IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

CONTENTS

****

http://omsz.met.hu/omsz.php?almenu_id=omsz&pid=references&pri=2

VOL. 109 * NO. 3 * JULY – SEPTEMBER 2005

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: P.O. Box 39, H-1675 Budapest, Hungary or
Gilice tér 39, H-1181 Budapest, Hungary
E-mail: bozo.l@met.hu or *antal.e@met.hu*
Fax: (36-1) 346-4809

Subscription by

mail: IDŐJÁRÁS, P.O. Box 39, H-1675 Budapest, Hungary;
E-mail: bozo.l@met.hu or *antal.e@met.hu*; *Fax: (36-1) 346-4809*

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 109, No. 3, July–September 2005, pp. 143–155

Radiance-bias correction for a limited area model

Roger Randriamampianina

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

(Manuscript received in final form August 1, 2005)

Abstract—Direct assimilation of satellite measurements requires correction of biases caused by both measurement problems and errors in the radiative transfer model. The bias correction is based on finding the difference between the observed radiances and those simulated from the model states. Since the bias correction method was originally developed for global models, its adaptation to limited area models (LAMs) raises further questions. The quality of the bias correction coefficients – scan-angle biases and coefficients for air-mass predictors – depends on the sample of the observation-minus-model-first-guess, obtained at each satellite (AMSU-A) scan position. The amount of satellite measurements along the scan line is much smaller in case of a limited domain (LAM) compared to global models. This can cause problems when evaluating the scan-angle biases for a limited area model. This paper investigates different bias correction coefficients in order to find the best method for processing satellite data in the ALADIN limited area model. Bias correction coefficients, computed for the French global (ARPEGE) model, and one computed for the ALADIN limited area model, and many of their combinations were tested. The impact of the bias correction coefficients computed for the ALADIN limited area model was found to be more "stable" in both the analysis and short-range forecasts. The impact of the bias correction coefficients computed for the global model depended on the synoptic situation of the investigated period, especially in the important for the synoptic meteorology 850 to 500 hPa layer.

Key-words: limited area model, radiance-bias correction, data assimilation, ATOVS/AMSU-A.

1. Introduction

In most numerical weather prediction (NWP) centres, satellite data are assimilated in the form of raw radiances. For the efficient use of raw radiances (in our case from ATOVS), biases between the observed radiances and those simulated from the model states (first-guess) must be removed.

Many investigations were carried out on the removal of these biases. *Eyre* (1992) introduced the radiance bias as the combination of the scan-angle dependent (originating form the measurement quality) and air mass dependent errors. *Harris* and *Kelly* (2001) showed that scan angle biases vary with the geographical latitude bands. *Dee* (2004) proposed an adaptive bias correction scheme that can automatically sense the change in the bias of a given channel and responses correspondingly. The bias parameters are then updated jointly and simultaneously with the model state during the variational analysis, and are fully consistent with all observational information available to the analysis. *Watts* and *McNally* (2004) introduced a bias correction scheme, which is based on a modification of the transmittance coefficients in the radiative transfer model (RTTOV), involving two global parameters for each channel that can be adjusted to reduce the systematic errors in the RTTOV calculations.

The proposed bias correction schemes, however, were developed for global models. Thus, their adaptation to limited area models (LAMs) raises further questions. The quality of the bias correction coefficients – scan-angle biases and coefficients for air-mass predictors – depends on the amount of the observation-minus-model-first-guess, obtained at each satellite (AMSU-A) scan position. The amount of satellite measurements along the scan line is much smaller in case of a limited domain (LAM) compared to global models, because satellite paths are likely to be cut at different scan positions during their pre-processing. This can cause problems when evaluating the scan-angle biases for a limited area model.

In ARPEGE/ALADIN model (*Horányi et al.*, 1996), the method described by *Harris* and *Kelly* (2001) is used for correcting radiance-biases (see Section 2.2). Scan-angle biases depend on the number of samples obtained at each scan position. When computing scan angle biases using a limited area model (LAM), it is not likely to have the same number of samples for all scan positions in a given channel. To illustrate this, two satellite paths – a complete one on the right and a portion of a second path on the left side of the domain – are shown in *Fig. 1*. The inadequacy in the number of samples leads to fluctuating bias curves along the scan-lines (*Fig. 2a)* instead of well-smoothed ones. Due to sufficient number of samples, this problem does not appear when computing the scan-angle biases for global models.

Fig. 2a demonstrates the statistics computed for a one month period for the old domain (*Fig. 3a*) of the ALADIN Hungary (ALADIN/HU) model, which is relatively small compared to the new one (*Fig. 3b*). Enlarging the domain, smoother curves were obtained (*Fig. 2b*). Less but valuable fluctuation, however, was still observed for several channels – see, for example, the curve representing the scan-angle bias for channel 9 of AMSU-A

(triangles in *Fig. 2b*). This indicates, that further efforts have to be done to improve the bias correction method for the ALADIN/HU LAM model.

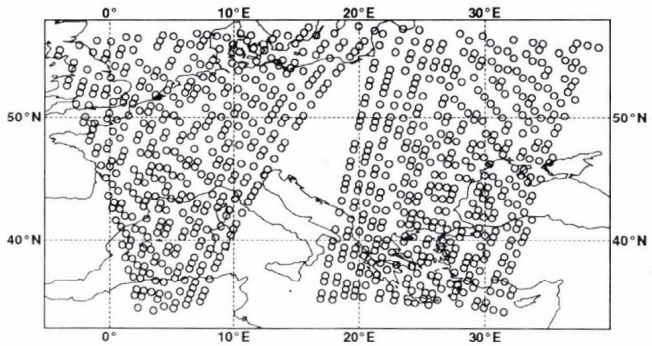

Fig. 1. Example of satellite paths inside the ALADIN/HU domain observed on April 23, 2003, 00 UTC.

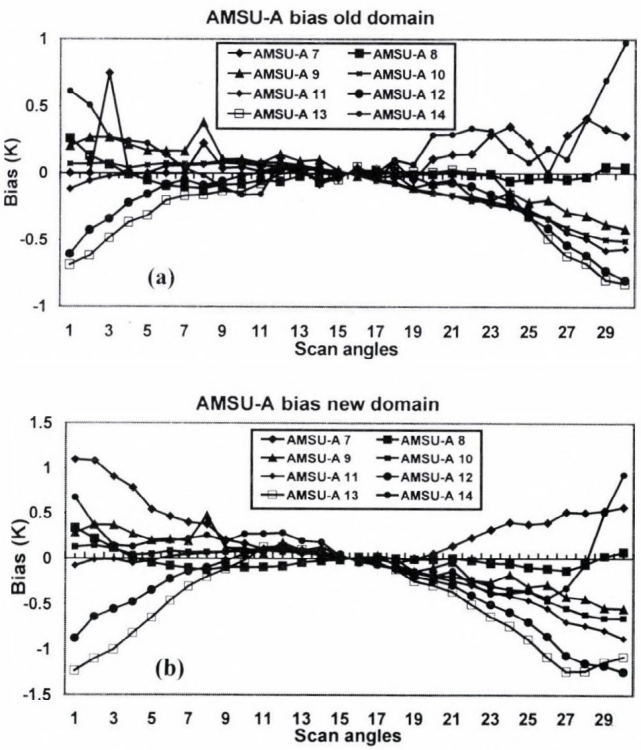

Fig. 2. Scan-angle bias computed for the old (*a*) and new (*b*) ALADIN/HU domains. Note that the domains of the ALADIN/HU model are presented in Figs. 3a and 3b.

This paper investigates different bias correction coefficients in order to find the best method for processing raw radiance satellite data in the ALADIN limited area model.

Bias correction coefficients, computed for the French global (ARPEGE) model and for the ALADIN limited area model, and many of their combinations were tested. Bias-correction coefficients computed for the restricted LAM domain were then compared with those, calculated for the coupling[1] global model. The need on removing air-mass related biases when assimilating the ATOVS observations in a limited area model was also investigated.

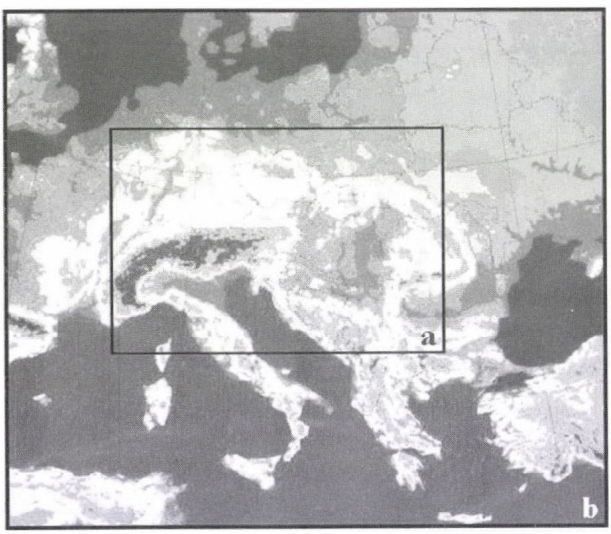

Fig. 3. Topography of the old (*a*) and new (*b*) ALADIN/HU domains, respectively.

Section 2 describes the main characteristics of ALADIN/HU model and its assimilation system. Section 2.1 illustrates the local pre-processing of satellite data, while Section 2.2 provides a short description of the bias correction method used in ALADIN/HU. Section 3 gives a detailed description of the experiments performed with various bias correction files. Section 4 reviews the results of the experiments, and in Section 5 we draw some conclusions of the results presented in this paper.

1 The integration of a limited area model needs information about its lateral boundary conditions – the coupling files. In the case of ALADIN model, we use a file from the global ARPEGE model, which is referred here as a coupling model.

2. *The ALADIN/HU model and its assimilation system*

At the Hungarian Meteorological Service (HMS), the ALADIN/HU model runs in its hydrostatic version. The model used in this investigation was the al15/cy24t1 version of the ARPEGE/ALADIN codes (http://www.cnrm.-meteo.fr/aladin/concept/historycycles.html). In this study we used the model with 12 km horizontal resolution (*Fig. 3b*) and 37 vertical levels from the surface up to 5 hPa. The three-dimensional variational data assimilation (3D-Var) system was applied to assimilate both conventional (SYNOP and TEMP) and satellite (ATOVS) observations. As the variational technique computes the observational part of the cost function in the observational space, it is necessary to simulate radiances from the model parameters. In ARPEGE/ALADIN (al15/cy24t1) we use the RTTOV6 radiative transfer code to perform this transformation (*Saunders et al.,* 1998), which has 43 vertical levels. Above the top of the model, an extrapolation of the profile is performed using a regression algorithm (*Rabier et al.,* 2001). Below the top of the model, profiles are interpolated to RTTOV pressure levels. A good estimation of the background error covariance matrix is also essential for the variational technique to be successful. The background error covariance – the so-called "B" matrix – is computed using the standard NMC method (*Parrish* and *Derber,* 1992; *Široká et al.,* 2003). The specific humidity was assimilated in univariate form to avoid certain problems, related to its assimilation (see *Randriamampianina* and *Szoták,* 2003 for more details). An optimal interpolation scheme was used to analyze the surface fields (*Radi* and *Issara,* 1994). The AMSU-A data were assimilated at 80 km resolution. The 3D-Var is running in 6-hour assimilation cycle generating an analysis at 00, 06, 12, and 18 UTC. In this study, we performed a 48-hour forecast once a day, starting from 00 UTC.

2.1 Pre-processing of satellite data

The ATOVS data are received through our HRPT antenna and pre-processed with the AAPP (ATOVS and AVHRR Pre-processing Package) software package. We used AMSU-A, level 1-C radiances in our experiments.

For technical reasons our antenna is able to receive data only from two different satellites. To acquire the maximum amount of satellite observations, the NOAA-15 and NOAA-16 satellites were chosen, which have orbits perpendicular to each other and pass over the ALADIN/HU domain at about 06 and 18 UTC, and 00 and 12 UTC, respectively.

For each assimilation time we used the satellite observations that were measured within ± 3 hours. The number of paths over the ALADIN/HU domain within this 6-hour interval varied up to three.

2.2 Bias correction

The direct assimilation of satellite measurements requires the correction of biases computed as differences between the observed radiances and those simulated from the model first guess. These biases, arising mainly from instrument characteristics or inaccuracies in the radiative transfer model, can be significant. The method developed by *Harris* and *Kelly* (2001) was used to remove this systematic error. This scheme is based on separation of the biases into scan-angle dependent bias and state dependent components. The air-mass dependent bias is expressed as a linear combination of the set of state-dependent predictors.

Four predictors computed from the first-guess fields were selected (p1 – the 1000–300 hPa thickness, p2 – the 200–50 hPa thickness, p3 – the skin temperature, and p4 – the total column water) for the AMSU-A data used in our experiments.

A carefully selected sample of background departures for the AMSU-A and channel set was used to estimate the bias, in a two-step procedure. First, scan bias coefficients were computed by separating the scan-position dependent component of the mean departures in latitude bands. Secondly, after removing the scan bias from the departures, the predictor coefficients for the state-dependent component of the bias were obtained by linear regression. At the end of this estimation procedure, bias coefficients for the AMSU-A were stored in a file. The data assimilation system could then access the coefficients in order to compute bias corrections for the latest observations, using update state information for evaluating the air-mass dependent component of the bias. The brightness temperatures were corrected accordingly, just prior to assimilation.

As ARPEGE model uses every second pixel of ATOVS measurements, it has zero scan-angle coefficients at non-used pixels, which may cause a large remaining bias when using one by one field of view of the AMSU-A data. To overcome this problem, the values of the two adjacent pixels were interpolated into pixels with zero coefficients.

3. Description of the experiments

In order to estimate the impact of different bias correction coefficients on the model analysis and forecasts, the scores of different experiments were compared with those from the run (NT80U) performed using the bias correction file, computed for the ALADIN/HU LAM model. The scores of each run were evaluated objectively. The bias and root-mean-square error

(RMSE) were computed from the differences between the analysis/forecasts and observations (surface and radiosondes).

A twenty-day period (April 18, 2003–May 07, 2003 – denoted as first period later on) was used for the first impact study that consisted of four experiments. A fifteen-day period (February 20, 2003–March 06, 2003 – denoted as second period later on) was chosen for the second impact study in order to confirm the main results of the first one by repeating some of the experiments.

The radiosonde (TEMP), surface (SYNOP), and AMSU-A observations were used in all the experiments, applying different bias correction methods:

NT80U: The bias correction coefficients were computed for the ALADIN/HU domain (control run).

T8B1I: The bias correction coefficients were computed for the ARPEGE model.

T8B2I: The scan angle coefficients were computed for the ARPEGE model, but no air-mass correction was applied.

T8B3I: The ARPEGE scan-angle coefficients and the air-mass bias correction coefficients computed for the ALADIN/HU were used.

NOT8U: The same as NT80U for the second period.

OT8B1I: The same as T8B1I for the second period.

OT8B3I: The same as T8B3I for the second period.

4. Results and discussion

Bias correction coefficients computed for the global ARPEGE and limited area ALADIN/HU models and their combinations were compared in order to find the best solution for processing the AMSU-A data in the ALADIN/HU model. Almost neutral impact of bias correction methods on the analysis and forecast of relative humidity, geopotential height, and wind speed was found.

Concerning the impact on the temperature, the results are classified as follows.

4.1 Comparison of biases using different bias correction files

The particularity of the data assimilation system at the HMS is that it has different (positive or negative) bias on temperature profile at different model levels. For example, clear positive and negative bias can be observed at the 1000 hPa and 850 hPa levels, respectively (*Fig. 4*). The bias on humidity profile is slightly positive for all the model levels (not shown).

According to our results, the bias coefficients for the global ARPEGE model (mentioned as global bias correction file later on) have a heating effect above and a cooling effect under the 500 hPa level (*Fig. 4*) compared to the control run. Our verification concerns only the levels below 100 hPa.

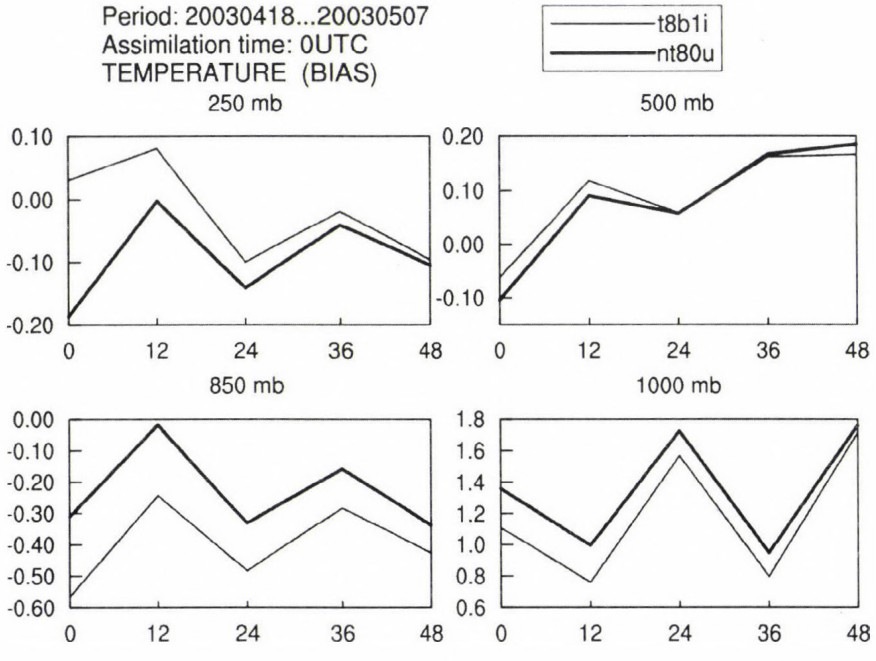

Fig. 4. Temperature biases, computed using the global (ARPEGE) bias correction coefficients (T8B1I) against biases, computed using the LAM coefficients (NT80U) for the first period.

4.2 Impact of the global bias correction file

Though the ALADIN/HU model has different biases of temperature in different model layers, the systematic cooling or heating does not necessarily yield an overall positive impact on temperature forecasts. For example, a clear positive impact on the forecast of temperature can be observed in the troposphere (500 hPa level) during the second period, although there was a negative impact at 850 hPa during the first period (*Fig. 5*). Thus, the behavior of the limited area model is not fully "controllable" when applying the global bias correction file in the assimilation system to process satellite observations. Consequently, no stable impact on the model analysis and forecast can be obtained.

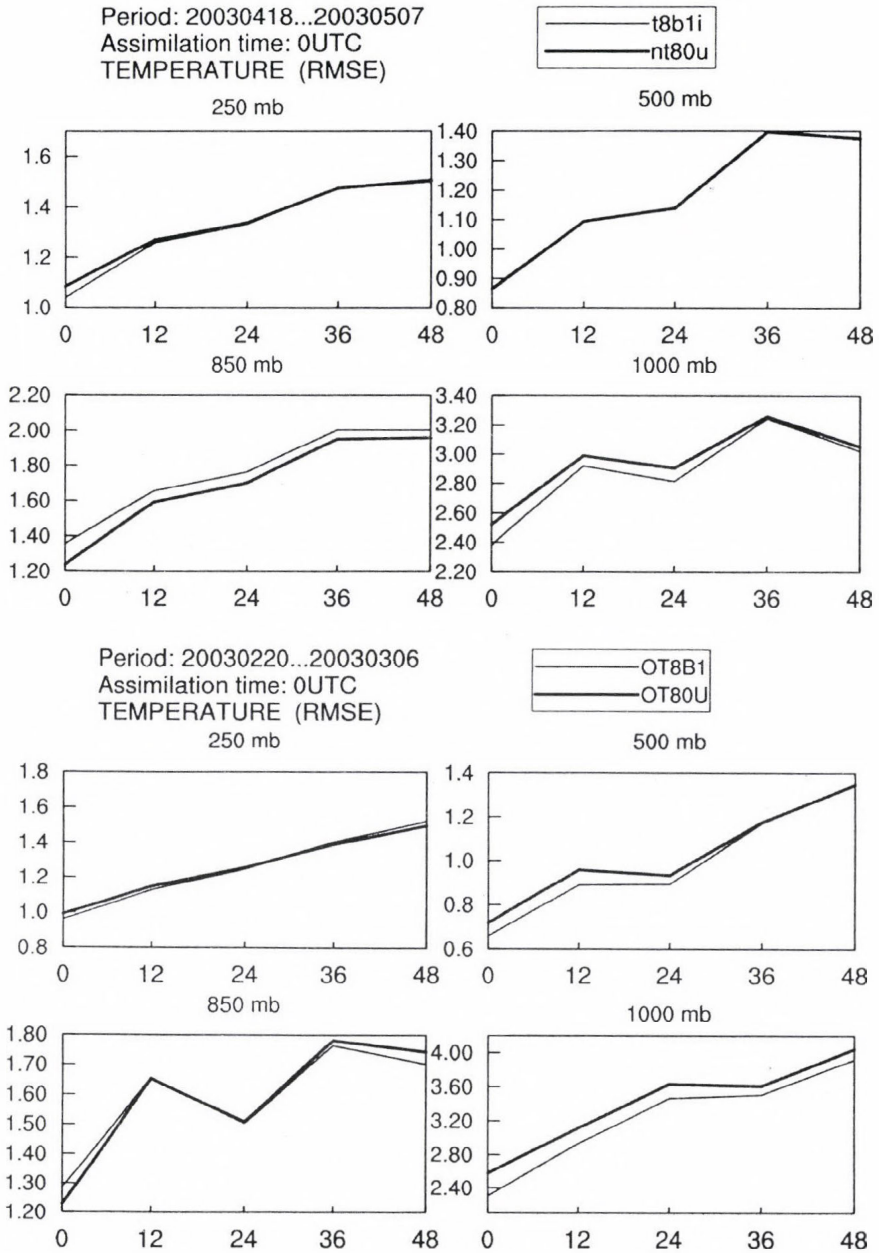

Fig. 5. Temperature root-mean-squares errors (RMSEs) for run with global bias correction coefficients (ARPEGE) (T8B1I and OT8B1I, for the first and the second period, respectively) against run with LAM coefficients (NT80U and OT80U, for the first and the second period, respectively).

4.3 Impact of no air-mass bias correction in the processing of AMSU-A

In order to assess the importance of air-mass bias correction, model runs with and without application of air-mass correction were compared. Thus, in the experiment T8B2I, no more than the interpolated ARPEGE scan-angle bias correction was used, since using a global model we can compute better representation of the scan-angle bias.

Without air-mass bias correction, satellite measurements warmed the model fields to a larger extent, which indicates that there was a residual bias in the temperature field shifted by satellite data (not shown). Accordingly, the verification scores showed a slightly negative or neutral impact on all the variables, including temperature forecast, in which the positive impact completely disappeared (*Fig. 6*). It seems likely that we need air-mass bias correction to assimilate radiances, since the ARPEGE scan-angle bias correction itself was not satisfactory.

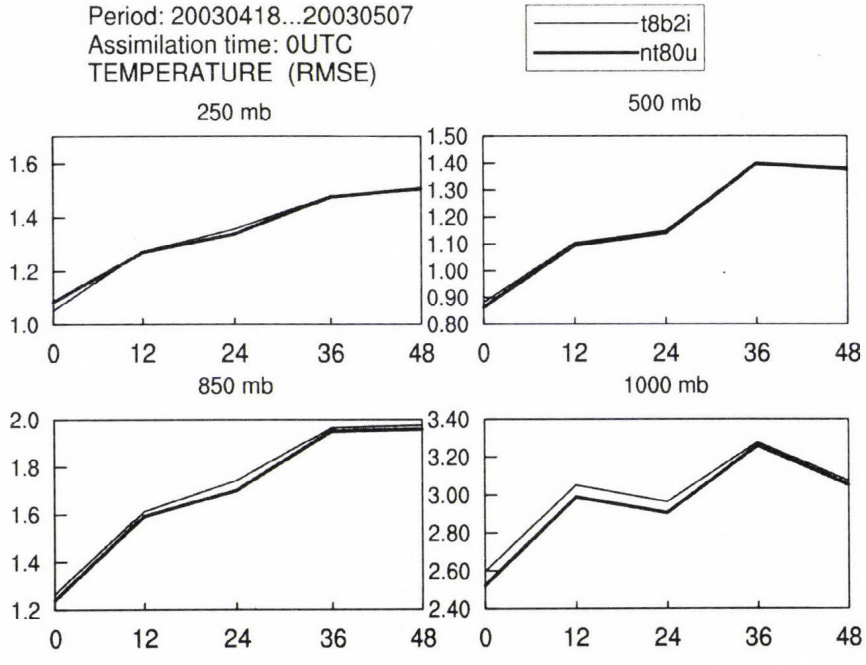

Fig. 6. Temperature root-mean-square errors (RMSEs) for run with global bias correction coefficients (ARPEGE) (T8B2I – no air-mass bias correction) against run with LAM coefficients (NT80U).

4.4 Combining the scan-angle bias correction of the global model with the air-mass bias coefficients of the LAM

Assuming that the air-mass bias correction is important, we combined the interpolated ARPEGE scan-angle bias correction with the ALADIN/HU air-mass bias correction in the experiment T8B3I. The combination of the global and local bias correction coefficients showed structurally similar results to those obtained in the experiment with ARPEGE bias correction file only (see *Fig. 5*), but both negative and positive impacts were negligible (*Fig. 7*). This reveals that using the global scan-angle bias correction with LAM air-mass bias correction coefficients did not improve the impact significantly.

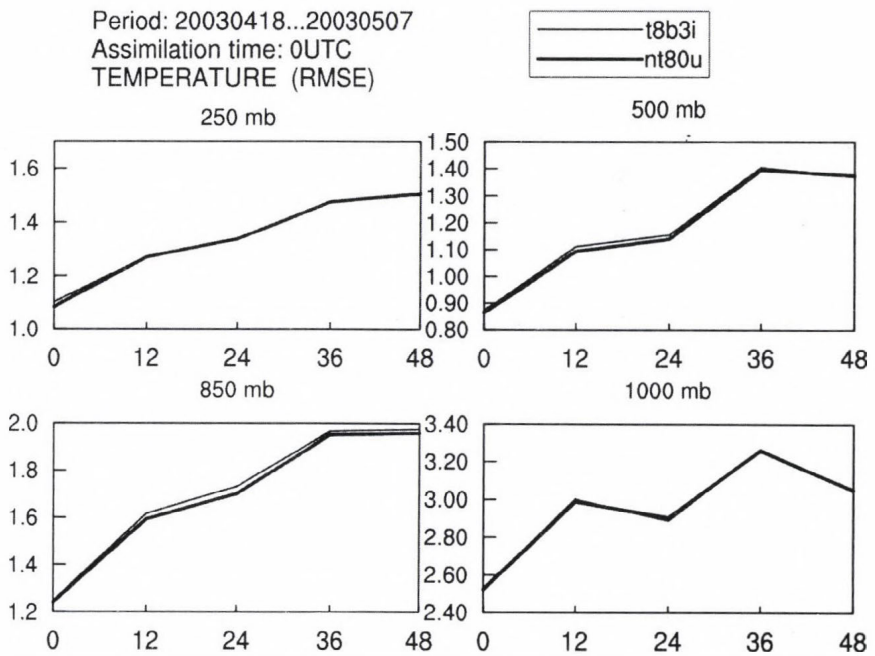

Fig. 7. Temperature root-mean-square errors (RMSEs) for run with global (ARPEGE) scan-angle bias correction coefficients and with LAM air-mass bias correction coefficients (T8B3I) against run with LAM bias correction coefficients (NT80U).

The sensitivity of channels 5, 6, 7, 10, 11, and 12 to the bias correction files was evaluated analyzing the number of assimilated satellite data (*Fig. 8*). More observation was available in the troposphere (channels 5, 6, and 7), while less data were used for channels 10, 11, and 12 when applying the

global air-mass bias coefficients in data processing. We assume, that the use of channels 5–7 was more efficient when applying the global bias coefficients compared to the local ones, probably because the analysis of the surface fields in the ARPEGE model is more accurate than that in the LAM.

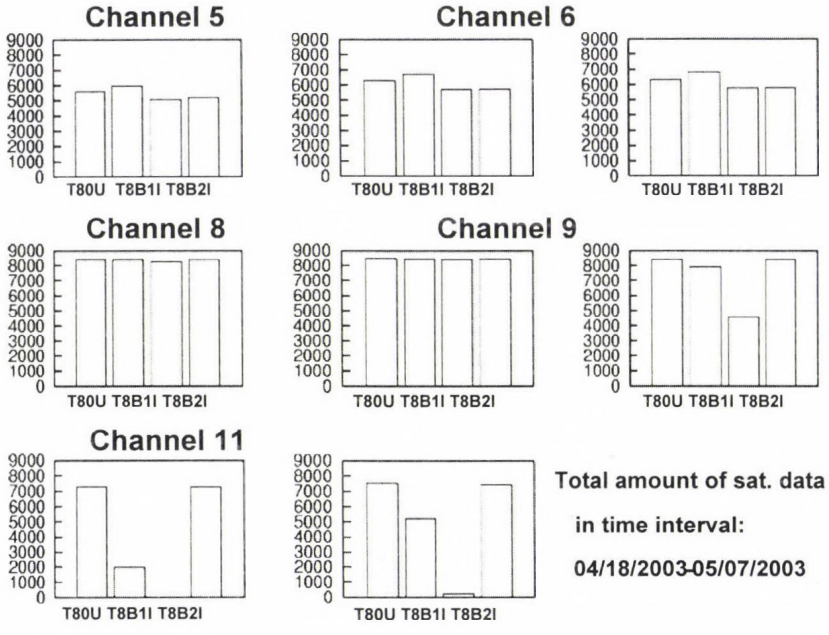

Fig. 8. Total number of assimilated satellite observations (active data) for the period April 18, 2003–May 07, 2003.

5. Conclusions

Our experiments show the importance of bias correction coefficients in the processing of AMSU-A data in the ALADIN/HU limited area model.

The use of the global bias correction file showed different impacts on short-range forecasts, especially in the lower troposphere, which is very important for synoptic meteorology. LAM bias correction coefficients provide a "stable" impact on the analysis as well as on the short-range forecasts.

Although the ARPEGE and the ALADIN models use basically the same parameterization of physical processes, and the bias correction coefficients are available from the global model, it is recommended to use bias correction, computed separately for the ALADIN model to ensure better processing of the AMSU-A data in the analysis system. It was found, that despite of smaller observation-

minus-first-guess samples, bias correction coefficients computed for the limited area are more suitable and reliable when assimilating radiances in a LAM.

It was proved, that the air-mass bias correction must be included in the processing of AMSU-A data in the limited area model.

It seems that the processing of the channels 10–12 in LAM is very sensitive to the bias coefficients computed for a global model.

At the Hungarian Meteorological Service, a 3D-Var system became operational this year (from May 17, 2005). In this system the pre-processing of the AMSU-A data uses bias correction coefficients computed locally according to the method, presented in this paper.

Acknowledgements—The research for this paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the Hungarian National Scientific Foundation (OTKA T049579). The helps from *Philippe Caille* and *Élisabeth Gérard* of the Météo France and the fruitful discussion with colleagues of the numerical prediction division, department research and development of the Hungarian Meteorological Service are highly acknowledged. The author recognizes and appreciates the participation of *Regina Szoták* in this investigation.

References

Dee, D.P., 2004: Variational bias correction of radiance data in the ECMWF system. *Proc. of the ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP*. 28 June – 1 July 2004, Reading, U.K.

Eyre, J.R., 1992: A bias correction scheme for simulated TOVS brightness temperatures. *ECMWF Technical Memorandum*, 176.

Harris, B.A. and *Kelly, G.*, 2001: A satellite radiance-bias correction scheme for data assimilation. *Q. J. Roy. Meteor. Soc.*, 1453-1468.

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

Parrish, D.F. and *Derber, J.C.*, 1992: The National Meteorological Centre's spectral statistical interpolation analysis system. *Mon. Weather Rev. 120*, 1747-1763.

Rabier, F., Gérard É., Sahlaoui Z., Dahoui M. and *Randriamampianina R.*, 2001: Use of ATOVS and SSMI observations at Météo-France. *11th Conference on Satellite Meteorology and Oceanography*. Madison, WI, 15-18 October 2001 (preprints). Boston, MA, American Meteorological Society. 367-370.

Radi, A. and *Issara, S.*, 1994: Introduction de l'analyse canari du modèle global arpège dans le modèle à domaine limité aladin. *Master's Thesis*. Ecole Nationale de la Météorologie, Toulouse.

Randriamampianina, R. and *Szoták, R.*, 2003: Impact of the ATOVS data on the Mesoscale ALADIN/HU Model. *ALADIN Newsletter*, N24, also at http://www.cnrm.meteo.fr/aladin/newsletters/newsletters.html.

Saunders, R., Matricardi, M., and *Brunel, P.*, 1998: An improved fast radiative transfer model for assimilation of satellite radiance observations. *Q. J. Roy. Meteor. Soc. 125*, 1407-1425.

Široká, M., Fischer, C., Cassé, V., Brožková, 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.

Watts, P.D. and *McNally, A.P.*, 2004: Identification and correction of radiative transfer modelling errors for atmospheric sounders: AIRS and AMSU-A. *Proc. of ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP*. 28 June – 1 July 2004, Reading, UK.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 109, No. 3, July–September 2005, pp. 157–172

Regulatory modeling in Hungary – the AERMOD model Part I. Description and application

Roland Steib and **Krisztina Labancz**

Hungarian Meteorological Service
P.O. Box 39, H-1675 Budapest, Hungary; E-mail: steib.r@met.hu

(Manuscript received in final form July 22, 2005)

Abstract—This paper presents the adaptation of a second-generation local-scale dispersion model, called AERMOD, at the Hungarian Meteorological Service (HMS). The aim of this work is to develop a method suitable for calculations of air pollutant concentrations for regulatory purposes. AERMOD was designed to introduce current planetary boundary layer concepts into regulatory dispersion models. In this paper the main features of the AERMOD modeling system are described. New developments, implemented during the adaptation work, are discussed in detail. Finally, our experiences running the model and modeling results are described. Test runs and the validation of the model will be described in Part II of this paper, where the sensitivity of the model to the meteorological and surface parameters will be examined, and with some case studies we would like to show how the stability of the atmosphere influences the distribution of this pollutants, calculated by the model.

Key-words: regulatory modeling, local-scale dispersion model, planetary boundary layer, bi-Gaussian distribution, elevated terrain, terrain height scale, dividing streamline height.

1. Introduction

In 1991, the American Meteorological Society (AMS) and the U.S. Environmental Protection Agency (EPA) initiated a formal collaboration with the goal of introducing current planetary boundary layer (PBL) concepts into regulatory dispersion models. A working group (AMS/EPA Regulatory Model Improvement Committee, AERMIC) comprising AMS and EPA scientists was formed for this collaborative effort.

EPA's regulatory platform for near-field modeling, during 25 years, has with few exceptions, remained fundamentally unchanged. During this period,

the Industrial Source Complex Model (ISC3) was the workhorse regulatory model. Therefore, AERMIC selected the EPA's ISC3 Model for a major overhaul. AERMIC's objective was to develop a complete replacement for ISC3 by: (1) adopting ISC3's input/output computer architecture; (2) updating, where practical, antiquated ISC3 model algorithms with newly developed or current state-of-the-art modeling techniques; and (3) insuring that the source and atmospheric processes presently modeled by ISC3 will continue to be handled by the AERMIC Model (AERMOD), albeit in an improved manner.

AERMOD underwent a comprehensive performance evaluation (*Brode et al.*, 1998) designed to assess how well AERMOD's concentration estimates compare against a variety of independent data bases, and to assess the adequacy of the model for use in regulatory decision making. AERMOD was evaluated against five independent data bases (two in simple terrain and three in complex terrain), each containing one full year of continuous SO_2 measurements. Additionally, AERMOD's performance was compared against the performance of four other applied regulatory models: ISC3 (*U.S. Environmental Protection Agency*, 1995), CTDMPLUS (*Perry*, 1992), RTDM (*Paine* and *Egan*, 1987), and HPDM (*Hanna* and *Paine*, 1989; *Hanna* and *Chang*, 1993). The performance of these models against AERMOD has been compared using the procedures in EPA's "Protocol for Determining the Best Performing Model" (*U.S. Environmental Protection Agency,* 1992). In April 21, 2000, EPA proposed that AERMOD be adopted as a replacement to ISC3 in Appendix A of the Guideline on Air Quality Models (*Code of Federal Regulations*, 1997). As such, upon final action, AERMOD would become EPA's preferred regulatory model for both simple and complex terrain. The newest performance of AERMOD is documented in *Perry et al.* (2003).

In 2002, the Hungarian Meteorological Service decided to replace its old dispersion model, a very simple Gaussian model by AERMOD. The new model, working with modern concepts of boundary layer and treating complex terrain, is a much more effective tool in regulatory modeling calculations.

2. The AERMOD modeling system

The AERMOD modeling system consists of three pre-processors and the dispersion model. The AERMIC meteorological preprocessor (AERMET) provides AERMOD with the meteorological information it needs to characterize the planetary boundary layer (PBL). The AERMIC terrain pre-processor (AERMAP) both characterizes the terrain, and generates receptor grids for the dispersion model (AERMOD). INTERFACE, which is inside AERMOD is a preprocessor, that makes connection between AERMET and

AERMOD. *Fig. 1* shows the flow and processing of information in AERMOD, where NWS means National Weather Service, ONSITE means on-site measurement, PASS OBS means passing observation, PBL PARAM means planetary boundary layer parameters, u is wind-vector, *turb* is turbulence, dT/dz is temperature gradient, X, Y, Z are coordinates of the receptor points, and h_c is the effective height scale.

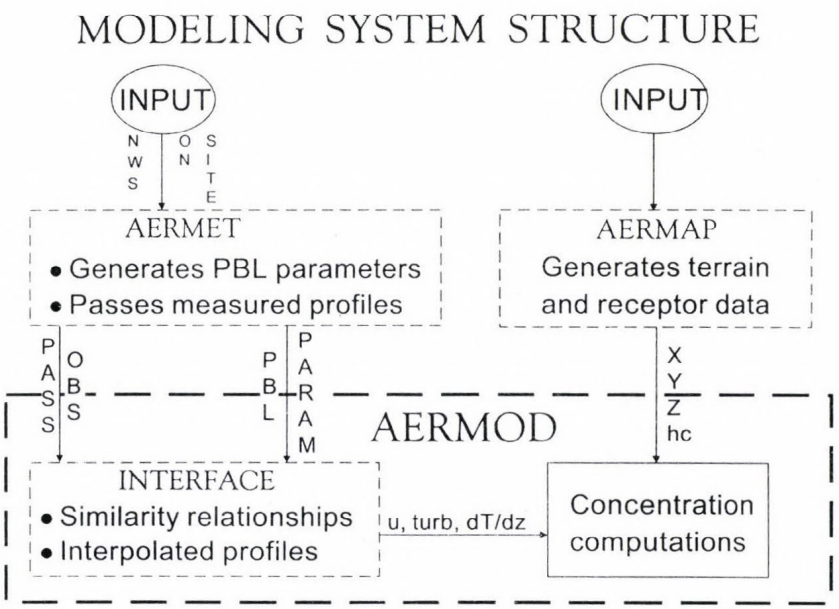

Fig. 1. Data flow in the AERMOD modeling system.

2.1 AERMOD – the dispersion model

AERMOD is a steady-state plume model. In the stable boundary layer (SBL), it assumes the concentration distribution to be Gaussian in both the vertical and horizontal. In the convective boundary layer (CBL), the horizontal distribution is also assumed to be Gaussian, but the vertical distribution is described with a bi-Gaussian probability density function (pdf). This behavior of the concentration distributions in the CBL was demonstrated by *Willis* and *Deardorff* (1981) and *Briggs* (1993).

Using a relatively simple approach, AERMOD incorporates current concepts about flow and dispersion in complex terrain. Where appropriate, the plume is modeled as either impacting and/or following the terrain. This

approach has been designed to be physically realistic and simple to implement while avoiding the need to distinguish among simple, intermediate, and complex terrain, as required by other regulatory models. As a result, AERMOD removes the need for defining complex terrain regimes. All terrain is handled in a consistent and continuous manner while considering the dividing streamline concept (*Snyder et al.*, 1985) in stably stratified conditions.

One of the major improvements that AERMOD brings to applied dispersion modeling is its ability to characterize the PBL through both surface and mixed layer scaling. AERMOD constructs vertical profiles of required meteorological variables based on measurements and extrapolations of those measurements using similarity (scaling) relationships. Vertical profiles of wind speed, wind direction, turbulence, temperature, and temperature gradient are estimated using all available meteorological observations. AERMOD is designed to run with a minimum of observed meteorological parameters. As a replacement for the ISC3 model, AERMOD requires only a single surface measurement of wind speed (measured in the layer between 7 times of the surface roughness height and 100 meters above the surface), wind direction, and ambient temperature. Like ISC3, AERMOD also needs observed cloud cover. However, if cloud cover is not available (e.g., from an on-site monitoring program), two vertical measurements of temperature (typically at 2 and 10 meters), and a measurement of solar radiation can be substituted. A full morning upper air sounding (RAWINSONDE) is required in order to calculate the convective mixing height throughout the day. Surface characteristics (surface roughness, Bowen ratio, and albedo) are also needed in order to construct similarity profiles of the relevant PBL parameters.

Unlike existing regulatory models, AERMOD accounts for the vertical inhomogeneity of the PBL in its dispersion calculations. This is accomplished by "averaging" the parameters of the actual PBL into "effective" parameters of an equivalent homogeneous PBL.

2.2 AERMET – the meteorological preprocessor

AERMOD requires a preprocessor that organizes and processes meteorological data and estimates the necessary boundary layer parameters for dispersion calculations. The meteorological preprocessor that serves this purpose is AERMET.

AERMET is designed to be run as a three-stage process (*Fig. 2*), where *QA* means quality assessment, *stage1n2* and *stage3* are the executable programs. AERMET operates on three types of data – hourly surface observations, daily upper air soundings, and data collected from an on-site measurement program such as an instrumented tower. The first stage extracts

data and assesses data quality. The second stage combines the available data for 24-hour periods and writes these data to an intermediate file. The third and final stage reads the merged data file and develops the necessary boundary layer parameters for dispersion calculations by AERMOD.

AERMET can extract data from several standard NCDC formats. These include the TD-6201 format for upper air sounding data, hourly surface weather observations in the CD-144 format, which is a time-based (i.e., by hour) format, and the TD-3280 format, which is an element-based (i.e., by variable) format for surface data.

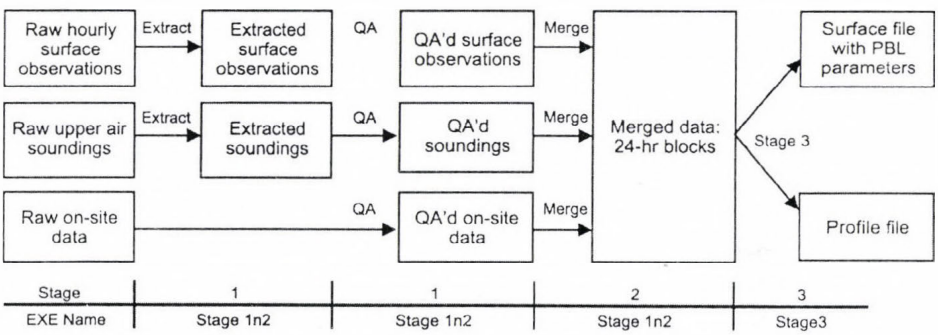

Fig. 2. The 3 stages of the AERMET processing.

The temperature structure of the atmosphere prior to sunrise is required by AERMET to estimate the growth of the convective boundary layer for the day. Currently, AERMET uses the 12:00 UTC upper air sounding for this purpose. A one-hour window of time in the search allows for the possibility of an early or late launch time. This search restricts the applicability of AERMET to those longitudes where the 12:00 UTC sounding corresponds to the early morning hours, i.e., the Western Hemisphere.

Surface characteristics in the form of albedo, surface roughness, and Bowen ratio, plus standard meteorological observations (wind speed, wind direction, temperature, and cloud cover), are input to AERMET. AERMET then calculates the PBL parameters: friction velocity (u^*), Monin-Obukhov length (L), convective velocity scale (w^*), temperature scale (θ^*), mixing height (z_i), and surface heat flux (H). These parameters are then passed to the INTERFACE (which is within AERMOD), where similarity expressions (in conjunction with measurements) are used to calculate vertical profiles of wind speed (u), lateral and vertical turbulent fluctuations (σ_V, σ_W), potential temperature gradient ($\mathrm{d}\theta/\mathrm{d}z$), and potential temperature (θ).

2.3 AERMAP – the terrain preprocessor

The AERMAP terrain pre-processor uses gridded terrain data to calculate a representative terrain-influence height (h_c), also referred to as the terrain height scale. The terrain height scale h_c, which is uniquely defined for each receptor location, is used to calculate the dividing streamline height. The gridded data needed by AERMAP is selected from Digital Elevation Model (DEM) data. AERMAP is also used to create receptor grids. The elevation for each specified receptor is automatically assigned through AERMAP. For each receptor, AERMAP passes the following information to AERMOD: the receptor's location (x_r, y_r), its height above mean sea level (z_r), and the receptor specific terrain height scale (h_c).

2.4 Output options

The AERMOD model provides various output options for regulatory purposes: summaries of high values (highest, second highest, etc.) by receptor for each averaging period (1-hour, 3-hour, 8-hour, 24-hour, monthly, yearly, average for the entire data period) and source group combination; summaries of overall maximum values for each averaging period and source group combination; tables of concurrent values summarized by receptor for each day of data processed.

Another output option of the AERMOD model is to give all occurrences when a concentration value equals or exceeds a user-specified threshold. Separate results are generated for only those combinations of averaging period and source group that are of interest.

In air quality dispersion modeling applications, there is a need to know the contribution that a particular source makes to an overall concentration value for a group of sources. The AERMOD model provides the option to perform the source attribution analyses. For example, the user can specify source groups for which the model calculates high values independently. In addition, when maximum values for individual sources are output, the user has the option of specifying whether the maximum source values are to be the maximum values for each source independently, or the contribution of each source to the maximum group values, or both.

3. Adaptation of the AERMOD system at the HMS

At the Hungarian Meteorological Service we do not have the data file formats, that AERMET and AERMAP can use as inputs. We had to create a program

that gives us the meteorological parameters in the extracted file formats for AERMET. So when we run stage1 of AERMET, we only have to run the second part (quality assessment) of this stage. Instead of AERMAP we also created our own terrain pre-processor to get the needed terrain parameters for AERMOD.

We had two other, major problems to solve. The first was that AERMET needs cloud cover to calculate planetary boundary layer parameters, but there are few stations with visual observations in Hungary. The second problem was that AERMET needs an early morning upper air sounding in order to calculate the convective boundary layer parameters. In Hungary we only have 00:20 LST (Budapest and Szeged) and 12:20 LST (Budapest) upper air soundings, which cannot be used in AERMET.

We thought that the best way to solve these problems is to use two- and three-dimensional analysis fields of the necessary meteorological parameters. The analysis fields are calculated by the MM5 model, a numerical weather prediction model developed especially for nowcasting purposes, which operationally runs at the Hungarian Meteorological Service. In these data files the meteorological parameters are represented in grid points. Cloud cover (SatCloudi) is calculated from MSG satellite data. Two and three-dimensional MM5 analysis data are available for every hour. We thought that the 05:00 UTC three-dimensional MM5 analysis data could represent well the early-morning vertical structure of the planetary boundary layer. In consequence, we had to change an original subroutine in AERMET (Mppbl.for). This original subroutine could read only the 11:00, 12:00, and 13:00 UTC upper air data. We changed the source code, so the new one is able to read the 05:00 UTC upper air data.

3.1 Structure of meteorological data

We construct the meteorological data files from the two- and three-dimensional MM5 analysis netCDF files. These netCDF files can be found on the *lamb* server of the Hungarian Meteorological Service. The resolution of the two dimensional files is $1/60 \times 1/50$ degrees, and that of the three dimensional files is 0.06×0.05 degrees. From the two-dimensional netCDF files we read the following parameters and store them in ASCII files: u-component of wind at 10 meters, v-component of wind at 10 meters, temperature at 2 meters, relative humidity at 2 meters, surface pressure, and cloud cover. From the three-dimensional netCDF files we read the following parameters (at the 1000, 950, 925, 900, 850, 800, 750, 700, 650, 600, 550, and 500 hPa isobaric levels) and store them in ASCII files: u-component of wind, v-component of wind, temperature, relative humidity, and geopotential height.

3.2 Constructing the meteorological inputs for AERMET

In our opinion, using the meteorological parameters from one grid point (nearest to the pollution source) of the MM5 analysis data files is not the best way to represent the meteorological situation in a bigger area, so we use meteorological parameters averaged over a grid, approximately 30 × 30 km. We have already seen that we have to construct the meteorological data in exactly that file format that the AERMET meteorological pre-processor gives after running stage1. We do this work in 3 steps. The first step is a small program which calculates the nearest grid point to the pollution source in the meteorological data with the help of the source coordinates. A 2D meteorological grid file consists of about 440 zonal points and 178 meridional points. A 3D grid file contains about 150 zonal points and 60 meridional points. The second step is a shell-script in which an awk-script is embedded. The shell-script does the following work: copies data files in the working directory, uncompresses them, runs the awk-script and writes the results in intermediate meteorological files. The awk-script reads the meteorological data from a 30 × 30 km domain around the source of the 2D and 3D files, and calculates the averages over this area. The third part is a program which reads the meteorological data from the intermediate files and converts them to the file format that AERMET can read.

3.3 Constructing the terrain inputs

The terrain preprocessor of AERMOD uses gridded terrain data to calculate the elevation and representative terrain-influence height (referred to as terrain height scale) for each specified receptor point. The gridded data originally was selected from the U.S. Digital Elevation Model (DEM) database. Because DEM does not contain terrain data for Europe, we used similar elevation data files consisting of a regular array of elevations referenced horizontally on the UTM longitude/latitude coordinate system calculated on the World Geodetic System (WGS-84) ellipsoid. Therefore, we had to write a new terrain preprocessor.

First, we have to define a modeling domain. A modeling domain is defined as the area that consists of all the receptors and sources being modeled. For regulatory purposes, we use a 30 × 30 km domain around the source point, where the minimal receptor point distance could be 500 m. Then we define the receptor point(s). The receptors are specified in a manner identical to the AERMOD dispersion model. One can specify discrete receptors as well as Cartesian and polar grid networks. However, there are some special considerations, e.g., when specifying discrete polar receptors, it

is necessary to specify the position of a source relative to which the receptor is assigned.

In addition to the location and height above mean sea level of each receptor, the AERMOD model requires the receptors' specific terrain height scale, which is used to calculate the critical dividing streamline height. Our terrain preprocessor contains a new calculation method for determining the height scale. First, the method calculates the maximum elevation difference in the domain, then the relief height of the terrain points with respect to the height of the receptor, and the slope are determined. If the slope is 10% or greater, the height scale for the receptor is updated.

3.4 Post-processing

The post-processing part of the AERMOD modeling system contains two programs. The first program constructs a tabular overview of the concentrations in the receptor points. First the period average concentrations, then the 24-hour maximum concentrations, and at last the 1-hour maximum concentrations are constructed. This program also defines the "efficient" area. The "efficient" area is a domain around the pollutant source where the concentration is higher than the 80% of the maximum 1-hour concentration or the concentration is higher than the 10% of the maximum 1-hour threshold value of a pollutant. The second program constructs the graphical overview of the period average, the 24-hour maximum and the 1-hour maximum concentration distribution. This step is done with the DISLIN graphical program, which produces three pictures in PNG file format.

4. Results

In this paper we present the results of two different test runs to illustrate the benefits of adaptation of the AERMOD modeling system at the Hungarian Meteorological Service. In the first test we made a comparison between AERMOD and the formal dispersion model, called TRANSMISSION 1.0 (*Szepesi et al.*, 1995), which is a very simple Gaussian model. Since TRANSMISSION can not treat complex terrain, in the first test we made calculations over flat terrain to compare the basic characteristics of the two models. In the second test we examined the effects of complex terrain on the results of AERMOD. In all our tests we used the following parameters:

- one point source (stack),
- location of the source: 46.06°N, 18.26°E (south-western part of Hungary),

- type of pollutant: NO_x,
- receptor resolution: ~ 500 m,
- number of receptors: 59 × 59,
- domain size: ~ 30 × 30 km,
- emission rate: 28 g/s,
- release height above ground: 82 m,
- stack gas exit temperature: 373 K,
- stack gas exit velocity: 13 m/s,
- stack inside diameter: 6 m.

4.1 Test 1. Dispersion distribution over flat terrain

In our first test we wanted to examine AERMOD's behavior in case of flat terrain, because our former dispersion model could only treat flat terrain. We wanted to see the differences or similarities in the concentration distributions between the two models. The yearly average concentration distribution of the two models showed similarities. *Fig. 3* shows the yearly average NO_x concentration field calculated by the TRANSMISSION 1.0 model, and *Fig. 4* shows the yearly average NO_x concentration field calculated by the AERMOD model over the above characterized domain and source parameters. In all graphics the cross-shaped symbol represents the place of the source. When long averaging period (1 year) was used, we found that the concentration field was mostly governed by the average wind direction of the domain. In this area the wind usually blows from north-west, so we could find the highest concentration values in both cases to the south-east direction of the source. The difference between the two models was that the concentration field calculated by AERMOD spreaded over a smaller area around the source giving higher yearly average concentration values than TRANSMISSION. When we used short averaging period (1 hour), the results between the two models were much more significant. *Fig. 5* shows the maximum 1-hour NO_x concentration field of TRANSMISSION and *Fig. 6* the max 1-hour NO_x concentration field of AERMOD. The concentration distribution in *Fig 5* is totally concentric, while *Fig. 6* also shows some concentricity, but the concentration distribution is much more irregular. This characteristic of the concentration field can be explained by that over a short averaging period the direction of the local wind is fluctuating, and any direction of the wind can result in high concentration values. In our opinion the differences between the two models in the case of short averaging periods is not surprising. Probably it comes from the different structures of the two models. The simple Gaussian model calculates the planetary boundary layer parameters empirically, while AERMOD calculates unique

parameters of the planetary boundary layer for every hour, so the boundary layer changes continuously during the day. In case of the 1-hour averaging period both models gave almost the same maximum concentration values. The AERMOD model usually gave high concentration values in calm wind situations.

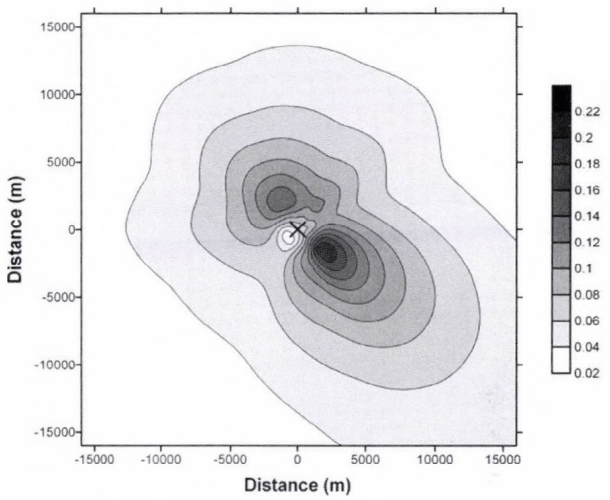

Fig. 3. Yearly average NO_x concentration of TRANSMISSION (flat terrain). Units: $\mu g/m^3$.

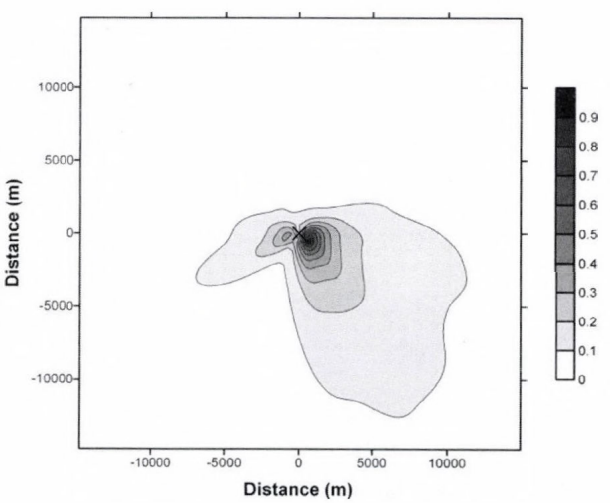

Fig. 4. Yearly average NO_x concentration of AERMOD (flat terrain). Units: $\mu g/m^3$.

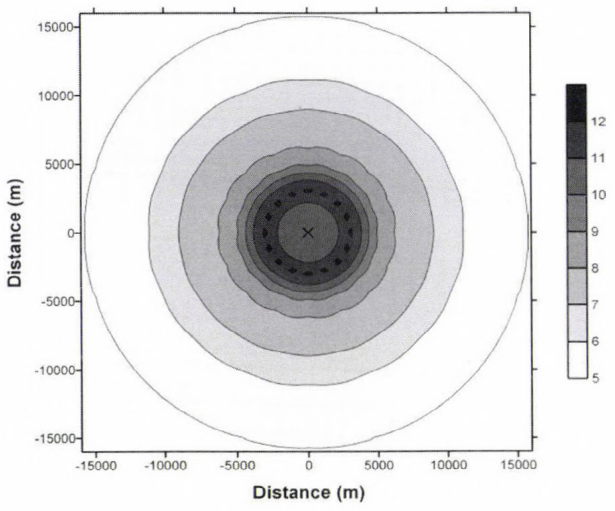

Fig. 5. 1-hour maximum NO_x concentration of TRANSMISSION (flat terrain). Units: $\mu g/m^3$.

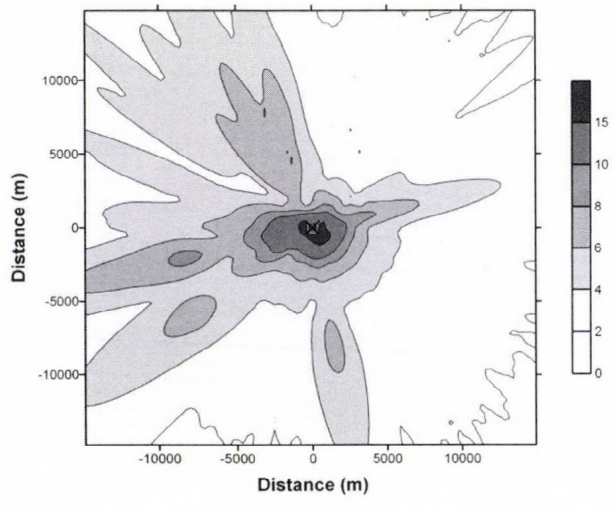

Fig. 6. 1-hour maximum NO_x concentration of AERMOD (flat terrain). Units: μg / m^3.

4.2 Test 2. Dispersion distribution over complex terrain

In this test we wanted to see the changes in the concentration field when the AERMOD model used elevated terrain. By short and also by long averaging periods we experienced big differences in the concentration fields between

calculations with elevated terrain and flat terrain. *Fig. 7* illustrates concentration distribution of AERMOD in the case of the yearly averaging period, while *Fig. 8* shows the same but with 1-hour averaging period. By long time averaging period we can notice two maximum places of the concentration distribution. The first maximum is at the same place as it was at the flat terrain. But also a second maximum can be seen at the place where the terrain begins to elevate at the feet of the mountains. This second maximum can be found near the centreline height of the plume. We can also see that the plume flows around the mountains, and at the top of the mountains (500–600 m) the concentration values are lower again.

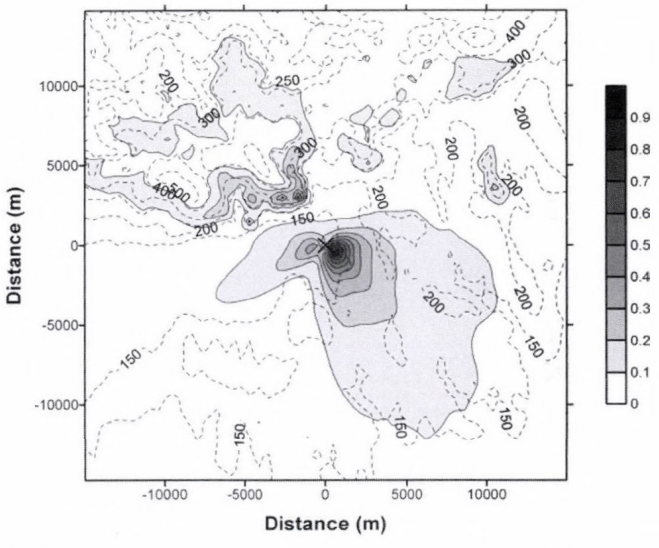

Fig. 7. Yearly average NO_x concentration of AERMOD (elevated terrain). Units: $\mu g/m^3$. Dashed isolines represent the terrain height. Contour intervals are at every 50 m.

By the 1-hour averaging period, the highest concentration values can be found at the feet of the nearest mountain to the source, and also at the height of the plume centerline. Thus, when short averaging period was used, the elevated terrain had even more influence on the concentration distribution than by the yearly averaging period.

We also have to notice that the maximum value of the yearly average concentration of AERMOD is almost the same with flat- and elevated terrain, but the maximum value of the 1-hour concentration is about 10 times higher

with elevated terrain than with flat terrain. This result makes us clear that the elevated terrain can very much change the concentration distribution of a pollutant.

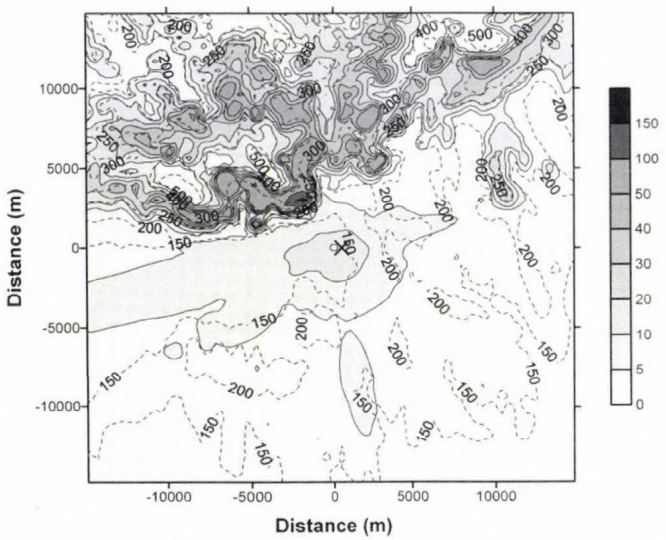

Fig. 8. 1-hour maximum NO_x concentration of AERMOD (elevated terrain). Units: $\mu g/m^3$. Dashed isolines represent the terrain height. Contour intervals are at every 50 m.

We have already mentioned in the abstract that in Part II of this paper we want to examine the model's sensitivity to meteorological and surface parameters. We also want to examine how the stability of the atmosphere influences the concentration distribution. We would like to show the differences in the concentrations between the stable and convective boundary layer. We also would like to examine in which way the height of the stable and convective boundary layer affects the concentration field.

5. Conclusions and future plans

In this paper a second-generation local-scale dispersion model was presented. The model, developed at the Environmental Protection Agency, U.S. for regulatory purposes, works with modern concepts of boundary layer and treats

complex terrain. The AERMOD model was adapted and tested at the Hungarian Meteorological Service, it was comparised with the old Gaussian model TRANSMISSION 1.0. One of the advantages of AERMOD is that it can treat complex terrain giving much more realistic figures of the distribution of pollutants. In addition to this, AERMOD gave better concentration distributions over flat terrain for the 1-hour averaging period. The difference was remarkable in the case of short averaging periods caused by the different structure of the two models. The simple Gaussian model calculates the planetary boundary layer parameters empirically, while AERMOD calculates unique parameters of the planetary boundary layer for every hour, so the boundary layer changes continuously during the day.

So far we could not afford to evaluate the performance of AERMOD by using measured data. In the U.S., AERMOD underwent a comprehensive performance evaluation against five independent data bases (two in simple terrain and three in complex terrain), each containing one full year of continuous SO_2 measurements. On the basis of our co-operation with Hungarian environmental agencies, we hope to evaluate our model version against measured data bases in the future.

The Hungarian Meteorological Service co-operates with local Hungarian environmental agencies in air pollution regulatory tasks. Our dispersion models are run for local, special environmental problems, for which the AERMOD model system has been proven to be a very efficient tool. This model version of AERMOD does not take into account dry and wet deposition. This can be the reason why the model usually gives slightly higher concentration values than the measurements or other dispersion models. The next version of the model is under construction. The new version of the model will be able to handle dry and wet deposition of gases and particles, so we can hope that we will be able to make even better concentration estimations.

References

Briggs, G.A., 1993: Plume dispersion in the convective boundary layer. Part II: Analysis of CONDORS field experiment data. *J. Appl. Meteorol. 32*, 1388-1425.

Brode, R.W., Perry, S.G., Cimorelli, A.J., Weil, J.C., Venkatram, A., Paine, R.J., Wilson, R.B., Lee, R.F., and *Peters, W.D.* 1998: Model Evaluation Results of AERMOD. EPA web site (http://www.epa.gov/scram001/tt26.htm#aermod).

Code for Federal Regulations, 1997: Service of the U.S. Government Printing Office, U.S. GPO web site (http://www.gpoaccess.gov/cfr).

Hanna, S.R. and *Paine, R.J.*, 1989: Hybrid Plume Dispersion Model (HPDM) development and evaluation. *J. Appl. Meteorol. 28*, 206-224.

Hanna, S.R. and *Chang, J.S.*, 1993: Hybrid Plume Dispersion Model (HPDM), improvements and testing at three field sites. *Atmos. Environ. 27A*, 1491-1508.

Paine, R.J. and *Egan, B.A.*, 1987: User's Guide to the Rough Terrain Diffusion Model (RTDM) - Rev. 3.20. ERT Document PD-535-585, ENSR, Acton, MA.

Perry, S.G., 1992: CTDMPLUS: A dispersion model for sources in complex topography. Part I: Technical formulations. *J. Appl. Meteorol. 31*, 633-645.

Perry, S.G., Cimorelli, A.J., Weil, J.C., Venkatram, A., Paine, R.J., Wilson, R.B., Lee, R.F., and *Peters, W.D.*, 2003: AERMOD: Latest Features and Evaluation Results. EPA web site (http://www.epa.gov/scram001/tt26.htm#aermod).

Snyder, W.H., Thompson, R.S., Eskridge, R.E., Lawson, R.E., Castro, I.P., Lee, J.T., Hunt, J.C.R., and *Ogawa, Y.*, 1985: The structure of the strongly stratified flow over hills: Dividing streamline concept. *J. Fluid Mech. 152*, 249-288.

Szepesi, D.J., Fekete, K.E., Gyenes, L., 1995: Regulatory Models for Environmental Impact Assessment in Hungary. *Int. J. Environ. and Pollution 5*, 497-507.

U.S. Environmental Protection Agency, 1992, Protocol for Determining the Best Performing Model. EPA-454/R-92-025, U.S. Environmental Protection Agency, RTP, NC.

U.S. Environmental Protection Agency, 1995, User's Guide for the Industrial Source Complex (ISC3) Dispersion Models (revised) Volume I - User Instructions. EPA-454/b-95-003a, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Willis, G.E. and *Deardorff, J.W.*, 1981: A laboratory study of dispersion in the middle of the convectively mixed layer. *Atmos. Environ. 15*, 109-117.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 109, No. 3, July–September 2005, pp. 173–188

Analysis of the splitting error for advection-reaction problems in air pollution models

Tamás Ladics

Department of Meteorology, Eötvös Loránd University
P.O. Box 32, H-1518 Budapest, Hungary
and
Department of Analysis, Budapest University of Technology and Economics
P.O. Box 91, H-1591 Budapest, Hungary
E-mail: tladics@math.bme.hu

(Manuscript received in final form October 5, 2004)

Abstract—Operator splitting is a widely used method in the numerical solution of air pollution transport equations. This paper presents an analysis of splitting error for advection-reaction problems. Our investigations are focused on Molenkampf–Crowley advection and chemical reactions. A necessary and sufficient condition is derived under which the splitting error vanishes. Some numerical experiments are presented for two-dimensional models.

Key-words: Molenkampf–Crowley advection, chemical reactions, numerical solution of partial differential equations, splitting error, operator splitting, air pollution modeling.

1. Introduction

Air pollution has become one of the main environmental problems due to increasing human activity in the fields of transport and industry (*Havasi et al.*, 2001). Reducing the amount of air pollutants emitted into the atmosphere is an important task. The efficient treatment of this problem requires the study of the mathematical models of air pollution transport. In the description of this process, *advection-diffusion-reaction equations* are widely used. Models describing the evolution of pollutant concentrations are based on the mass conservation law and represent different processes acting in the atmosphere (namely advection, diffusion, emission, chemical reactions, and deposition) in the form of a system of partial differential equations. To find the symbolic

solution of these equations is practically impossible, therefore, we use a numerical method to obtain an approximate solution of the equations. However, even the numerical treatment of the problem is complicated. To get a sufficiently precise approximation in reasonable time, we apply the *splitting method* (see, e.g., *Zlatev*, 1995). Operator splitting is a tool to make the numerical treatment simpler and to use our numerical methods more efficiently.

The technique of splitting to help us solve this system of equations was used for a long time in different ways. A kind of special splitting was introduced by *Marchuk* (1968) and *Strang* (1968), which has higher order accuracy. A splitting method was used in the solution of the diffusion equation by *McRae et al.* (1984). For advection-diffusion-reaction problems *Lanser* and *Verwer* (1999) analyzed the error caused by splitting. A study on splitting techniques applicable for problems with stiff operators can be found in *Sportisse* (2000).

The theoretical background of operator splitting and a detailed study of the splitting error in the case of air pollution models have been presented by *Havasi et al.* (2001). Here we strongly rely on their results especially on the condition for the splitting error to be zero.

In this paper we focus on the advection-reaction problem. We apply the result of *Havasi et al.* (2001) on L-commutation, and we determine the condition of L-commutativity in the case of advection and reaction operators. Furthermore, we present the results of some numerical computation for advection-reaction models. Section 2 introduces the transport equation used in air pollution modeling. Section 3 contains a review of the splitting method and a definition of the splitting error. In Section 4 we give a characterization of the splitting error and derive the error for the advection-reaction problem in general. In Section 5 we derive a necessary and sufficient condition for a two-dimensional advection-reaction model with Molenkampf–Crowley advection under which the operators (advection and reaction) commute, i.e., the splitting error is zero. In Section 6 we present our numerical results for this advection-reaction model. Section 7 discusses results and further plans.

2. The air pollution transport equation

The system of partial differential equations to describe advection, diffusion, and reaction in air pollution modeling is as follows:

$$\begin{aligned}\frac{\partial c_i(\mathbf{x},t)}{\partial t} &= -\nabla\cdot\left(\mathbf{u}(\mathbf{x},t)c_i(\mathbf{x},t)\right)+\nabla\cdot\left(\mathbf{K}_i(\mathbf{x},t)\nabla c_i(\mathbf{x},t)\right)+r_i\left(\mathbf{x},t,\mathbf{c}(\mathbf{x},t)\right)\\ &\quad +E_i\left(\mathbf{x},t\right)+\sigma_i c_i(\mathbf{x},t), \qquad i=1,2,\ldots,n;\end{aligned} \tag{1}$$

see, e.g., *Zlatev* (1995). Here

- $\mathbf{c} = (c_1,\ldots,c_n)$ is a vector-valued function and contains the concentration c_i of each species as a function of space ($\mathbf{x}$) and time (t);
- the function $\mathbf{u}$ denotes the velocity of flow depending on space and time;
- $\mathbf{K}_i$ is the diffusion coefficient matrix of pollutant i, it also depends on space and time;
- $r_i(\mathbf{x},t,\mathbf{c})$ represents the formation rate of pollutant i as a result of chemical reactions at place $\mathbf{x}$ and time t. The function r_i is usually nonlinear and may depend on the concentration of all species. (This dependence is polynomial, if *mass action type kinetics* is assumed, which is the most frequent case, or r_i are rational functions, e.g., in the cases when enzymatic or saturation processes are present, or if we are dealing with binding kinetics.) Therefore, r_i links the equations belonging to the different pollutants; further, this term leads to the greatest difficulties in handling the model equations;
- $E_i(\mathbf{x},t)$ represents the emission of pollutant i;
- $\sigma_i \leq 0$ is the sedimentation coefficient.

3. The splitting method

For simplicity we introduce the splitting procedure on a finite dimensional problem, i.e., on the example of a system of ordinary differential equations. Let us consider the initial value problem

$$\frac{\mathrm{d}\mathbf{c}(t)}{\mathrm{d}t} = A(\mathbf{c})(\mathrm{t}) \quad \mathbf{c}(0) = \mathbf{c}_0, \tag{2}$$

where $\mathbf{c}$ is the unknown vector function and A can be written as a sum of two simpler operators A_1 and A_2 : $A = A_1 + A_2$. Now let us take a short time τ, and let us split the problem into the following simpler problems:

$$\frac{\mathrm{d}\mathbf{c}^{(1)}(t)}{\mathrm{d}t} = A_1(\mathbf{c}^{(1)})(t) \quad \mathbf{c}^{(1)}(0) = \mathbf{c}_0, \tag{3}$$

and

$$\frac{\mathrm{d}\mathbf{c}^{(2)}(t)}{\mathrm{d}t} = A_2(\mathbf{c}^{(2)})(t) \quad \mathbf{c}^{(2)}(0) = \mathbf{c}^{(1)}(\tau). \tag{4}$$

The solution of the splitted problem at time τ is $\mathbf{c}_{\mathrm{sp}}(\tau) = \mathbf{c}^{(2)}(\tau)$.

The basic idea of splitting is to decompose the operator on the right hand side into the sum of simpler operators, and to solve the subproblems corresponding to the operators successively in each time step. More accurately, we solve the equation only with operator A_1 until time τ (as if only the subprocess represented by A_1 were present), and the solution in time τ will be the initial condition of the equation with A_2. It means that we return to the initial time and solve the equation with A_2, as well. The solution of the second equation in time τ will be the approximate solution of the original problem in time τ. This procedure is then repeated on the interval, $[\tau,2\tau]$, etc. Thus, the simpler subproblems are connected to each other through the initial conditions.

It is clear, that the numerical treatment of the separate subproblems is simpler. The most significant advantage of splitting is that we can exploit the special properties of different subproblems and apply the most suitable numerical method for each of them. Thus, we can obtain a more precise solution in a shorter time.

The splitting error at time τ can be defined as a norm of the difference between the solutions of the exact and splitted problems at time τ:

$$\mathrm{Err}_{sp}(\tau) = \| \mathbf{c}(\tau) - \mathbf{c}_{sp}(\tau) \| . \tag{5}$$

An operator can obviously be splitted into more than two summands, and this way the method can be extended. If we apply the method to our more complicated Eq. (1), then a possible decomposition of the operator corresponding to the right hand side might correspond to terms describing advection, diffusion, formation rate (chemical reactions), emission, and sedimentation of pollutants.

In the case of partial differential equations like Eq. (1) the splitting procedure is formally the same, but the operators act on functions with values in an *infinite dimensional* (as opposed to finite dimensional above) space. If we consider the air pollution transport equation in Section 2, the operator A in Eq. (2) corresponds to the whole right hand side of Eq. (1), that is A represents all the effects in the transport process which are presented in Eq. (1). We can split A to more, simpler operators corresponding to the physical effects that are presented in the equation. One possible splitting is: advection (A_1), diffusion (A_2), emission (A_3), reaction (A_4), sedimentation (A_5); i.e., there will be 5 subproblems corresponding to the 5 subprocesses. Of course there are other possible splittings of the problem, but it is not our aim to go into details in this question.

4. *Investigation of the splitting error*

For the investigation of the splitting error, Lie-operators can give us some useful help. In this section we define the basic concepts of the theory of Lie-operators, which play role in the determination of the splitting error. With the help of Lie-operators we can give a condition for zero splitting error, which is the main result of *Havasi et al.* (2001). At the end of this section we analyze this condition for the special case of advection and chemistry operators, presented in Eq. (1).

4.1 The splitting error

The zero splitting error can be characterized via the Lie-operator, see also *Varadarajan* (1974). The *Lie-operator* $\mathbf{A}$ of the operator A is a linear operator acting on the space of differentiable operators, and it is defined for a differentiable operator G as follows:

$$\big(\mathbf{A}(G)\big)(\mathbf{c}) := \big(G'(\mathbf{c}) \circ A\big)(\mathbf{c}) \qquad \text{for} \quad \forall \mathbf{c} \in D_G \cap D_A,$$

where ' denotes differentiation with respect to $\mathbf{c}$. If evaluated at the identity operator, it is actually the original operator. It is not the aim of this paper to give a very detailed description of Lie-operators and their connection with the splitting error. A more detailed discussion of this issue can be found in *Havasi et al.* (2001). Here we just mention the main points of the argument for the condition of zero splitting error.

For the solution $\mathbf{c}(t)$ of Eq. (2) using the Taylor-series expansion, we can write:

$$\mathbf{c}(t+\tau) = (e^{\tau \mathbf{A}}(I))(\mathbf{c})(t). \tag{6}$$

For the solution of the splitted problem in the same way, we obtain:

$$\mathbf{c}_{\mathrm{sp}}(t+\tau) = (e^{\tau \mathbf{A}_2} e^{\tau \mathbf{A}_1}(I))(\mathbf{c}_{\mathrm{sp}})(t), \tag{7}$$

where $\mathbf{A}$, $\mathbf{A}_1$, and $\mathbf{A}_2$ are the Lie-operators of the operators A, A_1, and A_2 which are presented in Eqs. (1), (2), and (3). Taking the difference of Eqs. (6) and (7) and applying simple changes we get

$$\mathbf{c}(t+\tau) - \mathbf{c}_{\mathrm{sp}}(t+\tau) = (e^{\tau(\mathbf{A}_1+\mathbf{A}_2)}(I))(\mathbf{c}-\mathbf{c}_{\mathrm{sp}})(t) + (e^{\tau(\mathbf{A}_1+\mathbf{A}_2)} - e^{\tau \mathbf{A}_2} e^{\tau \mathbf{A}_1})(I)(\mathbf{c}_{\mathrm{sp}})(t). \tag{8}$$

The second term of the right hand side is the *local splitting error*, which arises through the use of splitting on the time interval $[t,t+\tau]$. The first term expresses the error coming from the local splitting error of earlier time steps. Considering Eq. (7), we see that the splitting error vanishes for all functions $\mathbf{c}$ if and only if

$$e^{\tau(\mathbf{A}_1+\mathbf{A}_2)}(I) = e^{\tau\mathbf{A}_2}e^{\tau\mathbf{A}_1}(I). \tag{9}$$

Let us define the *commutator* of $\mathbf{A}_1$ and $\mathbf{A}_2$ as

$$[\mathbf{A}_1,\mathbf{A}_2] := \mathbf{A}_1 \circ \mathbf{A}_2 - \mathbf{A}_2 \circ \mathbf{A}_1,$$

where $\mathbf{A}_1 \circ \mathbf{A}_2$ denotes the composition of the operators $\mathbf{A}_1$ and $\mathbf{A}_2$. *Havasi et al.* (2001) have shown by the investigation of the Baker–Campbell–Hausdorff-formula that Eq. (9) holds if and only if the commutator of the operators $\mathbf{A}_1$ and $\mathbf{A}_2$ evaluated at the identity operator I is zero, that is the operators $\mathbf{A}_1$ and $\mathbf{A}_2$ commute on the operator I. Applying the definition of the Lie-operator, we obtain the following form of the necessary and sufficient condition:

$$0 = ([\mathbf{A}_1,\mathbf{A}_2](I))(\mathbf{c}) = \big((\mathbf{A}_1 \circ \mathbf{A}_2 - \mathbf{A}_2 \circ \mathbf{A}_1)(I)\big)(\mathbf{c}) =$$
$$= \big(A'_2(\mathbf{c}) \circ A_1\big)(\mathbf{c}) - \big(A'_1(\mathbf{c}) \circ A_2\big)(\mathbf{c}).$$

Hereafter we will use this condition to analyze the splitting error of a given splitting procedure. Henceforth, if the Lie-operators of A_1 and A_2 commute, we say that the operators A_1 and A_2 L-commute, furthermore, we will use the notation:

$$[A_1,A_2]_L := [\mathbf{A}_1,\mathbf{A}_2](I).$$

We remark that if A_1 and A_2 are linear operators then the L-commutation is equivalent to the usual commutation, although $[A_1,A_2]_L = -[A_1,A_2]$.

4.2 L-commutativity of the advection and chemistry operators

Now we investigate the condition of L-commutativity in the case of advection and chemistry operators. We study the conditions under which the advection and chemistry operators presented in Eq. (1) L-commute. Related studies can be found in *Dimov et. al.* (1999, 2002), *Faragó* and *Havasi* (2001), and *Hundsdorfer* and *Verwer* (1994). Let us denote the advection operator from now on by A. It is defined by

$$(A\mathbf{c})(\mathbf{x},t) := -\big(\nabla\cdot\big(\mathbf{u}(\mathbf{x},t)c_1(\mathbf{x},t)\big),\nabla\cdot\big(\mathbf{u}(\mathbf{x},t)c_2(\mathbf{x},t)\big),...,\nabla\cdot\big(\mathbf{u}(\mathbf{x},t)c_n(\mathbf{x},t)\big)\big).$$

Let the chemistry operator $R=(R_1,R_2,...,R_n)$ be defined by

$$(R(\mathbf{c}))(\mathbf{x},t) := \big(r_1(\mathbf{x},t,\mathbf{c}(\mathbf{x},t)),r_2(\mathbf{x},t,\mathbf{c}(\mathbf{x},t)),...,r_n(\mathbf{x},t,\mathbf{c}(\mathbf{x},t))\big),$$

or using another notation emphasizing that R acts on the space of functions:

$$R(\mathbf{c}) := \big(r_1\circ(\mathrm{id},\mathbf{c}),r_2\circ(\mathrm{id},\mathbf{c}),...,r_n\circ(\mathrm{id},\mathbf{c})\big).$$

Since the operator A is linear, its derivative is A itself. The derivative of the operator R evaluated at the function $\mathbf{c}$ is an $n\times n$ matrix, with the element $\dfrac{\partial r_i}{\partial c_j}$ in the ith row and jth column. Then

$$[A,R]_L(\mathbf{c}) = \big(R'(\mathbf{c})\circ A\big)(\mathbf{c}) - \big(A'(\mathbf{c})\circ R\big)(\mathbf{c})$$

is an n-dimensional vector valued function whose ith coordinate function is given by

$$\Big([A,R]_L(\mathbf{c})\Big)_i = -\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\nabla\cdot(\mathbf{u}c_j)+\nabla\cdot(\mathbf{u}r_i).$$

We remark that for the scalar function $r_i(\mathbf{x},t,\mathbf{c}(\mathbf{x},t))$ and vector function $\mathbf{u}(\mathbf{x},t)$, the following relation holds

$$\nabla\cdot(\mathbf{u}r_i) = r_i\nabla\cdot\mathbf{u}+\mathbf{u}\cdot\left(\nabla_{\mathbf{x}}r_i+\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\nabla c_j\right),$$

where

$$\nabla_{\mathbf{x}}r_i(x,y,z,t,c_1,c_2,...,c_n) = \left(\frac{\partial r_i}{\partial x},\frac{\partial r_i}{\partial y},\frac{\partial r_i}{\partial z}\right).$$

Applying that identity, we obtain

$$\Big([A,R]_L(\mathbf{c})\Big)_i = -\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}c_j\nabla\cdot\mathbf{u}-\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\mathbf{u}\cdot\nabla c_j+r_i\nabla\cdot\mathbf{u}+\mathbf{u}\cdot\nabla_{\mathbf{x}}r_i+\mathbf{u}\cdot\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\nabla c_j =$$

$$= -\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}c_j\nabla\cdot\mathbf{u}+r_i\nabla\cdot\mathbf{u}+\mathbf{u}\cdot\nabla_{\mathbf{x}}r_i. \qquad (10)$$

The operators L-commute if and only if the expression given by Eq. (10) is equal to zero. A sufficient condition for this is

$$\nabla \cdot \mathbf{u} = 0 \quad \text{and} \quad \mathbf{u} \cdot \nabla_{\mathbf{x}} r_i = 0,$$

which holds for all $i = 1,2,\ldots,n$. Obviously, if the flow is divergence free and the rate of chemical reactions depends only on the concentrations and time, the above sufficient conditions are satisfied.

5. An advection-reaction model with Molenkampf–Crowley advection

In this section we present our results for a specialized two-dimensional model, henceforth, $\mathbf{x}$ denotes the two spatial coordinates (x, y). We study a problem with special advection, called *Molenkampf–Crowley* advection and chemistry processes:

$$\frac{\partial c_i(\mathbf{x},t)}{\partial t} = y\frac{\partial c_i(\mathbf{x},t)}{\partial x} - x\frac{\partial c_i(\mathbf{x},t)}{\partial y} + r_i\left(\mathbf{x},t,\mathbf{c}(\mathbf{x},t)\right), \qquad i = 1,2,\ldots,n. \tag{11}$$

We give necessary and sufficient conditions under which the operator of the Molenkampf–Crowley advection and the operator of chemical reactions L-commute.

5.1 Molenkampf–Crowley advection

The operator A of the advection is linear, its derivative is A itself. For n species the operator is defined as

$$A\mathbf{c} = \begin{pmatrix} y\dfrac{\partial c_1}{\partial x} - x\dfrac{\partial c_1}{\partial y} \\ \vdots \\ y\dfrac{\partial c_n}{\partial x} - x\dfrac{\partial c_n}{\partial y} \end{pmatrix}$$

This is a special case of the general advection presented in Eq. (1). We obtain it if Eq. (1) is reduced into two spatial dimensions, and the function $\mathbf{u}$ – which has now two coordinate functions u_1 and u_2 – equals to $(-y, x)$, that is $u_1(x,y,t) = -y$ and $u_2(x,y,t) = x$.

The Molenkampf–Crowley advection is often used as a test in the investigation of different methods, see, e.g., in *Zlatev* (1995). It is also known as *cone-test*. As a simple advection, it is suitable for us to investigate the behavior of the splitting error. The solution of the problem

$$\frac{\partial c_i}{\partial t} = y\frac{\partial c_i}{\partial x} - x\frac{\partial c_i}{\partial y}, \qquad i = 1,2,...,n, \tag{12}$$

with the initial condition $c_i(x,y,0) = x^2 + y^2$, forms a cone for every $i = 1,2,...,n$ and keeps its form during the time interval of the integration.

5.2 Main result: necessary and sufficient condition of L-commutativity

Now we show which special type of chemistry operators L-commute with the Molenkampf–Crowley advection. The operator of the chemical reactions R has the same form as defined before. Let us consider the ith coordinate of the L-commutator of A and R in two spatial dimensions:

$$\left([A,R]_L(\mathbf{c})\right)_i = u_1\frac{\partial r_i}{\partial x} + u_2\frac{\partial r_i}{\partial y} - \sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\left(u_1\frac{\partial c_j}{\partial x} + u_2\frac{\partial c_j}{\partial y}\right). \tag{13}$$

The notation $\frac{\partial r_i}{\partial x}$ and $\frac{\partial r_i}{\partial y}$ can be misleading. The function r_i has $n+3$ variables $(x,y,t,c_1,...,c_n)$ and depends on space and time also through the concentrations $(c_1,...,c_n)$. Thus, the derivative of the function r_i with respect to x is $\frac{\partial r_i}{\partial x} + \sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial x}$, where the notation $\frac{\partial r_i}{\partial x}$ is correct in a sense. So we study the expression

$$\left([A,R]_L(\mathbf{c})\right)_i = u_1\left(\frac{\partial r_i}{\partial x} + \sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial x}\right) + u_2\left(\frac{\partial r_i}{\partial y} + \sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial y}\right) -$$

$$-\sum_{j=1}^{n}\left(u_1\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial x} + u_2\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial y}\right) =$$

$$= u_1\frac{\partial r_i}{\partial x} + u_1\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial x} + u_2\frac{\partial r_i}{\partial y} + u_2\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial y} - u_1\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial x} - u_2\sum_{j=1}^{n}\frac{\partial r_i}{\partial c_j}\frac{\partial c_j}{\partial y} =$$

$$= u_1 \frac{\partial r_i}{\partial x} + u_2 \frac{\partial r_i}{\partial y}. \tag{14}$$

Substituting the functions $u_1(x,y,t) = -y$ and $u_2(x,y,t) = x$, we obtain that the splitting error is zero for any initial functions if and only if the equality

$$-y \frac{\partial r_i}{\partial x} + x \frac{\partial r_i}{\partial y} = 0 \tag{15}$$

holds for $i = 1,2,...,n$. This is a first order homogeneous linear partial differential equation for r_i. A general solution of Eq. (15) can be obtained as:

$$r_i(x,y,t,\mathbf{c}) = \varphi_i\left(x^2 + y^2, t, \mathbf{c}\right), \tag{16}$$

where φ_i is an arbitrary continuously differentiable function. This provides a necessary and sufficient condition for the L-commutation of Molenkampf–Crowley advection and chemistry. As we have mentioned, the chemistry operator represents the rate of the reactions between the pollutants. It usually depends on the concentrations and on a coefficient, which is modeled by the Arrhenius-equation:

$$k = A\mathrm{e}^{-\frac{E_a}{RT}}, \tag{17}$$

where T is the temperature and A, R, E_a are physical constants (see in *Atkins*, 1998; *Seakins* and *Pilling*, 1996). Hence, the coefficient k depends on the spatial coordinates and time through temperature. Thus, the condition of L-commutativity of the operators A and R is

$$T(x,y,t) = f\left(x^2 + y^2, t\right). \tag{18}$$

This means geometrically that the isotherms form concentric circles in space.

6. Numerical examples

In this section we present some numerical results for models with Molenkampf–Crowley advection and chemistry operators. Thus, the problem is two-dimensional and we have only one species, so $\mathbf{c} = (c_1,...,c_n)$ reduces to c. We study the following initial value problem:

$$\frac{\partial c(x,y,t)}{\partial t} = y\frac{\partial c(x,y,t)}{\partial x} - x\frac{\partial c(x,y,t)}{\partial y} + r\big(x,y,t,c(x,y,t)\big) \quad c(x,y,0) = c_0. \tag{19}$$

We will call this full equation since both operators are present. The splitting problem in the first time step will be as follows:

$$\frac{\partial c^{(1)}(x,y,t)}{\partial t} = y\frac{\partial c^{(1)}(x,y,t)}{\partial x} - x\frac{\partial c^{(1)}(x,y,t)}{\partial y}, \quad c^{(1)}(x,y,0) = c_0,$$

and

$$\frac{\partial c^{(2)}(x,y,t)}{\partial t} = r\Big(x,y,t,c^{(2)}(x,y,t)\Big), \quad c^{(2)}(x,y,0) = c^{(1)}(x,y,\tau). \tag{20}$$

We solve both the full and splitting problems numerically on a bounded domain: on a square. We choose three different functions r to present the chemical reactions, and solve the problem with each of them. However, our r functions typically do not represent real-life processes, but they are adequate to analyze the splitting error arising in the solution of the problems. We choose the functions for r in such a way that we know the exact solution of the problem, hence, we can compare it with the approximate solutions. Thus, we can calculate the errors of the approximations, which are the maximum norm of the difference between the exact solution and the approximation in every given case. We compare the solution of the full and splitting problems to the exact solution in each time step.

6.1 Numerical experiments

The numerical treatment requires temporal and spatial discretisation of the equations. See *Stoer* and *Bulirsch* (1993). Let us consider the following division of our domain:

$$\{x_i = ih,\ i = 1,2,\ldots,10;\ y_j = jh,\ j = 1,2,\ldots,10\}.$$

Hence, we obtain a 10×10 square grid. The grid parameter h was 1.0 in each problem. The time step of the numerical scheme and the time step of the splitting were equal to 0.015. We approximated the derivatives with first order finite differences. To approximate the spatial derivatives, in the advection term we used central scheme, and on the boundary of the domain we used the appropriate *forward* or *backward* Euler scheme. To approximate the time derivative on the left side, we used the *Euler forward* scheme. We solved the

equations by applying both explicit and implicit schemes. As initial value, we chose the value of the exact solution at time $t = 0$. In every problem we calculated the relative error as the maximum norm of the difference between the exact and the numerical solution. We compared these errors in different ways to analyze the effect of splitting. In the figures we plotted the errors against time steps.

In the first problem we had $r(x,y,t,c) = c - xy + x^2 - y^2 - t + 1$. The exact solution is $c(x,y,t) = xy + t$, with the corresponding initial condition $c(x,y,0) = xy$. As we can see it easily, the condition of L-commutativity does not hold in this problem. In the figures we can see that the error in the solution of the splitting problem is greater than in the solution of the full problem. *Fig. 1* shows the relative errors of the solutions of the *full* and *splitting* problems. Both of the problems were solved numerically by an *explicit Euler* scheme. *Fig. 3* shows the same calculations by *implicit Euler* schemes. *Figs. 2* and *4* show the difference between the relative errors of the splitting and full solutions in order generated by an explicit scheme and implicit schemes. The implicit scheme provides a more accurate solution compared to the explicit scheme. We can observe that the implicit scheme equilibrates the effect of splitting and its error much more than the explicit scheme. *Fig. 5* shows the difference between the relative error of the full and splitting problem solutions generated by explicit and implicit schemes. In *Fig. 5* we can see how the splitting affects the accuracy with implicit and explicit schemes. See also *Fig. 6*, which shows the relative error of the solution of the splitting problem generated by explicit and implicit schemes.

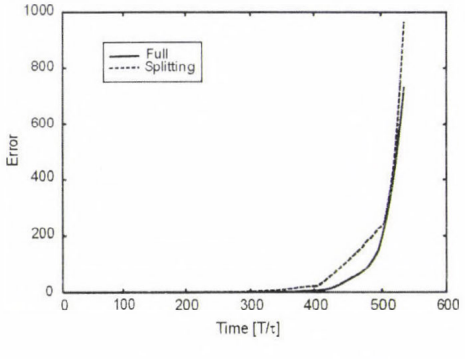

Fig. 1. The relative errors of the numerical solutions of the first problem are plotted against the number of time steps.

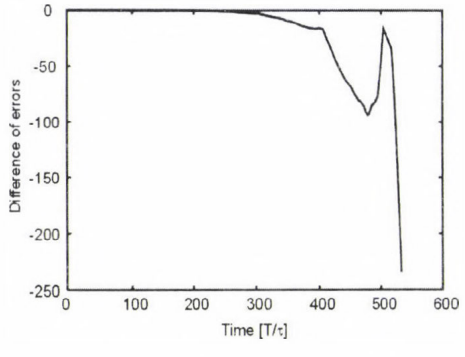

Fig. 2. The difference between the relative errors of the numerical solutions of the *full* and *splitting* problems is plotted against the number of time steps.

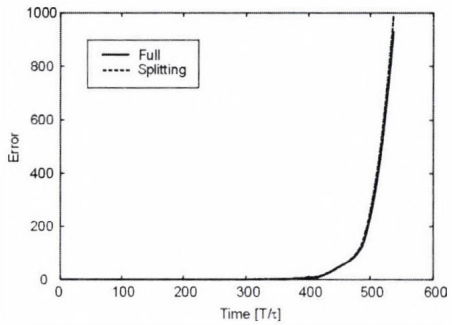

Fig. 3. The relative errors of the numerical solutions of the first problem are plotted against the number of time steps.

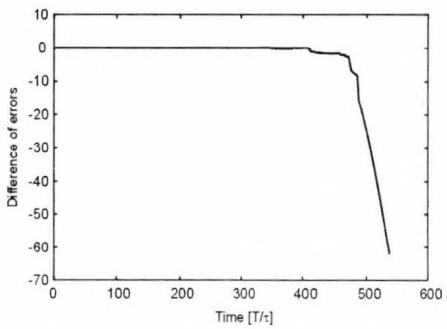

Fig. 4. The difference between the relative errors of the numerical solutions of the *full* and *splitting* problems is plotted against the number of time steps.

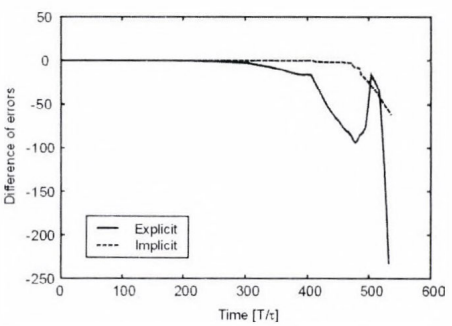

Fig. 5. The differences between the relative errors of the numerical solutions of the *full* and *splitting* problems are plotted against the number of time steps.

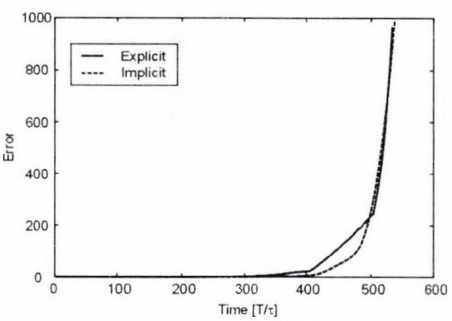

Fig. 6. The relative errors of the numerical solutions of the splitting problem are plotted against the number of time steps.

In the second problem we had $r(x,y,t,c)=c-2y$. The splitting error is not zero. The exact solution of this problem is $c(x,y,t)=x+y$. In this special case, the applied finite difference schemes provide the exact solution for the full problem. Using an implicit scheme we obtain the exact solution with splitting, too. The explicit scheme provides a solution different from the exact solution. See *Fig. 7*, which shows the relative errors of the solutions of the full and splitting problems. Both of the problems were solved numerically by an explicit Euler scheme. The solution of the full problem was equal to the exact solution; so its relative error is constantly zero. We may observe significant

the splitting problem. We can say that the splitting procedure is sensitive to the choice of the applied numerical scheme. The splitting error does not appear in the solution generated by the implicit scheme.

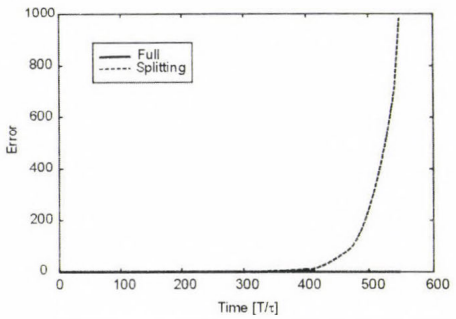

Fig. 7. The relative errors of the numerical solutions of the second problem are plotted against the number of time steps.

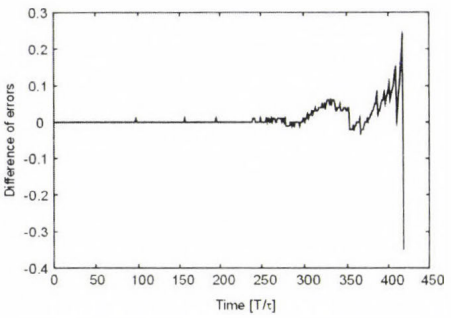

Fig. 8. The difference between the relative errors of the numerical solutions of the *full* and *splitting* problems is plotted against the number of time steps.

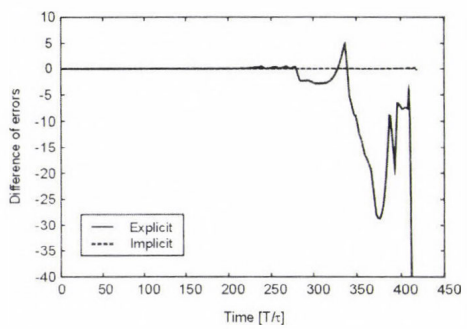

Fig. 9. The differences between the relative errors of the numerical solutions of the *full* and *splitting* problems are plotted against the number of time steps.

In the third problem we had $r(x,y,t,c) = c - x^2 - y^2 - t + 1$. Note that in this problem the condition of L-commutativity is satisfied. The function of chemical reaction rate has a form $f(x^2 + y^2, t, c)$. The exact solution of the problem is $c(x,y,t) = x^2 + y^2 + t$. Comparing to the previous problems, we can see that we obtain more accurate solutions of the splitting problem compared to the full one than before. There are intervals in which the solution of the splitting problem is closer to the exact solution than the solution of the full problem. See *Fig. 8*, which shows the difference between the relative errors of the solutions of the full and splitting problems. Both of the problems were solved numerically by an implicit Euler scheme. Also, in this third problem we

can say that the implicit scheme provides a more accurate solution than the explicit scheme. Furthermore, the implicit scheme equilibrates the effect of splitting and its error much more than the explicit scheme. See also *Fig. 9*, which shows the difference between the relative error of the full and splitting problem solutions generated by explicit and implicit schemes.

6.2 Conclusion of the numerical experiments

The solutions of different problems show that the errors generated by splitting are sensitive to the choice of the numerical method. We can say in general that the implicit scheme equilibrates the effect of splitting and its error much more than the explicit scheme. We also obtained a result, where the theoretically non-zero splitting error did not appear in the solution.

7. Discussion and perspectives

In this paper the general result by *Havasi et al.* (2001) was on zero splitting error. The condition of L-commutativity was investigated in the case of Molenkampf–Crowley advection and chemistry operators. A necessary and sufficient condition on the L-commutation of the advection and chemistry operators was derived. Numerical experiments were performed with Molenkampf–Crowley advection and different chemistry operators.

After splitting the problem, a numerical method is used to solve the subproblem. It means, that the subproblem is not solved exactly, so the L-commutativity does not provide direct information about the error of the obtained solution. To have a quantitative estimation of the error we need to study the approximation operators defined by the numerical methods. In all of the three studied problems we obtained very different results regarding the effect of splitting when using different numerical methods.

Our future plan is to make a systematic study on reaction diffusion equations. We would like to study the effect of splitting in reaction diffusion equations. Our main aim is to characterize the splitting error if it is not zero to map the advantages of splitting in running time and accuracy.

Acknowledgement—The present work has partially been supported by the National Science Foundation, Hungary (T037491).

References

Atkins, P.W., 1998: *Physical Chemistry* (6th ed.), Vol. III., *Change*. W.H. Freeman Co., New York.

Dimov, I., Faragó, I., and *Zlatev, Z.*, 1999: Commutativity of operator splitting methods for air pollution models. *Technical Report 04/99*. Bulgarian Academy of Sciences.

Dimov, I., Faragó, I., Havasi, Á., and *Zlatev, Z.*, 2002: L-commutativity of the operators in splitting methods for air pollution models. *Annal. Univ. Sci. Sec. Math. 44*, 127-148.

Faragó, I. and *Havasi, Á.*, 2001: The mathematical background of operator splitting and the effect of non-commutativity. *Proc. of the Conference on Large-scale Scientific Computations*. Springer, 264-271.

Havasi, Á., Bartholy, J., and *Faragó, I.*, 2001: Splitting method and its application in air pollution modeling. *Időjárás 105*, 39–58.

Hundsdorfer, W. and *Verwer, J.G.*, 1994: A note on splitting errors for advection-reaction equations. *Appl. Math. 18*, 191-199.

Lanser, D. and *Verwer, J.G.*, 1999: Analysis of operator splitting for advection-diffusion-reaction problems from air pollution modelling. *J. Comput. Appl. Math. 111*, 201–216.

Marchuk, G.I., 1968: Some application of splitting-up methods to the solution of mathematical physics problems. *Applik. Mat., 13*.

McRae, G.J., Goodin, W.R., and *Seinfeld, J.H.*, 1984: Numerical solution of the atmospheric diffusion equations for chemically reacting flows. *J. Comput. Phys. 45*, 1–11.

Seakins, P.W. and Pilling, M.J., 1996: *Reaction Kinetics*. Oxford University Press, Oxford.

Sportisse, B., 2000: An analysis of operator splitting techniques in the stiff case. *J. Comput. Phys. 161*, 140–168.

Stoer, J. and *Bulirsch, R.*, 1993: *Introduction to Numerical Analysis*. Springer-Verlag.

Strang, G., 1968: On the construction and comparison of difference schemes. *SIAM J. Numer. Anal. 5*, 505–517.

Varadarajan, V.S., 1974: *Lie-groups, Lie-Algebras and Their Representations*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Zlatev, Z., 1995: *Computer Treatment of Large Air Pollution Models*. Kluwer Academic Publishers, Dordrecht-Boston-London.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 109, No. 3, July–September 2005, pp. 189–202

Long-term variations of temperature, wind, and precipitable water in the troposphere and lower stratosphere over Budapest, Hungary

Sándor Cseh[1] and **Pál Bencze**[2]

[1]*Apáczai Csere János Teacher's Training Faculty, University of Western Hungary, P.O. Box 204, H-9002 Győr, Hungary*
[2]*Geodetic and Geophysical Research Institute, Hungarian Academy of Sciences, P.O. Box 5, H-9401 Sopron, Hungary; E-mail: bencze@ggki.hu*

(Manuscript received in final form May 27, 2005)

Abstract—Long-term variations of the height of different isobaric levels, temperature, wind, and precipitable water in the troposphere and lower stratosphere have been studied for determination of the temporal trend of these parameters at mid-latitudes over Budapest, Hungary. One of the reasons of this study was to establish the height, where the trend of temperature changes sign from positive (increasing) in the lower troposphere due to the global warming to negative (decreasing) in the upper part of the troposphere over Budapest, Hungary. For the analyses of the behavior of temperature with increasing altitude, radiosounding data were used. It has been found that the trend of temperature is positive from the ground to the height of the 500 hPa isobaric level (~5.6 km), then it turns into negative. The trend has been determined by using linear regression. The results show that the lower troposphere is warming from the ground to a height of about 6 km, while the region above this height is cooling. The precipitable water in a column of unit area extending from 850 hPa (1.5 km) to 200 hPa (12 km) indicates a significant negative (decreasing) trend, suggesting a drying tendency in the troposphere.

Key-words: temperature in the troposphere, precipitable water, warming of the atmosphere, long-term change of meteorological parameters.

1. Introduction

The atmosphere is a continuous substance. Effects of atmospheric processes taking place in the lower atmosphere are not limited to the lower atmosphere, and inversely, consequences of atmospheric phenomena in the upper

atmosphere are not limited to the upper atmosphere. Such phenomena are, e.g., the origin of the atmospheric waves, which can also be observed in the upper atmosphere, or absorption of the solar electromagnetic radiation in the upper atmosphere, as well as modification of the energy spectrum of energetic particles of galactic origin by the upper atmosphere. This circumstance should also be considered, if regional climate scenarios are drawn up, as it is done in case of the Carpathian Basin (*Mika,* 1988, 1994).

An important issue of the study of climate change is the determination of the proportion, by which different processes can affect the global change. These processes generally considered are the greenhouse effect, aerosol particles, cloudiness, solar radiation (solar activity), geomagnetic activity, and galactic cosmic rays. Investigations are partly carried out taking into account one of the processes. There is also difference in the length of time, to which these investigations are extended, short period or long-term changes of the processes.

Concerning the greenhouse effect, it is due to the so-called greenhouse gases, which absorb and reradiate the long wave temperature radiation of the Earth's surface and the overlying atmospheric layers in all directions. First of all, these greenhouse gases are responsible for global warming (*Lean* and *Rind,* 1999). The climate is especially sensible to enhancement of the concentration of carbon dioxide. This circumstance is related to the coincidence of the wavelength of the maximum long wave radiation of the ground with the main wavelength band of absorption of CO_2. A study of the effect of greenhouse gases extending to both of the middle and upper atmosphere has been carried out by *Roble* and *Dickinson* (1989). Model results of these authors predicted a long-term decrease of the NO concentration, which indicated a shrinking of the atmosphere that is cooling at stratospheric and mesospheric heights. Decrease of the NO concentration has been experimentally proved by analysis of the reflection height of low frequency (LF) radio waves (*Bremer* and *Berger*, 2002). Model results have shown that the experimental temperature trends can qualitatively be explained by an increasing atmospheric greenhouse effect (*Bremer* and *Berger*, 2002; *Bremer*, 2003). According to a review of the results referring to long-term changes in the upper atmosphere (*Bremer et al.*, 2004), an increasing atmospheric greenhouse effect can also be shown in the height range of 90–150 km. An increasing effect of greenhouse gases can be traced at altitudes from 150 to 180 km, too. However, above these altitudes no sure effect of greenhouse gases could be revealed. It is obvious that this circumstance is related to the decreasing concentration (rarefaction) of greenhouse gases.

The role of the cloudiness is a complex problem. Clouds can influence the global change by extension of the cloud cover depending on optical thickness, reflectivity, and height of the clouds. The extension of the cloud cover depends

on temperature and humidity of near surface atmospheric layers, as well as on many other factors, as, e.g., topography, instabilities.

Concerning the solar radiation, the total solar irradiance changes over periodicities from minutes to the eleven-year solar cycle (*Pap* and *Fröhlich,* 1999; *Rottman,* 1999). Observations of the intensity of solar radiation carried out during solar flares in the past have shown that the intensity of radiation in the EUV and X-ray bands increases for a period of about one hour, and the extent of the increase is reduced with increasing wavelength. The enhancement of radiation due to solar flares may not have a significant effect on the atmosphere because of the short time and changes limited to the EUV and X-ray bands, which can only penetrate to the height of about 60 km. Thus, this radiation increase is insignificant from point of view of the total irradiation. Intensity of the ultraviolet irradiance is varying due to the solar rotation related to the non-uniform distribution of active areas on the solar disk. In this case the magnitude of irradiance also decreases with increasing wavelength to about 300 nm (*Rottman,* 1999). The solar cycle variation of the irradiance is also limited to a wavelength of less than 300 nm. The 27-day variation expressed by the ratio of the solar spectrum referring to an active day to the spectrum related to a less active day shows a value of about 20% at the Lyman α wavelength. The ratio of the solar spectrum in the maximum solar activity period to the spectrum in the minimum period indicates an enhancement about 70% at the Lyman α wavelength (*Rottman,* 1999). Nevertheless, assuming a climate sensitivity of 1 °C W^{-1} m^{-2}, the enhancement of the total solar irradiance of 0.14% from 1900 to 1990 can only be responsible for 0.25 °C of the 0.6 °C surface warming (*Lean* and *Rind,* 1999).

Geomagnetic activity plays also an indirect role in the change of the global climate. Geomagnetic activity is the consequence of the solar activity in the Earth's environment. As a consequence of heat input during geomagnetic disturbances, the atmosphere may also be affected by creating and strengthening the reverse circulation at high latitudes in the middle atmosphere, as well as extending it to mid-latitudes. The heat input related to geomagnetic disturbances is due to precipitation of energetic electrons ($>$ 30 keV) into the auroral zone, increasing ionization there, and thus, increasing electric conductivity in the height range of 90–120 km. Enhancement of the conductivity results in the intensification of the auroral electrojet and in the enhancement of the Joule heat by the increased current intensity.

Effect of the galactic cosmic rays are taken into account, since they are the main ionization radiation below about 90 km in the atmosphere. The flux of galactic cosmic radiation depends on the state of the interplanetary magnetic field. On the other hand, the state of the interplanetary magnetic field is controlled by the solar activity increasing the production of magnetic field

irregularities (piling up of magnetic field lines in front of plasma clouds and production of shock fronts ahead and behind corotating high speed plasma streams). Galactic cosmic ray particles are scattered by these irregularities, and thus, their flux decreases, which is called Forbush decrease. The effect of irregularities of the interplanetary magnetic field is the greater, the smaller is the particle energy, and it is inversely proportional to the solar activity. Galactic cosmic rays may affect the global change indirectly by ionization influencing the effectiveness of the transformation of supercooled drops into ice (tropospheric aerosols).

On the one hand, supercooled drops are thermodynamically unstable, because they are in transitional state between the liquid and solid states. As the transformation of supercooled drops to ice is finished, these particles begin to grow by the deposition of water vapor to the surface of ice crystals. Both transition of the supercooled drops to ice and deposition of the water vapor to ice crystals are related to the release of latent heat. The ice grains begin to sink as soon as they reach the corresponding dimension. Thus, on the other hand, this system is also gravitationally unstable. Ice grains may decrease the water vapor concentration in the middle part of clouds, too, by deposition of water vapor to ice gains during their fall causing further release of latent heat.

The above described process would mean no unusual matter as compared with processes known so far. However, theoretical and experimental investigations have shown that if supercooled drops have elecrical charge, transition of the supercooled drops into ice is accelerated (*Pruppacher* and *Klett*, 1997). Supercooled drops originating in the upper part of the clouds become positively charged due to the positive charge accumulation as a result of polarization of the clouds. Polarization of clouds embedded in the atmospheric electric circuit is produced by the vertical air-earth current. Electric conductivity of the clouds is smaller than that of the environment. Charge accumulation in the upper part of the clouds depends on the intensity of the vertical air-earth current. However, intensity of the vertical air-earth current is also proportional to the electric conductivity of the air, which is determined by the ion concentration and mobility of ions. Thus, the degree of the charge accumulation can be changed by the effect of the galactic cosmic rays on the ionization, affecting the ion concentration. Estimations referring to the latent heat released in course of the above discussed cloud physical processes have shown the following results. The water leaving the cloud, as precipitation is only about 10% of the total cloudwater. The remaining part of the cloudwater evaporates again. By the increase of the effectivity of the precipitation formation, the water leaving the cloud as precipitation increases from 10% to 11%, taking a speed of the rainfall of 1 cm $hour^{-1}$, and it would produce an energy flux of about 10^{-1} W cm^{-2} for the heating of the atmosphere. Considering

the limited spatial extension and period of the rainfall, this energy flux would represent an increase of the energy flux supplied to the atmosphere by the galactic cosmic ray flux of 10^{-10} W cm^{-2} by a factor of 10^9 (*Tinsley*, 1996).

2. *Method and data*

Monthly mean values of pressure, temperature, humidity, and wind are measured by radiosounding at different altitudes at the Aerological Observatory of the Hungarian Meteorological Service, Budapest (47°26'N; 19°11'E). Measured profiles are published in yearbooks. Yearly means were calculated from the monthly mean values using the data of midday ascents. Linear regression analyses were used for determination of the trend of the different parameters. Correlation between the yearly mean values and the regression line has also been computed and indicated in the figures. The data series comprised data of 33 years from 1962 to 1994, referring to the middle atmosphere. Thus, trends of the height of selected isobaric levels, that of the temperature and wind at these levels, and the trend of precipitable water were determined. In case of such a long interval, it must be considered that the measuring equipment and data processing have been changed, furthermore, the measured data include smaller or larger errors, however, computation of yearly means may significantly decrease the errors.

According to previous investigations aimed at the determination of reliable trends, it is necessary to have a data series comprising at least three solar cycles (*Bremer,* 1992). *Bremer* (1992) found that the sign of the trend fluctuated strongly between positive and negative values, if the length of the data series is less than three solar cycles. The length of the period, within which the sign of the trend does not change already, depends on the season. Sign of the trend is most quickly stabilized in the summer months (less than two solar cycles), and the final value is most slowly reached in the winter months. This difference might be attributed to the variability of the data within the data series.

Concerning the significance of the correlation coefficient, it is to be noted that monthly mean values of 33 years are used, that is, nearly 400 values. The value of the significant correlation coefficient decreases with increasing magnitude of the sample (*Taubenheim*, 1969). In this case all correlation coefficients are significant at the 95% level of confidence, if their values are greater than 0.18. The coefficient of determination, R^2, is shown in the figures. The coefficient of the determination is equal to the square of the correlation coefficient. Thus, the correlation coefficient is simply its square root.

3. *Results and discussion*

The trend of the height of isobaric levels is positive. This means increasing height with time at the isobaric levels of 850, 500, 250 hPa, and at the tropopause (*Fig. 1*). The sign of the trend changes above the isobaric level 150 hPa, this level is indicating the transition. The trend of the height of isobaric levels is already negative, the height decreases with time in case of the isobaric levels of 150, 30, 20, and 10 hPa (*Fig. 2*). If we study the trend of the temperature at different isobaric levels, it may be established that trends of the height of isobaric levels follow the behavior of the trends of temperature at levels to a height of about 250 hPa. It can be seen in *Fig. 3*, that the temperature shows positive trend at the ground and at the isobaric levels of 900, 850, 700, 600, and 500 hPa. Change of the sign of the trend occurs below about 250 hPa, at this height the trend of the temperature is already negative, the temperature decreases with time (*Fig. 3*). Trends of the temperature are negative at the isobaric levels of 150, 100, 50, 30, and 20 hPa, but positive at the isobaric level of 10 hPa (*Fig. 4*).

The approximately similar behavior of the trends of the height of isobaric levels and temperature may be explained by the change of temperature resulting in the expansion or contraction of the atmosphere. The height of the isobaric levels increases with time, where the temperature rises with time, as in the troposphere. Similarly, the height of the isobaric levels decreases with time, where the temperature sinks, as in the lower stratosphere. Change of the positive sign of the trends below about 250 hPa into negative sign and then, its changing into positive sign at heights above 20 hPa hints at a possible effect of ozone. The heating in the stratosphere is associated with absorption of the solar electromagnetic radiation by ozone, mainly in the wavelength band of 200–300 nm. It may be assumed that the reduced ozone concentration is attributable to the human activity, and the decreased heating due to the lowered ozone concentration in the past century results in the negative trend of temperature in the lower stratosphere from the isobaric level of 250 hPa (10 km) to about 20 hPa (27 km). It is to be noted that the ozone depletion is the largest in the lower stratosphere (*Bencze*, 1991).

In the troposphere the height of isobaric levels increases with time, but in different degree. The positive trend decreases with decreasing pressure, the height difference between isobaric levels diminishes with time. This means that the vertical pressure gradient is growing with time. According to the equation of pressure of the kinetic gas theory, the vertical gradient of pressure is proportional to the vertical gradient of temperature. As we have seen, the degree of the trend of the temperature at isobaric levels decreases also with decreasing pressure in the troposphere. Thus, the vertical gradient of

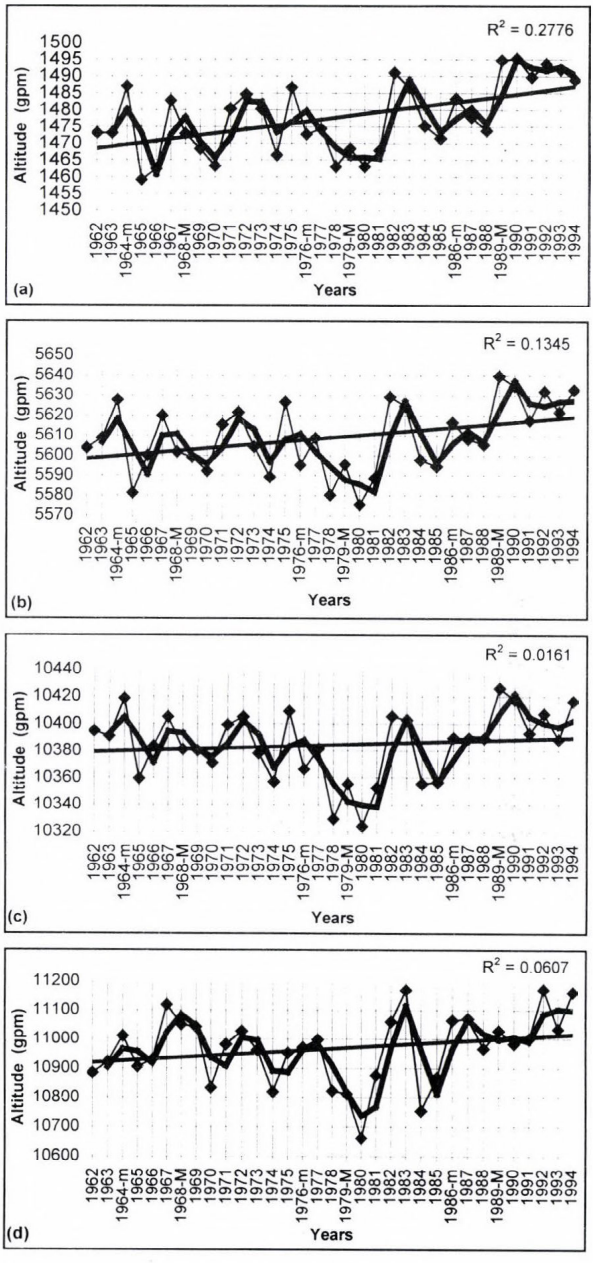

Fig. 1. Long-term change of the altitude of the 850 hPa (*a*), 500 hPa (*b*), 250 hPa (*c*) isobaric levels, and the tropopause (*d*) above Budapest. Square of the correlation coefficient is printed in the right upper corner. Thick curve shows the two-year sliding average. *m* and *M* indicate solar activity minimum and maximum, respectively.

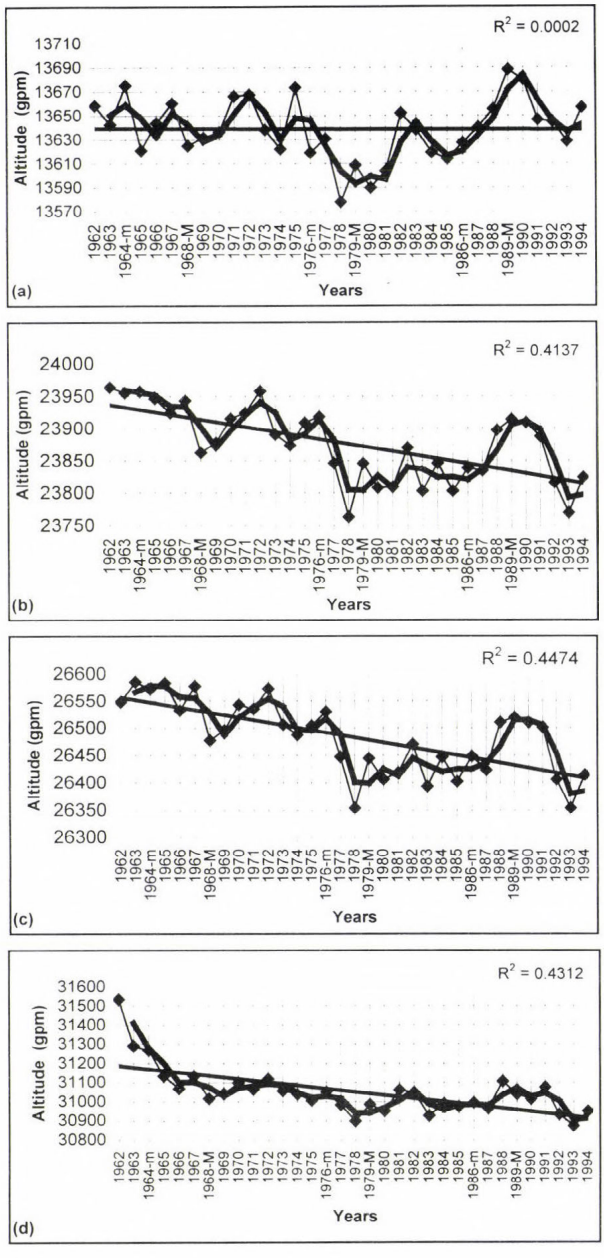

Fig. 2. Long-term change of the altitude of the 150 hPa (*a*), 30 hPa (*b*), 20 hPa (*c*), and 10 hPa (*d*) isobaric levels above Budapest. Square of the correlation coefficient is printed in the right upper corner. Thick curve shows the two-year sliding average. *m* and *M* indicate solar activity minimum and maximum, respectively.

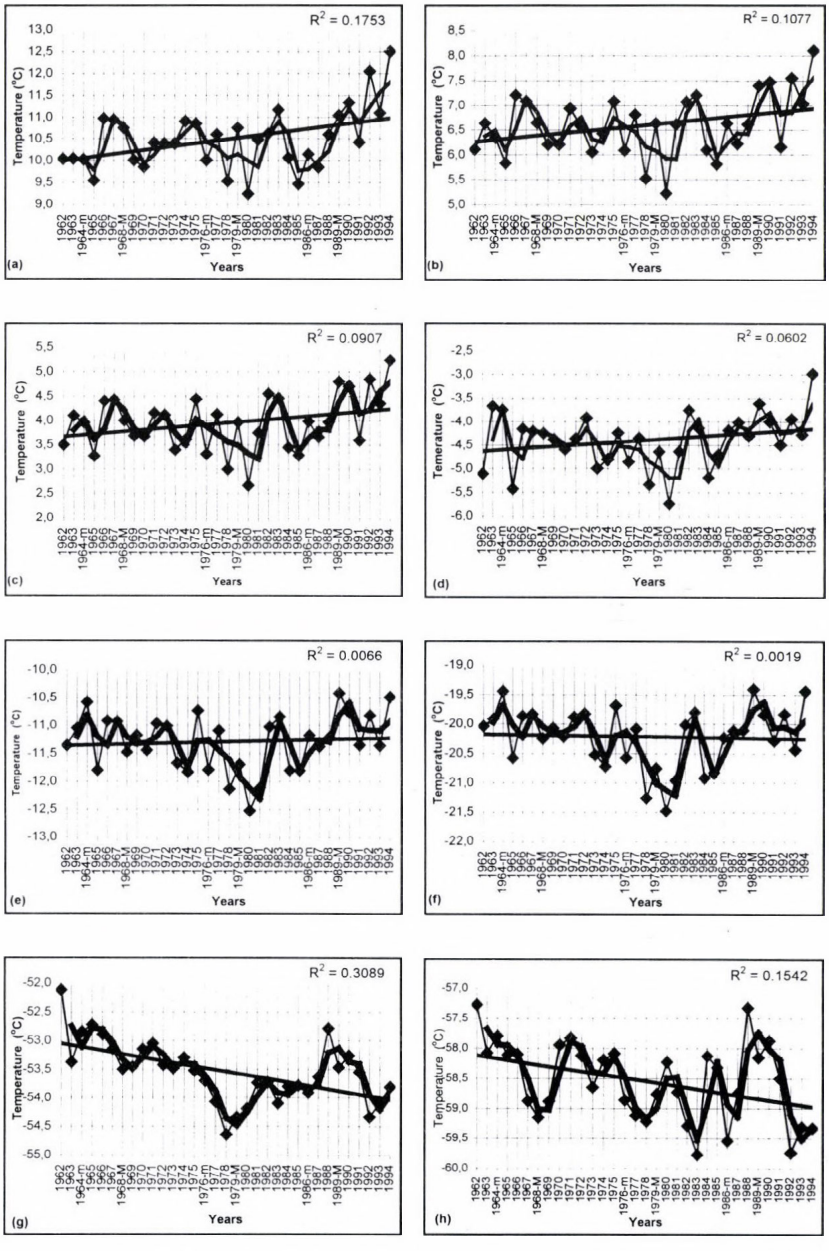

Fig. 3. Long-term change of the temperature at the ground (*a*), at the isobaric levels of 900 hPa (*b*), 850 hPa (*c*), 700 hPa (*d*), 600 hPa (*e*), 500 hPa (*f*), 250 hPa (*g*), as well as at the tropopause (*h*) above Budapest. Square of the correlation coefficient is printed in the right upper corner. Thick curve shows the two-year sliding average. *m* and *M* indicate solar activity minimum and maximum, respectively.

temperature increases in accordance with growing vertical pressure gradient. Considering the lower stratosphere, both the trend of the height of the isobaric levels and trend of the temperature are negative, the degree of the trend is increasing with falling pressure. This means that both the vertical gradient of pressure and that of the temperature decrease with time in accordance with the equation of pressure of the kinetic gas theory.

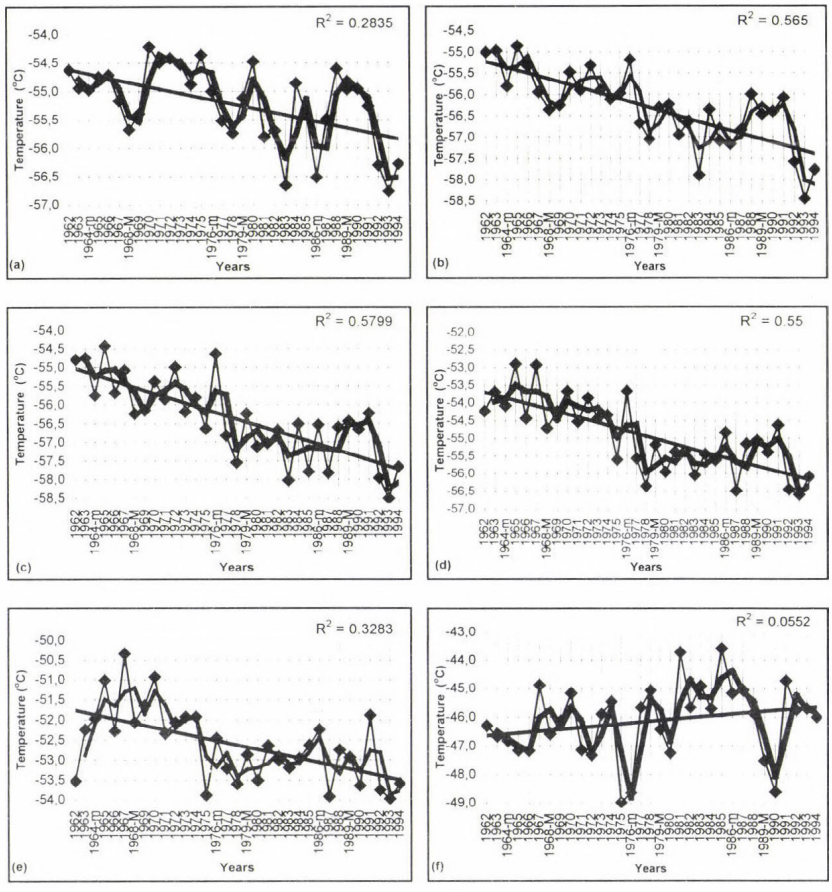

Fig. 4. Long-term change of the temperature at the isobaric levels of 150 hPa (*a*), 100 hPa (*b*), 50 hPa (*c*), 30 hPa (*d*), 20 hPa (*e*), and 10 hPa (*f*) above Budapest. Square of the correlation coefficient is printed in the right upper corner. Thick curve shows the two-year sliding average. *m* and *M* indicate solar activity minimum and maximum, respectively.

Trends of the wind at the isobaric levels also show some regularities. In the troposphere, for example, at the isobaric level of 500 hPa, the trend of the wind speed (speed of the prevailing wind) is positive, speed increases with

time (*Fig. 5*). Though, the change of the wind speed is not significant, nevertheless, it indicates an enhancement. However, in the lower stratosphere, at the isobaric level of 150 hPa, the trend of the wind speed is already negative, indicating a decrease of the wind speed with time (*Fig. 6*). In this case, the decrease of the wind speed with time is significant at the 95% level of confidence, according to both the t-test and F-test of the level of confidence for correlation coefficients. The behavior of the trend of the wind speed may be explained by considering the trend of the wind speed at different isobaric levels. The trend of the wind speed at different isobaric levels in the troposphere decreases with increasing height, thus, the vertical gradient of the wind speed is enhanced with time. This might be related to the similar change of the meridional temperature gradient (thermal wind equation), that is the latitudinal gradient of temperature increases with time increasing the speed of the zonal wind component with time. In the lower stratosphere one can meet the opposite situation.

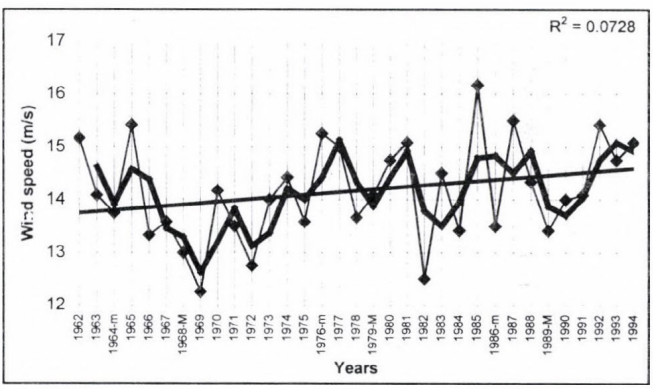

Fig. 5. Long-term change of the wind speed at the isobaric level of 500 hPa above Budapest. Square of the correlation coefficient is printed in the right upper corner. Thick curve shows the two-year sliding average. *m* and *M* indicate solar activity minimum and maximum, respectively.

The study of the precipitable water can be limited to the troposphere, since the predominant part of the water vapor is concentrated in the troposphere. Thus, the precipitable water in a column of unit area between the levels of 850 hPa and 200 hPa has been determined. The trend of the precipitable water is negative indicating the decrease of precipitable water in the troposphere with time (*Fig. 7*). This trend is significant according to both the t-test and F-test of the significance level of correlation coefficients. Though

the measuring equipments and data processing could be changed during this interval, as well as measurements include smaller or larger errors, yearly means enable us at least to determine the trend of the long-term variation. This would show the drying out of the troposphere. Other investigations related to the precipitable water indicate both positive and negative trends referring largely to the same period (1973–1995) (*Ross* and *Elliot*, 2001). Increase of the precipitable water from the surface to 500 hPa has been found over North America except of the north-eastern part of Canada. In Eurasia only China and the Pacific Islands show a regional increase of precipitable water. However, Europe, extending to the Ural Mountains, is an area of decreasing precipitable water. Thus, there are significant regional differences in the trends of precipitable water. Our results referring to Budapest (Hungary) fit very well into the spatial distribution of the trend of precipitable water.

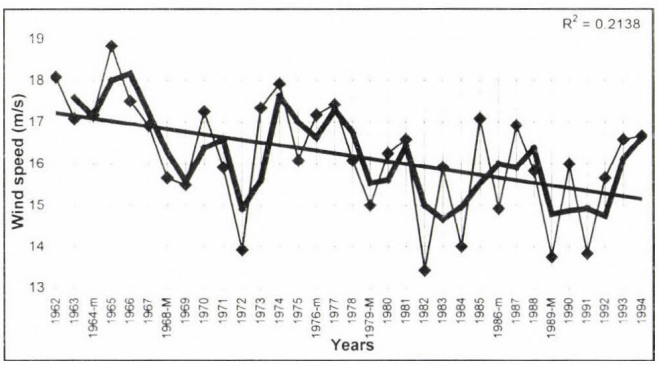

Fig. 6. Long-term change of the wind speed at the isobaric level of 150 hPa above Budapest. Square of the correlation coefficient is printed in the right upper corner. Thick curve shows the two-year sliding average. *m* and *M* indicate solar activity minimum and maximum, respectively.

Drying out of the troposphere might be attributed to the long-term changes of the total cloud amount over land (*IPCC*, 2001). Growing total cloud amount may be accompanied, on the one hand, by increasing precipitation reducing the precipitable water in the troposphere. On the other hand, the solar activity has been indicating an upward tendency since the end of the 19th century (*Reid*, 1999). This assumption is also in agreement with the enhanced sea-surface temperature (*Reid*, 1999). According to relevant investigations, increase of the sea-surface temperature is accompanied also by growth of the global thunderstorm activity. Thus, it may be assumed that drying out of the troposphere is also related in this way to enhancement of the global thunderstorm activity.

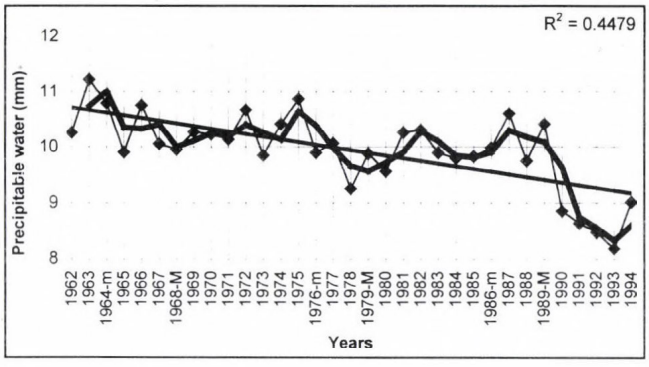

Fig. 7. Long-term change of the precipitable water referring to a cloumn of unit area extending between the isobaric levels of 850 hPa and 200 hPa above Budapest. Square of the correlation coefficient is printed in the right upper corner. Thick curve shows the two-year sliding average. *m* and *M* indicate solar activity minimum and maximum, respectively.

The relation between trends of the height of isobaric levels, temperature, and wind, as well as solar activity might also be studied. Earlier investigations have shown that the solar activity may affect indirectly the meteorological parameters. Correlations of atmospheric dynamics with solar activity evidenced by a connection via the solar wind (interplanetary magnetic field), atmospheric electricity, and cloud microphysics have been found (*Tinsley* and *Heelis,* 1993). This is confirmed by finding the solar wind modulation of the global electric circuit and apparent effects on cloud microphysics, latent heat release, and tropospheric dynamics (*Tinsley,* 1996), by demonstrating the signal of the 11-year solar cycle in the global stratosphere (*Labitzke* and *van Loon,* 1998) and in the total ozone content (*Labitzke* and *van Loon,* 1995). As a further indirect effect of the solar activity, the work of *Svensmark* and *Friis-Christensen* (1997) might be mentioned, who revealed relation between variations of the cosmic ray flux and the global cloud coverage – as a missing link in solar-climate relationship. According to these results, long-term change of the meteorological parameters studied in this paper might partly also be influenced by solar activity, though the main reason seems to be the greenhouse effect. Studies related to estimation of the magnitude of the effects attributed to various sources of a global temperature increase indicate that the solar variability may be responsible for 0.25 °C of the 0.6 °C warming from 1900 to 1990 (*Lean* and *Rind,* 1999).

Acknowledgements—The authors thank to the Hungarian Meteorological Service for using the radiosonde data. Also, thanks to the two unknown referees for their useful comments.

References

Bencze, P., 1991: Development of the concentration of ozone in the middle atmosphere (in Hungarian). *Időjárás 95*, 77-93.

Bremer, J., 1992: Ionospheric trends in mid-latitudes as a possible indicator of the atmospheric greenhouse effect. *J. Atmos. Terr. Phys. 54*, 1505-1511.

Bremer, J., 2003: Mesospheric temperature trends at mid-latitudes derived from LF radio propagation experiments. *Adv. Space Res. 32*, 1653-1662.

Bremer, J. and *Berger, U.*, 2002: Mesospheric temperature trends derived from ground-based LF phase-height observations at mid-latitudes: comparison with model simulations. *J. Atmos. Sol.-Terr. Phys. 64*, 805-816.

Bremer, J., Alfonsi, L., Bencze, P., Lastovička, J., Mikhailov, A.V., and *Rogers, N.*, 2004: Long-term trends in the ionosphere and upper atmosphere parameters. *Ann. Geophys. 47*, 1009-1029.

IPCC, 2001: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (eds.: *J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, C.A. Johnson*). Cambridge University Press, Cambridge, U. K.

Labitzke, K. and *van Loon, H.*, 1995: Total ozone and the 11-year sunspot cycle. *J. Atmos. Sol.-Terr. Phys. 59*, 9-19.

Labitzke, K. and *van Loon, H.*, 1998: The signal of the 11-year solar cycle in the global stratosphere. *J. Atmos. Sol.-Terr. Phys. 61*, 53-61.

Lean, J. and *Rind, D.*, 1999: Evaluating sun-climate relationships since the Little Ice Age. *J. Atmos. Sol.-Terr. Phys. 61*, 25-36.

Mika, J., 1988: Regional features of global warming in the Karpathian-Basin (in Hungarian). *Időjárás 92*, 178-189.

Mika, J., 1994: Regional climate change scenaria. In *Climate and Agroclimatic Variability in Central and Southeastern Europe* (ed.: *N.R. Delazios)*. University of Thessaly, Volos, Greece, pp. 53-81.

Pap, J.M. and *Fröhlich, C.*, 1999: Total solar irradiance variations. *J. Atmos. Sol.-Terr. Phys. 61*, 15-24.

Pruppacher, H.R and *Klett, J.D.*, 1997: *Microphysics of Cloud and Precipitation*. Kluwer, Dordrecht, The Netherlands.

Reid, G.C., 1999: Solar variability and its implications for the human environment. *J. Atmos. Sol.-Terr. Phys. 61*, 3-14.

Roble, R G. and *Dickinson, R.E.*, 1989: How will changes of carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? *Geophys. Res. Lett. 16*, 1441-1444.

Ross, R.J. and *Elliott, W.P.*, 2001: Radiosonde-based Northern-Hemisphere tropospheric water vapour trends. *J. Climate 14*, 1602-1612.

Rottman, G., 1999: Solar ultraviolet irradiance and its temporal variation. *J. Atmos. Sol.-Terr. Phys. 61*, 37-44.

Svensmark, H. and *Friis-Christensen, E.*, 1997: Variations of cosmic ray flux and global cloud coverage – a missing link in solar-climate relationships. *J. Atmos. Sol.-Terr. Phys. 59*, 1225-1232.

Taubenheim, J., 1969: *Statistische Auswertung geophysikalischer und meteorologischer Daten*. Akademische Verlagsgesellschaft Geest& Portig K.-G., Leipzig.

Tinsley, B.A., 1996: Solar wind modulation of the global electric circuit and apparent effects on cloud microphysics, latent heat release and tropospheric dynamics. *J. Geomagn. Geoelectr. 48*, 165-175.

Tinsley, B.A and *Heelis, R.A.*, 1993: Correlations of atmospheric dynamics with solar activity: evidence of a connection via the solar wind, atmospheric electricity and cloud microphysics. *J. Geophys. Res. 98*, 10375-10384.

Yearbooks of the Central Institute of Meteorology of the Hungarian Meteorological Service. Upper-Air Observations, Budapest, 1962-1994.

BOOK REVIEW

Houghton, J., 2002: **The Physics of Atmospheres. Third Edition.** Cambridge University Press, Cambridge. 320 pages, 135 figures.

It is not easy to explain all the basic principles of a scientific area to non-professionals, and it is especially a challenging task in the case of meteorology because of the stereotypes haunting the "science of weather-wises". Even some experts of physics and mathematics believe that the science of the atmosphere lacks established theoretical foundations. These facts will no doubt give a lot to the importance of meteorological textbooks dedicated to non-meteorologists.

The first version of *The Physics of the Atmospheres* was published by Professor Houghton in 1976. It was recommended for students of physics both undergraduate and graduate. The material of the book had been collected since 1958 at the University of Oxford, where Dr. Houghton served as Lecturer, later Reader, and since 1976 as Professor of Atmospheric Physics. Professor Houghton's monograph became very soon an internationally acknowledged volume of the scientific literature of meteorology. The second edition was published in 1986, and in 2002, the Cambridge University Press published a third, revised and extended edition.

Even in the first edition, Professor Houghton drew attention to the increase of interest in the atmospheric sciences. The book gives fundamental material on atmospheric radiation, radiation transfer, thermodynamics, statics, and the dynamics of the lower, middle, and upper atmosphere, turbulence, and atmospheric waves as well. The second edition was extended by chapters on general circulation, numerical modeling, observation, and climate, highlighting the modern developments and the increasing interest in these topics. The publication of the third edition was motivated by the large scientific and political interest in the possibility of an anthropogenic climate change on a global scale. Substantially, new chapters have been added to the monograph – on predictability and climate change. In the new chapters the author gives the basic physical foundations for these "hot topics", and discusses all questions addressed to the meteorological community in our days.

The Physics of Atmospheres can be recommended not only for undergraduate and graduate students of physics and meteorology, but also for physicists, meteorologists, geographers, and anyone interested in the foundations of meteorology determining the possibility of atmospheric and, especially, climate predictions.

Gy. Gyuró

GUIDE FOR 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

in three identical printed copies including all illustrations. 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 part: 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 double spacing with wide margins on one side of an A4 size white paper. 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 a *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 in camera-ready form in separate sheets with their numbers and captions given below them. Good quality laser printings are preferred.

The text should be submitted both in manuscript and in electronic form, the latter on diskette or in E-mail. Use standard 3.5" MS-DOS formatted diskette or CD for this purpose. MS Word format is preferred.

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

Information on the last issues: http://omsz.met.hu/irodalom/firat_ido/ido_hu.html

Published by the Hungarian Meteorological Service

Budapest, Hungary

INDEX: 26 361

HU ISSN 0324-6329