolation technique. The agreement between simulated and observed fields was estimated using categorical and continuous verification indices. Significance tests were done to estimate how large (important) or small the obtained differences were.

The main results are as follows. In average, 20 percent relative differences in the θ_f resulted in about 5–10 percent relative differences in the true skill statistic (*TSS*) categorical verification index. In several cases these *TSS* differences are significant on the 10% level. Similarly, in average, 50 percent relative differences in the θ_w resulted in about 10–50 percent relative differences in the *TSS*. These *TSS* differences were significant on the 10% level on eight from the nine days.

The results obtained can be useful in quantifying the strength of the soil moisture – precipitation feedback mechanisms.

Key-words: convective precipitation, field capacity, wilting point, MM5 modeling system, Hungary, true skill statistic categorical verification index, significance test

1. Introduction

The impact of biophysical mechanisms on atmospheric processes is revealed by *Charney* (1975). Since then an array of different sensitivity studies was performed to show the type and rate of the land-surface-atmosphere interaction processes (e.g., *Wolker* and *Rowntree*, 1977; *Deardorff*, 1978; *Shukla* and *Mintz*, 1982; *Sud et al.*, 1988; *Xue*, 1997; *Betts*, 2001; *Betts* and *Viterbo*, 2005). Recently, there are serious efforts not only to explain, but also to quantify the strength of the feedback mechanisms. Among these studies, it could be mentioned the numerical experiment performed by *Koster et al.* (2004) for present-day climate conditions and the study of *Seneviratne et al.* (2006) which was performed for near-future climate conditions in Central Europe.

Among feedback mechanisms, the soil moisture – precipitation feedback is one of the most important (e.g., *Yeh et al.*, 1984; *Oglesby* and *Erickson*, 1989; *Beljaars et al.*, 1996). For a complete analysis, all relevant processes have to be considered: processes occuring in the soil (water transport) and in the planetary boundary layer (triggering of convection), as well as the cloud physical (convection and microphysic) processes (*Hurk van den*, and *Blyth*, 2008). It could not be said that any of these processes is more important than the other one; this hierarchy of the processes according to the importance is highly case-dependent. The importance of soil hydraulic processes is well recognized (*Seneviratne et al.*, 2006), but there is no fully comprehensive analysis. An extensive analysis is made by *Guo et al.* (2006). They argued that the effect of soil moisture (θ) on precipitation should be examined in two distinctly separated parts: (1) when soil moisture affects evaporation and (2) when evaporation affects precipitation. According to *Guo et al.* (2006), the model results are highly dependent on the processes related to the first part. This statement is also confirmed by *Ács* (2002, 2003). Therefore, in this study we investigated the soil – convective precipitation relationship.

From the meteorological point of view, among soil features, soil hydraulic properties are the most important. Among them, the field capacity (θ_f) and the wilting point (θ_w) soil moisture contents have a special role. They depend not only on soil type, but also on the parameterization of soil hydraulic functions (soil water retention and soil water conductivity functions) and the criteria applied in their estimation procedure (*Ács*, 2005). At the same time, θ_f and θ_w are relevant for estimating evapotranspiration (*Chen* and *Dudhia*, 2001). One can pose the question whether local convective precipitation – via sensitivity of evapotranspiration – is also sensitive to the parameterization of θ_f and θ_w. To the best of our knowledge, this question is not discussed to date in the scientific community.

The investigation was made by performing a total of nine one-day case studies during a storm season period in Hungary. Simulations were made by the Penn State – NCAR MM5 modeling system (fifth generation mesoscale model).

The results obtained are statistically evaluated and validated with rain gauge measurements.

2. Model description

2.1. General characteristics of the nonhydrostatic model MM5

Numerical simulations were performed by the version 3 of the Penn State – NCAR MM5 (Fifth-generation Mesoscale Model) modeling system (*Dudhia*, 1993). The model applies a terrain-following sigma coordinate system. The predictive variables are: pressure perturbation, the three momentum components, temperature, specific humidity, and the mixing ratio of the different types of hydrometeors (cloud water, cloud ice, rain, snow, and graupel particles). The partial differential equation system is solved by using a relaxation lateral boundary condition and a radiation upper boundary condition. Model runs were performed using horizontal resolution of 6×6 km, and 26 vertical levels.

Scheme of *Grell et al.* (1994) was applied for parameterization of convection based on rate of destabilization or quasi-equilibrium. A simple single-cloud scheme with updraft and downdraft fluxes and compensating motion determining heating/moistening profile was used. Formation of cloud and precipitation elements are simulated with an explicit bulk microphysical scheme (*Reisner et al.*, 1998) with five different types of hydrometeors: cloud water, cloud ice, rain, snow, and graupel particles. Collision coalescence processes between different types of hydrometeors, furthermore, the diffusion of vapor, freezing of liquid elements, and melting of ice particles are simulated. The equation of conservation is not only solved for the mixing ratios of hydrometeors but for the concentration of cloud ice as well.

The planetary boundary layer (PBL) is described by the local mixing PBL scheme based on the work of *Janjic* (1990, 1994). Compared with other non-local or high-order closure schemes, this PBL scheme proved to be more efficient, because it needs less computer capacity.

2.2. Basic characteristics of the OSU LSM

Land-surface processes are simulated by the OSU LSM (Oregon State University Land-Surface Model). The model consists of a multilayer soil model (*Mahrt* and *Pan*, 1984), a single-layer snow model (*Chen* and *Dudhia*, 2001), and a canopy model (*Pan* and *Mahrt*, 1987).

Evapotranspiration of soil-plant system is estimated as the sum of transpiration and evaporation from the soil and the wetted parts of vegetation. Both transpiration and soil evaporation depend on θ_f and θ_w. They are estimated from soil hydraulic functions (soil water retention (Ψ) and conductivity (K) functions) parameterized after *Campbell* (1974) using *Clapp* and *Hornberger's*

(1978) parameter values, and supposing inhomogeneous areal distribution of soil moisture content.

Detailed description of the model can be found in the work of *Chen* and *Dudhia* (2001).

2.3. Parameterization of θ_f and θ_w

θ_f and θ_w depend not only on soil texture but also on many other factors which are relevant in their estimation: for instance, soil type, transpiration, areal distribution of soil moisture content, and parameterization of soil hydraulic functions. Therefore, their values are quite uncertain.

In this study, soil hydraulic functions referring to Hungarian soils were estimated using *Campbell's* (1974) parameterization and *Clapp* and *Hornberger*'s (1978) parameter values. θ_f values were estimated in two steps as in *Chen* and *Dudhia* (2001): in the first one, the so-called homogeneous (areal distribution of soil moisture content is homogeneous) θ_f values (θ_f^h) were calculated according to the following criteria:

$$pF = \log_{10}[\Psi(\theta_f^h)] = 2.3, \tag{1}$$

$$K_S \left(\frac{\theta_f^h}{\theta_S} \right)^{(2b+3)} = 0.5 \text{ mm day}^{-1}, \tag{2}$$

where θ_S is the saturated soil moisture content, K_S is the saturated soil water conductivity, and b is a dimensionless fitting parameter. The pF value and the corresponding θ_f^h value in Eq. (1) were experimentally determined using the method of *Várallyay* (1973). This criterion applied together with the Várallyay method was common for estimating of θ_f^h in Hungary. Eq. (2) is used for USA soils by *Chen* and *Dudhia* (2001). The effect of the inhomogeneous areal distribution of θ is taken into account as follows:

$$\theta_f = \theta_f^{inh} = \frac{\theta_S}{3} + \frac{2}{3} \cdot \theta_f^h, \tag{3}$$

where Eq. (3) implicitly expresses the fact that the slope of transpiration curve obtained for inhomogeneous distribution of θ is considerably smaller than the slope of transpiration curve obtained for homogeneous distribution of θ (*Ács*, 2003).

θ_w values were also estimated in two steps: as above, the so-called homogeneous θ_w values were calculated according to the following criteria:

$$pF = \log_{10}[\Psi(\theta_w^h)] = 4.2, \tag{4}$$

$$\theta_w^h = 5\rho_b hy, \tag{5}$$

where ρ_b is the bulk soil density and hy is the so-called hygroscopic coefficient. hy is determined by exposing a thin soil layer to an atmosphere saturated with water vapor for a period of about a day. As in Eq. (1), the pF value and the corresponding θ_w^h value in Eq. (4) were experimentally determined using the Várallyay method. In most cases, this expression serves for estimating θ_w^h. If there are no $\Psi(\theta)$ measurements, θ_w^h is usually estimated by expressions like Eq. (5) (*Stefanovits et al.*, 1999). Similarly to the former case, the effect of the inhomogeneous areal distribution of θ is taken into account as

$$\theta_w = \theta_w^{inh} = 0.5\theta_w^h. \tag{6}$$

Eq. (6) has the same meaning as Eq. (3). It expresses the fact that the slope of transpiration curve obtained for inhomogeneous distribution of θ is considerably smaller than the slope of transpiration curve obtained for homogeneous distribution of θ (*Ács*, 2003).

3. Data and numerical experiments

The nine days chosen for analysis include: June 27, 2006, August 1, 2006, August 7, 2006, May 5, 2007, May 22, 2007, June 1, 2007, June 2, 2007, August 10, 2007, and August 20, 2007. On all nine days the weather prevailed was favorable for evolution of local convective storms when the land-surface – air interaction is strong.

3.1. Rain gage data

24-hour accumulated precipitation data for Hungary was acquired from the Hungarian Meteorological Service (HMS). Information was created using data of stations whose type and number varied. In two cases only the main, approximately 100 stations' data were used, in the remaining 7 days the non-automatic gage (about 560 stations) data were incorporated if there was a measurement, otherwise the precipitation was taken as 0. To avoid errors from the rounding and calculation of coordinates, only those stations were added that were more than approximately 6 km from the actual measurements. Overall, on all days the Kriging interpolation error was below 0.5% at the gridcells located in Hungary.

3.2. Basic land-surface parameters

In this study, the USGS-25 category land use dataset was used. Vegetation characteristics in the model domain for May, June, and August are as follows: vegetation cover veg (vegetated part of one complete grid cell) changes between 0.5 and 1.0. In the Carpathian Basin, the prevailing vegetation type is type 2, "dryland, cropland, and pasture". The lowest veg values can be found in the case of type 2. Towards mountain regions, the "grassland" type increases. *veg* is estimated on the basis of leaf area index (LAI). Details concerning the specification of vegetation parameters are described in the work of *Chen* and *Dudhia* (2001).

Table 1. Main physical characteristics of the Hungarian soils

Soil texture	No. of samples	Clay fraction (%)	Sand fraction (%)	Silt fraction (%)	ρ_b (g cm^{-3})
Sand	7	3.84	91.61	4.54	1.32
Loamy sand	20	8.03	78.28	13.69	1.09
Sandy loam	32	14.39	60.29	25.33	1.44
Loam	28	25.36	41.18	33.47	1.45
Sandy clay loam	2	28.80	57.80	13.40	1.55
Clay loam	3	41.53	26.53	31.93	1.3
Clay	37	51.21	6.63	42.19	1.49
Heavy clay	24	60.20	5.37	34.43	1.32
Silt loam	181	19.42	23.33	57.26	1.41
Silty clay loam	154	36.87	9.05	54.08	1.46
Silt	6	5.77	8.05	86.18	1.46

Soil hydraulic properties are obtained from a Hungarian soil data set. Dataset comprises soil data from 576 samples collected in Hungary. Beside hydraulic properties, some other important physical characteristics are also measured. Their mean values for ten different soil textural classes are presented in *Table 1*. Soil textural classes are determined according to criteria given by *Filep* and *Ferencz* (1999). Bulk soil density varies from 1.09 g cm^{-3} for loamy sand to 1.55 g cm^{-3} for sandy clay loam soils. Note, that in the Hungarian soil database there are only two samples for sandy clay loam, three samples for clay loam, and six samples for silt. Noteworthy, that there are only three clay loam samples from the clay loam areas which are quite large in Hungary. Dataset does not contain information about the single map units, that is, we do not know how large the area is to which they are representative. Therefore, it is impossible to upscale this data. Soil hydraulic functions were obtained using results of *Nemes* (2003) and *Fodor* and *Rajkai* (2005).

In this study, θ_f and θ_w values were estimated using all equations. Equation combinations are presented and discussed in the section „Experimental design". The obtained θ_f and θ_w values are presented in *Table 2*. All other soil hydraulic

parameters used in this study are presented in *Table 3*. Note that these soil parameter values can slightly differ from the ones given in the work of *Horváth et al.* (2007). This is caused since slightly different soil data sets and estimation procedures were applied. Soil parameter values represent grid point values. Grid cell values were calculated from the grid point values, but no subgrid-scale variability was considered.

Table 2. Field capacity and wilting point soil moisture content values and the corresponding relative differences for different soil textures estimated by different equation combinations

Soil texture	Θ_f ($m^3\ m^{-3}$)			Θ_w ($m^3\ m^{-3}$)		
	Eqs (1), (3)	Eqs (2), (3)	Rel. diff.	Eqs (4), (6)	Eqs (5), (6)	Rel. diff.
Sand	0.325	0.272	0.163	0.029	0.068	1.345
Loamy sand	0.479	0.348	0.273	0.080	0.141	0.763
Sandy loam	0.379	0.288	0.240	0.064	0.045	0.297
Silt loam	0.408	0.316	0.225	0.080	0.043	0.463
Silt	0.437	0.312	0.286	0.069	0.027	0.609
Loam	0.406	0.296	0.271	0.088	0.044	0.500
Sandy clay loam	0.354	0.273	0.229	0.061	0.032	0.475
Silty clay loam	0.435	0.355	0.184	0.119	0.071	0.403
Clay loam	0.479	0.387	0.192	0.139	0.081	0.417
Sandy clay	0.340	0.330	0.029	0.055	0.056	0.018
Silty clay	0.340	0.273	0.197	0.113	0.032	0.717
Clay	0.489	0.406	0.170	0.147	0.132	0.102
Average	**0.406**	**0.321**	**0.205**	**0.087**	**0.064**	**0.509**

Table 3. Hungarian soil hydraulic parameters as used in the OSU LSM. θ_S = saturated soil moisture content, Ψ_S = saturated soil water retention, K_S = saturated water conductivity, b = dimensionless fitting parameter

Soil texture	b	ψ_s (m)	K_s ($m\ s^{-1}$)	Θ_s ($m^3\ m^{-3}$)
Sand	3.02	0.060	3.26E-05	0.507
Loamy sand	3.90	0.126	2.52E-05	0.598
Sandy loam	3.99	0.143	1.14E-05	0.476
Silt loam	4.18	0.182	2.73E-06	0.487
Silt	3.54	0.223	2.00E-06	0.496
Loam	4.20	0.224	4.58E-06	0.468
Sandy clay loam	4.21	0.132	7.98E-06	0.439
Silty clay loam	5.04	0.234	6.20E-07	0.491
Clay loam	4.74	0.207	3.05E-06	0.580
Sandy clay	3.58	0.890	4.58E-06	0.500
Silty clay	4.06	0.324	1.05E-06	0.453
Clay	6.21	0.228	8.00E-07	0.541

In all runs, that areal distribution of soil texture was used which was originally implemented in the MM5. According to this distribution six textural classes are used from the ten textural classes presented in *Table 1*. They are as follows: loamy sand, sandy loam, loam, sandy clay loam, clay loam and clay.

3.3. Initializations

Initial conditions for MM5 model runs were obtained from the ECMWF analysis and forecast. 00 UTC data were used for all nine days. 100 hPa layer was chosen for the top level of the MM5. Initial soil data (temperature and soil moisture) were also taken from the ECMWF analysis. According to soil moisture content values (about 180–300 mm m^{-3}) in the surface layer, on all nine days $\theta_w < \theta < \theta_f$.

3.4. Experimental design

In the comparative analyses, three runs were performed. The conditions used in the runs are presented in *Table 4*. Run 1 is the reference run. In this run, θ_f was parameterized using Eqs. (1) and (3), while θ_w by Eqs. (4) and (6). Run 2 differs from run 1 only in the parameterization of θ_f. In run 2, θ_f was parameterized using Eqs. (2) and (3). Run 3 and run 1 are different only in the parameterization of θ_w. In run 3, θ_w was parameterized by Eqs. (5) and (6). Comparing the results of runs 1 and 2 (comparison 1), we can analyze how important the parameterization of θ_f is. Comparing the results of runs 1 and 3 (comparison 2), we get an insight into how important the parameterization of θ_w is.

Table 4. Brief description of the conditions in the runs used in the study

	Definition of Θ_f	**Definition of Θ_w**
Run 1	Eqs. (1), (3)	Eqs. (4), (6)
Run 2	Eqs. (2), (3)	Eqs. (4), (6)
Run 3	Eqs. (1), (3)	Eqs. (5), (6)

4. Validation, comparison, and discussion

Storm events occurring on all nine days were analyzed. The model was running for 30 hours to ensure a 6-hour spin-up time. In all cases, the model was started at 00 UTC. 24-hour accumulated precipitation fields were obtained from +6 and +30 hours accumulated precipitation fields. The domain considered has 115 × 49 grid points (in total 5635 grid points). Precipitation fields were statistically analyzed in detail. Statistical analysis is performed only for the territory of Hungary (area of Hungary: 2288 grid points). First, the simulated and observed

fields were compared. Observed fields were estimated from rain gage data applying the ordinary block kriging interpolation technique. In the second step, the agreement between simulated and observed fields was estimated using a categorical and a continuous verification index. At the end, significance tests were also done to estimate how large (important) or small the obtained differences are.

Kriging interpolation technique was used because this method is an approved method for interpolating precipitation data (*Rubel*, 1996). Among the various categorical verification indices, true skill statistics (*TSS*) was chosen. It is based on the two dimensional contingency table (*Table 5*). *TSS* is computed as

$$TSS = \frac{hz - fm}{(h+m)(f+z)}. \tag{7}$$

TSS changes between –1 and +1. For perfect simulation, it tends towards +1. This is a perfect verification measure, because it is independent from the fraction of rain/no rain events. Root mean square error (*RMSE*) was chosen as the continuous verification index. Significance tests were made by using the Student t-test hypothesis. This is appropriate for independent, small-number samples with Gaussian distribution. In the following, we will be dealing with the validation and sensitivity experiments.

Table 5. Contigency table for the calculation of true skill statistics. Symbols: h = hits, f = false, m = misses, z = zero, n = h + f + m + z

	Observation YES	**Observation NO**	**Forecast TOTAL**
Forecast YES	h	f	h + f
Forecast NO	m	z	m + z
Observation TOTAL	h + m	f + z	n

4.1. Validation results

Distribution of the nine-day mean values of the true skill statistics obtained by run 1 (reference run) for different precipitation threshold limits is presented in *Fig. 1a*. Threshold limits denote precipitation intervals from the threshold to its maximum value. Three ranges can be separated. For small threshold limits (small and large precipitation together), the *TSS* values are quite low, about 0.2. For large threshold limits (referring to large precipitation events), the *TSS* values obtained are also small, between 0.05 and 0.2. The distinct break in the values of *TSS* at 13 mm·day^{-1} is caused by the dropout of one day, since there were no precipitation over that threshold. In between (from about 5 to 12 mm day^{-1}), the

TSS values estimated are somewhat larger, about 0.3. The same distribution but obtained on May 5, 2007 is presented in *Fig. 1b*. The main difference with respect to *Fig. 1a* can be observed for small threshold limits. In this case, the *TSS* values are larger, somewhere between 0.4 and 0.7. Note, that abrupt decrease from threshold limit of 13 mm day^{-1} does not exist. Distribution of the nine-day average values of the root mean square errors (*RMSE*) obtained using the reference run for different precipitation threshold limits is presented in *Fig. 2a*. *RMSE* values are smaller (about 10 mm day^{-1}) for larger and larger (about 11–12 mm day^{-1}) for smaller *TSS* values. The same behaviour can also be observed for *RMSE* values calculated on May 5, 2007. This is presented in *Fig. 2b*. In this case, the largest *RMSE* values are about 14 mm day^{-1}.

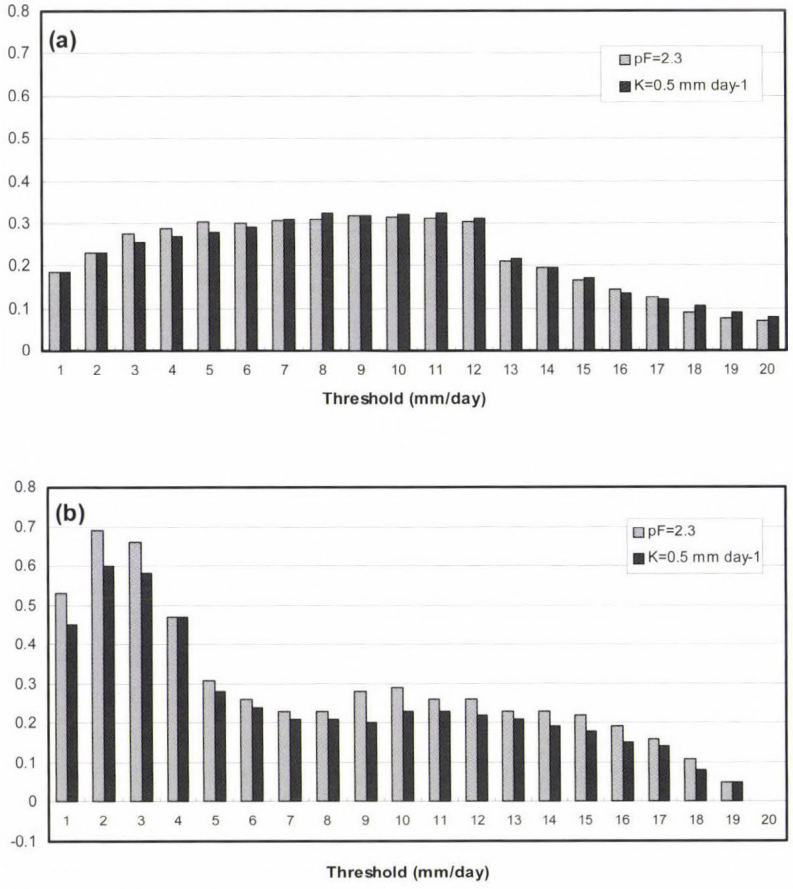


Fig. 1. Distribution of (a) nine-day mean values and (b) actual values on May 5, 2007 of the true skill statistics obtained using run 1 ($pF(\theta_f)$ = 2.3 and $pF(\theta_w)$ = 4.2) and run 2 ($K(\theta_f)$ = 0.5 mm day^{-1} and $pF(\theta_w)$ = 4.2) for different precipitation threshold limits.

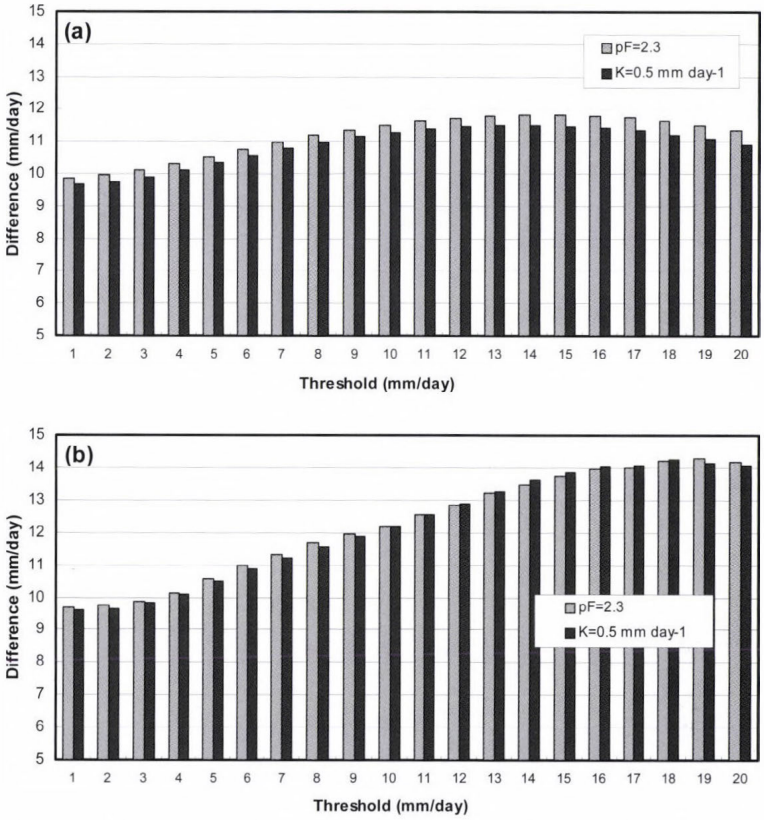


Fig. 2. Distribution of the of (a) nine-day mean values and (b) actual values on May 5, 2007 of the root mean square error obtained using run 1 ($pF(\theta_f) = 2.3$ and $pF(\theta_w) = 4.2$) and run 2 ($K(\theta_f) = 0.5$ mm day^{-1} and $pF(\theta_w) = 4.2$) for different precipitation threshold limits.

4.2. Sensitivity to θ_f parameterization

Comparing results of runs 1 and 2 (comparison 1), we can analyze the effect of θ_f parameterization upon precipitation field prediction. This is presented in *Fig. 1* for *TSS* and in *Fig. 2* for *RMSE* values. According to *Fig. 1*, the sensitivity (difference between the gray and dark columns) seems to be negligible. In average, the relative differences ($[TSS^{pF}–TSS^{K}]/TSS^{pF}$) are under 10 percent. However, according to significance test results, the obtained differences were significant on the 10% level on six from the nine days. That is, relative differences which seem to be small, can be significant. It is also obvious that there is no tendency in the distribution of the *TSS* values obtained by using runs 1 and 2. TSS^{pF} values were higher than TSS^{K} values on three from the six significant days. The sensitivity on a day, when the differences were significant

on the 10% level is presented in *Fig. 1b*. In some cases, for instance for threshold values of 9 and 10 mm day^{-1}, relative difference between the *TSS* values is about 20 percent. Though, it is obvious that these cases are rare. The sensitivity of *RMSE* to θ_f parameterization seems to be also small. In these cases no significance tests were performed.

4.3. Sensitivity to θ_w parameterization

Comparing results of runs 1 and 3 (comparison 2), we can analyze the effect of θ_w parameterization upon precipitation field prediction. Distribution of the nine-day mean values of the true skill statistics obtained using runs 1 and 3 for different precipitation threshold limits is presented in *Fig. 3a*.

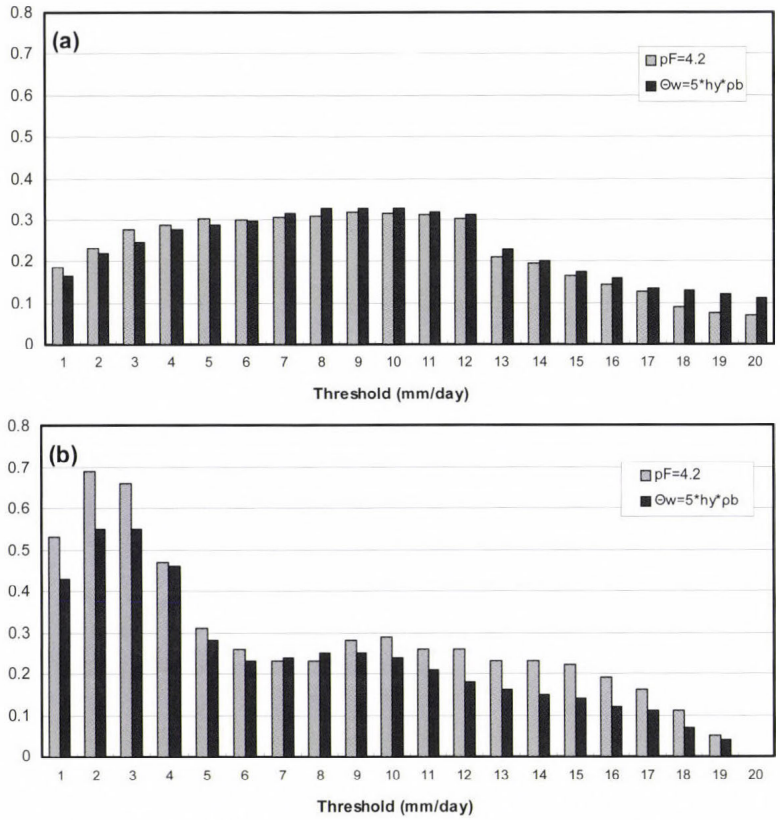


Fig. 3. Distribution of (a) nine-day mean values and (b) actual values on May 5, 2007 of the true skill statistics obtained using run 1 ($pF(\theta_f)$ = 2.3 and $pF(\theta_w)$ = 4.2) and run 3 ($pF(\theta_f)$ = 2.3 and θ_w estimated by *hy*) for different precipitation threshold limits.

As in the former case, the obtained relative differences ($[TSS^{pF}-TSS^{K}]/TSS^{pF}$) seem to be small (under or about 10 percent), but in some cases they were

especially large. For instance, for large threshold values (above 17 mm day^{-1}), relative differences amount 50 percent. According to significance test results, the obtained differences were significant on the 10% level on eight from the nine days, that is, on almost all days. TSS^{hy} values were higher than TSS^{pF} values on five from the eight significant days. The same distribution but obtained on May 5, 2007, when the differences between the *TSS* values were significant, is presented in *Fig. 3b*. In this case, there are large relative differences not only for large (about 30 percent) but also for small threshold values (about 15 percent). Distribution of the nine-day mean values of the *RMSE* for different precipitation threshold limits is presented in *Fig. 4a*. The obtained differences are very small. Distribution of the *RMSE* on May 5, 2007 is presented in *Fig. 4b*. Differences obtained in this case are obviously larger with respect to the differences obtained in *Fig. 4a*. As in the former case, these differences were not analyzed from the point of view of significance tests.

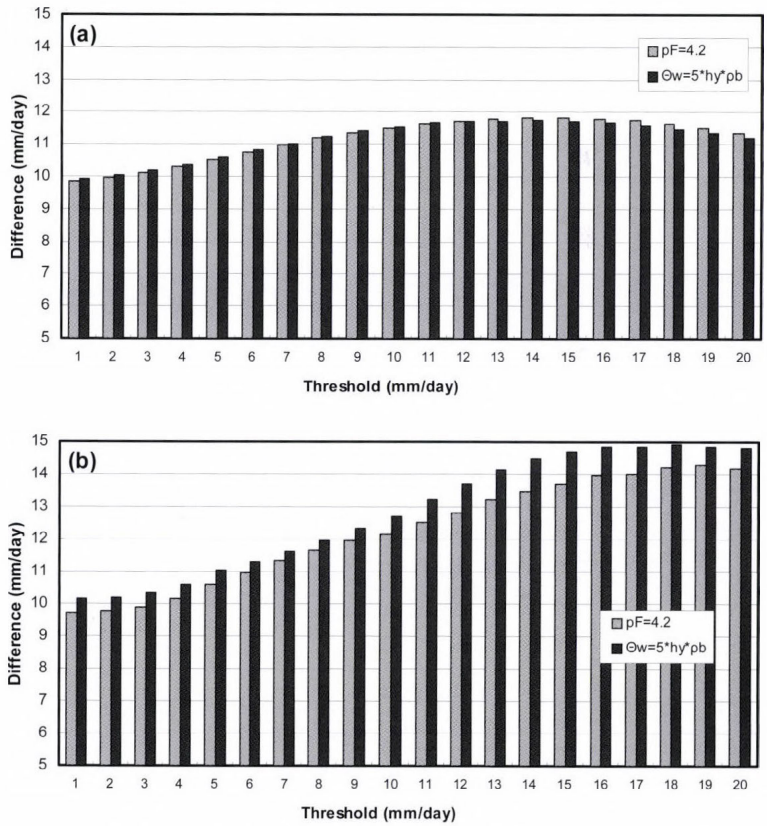


Fig. 4. Distribution of (a) nine-day mean values and (b) actual values on May 5, 2007 of the the root mean square error obtained using run 1 ($pF(\theta_f) = 2.3$ and $pF(\theta_w) = 4.2$) and run 3 ($pF(\theta_f) = 2.3$ and θ_w estimated by *hy*) for different precipitation threshold limits.

4.4. Significance tests applied to other skill measures

Since *TSS* differences were significant on the level of 10% in many cases (for θ_f in six, while for θ_w in eight from the nine days), we decided to performe significance tests for various skill measures. Differences on the level of 10 percent were evaluated for each day separately for in total nine skill measures. The skill measure used were as follows: *POD* = probability of detection ($POD = h/(h + m)$), *FAR* = false alarm ratio ($FAR = f/(h + f)$), *TSS* = true skill statistics, *PC* = proportion of correct hits and misses ($PC = (h + z)/n$), *CSI* = critical success index ($CSI = h/(h + f + m)$), *ORSS* = odds ratio skill score ($ORSS = (hz - fm)/(hz + fm)$), $CHI = \chi 2$-measures of association ($CHI = (hz - fm)^2/((h + f)(m + z)(h + m)(f + z))$, *CR* = correspondence ratio, and *HSS* = Heidke skill score ($HSS = 2(hz - fm)/((h + m)\cdot(m + z) + (h + f)\cdot(f + z))$.

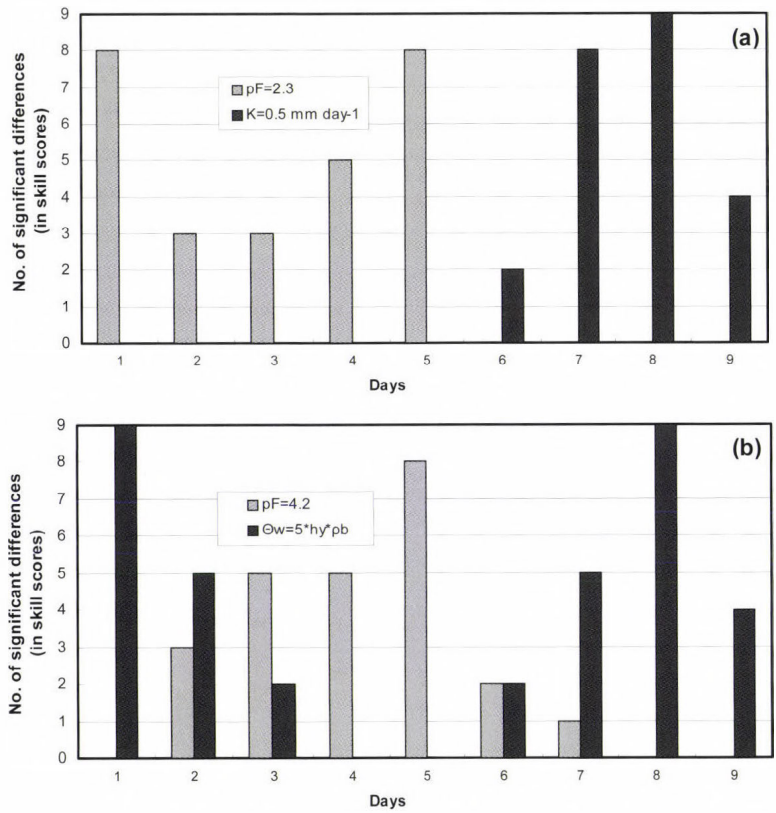


Fig. 5. Number of significant differences on the level of 10% between the skill scores (in total nine) obtained by using conditions (a) in run 1 and run 2 and (b) in run 1 and run 3 for each day separately. Grey column: simulated precipitation obtained by run 1 is closer to the observation, dark column: simulated precipitation obtained by run 2 is closer to the observation.

Number of significant cases for each day separately obtained analyzing the effect of the parameterization of θ_f is presented in *Fig. 5a*. All cases were significant on the eight day (this is August 10, 2007). In all nine cases, results obtained in run 2 were better with respect to the results obtained in run 1. On the first, fifth, and seventh days, the differences were in eight cases significant. Nevertheless, on the 1st and 5th days, the results obtained in run 1 were better than those obtained in run 2. Note that on the sixth day, the differences were only in two cases significant. The same distribution but for θ_w is presented in *Fig. 5b*. The number of significant differences is obviously higher than in the former θ_f case. Differences between all skill measures were significant on the first and the eighth days. The lowest number of the significant differences (in total four) can be observed on the 6th and 9th days. There is no tendency in the distribution of the number of significant cases obtained by using runs 1 and 3.

5. Conclusions

The impact of θ_f and θ_w parameterization upon convective precipitation fields was analyzed. The analysis was performed using simulation results of the Penn State – NCAR MM5 Modeling System obtained on nine days in 2006 and 2007. On all days, actual soil moisture content in the surface layer was higher than θ_w but lower than θ_f.

The considered domain covers the territory of Hungary. Convective precipitation fields were statistically analyzed in detail. First, the simulated and observed fields were compared. The observed fields were obtained from rain gage data applying a kriging interpolation technique. Agreement between the simulated and observed fields was estimated using categorical and continuous verification indices. Significance tests were done to estimate how large (important) or small the obtained differences are.

According to the results, MM5 simulates in an acceptable manner the formation of local convective precipitation field in Hungary. *TSS* values obtained are about 0.2–0.3 on average. We also showed that local convective precipitation is to some degree sensitive to the parameterizations of θ_f and θ_w. What does "to some degree" mean? It can be seen that on average, 20 percent relative differences in the θ_f ($[\theta_f^{pF} - \theta_f^{K}]/\theta_f^{pF}$) (see *Table 2*) caused on average about 5–10% relative differences in the *TSS* ($[TSS^{pF}–TSS^{K}]/TSS^{pF}$). However, in many cases these *TSS* differences were significant on the level of 10%. Similarly, on average, 50 percent relative differences in the θ_w ($[\theta_w^{pF} - \theta_f^{hy}]/\theta_f^{pF}$) (see *Table 2*) caused about 10–50 percent relative differences in the *TSS* ($[TSS^{pF}–TSS^{hy}]/TSS^{pF}$). These *TSS* differences were significant on the level of 10% on eight days. This latter result suggests that sensitivity to θ_w seems to be somewhat higher than sensitivity to θ_f.

Results obtained can be useful in quantifying of the strength of the soil moisture – precipitation feedback mechanisms. It should be mentioned that models used for scientific and operative purposes have completely different soil parameterizations, so different θ_f and θ_w calculation procedures. This study showed that uncertainties in the parameterization of θ_f and θ_w could considerably contribute to scatter of the simulation results (*Koster et al.*, 2004).

Acknowledgements—We thank the help of *K. Rajkai* who gave many remarkable suggestions related to soil physical aspects. The study is financially supported by the National Program for Research and Development (project number NKFP3-00022/2005).

References

Ács, F., 2002: On the relationship between transpiration and soil texture. *Időjárás 106*, 277-290.

Ács, F., 2003: On the relationship between the spatial variability of soil properties and transpiration. *Időjárás 107*, 257-272.

Ács, F., 2005: On transpiration and soil moisture content sensitivity to soil hydrophysical data. *Bound-Lay Meteorol 115*, 473-497.

Beljaars, A.C.M., Viterbo, P., Miller, M.J. and *Betts, A.K.*, 1996: The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil-moisture anomalies. *Mon Weather Rev 124*, 362-383.

Betts, A.K., 2001: Biogeophysical impacts of land use on present-day climate near-surface temperature change and radiating forcing. *Atmos Sci Lett 2*, 39-51.

Betts, A.K. and *Viterbo, P.*, 2005: Land-surface, boundary layer, and cloud-field coupling over the south-western Amazon in ERA-40. *J Geophys Res.110*, p D14108, doi:10.1029/2004JD005702.

Campbell, G.S., 1974: A simple method for determining unsaturated conductivity from moisture retention data. *Soil Sci 117*, 311-314.

Charney, J.G., Stone P.H. and *Quirk, J.W.*, 1975: Drought in the Sahara: A biophysical feedback mechanism. *Science 187*, 434-435.

Chen, F. and *Dudhia, J.*, 2001: Coupling an advanced land surface-hydrology model with the Penn State – NCAR MM5 Modeling System. Part I: Model implementation and sensitivity. *Mon Weather Rev 129*, 569-585.

Clapp, R.B. and *Hornberger, G.M.*, 1978: Empirical equations for some hydraulic properties. *Water Resour Res 14*, 601-604.

Deardorff, J., 1978: Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation. *J Geophys Res 83*, 1889-1903.

Dudhia, J., 1993: A non-hydrostatic version of the Penn State-NCAR Mesoscale Model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon Weather Rev 121*, 1493-1513.

Filep, Gy. and *Ferencz, G.*, 1999: Recommendation for improving the accuracy of soil classification on the basis of particle composition (in Hungarian). *Agrokem Talajtan 48*, 305-317.

Fodor, N. and *Rajkai, K.*, 2005: Estimation of physical soil properties and their use in models (in Hungarian). *Agrokem Talajtan 54*, 25-40.

Grell, G., Dudhia, J. and *Stauffer, D.*, 1994: A description of the fifth generation Penn State/NCAR Mesoscale Model. *NCAR Tech. Note NCAR/TN-398+STR*, 117 pp.

Guo, Y., Dirmeyer, P.A., Koster, R.D., Sud, Y.C., Bonan, G., Oleson, K.W., Chan, E., Verseghy, D., Cox, P., Gordon, C.T., McGregor, J.L., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Mitchell, K., Malyshev, S., McAwaney, B., Oki, T., Yamada, T., Pitman, A., Taylor, C.M., Vasic, R. and *Xue, Y.*, 2006: GLACE: The global land-atmosphere coupling experiment. Part II: Analysis. *J Hydrometeor 7*, 611-625.

Horváth, Á., Ács, F., and *Geresdi, I.*, 2007: Sensitivity of severe convective storms to soil hydrophysical characteristics: A case study for April 18, 2005. *Időjárás 111*, 221-237.

Hurk, B.J. van den and *Blyth, E.*, 2008: WATCH/LoCo Workshop. *GEWEX NEWS 18*, 12-14.

Janjic, Z.J., 1990: The step-mountain coordinate–physical package. *Mon Weather Rev 118,* 1429-1443.

Janjic, Z.J., 1994: The step-mountain Eta coordinate model. Further developments of the convection, viscous sublayer and turbulent closure schemes. *Mon Weather Rev 112*, 927-945.

Mahrt, L. and *Pan, H.L.*, 1984: A two-layer model of soil hydrology. *Bound-Lay Meteorol 29*, 1-20.

Nemes, A., 2003: Multi-scale hydraulic pedotransfer functions for Hungarian soils. *Ph. D. Dissertation, Wageningen University*, ISBN 90-5808-804-9 143 pp.

Koster, D.R., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C.T., Kanae, S., Kowalczyk, A., Lawrence, D., Liu, P., Lu, C-H., Malyshev, S., McAveney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y.C., Taylor, C.M., Versegy, D., Vasic, R., Xue, Y., and Yamada, T., 2004: Regions of strong coupling between soil moisture and precipitation. *Science 305*, 1138-1140.

Oglesby, R.J. and *Ericson, D.J.*, 1989: Soil moisture and the persistence of North American drought. *J Climate 2*, 1362-1380.

Pan, H.L. and *Mahrt, L.*, 1987: Interaction between soil hydrology and boundary-layer developement. *Bound-Lay Meteorol 38*, 185-202.

Reisner, J., Rasmussen, R.M. and *Bruintjes, R.T.*, 1998: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Q J Roy Meteor Soc 124*, 1071-1107.

Rubel, F., 1996: PIDCAP-Quick look precipitation atlas. *Österreichische Beitraege zu Meteorologie und Geophysik 15*, 97 pp.

Shukla, J. and *Mintz, Y.*, 1982: Influence of land-surface evapotranspiration on the earth's climate. *Science 215*, 1498-1501.

Seneviratne, S.I., Lüthi D., Litschi, M. and *Schär, C.*, 2006: Land-atmosphere coupling and climate change in Europa. *Nature 443*, 205-209.

Seneviratne, S.I., Koster, R.D., Guo, Z., Dirmeyer, P.A., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C-H., Mocko, D., Oleson, K.W., Verseghy, D., 2006: Soil moisture memory in AGCM simulations: Analysis of Global Land-Atmosphere Coupling Experiment (GLACE) data. *J Hydrometeorology 7*, 1090-1112.

Stefanovits, P., Filep, Gy., Füleky, Gy., 1999: Soil science (in Hungarian), revised 4th edition. *Mezőgazda*, Budapest, pp. 472.

Sud, Y.C., Shukla, J. and *Mintz, Y.*, 1988: Influence of land-surface roughness on atmospheric circulation and rainfall: a sensitivity study with a general circulation model. *J Appl Meteorol 27*, 1036–1054.

Várallyay, Gy., 1973: Soil moisture potential and new apparatus for the determination of moisture retention curves in the low suction range (0-1 atmospheres). *Agrokem Talajtan 22*, 1-22.

Wolker, J. and *Rowntree, J.R.*, 1977: The effect of soil moisture on circulation and rainfall in a tropical model. *Q J Roy Meteorol Soc 103*, 29-46.

Xue, Y., 1997: Biosphere feedback on regional climate in tropical north Africa. *Q J Roy Meteorol Soc 123*, 1483-1515.

Yeh, T.-C., Wetherald, T. and *Manabe, S.*, 1984: The effect of soil moisture on the short-term climate and hydrology change – A numerical experiment. *Mon Weather Rev 112*, 474-490.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–June 2010, pp. 57–77

Validation of the dynamically adapted high-resolution wind forecasts for the wind power stations in Hungary

Gabriella Szépszó* and **András Horányi**

Hungarian Meteorological Service
P.O. Box 38, H-1525 Budapest, Hungary
E-mails: szepszo.g@met.hu; horanyi.a@met.hu

**Corresponding author*

(Manuscript received in final form March 15, 2010)

Abstract—In the recent years, the interest has been increasing regarding the more and more precise prediction of the low-level atmospheric wind characteristics over Hungary. At the Hungarian Meteorological Service this challenge had been completed with the application of the high-resolution dynamical adaptation originally developed in the framework of the ALADIN cooperation in the late 90s. The dynamical downscaling of the coarser wind fields is realized with the ALADIN meso-scale numerical weather prediction model by a half-hour model integration on 5 km horizontal and 15-level vertical resolution with simplified physical parameterization package. It is shown that this dynamical adaptation step improves the original 10-meter wind forecasts obtained with the 8 km operational version of the ALADIN model. The performance of the method was intensively verified with respect to the local observational data at 80 meters for a wind power station situated at the northwestern part of Hungary. The validation results indicate that the dynamically post-processed forecasts do not have systematic errors, however, the diurnal wind cycle is not properly simulated. In the seven-month evaluation period the low and high wind speeds are overrepresented, whereas the occurrences of intermediate velocities are underestimated. The results are rather satisfactory for the investigated location, however, ideas are also presented for further improvements of the wind predictions.

Key-words: numerical weather prediction, ALADIN limited area meso-scale model, dynamical adaptation, low-level wind, verification

1. Introduction

Precise wind forecasts for the near-surface are crucial for the wind power stations, since they are obliged to make a priori estimation of their daily energy production. For this purpose the only solution is provided by the application of

short-range numerical weather prediction (NWP) models. Nevertheless, the accurate wind information is desired in such detail, which is not ensured by today's typical operational high- (8–10 km) resolution weather predictions. Therefore, the operational forecasts need to be further enhanced by appropriate (dynamical) methods.

Increasing the spatial resolution of the limited area models (LAMs) seems to be a natural and simple way to improve the quality of the predictions. Nevertheless, the horizontal mesh cannot be always excessively refined without reconsideration of the physical parameterization and dynamical core of the NWP models. One obvious and essential change together with the resolution increase is the abandonment of the hydrostatic approximation. The most important phenomena to be more precisely or explicitly described by the non-hydrostatic models are the cloud-, microphysics, and micro-scale processes (e.g., convection), which play key role on the 2–5 km horizontal resolution. The complexity of this kind of models is significantly higher, as besides the vertical velocity their microphysical part is also extended with further prognostic variables associated with the different phases of water. Thanks to these enhancements, not only the precipitation-related processes are better simulated in the non-hydrostatic manner, but also the strong wind events accompanying some of the extreme weather situations.

Today, non-hydrostatic modeling is an area of intensive research and development of numerical weather prediction; however, its computational demands are still too high for its widespread applications. The dynamical adaptation (DADA) procedure (*Žagar* and *Rakovec*, 1999) presents a brilliant (simple and elegant) solution for the more accurate description of the small-scale wind characteristics of the atmosphere. The main advantages of the method are, that it does not require huge efforts on model developments (like for non-hydrostatic models or for the implementation of sophisticated physical parameterization schemes), and additionally, it can be realized with modest computational resources. From the practical point of view, the first step of the DADA method is the interpolation of the coarser-resolution forecast onto a finer-resolution grid (not bringing any new information into the forecast), which is followed by a short model integration on this very high resolution for the adaptation of the large-scale fields to the detailed surface characteristics. This model running takes only 30–45 minutes, while the wind is adapted to the new representation of the orography and possibly to other surface features (*Žagar* and *Rakovec*, 1999). This integration does not use the full, complex numerical model: only those processes are taken into account, which can affect the near-surface wind field and those ones are excluded, which would need more time to develop than the applied 30–45 minutes (for instance, the diabatic processes as cloud water vapor condensation, precipitation formation, or temperature changes due to radiation). The simplification concerns also the vertical levels: since the wind is influenced by the topography mainly in the lower part of the

atmosphere, the number of vertical layers is reduced in the stratosphere and upper troposphere. Further necessary precondition for the success of the method is the capability of the driving model to provide realistic forecasts for the target area, because DADA cannot cure the deficiencies of the original (large-scale) prediction, it only adds some fine-scale details developing due to the more precise surface conditions.

The dynamical adaptation method is a widely used efficient technique to improve the predicted wind field; its positive impacts are anticipated especially over terrains with complex topography (e.g., in mountains or at coastal regions), in case of strong wind events, and when the dynamical forces are determining the flow. Applicability of the dynamical adaptation in the prediction of extreme winds was investigated in the case studies described by *Ivatek-Sahdan* and *Tudor* (2004). They were studying two *bora* events, which usually evolve in winter due to the temperature gradient between the cold continental and mild sea surface at the eastern side of the Adriatic Sea and it is accompanied by strong wind speeds exceeding even 15 m/s in average. The dynamical adaptation procedure was applied on 2 km horizontal resolution for the original NWP products available on an 8 km grid. The "low-resolution" model was capable of predicting the start and end of the extreme events, however, their strengths were underestimated. DADA presented more realistic wind speeds, moreover, it was also shown, that neither the dynamical adaptation with the use of complete physical parameterization set nor the full model forecast on 2 km resolution could outperform it.

Combination of the method with a non-hydrostatic model was explored by *Alexandru* (2004). The 10 km resolution forecast for Romania was dynamically downscaled to 2.5 km, on the one hand, by the non-hydrostatic version of the ALADIN/Romania model, and on the other hand, with its hydrostatic version. The non-hydrostatic downscaling solution was not able to provide additional important details, which would justify its increased computational costs. In another investigation (*Vasiliu*, 2005), a thorough verification of the results produced by the hydrostatic setup was carried out for a 5-month period concentrating on two areas of interest: the Black Sea coast and a mountainous region of Romania. The improvement of the wind fields could be noticed over the mountain region, while at the coastal territory only minor advantages could be detected.

A new scope of the method was presented at the University of Vienna (*Beck et al.,* 2005), at the Hungarian Meteorological Service (HMS) (*Kertész et al.*, 2005), and in Ljubljana (*Žagar et al.*, 2005): dynamical downscaling of the ERA-40 re-analysis data (*Simmons* and *Gibson*, 2000) in order to produce high-resolution wind climatology over the Alpine region, Hungary, and Slovenia, respectively. The fact, that the coarse-resolution re-analysis dataset describes only the large-scale patterns of the flow, justified the relevance of the use of dynamical adaptation. Since the resolution difference between the initial global and the target domains was quite significant (125 km versus 12, 5, and 2.5 km,

respectively), the downscaling was designed in a hierarchical structure with two nested short-range ALADIN model integrations with increasing resolution. In the final step, the special dynamical adaptation configuration of the ALADIN model (i.e., DADA as described above) was applied to reach the desired wind climatology. The results were validated against observations (*Beck* and *Ahrens*, 2006; *Kertész et al.*, 2005; *Žagar et al.*, 2006), and the verification indicated that the dynamical downscaling is able to produce improved and more detailed wind climatology than the initial large-scale data.

In the last decade, the technique is applied at the Hungarian Meteorological Service in order to operationally produce precise wind fields for aviation weather forecasts and in the last few years for the wind power stations to their energy production estimates. From the point of view of the wind energy potential, the northwestern region of Hungary possesses the most advantageous climatological characteristics: this territory lies along the relatively strong and dominant northwesterly flow (*Kertész et al.*, 2005). Some favorable area can be also found over the eastern part of the country, where the northeasterly winds are typical. Until recently, in Hungary the dynamically adapted wind fields could be verified only against 10-meter wind observations, but since the end of 2008, a cooperation has been started with the wind power stations: they continuously provide their measurement data valid at the hub height of the power plant, which make a comprehensive verification of the simulated higher-level wind results possible. The novelty of the utilization of this new source of information is, that the model performance will be known not only on the levels of SYNOP measurements, but also at the heights, where the results are directly used. Therefore, the present study is dedicated to the validation of the dynamical adaptation method applied at HMS for the prediction of the low-level wind fields.

After this introduction, Section 2 describes the most important characteristics of the wind forecasts and observational dataset used as input information for the evaluation, and the employed verification method is also briefly presented. Section 3 is devoted to the thorough analysis of the results obtained by the application of upper-level observational data at Mosonszolnok-Levél. In Section 4, several open issues are addressed and discussed together with those major conclusions, which can be drawn based on the 7-month validation.

2. Methodology

2.1. High-resolution dynamical adaptation of wind forecasts

In the operational practice, the short-range weather forecasts of the Hungarian Meteorological Service are based on the ALADIN meso-scale numerical weather prediction model. ALADIN (*Horányi et al.*, 2006) is a spectral limited area model, where the horizontal meteorological fields are represented by full harmonic functions (2-dimensional Fourier decompositions). In vertical a hybrid

coordinate system is defined (*Simmons* and *Burridge*, 1981): terrain following at the lower model levels and pressure-type for the upper atmosphere. For the vertical equation of motion the model uses the hydrostatic assumption, its prognostic variables are the temperature, horizontal wind components, specific humidity on the model levels, and the surface pressure. Due to the combined semi-implicit and semi-Lagrangian schemes (*Temperton* and *Staniforth*, 1987) applied for the temporal integration, a very high computational efficiency can be realized: for instance, at 8 km horizontal resolution 5-minute integration time step can be used.

Presently, the ALADIN weather forecasts are produced operationally four times a day: at 00, 06, 12, and 18 UTC for 54-, 48-, 48-, and 36-hour periods, respectively. The integration domain covers mainly continental Europe with 8 km horizontal resolution and 49 vertical levels (see *Fig. 1* and *2*). The initial condition for the integration is provided by the 3-dimensional variational data assimilation procedure (*Bölöni*, 2006) developed for the ALADIN model. The time-dependent lateral boundary conditions in 3-hour frequency are ensured by the global NWP model of the European Centre for Medium-Range Weather Forecasts (ECMWF).

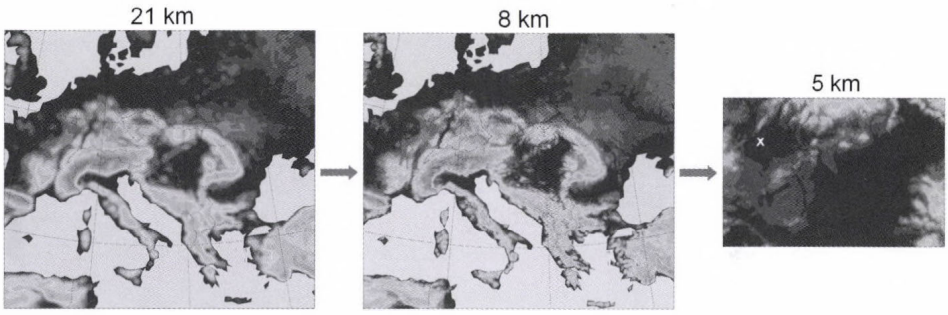


Fig. 1. Orography of the nested model integrations with increasing horizontal resolution: approximately 21 km for the global ECMWF grid (left; in the figure only the selected target integration domain can be seen), 8 km for the operational ALADIN forecast (middle), and 5 km over Hungary for the dynamical adaptation (right). The white "x" on the right represents the selected verification point, Mosonszolnok-Levél.

The raw wind predictions are dynamically refined in the planetary boundary layer, i.e., an additional dynamical adaptation step is performed on the operational model outputs. In practice, the original 8 km resolution information is interpolated onto a 5 km resolution grid (see *Fig. 1*), and the number of vertical layers is reduced in the upper atmosphere resulting in 15 model levels focusing on the planetary boundary layer (see *Fig. 2*). Using these fields as initial and lateral boundary conditions, a simplified model integration lasting 29 minutes (with 60-second time steps) is carried out applying DADA mentioned

in the introduction (*Žagar* and *Rakovec*, 1999). This short dynamical post-processing method adapts the near-surface wind to the more detailed topography taking the "large-scale" forcing from the operational ALADIN model. The final output fields are obtained through vertical interpolation or extrapolation on the desired height coordinates, which are defined at every 10 meters between 10 and 500 m. Vertical post-processing (*Yessad*, 2009) consists of either a linear interpolation for the layers positioned between two model levels or an extrapolation (with the application of logarithmic profile) for heights below the lowest model level (e.g., at 10 meters).

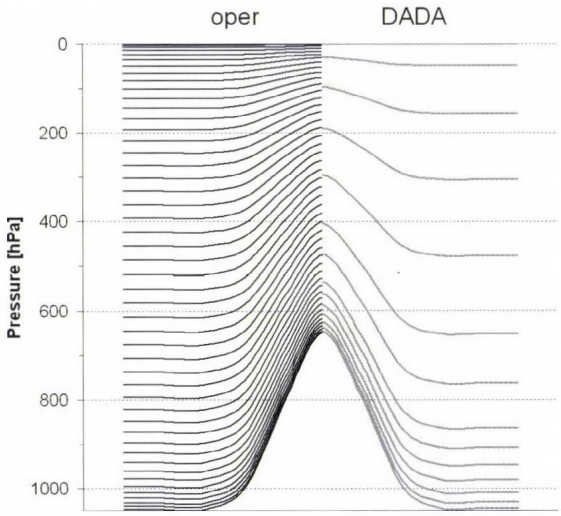


Fig. 2. Hybrid vertical coordinate levels used at the operational forecast ("oper", black curves on the left) and the dynamical adaptation ("DADA", gray curves on the right) with 49 and 15 levels, respectively.

2.2. Input data for the verification

The Hungarian Meteorological Service produces wind information operationally for the wind power plants based on its high-resolution dynamically adapted forecasts starting at 00 UTC and valid for 54 hours. The nearest grid point and height to the given power station is selected as the predicted value, therefore, the 5 km horizontal grid spacing applied at the final dynamical adaptation step allows 3.5 km precision, whereas in vertical the final levels at every 10 meters ensure 5-meter accuracy. The present wind forecast evaluation was carried out for the location of Mosonszolnok-Levél, which is situated in the northwestern part of Hungary at approximately 120-meter height above the sea level (see the white "x" in *Fig. 1*). This power station park stands in a relatively flat area in the gate of the northwesterly flow zone, which is a climatologically favorable

position from the aspect of wind energy production. In this location the wind speed and direction forecasts are provided at the hub-height of the power plants, i.e., at 80 meters above the surface. The forecasts are started from 00 UTC, however, they are used from 9 a.m. (in local time) onwards until 1 a.m. (also local time) of the 3rd forecast day. The output frequency of the forecasted wind data is 15 minutes, and the model data describe the instantaneous wind components of the flow in these time steps. From the observation side, the measurements were available at every 10 minutes for the period of April 19–November 17, 2008. The anemometer is located on an independent tower, at the same height as the hubs of power plants, i.e., in around 78 meters above the surface, and it provides the wind speed as an average over the preceding 10-minute period. It can be read from *Table 1*, that the ALADIN model precisely represents the elevation of the selected point, e.g., its altitude in the model is almost equal with its real height above the sea level.

Table 1. Main characteristics of the input wind datasets: provided wind information, temporal range of the forecasts (LT: in local time), temporal resolution of the produced data, height above the sea level and surface (respectively), and geographical location of the selected point

	Data	**Range**	**Frequency**	**Height**	**Coordinates**
Forecast	Instantaneous wind speed	40 hours from 9 a.m. LT	15 min	125 + 80 m	47.891°N; 17.166°E
Measurement	Mean wind speed (over the last 10 minutes)	-	10 min	124 + 78 m	47.887°N; 17.176°E

Since the data originating from two sources do not correspond to the same type of wind information (i.e., mean and instantaneous values), and moreover, they are available for different time intervals, their direct comparison is not possible. Therefore, the verification was realized on the basis of averages at every half an hour, which were calculated from the wind speed of two time steps in the forecast case and three measurements in the case of observational data. (For instance, for the evaluation at 10:30 UTC, the measurements at 10:10, 10:20, and 10:30 UTC, and the forecasts at 10:15 and 10:30 UTC were respectively averaged, then compared.)

2.3. Verification technique

Evaluation of the wind forecasts for Mosonszolnok-Levél was accomplished for the available seven-month period. On the one hand, the general flavor of the wind predictions during the verification interval is obtained by the unified assessment of all forecast ranges, and on the other hand, the quality of the simulated results is also investigated forecast range by forecast range,

separately. Considering all time steps together, different pairs of forecasts, measurements, and forecast errors are examined in scatter plot diagrams; empirical distribution and density functions were generated on the basis of the simulated and observed data; histograms were also prepared about the forecast errors in function of their signed magnitude. The mean systematic and root mean square (RMS) errors of the forecasts are calculated for all time steps separately and together, as well; the temporal evolution of absolute errors is analyzed for each time step; while the Taylor diagram is used only for selected forecast ranges in every 3 hours. Hereafter, most of the applied verification scores and tools are supposed to be well-known (*Nurmi*, 2003), only the main features of the Taylor diagram are briefly detailed.

The Taylor diagram (*Taylor*, 2001) provides useful information about how well the reference (measured) and the test (simulated) patterns (wind speeds) match each other in terms of their statistical correlation, root mean square (RMS) difference, and ratio of their variances. The geometric relationship between these three statistics allows to plot them into the same diagram. The two (reference and test) samples are represented by two points in the diagram. The radial distances from the origin to each point are proportional to the pattern standard deviation normalized by the observational variance, therefore, the reference point is positioned usually at value 1. The azimuthal position gives the correlation coefficient between the simulated and reference time series. Finally, the distance from the reference point (along concentric circles around it) measures the centered RMS difference between the two fields (as a consequence of the relationship between the three statistics). In the present verification the normalized Taylor diagram was constructed, i.e., the standard deviation of the simulations and the centered RMS difference were standardized with the variance of the observations. The more accurate and consistent the forecast is, the closer its point is positioned to the reference point.

3. Results

First of all, an obvious question is, whether the high- (5 km) resolution dynamical adaptation is able to improve the wind predictions produced on the 8 km operational resolution. For the correct evaluation of this issue, the operational forecasts and the dynamically adapted wind values should be systematically compared at 80-meter height for the 7-month verification period. However, in practice it would mean the necessity of the re-running the ALADIN model for the investigated period. Therefore, it was chosen to study the question by the validation of the 10-meter wind speed values with respect to two nearby observational points (Győr and Mosonmagyaróvár) for an "independent" 3-month period. The 10-meter wind speeds based on the operational 8 km forecasts and the dynamically adapted 5 km ones were assessed in terms of bias and RMS error (*Fig. 3*).

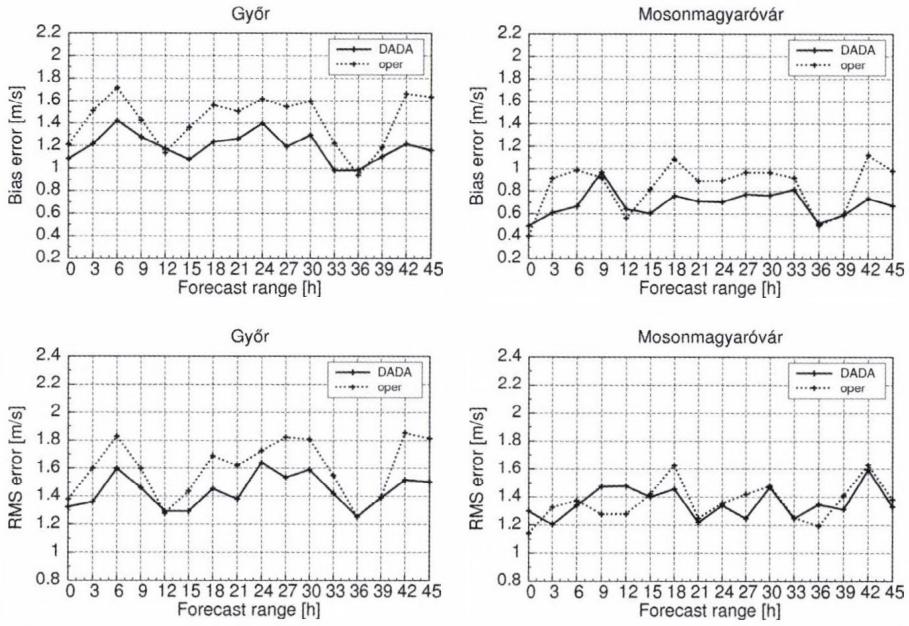


Fig. 3. Mean error (top) and root mean square error (bottom) of the operational ALADIN wind forecasts on 8 km horizontal resolution (dotted curve, "oper") and the 5 km resolution wind forecasts provided by the dynamical adaptation (solid curve, "DADA") for the different time steps (the time steps are considered from the analysis time at 00 UTC) with respect to two SYNOP stations close to Mosonszolnok-Levél: Győr and Mosonmagyaróvár. Verification period: December 1, 2008 – March 1, 2009; level: 10 m.

According to the 3-month verification, it can be generally concluded that the dynamical post-processing procedure improves the wind predictions. The results indicated, that the ALADIN 10-meter wind speed forecasts are hampered by a general overestimation (with 1.4 and 0.8 m/s in average for Győr and Mosonmagyaróvár, respectively), which cannot be fully cured by DADA (the corresponding bias values are 1.2 and 0.7 m/s), since the success of the dynamical adaptation process also depends on the quality of the "large-scale" constraints, which are provided by the 8 km resolution ALADIN forecasts. On the other hand, it is clear that in most time steps DADA outperformed the competing operational forecast in both error characteristics. It is remarked here, that there are some exceptions at the 12h and 36h ranges (i.e., at noon), when the operational results are characterized by identical bias and lower RMS errors than that of the dynamical adaptation (however, the differences between them are rather negligible). These relatively larger reductions in the added value of DADA compared to ALADIN might be associated with the fact, that there are several processes which cannot be more precisely described by this method. Such typical phenomenon is the convection being the most intensive at noon and early

afternoon: since DADA uses only part of the physical parameterization schemes excluding convection, therefore, it has no chance to enhance the results prescribed by the 8 km ALADIN version in cases where these processes have important impact on the wind field. Nevertheless, all this concludes that the application of additional dynamical adaptation on the operational forecasts is useful to provide more accurate 10-meter wind predictions over this territory. Certainly, this conclusion cannot be directly interpreted at higher levels, because the 10-meter wind is an extrapolated diagnosed quantity based on fitting the logarithmic wind profile, whereas the 80-meter wind is rather determined by the dynamics.

Henceforth, the evaluation is concentrating on the 80-meter height, where the wind forecasts and measurements are available for Mosonszolnok-Levél. First, a general overview is given about the main features of the predictions in terms of "lead time independent" indicators, and then the quality is separately investigated in each forecast step. The empirical distribution and density functions (*Figs. 4* and *5*), where the latter ones were constructed by dividing the wind speeds between 0 and 20 m/s into bins of 0.5 m/s for both the model and observational data, provide information about statistical properties of the observed and simulated wind climatologies. (Certainly these are not climatological characteristics in its classical sense, because they are valid only for the investigated period.) It has to be remarked that for the distribution functions the inverse ones are computed (where the relative frequencies of the wind speed occurrences are inverted at the y-axis), since by this visualization the model over- and underestimations, with respect to the observed quantities, can be seen more transparently.

Comparing the histograms and distribution functions for the observations and forecasts, some similarities and also a few differences can be assessed in the simulated and observed wind climatology at the location of Mosonszolnok-Levél:

- Both the measured and forecasted wind speeds cover the range between 0 and 16 m/s.
- The median value (denoted in *Fig. 5*) is similarly around 5 m/s (for the exact values see *Table 2*), i.e., the half of the wind speeds exceeds the 5 m/s threshold in both datasets.
- In the measurements the dominant wind category is between 2.5 and 3 m/s with approximately 7% of relative frequency, in addition, the occurrence of the wind between 2 and 6 m/s exceeds 6% for every bin. The predictions are rather hampered by frequency underestimation between 2 and 7 m/s, which is most characteristic in the 2–3 m/s bins. All this results that the prevailing forecasted wind speeds are in the 5–5.5 m/s category.
- It is also interesting to see the sharp change in the error sign (from occurrence overestimation to underestimation) at the 2 m/s threshold.

(This feature might query the quality of the observations, however, such conclusions are not considered for this short verification period.) The weaker wind speeds below 2 m/s are overrepresented in the simulated results, i.e., the probability of wind speeds from every bin between 0 and 2 m/s is higher in the forecasts than in the reality. Although the level of the overestimation can reach even 80%, the related absolute frequency values are rather small (especially in the category between 0 and 0.5 m/s).

- For the wind speeds above 8 m/s the ALADIN model gives a bit high relative occurrences (though the magnitudes of these errors are even smaller than it was indicated at the lower and intermediate bins).

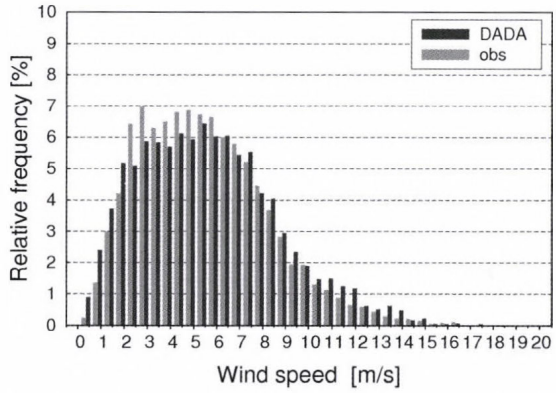


Fig. 4. Discrete density functions of the wind speeds for the model results (black) and the observations (gray). The range of wind speeds are divided into bins of 0.5 m/s. Verification period: April 19–November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

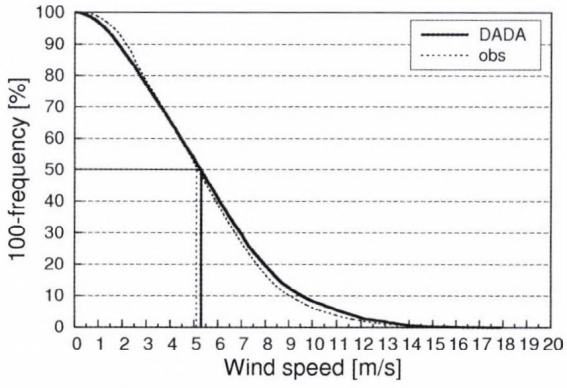


Fig. 5. Inverse empirical distribution functions of the wind speeds for the model results (solid curve) and the observations (dotted curve). Straight lines represent the corresponding median values. Verification period: April 19–November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

Table 2. Mean forecast errors and wind speed characteristics in the simulated and observed datasets (in m/s)

	DADA	Observation
Mean wind speed	5.5	5.4
Median value	5.3	5.1
Mean bias	0.1	-
Mean RMSE	2.2	-

The conclusions drawn on the basis of the empirical density and distribution functions (i.e., the frequency exaggeration of the lower and higher wind speeds and underestimation in the intermediate wind speed intervals) can be also identified in the scatter plot diagrams with several additional details. Looking at the diagram based on the forecast-observation pairs (first panel of *Fig. 6*), at the first glance it can be noticed that most points are situated rather symmetrically around the diagonal, which indicates that there are no systematic errors. This fact is also proved by the slight 0.1 m/s bias value calculated for the 80-meter level (see *Table 2*), and moreover, comparing it with the values for 10 meters, the performance of DADA is significantly improved with the altitude in terms of bias.

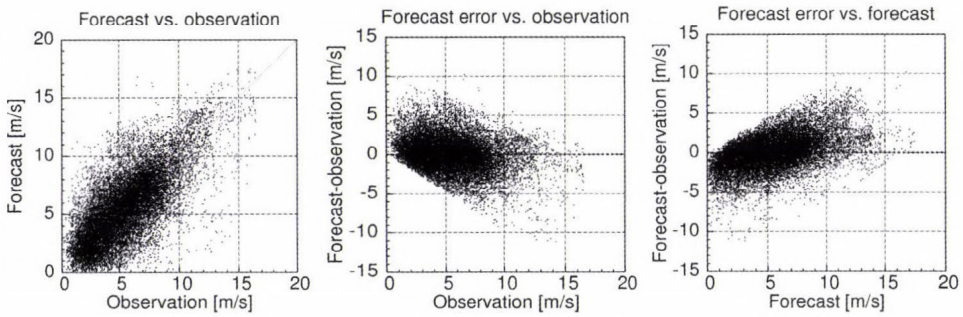


Fig. 6. Scatter plot diagrams (left: forecasts vs. observations, middle: forecast errors vs. observations, right: forecast errors vs. forecasts) for the 5 km resolution wind forecasts provided by the dynamical adaptation procedure. The plots were generated with the use of all time steps. Verification period: April 19–November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

Nevertheless, with a more careful look, also several asymmetric features can be found in *Fig. 6*. For the higher observed wind speeds (exceeding 10 m/s) the model tends to have underestimation, sometimes even with 5–10 m/s (the size of the error is naturally bounded above by the magnitude of the wind speed). This finding is confirmed by the similar diagram for the forecast errors as function of the observations (middle panel of *Fig. 6*), which additionally

shows that the weaker winds (between 0 and 5 m/s) are simulated too large by ALADIN. At the right panel of *Fig. 6* it is also clearly visible that the simulated wind speeds reaching 10 m/s often exceed the measured values. All this is not in contradiction with the conclusions drawn from the histograms and distribution functions (which do not compose pairs from the simulated and observed values): the model is able to predict stronger winds, though these cases are not always at their real occurrences. In other words: when strong winds happen, the model underestimates them, whereas the weaker observed winds are regularly overestimated.

Investigating the scatter plot diagrams in the different time steps (not shown), it can be concluded that wind speeds above 5 m/s are mainly underestimated by the model at around the 12- and 36-hour forecast ranges (at around noon time), and for the interval of 5 and 10 m/s the overestimation has maximum at the 18- and 42-hour forecast ranges (i.e., in the evenings). Based on these features it can be suspected, that the diurnal cycle of the wind speed is not well represented in the simulated results, e.g., it might be shifted with a few hours. This fact can be analyzed in detail in *Fig. 7*, where the mean observed and simulated wind speeds with respect to the forecast range are displayed. (Forecasts over the 47-hour range are neglected since the sample from these time steps was small due to the change from daylight-saving time to the normal one at the end of October.)

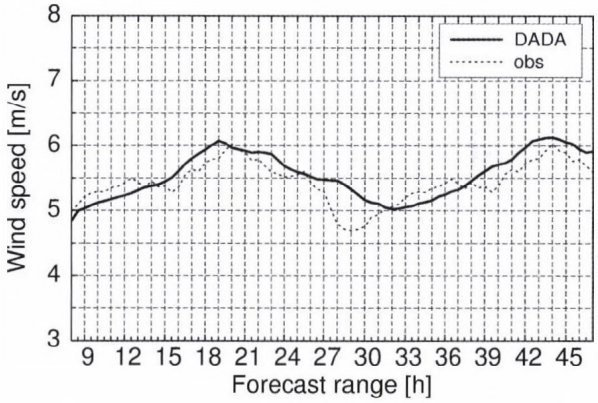


Fig. 7. The simulated (solid curve) and observed (dotted curve) mean wind speed for the different time steps (the time steps are calculated from the analysis time at 00 UTC). Verification period: April 19–November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

It can be easily seen that the displacement of the simulated values is not constant all along the forecast range. Overestimations occur from the afternoons to the mornings and underestimations from the mornings to the very early afternoons. It implies that the afternoon intensification of the wind speed starts

approximately one hour earlier in the model, therefore, the daily maxima are also reached one hour prior (at 19- and 43-hour forecast ranges instead of 20- and 44-hour ones). After getting the maximum, the wind weakening begins later and this process is slower in the simulations than in the reality resulting in a 3-hour delay of the minimum (at 32 hours forecast instead of its typical real occurrence at 29-hour range). Possible reasons behind the introduced error properties can be the insufficient parameterization of the convective processes as well as the weak description of the atmospheric stratification (however, these two factors are not completely independent from each other). Nevertheless, this hypothesis should be checked in the full ALADIN model at first, since as mentioned, DADA is unable to improve those processes which were insufficiently described by the "host" model. Inevitably, more experimentation would be needed to completely understand the physical background of this deficiency.

According to the systematic errors (*Fig. 8*), it can be seen, that in the wind speed forecasts the overestimation is somewhat more dominant than the underestimation resulting in a small (0.1 m/s) positive error for the complete verification period. This fact is supported by the empirical density function of the signed errors (*Fig. 9*), which indicates an almost symmetric pattern. It can be also seen, that the spectrum of the errors covers the range between –8.5 and 8.5 m/s (though there are only few cases with error exceeding 7 m/s), i.e., the negligible mean systematic error is resulted by the balance between the positive and negative differences, which compensate each other. During the two forecasting days, the main systematic behavior of the predictions depends on the actual time steps (*Fig. 8*): the interval between the 8 and 13 hours at the beginning of the forecast period and between the 32 and 37 hours (i.e., 24 hours later) are characterized by small underestimation. The intermediate periods are mostly exacerbated by overestimation, which reaches its maximum (0.7 m/s) at the 28.5-hour forecast range. Looking at the mean and root mean square differences between the simulated and observed wind speeds, it is remarkable, that the model errors do not grow by the forecast range (excluding the first few hours). In each investigated time step the magnitudes of the mean errors basically do not exceed 0.5 m/s; the only exception is the abovementioned maximum at around 28 hours.

Nevertheless, the fact that the mean systematic error is small does not suggest, that the prediction would have equal accuracy in every forecast range, additionally, the minor bias in given time step does not guarantee the perfect prediction, since the positive and negative errors might compensate each other during a longer period. The mean RMSE value is 2.2 m/s for the entire forecast range (*Table 2*). The temporal evolution of the root mean square differences (*Fig. 8*) reveals that the errors are larger (reaching even 2.5 m/s value) when the overestimation is dominant, whereas the model performance is better in the periods of underestimation (the RMSE value is approximately 1.7 m/s in the

8–12 hours and 32–36 hours ranges). One reason might be behind the lower level of accuracy at the overestimation: since the wind speed is always positive, it can be underestimated only with its magnitude (i.e., the size of the negative errors has upper bound), whereas in the case of overestimation, the errors have basically no bound. The ranges of 9 hours and 33 hours are characterized not only by the smallest RMSE values, but they are also in best correlation (approximately 0.85 according to the Taylor diagram at *Fig. 10*) with the observation time series.

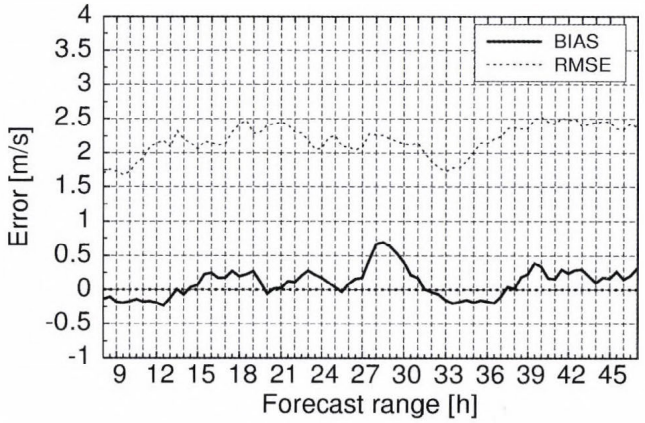


Fig. 8. Mean systematic error (solid curve) and root mean square error (dotted curve) for the different time steps (the time steps are calculated from the analysis time at 00 UTC). Verification period: April 19–November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

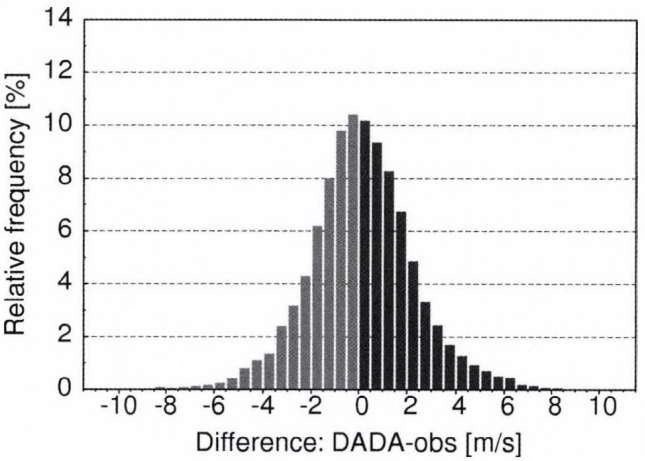


Fig. 9. Discrete density function for the signed model errors. Verification period: April 19–November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

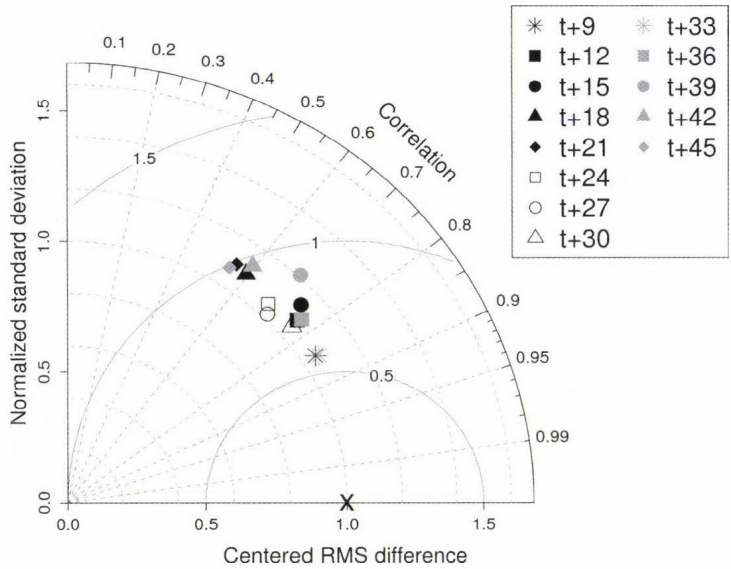


Fig. 10. Normalized Taylor diagram for the model results in different time steps. Verification period: April 19–November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

The temporal variation of the mean absolute error during the seven-month verification period was also examined for the different forecast ranges (only the 9- and 42-hour time steps are presented in *Fig. 11* as examples for the better and worse predictions). Rather large absolute simulation errors can be found in certain days: the deviations can reach even 6–8 m/s (although these errors occur quite seldom). Nonetheless, according to *Fig. 12* one can say, that the errors are under 1 m/s in around 40% of the cases, and the relative frequency of the errors exceeding 5.5 m/s does not reach 1%. The lowest absolute departures from the observations and the smallest error variability can be identified in the time step of 9 hours (*Fig. 11*), the errors mainly remain under 4 m/s all along the period. There are some forecast ranges when the larger errors are more frequent, for instance, the 42-hour step (right panel of *Fig. 11*) when the RMSE is also relatively high. Such dissimilarity between the 9 and 42 hours cannot be explained by the difference in the forecast ranges (i.e., the longer-range forecast is worse), because as mentioned earlier, the error patterns are rather uniform with respect to the lead time (*Figs. 8* and *10*). A more possible and plausible explanation might be again the imprecise description of the daily cycle in the model, which is also confirmed by the similar error characteristics for the 18-hour forecast range, i.e., in the evening one day before. In the Taylor diagram (*Fig. 10*), the points for the time steps of 18h, 21h, 42h, and 45h are positioned in a small group with somewhat negative properties: the predictions in these steps (i.e., in

the evenings) are correlated in a weakest way with the observational time series accompanied by around 0.55–0.6 values; at the same time, they are situated in the largest distance from the reference point bringing the highest RMSE values during the 2-day forecast range. Regarding the temporal evolution of the scores, at some particular time steps (e.g., at 12h, 30h, 36h, not shown) one can conclude some temporary error reduction at the end of summer and the beginning of autumn. These features might be attributed to the particular meteorological situations, nevertheless, it is hard to find any seasonal behavior in the simulation results due to the short investigation period. For more detailed examination of the seasonal characteristics, it would be worthwhile, on the one hand, to extend the verification length, and on the other hand, to perform some smoothing and/or temporal averaging on the data in order to remove its large variability.

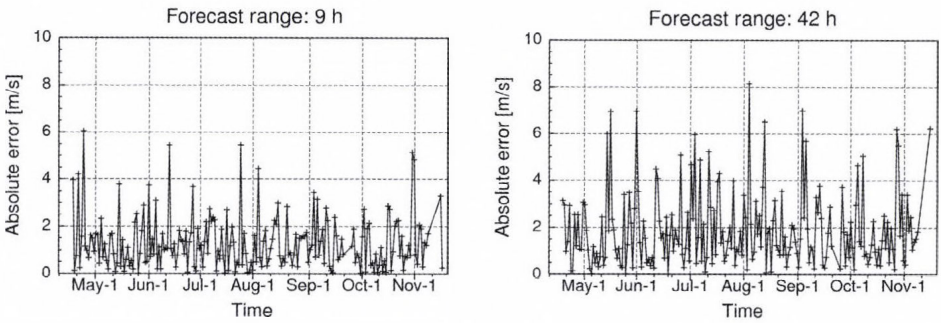


Fig. 11. Temporal evolution of the absolute error during the complete verification period in different forecast ranges (the time steps are calculated from the analysis time at 00 UTC): 9h on the left and 42h on the right. Verification period: April 19 – November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

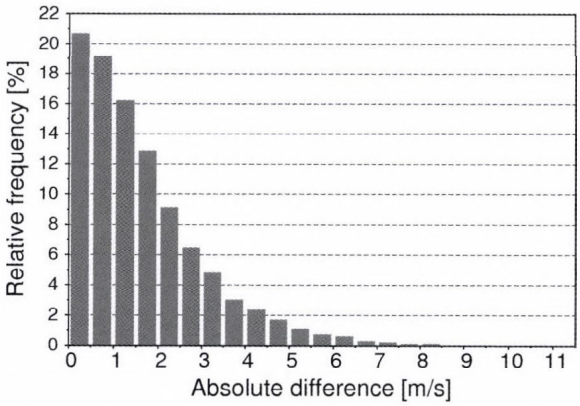


Fig. 12. Discrete empirical density function for the absolute errors. Verification period: April 19 – November 17, 2008; location: Mosonszolnok-Levél; height: 80 m.

4. Summary, conclusion, discussion, and future plans

In this article an overview was given about the applicability of the dynamical adaptation method originally developed by *Žagar* and *Rakovec* (1999) for wind power stations in Hungary. At the Hungarian Meteorological Service this special procedure is applied with the aim of improving operational wind predictions and providing more precise wind speed forecasts required by the wind power stations. The present paper is focusing on the validation results of the method: the dynamically adapted wind forecasts were compared and assessed with measurements at 80-meter height for a single location (namely Mosonszolnok-Levél) in a seven-month period. The selected point to be investigated in detail is situated at the northwestern part of Hungary, in approximately 120-meter height above the sea level. The anemometer stands in 78-meter height above the surface at a relatively strong wind tunnel, where the prevailing northwesterly flow streams over this flat area without barrier.

As far as the general quality of the operational forecasts is concerned, based on the verifications at two nearby SYNOP stations, it can be said that in the northwestern corner of Hungary the wind speeds are principally overestimated at the 10-meter height. The special dynamical downscaling improves the performance of the operational wind predictions, however, it cannot fully eliminate the positive bias. Regarding the main flavor of the fine-scale wind predictions at the higher 80-meter level for Mosonszolnok-Levél, it could be seen as follows:

- The wind speed forecasts during the 7 months are equally characterized by over- and underestimations resulting in almost no (0.1 m/s) systematic error;
- The intervals of 8–13 and 32–37 hours are characterized by small systematic underestimations, which are accompanied with relatively minor RMSE values;
- The intermediate periods are mostly exacerbated by overestimation and higher RMSE values;
- The errors do not have temporal evolution, i.e., they basically do not grow by the forecast range;
- Concerning the climatological features of the wind datasets, the low and high wind speeds (below 2 m/s and over 7 m/s, respectively) are over-represented in the model, whereas the occurrences of intermediate velocities are rather underestimated.

All this can be attributed to the fact, that basically DADA is unable to correctly represent the diurnal wind cycle. This weaker description of the daily velocity cycle might be associated with the deficiencies of the driving model, however, this statement should be thoroughly examined in the future. It would be worthwhile to further extend the investigations with the study of the seasonal behavior of the model (this certainly requires longer verification period),

additionally, the results should be scrutinized with respect to the wind direction, which could supply further hints regarding the strengths and weaknesses of the methodology in general and the ALADIN model in particular.

In spite of some identified deficiencies, the dynamical adaptation method provides not only an efficient, but also a reliable tool for the preparation of accurate wind predictions. At the same time, there are several further possibilities to "tune" the method in Hungary. The first "trivial" solution can be enhancing the spatial resolution of the target dynamical adaptation domain. The choice of 5 km horizontal grid spacing was motivated by that when the method was first tried, the topography was only available on this resolution, therefore, a further refinement of the applied grid would not make sense in the lack of more detailed surface description. Wind predictions are always valid at the gridpoint nearest to the given power plant, i.e., this resolution allows 3.5 km precision. It was indicated in *Table 1* that the orography on this resolution is relatively well described by the model in the investigated point: 125 versus 124 meters. Nevertheless, presently in the operational ALADIN model, an even finer topography is used, namely the GTOPO30 database (*Bliss* and *Olsen*, 1996) having 1 km resolution. The implementation of this detailed surface representation into the dynamical adaptation configuration can promise even better results. As far as the vertical resolution is concerned, the number of model levels could be still extended from the actually used 15 levels putting even more emphasis on the lower part of the atmosphere.

Regarding the Hungarian implementation of the dynamical adaptation procedure, another obvious concern is, whether the application of the half-hour model integrations for the operational outputs provided at every 15 minutes is correct from the aspect of temporal representativeness of the data. I.e., during this construction, the half-hour periods overlap each other, therefore, it is not evident on what time step the output of the DADA procedure can be considered (however, the same boundary conditions at the beginning and end of the short integration suppose that the forecast is valid for the initial time). Originally, the dynamical adaptation post-processing method was developed, when the model outputs were available only with 6- or maximum 1-hour frequency. But today the demands of the partners require the production of forecasts with denser temporal frequency.

One constraint of the method is that it is not able to attenuate the deficiencies of the applied initial and lateral boundary conditions, i.e., only the good-quality large-scale forcings can be improved by the dynamical adaptation procedure. On the other hand, even in case of correct driving fields, there are some small-scale phenomena that cannot be accounted by the larger-scale model, and the simplified dynamical adaptation is unable to additionally reflect them. In these processes other factors than the dynamic forcing are dominating (for instance, local thermal circulations or inversion situations), or for their accurate prediction the complete physical parameterization would have been needed on higher horizontal resolution (*Žagar* and *Rakovec*, 1999). Therefore, it is indispensable

to simultaneously develop the driving model, as well. In the last years, the operational ALADIN model applied at the Hungarian Meteorological Service has been undergoing continuous improvements and there are also some additional plans for enhancing the low-level wind forecasts of the model. The main realized and intended developments are briefly listed hereafter:

- The three-dimensional variational data assimilation scheme of ALADIN (*Bölöni*, 2006) were improved from the point of view of background error computations (*Bölöni* and *Horvath*, 2010) and the inclusion of new and emerging observations types into the assimilation process. The near-future activities focus on the increase of the data assimilation cycling frequency (to 3 hours) and also the computation of flow-dependent background errors with the establishment of an ensemble data assimilation system (*Adamcsek et al.*, 2010).
- The lateral boundary conditions of the ALADIN model were recently updated by the use of the ECMWF/IFS (Integrated Forecast System), which resulted in essential positive impact on the performance of the ALADIN model (*Bölöni et al.*, 2009).
- Wind forecasts can be further enhanced with the application of probabilistic information, which can be obtained by the operational ensemble prediction system of HMS. This short-range limited area ensemble system consists of downscaling the 11-member global ARPEGE-based ensemble of Météo France with the ALADIN meso-scale model (*Hágel*, 2010). Currently, experiments are ongoing for generation of local initial condition perturbations by ALADIN (with the method of singular vectors as employed also in ARPEGE).
- The latest developments are related to the installation of the non-hydrostatic AROME model in light of its near-future operational introduction at HMS (*Horányi et al.*, 2006). At the moment, the model is exploited quasi-operationally providing 36-hour forecast over a domain covering Hungary with 2.5 km horizontal and 60-level vertical resolution.

All these developments contributed and will contribute to the further enhancements of the operational version of the ALADIN model, and consequently, that of its wind predictions. Finally, it is mentioned here that some wind forecast improvements can be also assessed with the statistical post-processing of the raw wind prediction information, which can account for the elimination of systematic model errors. Nevertheless, in the future we are going to concentrate more on the dynamical refinements of the wind forecasts (as listed above) instead of the statistical approach.

Acknowledgements — The authors would like to express their special thanks to *Antal Balogh*, for his invaluable help in the establishment of the mutual beneficial cooperation with the power plant station at Mosonszolnok-Levél. This work was partly supported by the Hungarian National Office for Research and Technology (NKFP, grant No. 3A/0038/2002).

References

Adamcsek, E., Bölöni, G., Csomós, P., and *Horányi, A.,* 2010: The application of the Ensemble Transform Kalman Filter technique at the Hungarian Meteorological Service: Preliminary results. *Időjárás 114,* 21-37.

Alexandru, S., 2004: High-resolution dynamical adaptation of the wind forecast using the non-hydrostatic version of the ALADIN model. *ALADIN/ALATNET Newsletter 27,* 96-100.

Beck, A., Ahrens, B., and *Stadlbacher, K.,* 2005: Impact of nesting strategies in dynamical downscaling of reanalysis data. *Geophys Res Lett 31,* L19101, doi:10.1029/2004GL020115.

Beck, A. and *Ahrens, B.,* 2006: Third Interim Report to ECMWF Special Project – SPATAB01: Alpine regional downscaling of reanalysis data using the LAM ALADIN. Available at http://www.ecmwf.int/about/special_projects/finished_projects/beck_downscaling_reanalysis_data/report_2006.pdf.

Bliss, N.B. and *Olsen, L.M.,* 1996: Development of a 30-arc-second digital elevation model of South America. In *Pecora Thirteen, Human Interactions with the Environment – Perspectives from Space, Sioux Falls, South Dakota.*

Bölöni, G., 2006: Development of a variational data assimilation system for a limited area model at the Hungarian Meteorological Service. *Időjárás 110,* 309-328.

Bölöni, G. and *Horvath, K.,* 2010: Diagnosis and tuning of background error statistics in a variational data assimilation system. *Időjárás 114,* 1-19.

Bölöni, G., Kullmann, L., and *Horányi, A.,* 2009: Use of ECMWF lateral boundary conditions and surface assimilation for the operational ALADIN model in Hungary. *ECMWF Newsletter 119,* 29-35.

Hágel, E., 2010: The quasi-operational LAMEPS system of the Hungarian Meteorological Service. *Időjárás 114,* 121–133.

Horányi, A., Kertész, S., Kullmann, L., and *Radnóti, G.,* 2006: The ARPEGE/ALADIN mesoscale numerical modeling system and its application at the Hungarian Meteorological Service. *Időjárás 110,* 203-228.

Ivatek-Sahdan, S. and *Tudor, M.,* 2004: Use of high-resolution dynamical adaptation in operational suite and research impact studies. *Meteorol Z 13,* 99-108.

Kertész, S., Szépszó, G., Lábó, E., Radnóti, G., and *Horányi, A.,* 2005: Dynamical downscaling of the ECMWF ERA-40 re-analyses with the ALADIN model. *ALADIN/ALATNET Newsletter 28,* 78-83.

Nurmi, P., 2003: Recommendations on the verification of local weather forecasts. *Technical Memorandum 430,* European Centre for Medium-range Weather Forecasts, Reading, UK.

Simmons, A.J. and *Burridge, D.M.,* 1981: An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon Weather Rev 109,* 758-766.

Simmons, A.J. and *Gibson, J.K.,* 2000: The ERA-40 Project Plan. *ERA-40 Project Report Series 1.*

Taylor, K.E., 2001: Summarizing multiple aspects of model performance in single diagram. *J Geophys Res 106,* D7, 7183-7192.

Temperton, C. and *Staniforth, A.,* 1987: An efficient two-time-level semi-Lagrangian semi-implicit integration scheme. *Q J Roy Meteorol Soc 113,* 1025-1039.

Vasiliu, S., 2005: High-resolution dynamical adaptation of the ALADIN/Romania model's surface wind forecast. *Romanian Journal of Meteorology 7,* 29-40.

Yessad, K., 2009: Full-pos in the cycle 36 of ARPEGE/IFS. Technical documentation, Météo France. Available at http://www.cnrm.meteo.fr/gmapdoc/IMG/ps_ykfpos36.ps.

Žagar, M. and *Rakovec, J.,* 1999: Small-scale surface wind prediction using dynamic adaptation. *Tellus 51A,* 489-504.

Žagar, M., Žagar, N., Cedilnik, J., Gregorič, G., and *Rakovec, J.,* 2005: High-resolution wind climatology from ERA-40. *ALADIN/ALATNET Newsletter 28,* 101-104.

Žagar, N., Žagar, M., Cedilnik, J., Gregorič, G., and *Rakovec, J.,* 2006: Validation of mesoscale low-level winds obtained by dynamical downscaling of ERA40 over complex terrain. *Tellus 58A,* 445-455.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–June 2010, pp. 79–100

Impact of extreme rainfall and soil moisture on flash flood generation

Szabolcs Czigány*, Ervin Pirkhoffer, and **István Geresdi**

Department of Soil Sciences and Climatology, Institute of Environmental Sciences, Faculty of Sciences, University of Pécs
Ifjúság u. 6, H-7624 Pécs, Hungary; E-mail: szczigany@gamma.ttk.pte.hu

**Corresponding author*

(Manuscript received in final form January 28, 2010)

Abstract—Convective rainstorms frequently result in floods in small mountainous watersheds. This type of floods, called flash floods, may pose significant threat to lives and properties and present a particularly challenging natural phenomenon to predict with a sufficient lead time.

Our approach to predict flash floods is through runoff modeling. Hydrological models, however, require a plethora of input data. Some temporally variable data, such as precipitation intensity and soil moisture, significantly influence the calculated cumulative outflow and the forecasted peak discharge. In the present paper, the spatio-temporal behavior of soil moisture is studied with the means of time domain reflectrometry (TDR) technique and ArcGIS tools at 14 monitoring stations in an experimental watershed, SW Hungary. Based on field measurements, two minor discharge peak (MDP) events of the studied Sás Stream are simulated using the HEC Hydrologic Modeling System (HEC-HMS). Under the aforementioned environmental conditions, the largest flood event of 2008 on Sás Stream is simulated by a hindcast method. As no environmental data were available for this flood event, the majority of the parameters were estimated from monitored data and other data resources.

Soil moisture varies considerably in both space and time, but the variations among the measurement stations are predictable. Thus, we believe that from measurements at one location of the watershed, soil moisture can be estimated for the entire watershed.

Rainfall varies greatly, which constrains predictability. A possible solution to obtain the rainfall forecasts is the application of radar observation and meteorological models.

We found that HEC-HMS is suitable for model simulations as well as for the generation of a nationwide flash flood warning system. The HEC-HMS is only restrictedly applicable to reproduce flood events with low peak discharge as well as floods with multiple peaks. We also concluded that HEC-HMS exaggerates the impact of volumetric water content on flash flood generation.

Key-words: flash flood, hazard, soil moisture, monitoring, modeling, HEC-HMS

1. Introduction

In Hungary, weather phenomena often cause disasters (hail storms, floods, and mudflows) that subsequently generate considerable economic loss and may jeopardize human life. The most severe natural hazards in the country are associated with atmospheric convections and storms (*Horváth*, 2005). Intense upward convection triggers various phenomena ranging from small cumulus clouds to devastating supercells. Their devastation is further exacerbated by prediction challenges, as their magnitude and exact locality varies considerably in space. Convective processes develop rapidly and, with their associated features and consequences, may create tremendous damage and disastrous corollaries. Typical observed phenomena related to convective storms in hilly and mountainous regions are flash floods (*Horváth*, 1999).

According to the report of the Environmental Protection Agency of the European Union, floods generate the largest economic loss in Europe. The average annual economic loss caused by any type of natural hazards is estimated at 40 billion EUR: the Gard 2002 and the Aude 1999 flash flood events alone caused 1.2 and 3.3 billion EUR economic losses, respectively (*Lefrou et al.*, 2000; *Huet et al.*, 2003; *Gaume et al.*, 2009). Although the majority of the losses are caused by „conventional" large-river floods, over the past decades, floods more frequently occur on small streams located in small (10 to 100 km^2) mountainous watersheds. In Europe, lowland floods are rarely associated with fatalities; in contrast, often demand loss of life (*Gaume et al.*, 2009). Flash floods are generated 0.5 to 5 hours after an intense rainfall event and usually last for a few hours (in extreme cases up to a day). In certain cases, however, snowmelt may also contribute to the generation of flash floods, hence low-intensity rainfall, amid ideal environmental settings may also trigger flash floods (*Pirkhoffer et al.*, 2008). A third, recently more frequent type of flash flood occurs in heavily urbanized areas, where paved surfaces are impervious and, in general, runoff is affected by various human factors (*Gyenizse*, 2009). This latter type of floods is called urban floods or pluvial floods; however, some authors clearly differentiate them from typical flash floods (*Cobby et al.*, 2008).

Due to the geographical settings of western and northern Hungary, flash floods are frequent phenomena and have been reported several times over the past decades. However, prediction is rather challenging, due to the large spatial and temporal variability of the aforementioned convective rainfall events and the mosaical and heterogeneous pattern of topography, land use, and soil types. In addition, prediction uncertainty is very high due to the available rainfall forecasting methods, and the localized characteristics of the precipitation.

The majority of flash floods, at least in Hungary, occur between March and mid-October. Torrential, high-intensity precipitation caused significant economic loss in the hilly and low-mountain parts of Hungary. For example, the Által-ér in north-central Hungary inundated its valley following a 253 mm torrential

rainfall on June 4, 1953 (*Szilágyi*, 1954). On June 27, 1987, several houses and part of the railroad were washed away in the Bükkösd Valley (Mecsek Mountains, SW Hungary) when 71 to 88 mm rain fell during a 6-hour period (*Gyenizse* and *Vass*, 1998). In the latter case, the largest economic losses were reported from Hetvehely, a village that is located just upstream from the confluence of the Bükkösdi-víz and the Sás Streams. *Vass* (1998) concluded that, due to its rugged watershed topography, the Sás Stream is likely to be one of the major contributing and triggering factors for the extreme flash flood events of the trunk river. Perhaps the largest economic loss was associated with flash floods in Mátrakeresztes, when a flash flood inundated the valley of the Csörgő and Kövicses Streams on April 18, 2005 (*Horváth*, 2005). Economic loss was estimated to reach 1 billion HUF (approximately 5 million USD) there. The city of Kaposvár was flooded by the Kapos Stream on August 21, 2008 (*Hizsák*, 2005) when 105 mm rain fell in 3 hours. However, we need to emphasize that in terms of discharge, the flash floods of Hungary are not comparable to those documented for the Alps (e.g., *Ranzi et al.*, 2007), although in terms of general characteristics, behavior, and economic loss caused the Hungarian flood events bear a close resemblance with the Alpine floods.

Rainfall is considered to be the primary triggering factor for flash floods, but certain environmental temporally-variable boundary conditions, at least by many authors, are also considered to be crucial. These environmental factors include canopy cover (land use) and antecedent soil moisture content (e.g., *Cassardo et al.*, 2002; *Smith et al.*, 2002; *Le Lay* and *Saulnier*, 2007; *Jessup* and *DeGaetano*, 2008). Other authors, however, found little correlation between magnitude of runoff and land use or soil type (*Merz et al.*, 2006), or found peak discharge unaffected by antecedent soil moisture content (*Cras et al.*, 2007). Volumetric water content (hereafter VWC), on the other hand, not only affects runoff, but may also influence rainfall generation (*Cassardo et al.*, 2002).

These environmental factors affect the time of concentration, and indirectly the prediction time lead. Prediction time lead, in turn, will strongly influence the general character of analysis, prediction, prevention, and evacuation. To overcome the challenging prediction issues in unexplored and (quasi-) ungaged small watersheds, hydrological models may be an appropriate tool to forecast flash flood events in sufficient lead time (e.g., *Georgakakos*, 1986; *Ogden* and *Julien*, 2002; *Georgakakos*, 2002; *Javier et al.*, 2007). However, hydrological models of this type require a plethora of input data. In this case, not only the precipitation is essential as input data, but the model processes topographic data (digital elevation model, hereafter DEM), land use and land cover, and various soil properties, such as infiltration rate, topsoil thickness and soil moisture content.

To simulate selected flood events, we employed the HEC-Hydrologic Modeling System (hereafter HEC-HMS, developed in Davis, CA, United States) in the present study. Based on the available environmental data of a 6.7 km^2 watershed in SW Hungary, we verified the suitability of the HEC-HMS model

for small watersheds (and in the future, for the entire hilly and low-mountain areas of Hungary). Over the verification process, we numerically simulated two MDP events that occurred on the Sás Stream during fall, 2008. Subsequently, we aimed at numerically simulating a large flash flood event that occurred on June 4, 2008. As the watershed was ungaged at that time, i.e., neither on-the-spot precipitation data, nor VWC data were available for this particular event, our objective was to reconstruct both antecedent soil moisture content and changes in rainfall intensity during the precipitation event based on data of other nearby rain gages, radar observations, and available air temperature.

2. Materials and methods

2.1. General description of the study site

The study area is one of the headwater subbasins of the Sás Stream and is located in the NW part of the Mecsek Hills, SW Hungary and on the eastern edge of the Zselic Hills (*Fig. 1*). The watershed covers an area of 6.7 km^2 and belongs to the broader catchment of the Bükkösdi-víz. The watershed of the Sás Stream has a typical low mountain character with a steep V-shaped gorge in the headwaters of the watershed and is bordered with slopes generally exceeding 20°. The 14 monitoring stations are located on the smaller western watershed of the Sás Stream watershed, covering a mere 1.7 km^2. Based on the Corine Land Cover 2000 database, 90.6% of the watershed is covered with deciduous forest primarily of beech (*Fagus sylvatica*). The watershed exhibits frequent, but randomly distributed clearings (*Fig. 1*).

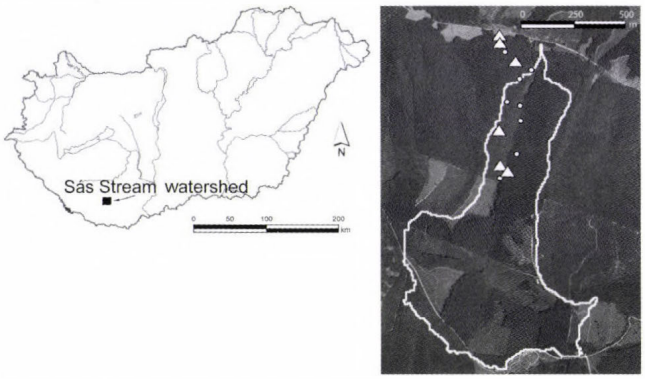


Fig. 1. Location of the studied watershed, soil moisture monitoring stations (dots and triangles), and combined soil moisture stations with rain gages (triangles)

Stream discharge ranged from 0.001752 m^3 s^{-1} to 0.9351 m^3 s^{-1} over 2008. Peak discharge was observed on June 4, at 22:00 CET. For simulation purposes,

during the period of our monitoring project (between September 5 and December 5, 2008), we selected two MDP events: the first one occurred on October 4 and a second on October 17.

2.2. Determination of soil physical type

In order to determine the soil physical type of the monitoring stations, topsoil samples were taken from 13 monitoring stations. One station located on the very edge of the alluvial plain was ignored. Soil samples were treated with 10% HCl and H_2O_2 in order to remove $CaCO_3$ and organic matter, respectively. Following the treatments, soil samples were diluted with DI water and subsequently were centrifuged at 3,000 rpm. After centrifugation, supernatant was decantated and were ovendried at 60 °C until a paste was obtained. Particle size was then determined with static light scattering method using a Fritsch Analysette 22 (Fritsch GmbH, Idar-Oberstein, Germany) instrument.

2.3. Monitoring of VWC

Due to the large input data demand of the selected HEC-HMS runoff model for the simulation of flash flood events, the measurement, quantification, and estimation of several environmental factors were essential. Owing to its geographical proximity, appropriate environmental settings, and its significant contribution to the floods of the trunk river, we selected the western watershed of the Sás Stream (left tributary of the Bükkösdi-víz, SW Hungary) for simulation purposes (*Fig. 1*). Here regular environmental monitoring was undertaken for the period of September 5, 2008 – January 5, 2009. During this period we regularly monitored VWC (θ_v) and canopy cover at 14 monitoring stations (*Fig. 1*). These two types of data, among many other ones, are crucial input for the HEC-HMS modeling software applied. As the output data of the HEC-HMS are highly sensitive to VWC, our measurements primarily focused on the determination of soil moisture content over the watershed. Another explanation to focus on VWC in the present study is its large spatial and temporal heterogeneity as well as its predominant role in flash flood generation.

Soil moisture content was measured with the time domain reflectrometry (TDR) technique. The applied TDR-300 instrument (Spectrum Inc., Planfield, Illinois, United States) was equipped with 20 cm long stainless steel rods. Soil moisture content was measured in a circle of 1.5 m radius at each monitoring station. At each station three repeated measurements were taken, and subsequently measured values were averaged. At the beginning of the measurement period, two plots of about 2 m by 1 m were created at each monitoring station, one plot covered with forest litter, while from the other plot forest litter was removed. However, as no significant variation in VWC was found among the two types of plots, litter-free measurements were then terminated. Measured data was then

plotted in ArcGIS 9.2 and interpolated with the inverse distance weighted function to determine the general trend of spatial soil moisture pattern. As the HEC-HMS model requires soil saturation (in percentage), measured volumetric water content values were converted to saturation percentage based on Eq. (1).

$$S=\varepsilon\cdot\theta_v\cdot 100, \tag{1}$$

where S is the saturation, ε is the porosity and θ_v is the VWC. This conversion requires the determination of soil bulk density and specific density, then, subsequently porosity is calculated as follows:

$$\varepsilon=1-\left(\frac{\rho_b}{\rho_s}\right), \tag{2}$$

where ρ_b is the bulk density and ρ_s is the specific density of soil. The average porosity of the collected soil samples was determined with a pycnometer, according to *Flint* and *Flint* (2002). We obtained an average porosity of 0.48 of 14 collected topsoil samples, thus, we received a conversion factor of 208 between VWC and saturation percentage. This soil porosity value was applied for the entire studied watershed, as well as all monitoring stations, as soil physical type does not vary significantly over the area of the watershed (*Table 1*).

Table 1. Particle size distribution of topsoils at 13 monitoring stations

Monitoring station ID	d_{mean} (μm)	Clay (%)	Silt (%)	Sand (%)	Physical soil type
2	19.479	10.29	88.18	1.53	silt
3	37.361	5.42	80.33	14.25	silty loam
4	30.156	6.66	83.65	9.69	silty loam
5	29.726	7.77	82.31	9.92	silty loam
6	31.58	6.72	82.44	10.84	silty loam
7	25.435	7.44	87.04	5.52	silty loam
8	29.12	8.25	81.37	10.38	silty loam
9	33.98	6.9	79.59	13.51	silty loam
10	29.429	6.85	83.79	9.36	silty loam
11	32.435	6.66	81.88	11.46	silty loam
12	36.368	4.71	82.21	13.08	silty loam
13	36.555	4.54	77.22	18.24	silty loam
14	27.205	8.3	85.32	6.38	silty loam

2.4. Employed precipitation data for the HEC-HMS model

Besides the VWC measurements at the 14 monitoring stations, between October 10, 2008 and October 17, 2008 daily rainfall measurements were carried out at 6

selected monitoring stations (out of the 14 stations used for VWC determination) using Hellmann type rain gages. The locations of the deployed rain gages are shown in *Fig. 1*.

To reconstruct rainfall intensities during the two MDP event, we used the following rainfall data observed at rain gages located in the relative proximity of the Sás Stream:

- Pécs, at University of Pécs, Ifjúság Street Campus, located at 46°04'40.47" N, 18°12'22.68" E (intermittent data between September 1 and December 5, both tipping bucket and Hellmann type rain gages).
- Central Pécs, located at 46°03'34.31" N, 18°13'39.35" E, daily cumulative precipitation data from June 15, Hellmann type rain gage.
- Orfű automated rain gage located at 46°08'23.77" N and 18°13'39.35" E, 10-minute cumulative rainfall data from September 22, 2008, tipping bucket rain gage.
- Hetvehely automated rain gage, located at 46°07'39.19" N and 18°02'04.22" E about 3 km west of the studied watershed, tipping bucket rain gage.
- Pogány automated meteorological station located at 45°59'41.22" N and 18°14'05.71" E, daily cumulative data, obtained from the publicly available website of the National Oceanic and Atmospheric Administration (NOAA) database, accessible at http://www.ncdc.noaa.gov.

2.5. Applied air temperature data

Air temperature data was obtained from two sources:

- Mean daily air temperature was obtained from the publicly available NOAA database, accessible at http://www.ncdc.noaa.gov.
- 1-minute and 10-minute air temperature data was observed at University of Pécs, Ifjúság Street Campus (exact location is the same as for the rain gage).

2.6. Applied stream discharge data

Water level data of 30-minute temporal resolution for the Sás Stream were measured in the central part of the Sás Stream watershed. Stream gage is located at 46°07'12.29" N and 18°04'07.05" E. Discharge data were received from the MecsekÉrc Ltd., a daughter company of the former Uranium Ore Mining Company.

2.7. Model setup

Runoff was calculated by the HEC Hydrologic Modeling System (HEC-HMS, developed in Davis, CA, United States) surface runoff model that calculates

stream discharge as a function of input precipitation and various environmental parameters. The HEC-HMS is designed to simulate the precipitation-runoff processes of dendritic watershed systems. It is designed to be applicable in a wide range of geographic areas for solving the widest possible range of problems. This includes large river basin water supply and flood hydrology, and small urban or natural watershed runoff (*US Army Corps of Engineers*, 2005).

For simulation purposes we selected the two afore mentioned MDP events. The reason for selecting these two, rather low-discharge floods, was sufficient data availability (10-minute rainfall, soil moisture, and canopy cover) for this period. The simulation of the selected two MDP events were extremely challenging as the peak discharge values only slightly exceeded the average baseflow, and the event observed on October 17, 2008 showed a fluctuating sawtooth-type, multi-peak pattern.

To set up the model, we employed the soil moisture accounting (SMA) module of HEC-HMS that determines the ratio of runoff and infiltration based on set parameters shown in *Table 2*. This method is a built-in function of the HEC-HMS.

Table 2. Value and quantification methods of the input environmental parameters of the October 4 and October 17 MDP events for the HEC-HMS model. Note, that not all the parameters of the SMA module are shown

Input parameter	**Value used for best simulation**	**Method of quantification**
Canopy saturation (%)	0	Estimated and measured leaf wetness
VWC ($m^3\ m^{-3}$)	0.269/0.2836/0.192[1,2]	Measured (TDR)
Soil saturation (%)	56/59/40[1]	Measured (TDR)
Canopy storage (mm)	2	Estimated, based on literature (e.g., *Gash et al.*, 1980)
Surface storage	2	Estimated from DEM
Maximum infiltration rate ($mm\ h^{-1}$)	26	Particle size distribution measured, then infiltration rate was based on literature
Soil thickness and water storing capacity (mm)	220	Based on field measurements and the AGROTOPO soil database

[1]June 4, 2008/October 4, 2008/October 17, 2008

[2]The reason for showing both VWC and water saturation is that the TDR measures in VWC, while HEC-HS requires water saturation as input data

To set the input hydraulic parameters, we employed the Clark Unit Hydrograph function. Here time of concentration was determined from the time

lag between the precipitation event and the peak discharge (*Fig. 2*). As this parameter highly depends on actual rainfall intensity, the time of concentration needs to be determined for each precipitation-discharge pair and adjusted according to the total contributing land area.

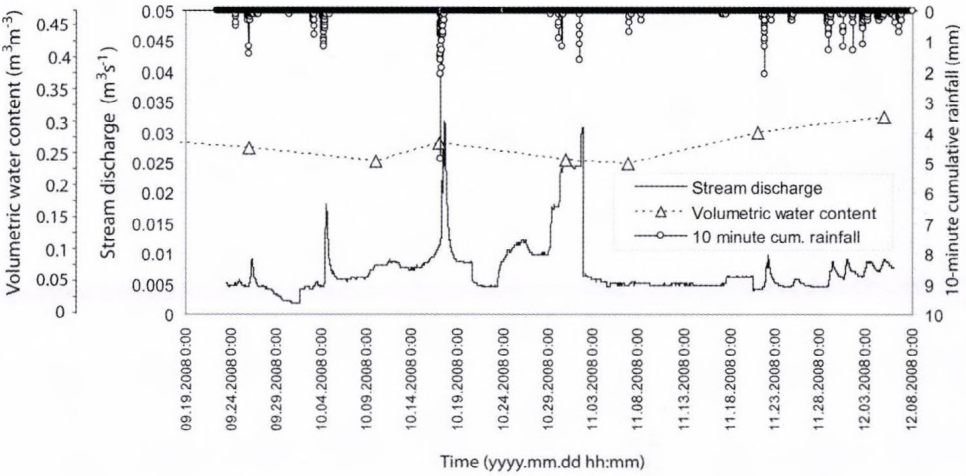


Fig. 2. Correlation between stream discharge of the Sás Stream, VWC and 10-minute cumulative rainfall (observed in Orfű) between September 5 and December 5, 2008.

3. Results

3.1. Input data acquisition

The smallest difference between the observed and simulated flow was found at SMA settings shown in *Table 2.* Parameterizations for each individual SMA are also shown in *Table 2.*

Input soil moisture, as it is indicated in *Table 1,* was obtained from the area-averaged TDR measurements (and temporal interpolation) in our study site. Over the period of measurement, VWC ranged between 0.148 $m^3 m^{-3}$ (September 5, 2008) and 0.324 $m^3 m^{-3}$ (December 5, 2008) (*Fig. 2*). At the measured porosity values, these volumetric water contents equal to 30.72% and 67.4% average water saturation, respectively. In general, topography (besides the slight impact of coarse rocky fragments) overwhelmingly impacted the spatial distribution of soil moisture content. Based on ArcGIS interpolation (using the inverse distance weighted function), highest soil moisture values were always detected at the perimeter of the alluvial plain (lowest elevation), while lower value were observed with increasing altitude (*Fig. 3*).

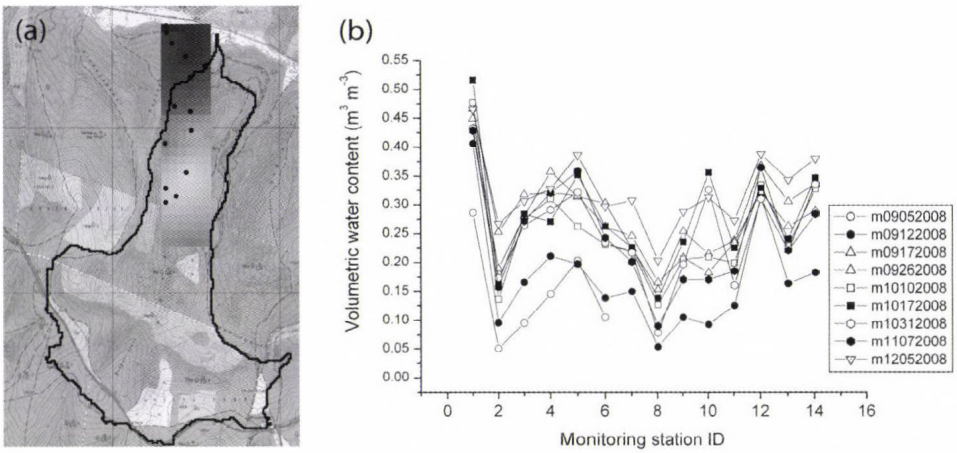


Fig. 3. Raster-based interpolation of VWC based on the October 17, 2008 TDR measurements. Darker pixels indicate higher soil moisture content.

Both mean and extreme soil moisture values varied greatly among the monitoring stations, but behaved spatially consistently over the small (1.7 km^2) watershed (*Fig. 3*). This observation was clearly exemplified during the September 26, 2008 precipitation event, when daily cumulative precipitation ranged between 13.6 and 14.6 mm over the 14 monitoring stations. Despite the relatively uniform spatial distribution of rainfall, VWC ranged between 0.166 and 0.455 m^3 m^{-3}.

Minimum measured values for the entire monitoring period were usually observed at monitoring station 8. At this point, shallow soil and frequent occurrence of coarse fragments are likely to be responsible for the low degree of saturation. On average, the second lowest soil moisture content was measured at monitoring station 2. Here, the low moisture content is explained by the dense rhizosphere where preferential flowpaths are responsible for the high initial infiltration rate. This was further confirmed by the authors' on-site dye experiment. Additionally, at monitoring station 2, topsoil has an extremely granular structure, which further enhances infiltration. Monitoring station 10 is located at the edge of a clearing, thus, long-term water balance is typically characterized with extreme VWC. However, the consistency of the spatial correlation of soil moisture content among the 14 monitoring stations is noteworthy and clearly shown in *Fig. 3*. Besides the changes of physical soil type, steeper slopes and decreasing soil depth may contribute to decreasing soil moisture at higher elevation.

Nonetheless, the literature-based strong impact of VWC on runoff (e.g., *Jessup* and *DeGaetano*, 2008) is not entirely reflected in the discharge/soil moisture content chart (*Fig. 2*). This low-degree correlation may be attributed to

the low temporal resolution of our soil moisture measurements (weekly and biweekly intervals), the rugged topography of the studied watershed (thus limited infiltration), and the dominant effect of temperature on evaporation rate. The timing of the soil moisture measurement is also crucial, most accurate values are obtained when the date of measurement immediately precedes the precipitation event.

Maximum infiltration rate, as input data for the hydrological model, was based on literature data (e.g., *Jury et al.*, 1991; *Hillel*, 1998) and the actual physical soil type (particle size distribution) collected from the uppermost soil horizons at each monitoring stations, with the exception of monitoring station 1 (located on the edge of the alluvial valley floor). Based on the particle size distribution, the topsoils of the studied catchment, with one exception, belong to the silty loam soil physical type (*Table 1*). Thus, based on literature, initial infiltration rate was set to 26 mm h^{-1}.

As clearly exemplified here by the identical soil physical types, soil type alone does not account for the varying water budget of soils. It needs to be emphasized, however, that our particle size measurement was applied for particles smaller than 300 μm. At higher elevation (monitoring stations 8 and 9) the increased frequency of coarse rocky fragments is detectable, assuming the presence of preferential flowpath at the edge of the coarse fragments, while the upper surface of this rock fragment may behave as an impervious surface.

3.2. Reconstruction of the October 4, 2008 rainfall event

To reconstruct the rainfall for the October 4, 2008 MDP event, we used three rainfall time series sources: (a) 10-minute cumulative rainfall series from the automated Orfű rain gage, (b) 10-minute cumulative rainfall series from the automated Hetvehely rain gage, (c) the authors on-site measurement by Hellmann-type rain gages, and (d) radar images.

Only a minor difference was detected among the six-station-averaged rainfall measured in the watershed of the Sás Stream (with a mean value of 15.2 mm, ranging between 14.6 and 16.7 mm), the Hetvehely rain gage (14.2 mm), and the Orfű automated rain gage (15.6 mm). To reconstruct the temporal fluctuation of rainfall intensities in 10-minute time intervals, we employed the data supplied by the Hetvehely rain gage. Intensity in general was low and relatively even for this event, reaching a maximum intensity of 1.2 mm h^{-1} between 7:50 and 8:00 A.M. (*Fig. 4c*). The 10-minute cumulative data series of the Hetvehely data was then multiplied by a factor of 0.934 (i.e., 14.2/15.2). By doing this, we obtained a temporal distribution identical to that of the Hetvehely rain gage, while the cumulative rainfall was equal to that measured in the watershed of the Sás Stream.

However, despite the non-convective character of the rainfall event, total cumulative rainfall differed largely within an area of about 10 km in radius.

Cumulative daily rainfall totaled 8.6 mm at the automated Pogány meteorological station, 12.3 mm in central Pécs, and 11.5 mm at the Ifjúság Street Campus of University of Pécs.

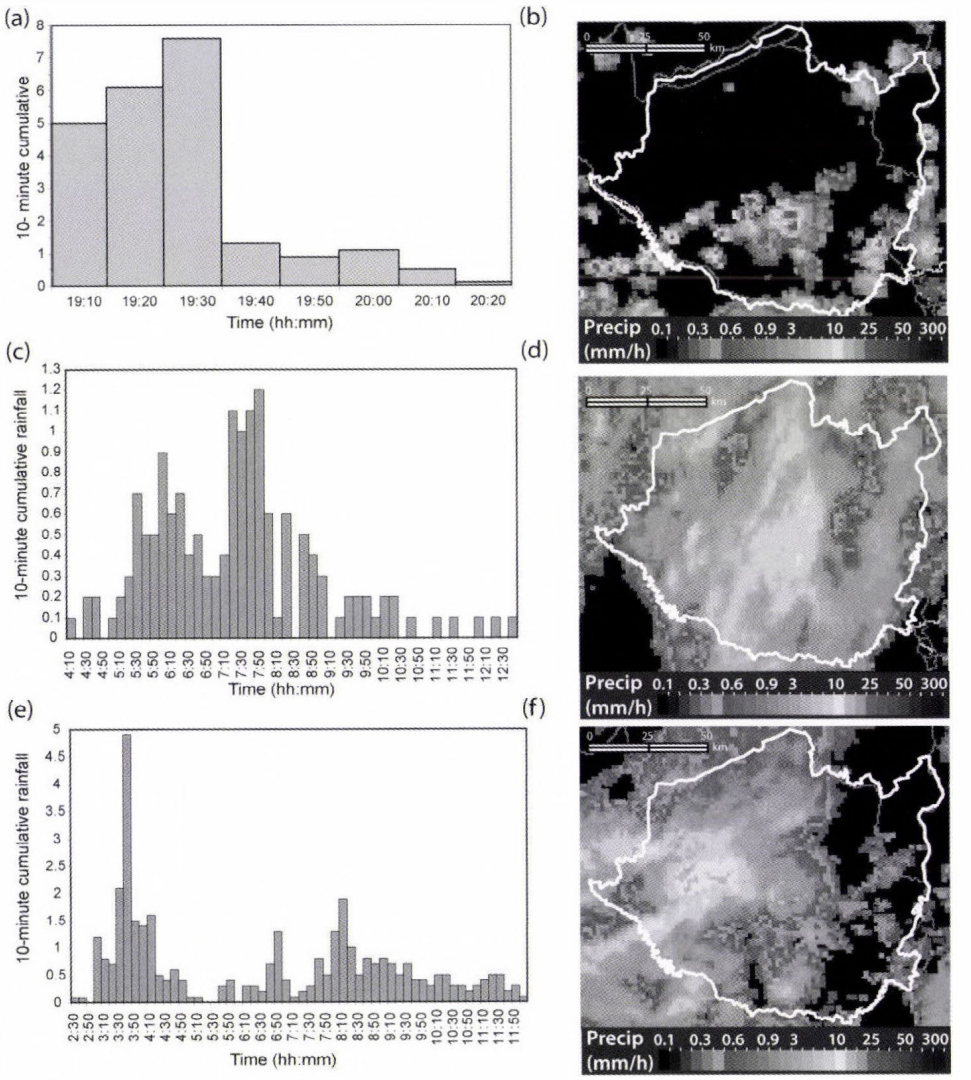


Fig. 4. (a) 10-minute cumulative rainfall observed in Hetvehely during June 4, 2008; (b) characteristic radar image taken during the June 4, 2008 rainfall event; (c) 10-minute cumulative rainfall observed in Hetvehely during October 4, 2008; (d) characteristic radar image taken during the October 4, 2008 rainfall event; (e) measured total daily rainfall values in Hetvehely (NOAA data) on October 17, 2008; (f) characteristic radar image taken on October 17, 2008.

Interpolated and estimated volumetric water content for October 4, 2008 was 0.260 $m^3 m^{-3}$ (54.1% water saturation). In the case of the latter event, the simulation best reproduced the observed hydrograph at 56% saturation. This difference is less than the measurement accuracy of the employed TDR instrument (± 0.01 $m^3 m^{-3}$). The measured average VWC of 0.282 $m^3 m^{-3}$ (58.6% water saturation) for October 17, 2008 was only slightly less than the 0.284 $m^3 m^{-3}$ (59%) used as input value in the HEC-HMS.

3.3. Reconstruction of the October 17 rainfall event

Similarly to the MDP event discussed in the previous section, to reconstruct the October 17, 2008 rainfall time series, we used three types of data: (a) 10-minute cumulative rainfall series from the automated Orfű rain gage, (b) 10-minute cumulative rainfall series from the automated Hetvehely rain gage, (c) the authors on-site measurement by Hellmann-type rain gages, and (d) radar images.

Our field rainfall measurements at six of the total of 14 monitoring stations resulted in 16.3 mm cumulative precipitation on average (ranging between 13.0 mm and 18.2 mm, depending on the location of the monitoring station), which is significantly less than that of measured by the Orfű rain gage located second closest to the study area in a distance of 6.70 km (36 mm). During most of the rainfall event, intensity was low, with the exception of a 10-minute period, when cumulative rainfall totaled 4.9 mm (*Fig. 4e*). At the Hetvehely rain gage, a cumulative rainfall of 24.7 mm was observed. This large difference in cumulative rainfall among the observation sites is attributed to the convective character of the rainfall event (indicated by the radar image in *Fig. 4f*), the rugged topography (orographic effects), and interception. However, the two automated rain gages are not covered by canopy, still it is unlikely, at least based on the available literature, that interception alone would account for that considerable difference (e.g., *Rutter et al.*, 1975; *Gash* and *Morton*, 1978; *Gash et al.*, 1980; *Klaassen et al.*, 1998, *Link et al.*, 2004). This latter statement is further confirmed by the authors' rainfall measurement on October 4, 2008, when a mere 0.4 mm difference in rainfall amounts was observed between the Sás Stream observation sites and the Orfű rain gage (see previous section 3.2). The radar based precipitation intensity, however, probably overestimated rainfall intensities, and care needs to be taken when radar data is used in rainfall forecast or prediction of flood levels.

As the model requires temporal changes of rainfall intensity during a given rainfall event, 10-minute rainfall intensities of the Hetvehely rain gage data (located closest to the studied area) were multiplied by a factor of 24.7/16.3 (i.e., 1.51) to obtain a cumulative value measured in the watershed, which were then employed in the model simulation runs. We note that the figure for cumulative daily rainfall we obtained on-site (i.e., in the Sás Valley) only slightly differed from that measured in central Pécs, at a distance of 15.05 km (16.0 mm), and

were somewhat higher than that observed in Pogány at a distance of 19.64 km (11.9 mm). For the period of October 6 – October 29 a total of 22.5 mm rainfall was measured at the Ifjúság Street Campus of University of Pécs (at a distance of 12.21 km).

Rainfall observations from the Pogány meterological station have been available since 1953 (longest record in Baranya County) and could be potentially used as input time series to simulate flash flood events on the adjacent streams. However, based on the poor correlation among the Orfű, Hetvehely, and Pogány daily precipitation data, we conclude that the 10-minute resolution Pogány precipitation data, mainly for convective precipitation, should be used with extreme caution for the estimation of total rainfall for the watershed of the Sás Stream. To overcome such difficulties, data of rain gages located closer to the studied streams should be used. The Orfű and Hetvehely rain gages, which were put to operation recently, may offer substitute solutions to aid the estimation of actual cumulative rainfall and precipitation intensities in the watersheds of the western Mecsek Hills.

The simulated results for the two MDP events, at least considering the shape of the curve, is similar to the observed hydrograph. The cumulative simulated outflow was considerably smaller than the total observed outflow in the case of the October 4, 2008 event (*Table 3*). However, in the case of the October 17, 2008 event, the simulated outflow exceeded the simulated value. Furthermore, the HEC-HMS was unable to reproduce the multi-peak pattern of the October 17, 2008 MDP event. In the latter case, the volume of the simulated cumulative outflow was 1.19 times larger than the observed one, while peak discharges hardly differed (*Table 3*). This large discrepancy is probably explained by (i) the extremely low peak discharge values, and (ii) the sawtooth pattern of the discharge peaks.

Table 3. Peak discharge and cumulative outflow values for the simulated and observed outflows of the October 4, 2008 and October 17, 2008 MDP events

	Peak discharge ($m^3 s^{-1}$)	**Cumulative outflow (1000 m^3)**	**Outflow difference compared to observed (%)**
Observed, October 4, 2008	0.02	1.23	
Simulated, October 4, 2008	0.02	0.73	– 40.65
Observed, October 17, 2008	0.03	1.56	
Simulated, October 17, 2008	0.03	1.85	+18.59

3.4. Simulation results of the June 4, 2008 flood event

To verify our results and settings obtained by the simulation of the October 4, 2008 and October 17, 2008 MDP events, we simulated the largest flood event on the Sás Stream in 2008. This flood event took place in the evening of June 4, 2008, when peak discharge reached 0.93 $m^3 s^{-1}$ as a result of a 22.6 mm rainfall (measured in Hetvehely) that reached its highest rainfall intensity between 19:20 and 19:30 (*Fig. 4a*). According to the radar images, the storm was located over the watershed of the Sás Stream between 18:45 and 20:00 CET. Since no rainfall data was available from the immediate vicinity of the watershed, it was essential to reconstruct the flood-generating rainfall based on precipitation data of other rain gages and radar images.

To reconstruct the total precipitation and rainfall intensity for June 4, 2008, we used the Hetvehely automated rain gage data and 15-minute resolution radar images obtained from the Hungarian Meteorological Services. The input precipitation data was aided by radar based rainfall intensities and timing (*Fig. 4*).

Environmental settings for the SMA loss method and the Clark Unit Hydrograph functions were identical to those used in the two MDP simulations, with the exception of soil moisture content.

As no measured soil moisture data is available for June 4, 2008, it was necessary to reconstruct the soil moisture content based on observed mean daily air temperature and field measurements on June 9 and June 10, 2008 and in June 2009. Air temperature is a suitable indicator of soil moisture changes, which fairly well reflects temporal changes of soil wetness. In our studied watershed, mean daily temperature measured in Pogány and VWC show a strong reverse correlation ($r^2 = 0.9629$), which further supports the suitability of air temperature for the estimation of the actual VWC. We partly estimated the soil moisture content based on the mean daily temperatures of NOAA for Pogány (*Fig. 5*).

Over this 15-day period, average air temperature was 18.2 °C. Based on the equation of the trendline in *Fig. 5*, the approximate VWC was only 0.210 $m^3 m^{-3}$, if the average temperature of the preceding 15-day (18.2) is applied. If soil moisture is calculated with the mean temperature of June 4, 2008 (17.9 °C), then the obtained VWC is 0.212 $m^3 m^{-3}$. The last two rainfall events prior to June 4, 2008 occurred on May 20 and May 21, 2008, with a cumulative rainfall of 22.1 and 10.2 mm. On June 9 and June 10, 2008, the average, however random (i.e., not at the monitoring stations) TDR measured VWC values in the Sás watershed reached 0.341 and 0.327 $m^3 m^{-3}$, respectively. The large VWC values measured on June 9 and June 10 are likely explained by daily rainfalls of 2.2 mm, 10.1 mm, 3.1 mm, 2.1 mm, and 1.5 mm (observed in Hetvehely) between June 5, 2008 and June 9, 2008, respectively. The average VWC for the same period in 2009 (measured on June 5) ranged between 0.17 and 0.45 $m^3 m^{-3}$ in the Sás watershed with a mean value of 0.28 $m^3 m^{-3}$.

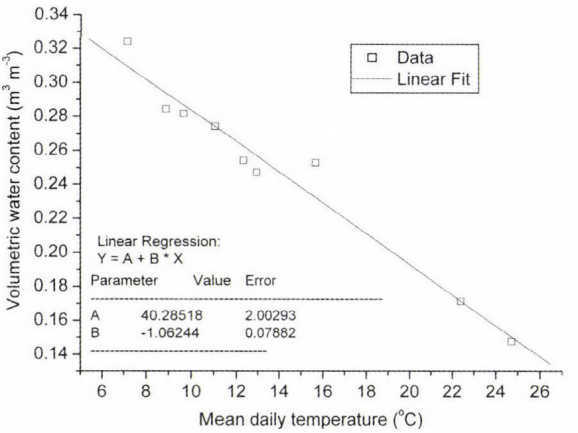


Fig. 5. Effect of mean daily temperature (observed in Pogány) on VWC between September 5, 2008 and December 5, 2008.

Based on the aforementioned environmental factors, we concluded that, preceding the rainfall event of June 4, 2008, the mean VWC of the Sás watershed ranged between about 0.150 and 0.300 $m^3 m^{-3}$ in the watershed of the Sás Stream. The exact average value was then determined by numerical simulation, using the HEC-HMS.

Table 4. Observed and simulated peak discharge and cumulative outflow values for the June 4, 2008 flash flood event at various VWC values

	Peak discharge ($m^3 s^{-1}$)	**Difference compared to observed (%)**	**Cumulative outflow (1000 m^3)**	**Difference compared to observed (%)**
Observed	0.93		7.79	
Simulated at a VWC of 0.168 $m^3 m^{-3}$	0.43	–53.76	3.80	–51.22
Simulated at a VWC of 0.192 $m^3 m^{-3}$	0.86	–7.52	7.60	–2.50
Simulated at a VWC of 0.216 $m^3 m^{-3}$	1.30	+39.78	11.40	+46.34

The best simulation results were obtained at a VWC of 0.192 $m^3 m^{-3}$ (40% water saturation). In this case the total volume of the observed outflow reached 7,100 m^3, while in the case of the simulated outflow it reached 7,600 m^3, when peak discharges were 0.93 $m^3 s^{-1}$ and 0.86 $m^3 s^{-1}$ for the observed and simulated outflows, respectively (*Table 4*). Due to the more regular shape of the June 4, 2008 flood event, the shape of the simulated curve was closer to the observed curve than in the cases of the MDP events (*Fig. 6c*). Similarly, in the case of the

June 4, 2008 flood event, the simulated and observed hydrographs bore a closer resemblance than in the case of the two simulated MDPs. This difference is likely caused by the irregular shape of the MDPs and the unsuitability of HEC-HMS to reproduce flash flood events of low peak discharge and cumulative outflow values.

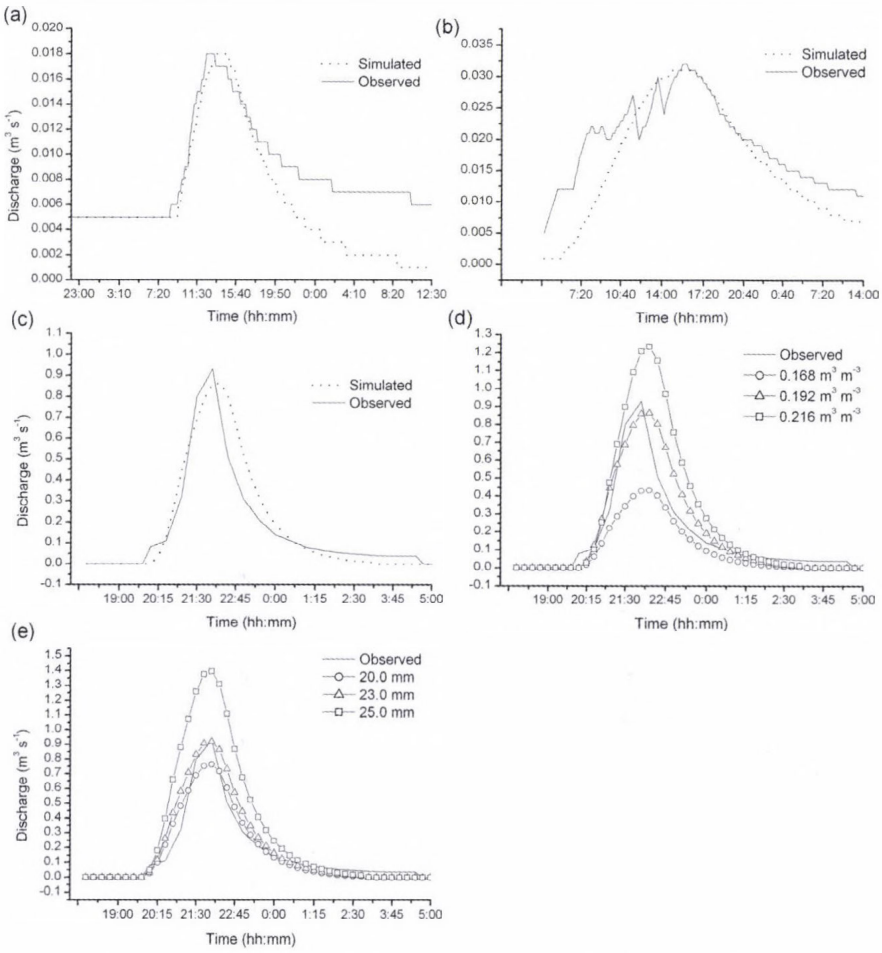


Fig. 6. Observed and simulated hydrographs of the flash flood event of the Sás Stream on (a) October 4, 2008, (b) October 17, 2008, and (c) June 4, 2008; (d) effect of various VWCs on modeled outflow generated by the HEC-HMS at a cumulative rainfall of 22.6 mm for the June 4, 2008 flood event; (e) effect of various cumulative rainfall on modeled outflow generated by the HEC-HMS at a VWC of 0.192 m^3 m^{-3} for the June 4, 2008 flood event.

When the input VWC was set to 0.168 m^3 m^{-3} (0.350% water saturation), simulated cumulative outflow was 51.22% less than the observed outflow (*Fig.*

6d and *Table 4*). Best fit was obtained at a VWC of 0.192 m^3 m^{-3} (45% water saturation) when the model underestimated the peak discharge by only 7.52%, while the cumulative outflow differed by only 2.5% from the observed outflow. At the best fit, however, cumulative outflow difference was less than in the case of the two MDP events. However, when input VWC was set to 0.216 m^3 m^{-3}, the model overestimated both peak discharge and cumulative outflow by 39.78% and 46.34%, respectively (*Table 4*).

Due to the lack of actual rainfall data on the Sás watershed (closest observation is about 3 km to the west), we also simulated the effect of various input precipitation for the June 4, 2008 event at a VWC of 0.192 m^3 m^{-3} (*Fig 6e* and *Table 5*). Best correlation between the observed and simulated output was obtained at 23 mm rainfall (i.e., 0.4 mm more than observed at the Hetvehely rain gage) if a VWC value of 0.192 m^3 m^{-3} was applied. In this case, observed and simulated peak discharge was almost identical (with a difference of only 1.07%), while the simulated cumulative outflow exceeded the observed value by 14.63% (*Table 5*). Both cumulative outflow (+46.34%) and peak discharge (+39.78%) greatly increased when a cumulative rainfall of 25 mm was applied in the model. Here we may conclude that the role of precipitation, similarly to VWC, is also exaggerated in the HEC-HMS.

Table 5. Observed and simulated peak discharge and cumulative outflow values for the June 4, 2008 flash flood event at various cumulative rainfall values at a VWC of 0.192 m^3 m^{-3}

	Peak discharge (m^3 s^{-1})	**Difference compared to observed (%)**	**Cumulative outflow (1000 m^3)**	**Difference compared to observed (%)**
Observed	0.93		7.79	
Simulated at a cumulative rainfall of 20 mm	0.76	–18.28	7.41	–5.12
Simulated at a cumulative rainfall of 23 mm	0.92	–1.07	8.93	+14.63
Simulated at a cumulative rainfall of 25 mm	1.40	+50.54	13.53	+73.68

4. Conclusions

4.1. Estimation of soil moisture

VWC as an input data for runoff models is hard to estimate, as it varies greatly in both space and time. However, VWC values vary with the measurement sites in the studied watershed in a predictable way. This consistent behavior is explained by the strong impact of soil properties and topography and the continuous

canopy cover. Thus, for a given, relatively small watershed, a measurement at one particular site would be sufficient for the estimation of soil moisture values at other sites of the watershed, once the relative differences have been measured and calibrated by either an on-site instrument or a single measurement taken by remote-sensing tools, and the measured value could be extrapolated to other parts of the watershed. We believe that regular on-site observations and (preferably automated) soil moisture measurements are essential on the most hazardous watersheds to provide VWC as input parameter for the real-time runoff models.

When regular soil moisture measurements are unavailable, temporal interpolation is required. To achieve appropriate temporal interpolation under forest cover, air temperature (mean daily) and cumulative rainfall time series have to be employed. In our case study, mean temperature showed a good correlation with daily VWC values. However, such correlations need to be developed site-specifically. Additionally, for non-forested portions of the watersheds, relative humidity and wind speed need also to be considered.

In summary, we believe that the accuracy of the estimation of the actual soil moisture content will increase by estimating the soil moisture loss to evaporation. Thus, comprehensive watershed-based meteorological and hydrological monitoring is desirable to reliably predict flood potential. Finally, we conclude that for studies of this type a numerical model need to be applied to generate real-time VWC values.

4.2. Uncertainty in rainfall estimation

Forecast of actual rainfall intensities and cumulative rainfall poses further challenges for the estimation of expected flood levels, as the applied model is highly sensitive to input cumulative precipitation data as well as rainfall intensity. As flash floods are primarily triggered by convective precipitation in orographically diverse areas, spatio-temporal variability in both cumulative precipitation and rainfall intensities may be considerable. To obtain more accurate rainfall prediction and lead time (at least 3 hours), meteorological models and/or radar-based prediction offer appropriate tools. Model predictions may be suitable to delineate the broader area of the potential precipitation event a few hours prior to the actual flood event. For the subsequent refinement of the location, radar observation can be applied. However, in radar observation a considerable error of the measurement of precipitation intensity is inherent.

The present-day radar observations in Hungary have a 2 km spatial and 15-minute temporal resolution. Considering the size of the individual watersheds in the hilly and low-mountain areas of Hungary, we may obtain a sufficient-resolution radar measurements for a given watershed. Estimation of the impact of rainfall is further complicated by rain-on-the-snow events (*Pirkhoffer et al.*, 2007;

Czigány et al., 2009). In summary, close collaboration between the radar data providing meteorological services and the flash flood warning system is necessary in order to obtain sufficient lead time for timely warning of the end users.

4.3. Suitability of HEC-HMS runoff model for numerical simulations

The HEC-HMS is a suitable tool to reproduce large flash flood events on small mountain streams. However, based on the present study, it is probable that both the VWC and rainfall exaggeratedly affect outflow values. Still, we believe that the HEC-HMS would be an appropriate means to determine and predict pronounced flood peaks, and could also be used to determine threshold runoff values, as well as the threshold cumulative precipitation that can trigger floods. We note, however, that when HEC-HMS is used to reproduce high-flow events, significant discrepancy between the simulated and observed values may occur. However, this difference decreases when peak discharge increases.

The applicability of the HEC-HMS model for the development of a nationwide vulnerability map is sufficient, however, in certain aspects, is limited. We believe that the approximate amount (or range) of threshold rainfall can be determined if environmental parameters are available with sufficient accuracy. However, as already pointed out, certain environmental factors may strongly vary in both space and time.

The HEC-HMS tends to be over-sensitive to soil moisture changes. Soil moisture is a highly variable and challenging input parameter for the HEC-HMS runoff model and often represents an average value for large and heterogeneous watersheds. Due to its considerable spatial heterogeneity, the input data may represent a broad range of soil moisture values. In the case of the HEC-HMS, for each watershed, regardless of their area, only a single soil moisture value is provided and, consequently, that does not reflect the above described spatial variance. Multiple model simulations are suitable tools to determine and reproduce field soil moisture values, but simulation errors may become large for future discharge predictions. Our findings suggest that potential discrepancies are further exaggerated by the HEC-HMS model and do not reflect flood level responses appropriately.

Additionally, hydrological models require the downscaling of watershed areas in order to use spatially homogenized input data, such as soil type or canopy cover. However, to approximate these environmental factors at this scale, databases (e.g., soil and land use data) of high resolution are essential. Simultaneously, radar data of at least 15-minute temporal resolution are required to properly model flash flood events. Based on radar predicted precipitation data, available hydrological models (e.g., HEC-HMS) are suitable to predict flood events on small hilly and low-mountain catchments in our study areas and are likely to be capable to form the foundation of a nationwide, flow chart based flash flood warning system.

Acknowledgements—This research was supported by the Jedlik Ányos Foundation (Grant No. NKFP3-00022/2005) and by the Bolyai János Research Scholarschip of the Hungarian Academy of Sciences. The authors are grateful to *Dénes Lóczy* for reviewing the entire manuscript, and to *Ákos Horváth, Gábor Horváth, Roland Vendégh, Gábor Földing, Emil Horváth*, and *Gábor Antal* for their help and contribution to the present work.

References

Cassardo, C., Balsamo, G.P., Cacciamani, C., Cesari, D., Paccagnella, T. and *Pelosini, R.*, 2002: Impact of soil surface moisture initialization on rainfall in a limited area model: a case study of the 1995 South Ticino flash flood. *Hydrol Process 16*, 1301-1317.

Cras, A., Marc, V., and *Travi, Y.*, 2007: Hydrological behaviour of sub-Mediterranean alpine headwater streams in a badlands environment. *J Hydrol 339*, 30-144.

Czigány, S., Pirkhoffer, E. and *Geresdi, I.*, 2009: Environmental impacts of flash floods in Hungary. In *Flood Risk Management: Research and Practice* (eds.: *P. Samuels, S. Huntington, W. Allsop,* and *J. Harrop*). Taylor and Francis Group, London, pp. 1439-1447.

Flint, A.L. and *Flint, L.A.*, 2002: Particle density. In *Methods of Soil Analysis Part 4 Physical Methods* (eds.: *J.H. Dane* and *G.C. Topp*). Soil Science Society of America Inc, Madison, WI.

Gash, J.H.C. and *Morton, A.J.*, 1978: An application of the Rutter model to the estimation of the interception loss from Thetford forest. *J Hydrol* 38, 49-58.

Gash, J.H.C., Wright, I.R. and *Lloyd, C.R.*, 1980: Comparative estimates of interception loss from three coniferous forests in Northern Gemany. *Agr Forest Meteorol* 79, 131-148.

Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaškovičová, L., Blöschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A., Kohnova, S., Koutroulis, A., Marchi, L., Matreata, S., Medina, V., Preciso, E., Sempere-Torres, D., Stancalie, G., Szolgay, J. Tsanis, I., Velasco, D., and *Viglione, A.*, 2009: A compilation of data on European flash floods. *J Hydrol 367*, 70-78.

Georgakakos, K.P., 1986: On the design of national, real-time warning systems with capability for site specific flash flood forecasts. *B Am Meteorol Soc 67*, 1233–1239.

Georgakakos, K.P., 2002: Hydrometeorlogical models for real time rainfall and flow forecasting. In *Mathematical Models of Small Watershed Hydrology and Application* (eds.: *V.P. Singh* and *D. Frevert*). Water Resources Publications, LLC, Highlands Ranch, Colorado, pp. 593-655.

Gruntfest, E. and *Ripps, A.*, 2000: Flash floods: Warning and mitigation efforts and prospects. In Floods (ed.: *D.J. Parker*). Vol. 1. Routledge, London, pp. 377-390.

Gyenizse, P., 2009: Geoinformatical studies in Pécs. Studies on physical and social factor affecting the urban development of Pécs. A case study (in Hungarian). *Geographia Pannonica Nova 7*, Publikon Kiadó, Pécs, p. 110.

Gyenizse, P. and *Vass, P.*, 1998: The impact of phyiscal environment on the formation and developemnt of the settlements of the Mecsek Hills (in Hungarian). *Földrajzi Értesítő XLVII*(2), 131-148.

Hillel, D., 1998: *Environmental Soil Physics*. Academic Press, San Diego, CA.

Hizsák, I., 2005: Flooding of the River Kapos (in Hungarian). A Dél-Dunántúli Környezetvédelmi és Vízügyi Igazgatóság, valamint a Dél-Dunántúli Környezetvédelmi, Természetvédelmi és Vízügyi Felügyelőség időszaki lapja, *2*, 6-8.

Horváth, E., 1999: Simulation of extreme hydrological events in unexplored small drainage basins (in Hungarian with English summary). *Vízügyi Közlemények 81*, 486-497.

Huet, Ph., Martin, X., Prime, J.-L., Foin, P., Laurain, Cl., and *Cannard, Ph.*, 2003: Retour d' expériences des crues de septembre 2002 dans les departments du Gard, de l'hérault, du Vaulcluse, des bouches du Rhone, de l'Ardéche et de la Drome. Inspection générale de l'Environnement. Paris, France, 124 p.

Javier, J.R.N., Smith, J.A., Meierdiercks, K.L., Baeck, M.L. and *Miller, A.J.*, 2007: Flash flood forecasting for small urban watersheds in the Baltimore metropolitan region. *Weather Forecast 22*, 1331-1344.

Jessup, S.M. and *DeGaetano, A.T.*, 2008: A statistical comparison of the properties of flash flooding and nonflooding precipitation events in portions of New York and Pennsylvania. *Weather Forecast 23*, 114-130.

Jury, W.A., Gardner, W.R. and *Gardner, W.H.*, 1991: *Soil Physics*, 5th edition. John Wiley and Sons Inc., New York.

Klaasen, W., Bosweld, F. and *de Water, E.*, 1998: Water storage and evaporation as constituents of rainfall interception. *J Hydrol 212/213*, 36-50.

Lefrou, Cl., Martin X., Labarthe, J-P., Varret, J., Maziére, B., Tordjman, R., and *Feunteun, R.*, 2000: Les crues des 11, 12 et 13 novembre 19999, dans les departments de l'Aude, l'Hérault, les Pyrennées Orientales et du Tarn (in French). Inspection générale de l'Environnement. Paris, France, 140 p.

Le Lay, M. and *Saulnier, G.M.*, 2007: Exploring the signature of climate and landscape spatial variabilities in flash flood events: Case of the 8-9 September 2002 Cevennes-Vivarais catastrophic event. *Geophys Res Lett 34*(13) Article No.: L13401.

Link, E.T., Unsworth, M. and *Marks, D.*, 2004: The dynamics of rainfall interception by a seasonal temperate rainforest. *Agr Forest Meteorol 124*, 171-191.

Merz, R., Bloschl, H. and *Parajka, J.*, 2006: Spatio-temporal variability of event runoff coefficients. *J Hydrol 331*, 591-604.

Ogden, F.L. and *Julien, P.Y.*, 2002: CASC2D: a two-dimensional, physically-based, Hortonian hydrologic model. In *Mathematical Models of Small Watershed Hydrology and Applications* (eds.: *V.P. Singh* and *D. Frevert*), pp. 69-112.

Pirkhoffer, E., Czigány, S. and *Geresdi, I.*, 2008: Modeling of flash flood events in a small low-mountain watershed in SW Hungary. In *Proceedings of the joint MAP D-PHASE Scientific Meeting COST 731 mid-term seminar*. Challenges in hydrometeorological forecasting in complex terrain (eds.: *A. Montani, P.P. Alberoni, A. Rossa, M.W. Rotach, A. Buzzi,* and *S. Daviolo*). http://www.smr.arpa.emr.it/dphase-cost/master_proceeding_final.pdf

Pirkhoffer, E., Czigány, S., Geresdi, I. and *Nagyváradi, L.*, 2007: Effect of topography and rainfall pattern on the occurence of flash floods in Hungary. Carpatho-Balkan-Dinaric Conference on Geomorphology, October 24-28, 2007, Pécs, Hungary (poster presentation).

Ranzi, R., Zappa, M. and *Bacchi, B.*, 2007: Hydrological aspects of the Mesoscale Alpine Programme: Findings from field experiments and simulations. *Q J R Meteor Soc 133*, 867-880.

Rutter, A.J., Kershaw, K.A., Robins, P.C. and *Morton, A.J.*, 1971: A predictive model of rainfall interception in forests, 1. Derivation of the model from observations in a plantation of Corsican pine. *Agr Meteorol 9*, 367-384.

Smith, J.A., Baeck, M.L., Morrison, J.E., Sturdevant-Rees, P., Turner-Gillespie, D.F. and *Bates, P.D.*, 2002: The regional hydrology of extreme floods in an urbanizing drainage basin. *J. Hydrometeor 3*, 267-289.

Szilágyi, J., 1954: Extreme flooding of the Által-ér and the Váli víz on June 9, 1953 (in Hungarian). *Vizügyi Közlemények 36*, 169-176.

US Army Corps of Engineers, 2005: *Hydrologic Modeling System HEC-HMS*. User's Manual Version 3.0.0.

Vass, P. 1997: Floods in the headwaters of the Bükkösd Stream (in Hungarian, English summary). In *Földrajzi tanulmányok a pécsi doktoriskolából* I. (eds.: *R. Tésits* and *J. Tóth*). Bornus Nyomda, Pécs, 261-285.

Internet sources

Horváth, Á., 2005: Meteorological aspects of the April 18, 2005 flood event of Mátrakeresztes in Hungarian) http://www.met.hu/pages/vihar 20050418.html, last accessed: April 29, 2008.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–June 2010, pp. 101–120

Simulation of accidental release using a coupled transport (TREX) and numerical weather prediction (ALADIN) model

Róbert Mészáros[1*], **Csilla Vincze**[1], and **István Lagzi**[2]

[1]*Department of Meteorology, Eötvös Loránd University*
P.O. Box 32, H-1518 Budapest, Hungary; E-mail: mrobi@nimbus.elte.hu

[2]*Department of Chemical and Biological Engineering, Northwestern University*
2145 Sheridan Road, Evanston, Illinois 60208, U.S.A.

[*]*Corresponding author*

(Manuscript received in final form March 18, 2010)

Abstract—A dispersion model-system called TREX (TRansport–EXchange) was developed at the Eötvös Loránd University (Budapest, Hungary) for modeling the transport and deposition processes of air pollutants originated from either continuous or accidental releases. In this study, a multi-layered, Eulerian version of the TREX model is presented that can simulate transport, transformation, and deposition processes of radionuclides or chemically toxic substances over Central Europe. The TREX model was coupled to the ALADIN mesoscale limited area numerical weather prediction model used by the Hungarian Meteorological Service to predict the path of nuclear contamination in the atmosphere. Two case studies – hypothetical accidents at the Paks Nuclear Power Plant (Paks NPP) – are presented to show how the coupled model can capture the dispersion of radionuclides from a single point source. Another effort of this study is to present a new method, a simple spatial semi-adaptive grid algorithm, which could be an effective tool for accelerating simulations. In this method, model calculations are performed only in the grid cells where concentrations in the previous time step are higher than a user defined threshold concentration. Reduction of the computational time depends on this threshold value, but decreases of more than 60% are found without decreasing the accuracy of the results. The efficiency of this method and a comparative analysis of model estimations with and without this method are also presented in this study.

Key-words: dispersion model, numerical weather prediction model, accidental release, semi-adaptive grid

1. Introduction

The potential for an accident in a nuclear power plant requires continuous developments of atmospheric dispersion models and widespread simulations of an accidental release of radionuclides with these models. Since the accident at the nuclear power plant in Chernobyl in 1986, an increasing demand is observable on the part of countries for construction of sophisticated dispersion models. On the basis of accurate model calculations, the decision makers have to make important arrangements, which can save human lives. In the last two decades, several accidental release models for different spatial and time scales have been developed. Some of these models are based completely or in part on the traditional Gaussian formulation. The DERMA (Danish Emergency Response Model of the Atmosphere) model uses a hybrid stochastic particle-puff diffusion description (*Sørensen et al.*, 1998, 2007). In the horizontal, a Gaussian distribution of the concentration is assumed for each puff. For puffs inside the boundary layer, an assumption of complete mixing is applied in the vertical, while for puffs above the boundary layer, a Gaussian distribution is used. The RIMPUFF (Risø Mesoscale PUFF) model (*Mikkelsen et al.*, 1997) was applied for example, in the study of *Gultureanu et al.* (2000). The Hungarian Meteorological Service has also adapted the RIMPUFF model for mesoscale simulations of a plume. The NAME model of the UK MET Office (the British Meteorological Service) and the Norwegian SNAP model apply Lagrangian description which simulates numerous particles that provide facility to take account of the effect of the fluctuation in the meteorological data (*Ryall* and *Maryon*, 1998; *Saltbones et al.*, 1998; *Bartnicki et al.*, 2003; *Jones et al.*, 2007). The Austrian emergency response modeling system, TAMOS (*Pechinger et al.*, 2001) is based on the Lagrangian particle dispersion model FLEXPART (*Stohl et al.*, 1998, 2005; *Srinivas et al.*, 2006) and the trajectory model FLEXTRA (*Baumann* and *Stohl*, 1997; *Stohl* and *Wotawa*, 1997). The Hungarian Meteorological Service has also adapted FLEXRTA/FLEXPART for calculating the long-range transport of particles from a nuclear power plant.

The Lagrangian models have the advantage that they can afford to use high spatial resolution, although they rely on the interpolation of meteorological data. Their potential disadvantages are that in some cases they neglect important physical processes and often experience problems when strongly diverging flows lead to uncertainties in long-range trajectories. In contrast to the Lagrangian approach, the Eulerian models use grid based methods and have the advantage that they may take into account fully 3D descriptions of the meteorological fields rather than single trajectories. An Eulerian dispersion model, MEDIA was joined to the emergency response system of the Hungarian Meteorological Service (*Ferenczi* and *Ihász*, 2003). However, Eulerian models show difficulty in resolving steep gradients, when fixed meshes are used. This causes particular problems for resolving dispersion from a single point source,

which creates very large gradients close to the point of release. For a coarse Eulerian mesh, the release is immediately averaged into a large area, which smears out the steep gradients and creates a large amount of numerical diffusion (*Lagzi et al.*, 2004). Therefore, in some cases a mixed approach is used. In these models (e.g.: DREAM – Danish Rimpuff and Eulerian Accidental release Model, or MATCH – Multiscale Atmospheric Transport and Chemistry model), close to the source, a Lagrangian description is used, while an Eulerian model calculates the long-range transport (*Langnera et al.*, 1998; *Brandt et al.*, 2000, 2002). The accidental release models are usually coupled with a nuclear decision-support system, e.g., the DERMA model has become interfaced with the Accident Reporting and Guidance Operational System – ARGOS (*Baklanov et al.*, 2006), or RIMPUFF was jointed to RODOS (Realtime Online DecisiOn Support) system (*Ehrhardt et al.*, 1997; *Mikkelsen et al.*, 1997).

The European real time modeling exercise RTMOD was developed to the comparison of long-range transport and dispersion models and support of decision making. It gathers the results of more than 20 models, and it provides not only one-to-one model comparison but ensemble dispersion forecasts (*Bellasio et al.*, 1999; *Galmarini et al.*, 2001, 2004).

Although several different kinds of dispersion models are available, further and continuous developments of simulation techniques are required. Based on the experience of our previous investigations (*Lagzi et al.*, 2001, 2004, 2006; *Lovas et al.*, 2006; *Mészáros et al.*, 2006; *Dombóvári et al.*, 2008), a newly developed Eulerian dispersion model is presented in this study. Model simulations and tests were carried out to estimate the long-range transport of radionuclides from a nuclear power plant (Paks NPP).

Dispersion simulations must have a high degree of accuracy and must be achieved faster than real time to use it for decision making strategy. There are several well defined methods and techniques to decrease the calculation time of the applications. One useful solution is the parallelization of source code and the application of the supercomputers, clusters, and grid systems to solve these tasks (*Dabdub* and *Seinfeld*, 1996; *Martin et al.*, 1999; *Larson* and *Nasstrom*, 2002; *Alexandrov et al.*, 2004; *Dimov et al.*, 2004; *Martín et al.*, 2004; *Ostromsky et al.*, 2005; *Singh et al.*, 2006). A new method for parallelization using newly developed video-card in CUDA environment is presented in *Molnár et al.* (2010). On the other hand, there are several numerical models, in which adaptive gridding techniques have been implemented (*Lagzi et al.*, 2004, 2006, 2009; *Zegeling* and *Kok*, 2004). The numerical algorithm automatically places a finer resolution grid in regions characterized by high spatial numerical errors. Therefore, the fine resolution grid follows the plume of the air pollutants.

In this study, a new semi-adaptive calculation method was introduced to reduce the CPU time for simulating dispersion processes. Using this simple algorithm, faster simulations are achievable due to decreasing of the active cells determined by a user-defined critical (or threshold) concentration. Numerical

simulations are performed only in these active cells, where the concentration of air pollutant in the previous time step exceeded the predefined level. The effectiveness of this method (how fast and precise the model) depends on this critical value. Very large critical concentration may cause uncertainty in the simulation, and very low critical concentration cannot call forth effective acceleration. However, the optimal choice of the critical value decreases the computational time without decreasing the quality of the results.

2. *The TREX-Euler model*

2.1. Model description

As a part of TREX model-system, developed at the Eötvös Loránd University, Hungary, the TREX-Euler is a multi-layered, Eulerian passive tracer dispersion model to describe transport and deposition of air pollutants over the Central European region. The model was developed in a flexible framework, and therefore, can simulate both single source accidental releases and photochemical air pollution. In this study we only focus on the dispersion of radionuclides from a nuclear power plant. All required meteorological data were obtained from the ALADIN mesoscale numerical weather prediction model (*Horányi et al.*, 1996, 2006). In the TREX-Euler model, horizontal dispersion of radionuclides is described within a regular Eulerian grid framework. The vertical mixing is parameterized using K-theory. Dry and wet deposition and also the transformation of radionuclides by nuclear decay are considered in each grid cell. The simulation of the dispersion is based on the following atmospheric diffusion equation:

$$\frac{\partial \mathbf{c}_i}{\partial t} = -\underline{\mathbf{V}}\nabla \mathbf{c}_i + \nabla \underline{\underline{\mathbf{K}}} \nabla \mathbf{c}_i - (k_{c\,i} + k_{d\,i} + k_{w\,i})\mathbf{c}_i + \mathbf{E}_i, \tag{1}$$

which describes the advection, diffusion, and source and sinks of the radionuclides. In the equation $\mathbf{c}_i$ is the concentration of the ith radionuclides, $\underline{\mathbf{V}}$ is the three-dimensional velocity vector, $\underline{\underline{\mathbf{K}}}$ is the tensor of turbulent diffusions coefficient, $k_{c\,i}$ is the coefficient of radioactive decay, $k_{d\,i}$ and $k_{w\,i}$ are the dry and wet deposition coefficients, and $\mathbf{E}_i$ is the emission of radionuclides, respectively.

This partial differential equation has been solved by 'method of lines' technique. The two main components of the method of lines are spatial discretization followed by time integration. A second-order central difference stencil and upwind approximation are used for spatial discretization of the turbulent diffusion and advection. The model uses the forward Euler method to solve the obtained original ordinary differential equations. However, this

method is robust only with the appropriately chosen grid spacing (Δx) and time step (Δt). An accidental release model must be precise and fast at the same time. It is difficult to achieve both, because the more accurate and complex the model is the more computational time is required. The accuracy and the CPU time depend on model algorithms and also on the spatial and time resolution. For example, a shorter time step increases the accuracy of solution, at the same time, the computational efficiency decreases. To preserve the stability of the solution, the following conditions must be realized between the Δx spatial resolution and Δt time step for advection and diffusion, respectively:

$$\frac{|\mathbf{V}|\Delta t}{\Delta x} \leq 1, \tag{2}$$

$$\frac{2K\Delta t}{\Delta x^2} \leq 1, \tag{3}$$

where $|\mathbf{V}|$ is the amplitude of wind velocity vector and K is the horizontal turbulent diffusion coefficient. The spatial resolution is same as the resolution of the meteorological data from the ALADIN model, therefore, the model calculation was discretized on an equidistant rectangular grid with resolution of 0.0375×0.025 degrees (~ 2.5 km × 2.5 km). For this spatial resolution the time step Δt was set to 10 s, which assures the stability of solution.

In the vertical direction the model contains 32 levels and the resolution increases exponentially with height. Each vertical level was determined with the barometric formula (assuming isothermal layers) in a hypothetical air column, where the surface pressure is equal to 1013.25 hPa. Up to 200 m above the surface, the column was divided into 12 levels with equidistant pressure difference (197 Pa in isothermal layers), up to 3000 m the layers were thicker with higher pressure difference (1514 Pa), and these two resolutions were joined together. Two additional levels were added to the bottom and the top of column to ensure the boundary conditions.

The horizontal turbulent diffusion coefficients (K_x and K_y) were considered with a constant value equal to 10^4 m^2 s^{-1} (*Brandt*, 1998). In regional scale, in the function of the wind speed, the effect of the advection on dispersion is one or two orders of magnitude higher than the effect of the horizontal turbulent diffusion. Vertical transport was parameterized using the K-theory. The vertical turbulent diffusion coefficient (K_z) at a given level z has spatial and temporal variation:

$$K_z = \frac{k u_* z}{\Phi_T}\left(1 - \frac{z}{H_{mix}}\right)^2, \tag{4}$$

where k is von Kármán's constant taken to be 0.41, u_* is the friction velocity, z is the height above the surface, Φ_T is the similarity function for heat, and H_{mix} is the height of the mixing layer. Similarity functions for stable and unstable stratifications were used after *Arya* (1988). The value of H_{mix} was obtained from ALADIN model. The friction velocity was calculated by the following equation:

$$u_* = \frac{ku}{\log\frac{z_{ref}}{z_0} - \Psi_m}, \tag{5}$$

where z_{ref}, u, z_0, and Ψ_m are the reference height, wind velocity at this height, roughness length, and integral form of universal stability correction functions for the momentum, respectively.

Universal function in stable stratification was estimated after *Businger et al.* (1971) in case of $z_{ref} / L \leq 0.5$, and an empirical relationship of *Holtslag* and *de Bruin* (1988) was used in other cases, where L is the Monin-Obukhov length. In case of unstable stratification, the stability function was based on *Paulson* (1970). The stratification of the layer was determined iteratively by the Monin-Obukhov length:

$$L = -\frac{Tu_*^3}{\frac{gk}{\rho c_p}H}, \tag{6}$$

where T is the air temperature at 2 m (data from ALADIN model), g is the acceleration of gravity, ρ is the air density, c_p is the specific heat at constant pressure, and H is the sensible heat flux. In this study, this latter term was calculated after *Brandt* (1998).

The vertical dispersion was calculated with a Lagrangian stochastic, random displacement method. In this method, the particles, which are in a cell, are divided into a number of groups or packages (10 packages in this study) with equal number of particles, and the 'packages' are moving together. The vertical displacement of a package l_n is calculated as follows:

$$l_n = R\sqrt{K_z(n)\Delta t}, \tag{7}$$

where R is a random number with normal distribution and $K_z(n)$ is the vertical turbulent diffusion coefficient at the nst level (z_n). During a time step, a particle package can move only into the adjacent cells (the direction of the movement depends on the sign of the l_n). The ith (i = 1 to 10 in this study) package from the level

z_n gets into the level z_{n+1} (upward) if $l_n(i)$ is positive and $l_n(i) > (z_{n+1} - z_n)/2$, and into the level z_{n-1} (downward) if $l_n(i)$ negative and $-l_n(i) > (z_n - z_{n-1})/2$. If the condition is not met, the ith package stays at the level z_n. From the top and bottom levels the vertical transport is one way.

During current model simulations, the dispersion of a radionuclide – ^{131}I – was considered. The emission was assumed to be from the Paks NPP (46°34'N, 18°51'E) as a point source. The radioactive decay and the change of activity of ^{131}I were simulated. Chemical reactions were not considered. However, the structure of the program allows simultaneous simulations of the dispersion of several hundred isotopes, taking into account all radioactive decays and reactions.

The deposition is handled as a first-order reaction, which decreases the amount of the radionuclide in the atmosphere. Dry deposition of radionuclides from the bottom layer is parameterized by a constant deposition coefficient (3×10^{-6} s^{-1} for ^{131}I) based on *Baklanov* and *Sørensen* (2001). The wet deposition velocity was parameterized using a simple scheme to calculate the wet deposition coefficient k_w based on the relative humidity RH with the following parameterization (*Pudykiewicz*, 1989, 1991; *Brandt*, 1998):

$$k_w = 0, \quad \text{if } RH < RH_t$$
$$k_w = 3.5\times10^{-5}\left(\frac{RH - RH_t}{RH_s - RH_t}\right), \quad \text{if } RH \geq RH_t, \tag{8}$$

where RH is the current relative humidity, RH_s is the saturation relative humidity (100%), and RH_t is a threshold value of relative humidity (80%); above which the wet deposition of radionuclides from the bottom layer occurs. The radioactive decay is calculated by a constant rate: $k_c = \log 2 / t_{1/2}$, where the radioactive half life ($t_{1/2}$) of ^{131}I is 6.948×10^5 s.

During the model computations, each process – horizontal spreading (advection and diffusion), vertical dispersion, source and sinks terms (deposition, radioactive decay, and emission processes) – are calculated separately using the operator-splitting approach. The model was coupled with the ALADIN mesoscale limited area numerical weather prediction model used by the Hungarian Meteorological Service.

2.2. The semi-adaptive method

In a previous investigation (*Lagzi et al.*, 2006), an adaptive grid model has been presented to describe the formation and transformation of air pollutants based on triangular unstructured grids. This technique is based on the calculation of

spatial error estimations. The algorithm can then choose to refine the grid in regions of high spatial error by comparison with a user defined tolerance for radionuclides. Therefore, depending on the numerical error, the spatial resolution was varied. In this study, another effective way is presented to reduce the CPU time of model calculations. The model applies a semi-adaptive method to perform faster simulations. In contrast to the adaptive method, in the newly introduced semi-adaptive method, the spatial resolution remains constant but a concentration test is applied to determine if model calculations are to be performed in particular grid cells.

The main concept of the semi-adaptive method is that the model calculation is performed only in the grid cells, where a predefined condition is met. This condition is that the concentrations in the grid point in the previous time step must be greater than a user defined critical concentration. By this method, the accuracy and the CPU time depend on the critical concentration and the simulation period. At the beginning of the simulation, immediately after the accidental release, the efficiency of the acceleration is higher, since model calculations are performed only in a few grid points. With the development of the plume, the efficiency of the method decreases as the number of affected grid points increases. Through the choice of the critical concentration, the CPU time and the error of the method can be adjusted. By increasing the critical concentration, the CPU time can be significantly decreased, however, the calculation error increases.

3. Results and discussion

3.1. Modeling accidental release

The features of the transport model and the efficiency of the semi-adaptive method are illustrated by two simulations of a hypothetical nuclear accident in the Paks Nuclear Power Plant, in the central part of Hungary (46°34'N, 18°51'E). During the simulations, ^{131}I isotope was emitted into the atmosphere from the point source to the given grid cell with the rate of 10^{10} cm^{-3} s^{-1} between 00 and 12 UTC on December 2, 2005 and March 24, 2006. Meteorological fields (wind speed, wind direction, temperature, relative humidity, cloudiness, and mixing layer height) were obtained from ALADIN weather prediction model. Simulations were performed for 48 hours after the hypothetical accidents. The modeled area covered the region around Hungary (10.85°–25.1°E, 41.3°–52°N) and contained 8640 grid cells on each 32 vertical levels.

Fig. 1 depicts the horizontal structures of the radioactive plume originating from the Paks Nuclear Power Plant at the emission level (120 m above the surface), 12, 24, 36, and 48 hours after a hypothetical accident on December 2, 2005. The plume structure is predominantly determined by the wind field. However, other meteorological parameters can also influence the dispersion

(e.g., vertical temperature gradient, planetary boundary layer height, etc.). During this simulation period, the Carpathian Basin was affected by a Mediterranean cyclone and a warm front caused rainy weather. The variable wind field caused the plume to spread in south, south-western direction in the first few hours of the simulation, but as the wind direction changed, the radioactive cloud had moved in eastern, north-eastern, and finally, northern direction. Due to this variable wind field, a wide region (including all of Hungary) was affected by the plume. However, as the concentration was highest after the release, when the plume had moved in southern direction, the highest cumulative dry deposition of ^{131}I to the surface appeared to the south of the point source (*Fig. 2a*). The cumulative wet deposition (*Fig. 2b*) was also significant, because more than 15 mm precipitation occurred in some parts of central Hungary during two days after the hypothetical accident. Highest wet deposition was found around the point source (the maximum was higher than 10^9 cm^{-2}), where the largest amount of precipitation was fallen during 48 hours.

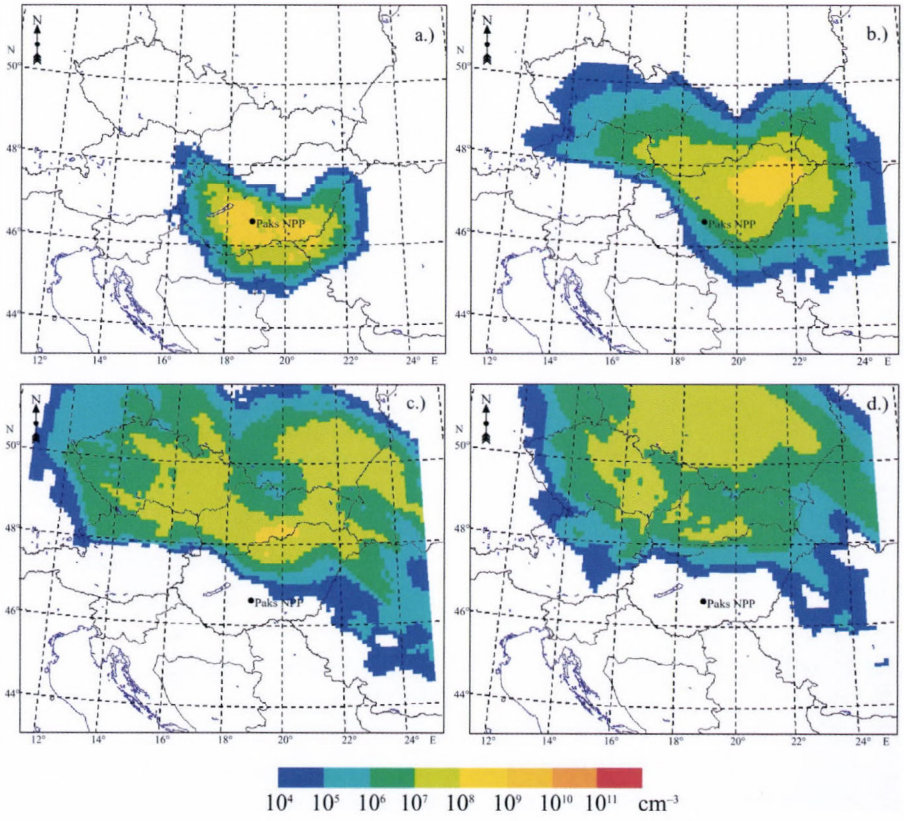


Fig. 1. Horizontal distribution map of ^{131}I concentration at 120 m above the surface, (a) 12, (b) 24, (c) 36, and (d) 48 hours after the hypothetical accident at 00 UTC, December 2, 2005.

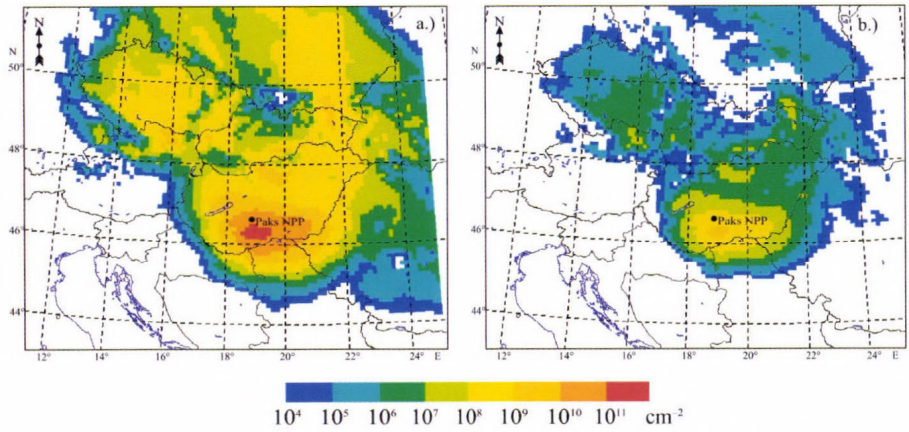


Fig. 2. Horizontal distribution map of cumulative (a) dry and (b) wet deposition of ^{131}I to the surface, 48 hours after the hypothetical accident at 00 UTC, December 2, 2005.

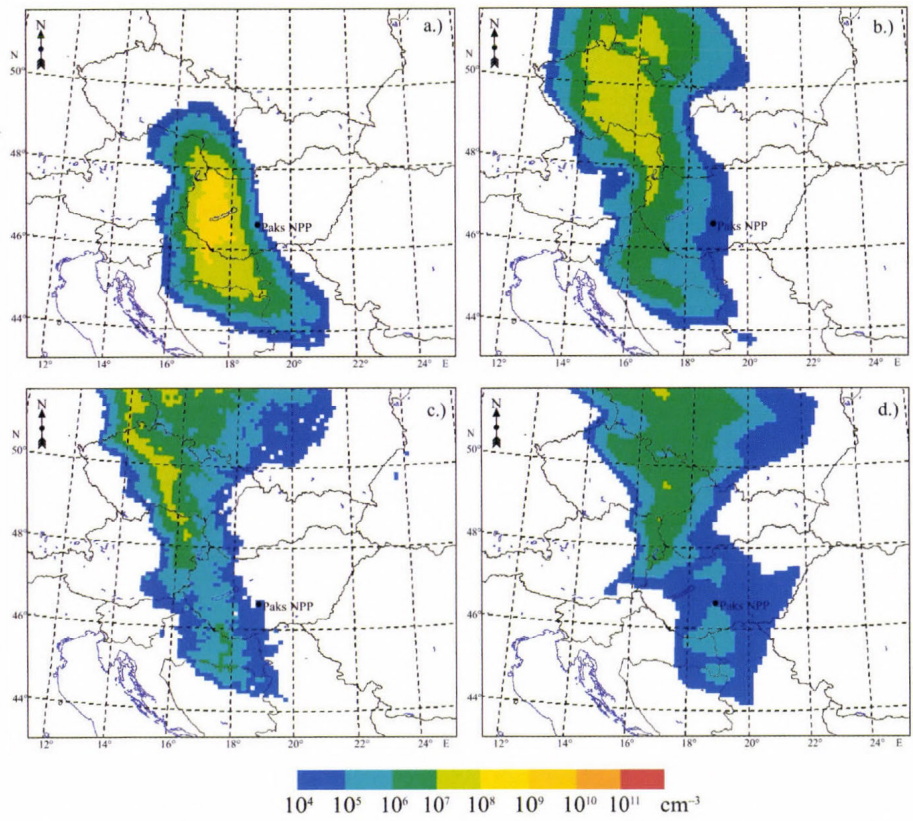


Fig. 3. Horizontal distribution map of ^{131}I concentration at 120 m above the surface, (a) 12, (b) 24, (c) 36, and (d) 48 hours after the hypothetical accident at 00 UTC, March 24, 2006.

During the period of the second simulation (March 24–25, 2006), an anticyclone was located over Central Europe, which caused sunny, dry weather. In the first 12 hours of the simulation, the plume was spread in a wide south-north belt, which had moved to the North and finally to the East (*Fig. 3*). Significant cumulative dry deposition was observed in south-northern axes (*Fig. 4a*), however, wet deposition (*Fig. 4b*) was very minor due to the dry weather conditions.

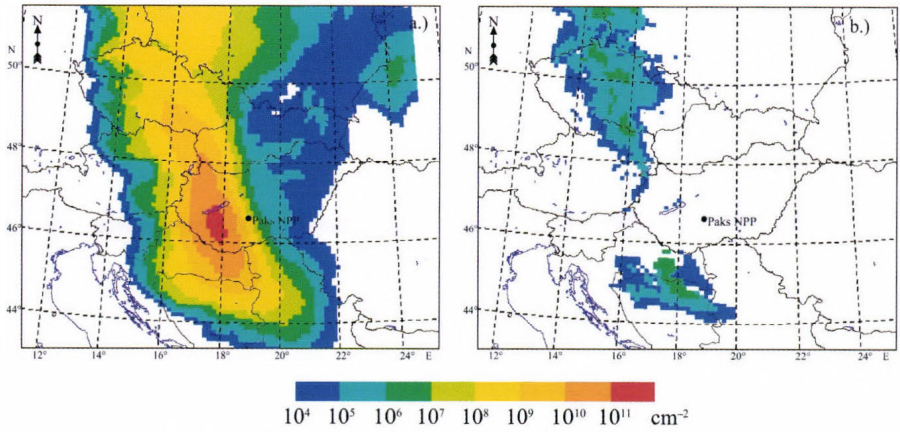


Fig. 4. Horizontal distribution map of cumulative (a) dry and (b) wet deposition of ^{131}I to the surface, 48 hours after the hypothetical accident at 00 UTC, March 24, 2006.

3.2. Efficiency of the semi-adaptive method

Efficiency of the semi-adaptive method is illustrated by running a set of simulations with the same meteorological and emission conditions used in the standard model runs described above, but during these simulations, the calculations have only been made on grid columns (on active cells), when the concentration of ^{131}I, at least at one grid cell in any of the 32 layers of a given grid column, exceeds a previously defined critical concentration. *Fig. 5* shows the number of active cells as a function of time, and critical concentrations. The higher the critical concentration is, the lower the number of active cells are. By increasing this critical concentration, higher acceleration can be achieved, since fewer calculations must be performed over the domain. In *Fig. 6*, the relative CPU time and the ratio of active cells can be seen as a function of the critical concentration for a 48-hour simulation (00 UTC, March 24, 2006 – 00 UTC, March 26, 2006). In case of relative CPU time, 1 refers to the CPU time of the simulation without the semi-adaptive method. The ratio of active cells is the maximum number of calculated horizontal grid cells during the whole simulation divided by the total number of horizontal grid cells 8640 (96×90). One can see that both the ratio of active cells and the relative CPU time

decreased with the increase of the critical concentration, but the size of the decreases are different. For lower critical concentrations (up to 10^7 cm^{-3}), the relative CPU time decreases rather than the maximum number of active cells during the simulation. The reason for this difference is that the number of active cells increases during the simulation. Accordingly, the acceleration is most effective after starting the simulation and decreases with time as the number of active cells increases.

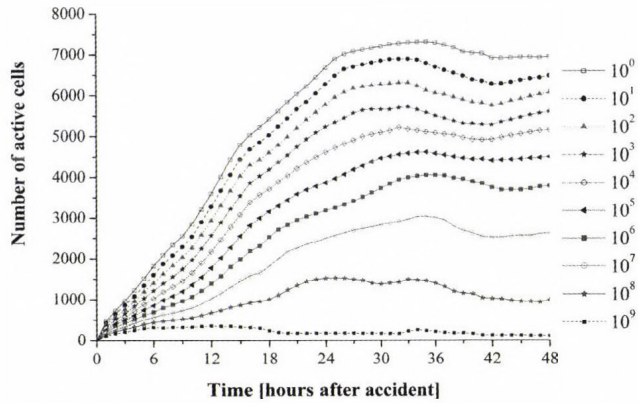


Fig. 5. Number of active cells during a 48-hour simulation after the hypothetical accident at 00 UTC, March 24, 2006 with different critical concentrations (in cm^{-3}).

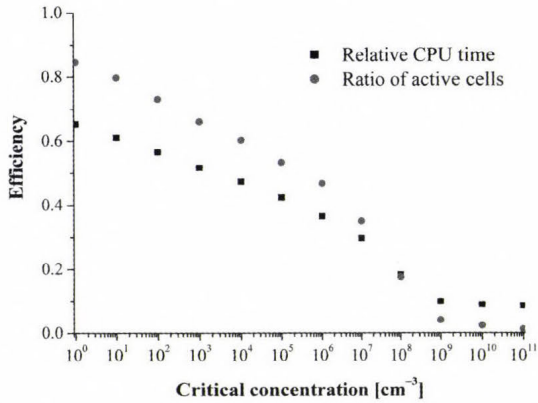


Fig. 6. Relative CPU time and ratio of active cells as a function of critical concentration. Test data: March 24–25, 2006.

In *Fig. 6*, the ratio of maximum number of active cells are presented, however, the relative CPU time corresponds to the whole period. In case of higher critical concentrations (from 10^9 cm^{-3}), with increase of the critical concentration the number of active cells slightly increases further, but it seems

that there is no further significant changes in the effective CPU time. This result suggests that there is a possible limit of the acceleration and any further increase of critical concentration will not increase the efficiency of the simulation. We should note that the CPU time also contains the time of unpacking, reading, and repacking of all meteorological datasets from the ALADIN model output files, and this time cannot be shortened by the semi-adaptive method.

The decrease of the simulation time of an accidental release model is crucially important and, at the same time, the method applied for the acceleration must be reliable. In *Fig. 7*, the horizontal distribution of the plume at the surface can be seen after 12 hours of a hypothetical accident using different critical concentrations. It appears that there is good agreement between maps with no critical concentration and lower critical concentrations (10^4–10^6 cm^{-3}). The polluted area and region with the highest concentrations can be captured similarly in all cases. However, some differences appear in the distribution, and these differences increase with increases in the critical value. For higher critical concentrations (10^7–10^9 cm^{-3}) the differences became significant, and only the most polluted area appears in the maps. Additionally, in some part of these regions, higher concentrations can be found as in case of standard model runs without the semi-adaptive method. Nevertheless, simulations using high critical concentrations can rapidly provide information about the direction of plume transport.

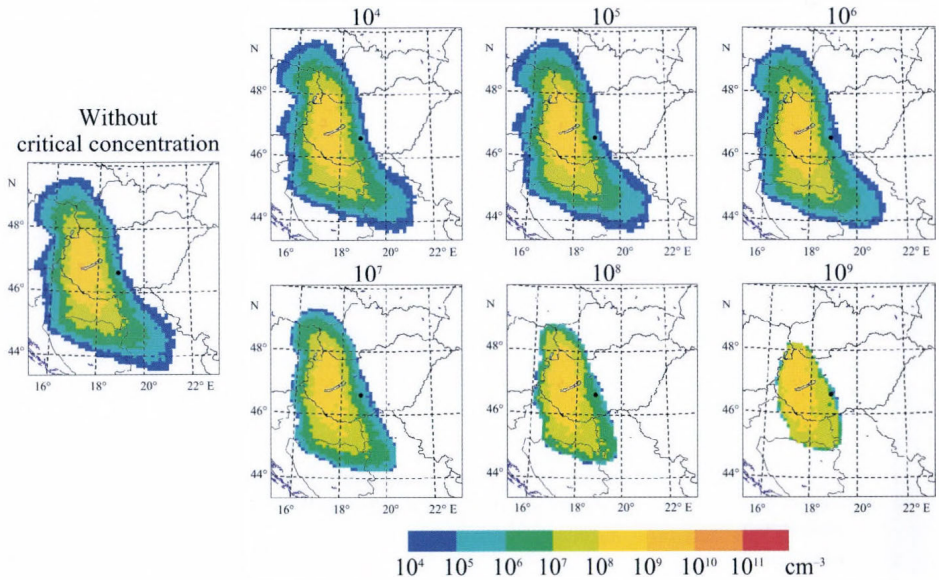


Fig. 7. Horizontal distribution map of ^{131}I concentration at the surface 12 hours after the hypothetical accident at 00 UTC, March 24, 2006 without and with different critical concentrations (in cm^{-3}). Same input meteorological and emission data were used in all cases. The black point denotes Paks NPP.

The differences depend not only on the choice of critical concentration but also on the simulation time. *Figs. 8a* and *b* present the relative differences of average and maximum concentrations of ^{131}I at the surface calculated without (c_1) and with (c_2) a critical concentration.

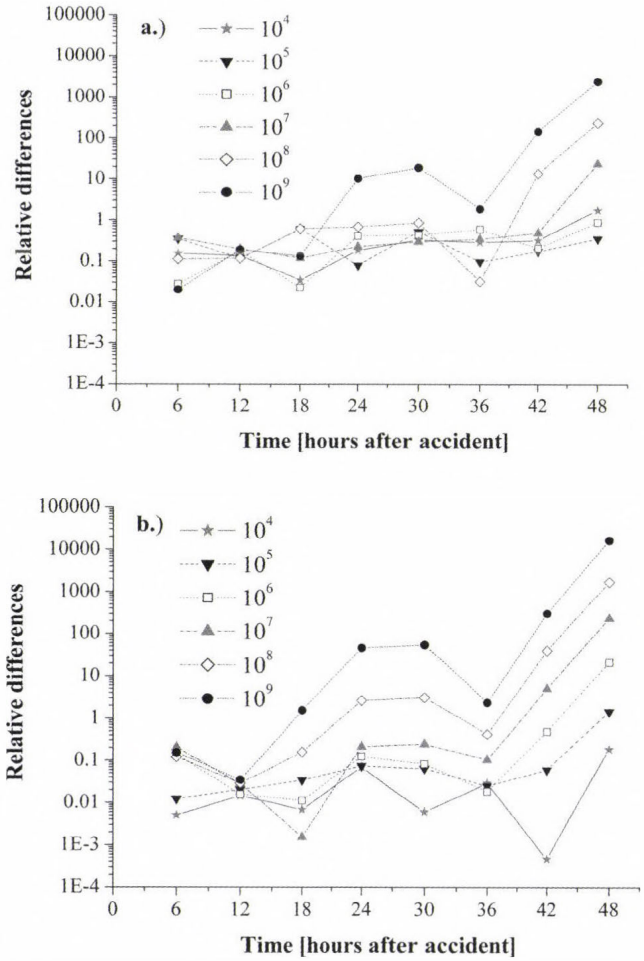


Fig. 8. (a) Relative differences of average ^{131}I concentration at the surface during a 48-hour simulation after the hypothetical accident at 00 UTC, March 24, 2006 in 6-hour time step. Values are calculated over the whole domain with and without semi-adaptive method with different critical concentrations (in cm^{-3}). (b) Same as Fig. 8a, but for maximum values.

The relative difference was calculated as $diff = |c_1 - c_2| / c_1$. Relative differences of average values (*Fig. 8a*) varied around 0.1 in general, but there are greater differences in case of 10^9 cm^{-3} critical concentration in the first

simulation day. From the 36th hour of the simulation, the differences became larger and the highest differences were found in case of higher critical concentrations. It is particularly noticeable, that differences decreased at 12 and 36 hours after the hypothetical accident (at 00 UTC). This is because of more turbulent conditions during daytime. Relative differences of maximum values (*Fig. 8b*) show greater variability. Nevertheless, if the critical concentration remains lower than 10^8 cm^{-3} till the 36th hour of the simulation, model estimation gives similar maximum values (the relative differences are small). It should be noted, that for a purpose of a shorter (for 12 hours) prediction of the maximum value, higher critical concentrations (10^8–10^9 cm^{-3}) should also provide appropriate results.

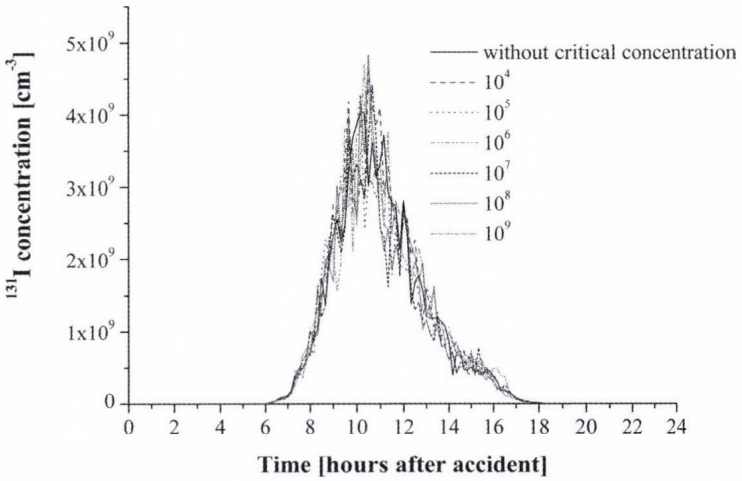


Fig. 9. Temporal variation of ^{131}I concentration at the surface over the city of Veszprém (47.1°N, 17.92°E) after the hypothetical accident at Paks NPP at 00 UTC, March 24, 2006. Values are calculated without and with different critical concentrations (in cm^{-3}).

The semi-adaptive method was also tested over given points. *Fig. 9* shows the temporal variations of ^{131}I concentration at the surface in Veszprém (47.1°N, 17.92°E, about 90 km to the NW direction of the point source) after a hypothetical accident in Paks NPP at 00 UTC, March 24, 2006, calculated with different critical concentrations. The plume passed over the city in the same interval, from 06 UTC to 18 UTC, and maximum concentrations appeared between 10 and 11 UTC in case of all different model runs with different critical concentrations. In the time series, about 20% variations on average can be found between values calculated with and without the semi-adaptive method. However, these differences are mainly caused by the stochastic calculation of the vertical transport term and not due to the different critical concentrations. This can be seen in *Fig. 10,* where two standard model runs without the semi-

adaptive method are compared. In spite of that in both cases the same meteorological and emission input fields were used, similar differences in concentration values can be seen similarly to *Fig. 9*, where model runs with different critical concentrations are presented. The differences between any two simulations should decrease with the use of more 'packages' in stochastic diffusion. These results also underline that the semi-adaptive method should be an effective way to accelerate the calculations.

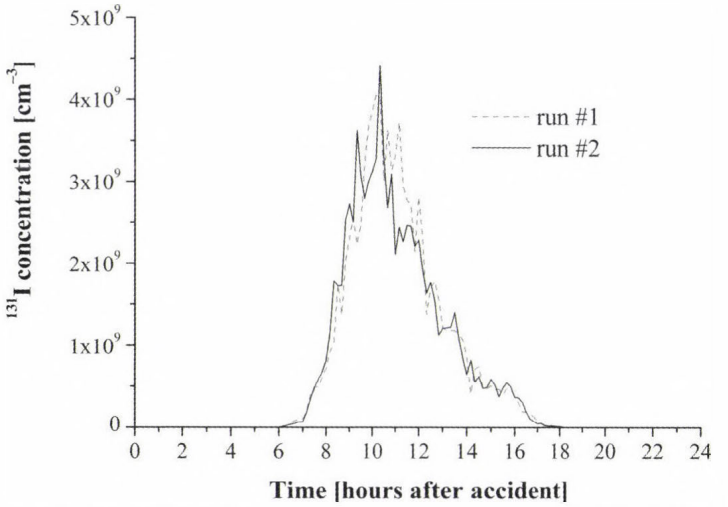


Fig. 10. Same as Fig. 9, but both time series were calculated without semi-adaptive method.

4. Conclusion

Modeling of the accidental release of radionuclides or highly toxic air pollutants requires fast and precise model estimations. In this study an Eulerian version of the TREX transport-exchange model is presented. We have also provided here a new method to accelerate the simulation, which is crucially important if a nuclear accident occurs, as we can gain time to prepare for the effects of the plume. With the semi-adaptive method, the efficiency of model simulations could be increased.

Two test simulations have been carried out for 48 hour periods. In case of each simulation, a hypothetical accident was assumed at the Paks Nuclear Power Plant, and the transport and deposition processes of ^{131}I have been estimated over a domain of $96 \times 90 \times 32$ grid cells. All meteorological fields for the simulations were obtained from the ALADIN numerical weather prediction model. Concentration fields, dry and wet deposition have been presented for two different meteorological conditions.

To apply the semi-adaptive method, the model simulations have been performed only on grid columns, where the concentration exceeds a predefined level in any of 32 layers in the column. The more efficient the method is, the most inactive cells are there, so the model became faster than without the semi-adaptive method, because less calculation is needed. The calculation time decreases with increasing critical concentration.

Based on our results, it seems that the semi-adaptive method could be an appropriate tool for accelerating the simulations. With high critical concentration, 10 times acceleration could be achieved compared to simulation without semi-adaptive method. However, there are some limits to this process. The higher the critical concentration is, the greater the differences are in the structure of the plume and in the estimated concentrations compared to the results obtained from normal simulation. The agreements between two cases are better in earlier part of simulation and generally decline with time, but more intense turbulence in daytime should be diminish the differences either during the latter part of simulation. Although high critical values (more than 10^7 cm^{-3} – compared to the emission rate used in this study) may cause large differences in simulated concentrations, it also should be usable for a fast, rough estimate for the direction of travel of the plume and the order of magnitude of the concentration. Determination of the critical concentration plays an important role of this method in decision support. It should be noted that further investigations should precede the implementation and any operational use of the semi-adaptive method in simulation of accidental release. Results of this study contribute to the further development of TREX dispersion model system.

Acknowledgements—The authors would like to thank *András Horányi* and *László Kullmann* for the ALADIN data. This research is supported by Hungarian Research Found (OTKA K68253, OTKA K81933 and OTKA K81975) and by SEE-GRID-SCI (SEE-GRID eInfrastructure for regional eScience) project, funded by the European Commission through the contract nr RI-211338.

References

Alexandrov, V.N., Owczarz, W., Thomson, P.G. and *Zlatev, Z.*, 2004: Parallel runs of a large air pollution model on a grid of Sun computers. *Math Comput Simulat 65,* 557-577.

Arya, S.P., 1988: Introduction to Micrometeorology. Academic Press, INC, San Diego. pp. 307.

Baklanov, A. and *Sørensen, J.H.*, 2001: Parameterisation of Radionuclide Deposition in Atmospheric Long-Range Transport Modelling. *Phys Chem Earth (B) 26*, 787-799.

Baklanov, A., Sørensen, J.H., Hoe, S.C. and *Amstrup, B.*, 2006: Urban meteorological modelling for nuclear emergency preparedness. *J Environ Radioactiv 85*, 154-170.

Bartnicki, J., Saltbones, J. and *Foss, A.*, 2003: Performance of the SNAP model in an ENSEMBLE exercise of simulating real-time dispersion from a nuclear accident. *Int J Environ Pollut 20*, 1-6, 22-32.

Baumann, K. and *Stohl, A.*, 1997: Validation of a long-range trajectory model using gas balloon tracks from the Gordon Bennett Cup 95. *J Appl Meteor 36*, 711-720.

Bellasio, R., Bianconi, R., Graziani, G. and *Mosca, S.*, 1999: RTMOD: an Internet based system to analyse the predictions of long-range atmospheric dispersion models. *Computers & Goesciences 25*, 819-833.

Brandt, J., 1998: Modelling transport, dispersion and deposition of passive tracers from accidental releases. PhD thesis. Ministry of Environment and Energy National Environmental Research institute and Ministry of Research and Information Technology Risø National Laboratory, Denmark, Roskilde. pp 307.

Brandt, J., Christensen, J.H., Frohn, L.M. and *Zlatev, Z.,* 2000: Numerical modelling of transport, dispersion, and deposition – validation against ETEX-1, ETEX-2 and Chernobyl. *Environ Model Softw 15,* 521-531.

Brandt, J., Christensen, J.H. and *Frohn, L.M.,* 2002: Modelling transport and deposition of caesium and iodine from the Chernobyl accident using the DREAM model. *Atmosph Chem and Phys,* 397-417.

Businger, J.A., Wyngaard, J.C., Izumi, Y. and *Bradley, E.F.,* 1971: Flux-profile relationships in the atmospheric surface layer. *J Atmos Sci 28,* 181-189.

Dabdub, D. and *Seinfeld, J.H.,* 1996: Parallel computation in atmospheric chemical modelling. *Parallel Comput 22,* 111-130.

Dimov, I., Georgiev, K., Ostromsky, T. and *Zlatev, Z.,* 2004: Computational challenges in the numerical treatment of large air pollution models. *Ecol Model 179,* 187-203.

Dombvári, P., Ranga, T., Nényei, Á., Bujtás, T., Kovács, T., Jobbágy, V., Vincze, Cs. and *Molnár, F.,* 2008: Új terjedésszámító szoftver fejlesztése és bevezetése a Paksi Atomerőműnél. (in Hungarian) *Sugárvédelem I. 1,* 30-36.

Ehrhardt, J., Brown, J., French, S., Kelly, G.N., Mikkelsen, T. and *Müller, H.,* 1997: RODOS: Decision-making support for off-site emergency managment after nuclear accidents. *Kerntechnik 62,* 122-128.

Ferenczi, Z. and *Ihász, I.,* 2003: Validation of the Eulerian dispersion model MEDIA at the Hungarian Meteorological Service. *Időjárás 107,* 115-132.

Galmarini, S., Bianconi, R., Bellasio, R. and *Graziani, G.,* 2001: Forecasting the consequences of accidental releases of radionuclides in the atmosphere from ensemble dispersion modelling. *J Environ Radioactiv 57,* 203-219.

Galmarini, S., Bianconi, R., Klug, W., Mikkelsen, T., Addis, R., Andronopoulos, S., Astrup, P., Baklanov, A., Bartniki, J., Bartzis, J.C., Bellasio, R., Bompay, F., Buckley, R., Bouzom, M., Champion, H., D'Amours, R., Davakis, E., Eleveld, H., Geertsema, G.T., Glaabl, H., Kollax, M., Ilvonen, M., Manning, A., Pechinger, U., Persson, C., Polreich, E., Potemski, S., Prodanova, M., Saltbones, J., Slaperj, H., Sofiev, M.A., Syrakov, D., Sørensen, J.H., Van der Auwera, L., Valkama, I. and *Zelazny, R.,* 2004: Ensemble dispersion forecasting – Part I: concept, approach and indicators. *Atmos Environ 38,* 4607-4617.

Gultureanu, D., Gultureanu, B., Mikkelsen, T., Thykier-Nielsen, S. and *Matenciuc, M.,* 2000: Air pollution studies for the Romanian nuclear power plant – Cernavoda based on the RIMPUFF dispersion model. *Air Pollution Modeling and Its Applications XIII Book Series: NATO challenges of modern society 13,* 779-780.

Holtslag, A.A.M. and *De Bruin, H.A.R.,* 1988: Applied Modeling of the Nighttime Surface Energy Balance over Land. *J Appl Meteorol 27,* 689-704.

Horányi, A., Ihász, I., and *Radnóti, G.,* 1996: ARPEGE/ALADIN: A numerical weather predicition model for Central-Europe with the participation of the Hungarian Meteorological Service. *Időjárás 100,* 277-300.

Horányi, A., Kertész, S., Kullmann, L. and *Radnóti, G.,* 2006: The ARPEGE/ALADIN mesoscale numerical modeling system and its application at the Hungarian Meteorological Service. *Időjárás 110,* 203-227.

Jones, A., Thomson, D., Hort, M. and *Devenish, B.,* 2007: The UK Met Office's next-generation atmospheric dispersion model, NAME III. *Air Pollution Modeling and Its Applications XVII Book Series: NATO challenges of modern society 17,* 580-589.

Lagzi, I., Tomlin, A.S., Turányi, T., Haszpra, L., Mészáros, R. and *Berzins, M.,* 2001: The simulation of photochemical smog episodes in Hungary and Central Europe using adaptive gridding models. *Lect Notes Comput Sc 2074,* Part II, 67-70.

Lagzi, I., Kármán, D., Tomlin, A.S., Turányi, T. and *Haszpra, L.,* 2004: Simulation of the dispersion of nuclear contamination using an adaptive Eulerian grid model. *J Environ Radioactiv 75,* 59-82.

Lagzi, I., Mészáros, R., Ács, F., Tomlin, A.S., Haszpra, L. and *Turányi, T.,* 2006: Description and

evaluation of a coupled Eulerian transport-exchange model: Part I: model development. *Időjárás 110,* 349-363.

Lagzi, I., Tomlin, A.S., Turányi, T. and *Haszpra, L.,* 2009: Modelling photochemical air pollutant formation in Hungary using an adaptive grid technique. *Int J Environ Pollut 36,* 44-58.

Langnera, J., Robertsona, L., Perssona, C. and *Ullerstiga, A.,* 1998: Validation of the operational emergency response model at the Swedish Meteorological and Hydrological Institute using data from etex and the chernobyl accident. *Atmos Environ 32,* 4325-4333.

Larson, D.J. and *Nasstrom, J.S.,* 2002: Shared- and distributed-memory parallelization of a Lagrangian atmospheric dispersion model. *Atmos Environ 36,* 1559-1564.

Lovas, R., Patvarczki, J., Kacsuk, P., Lagzi, I., Turányi, T., Kullmann, L., Haszpra, L., Mészáros, R., Horányi, A., Bencsura, A. and *Lendvay, Gy.,* 2006: Air pollution forecast on the HUNGRID infrastructure. In *Parallel Comput: Current & Future Issues of High-End Computing,* (eds.: *G.R. Joubert, W.E. Nagel, F.J. Peters, O.G. Plata, P. Tirado, E.L. Zapata*). *John von Neumann Institute for Computing Series, Vol. 33, Jülich, Germany,* 121-128, ISBN 3-00-017352-8.

Martin, M., Oberson, O., Chopard, B., Mueller, F. and *Clappier, A.,* 1999: Atmospheric pollution transport: the parallelization of a transport & chemistry code. *Atmos Environ 33,* 1853-1860.

Martín, M.J., Parada, M. and *Doallo, R.,* 2004: High performance air pollution simulation using OpenMP. *J. Supercomput 28,* 311-321.

Mészáros, R., Lagzi, I., Juhász, Á., Szinyei, D., Vincze, Cs., Horányi, A., Kullmann, L. and *Tomlin, A.S.,* 2006: Description and evaluation of a coupled Eulerian transport-exchange model: Part II: Sensitivity analysis and application. *Időjárás 110,* 365-377.

Mikkelsen, T., Thykier-Nielsen, S., Astrup, P., Santabarbara, J.M., Sørensen, J.H., Rasmussen, A., Robertson, L., Ullerstig, A., Deme, S., Martens, R., Bartzis, J.G. and *Pasler-Sauer, J.,* 1997: MET-RODOS: A comprehensive atmospheric dispersion module. *Radiat Prot Dosim 73,* 45-56.

Molnár, F. Jr., Szakály, T., Mészáros, R. and *Lagzi, I.,* 2010: Air pollution modelling using a Graphics Processing Unit with CUDA. *Comput Phys Commun 181,* 105-112.

Ostromsky, T., Dimov, I.T. and *Zlatev, Z.,* 2005: Parallel implementation and one year experiments with the Danish Eulerian Model. *Lect Notes Comput Sc 3401,* 440-447.

Paulson, C.A., 1970: The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J Appl Meteorol 9,* 857-861.

Pechinger, U., Langer, M., Baumann, K. and *Petz, E.,* 2001: The austrian emergency response modelling system TAMOS. *Phys. Chem. Earth (B) 26,* 99-103.

Pudykiewicz, J., 1989: Simulation of the Chernobyl dispersion with a 3-D hemispheric tracer model. *Tellus 41B,* 391-412.

Pudykiewicz, J., 1991: Environmental prediction systems: Design, implementation aspects and operational experience with application to accidental releases. In *Air Pollution Modeling and Its Applications VIII. NATO challenges of modern society* (eds.: *H. van Dop* and *D.G Steyn*). *Vol. 15.* Plenum Press, New York and London, pp. 561-590.

Ryall, D.B. and *Maryon, R.H.,* 1998: Validation of the UK Met. Office's name model against the ETEX dataset. *Atmos. Environ. 32,* 4265-4276.

Saltbones, J., Foss, A. and *Bartnicki, J.,* 1998: Norwegian Meteorological Institute's real-time dispersion model SNAP (severe nuclear accident program): Runs for ETEX and ATMES II experiments with different meteorological input. *Atmos. Environ. 32,* 4277-4283.

Singh, D.E., Garcia, F. and *Carretero, J.,* 2006: Parallel I/O optimization for an air pollution model. *Parallel Comput: Current & Future Issues of High-End Computing. Proceedings of the International Conference ParCo 2005* (eds.: *G.R. Joubert, W.E. Nagel, F.J. Peters, O. Plata, P. Tirado, E. Zapata*), *33,* 523-530.

Sørensen, J.H., Rasmussen, A., Ellermann, T. and *Lyck, E.,* 1998. Mesoscale influense on long-range transport – evidence from ETEX modelling and observations. *Atmos Environ 32,* 4207-4217.

Sørensen, J.H., Baklanov, A. and *Hoe, S.,* 2007. The Danish emergency response model of the atmosphere (DERMA). *J Environ Radioactiv 96,* 122-129.

Srinivas, C.V., Venkatesan, R., Muralidharan, N.V., Das, S., Dass, H. and *Kumar, P.E.,* 2006: Operational mesoscale atmospheric dispersion prediction using a parallel computing cluster. *J Earth System Science 115,* 315-332.

Stohl, A. and *Wotawa, G.*, 1997: Validation of the Lagrangian particle model FLEXPART using ETEX data. ETEX syposium on long-range atmospheric transport, model verification and emergency response. *European Commission* EUR 17346,167-170.

Stohl, A., Hittenberger, M. and *Wotawa, G.*, 1998: Validation of the lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data. *Atmos Environ 32*, 4245-4264.

Stohl, A., Forster, C., Frank, A., Seibert, P. and *Wotawa, G.*, 2005: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos Chem and Phys 5,* 2461-2474.

Zegeling, P.A. and *Kok, H.P.*, 2004: Adaptive moving mesh computations for reaction-diffusion systems. *J. Comput Appl Math 168,* 519-528.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–June 2010, pp. 121–133

The quasi-operational LAMEPS system of the Hungarian Meteorological Service

Edit Hágel

Hungarian Meteorological Service
P.O. Box 38, H-1525 Budapest, Hungary; E-mail: hagel.e@met.hu

(Manuscript received in final form February 11, 2010)

Abstract — Dynamical downscaling of the global ARPEGE based ensemble prediction system PEARP is running quasi-operationally at the Hungarian Meteorological Service since February 2008. For the downscaling of the PEARP members, the ALADIN limited area model is used with 12 km horizontal resolution. Both systems have 10+1 ensemble members (one control member that starts from the unperturbed analysis and 10 members which start from perturbed initial conditions). At present, both the initial and lateral boundary conditions of the ALADIN runs are provided by the PEARP ensemble members, however, it is planned to generate the initial condition perturbations locally at a later stage. In the article this quasi-operational short-range limited area ensemble prediction system, as well as its performance through verification results will be presented.

Key-words: numerical weather prediction, limited area forecasting, predictability, ensemble method, verification

1. Introduction

In order to predict the future state of the atmosphere, the use of mathematical models is required. Forecasts are made by solving a set of partial differential equations, the so-called primitive equations. These equations are nonlinear and, therefore, impossible to solve analytically. Because of the nonlinear nature of the equations, the solution is highly dependent on the accuracy of the initial conditions. The problem is that the true state of the atmosphere cannot be known exactly. The reason of this is that the number of the observations is limited (smaller than the degrees of freedom in the models), their distribution is uneven around the globe, there are inevitable observational errors, and also errors in the data assimilation techniques. As a result, there will always be some uncertainty in the initial conditions of the numerical weather prediction (NWP) models.

(One should not forget that the initial condition error is only one source of the possible forecast errors, there are errors in the models themselves. However, in the system described in the present article we only dealt with the initial condition errors.) Possible solution of the above mentioned problem is to run a set, or as usually called, an ensemble of forecasts, each starting from a slightly different initial condition, thus, they are equally likely realizations of the "true" atmospheric state. The advantage of this method is clear: the spread of the ensemble members can provide useful information on the predictability of the atmospheric state, and a probability value can be assigned to different weather events. In other words, not only the future state of the atmosphere can be forecasted, but one can also predict the uncertainty related to this forecast.

Since its first operational application in 1992 (see *Buizza et al.*, 1993; *Toth* and *Kalnay*, 1997), ensemble forecasting has become a widely used technique by many meteorological services around the world. Despite its obvious benefits, it was used only on global scales and in the medium-range for a long time. In the last couple of years, intensive research has started to apply the ensemble method in short-range limited area forecasting as well (see, e.g., *Montani et al.*, 2003; *Jensen et al.*, 2006; *García-Moya et al.*, 2007). Motivated by the results of these experiments, research started at the Hungarian Meteorological Service (HMS) with the final aim to establish an operational LAMEPS[1] system for the Central European area, and to see how it can improve the predictions of the existing global forecasting systems. It was decided to start the experiments with the dynamical downscaling of global ensemble forecasts. Two possible choices were considered: downscaling of ARPEGE[2] ensemble forecasts (see *Hágel* and *Horányi*, 2006, 2007 for a detailed description of the experiments) and downscaling of ECMWF[3] EPS[4] members (for further information on ECMWF EPS downscaling see *Szintai* and *Ihász* (2006)).

Since February 2008, a LAMEPS system is running in quasi-operational status at the HMS. The system is based on the dynamical downscaling of the ARPEGE (global) ensemble system, PEARP[5]. The aim of this article is to present this quasi-operational short-range limited area ensemble prediction system. The characteristics of the system are described in Section 2. Verification results are presented in Section 3, and finally, conclusions and future plans are outlined in Section 4.

1 LAMEPS: Limited Area Model Ensemble Prediction System
2 ARPEGE: Action de Recherche Petite Echelle Grande Echelle (i.e., Research Project on Small and Large Scales)
3 ECMWF: European Centre for Medium-Range Weather Forecasts
4 EPS: Ensemble Prediction System
5 PEARP: Prévision d'Ensemble ARPege

2. *Characteristics of the system*

The 11-member short-range limited area ensemble prediction system of HMS has been running every day, in quasi-operational status, since February 2008. It is run with the ALADIN[6] limited area model and it is driven by the members of the global PEARP system. PEARP is the operational global ensemble prediction system of Météo-France. The initial perturbations of this global system are created as a combination of singular vectors (targeted over four different areas) and evolved perturbations of the previous PEARP run. PEARP has 10 + 1 members (10 perturbed and the unperturbed control member) and it is run once a day starting from the 18 UTC analysis.

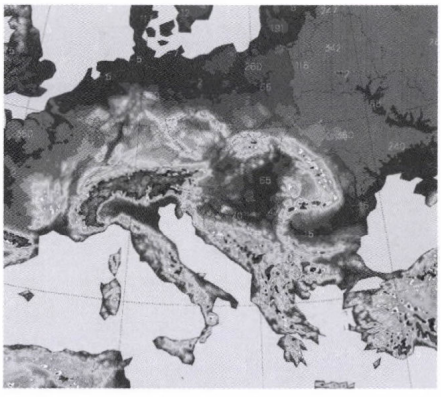


Fig. 1. The integration domain and the orography of the ALADIN LAMEPS system.

The members of the global PEARP system are downscaled at HMS with the limited area model ALADIN (for more information on the ALADIN model, see *Horányi et al.*, 2006). At present, no local data assimilation or generation of local perturbations are applied for the LAMEPS. Forecasts are made once a day starting from the 18 UTC data. In order to be able to use the outputs of the global model as initial and lateral boundary conditions, an interpolation is needed to the exact domain and resolution which is used for the model integration. Once the initial and lateral boundary conditions are in the proper format (resolution, domain, etc.), the integration of the model can start. The ALADIN ensemble system is running on a domain covering a large part of Continental Europe (*Fig. 1*) with a horizontal resolution of approximately 12 km. In the vertical 46 levels are used. Forecast length is 60 hours and the time step used for the integration is 450 seconds (7.5 minutes). Finally, post-processing is performed on the raw model outputs in order to support the application of the model results by forecasters or end-users. After performing post-processing to a

6 ALADIN: Aire Limitée Adaptation dynamique Développement International

latitude-longitude grid, the outputs of the LAMEPS system are mainly visualized using HAWK[7] which is used in the everyday work of forecasters to visualize the outputs of several NWP models (both deterministic and probabilistic), observations, radar and satellite data, etc. The available products from our LAMEPS system are the ensemble mean, the ensemble spread (computed around the mean), individual ensemble members, and probability fields for several parameters. The individual members can be visualized in the form of spaghetti diagrams. In addition, plume diagrams are also plotted for several parameters and selected Hungarian locations.

3. Verification results

Verification of the quasi-operational LAMEPS system was performed for almost nine months from March 10, 2008 to November 30, 2008 against the ECMWF analysis. Scores were computed for the whole period and also for the different seasons separately (only the results for the whole period are shown, not the separate seasons). The verified parameters were temperature, geopotential, and wind speed on several levels (500 hPa, 700 hPa, 850 hPa, 925 hPa, and 1000 hPa). The common LACE[8] verification package was used. This verification package was developed in collaboration with colleagues from other LACE countries (see *Hágel*, 2006; *Mladek*, 2006; *Kann*, 2007). For a detailed description of the verification methods, the reader is referred to *Jolliffe* and *Stephenson* (2003).

3.1. Comparing the error of the ensemble mean and the control member

As a first step, the error of the control member was compared with the error of the ensemble mean. Since the perturbations in the global PEARP system are symmetric around the unperturbed initial condition and have a small amplitude at initial time, the ensemble mean and the control forecast are almost identical in the early forecast ranges. As no local perturbations are added, the same behavior is assumed for the LAMEPS. The similarity of the ensemble mean and the control member means that their RMSE (root mean squared error) is also very similar. However, after the initial linear phase it is expected that the ensemble mean has lower RMSE values than the control forecast, since the averaging has the effect of filtering out the less predictable features and leaving only the more predictable ones, those that show agreement among the members of the ensemble.

For geopotential the RMSE of the control member and the RMSE of the ensemble mean were almost identical during the whole period. (*Fig. 2a* shows the results for 500 hPa.) For temperature (*Fig. 2b* shows the results for 500 hPa) and wind speed (see *Fig. 2c* for the scores on 1000 hPa) there was a difference

7 HAWK: Hungarian Advanced WorKstation (http://www.met.hu/nmo/hawk_en)
8 LACE: Limited Area modeling in Central Europe (www.rclace.eu)

between the ensemble mean and the control forecast already after the first 6 hours, with the ensemble mean having lower RMSE values. Results indicate that in the case of geopotential, the ensemble members remain centered around the control member during almost the whole forecast period, hence the RMSE of the ensemble mean and the control member are almost identical. For wind speed and temperature the ensemble mean and the control forecast start to differ from the early forecast ranges, which suggests that nonlinearity has a stronger effect on these parameters.

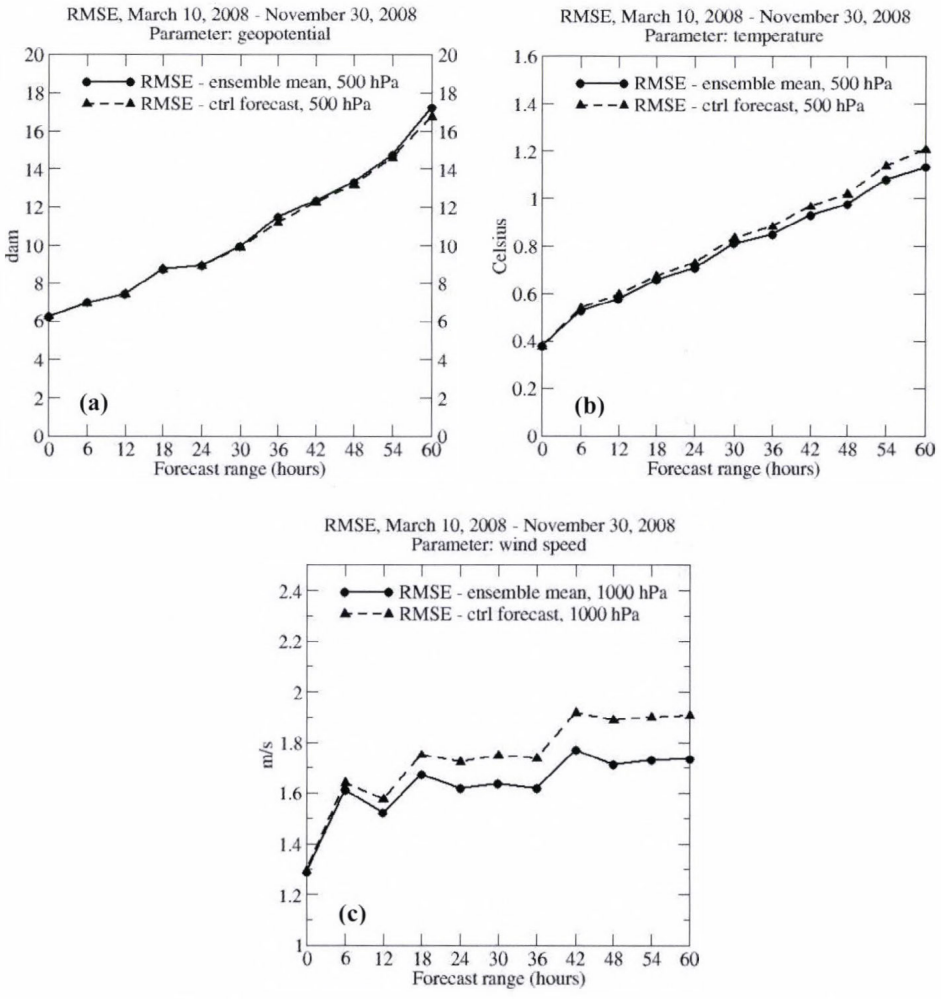


Fig. 2. RMSE of the ensemble mean and RMSE of the control member for (a) geopotential at 500 hPa, (b) temperature at 500 hPa, and (c) wind speed at 1000 hPa. The verification interval is March 10, 2008–November 30, 2008. Verification was performed against ECMWF analysis.

3.2. Spread-skill relationship, percentage of outliers

Another important feature of an ensemble system is the spread-skill relationship. The spread of the ensemble system (computed around the ensemble mean) should be in a good agreement with the forecast error (e.g., RMSE of the ensemble mean). In the case of large error, large spread is expected as a sign of high unpredictability. On the other hand, if the spread is small, it is expected that the situation has good predictability, therefore, the error should be small as well. If the spread is larger (smaller) than the error, then the system is said to be over- (under-) dispersive.

In *Fig. 3*, the spread-skill relationship is plotted for geopotential, temperature, and wind speed for two levels: 500 and 850 hPa. In all cases the system was found to be underdispersive, i.e., the spread was smaller than the RMSE of the ensemble mean.

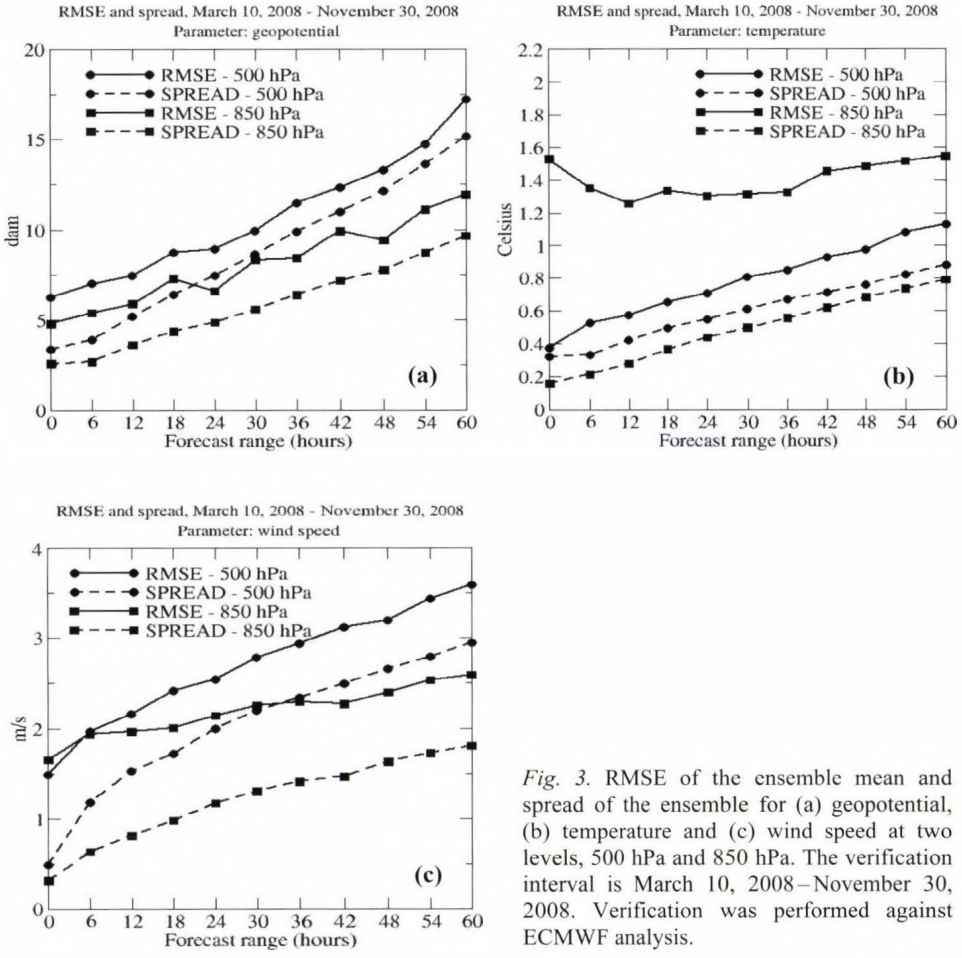


Fig. 3. RMSE of the ensemble mean and spread of the ensemble for (a) geopotential, (b) temperature and (c) wind speed at two levels, 500 hPa and 850 hPa. The verification interval is March 10, 2008–November 30, 2008. Verification was performed against ECMWF analysis.

Another way of analyzing the spread of an ensemble system is the use of percentage of outliers diagrams. These diagrams tell us how often the verifying analysis lies out of the interval defined by the (sorted) ensemble members. *Fig. 4* shows the percentage of outliers for geopotential, temperature, and wind speed at 500, 700, 850, 925, and 1000 hPa.

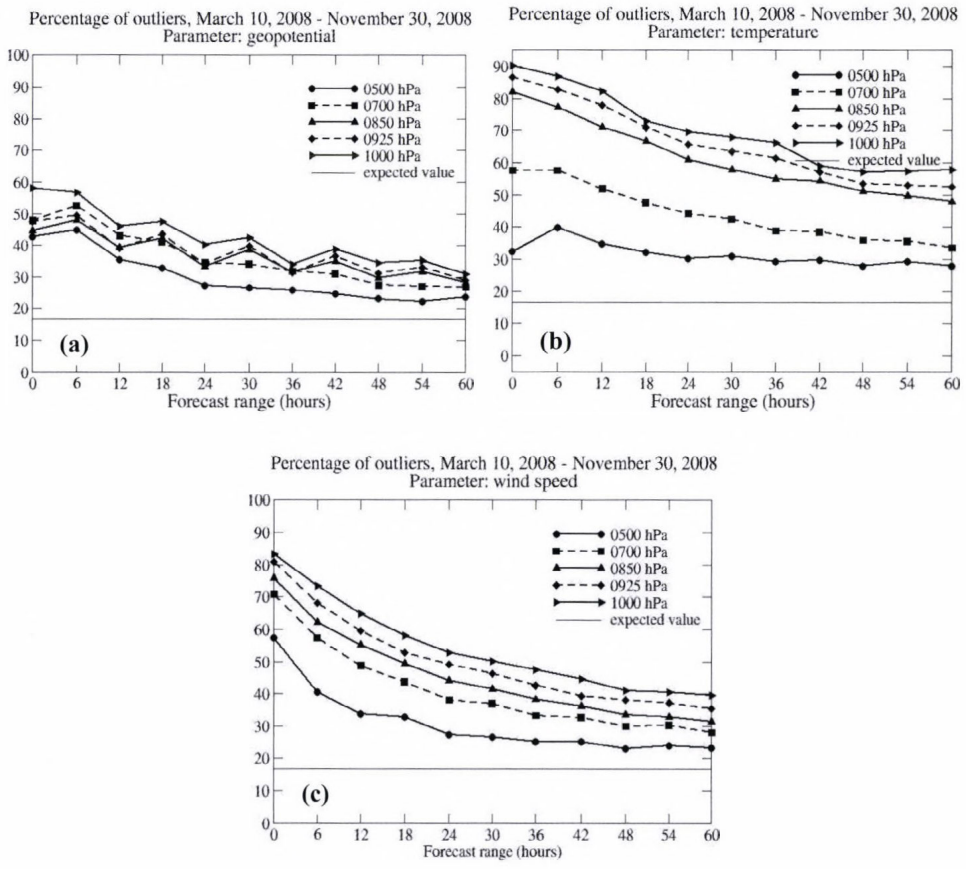


Fig. 4. Percentage of outliers diagrams for (a) geopotential (b) temperature, and (c) wind speed at 500 hPa, 700 hPa, 850 hPa, 925 hPa, and 1000 hPa. Values are in %. The verification interval is March 10, 2008 – November 30, 2008. Verification was performed against ECMWF analysis. The thin horizontal line is the expected value, i.e., $2 \cdot 100/(\text{ensemble members} + 1)$.

For all levels and parameters, the percentage of outliers was above the expected value. This means that the verifying analysis too often falls out of the interval defined be the (sorted) ensemble members, indicating that the spread of the ensemble is not sufficient. Both in case of the spread-skill relationship and the

percentage of outliers diagrams, best results were obtained for 500 hPa and got worse as going closer to the surface. This behavior might have several reasons, related to the perturbations used in the PEARP system:

- No perturbation of surface parameters, except for surface pressure.
- The maximum of the energy of the (global) singular vectors is located around 700 hPa and during their evolution the energy propagates upwards, rather than downwards.
- The uncertainties related to the physical parameterizations are not addressed at all during the singular vector computation.

This behavior (better results for higher levels) suggests that surface perturbations should also be included in the system. One can also conclude that even for 500 hPa, the spread of the ensemble is smaller than the RMSE of the ensemble mean, and the percentage of outliers is above the expected value. This indicates that local perturbations, targeted especially to the area of our interest would be needed to improve the quality of the system.

3.3. ROC and reliability diagrams

ROC[9] and ROC area diagrams represent the skill of the ensemble system compared to the use of climatological statistics. A ROC area of 1 represents a perfect system, while an area less than 0.5 means the forecasts have no skill compared to climatological data. Reliability diagrams are used to test the ability of the system to correctly forecast probabilities of a certain event. For that reason, forecast probabilities are plotted against conditional observed frequencies (with forecast probabilities on the x-axis and observed frequencies on the y-axis). For a perfect system the points lie along the diagonal.

ROC and reliability diagrams were plotted for wind speed only (for technical reasons), with thresholds 1, 2, 5, and 10 m/s on five levels (500, 700, 850, 925, and 1000 hPa). Comparison was made between (i) different thresholds, (ii) different forecast ranges, and (iii) different vertical levels. The comparison between the different thresholds reveals that results are better for higher wind speeds. For reliability diagrams, in the case of 1 m/s threshold, the system shows significant underestimation for low and middle probabilities and slight overestimation for high probabilities. For higher thresholds, the underestimation for low probabilities and overestimation for high probabilities remains, but the curves move significantly closer to the diagonal. (Fig. 6a shows examples for 1000 hPa and T+60 hours.) Similar results were obtained by analyzing the ROC diagrams (*Fig. 5a*): the ROC area was well above 0.5 for all thresholds, with more skillful forecasts for wind speeds larger than, e.g., 10 m/s than wind speeds exceeding 1 m/s.

9 ROC: Relative Operating Characteristics

The comparison between different forecast ranges shows that the system has similar skill throughout the whole forecast interval, however, results are somewhat worse in the early forecast ranges according to the reliability diagrams (this was more pronounced in autumn than during spring and summer). *Fig. 5b* and *Fig. 6b* show examples for 1000 hPa and 10 m/s as threshold.

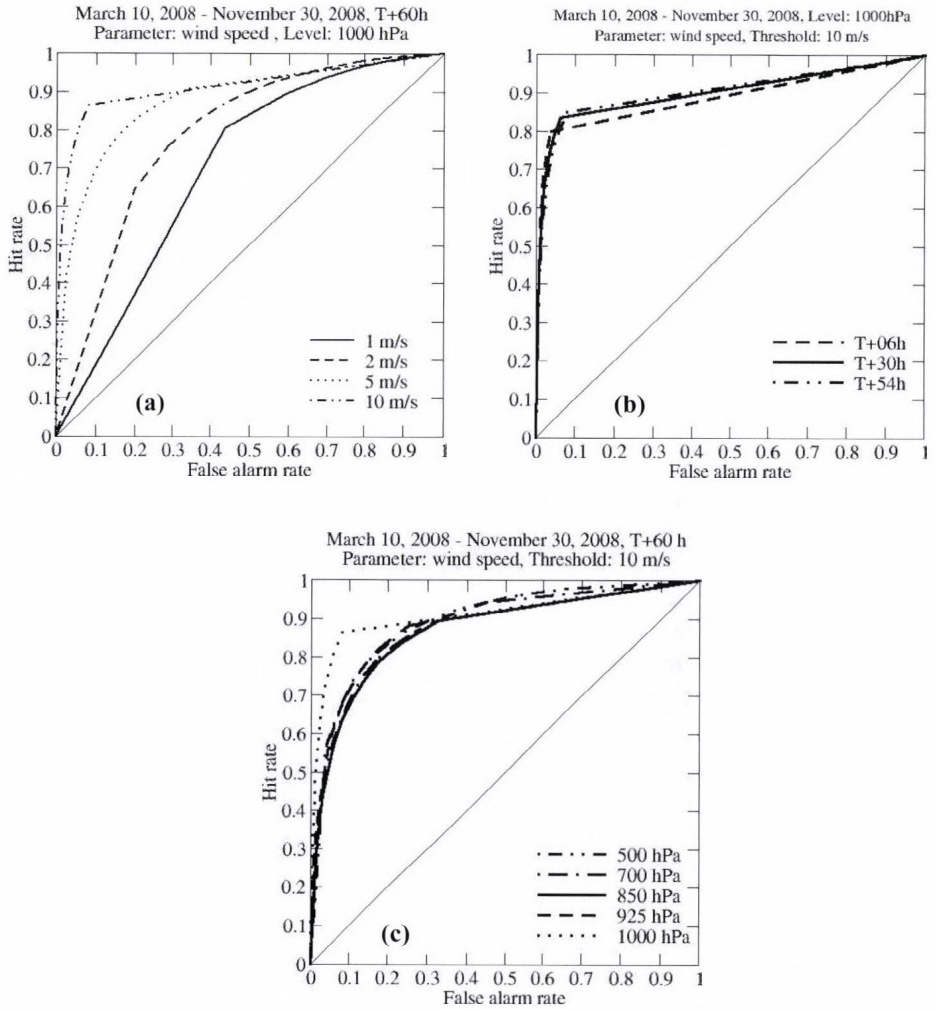


Fig. 5. ROC diagrams for wind speed. (a) Diagrams for thresholds 1 m/s, 2 m/s, 5 m/s, and 10 m/s at 1000 hPa and T+60 hours. (b) Diagrams for time range T+06 hours, T+30 hours, and T+54 hours at 1000 hPa and 10 m/s threshold. (c) Diagrams for 500 hPa, 700 hPa, 850 hPa, 925 hPa, and 1000 hPa at T+60 hours and 10 m/s threshold. The verification interval is March 10, 2008–November 30, 2008. Verification was performed against ECMWF analysis.

As regards the different vertical levels, ROC and reliability curves were plotted for different vertical levels as well. In terms of ROC, results are quite similar for all levels, the only exception is 1000 hPa where the shape of the curves was slightly different, but the ROC area was similar to the other levels. *Fig. 5c* shows examples for T+60 hours and 10 m/s as threshold. Reliability diagrams show that for low probabilities results are quite similar for all levels and show a very good reliability in case of higher thresholds. For high probabilities, differences are more significant and the curves lie somewhat farther from the diagonal with better results for higher levels. This is in agreement with the conclusions drawn from other verification measures. *Fig. 6c* shows examples for T+60 hours and a 10 m/s threshold.

3.4. Comparing the performance of the global and the limited area systems

When running limited area forecasts, it is always important to know whether the limited area model can improve the predictions of the global model or not. Verification results of the global and the limited area ensemble systems were compared to each other during the pre-operational experiments (*Hágel* and *Horányi*, 2006, 2007) for the period from January 15, 2005 to February 15, 2005. It can be said that by simply downscaling the global PEARP forecasts using the higher resolution ALADIN model, it is very difficult to achieve significant – overall – improvements. For some parameters and verification measures the limited area ensemble forecasts performed better, in other cases the global forecasts were more skillful. Also, in a couple of cases, the two models had nearly the same scores. Some aspects behind these results might be the relatively small resolution difference between the global and the limited area models (approximately 23 km and 12 km, respectively), or the too strong impact of the lateral boundary conditions.

As limited area models are running with a better horizontal resolution than their global counterparts, the representation of the orography is more realistic, which is very important in the case of, e.g., wind or precipitation forecasts. As a result of the increased resolution, limited area models can produce more reliable mesoscale structures. Therefore, it is important to run limited area models even if there is no significant – overall – improvement with respect to the global model providing the lateral boundary conditions.

In addition, one should not forget that it is a common phenomenon, that high resolution models might perform worse (on average, not for all individual cases) than the low resolution ones, when usual verification measures are applied. Although the increased resolution generally produces more realistic results, inevitable errors in timing and position can lead to larger RMSE values than for the smoother forecasts of the low resolution model. This is known as the *double penalty* problem. Therefore, the results presented above (i.e., no significant overall improvement by the limited area system) should be interpreted with care.

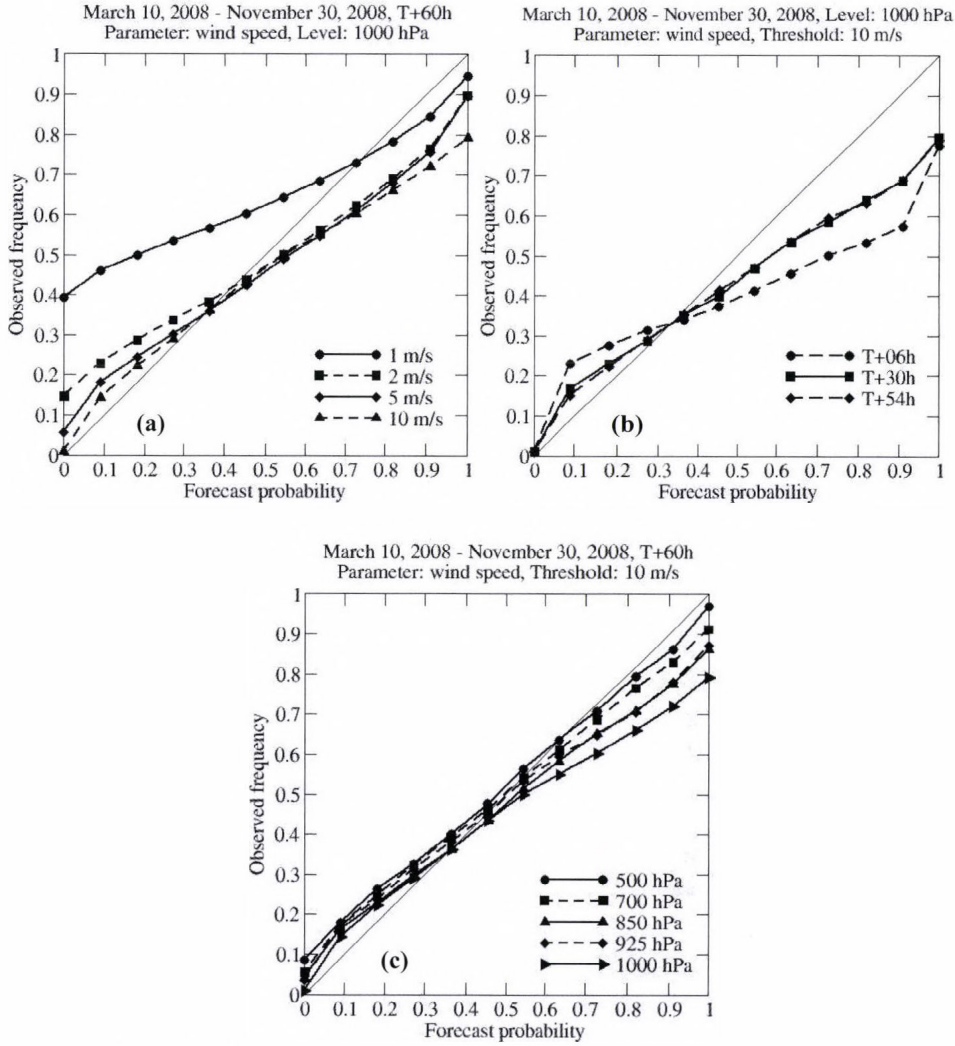


Fig. 6. Reliability diagrams for wind speed. (a) Diagrams for thresholds 1 m/s, 2 m/s, 5 m/s, and 10 m/s at 1000 hPa and T + 60 hours. (b) Diagrams for time range T + 06 hours, T + 30 hours, and T + 54 hours at 1000 hPa and 10 m/s threshold. (c) Diagrams for 500 hPa, 700 hPa, 850 hPa, 925 hPa, and 1000 hPa at T + 60 hours and 10 m/s threshold. The verification interval is March 10, 2008 – November 30, 2008. Verification was performed against ECMWF analysis.

4. Conclusions

In February 2008, a short-range limited area ensemble prediction system – based on the ALADIN model – was put into operations at HMS in order to gain experience not only from case studies and test periods, but on a day-to-day, real-

time basis. At present, the only operationally feasible solution is the direct downscaling of the PEARP members, therefore, this method is applied. Verification results of almost nine months (from March 10, 2008 to November 30, 2008) were analyzed using several verification measures for different parameters. As an overall conclusion it can be said, that the system behaved very similar in all three seasons (only the overall scores were shown). In spite of the problems with the spread-skill relationship (especially observed at lower levels and the surface), there are cases when the ensemble members show large spread, indicating the high uncertainty of the situation. In such cases the LAMEPS is definitely a useful complement of the operational deterministic ALADIN run. Results have shown that better scores are obtained for higher levels. The possible reasons of this behavior were described in Section 3. However, it is important to have skillful prediction of surface parameters as well. It would also be desirable to compute perturbations that are targeted for the area of our interest (Central Europe, particularly Hungary), in order to improve the spread-skill relationship and the quality of the forecasts as well. For these reasons it was decided to focus on the evaluation of local perturbations. Experiments have started to compute singular vectors with the ALADIN model (see *Hágel*, 2008). The aim of these experiments is to generate perturbations from the ALADIN singular vectors, and use them to perturb locally the initial conditions of the LAMEPS system. Perturbation of surface fields is also an important issue to be solved. A possible solution could be the method applied in the LAEF[10] system. Thus, the quasi-operational LAMEPS system of HMS is going to be developed and improved continuously, using the results of the ongoing researches.

Acknowledgments—The author wishes to thank to the team of Météo-France for making possible the operational use of the PEARP members as initial and lateral boundary conditions. The support and help of the members of the NWP group of the Hungarian Meteorological Service is also very much appreciated.

References

Buizza, R., Tribbia, J., Molteni, F. and *Palmer, T.*, 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus 45A*, 388-407.

García-Moya, J.A., Callado, A., Santos, C., Santos-Munoz, D. and *Simarro, J.*, 2007: El sistema de predicción por conjuntos para el corto plazo del INM. *AME Boletín 16*, 22-27.

Hágel, E., 2006: Report on work in Vienna on the common verification package (available at www.rclace.eu).

Hágel, E. and *Horányi, A.*, 2006: The development of a limited area ensemble prediction system at the Hungarian Meteorological Service: Sensitivity experiments using global singular vectors, preliminary results. *Időjárás 110*, 229-252.

Hágel, E. and *Horányi, A.*, 2007: The ARPEGE/ALADIN limited area ensemble prediction system: The impact of global targeted singular vectors. *Meteorol Z 16*, 653-663.

10 LAEF: Limited Area Ensemble Forecasting

Hágel, E., 2008: Singular vector experiments at the Hungarian Meteorological Service. *ALADIN Newsletter 33*, 52-59.

Horányi, A., Kertész, S., Kullmann, L., and *Radnóti, G.*, 2006: The ARPEGE/ALADIN mesoscale numerical modeling system and its application at the Hungarian Meteorological Service. *Időjárás 110*, 203-227.

Jensen, M.H., Frogner, I-L., Iversen, T. and *Vignes, O.*, 2006: Limited area ensemble forecasting in Norway using targeted EPS. *ECMWF Newsletter 107*, 23-29.

Jolliffe, I.T. and *Stephenson, D.B.* (eds.), 2003: *Forecast Verification: A Practitioner's Guide in Atmospheric Science.* Wiley.

Kann, A., 2007: Work on the Common Verification Package for the evaluation of ensemble forecasts (available at www.rclace.eu).

Mladek, R., 2006: Work on the Common Verification Package for the evaluation of ensemble forecasts (available at www.rclace.eu).

Montani, A., Capaldo, M., Cesari, D., Marsigli, C., Modigliani, U., Nerozzi, F., Paccagnella, T., Patruno, P. and *Tibaldi, S.*, 2003: Operational limited-area ensemble forecasts based on the 'Lokal Modell'. *ECMWF Newsletter 98*, 2-7.

Szintai, B. and *Ihász, I.*, 2006: The dynamical downscaling of ECMWF EPS products with the ALADIN mesoscale limited area model: Preliminary evaluation. *Időjárás 110*, 253-277.

Toth, Z., Kalnay, E., 1997: Ensemble Forecasting at NCEP and the Breeding Method. *Mon Weather Rev 125*, 3297-3319.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 1–2, January–June 2010, pp. 135–148

Fine scale simulation of turbulent flows in urban canopy layers

Miklós Balogh* and **Gergely Kristóf**

Department of Fluid Mechanics, Budapest University of Technology and Economics
Bertalan Lajos u. 4-6, H-1111 Budapest, Hungary
E-mails: baloghm@ara.bme.hu; kristof@ara.bme.hu

**Corresponding author*

(Manuscript received in final form October 22, 2009)

Abstract—The analysis of turbulent flows in urban areas is part of our research for simulating meso- and microscale atmospheric flows, since it is of great significance with respect to civil design and environmental protection. The application of general purpose computational fluid dynamics (CFD) solvers for the examination of smaller scale atmospheric phenomena, such as urban flows, has many advantages due to their flexible use on arbitrary meshes with complex geometries. A CFD solver has been adapted to atmospheric applications by using a purpose developed transformation method along with some volume sources active in the transport equations. In this way, the effects of the thermal stratification, adiabatic temperature variation, compressibility, and the Coriolis force were taken into account. Our present work is aimed at the development of a simulation methodology and software components for solving urban ventilation problems with special respect to the modeling of urban canopy layers. An advantage of the CFD based modeling technique is that the mesh size could change continuously in the computational domain according to the location of the building arrays, the road-system, and the examination area. This paper presents the modeling concept and its functionality in practice.

Key-words: CFD, urban canopy, turbulent flows, distributed drag force approach, two-equation model

1. Introduction

The detailed description of turbulence in the atmospheric boundary layer (ABL) is essential with regard to the dispersion and heat transfer processes, both within and above urban canopies. The realization of the accurate description is very difficult due to the complex structure of an urban canopy layer, which is a varied

system of bluff obstacles including trees and buildings. The examination of the urban climate, including the ventilation of the city, is possible by using statistical methods, such as roughness parameter mapping (*Gál* and *Unger*, 2009) or using dynamical methods. Though the detailed numerical simulation of flows in an urban canopy layer is possible by using CFD techniques, the higher numerical cost of the spatial discretization of the complex geometries makes this unrealizable in practice for very large domains. Finding the balance between the numerical cost and reasonable results is very important with regard to the civic design and environmental protection.

A potential technique, which keeps the balance, and furthermore, conforms to the variation of surface coverage, could be a hybrid method, such as a combination of the explicit and implicit description of the flow properties in an urban canopy. For simulating turbulent flows in the mostly exposed areas (e.g., some parts of the downtown area) where we need the most detailed results, the explicit modeling of the buildings could be used, and in the other parts of the examination area, the distributed drag force approach (*Green*, 1992; *Liu et al.*, 1996), as an implicit technique, could be applied with much lower numerical costs, although resulting in a lower resolution.

The efficient modeling of the flows in the atmospheric boundary layer, including turbulence, is feasible by solving the unsteady or steady state Reynolds-Averaged Navier-Stokes (URANS and RANS) equation due to the relatively low computational cost and reasonable accuracy. One of the most common turbulence model used in microscale investigations is the $k-\varepsilon$ two-equation model both for unsteady (URANS) and steady (RANS) simulations (*Hargreaves* and *Wright*, 2007). The main goal of this study is to develop an efficient hybrid method for simulating turbulent urban flows with CFD techniques.

2. *Application of a general purpose CFD solver for atmospheric simulations*

The use of a general purpose CFD solver adapted to atmospheric applications (*Kristóf et al.*, 2009) based on the realizable $k-\varepsilon$ model (*Shih et al.*, 1995) could be efficient for simulating urban ventilation and heat island problems in practice. The most common solvers are not capable of handling all of the physical processes in the arbitrarily stratified atmosphere, hence a mathematical transformation has been developed (*Kristóf et al.*, 2009). These are implemented in the ANSYS-FLUENT simulation system as a user defined function package, and validated with well documented laboratory experiments, analytical solutions, and widely accepted numerical results. This function package has been extended with functions for taking into account other important physical effects in connection with geophysical use, such as Coriolis force effects validated by analytical solutions. The results of the simulations of non-

hydrostatic phenomena, such as a descending cold air bubble, laboratory scale gravity waves, as well as urban heat island circulation agree well with references. The implemented physical processes, along with the distributed drag force approach, allow for the simulation of non-hydrostatic atmospheric flows in an urban environment. The present study focuses on modeling the impact of the obstacle arrays in the urban canopy layer, although, the examination of heat island phenomena is also in progress.

3. Distributed drag force approach

Recently, the application of CFD techniques is increasing in the field of micrometeorology, thus the knowledge of the flow structure and turbulence in an urban canopy layer is rapidly developing as well. Many scientists are working on the development of a distributed drag force approach, both for vegetated canopies (e.g., *Green*, 1992; *Liu et al.*, 1996) and for building arrays (*Lien* and *Yee*, 2004, 2005; *Lien et al.*, 2004, 2005; *Carissimo* and *MacDonald*, 2002). Therefore, it already has a solid theoretical background. This approach has already been used in practice for the urbanization of weather prediction models (*Hamdi* and *Masson*, 2008) to implicitly take into account the effect of buildings.

The essence of the drag force approach is an additional drag term in the momentum equation and two other terms in the transport equation of the turbulent kinetic energy (k) and the turbulent dissipation rate (ε). The drag term of the momentum equation is composed of the viscous and the form drags, while the value of the viscous component is much lower than the form component. Therefore, the former could be neglected. With this simplification, the source term of the momentum equation has a general form of

$$S_i = -\rho C_d A_i U u_i , \tag{1}$$

where ρ is the air density, C_d is the drag coefficient, A_i is the frontal area per unit volume normal to the ith direction, U is the velocity magnitude, and u_i is the velocity component in the ith direction. The unit of the momentum source is $\mathrm{N\,m^{-3}\,s^{-1}}$.

In the case of vegetated canopies, the obstacles (e.g., branches and leafs) convert the kinetic energy of the flow into wake turbulences with a smaller length scale than the shear-generated turbulence. Therefore, the canopy yields a net turbulent kinetic energy loss (*Green et al.*, 1995) instead of enhancing the wake production. This could be modeled with a source term in the following form,

$$S_k = \rho C_d A_f \left[\beta_p U^3 - \beta_d U k\right], \tag{2}$$

where A_f is the total frontal area per unit volume, k is the turbulent kinetic energy, β_p constant is a fraction of the mean flow kinetic energy produce k, and β_d is an empirical constant for short-circuiting the turbulent cascade (*Green*, 1992). The unit of the turbulent kinetic energy source is $\mathrm{kg\,m^{-3}\,s^{-1}}$.

The simplest model of the turbulent dissipation rate source term is based on the Kolgomorov's relation, which yields

$$S_\varepsilon = C_{\varepsilon 4} \frac{\varepsilon}{k} S_k \,. \tag{3}$$

The units for the turbulent dissipation rate source are $\mathrm{kg\,m^{-4}\,s^{-1}}$. In Eq. (3), ε is the turbulent dissipation rate and $C_{\varepsilon 4}$ is a constant. This relation was improved by *Liu et al.* (1996) providing a better fit to wind tunnel data. Accordingly, an alternative model could be defined as a more general form, which reads

$$S_\varepsilon = \rho C_d A_f \left[C_{\varepsilon 4} \beta_p \frac{\varepsilon}{k} U^3 - C_{\varepsilon 5} \beta_d U \varepsilon \right]. \tag{4}$$

In Eq. (4), the new constant $C_{\varepsilon 5}$ defines the mixing length anisotropy, if it is not equal to $C_{\varepsilon 4}$. Otherwise, the alternative model of Eq. (4) turns into the simpler Eq. (3) as noticed by *Sanz* (2003). It should be mentioned, that the source term of the turbulent dissipation rate is required in those microscale simulations that are based on two-equation turbulence models. In mesoscale models these terms can be neglected (*Otte et al.*, 2004). The coefficients in Eqs. (2)–(4) depend on the type of the turbulence model applied and the characteristics of the canopy layer. The relations between these constants and the constants of the k–ε model, together with the characteristics of the vegetated canopy were suggested by *Sanz* (2003) and were analyzed by *Katul et al.* (2004) and *Sanz* and *Katul* (2007).

In the present studies, the source terms of the momentum, turbulent kinetic energy, and turbulent dissipation rate were modeled in the above presented forms, applying the alternative model for describing S_ε. In vegetated areas, the properties of the drag terms were based on the works of *Balczó et al.* (2009), while in building arrays, the drag coefficient C_d was calculated as a function based on the volumetric porosity of the obstacle arrays (Appendix A) applied by *Coirier* and *Kim* (2006) in a similar application.

The goal of our method is to provide a mesh which follows the outlines of open areas, such as streets, squares, and open fields. The drag source terms should not be applied in these areas if the local mesh resolution is sufficiently high, thus explicit modeling is possible. Therefore, the impacts of these

cavities can be taken into consideration, although this strongly depends on mesh resolution.

4. An example on modeling turbulent flows in an urban canopy layer

In this section, we introduce a typical and practical application of CFD techniques in modeling turbulent atmospheric flows in urban areas. In this case study, the examination area is the 11th district of Budapest (the capital of Hungary), where a diversified landscape could be found, ranging from rural to downtown areas. Since this region is also surrounded by diversified regions, the examination area was extended with a relaxation zone for simplifying the setup of the lateral boundary conditions (Appendix B). In this way, the dimensions of the examination area in the x, y, and z directions were 9155, 7150, and 1800 m, respectively, while the computational domain extended by the relaxation zone (*Fig. 1*) is twice the original.

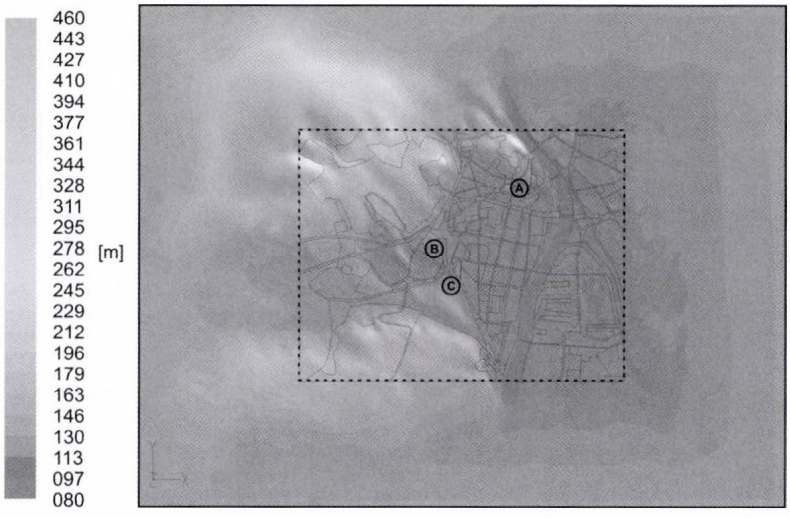


Fig. 1. Top view of the computational domain colored by the elevation. The similar surface coverage regions are highlighted (fine solid lines) in the examination area (thick dashed line), which is extended with the relaxation zone (thick solid line). The circles denote the sampling points of different canopy profiles used later.

The geometrical setup of the computational domain was based on the SRTM (Shuttle Radar Topography Mission) elevation database and a raster graphical map. It contains surface coverage and building cluster data in a simplified, type dependent form. The elevation data could be used after a coordinate transformation from WGS84 to a Cartesian frame of reference, namely the Uniform National Projection system (Hungarian abbreviation:

EOV). Nevertheless, the surface coverage data needed a pre-processing procedure for efficient use in modeling. As a result of these, the properties of the canopy layer such as the canopy layer height, the frontal area and solid fraction of the obstacles are available at each region of the computational domain. Since the domain contains several regions, we only introduce some typical parameters of urban type canopies (*Table 1*).

Table 1. Typical values of the total canopy height H, the solid fraction λ_f, and the frontal area per unit volume A_f in the urban type canopy, from different parts of the examination area

	Suburban (sample C)	**Block of flats (sample B)**	**Downtown (sample A)**
H [m]	7	20	15
λ_f [-]	0.3	0.4	0.6
A_f [$m^2 m^{-3}$]	0.55	0.63	0.77

Due to the finite volume method used in our simulation system, the spatial discretization of the domain, based on the computational mesh, is a system of wedge cells. This mesh was generated from the elevation data and the polygons are bounding the areas with the same type of surface coverage. The cells are triangular wedges with vertical orientation and their lateral edge length is varying between 8 and 160 meters in the examination area, growing up to 1000 meters in the relaxation zone, with a cell growth rate 1.3. For open fields, such as wider streets, parks, and the Danube River, the mesh size is the minimum possible. Therefore, the impact of these could be described at a higher resolution, as it can be seen in *Fig. 2*.

Vertically, the cells are ordered in layers and their height is increasing from 3 to 530 meters with the distance from the ground surface. The layers follow the terrain near the ground and become flat when approaching the upper boundary.

At the inlet boundaries the velocity components, turbulent kinetic energy, and dissipation rate were defined as pre-calculated profiles shown in *Fig. 3*. These were calculated by a one-dimensional, steady-state, realizable $k-\varepsilon$ model with the reference surface coverage conditions.

The wind climate in the district could be examined by averaging the results of the simulations executed for different boundary conditions. In the present example, eight runs were performed with different boundary conditions for the primary wind directions heaving the same velocity magnitude at 10 meters, which was 3 m s^{-1}. The simulations were steady state runs, while both the non-hydrostatic and Coriolis force effects were neglected. Therefore, only the source terms of the porous drag model and the non-reflective diffusion were enabled. The solver was used with second order upwind schemes for the spatial

discretization of the momentum equation, the turbulent kinetic energy and turbulent dissipation rate transport equations, and a second order scheme for the pressure equation. For better numerical stability, the SIMPLE pressure-velocity coupling was used, along with the node based Green-Gauss gradient scheme. With these settings, the simulations for the eight wind directions took only six hours using a Quad-Core computer.

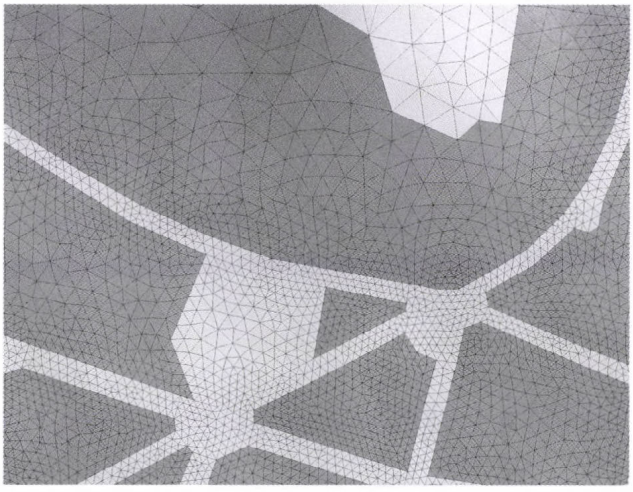


Fig. 2. The structure of the computational mesh of the downtown area of the 11th district, where the shaded areas show the building arrays, and the white ones denote uncovered areas such as streets, squares, and open fields, together with vegetated areas.

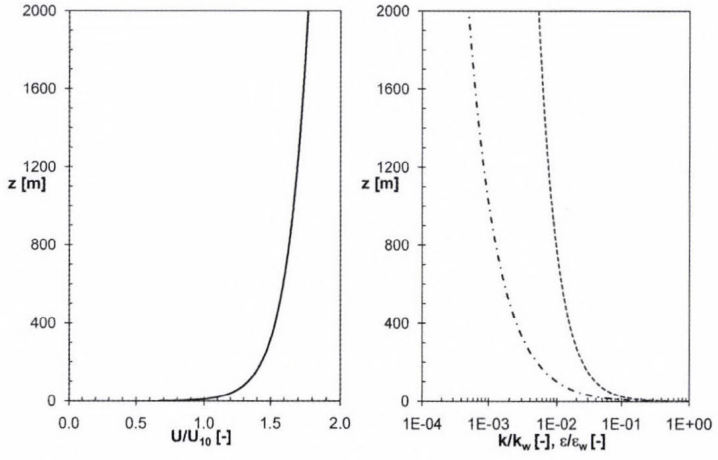


Fig. 3. Vertical profiles at the inlet boundaries, where the velocity magnitude (black) is scaled by the reference value at 10 meters above the ground. The turbulent kinetic energy (dashed) and dissipation rate (dashed-dotted) are normalized by the values next to the wall.

5. Results

Vertical profiles of the weighted average of the velocity magnitude and the turbulent shear stress, plotted along lines selected from different canopy regions from sampling points A, B, and C in *Fig. 1*, are verified using analytical canopy profiles. The analytical canopy profiles were published by *Finnigan* and *Belcher* (2004), calculating the velocity as

$$U(z)=\begin{cases}\dfrac{u_\tau}{\kappa}\ln\left(\dfrac{z-H+d}{z_0}\right), & \text{if } z>H \\ U_H e^{\beta(z-H)/l_m}, & \text{if } z\le H\end{cases}, \tag{5}$$

and the shear stress as

$$\tau(z)=l_m{}^2\left[\frac{dU}{dz}\right]^2, \tag{6}$$

where $U(z)$ and $\tau(z)$ are the velocity and the shear stress magnitudes, u_τ is the friction velocity, κ is the von Kármán constant, z is the height above the ground, H is the height of the canopy, U_H is the velocity magnitude at the top of the canopy, d is the displacement height, z_0 is the roughness height, l_m is the mixing length, and β is a constant of the profile. Note that u_τ, z_0, d, and l_m are functions of the canopy density. These fall within the range of the analytical results, for low and high canopy densities (*Fig. 4*), although some differences could be found in the shape of the profiles farther on the ground surface. The reason of these differences could be that the analytical profiles were calculated for flat surface, while our simulations were applied on complex terrain.

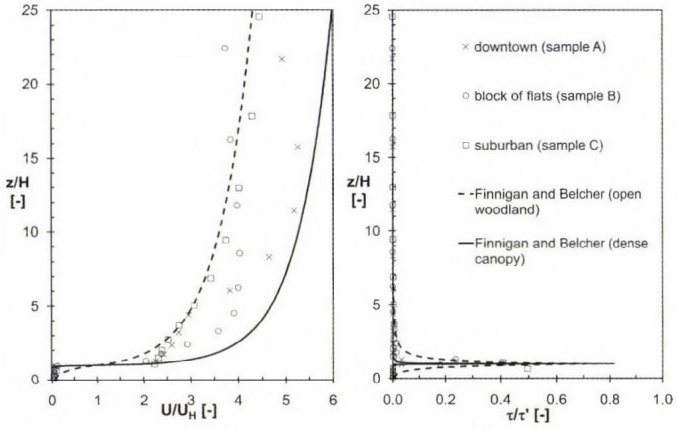


Fig. 4. Normalized velocity and shear stress profiles as a function of the height scale z/H, where U_H is the velocity at the top of the canopy and $\tau'=\rho u_\tau^2$. Sampling points A, B, and C can be seen in *Fig. 1*.

The averaged flow fields, which characterize the different regions of the district concerning the ventilation, are calculated from the results of the different cases weighted by the probability of the case dependent wind direction, both for the velocity magnitude and turbulent quantities. The average velocity magnitude as a quantitative parameter predicts the ventilation of the different regions, while the turbulent intensity contains useful information about the turbulent fluctuations. These were plotted at different heights above the ground, namely at 10 and 30 meters. The velocity and turbulent intensity fields are also scaled by the reference values of those calculated from the inlet profiles taking into consideration the local elevation. Since the inlet profiles are defined for an undisturbed free flow over a smooth surface, the scaled fields express the impact of the topography of the examination area and the topology of the canopy layer.

Fig. 5. Distribution of the velocity magnitudes scaled by reference values (inlet profiles) in the examination area, at 10 meters above the ground.

Near the surface, as shown in *Fig. 5*, over open areas, such as streets, squares, parks, and over the Danube River, the velocity magnitudes are significantly higher than the reference values due to the horizontal displacement of dense regions where the flow is moderated by obstacles. Moving away from the surface, the velocity is increasing as an effect of the canopy reducing blocking, although above the canopy layer height, the impact of the drag is still realizable (*Fig. 6*).

The scaled turbulence intensity has a local minimum near the surface (*Fig. 7*), since the reference turbulence intensity has the maximum next to the wall. The reference profiles were calculated with free flow conditions, thus, only

the wall has an impact on its turbulent properties. At 30 meters above the ground, shown in *Fig. 8*, the turbulence develops at the boundaries of the blocks, mainly where the canopy properties change suddenly. This effect is stronger on those side of the canopy blocks where the gradient of the porous drag has a high positive value in the streamwise direction of the locally dominant wind.

Fig. 6. Distribution of the velocity magnitudes scaled by reference values (inlet profiles) in the examination area, at 30 meters above the ground.

Fig. 7. Distribution of the turbulent intensity scaled by reference values (inlet profiles) in the examination area, at 10 meters above the ground.

Fig. 8. Distribution of the turbulent intensity scaled by reference values (inlet profiles) in the examination area, at 30 meters above the ground.

6. Conclusions and further developments

After the implementation of the distributed drag force parameterization, a practical application was executed, which demonstrates the capabilities of the CFD based approach in fields of the urban climatology and pollution control. The source term of the parameterization was applied within the areas where obstacles were found, thus, the impacts of the street canyons could also be considered. The properties of the canopy layer were also changed according to their type. The averaged results of the simulation were verified with analytical canopy profiles in representative points and good qualitative agreement has been found.

The results of the current study are useful for the further development in modeling stratified canopy layers, including the effects of the heat island phenomena and thermal convection. The realization of this requires the adaptation of the parameterization schemes modeling heat transfer and storage in the urban canopy layer (e.g., *Vu et al.*, 2002) with higher resolution.

Acknowledgements—This research was supported by the Hungarian Scientific Research Fund (Hungarian abbreviation: OTKA, grant No. T049573), and the wind climate examination supported by the Local Government of the 11th district of Budapest. The author would like to express special thanks for the data and information regarding the district, given free run by the Local Government.

Appendix A: Calculation of the drag coefficient

For the simplification of the drag coefficient C_d, it was calculated as a function of the volumetric porosity of the obstacle array applied by *Coirier* and *Kim* (2006) in a similar application. After adapting this function to our model, the drag coefficient could be defined in the following way

$$C_d = \begin{cases} \dfrac{\min(\Delta z, H - z_c + 0.5\Delta z)}{[1-\lambda_t]\Delta z}, & \text{if } z_c - 0.5\Delta z \le H \\ 0, & \text{if } z_c - 0.5\Delta z > H \end{cases}, \tag{A1}$$

where λ_t is the total solid volume per unit volume composed by the volume of the buildings and the vegetation, H is the canopy height, z_c is the height of the cell centroid above the ground, and Δz is the height of the cell.

Appendix B: Simplified specification of the boundary conditions

Both the elevation and surface coverage data were relaxed in the space to their reference values along the relaxation zone from the edge of the examination area to the lateral boundaries with the use of a dumping function Eq. (6). The reference value for the elevation was its spatial average at the lateral sides of the domain, while for the specification of the relaxed surface coverage parameters, the properties of the open grassland were used as a reference. For this reason, identical vertical profiles could be defined at every inlet boundary, calculated with the reference surface coverage properties.

$$\sigma(r) = \frac{1+\cos r}{2}, \quad r \in [0, \pi]. \tag{B1}$$

The dumping coefficient $\sigma(r)$ is a function of the normalized distance from the closest lateral boundary r. In the relaxation zone, the elevation and the characteristics of the canopy continuously approach a reference value defined on the boundary by Eq. (7).

$$\phi(r) = \sigma(r)\phi_{ref} + [1-\sigma(r)]\phi, \tag{B2}$$

where ϕ and ϕ_{ref} are the values of the relaxed parameter at the nearest part of the examination area and at the lateral boundary, respectively. The realization of the non-reflective boundary conditions could also be obtained by using

additional source terms (*Bodony*, 2006) in the relaxation zone as an analogy of Eq. (7), which is, written in a general form, is

$$S_r(r,\phi) = -\frac{\sigma(r)}{\Delta t}\left[\rho\phi - \rho_{ref}\phi_{ref}\right], \tag{B3}$$

where S_r is the non-reflective diffusion source term, Δt is the time-step size, ρ and ϕ are the current, ρ_{ref} and φ_{ref} are the reference values of the fluid density and the field variable of the transport equation, respectively. The field variable is the velocity in the source term of the momentum equation, the enthalpy in the energy equation, and φ equals 1 in the continuity equation. Note that in steady simulations, the value of the time-step size could be replaced by a time scale, while in incompressible cases ρ is equal to ρ_{ref}.

References

Balczó, M., Gromke, C., Ruck, B., 2009: Numerical modelling of flow and pollutant dispersion in street canyons with tree planning. *Meteorol Z 18*, 197-206.

Bodony, D.J., 2006: Analysis of sponge zones for computational fluid mechanics. *J Comput Phys 212*, 681–702.

Carissimo, B., Macdonald, R.W., 2002: A porosity/drag approach for the modeling of flow and dispersion in the urban canopy. In *Air Pollution Modeling and Its Application XV*. Springer, pp. 385-393.

Coirier, W.J., Kim, S., 2006: CFD modeling for urban area contaminant transport and dispersion: Model description and data requirements. *6th Symposium on the Urban Environment*. Atlanta, *GA. Amer Meteor Soc.*

Finnigan, J.J. and *Belcher, S.E.*, 2004: Flow over a hill covered plant canopy. *Q J Roy Meteor Soc130*, 1-29.

Gál, T., Unger, J., 2009: Detection of ventilation path using high-resolution roughness parameter mapping in a large urban area. *Build Environ 44*(1), 198-206

Green, S.R., 1992: Modelling turbulent air flow in a stand of widely-spaced trees. *PHOENICS Journal of Computational Fluid Dynamics and Its Applications 23*, 294–312.

Green, S.R., Grace, J., Hutchings, N.J., 1995: Observations of turbulent air flow in three stands of widely spaced Sitka spruce. *Agr Forest Meteorol 74*, 205-225.

Hamdi, R., Masson, V., 2008: Inclusion of a drag approach in the Town Energy Balance (TEB) Scheme: Offline 1-D evaluation in a street canyon. *J Appl Meteorol Clim 47*, 2627-2644.

Hargreaves, D.M., Wright, N.G., 2007: On the use of the $k-\varepsilon$ model in commercial CFD software to model the neutral atmospheric boundary layer. *J Wind Eng Ind Aerod 95*, 355–369.

Katul, G.G., Mahrt, L., Poggi, D., and *Sanz, C.*, 2004: One and two equation models for canopy turbulence. *Bound-Lay Meteorol 113*, 81–109.

Kristóf, G., Rácz, N., Balogh, M., 2009: Adaptation of pressure based CFD solvers for mesoscale atmospheric problems. *Bound-Lay Meteorol 131*, 85-103.

Otte, T.L., Lacser, A., Dupont, S., and *Ching, J.K.S.*, 2004: Implementation of an urban canopy parameterization in a mesoscale meteorological model. *J Appl Meteorol 43*, 1648-1665.

Lien, F-S., Yee, E., 2004: Numerical modelling of the turbulent flow developing within and over a 3-d building array, part I: A high-resolution Reynolds-averaged Navier-Stokes approach. *Bound-Lay Meteorol 112*, 427–466.

Lien, F-S., Yee, E., 2005: Numerical modelling of the turbulent flow developing within and over a 3-d building array, part III: A distributed drag force approach, its implementation and application. *Bound-Lay Meteorol 114,* 287–313.

Lien, F-S., Yee, E., Cheng, Y., 2004: Simulation of mean flow and turbulence over a 2d building array using high-resolution CFD and a distributed drag force approach. *J Wind Eng Ind Aerod 92,* 117-158.

Lien, F-S., Yee, E., Wilson, J.D., 2005: Numerical modelling of the turbulent flow developing within and over a 3-d building array, part II: A mathematical foundation for a distributed drag force approach. *Bound-Lay Meteorol 114,* 245–285.

Liu, Z., Chen, J.M., Black, T.A., and *Novak, M.D.*, 1996: $E-\varepsilon$ modelling of turbulent air flow downwind of a model forest edge. *Bound-Lay Meteorol 77*, 21–44.

Sanz, C., 2003: A note on $k-\varepsilon$ modelling of vegetation canopy air-flows. *Bound-Lay Meteorol 108*, 191–197.

Sanz, C., Katul, G.G., 2007: Dual length scale two-equation modelling of the canopy turbulent kinetic energy wake budget. *C. R. Mecanique 335,* 685–690.

Shih, T.-H., Liou, W.W., Shabbir, A., Yang, Z., and *Zhu, J.*, 1995: A new $k-\varepsilon$ Eddy-viscosity model for high Reynolds number turbulent flows – Model development and validation. *Comput Fluids 24*, 227-238.

Vu, T.C., Ashie, Y., and *Asaeda, T.*, 2002: A $k-\varepsilon$ turbulence closure model for the atmospheric boundary layer including urban canopy. *Bound-Lay Meteorol 102*, 459–490.

In memoriam: Dezső Dévényi (1948–2009)

By now, we all know the stubborn facts: on November 26 *Dezső Dévényi* suddenly passed away of a heart attack in Boulder, Colorado. Maybe we are over the first shock, but it is still impossible to understand and comprehend that Dezső is not among us any more, and surely we are unable to account how much Hungarian and international meteorology lost with his death.

Dezső was born in Keszthely, attended the Eötvös Loránd University in Budapest, and he received his MSc degree in meteorology and teacher of mathematics/physics in 1973. During his entire career his scientific interest was devoted to the mathematical problems of meteorology. His basic interest towards atmospheric data assimilation (the statistical aspects at that time) started in the mid-70s, when he had several visits to the Soviet Union (in Leningrad, now St. Petersburg), and he had the opportunity to meet and work with *Lev Gandin*, the father of atmospheric optimal interpolation. This interaction probably determined his later scientific interest within meteorology, which was numerical weather prediction and mainly data assimilation. Later he co-authored (with *Ottó Gulyás*) a book entitled "Mathematical Statistical Methods in Meteorology". This is still an important reference in statistical training for Hungarian students reading

meteorology. In the second part of the 1980s he played a major role in the establishment of operational numerical weather prediction in Hungary by initiating cooperation with the Swedish Meteorological and Hydrological Institute (SMHI) in order to adapt their limited area model. Besides putting the SMHI model into operations, this adaptation work also led to the establishment of a small NWP team (consisting of 3–4 scientists at that time), which was the nucleus of the recent (much larger) NWP team of the Hungarian Meteorological Service. He was one of the early pioneers who visited Météo France in March 1991, to discuss and assess the feasibility of the LAM-ARPEGE project, which later became known as ALADIN. (And now it is an essential numerical weather prediction project not only in Central Europe, but also including the HIRLAM countries, who are now working on code cooperation with the ALADIN group.) He received his scientific degree at the Hungarian Academy of Sciences (which was later recognized as a PhD at the Eötvös Loránd University) in 1991 with a thesis entitled "The Application of Satellite Data in the Objective Analysis of Meteorological Fields". Dezső joined NOAA's Forecast Systems Laboratory in Boulder, Colorado for the first time between 1991 and 1993. After a short break for a return to Hungary, he continued his work in the US from 1995 to 1999 and, after another "Hungarian" break, from 2000 onwards. His work in the US was closely related to data assimilation with the development of the Rapid Update Cycle (RUC) and later Rapid Refresh (RR). During his years in Hungary he had several managerial positions (the highest being vice president) at the Hungarian Meteorological Service, but his main interest focused on the scientific aspects (rather than the administrational ones) of meteorology. In 1996, he was awarded the Doctor Habilitationis Degree by the Eötvös Loránd University, where he became an Associate Professor between 1999 and 2004.

Besides his official positions at the Hungarian Meteorological Service, Dezső was active in the social life of meteorology as well. He was a member of the Hungarian Meteorological Society, the Mathematical Society of Hungary, and lately the American Meteorological Society, too. He attained membership and became elected chairman of various Working Committees of the Hungarian Academy of Sciences (for Observations and Data Assimilation, Atmospheric Dynamics, Climate). He was member of the Editorial Board of the IDŐJÁRÁS (Quarterly Journal of the Hungarian Meteorological Service), where he made important contributions both as author and reviewer of numerous scientific papers.

Dezső was also famous for his passion of reading professional (during the early mornings) and recreational literature (in the evenings). He had a very extensive library on various topics of mathematics and physics, but he was also a fan of science fiction literature. He had a keen interest in sports, especially football, and avidly followed European sport events, football championships, and major international tournaments as much as time permitted. He very much regretted that in the United States opportunities to watch good football matches

were limited. This was just another reason, why, in the last years, he planned his final return to Hungary.

Dezső was a well-known and recognized person in the Hungarian meteorological society in spite of the fact that for the last 15 years he had been mostly living in the United States. His special, sometimes ironic and sarcastic sense of humor was highly appreciated. He kept in contact with many of us, and when he came home for a short visit, he never forgot to meet with old colleagues and friends and regularly presented his scientific achievements at the Hungarian Meteorological Society. Dezső was a great teacher of numerical weather prediction for an entire generation of Hungarian meteorologists. Many of us now working in numerical weather prediction or other related fields recognize him as our mentor. He was much more than that for some of us, for those having everyday contacts even when he was far away in space, but not in thoughts. It just means that many of us miss him tremendously, but we will always remember him as the "father" of numerical weather prediction in Hungary, a colleague who never forgot his roots in Hungary, and a friend to whom we could always address questions, being sure to get the most appropriate answers almost immediately.

Although this remembrance is a farewell to Dezső on the pages of IDŐJÁRÁS, we will never forget him, and we will follow the tracks he laid down for numerical weather prediction.

András Horányi and *Gábor Radnóti*

INSTRUCTIONS TO AUTHORS OF *IDŐJÁRÁS*

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

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

Some issues contain "News" and "Book review", therefore, such contributions are also welcome. The papers must be in American English and should be checked by a native speaker if necessary.

Authors are requested to send their manuscripts to

Editor-in Chief of IDŐJÁRÁS
P.O. Box 39, H-1675 Budapest, Hungary
E-mail: antal.e@met.hu

including all illustrations. MS Word format is preferred in electronic submission. Papers will then be reviewed normally by two independent referees, who remain unidentified for the author(s). The Editor-in-Chief will inform the author(s) whether or not the paper is acceptable for publication, and what modifications, if any, are necessary.

Please, follow the order given below when typing manuscripts.

Title page: should consist of the title, the name(s) of the author(s), their affiliation(s) including full postal and e-mail address(es). In case of more than one author, the corresponding author must be identified.

Abstract: should contain the purpose, the applied data and methods as well as the basic conclusion(s) of the paper.

Key-words: must be included (from 5 to 10) to help to classify the topic.

Text: has to be typed in single spacing on an A4 size paper using 14 pt Times New Roman font if possible. Use of S.I. units are expected, and the use of negative exponent is preferred to fractional sign. Mathematical formulae are expected to be as simple as possible and numbered in parentheses at the right margin.

All publications cited in the text should be presented in the *list of references*, arranged in alphabetical order. For an article: 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 photo-synthetically active radiation and cloud amount. *Időjárás 90,* 10-13. For a book: 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. Reference in the text should contain the name(s) of the author(s) in Italics and year of publication. E.g., in the case of one author: *Miller* (1989); in the case of two authors: *Gamov* and *Cleveland* (1973); and if there are more than two authors: *Smith et al.* (1990). If the name of the author cannot be fitted into the text: *(Miller,* 1989); etc. When referring papers published in the same year by the same author, letters a, b, c, etc. should follow the year of publication.

Tables should be marked by Arabic numbers and printed in separate sheets with their numbers and legends given below them. Avoid too lengthy or complicated tables, or tables duplicating results given in other form in the manuscript (e.g., graphs).

Figures should also be marked with Arabic numbers and printed in black and white or color (under special arrangement) in separate sheets with their numbers and captions given below them. JPG, TIF, GIF, BMP or PNG formats should be used for electronic artwork submission.

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 for authors is available: antal.e@met.hu

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/Journal-Idojaras.php

VOL. 114 * NO. 3 * JULY – SEPTEMBER 2010

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: Gilice tér 39, H-1182 Budapest, Hungary
P.O. Box 39, H-1675 Budapest, Hungary
E-mail: bozo.l@met.hu or antal.e@met.hu
Fax: (36-1) 346-4809

Indexed and abstracted in Science Citation Index Expanded™ and Journal Citation Reports/Science Edition
Covered in the abstract and citation database SCOPUS®

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary
E-mail: kenderesy.k@met.hu or antal.e@met.hu

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 3, July–September 2010, pp. 153–168

Neural networks for precipitation forecasting in Bulgaria

Ivan V. Tsonevsky* and **Valery G. Spiridonov**

National Institute of Meteorology and Hydrology
66 Tsarigradsko Chaussee, Sofia 1784, Bulgaria; E-mail: Ivan.Tzonevski@meteo.bg

**Corresponding author*

(Manuscript received in final form October 15, 2009)

Abstract—The application of neural networks for forecasting precipitation in Bulgaria is discussed in this note. The problem is to determine whether a given forecasted precipitation event would exceed the operational criteria for a heavy-snow warning in the cool season (October–March) or a heavy-rain warning in the warm season (April–September) (10 mm and 15 mm in 12 hours, respectively). The inputs to two sets of four multilayer perceptrons (MLPs) were selected from among 27 variables extracted from the global numerical weather prediction model from the UK Met Office. These variables (e.g., 850 hPa and 500 hPa temperatures and wind) were picked, because they are more predictable than the amount of precipitation. Bulgaria was divided into six geographic regions, and all MLPs were trained independently for each region. The training dataset consisted of all 12-hour periods with precipitation of the years 1998–2004. The MLPs performed better during the cool season; the mean of the correct classifications across all regions was 75% in the cool season and 65% in the warm season. Sensitivity analysis enabled a determination of the relative importance of each of the 27 variables to the performance of the neural network. The most important variables were temperature and wind fields at both 850 and 500 hPa levels. A combined system of MLPs ensemble and operational limited area numerical weather prediction model ALADIN is presented that shows better skill in forecasting heavy precipitation events.

Key-words: artificial neural networks, weather forecasting, heavy precipitation, NWP models, weather warnings

1. Introduction

Numerical weather prediction (NWP) models have improved considerably during the last few decades. For example, the European Centre for Medium-Range Weather Forecasts model showed gains of one day in predictability of mean sea level pressure and 500 hPa height over the last decade of the twentieth century in the Northern Hemisphere (*Simmons* and *Hollingsworth*, 2002).

Precipitation forecasts showed the same improvement over Europe (*Ghelli*, 2006). Because precipitation has an inherently small spatial scale and low predictability (*Rodwell*, 2006), precipitation events exceeding 60 mm/day are poorly represented by NWP models (e.g., *Lalaurette* and *van der Grijn*, 2006). Many statistical methods can be used to improve the raw output of NWP models such as model output statistics (MOS; *Glahn* and *Lowry*, 1972) and downscaling (*Wilby* and *Wigley*, 1997).

Another way to improve precipitation forecasts is using artificial neural networks (ANN). Neural networks have been applied in meteorology for classification and prediction by many authors (e.g., *Banket*, 1994; *Xiao* and *Chandrasekar*, 1997; *Hall et al.*, 1999; *Marzban*, 2000; *Chevallier et al.*, 2000; *Zhanqing et al.*, 2001; *Roebber et al.*, 2003).

The purpose of our investigation is to develop a tool for forecasting of heavy precipitation events in Bulgaria over six different geographically distinct regions. NWP models explicitly predict the variables included in the model equations, more accurately than the precipitation. On the other hand, experienced forecasters can draw upon pattern recognition to skilfully predict dry and wet situations. We used some of the most predictable variables to train a number of neural networks specifying the regions with heavy precipitation.

The main objective of this investigation is to find an appropriate modern tool, such as neural networks, for applying the classical synoptic rules for improving NWP forecast of heavy precipitation. Hence, our goal is to give forecasters additional information to the available operational NWP predictions (deterministic and probabilistic) in order to increase their confidence treating heavy precipitation situations.

In Section 2, the datasets used to train the neural networks are described. In Section 3, the structure of two sets of four multilayer perceptrons (MLPs) are presented and tested on precipitation data of the years 1998–2004. In Section 4, ALADIN NWP model predictions and those of a combined system (ALADIN+ANN ensemble) are compared using a dataset of precipitation events from 2005, and Section 5 concludes this paper.

2. Data

Neural networks have to learn the input–output relationship through a training process. In supervised learning, the training dataset contains examples of inputs with corresponding outputs. We chose the input variables from archived analysis of the Met Office's global NWP model in GRID code with spatial resolution of 2.5° latitude × 2.5° longitude at three levels: surface, 850 hPa, and 500 hPa. The following 27 variables served as inputs to the neural networks: latitude and longitude of the centers of the closest cyclone and anticyclone at each level to Bulgaria (12 variables in total); mean sea level pressure and geopotential heights at 850 and 500 hPa of these centers (6 variables); wind speed and wind direction

at 850 and 500 hPa at the model grid point closest to the geographic region in Bulgaria (4 variables); temperatures at 850 and 500 hPa, as well as the difference between these temperatures ($\Delta T_1 = T_{850}(\varphi_r,\lambda_r) - T_{500}(\varphi_r,\lambda_r)$) at the same grid point with geographical coordinates (φ_r,λ_r); and two horizontal temperature differences at 850 hPa over Bulgaria: $\Delta T_2 = T_{850}(\varphi_1,\lambda_1) - T_{850}(\varphi_2,\lambda_2)$ and $\Delta T_3 = T_{850}(\varphi_3,\lambda_3) - T_{850}(\varphi_4,\lambda_4)$, where φ_1= 45.0° N, λ_1= 30.0° E, φ_2= 42.5° N, λ_2= 22.5° E, φ_3= 45.0° N, λ_3= 20,0° E, φ_4= 42.5° N, λ_4= 27.5° E are the geographical coordinates of the grid points. The horizontal temperature differences at 850 hPa are a measure of the presence of frontal zones and represent the dominant directions of the movement of the fronts over Bulgaria (ΔT_2 and ΔT_3 represent the northeast–southwest and northwest–southeast directions, respectively). Our choices of input variables were based on the following two criteria: their availability in the archive and their predictability. The selected variables are much more predictable than the precipitation, thus serving to represent the large-scale pattern. We tried to use the most predictable, dynamic variables which allow us to extend the neural network approach beyond the first 12 hours even in the medium range keeping high confidence of the forecasts. Here, the neural network technique could be considered as the development of the classical synoptic method, which uses the evolution of the weather systems (such as cyclones and anticyclones) to predict the precipitation over some region not including explicitly the humidity. The reason is that the humidity in NWP models depends on the fluxes from the surface and the evaporation of precipitation. These processes are parameterized and hence the humidity is less predictable than the temperature, geopotential, or mean sea-level pressure, which are product of the physical equations in the dynamical component of the model. On the other hand, since the humidity or precipitable water and the precipitation amounts are very closely connected, we may expect that adding the humidity as an input would dominate the other more predictable variables such as geopotential and temperature. Therefore, neural networks are not used here for statistical post-processing of NWP model's output but are used as a model that include classical synoptic rules which, in combination with NWP model, gives better results.

The difference ΔT_1 between temperatures at 850 and 500 hPa is a measure of atmospheric stability, included in all neural networks in the warm season when convection is a dominant precipitation mechanism. In future investigations, the influence of other measures of stability, such as convective available potential energy (CAPE) and lifted index, could be examined.

The centers of cyclones and anticyclones are determined automatically and if there is not any close circulation system, the minimum and, respectively, the maximum value over the boundaries of the European region and their geographical coordinates are taken as inputs. We calculate the distance between the closest circulation centers and a point situated at the center of Bulgaria.

The output variable is the precipitation class, indicating whether the precipitation amount for the 12-hour period exceeded a given seasonal threshold. The precipitation class is a nominal variable that has two values: '0' if the precipitation amount is below the threshold and '1' if the precipitation amount is above this threshold. Since the dominant precipitation mechanism in the cool season is generally different than that in summer, we created two training sets with two different thresholds. In the warm season (April–September), when convective precipitation dominates, we chose the threshold to be 15 mm/12 hours, consistently with the operational criterion for heavy-rain warning in Bulgaria. Similarly, in the cool season (October–March), stratiform precipitation dominates and usually falls as snow. Therefore, we chose the threshold to be 10 mm/12 hours, or about 10 cm snow/12 hours, which is the operational criterion for heavy-snow warning in Bulgaria.

The design of the neural network suggests that a relationship exists between the inputs chosen above and the output precipitation class. In practice, however, we do not know the exact relationship between the inputs and output. The training datasets contain all the cases of the years 1998–2004 when precipitation was measured; hence they are representative and balanced datasets. This means that the proportion of the different classes in the dataset is the same as it was measured, which is of critical importance to the training. Precipitation amounts are taken from the daily standard observations from the network of 34 synoptic stations over Bulgaria (*Fig. 1*). In order to obtain the nominal output (i.e., the precipitation class), the maximum precipitation amount for each region is selected.

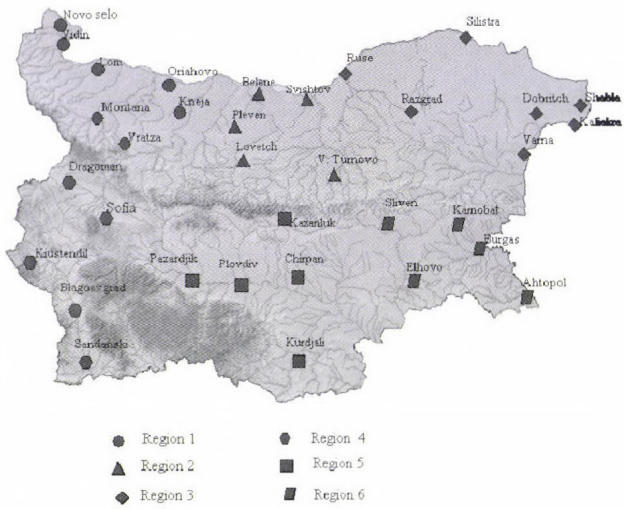


Fig. 1. Distribution of the synoptic stations over the territory of Bulgaria used for dataset construction (different marks of the stations represent their belonging to different regions).

3. Construction and training of the neural network sets

Classification can be performed using a wide range of neural network types. We used the multilayer perceptron (MLP), because this is perhaps the most popular network architecture in use today, providing wide applicability and ability to model very complex nonlinear functions. Moreover, MLP can be trained by various training algorithms such as the back propagation (e.g., *Werbos*, 1974; *Parker*, 1985; *Rumelhart et al.*, 1986), conjugate gradient descent (e.g., *Bishop*, 1995; *Shepherd*, 1997), and Levenberg–Marquardt algorithms (*Levenberg*, 1944; *Marquardt*, 1963). We used the first two algorithms in our training procedure. MLP has a simple interpretation as an input-output model with one or more hidden layers, each of them consists of a number of hidden units. A network with a single hidden layer is in practice capable of modeling most real classification problems.

The first step in constructing the MLPs was to divide Bulgaria into an appropriate number of regions. Taking into account the orography and some geographical features, we divided Bulgaria into six regions (*Fig. 1*):

Region 1 = northwestern part of Bulgaria,
Region 2 = central part of northern Bulgaria,
Region 3 = northeastern part of Bulgaria,
Region 4 = southwestern part of Bulgaria containing Sofia, the capital of the country,
Region 5 = central part of southern Bulgaria,
Region 6 = southeastern part of Bulgaria.

To design MLPs, we used STATISTICA Neural Networks, a software package for neural-network applications. Using the options in STATISTICA Neural Networks, we randomly divided each dataset into three sections necessary for the training process: a training set, a verification set, and a test set. The verification set is used to track the network's error performance, to identify the best network, and to stop training if over-learning occurs. The test set is not used in training at all, and is designed to give an independent assessment of the network's performance when an entire network design procedure is completed. We selected initial configurations of MLPs by choosing one hidden layer, because experiments with two and more hidden layers did not lead to better performance of the networks. Conducting sensitivity analysis, we selected three configurations with different input variables for each season and region. Sensitivity analysis, discussed in more detail in Section 3.3, rates the importance of variables. Then, we conducted a large number of experiments with different configurations with a different number of hidden units to obtain the network with the lowest verification error. Determining the hidden units is important, but not easy. If the number of the hidden units is too large, overfitting may occur, and the verification error rises while the training error decreases. In this context,

we performed a huge number of experiments to determine the number of hidden units starting with a half of the number of input units. Varying the number of hidden units helped to choose the neural network with the best performance. Following the training process, each network was run on the independent test dataset that is not used in the training, giving an independent assessment of the networks' performance. The activation function of the hidden and output layers is the logistic function. The latter guarantees an output lying in the range (0, 1). The target output '1' indicates membership in one class and '0' represents membership in the other. The error function used in the training is very important (*Bishop*, 1995). We used sum-squared error function appropriate for classification.

The selected configurations were trained many times to avoid local minima. The configuration with all 27 input variables was also trained. This training process was repeated for each region independently. Finally, for each region, the result was two neural network sets (cool and warm season) consisting of four MLPs. These networks are combined into an ensemble.

3.1. Neural network set for the cool season

(a) MLP (7, 8)

The first MLP from the cool-season network set consists of 7 input variables and 8 hidden units: MLP (7, 8). The input variables are: latitude and longitude of the centers of the closest cyclone to Bulgaria at 850 hPa; wind speed and wind direction at 500 hPa at the model grid point closest to the examined region; temperature at 850 hPa and vertical temperature difference (ΔT_1) at the same grid point; temperature difference in the northwest-southeast direction (ΔT_2).

(b) MLP (8, 11)

The second MLP has 8 inputs and 11 hidden units. The input variables are: latitude and longitude of the centers of the closest cyclone to Bulgaria at 850 hPa; wind speed and wind direction at 500 hPa at the model grid point closest to the examined region; temperature differences at 850 hPa in the northeast-southwest and northwest-southeast directions (ΔT_2 and ΔT_3, respectively); wind speed and wind direction at 850 hPa at the model grid point closest to the examined region.

(c) MLP (5, 16)

The third MLP of the ensemble has 5 inputs and 16 hidden units. The input variables are: wind speed and wind direction at 500 hPa in the model grid point closest to the examined region; vertical temperature gradient ΔT_1 in the same grid point; temperature differences ΔT_2 and ΔT_3 at 850 hPa in northeast-southwest and northwest-southeast direction, respectively.

(d) MLP (27, 14)

The last MLP from the ensemble contains all 27 variables described in Section 2 and has 14 hidden units.

The performance is a measure of the success of the network and gives the percentage of cases correctly classified. The performance of the neural networks during the cool season varies from region to region. The mean for all regions is 70%.

3.2. Neural network set for the warm season

There is also a set of four MLPs for the warm season.

(a) MLP (4, 14)

The first neural network is an MLP with 4 inputs and 14 hidden units. The input variables are: wind direction at 500 hPa in the model grid point closest to the examined region; vertical temperature gradient ΔT_1 in the same grid point; temperature differences ΔT_2 and ΔT_3 at 850 hPa in northeast-southwest and northwest-southeast direction, respectively.

(b) MLP (6, 8)

The second MLP consists of 6 inputs and 8 hidden units. Here, input variables are: latitude and longitude of the centers of the closest cyclone to Bulgaria at 850 hPa; wind speed at 500 hPa in the model grid point closest to the examined region; vertical temperature gradient ΔT_1 in the same grid point; temperature differences ΔT_2 and ΔT_3 at 850 hPa in northeast-southwest and northwest-southeast direction, respectively.

(c) MLP (5, 7)

The third MLP has 5 inputs and 7 hidden units. Variables here are: latitude and longitude of the centers of the closest cyclone to Bulgaria at 500 hPa; wind speed and wind direction at 500 hPa in the model grid point closest to the examined region; vertical temperature gradient ΔT_1 in the same grid point.

(d) MLP (27, 14)

MLP with all 27 variables and 14 hidden units was trained as well.

The performance of the neural networks for the cool season (*Table 1*) is higher than that of the warm season (*Table 2*), varying from 60 to 70%, with a few exceptions below and above but not far from this range. The mean for all regions is 64%. The lower performance during the warm season is consistent

with previous neural network studies of precipitation (e.g., *Hall et al.*, 1999) and other forecast systems (e.g., *Fritsch* and *Carbone*, 2004).

Table 1. Performance (%) of the MLPs for all regions for the cool season of year 2005 (from January to March and from October to December)

Regions	Performance (%)			
	MLP (7, 8)	MLP (8, 11)	MLP (5, 16)	MLP (27, 14)
1	75	69	75	74
2	76	63	78	72
3	68	60	64	69
4	74	69	76	80
5	75	69	74	73
6	66	54	65	64

Table 2. Performance (%) of the MLPs for all regions for the warm season of year 2005 (from April to September)

Regions	Performance (%)			
	MLP (4, 14)	MLP (6, 8)	MLP (5, 7)	MLP (27, 14)
1	69	62	69	62
2	63	58	63	63
3	62	58	55	55
4	70	67	77	65
5	70	68	64	58
6	65	60	68	59

3.3. Sensitivity analysis and skill measures

An interesting problem is to find the input variables, which are considered the most important by a particular neural network. The STATISTICA Neural Networks software provides an opportunity for us to address this problem by conducting sensitivity analysis. Sensitivity analysis can give important insights into the importance of individual variables, although it must be employed with some care, because the sensitivity analysis does not rate the usefulness of variables in an absolute sense. Consider, for example, the case where two input variables are not independent. Then, if either is eliminated the model may compensate, because the other still provides the key information. In this case, the sensitivity analysis may rate these variables as of low sensitivity, even

though they might encode key information. Similarly, a variable that encodes relatively unimportant information may have higher sensitivity than any other variables that encode more important information. If variable always has high sensitivity in various models, we can try to find some physical reason why this variable is so important for the model and what the nature of the dependence is between the input and output variables. Because of the complex relationships between the large-scale variables from NWP models and small-scale precipitation amounts, caution must be employed with sensitivity analysis.

Using sensitivity analysis, we selected a number of input variables for the different MLPs. For the cool season, 500 hPa wind speed and direction are presented in all MLPs. For the warm season, variables that are selected in all MLPs are the temperature difference between 850 and 500 hPa (ΔT_1) and the wind direction at 500 hPa. The vertical temperature difference (ΔT_1) is a measure of the static stability and the ingredients of the heavy precipitation (e.g., *Doswell et al.*, 1996).

Sensitivity analysis has been done on MLP (27, 14) as well, for the cool and warm seasons. The most important of all the 27 variables for classifying the precipitation cases obtained by averaging over all the regions are shown in *Table 3*. There is some difference for the cool and warm season, but temperature and wind variables seem to be the most important for the classification.

Table 3. Most important variables for the cool season and warm season

Cool season
Temperature difference at 850 hPa over Bulgaria in northwest-southeast direction
Temperature at 500 hPa in the model grid point closest to the geographic region
Wind direction at 850 hPa in the model grid point closest to the geographic region
Temperature at 850 hPa in the model grid point closest to the geographic region
Wind direction at 500 hPa in the model grid point closest to the geographic region
Longitude of the anticyclone closest to Bulgaria at 500 hPa
Wind speed at 850 hPa in the model grid point closest to the geographic region
Temperature difference at 850 hPa over Bulgaria in northeast-southwest direction
Temperature difference between 850 and 500 hPa in the model grid point closest to the geographic region
Latitude of the anticyclone closest to Bulgaria at sea level
Warm season
Wind direction at 500 hPa in the model grid point closest to the geographic region
Temperature difference at 850 hPa over Bulgaria in northwest-southeast direction
Temperature difference at 850 hPa over Bulgaria in northeast-southwest direction
Latitude of the closest to Bulgaria cyclone at 500 hPa
Mean sea level pressure at the center of the anticyclone closest to Bulgaria
Longitude of the cyclone closest to Bulgaria at 500 hPa
Longitude of the anticyclone closest to Bulgaria at sea level
Latitude of the anticyclone closest to Bulgaria at 500 hPa
Longitude of the anticyclone closest to Bulgaria at 500 hPa
Longitude of the anticyclone closest to Bulgaria at sea level

Hall et al. (1999) found, that the most significant variables in their neural network for precipitation forecasting for Dallas-Fort Worth, Texas, were the precipitable water and a number of variables closely related to the temperature and wind fields. Thus, our results are similar to that obtained in that study, taking into account the absence of the humidity as an input in our neural networks.

A receiver operating characteristic (ROC) curve (e.g., *Zweig* and Campbell, 1993; *Masters*, 1993) measures the performance of a two-class classifier. A ROC curve plots the hit rate versus the false alarm ratio. A random classifier, with no ability to discriminate between two classes, has a diagonal line ROC curve from (0,0) to (1,1). The classifiers with curves above the diagonal are skilled. The area under the curve is a scalar measure of the classifier's performance. The random classifier has an area of 0.5 under the ROC curve, whereas ideal classifiers have an area of 1 under the ROC curve.

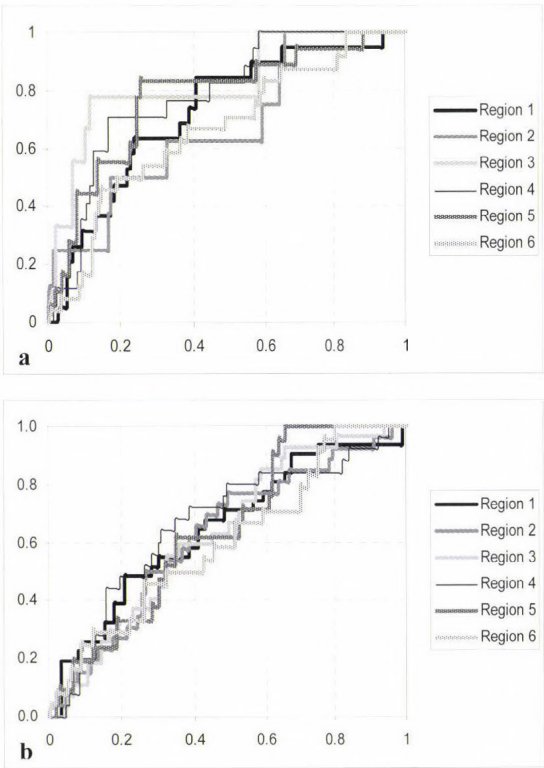


Fig. 2. ROC curves of the MLP (27, 14) for the cool (a) and warm (b) season of 2005.

ROC curves of the MLP (27, 14) for all regions in the cool and warm seasons (*Fig. 2*) show that the performance of the network in the cool season is much better than in the warm. For the cool season, ROC curves are closer to the

perfect classifier and the areas under the curves are larger than in the warm season. The performance of MLP (27, 14) is not the same for all regions. Areas under the ROC curves vary from region to region and from season to season (*Table 4*). The area averaged over all the regions for the cool season is 0.74 and for the warm season it is 0.64.

Table 4. Areas under the ROC curves of MLP (27, 14) for the cool and warm seasons of 2005

Region	Cool season	Warm season
1	0.72	0.65
2	0.68	0.63
3	0.80	0.63
4	0.78	0.66
5	0.78	0.65
6	0.67	0.61
Average	**0.74**	**0.64**

We defined an ensemble of the four MLPs for each season in the following way: when three of the four MLPs predict the precipitation amount to exceed the seasonal dependent threshold, then the ensemble issues the warning, otherwise there is no warning.

4. ALADIN NWP model and a combined system (ALADIN+ANN ensemble)

The total amount of precipitation in 2005 at the 40 Bulgarian stations was 130–170% of the annual average measured in the years 1961–1990 (*Fig. 3*).

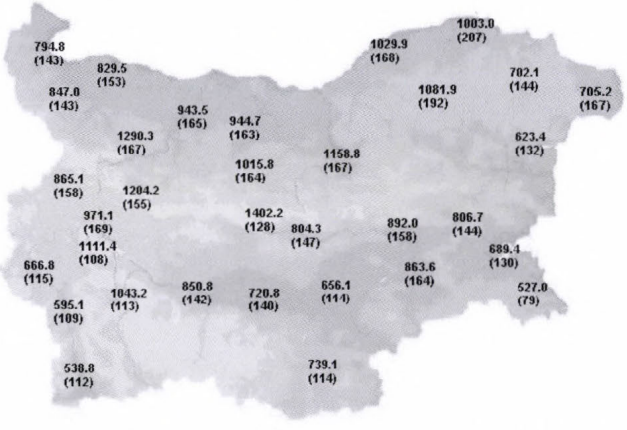


Fig. 3. Annual amount of precipitation (mm) for 2005 over Bulgaria and the percentage (in brackets) of the annual average for the period 1961–1990.

Heavy precipitation produced many flash floods, affecting most of the country, especially during the summer. Damage was extensive, and there were even casualties (*Gospodinov et al.*, 2006; *Stoicheva* and *Latinov*, 2006; *Hristov* and *Latinov*, 2006; *Georgiev* and *Santurette*, 2006). June and August were extremely wet months with monthly precipitation amounts exceeding 2–3 times the average values. The second wettest May–August period in Bulgaria since 1961 was 2005 (*Simeonov et al.*, 2006). Some verification results of precipitation event forecasting are presented here for ALADIN NWP model and for a combined system using ANN technique for the year 2005.

The ALADIN limited area NWP model has been built by the international collaboration between Météo-France and National Meteorological Services of Central and Eastern European countries and has been in operation in Bulgaria since 1999. Verification results for the year 2005 of ALADIN NWP model for precipitation predictions are presented in *Table 5*. Four skill scores are presented here. The first index presented here is the bias score (*BIAS*) defined as

$$BIAS = \frac{H + FA}{H + M}, \quad 0 \leq BIAS \leq \infty, \tag{1}$$

where H is the number of hits (the forecast system predicted the heavy precipitation event and it occurred), FA is the false alarms (the forecast system predicted the heavy precipitation event but it did not occurred), and M is the misses of the contingency table. This score measures the ratio of the frequency of forecasted events to the frequency of observed events.

Table 5. Bias score (*BIAS*, the perfect score is 1), Equitable threat score (*ETS*, the perfect score is 1), Hanssen and Kuipers discriminant (*HK*, the perfect score is 1), and odds ratio (*OR*, the perfect score is infinity) of the ALADIN NWP model (ALAD) and the combined system (ALAD + ANN) for the six regions in 2005

Region	BIAS		ETS		HK		OR	
	ALAD	ALAD+ANN	ALAD	ALAD+ANN	ALAD	ALAD+ANN	ALAD	ALAD+ANN
1	0.58	0.79	0.11	0.59	0.16	0.69	3.5	108.1
2	0.82	0.82	0.12	0.69	0.20	0.75	4.2	318.5
3	0.74	1.05	0.21	0.57	0.30	0.74	8.6	78.8
4	0.88	1.00	0.21	0.57	0.33	0.72	8.0	77.2
5	0.83	1.00	0.24	0.57	0.36	0.73	8.3	66.5
6	0.66	0.83	0.27	0.52	0.38	0.64	10.6	44.4
Avg.	**0.75**	**0.92**	**0.19**	**0.59**	**0.29**	**0.71**	**7.2**	**115.6**

The equitable threat score (Gilbert threat score) (*ETS*; *Gandin* and *Murphy*, 1992) is often used for verification of precipitation:

$$ETS = \frac{H - H'}{H + FA + M - H'}, \quad -1/3 \leq ETS \leq 1, \tag{2}$$

and it measures the fraction of observed and/or forecast events that were correctly predicted, adjusted for hits associated with random chance H'. H' is given as

$$H' = \frac{(H + FA)(H + M)}{N}, \tag{3}$$

where N is the total number of cases. The perfect score is $ETS=1$, while $ETS=0$ indicates no skill.

The Hanssen and Kuipers discriminant (*HK*; *Hanssen* and *Kuipers*, 1965) is defined as

$$HK = \frac{H}{H + M} - \frac{FA}{FA + CN}, \quad -1 \leq HK \leq 1, \tag{4}$$

where *CN* is the number of correct negatives (the forecast system predicted that the precipitation would be below the threshold and it occurred) of the contingency table. This score can be interpreted as a_1+a_2-1, where a is the accuracy for heavy precipitation events and a_2 is the accuracy for non-heavy precipitation events, and it shows how well a given forecast system separates the events of the two classes. The perfect score is $HK=1$, while $HK=0$ indicates no skill.

The last skill score presented here is the odds ratio (*OR*; *Stephenson*, 2000) defined as

$$OR = \frac{H \cdot CN}{M \cdot FA}, \quad -\infty \leq OR \leq +\infty, \tag{5}$$

measuring the odds of making a hit to the odds of making a false alarm.

To obtain better skill in forecasting heavy precipitation events we defined an ANN ensemble of the MLPs in the following way: when three of the four MLPs predict the precipitation amount to exceed the seasonal dependent threshold, then the ensemble issues the warning, otherwise there is no warning.

The second step is to create a combined system consisting of the ALADIN and ANN ensemble. This system works in the following way. First, the ALADIN model determines the precipitation events as those, when the model predicts precipitation amounts greater than a given threshold, here greater than 3 mm/12 hours. Since ALADIN overestimates precipitation events, the threshold is defined to minimize false alarms, and mountains are excluded from consideration as well. Second, as the precipitation event is predicted, the prediction of ANN ensemble is taking into account.

The combined system gives considerably better results than the ALADIN model alone in forecasting heavy precipitation events. The ALADIN NWP model underestimates heavy precipitation events. Although the combined system also shows tendency to underforecast heavy precipitation on average, the *BIAS* is much closer to 1. *ETS* gives clear indications that the combined system predicts more accurately heavy precipitation events than the ALADIN. *HK* shows that the combined system exceeds considerably the results of the ALADIN model in separating the two classes: heavy and non-heavy precipitation. Results for the *OR* indicate that the odds of correct heavy precipitation forecast for the combined system is several times greater than the odds calculated for the ALADIN model alone.

5. Conclusions

Successful application of the neural network technique for precipitation class distinction was presented. Two four-member sets of MLPs for cool and warm seasons were designed. The ROC curves showed good skill of the trained MLPs, although better in the cool season than in the warm season, consistently with previous results. Two reasons are possible for this performance. First of all, the dominant precipitation mechanisms during these two periods of the year are different. In summer, convective precipitation is dominant, in contrast to winter, when the precipitation is mainly stratiform. Second, convective precipitation might be local, but torrential. If the observing network is sparse, then such local precipitation might not be measured.

We found close relationship between certain input variables and the precipitation class through sensitivity analysis. The most important variables for cool and warm seasons are different, but temperature and wind variables are dominant.

We have showed that the combination of the ALADIN NWP model and the ANN ensemble gives better results than the NWP model alone. In this context, we succeeded in our objective to improve the forecasting of heavy precipitation combining NWP predictions with the classical synoptic approach.

Our technique could be improved in the following ways. First, the combinations of the input variables might not be unique. Perhaps, other combinations are possible. Other variables probably exist, which are not among these 27 inputs, but which may be very important for the precipitation process.

This paper is the first attempt to use neural networks for weather forecasting in Bulgaria, and the MLPs described in this paper could be applied to operational forecasting. This method should be used together with other forecasting data, such as the input from limited area NWP models, for achieving better results. Including the neural network technique in the forecasting process will help forecasters be more confident of issuing weather warnings. However,

because the neural network approach is not perfect, there is still a role for humans in the forecast process.

Acknowledgments—Thanks to *David Schultz* for providing very helpful comments to improve this manuscript. Thanks to *Pravda Dimitrova* for her courtesy in providing information about the annual precipitation amounts for 2005. We are grateful to *Plamen Neitchev* and *Anna Ghelli* who helped us improve this manuscript.

References

Banket, R., 1994: Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network. *J Appl Meteorol 33,* 909–917.

Bishop, C., 1995: *Neural Networks for Pattern Recognition.* Oxford University Press, 504 pp.

Chevallier, F., Morcrette, J.-J., Chéruy, F. and *Scott, N.A.*, 2000: Use of a neural-network-based long-wave radiative-transfer scheme in the ECMWF atmospheric model. *Q J Roy Meteorol Soc 126,* 761-776.

Doswell, C.A. III, Brooks, H.E. and *Maddox, R.A.*, 1996: Flash flood forecasting: An ingredients-based methodology. *Weather Forecast 11,* 560-581.

Fritsch, J.M. and *Carbone, R.E.*, 2004: Improving quantitative precipitation forecasts in the warm season: A USWRP research and development strategy. *B Am Meteorol Soc 85,* 955-965.

Georgiev, C. and *Santurette, P.*, 2006: MSG WV-channels. Exercises. Diagnosis of mid- to upper-level convection environment by using 6.2 μm and 7.3 μm images. *NIMH–EUMETSAT Workshop on MSG Applications in Nowcasting*, Sofia, Bulgaria, EUMETSAT, CD-ROM. [Available from C. Georgiev, Forecasting Department, National Institute of Meteorology and Hydrology, 66 Tsarigradsko Chaussee, Sofia 1784, Bulgaria.]

Ghelli, A., cited 2006: Verification of weather parameters. Available online at http://www.ecmwf.int/newsevents/meetings/forecast_products_user/Presentations2006/Ghelli.pdf.

Glahn, H.R. and *Lowry, D.A.*, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J Appl Meteorol 11*, 1203-1211.

Gandin, L.S. and *Murphy, A.H.*, 1992: Equitable skill scores for categorical forecasts. *Mon Weather Rev 120*, 361-370.

Gospodinov, I., Stoyanova, S. and *Dimitrova, P.*, 2006: Flood event in Bulgaria in August 2005 (the Ichtiman Cyclone). *Conf. on Water Obs. and Information Sys. for Decision Support.* Ohrid, Macedonia, BALWOIS, CD-ROM, A-204.

Hall, T., H. E. Brooks Doswell III, C.A.. 1999: Precipitation forecasting using a neural network. *Weather Forecast 3,* 338-345.

Hanssen, A.W. and *Kuipers, W.J.A.*, 1965: On the relationship between the frequency of rain and various meteorological parameters. *Meded. Verhand. K. Nederlands Meteor. Inst. 81,* 2-15.

Hristov, H. and *Latinov, L.*, 2006: Three situations in summer 2005 led to floods in Bulgaria (in Bulgarian). *Proc. First National Research Conf. on Emergency Management and Protection of the Population.* Sofia, Bulgaria, Bulgarian Academy of Sciences, 134-141

Lalaurette, F. and *van der Grijn, G.*, 2006: Ensemble forecasts: Can they provide useful early warnings? In *Predictability of Weather and Climate* (eds.: *T. Palmer* and *R. Hagedorn*). Cambridge University Press, 614-627.

Levenberg, K., 1944: A method for the solution of certain non-linear problems in least squares. *Q J Appl Math 2,* 164-168.

Marquardt, D.W., 1963: An algorithm for least-squares estimation of non-linear parameters. *J Soc Indust Appl Math 11,* 431-441.

Marzban, C., 2000: A neural network for tornado diagnosis. *Neural Comput Appl 9,* 133-141.

Masters, T., 1993: *Practical Neural Network Recipes in C++*. Academic Press, San Diego, 493 pp.

Parker, D.B., 1985: Learning logic. *Tech. Rep. TR-47,* Center for Comp. Research in Economics and Management Science, Massachussetts Institute of Technology, Cambridge, Massachussetts.

Rodwell, M.J., 2006: Comparing and combining deterministic and ensemble forecasts: How to predict rainfall occurrence better. *ECMWF Newsletter 106*, 17-23.

Roebber, P.J., Bruening, S.L., Schultz, D.M. and *Cortinas, J.V. Jr.*, 2003: Improving snowfall forecasting by diagnosing snow density. *Weather Forecast 18,* 264-287.

Rumelhart, D.E., Hinton, G.E. and *Williams, R.J.*, 1986: Learning internal representations by error propagation. In *Parallel Distributed Processing, Vol. 1* (eds.: *D.E. Rumelhart* and *J.L. McClelland*). MIT Press, 318-362.

Shepherd, A.J., 1997: *Second-Order Methods for Neural Networks: Fast and Reliable Training Methods for Multi-Layer Perceptrons*. Springer-Verlag, 145 pp.

Simeonov, P., Bocheva, L. and *Marinova, T.*, 2006: Assessment of the risk of hazardous convective weather during the warm half of the year (in Bulgarian). *Proc. First National Research Conf. on Emergency Management and Protection of the Population*. Sofia, Bulgaria, Bulgarian Academy of Sciences, 88-96.

Simmons, A.J. and *Hollingsworth, A.*, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Q J Roy Meteor Soc. 128,* 647-677.

Stephenson, D.B., 2000: Use of the "odds ratio" for diagnosing forecast skill. *Weather Forecast 15*, 221-232.

Stoicheva, A., and *Latinov, L.*, 2006: Synoptical conditions in winter 2005 led to floods and snowstorms in particular regions of Bulgaria (in Bulgarian). *Proc. First National Research Conf. on Emergency Management and Protection of the Population*. Sofia, Bulgaria, Bulgarian Academy of Sciences, 211-219.

Werbos, P.J., 1974: Beyond regression: New tools for prediction and analysis in the behavioural sciences. *Ph.D. thesis*, Harvard University.

Wilby, R.L. and *Wigley, T.M.L.*, 1997: Downscaling general circulation model output: a review of methods and limitations. *Prog Phys Geog 21*, 530-548.

Xiao, R. and *Chandrasekar, V.*, 1997: Development of a neural network based algorithm for rainfall estimation from radar observations. *IEEE T Geosci Remote 35,* 160-171.

Zhanqing, L., Khananian, A., Fraser, R.H. and *Cihlar, J.*, 2001: Automatic detection of fire smoke using artificial neural networks and threshold approaches applied to AVHRR imagery. *IEEE T Geosci Remote 39,*1859-1870.

Zweig, M.H. and *Campbell, G.*, 1993: Receiver-Operating Characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem 39,* 561-577.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 3, July–September 2010, pp. 169–185

Satellite-derived normalised difference vegetation index for monitoring climate impacts on European agriculture

Leonidas Toulios[1], **Gheorghe Stancalie**[2], **Elena Savin**[2], **F. Mark Danson**[3], **Piotr Struzik**[4], **Zoltán Dunkel**[5], and **János Mika**[6,7]

[1]*National Agricultural Research Foundation, 413 35 Larissa, Greece; E-mail: ltoulios@nagref.gr*
[2]*National Meteorological Administration*
013686, Bucharest, Romania; E-mail: gheorghe.stancalie@meteoromania.ro
[3]*Centre for Environmental Systems Research, University of Salford*
Manchester M5 4WT, UK; E-mail: f.m.danson@salford.ac.uk
[4]*Institute of Meteorology and Water Management*
P. Borowego Str. 14, 30-215 Krakow, Poland; E-mail: piotr.struzik@imgw.pl
[5]*Hungarian Meteorological Service*
P.O. Box 39, H-1675 Budapest, Hungary; E-mail: dunkel.z@met.hu
[6]*Hungarian Meteorological Service*
P.O. Box 38, H-1525 Budapest, Hungary; E-mail: mika.j@met.hu
[7]*Eszterházy Károly College, Department of Geography, Leányka u. 6, H-3300 Eger, Hungary*

(Manuscript received in final form September 21, 2009)

Abstract—Some of the climate and biophysical variables essential for understanding and monitoring the climate system and the impact of climate change on agriculture can be efficiently observed from orbital platforms, providing global data sets continuously and consistently. In order to describe the status of satellite–derived data useful for monitoring climate impacts on European agriculture, an initiative was started within the EU COST Action 734, for the registration and analysis of the relevant satellite data records, based on a specific questionnaire. It was noted that among European countries there are great differences concerning the climate and biophysical data derived from satellite measurements in terms of type, collecting period, spatial and spectral characteristics. However, in many countries satellite data have been collected systematically for several years, and these data records have proved to be useful for climate change impact studies in agriculture. The main variables that are collected in an operational or experimental way are land surface temperature and normalized difference vegetation index (NDVI). This paper presents case studies in three European countries, representing the current application stages in the field of the use of NDVI data in impact studies of climatic change on agriculture. Future research will assess the availability and quality of time series of spectral and biophysical data available from current satellite sensors. The main goal of this paper is to demonstrate the common research activity focused on NDVI and carried out in the frame of EU COST Action 734.

Key-words: NDVI, NOAA/AVHRR, SPOT/VEGETATION, vegetation state monitoring, drought monitoring, regionalization, European survey, COST/CLIVAGRI

1. Introduction

The main objective of EU COST Action 734 "Impacts of Climate Change and Variability on European Agriculture – CLIVAGRI" is the evaluation of possible impacts from climate change and variability on agriculture, and the assessment of critical thresholds for various European areas. This goal will be achieved through the accomplishment of intermediate aims, in order to assess the current and future levels of critical thresholds and hazards for agricultural activity and environmental resources. The activities of four working groups have been structured as a matrix, with the rows representing the methods of analysis and the columns the phenomena and hazards. Each intersection point describes the evaluation of past, present, and future trends of climate and so the impacts on agriculture. Based on these results, possible actions (specific recommendations, suggestions, warning systems) will be defined and proposed to the end users, depending on their needs. More details are available at: http://www.cost734.eu.

An important objective of Action 734 is to study the benefits of satellite remote sensing to assess climate change and variability impacts on agriculture. The research effort is directed towards the analysis of the role of satellite data in models and indices for assessing the impacts of climate change and variability on European agriculture. The type and distribution of vegetation in a given geographical region is diagnostic of the area's climate. This is because vegetation integrates the effects of precipitation and temperature over all time frames longer than a few days. In addition, the vegetation feeds back into climate through its effect on the surface energy budget and moisture balance and its impact on surface roughness and albedo. Vegetation index values, specifically the NDVI, calculated from reflectance recorded in red and NIR wavebands are often referred to as greenness values, because they are strong indicators of vegetation condition and quantity. For these reasons, NDVI time series allow monitoring of vegetation changes over seasonal to inter-annual time frames and over long periods, to monitor climate impacts on agriculture. The objectives of this study are to provide an overview of the underpinning principles behind the application of vegetation indices to assess climate change impacts, and to assess the current operational application of such methods at a European scale.

2. COST 734 pan-European survey

In order to describe the status of satellite climatic and biophysical data that are used for warning purposes for agriculture in Europe, an inventory (http://www.cost734.eu/working-group-2) was created through a questionnaire disseminated to the national delegates of COST 734 member countries (*Toulios et al.*, 2008). The analysis, based on the questionnaires returned by the COST 734 national delegates (15 European countries), shows a general interest in using

satellite climate and biophysical data and products to understand better how climate affects crop growth and yield, and for warning purposes in agriculture. Among European countries there is a great heterogeneity concerning the climatic and biophysical data received from satellite sensors or collected as satellite-derived ready products. Some countries have been collecting satellite data systematically for many years, and these data records could be useful for modeling in climate change impacts studies. Based on the distributed questionnaire, the 12 countries that use the variable NDVI are Austria, Bulgaria, France, Germany, Greece, Hungary, Italy, Poland, Romania, Slovenia, Spain, and Switzerland. The type of satellites and associated instruments are MODIS/TERRA-AQUA, AVHRR/NOAA, VEGETATION/SPOT, TM/LANDSAT, and SEVIRI/-METEOSAT.

Climate and biophysical variables are very relevant for monitoring vegetation status and predicting the possible impacts on crops. Some efforts are currently made by the remote sensing community to provide such products from medium spatial resolution satellite observations operationally available, including VEGETATION, MODIS, and AVHRR. However, the products are currently poorly validated because of the lack of ground measurements. In addition, most of the validation activities correspond to 'one shot' ground measurements, precluding access to the seasonality and associated phenology of the crops, even though most of the information on vegetation functioning lies in its dynamics.

Since the COST 734 survey revealed NDVI as one of the most widely used satellite-based variable in many European countries, this paper focuses on the use of NDVI time series imagery for monitoring climate impacts on agriculture. In this respect case studies in three European countries are presented.

3. Current status of satellite-derived NDVI in impact studies of climatic change on agriculture

Observations of spectral reflectance from remotely sensed data allow the quantitative characterization of terrestrial vegetation canopies. Efforts have been made to develop algorithms deriving surface properties from remotely sensed spectral reflectance or their end products. These include both empirical schemes, such as the spectral vegetation index (*Sellers et al.*, 1994; *Hall et al.*, 1995) and model inversion methods (*Pinty et al.*, 1990; *Privette et al.*, 1995; *Gao* and *Lesht*, 1997; *Danson et al.*, 2003).

The most relevant variables measured over land are solar radiation, daily global albedo, leaf area index (LAI), land surface temperature (LST), rainfall, evapotranspiration, cloud cover, fires and burned areas, snow cover, digital elevation maps of ice sheet surfaces, glacier evolution, and land cover. Some of these variables are required as inputs to give an immediate view of climate

change impacts, for example (*Struzik et al.*, 2008). They are also widely used as inputs to agrometeorological models (*Struzik*, 2005; *Stancalie et al.*, 2006). The most important parameters derived from vegetation indices are the maximum greenness during the growing season, total greenness during the growing season, fraction of photosynthetically active radiation (FPAR), absorbed photosynthetically active radiation (APAR), and leaf area index (LAI).

Various remote-sensing-based studies have revealed persistent spectral relationships between the red and near-infrared (NIR) reflectance and the amount of green vegetation. Due to vegetation pigment absorption (chlorophyll, proto-chlorophyll), the reflected red energy decreases, while the reflected NIR energy increases as a result of the strong scattering processes of healthy leaves within the canopy. Directly using the amount of reflected red and/or NIR radiation to study the biophysical characteristics of vegetation is normally inadequate, for reasons rooted in the intricate radiative energy interactions between the canopy, soil background, atmosphere, and the non-uniqueness of the signatures. When, however, two or more bands are combined into a vegetation index (VI), the vegetation related component of the signal is boosted and the information becomes more useful. VIs can therefore be used as an indirect measure of vegetation activity.

Many VIs have been reported in the literature and some have proved to be well correlated with vegetation biophysical parameters. Tremendous efforts have been devoted to improve VI and render them insensitive to variations in illumination conditions, observing geometry, and soil properties. Even though the external perturbing factors related to changes in soil brightness and atmospheric conditions may be taken into account, VI still have definite intrinsic limitations; they are not a single measure of a specific variable of interest such as pigment content, plant geometry, or canopy architecture. So far, it has not been possible to design an index which is sensitive only to the desired variable and totally insensitive to all other vegetation parameters.

Therefore, different indices have been defined for different purposes, and optimized to assess a process of interest. For instance, some spectral indices have been proposed to capture the photochemical processes associated with photosynthetic activity such as light use efficiency or to estimate leaf pigment content (*Kim et al.*, 1994; *Broge* and *Leblanc*, 2000; *Daughtry et al.*, 2000; *Haboudane et al.*, 2002). A major problem in the use of these indices arises from the fact that canopy reflectance, in the visible and NIR, is strongly dependent on both structural (e.g., LAI) and biochemical properties (e.g., chlorophyll content) of the canopy (*Jacquemoud et al.*, 2000; *Zarco-Tejada et al.*, 2001).

LAI and chlorophyll content have similar effects on canopy reflectance, particularly in the spectral region from the green (550 nm) to the red edge (750 nm). To uncouple their combined effect, recent research studies (*Daughtry et al.*, 2000; *Haboudane et al.*, 2002) have demonstrated that leaf chlorophyll

content can be estimated with minimal confounding effects due to LAI through a combination of two kinds of spectral indices: indices sensitive to pigment concentration and indices resistant to soil optical properties influence.

Other studies have dealt with modifying VIs to improve their linearity with, and increase their sensitivity to, LAI (*Haboudane et al.*, 2004a). Consequently, some indices have been identified as 'best' estimators of vegetation biophysical parameters, because they are less sensitive to the variation in external parameters affecting the spectral reflectance of the canopy, namely soil optical properties and atmospheric conditions (*Broge* and *Leblanc*, 2000), as well as to changes in leaf intrinsic properties such as chlorophyll concentration (*Haboudane et al.*, 2004b).

The most widely used form of VI, the normalized difference vegetation index, was introduced by *Deering* (1978) and *Tucker* (1979) and is the ratio of the difference of the NIR and red reflectance divided by their sum. The NDVI mitigates a large part of the variations that result from the overall remote-sensing system (radiometric, spectral, calibration, noise, viewing geometry, and changing atmospheric conditions). NDVI is often used as a monitoring tool for vegetation health and dynamics, enabling easy temporal and spatial comparisons. In order to make effective use of NDVI data, issues related to the remote-sensing system need to be addressed. Most serious are clouds, which render any observation unusable by obstructing the target, and, to a lesser degree, the effects of the bidirectional reflectance distribution function (BRDF). To overcome these impediments, the maximum value compositing was developed as an operational approach to producing cloud free consistent NDVI images. Multiple daily images are processed to create a representative, cloud-free image with the least atmospheric attenuation and viewing geometry effects (*Holben*, 1986). The maximum value compositing (MVC) technique is the most widely used method based on maximizing the NDVI signal over a preset period of time (e.g., 16 days). Whilst the MVC helps screen for clouds, it has been found also to favor extreme viewing geometry (large solar zenith angles and large view angles in the forward scatter direction) and, to a lesser extent, cloud shadow. Several studies have attempted to address these issues (*Roujean et al.*, 1992; *Meyer et al.*, 1995; *Leeuwen et al.*, 2006).

Several factors can influence differences in NDVI computed from different satellite sensors. Impacts from BRDF are well documented (*Roujean et al.*, 1992; *Walter-Shea et al.*, 1997) and in addition, the spectral response functions for different sensors can lead to systematic differences in NDVI (*Trishchenko et al.*, 2002). Each sensor has its own instantaneous field of view, swath width, and orbiting geometry. Adjusting for these known differences is somewhat problematic and, in many cases, beyond the capability or resources available to many users (*Linthicum et al.*, 1999). *Table 1* presents the most widely used NDVI datasets, including information on sensor type, data source, spatial and temporal resolutions, equatorial crossing time, and field of view for each sensor.

Table 1. The most widely used NDVI datasets

Sensor	AVHRR	SPOT/ VEGETATION	MODIS	SeaWiFS	LANDSAT ETM+
Data source	GIMMS NDVIg operational dataset	FAS-GIMMS VITO	MODIS Land and Vermote/ Saleous	SeaWiFS/GS F/GIMMS	EOS web
Spatial resolution	800 m and 1 degree	1000 m and 1 degree	500 m 5000 m and 1 degree	4633 m 1 and 1 degree	30 m
Temporal resolution	15 days and monthly	10 days and monthly	16 days and monthly	monthly	16 days and monthly
Equatorial crossing	9 AM - 6 PM	10:30 AM	10:30 AM	12:05 PM	10 AM
Field of view	+/- 55.4°	+/- 101°	+/- 55°	+/- 58.3°	+/- 15.5°

An analysis made by *Brown et al.* (2007) revealed that, although relatively large differences existed between the mentioned NDVI datasets, the NDVI anomalies exhibited similar variances. Composite NDVI images are quite robust, which can be seen when comparing time series data with NDVI from LANDSAT ETM+ images that have been corrected for atmospheric effects.

Bai et al. (2004) generated an NDVI data set to provide a satellite record of monthly changes in terrestrial vegetation, using NOAA/AVHRR data. The NDVI trend in Europe was analyzed by taking representative pixels of each land use type. Generally, an increasing trend in NDVI in Europe is shown over the past 21 years (1982–2002). The increase of NDVI ranges from 0.03% to 0.2% per year (*Fig. 1*). The variations in NDVI of the different land use types over certain months in a year were observed. The variations were mainly due to the annual growth cycle of the vegetation. Characteristic features of specific land use types were visible in the variation of NDVI, like the harvesting of agricultural land, the loss of leaves in broad-leaved forests, and the persistent greenness of evergreen coniferous forests.

Satellites have different repeat cycles, which means that some satellites may provide only two images per month, whereas others can produce an image every day. This can have considerable impact on agricultural applications, since the repeat cycle characteristics of satellites are, for example, one of the determinants for forecasting yield using the NDVI. In addition, these high repeat cycle platforms usually have lower spatial resolution, influencing the appropriateness of its data to fine-scale applications. Another major cause of variation in the availability of optical data is cloud coverage. This could be problematic when the timing of image acquisition is critical as in crop identification. However, it may be avoided by using short-time compositing

synthesis that is available. In the case of NDVI, the highest value of the considered time period is chosen.

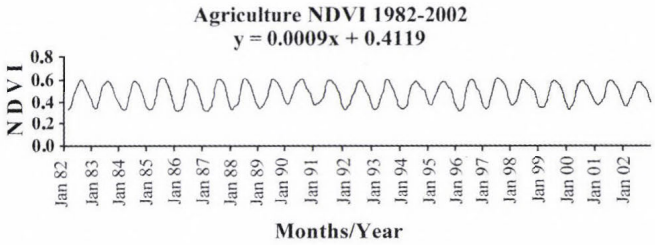


Fig. 1. Agriculture NDVI changes in Europe during 1982–2002 (*Bai et al.*, 2004).

3.1. Temporal resolution

It is important to get the right balance between the temporal resolution of data and the timescale of variation in the quantity measured. For example, maximum value composites of NDVI data are rarely less than decadal. Consequently, NDVI MVC should not be used to investigate short-term events. It is more appropriate to use NDVI data to monitor events of longer duration like the growth of vegetation through a season.

3.2. Atmospheric interference

Light scattering tends to increase the amount of red radiation received by the satellite as red is more readily scattered in the atmosphere than near infrared. This induces a reduction in the NDVI values. For short periods, this problem is minimized by taking the maximum value composite, because each pixel's maximum value is likely to have occurred when scattering was at a minimum.

3.3. Land cover type

With the exception of very large irrigation schemes and large area commercial agriculture, an NDVI pixel rarely covers a single, homogeneous cover type of agricultural land. Instead it may cover roads, buildings, bare soil, small water bodies, natural vegetation, and different agricultural crops, all within one pixel. In an NDVI image, one pixel is the weighted sum of the radiation reflected from all the land cover types within the area covered by the pixel. NDVI is an indicator of the condition of the overall vegetation in an area, including natural vegetation and crops. In rain-fed agriculture, natural vegetation may follow patterns similar to those of the crops. More often, however, agricultural crops are more susceptible to adverse conditions and follow different growth cycles.

3.4. Soil-type

Light reflected from the soil can have a significant effect on NDVI values (changing the values by up to 20%). Generally, the greater the reflectance of the soil is, the lower the NDVI values are. Soil-type is, therefore, an important factor. Given two soil types, the one with a greater reflectivity but with similar vegetative conditions, will on average produce lower NDVI values.

3.5. Off-nadir viewing effects

As the radiometer scans across the Earth, there is only one point, in the centre of the scan that is directly underneath the radiometer (called the sub-satellite point or nadir). The distance from the radiometer to the ground increases away from the sub-satellite point. This results in increased atmospheric interference as the light must pass through more atmosphere before reaching the radiometer, and, therefore, in reduced NDVI values. These effects are minimized by simply dropping the pixels at large off-nadir view angles.

4. Case studies

4.1. Time series of SPOT/VEGETATION NDVI for monitoring crop status in Romania

The agrometeorological activity of the Romanian National Meteorological Administration (RNMA) integrates complex topics concerning the current and future development of the phenological stages of the crops and soil moisture supply with respect to trends in meteorological parameters. This is a particularly important activity whose final objective is to provide information to the agro-meteorological bulletins and disseminate information to the level of the agricultural decision makers.

From 2005, the RNMA has benefitted from data provided by the SPOT/VEGETATION sensor (transmitted in quasi real time via FTP, as well as archived data) for better operational surveillance of agricultural areas. VGT-S10 products (ten-day synthesis) are compiled by merging segments acquired during a ten-day period. These products provide data from all spectral bands, the NDVI and auxiliary data on image acquisition parameters. An MVC synthesis can be also obtained at several different spatial resolution levels (1×1 km^2, 4×4 km^2, or 8×8 km^2).

The VGT-S10 products, namely decadal NDVI and color synthesis of composites, allow for crop state monitoring and analysis by using NDVI time series. Other more complex vegetation indices are derived, like the modified soil vegetation index (MSAVI) and vegetation condition index (VCI), by correlating

with various agrometeorological parameters provided by the national weather network.

- The key stages for crop state analysis, based on NDVI data series are: overlay of a mask in order to isolate agricultural land,
- overlay of reference points,
- selection of representative crop-cultivated areas with at least 4 km2 surfaces (equivalent of 4 pixels) using terrain information or high resolution satellite-derived land cover maps,
- accurate localization of the reference points using GPS,
- extraction of NDVI values.

The method developed for agricultural drought monitoring includes the following main steps:

- Deriving the NDVI values for each decade for each study area.
- Smoothing the data series using a moving temporal window and linear regression.
- Averaging the value of this linear function using various interpolation functions of a higher order.
- Comparing the yearly data series to observe a trends caused, for example, by the growth of the vegetation. Small differences caused by long-term fluctuations of climatic conditions in some regions can be highlighted as well.
- Using the multiyear NDVI minima and maxima, the values are forced into the same range and are made quantitatively comparable afterwards. This complex parameter is the vegetation condition index (VCI) given by

$$VCI=(NDVI-NDVI_{min})/(NDVI_{max}-NDVI_{min})\cdot 100. \quad (1)$$

The whole algorithm is applied for every pixel of the satellite image. In research by the Remote Sensing Team of the Romanian National Meteorological Administration (*Kianicka et al.*, 2007), this algorithm was applied using a time series of SPOT/ VEGETATION NDVI for a test area, situated in Arges County located in the southern part of Romania.

The latest generation of SPOT satellites is well known for their high-resolution imagery that is useful in agriculture, urban planning, forest management, monitoring and surveillance, etc. Considering the need for data with higher temporal resolution covering the whole globe, the new generation of satellites carries the VEGETATION instrument with a spatial resolution of 1 km and a

swath width of 2250 km which is able to image the whole of the Earth's surface on a daily basis. With its four channels, the sensor may be used to derive various secondary products. They are well suited to mapping vegetation over large areas and for global change studies.

In this application, the raw data are geometrically and atmospherically corrected and geo-referenced. The ten-day syntheses are also available. The original data are contaminated by clouds and weather fluctuations. In order to avoid this phenomenon as much as possible, short-time syntheses are generated. In the case of NDVI, the highest value of ten available images is chosen. Nevertheless, the cloudy parts are still presented, but are assigned specific values.

First, the data are smoothed, VCI indices are computed for each year, and the differences from the average value in this case are evaluated. These differences generate the first information about the behavior of drought in the study area. Applying a threshold, the potential map of drought stress can also be derived. *Fig. 2* presents the co-plot of VCI annual curves of the 1998–2005 decadal values, for a test area situated in Arges County. The computation of VCI has been restricted only to the crop growing season from April to the beginning of November.

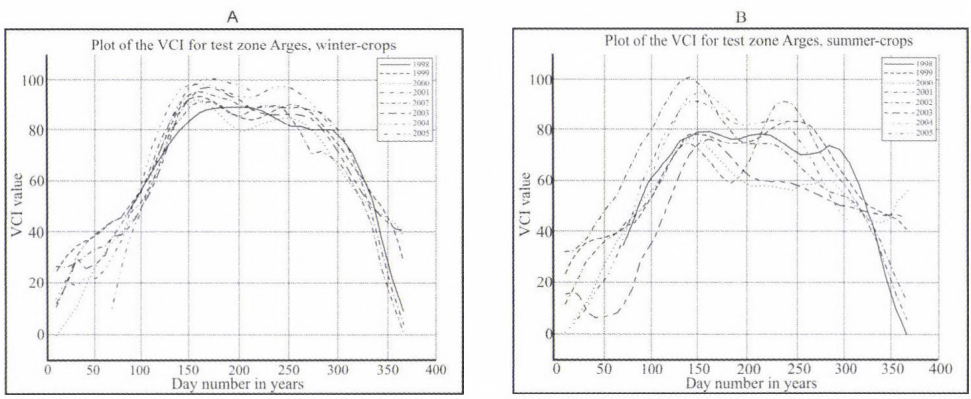


Fig. 2. Co-plot of annual VCI curves for winter (A) and summer crops (B).

The difference between winter and summer crops is clear; the summer crops have a more marked dispersion of values. In the drought years, 2000 and 2003, the cereal production of Romania was seriously reduced by drought, similarly to the reduced VCI values during summer. The co-plot of VCI curves and, consequently, the co-plot of differences regarding the linear-regression-in-time values (*Fig. 3*) confirm the results connected to monitoring of cereal output, temperature, and soil moisture.

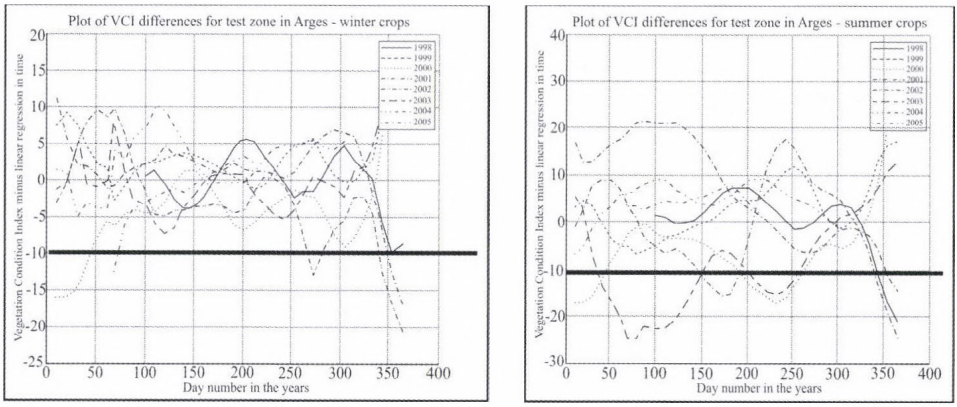


Fig. 3. Differences between original decadal value and linear-regression-in-time values for the winter and summer crops.

4.2. Vegetation indices for drought monitoring and crop damage assessment in Poland

Drought is the most important weather-related natural disaster. Droughts cannot be avoided, but although predictions will never be perfect, we can reduce their impacts. Large-scale intensive droughts have been observed on all continents, leading to massive economic losses, destruction of ecological resources, food shortages, and starvation of millions of people.

Reduction of drought consequences requires proper drought prediction, early warning and monitoring, to help with proper agricultural practices and decisions. Monitoring the impacts and extent of drought is also important for insurance companies. Lessons learned during an event may help in proper planning of irrigation systems. Thus, satellite information may have a substantial impact on improved drought management strategies (*Struzik*, 1991 and 1992). Its use in drought monitoring and early warning requires retrieval and monitoring of parameters such as:

- Rainfall, surface wetness, and temperature monitoring. Multi-channel and multi-sensor data sources from geostationary platforms (such as GOES, METEOSAT, GMS) and polar orbiting satellites (such as NOAA, DMSP SSMI, METOP, SMOS, GPM) have been used. Estimated indices describing actual atmosphere and surface conditions and also quantities useful for forecasting are precipitation intensity, amount, and coverage, surface soil wetness indices, atmospheric moisture content, wind speed and direction.

- Vegetation condition monitoring is currently possible, ranging from NOAA AVHRR data at 1.1 km resolution on a daily revisit from each satellite, to environmental satellites (LANDSAT etc.) on a 8–16 days revisit with a 10–30 m spatial resolution or even very high spatial resolution missions like IKONOS, QuickScat, etc. with resolutions of 0.6–4 m. The normalized difference vegetation index (NDVI), associated vegetation conditions index (VCI), and temperature condition index (TCI) derived from the satellite data are accepted worldwide for regional monitoring.
- Assessment of drought impact by high-resolution satellite sensors from LANDSAT, SPOT, and other high resolution satellites are being used for the assessment of impacts, but in most cases this is not a country-wide activity, mainly due to high costs of such information.

Examples of vegetation anomalies in Poland connected with severe drought conditions are presented in *Fig. 4*. NDVI indices were calculated from NOAA/AVHRR data, using the MVC for selected periods. Year 1992 was extremely dry in Poland, preceded by two dry years 1990 and 1991. Monthly precipitation amounts in Wielkopolska and South Pomeranian regions (dark on satellite NDVI image) were well below the long term monthly means. At some stations, precipitation in all months of the year was well below the long term mean (*Table 2*). June precipitation was less than 10% of the monthly mean followed by a very dry July, which resulted in severe drought lasting into late summer. The observed satellite-derived NDVI reached a minimum in August after four dry months.

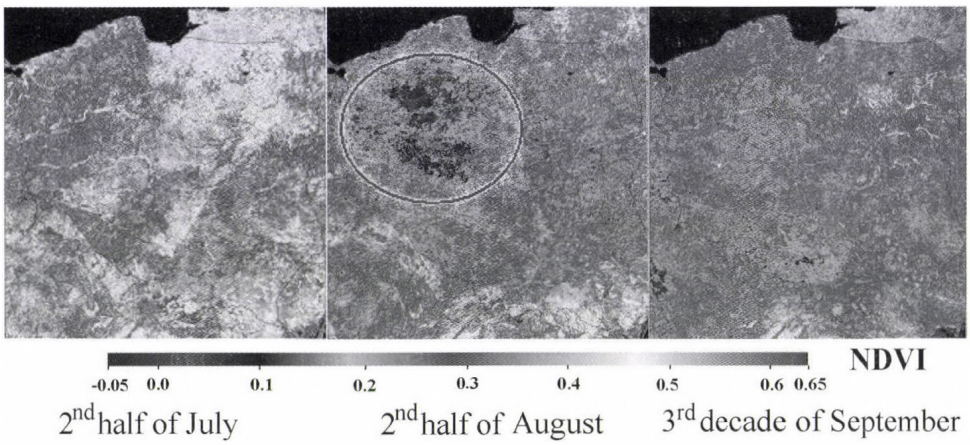


Fig. 4. Vegetation indices during the severe drought in Poland (1992).

Table 2. Monthly precipitation recorded at three selected synop stations during the drought in 1992, in comparison with climatological means for this region

Month	Monthly precipitation in 1992 for selected stations [mm]			Climatological mean for region [mm]
	Poznan	Chojnice	Pila	
April	18.3	31.7	16.4	31.3
May	27.6	31.0	25.7	47.1
June	3.4	36.8	7.9	61.9
July	24.7	36.7	23.1	76.0
August	45.6	43.2	39.6	55.4
September	15.4	32.4	26.4	43.6
Amount for season April–September	**135.0**	**211.8**	**139.1**	**315.3**

4.3. Using vegetation indices for objective regionalization in Hungary

The NDVI was derived from the NOAA/AVHRR sensor for a 14-year period, 1985–1998. The March–October vegetation growth period was further selected representing two-thirds of the year. Since yield data were available only on a county-wide basis, the NDVI series were also averaged over the 19 administrative counties, characterized by 50–90 km of linear size. The integration time step was one week, since under the climatic conditions of the country there is likely to be one cloud free day per week (*Mika et al.*, 2002).

Weekly NDVI data during the vegetation growth periods of 14 years from each of these 19 counties were subjected to a cluster analysis to characterize similarity or dissimilarity among the 19 administrative counties of Hungary. The aim was to find coherent groups of counties to unify them into larger, cumulative samples. More specifically, hierarchical joining based on Euclidean distance was used. The amalgamation was performed by the Ward's method (*Ward*, 1963), which used an analysis of variance approach to evaluate the distances between clusters. The result of the cluster analysis did not yield a strictly determined number of classes, but it is reasonable to rely on the stability of the classification which can be characterized by the average linkage distance between the classes. When a change in this linkage distance does not change the number of classes, the number of classes may be considered as stable. In this case, of the 19 counties, the most stable number of classes was 3 (*Fig. 5*). The remarkable feature of this objective classification is the west-east differentiation, with no clear north-south differences over counties in the NDVI values.

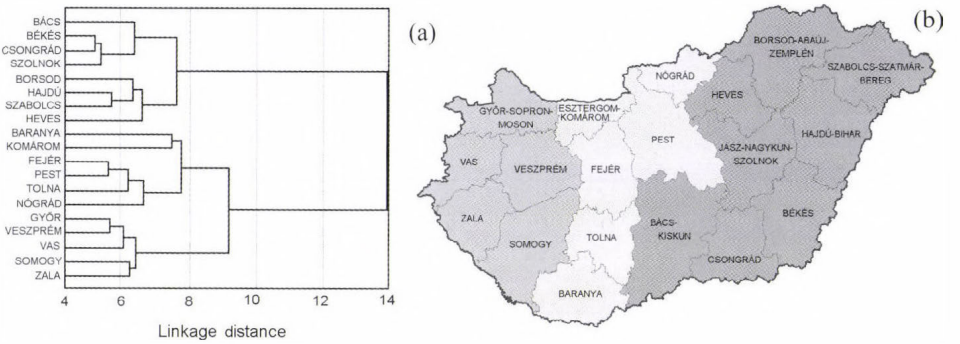


Fig. 5. Results of cluster analysis to determine objective vegetation regions in Hungary. The left axis of part (a) indicates the clustering process with the short names of the administrative counties in the order of their amalgamation. Part (b) indicates the three regions for which a stable linkage distance can be obtained.

5. Conclusions

For the past decade, newer and more sophisticated remote sensing systems have become operational providing biophysical measurements that are aimed at addressing monitoring climate impacts in European agriculture. Based on these advances, the use of multispectral satellite data allows improvement of the classical determination methods for agrometeorological parameters of interest. The most important advantages are related to the improvement in the spatial resolution in the range of meters to kilometers, as well as the frequency of updating with a temporal resolutions varying from hours to seasons.

A new generation of EUMETSAT space sensor systems presents a real opportunity to improve our knowledge of surface processes on a short-term basis. Instruments (like SEVIRI on the board of MSG, AVHRR-3 on the board of NOAA, and future EPS/METOP) will offer the opportunity to follow vegetation changes on a daily time scale, due to a high temporal resolution and more appropriate spectral bands to assess vegetation status. They will also provide data with multiple surface illumination angles that will bring new insights to our knowledge of the anisotropic surface reflectance properties. However, to achieve any meaningful monitoring of the land surface vegetation, stable, inter-calibrated long-term vegetation records (a decade or longer) are a key requirement. Progress can be made towards a unified NDVI dataset given that absolute variances across sensors are relatively similar, especially when seasonality is removed. Examples of the successful use of long time series of satellite data for monitoring climate impacts on agriculture in three European countries have been presented here together with an assessment of the limitations of the technology.

There is a need for additional research to improve the long-term data series record and addressing cross-sensor NDVI continuity. Good examples of the potential application could be the satellite image series of summer droughts of 2003, 2005, and 2006 in France. Among others, the NDVI deviations based on VEGETATION/SPOT 5 data could give very useful time series. Another good example would be the 2003 heatwave at the European scale (July 2003/July 2002). The initial results of the COST 734 inventory show an incomplete snapshot of the situation in the different countries across Europe. The gaps show the importance of pan-European collaboration on the collection, archiving and analysis of satellite derived data on variables related to climate change and agriculture.

References

Bai, Z., van Barneveld, B., Laita, E.L., Mengesha, T. and *Morelissen, B.*, 2004: Trends in NDVI changes in Europe using a 21 year satellite record. Poster submitted in partial fulfilment of the course Academic Master Cluster Environmental Sciences, part II. Wageningen UR, May 2004.

Broge, N.H. and *Leblanc, E.*, 2000: Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens Environ 76*, 156-172.

Brown, M.E., Pinzón, J.E., Didan, H., Morisette, J.T. and *Tucker, C.J.*, 2007: Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-vegetation, SeaWiFS, MODIS, and LandSAT ETM+ sensors. *IEEE T Geosci Remote 44*, 1787-1793.

Danson, F.M., Rowland, C.S. and *Baret, F.*, 2003: Training a neural network to estimate crop leaf area index. *Int J Remote Sens 24*, 4891-4905.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., Brown de Colstoun, E. and *McMurtrey III, J.E.*, 2000: Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens Environ 74*, 229-239.

Deering, D.W., 1978: *Rangeland Reflectance Characteristics Measured by Aircraft and Spacecraft Sensors*. College Station, TX, Texas A&M University, p. 338.

Gao, W. and *Lesht, B.M.*, 1997: Model inversion of satellite-measured reflectances to obtain surface biophisical and bidirectional reflectance characteristics of grassland. *Remote Sens Environ 59*, 461-471.

Haboudane, D., Miller, J.R., Pattey, E., Strachan, I. and *Zarco-Tejada, P.J.*, 2002: Effects of chlorophyll concentrations on green LAI prediction in crop canopies: modelling, validation and heterogeneity assessment. *First International Symposium on the Recent Advances In Quantitative Remote Sensing*. Valencia, Spain, 16-20 September 2002.

Haboudane, D., Miller, J.R., Tremblay, N., Pattey, E. and *Vigneault, P.*, 2004a: Estimation of leaf area index using ground spectral measurements over agriculture crops: prediction capability assessment of optical indices. *Proc. of the ISPRS Symposium*, Istanbul, Turkey.

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J. and *Strachan, I.*, 2004b: Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. *Remote Sens Environ 90*, 337-352.

Hall, F.G., Townshend, J.R. and *Engman, E.T.*, 1995: Status of Remote Sensing Algorithms for Estimation of Land Surface State Parameters. *Remote Sens Environ 51*, 138-156.

Holben, B., 1986: Characteristics of maximum-value composite images from temporal AVHRR data. *Int J Remote Sens 7*, 1417-1434.

Jacquemoud, S., Bacour, C., Poilve, H. and *Frangi, J.-P.*, 2000: Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sens Environ 74*, 417-481.

Kianicka, J., Savin, E., Flueraru, C. and *Craciunescu, V.*, 2007: Remote sensing data analyses for the study of drought. *Symposium GIS,* Ostrava, 28-31 Jan. 2007.

Kim, M.S., Daughtry, C.S.T., Chappelle, E.W., McMurtrey III, J.E. and *Walthall, C.L.*, 1994: The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (Apar). *Proc. of the 6th Symp. on Physical Measurements and Signatures in Remote Sensing.* Val D'Isere, France, Jan. 17-21, 1994, 299-306.

Leeuwen, W., Orr, B.J., Marsh, S.E. and *Herrmann, S.M.*, 2006: Multisensor NDVI data continuity: uncertainties and implications for vegetation monitoring applications. *Remote Sens Environ 100*, 67-81.

Linthicum, K.J., Anyamba, A., Tucker, C.J., Kelley, P.W., Myers, M.F. and *Peters, C.J.*, 1999: Climate and satellite indicators to forecast Rift Valley Fever epidemics in Kenya. *Science 285*, 397-400.

Meyer, B., Verstraete, M.M. and *Pinty, B.*, 1995: The effect of surface anisotropy and viewing geometry on the estimation of NDVI from AVHRR. *Remote Sens Rev 12*, 3-27.

Mika, J., Kerényi, J., Rimóczi-Paál, A., Merza, Á., Szinell, C. and *Csiszár, I.*, 2002: On correlation of maize and wheat yield with NDVI: example of Hungary (1985–1998). *Adv Space Res 30*, 2399-2404.

Pinty, B., Verstraete, M.M. and *Dickinson, R.E.*, 1990: A physical model of the bidirectional reflectance of vegetation. *J Geophys Res 95*, 11767-11775.

Privette, J.L., Myneni, R.B., Emery, W.J. and *Pinty, B.*, 1995: Inversion of a Soil Bidirectional Reflectance Model for use with vegetation reflectance models. *J Geophys Res 100*, 25497-25508.

Roujean, J.-L., Leroy, M., Podaire, A. and *Deschamps, P.Y.*, 1992: Evidence of surface reflectance bidirectional effects from a NOAA/AVHRR multitemporal data set. *Int J Remote Sens 13*, 685-698.

Sellers, P.J., Los, S.O., Tucker, C.J., Justice, C.O., Dazlich, D.A., Collatz, G.J. and *Randall, D.A.*, 1994: A global 1 by 1 degree NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *Int J Remote Sens 15*, 3519-3545.

Stancalie, G., Struzik, P. and *Toulios, L.*, 2006: Use of data from remote sensing as input for agrometeorology. In *COST Action 718, Metorological Applications for Ariculture*, COST Office, Brussels, 47-80.

Struzik, P., 1991: Seasonal changes of the NDVI over Poland. *Proc. of Fifth AVHRR Data Users Meeting*. Tromso 25-28, June 1991. EUMETSAT P 09, ISBN 92-9110-003-X, 77-82.

Struzik, P., 1992: Application of AVHRR/NOAA satellite information for monitoring of river basins surface parameters. *Proc. of Int. Conf. Operational Hydrology*. Wola Zręczycka, 21-22 Sept. 1992, 65-74.

Struzik, P., 2005: Possibilities for use of satellite information as an input for CROPWAT software - preliminary study. In *Use and Availability of Meteorological Information from Different Sources as Input in Agro-meteorological Models*. European Commission, DG XII, COST Action 718, 339-361.

Struzik, P., Toulios, L., Stancalie, G., Danson, M., Mika, J. and *Domenikiotis, C.*, 2008: Satellite remote sensing as a tool for monitoring climate and its impact on the environment - Possibilities and limitations. In *Survey of Agrometeorological Practices and Applications in Europe Regarding Climate Change Impacts*. COST, European Science Foundation, 205-236.

Toulios, L., Stancalie, G., Struzik, P., Danson, M., Mika, J., Dunkel, Z. and *Tsiros, E.*, 2008: Satellite spectral, climatic and biophysical data for warning purposes for european agriculture. In *Survey of Agrometeorological Practices and Applications in Europe Regarding Climate Change Impacts*. COST, European Science Foundation, 163-204.

Trishchenko, A.P., Cihlar, J. and *Li, Z.*, 2002: Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. *Remote Sens Environ 81*, 1-18.

Tucker, C.J., 1979: Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens Environ 8*, 127-150.

Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H. and *Sampson, P.H.*, 2001: Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content

estimation in closed forest canopies with hyperspectral data. *IEEE Trans Geosci Remote Sens 39*, 1491-1507.

Walter-Shea, E.A., Privette, J.L., Cornell, D., Mesarch, M.A. and *Hays, C.J.*, 1997: Relations between directional spectral vegetation indices and leaf area and absorbed radiation in alfalfa. *Remote Sens Environ 61*, 162-177.

Ward, J.H., 1963: Hierarchical grouping to optimize an objective function. *J American Stat Assoc 58*, 236-300.

Internet sources

http://www.cost734.eu and http://www.cost734.eu/working-group-2 (both available on 15 September, 2009)

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 3, July–September 2010, pp. 187–201

Analysis of precipitation on Lake Balaton catchments from 1921 to 2007

Angéla Anda[*1] and **Balázs Varga**[2]

[1]*Department of Meteorology and Water, University of Pannonia Georgikon Faculty H-8360 Keszthely, Hungary; E-mail: anda-a@georgikon.hu*

[2]*Agricultural Research Institute of the Hungarian Academy of Sciences, Martonvásár, Hungary*

[*]*Corresponding author*

(Manuscript received in final form November 20, 2009)

Abstract—The aim of this analysis is to provide an overview of the precipitation conditions of the Balaton catchments. Because Lake Balaton has shallow water, it responds sensitively to any change in the environment, including precipitation. The influence of the dry period on the lake's water level in the years 2000–2003 serves as a good example for this. Data series of 25 precipitation gauging stations of catchments, recorded between 1921 and 2007, were investigated in this study.. Annual precipitation sums, monthly distribution of rainfall, and the influence of rainfall on natural change in water storage were studied. The whole catchment area was divided into different parts depending on geographical locations and climatic conditions. We found significant differences in the regional precipitation amounts. The western part of the catchments (half of the Zala River basin) has the highest annual precipitation sum, while going towards the east and in the "cut" of the lake, the amount of rainfall decreases. At the majority of the examined stations the decrease in precipitation was also detectable on the basis of WMO climate means, but we did not find explicit decreasing tendencies in the case of several stations. Surface runoff moved collaterally with precipitation decreases, and its impact on natural change in water storage was even more pronounced. The natural change in water storage of the lake is a good indicator reflecting the drying tendency of the recent past, and it calls attention to the necessity of more deliberated water management of the Lake Balaton. Presumably, the precipitation phenomenon, similarly to the past, will also vary in the future. Analysing the data of the past decades may help in making more established decisions.

Key-words: precipitation, Balaton catchments, natural change in water budget

1. Introduction

Investigations of the water budget of Lake Balaton, the largest lake of Central Europe (46°42'6"–47°3'50" N and 17°14'58"–17°14'58" E, its surface is about 600 square kilometers, and the average depth of it is ~3 m) has been ongoing since the 1870s, because of the construction works of the local southern railway. The scientific investigation of the lake (covering gage measurements) was established in the 1890s. However, reliable data series of water budget are available only from the beginning of the 1920s. Besides the most important water supplier of the lake, the river Zala, there are about 50 watercourses. The inflow to the lake is mainly determined by the amount of precipitation. In the last decade, investigations on a local level have been devoted to precipitation depression as a main issue of global warming (*Bartholy et al.*, 2007a; *Kertész* and *Mika*, 1999; *Varga et al.*, 2007). According to the latest PRUDENCE simulations, the annual precipitation sums of Balaton watershed are not expected to change (A2 and B2 scenarios) significantly, but it is not valid to seasonal distribution of rainfall (*Bartholy et al.*, 2009). Decreases in summer precipitation for 2071–2100 are 35% and 20% at A2 and B2 scenarios, respectively. The increase in winter precipitation is the same as the projected rainfall decline in summer. Earlier results of the authors (*Bartholy et al.*, 2001) harmonize with their last projection.

As a result of increased air temperature and precipitation decline, an unprecedented decrease in the water level of the lake was measured around 2000. The former water shortage was restored without any controlled human intervention by the winter of 2003/2004. Similar fluctuations in lake water levels were also observed in other parts of the world (*Gianniou* and *Antopoulos*, 2007; *Mercier et al.*, 2002; *Schindler*, 2001; *Winter et al.*, 2003). Decreases in lake water levels are not independent of global climate change (*Lenters et al.*, 2005). Aspects of climate variability on the lake water level of different sites were reviewed by *Greenland* and *Kittel* (2002).

The most serious problem with handling of Balaton water level is that only reduction of excess water can be controlled, by discharging it through the Sió sluices, but it is not possible to provide a controlled water supply. The inflow to the lake is totally exposed to the precipitation and water amount delivered by surface runoff determined thereby. Any precipitation deficit is immediately reflected by the water level of the lake, like a mirror. *Somlyódi* and *Honti* (2005) projected warmer (0.5–1.5°C) and drier (5–15% less precipitation) climate conditions for Balaton watershed at doubled CO_2 concentration for 2100.

Direction of change in weather of different projections is similar, but the size of change may vary in every scenario. The uncertainty seems to be extremely high regarding the precipitation amount and its distribution.

Antal (1974) determined the multi-year average precipitation amount of the entire watershed as 700 mm in 1951–80. Later it was determined more precisely

broken down to sub-basins; according to this, the source region of Zala River has the most precipitation with an annual amount of more than 800 mm, and the "cut" along the lake proved to be the driest with an annual amount of 700 mm. Above the lake this feature is even less: precipitation of about 600 – 650 mm is expected. *Varga et al.* (2006) determined the annual average precipitation of 619 mm based on measurements performed between 1921 and 2005. The differing values can be explained by the differences in the period examined. Analysis of the long-term annual precipitation sums shows negative trend (– 90 mm/100 year) in more than half of the 17 stations of Hungary. The cited precipitation data were homogenized and controlled by the Hungarian Meteorological Service. Without this thorough check-up, the conclusion will not be adequate for later processing.

The aim of this analysis is to provide an overview of the precipitation conditions of the Balaton catchments for the period from the beginning of reliable precipitation gauging (1921) until the end of 2007. These results are more complex, containing more information than the earlier findings regarding the Balaton watershed or the lakeshore alone (*Béll* and *Takács*, 1974; *Bartholy et al.*, 1995). The Balaton watershed and its surroundings is one of the two most vulnerable regions of Hungary. Although the amount and occurrence of precipitation are important measures, our investigation starts with analysis of monthly and annual precipitation sums. Some of the elements of lake water budget influenced by precipitation were also included in the study. We used the simplest statistical method, the linear trend analysis in our work. This procedure results enough information about the spatial and temporal changes in precipitation amounts. The former precipitation observations generally tried to draw a conclusion from shorter time intervals. The change in precipitation that is analyzed can be utilized in the future for local climate change evaluation of lake's watershed or water budget more precisely.

2. Material and methods

The catchment area of Balaton was divided into five sub-catchments (*Table 1* and *Fig. 1*). Besides geographical position of the sub-catchments, their climate was also taken into consideration. Zala River, the main water supplier, enters the lake on the western side of Balaton, providing 55% of the total water input.

The basin of Zala River was divided into two parts due to the difference in the amount of precipitation. Part I is the western sub-basin of the river, the wettest part of the whole region considered, and includes 6 precipitation gauging stations. The eastern sub-basin of the river is closer to the lake and it is warmer and drier than the western sub-basin. The determination of the remaining three sub-basins was performed on the basis of the former classifications. Parts III–V cover the three corners around the lake: the northern, southern, and eastern sub-basins of Lake Balaton, with 5, 6, and 3 precipitation gauging stations,

respectively. Since the warmest and driest sub-catchment of the region was the eastern one (part V) it was separated from the other sub-regions. As the number of gages was limited on the fifth sub-catchment, outside gages were also included in the study. The largest distance of the outside stations should not exceed 10–15 kilometers from the catchment's border. The precipitation of sub-catchment was equal to the average rainfall measured at different precipitation stations.

Table 1. Main information for sub-catchments of Lake Balaton (* based on Climate Atlas of Hungary, 2002)

Name of sub-catchment areas	Number of stations	Area (km^2)	Annual mean temperature* (°C)	Measured precipitation (mm)	Indicator
River Zala, western	6	1678	9.0– 9.5	650–750	I
River Zala, eastern	5	944	9.5–10.5	630–700	II
Balaton, northern	5	820	9.0–10.5	580–660	III
Balaton, southern	6	1272	10.0–10.5	650–700	IV
Balaton, eastern basin	3	36	10.0–10.5	580–620	V

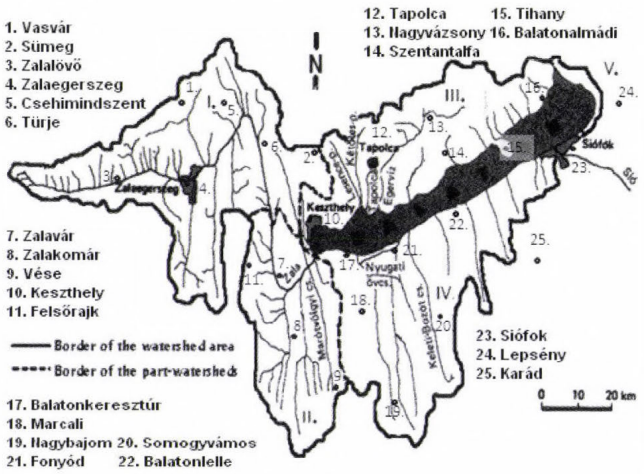


Fig. 1. Catchments of Lake Balaton were divided into five sub-catchments (parts I–V).

The 25 precipitation gauging stations selected for the description of the catchments are part of the precipitation gauging network of the Hungarian Meteorological Service consisting of about 600 stations. At the beginning, gauging stations were equipped exclusively by Hellmann type precipitation gauging units (up to about the middle of the 1990s), that were partly replaced by tipping-bucket automatic devices. The automatic devices were mainly deployed to bigger settlements (Keszthely, Zalaegerszeg, Siófok, Fonyód), and the majority of the stations are still equipped with Hellmann type devices. In the procession,

the monthly and annual amounts were derived from the daily amount of precipitation gaged at 07:00 CET.

Regression analysis of long time series was included in the study. The level of significance referred to in the results was less than 0.05. The critical R^2 value for 87 elements was 0.0446 (linear regression). The statistical evaluation was performed by the free version of STATA 5.0 (1996) program package.

Water budget of Balaton balances the flows in (precipitation, surface water, and groundwater inflows) and out (evaporation, surface and groundwater outflows, surface water withdrawal) of the lake, and the natural change in storage of the lake water (ΔS) is:

$$Inflow - Outflow = \Delta S. \tag{1}$$

In the case of artificially regulated lakes as Lake Balaton, the change in natural water storage is

$$\Delta S = (P + I) - E, \tag{2}$$

where P is the precipitation (mm), I is the surface runoff (mm), and E is the evaporation (mm). Change in water storage represents modification in the lake water level. To characterize S, the lake evaporation (E in mm) was calculated by using the classical Meyer formula between 1921 and 1985. Regarding the relatively large error in evaporation estimates, the so-called Balaton formula was introduced after 1985 (*Antal*, 1963). Partial analysis of evaporation was excluded from the study.

The surface runoff was calculated statistically by using the measured discharge of principle streams. More details can be found in *Varga* (1986).

3. Results and discussion

3.1. The annual amount of precipitation over the Balaton catchments

According to the PRUDENCE results for 2071–2100, regional warming with a mean temperature rise of 3.2–4°C is projected in the Carpathian Basin (*Bartholy et al.*, 2009). It is expected that drying will not be even and will be concentrated in the summertime (25–35% decline in the amount of rainfall). In our investigation we reviewed the changes that have already occurred in the amount of precipitation first; for other experts it can serve as a good starting point for establishing future scenarios. Future precipitation prognosis is excluded from our study.

The regional average of the annual precipitation amount calculated for the entire catchment based on the data of the 25 stations for the period of 1921–2007 is 684.4 ± 111.4 mm. The year with the highest amount of precipitation was 1965 with 986 mm. The driest year of the observed period was 2000, with 458 mm. These two extreme values well demonstrate the eccentric precipitation

regime of the Carpathian Basin, where the ever measured wettest year may gain three times more precipitation than the driest year.

In spite of consistent negative signs in the regression coefficient of annual precipitation sums, the decline was not always significant at every station. We found a significant decrease in annual precipitation sum (that covered several stations) only in the case of the western sub-catchments of Zala River (part I), where a drop in precipitation (1.1–1.7 mm in annual average) was justifiable during the period from the beginning of gauging to date in the case of four out of six stations (*Table 2*). For the whole investigation period it means 95–146 mm precipitation depression depending on the place of the station. The confidence intervals in *Table 2* have got meaning only when the rainfall change is significant. Values of regression coefficient have to fall in the confidence intervals on 95% probability level.

Table 2. Development of the annual sums of precipitation at stations for 1921–2007. Stations showing significant changes in rainfall are in bold (*station excluded from the watershed)

Station	Regression coefficient (mm/year)	R^2	Confidence intervals	
			Lower	Upper
Vasvár	–0.056	0.000	–1.0902	0.9796
Sümeg	**–1.128**	**0.048**	**–2.2087**	**–0.0477**
Zalalövő	**–1.336**	**0.048**	**–2.6217**	**–0.0519**
Zalaegerszeg	**–1.736**	**0.109**	**–2.8066**	**–0.6668**
Csehmindszent	–1.004	0.039	–2.0752	0.0671
Türje	**–1.103**	**0.050**	**–2.1354**	**–0.0692**
Zalavár	–0.942	0.029	–1.9491	0.257
Zalakomár	–0.771	0.018	–1.9884	0.4457
Vése	–1.049	0.028	–2.3673	0.0269
Keszthely	**–1.480**	**0.073**	**–2.6127**	**–0.3471**
Felsőrajk	–1.020	0.033	–2.2161	0.1753
Tapolca	–0.967	0.037	–2.0257	0.0899
Nagyvázsony	**–1.323**	**0.066**	**–2.3976**	**–0.2493**
Szentantalfa	–0.303	0.003	–1.3639	0.7578
Tihany	–0.787	0.028	–1.7765	0.2019
Balatonalmádi	–0.944	0.038	–1.9609	0.072
Balatonkeresztúr	–0.655	0.015	–1.7728	0.4612
Marcali	–0.071	0.000	–1.3556	1.2141
Nagybajom	–0.383	0.004	–1.6584	0.8908
Somogyvámos	–0.867	0.025	–2.0338	0.3009
Fonyód	**–1.385**	**0.054**	**–2.6252**	**–0.1453**
Balatonlelle	–0.020	0.000	–1.0081	0.9651
Siófok	**–1.404**	**0.087**	**–2.3817**	**–0.4276**
Lepsény*	–0.47	0.010	–1.4403	0.4989
Karád*	–0.392	0.005	–1.5305	0.7469

In the case of the other four sub-catchments (though the regression coefficients of the trend lines are negative), we found a significant decrease in the annual amount of precipitation time series at one station from each (Siófok, Fonyód, Nagyvázsony, Keszthely). At the majority of the stations, in the case of three out of four sub-catchments showing changes in precipitation that cannot be statistically justified – in an interesting way – the only gauging location showing significant decrease was at the lakeshore (Keszthely, Fonyód, and Siófok).

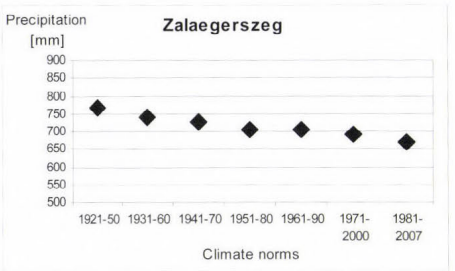
The average rate of precipitation decline computed from the regression coefficients for the whole watershed is 86 mm/100 year. Previous work on precipitation trend of whole Hungary (*Koflanovits-Adámy* and *Szentimrey*, 1986) for the time period of 1901–1984 provided almost the same result (–90 mm/100 year). It is interesting that decline in precipitation of the past two decades did not influenced strongly the regression coefficient. The possible reason might have been the yearly precipitation sums used as a basis in this study.

On the basis of the recommendation of the WMO, we created 10-year running climate norms (*Fig. 2*). In *Fig. 2*, the decreasing precipitation characteristic of the majority of the stations is demonstrated by the data of Zalaegerszeg station located at the centre of the catchments (sub-catchments I) (*Fig. 2a*). All the other sub-figures contain the exceptions. The constant amount of rainfall for Tihany compared to the other stations is remarkable; it is in connection with the geographical location of the gauging station. Tihany is located on the only peninsula pushing out into the lake (*Fig. 2b*). The relatively less yearly amount of rainfall in Tihany is closer to the lower amount of precipitation falling to the lake surface.

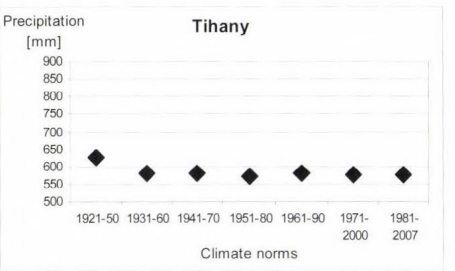
The amount of precipitation measured at the other stations show fairly diversified pictures in time, but at 21 out of 25 stations the period between 1921 and 1970 was wetter, followed by a drying period. The abovementioned tendency did not emerge at four stations; 3 of them are located in the southern river basin of Lake Balaton (Balatonlelle, Marcali, and Nagybajom); here there was a modest increase in precipitation in the period up to 1970, then came the decrease lasting up to the present (*Figs. 2c–e*).

There were only three out of the 25 precipitation gauging stations (Marcali, Zalalövő, and Csehimindszent) where the last two climate norms (the period between 1971 and 2007) were not the two driest ones (*Figs. 2d–g*). Looking for the reasons we did not find any change in location or environmental conditions of the stations. The amount of rainfall slightly increased until the 1960s, and from the 1980s to present a decreased tendency manifested. Our multi-year precipitation data for the largest part of the Lake Balaton catchments correspond well to the studies demonstrating this drying tendency of the last decades (*Bartholy et al.*, 2007b; *Kertész* and *Mika*, 1999; *Nováky*, 1991).

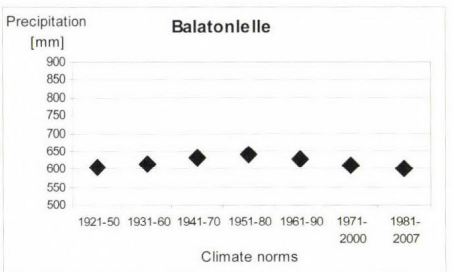
The mean rate of precipitation decline for the whole watershed is 83.4 mm/-100 year (*Fig. 3*). Until now, the decrease was not significant at less than 0.05 significance level.



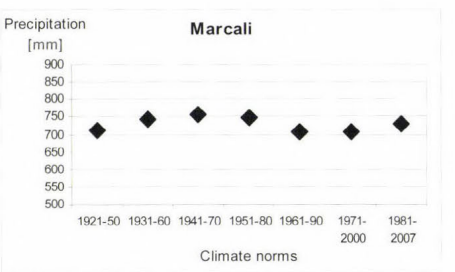
(a)



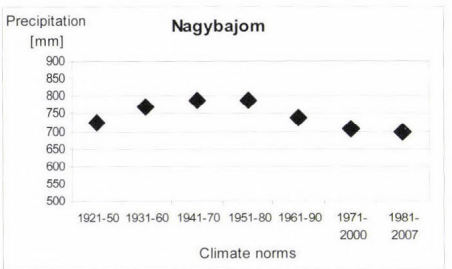
(b)



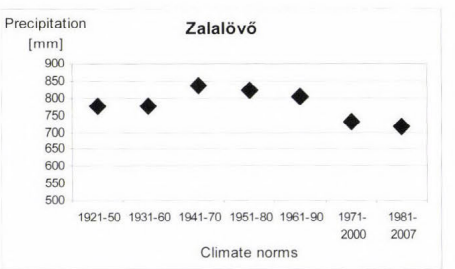
(b)



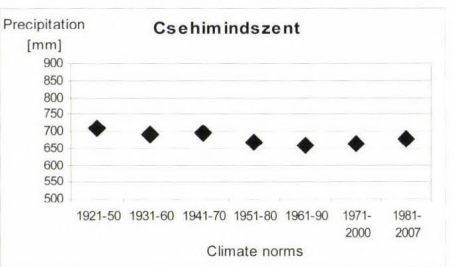
(d)



(e)



(f)



(g)

Fig. 2. Mean annual precipitation sum for 30-year-long time periods at some selected stations of the Balaton catchment area

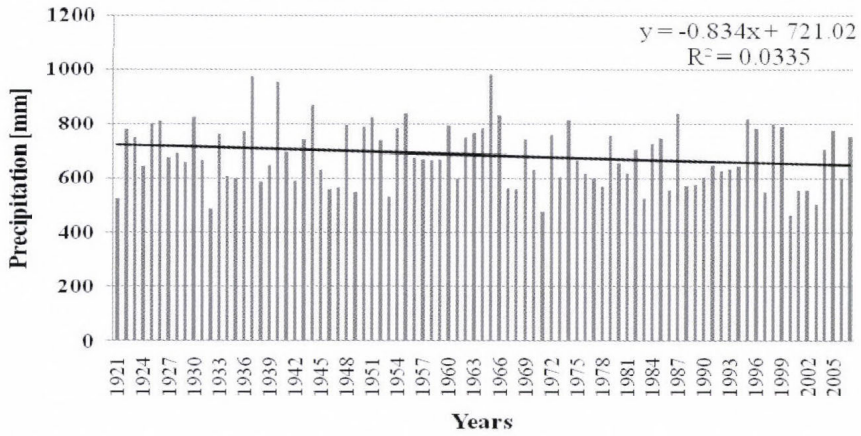


Fig. 3. Variation in yearly precipitation sums for the Balaton catchments (1921–2007)

From the point of view of the development of the lake's water budget, the annual distribution of precipitation is also important. Earlier studies pointed out two maximum points in the annual variation of monthly precipitation sums in the catchments of the lake (*Bartholy et al.*, 2001). The main maximum point is the consequence of wet air inflowing from the Atlantic Ocean at the beginning of summer. The second maximum point is in October-November, owing to the intensifying cyclone activity coming from the Mediterranean basin (*Antal*, 1974). Several studies analyzing Hungary's precipitation conditions found moderate spatial and temporal shift in annual distribution of rainfall sums (*Ambrózy et al.*, 1990). Less than 10% change was observed in the southwestern region of the country, like in the area of the Balaton catchment. *Bartholy et al.* (1995) published that the modification in annual rainfall distribution is one of the local consequences of global warming. Significant decrease in the second peak of precipitation amounts was not confirmed by the precipitation data of the 25 stations located in the catchments of Lake Balaton. On the contrary, in the case of 10 out of the 25 stations (Sümeg, Zalavár, Zalakomár, Vése, Tapolca, Balatonkeresztúr, Marcali, Fonyód, Balatonlelle, Siófok), the ratio of precipitation in August has risen significantly, by 5–10%, compared to the main maximum point in the past 30 years. The uncertainty of the precipitation in the winter months and the higher variability of the amount of precipitation can be experienced. Our observations correspond to the outcomes published to Balaton watershed by *Bartholy et al.* (1995) and *Varga et al.* (2004).

3.2. Water budget elements closely related to precipitation: surface runoff and release through the Sió sluices

Among the two referred elements, the release is intensively affected not only by the rainfall, but by the amount of evaporation. The aim of our investigation was

the study of precipitation and strongly concerned elements as surface runoff and release in the vicinity of Lake Balaton. Knowing the importance of evaporation in lake water budget, we took it into account where it was necessary, but we did not focus on it, discussion of evaporation is excluded from this work.

Watercourses deliver the larger part (about 60%) of the entire lake water acquisition. In the period studied, the annual average surface runoff was 873.9 ± 300 mm (*Fig. 4*) with the maximum of 1974 mm measured in 1965, and the minimum of 293 mm determined both in 2002 and 2003. There were 26 out of 87 years of the whole time series (about one-third) when the surface runoff exceeded 1000 mm, and only two of them were after the year 2000 (dry spell). The average runoff fell back to 501 mm in 2000 and remained on this value until 2004. The moderate decrease in precipitation leads to higher depression in surface runoff. Even in the case of non-significant precipitation changes, 42.7% reduction in runoff seriously affected the lake water level in the beginning of the 2000s. Recovery of the lake water level was late. This is in contrast with other semi-arid regions, where surplus water after drought is mostly used for infiltration and replenishment of groundwater resources, and the surface runoff to the stream-network appears only at a later stage (*Varga et al.*, 2007). The authors mentioned that despite the humid years of 2004 and 2005, the runoff depression still was one third of the average.

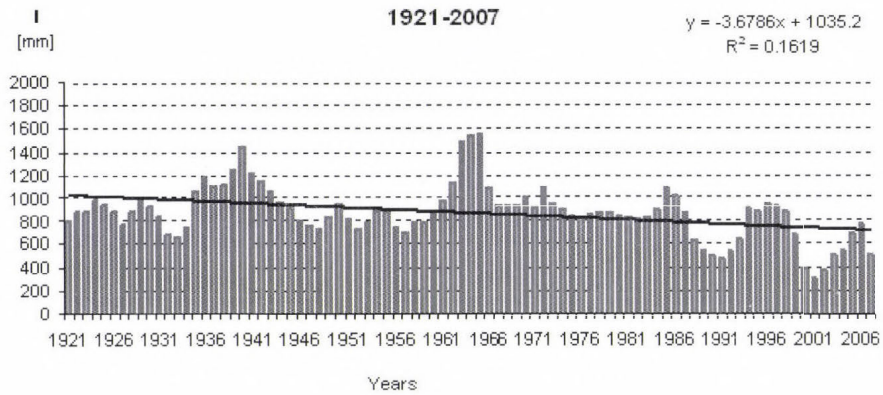


Fig. 4. Annual variation in inflow (*I* mm) for the investigation period (based on data provided by the VITUKI). The change is significant at less than 0.05 significance level.

Signs of change in the annual variation of surface runoff were the same as for precipitation, and water depression was even significant at less than 0.05 significance level. From the beginning of measurements (1921) to the present, the surface runoff has fallen by 315.6 mm (*Fig. 4*). Similar results were published by *Nováky* (1991).

Because of the precipitation pattern and crop cover modifications in Hungary, the runoff has special seasonal variability (*Table 3*).

Table 3. Monthly averages of runoff in the Balaton region. Months having significant change below the 0.05 significance level are in bold

Month	Average runoff (mm/month)	Direction of change	Degree of change (mm/year)
January	81.6	not changed	0.008
February	**99.2**	**decreasing**	**–0.481**
March	121.1	not changed	–0.637
April	96.4	not changed	–0.297
May	**74.5**	**decreasing**	**–0.312**
June	**64.4**	**decreasing**	**–0.411**
July	**54.7**	**decreasing**	**–0.400**
August	44.7	not changed	–0.222
September	43.0	not changed	–0.126
October	49.6	not changed	–0.092
November	66.7	not changed	–0.037
December	78.0	not changed	0.044

The highest values are observed in early spring (February, March, and April). In winter time, there is an increase in the development of surface runoff. Out of the monthly figures of the decrease in water input caused by runoff, the following decreases were significant: at the end of winter in February, then from the end of spring in May, June, and July. The phenomenon experienced in this latter period is doubly dangerous; first, the statistically justified failure in lake water input occurred in summer, in the hottest period with the highest evaporation level; second, the amount of water coming from the catchments and feeding the lake fell back in several subsequent months.

The release of excess water through the Sió sluices is the only water budget member that is a result of controlled human intervention. The aim is to control the lake water level that is regulated by actual law whose borders are not constant in time. The desired level of Lake Balaton is between 0.7 and 1.1 m regulated in 1997 for the last time.

Because of precipitation deficit and surface runoff depression, the quantity of release dropped drastically in the past decade (*Fig. 5*). During the drought of 2000–2003, the Sió sluices were closed, but in 2006 and 2007 they were opened again. Note, that there is an other important affecting factor, the evaporation, but since this is not directly influenced by the precipitation, it is excluded from this study.

Cumulative precipitation and surface runoff shortage were the causes of the change of sign in natural water storage from positive into negative from 2000 to 2003 (*Fig. 6*). During this unique dry period, the water input was not enough to

cover the water necessity of lake evaporation. This negative phenomenon left a mark on ΔS as well. The most variable indicator of Lake Balaton is change in natural water storage, whose average is 592.8 ± 400 mm. Its highest value reached 2031 mm in 1965, and the lowest one was only –180 mm in 2003. The cumulative water deficit of the lake broke the record between 2000 and 2004.

Analysis of the annual variation in ΔS statistically proves a decreasing tendency at less than 0.05 significance level (389.9 mm/100 years). Similarly to the other factors at the input side of the lake water budget, the values of the past two decades falling behind average are worrisome.

Decline in natural change in water storage is more pronounced during summer. In most cases the ΔS increased in winter.

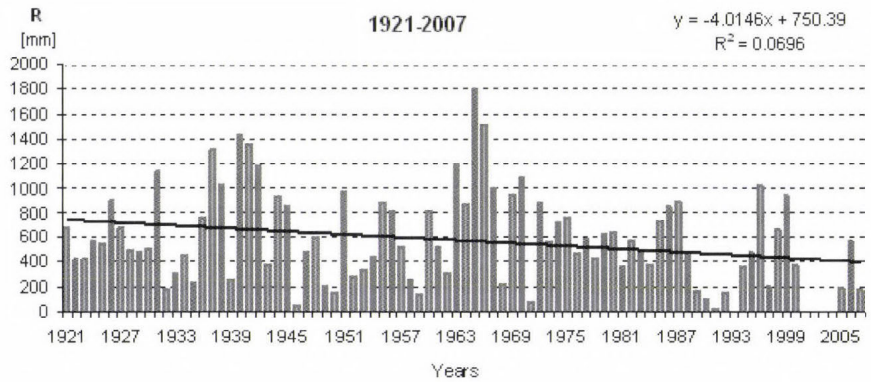


Fig. 5. Annual variation in release (R mm) through the Sió sluices from 1921 until present (based on data provided by the VITUKI).

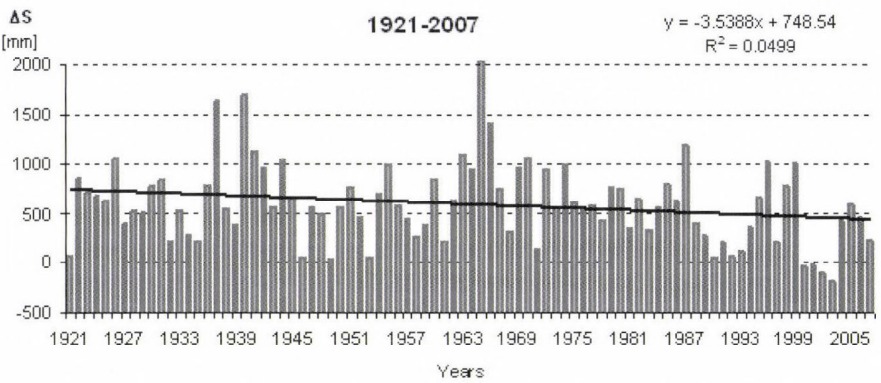


Fig. 6. Natural change in water storage (ΔS mm) from the beginning of the investigation period ((based on data provided by the VITUKI)). The change is significant at less than 0.05 significance level.

4. Conclusions

In the period between 1921 and 2007, the annual precipitation showed a decreasing tendency in the catchments of Lake Balaton, though this change could only be significantly justified in the western part of the basin of Zala River and at several stations close to the lakeshore. At the same time the western side of Zala River is the wettest part of the whole catchments regarding the annual amounts of rainfall. The statistically justifiable decrease in precipitation of the lakeshore stations of the other sub-catchments renders lesser danger, since these contribute to the surface runoff to a smaller extent, and the occurrence of less, and less "disturbing" precipitation is unanimously positive for the bathing holiday-makers.

Distribution of precipitation has not significantly changed at the 25 stations of the lake's catchments. Examining a longer time period, the two maximum points can still be followed. It is a positive phenomenon that a 5–10% rise in the precipitation of August can be experienced. This event is beneficial, since this is the period when the intensified evaporation owing to the higher summer temperature can seriously decrease the water level.

An unfavorable sign considering the water level of the lake is that the runoff has significantly decreased, mainly in the past two decades. In the decades before 1990, the runoff seemed to be a more stable element based on yearly means. The reduction of the lake's water inflow was compensated by the reduction of release in the period after the 1990s. Release through the Sió sluices is the only controlled factor to regulate the water level of Lake Balaton. Similarly to the reduction of inflow, release has also decreased over the past several decades. This human intervention should be implemented with special care, mainly with the knowledge of the reduced amount of inflow during the summer months.

A good example of the local variability caused by global climate change is the annual rainfall sum recorded at various meteorological stations in the watershed of Lake Balaton. The most sensitive region of the watershed is the western part containing the river Zala, where, although the mean drop in annual rainfall quantity since 1921 was only 1.5 mm, the annual decrease in runoff, which is also influenced by rainfall, was more than twice as great. Earlier observations suggest that this is especially dangerous, because when wet years follow dry periods, the infiltration and replenishment of groundwater resources causes a further delay in the runoff (*Varga et al.*, 2007). One positive observation is that, despite the decline in mean annual rainfall over the last 30 years, a 10–15% increase in monthly precipitation has been noted in August at most stations. This could compensate the water supplies of the lake to some extent for the increased evaporation experienced during the summer months.

Greater changes in the rainfall sums recorded at the various stations were observed at stations situated in the immediate vicinity of the lake (Keszthely,

Siófok, Fonyód). This is important, because these are the most frequented regions of the watershed, and thus, they are the most exposed to environmental load. This should be emphasized when compiling long-term land-use plans.

Acknowledgement—We express our gratitude to colleagues of the Central-Transdanubian Environmental and Water Authority (VITUKI), especially *György Varga*, for providing data used in this analysis.

References

Ambrózy, P., Koflanovits, E. and *Kövér, B.*, 1990: Shift in temporal distribution of precipitation amounts in Hungary (in Hungarian). *Időjárás 94*, 156-167.

Antal, E., 1974: Characterisation of the members of heat and water balances in Balaton region. In *The Climate of Balaton* (eds.: *B. Béll* and *L. Takács*) (in Hungarian). Országos Meteorológiai Intézet, Budapest, 188-204.

Antal, E., 1963: Evaporation of Lake Balaton (in Hungarian). *Időjárás 67*, 290-297.

Bartholy, J., Pongrácz, R., Gelybó, Gy. and *Szabó, P.*, 2009: Analysis of expected climate change in the Carpathian Basin using the PRUDENCE results. *Időjárás 112*, 249-265.

Bartholy, J., Pongrácz, R. and *Gelybó, Gy.*, 2007a: Regional climate change expected in Hungary 2071–2100. *Applied Ecol and Environ Res 5*, 1-17.

Bartholy, J. and *Pongrácz, R.*, 2007b: Regional analysis of extreme temperature and precipitation indices for the Carpathian Basin from 1946 to 2001. *Global Planet Change 57*, 83-95.

Bartholy, J., Matyasovszky, I. and *Weidinger, T.*, 2001: Regional climate change in Hungary: a survey and stohastic downscaling method. *Időjárás 105*, 1-17.

Bartholy, J., Bogárdi, I. and *Matyasovszky, I.*, 1995: Effects of climate change on regional precipitation in Lake Balaton watershed. *Theor Appl Climatol 51*, 237-250.

Climate Atlas of Hungary (in Hungarian), 2002, Országos Meteorológiai Szolgálat, Budapest.

Gianniou, S.K. and *Antonopoulos, V.Z.*, 2007: Evaporation and energy budget in Lake Vegoritis, Greece. *J Hydrol 345*, 212-223.

Greenland, D. and *Kittel, T.G.F.*, 2002: Temporal variability of climate at US Long-Term Ecological Research (LTER) sites. *Climat Res 19*, 213-231.

Kertész, A. and *Mika, J.*, 1999: Climate Change in South-Eastern Europe. *Phys Chem Earth (A) 24*, 913-920.

Koflanovits-Adámy, E. and *Szentimrey, T.*, 1986: The variations of precipitation amounts in the Carpathian Basin during the present century (in Hungarian). *Időjárás 90*, 206-216.

Lenters, J.D., Kratz, T.K. and *Bowser, C.J.*, 2005: Effects of climate variability on lake evaporation: Result from a long term energy budget study of Sparkling Lake, northern Wisconsin (USA). *J Hydrol 308*, 168-195.

Mercier, F., Cazenave, A. and *Maheu, C.*, 2002: Interannual lake level fluctuations (1993–1999) in Africa from Topex/Poseidon connections with ocean-atmosphere interactions over the Indian Ocean. *Global Planet Change 32*, 141-163.

Nováky, B., 1991: Climate effects on runoff conditions in Hungary. *Earth Surface and Landforms 16*, 593-601.

Schindler, D.W., 2001: The cumulative effects of climate warming and other human stresses on Canadian freshwaters in the new millenium. *Can J Fish Aquat Sci 58*, 18-29.

Somlyódi, L. and *Honti, M.*, 2005: The water substitution of Lake Balaton (in Hungarian). *Magyar Tudomány 5*, 570-581.

STATA 5.0, 1996: Stata Corporation LP Texas, USA, www.stata.com

Varga, Gy., 1986: Developing calculation of water budget of Lake Balaton (in Hungarian). *Report on Research carried out on Lake Balaton by VITUKI*, Budapest.

Varga, Gy., Papp, U.J. and *Bálint, G.*, 2007: Recent extremes in the water budget of Lake Balaton. *Georgikon for Agric 10*, 25-40.

Varga, B., Boldizsár, A. and *Gimesi, L.*, 2006: Some remarks on water balance components of Lake Balaton. *Georgikon for Agric 10*, 69-82.

Varga, Gy., Papp, U.J., Mika, J., Pálfy, L. and *Bálint, G.*, 2004: The 2000–2003 extremes in the water budget of Lake Balaton. *Geophys Res Abstract Vol. 6 07795*, SRef-ID: 1607-7962/gra/EGU04-A-07795.

Winter, T.C., Buso, D.C., Rosenberry, D.O., Likens, G.E., Sturrock, A.M. and *Mau, D.P.*, 2003: Evaporation determined by the energy budget method for Mirror Lake, New Hampshire. *Limnol Oceanogr 48*, 995-1009.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 3, July–September 2010, pp. 203–215

Surface resistance estimation of some crops using different climate, soil-, and vegetation-specific data

Hajnalka Breuer and **Ferenc Ács**

Department of Meteorology, Eötvös Loránd University
P.O. Box 32, H-1518 Budapest, Hungary;
E-mails: bhajni@nimbus.elte.hu, acs@caesar.elte.hu

(Manuscript received in final form May 25, 2010)

Abstract — Annual course of different crop surface resistance (r_c) parameters at Szarvas was analyzed. In the analysis, a set of independent climatological, satellite, soil- and vegetation-specific data was used. Soil- and plant-specific data were taken from the literature. Of plant-specific data, the plant coefficient data determined by evapotranspirometer measurements were used. The soil texture at Szarvas is clay. The estimation method of r_c is based on calculating the ratio between the actual and potential evapotranspiration, which is a proven method for climatological applications.

It is found that r_c is mainly governed by the changes of the available water in the soil. It is also found that r_c variations induced by variations of the available water in the soil are much larger than r_c variations induced by plant species changes represented by plant coefficients. The analysis showed that some results obtained in the faraway past can also be successfully used for "state of the art" biophysical modeling purposes in the meteorology.

Key-words: climate data, soil data, evapotranspirometer measurements, plant coefficients, actual and potential evapotranspiration, crops, surface resistance

1. Introduction

The first evapotranspirometer measurements in Hungary started in the 1960s. Daily potential evapotranspiration was measured using Thornthwaite-Mather evapotranspirometers (*Thornthwaite* and *Mather*, 1955) above different vegetation types like short grass, potato, apple, winter wheat, and corn to work out irrigation scheduling for each vegetation type separately. The first description and user guide of this apparatus in Hungary was given by *Antal* (1966). These measurements served agrometeorological goals. It should be mentioned that Thornthwaite's scientific influence in Hungary was strong; this could be seen not only via using Thornthwaite's type evapotranspirometer (e.g., *Antal*, 1968;

Posza and *Stollár*, 1983) but also via application of Thornthwaite's climate classification system (*Szesztay*, 1958; *Kakas*, 1960).

In about the same period, the first mesoscale modeling systems were applied (*Deardorff*, 1977, 1978). These models contain complex biophysical modules for simulating exchange processes at the interface between the land surface and the atmosphere. The exchange processes could be described only by characterizing albedo, roughness length, and the surface resistance of the soil-vegetation system. Albedo is the most important radiation transfer parameter. Roughness length parameter determines strongly the intensity of the turbulent transfer. Surface resistance (r_c) is a basic parameter for characterizing water transfer in the soil-vegetation system which was introduced by *Monteith* (1965) analyzing the process of evapotranspiration from vegetated surfaces. Of the three parameters mentioned, the latter one is the most changeable, so the uncertainties related to it are also the largest.

In this study, we will focus on the estimation of r_c of some chosen crops which are common in Hungary. r_c is estimated using a common climatological method based on calculation of the ratio between the actual and potential evapotranspiration. The aim of this study is to determine and analyze the annual course of r_c of some chosen crops at Szarvas using the above mentioned method. The analysis was performed using plant coefficient data obtained from evapotranspirometer measurement results (*Antal*, 1968; *Posza* and *Stollár*, 1983) and some other needed environmental data (*Mitchell et al.*, 2004; *Bartholy et al.*, 2004). All data refer to Szarvas, but they are representative for different time periods. This shortcoming is less important in the case of climatological analysis. The novelty of the study does not lie in the methodology applied but rather in the combination of the data which will be presented below. It should be mentioned that, to the best of our knowledge, there is not any study in Hungary in which r_c would have been analyzed from either meteorological or climatological point of view.

2. Data

Climate and soil data are used in this study. Climate data can be classified into three groups: meteorological data, satellite data, and in-situ vegetation specific data obtained by using evapotranspirometer measurements.

2.1. Meteorological and satellite data

Mean monthly precipitation (P), air temperature (T), and diurnal temperature range data constitute meteorological data. They were taken from the CRU (**C**limate **R**esearch **U**nit), which is one of the well-known climate data centers. The CRU TS 1.2 (*Mitchell et al.*, 2004) database contains four observed state variables and one parameter (2 m air temperature, diurnal temperature range,

precipitation, cloud cover, and vapor pressure) interpolated to a 10' (≈ 11 km) resolution grid over Europe from 1901 to 2000 on a monthly scale. The interpolation was made using the measurements of meteorological stations in Europe. Naturally the accuracy of the fields is increasing in time, as the number of measurements is increasing. The spatial resolution is high enough to be taken as a reference measurement field used for verification and validation of climate models. From this database the nearest point to Szarvas was taken, which is only 0.02° away.

Satellite measurement of ten-day composite *NDVI* (normalized difference vegetation index) data (*Bartholy et al.*, 2004) is freely available from 1982 up to 2000 at a 8 km horizontal resolution from the NASA Goddard Space Flight Center. Naturally, we could not acquire such data to 1960 as there were not any kind of measurements of *NDVI*. For our purposes, we calculated the average *NDVI* of the 1982–2000 period for Szarvas. These data together with *P* and *T* are presented in *Table 1*.

Table 1. Monthly mean values of *NDVI*, *T*, and *P* referring to period 1982–2000 as used for Szarvas

Month	*NDVI*	*T* (°C)	*P* (mm)
January	0.147	–0.866	27.544
February	0.160	1.097	30.823
March	0.231	5.789	24.359
April	0.412	11.995	47.269
May	0.547	17.038	58.081
June	0.507	20.055	64.794
July	0.441	22.005	49.754
August	0.401	21.745	45.933
September	0.346	17.395	43.840
October	0.298	11.986	35.553
November	0.262	5.506	43.594
December	0.157	0.714	41.877

2.2. Plant coefficients

In this study, five-day average plant coefficient values (*b*), published by *Posza* and *Stollár* (1983), were used. Seasonal changes of the values of *b* for different plant species are presented in *Fig. 1*. The values were obtained by using measurement results of Thornthwaite-Mather's circular evapotranspirometers (surface equals to 4 m^2, depth is 70 cm), in which the level of water was kept 40 cm below the surface by an automatic level control. The measurements were performed at different locations in Hungary, but mostly in Szarvas. The time interval of the measurements was different for different vegetation types. It is to

be noted that important details concerning the measurements were described by *Antal* (1968) and *Posza* and *Stollár* (1983). Inspecting *Fig. 1*, some interesting features could be noted. Values of the *b* coefficient referring to grass show a step-like change; this is caused by its systematic cutting. It could also be seen that the growing season of winter wheat lasts until the middle of July. This could be true, but only under measurement conditions which insured always sufficient water. So, values of *b* for winter wheat in July do not reflect the real state of winter wheat which prevails on the ploughlands. The same is valid for the *b* values of apple at the end of October. It is obvious that the *b* values obtained refer to the conditions when the water supply is good.

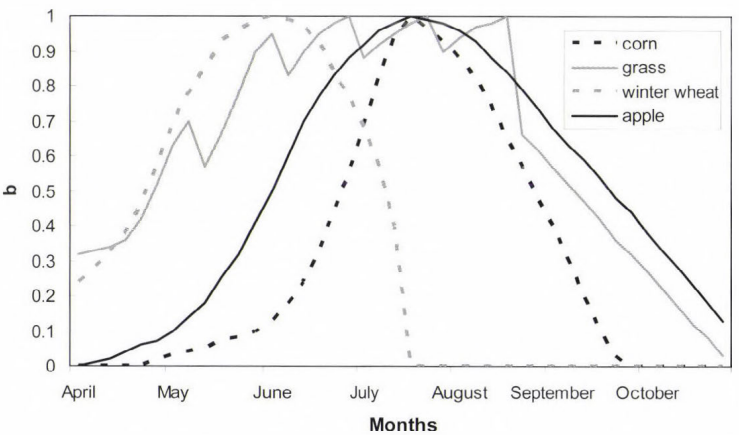


Fig. 1. Seasonal changes of plant coefficients for different plant species as reported by *Posza* and *Stollár* (1983). Monthly values are obtained by using five-day averages.

2.3. Soil data

For calculating actual evapotranspiration, soil data are also needed. At Szarvas, clay is the prevailing soil texture (*Várallyay et al.*, 1980). The most important soil hydraulic parameters of this soil texture are presented in *Table 2*. The values presented were obtained using the results of *Nemes* (2003) and *Fodor* and *Rajkai* (2005).

Table 2. Soil hydrophysical properties used in calculation (saturated, field capacity, and wilting point soil water contents, Θ_s, Θ_f, Θ_w, respectively; Clapp-Hornberger pore size index, b_{CH}; saturated soil water conductivity, K_s)

Θ_s (mm m^{-1})	Θ_f (mm m^{-1})	Θ_w (mm m^{-1})	b_{CH}	K_s (m s^{-1})
527.94	469.55	140.24	6.15	$8 \cdot 10^{-7}$

3. Method

3.1. Basic equation

The ratio of actual (E) and potential (E_p) evapotranspiration (often noted as β) using Penman-Montieth's equation (*Monteith*, 1965) can be expressed as

$$\beta = \frac{E}{E_p} = \frac{\Delta + \gamma}{\Delta + \gamma(1 + r_c / r_a)}, \tag{1}$$

where Δ is the slope of the saturated vapor pressure versus air temperature T, γ is the psychrometric constant (0.65 hPa K^{-1} value is used in this study), r_c is the surface resistance of the soil-vegetation system, and r_a is the aerodynamic resistance characterizing turbulent transfer above the soil-vegetation system.

β can also be expressed by using an empirical formula. *Antal* (1968) presents a formula based on Thornthwaite-Mather's evapotranspirometer measurements:

$$\beta_{hun} = \frac{F_{ma} + b}{1 + b} F_{ma}, \tag{2}$$

where $F_{ma} = (\Theta - \Theta_w)/(\Theta_f - \Theta_w)$, Θ is the actual soil moisture content [$m^3\ m^{-3}$], Θ_f is the field capacity soil moisture content [$m^3\ m^{-3}$], Θ_w is the wilting point soil moisture content [$m^3\ m^{-3}$], and b is the plant coefficient (*Posza* and *Stollár*, 1983). Since the b values obtained refer to the conditions when the water supply is good the formula of *Antal* (1968) is more suitable for describing evapotranspiration in wet than in dry conditions.

Equating Eqs. (1) and (2), it is easy to obtain an expression for r_c:

$$r_c = \frac{1 - \beta_{hun}}{\beta_{hun}} \frac{\Delta + \gamma}{\gamma} r_a. \tag{3}$$

According to *Szeicz* and *Long* (1969), Eq. (3) is suitable for estimating monthly values of r_c in the growing season when monthly values of r_a are available. This opinion can be accepted knowing that the relationship between the stomata and the near surface atmosphere is the strongest when unstable stratification is prevailing (*Jarvis* and *McNaughton*, 1986; *Ács*, 2008; *Breuer*, 2009). Since during evapotranspirometer measurements there were no measurements for determining r_a, the range of the changes of r_a has to be estimated after some assumptions regarding climatic (wind) conditions and vegetation morphology (roughness length) characteristics. We supposed that the wind speed varies mainly between 1–4 $m\ s^{-1}$ in the growing season and that the roughness length

of the investigated plant species changes between 0–0.4 m. For above mentioned conditions, the typical values of r_a in unstable stratification vary between about 5–50 s m^{-1}. The lowest values refer to apple (large roughness), while the highest values refer to grass (small roughness). For stable stratification, the r_c *versus* r_a relationship is weak and it becomes always weaker with the growing atmospheric stability (*Ács*, 2008), therefore, Eq. (3) cannot be applied for these conditions at all. It is to be mentioned, that these considerations refer only to vegetation, namely, the bare soil surface resistance does not depend upon r_a at all.

3.2. Estimating of soil moisture content and leaf area index

For estimating r_c, it is necessary to calculate Θ. Θ was calculated by using a 1 m deep bucket type model. The potential evapotranspiration (*PET*) was calculated after *Turc* (1961). The radiation in Turc's formula was estimated by using Hargreaves expression (*Hargreaves* and *Samani*, 1982). Using F_{ma}, the partitioning of the *PET* to the actual evapotranspiration (*AET*) was calculated. The model estimates water deficiency and surplus from *P* and *AET*. An estimation of gravitational runoff (*Clapp* and *Hornberger*, 1978) was also incorporated.

For estimating the minimum stomatal resistance parameter (r_{stmin}), it is necessary to know the leaf area index (*LAI*) values. There are many methods for estimating *LAI*; we used a formula which is based on normalized difference vegetation index (*NDVI*) estimations. According to *Zhangshi* and *Williams* (1997),

$$LAI_i = LAI_{\max} \frac{NDVI_i - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}}, \tag{4}$$

where max, min, and *i* subscripts denote maximum, minimum, and actual values, respectively. The *NDVI* values represent observed values. The formula assumes that *NDVI*/*LAI* relationship is linear (*Wiegand et al.*, 1979; *Wardley* and *Curran*, 1984); and that the maximum *NDVI* value in a season correspond to the maximum *LAI* of vegetation cover (*Justice*, 1986). This latter fact does not mean that $LAI_{\max}$ has to be determined by all means from $NDVI_{\max}$; this could also be achieved using independent sources.

4. Results

The results obtained will be analyzed in three steps. First, the β_{hun} values shall be considered. After this, an analysis of r_c will be given. Lastly, we shall try to give an estimate for r_{stmin} because of its importance in the mesoscale models. We will show that this is possible only if enough accurate *LAI* values are available.

4.1. Monthly values of β_{hun}

β_{hun} represents the ratio between the actual and potential evapotranspiration. It is parameterized after *Antal* (1968), (Eq. 2); it depends upon F_{ma} and b. Note that F_{ma} is estimated using a bucket type model, while b is taken from *Posza* and *Stollár* (1983). When no b values were available, for instance, in the time period after growing season, then the b values used were put to zero. Seasonal changes of β_{hun} are presented in *Fig. 2*. β_{hun} is plotted as the plant species mean (β^{mean}_{hun}) and for all plant species separately. It has to be said that β_{hun} refers to Szarvas though b values of some plant species (for instance apple) were determined by performing evapotranspirometer measurements outside Szarvas. It is essential, that the course of F_{ma} and β_{hun} is very similar. There is a maximum in March and a minimum in September or October. The variability of β_{hun} according to plant species is observable. Nevertheless, the largest β_{hun} differences caused by b differences are as large as the standard deviation of β^{mean}_{hun}. This difference is the largest between the apple and winter wheat in the middle of August. It should be underlined that β_{hun} variations caused by b differences are equivocally smaller than those β_{hun} changes which were caused by F_{ma} changes.

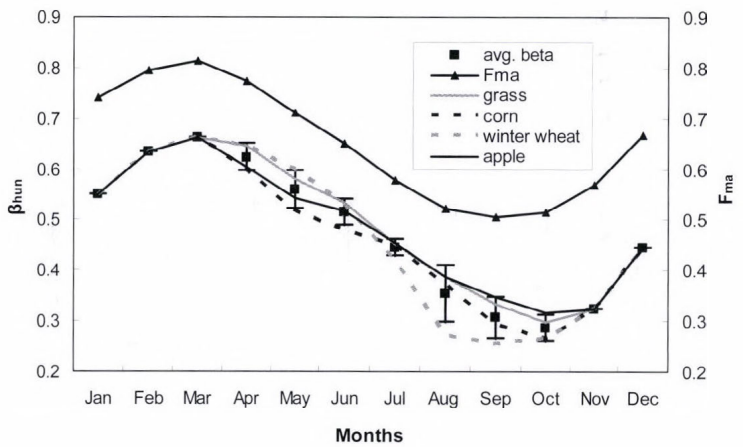


Fig. 2. Annual course of relative water-holding capacity (F_{ma}) and the ratio between the actual and potential evapotranspiration (β_{hun}) for different plant species.

4.2. Monthly values of the land-surface resistance

Land-surface resistance (r_c) is a weighted mean of soil and vegetation canopy resistances. For closed vegetation canopy, r_c converges to the canopy resistance. According to the model (Eq. 3), r_c depends upon β_{hun}, air temperature (via Δ), and atmospheric stratification (via r_a). Note that there is a linear dependence between the r_c and r_a, though in reality r_c is weakly and indirectly dependent

upon atmospheric stratification. This could be checked by using those models, where r_c is parameterized as a function of leaf water potential (*Ács*, 1994; *Ács* and *Hantel*, 1998). Despite its former weakness, Eq. (3) could be successfully used for determining r_c if it serves climatological purposes, for instance, when a long series of evapotranspiration measurements are available (*Szeicz* and *Long*, 1969). Dependence between r_c and r_a for different plant species and months is presented in *Fig. 3*.

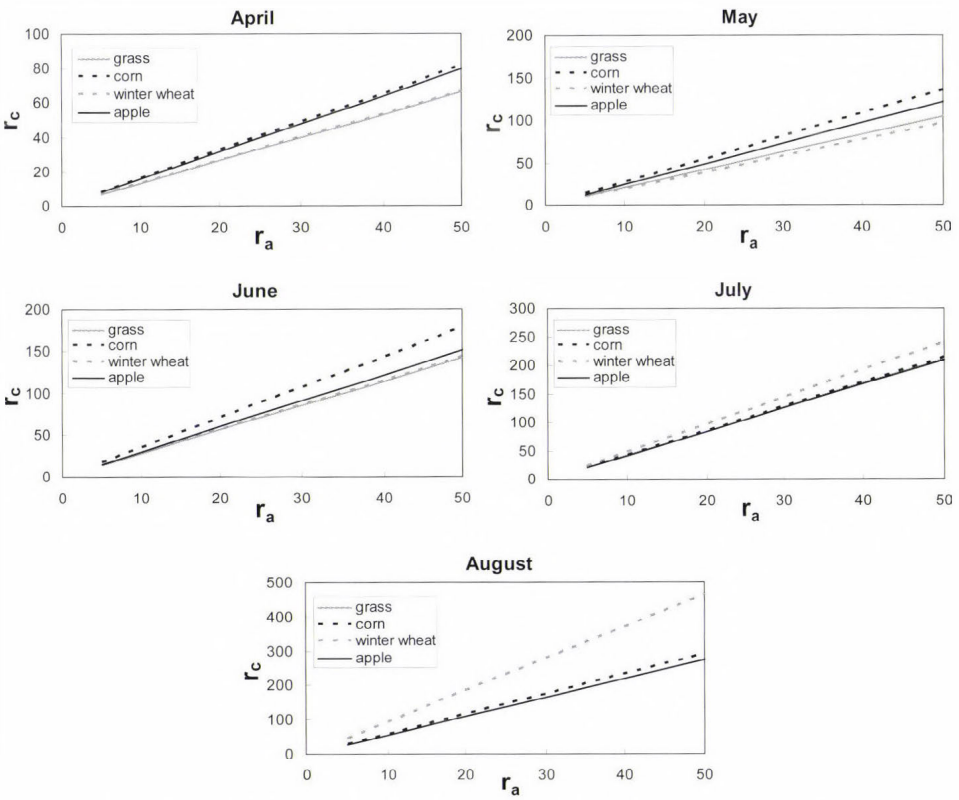


Fig. 3. Dependence between the surface resistance of the land-surface and the aerodynamic resistance for different plant species and months (for the period of April–August).

Note that r_a is estimated as it was described in Section 3, while Δ and β_{hun} were calculated from climate, plant coefficient, and soil data. Two basic features can be observed: first, there is a variability according to the plant species; this is easy to observe, for instance, in May and August, and second, by decreasing r_a values, the obtained plant-specific lines converge to each other into a fix and low r_c value. That is, according to this model, r_c is also weakly dependent on r_a, when the turbulent mixing is large. Annual course of r_c for $r_a = 5$ s m^{-1}, when

turbulent mixing is large enough for all vegetation types considered, is presented in *Fig. 4*. Note that r_c reaches its minimum value when β_{hun} has a maximum, and vice versa, r_c has its maximum value when β_{hun} has a minimum. Furthermore, r_c is mainly governed by F_{ma}, that is by the actual soil moisture content. This is in accordance with the results obtained by *Alfieri et al.* (2008). Plant specific variations of r_c could be observed in May, June, August, September, and October. It is obvious that r_c variations induced by changes of F_{ma} are much larger than the ones induced by plant species changes of b at least in the investigated cases. b-induced differences in May and June are smaller than those obtained in August, September, and October. Note that until June the values of r_c are smaller than 40 s m^{-1}. The high values of r_c for winter wheat in the period after its harvest (August, September, October) refer to bare soil conditions (arable land). It has to be said that our results are in agreement with the results obtained by *Szeicz* and *Long* (1969) as well as *Löpmeier* (1983). *Szeicz* and *Long* (1969) used different methods, among them also the method applied by us, and investigated land-surfaces covered by barley crop and tall grasses. *Löpmeier* (1983) applied lysimeter measurements investigating spring wheat and rye crops.

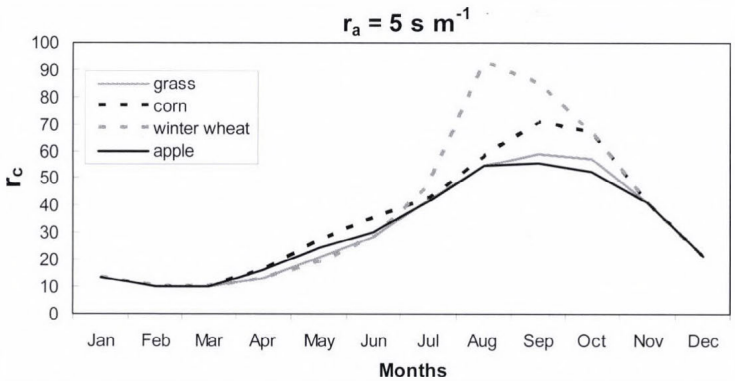


Fig. 4. Annual course of the surface resistance of the land-surface covered by different plant species when $r_a = 5$ s m^{-1}.

4.3. Minimum stomatal resistance values

To estimate r_{stmin}, we shall use the big leaf model of *Jarvis* (1976):

$$r_c = \frac{r_{st\min} \cdot F_{ad}}{LAI \cdot GLF \cdot F_{ma}}, \tag{5}$$

where F_{ad} is the relative stomatal conductance expressing the atmospheric (global radiation, air temperature, and humidity) impact upon stomatal

functioning and *GLF* is the green leaf fraction. Knowing r_c and F_{ma} and making some assumptions for *LAI*, *GLF*, and F_{ad}, r_{stmin} can be inferred. In doing so, first of all, we shall suppose a closed vegetation canopy (this is fulfilled when *LAI* is larger or about 4 $m^2 m^{-2}$), and that *GLF* and F_{ad} equal 1. The latter assumption is quite crude, nevertheless, there are examples in the literature for its use (e.g., *De Ridder* and *Schayes*, 1997). In this case, the stomata are open at a maximum rate (*Ács*, 2008), that is, there is no atmospheric regulation of stomatal functioning. The former assumption that $GLF = 1$ is not rare, this is surely fulfilled at the beginning and in the middle of the growing season. Since direct observations of *LAI* were not available, the course of *LAI* is estimated using Eq. (4). *NDVI* values are taken from *Bartholy et al.* (2004). The LAI_{max} values are chosen after *Masson et al.* (2003). They were 2 and 4 $m^2 m^{-2}$ for tall grass and other vegetation cultures, respectively. These courses of *LAI* for tall grass and other plant species are presented in *Fig. 5*. Since *LAI* is about 4 $m^2 m^{-2}$ only in May and June, r_{stmin} shall be estimated only for these two months.

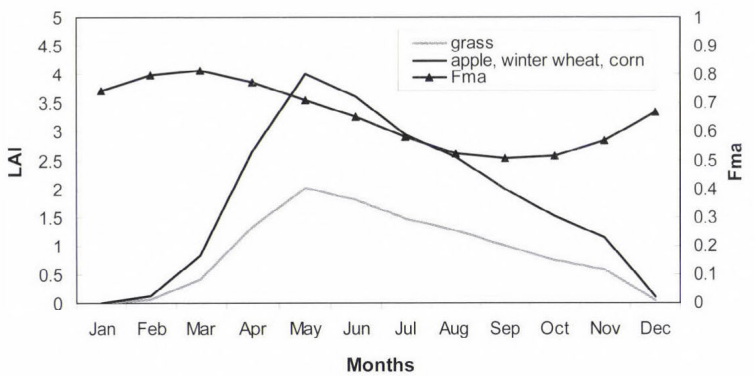


Fig. 5. Seasonal changes of leaf area index for different plant species

The estimated r_{stmin} values for $r_a = 5$ s m^{-1} are presented in *Table 3*. Note that r_c does not depend upon low r_a values. Since r_c and F_{ma} are about equal for both months, r_{stmin} values differ only because of the different *LAI* values. For tall grass r_{stmin} is about 15 s m^{-1}, while for other plant species it is about 35 s m^{-1}. According to *Alfieri et al.* (2008), the r_{stmin} value of grass seems to be too low. They obtained a value of for r_{stmin} about 90 s m^{-1} with a standard deviation of about 40 s m^{-1}. Nevertheless, their grass is wild, while our grass is cultivated. It is well known that r_{stmin} values of the cultivated species are much lower than the r_{stmin} values of the same wild species. r_{stmin} value of 35 s m^{-1} estimated for other plant species is in accordance with the r_{stmin} value obtained by *Alfieri et al.* (2008) referring to croplands. Their suggested r_{stmin} value for croplands is about 20–30 s m^{-1} which is close to our value. It has to be noted that the estimations of *Alfieri et al.* (2008) are based on the field measurements and on the use of the

Jarvis' formula. They also showed that r_{stmin} values are highly changeable in the space and time, which is not a reassuring fact from the point of view of the mesoscale meteorological models which apply Jarvis' formula.

Table 3. Estimated values of canopy vegetation and minimum stomatal resistances for grass and other cultivated vegetation (*other* means the average value of apple, winter wheat, and corn) in the months May and June. F_{ma} indicates the relative stomatal conductance caused by soil moisture deficit and *LAI* is the leaf area index

Month	F_{ma}	r_c		*LAI*		r_{stmin}	
		grass	other	grass	other	grass	other
May	0.71	10.44	11.78	2.0	4.0	14.85	33.53
June	0.65	14.18	15.69	1.8	3.6	16.61	36.76

5. Conclusions

In biophysical modeling, the resistance/flux relationship is usually used in a "forward mode" when flux has to be calculated from the known resistance. In contrast to the former case, there is also a "backward mode" calculation when resistance has to be estimated supposing that the flux is known. In this study, a "backward mode" treatment for estimating monthly mean values of crop surface resistance is presented. r_c is calculated estimating the ratio of actual and potential evapotranspiration, β_{hun}, from simulated F_{ma} and observed b values (Eq. 2). r_c is estimated only for small r_a values (strong unstable stratification), when the relationship between the stomata and the near surface atmosphere is strong. In this case r_c is only weakly dependent upon r_a. To the best of our knowledge, this study is the first "backward mode" type study in Hungary in the subject of biophysical modeling.

To apply the method, a set of meteorological, satellite, soil- and vegetation-specific data was used. Meteorological data are taken from CRU, satellite data are obtained from *Bartholy et al.* (2004), location-specific soil data are determined by using the studies of *Nemes* (2003) and *Fodor* and *Rajkai* (2005), while vegetation-specific data are taken from *Posza* and *Stollár* (1983). These latter data are obtained from evapotranspirometer measurements performed at different locations but mostly in Szarvas. The novelty of the analysis does not lie in the methodology applied, but rather in the combination and application of the relevant input data mentioned above. It is shown that r_c is mainly governed by F_{ma}, that is, by the amount of water available for evapotranspiration. It is also shown that r_c variations induced by changes of F_{ma} are much larger than those ones induced by plant species changes represented by b. The results obtained for r_c are in accordance with results published in the scientific literature (e.g., *Szeicz* and *Long*, 1969). We tried to estimate r_{stmin} as well, by using the formula of *Jarvis* (1976) and applying different more or less crude assumptions. The results

obtained in this way are in a quite good agreement with the results obtained by *Alfieri et al.* (2008). This analysis revealed also that r_{stmin} is impossible to estimate without precise knowledge of *LAI*.

The analysis used some "old results" obtained from evapotranspirometer measurements for purposes of biophysical modeling. The opportunity of such an application has been written by *Kozmáné Tóth* and *Hunkár* (1995). The results obtained are instructive in the field of land-surface/atmosphere interaction modeling.

Acknowledgements—This work was motivated by the inspiration that we got by investigating evapotranspirometer measurements done by *Emánuel Antal*, *István Posza*, *András Stollár* and others. Thank you for this motivation.

References

Alfieri, J.G., Niyogi, D., Blanken, P.D., Chen, F., LeMone, M., Mitchell, K.E., Ek, M.B. and *Kumar, A.*, 2008: Estimation of the minimum canopy resistance for croplands and grasslands using data from the 2002 international H_2O project. *Mon Weather Rev 136*, 4452-4469.

Antal, E., 1966: Potential evapotranspiration of some agricultural plant species (in Hungarian). *Öntözéses gazdálkodás 4*(1), 69-86.

Antal, E., 1968: Irrigation schedule on the basis of meteorological data (in Hungarian). *CSc Thesis*, Magyar Tudományos Akadémia, Budapest, 147 pp.

Ács, F., 1994: A coupled soil-vegetation scheme: Description, parameters, validation and sensitivity studies. *J Appl Meteorol 33*, 268-284.

Ács, F. and *Hantel, M.*, 1998: The land-surface flux model PROGSURF. *Global Planet Change 19*, 19-34.

Ács, F., 2008: Meteorological modeling of the soil-vegetation system processes. Applications in science and in education (in Hungarian). *ELTE Eötvös Kiadó*, Budapest, 249 pp.

Bartholy, J., Pongrácz, R., Barcza, Z. and *Dezső, Zs.*, 2004: Aspects of urban/rural population migration in the Carpathian basin using satellite imagery. In *Environmental Change and Its Implications for Population Migration* (eds.: *J.D. Uruh, M.S. Krol* and *N. Kliot*). Kluwer Academic Publishers, 289-313.

Breuer, H., 2009: The strength of the coupling between the vegetation and the atmosphere (in Hungarian). *Légkör 54*(3), 8-11.

Clapp, R.B. and *Hornberger, G.M.*, 1978: Empirical equations for some hydraulic properties. *Water Resour Res 14*, 601-604.

Deardorff, J.W., 1977: A parameterization of the ground surface moisture content for use in atmospheric prediction models. *J Appl Meteorol 16*, 1182-1185.

Deardorff, J.W., 1978: Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation. *J Geophys Res 83*, 1889-1903.

De Ridder, K. and *Schayes, G.*, 1997: The IAGL Land Surface Model. *J Appl Meteorol 36*, 167-183.

Fodor, N. and *Rajkai, K.*, 2005: Estimation of physical soil properties and their use in models (in Hungarian). *Agrokem Talajtan 54*, 25-40.

Hargreaves, G.H. and *Samani, Z.A.*, 1982: Estimating potential evapotranspiration. *J Irrig Drain E-ASCE 108*, 225-230.

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.

Jarvis, P.G. and *McNaughton, K.G.*, 1986: Stomatal Control of Transpiration: Scaling Up from Leaf to Region. *Adv Ecol Res 15*, 1-49.

Justice, C.O., 1986. Monitoring East African Vegetation using AVHRR data. *Int J Remote Sens 6*, 1335-1372.

Kakas, J., 1960: Annual potential evapotranspiration. Annual water excess. Annual water deficit (in Hungarian). Magyarország Éghajlati Atlasza. *Akadémiai Kiadó*, Budapest.

Kozmáné Tóth, E. and *Hunkár, M.*, 1995: Agrometeorology (in Hungarian). In Fejezetek a magyar meteorológia történetéből 1971-1995 (eds.: *A. Simon* and *T. Tänczer*). Országos Meteorológiai Szolgálat, Budapest, 235-257.

Löpmeier, F.-J., 1983: Agrarmeteorologisches Modell zur Berechnung der aktuellen Verdunstung (AMBAV). *Beiträge zur Agrarmeteorologie 7/83*. Deutscher Wetterdienst, Offenbach.

Masson, V., Champeaux, J-L., Chauvin, F., Meriguet, C., and *Lacaze, R.*, 2003: A global database of land surface parameters at 1-km resolution in meteorological and climate models. *J. Climate 16*, 1261-1282.

Mitchell, T., Carter, T.R., Jones, P., and *Hulme, M.*, 2004: A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: The observed record (1901-2000) and 16 scenarios (2001-2100). *Tyndall Centre Working Paper* 55, 30 p.

Monteith, J.L., 1965: Evaporation and environment. *Proc Symp Soc Exp Biol 19*, 205-234.

Nemes, A., 2003: Multi-scale hydraulic pedotransfer functions for Hungarian soils. *Ph.D. Dissertation*, Wageningen University.

Posza, I. and *Stollár, A.*, 1983: Plant constants evaluated by several years' measurements for the calculation of effective evaporation (in Hungarian). *Időjárás 87*, 170–177.

Szeicz, G. and *Long, I.F.*, 1969: Surface resistance of crop canopies. *Water Resour Res 5*, 622–633.

Szesztay, K., 1958: Estimation of water balance of catchment areas in Hungary. *Időjárás 62*, 313-328.

Thornthwaite, C.W. and *Mather, J.R.*, 1955: The water balance. *Drexel Inst. of Techn., Lab. of Clim. VIII.* No. 1, 1–104, N.J.

Turc, L., 1961: Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. *Annals of Agronomy 12*, 13-49.

Wardley, N.W. and *Curran, P.J.*, 1984: The estimation of green leaf area index from remotely sensed airborne multiespectral scanner data. *Int J Remote Sens 5*, 671-679.

Wiegand, C.L., Richardson, A.J. and *Kanemasu, E.T.*, 1979: Leaf area estimates for wheat from landsat and their implications for evapotranspiration and crop modeling. *Agron J 71*, 336–342.

Zhangsi, Y. and *Williams, T.H.L.*, 1997: Obtaining spatial and temporal vegetation data from landsat MSS and AVHRR/NOAA satellite images for a hydrologic model. *Photogramm Eng Rem S 63*, 69-77.

Várallyay, G., Szűcs, L., Rajkai, K., Zilahy P. and *Murányi, A.*, 1980: Hydro-physical classification and 1:100 000 scale maps of Hungarian soils (in Hungarian). *Agrokem Talajtan 29*, 77-112.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 3, July–September 2010, pp. 217–227

Modeling of redline dayglow emission

Vir Singh[1*], A. K. Upadhayaya[2], and M. V. Sunil Krishna[1]

[1]*Department of Physics, Indian Institute of Technology Roorkee*
Roorkee – 247 667, India

[2]*Radio and Atmospheric Sciences Division, National Physical Laboratory,*
New Delhi – 110 012, India

[*]*Corresponding author; E-mail: virphfph@iitr.ernet.in*

(Manuscript received in final form August 8, 2008)

Abstract — The present paper deals with the morphological study of redline dayglow emission. The morphology is obtained from the profiles of redline dayglow emission using the updated Glow model. The Glow model is updated in terms of measured various collisional cross sections and reaction rate coefficients. The volume emission rate is obtained for some specific cases (18.3°S, 99.0°E, at 7:33 a.m. on February 1, 1993, 1.3°N, 139.0°E, at 11:30 a.m. on February 11, 1993, 35.2°S, 197.0°E, at 3:13 p.m. on April 2, 1993, and 52.9°S, 207.0°E, at 3:45 p.m. on April 9, 1993) using updated Glow model, and the results are compared with WINDII measurements. The modeled emission rate is found in good agreement with the WINDII observations at all altitudes except in the peak region where model underestimates the WINDII observations within 25%. The updated Glow model is further used to obtain the morphology of redline emission under equinox conditions (for the months of March and April, 1993) between 50°S and 50°N latitudes. It has been found that this emission shows an asymmetry between the Northern and Southern Hemispheres under equinox conditions. The asymmetry in the thermospheric region is likely due to the changing contributions with altitude of the various sources which are responsible for the production of $O(^1D)$ during daytime. The daytime intensity variation is found quite consistent with the WINDII measurements under equinox conditions.

Key-words: airglow, satellite observations, photochemistry, global emission intensity

1. Introduction

Redline dayglow emission is possibly the most extensively observed emission in the dayglow (*Noxon* and *Johanson*, 1972; *Narayan et al.*, 1989; *Sridharan et al.*, 1992; *Shepherd et al.*, 1993). It has now been well established that the redline dayglow emission shows a peak in upper thermospheric region (200–230 km).

The well-known identified sources of redline dayglow emission are dissociative recombination, photoelectron impact on atomic oxygen, and photodissociation of O_2. However, the contribution of reaction of $N(^2D)$ with O_2 to the redline emission in airglow is still a matter of discussion (*Torr et al.*, 1981; *Link*, 1982; *Singh et al.*, 1996; *Link* and *Cogger*, 1988). Although, the altitude range of the production of $O(^1D)$ state extends from 100 to 300 km, the emission is mainly observed above 120 km, as below 120 km the $O(^1D)$ atoms are quenched by molecular nitrogen and molecular oxygen. A number of model calculations (*Hays et al.*, 1978; *Torr et al.*, 1981; *Bates*, 1990; *Singh et al.*, 1996; *Wittase et al.*, 1999) have been carried out to obtain volume emission rates (VER) of redline dayglow emission. These studies have only discussed the volume emission rate profile of 6300 Å in dayglow emission in the light of above mentioned production sources. From these studies one cannot have the idea about the global distribution of redline emission. The wind imaging interferometer (WINDII) has provided numerous data (*Shepherd et al.*, 1993) on redline dayglow emission along with the global distribution between 50°S and 50°N latitudes. WINDII measurements have thus provided an opportunity to study the morphology (volume emission rate as a function of altitude and latitude) of redline dayglow emission using theoretical models. The morphological study of 6300 Å would provide us valuable information about the relative distribution of this emission in both the hemispheres. The most reliable model that has been used to study the airglow emissions is the Glow model developed by *Solomon* (1992). After the development of this model, many input parameters, such as collisional cross-sections, reaction rate coefficients, and quantum yields, have been reexamined experimentally. Consequently the Glow model needs upgrade in light of the newly evaluated parameters.

In the present paper, the morphology of redline dayglow emission is studied in the altitude region of 120 to 300 km under equinox conditions at 10:00 a.m. in local time. The emission profiles are obtained using the Glow model of *Solomon* (1992). The Glow model is updated by using more recent cross-section data, reaction rate coefficients, and variation of solar fluxes. A comparison between the modeled morphology and the observed morphology from the WINDII measurements is made for a specific case for which WINDII data are available. The present morphology is used to discuss the relative variation of 6300 Å dayglow emission in both hemispheres.

2. Model

Mechanisms for the production of redline dayglow emission have been discussed by several workers (*Bates*, 1990; *Link* and *Swaminathan*, 1992; *Tyagi* and *Singh*, 2000), and the following reactions have been identified as the potential sources of redline emission in dayglow:

$$O(^3P) + e_{ph} \rightarrow O(^1D) + e_{ph}, \quad (1)$$

$$O_2^+ + e_{th} \xrightarrow{\eta_D k_2} O(^1S, ^1D, ^3P) + O(^3P), \quad (2)$$

$$N(^2D) + O_2 \xrightarrow{\eta_{2D} k_3} O(^1D) + NO \text{ (not confirmed)}, \quad (3)$$

$$O_2 + h\upsilon \rightarrow O(^1D) + O(^3P), \quad (4)$$

$$O(^1S) \xrightarrow{A_5} O(^1D) + h\upsilon\ (5577\ \text{Å}). \quad (5)$$

In Eqs. (1) and (2), e_{ph} and e_{th} represent photoelectrons and thermal electrons, respectively. A_i and k_i represent the Einstein's coefficient and rate coefficients of the reactions. The production of $O(^1S)$ in Eq. (5) has been discussed by a number of workers (*Tyagi* and *Singh*, 1998; *Singh* and *Upadhayaya*, 2004) in detail. The $O(^1D)$ atoms are quenched by the following reactions:

$$O(^1D) + N_2 \xrightarrow{k_6} O(^3P) + N_2, \quad (6)$$

$$O(^1D) + O_2 \xrightarrow{k_7} O(^3P) + O_2, \quad (7)$$

$$O(^1D) \xrightarrow{A_8} O(^3P) + h\upsilon(6300\ \text{Å}). \quad (8)$$

The Glow model developed by *Solomon* (1992) is used in present calculations, and all the above sources of $O(^1D)$ have been included in the model. The neutral densities and neutral temperature have been used from MSISE-90 neutral atmosphere model (*Hedin*, 1991). The Glow model is updated by using more appropriate O excitation and ionization cross-sections, as given by *Kanik et al.* (1993). The total electron impact cross-section for O is taken as given by *Laher* and *Gilmore* (1990). The transport of photoelectrons and conjugate point effects as given by *Banks* and *Nagy* (1970) and *Nagy* and *Banks* (1970), respectively, have been included in the present calculations. The solar flux values are based on the full F74113 reference solar spectrum of *Hinteregger et al.* (1981), which is scaled using parameterization method based on F10.7 (daily 10.7 cm solar flux) and F10.7A (81 days average of the 10.7 cm solar flux). For ionizing EUV, the bin structure method of *Torr* and *Torr* (1985) is used in the present model. More details about the scaling techniques can be seen in *Upadhayaya* and *Singh* (2002) and *Tyagi* and *Singh* (2000). The total production rate of $O(^1D)$ at an altitude z is given by the following equation:

$$R_z[O(^1D)] = R_{EI}[O(^1D)] + R_{DR}[O(^1D)] + R_{N2D}[O(^1D)] + R_{DS}[O(^1D)] + R_{CAS}[O(^1D)], \quad (9)$$

where R_{EI}, R_{DR}, R_{N2D}, R_{DS}, and R_{CAS} are production rates of $O(^1D)$ due to individual sources for the reactions of Eqs. (1)–(5). Contributions for these individual sources have been discussed in more detail by *Tyagi* and *Singh* (2000). The volume emission rate of $O(^1D)$ at a particular altitude z is given by

$$V_z[O(^1D)] = Q_D R_z [O(^1D)], \quad (10)$$

Q_D is the quenching factor of $O(^1D)$ given by

$$Q_D = \frac{A_8}{A_8 + k_6[N_2] + k_7[O_2]}, \quad (11)$$

where A_8 is the Einstein coefficient for $O(^1D)$ state. The reactions and rate coefficients are given in *Table 1*.

Table 1. Reactions and rate coefficients

Reaction	Rate coefficient ($cm^3 s^{-1}$)	References
$O(^3P) + e_{ph} \rightarrow O(^1D) + e_{ph}$	Impact cross-section	*Laher* and *Gilmore* (1990)
$O_2^+ + e_{th} \rightarrow O(^1D) + O(^3P)$	$k_2 = 1.6 \times 10^{-7}(300/T_e)^{0.5}$	*Walls* and *Dunn* (1974)
$N_2(^2D) + O_2 \rightarrow O(^1D) + NO$	$k_3 = 6.0 \times 10^{-12}$	*Lin* and *Kaufman* (1971)
$O(^1S) \rightarrow O(^1D) + h\upsilon$ (5577 Å)	$A_5 = 1.18\ s^{-1}$	*Nicolaides et al.* (1971)
$O(^1D) + N_2 \rightarrow O(^3P) + N_2$	$k_6 = 3 \times 10^{-11}$	*Hays et al.* (1978)
$O(^1D) + O_2 \rightarrow O(^3P) + O_2$	$k_7 = 2.9 \times 10^{-11}\exp(67.5/T_n)$	*Streit et al.* (1976)
$O(^1D) \rightarrow O(^3P) + h\upsilon$ (6300 Å)	$A_8 = 9.1 \times 10^{-3}\ s^{-1}$	*Nicolaides et al.* (1971)

3. *Results and discussion*

The 6300 Å dayglow volume emission rate profiles are obtained using the updated Glow model for several cases. However, to validate the model results, four cases are shown in *Fig. 1*. *Fig. 1a* shows the results for February 1, 1993 (35.2°S, 197.0°E, at 3:13 p.m., F10.7 = 125.1), *Fig. 1b* shows the results for February 11, 1993 (1.3°N, 139.0°E, at 11:30 a.m., F10.7 = 173.2), *Fig. 1c* shows the results for April 2, 1993 (18.3°S, 99.0°E, at 7:33 a.m., F10.7 = 120.8), and *Fig. 1d* shows the results for April 9, 1993 (52.9°S, 207.0°E, at 3:45 p.m., F10.7 = 135.5). WINDII data are available for these cases. It is quite clear from the values of F10.7 that these cases are for active days.

It is noticeable from *Fig. 1*, that the modeled results are in good agreement with the WINDII measurements except in the peak region, where the model underestimates the WINDII observations within 25%. Calculations have been performed for several other cases as well, and similar agreement is found with the WINDII observations. *Culot et al.* (2004) have also studied the redline emission using the TRANSCAR model and have compared their results with the WINDII observations. The TRANSCAR model also underestimates the WINDII observations, when the solar zenith angle (SZA) is less than 40°. The two cases for February 1, 1993 (*Fig. 1a*) and February 11, 1993 (*Fig. 1b*) have SZA less than 40°, where the model underestimates the WINDII observations within 25%. The other two cases for April 2, 1993 (*Fig. 1c*) and April 9, 1993 (*Fig. 1d*) have SZA greater than 40°, where the model is in very good agreement with the WINDII observations.

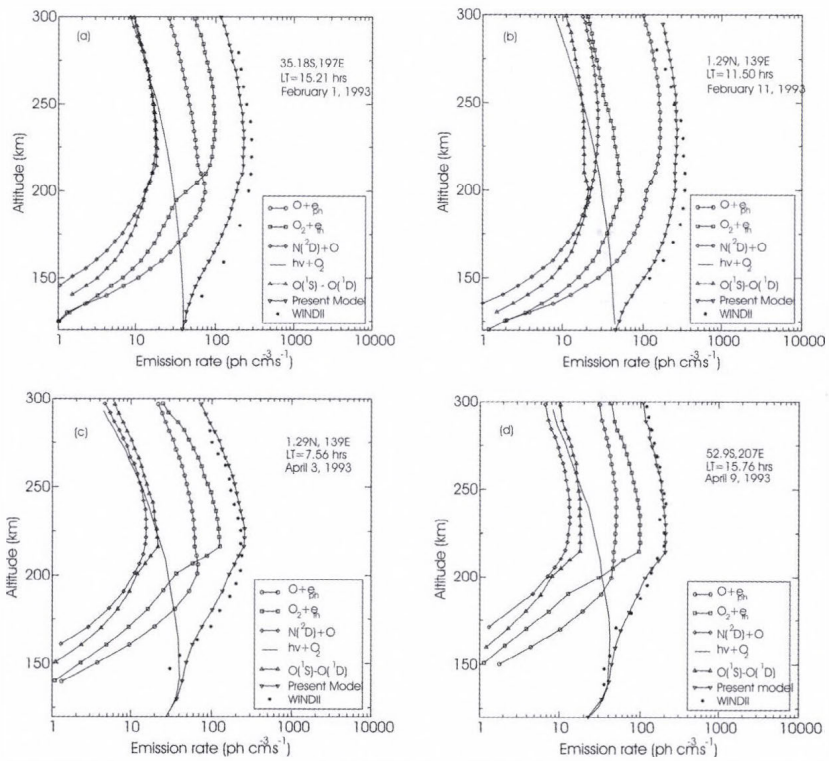


Fig. 1. Modeled and measured redline dayglow emission profiles for selected observing (WINDII) conditions with the various contributions of O(^{1}D) production processes.

The Glow model and the TRANSCAR model are thus quite consistent with each other. Volume emission rates of redline dayglow emission at various latitudes have been calculated along the track of satellite to obtain the

morphology of this emission on February 11, 1993, for which WINDII data are available. On February 11, 1993 the WINDII started measurements of redline emission at 40°S, 2:07 UTC and ended the measurements at 50°N, 2:56 UTC. These observations are shown in *Fig. 2a*. It is quite evident from *Fig. 2a* that the WINDII observations show an asymmetry between the Northern and Southern Hemispheres for redline emission. The modeled morphology of redline emission for this case is shown in *Fig. 2b*. One can notice from *Fig. 2b* that the modeled morphology also shows the asymmetry between the Northern and Southern Hemispheres for the redline emission. It is relatively higher in the Northern Hemisphere than in the Southern Hemisphere between the 30° and 40° latitudes. *Fig. 2c* shows the percentage difference ((Glow model – WINDII / WINDII) × 100) between the modeled results and the WINDII observations for volume emission rate as a function of altitude and latitude for the above case. A close examination of *Fig. 2c* reveals that the modeled results are within 25% agreement with the WINDII observations in the altitude region of 180–220 km. At other altitudes the percentage difference between modeled and WINDII observation is very narrow (much smaller than 25 percent). Consequently, on the basis of the results shown in *Fig. 2c*, one may consider that the model is in good agreement with the WINDII observations.

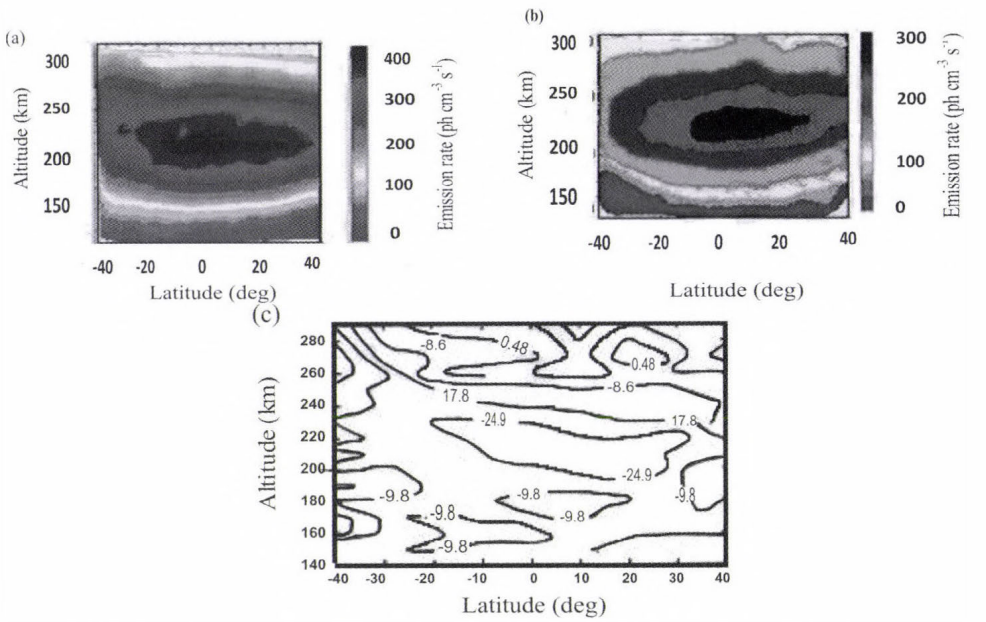


Fig. 2. (a) The altitude/latitude variation of volume emission rates as obtained from WINDII observations on February 11, 1993 (by courtesy of *G.G. Shepherd*). (b) The altitude/latitude variation of volume emission rates as obtained from model on February 11, 1993. (c) The percentage difference (((Glow model – WINDII)/WINDII) × 100) between the model calculations and WINDII observations on February 11, 1993.

The asymmetry of 6300 Å dayglow emission in both hemispheres can be best solved by studying the redline dayglow emission under equinox conditions (for the months of March and April). The reason to choose equinox case is that the sun is closer to the equator (approximate declination of 3°N). Note that this asymmetry in the illumination of the hemispheres due to the present location of the sun does not show any appreciable change (less than 3 percent) over the volume emission rates of redline dayglow emission obtained from the model in both hemispheres under the zero degree declination condition. Consequently, we may use this situation to identify the asymmetry of redline dayglow emission in both hemispheres. The 6300 Å dayglow volume emission rate profiles are obtained at various latitudes between 50°S and 50°N using the Glow model under equinox conditions. The volume emission rate profiles are obtained every fifth day starting from March 1, 1993 to April 30, 1993 at an interval of 5° latitude starting from 50°N, through the equator to 50°S at 10:00 a.m. using the Glow model. Averaging of volume emission rate is done for the above mentioned days for two months (March and April) at various altitudes for a fixed latitude. It would be worthwhile to mention here that the majority of the days during March and April 1993 were having the value of daily F10.7 solar flux greater than 110.

Fig. 3a shows the morphology of redline dayglow emission for 10:00 a.m. A close examination of the contours in *Fig. 3a* reveals that there is a clear asymmetry between the Northern and Southern Hemispheres in the thermospheric peak region (200–240 km) between the 20° and 50° latitudes. This asymmetry is clearly depicted in *Fig. 3b*, where the ratio of Northern to Southern Hemisphere volume emission rates is plotted as a function of altitude and latitudes. It is quite evident from *Fig. 3b* that the redline emission is about 15–20% higher in the Northern Hemisphere in comparison to the Southern Hemisphere between the 30° and 50° latitudes in the thermospheric peak region (200–240 km). The processes of dissociative recombination and photoelectron excitation of atomic oxygen are the dominating sources of this emission in this altitude region, and this, in, turn implies that the dissociative recombination and photoelectron excitation processes seem to be contributing more to redline emission in the Northern Hemisphere in comparison to the Southern Hemisphere. However, one can notice from *Fig. 1* that the contribution to redline emission due to dissociative recombination dominates over the photoelectron excitation contribution at midlatitudes. The TRANSCAR model (*Culot et al.*, 2004) also shows that the relative contribution due to the dissociative recombination process dominates over the photoelectron excitation contribution at midlatitudes. The production rate of $O(^1D)$ atoms due to the dissociative recombination process is proportional to the product of O_2^+ and thermal electron densities. Since, dissociative recombination process is the dominating source at midlatitudes, the asymmetry in redline emission may possibly be attributed due to higher densities of O_2^+ or thermal electrons in the

Northern Hemisphere in comparison to the Southern Hemisphere. However, this fact can only be ascertained if the simultaneous measurements of O_2^+ and thermal electron densities are available in both the hemispheres. It will be worthwhile to mention here, that the transport of photoelectrons is not symmetric about the geographic equator, because photoelectron fluxes are paired by conjugate points, which are not symmetric about geographic equator. Due to this fact, the production of redline emission due to photoelectron excitation process may vary from one hemisphere to another. This asymmetry may be a function of longitude. The results may vary from longitude to longitude. However, this longitudinal variation is not very prominent (the variation is within 5%) and has no strong bearing on the asymmetry of redline emission.

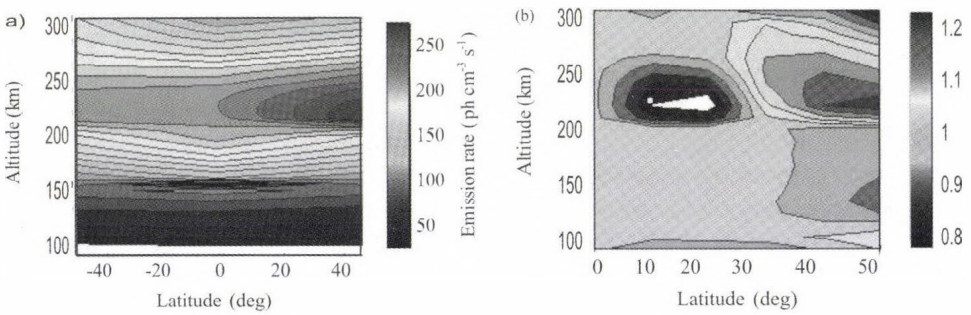


Fig. 3. (a) The altitude/latitude variation of two monthly (March–April 1993) averaged volume emission rates as obtained from model at 10:00 a.m. (b) The ratio of volume emission rate of the Northern to Southern Hemisphere as obtained from the model at 10:00 a.m.

There is no noticeable variation in the ratio of volume emission rates in the lower altitude region (120–180 km). In this altitude region, the photodissociation of O_2 is the dominating source of the production of $O(^1D)$ atoms. It indicates that the contribution to $O(^1D)$ by photodissociation process is more or less uniform in both hemispheres. In *Fig. 4*, the averaged latitudinal variation of intensity for redline dayglow emission is shown under equinox conditions at various local times. The intensities are obtained by integrating the volume emission rate over the vertical column of the altitude for every fifth day starting from March 1, 1993 to April 30, 1993 at an interval of 5° latitude starting from 50°N, through the equator, to 50°S at 8:00, 10:00, 12:00 a.m., 14:00 and 16:00 p.m. in local time. The averaging of intensities is done for the above mentioned days for two months (March and April) at various latitudes. The intensity reaches its maximum at noon, its magnitudes are quite similar at 10:00 a.m. and 2:00 p.m., and it obtaines relatively lower values at 8:00 a.m. and 4:00 p.m.

A variation in intensity from 1.6 KR to 2.1 KR depending on latitude and local time is seen for redline dayglow emissions. This is quite consistent with WINDII measurements as reported by *Zhang* and *Shepherd* (2004).

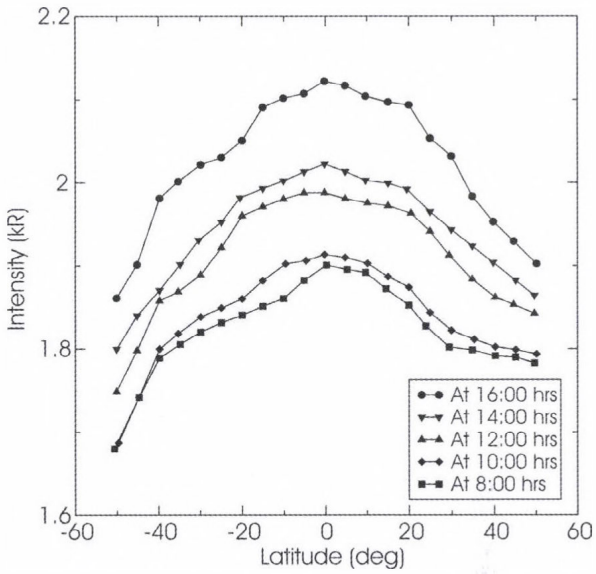


Fig. 4. Averaged latitudinal variation of intensity of 6300 Å dayglow emission under equinox conditions (March–April 1993) at various local times as obtained from the Glow model.

4. Conclusions

The morphology of redline dayglow emission has been studied using the updated Glow model between the 50°S and 50°N latitudes under equinox conditions (for the months of March and April 1993). It has been found that this emission shows an asymmetry between the Northern and Southern Hemispheres. It may be concluded that the asymmetry in the upper thermospheric region is likely due to the changing contributions of the dissociative recombination and photoelectron excitation processes with altitude and latitude. However, the dissociative recombination process seems to be more accountable for the asymmetry found in redline dayglow emission. The present study also indicates that the contribution to $O(^1D)$ by photodissociation process is more or less uniform in both hemispheres. A variation of 1.6 KR to 2.1 KR in intensity depending on latitude and local time is found for this emission, which is quite consistent with the WINDII observations.

Acknowledgement—The authors thank *Prof. G.G. Shepherd*, York University Canada for providing the WINDII data.

References

Banks, P.M. and *Nagy, A.F.*, 1970: Concering the influence of elastic scattering upon Photoelectron transport and escape. *J Geophys Res 75*, 1902-1910.

Bates, D.R., 1990: Oxygen green and redline emission and O_2^+dissociative recombination. *Planet Space Sci 38*, 889-902

Culot, F., Lathuillere, C. and *Lilensten, J.*, 2004: The OI 630.0 and 557.7 nm dayglow measured by WINDII and modeled by TRANSCAR. *Ann Geophys 22*, 1947-1960.

Hays, P.B., 1978: The OI(6300) airglow. *Rev Geophys Space Phys 16*, 225-232.

Hedin, A.E., 1991: Extension of the MSIS thermosphere model into the middle and lower atmosphere. *J Geophys Res 96*, 1159-1172.

Hinteregger, H., Fukui, K. and *Gilson, B.R.*, 1981: Observational references and model data on solar EUV, from measurements on AE-E. *Geophys Res Lett 8*, 1147-1150.

Kanik, K., Trajmar, S. and *Nickel, J.C.*, 1993: Total electron scattering and electronic state excitations cross sections for O_2, CO and CH_4. *J Geophys Res 98*, 7447-7460.

Laher, R.R. and *Gilmore, F.R.*, 1990: Updated excitation and ionization cross sections for electron impact on atomic oxygen. *J Phys Chem Ref Data 19*, 227-305.

Lin, C.-L. and *Kaufman, F.*, 1971: Reactions of metastable nitrogen atoms. *J Chem Phys 55*, 3760-3771.

Link, R., 1982: Dayside magnetosphere cleft auroral processes. *Ph.D. Thesis*, York University, Canada.

Link, R. and *Cogger, L.L.*, 1988: A reexamination of the OI 6300 nightglow. *J Geophys Res 93*, 9883-9892.

Link, R.J. and *Swaminathan, P.K.*, 1992: $N(^2D) + O_2$: A source of thermospheric 6300 Å emission. *Planet Space Sci 40*, 699-705.

Nagy, A.F. and *Banks, P.M.*, 1970: Photoelectron fluxes in the ionosphere. *J Geophys Res 75*, 6261-6270.

Narayanan, R. Desai, J.N. Modi, N.K. Raghavarao, R. and *Sridharan, R.*, 1989: Dayglow photometry: A new approach. *Appl Opt 28*, 2138-2142.

Nicolaides, C. Sinanoglu, O. and *Westhouse, P.*, 1971: Theory of atomic structure including electron correlation. IV. Method for forbidden transition probabilities with results for [OI], [OII], [OIII], [NI], [NII] and [CI]. *Phys Rev A4*, 1400-1410.

Noxon, J.F. and *Johanson, A.E.*, 1972: Changes in thermospheric molecular oxygen abundance inferred from twilight 6300 Å airglow. *Planet Space Sci 20*, 2125-2151.

Shepherd, G.G. et al., 1993: WINDII, the wind imaging interferometer on the upper atmosphere research satellite. *J Geophys Res 98*, 10725-10750.

Singh, V. and *Upadhayaya, A.K.*, 2004: Greenline dayglow emission under equinox conditions. *J Geophys Res 109*, A01308.

Singh, V., McDade, I.C., Shepherd, G.G., Solheim, B.H. and *Ward, W.E.*, 1996: The $O(^1D)$ dayglow emissions as observed by the WIND imaging interferometer on the UARS. *Adv Space Res 17*, 11-14.

Solomon, S., 1992: *Glow Model Version 0.95*. LASP, University of Colorado, Boulder, USA.

Sridharan, R., Haider, S.A., Gurubaran, S., Sekar, R. and *Narayanan, R.*, 1992: OI 630.0nm dayglow in the region of equatorial ionization anomaly: Temporal variability and its causative mechanism. *J Geophys Res 97*, 13715-13721.

Streit, G.E., Carleton, J.H., Schmeltekopf, A.L., Davidson, J.A. and *Schiff, H.I.*, 1976: Temperature dependence of $O(^1D)$ rate constants for reactions with O_2, N_2, CO_2,O_3 and H_2O. *J Chem Phys 65*, 4761-4764.

Torr, M.R. and *Torr, D.G.*, 1985: Ionization frequencies for solar cycle 21: revised. *J Geophys Res 90*, 6675-6678.

Torr, D.G., Richards, P.G., Torr, M.R. and *Abreu, V.J.*, 1981: Further quantification of the sources and sinks of thermosphere $O(^1D)$ atoms. *Planet Space Sci 29*, 595-600.

Tyagi, S. and *Singh, V.*, 1998: The morphology of oxygen greenline dayglow emission. *Ann Geophys 16*, 1599-1606.

Tyagi, S. and *Singh, V.*, 2000: Re-examination of reaction of $N(^2D)$ with O_2 as a source of $O(^1D)$ in dayglow emission. *Indian J Radio Space Phys 29*, 291-295.

Upadhayaya, A.K. and *Singh, V.*, 2002: Effects of temperature dependence of reaction $N_2(A^3\Sigma_u^+)$ + O on greenline dayglow emission. *Ann Geophys 20*, 2039-2045.

Walls, F.L. and *Dunn, G.H.*, 1974: Measurement of total cross sections for electron recombination with NO^+ and O_2^+ using ion storage techniques. *J Geophys Res 79*, 1911-1915.

Wittase, O., Lilensten, J., Lathuilliere, C. and *Blelly, P.L.*, 1999: Modeling the OI 630.0 and 557.7 nm thermospheric dayglow during EISCAT-WINDII coordinated measurements. *J Geophys Res 104*, 24.639-24.656.

Zhang, S.P. and *Shepherd, G.G.*, 2004: Solar influence on $O(^1D)$ dayglow emission rate: Global -scale measurement by WINDII on UARS. *Geophys Res Lett 31*, L07804.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 3, July–September 2010, pp. 229–234

Short communication

The virial theorem and planetary atmospheres

Viktor T. Toth

Ottawa, Ontario K1N 9H5, Canada; E-mail: vttoth@vttoth.com

(Manuscript received in final form February 19, 2010)

Abstract—We derive a version of the virial theorem that is applicable to diatomic planetary atmospheres that are in approximate thermal equilibrium at moderate temperatures and pressures and are sufficiently thin such that the gravitational acceleration can be considered constant. We contrast a pedagogically inclined theoretical presentation with the actual measured properties of air.

Key-words: atmosphere, degrees of freedom, hydrostatic equilibrium, thermodynamics, virial theorem

In his widely discussed article, *Miskolczi* (2007) postulates that the virial theorem, which relates the average kinetic and average potential energies of a bound mechanical system (see *Landau* and *Lifshitz* (1972) for a thorough introduction), can be applied to a planetary atmosphere in equilibrium in the planet's gravitational field.

To investigate if Miskolczi's postulate is correct (whether or not the postulate was correctly applied is a question beyond the scope of the present paper), let us first consider the case of a bouncing ball in a homogeneous gravitational field. (The gravitational field within the Earth's atmosphere, the thickness of which is small compared to the Earth's radius, is approximately homogeneous. However, the same results presented here could also be obtained using a Newtonian gravitational potential (*Pacheco* and *Sañudo*, 2003).)

Consider dropping the ball from a height h_b above the surface, and assume that it falls without air resistance, and bounces back from the ground with no loss of kinetic energy. We wish to calculate its average kinetic and average potential energy. It is sufficient to compute these averages for the first part of the ball's motion, as it falls to the ground; the bounce-back is just a time-reversed copy of its initial drop, and afterwards, in the absence of dissipative losses, the ball repeats the same motion *ad infinitum*.

We presume that the ball was dropped at $t = 0$. At any other time $t > 0$, before the ball hits the ground, its height will be

$$h = h_b - \frac{1}{2} g t^2, \tag{1}$$

where g (≈ 9.81 m/s^2 on the Earth) is the surface gravitational acceleration. From this,

$$t = \sqrt{\frac{2(h_b - h)}{g}}, \tag{2}$$

and notably, the time it takes to reach the ground ($h = 0$) is

$$t_0 = \sqrt{\frac{2h_b}{g}}. \tag{3}$$

The velocity of the ball at time t ($0 \le t \le t_0$) is

$$v = gt = \sqrt{2g(h_b - h)}. \tag{4}$$

The kinetic energy K and potential energy U of the ball are calculated the usual way:

$$K = \frac{1}{2} m v^2 = \frac{1}{2} m g^2 t^2, \tag{5}$$

$$U = mgh = mgh_b - \frac{1}{2} m g^2 t^2. \tag{6}$$

According to the virial theorem, for a bound mechanical system with kinetic energy K,

$$2\langle K \rangle = \left\langle \sum_a \mathbf{r}_a \mathbf{F}_a \right\rangle, \tag{7}$$

where the angle brackets denote time averaging and $\mathbf{r}_a$ and $\mathbf{F}_a$ are the position of, and the force acting on, the ath particle that constitutes the system. In the case of a system in which the potential energy is a homogeneous function of degree k of the coordinates, we get

$$2\langle K \rangle = k \langle U \rangle. \tag{8}$$

For Eq. (6), $k = 1$. The time averages of the kinetic energy, Eq. (5), and potential energy, Eq. (6), between $t = 0$ and $t = t_0$ can be calculated as

$$\langle K\rangle=\frac{1}{t_0}\int_0^{t_0} K=\frac{1}{6}mg^2t_0^2=\frac{1}{3}mgh_b, \tag{9}$$

$$\langle U\rangle=\frac{1}{t_0}\int_0^{t_0} U=mgh_b-\frac{1}{3}mgh_b=\frac{2}{3}mgh_b. \tag{10}$$

Hence, $2\langle K\rangle=\langle U\rangle$ and the virial theorem for the potential energy, Eq. (6), appears satisfied. This was made possible, in part, by referencing the height h to ground level; this allowed us to ignore the effects of the ground surface in Eq. (7), because at the ground, $\mathbf{r}_a=0$, as observed also by *Pacheco* and *Sañudo* (2003).

How about a column of atmospheric gas? We assume a column of gas in hydrostatic equilibrium standing over a unit surface area in a homogeneous gravitational field. Its density will be a function of height h above ground:

$$\rho=\rho(h). \tag{11}$$

The pressure at h is equal to the weight of gas situated at heights above h:

$$p(h)=\int_h^{\infty} g\rho(h')dh'. \tag{12}$$

We assume that the gas is in thermal equilibrium (the real atmosphere is not in thermal equilibrium, but that is another story), so its temperature is constant:

$$T=T_0. \tag{13}$$

We also assume that the gas obeys the ideal gas law (this is a valid approximation for air at room temperature and sea level pressure), hence

$$pV=nRT, \tag{14}$$

where V is the volume of n moles of gas, and $R\approx 8.31\ \mathrm{J\,K^{-1}\,mol^{-1}}$ is the ideal gas constant. The mass of n moles of gas is nM_n where M_n (≈ 0.029 kg/mol for air) is the molar mass of the gas; its density is $\rho=nM_n/V$, hence $V=nM_n/\rho$. We can thus rewrite the ideal gas law in the form

$$p=\frac{RT_0}{M_n}\rho. \tag{15}$$

Using this in Eq. (12), we obtain

$$\frac{RT_0}{M_n}\rho=\int_0^{\infty} g\rho(h')dh'-\int_0^{h} g\rho(h')dh', \tag{16}$$

or, in differential form,

$$\frac{RT_0}{gM_n}\frac{d\rho}{dh}=-\rho, \tag{17}$$

which can be solved trivially:

$$-\frac{RT_0}{gM_n}\frac{1}{\rho}d\rho = dh, \tag{18}$$

$$-\frac{RT_0}{gM_n}\log\rho = h + C, \tag{19}$$

or

$$\rho = \rho_0 e^{-gM_n h/RT_0}, \tag{20}$$

which we can also write in the form

$$\rho = \rho_0 e^{-h/h_0}, \tag{21}$$

$$h_0 = \frac{RT_0}{gM_n}. \tag{22}$$

For air at $T_0 = 273$ K, we get

$$h_0 = \frac{8.31\times 273}{9.81\times 0.029} \approx 8\text{km}, \tag{23}$$

which agrees well with the observed properties of the atmosphere.

So what about the virial theorem? Going back to the bouncing ball for a moment, we can immediately spot a potential problem: what if the ball is moving horizontally as well? Indeed, it can move horizontally at an arbitrary velocity, yet its potential energy will be no different, hence the virial theorem fails. We must make sure that we only consider the vertical component of the ball's velocity before the theorem can be considered valid. The velocity of the ball can be written in rectilinear form as $v^2 = v_x^2 + v_y^2 + v_z^2$, but we are only interested in the vertical component. In the specific case when the average velocities in the x, y, and z direction are the same, we get $\langle v_z^2 \rangle = \frac{1}{3}\langle v^2 \rangle$. Accordingly, the virial theorem in this case reads

$$\frac{2}{3}\langle K \rangle = \langle U \rangle. \tag{24}$$

This result can also be obtained using another argument, presented by *Pacheco* and *Sañudo* (2003): rather than allowing the coordinates to remain unbounded in the horizontal plane, we can consider confining the ball to within

a cylinder of unit radius, integrating and time averaging the force acting on the ball as it hits the cylinder walls, in order to obtain the right-hand side of Eq. (7).

An atmosphere, unfortunately, is not made of bouncing balls, however appealing that picture might appear. Air, in particular, is composed mainly of diatomic gases (notably N_2 and O_2), which at room temperature have two rotational degrees of freedom in addition to the three translational degrees of freedom discussed above. (At higher temperatures, vibrational modes also play a role.) The kinetic energy of a column of gas is its internal thermal energy. The principle of equipartition of energy states that internal thermal energy is distributed equally between all degrees of freedom. Therefore, the virial theorem now reads

$$\frac{2}{5}\langle K\rangle=\langle U\rangle. \tag{25}$$

This is our main result, valid for any diatomic atmosphere that obeys the ideal gas law in an homogeneous gravitational field at moderate temperatures.

For a column of gas over a unit surface area, the thermal kinetic energy is

$$K=\int_0^\infty c_V T_0 \rho\, dh=c_V \rho_0 T_0 \int_0^\infty e^{-h/h_0}\, dh=c_V \rho_0 T_0 h_0, \tag{26}$$

where c_V is the specific heat of the atmosphere at constant volume. The potential energy is just the gravitational potential energy:

$$U=\int_0^\infty g\rho h\, dh=g\rho_0 \int_0^\infty h e^{-h/h_0}\, dh=g\rho_0 h_0^2. \tag{27}$$

The ratio of the two is

$$\frac{U}{K}=\frac{gh_0}{c_V T_0}. \tag{28}$$

Using Eq. (22), we obtain

$$\frac{U}{K}=\frac{R}{c_V M_n}. \tag{29}$$

Given Eqs. (25) and (29), we can calculate the specific heat c_V. For air, we obtain

$$c_V=\frac{5R}{2M_n}\approx 716\frac{\mathrm{J}}{\mathrm{K\cdot kg}}, \tag{30}$$

a value that agrees well with the known properties of air ($c_V = 718\ \mathrm{J\ K^{-1}\ kg^{-1}}$).

In this derivation, we assumed that $T = T_0$ is constant. Our result, however, remains valid even when T is not constant, so long as the gas is in "local thermodynamic equilibrium", which ensures that the principle of equipartition remains valid and that thermodynamic quantities, such as temperature or specific heat, remain well-defined. To see this, we first rewrite the condition of hydrostatic equilibrium, Eq. (12), in differential form:

$$dp = -g\rho\, dh. \tag{31}$$

This allows us to write the potential energy of the gas, Eq. (27), as

$$U = \int_0^\infty g\rho h\, dh = -\int_0^\infty h\, dp = -\int_0^\infty h \frac{dp}{dh} dh = \int_0^\infty p\, dh, \tag{32}$$

where the last step was taken by integrating in parts and using $p(\infty) = 0$. On the other hand, the thermal kinetic energy, Eq. (26), can be rewritten as

$$K = c_V \int_0^\infty T\rho\, dh = \frac{c_V M_n}{R} \int_0^\infty p\, dh. \tag{33}$$

The ratio of Eqs. (32) and (33) remains the same constant ratio, Eq. (29), that we obtained in the $T = T_0$ case:

$$\frac{U}{K} = \frac{R}{c_V M_n}, \tag{34}$$

even as T varies with altitude. Therefore, even as we allow T to be a function of h, Eq. (25) remains satisfied.

Hence we were able to demonstrate, without having to invoke concepts such as "hard core" potentials or intramolecular forces, that the virial theorem is indeed applicable to the case of an atmosphere in hydrostatic equilibrium. However, it must be "handled with care": the nature of the atmosphere and the fact that the horizontal (translational) and internal (rotational) degrees of freedom of the gas molecules are unrelated to the gravitational potential cannot be ignored.

References

Landau, L.D. and *Lifshitz, E.M.*, 1972: *Theoretical Physics. Vol. I: Mechanics.* Nauka, Moscow.

Miskolczi, F.M, 2007: Greenhouse effect in semi-transparent planetary atmospheres. *Időjárás 111*, 1-40.

Pacheco, A.F. and *Sañudo, J.*, 2003: The virial theorem and the atmosphere. In *Nuovo Cimento C Geophysics Space Physics C, 26*, 311-316.

NEWS

Gabriella Szépszó won the 2010 WMO Research Award for Young Scientists

The Executive Council of the World Meteorological Organization (WMO) annually awards prizes for outstanding contributions in the fields of meteorology, climatology, hydrology, and related sciences. The prizes include the publication award Research Award for Young Scientists. This prize was conferred this year to two young researchers, one of them being *Gabriella Szépszó* of Hungary. Gabriella was awarded for the paper entitled "Transient simulation of the REMO regional climate model and its evaluation over Hungary". The paper was published in 2008 in IDŐJÁRÁS by Gabriella and *András Horányi*.

Gabriella is a gifted young scientist, who has been working for the Numerical Modeling and Climate Dynamics Division of the Hungarian Meteorological Service since graduating from the Eötvös Loránd University of Budapest in 2003. I had the opportunity to assist to the first steps of Gabriella in the field of numerical weather prediction, when she started to work at the Hungarian Meteorological Service. Initially she has worked on ensemble prediction systems, but soon she changed her orientation towards climate modeling. She spent several months at the Max Planck Institute for Meteorology in Hamburg as a visiting scientist, and during this stay she deepened her main scientific interest in the field of regional climate modeling and research. As a continuation of her work in Hamburg, she successfully adapted the REMO regional climate model for the Carpathian Basin, and this version has been serving as a research tool for dynamical climate downscaling for the territory of Hungary. Gabriella, together

with her PhD supervisor and coauthor András Horányi, have been carrying out outstanding research in this area in the recent years and with this prize their work is now recognized internationally.

Gabriella Szépszó has been earlier awarded as Prima Junior in the category of science. This prestigious Hungarian prize was conferred to her in 2007 in recognition of both her scientific talent and the importance of the topic of her interest.

It is the second time that a young Hungarian scientist is recognized by WMO for their outstanding research. This is, however, the first occasion that the award is given to the scientist for a publication in our Quarterly Journal IDŐJÁRÁS, which can also indicate the further improving quality of the journal of the last few years. This fact makes also us, the members of the Editorial Board of IDŐJÁRÁS particularly proud and honored.

I am convinced that this prize is a very promising start of a brilliant scientific career, where more similar scientific recognitions are anticipated in the forthcoming decades. I would like to congratulate Gabriella Szépszó on winning this prestigious award and I wish her a very successful future scientific career.

Gábor Radnóti

INSTRUCTIONS TO AUTHORS OF *IDŐJÁRÁS*

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

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

Some issues contain "News" and "Book review", therefore, such contributions are also welcome. The papers must be in American English and should be checked by a native speaker if necessary.

Authors are requested to send their manuscripts to

Editor-in Chief of IDŐJÁRÁS
P.O. Box 39, H-1675 Budapest, Hungary
E-mail: antal.e@met.hu

including all illustrations. MS Word format is preferred in electronic submission. Papers will then be reviewed normally by two independent referees, who remain unidentified for the author(s). The Editor-in-Chief will inform the author(s) whether or not the paper is acceptable for publication, and what modifications, if any, are necessary.

Please, follow the order given below when typing manuscripts.

Title page: should consist of the title, the name(s) of the author(s), their affiliation(s) including full postal and e-mail address(es). In case of more than one author, the corresponding author must be identified.

Abstract: should contain the purpose, the applied data and methods as well as the basic conclusion(s) of the paper.

Key-words: must be included (from 5 to 10) to help to classify the topic.

Text: has to be typed in single spacing on an A4 size paper using 14 pt Times New Roman font if possible. Use of S.I. units are expected, and the use of negative exponent is preferred to fractional sign. Mathematical formulae are expected to be as simple as possible and numbered in parentheses at the right margin.

All publications cited in the text should be presented in the *list of references*, arranged in alphabetical order. For an article: 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 photo-synthetically active radiation and cloud amount. *Időjárás 90,* 10-13. For a book: 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. Reference in the text should contain the name(s) of the author(s) in Italics and year of publication. E.g., in the case of one author: *Miller* (1989); in the case of two authors: *Gamov* and *Cleveland* (1973); and if there are more than two authors: *Smith et al.* (1990). If the name of the author cannot be fitted into the text: *(Miller,* 1989); etc. When referring papers published in the same year by the same author, letters a, b, c, etc. should follow the year of publication.

Tables should be marked by Arabic numbers and printed in separate sheets with their numbers and legends given below them. Avoid too lengthy or complicated tables, or tables duplicating results given in other form in the manuscript (e.g., graphs).

Figures should also be marked with Arabic numbers and printed in black and white or color (under special arrangement) in separate sheets with their numbers and captions given below them. JPG, TIF, GIF, BMP or PNG formats should be used for electronic artwork submission.

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 for authors is available: antal.e@met.hu

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/Journal-Idojaras.php

VOL. 114 * NO. 4 * OCTOBER – DECEMBER 2010

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: Gilice tér 39, H-1182 Budapest, Hungary
P.O. Box 39, H-1675 Budapest, Hungary
E-mail: bozo.l@met.hu or antal.e@met.hu
Fax: (36-1) 346-4809

Indexed and abstracted in Science Citation Index Expanded™ and Journal Citation Reports/Science Edition
Covered in the abstract and citation database SCOPUS®

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary
E-mail: kenderesy.k@met.hu or antal.e@met.hu

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 4, October–December 2010, pp. 237–250

A new bias correction scheme for assimilating GPS zenith tropospheric delay estimates

Andrea Storto[1,2*] and **Roger Randriamampianina**[1,3]

[1]*Norwegian Meteorological Institute,*
P.O. Box 43, Blindern, N-0313 Oslo, Norway

[2]*Euro-Mediterranean Centre for Climate Change,*
viale Aldo Moro 44, I-40127 Bologna, Italy; E-mail: andrea.storto@cmcc.it

[3]*Hungarian Meteorological Service,*
P.O. Box 38, H-1525 Budapest, Hungary

[*]*Corresponding author*

(Manuscript received in final form August 10, 2010)

Abstract—The impact of GPS zenith tropospheric delay estimated from the ground-based network of GPS receiving stations in the Norwegian regional model is assessed in this work. The data are screened from the stations with irregular or insufficient dissemination of observations. A normality test on the observation minus model-equivalent of the remaining dataset led to 29 assimilated pairs of stations/processing centers from the initial 668. A site-dependent multi-linear scheme that uses model surface pressure, skin temperature, integrated water vapor, and tropospheric thickness as predictors to correct the observation departure biases has been implemented. This scheme is proved to successfully remove the observational error bias. A two-month impact study during the winter of 2008 showed that the selected and bias-corrected dataset was of benefit for geopotential and short-range humidity fields. The impact on temperature fields and precipitation forecasts was neutral.

Key-words: data assimilation, 3d-var, GPS, bias correction

1. Introduction

The zenith tropospheric delay (ZTD) is the converted-to-distance time delay of the Global Positioning System (GPS) signal transmitted by the Global Navigation Satellite System (GNSS) constellation. The time delay is mainly due to the constituents of the troposphere and stratosphere, and GNSS data processing provides estimates of the excess length, rather than a delay, along the zenithal path above the receiving station (*Bevis et al.*, 1992).

The tropospheric delay at the radio frequencies depends on the refractivity $N=10^6(n-1)$, n being the refractive index, integrated over the vertical column from the receiver altitude h to the top of the atmosphere (TOA), namely

$$ZTD = \int_{z=h}^{TOA} [n(z)-1]dz. \tag{1}$$

It is well-known (e.g., *Smith* and *Weintraub*, 1953) that the refractivity, neglecting the Lorentz polarization, can be expressed as

$$N = p_d \frac{K_1}{T} + p_v \left(\frac{K_2}{T} + \frac{K_3}{T^2}\right), \tag{2}$$

where p_d is the partial pressure of dry air, p_v is the partial pressure of water vapor, T is the temperature, and K_1, K_2, and K_3 are three constants. According to the previous formulation of Eq. (2), the refractivity within a zenithal atmospheric path depends on the temperature and specific humidity profiles. It is common practice in GPS meteorology to separate the total delay into a part which depends on surface pressure (called zenith hydrostatic delay, ZHD) and a part which is function of the water vapor distribution (called zenith wet delay, ZWD). The latter can be empirically related to the integrated water vapor (IWV) (*Bevis et al.*, 1992).

Zenith estimates of atmospheric delays have been extensively exploited for mapping the vertical water vapor and the precipitable water content. Since 2005, the EUMETNET EIG GNSS Water Vapour Programme (E-GVAP) has been devoting many efforts in building a homogeneous network of GPS stations over Europe to provide near real-time ZTD estimates to be used in numerical weather prediction (NWP) and nowcasting (*de Haan et al.*, 2004; *de Haan*, 2006).

The comparison of GPS delay data with NWP models and collocated radiosonde observations (*Haase et al.*, 2003; *Deblonde et al.*, 2005) has suggested that GPS zenithal data of delay can be fruitfully used for improving moisture forecasts. Similarly, *Cucurull et al.* (2004) showed that a dense network of stations is able to improve precipitation forecasts in mesoscale models.

Computing temperature and moisture profiles from integrated refractivity data is an inverse problem, as the delay data represent a measure integrated over the vertical. Modern atmospheric data assimilation methods, like the variational ones, are based on the computation of the model departures from the observations in observation space. Hence, they require the computation of the zenith tropospheric delay from the model temperature and humidity profiles, which is a direct problem. In past years, several national weather services have studied the impact of assimilating ZTD data. In the frame of COST Action 716, an improvement of geopotential height has been found at the Danish

Meteorological Institute in the regional three-dimensional data assimilation system (*Vedel* and *Huang*, 2004). At the Swiss Federal Office of Meteorology and Climatology, *Guerova et al.* (2006) used the nudging technique for correcting specific humidity in the low and middle troposphere through IWV synthetic observations derived from the ZWD. They found a positive impact of GPS data during the summer period on the forecasts of humidity, though precipitation forecasts scores were not completely satisfactory. At ECMWF and Météo-France, *Poli et al.* (2007) assimilated ZTDs in global four-dimensional variational (4d-var) schemes with a constant site-dependent bias correction, obtaining promising results in terms of both short-range quantitative precipitation forecasts and medium-range circulation.

In the present implementation, GPS data are assimilated in terms of ZTD, rather than ZWD or IWV, in the HARMONIE-Norway data assimilation system, whose theoretical formulation is briefly reviewed in Section 2. It is important to underline that the quality of the data coming from different processing centers may vary significantly (*Vedel* and *Huang*, 2004); we adopted a strategy consisting in carefully selecting the stations suitable for data assimilation, rather than using a denser but qualitatively inhomogeneous network of observations. Therefore, we have dedicated particular emphasis on the data selection, and implemented an original multi-linear bias removal procedure, which are described in Section 3. Results from a two-month impact study are presented in Section 4. Finally, Section 5 summarizes the main achievements.

2. The data assimilation system

The HARMONIE-Norway data assimilation system, used in this study, is based on the ALADIN regional model data assimilation and forecast system (*Randriamampianina* and *Storto*, 2008), which is developed by the ALADIN consortium (e.g., *Bölöni* and *Horvath*, 2010; *Bubnová et al.*, 1993; *Fischer et al.*, 2005; *Radnoti*, 1995). The HARMONIE-Norway model runs in experimental mode at the Norwegian Meteorological Institute at a resolution of 11 km. The upper-air analysis is performed through a spectral three-dimensional variational (3d-var) data assimilation system, which consists of minimizing a cost function (see e.g., *Courtier*, 1997) given by

$$J=\frac{1}{2}(x-x_b)^T\,\mathbf{B}^{-1}(x-x_b)+\frac{1}{2}(y-\mathbf{H}(x-x_b)-H(x_b))^T\,\mathbf{R}^{-1}(y-\mathbf{H}(x-x_b)-H(x_b)) \quad , \tag{3}$$

where x is the analysis at the minimum of J, x_b is the background, which is a prior estimate of the atmospheric state (a 6-hour forecast), y is the vector of the

observations, H is the fully non-linear *observation operator* which projects the state of the atmosphere onto the space of the observations, **H** is the tangent-linear version of the *observation operator* and **B** and **R** are the covariance matrices of the background and observational errors, respectively. In the formulation of Eq. (3), the fully non-linear observation operator is used only once for computing the initial departures. Conversely, the tangent-linear model is used for updating the cost function at each iteration according to the new model state. This ensures the quadratic form of the cost function. The observation operator for the GPS delay data (*Poli et al.*, 2007) is given by Eqs. (1) and (2) which are computed separately for each model layer and summed up to give the total atmospheric delay along the zenithal path. In particular, to account for possible inconsistencies between the model orography and the actual receiving station altitude, a criterion based on the comparison between the pressure corresponding to the station altitude (computed from the model first-guess fields) and that of the model layers is used: (i) when the station pressure is smaller than the layer top pressure, that layer does not contribute to the total ZTD; (ii) when the station pressure lies in a layer, only the contribution from the station pressure to the layer top pressure is considered; (iii) when the station pressure is greater than the bottom lowest layer pressure, the bottom pressure is changed to that of the station and values of temperature and specific humidity are assumed constantly equal to the layer values between the layer bottom and the station (see *Poli et al.*, 2007 for more details).

In its standard configuration (*Randriamampinina* and *Storto*, 2008), the HARMONIE-Norway supports the assimilation of synoptic station observations (SYNOP), radiosondes (TEMP), surface data from oceanographic buoys (BUOY), wind profilers (PILOT), aircraft observations (AIREP), satellite infrared radiances from MSG/SEVIRI, and microwave radiances from the NOAA and Metop polar-orbiting satellites. The latter satellite and the radiosondes have been proved to be the observation network with largest impact on the forecasts (*Storto* and *Randriamampianina*, 2010a). Background-error covariances are calculated from ensemble global perturbed analysis (*Isaksen et al.*, 2007) downscaled to the regional domain and projected to a 6-hour forecast (*Storto* and *Randriamampianina*, 2010b). The balances are purely statistical, and computed through multi-variate linear regression (*Berre*, 2000).

3. Data processing

3.1. Selection of GPS receiving stations

Usual NWP implementations of the variational theory require the observational errors to follow a normal distribution. The test used for verifying the normality of the observations was the "Omnibus D'Agostino K2 test" (*D'Agostino* and *Pearson*, 1973; *D'Agostino et al.*, 1990), which is a versatile and powerful tool

for both small and large sample sizes. The test has been conducted for all the available couples of station/processing centers using the observation minus first guess (OMF) data for the 6-hour assimilation window in the period of December 1, 2007– April 15, 2008, after having rejected observations suspicious to be affected by gross errors (i.e., with a square OMF departure three times larger than the sum of the observation and background error variances, which for simplicity are assumed constant). Only the stations with a p-value greater than 0.1 have been retained. *Fig. 1* shows the occurrences of the test applied to the OMF data for two couples of station/processing centers; the station on the left panel (SAFONGAA) has been rejected by means of the normality test, while the station ALKMKNM1 (right panel) has been retained. From the initial set of 668 pairs of stations/processing centers, only 37 presented a temporally regular dissemination of the data and have therefore been considered for this impact study. In particular, only an average number of 1 ZTD datum per day within the study period was considered the minimum threshold to include a station among those with sufficient amount of data. Eight stations have been rejected by means of the normality tests and two pairs have been filtered out because of duplication of the ground-based station, leading to a number of retained pairs equals to 29.

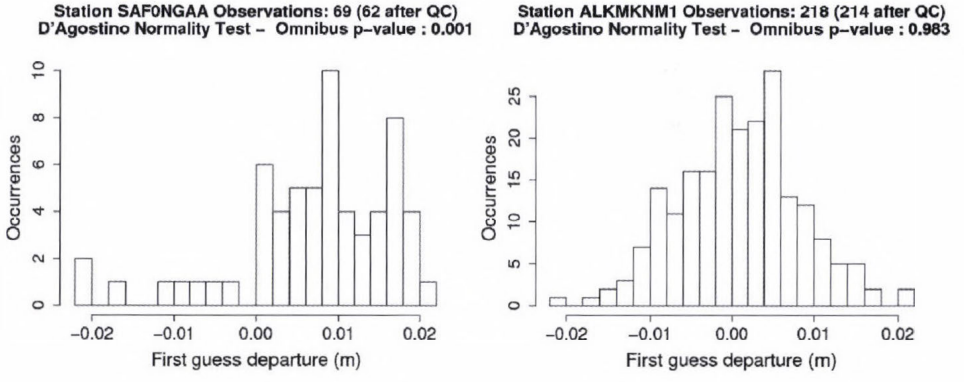


Fig. 1. Occurrences of first guess departures (observation minus first guess) for the station SAFONGAA (12.935° E, 59.136° N), rejected by means of the "Omnibus D'Agostino K2test", and the station ALKMKNM1 (4.728° E, 52.640° N), retained.

Fig. 2 shows the location of the selected stations. Most of them are located in the United Kingdom, Belgium, and the Netherlands, while a few of them are situated in southern Sweden and France. No horizontal thinning of the stations was considered in this study, because of their small amount.

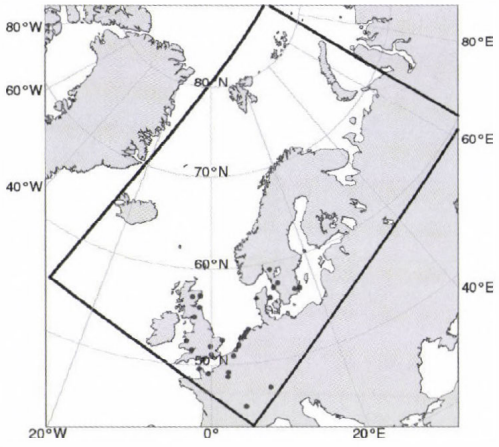


Fig. 2. Geographical distribution of the ground-based GPS stations retained for the assimilation of zenith tropospheric delays. Black lines represent the domain boundaries.

3.2. The ZTD bias correction scheme

In general, the retention of the observational bias, which usually arises from systematic errors in data processing or forecast model, is dangerous as it strengthens and propagates the bias in the analysis and in the forecasts (*Dee*, 2005). We have applied a bias removal that accounts for both GPS processing, observation operator, and NWP model systematic errors. The scheme was inspired by the bias correction scheme proposed by Harris and Kelly (*Harris* and *Kelly*, 2001). This scheme is largely adopted for bias correcting satellite radiances in NWP models. The OMF data can be thought of as the sum of different terms:

$$y - H(x) = \varepsilon_b + \varepsilon_o, \tag{4}$$

where ε_b and ε_o are the background and observational error, respectively. By applying the expectation operator, Eq. (4) becomes

$$E[y - H(x_b)] = b_o + b_b, \tag{5}$$

where b_o and b_b are the biases associated with the observation and the background errors, respectively. We assume that a linear bias model of the form

$$E[y - H(x_b)] = \beta_o + \sum_i \beta_i p_i \tag{6}$$

is able to represent the link between the model parameters and the OMF distribution. In Eq. (6), p_i is the ith predictor of the bias, β_i is its associated

coefficient, and β_0 is the offset. It is clear from Eqs. (5) and (6) that the bias coefficients can be calculated from the OMF dataset through multi-linear regression, if one knows the values of the model predictors corresponding to each observation. In the case of GPS ZTD, the sources of systematic bias in the OMF distribution are represented by errors in GPS data processing (formulation of hydrostatic delay, use of the mapping functions, errors in the ray-path to zenith delay conversion), systematic errors in NWP model fields, and systematic errors in the ZTD observation operator (such as, for instance, the contribution to the total atmospheric delay from the part of the atmosphere above the model top or the mismatch between model orography and station altitude). In addition, *Haase et al.* (2003) demonstrated how the station altitude and the moisture content may affect the bias between the GPS stations and the model-equivalents. In order to account for these several sources of systematic errors, we chose the skin temperature, surface pressure, integrated water vapor (IWV), and thickness of the layer between 950 and 200 hPa as predictors. This choice relies on the idea of representing all of the most significant sources of bias in a simple model: surface pressure and skin temperature are indexes of large-scale regime and hydrostatic delay calculation, and near-surface heating, respectively, and IWV and tropospheric thickness correlation may be thought of as model bias indexes for vertically integrated mass fields and, in general, accuracy of the observation operator. The offset (i.e., the intercept of the regression) also explains systematic errors in data preprocessing. Such a bias correction scheme has been applied independently for each selected pairs of station/processing centers. The calculation of the bias coefficients on an individual station basis stems from the evidence that the correlation between the observational bias and the predictors varies notably between different stations. The OMF dataset was derived from an experiment without GPS ZTDs assimilation but with the on-line computation of the ZTD observation minus first guess. The surface pressure results the parameter most correlated with the observations, with a negative correlation often of the order of -0.4 or -0.5. This seems to be related to the fact that the ZHD calculation, which dominates the total atmospheric delay, relies on the surface pressure fields. The relative importance of the other parameters in explaining the OMF bias depends strongly on the station, and the correlation is globally positive for the skin temperature and the integrated water vapor, whereas negative for the tropospheric thickness. In *Fig. 3*, the correlation coefficients between the predictors and the bias are shown for all the selected stations, together with the station altitude. In addition, also the correlation values between the bias and the stratospheric thickness (between 200 and 50 hPa) are plotted, although this parameter has been excluded from the set of predictors used in the impact study (in the next Section), as it presents a globally smaller correlation. According to *Fig. 3*, it is not possible to see any link between the station altitude and the average integrated moisture content, unlike similar studies (e.g., *Haase et al.*, 2003). The new resulting distribution of the observation minus guess is

unbiased and shows a reduction of standard deviation of the order of 1.5 mm in most cases. Despite the low correlation between the observation bias and the predictors, this bias correction scheme is efficient to correct the GPS bias.

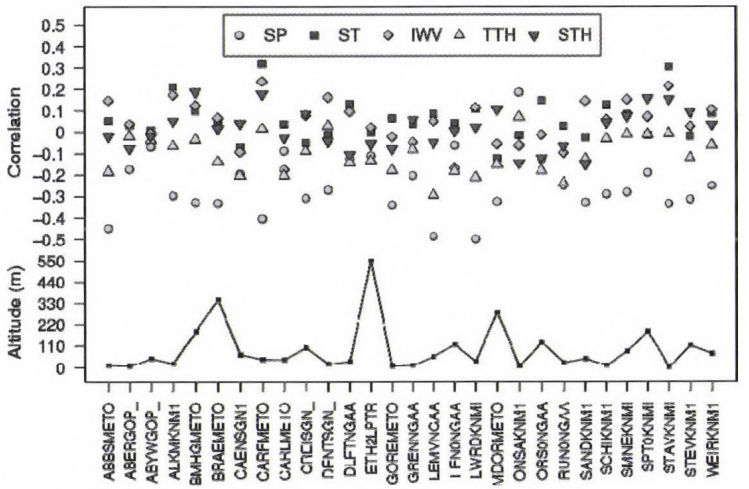


Fig. 3. Correlation values between observational bias and the predictors plus the stratospheric thickness, for all of the selected stations. In the bottom part of the figure, the station altitude is also plotted. The predictors are the surface pressure (SP), skin temperature (ST), integrated water vapor (IWV), tropospheric thickness (TTH, defined as thickness between 950 and 200 hPa pressure levels), and stratospheric thickness (STH, defined as thickness between 200 and 50 hPa pressure levels).

Table 1 demonstrates the impact of our multi-linear bias correction scheme. It reports the mean, the standard deviation and the maximum absolute value of the differences between the constant and the varying part of the bias correction scheme, i.e., between the first and the second term of the right hand side of Eq. (6), along with the station-dependent OMF statistics. Apart from two stations (ONSAKNM1 and STAVKNMI) which present very large differences between the two contributors to the bias correction of the order of tenths of centimeters, most stations show an average difference between the two terms of Eq. (6) of the order of 1 to 6 mm with a standard deviation between 1 and 4 mm. Maximum values are often around 1 cm, which is a value larger than the estimated accuracy of the ZTD (estimated to about 7 mm in both *Haase et al.* (2003) and *Poli et al.* (2007)), thus denoting that the varying part of the bias correction scheme effectively impacts the bias corrected values of the ZTD estimates. This may also be highlighted by comparing the OMF standard deviations with the standard deviation of the flat and the varying part of the bias correction scheme (in *Table 1*), where the former is in general between 2 and 10 times larger than the latter.

Table 1. Station-dependent observation minus guess statistics and difference between the constant and the varying part of the bias correction scheme (in Eq. (6) β_o and $\sum_i \beta_i p_i$, respectively). The table reports the mean the standard deviation of the observation minus guess departures and the mean, the standard deviation and the maximum absolute of the differences of the two components of the bias correction scheme in millimeters within the study period

Station	Observation – guess		Flat – varying		
	Mean	Standard deviation	Mean	Standard deviation	Maximum
ABBSMETO	2.2	10.8	1.1	3.8	9.9
ABERGOP_	7.2	7.0	6.4	1.3	9.4
ABYWGOP_	1.8	8.4	1.4	0.6	3.2
ALKMKNM1	0.8	8.2	0.8	2.5	6.4
BMHGMETO	-1.3	8.5	-1.3	2.8	6.4
BRAEMETO	1.8	7.3	1.7	3.1	10.6
CAENSGN1	8.5	68.0	2.3	2.5	7.2
CARFMETO	-0.4	9.1	-0.4	4.2	9.7
CARLMETO	-1.6	7.6	-1.5	1.9	5.9
CREISGN_	1.5	9.9	1.3	2.6	9.5
DENTSGN_	-0.4	25.9	3.5	2.4	9.9
DLFTNGAA	12.9	93.0	4.4	1.9	8.6
ETH2LPTR	8.7	10.1	6.5	1.9	10.8
GOREMETO	-0.1	10.3	-0.4	3.8	7.9
GRENNGAA	1.6	17.4	0.5	1.9	4.2
LEMVNGAA	1.2	17.9	-1.4	5.5	11.7
LEN0NGAA	8.0	15.4	6.5	1.5	10.4
LWRDKNMI	-0.8	9.5	-0.6	4.9	10.7
MDORMETO	2.5	8.1	2.5	3.0	10.1
ONSAKNM1	-2.5	9.4	-398.9	1.8	402.2
ORS0NGAA	2.2	18.3	4.2	3.0	11.6
RUN0NGAA	8.0	37.0	5.3	2.4	11.0
SANDKNM1	-3.3	8.0	-2.9	2.5	7.5
SCHIKNM1	0.4	9.8	0.2	2.7	7.3
SMNEKNMI	4.9	7.8	4.7	2.1	9.2
SPT0KNMI	4.6	7.0	4.5	1.5	7.8
STAVKNMI	2.3	10.4	-696.9	3.7	704.3
STEVKNM1	0.2	8.6	0.5	2.5	8.0
WEIRKNM1	1.2	9.2	0.9	2.6	6.8

4. Results

The wintertime impact study has been run for the period including February and March 2008, with four daily assimilation and forecast cycles (initialized at 00, 06, 12, and 18 UTC). Basic diagnostic statistics for the assimilation of GPS

delays are shown in *Fig. 4*. In most of the assimilation cycles, the OMF bias varied between –5 and 5 mm, with an average value of 0.6 mm. With respect to the observation minus analysis (OMA) bias, the average value is equal to 0.1 mm. Global standard deviation values of the OMF and OMA were equal to 8.7 and 2.7 mm, respectively. These values are smaller than the standard deviation of the difference between delay data and model-equivalents given by *Haase et al.* (2003), which was equal to 18 mm for the case of the HIRLAM model equivalents. Although this may be seen as a better fit of the HARMONIE-Norway first guess fields with the observations, it should be however mentioned that *Haase et al.* (2003) calculated the statistics within a summertime study period and for a domain covering southern Europe, as both these factors are important for increasing the ZTD OMF standard deviation, due to a larger moisture amount.

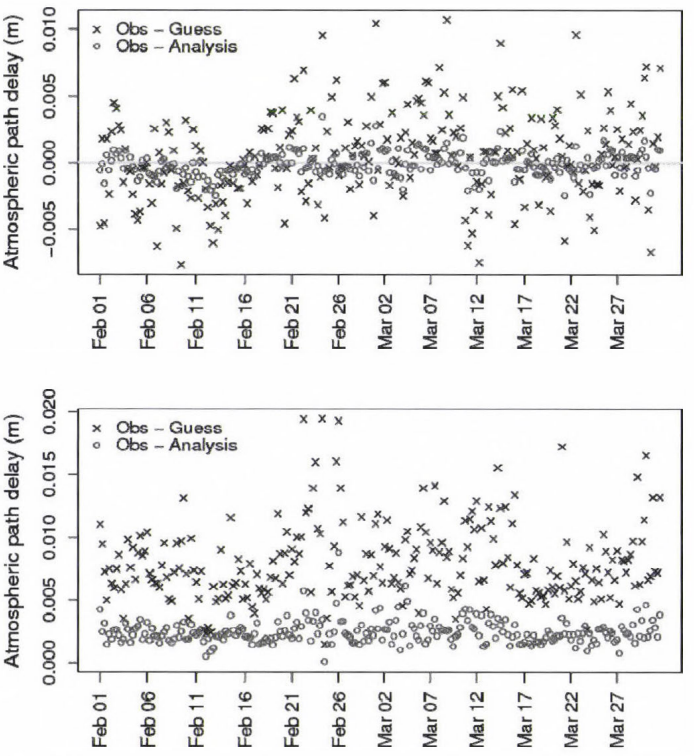


Fig. 4. Bias (top panel) and standard deviation (bottom panel) of observations minus first guess (black crosses) and observations minus analysis (red circles) for the six-hour GPS ZTD assimilation.

To quantify the impact of the GPS delay observations, the bias between the analysis increments of the GPS experiment and the analysis increments of the experiments without GPS data ("GPS" and "NOGPS", respectively) has been

calculated. Most of the signal coming from the new observations is transferred to dynamical parameters, especially for geopotential height (with overall bias reaching 0.3 $m^2 s^{-2}$ in low tropospheric levels), and velocity fields (up to 0.2 $m s^{-1}$), especially in the areas including the Norwegian Sea, the northern part of the Scandinavian peninsula, and the Danish peninsula. Effects of GPS observations on the humidity fields occur mostly in the Belgium and UK areas further to the Norwegian Sea, while the bias of the temperature analysis increment differences is very small. In general, the impact of the GPS observations can be observed rather in lower and middle than in upper tropospheric levels. In *Fig. 5*, the bias of the analysis increments between the two experiments is shown for 850 hPa specific humidity.

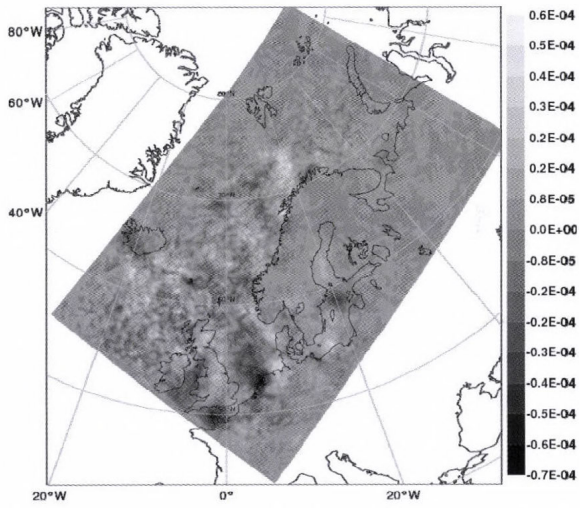


Fig. 5. Bias of analysis increments (analysis minus first guess) of GPS experiment minus NOGPS experiment, averaged over the two-month simulation period for 850 hPa specific humidity (kg/kg).

Verification skill scores against independent radiosonde measurements have been calculated for the overall impact period and compared against the NOGPS experiment. In *Fig. 6*, the contours of the differences between the root mean square error (RMSE) of the NOGPS and the RMSE of the GPS experiment as a function of forecast time and pressure are shown. The overall mean decrease of the error is particularly notable for the geopotential fields at all forecast ranges and below 300 hPa. The positive impact of the ZTD data increases with the forecast length (*Fig. 6*). *Fig. 7* shows the significance test[1] applied to the forecasts differences for the geopitential fields in middle tropospheric

[1] The significance test is based on statistical t-test on the difference in the expected values of the RMSE scores of the compared experiments.

levels (700 hPa). The impact on temperature is neutral to slightly positive, the impact on moisture fields is significantly positive on short-range forecasts throughout the troposphere, while the impact becomes negative on middle-range lower tropospheric levels (*Figs. 6* and *7*). We observed as well small, but significant positive impact of GPS data on two-day forecast of low tropospheric wind speed (not shown). For surface parameters (not shown), verified against independent observations from the EWGLAM network of synoptic stations, the effect of the GPS-ZTD assimilation is slightly positive for surface pressure and negligible for 2-meter temperature and humidity, and 10-meter wind. Unlike other impact studies, we did not find any significant impact on precipitation skill scores.

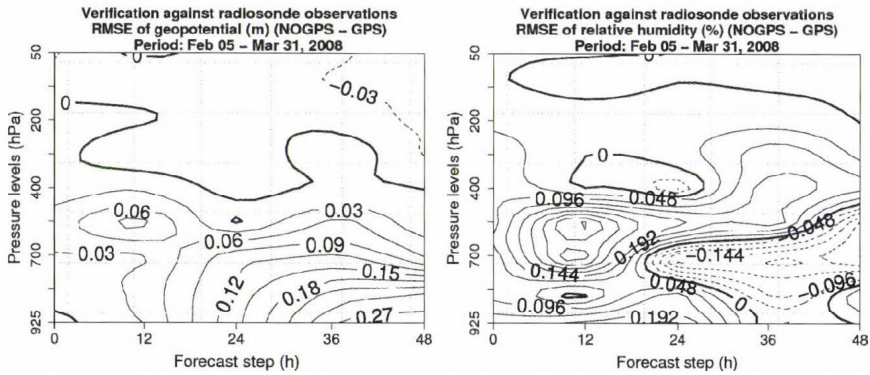


Fig. 6. Radiosonde verification skill scores: contour of differences between RMSE of the reference experiment (NOGPS) and RMSE of GPS experiment as a function of pressure level and forecast length. Positive values indicate that the assimilation of GPS-ZTD decreases the root mean square error; (left) geopotential, (right) relative humidity.

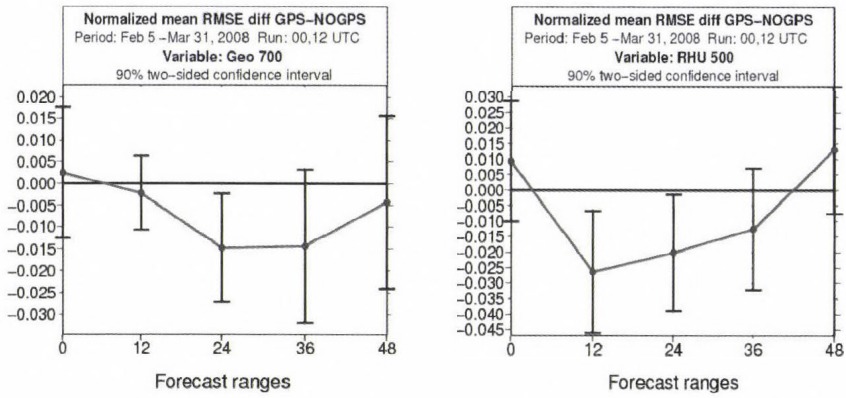


Fig. 7. Significance test: geopotential (left) and relative humidity (right) normalized mean RMSE differences at 700 hPa and 500 hPa, respectively The RMSE was computed from differences against observations.

5. Conclusions

At present, HARMONIE-Norway supports the assimilation of zenith tropospheric delay data from the ground-based GPS stations directly in the 3d-var assimilation system. Model-equivalents of the zenith delays are calculated as refractivities integrated over the vertical column. The strategy adopted in this study aims to evaluate the impact of those few stations, which are considered reliable according to (i) a criterion of regularity in the data availability and (ii) to the "Omnibus D'Agostino K2 test" of Normality of the OMF distribution. Instead of a constant site-dependent bias correction procedure, we noted a non-negligible correlation between the observation minus guess values and several surface and upper-air integrated parameters. This approach facilitated the identification of some sensitive predictors for a bias correction scheme, through multi-linear regression of OMF, independently for each station. Such a strategy allowed to have unbiased observation departures and, in most cases, a sensitive reduction of the OMF standard deviation. As point of discussion, the hierarchy of correlation between OMF bias and model predictors may largely vary among the GPS stations, except for the surface pressure which shows an important correlation for almost all the stations. As a matter of course, we stress the importance of the station-dependent definition of the bias correction scheme, which we kept in the data assimilation system. A wintertime impact study during February and March 2008 proved small but promising positive impact of GPS data on the regional HARMONIE-Norway assimilation and forecast system, especially for geopotential fields and short-range moisture forecasts. Negligible impact on temperature fields, skin parameters, and precipitation forecasts was found.

Acknowledgements—The authors are grateful to *Henrik Vedel* (Danish Meteorological Institute) and *Paul Poli* (Météo-France) for their support in the use of ZTD data. Two anonymous reviewers and the Executive Editor are acknowledged for the important help in improving the quality of this manuscript.

References

Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Mon. Weather Rev. 128*, 644-667.

Bevis, M., Businger, S., Herring, T., Rocken, C., Anthes, R., and *Ware, R.H.*, 1992: GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *J. Geophys. Res. 97*, 15787-15801.

Bölöni, G. and *Horvath, K.*, 2010: Diagnosis and tuning of background error statistics in a variational data assimilation system. *Időjárás 114*, 1-19.

Bubnová, R., Horányi, A., and *Malardel, S.*, 1993: International project ARPEGE/ALADIN. In *Newsletter 22*, EWGLAM, 117–130.

Courtier, P., 1997: Variational methods. *J. Meteorol. Soc. Japan 75*, 211-218.

Cucurull, L., Vandenberghe, F., Barker, D., Vilaclara, E. and *Rius, A.*, 2004: Three-dimensional variational data assimilation of ground-based GPS ZTD and meteorological observations during the 14 December 2001 storm event over the western Mediterranean Sea. *Mon. Weather Rev. 132*, 749-763.

D'Agostino, R.B. and *Pearson, E.S.*, 1973: Testing for departures from normality. *The American Statistician 60*, 613-622.

D'Agostino, R.B., Belanger, A. and *D'Agostino, R.B.J.*, 1990: A suggestion for using powerful and informative tests of normality. *The American Statistician 44*, 316-321.

Deblonde, G., MacPherson, S., Mireault, Y. and *Héroux, P.*, 2005: Evaluation of GPS Precipitable Water over Canada and the IGS Network. *J. Appl. Meteorol. 44*, 153-166.

Dee, D., 2005: Bias and data assimilation. *Q. J. Roy. Meteor. Soc. 131*, 3323-3343.

de Haan, S., 2006: Measuring atmospheric stability with GPS. *J. Appl. Meteorol. 45*, 467-475.

de Haan, S., Barlag, S., Baltink, H. and *Debie, F.*, 2004: Synergetic use of GPS Water Vapor and Meteosat images for synoptic weather forecasting. *J. Appl. Meteorol. 43*, 514-518.

Fischer, C., Montmerle, T., Berre, L., Auger, L., and *Stefănescu, S.*, 2005. An overview of the variational assimilation in the ALADIN/FRANCE NWP system. *Q. J. Roy. Meteor. Soc. 131*, 3477-3492.

Guerova, G., Bettems, J.-M., Brockmann, E. and *Matzler, C.*, 2006: Assimilation of COST 716 near-real time GPS data in the nonhydrostatic limited area model used at MeteoSwiss. *Meteorol. Atmos. Phys. 91*, 149-164.

Haase, J., Maorong, G., Vedel, H. and *Calais, E.*, 2003: Accuracy and variability of GPS tropospheric delay measurements of water vapor in the Western Mediterranean. *J. Appl. Meteorol. 42*, 1547-1568.

Harris, B. A. and *Kelly, G.*, 2001: A satellite radiance-bias correction scheme for data assimilation. *Q. J. Roy. Meteor. Soc. 127*, 1453-1468.

Isaksen, L., Fisher, M. and *Berner, J.*, 2007: Use of analysis ensembles in estimating flow-dependent background error variance. In *ECMWF Workshop on Flow-dependent Aspects of Data Assimilation*. 11–13 June 2007, Reading, United Kingdom.

Poli, P., Moll, P., Rabier, F. Desroziers, G., Chapnik, B., Berre, L., Healy, S., Andersson, E., and *El Guelai, F.-Z.*, 2007: Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. *J. Geophys. Res. 112*, 644-667.

Radnóti, G., 1995: Comments on "A spectral limited-area formulation with time-dependent boundary conditions applied to the shallow-water equations". *Mon. Weather Rev. 123*, 3122-3123.

Randriamampianina, R. and *Storto, A.*, 2008: ALADIN-HARMONIE/Norway and its assimilation system: the implementation phase. *HIRLAM Newsletter 54*, 20-30.

Smith, E. and *Weintraub, S.*, 1953: The constants in the equation for atmospheric refractive index at radio frequencies. In *Proceedings of IRE*, 1035-1037.

Storto, A., and *Randriamampianina, R.*, 2010a: The relative impact of meteorological observations in the Norwegian regional model as determined using an energy norm-based approach. *Atmos. Sci. Let. 11*, 51-58.

Storto, A. and *Randriamampianina, R.*, 2010b: Ensemble variational assimilation for the representation of background-error covariances in a high-latitude regional model. *J. Geophys. Res. Atmospheres 115*, D17204, doi: 10.1029/2009JD013111.

Vedel, H. and *Huang, X.-Y.*, 2004: Impact of ground based GPS data on Numerical Weather Prediction. *J. Meteorol. Soc. Japan 82*, 459-472.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 4, October–December 2010, pp. 251–261

Analysis of climate change in Hungary according to an extended Köppen classification system, 1971–2060

Áron Péter Fábián[1] and **István Matyasovszky**[2]

Department of Meteorology, Eötvös Loránd University,
P.O. Box 32, H-1518 Budapest, Hungary
E-mails: [1]*csigaa@gmail.com,* [2]*matya@ludens.elte.hu*

(Manuscript received in final form August 31, 2010)

Abstract—The purpose of this paper is to provide information on tendencies of the climate change in the area of Hungary (48.3°–45.5° W; 16.0°–22.3° E) between 1971 and 2060. Future climate change results obtained from climate models are not used directly due to their uncertainties and inaccuracies, but spatio-temporal distribution of climate classes of the Köppen climate classification system is analyzed. It has become clear that the Köppen system is unable to indicate the expected warming of the summer in certain cases, thus, two additional climate classes are introduced. Two datasets are used, the CRU TS 1.2 (with a spatial resolution of $1/6 \times 1/6$ degrees) for observed temperature and precipitation, and TYN SC 1.0 (same resolution) for climate model generated data. In the latter case, all the four available emission scenarios (A1FI, A2, B1, B2) are analyzed. The climate class of each grid box is calculated for each year from temperature and precipitation data, then the changes of areas covered by the classes resulted from the previous step are analyzed. The evaluation of tendencies is based on 10- and 30-year climate average maps and on the extremes of yearly values of each class. It is found that climate extremes will strengthen in the future during a nearly constant warming. Additionally, the more and more frequent appearance of the steppe climate and the constant presence of the dry-summered class at the end of the period show growing aridity in the area.

Key-words: climate change, climate model, Köppen climate system, warming, drought

1. Introduction

Assessment of both the climate change experienced in recent decades and regional characterization of the anticipated global climate change is in the focus of research. The purpose of this paper is to provide information on tendencies of the climate change in the area of Hungary between 1971 and 2060. As the global or even regional climate models reproduce the observed climate with high

inaccuracies (*Bartholy et al.*, 2006; *Bartholy et al.*, 2008; *Szépszó* and *Horányi*, 2008) over the area of Hungary, future climate change results obtained from these models are not used directly but spatio-temporal distribution of climate classes of the Köppen climate classification system is analyzed.

Because model generated values of meteorological variables are used to determine these climate classes, their inaccuracies might result in climate class uncertainties. However, the applied classification methods are quite insensitive to deviations and errors (*Beck et al.*, 2006).

2. Data and methods

2.1. The Köppen climate classification system

The Köppen climate classification system is a so-called effective system. It classifies the climate by environmental characteristics and given threshold values of given climate components. The components used for classification are: annual mean temperature (t), annual precipitation amount (r), monthly average temperatures of the summer and winter months, monthly precipitation amounts of the summer and winter months, mean temperature of the summer and winter, precipitation amounts of the summer and winter, monthly mean temperatures of the 4 warmest months, monthly mean temperature of the coldest month, and elevation above the sea level. Threshold values of the climate classes which are likely to appear in Hungary are described in *Tables 1–3* according to *Critchfield* (1983).

Table 1. Arid Köppen climate classes according to *Critchfield* (1983)

Class	Description	Precipitation limit r (cm)	Temperature limit t (°C)
Arid with winter precipitation: at least 70% of the annual precipitation falls in the 6 coldest months, and $r < 2t$ applies			
BW	Desert	$r < t$	-
BSh	Warm steppe	$r > t$	$t > 18$
BSk	Cold steppe	$r > t$	$t < 18$
Arid with summer precipitation: at least 70% of the annual precipitation falls in the 6 warmest months, and $r < 2(t + 14)$ applies			
BW	Desert	$r < t + 14$	-
BSh	Warm steppe	$r > t + 14$	$t > 18$
BSk	Cold steppe	$r > t + 14$	$t < 18$
Arid with balanced precipitation: less than 70% of the annual precipitation falls in either halves of the year, and $r < 2(t + 7)$ applies			
BW	Desert	$r < t + 7$	-
BSh	Warm steppe	$r > t + 7$	$t > 18$
BSk	Cold steppe	$r > t + 7$	$t < 18$

Table 2. Temperate rain Köppen climate classes according to *Critchfield* (1983)

Class	Description	Summer (warmest month)	Winter (coldest month)	Precipitation
Cs	Arid in summer	Mean temperature over 10 °C	Mean temperature over -3 °C	Sum of the most arid summer month is less than 4 cm, and also less than the tierce of the precipitation sum of the most humid winter month
Cw	Arid in winter	Mean temperature over 10 °C	Mean temperature over -3 °C	Sum of the most arid winter month does not exceed the tenth of the precipitation sum of the most humid summer month
Cfa	Balanced precipitation	Mean temperature over 22 °C	Mean temperature over -3 °C	Balanced
Cfb	Balanced precipitation	Mean temperature over 10 °C for 4 months	Mean temperature over -3 °C	Balanced
Cfc	Balanced precipitation	Mean temperature over 10 °C	Mean temperature over -3 °C	Balanced

Table 3. Boreal forest and snow Köppen climate classes according to *Critchfield* (1983)

Class	Description	Summer (warmest month)	Winter (coldest month)	Precipitation
Ds	Arid in summer	Mean temperature over 10 °C	Mean temperature below -3 °C	Sum of the most arid summer month is less than 4 cm, and also less than the tierce of the precipitation sum of the most humid winter month
Dw	Arid in winter	Mean temperature over 10 °C	Mean temperature below -3 °C	Sum of the most arid winter month does not exceed the tenth of the precipitation sum of the most humid summer month
Dfa	Balanced precipitation	Mean temperature over 22 °C	Mean temperature below -3 °C	Balanced
Dfb	Balanced precipitation	Mean temperature over 10 °C for 4 months	Mean temperature below -3 °C	Balanced
Dfc	Balanced precipitation	Mean temperature over 10 °C	Mean temperature below -3 °C	Balanced
Dfd	Balanced precipitation	-	Mean temperature below -38 °C	Balanced

The original classification system is inappropriate to represent warming trends in such environments where the thresholds of Cfa or Dfa apply, but the highest monthly mean temperature is far above 22 °C. There is a need, therefore, to introduce additional climate classes to identify the warming of the climate regions with already hot summer in the beginning of the time series. Of course, the original methodology of the system is kept in sight. The newly-introduced classes are:

- **Cfa+**: the monthly mean temperature exceeds 22 °C for at least 4 months,
- **Dfa+**: the monthly mean temperature exceeds 22 °C for at least 4 months.

2.2. The CRU TS 1.2 data set

The *CRU TS 1.2* (*Mitchell et al.*, 2004) is a gridded data set of monthly means and sums of several climate variables for the period of 1901–2000 with a resolution of $1/6 \times 1/6$ degrees. Its spatial coverage is defined by the 11° W–32° E longitudes and 34°–72° N latitudes including the entire continent of Europe. Actually, temperature means and precipitation totals are used over an area with latitudes 48.3°–45.5° N and longitudes 16.0°–22.3° E represented by $18 \times 40 = 720$ grid boxes.

2.3. The TYN SC 1.0 data set

The *TYN SC 1.0* (*Mitchell et al.*, 2004) is a gridded data set of monthly means and sums of several climate variables for the period of 2001–2100 with a spatial coverage and resolution as in the case of the CRU TS 1.2 data set. It does not contain raw model output, but its values are derived from the means of 1961–1990, the variability of 1961–1990, and a model provided response pattern to the rise of temperature. This method, although does not use downscaling of the model-generated data, can give more accurate results than raw model output as it bypasses the inaccurate prediction process of mean and variability, and replaces them by real observed data (*Mitchell et al.*, 2004).

Temperature means and precipitation amounts obtained from the HadCM3 (*Gordon et al.*, 2000) climate model with 4 scenarios (A1FI, A2, B1, B2) (*IPCC*, 2000) are used. The A1FI scenario expects a future with fast technological development and convergence between regions, but strong reliance on fossil fuels. The A2 describes a more heterogeneous, less technically developed world. The B1, like A1, assumes convergence, but along with changes in economical structure and more usage of resource-effective technologies. The B2 works with slower, more locally focused development, but also optimal environmental politics, unlike A2.

2.4. Software used for data processing

Calculations have been performed with an own developed software written in the Python programming language. A scheme of its operation is shown in *Fig. 1*,

a more detailed discussion can be found in the Appendix. Three kinds of products were made: climate maps for each analyzed periods (10-year average maps from the observed and 30-year average maps from the predicted datasets), extremes plots showing the minimum and maximum gridbox numbers covered by all climate classes in all years of the analyzed period, and annual spatial mean temperature diagrams for each analyzed period.

The analysis of climate change was mainly based on the average maps; the extreme plots were used to indicate the variability during the analyzed periods. For instance, an arid climate class with relatively high maximum value indicates that although this class may not appear on the average map (wetter years compensate it), there were droughts during the analyzed period.

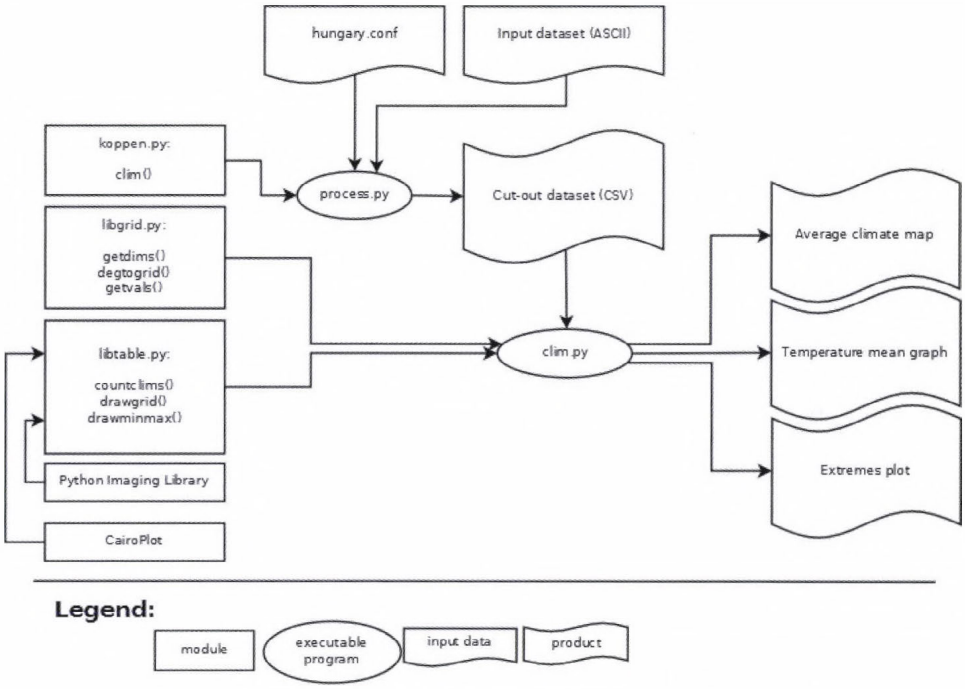


Fig. 1. A scheme of the software used for Köppen climate classification.

3. Results

3.1. Observed data

The climate of Hungary, according to the averages of the period 1951–2000 is classified as mild wintered and moderately warm summered Cfb class (*Kottek et al.,* 2006). However, the analysis of the years from 1971 to 2000 with the above

mentioned higher spatial resolution produces more precise and quite different results. The time series of annual spatial mean temperatures (*Fig. 2*) shows a quasi-constant rise with minor declines that can also be observed in the occupied areas of the climate classes through the period.

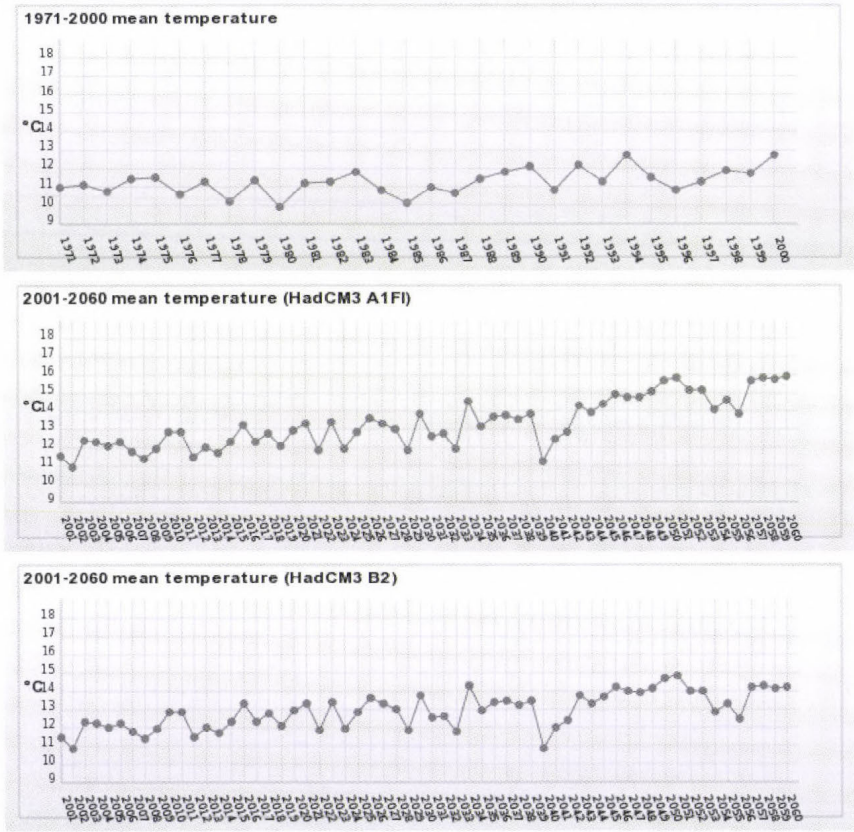


Fig. 2. Spatially averaged annual mean temperature for the periods of 1971–2000 (top) and 2001–2060 for scenarios A1FI (middle) and B2 (bottom).

3.1.1. The 1970s (1971–1980)

The climate average map of the 1970s (*Fig. 3*) shows an almost concentric spatial distribution: the middle of the basin is dominated by the hot summered and mild wintered Cfa, the hill and mountain regions are covered by the cooler summered Cfb, while the north is accompanied by the colder wintered Dfb. The extremes plot shows remarkable Cw-Cs and Cfb-Dfb fluctuations. The prior indicates intense oscillations in the distribution of precipitation through the years; the latter shows the same in mean temperature of the winter.

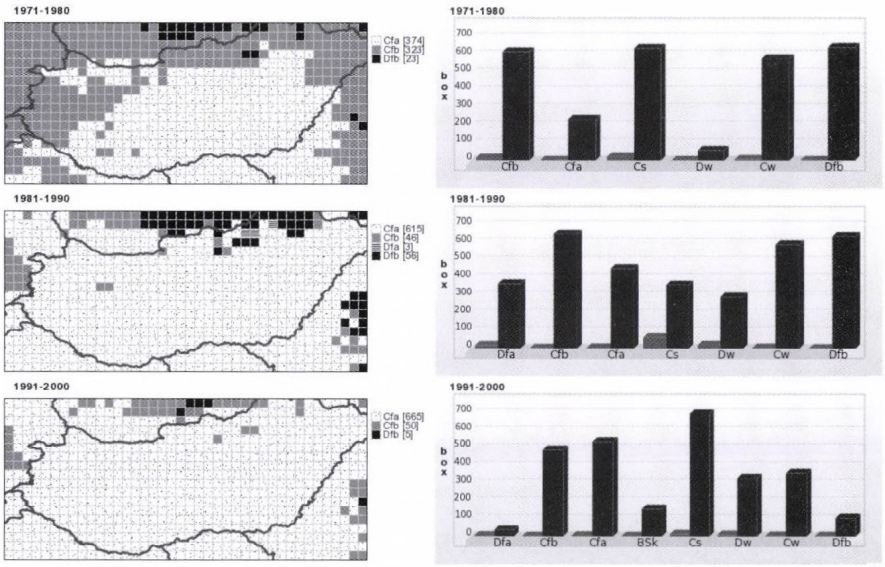


Fig. 3. Spatial distribution of Köppen climate classes with numbers of grid boxes (in parentheses) corresponding to different classes (left), and maximum (dark) and minimum (bright) numbers of grid boxes corresponding to different climate classes in individual years of the ten-year periods of 1971–2000 (right).

3.1.2. The 1980s (1981–1990)

On the average map of the 1980s (*Fig. 3*), the area of Cfa grows remarkably on the expense of Cfb (warming of summer), and the border of the northern Dfa zone expands to south suggesting the cooling of winters in these territories. The extremes plot shows again the Cfb-Dfb fluctuation started in the previous decade, but with lesser disturbances of the annual temporal distribution of precipitation (reduction of Cs area) and with somewhat major fluctuations in the mean temperature (Cw-Dw – mild or cold winter) as compared to the 1970s. These changes show a cooling in this decade identified also in the 30-year spatial mean temperature plot. The lower summer mean temperatures explain the phenomenon (*Szalai et al.*, 2005).

3.1.3. The 1990s (1991–2000)

In the 1990s (*Fig. 3*), the expansion of the Cfa area and the narrowing of colder climates continues. After the temporary decline during the previous decade, the rise of mean temperature returns. This tendency can also be observed in the extremes plot where the D classes are almost missing. Additionally, the considerably high maximum of the dry-summered Cs class with appearance of the steppe (BSk) show rising aridity with general and summer-specific precipitation reduction in the area.

3.2. Climate model generated data

The period 2001–2060 was analyzed according to the four emission scenarios. The scenarios generated similarly fluctuating spatial mean temperature curves (*Fig. 2*), but the slopes differ in each scenario. The entire period was divided into two intervals including the years 2001–2030 and 2031–2060, respectively. All scenarios produced similar results for the near-future period (2001–2030), so only one scenario (A1FI) is presented in this case. However, two scenarios – one with the least extreme changes (B2) and one with the most extreme changes (A1FI) – are presented for the latter period (2031–2060).

3.2.1. The period 2001–2030

The climate average map of this period is similar to the previous one (1990s) but with more Cfb in the borders of the basin indicating colder summers in these areas (*Fig. 4*). Additionally, the Dfb almost vanishes in this area showing the warming of the winters. In the middle of the basin, the mild-wintered, hot-summered Cfa still dominates. The extremes diagram is also very similar to the 1990s. The Cw and BSk maximums are higher on the expense of Cs, so the rise of aridity continues throughout this period.

3.2.2. The period 2031–2060 (A1FI)

The most intense climate change was obtained with scenario A1FI in this period (*Fig. 4*). On the average map, the Dfb fully and the Cfb almost vanishes. In the southern region, a big area of the very hot summered Cfa+ emerges. These changes unanimously indicate the warming of the summer. The majority of the basin remains Cfa. On the extremes plot, the maximum levels of BSk and Cs are remarkably higher than in the previous period, so very dry years can be counted on. As the Cw reduces the precipitation distribution has changed throughout the year. The relatively high maximum and minimum of Dfa+ (cold winter, very hot summer) shows extreme fluctuations of the temperature throughout the year in some regions.

3.2.3. The period 2031–2060 (B2)

The B1 scenario shows more similar picture to the previous period (*Fig. 4*). On the average climate map, the Cfb area reduces remarkably, although some Dfb remains. The coverage of Cfa+ in the south is also smaller. These effects involve a more moderate warming compared to A1FI. The extremes plot also shows milder changes: the maximum of BSk reduces, but the rising of Cs (observed in A1FI) is also unharmed here. The reduction of Cw is also observable, but the Dfa+ is weaker as compared to A1FI. To summarize, the B2 scenario shows almost the same changes as A1FI, but with a remarkably smaller level.

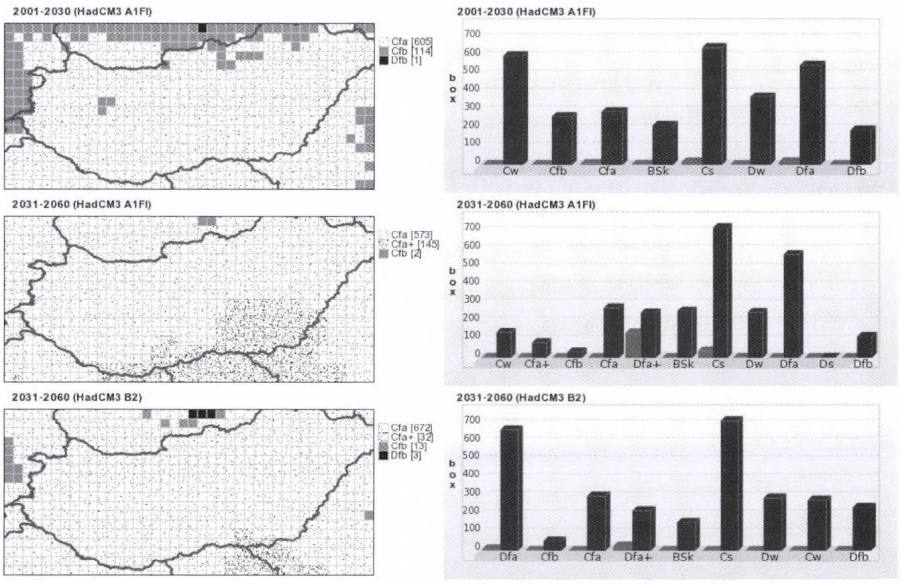


Fig. 4. Spatial distribution of Köppen climate classes with numbers of grid boxes (in parentheses) corresponding to different classes (left), and maximum (dark) and minimum (bright) numbers of grid boxes corresponding to different climate classes in individual years of the thirty-year periods of 2001–2060 (right).

4. Conclusions

The dominant climate class of the observed period 1971–2000 was Cfa in contrary to *Beck*'s (2006) results. Although the yearly results showed fluctuations with different magnitudes of the mean temperature and precipitation distribution, the area of Cfa extended on expense of the colder wintered and summered climate classes. This tendency fully coincides with the almost constant rise of mean temperatures observed in the period. Additionally, the emergence of steppe (BSk) in the last decade means minor drying.

In the climate model generated datasets, a warming tendency evidently exists. Also, the summer precipitation unanimously reduces in some years (see the high Cs maxima) and the occurrences of steppe (BSk) are also unharmed, so occasional droughts can be counted on. The different scenarios show these same changes but with different magnitudes, the most extreme is A1FI, and the least is B2. In A1FI, the strong presence of the Dfa+ (a class with cold winter and very hot summer) predicts extreme fluctuations of temperature in some years of the period.

The above discussed qualitative results coincide with the conclusions of IPCC for the future climate of Europe (*Alcamo et al.*, 2007). As the Köppen classification system is based on vegetation groups (*Peel et al.*, 2007), the results of this study can be used for various environmental, climatological, and agricultural purposes.

Also, a survey suggested that maps illustrating projected shifts of Köppen climatic zones are an effective visualization tool for disseminating climate change information (*Jylhä et al.*, 2010). Previous similar studies working on global (*Rubel* and *Kottek*, 2010) or continental (*Jylhä et al.*, 2010) scales give climate change information on relatively coarse spatial resolutions. For instance, an analysis of *Jylhä et al.* (2010) for Europe shows major future climate changes in the Iberian Peninsula, around the Black Sea and in the Alps. They found the regions with temperate (Cf, Cs) and dry climates (BW, BS) expanding. However, the entire area of Hungary is assigned here to Cfb class using data of the last five decades, and the projected future shifts of Köppen climatic zones are roughly represented as compared to our fine resolution approach. Also, presentation of extreme appearance of the different climate classes within the periods examined is an added product.

Since model generated values of meteorological variables are used to determine the climate classes, model output inaccuracies might result in climate class uncertainties. However, the applied classification methods are quite insensitive to deviations and errors (*Beck et al.*, 2006); see the threshold widths of the climate classes. Also, the raw model output is adjusted in the TYN 1.0 data set with a procedure reproducing much better the observed climate, and this procedure is then applied to model outputs corresponding to the future climate. This adjustment is expected to deliver more reliable model generated data (*Mitchell et al.*, 2004).

References

Alcamo, J., Moreno, J.M., Nováky, B., Bindi, M., Corobov, R., Devoy, R.J.N., Giannakopoulos, C., Martin, E., Olesen, J.E., and *Shvidenko, A.*, 2007: In *Climate Change 2007: Impacts, Adaptation and Vulnerability.* Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, UK, 541-580.

Bartholy, J., Pongrácz, R., Torma, Cs., and *Hunyady, A*, 2006: *Modelling Regional Climates and Adapting the PRECIS Climate Model* (in Hungarian). Országos Meteorológiai Szolgálat, Budapest, pp. 228.

Bartholy, J., Pongrácz, R., Gelybó, Gy., and *Szabó, P.*, 2008: Analysis of expected climate change in the Carpathian Basin using the PRUDENCE results. *Időjárás 112*, 249-264.

Beck, C., Grieser, J., Kottek, M., Rubel, F., and *Rudolf, B.*, 2006: Characterizing global climate change by means of Köppen climate classification. *Klimastatusbericht KSB* 2005, 139-149.

Critchfield, H. J., 1983: *General Climatology*. 4th ed. Prentice Hall, 453 pp.

Gordon, C., Cooper, C., Senior, C.A., Banks, H., Gregory, J.M., Johns, T.C., Mitchell, J.F.B., and *Wood, R.A.*, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Clim. Dynam. 16,* 147-168.

IPCC, 2000: *Special Report on Emission Scenarios*. Cambridge University Press, UK, 570 pp.

Jylhä, K., Tuomenvirta, H., Ruosteenoja, K., Niemi-Hugaerts, H., Keisu, K., and *Karhu, J.A.*, 2010: Observed and projected future shifts of climatic zones in Europe, and their use to visualize climate change information. *Wea. Climate Soc. 2,* 148-167.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and *Rubel, F.*, 2006: World Map of Köppen-Geiger Classification Updated. *Meteorol. Z. 15*, 1-5.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M., and *New, M.*, 2004: A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000)

and 16 scenarios (2001–2100). *Tyndall Centre Working Paper No. 55.*, URL: *http://www.ipcc-data.org/docs/tyndall_working_papers_wp55.pdf.*

Peel, M.C., Finlayson, B.L., and *McMahon, T.A.*, 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci. Discuss. 4,* 439–473.

Rubel, F. and *Kottek, M.*, 2010: Observed and projected shifts 1901–2100 depicted by world maps of Köppen-Geiger climate classification. *Meteorol. Z. 19,* 135-141.

Szalai, S., Konkolyné Bihari, Z., Lakatos, M., and *Szentimrey, T.*, 2005: *Some Characteristics of the Climate of Hungary from 1901* (in Hungarian). Országos Meteorológiai Szolgálat, Budapest.

Szépszó, G. and *Horányi, A.*, 2008: Transient simulation of the REMO regional climate model and its evaluation over Hungary. *Időjárás 112*, 203-231.

Appendix

The software group used for data processing consists of three common libraries (the koppen module for climate classification, the libgrid module for database processing, and the libtable module for producing climate maps and charts), a process program creating a sub-area including Hungary from the full dataset, and the clim program used to perform the analyses.

The koppen module's only function is clim(). This function is used to classify the climate of a year according to the Köppen system based on the given monthly mean temperatures and monthly precipitation totals, hemisphere (determining which months belong to winter and summer) and elevation data.

The functions of the libgrid module are the getdims() that reads the needed information from the headers of the data files, the degtogrid() that converts degree values to grid coordinates, and the getvals() that extracts the time series of the given grid boxes for the given dates.

The libtable module consists of countclims() creating the occurrence frequencies of the climates in a table, drawgrid() drawing a grid map using the values of the table, drawminmax() plotting the maximal and non-zero minimal area coverage data of all the climate classes in a given dataset. The Python Imaging Library and the CairoPlot module were used for drawing and plotting, respectively.

The process program creates a cut-out from the dataset according to the given geographical and temporal coordinates recorded in the configuration file hungary.conf, classifies the climates of all grid boxes in the cut-out, and saves them into a CSV data table. The difference between the process programs used for analyzing the observed and model generated datasets is that the program for the model data sets needs the name of the emission scenario (A1FI, A2, B1, B2) in command line argument; in addition, it indicates the scenario in the output file names.

The clim program uses the output files of the process. It draws the decadal mean climate maps, the spatial mean temperature graphs for the entire period, and the bar plots with the functions of the libtable. The program for the model generated datasets also needs the name of the emission scenario (A1FI, A2, B1, B2) in command line argument.

These programs are not actually available for public use, but development of their user friendly version is in progress.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 4, October–December 2010, pp. 263–273

Relation analysis between biomass and yield in maize seed production

Kálmán Kovács[1*], **Attila Csaba Dobos**[2], **Róbert Víg**[2], and **János Nagy**[2]

[1]*Federated Innovation and Knowledge Center,*
Budapest University of Technology and Economics,
Egry J. u. 18, bldg VI 526, H-1111 Budapest, Hungary; E-mail: kovacsk@mail.bme.hu

[2]*Centre for Agricultural and Applied Economic Sciences,*
Institute of Land Utilisation, Technology and Regional Development,
University of Debrecen,
Böszörményi út 138, H-4032 Debrecen, Hungary

[*]*Corresponding author*

(Manuscript received in final form November 3, 2010)

Abstract—The fact and volume of global warming in Hungary is proven by accurate meteorological measurements from the previous century. In agriculture and cultivation one should count with frequent changes of weather conditions due to climate change. Although the tendency of 0.6 °C raise in the annual average of temperature in the past one hundred years does not risk our cultivation so far, the high fluctuation of temperature, the winters, and the shorter transition seasons cause real problems. Similar potential threads are the decreasing trend of precipitation volume, the longer drought periods, and more frequent sudden heavy rainfalls. To provide a safe and well scheduled production, it is essential to maintain a decision support monitoring system based on up-to-date and accurate weather data. It is extraordinary needed in site-specific precision farming systems, where the aim is to produce goods of the highest quality. Nowadays, remote sensing is one of the basic instruments of continuous monitoring. Processing the collected data provides information for many fields of cultivation from technological parts to decision support processes.

The survey the vegetation state of plants, a measure can be calculated from remote sensing data, the so-called normalized vegetation index (NDVI).The aim of our research is to prove that satellite records can be a useful tool for maize seed production in two aspects: first, to find correlation between the NDVI and the weather conditions of a given year, and second, to find the optimal time interval for NDVI based crop estimation method.

Key-words: NDVI, yield estimation, climate change, remote sensing, crop of maize (*Zea mays L.*)

1. Introduction

To do successful cultivation, one should be aware of the interactions between the environmental factors and the cultivated plant as well as the crop fluctuations in time and space on account of these factors.

Hungary is on the edge of temperate continental climate with a high frequency of drought periods during the summer. The volume of precipitation, especially in the middle and southern parts of the country, decreased a rough 50 mm in the past 100 years, which is an extremely high degradation in respect to water needs. Meanwhile, the periods without precipitation have grown, and drought seasons have become more frequent (*Nagy*, 2006). Drought, as the most acute water shortage, can cause an economical catastrophe in cultivation. Fluctuation of precipitation in time and location is various in the territory of the country, thus periods and intensity of drought itself are also frequent. This aspect of the hypothetic climate change has a negative effect, extremely on the plants with high water needs. Further problems arise because of the sudden high volume falls (*Szász*, 2002). Naturally, the northern Great Plain region is not an exception either from climate change. Although the degradation of precipitation volume is rather timid, the annual average has declined 30–40 mm in the past 100 years. The question is how effective we are in eliminating the negative effects of climate change with irrigation, and what else we can do for this purpose.

Maize is a heat-sensitive field-corn, but the geographical edge of its potential of growing runs more and more to the north due to the result of plant breeding. The effective growing degree days (GDD) in Hungary are between 900 and 1500 °C, however, only maize hybrids between FAO 290 and FAO 540 maturity groups can be grown successfully (*Ángyán*, 1985). The minimum temperature of sprouting is 8–12 °C. If the soil temperature exceeds the minimum temperature of maize, the sprouting begins and its duration changes with regard to the ground temperature (*Keszthelyi*, 2005). The optimum temperature of shoot growth is 25–35 °C, therefore, maize can be grown in the most effective way where the average temperature of the hottest summer months is 21–27 °C. It needs an average temperature of 24–26 °C from detasseling period to the milk stage, however, in the period of pollen dispersal, too high temperature with meteorological drought can cause deficient fructification (*Shaw*, 1977 cit. *Turcsányi*, 1985). Maize is a good water utilizing plant, but it requires high quantity of water for intensive biomass production. The water need of maize is 420–440 mm in the growing season, but it can cope with a 250 mm lack of water compared to the optimum, since it can compensate the lack with the water stored in the ground, the groundwater, and the dew precipitation (*Bocz*, 1996). In Hungary, maize cultivation must cope with water shortages, as the difference between the water need of maize and the precipitation of the growing season is 40–150 mm on the maize croplands. The lack of water is the lowest (40–80 mm)

in the south-western Transdanubia, and the highest in the middle and southern parts of the Great Plain (100–150 mm) (*Nagy*, 2007).

2. *Material and methods*

2.1. Data acquired by on-site monitoring

On-site survey and monitoring were carried out in a systematic, cyclical way in the nearby maize seed producing area of Hajdúszoboszló between the years of 2004 and 2007. The net area of the agrarian fields was measured in the 2–3 leaf states of the vegetation, hence the not-sown fields (fallow fields, polders, etc.) were filtered out of further examinations. The homogeneous area samples, serving as the basis of the phenological observations, were chosen during the monitoring process. In the years of examination, 68 agrarian units were observed phenologically, twice a week on the average. The confinement of the fields was made by means of ProXH GPS and ArcPad x.x data collection softwares. The net area (the area covered with vegetation) was measured due to the unique cropping technology of the observed crop. The spatial uncertainty (2–5 m) was reduced by post-processing of the *.ssf files collected by the GPSCorrect-for-ArcPad auxiliary module. The accuracy was between 0.3 – 0.5 m.

2.2. Climatic data

The determination of climatic data was made by local measurements provided by the Hungarian Meteorological Service. Precipitation and conformation of global radiation and effective heat amount in the vegetative and generative periods were taken into consideration from the agrometeorological parameters. The GDD (1377–1660 °C) and average temperature (17.3 –18.7 °C) of the vegetative period in the years under examination proved to be optimal (*Table 1*).

Table 1. Climate characteristics of Hajdúszoboszló in the years of monitoring

	Favorable values	**2004**	**2005**	**2006**	**2007**
Growing degree days (GDD) (°C)	1250 – 1750	1377	1388	1442	1660
Average temperature of the growing season (°C)	16.8 – 19.0	17.5	17.3	17.6	18.7
Precipitation of the growing season (mm)	270 – 410	378	**455**	348	394
Precipitation of the winter term (mm)	230 – 310	**224**	**186**	237	151
Yearly precipitation (mm)	500 – 720	602	641	585	545
Heat unit – precipitation ratio (°C/mm)	1.9 – 3.1	2.3	2.2	2.5	3.0
Average humidity in July (%)	-	70.2	82.3	73.1	66.6

GDD are generally calculated by taking the average of the daily maximum and minimum temperatures compared to a base temperature, T_{base}. We note, that we used 8 temperature data measured by every 3 hours to improve the calculation of daily average, working with T_{base}=8 °C. The yearly precipitation amount fluctuated between 545 and 641 mm. The amount of precipitation in the growing season was optimal every year; however, the amount was low in the winter terms (151 and 224 mm) except the year of 2006. The heat unit–precipitation ratio (2.2–3.0 °C/mm) was in the optimum range in all the three years. The years 2006 (2.5 °C/mm) and 2007 (3.0 °C/mm) were dry, the years 2004 and 2005 (2.2–2.3 °C/mm) were wet for the maize on the basis of the heat unit–precipitation ratio.

In the period of pollen dispersal (July), the average humidity was pretty high (82.3%) in 2005, it was lower in the years of 2004 and 2006 (70.2% and 73.1%), and it was the lowest in 2007 (66.6%).

The monthly amount of precipitation was fairly diverse during the examination (*Fig. 1*):

- In 2004, the amount of precipitation of April, May, and September (40, 24, and 30 mm) was below, that of July (150 mm) was above the optimum value.
- In 2005, the amount of precipitation of June and July (64 and 68 mm) was below the optimal value; however, the quantity of all the other months exceeded its upper limit.
- In 2006, the sum of precipitation in June and August was above (84 and 82 mm) the upper limit of the optimum interval, but that of September was far below the lower limit.
- In 2007, April and June were featured by low precipitation (1 and 12 mm), while August and September were characterized by high precipitation (160 and 90 mm).

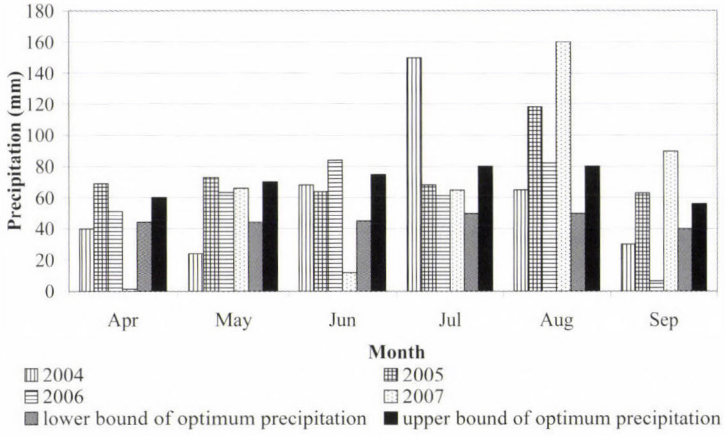


Fig. 1. Monthly amount of precipitation

2.3. Remote sensing data and method of evaluation

Landsat 5TM recordings transformed geometrically were provided by the Institute of Geodesy, Cartography and Remote Sensing, Budapest (see *Table 2*).

Table 2. Time of applied satellite recordings in the sample regions

Serial number of observation	2004	2005	2006	2007
1	May 3	May 31	May 11	May 21
2	June 20	June 16	June 26	June 22
3	July 22	July 25	July 12	July 15
4	August 23	August 10	August 22	August 16
5		September 4	September 14	

During quantitative analyses, information about the particular objects are the given radiation quantity of their characteristics as well as parameters and indicators calculated from them, thus we need to conduct the radiometric correction before processing data. COST model was used to make the atmospheric and radiometric correction (*Chavez*, 1996). In the model, firstly, radiance rates of the sensor have to be determined from the digital numerical values, then, in view of secondary data (zenith angle, Earth-Moon distance), the reference factor measured at the sensor can be calculated. During the correction process, the relief-deforming effect was ignored, since no significant difference was found in the altitudinal value of the territory examined. During pre-processing of the corrected recordings, thresholds of the cloudy area were defined based on the diffusion diagrams of the red channel, above which thresholds pixels have to be masked. Shades of the individual cloudy areas were picked in view of the particular areas.

2.4. Evaluation methodology

To survey the vegetation state of plants, the normalized vegetation index (NDVI) was calculated from the reflectance values of close infra and red channels (*Rouse et al.*, 1974). The possible instability in the NDVI values is accounted for the variance of field color (on surfaces with low vegetation) and humidity content, and also the most important climate attributes (precipitation, temperature, evapotranspiration, sunlight, and relative humidity). In the interval [–1,+1] the fields with healthy vegetation have their typical NDVI values between 0.2 and 0.8 (*Huete*, 1988), but on the spot, sampling and recordings are needed to evaluate the index values. Processing and evaluation using geographical information systems has been done using ArcGIS ArcInfo 9.1 and Erdas Imagine Professional 8.7 software environments.

The correlation between the NDVI and the heat unit is examined yearly, while the correlation between the NDVI and the yield has been further separated

to examination periods. For both cases regression analysis has been carried out. The correlation strength has been determined using the correlation coefficient R and determination coefficient R^2, and the validation of the regression model has been evaluated with the F-statistics of the variance of the estimations and residuals with a significance level of 5%. We have found correlation between NDVI and yield in every year but 2007. Statistical analysis is carried out using SPSS 13.0 program package.

3. *Results and discussion*

3.1. Evaluation of the correlation between the GDD and the NDVI

It is not possible to detect an unequivocal connection between growing seasons and calendar days. Instead of this, it is more suitable to choose the environmental parameter determining the growth of the crop. We have analyzed the change of the NDVI values depending on GDD.

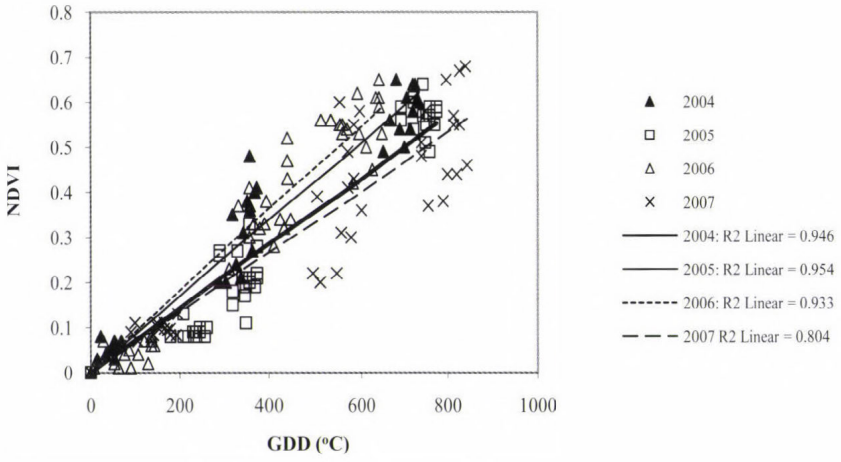


Fig. 2. Correlation between the GDD and the NDVI

The examination of the correlation between the GDD and the NDVI ended at the maximum point of the NDVI. It can be ascertained with the help of graphic trend analysis (*Fig. 2*), that the correlation between the GDD and the NDVI is positive in all the four years. The steepness of the trend lines changed yearly, it was the steepest in 2006, and the least steep in 2007. Examining the validity of the correlations between the GDD and the NDVI with regression analysis, a strong correlation for all the four years is obtained. The correlation coefficient was the largest in 2005 ($R = 0.977$), and the smallest in 2007

($R = 897$). The NDVI depended upon the GDD by 94.6 per cent in 2004 ($R^2 = 0.946$), 95.4 per cent in 2005 ($R^2 = 0.954$), 93.3 in 2006 ($R^2 = 0.933$), and 80.4 in 2007 ($R^2 = 0.804$) (*Table 3*). The variance of the residuals and the estimated values differed significantly in all the four years, so regressions are verifiable. The F-statistics has the biggest value in 2005 ($F = 1078$), and the smallest in 2007 ($F = 164$), i.e., the correlation between the GDD and the NDVI-index was the strongest in 2005.

Table 3. Results of the regression analysis between the GDD and the NDVI (R = correlation coefficient, R^2 = determining coefficient, F = F-statistics, *Sig* = significance, (*) = the variance of estimated and residual values differs on a 5% level of significance)

Year	*R*	R^2	*F*	*Sig*
2004	0.973	0.946	648.007	0.000 (*)
2005	0.977	0.954	1078.400	0.000 (*)
2006	0.966	0.933	630.197	0.000 (*)
2007	0.897	0.804	163.792	0.000 (*)

3.2. Evaluation of the correlation between the NDVI and the yield (t/ha)

On the basis of the graphic trend analysis (*Fig. 3*), the correlation was positive in all the seven cases, i.e., the yields increased with the growth of the NDVI. The correlation between the NDVI and the yield was most significant in the fourth observation time in 2004, in the third and fourth in 2005, and in the third in 2006. The steepness of the trend lines changed yearly and with observation times; it was the steepest in the third observation time in 2006.

From *Table 4* it can be seen that in 2004, observations 1 and 2 showed weak correlation ($R = 0.199$ and 0.287) between the NDVI and the yield, stronger ($R = 0.330$) for observation 3, but the strongest ($R = 0.687$) for observation 4. The values of the F-statistics showed significant and real correlation (*Sig* = 0.009) only for observation 4, while the values are higher than the chosen significance level (0.05) for the other ones. In our tables we detail the results for observations with significant correlation (for example in *Table 4* for observation 4). In observation 4, 47.2% ($R^2 = 0.472$) of the yield is described by the NDVI, and based on the standard error of estimation ($SEE = 1.01$) it can be estimated with an accuracy of 1.01 t/ha. The regression coefficient significantly ($Sig = 0.009$) differs from zero, whereas the regression constant does not ($Sig = 0.405$), therefore, the regression equation is $Y = 10.178 \cdot X$, where Y is the estimated yield and X is the NDVI value.

In 2005, the correlation coefficient is rather small for the first two observations ($R = 0.222$ and 0.067) but stronger for observations 3, 4, and 5 ($R = 0.638$, 0.673, and 0.410). The variance of the estimated values and residuals differed significantly from zero only for observations 3 and 4 ($Sig = 0.004$ and

0.002). In observation 3, the 40.7% of the yield is described by the NDVI, and it can be estimated with an accuracy of 0.98 t/ha (see *Table 5*). The slope and the regression coefficient significantly differ from zero, thus we can accept that the regression equation $Y = 22.477 \cdot X - 9.649$ is valid. In observation 4, the 45.3% of the yield is described by the NDVI with an accuracy of 0.94 t/ha. The regression equation is $Y = 23.479 \cdot X - 10.221$.

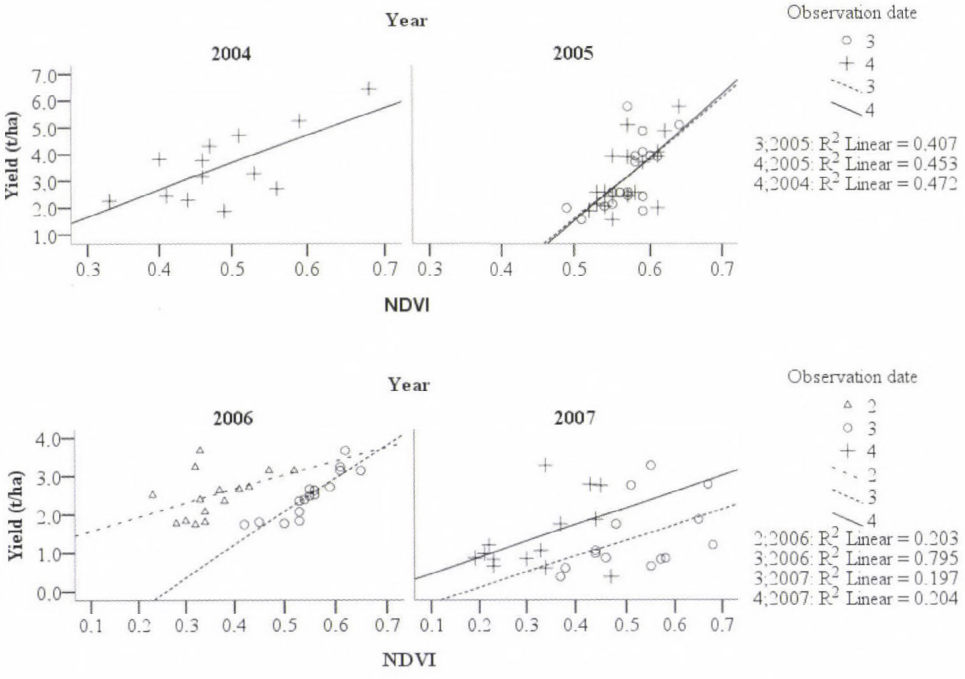


Fig. 3. Correlation between the NDVI and the yield (2004–2007)

Table 4. Regression analysis table between the NDVI and the yield (t/ha) for the year 2004 (*B* = non-standardized coefficient, *SE* = standard error, *T* = t-statistics)

Observation date	***R***	***R²***	***SEE***	***F***	***Sig***
1	0.199	0.039	-	0.451	0.516
2	0.287	0.082	-	0.985	0.342
3	0.330	0.109	-	1.344	0.271
4	0.687	0.472	1.01	9.827	0.009 (*)
Observation 4		***B***	***SE***	***T***	***Sig***
Regression constant (slope)		-1.390	1.606	-0.866	0.405
Regression coefficient		10.178	3.247	3.135	0.009

Table 5. Regression analysis table between the NDVI and the yield (t/ha) for the year 2005

Observation date	***R***	R^2	***SEE***	***F***	***Sig***
1	0.222	0.049	-	0.829	0.376
2	0.067	0.004	-	0.071	0.793
3	0.638	0.407	0.98	10.986	0.004 (*)
4	0.673	0.453	0.94	13.233	0.002 (*)
5	0.410	0.168	-	3.230	0.091
Observation 3		***B***	***SE***	***T***	***Sig***
Regression constant (slope)		-9.649	3.88	-2.487	0.024
Regression coefficient		22.477	6.781	3.314	0.004
Observation 4		***B***	***SE***	***T***	***Sig***
Regression constant (slope)		-10.221	3.693	-2.768	0.014
Regression coefficient		23.479	6.454	3.638	0.002

In 2006, observations 1, 4, and 5 show weak correlation ($R=0.186$, 0.219, and 0.237) between the NDVI and the yield, stronger ($R=0.451$) for observation 2, but the strongest ($R=0.892$) for observation 3 (see *Table 6)*. The values of the F-statistics showed significant and real correlation only for observation 3, while the values are higher than the chosen significance level for all the other ones. In observation 3, 79.5% of the yield is described by the NDVI and it can be estimated with an error of 0.27 t/ha. The slope and the regression coefficient significantly differ from zero, thus we can accept that the regression equation $Y=8.625\cdot X-2.232$ is valid.

Table 6. Regression analysis table between the NDVI and the yield (t/ha) for the year 2006

Observation date	***R***	R^2	***SEE***	***F***	***Sig***
1	0.186	0.035	-	0.502	0.490
2	0.451	0.203	-	3.314	0.092
3	0.892	0.795	0.27	54.336	0.000 (*)
4	0.219	0.048	-	0.654	0.433
5	0.237	0.056	-	0.836	0.376
Observation 4		***B***	***SE***	***T***	***Sig***
Regression constant (slope)		-2.232	0.647	-3.448	0.004
Regression coefficient		8.625	1.170	7.371	0.000

In 2007, observations 1 and 2 show weak correlation ($R=0.078$ and 0.244) between the NDVI and the yield, while stronger ($R=0.444$ and 0.452) for observations 3 and 4. The variance of the estimated values and residuals does

not differ significantly in either case, so no evidence for the correlation between NDVI and the yield from 2007 has been found (*Table 7*).

Table 7. Regression analysis table between the NDVI and the yield (t/ha) for the year 2007

Observation date	R	R^2	F	*Sig*
1	0.078	0.006	0.073	0.792
2	0.244	0.060	0.762	0.400
3	0.444	0.197	2.942	0.112
4	0.452	0.204	3.073	0.105

In *Table 8* we can see that the standard variation of the yield was the smallest (0.226) in 2006, higher in 2004 and 2005 (0.364 and 0.379 respectively), and the highest (0.632) in 2007 saying that the crop was the most homogenous in 2006 and the least homogenous in 2007. According to the investigations, the higher the standardized deviation the weaker the correlation between the NDVI and the yield. In 2007, the low yield and the heterogeneity of it can be explained by a drought period and high temperature during blooming that caused low fertility.

Table 8. The standard variation of the yield and the correlation between the NDVI and the yield

Year	Standard variation	R^2
2006	0.226	0.892
2005	0.379	0.673
2004	0.364	0.687
2007	0.632	0.452

4. Conclusions

It has been shown that estimating plant growth using satellite records can be applied in a new field, namely in maize seed production. Estimated production results of crop fields were derived from processed data. During growing season, especially at blooming, one can examine the heterogeneity and the fertility of crop mass using the files provided. Applying the method for longer time series, the most suitable fields for crop production and high productivity cultures can be identified with high probability. Providing exact information quickly has become a basic requirement of reliable agricultural production. Data collected by remote sensing provides information that can be used widely in cultivation both in the technological and decision support processes areas. The technologies in industrial processing require constant quality of input goods. It is essential to

set up and improve the monitoring system of production (optimal selection of fields, yield estimation, and effective identification of negative processes during plantation growth).

Acknowledgements—This work is connected to the scientific program of the "Development of quality-oriented and harmonized R+D+I strategy and functional model at BME" project. This project is supported by the New Hungary Development Plan (Project ID: TÁMOP-4.2.1/B-09/1/KMR-2010-0002).

References

Ángyán, J., 1985: Mass production forage maize production – Geographical location of maize production (in Hungarian). In *A kukoricatermesztés kézikönyve* (ed.: *Z. Menyhért*). Mezőgazdasági Kiadó, Budapest, 199-228.

Bocz, E., 1996: Maize – Climatic requirements, soil requirements, soil tillage, crop rotation (in Hungarian). In *Szántóföldi növénytermesztés* (eds.: *E. Bocz, I. Késmárki, A. Kováts, L. Ruzsányi, M. Szabó*). Mezőgazdasági Kiadó, Budapest, 379-387.

Chavez, P.S., 1996: Image-based atmospheric corrections – Revisited and improved. *Photogrammetric Engineering and Remote Sensing 62*, 1025-1036.

Huete, A.R., 1988: A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ. 25*, 295-309.

Keszthelyi, S., 2005: Effect of the climatic factors of year 2004 on the development of maize, and the appearance of its pests and their damage (in Hungarian). *Agrofórum 16*, 10, 3-7.

Nagy, J., 2006: Multifunctional agriculture (in Hungarian) In *Területfejlesztés, agrárium és regionalitás Magyarországon* (eds.: *B. Baranyi, J. Nagy*). DE-AMTC & MTA-RKK, Debrecen, 191-206.

Nagy, J., 2007: *Maize Production* (in Hungarian) *Kukoricatermesztés*. Akadémiai Kiadó, Budapest.

Rouse, J.W., Haas, R.H., Schell, J.A., 1974: Monitoring vegetation systems in the Great Plains with ERTS. In *Third Earth Resources Technology Satellite-1 Symposium.* December 10-14, 1973, NASA SP-351, Washington, D.C. (eds.: *S.C. Freden, E.P. Mercanti, M.A. Becker*). NASA STIO, Goddard SFC and Washington, D.C., 309-317.

Szász, G., 2002: Utilization of climatic natural resources in domestic crop production (in Hungarian). *Agrártudományi Közlemények – Acta Agraria Debreceniensis 9*, 101-106.

Turcsányi, G., 1985: Maize morphology, histology and physiology – Maize physiology (in Hungarian). In *A kukoricatermesztés kézikönyve* (ed.: *Z. Menyhért*). Mezőgazdasági Kiadó, Budapest, 58-76.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 4, October–December 2010, pp. 275–286

Ensemble calibration of ECMWF's medium-range forecasts

István Ihász[1*], **Zoltán Üveges**[1], **Máté Mile**[1], and **Csilla Németh**[2]

[1]*Hungarian Meteorological Service,*
P.O. Box 38, H-1525 Budapest, Hungary; E-mail: ihasz.i@met.hu

[2]*Department of Meteorology, Eötvös Loránd University,*
P.O. Box 32, H-1518 Budapest, Hungary; E-mail: csidori@gmail.com

[*]*Corresponding author*

(Manuscript received in final form November 4, 2010)

Abstract — A post-processing method is presented, that is aimed at enhancing the value of ensemble forecasts through ensemble calibration. This approach can reduce model errors. In March 2008, ECMWF introduced the VarEPS/Monthly system. Part of this service provides a weekly reforecast dataset for the last 18 years. This dataset has been successfully applied for ensemble calibration at the Hungarian Meteorological Service. The calibration was performed for surface meteorological variables such as 2m temperature, minimum and maximum temperature, 10m wind speed, and daily precipitation. An objective verification was performed on the basis of several deterministic and ensemble verification scores. The calibration method was introduced operationally in mid-2009.

Key-words: ensemble calibration, medium–range weather forecasts, reforecast dataset, model climate, ensemble verification

1. Introduction

The ECMWF, established by 18 European countries in 1973, is the world leader of numerical weather prediction providing highly reliable medium-range forecasts. In 1994, Hungary was the first Eastern European country to sign a co-operation agreement with the ECMWF (*Woods*, 2006) and the Hungarian Meteorological Service (HMS) has been using the full set of operational forecasts ever since. In the last 15 years, the HMS made several local developments to provide additional tools for forecasters and users to help decision making and improving the quality of the forecasts, especially in

medium-range and extended medium-range scales (*Ihász*, 2008). In this paper a new development, the so-called ensemble calibration post-processing method is presented.

Calibration of ensemble forecasts is a new challenge in meteorology (*Hamill* and *Whitaker*, 2007), and in most cases, different techniques are used for different types of meteorological variables depending on their probability distribution. Even though a number of derived ensemble products (e.g., forecast probabilities) are provided directly by the ECMWF, its Member States are encouraged to make local developments. Therefore, since 2008 reforecast products have been available for calibration purposes (*Hagedorn*, 2008), whilst the calibration itself is not performed by the ECMWF. We have successfully applied a simple calibration method for several meteorological variables such as 2m temperature, daily precipitation, 10m wind speed, which have completely different distribution functions. The method and the verification results are presented in Sections 2, 3, and 4 of this paper. This method was introduced operationally in the second part of 2009 at the HMS.

2. Calibration of ensemble forecasts

Although numerical forecasting models have gone under huge development, their raw forecasts still contain errors stemming from a wide range of sources. The influence of these errors can be reduced to some extent by post-processing methods. In this part of the paper, a quite simple but effecient calibration method will be shown. This method has been developed and applied at the Hungarian Meteorological Service in order to make ECWMF's medium-range ensemble forecasts more reliable.

At first we take a deeper look into the general procedure of calibration, which is based on the comparison of statistical distributions of the model variables and observations. In the second part, deep investigations will be presented on the distribution functions of model forecasts in order to size up the model's behavior in case of different meteorological parameters. Then we will try to demonstrate the effectiveness of calibration by representing a couple of examples of calibrated and verified model results as well as a case study.

3. Method of the calibration

As it was mentioned above, the chosen calibration method is based on the comparison of empirical distribution functions. Thus, the adequate construction method and data quality of these functions have primary importance.

The classical and the most simple method is to create a model climate using ensemble forecasts of the past couple of years. As a matter of fact, the model climate is a collection of forecasts with start and prediction dates from the past,

containing data for a considerable number of years (*Hamill* and *Whitaker*, 2007; *Hagedorn*, 2008).

However, it is important to note that there are always some differences between past and recent behavior of the model due to model developments. To avoid difficulties caused by this fact, reforecast datasets can be generated operationally at the largest centers (*Hagedorn*, 2008). Let us see a short summary of the model climates generated at the ECMWF.

The necessity of generating and using a set of runs to define the model climate was addressed more than fifteen years ago. Firstly, the reanalysis projects provided model climates. At the ECMWF, the ERA-15 (*Gibson et al.*, 1997), ERA-40 (*Uppala et al.*, 2005), and ERA-Interim (*Simmons et al.*, 2007) products provided model climate for the deterministic model, but they were not relevant for ensemble forecasts. In the beginning of the first decade of the 21st century, a new product, the Extreme Forecast Index (EFI) was developed (*Lalaurette* and *Grijn*, 2003), which also needed model climate. The first version of EFI used the classical method for generating model climate between the late 2003 and early 2006. In November 2006, daily operational forecasts were extended from 10 to 15 days by the introduction of the VarEPS system (*Buizza et al.*, 2006). Between 2006 and 2008, the model climate was operationally generated daily, by running a reforecast of the control model of ensemble system only up to +48 hours for the period of 1971–2000. This was the first model climate that was adequate enough to be used for calibration purposes. By the middle of March 2008 VarEPS was merged with the former monthly forecast. At the same time, a new reforecast method was introduced, i.e., a 5-member ensemble reforecast dataset was operationally produced weekly for the last 18 years (*Hagedorn*, 2008). To increase the representativeness, the distributions of model climates can be generated from 5 consecutive weekly reforecast datasets that can be obtained from ECMWF's Meteorological Archival and Retrieval System (MARS) (*Raoult*, 2001).

Hagedorn (2008) proved the usefulness of calibration for 2m temperature, and this encouraged the members to develop and apply calibration procedures for their local needs. *Hagedorn* (2008) used non-homogeneous Gaussian regression method (*Hamill* and *Whitaker*, 2007) based on daily values of archived observations and model variables. Verification results showed that the success of calibration is partly originated from simple bias correction, but additional improvements could be achieved by ensemble calibration. The 2m temperature ensemble forecast was calibrated for 250 European stations. A verification measure, the continuous rank probability score (CRPS) showed that calibrated EPS forecasts for day 4 are as good as uncalibrated forecasts for day 1. The disadvantage of this method is that it can only be used for meteorological variables, which have Gaussian distribution. Another type of calibration method must be applied for non-Gaussian meteorological variables, e.g., precipitation (*Hamill et al.*, 2007).

Since national meteorological services have real-time access to data coming from the most dense network of reliable observations, the best approach is to make calibration at the national level. Beside the model climate, a reliable observed climate is also needed for successful calibration which can be provided by the national meteorological services. It has been noticed that it is essential to match the reference interval of the model and observed climates since different reference intervals. Serious inaccuracies can be caused in calibrated forecasts as a result of slight climate changes. Model climate can be locally calculated at meteorological services with the most appropriate method for their local needs.

In the beginnings of 2007, HMS decided to develop a calibration algorithm for EPS forecasts, to test and verify the system as well. In the first stage of the work, model climate was created from the +48-hour reforecasts of the control model, for the period of 1971-2000 *(Ihász*, 2007; *Ihász* and *Mile*, 2008). Forecast calibration was aimed only between +24 and +120 hours (*Mile*, 2008a, 2008b). The calibrations were made on five meteorological parameters: 2m minimum and maximum temperatures, 2m temperature at 12 UTC, daily accumulated precipitation, 10m wind speed at 12 UTC. Nine reliable Hungarian synoptical stations (Békéscsaba, Budapest, Miskolc, Nagykanizsa, Nyíregyháza, Pécs, Siófok, Szeged, and Szombathely) were selected for the investigations.

It has been mentioned that in the middle of March 2008, a brand new EPS system was launched. In addition, a new model climate with totally different properties (such as reference interval, number of ensemble members, forecast period, etc.) became available. It was worth to include products of the new reforecast system into our calibration process.

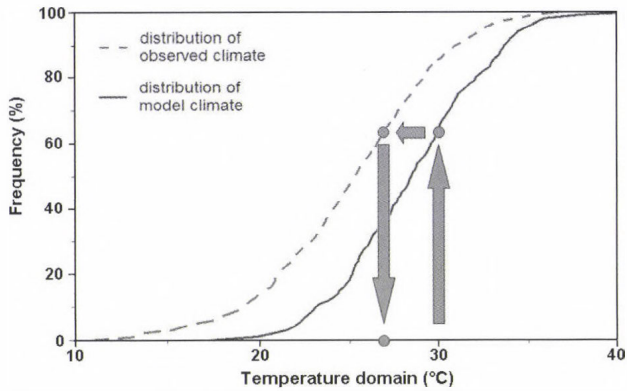


Fig. 1. Example of calibration method.

To understand our calibration method itself, procedure is illustrated in *Fig. 1.* Suppose that one of our ensemble members predicts a temperature of 30 °C. Firstly, we have to find the frequency in the model climate that belongs to the value of 30 °C. Then, this frequency must be projected onto the observed

climate. And finally, we get a value belonging to this frequency which is the new value of our ensemble member. As a matter of fact, we assume that the distribution of observed values describes weather conditions more accurately in a given place than the distribution of values forecasted by a numerical model. Thus, instead of a simple shifting of the curve of distribution functions, like in case of bias correction, we look for the frequency of each EPS value both in the model- and observed climates.

4. Comparison of modeled and observed climate distributions

Before calibrating forecasts, deeper investigations were made on the model climates in order to scale the model's behavior in case of different parameters.

4.1. Comparison of recent and former model climates

The model climates were compared (the recent one having a reference interval of 1990–2007 and the former one with an interval of 1971–2000) so that the climate changes as well as the differences between model versions could be found (*Üveges*, 2009a, 2009b). It can be clearly seen that the difference is in some cases is greater than 1.5 °C.

4.2. Comparison of model climate and observations

In the following part of the investigation, the systematic errors of the model could be demonstrated by comparing the model climate with observations at synoptic stations. All of the graphs were made from reforecast climates with a time step of 180 hours and from the corresponding observated climates. The left side of *Fig. 2* shows an example for distribution of daily precipitation.

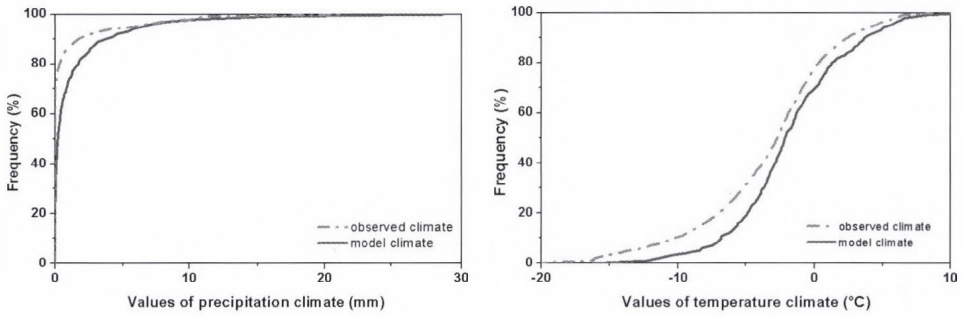


Fig. 2. Monthly distributions of 24-hour precipitation for Szombathely in January (left) and monthly distributions of daily minimum temperature for Nagykanizsa in January (right).

Overestimation can be observed in daily precipitation in January in the case of Szombathely, a town near the Austrian border at the foot of the Alps. In summer, the model climate is colder than the observed one. This difference may be caused by convective precipitation. Lowest temperatures can be seriously affected by local conditions which cause significant systematic errors in the forecasts (see the curves showed on the right side of *Fig. 2*).

5. Comprehensive verification of uncalibrated and calibrated ensemble forecasts

A wide range of methods is available for ensemble verification (*Persson* and *Grazzini*, 2007). During our investigation, the reliability diagram, ROC diagram, Brier (skill) score, and calculation of outliers were used.

Since the monthly ensemble forecasts are produced only once a week, and during our study they were available altogether from April 2008 till March 2009, the amount of forecast values was insufficient to get reliable verification results in some cases. Thus, only 15-day forecasts were verified and joint verification was made.

5.1. 2m temperature at 12 UTC

The first calibrated and verified parameter is temperature at 12 UTC. The results are displayed among others in Talagrand diagrams, which provide useful information about the relationship of ensemble spread and observations. In uncalibrated cases, J-shape diagrams can be seen on both graphs caused by accumulation of the observed values in the lower temperature domain (*Fig. 3*). Thus, the ensemble system makes a constant overestimation of temperature at 12 UTC. In the case of calibrated forecasts, the distributions of observations are much flatter, i.e., the calibration managed to reduce the error discussed above.

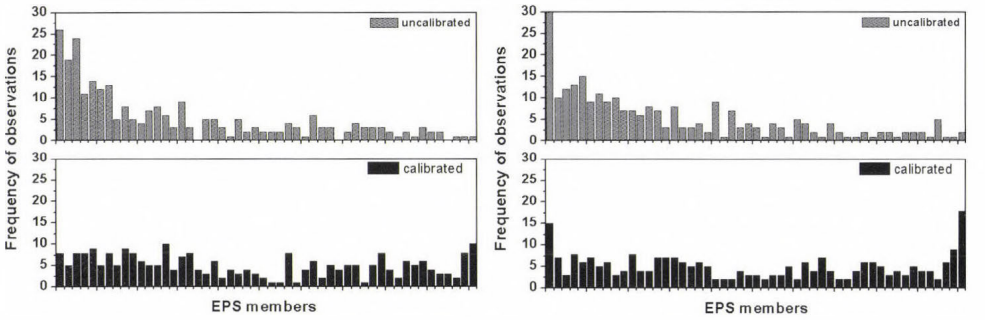


Fig. 3. Talagrand diagrams of 2m temperature forecasts at 12 UTC for Miskolc (left) and Siófok (right).

5.2. *Daily maximum temperature*

Probabilistic forecasts can be transformed into a categorical yes/no forecast defined by some probability thresholds as well as meteorological events. For different thresholds the corresponding hit rates and false alarm rates can be computed. A very powerful way to display and interpret this information is the two-dimensional relative operating characteristics or ROC diagram. A point in the ROC diagram is defined with the POFD (or False Alarm rate) on the x-axis and POD (or HR) value on the y-axis. The upper left corner of the ROC diagram represents a perfect forecast system where there are no false alarms and only hits. The closer the point is to this upper left corner the higher the skill is. The lower left corner, where both hit and false alarm rates are zero, represents a system which never warns of a predefined event.

Although many kinds of verification methods have been tried, maximum temperature is the only parameter for which no significant improvement has been observed.

5.3. *Daily minimum temperature*

One of the best methods for displaying the verification results of an ensembe forecast is making reliability diagrams. They are simple graphs of the observation frequency (y-axis) of an event against the forecast probability of the same event (x-axis). This graph effectively tells how often a forecast probability actually occurred. In a perfect forecast system, for example, a meteorological event forecasted with a probability of 10%, should occur one out of 10 times. The curves of the graph connect the intersections of probabilities and the corresponding observation frequencies. Thus, in case of a reliable ensemble forecast system, the curves have low deviation from the diagonal of the graph box. One of the predefined events was the minimum temperature value exceeding 5 °C. The results for two locations are displayed in *Fig 4.*

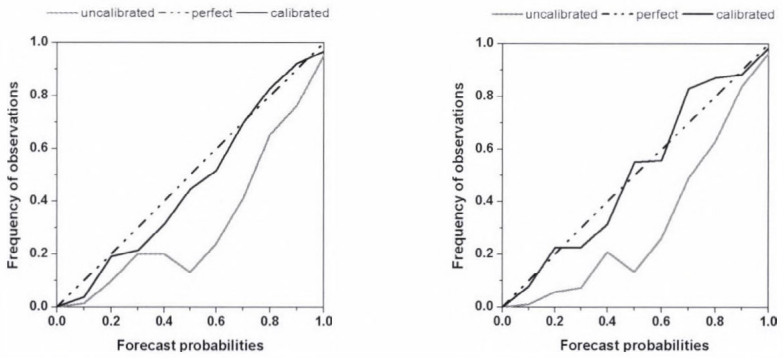


Fig. 4. Reliability diagrams of daily minimum temperature forecasts for Budapest (left) and Siófok (right).

In both cases, uncalibrated forecasts show serious overestimation of the event. However, the calibrated forecasts illustrated with red lines are much closer to the diagonal, which means that the calibration resulted in much more reliable forecasts. The synoptic station of Siófok is located on the shore of the Lake Balaton, the other one in Budapest is located in a quite dense urban area. It seems like the local conditions are taken into account more precisely in the case of the calibrated forecasts.

5.4. Daily precipitation

As it is well known, precipitation is much more variable in space and time than temperature. Precipitation is a non-continuous meteorological variable, therefore, making a successful forecast and verification is quite challenging, especially in summer, when convection can be quite intense.

Brier skill scores (BSS) were calculated for verifying precipitation forecasts. If BSS equals 1, the forecast is perfect. If it equals 0, the forecast has no skill compared to a reference forecast, in this case to the uncalibrated forecast. In case of negative BSS values, the examined forecast is worse then the reference. The defined meteorological event was the occurence of daily precipitation greater than 1 mm. As we saw at the examples of systematic errors, EPS forecasts are likely to overestimate the amount and frequency of precipitation. Since positive BSS values can be seen on the graphs in all cases, the calibration process reduced the error (*Fig 5*).

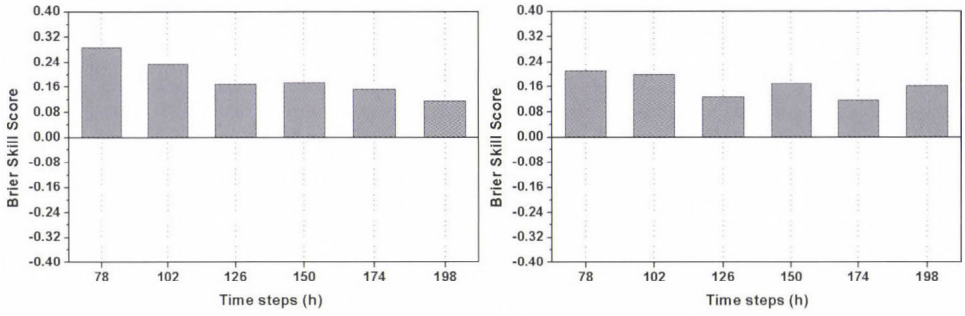


Fig. 5. Computed BSS values of daily precipitation forecasts for Pécs (left) and Szeged (right).

5.5. 10m average wind speed at 12 UTC

The surface wind speed strongly depends on surface conditions. For this variable, calibrated wind speed forecasts showed the greatest improvement. This is demonstrated by the graphs displaying mean square error (MSE) values. MSE is the squared deviation of a forecast value from the observation.

Reduction in MSE values of wind forecasts were more succesful where the surface conditions are highly variable. The ensemble forecast of ECMWF takes into account the mountains and hills to an unsufficient extent because of its horizontal resolution. This is the reason for the big improvement of the forecasts at Miskolc at the foot of the Northern Mountains (*Fig. 6*). On the contrary, calibration has hardly any effect in Nyíregyháza, which is located in the Great Plain.

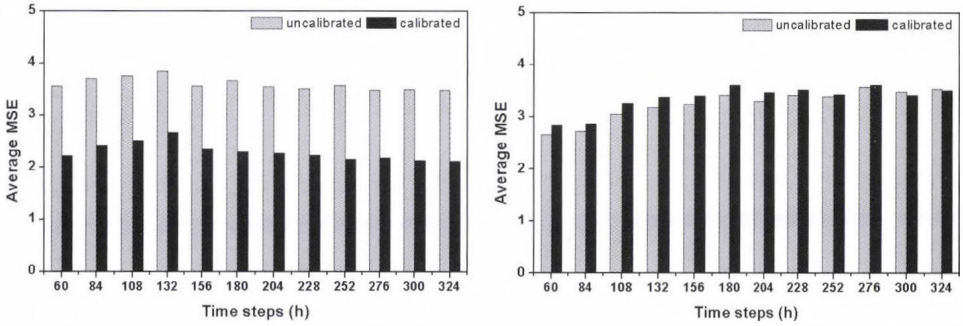


Fig. 6. MSE values of the 10m average wind speed forecasts at 12 UTC for Miskolc (left) and Nyíregyháza (right).

5.6. Case study

Case studies were also made in order to prove the effectivenes of calibration. At the beginning of 2009, a strong and stable inversion layer developed over the Carpathian Basin. Consequently, the air pollution became more and more critical during the period, and the mayor of Budapest had to issue a smog alert in the city. It can be clearly seen in *Fig. 7*, that the observations marked with red points are located much closer to the ensemble mean in the case of calibrated forecasts, i.e., if the process of calibration had been used operationally at the HMS, much better wind speed forecast could have been done for the critical period.

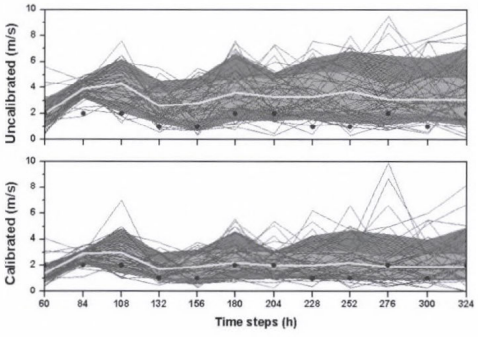


Fig. 7. EPS plumes of the 10m average wind speed forecasts at 12 UTC, January 1, 2009, in Budapest.

5.7. Grid based calibration

After having encouraging verification results concerning to the selected synoptical stations, it was considered to extend calibration for the regular 0.5 by 0.5 degrees grid belonging to EPS model resolution valid in 2009. The area of the country is 93,030 km^2, it is covered by approximately 70 grid points, so 70 stations were selected for providing 'observed' climate distributions for all grid points.

The largest part of the country is flat, while in the mountainous regions the observation density is not completely enough for providing perfect interpolation for ensemble grid, therefore, the observed climate distribution of each gridpoints is represented by distribution of the closest observation. The method of the calibration was exactly the same as in case of station based calibration. An important advantage of the grid based calibration is that uncalibrated and calibrated meteorological fields can be easily visualized, and local forecasts can be easily derived for end users.

Firstly, the 2m temperature was chosen to test the new grid based calibration, because this variable is quite sensitive to the influence of the orography (*Németh*, 2010). Verification of calibration was made for 2009. Calculation of the monthly mean of raw and calibrated ensemble forecasts in addition to observed values was made for each month of 2009. Maps belonging to each month showed that calibration forecasts are definitely closer to the reality, especially in the mountainous regions (*Fig. 8*).

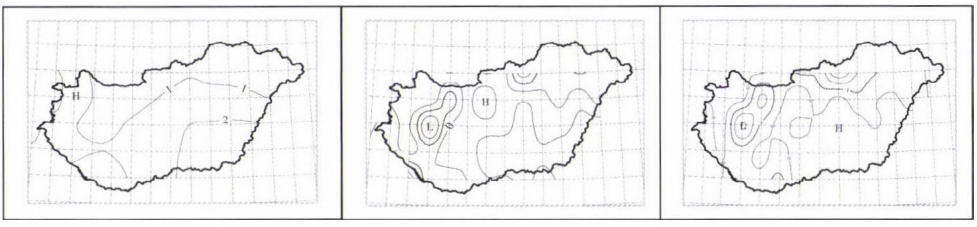


Fig. 8. Monthly mean chart of raw (left), calibrated (middle), and observed (right) values for Hungary in January 2009.

Mean error and root mean square error maps were also made between the raw and observed values beside the calibrated and observed values. Finally, outlier maps were used to clearly show the regions where under or overestimation appeared.

The ensemble spread was found not sufficiently wide at lower time steps, it was good just after an approximately 3-day lead time. At all time steps between +12 and +180 hours, critical mountainous regions show systematic overestimation of the uncalibrated forecast, that could be eliminated well by calibration.

To demonstrate that ensemble calibration could particularly improve the forecast in mountainous regions. a grid point was chosen in the area in Bakony hill in the transdanubian region. It can be clearly seen that calibration successfully eliminates the systematic error (*Fig. 9*).

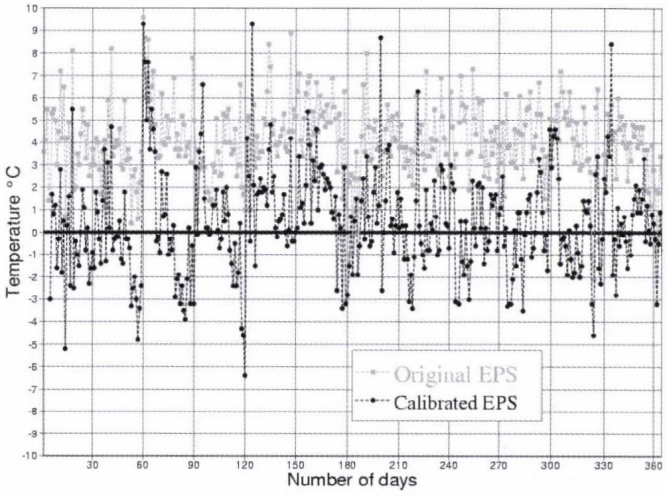


Fig. 9. Time series of forecasted 12 UTC temperature based on raw (red) and calibrated (green) temperature at a selected mountainous grid point (47.0° N, 17.5° E).

6. Summary, conclusions

In this paper, a calibration method developed and used at HMS for post-processing of ensemble forecasts was presented. It has been shown that the method significantly improves the forecast efficiency on regional and local scales. This method was intensively investigated and verified for five important meteorological variables: 2m temperature, daily minimum and maximum temperatures, daily precipitation, and 10m average wind speed.

In the new calibration system, a new type of reforecast datasets was used. The model climate based on 5 member EPS reforecasts was compared to an older one and the observed climate as well, in order to examine the special characteristics over the territory of Hungary. Verification of calibrated forecasts shows that in most cases the calibrated products are more skillful than the raw forecasts.

Acknowledgements — Authors are very grateful to *László Hercsényi*, *Rita Hodossy,* and *Zsuzsa Kövér* for kindly providing the monthly observed distribution dataset by Software Development Division of Telecommunication and Informatics Department. We also thank *Gábor Radnóti* and *Miklós Vörös* for useful suggestions and comments on work and reviewing the manuscript. This work was partly supported by the Hungarian National Office for Research and Technology (NKFP, grant No. 3A/051/2004).

References

Buizza, R., Bidlot, J-R., Wedi, N., Fuentes, M., Hamrud, M., Holt, G., Palmer, T., Vitart, R., 2006: The ECMWF Variable Resolution Ensemble System (VAREPS). *ECMWF Newsletter 108,* 14-20.

Gibson, J.K., Kallberg, P., Uppala, S., Hernandez, A.., Nomura, A. and *Serrano, E.,* 1997: ERA-15 description. *ERA-15 Project series 1.* ECMWF.

Hamill, T.M. and *Whitaker, J.S.,* 2007: Ensemble calibration of 500 hPa geopotential height and 850 hPa and 2-metre temperature using reforecasts. *Mon. Weaather Rev. 135,* 3273-3280.

Hamill, T.M., Hagedorm, R. and *Whitaker, J.S.,* 2007: Probalistic Forecast Calibration: Using ECMWF and GFS Ensemble Reforecasts. Part II: Precipitation. *Mon. Weather Rev. 135,* 2620-2632.

Hagedorn, R., 2008: Using the ECMWF reforecast dataset to calibrate EPS forecasts. *ECMWF Newsletter 117,* 8-13.

Ihász, I., 2007: Experiences using VarEPS products at the Hungarian Meteorological Service. *Proceedings of the Eleventh ECMWF Workshop on Meteorological Operational Systems.* Reading, UK, 14-18 November 2007, 130-135.

Ihász, I. and *Mile, M.,* 2008: Calibration of ECMWF ensemble precipitation forecasts for hydrological purposes at the Hungarian Meteorological Service. *Proceedings of the XXIV Conference of the Danubian Countries.* 2-4 June 2008, Bled, Slovenia.

Ihász, I., 2008: Model products of medium and long range forecasts and possibilities for applications (in Hungarian). *Proceedings of 34th Scientific Meteorological Days at Hungarian Scientific Academy,* 20-21 November 2008. Budapest, Hungary.

Lalaurette, F. and *van der Grijn, G.,* 2003: Ensemble forecasts: Can they provide useful early warnings? *ECMWF Newsletter 96,* 10-18.

Mile, M., 2008a: Subsequently calibration of ensemble forecasts of the ECMWF (in Hungarian). *Proceedings of the Scientific Student Conference on Environment.* 25-26 March 2008, Nyíregyháza, Hungary.

Mile, M., 2008b: Subsequently calibration of ensemble forecasts of the ECMWF (in Hungarian). *Master Thesis,* Eötvös Loránd University, Budapest, Hungary.

Németh, Cs, 2010: Verification of the calibrated probability forecasts of ECMWF's ensemble model (in Hungarian). *BSc Thesis,* Eötvös Loránd University, Budapest, Hungary.

Persson, A. and *Grazzini, F.,* 2007: *User Guide to ECMWF Forecasts Products: Meteorological Bulletin, 3.2 (version 4.0),* ECMWF, Reading, UK.

Raoult, B., 2001: MARS on the Web: a virtual tour. *ECMWF Newsletter 90,* 9-17.

Simmons, A., Uppala, S., Dee, D. and *Kobayashi, S.,* 2007: ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter No. 110,* 25-35.

Uppala, S.M., Kallberg, P.W. Simmons, A.J., Andrae, U., Bechtold, V.D., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A.;Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.F., Morcrette, J.J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., Woollen, J., 2005: The ERA-40 Reanalysis. *Q. J. Roy. Meteor. Soc. 131,* 2961-3013.

Üveges, Z., 2009a: Calibration of monthly forecasts of the ECMWF (in Hungarian). *Proc. of the 29th Scientific Student Conference (OTDK),* Szombathely, Hungary. March 23-24, 2009.

Üveges, Z., 2009b: Calibration of monthly forecasts of the ECMWF.(in Hungarian). *Master Thesis,* Eötvös Loránd University, Budapest, Hungary.

Woods, A., 2006: *Medium-Range Weather Prediction The European Approach.* Springer.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 4, October–December 2010, pp. 287–302

Modeling of the urban heat island pattern based on the relationship between surface and air temperatures

J. Unger[1*], T. Gál[1], J. Rakonczai[2], L. Mucsi[2], J. Szatmári[2], Z. Tobak[2], B. van Leeuwen[2], and K. Fiala[3]

[1]*Department of Climatology and Landscape Ecology, University of Szeged, P.O. Box 653, H-6701 Szeged, Hungary; E-mail: unger@geo.u-szeged.hu*

[2]*Department of Physical Geography and Geoinformatics, University of Szeged, P.O. Box 653, H-6701 Szeged, Hungary*

[3]*Directorate for Environmental Protection and Water Management of Lower Tisza District, P.O. Box 390, H-6701 Szeged, Hungary*

[*]*Corresponding author*

(Manuscript received in final form July 27, 2009)

Abstract—The aim of this study is to develop a new – and easy to use – method for early night-time near-surface air temperature pattern estimation based on surface temperature data in urban areas. The surface temperature data have been collected by an airborne thermal infrared sensor at an altitude of 2000 m above ground level. The study area was covered by hundreds of images with a spatial resolution of about 2 m. The measured values were calibrated with data of in situ surface measurements of different land use types. Simultaneous air temperature measurements were carried out using a car-based temperature sensor along an almost 12 km long N-S urban transect. The measured points were located using a GPS device. Data were processed with GIS methods, including newly developed algorithms. In order to find the relationship between air and surface temperatures a wider environment, the source area which determines the air temperature at a given point and time was taken into account. Using a source area with a radius of 500 m, a strong relationship was detected between the two parameters. Namely, the temperatures of the surfaces found in the surroundings (weighted by the distance) determine the temperature of the air parcel located at a given point. The obtained regression equation was applied to extend our results in order to model the air temperature field in a larger urban area.

Key-words: urban environment, surface and air temperatures, remote sensing, source area, Szeged, Hungary

1. Introduction

The thermal features of settlements are different from their natural surroundings. This phenomenon is related to the alteration of the original surface (material, geometry) and the by-products of human activity (heat, water vapor, pollution). As a consequence, a temperature excess develops in our living space in the near-surface air layer (compared to the rural areas). This is the so-called 'classical' *urban heat island* (UHI), already recognized at the end of the 18th century but named only much later by *Balchin* and *Pye* (1947).

As subsequent studies revealed, many kinds of UHIs can be defined. They include those defined according to the target medium (air, surface, sub-surface), the location (surface nature, height of measurement), and the type of sensor (*Roth et al.*, 1989). Considering the thermal effect of cities according to these new approaches, they do not appear in an 'island' structure. Moreover, the investigations are often directed on the intra-urban variations of the temperature. Therefore, in these cases it is more appropriate to talk about the *urban temperature field* or *pattern*. Besides 'classical' UHI, this study concentrates on urban surface temperature patterns.

Cities present an almost limitless array of surface configurations (*Soux et al.*, 2004). The detemination of the surface temperature in cities is difficult because of the complex structure of the urban-atmosphere interface (*Voogt* and *Oke*, 1997). In large urban areas it is usually measured indirectly by remote sensing technology mounted on satellite or aircraft platforms, but the sensor can be carried by car or manually too. In this case the problem is that the camera does not 'see' the total active surface because of the obstructions present on the 3D surface (*Roth et al.*, 1989; *Soux et al.*, 2004).

The temperature of the air (T_a) among the buildings is affected by the temperatures (T_s) of both horizontal and vertical surfaces (e.g., roofs, roads, tree tops, ground, walls) (*Voogt* and *Oke*, 1998). This multiple impact and the magnitudes of the effects of individual factors are very difficult to determine. *Voogt* and *Oke* (1997) introduced the concept of *complete surface temperature* which cannot be measured directly, but it can be calculated or estimated as a result of the radiation originating from all of the (horizontal and vertical) surfaces. Since such a detailed survey is extremely time-consuming, it cannot be applied in a larger area.

Roth et al. (1989) came up with the question of what the relationship is between surface and air temperature patterns. It is known that the nocturnal intra-urban variability in T_s is much smaller than the diurnal variability, while the opposite is true for T_a. As the near-surface climates are directly connected to the active surface (if the larger scale weather situation is favorable) there seems to be a contradiction. According to *Roth et al.* (1989), this could be due to the following: (i) lack of simple connection between the T_s and T_a values (implied also by *Goldreich* (1985)), (ii) remote sensors do not perceive the full active

surface, (iii) failure to recognize the different scales of climatic phenomena in the urban environment.

In the frame of a small review, we only go through the results of some earlier studies, which partly dealt with T_s–T_a relationship. Because of the better geometric resolution, in these studies the surface infrared images were taken from airborne platforms (with one exception). The T_a data were obtained by mobile transects or fixed stations at 1.5–2 m above the ground. The observations of both parameters were taken at 'ideal' (calm and clear) weather conditions in the evening or at night. The parameters of the mentioned studies are summarized in *Table 1*.

Table 1. Parameters of nocturnal surface and air temperature measurements of some earlier studies

City	Thermo camera platform, altitude of flying	Pixel size	Type of stations	Source
Johannesburg	Aircraft, 400 m	2.5 m^2	Mobile	*Goldreich* (1985)
Malmö	Aircraft, -	1 × 2 m	Mobile	*Bärring et al.* (1985)
Göteborg	Aircraft, 600 m	2 × 2 m	Fixed, mobile	*Eliasson* (1992)
Göteborg	Manual	-	Fixed, mobile	*Eliasson* (1996)
Tel-Aviv	Helicopter, 2300 m	2 × 2 m	Mobile	*Ben-Dor* and *Saaroni* (1997)
Tel-Aviv	Helicopter, 2300 m	2 × 2 m	Mobile	*Saaroni et al.* (2000)

The investigation made by *Goldreich* (1985) is among the first of our subject. Results of both measurement methods indicate one nocturnal UHI core and a strong temperature gradient at about 600 m from the core, so the two temperature patterns overlap.

Bärring et al. (1985) found that high T_s can also be observed in suburban streets, provided that the streets are narrow enough. In contrast, T_a values decrease from the city center to the outskirts. According to their explanation, the air temperature is influenced both by the local street surface temperature, which is regulated by street geometry, and by the general thermal level of the surrounding city area determined by its general geometry and other UHI generating factors (air pollution, anthropogenic heat release).

Eliasson (1992) compared the two temperatures measured along a route in streets and squares in a central urban area. While she observed 5 °C differences on the surface, the T_a variation was well-balanced with a range of only a few tenths of degrees. Later *Eliasson* (1996) draws attention to the fact that many studies do not make a clear distinction between surface and air temperatures. For example, the results of models predicting T_s are often compared or validated with field studies of T_a. She compared the variations in the two parameters along urban transects and found that the intra-urban surface and air temperature patterns are very different. T_s is influenced by the immediate urban structure (geometry), while T_a is not, so the former shows far larger fluctuations than the latter.

Ben-Dor and *Saaroni* (1997) compared surface and air temperature patterns measured separately along four N-S transects in a large urban area. The T_s values were computed as averages of about 40 pixel values around the spots of the T_a recordings. On and over uniform (asphalt) surfaces, they found strong relationships between the two temperature variations for all transects. *Saaroni et al.* (2000) generated isotherm maps from the measured T_s and T_a values along the four transects mentioned above. The obtained temperature fields were similar both in areal structure and magnitude.

The results of the mentioned studies are a bit contradictory. As *Voogt* and *Oke* (2003) summarized, the relations between the two parameters remained empirical and no simple general relation was found, but the correlations improved at night when microscale advection is reduced. It is clear that the complex interrelations between the two parameters in urban environments are not unambiguously detected or it is not even possible to describe them. Within this system, our research is focused on the area of impact which has an influence on the temperature of the near-surface air, presuming that a statistically based relationship can be established between the two temperatures if the size of this area is appropriately selected. Accordingly, the area of impact or *source area* is the place where the total impact of the physical features of its elements and their responses to the outer effects (heating-cooling, generation of turbulent processes) determine the temperature of a given air parcel.

The aim of this study is (i) to compare the air and surface temperatures developing in a complex urban environment, (ii) to reveal the relationship between them applying different source areas, and (iii) to generate an air temperature field for a large urban area with the help of the obtained relationship for two summer evenings, when the weather conditions in the preceding 36 hours were favorable for the microclimate modifying effects of the surface features.

2. Study area, weather situations

2.1. Szeged

Szeged is located in the southeastern part of Hungary (46°N, 20°E) at 79 m above sea level on a flat plain (*Fig. 1*). It belongs to the climatic region *Cf* according to Köppen's classification (temperate warm climate with uniform annual distribution of precipitation) (*Péczely*, 1979). Szeged provides a favorable background for urban climatological research, since its climate is free from the influences of orography and large water bodies, while its weather is characterized with high sunshine duration, low wind speed, and low precipitation (*Table* 2). Therefore, research on urban heat island and its consequences in Szeged has a several year long history (e.g., *Unger*, 1996; *Unger et al.*, 2001; *Kristóf et al.*, 2006).

Fig. 1. (a) City center, (b) housing estates, (c) detached houses with gardens, (d) industry and warehousing, (e) agricultural, green, and open area as land use types in Szeged.

The city's population of about 160,000 lives within an administration district of 281 km^2, but the highly urbanized area is restricted to an area of about 30–35 km^2. In the course of the 19th century, a structure of boulevards and avenues was established using the river Tisza as an axis. A large number of different land-use types are present including a densely built center with medium-wide streets and large housing estates of tall apartment buildings set in wide green spaces. There are zones used for industry and warehousing, areas occupied by detached houses, and considerable open spaces along the banks of the river, in parks, and around the city's outskirts (*Fig. 1*).

Table 2. Climate averages (1961–1990) in Szeged region (*WMO*, 1996)

Parameter	August	Year
Sunshine duration (h)	266	2023
Wind speed (m s^{-1})	2.8	3.4
Air temperature (°C)	20.2	10.5
Relative humidity (%)	69	75
Precipitation (mm)	57	495

2.2. *Weather situations before and during the measurements*

The observed T_a and T_s values depend not only on the momentary weather conditions but on the conditions prevailing in the previous (a few hours or days) time period as well. Therefore, a short description is given about the weather features of the 36-hour periods preceding the measurements of August 12 and 14, 2008 (between 18:30 and 19:30 UTC, see Section 3.1.).

The observations of the University station located at the edge of the city center were used to characterize the weather in the mentioned time period. Among them, the data of global radiation (G) and wind speed (v) are the most important, because these values reflect whether the weather was clear and calm (*Fig. 2*). According to the data, the insolation was undisturbed during the daylight hours (regular bell-shape G variation) with maximum values of 810 – 860 W m^{-2}. On the whole, the air movement was moderate (0 – 4.7 m s^{-1}) and during the measurements, it ranged between 0.8 and 3.1 m s^{-1}. Since the wind speed was measured at 26 m above ground level, it was likely to be significantly smaller at the street level. The days were rather warm with maximum values of 28 –36 °C and with minimum values of 17–22 °C at dawn.

Consequently, during the investigated period, the weather conditions promoted the microclimatic effects of the surface features in Szeged.

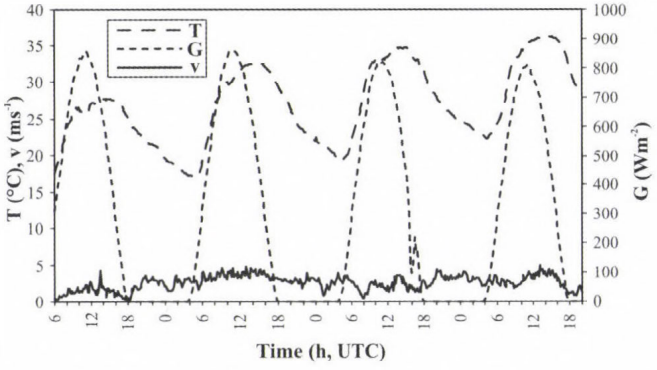


Fig. 2. Variation of global radiation (G), air temperature (T), and wind speed (v) at the University station during the measurements and the preceding one and a half days (06:00 UTC, August 11, 2008 – 20:00 UTC, August 14, 2008).

3. *Methods*

3.1. *Temperature measurements*

The measurements in August took place in the 1.5 –2-hour periods immediately after sunset, because there was no direct shortwave radiation disturbing the signals recorded by the camera at that time and the flight did not violate the flight ban

after 22:00 LST. As sunset was at 19:57 LST (17:57 UTC), the airborne images were taken between 18:15 and 19:45 UTC, the surface and near-surface measurements were performed between 18:30 and 19:30 UTC. Thus, 19:00 UTC (1 hour after sunset) can be regarded as a reference time of all measurements. It is worth to note that test measurements were taken on July 29, 2008 in order to optimize the flying parameters, spots, and route of the surface and near-surface observations, as well as to harmonize the work of the different measurement teams (flying – 3, mobile measurement – 2, measurements by hand – 4 people).

3.1.1. Mobile air temperature measurements along a N-S urban transect

Mobile measurements are widespread in studying urban climatic parameters (e.g., *Conrads* and *van der Hage*, 1971; *Eliasson*, 1996; *Voogt* and *Oke,* 1997; *Henninger* and *Kuttler*, 2007). In our case the temperature observations were carried out by an automatic radiation-shielded sensor (DCP D100089 HiTemp) connected to a digital data logger (LogIT DataMeter 1000) that provided data at every 2 seconds. In order to diminish the thermic disturbing effect of the car, the sensor was located on a bar 0.6 m before the car and 1.45 m above the ground. Due to the requirements of efficient ventilation and data density, the speed of the car was 20–25 km h^{-1}, and thus the data were gathered every 11–14 m. The locations of the data along the measurement route were recorded by GPS. Data measured at the compulsory stops (red light, barrier etc.) were later deleted from the database.

The 11.8 km long measurement route is a N-S transect of Szeged crossing the typical urbanized land use areas of the city (housing estates, detached houses with gardens, city center, industry, and warehousing). One-hour measurements were taken there and back along the route in order to make time adjustments to the reference time (19:00 UTC) assuming linear air temperature change with time at this period of the day (*Oke* and *Maxwell*, 1975). This linear change was also observed in the records of the University weather station (*Fig. 2*). In order to get uniform temperature distribution along the route, the route was divided by 15 m, then the averages of the mean values of the readings taken separately there and back on the 15 m long sections were computed (*Fig. 3*). Thus, there are altogether $2 \times 786 = 1572$ T_a values, which refer to the middle points of these sections.

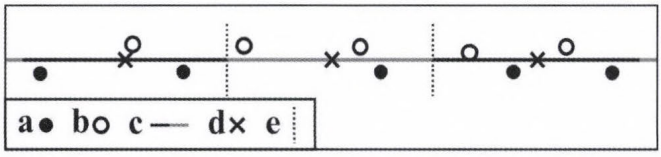


Fig. 3. An example on the spots of air temperature values regarded at the averaging route. (a) spots there, (b) spots back, (c) 15 m long sections, (d) section centers, (e) section borders.

The relative temperatures by points along the route (that is the UHI intensity) were determined by regarding the value of the minimum T_a as a UHI intensity of 0 °C.

3.1.2. Surface temperature measurements in the urban environment using an airborne thermal camera

The thermal measurements were carried out using a self-developed low-cost small-format digital imaging system. The current system is based on the experiences gained during earlier research using thermal (*Mucsi et al.*, 2004) and CIR (color-infrared) small format aerial photography (SFAP) data collected with a small airplane (*Rakonczai et al.*, 2003; *Szatmári et al.*, 2008). The system is based on a FLIR ThermaCam P65 thermal camera integrated with a navigation system and a GPS/GNSS receiver (Mobile Mapper CE).

The FLIR (forward looking infrared radiometer) is a compact single-band thermal camera with a temperature sensitivity of 0.08 °C. The instrument can measure temperatures between −40 °C and +500 °C and has a resolution of 320 × 240 pixels. The camera has germanium lenses to register the thermal radiation emitted by the object of investigation.

The flight plan determined the data collection density during the survey. Since a fixed photographic lens was used for the camera, the flying speed, height, and the distance between consecutive flight lines defined the overlap between adjacent images. For later post-processing purposes, an overlap of at least 20–30% was needed. To collect data at the required density, one image per 4 seconds was recorded while flying with a speed of 120 km h^{-1} at a height of 2000 m.

Before the flight, a detailed flight plan (*Fig. 4*) was created, and during the survey this plan was followed as accurately as possible. A Mobile Mapper CE GPS/GNSS receiver provided navigational aid during the flight and recorded the actual high precision GPS flight track, which was used for positioning the images during the post-processing phase.

After the survey, the acquired data needed to be post-processed. Using software specifically designed for the analysis of the thermal images, several parameters needed to be readjusted. Using the results of the fieldwork carried out at the same time as the flight, the data was inspected and calibrated (see Section 3.1.3.). The temperature data collected by the thermal camera was refined based on the time-synchronized field measurements of the same evening as the flight. After this, the thermographic data was stored in a file as a matrix, containing the calibrated temperatures.

The next step in the thermal processing chain was the extraction of the coordinates and time data from the GPS track log recorded during the flight. This data was connected to the camera log file that stores the precise time of every recorded image. With this operation, EOV coordinates were generated for

every image center. During the next step, the center coordinate and the image resolution were combined to create a so-called world file. A world file is a simple text file that was used to georeference the thermal images. Using this georeference, a first, coarse geometric correction was executed with an accuracy of about 100–150 meter, depending on the weather circumstances during the flight (*Fig. 5*). In the next step, the images were combined to one single image, resulting in a thermal mosaic with a spatial resolution of 2.5 m, covering the complete survey area (*Fig. 6*).

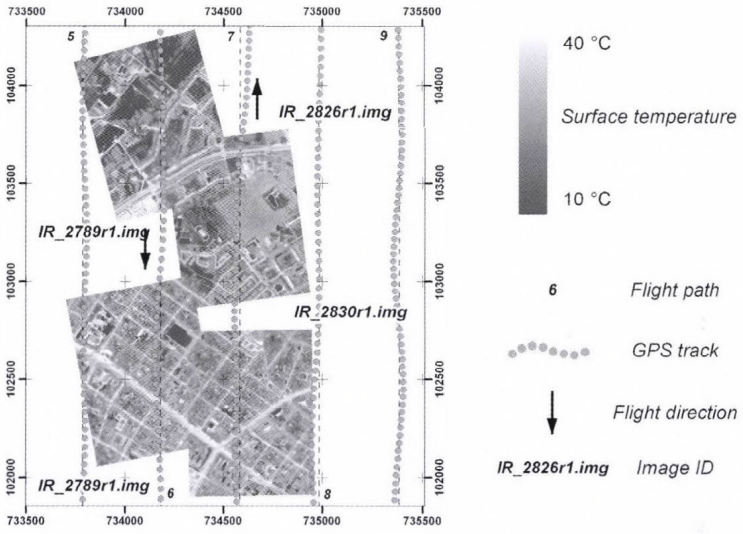


Fig. 4. Flight plan with real GPS track and calibrated sample thermal images of the city.

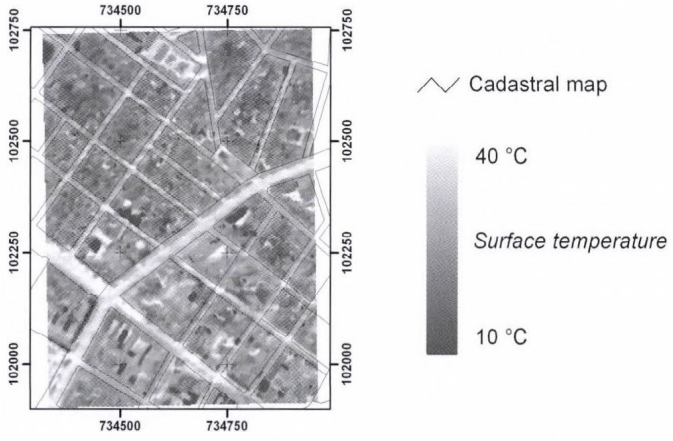


Fig. 5. Georeferenced (Hungarian EOV) thermal image with cadastral vector layer.

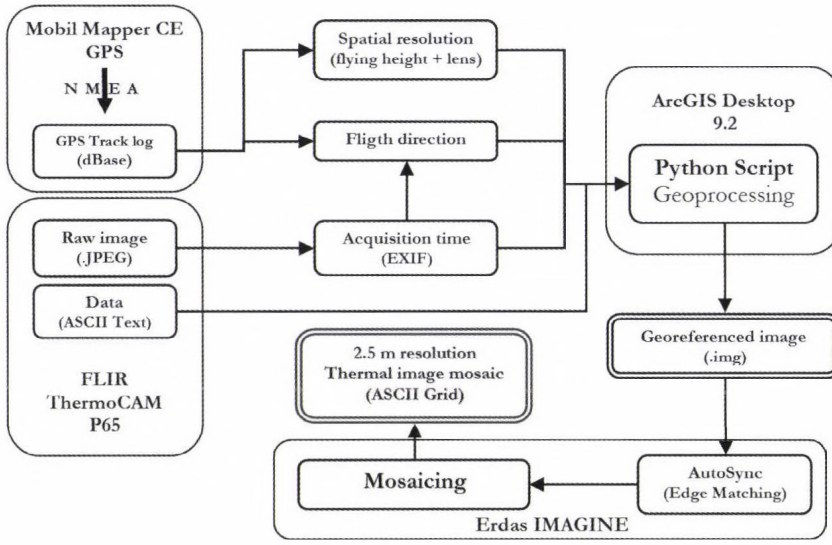


Fig. 6. Flowchart of the image processing.

It was assumed that some decrease can be found in the surface temperatures during the flight times (50 and 51 minutes on August 12 and 14, respectively). If this dependence from the time is significant, the pixel values have to be corrected to the middle time (19:00 UTC) of the survey according to the general cooling tendency. A statistically established trend was found only in the evening of August 14 (*Fig. 7*).

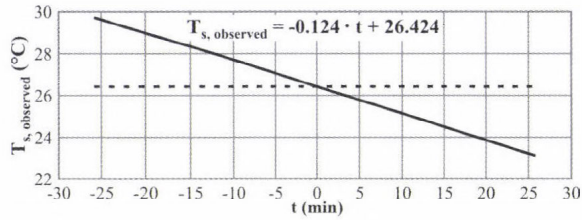


Fig. 7. Regression line of the pixel values ($T_{s,observed}$) as a function of time (t) in August 14, 2008.

In order to refer the pixel values to the middle time, these values were corrected by the obtained cooling rate: the values before the middle time were lowered; the ones after the middle time were increased:

$$T_{s,observed} = -at + b, \qquad T_s = T_{s,observed} + at,$$

where $T_{s,observed}$ is the measured pixel value, a is the slope of the regression line,

t is the time in minutes (–25.5 – 25.5), b is the value of the regression line at the middle time, and T_s is the corrected pixel value. Henceforth T_s means the uncorrected and corrected surface temperature in August 12 and 14, respectively.

3.1.3. In situ surface temperature measurements in selected points of representative urban surfaces

The temperature of the surface was measured directly by hand (DCP D100047 ProTemp sensor, LogIT DataMeter 1000 data logger) in the 40 selected points representing the urban surfaces in two sample areas (*Fig. 1*). These T_s values are necessary for the calibration of the airborne thermal images, namely for transforming the relative temperature scale of the images to an absolute scale.

3.2. Statistical method to reveal the relationship between air and surface temperatures

As mentioned in Section 1, the air temperature at a given point at 1.5 m above ground level depends not only on the immediate surface under it and its own temperature. The T_a of this point is a result of the total effects of the turbulent heat transports generated by the surrounding heated surfaces. Several studies looking for T_a–T_s relationship point out the importance of the micro-advection in the near-surface air layer (e.g., *Roth et al.*, 1989), which promotes the mixing of the thermal properties in a wider environment. Therefore, to investigate this connection a larger area, the source area and its thermal features have to be taken into account. According to the related literature, this source area covers an area with a radius of a few hundred meters around the measurement point, and in the case of calm weather it can be regarded as a circle. In case of a temperature sensor at 1.5–2 m above ground level in urban environment, the circle has a radius of maximum 0.5 km, but this is likely to depend on the building density (*Oke*, 2004).

These active surfaces are not only horizontal but also vertical (e.g., walls). In affecting the value of T_a, the role of the surfaces nearer to the point is more important compared to the ones farther away. It has to be taken into account that in the case of 1.5 m level, the role of the roofs is likely to be smaller than the role of the ground surfaces (pavements, roads, parking lots, grass areas, etc.). In our case, only the temperatures of the horizontal surfaces can be detected on the thermal images, thus we can only use these temperatures in searching for the relationship between T_a and T_s.

For the determination of the size of the source area for the T_a values along the transect and the distance-based weighting of the pixel values inside it, different approaches were applied. Only those are presented here which gave the best results in investigating the relationship and that are in accordance with the conditions mentioned above.

Among the results of the manual T_s measurements, the lowest value was 19.4 °C, however, the possibility of lower values occurring in the investigated area cannot be excluded. After the air temperature data was processed, it became clear that there are areas detected by the camera that have temperatures that are up to 15 degrees lower than the directly surrounding surface. The reason for this is the heterogeneity of the urban environment. The camera was setup to use normal calibration values. The physical properties of the roofs constructed from metal considerably differ from the standard calibration values. Highly reflective surfaces like aluminum, copper, or stainless steel work as a mirror resulting in false temperature measurements. The false temperatures were identified and filtered out of the images.

As a second step, according to the maximum size of the source area mentioned above (*Oke*, 2004), circles with a radius r (r = 100–500 m) around the points of T_a values were selected (see Section 3.1.1.). The (filtered) T_s pixel values of these circles were taken by weighting, the farthest (r) values with a factor of 0.5, the nearest ones (0 m) with a factor of 1, and the ones between them with a proportionality factor between 0.5 and 1. Thus, the weighted and averaged surface temperature ($T_s(wr)$) regarding a given point's surrounding with a radius of r is determined by the next formula:

$$T_S(wr)=\frac{\sum_{i=1}^{n} T_{Si}\cdot\left(1-\frac{D_i}{2r}\right)}{\sum_{i=1}^{n}\left(1-\frac{D_i}{2r}\right)}, \tag{1}$$

where T_{Si} is the ith pixel value, D_i is the distance of the ith pixel from the given point, and the summation refers to all (not filtered) pixels inside the circle with a radius of r. In order to automate this calculation for all points along the transect, an algorithm was developed in the ArcView Avenue script language.

4. Results

Fig. 8 shows the variations of the UHI intensity along the transect on the two evenings. The reference place (with the lowest T_a value) is located in an area with abundant vegetation almost at the south end of the transect. In cities in the temperate climate zone during 'ideal' weather conditions the maximum development of the UHI occurs 3–5 hours after sunset because of the different cooling rates of urban and rural areas (*Oke*, 1987). Therefore, in our case (1 hour after sunset), the obtained UHI intensities are likely to be smaller than the maximum values occurring during the same night.

In the evening of August 12 the intensity was relatively high (~2 °C) at the north end of the transect, wich can be attributed to the area characterized with buildings of a housing estate. Going towards the south, two maximums

reaching almost 3 °C occurred: the first one in an area with large artificial surfaces (shopping centers, parking lots), while the second one appeared in the densely built-up city center. From here the intensity decreases almost continuously towards the south end of the transect.

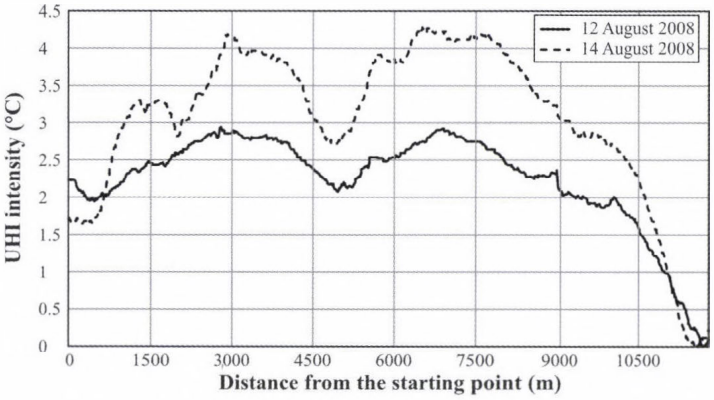


Fig. 8. UHI intensities along the N-S transect in Szeged (19:00 UTC, August 12 and 14, 2008), starting point at the north.

On August 14, the variation of the intensity was very similar to the previous one with the exceptions of the maximum values, which exceeded even 4 °C this evening.

Since the measurement route was located inside the urban area, the obtained intensity values reflect the intra-urban temperature variation rather than the real magnitude of the UHI. Namely, when we talk about the heat island, the urban temperatures are usually compared to the temperatures of the surrounding rural ('natural') areas (*Oke*, 1987; *Lowry*, 1977). This can also be an explanation for the moderate magnitude of the obtained maximum UHI values in this study.

According to the method described in Section 3.2, several relationships can be obtained between the T_a values along the transect and the $T_s(wr)$ values computed by Eq. (1) with different radii (two measurements – n = 1572). As *Table 3* shows, the relationships are significant in all cases (even at a level of 0.1% at this high number of element pairs), however, the larger the radius the stronger the connection between the two parameters (to the upper limit of r, see Section 3.2.). This can be explained by the thermal compensatory effect of the larger area and confirm the possible size of the source area related to the air temperature measured at a level of 1.5 m in urban environments. Since the best relationship is derived in the case of $r = 500$ m, henceforth the regression equation obtained in this case is applied to extend our results: to model the spatial distribution of T_a in a larger urban area in the case of the investigated two evenings.

Table 3. Relationships between T_s and T_a along the urban transect and their parameters regarding the surroundings with different radius r (R^2 is the determination coefficient, R is the correlation coefficient, σ_R is the standard deviation around the regression line) on August 12 and 14, 2008 (n = 1572)

r (m)	Regression equation	R^2	R	σ_R	Significance level
100	$T_a = 0.373 \times T_s(w100) + 17.691$	0.574	0.757	0.858	< 0.001
200	$T_a = 0.406 \times T_s(w200) + 16.898$	0.611	0.781	0.820	< 0.001
300	$T_a = 0.426 \times T_s(w300) + 16.453$	0.642	0.801	0.787	< 0.001
400	$T_a = 0.436 \times T_s(w400) + 16.228$	0.663	0.814	0.763	< 0.001
500	$T_a = 0.447 \times T_s(w500) + 15.982$	0.685	0.828	0.738	< 0.001

In the course of the extension of our results, surface temperatures in the whole study area were used as input data in a 100 m × 100 m mesh for modeling the air temperature fields in the evenings of August 12 and 14. Since for computing $T_s(w500)$ the pixel values around the mesh points ($r = 500$ m) are needed, the modeling area is smaller (~21 km^2) compared to the originally recorded area (marked by intermittent line in *Fig. 1*). The deficiencies at the corners are caused by some errors in the flying paths.

The modeled air temperature field on August 12 has large areas with a temperature greater than 26.5 °C. This area is located in the city center and streches out towards NW, where industrial and warehousing land use dominates (*Fig. 9a*). A temperature extension is present in NE direction, where the large housing estates can be found. The cooler areas of the low built-up districts, the outer green zones, and the belt of the Tisza River and its banks (at NW, SW, and SE) are also recognizable. On the whole, a difference of about 3 °C occurs in the area which is in accordance with the temperature range experienced along the transect in this evening (*Figs. 8* and *9a*).

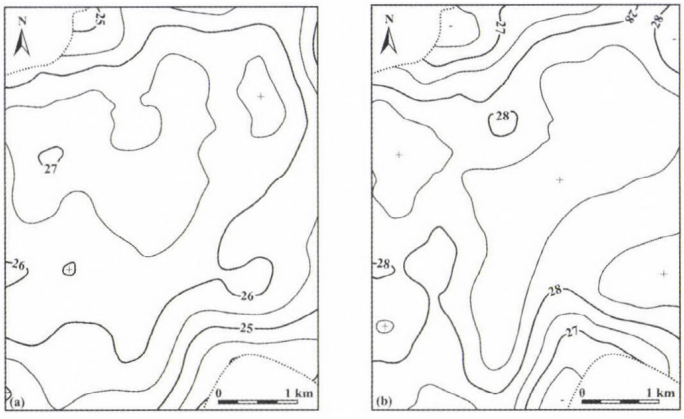


Fig. 9. Modeled air temperature patterns in Szeged at 19:00 UTC, (a) August 12 and (b) August 14, 2008 (location of the area is shown in *Fig. 1*).

The values of the modeled air temperature pattern on August 14 are higher than the values of the previous one (*Fig. 9b*). As on August 12, basically the warmest (>28.5 °C) areas can be found in the center, in the NW and NE parts, augmented towards the south. Broadly speaking, the cooler areas are also the same. On the whole, also a difference of about 3 °C occurs in the area, which is a bit lower than the temperature range measured along the transect of this evening (*Figs. 8* and *9b*).

5. Conclusions

In this study we tried to reveal a statistical (but physically supported) connection between the air and surface temperatures measured in an urban environment. During the data collection, airborne and surface-based remote sensing (indirect) and traditional (direct) tools and methods were applied. In order to search for the mentioned relationship, a wider environment, the source area was taken into account, which determines the air temperature at a given point and time. As the results show, using a source area with a radius of 500 m, a strong relationship was detected between the two parameters. Namely, the temperatures of the surfaces found in the surroundings (weighted with the distance) decisively influence the temperature of the air parcel located at a given point. The obtained regression equation was applied to extend our results in order to model the air temperature field in a larger urban area in the investigated two evenings.

By all means we should not forget that the obtained relationship is based on the data of only two, however complex, measurement campaigns. In the future, when using data of more measurements on days with similar environmental conditions to that of the investigated days, the result could be refined. Based on these new results we can make steps to the generalization of the operation mechanisms between the urban air and surface temperatures. In the frame of this study, a data collection in different seasons could also be a new direction, which can provide a possibility to examine the specific seasonal features and enable their comparison.

Acknowledgements—This work was supported by the Hungarian Scientific Research Fund (OTKA T048400, K-67626). The authors' special thanks are due to *G. Barna, K. Balogh, N. Kántor, Z. Ladányi, Z. Sümeghy, T. Unger* taking part in the measurements and to *E. Tanács* for the language revision of the manuscript.

References

Balchin, W.G.V. and *Pye, N.*, 1947: A micro-meteorological investigation of Bath and the surrounding district. *Q. J. Roy. Meteor. Soc. 73,* 297-319.

Bärring, L., Mattsson, J.O. and *Lindqvist, S.*, 1985: Canyon geometry, street temperatures and urban heat island in Malmö, Sweden. *J. Climatol. 5,* 433-444.

Ben-Dor, E. and *Saaroni, H.*, 1997: Airborne video thermal radiometry as a tool for monitoring microscale structures of the urban heat island. *Int. J. Remote Sens. 18,* 3039-3053.

Conrads, L.A. and *van der Hage, J.C.H.*, 1971: A new method of air-temperature measurement in urban climatological studies. *Atmos. Environ. 5*, 629-635.

Eliasson, I,, 1992: Infrared termography and urban temperature patterns. *Int. J. Remote Sens. 13,* 869-879.

Eliasson, I., 1996: Urban nocturnal temperatures, street geometry and land use. *Atmos. Environ. 30,* 379-392.

Goldreich, Y., 1985: The structure of the ground heat island in a central business district. *J. Clim. Appl. Meteorol. 24*, 1237-1244.

Henninger, S. and *Kuttler, W.*, 2007: Methodology for mobile measurements of carbon dioxide within the urban canopy layer. *Climate Res. 34,* 161-167.

Kristóf, G., Bányai, T. and *Rácz, N.*, 2006: Development of computational model for urban heat island convection using general purpose CFD solver. *Preprints 6th Int Conf on Urban Climate,* Göteborg, Sweden, 822-825.

Lowry, W.P., 1977: Empirical estimation of urban effects on climate: A problem analysis. *J. Appl. Meteorol. 16*, 129-135.

Mucsi, L., Kiss, R., Szatmári, J., Bódis, K., Kántor, Z., Dabis, G. and *Dzsupin, M.*, 2004: The analysis of contamination deriving from the leakage of subsurface pipeline networks via remote sensing (in Hungarian). *Geodézia és Kartográfia 56*/4, 3-8.

Oke, T.R., 1987: *Boundary Layer Climates.* 2nd edition. Routledge, London-New York.

Oke, T.R., 2004: Siting and exposure of meteorological instruments at urban sites. *27th NATO/CCMS Int Tech Meeting on Air Pollution Modelling and Application.* Kluwer, Banff, Canada, 14 p.

Oke, T.R. and *Maxwell, G.B.*, 1975: Urban heat island dinamics in Montreal and Vancouver. *Atmos. Environ. 9,* 191-200.

Péczely, G., 1979: *Climatology* (in Hungarian). Tankönyvkiadó, Budapest.

Rakonczai, J, Csató, Sz, Mucsi, L, Kovác,s F, Szatmári, J, 2003: Experiences on the evaluation of the 1999 and 2000 excess waters (in Hungarian). *Vízügyi Közlemények Különszám IV,* 317-336.

Roth, M., Oke, T.R. and *Emery, W.J.*, 1989: Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology. *Int. J. Remote Sens. 10,* 1699-1720.

Saaroni, H., Ben-Dor, E., Bitan, A. and *Potchter, O.*, 2000: Spatial distribution and microscale characteristics of the urban heat island in Tel-Aviv, Israel. *Landscape Urban Plan 48,* 1-18.

Soux, A., Voogt, J.A. and *Oke, T.R.*, 2004: A model to calculate what a remote sensor 'sees' of an urban surface. *Bound.-Lay. Meteorol. 111,* 109-132.

Szatmári, J., Tobak, Z., van Leeuwen, B., Olasz, A. and *Dolleschall, J.*, 2008: An effective and low-cost method to detect environmental contaminations: the promise of CIR small format aerial photography. *12th Geomathematics Conf.*, Mórahalom, Hungary.

Unger, J., 1996: Heat island intensity with different meteorological conditions in a medium-sized town: Szeged, Hungary. *Theor. Appl. Climatol. 54,* 147-151.

Unger, J., Sümeghy, Z. and *Zoboki, J.*, 2001: Temperature cross-section features in an urban area. *Atmos. Res. 58,* 117-127.

Voogt, J.A. and *Oke, T.R.*, 1997: Complete urban surface temperatures. *J. Appl. Meteorol. 36,* 1117-1132.

Voogt, J.A. and *Oke, T.R.*, 1998: Radiometric temperatures of urban canyon walls obtained from vehicle traverses. *Theor. Appl. Climatol. 60,* 199-217.

Voogt, J.A. and *Oke, T.R.*, 2003: Thermal remote sensing of urban climates. *Remote Sens. Environ. 86*, 370-384.

WMO, 1996: *Climatological normals (CLINO) for the period 1961-1990.* WMO/OMM-No. 847, Geneva.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 4, October–December 2010, pp. 303–318

Odor setback distance calculations around animal farms and solid waste landfills

Melinda Cseh[1], **Katalin F. Nárai**[1], **Endre Barcs**[1], **Dezső B. Szepesi**[1], **Dezső J. Szepesi**[1], and **James L. Dicke**[2]

[1]*Consultants on Air Resources Management (CARM) Inc., Katona József u. 41. V/25, H-1137 Budapest, Hungary; E-mail: szd12506@ella.hu*

[2]*Consulting Meteorologist, U.S.A.; E-mail: jldandpad@hotmail.com*

(Manuscript received in final form February 27, 2010)

Abstract—ODOR-TRANSMISSION (OT) is a regulatory model for estimating odor setback distances around livestock production farms and municipal solid waste landfills on virtual maps prepared by an interdisciplinary team. This paper describes the development of the model, odor emission and dispersion calculations, and visualization techniques. The atmospheric dispersion simulation portion is based on HNS-TRANSMISSION, an air quality model published by *Szepesi et al.* (2005), while the odor emission module is adapted mostly from U.S.A. results, e.g., *Guo et al.* (2004). As a new development, estimated setback distances from an odor source are presented on virtual maps. The main characteristics of the OT model can be found in the model catalogue system of the European Environmental Agency. A case study for a municipal solid waste landfill is also included. Furthermore, a comparison study with other odor setback distance calculation methods shows that reliable results can be obtained by using the OT model.

Key-words: odor, odor unit, transmission modeling, setback distance, exceedance probability, odor threshold

1. Introduction

Over the last decades, odor nuisances, e.g., from livestock production units, have become more frequent and intensive in our environment. Fundamental human rights include a pure, livable environment, and this includes odor-free air as well. To find solutions for odor problems, standard regulations must be adopted. This involves the need to determine setback distances around animal production farms, municipal solid waste landfills, wastewater treatment plants and factories that emit odorous gases. Most of the existing setback guidelines

were determined either by individual judgment and experience or by a combination of neighbor surveys and odor measurements, instead of calculations of dispersion models (*Guo et al.*, 2004). Determination of odor nuisances by olfactometria (testing and measurement of the sensitivity of the sense of smell) is essential, although the practice involves many subjective factors. Therefore, it is essential to establish a science-based model to predict setback distances more objectively. Such a model can become an effective tool for researchers, regulatory agency staff, and other decision-makers. The aim of this paper is to present the newly developed OT model for determining odor setback distances.

The OT model is an extended version of HNS-TRANSMISSION (*Szepesi et al.*, 2005), an air pollution regulatory model that includes a meteorological database for the calculations. The OT model consists of three main parts: (i) the odor emission module for estimating the odor source term; (ii) the HNS-TRANSMISSION model for calculating the atmospheric dispersion of the odorants, and; (iii) visualization techniques for presenting the setback distances on maps such as those obtained from Google®.

The odor emission module quantifies the odor emissions originating from animal farms, municipal solid waste landfills, and similar sources. Quantitative operating parameters of the facilities are necessary, e.g., types of animals, the area of the outdoor manure storage or the size of the municipal solid waste landfills, etc.

Odor dispersion in the outdoor air is simulated by the HNS-TRANSMISSION model. The OT model outputs are tables containing the number of odor exceedances vs the distance from each source by each of 16 wind directions. These results are analyzed according to the number of hours in a year, and thereby the area that is exposed to any given level of odor annoyance can be determined. The OT model produces setback distances along the 16 wind directions on a Google®-type map.

Both of these models are part of the model documentation system of EIONET program of the European Environmental Agency, (ODOR-TRANSMISSION: http://pandora.meng.auth.gr/mds/showlong.php?id=179, HNS-TRANSMISSION: http://pandora.meng.auth.gr/mds/showlong.php?id=48).

2. Estimating the odor source term

2.1. Odor

Odor is the property of a chemical substance or substance mixtures, dependent on the concentration, that activates the sense of smell, and thus starts an odor sensation (*Winneke* and *Steinheider*, 1994). As odor sensitivity differs in each person, a completely objective and analytical measurement cannot be done. However, it is possible to define the annoyance level of the odorant by several

methodologies. To measure odor, an olfactometry methodology is used that is based on a panel of human noses as sensors. For the determination of the quantity of odorants in the air, odor concentration is used – the dimension is European odor unit per m^3 (OU_E / m^3). The European odor unit by definition is: the amount of odorant(s) that, when evaporated into one cubic meter of neutral gas at standard conditions, elicits a physiological response from a panel (detection threshold) equivalent to that elicited by one European reference odor mass (EROM), evaporated in one cubic meter of neutral gas at standard conditions (*CEN*, 2003). EROM is the accepted reference value for the European odor unit, equal to a defined mass of a certified reference material. One EROM is equivalent to 123 mg *n*-butanol (CAS 71-36-3) evaporated into one cubic meter of neutral gas. This produces a concentration of 0.040 $mmol / m^3$ (*CEN*, 2003). These olfactometric measurements are needed to determine a reference scale for each odorant. However, for everyday regulations, these processes are too complicated and require considerable time. Therefore, more effective and straightforward techniques should be used, such as can be achieved by developing models based on the results of analytical measures. This is one aim of the OT modeling program.

2.2. Odors and their sources

The OT model was developed to address odors emitted from livestock operations, municipal waste landfills, and factories. The odor that is detected from a livestock operation is a complex mixture of gases. Most often the odor is a result of the uncontrolled anaerobic decomposition of manure. However, feed spoilage can also contribute to the odor. The odor that is detected by the human nose can be a combination of 60 to 150 different compounds. Some of the most important types of odor causing compounds are volatile fatty acids, mercaptans, esters, carbonyls, aldehydes, alcohols, ammonia, and amines. The odor strength of these compounds does not combine in an additive manner. That is, sometimes mixing several of these compounds can result in reduced odor by dilution of the strongest smelling compounds. In other instances, the mixture is worse than any of the individual compounds.

2.2.1. Animal farms

To determine the setback distance around an animal production farm, odor emissions must be estimated. Eq. (1) for calculating odor emissions is based on the Purdue model, which was developed by *Lim et al.* (2000) and combines features of Austrian and British setback guidelines (*Schauberger* and *Piringer*, 1997; *Williams* and *Thompson*, 1985):

$$OE = \Sigma (A_E \cdot N \cdot P \cdot B), \quad (1)$$

where

OE is the odor emission rate, OU/s,
A_E is the building odor abatement factor (0.30 –1.00),
N is the number of animals, and
P is the building odor emission factor, OU/s.

The second equation of the model is

$$B = M - D, \tag{2}$$

where

B is the building design and management factor,
M is the manure removal frequency (0.40 – 1.00), and
D is the manure dilution factor (0.00 – 0.20).

The OT model takes into consideration the following parameters: type of animals, number of animals, manure removal frequency, manure dilution factor, area of the storage building, wind speed above the ground, and odor abatement factors. The method for estimating odor emissions is based on the livestock unit. The livestock unit (LU) is a unit to compare or aggregate numbers of animals of different species or categories. Equivalences based on the food requirements of the animals are defined. While the Purdue model estimates the emissions for swine buildings, the OT model applies the methodology to other livestock units as well (*Table 1*). These livestock units are based on the field and laboratory measurements from the North Dakota State University, Dickinson Research Extension Center (*Manske*, 1998) and the Minnesota Department of Agriculture (*MDA*, 2009).

Concerning the outdoor manure system, an empirical equation (Eq. 3) is used based on the studies of *Heber et al.* (2002). This equation considers the area of the storage system and the wind speed above the ground.

$$OE_o = S \cdot 10^{-0.56+0.671V} \cdot A_s, \tag{3}$$

where

OE_o is the outdoor odor emission (OU/s),
S is the surface area of the outdoor manure storage, (m^2),
V is the wind speed, (m/s), and
A_s is the outdoor odor abatement factor (0.30 – 1.00).

After calculating the emissions for each subsystem of the animal production farm, the results are summed to obtain the final emission rate (Eq. 4):

$$OU/s = \Sigma OE + \Sigma OE_o. \tag{4}$$

Table 1. Livestock units for different animals (*Manske*, 1998; *MDA*, 2009)

Animal type	LU
Swine	
Swine heavier than 135 kg	0.400
Swine between 25 – 135 kg	0.300
Swine under 25 kg	0.050
Cattle	
Weaned cow lighter than 360 kg	0.750
Young cow between 360 – 405 kg	0.850
Cow between 405.–.495 kg with calf	1.000
Cow between 495 – 585 kg with calf	1.150
Cow heavier than 585 kg with calf	1.250
Bull lighter than 900 kg	1.500
Bull heavier than 900 kg	2.000
Calf	0.200
Horse	1.000
Lamb	0.100
Sheep	0.100
Chicken	
Broiler (liquid manure system)	0.033
Chicken heavier than 2.25 kg (dry manure system)	0.005
Chicken lighter than 2.5 kg (dry manure system)	0.003
Turkey	
Turkey heavier than 2.25 kg	0.018
Turkey lighter than 2.5 kg	0.005
Duck	0.010

2.2.2. Municipal solid waste landfills

Odor emission calculations for municipal solid waste landfills are based on Italian studies (*Sironi et al.*, 2005). For the estimation of odor emissions, the following parameters have to be taken into consideration: (1) annual waste acceptance, (2) waste density, (3) working days in one year, (4) height of the daily deposited waste layer, (5) surface of the active parcels, and (6) surface of the restored parcels. Eq. (5) expresses the relationships of these parameters:

$$OE_w = 59R/(\rho \cdot D \cdot H) + 8S_A + 4S_R, \tag{5}$$

where

OE_w is the odor emission from the municipal solid landfill, (OU/s),
R is the annual waste acceptance, (t/yr),
ρ is the waste density, (t/m^3),
D is the working days, (day/yr),

H is the height of daily deposited waste layer, (m/day),
S_A is the surface of active parcel, (m^2), and
S_R is the surface of the restored parcels, (m^2).

The results of the OT emissions module for all cases (animal production farms, manure storage systems, municipal solid waste landfills) are given in odor units/second. One odor unit (OU) in this paper is equal to one European Odor Unit (OU_E), which was defined above. The output of the OT emissions module is one of the data inputs to the HNS-TRANSMISSION model, which calculates the dispersion of odorants. HNS-TRANSMISSION normally expresses emissions in units of kilogram/hour, which is not equal to odor unit/second. The conversion is: 1 OU/s = $3.6 \cdot 10^{-6}$ kg/h.

3. Odor dispersion in the air

The OT model was developed from the HNS-TRANSMISSION model (*Szepesi et al.*, 2005), which is used for the calculation of pollutant dispersion in the atmosphere. The model is well-known and commonly used in air quality modeling practice in Hungary. The software considers meteorological parameters (wind speed, wind direction frequency, and atmospheric stability) and uses a long-period meteorological database for the whole area of Hungary. This database can be modified for use in other specific locations as well. Land-use and topographic conditions are taken into consideration by parameters specified in the software. The model calculates the concentrations of given pollutants at user-specified distances from the sources. Concerning odor nuisance modeling, the required data for the OT model include the following: (1) coordinates of the livestock sites (geometrical coordinates are used in order to determine the center of the emissions), (2) odor emissions (discussed above), and (3) properties of the emitted air (temperature, rate of flow). Based on these parameters, the OT model calculates odor concentrations along the 16 wind directions at user-specified distances from each source.

The HNS-TRANSMISSION model incorporates a Gaussian dispersion model. This is the most widely used model type for plume dispersion calculations. Its most attractive feature is that it fits what we see and experience in the real world for a range of conditions. In addition, the mathematics of the model is fairly straightforward. On the other hand, Gaussian models need significant empirical input to be used for practicable dispersion estimates, making the model results highly dependent on the conditions of the sampling used to derive the empirical values.

Similarly to the decision process used to select the appropriate model for regulatory purposes, the selection of the appropriate dispersion model for odor assessment starts with the source type and release scenario. In general, most sources can be categorized as point, area, or volume sources, with continuous or

instantaneous releases. The sources responsible for odor complaints are generally continuous sources, such as from stacks, scrubbers, or basins, although routine but instantaneous or very short-term releases (for example, from digester pressure release valves) can also pose problems at nearby receptors. Depending upon the rate of release relative to the short time frame of odor perception, intermittent sources can be classified as either continuous sources (release rate on the order of minutes or longer), or instantaneous sources (release rate on the order of seconds). Short-time concentrations are estimated from 30-minute maximum plume axis concentrations.

The OT model determines the number of odor exceedances at selected distances from sources along each of the 16 wind directions. The calculation requires the determination of the threshold limit for odors. In some European countries the odor threshold limit varies from 3 to 10 OU/m^3 (*Table 2*).

By setting a specific number of odor exceedances, the OT model calculates the setback distance along each wind direction around the virtual emission point. The setback distance represents the area within which odor nuisances are expected to occur more frequently than allowed in a year. For example, in an area where the odor exceedance probability is 1%, odor nuisances would be expected to occur on 88 hours per year. To develop a standard modeling system, an odor exceedance probability that is unacceptable needs to be determined by regulatory authorities, based on appropriate studies. Internationally, the most recently used value is 2%, which means that odor nuisances are expected to occur on 176 hours per year.

Table 2. Odor threshold limits in some European countries (*Ritvay* and *Kovács*, 2006)

Country	Odor threshold limit (OU/m^3)
Denmark	5 – 10
The Netherlands	5
Ireland	3 or 6
Norway	5 – 10
Hungary (suggested)	3 – 5

4. Presentation of setback distances on a virtual map

The OT model plots setback distances on Google®-type maps with an accuracy of ± 1 meter. In order to define the setback distance accurately, dimensions of the animal farm must be known. Based on these dimensions, the center, which is taken as the virtual emitting point, is defined. The geographical longitude and latitude coordinates of this point are the bases of the plotting system. The setback distances are measured from this point. By plotting the setback distance on each of the 16 wind direction axes, the odor impact zone can be illustrated on

an appropriate map. *Fig. 1* shows an example of the setback distance visualization around an animal farm. In this case the odor threshold limit was 3 OU/m^3, while the exceedance probability was 2%.

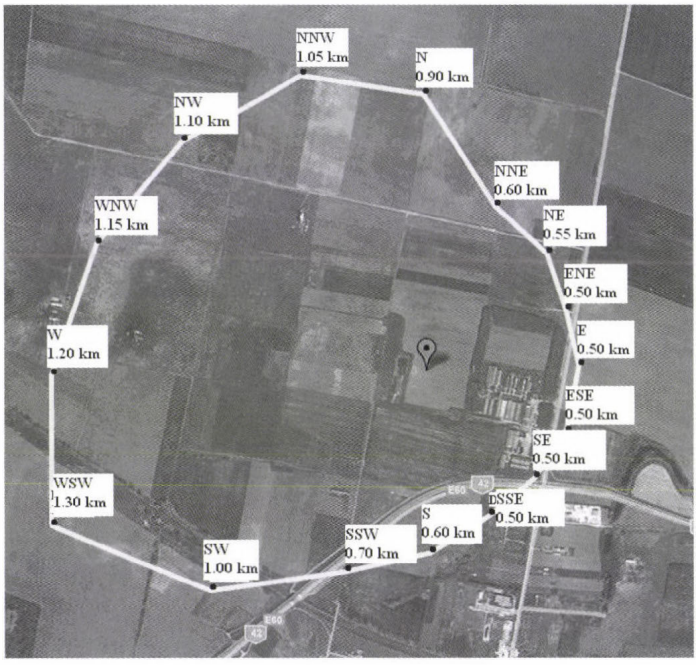


Fig. 1. Visualization on a Google map of the setback distances around an animal farm. The distances are measured from the virtual emitting point along the 16 wind directions.

This final visualization is a helpful tool for decision-makers, as the impact zone is well determined and visible. Furthermore, the mode of the map can be switched from any normal map to satellite-based (e.g., from Google®) and vice versa.

5. Setback distance determination around a municipal solid waste landfill

A case study was conducted to calculate odor emissions and setback distances around a municipal solid waste landfill in Dunakeszi, Hungary. The parameters of the landfill are given in *Table 3*.

Odor emissions are calculated based on Eq. (5). The output data of the emission calculations are also presented in *Table 3*. Based on the odor emissions, the frequency of odor exceedances was calculated using a specified 4 OU/m^3 threshold limit. The required meteorological parameters (wind direction frequency, wind speed, and atmospheric stability) for the dispersion calculations

are given in the database of the HNS-TRANSMISSION model. From the analysis of the frequency of exceedances for the number of hours in a year, the resulting setback distances along the 16 wind direction axes are plotted around the emitting point on a Google map (*Fig. 2*).

Table 3. Input and output data for odor emission calculations for the municipal solid waste landfill in Dunakeszi, Hungary

Input data:	
Annual waste acceptance	200,000 t/yr
Waste density	0.6 t/m^3
Working days	300 day/yr
Height of daily deposited waste layer	3 m/day
Surface of active parcels	3000 m^2
Surface of the restored parcels	22,000 m^2
Output data:	
Emissions from the daily deposited waste layer	21,852 OU/s
Emissions from active parcels	24,000 OU/s
Emission from restored parcels	88,000 OU/s
Total odor emissions (OU/s)	133,852 OU/s

Fig. 2. Odor setback distances along the 16 wind directions around a municipal solid waste landfill in Dunakeszi, Hungary.

6. *Model testing and validation*

For testing and validation purposes, the OT model was compared to three international odor models: W-T model, Purdue model, and Minnesota OFFSET model (*Guo et al.*, 2004). The first two models are based on empirical principles, while the OFFSET model incorporates an atmospheric dispersion module.

In this study, odor setback distance calculations were made for 12 swine farms in the USA. The livestock facilities were located in the State of Minnesota. The sizes of the swine farms were different, and the type and number of animals also differed as shown in *Table 4*.

Table 4. Characteristics of the 12 livestock operation farms (*Guo et al.*, 2004)

Farm	Types of swine	Odor sources	
		Buildings (m^2)	Outside manure storage
1	960 nursery to finishing	4 barns (735)	None
2	1720 finishing	2 barns (1637)	None
3	2500 nursery/finishing	7 barns (2725)	None
4	750 sows	2 barns (1869)	1 lagoon (91 × 91 m)
5	600 sows, 2500 nursery/finishing	6 barns (3450)	1 earthen basin (31 × 38 m)
6	2000 nursery, 1000 sows	3 barns (3534)	1 earthen basin (61 × 61 m)
7	1300 sows farrowing to weanling	3 barns (3348)	2 earthen basins (61 × 48 m, 61 × 61 m)
8	1300 sows, 4000 nursery	3 barns (4167)	2 earthen basins (58 × 58 m, 58 × 61 m)
9	1400 sows, 2800 nursery	4 barns (4508)	2 earthen basins (48 × 48 m, 48 × 76 m)
10	2400 sows farrowing to weanling	3 barns (6882)	1 tank (1116 m^2), 1 basin (61 × 76 m)
11	4600 sows farrowing to weanling	6 barns (13020)	2 tanks (1116 m^2), 1 basin (61 × 122 m)
12	3500 nursery, 3500 finishing	5 barns (4185)	2 earthen basins (61 × 152 m, 61 × 203 m)

Based on the given data concerning farm size, swine types, outdoor manure storage, and meteorological parameters measured at the Rochester, MN Airport, odor setback distances were calculated using the newly developed OT model and were compared to the results of other methods – the W-T model, the Purdue model, and the OFFSET model, as summarized in *Guo et al.* (2004). In this study, the terrain around each swine farm was assumed to be flat.

The calculated setback distances for the 4 models are shown in *Table 5*. Although the OT model calculates setback distance for the 16 wind directions, in this study the maximum setback distance was considered around each swine farm, irrespective of wind direction. The numbers (kilometers) are the maximum setback distances, considering the percent probability of a nuisance in a year. In the case of the OT and OFFSET models, the given percentages represent the percentage of hours per year when an odor threshhold would likely be exceeded. The two cases for the Purdue model are the farthest (f) and shortest (s) setback distances that represent the 'worst' and 'best' cases. The W-T model calculates only the maximum setback distance.

Considering the OFFSET model, which also uses an atmospheric dispersion model for the setback distance determination, the distances for the 2, 3, and 4% exceedance probabilities show good agreement with the OT model. In the case of the Purdue model, the agreement is best with the 3, 4, and 5% exceedance probabilities. In general, the best agreement between the OT and W-T models – which is based on an empirical formula – occurs, when comparing the 2% exceedance probability distances (*Fig. 3*).

Table 5. Setback distances (km) determined by four models (OT, OFFSET, W-T, and Purdue). Odor setback distance is shown for the 12 swine farms considering different exceedance probabilities (the percentages refer to the exceedance probabilities in terms of the hours in one year)

		Setback distances (km)											
		Farm											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
OT	1%	0.35	0.65	0.75	2.05	1.20	1.60	1.90	2.50	2.20	2.25	3.45	5.50
	2%	0.30	0.55	0.65	1.75	1.00	1.40	1.65	2.15	1.90	1.90	2.95	4.60
	3%	0.25	0.40	0.50	1.25	0.75	1.00	1.15	1.50	1.35	1.35	2.05	3.20
	4%	0.25	0.40	0.50	1.20	0.70	0.95	1.10	1.40	1.25	1.30	1.95	3.20
	6%	0.15	0.30	0.35	0.90	0.55	0.70	0.85	1.10	0.95	0.95	1.45	2.30
	9%	0.15	0.25	0.30	0.75	0.45	0.60	0.70	0.90	0.80	0.80	1.15	1.70
OFFSET	1%	1.09	1.55	2.03	2.31	2.43	2.57	2.73	2.77	3.08	3.43	4.52	4.61
	2%	0.63	0.92	1.21	1.39	1.46	1.55	1.66	1.68	1.88	2.10	2.80	2.86
	3%	0.40	0.59	0.78	0.89	0.94	0.99	1.06	1.07	1.20	1.34	1.80	1.84
	4%	0.28	0.43	0.58	0.67	0.71	0.76	0.81	0.82	0.93	1.05	1.44	1.47
	6%	0.20	0.31	0.41	0.48	0.51	0.54	0.58	0.59	0.67	0.75	1.03	1.05
	9%	0.14	0.21	0.29	0.34	0.36	0.39	0.42	0.43	0.49	0.55	0.78	0.80
W-T		0.37	0.56	0.77	1.55	1.09	1.43	1.77	1.71	1.80	2.06	2.72	3.83
Purdue	f	0.40	0.55	0.64	0.81	1.18	1.01	1.08	1.21	1.13	1.45	2.00	1.37
	s	0.18	0.24	0.42	0.40	0.77	0.66	0.71	0.81	0.74	0.95	1.31	0.97

Model results differ mainly because of the different calculation methods. The OFFSET model uses meteorological parameters and contains the

atmospheric dispersion model INPUFF-2 (*Petersen* and *Lavdas*, 1986). The OT model also incorporates a Gaussian dispersion model, however, the odor emission calculations differ between the two. Although the OT model incorporates the odor emission techniques in the Purdue model, wind frequency is the only meteorological parameter considered in the Purdue model. This affects the model results. Again, the W-T model is an empirical model, developed by correlating odor emissions with data on the spatial extent of odor complaints (*Guo et al.*, 2004).

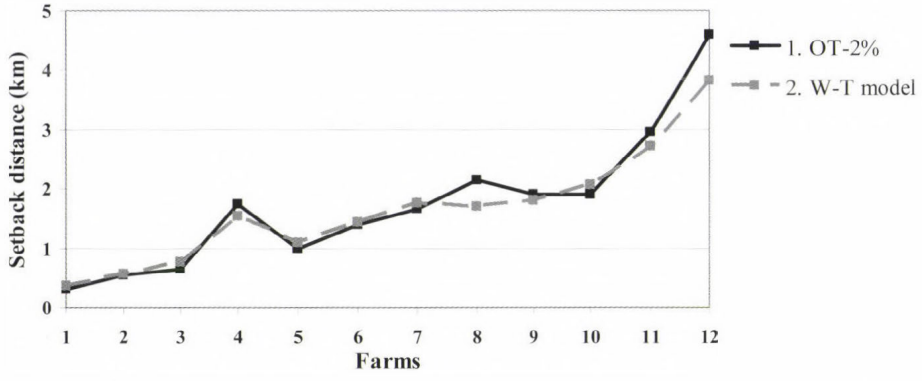


Fig. 3. Comparison of maximum setback distances: OT model and W-T model. The OT model distances are for the 2% exceedance probability case. The W-T model distances are the maxima calculated.

7. Sensitivity study

A sensitivity study was conducted in order to compare setback distance as a function of animal type, number of animals, and odor threshold limit. The study site is near the city of Kecskemét (Törökfái) in Hungary, and the exceedance probability considered was 2%. First, odor emission calculations were carried out for chickens, cattle, and ducks. Two cases were considered based on the number of animals. The results (*Table 6*) show that if the number of animals is doubled, the setback distance grows, but not in direct proportion to the changes in number of animals. Considering different threshold limits (3 OU/m^3, 4 OU/m^3, 5 OU/m^3), the setback distance is inversely proportional to the odor threshold limit.

The setback distance was also compared using a swine farm containing 10000 swine as a function of the odor threshold limit. The change in the setback distance was studied using 3, 4, and 5 OU/m^3. *Figs. 4*, *5*, and *6* clearly show that as the odor threshold limit increases, the maximum setback distance decreases. In the case of 3 OU/m^3 (*Fig. 4*), the maximum setback distance from the

emitting point is 1.75 km. This value is 1.5 km if the odor threshold limit is set at 4 OU/m^3 (*Fig. 5*), while for 5 OU/m^3(*Fig. 6*), the setback distance is just 1.3 km.

Table 6. Sensitivity analysis based on a livestock production farm near Kecskemét. Setback distance is analyzed as a function of odor emission and threshold limit

Animal	**Number of animals**	**Odor emissions**	**Setback distance (m) as a function of odor threshold limit**		
		OU/s	**3 OU/m^3**	**4 OU/m^3**	**5 OU/m^3**
Chickens	100000	107100	600	500	450
	200000	214200	950	850	700
Cattle	2000	74520	500	400	350
	4000	149040	800	650	550
Ducks	100000	32400	250	200	150
	200000	64800	450	350	300

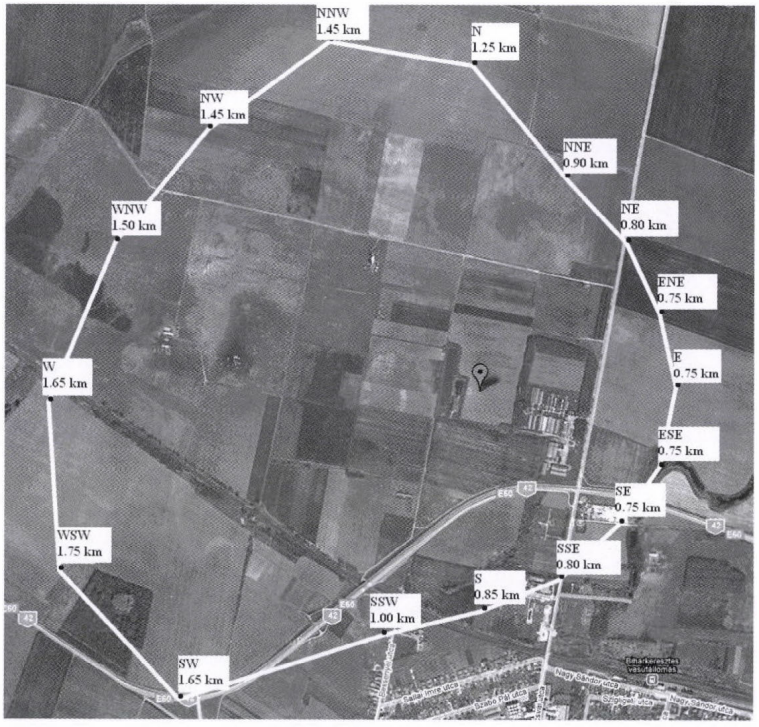


Fig. 4. Maximum setback distances along the 16 wind directions around a swine farm (10,000 swine) in the case of a 3 OU/ m^3 threshold limit. The maximum setback distance is 1.75 km.

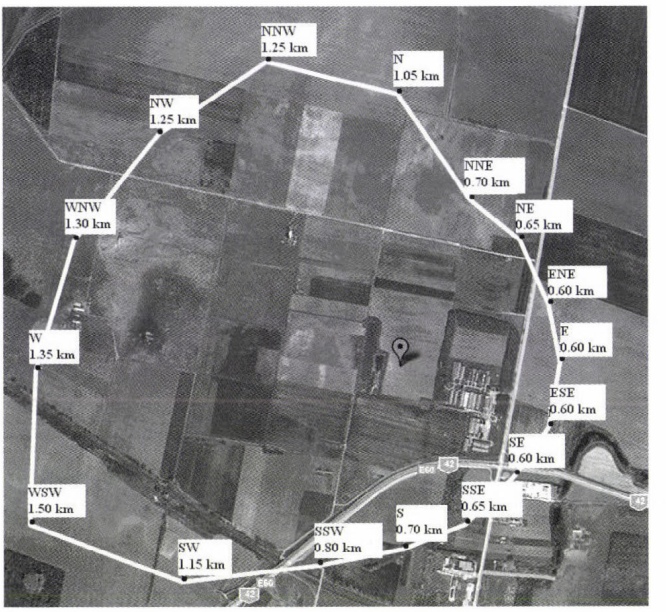


Fig. 5. Maximum setback distances along the 16 wind directions around a swine farm (10,000 swine) in the case of a 4 OU/ m^3 threshold limit. The maximum setback distance is 1.50 km.

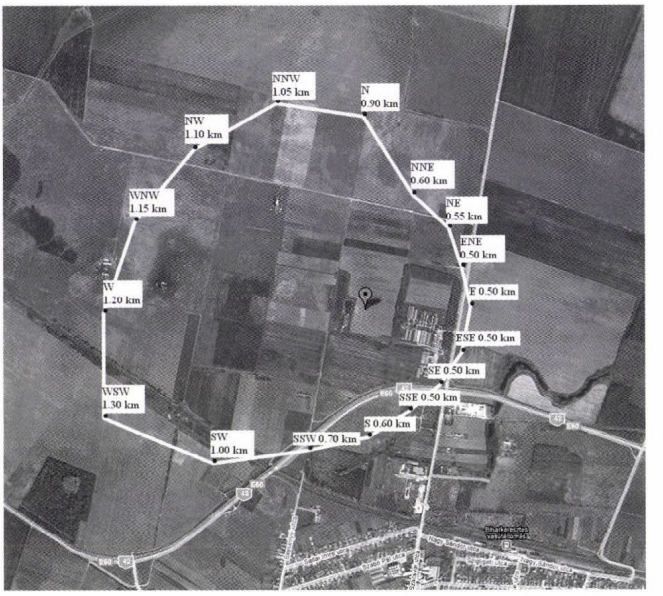


Fig. 6. Maximum setback distances along the 16 wind directions around a swine farm (10,000 swine) in the case of a 5 OU/ m^3 threshold limit. The maximum setback distance is 1.30 km.

8. Conclusions

The ODOR-TRANSMISSION (OT) model has been shown to be a very useful, objective method for determining setback distances around odor-emitting sources. Using the model is quite straightforward and does not require special skills. If the input data are correctly provided, then odor annoyances can be determined quantitatively and in a standardized way.

The odor emissions module in the OT model is based on several studies of animal farms and considers the livestock unit of each type of animals (*Manske*, 1998; *MDA*, 2009). The OT model considers not only animal type and number, but also manure handling methods, indoor or outdoor manure storage buildings, and odor abatement technologies as well. The OT model executes a Gaussian model to calculate the dispersion of odors from their emission point(s). The model currently contains a meteorological database for Hungary (wind direction frequency, wind speed, and atmospheric stability) and can also consider topography around the source. Using visualization techniques to portray model results on a map, e.g., those available via Google®, one can show, in accurate and understandable terms, the area around an emission point where odor annoyances are likely to occur. Model results can help decision-makers and other experts control, plan for, and regulate odor sources.

References

CEN, 2003: Air quality determination of odour concentration by dynamic olfactometry. Ref. No.: EN13725: 2003. Management Centre, Rue de Stassart 36, Brussels.

Guo, H., Jacobson, L.D., Schmidt, D.R., Nicolai, R.E. and *Janni, K.A.,* 2004: Comparison of five models for setback distance determination from livestock sites. *Canadian Biosystems Engineering 46,* 6.17-6.25.

Heber, J., Ni, J.-Q and *Lim, T.T.,* 2002: Odor flux measurements at a facultative swine lagoon stratified by surface aeration. *Applied Engineering in Agriculture 18*, 593-602.

Lim, T.T., Heber, A.J., Ni, J-Q., Grant, R. and *Sutton, A.L.,* 2000: Odor impact distance guideline for swine production systems. In *Water Environment Federation, Odors and VOC Emissions 2000.* Odors/VOC Emissions Conference, Cincinatti, OH, April 17-19, 2000.

Manske, L.L., 1998: Animal unit equivalent for beef cattle based on metabolic weight. *Research Report*, North Dakota State University, Dickinson Research Extension Center.

MDA, 2009: *Minnesota Department of Agriculture*: The Minnesota Livestock Producer's Feedlot Planning and Operations Manual. *Earth Tech, Inc.,* Minneapolis, Minnesota, http://www.mda.state.mn.us/en/sitecore/content/Global/MDADocs/animals/siting/feedlotguide.aspx.

Petersen, W.B. and *Lavdas, L.G.,* 1986: *INPUFF-2.0 A Multiple Source Gaussian Dispersion Algorithm User's Guide.* EPA/600-8-86-024. Washington, DC.

Ritvay, D. and *Kovács, D.*, 2006: The international and Hungarian practice of odor emission control (in Hungarian). National Environmental Conference, Siófok, Sept 19-21, 2006, http://www.vituki.hu/files/buz_szabgyak.pdf.

Schauberger, G. and *Piringer, M.,* 1997: Guideline to assess the protection distance to avoid annoyance by odour sensation caused by livestock husbandry. *Proceedings of the Fifth International Livestock Environment Symposium*, Bloomington (Minnesota), May 29-31, 1997, 170-178.

Sironi, S., Capelli, L., Céntola, P., DelRosso, R., and *Il Grande, M.,* 2005: Odour emission factors for assessment and prediction of Italian MSW landfills odour impact. *Atmos. Environ. 39,* 5387–5394.

Szepesi, D., Fekete, K., Büki,, R., Koncsos, L., and *Kovács, E.*, 2005: Development of regulatory transmission modeling in Hungary. *Időjárás 109*, 257-279.

Williams, M.L. and *Thompson, N.*, 1985: The effects of weather on odour dispersion from livestock buildings and from fields. In *Odor Prevention and Control or Organic Sludge and Livestock Farming* (eds: *V.C. Nielsen, J.H. Voorburg* and *P. L'Hermite*). Elsevier Applied Science Publishers, New York, 227-233.

Winneke, G. and *Steinheider, B.*, 1994: Exposure-response associations between environmental odours, traffic noise, annoyance and somatic complaints. In *Umwelthygiene – Supplement 2.* 1st Internationaler Kongress für Umwelt-medizin. Gesellschaft für Hygiene und Umweltmedizin/-Medizinisches Institut für Umwelthygiene an der Heinrich-Heine-Universität Düsseldorf (Hrsg.), 87-89.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 114, No. 4, October–December 2010, pp. 319–324

Comments on "Greenhouse effect in semi-transparent planetary atmospheres" by Ferenc M. Miskolczi

Henk A. R. de Bruin

*King's College, London, U.K.; Freelance Researcher, Bilthoven, The Netherlands**

(Manuscript received in final form November 12, 2010)

Abstract—This commentary is meant to show that several relationships derived in *Miskolczi* (2007) are debatable and, in my opinion, based on untenable physics.

Key-words: greenhouse effect, radiative equilibrium

1. Introduction

As lecturer emeritus at the University of Wageningen, the Netherlands, I was involved in introductory courses of meteorology, and I used simple 'flat earth greenhouse' models to illustrate the optical effects of greenhouse gases on climate. In these models the atmosphere is replaced with a hypothetic 'glass-plate' with a prescribed spectral emissivity $\varepsilon(\lambda)$ as function of wavelength λ mimicking the optical properties of the different greenhouse gases as well as the so-called atmospheric window. In the context of computer-aided-learning (CAL), I developed a number of interactive CAL modules with which the students could simulate the climate on Earth or Mars by selecting themselves appropriate functions for $\varepsilon(\lambda)$. I always stressed that these CAL modules were meant to clarify a single aspect of the climate system, notably, the optical effects of greenhouse gases, and that they are too simple to describe the actual climate, let alone climate sensitivity. Recently, I came across the papers by *Ferenc Miskolczi* (2007, 2010; hereafter M07 and M10, respectively) and by *van Andel* (2010; hereafter VA10) forcing me to reconsider my understanding of the optical effects of greenhouse gases on climate. I was challenged by van Andel who wrote: *The Miskolczi theory is so different from the usual stream of thought on climate change, that many are either confused, or revolt or embrace this*

*Associate Professor Emeritus Meteorology and Air Quality Group, Wageningen University, Wageningen, The Netherlands; E-mail: hardb@xs4all.nl

theory without really having penetrated it. In this commentary I try to penetrate parts of the Miskolczi theory. It is my objective to show that at several points this theory is untenable.

2. *Brief summary of M07*

The novel laws derived in M07 are:

- Referring to Kirchhoff's law and analyzing sets of radiosonde data with a line-by-line radiation transfer model by *Miskolczi* and *Mlynczak* (2004), it is found that

$$A_A = E_D. \tag{1}$$

- Referring to hydrostatic equilibrium, the virial theorem and identifying E_U as the total internal kinetic energy of the atmosphere, it is stated that

$$E_U = S_U/2. \tag{2}$$

- By applying the principle of conservation of energy and arguing that $S_U - (F^0 + P^0)$ and $E_D - E_U$ represent two flux terms of equal magnitude, propagating into opposite directions, while using the same F^0 and P^0 as energy sources, it is concluded that

$$S_U - (F^0 + P^0) + E_D - E_U = F^0 + P^0. \tag{3}$$

- Next it is stated that the latter equation poses a strict criterion on the global average S_U:

$$S_U = 3OLR/2. \tag{4}$$

- Furthermore, according to M07 this leads to

$$S_U - (F^0 + P^0) = R, \tag{5}$$

where R is the *pressure of thermal radiation at the surface,* which, according to M07 is given by

$$R = S_U/3. \tag{6}$$

- By arguing that the *virial term* $S_V = S_T/2 - E_D/10$ forces the hydrostatic equilibrium while maintaining the radiative balance, it is stated that

$$S_U + S_T/2 - E_D/10 = 3OLR/2. \tag{7}$$

- According to M07, this leads to

$$S_U - (F^0 + P^0) = 6 \text{ A } R/5. \tag{8}$$

- Finally, it is stated in M07 that, among others, the above yields

$$E_U / E_D = 3/5. \tag{9}$$

My comments concern these novel findings.

3. *My comments*

3.1. $A_A = E_D$

In a personal communication, dr. Miskolczi agreed that referring to Kirchhoff's law (and later in M10 to *Prevost*, 1791) can lead to misunderstandings, so I will not comment on these references, except that they are incorrect. The main evidence is to show that $A_A = E_D$ is provided by the calculations of A_A and E_D using the line-to-line radiation transfer model of *Miskolczi* and *Mlynczak* (2004). Arguing that there is *radiative exchange equilibrium between land-sea surface and atmosphere*, these calculations are performed using the boundary condition that the surface temperature equals the air temperature, t_A, at the lowest level in the radiosonde dataset, usually 1.5 or 2 m. Moreover, it is assumed that the surface is black for long-wave radiation. So, in the line-to-line calculations yielding A_A and E_D, the boundary conditions

$$S_U = \sigma t_a^4 \tag{10}$$

are used.

This choice or assumption explains why it is found that $A_A = E_D$. Experience shows that E_D can be approximated accurately with

$$E_D = \varepsilon_{atm} \sigma t_a^4, \tag{11}$$

where ε_{atm} is an apparent emissivity of the atmosphere. Accurate radiation transfer will yield an E_D value for a given radiosonde sounding, so this can be written as $E_D = \varepsilon_{atm} \sigma t_a^4$. In fact, the model yields a value for ε_{atm}. If Eq. (10) is used as boundary conditions for the calculation of A_A with the same model, one will find (due to Kirchhoff's law applied to each absorption line) that $\varepsilon_{atm} = \alpha_{atm} = A$ (apparent emissivity equals the apparent absorptivity), by which one obtains $E_D = A_A$. In fact, evidence given in Fig. 2 in M07 supports this argument. For a surface emissivity of 0.96, E_D appears to be not equal to A_A, because now $S_U \neq \sigma t_a^4$.

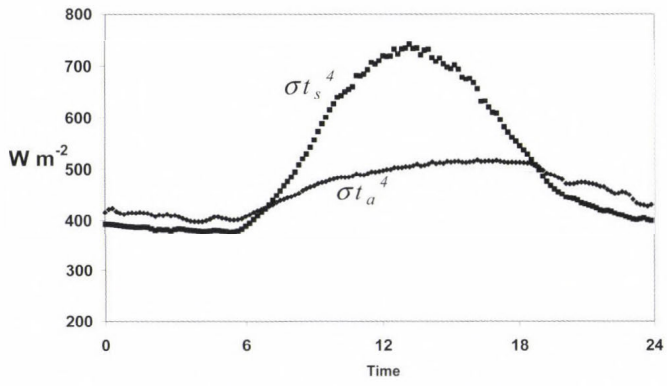


Fig. 1. Diurnal cycle of σt_a^4 and σt_s^4, observations during the MATDADOR field campaign in Arizona desert, t_s is observed with an IR-thermometer.

Is the boundary condition $S_U = \sigma t_a^4$ correct? I don't think so. In textbooks on micrometeorology (e.g., *Geiger*, 1971; *Oke*, 1987; *Garratt*, 1992) it is explained why, in general, the surface temperature is not equal to t_a. In a desert, the difference can exceed 40 °C, but also over grass in mid-latitude climate zones, the difference can be 5 to 10 °C. In the last decade, the sensors to measure surface temperature S_U and E_D are improved significantly, and in open literature many micrometeorological data sets collected over a wide range of surface types can be found. Several of these can be downloaded freely. Some of these are at my disposal and I have sent dr. Miskolczi some results showing clearly that $S_U \neq \sigma t_a^4$. One example is depicted in Fig. 2 of M07. Usually, S_U is greater than σt_a^4 during daytime. Although during nighttime the opposite occurs, the daily mean of the difference is not zero. See, for example, *Fig. 1*. Among others, this is due to non-linearity effects in atmospheric turbulence transfer and, not at least, because during daytime the surface is heated by solar radiation, whereas nighttime cooling is governed by entirely other physical processes. I conclude that, in general, $E_D \neq A_A$. The evidence provided in M07 to show that $E_D = A_A$ is based on the incorrect assumption that $S_U = \sigma t_a^4$. Consequently, all conclusions in M07 drawn from $E_D = A_A$ are untenable. Because in the later sections in M07 often the assumption that the surface temperature equals t_a is adopted, most results derived from that must be considered with caution also.

3.2. Comment on virial theorem and $E_U = S_U/2$

Firstly, I would like to remark that $E_U = S_U/2$ cannot be true in general. Suppose that the atmosphere consists of 20% O_2 and 80% N_2, so solely of diatomic gases, then there is no LW absorption or emission, so $E_U = 0$, whereas, because most

surface have emissivity close to 1, S_U will be greater than 0 due to solar heating. I give this simple example also to introduce my comment on the statement in M07: "*the total internal kinetic energy of the atmosphere, which – according to the virial theorem – hydrostatic equilibrium balances the total gravitational potential energy*", because the actual atmosphere consists primarily of O_2 and N_2, so of diatomic gases. The virial theorem is developed in astrophysics, where it is applied to, e.g., gaseous stars. Without any explanation, the virial theorem is applied to the atmosphere in the gravitation field of the Earth, which is quiet different from the gaseous star case. In stars one has to account for gravitation forces between different layers, whereas in the atmosphere these can be ignored, because they are much smaller than the gravitation force by the Earth. Thus, application of a virial theorem to climate topics is questionable. Anyhow, the virial theorem applied to gaseous stars deals with the potential and internal energy. In textbooks (*Holton*, 1979; *Iribarne* and *Godson*, 1986) one can find that for the atmosphere, in a gravity field that thermodynamic principles lead to $P=(\eta-1)U$, where P is the potential energy, U is the internal energy, and η is the ratio of specific heat at constant pressure and constant volume. Because U can be linked to *the total internal kinetic energy*, I presume that this is meant in M07. But for diatomic atmospheres $P=\frac{2}{5}U$, so P is not equal to U. This result was also found recently by *Toth* (2010). Let us ignore this factor 2/5 and also the question whether horizontal and vertical atmospheric motions of 'air parcels' are included in what is denoted as *total kinetic energy* (these are NOT covered by U, see, e.g., *Toth*, 2010), my main criticism concerns the fact that, without a single clue, in M07 U is linked to E_D. Note that E_D originates from LW radiation emission by three-atomic gases, and the above 'virial theorem' applies to diatomic atmospheres. I conclude that the derivation in M07 of $E_U = S_U/2$ is so unclear that, for the time being, it must be considered not to be based on the first physical principle. The same applies to all further formulas in M07 derived from the virial theorem and $E_U = S_U/2$.

3.3. Comments on $S_U - (F^0 + P^0) + E_D - E_U = F^0 + P^0$ and term R

This equation is based on the following arguments: "$S_U - (F^0 + P^0)$ and $E_D - E_U$ *represent two flux terms of equal magnitude, propagating into opposite directions, while using the same F^0 and P^0 as energy sources. The first term heats the atmosphere and the second term maintains the surface energy balance. The principle of conservation of energy dictates that*: $S_U - (F^0 + P^0) + E_D - E_U = F^0 + P^0$ ". I am sorry, this is entirely unclear to me. The conservation law has been applied before. Moreover, I cannot follow the algebra with which next leads to $S_U = 3OLR/2$ and $S_U - (F^0 + P^0) = R$. With $(F^0 + P^0) = OLR$ and $G = S_U - OLR$ the latter yields $R = G$.

The term R was introduced in M07 as *the pressure of thermal radiation at the surface.* Firstly, I have problems with the units. R is a pressure (force per m^2), whereas all other terms are energy flux densities??? But more importantly, in astrophysics one have to account for pressure of radiation indeed, but this pressure (*Jeans*, 1961, page 37) is given by $P_{rad} = \frac{1}{3} aT^4$, with $a = 4\sigma/c$, where c is the speed of light. In M07 it is stated that $R = \frac{1}{3} \sigma T^4$, which is greater than P_{rad} by a factor of $c/4$! The correct expression shows that LW radiation effects on atmospheric pressure can be ignored entirely. Consequently, any statement referring to P_{rad} must be considered inappropriate to explain atmospheric phenomenon. Statements such as "*G will always be equal to the radiation pressure of the ideal gas and the atmosphere will have a constant optical depth*" (page 8 in M07) and "*Formally, in the presence of a solid or liquid surface, the radiation pressure of the thermalized photons is the real cause of the greenhouse effect, and its origin is related to the principle of the conservation of the momentum of the radiation field.*" are incorrect.

3.4. Further comments

The algebra or physics leading to Eqs. (7)–(9) are not clear to me. In Eq. (7), suddenly the virial term $S_T/2 - E_D/10$ is introduced. Where does this come from? According to Eq. (4) this term is zero. Moreover, I do not understand the way Eqs. (8) and (9) are derived.

Acknowledgement — In content I disagree with his thoughts, but I would like to thank *dr. Miskolczi* for inspiring novel ideas about our climate system.

References

Garratt, J.R., 1992: *The Atmospheric Boundary Layer*. Cambridge University Press, 316 pp.

Geiger, R., 1971: *The Climate Near the Ground*. Harvard University Press, Cambridge, USA, 611 pp.

Holton, J.R., 1979: *An Introduction to Dynamic Meteorology*. Academic Press Inc., 391 pp.

Iribarne, J.V. and *Godson, W.L.*, 1986: *Atmospheric Thermodynamics*. Reibel Publ. Comp., Dordrecht, 259 pp.

Jeans, Sir J.H., 1961: *Astronomy* and *Cosmogony*. Dover, New York, 426 pp.

Miskolczi, F.M., 2007: Greenhouse effect in semi-transparent planetary atmospheres. *Időjárás 111*, 1-40.

Miskolczi, F.M., 2010: The stable stationary value of the Earth's global average atmospheric Planck-weighted greenhouse-gas optical thickness. *Energy and Environment 21*, 243-262.

Miskolczi, F.M. and *Mlynczak, M.G.*, 2004: The greenhouse effect and the spectral decomposition of the clear-sky terrestrial radiation. *Időjárás 108*, 209-251.

Oke, T.R., 1987: *Boundary Layer Climates*. Methuen, London and New York, 435 pp.

Prevost, P., 1791: Mémoire sur l'équilibre du feu. *Journal de Physique 30*, 314–323.

Toth, V.T., 2010: The virial theorem and planetary atmospheres, *Időjárás 114*, 229-234.

van Andel, N., 2010: Note on the Miskolczi theory. *Energy and Environment 21*, 277-292.

INSTRUCTIONS TO AUTHORS OF *IDŐJÁRÁS*

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

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

Some issues contain "News" and "Book review", therefore, such contributions are also welcome. The papers must be in American English and should be checked by a native speaker if necessary.

Authors are requested to send their manuscripts to

Editor-in Chief of IDŐJÁRÁS
P.O. Box 39, H-1675 Budapest, Hungary
E-mail: antal.e@met.hu

including all illustrations. MS Word format is preferred in electronic submission. Papers will then be reviewed normally by two independent referees, who remain unidentified for the author(s). The Editor-in-Chief will inform the author(s) whether or not the paper is acceptable for publication, and what modifications, if any, are necessary.

Please, follow the order given below when typing manuscripts.

Title page: should consist of the title, the name(s) of the author(s), their affiliation(s) including full postal and e-mail address(es). In case of more than one author, the corresponding author must be identified.

Abstract: should contain the purpose, the applied data and methods as well as the basic conclusion(s) of the paper.

Key-words: must be included (from 5 to 10) to help to classify the topic.

Text: has to be typed in single spacing on an A4 size paper using 14 pt Times New Roman font if possible. Use of S.I. units are expected, and the use of negative exponent is preferred to fractional sign. Mathematical formulae are expected to be as simple as possible and numbered in parentheses at the right margin.

All publications cited in the text should be presented in the *list of references*, arranged in alphabetical order. For an article: 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 photo-synthetically active radiation and cloud amount. *Időjárás 90,* 10-13. For a book: 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. Reference in the text should contain the name(s) of the author(s) in Italics and year of publication. E.g., in the case of one author: *Miller* (1989); in the case of two authors: *Gamov* and *Cleveland* (1973); and if there are more than two authors: *Smith et al.* (1990). If the name of the author cannot be fitted into the text: *(Miller,* 1989); etc. When referring papers published in the same year by the same author, letters a, b, c, etc. should follow the year of publication.

Tables should be marked by Arabic numbers and printed in separate sheets with their numbers and legends given below them. Avoid too lengthy or complicated tables, or tables duplicating results given in other form in the manuscript (e.g., graphs).

Figures should also be marked with Arabic numbers and printed in black and white or color (under special arrangement) in separate sheets with their numbers and captions given below them. JPG, TIF, GIF, BMP or PNG formats should be used for electronic artwork submission.

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 for authors is available: antal.e@met.hu

Published by the Hungarian Meteorological Service

Budapest, Hungary

INDEX 26 361

HU ISSN 0324-6329