Consortium for Small-Scale Modelling

Newsletter

July 2006

No. 6

www.cosmo-model.org

Editors: Ulrich Schättler (DWD); Andrea Montani (ARPA-SIM); Massimo Milelli (ARPA-Piemonte)
Printed at Deutscher Wetterdienst, PO. Box 100465, 63004 Offenbach am Main, Germany

Table of Contents

1 Introduction

This is the sixth Newsletter of the Consortium for Small-Scale Modelling (COSMO) and the first one with a different format. The main purpose of this Newsletter is to serve as proceedings from the 7th COSMO General Meeting, which has been held in Zürich, Switzerland, from 20-23 September 2005. In addition, there are some contributions (not presented at the General Meeting) from "external" users of the LM, who report about their work.

Because of the large number of contributions for this issue, we had to waive all the other information we usually give within the Newsletter, e.g. on the COSMO structure and the Operational Applications of the LM. These can be obtained within a short time from the COSMO Web Page `http://www.cosmo-model.org`.

Figure 1: Participants of the 7th COSMO General Meeting in Zürich

In 2005, the format of the General Meeting had been changed. To give the contributing scientists more time to present their work, parallel Working Group sessions have been held on the first day. Most articles in this Newsletter are write-ups from the corresponding presentations. To reflect the Working Group structure, the different contributions for this Newsletter are also grouped together under the corresponding Working Groups.

There are some other important changes to the structure of the COSMO activities. The procedure to define the work plans of the Working Groups has been revised to improve the efficiency of the COSMO cooperation. To reduce the number of Work Packages (some lasting for several years) and to emphasize important work, so-called *Priority Projects* have been introduced, to which every COSMO member must contribute 2 persons per year. These projects have been initiated to ensure that resources are available for necessary work to be done. It is planned to run them with as few as possible but as much as necessary formalism, to control the work done. More information about the Priority Projects can be found on the Web soon.

We would like to thank all the scientists who submitted articles to the Newsletter. Special thanks go to the external contributors, the AWI Bremerhaven, the DLR Oberpfaffenhofen, the CLM Community and the Universities of Frankfurt, Mainz and München, who provided papers about their work done with the LM. We would like to encourage also other groups to document their work, e.g. in form of a short progress summary or a longer report, to be included in the next Newsletters.

Finally we want to apologize for the late publishing of this Newsletter, which was planned to do much earlier. A couple of reasons could be given as excuse, but we are sure everybody understands that there is always too little time for such supporting activities. Nevertheless, we try to fasten up things for the next issue.

Ulrich Schättler `Ulrich.Schaettler@dwd.de`
Andrea Montani `amontani@arpa.emr.it`
Massimo Milelli `Massimo.Milelli@arpa.piemonte.it`

Tomographic Determination of the Spatial Distribution of Water Vapor Using GPS Observations

M. Troller[1], A. Geiger[1], E. Brockmann[2], J.-M. Bettems[3], B. Bürki[1], H.-G. Kahle[1]

[1] *Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology, Zürich, Switzerland*

[2] *Swiss Federal Office of Topography, Geodetic bases and permanent networks, Wabern, Switzerland*

[3] *Federal Office of Meteorology and Climatology, Zürich, Switzerland*

1 Introduction

GPS meteorology is a promising technique to estimate the total amount of water vapor in the troposphere on a continuous basis using satellite navigation systems, (e.g. Bevis et. al., 1992). With GPS tomography, it is aimed at determining also the vertical distribution of water vapor in the troposphere with a high temporal resolution (e.g. Flores et. al., 2000; Hirahara, 2000; Kruse, 2001; Troller, 2004). In this study, hourly resolution has been investigated, which represents a major improvement compared to time resolution obtained from operational ballon soundings.

2 Principle of GPS Tomography

GPS signals are significantly influenced by the atmosphere, especially the ionosphere and troposphere, along their path from the satellite to the GPS antenna. The satellites are transmitting at two different carrier frequencies in the radio L-band. As the ionosphere is a dispersive medium in the radio frequency range, its influence can be eliminated by composing a linear combination of the two GPS carrier frequencies. The remaining effect is caused by the delay of the signal due to the refractivity of the troposphere. Furthermore, this effect can be subdivided into a so-called dry and wet part, the latter being proportional to the integrated precipitable water vapor (e.g. Hofmann-Wellenhof et. al., 2001).

The main advantages of GPS are that

1. measurements can be carried out on a continuous basis at relatively low expense
2. GPS is an all-weather observing system, allmost insensitive to clouds, and the effect of liquid water can usually be neglected (Elgered, 1993).

The wet propagation delay Δ_{wet}^{PD} of a radio signal from the satellite to the receiver antenna is defined as:

$$\Delta_{wet}^{PD} = \int_{antenna}^{satellite} (n_{wet} - 1)\, ds = 10^{-6} \int_{antenna}^{satellite} N_{wet} ds \qquad (1)$$

n_{wet} represents the refractive index due to water vapor and N_{wet} the wet refractivity whereby latter is defined as $N_{wet} = 10^6 \cdot (n_{wet} - 1)$; ds is the length of a ray path element. The ray bending effect can be neglected, as the cutoff angle of 10 degrees (Mendes, 1999) has been applied throughout the presented evaluation.

However, the wet path delay reflects the integrated amount of water vapor, only. To achieve the spatial distribution, a tomographic approach was investigated. In the tomographic approach, a discretization of the atmosphere with a voxel model is used. The number of layers and the horizontal and vertical size of voxels are correspondingly defined. Ellipsoidal boundary surfaces of the layers account for the earth's curvature. At the horizontal boundaries in each layer open voxels are added, i.e., these voxels reach ad infinitum. Consequently, no rays are neither completely nor partially outside the model. Within each voxel, the wet refractivity N_{wet} is introduced as an unknown constant. Discretizing (1), the wet slant path delay from satellite r to receiver antenna p as observed is related to the unknowns (refractivity N_{wet}) as follows:

$$\Delta^{PD,r}_{wet,p} = 10^{-6} \cdot \sum_{i=1}^{k} N_{wet,i}\, \Delta s_{i,p}^{\ r} \tag{2}$$

where, $\Delta s_{i,p}^{\ r}$ represents the length of the path through voxel i, and k the number of voxels. Wet slant path delays can be derived from several remote sensing techniques such as GPS, water vapor radiometry and solar spectrometry. In the current approach, we focus on GPS double difference path delays. By following the well established concept of double differencing (e.g. Beutler et. al., 2001), we obtain from (2) the double difference wet slant path delays $\Delta^{2,PD,rs}_{wet,pq}$ of the satellites r and s and the stations p and q:

$$\Delta^{2,PD,rs}_{wet,pq} \;=\; (\Delta^{PD,r}_{wet,q} - \Delta^{PD,r}_{wet,p}) - (\Delta^{PD,s}_{wet,q} - \Delta^{PD,s}_{wet,p}) \tag{3}$$

The following approach has been applied to retrieve the double difference path delays: GPS data have been processed using the Bernese GPS Software (Beutler et. al., 2001) yielding GPS zenith path delays $\overline{\Delta^{PD}}$ and double difference phase residuals $\Delta^2\Phi$ as processing output. The double difference path delays $\Delta^{2,PD,rs}_{pq}$ of the satellites r, s and the stations p, q are reconstructed from the GPS-derived zenith path delays $\overline{\Delta^{PD}_p}, \overline{\Delta^{PD}_q}$, which have to be mapped back to the corresponding elevations using the corresponding mapping functions (airmass factors) $m(el^r_p), m(el^r_q), m(el^s_p), m(el^s_q)$. Furthermore, the double difference phase residual $\Delta^2\Phi^{rs}_{pq}$ is added:

$$\Delta^{2,PD,rs}_{pq} \;=\; \overline{\Delta^{2,PD,rs}_{pq}} + \Delta^2\Phi^{rs}_{pq} \tag{4}$$

where:

$$\begin{aligned}\overline{\Delta^{2,PD,rs}_{pq}} \;=\; & (\overline{\Delta^{PD}_q} \cdot m(el^r_q) - \overline{\Delta^{PD}_p} \cdot m(el^r_p)) - \\ & (\overline{\Delta^{PD}_q} \cdot m(el^s_q) - \overline{\Delta^{PD}_p} \cdot m(el^s_p))\end{aligned} \tag{5}$$

Using ground meteorological measurements, the zenith dry path delay can be determined by applying the formula of Saastamoinen (Saastamoinen, 1972; Troller, 2004). Subsequently, the double difference dry slant path delay has been constructed using the same approach as in (5). Finally, the wet part of the double difference slant path delay $\Delta^{2,PD,rs}_{wet,pq}$ is extracted by subtracting the dry part from the total amount.

The number of traversing rays per voxel depends on the geometry defined by the number of visible satellites, the distribution of the ground stations and the size of the voxels. GPS-tomography usually generates a partly ill-posed problem in the sense that only a portion of the unknowns (refractivity per voxel) can be determined whereas the other part is under-determined. Additional information is necessary to solve the equation system. In this approach, the uppermost layer is bounded by 8.000 m and 15.000 m lower and upper level, respectively. This allows for the assumption of a priori values of $N_{wet} = 0\,ppm$ in that level. In special cases, a priori values can be introduced for other voxels. This procedure is described in Section 4. In addition, the voxels are mutually coupled by using a realistic covariance function to limit the variation of the difference in refractivity of neighboring voxels. However, only direct neighboring voxels are correlated to allow for steep refractivity gradients.

Extensive simulations using various weather conditions show the feasibility of this approach and indicate the need of double difference measurements with noise not exceeding 5 mm (Troller et. al., 2002). By properly treating the GPS data a double difference noise limit of 4-5 mm can be achieved in most cases, thus allowing for a tomographic solution with sufficient accuracy.

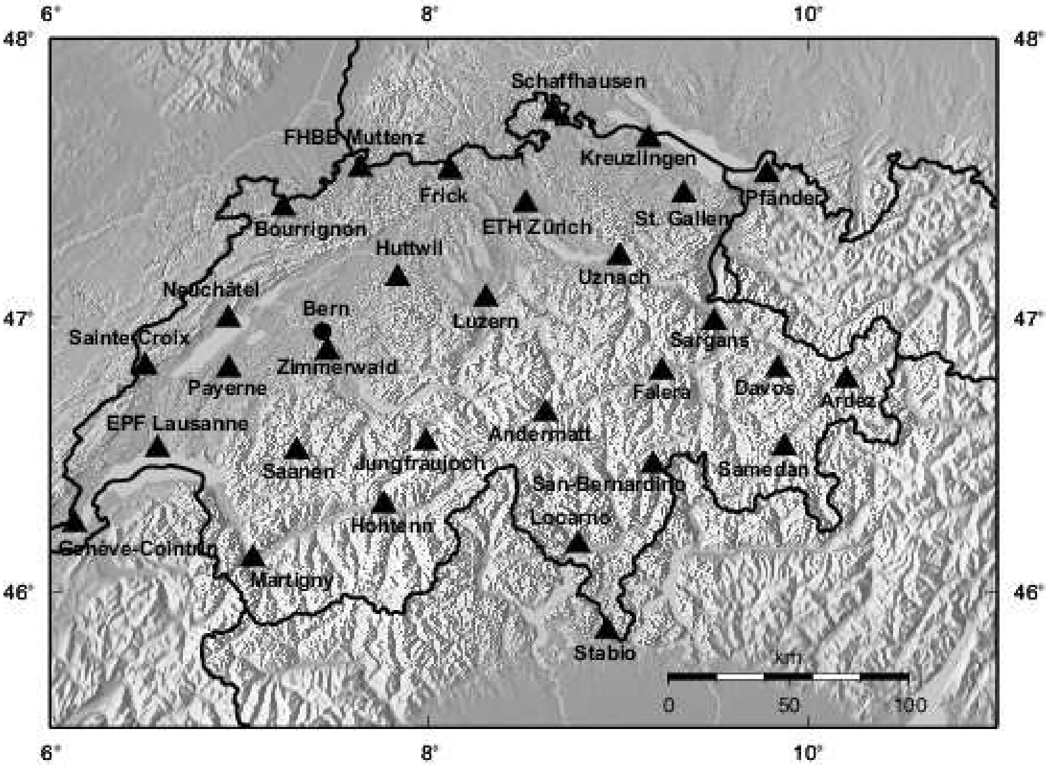


Figure 1: GPS stations of the Swiss Permanent GPS Reference Network AGNES (swisstopo). The network contains 30 AGNES stations, with a height distribution from 400 - 3.600 m. For the automated processing, 20 EUREF stations and 23 stations from other networks are included. The figure shows the stations of the Swiss territory only.

3 Description of the Experiment

The observations have been carried out in Switzerland. The data used are collected in the Swiss permanent GPS reference network (AGNES) of swisstopo. It covers the entire Swiss territory with a dense horizontal and vertical resolution (Fig. 1). Large height differences between the GPS stations are suitable for an accurate tomographic solution. An automated near real-time processing provides hourly means of zenith total delays, which are then used to derive instantaneous values of GPS path delays.

Data of the meteorological network ANETZ of MeteoSwiss (SMA, 1985) is used to decompose the total delays in its wet and dry part. ANETZ contains a total of 72 well distributed ground stations covering Switzerland and allows an accurate splitting of the zenith total delay.

A period of one week in November 2002 was chosen for the investigations. Rapid weather changes, heavy rainfall, clear conditions and sunny periods occurred during that week.

The voxel model above the Swiss territory is chosen with 6×3 voxels in horizontal (voxel side lengths $\sim 50\,km$) and 16 layers up to 15.000 m (Fig. 2). Water vapor profiles are retrieved on an hourly basis.

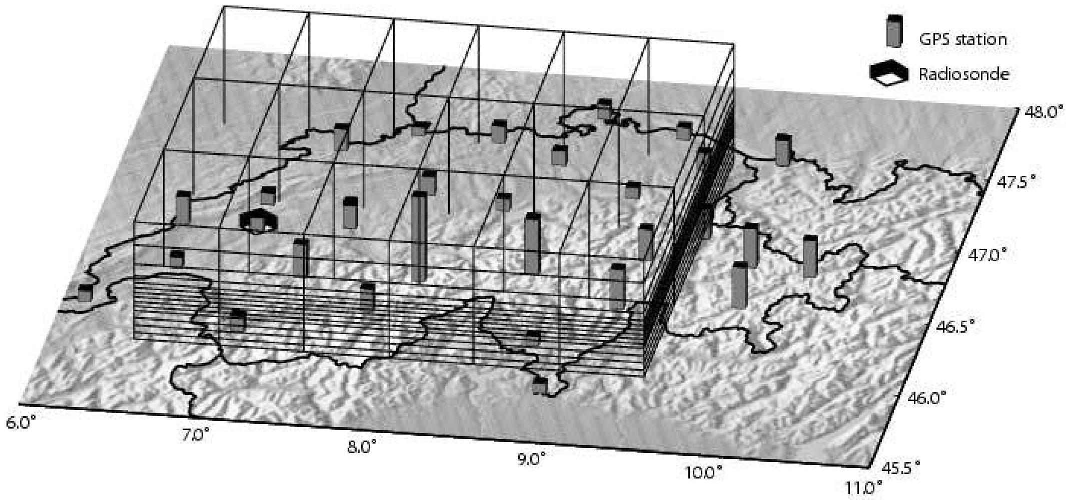


Figure 2: 3D view of the evaluation perimeter. The tomographic voxel model consists of 16 layers up to 15.000 m height. The borders of the layers are set at 0 m, 200 m, 600 m, 900 m, 1.200 m, 1.400 m, 1.600 m, 1.800 m, 2.000 m, 2.200 m, 2.400 m, 2.700 m, 3.200 m, 4.000 m, 5.000 m, 8.000 m, 15.000 m (the figure shows layers up to 5.000 m height only). Each layer contains 6 voxels in longitude and 3 voxels in latitude (spacing 0.5°) plus 22 outer voxels (not shown on the figure). The GPS stations are shown as column according to their station height. Radiosondes are launched from the radiosonde station Payerne (MeteoSwiss), shown as cuboid.

The tomographic results are compared with data of the operational high resolution NWP model of MeteoSwiss (aLMo).

4 Profile Determination and Evaluation

Double difference residuals and hourly means of zenith path delays are determined using the Bernese software package (Beutler et. al., 2001). Reconstructed double-difference wet slant delays are then introduced into the tomographic software package AWATOS.

Two types of tomographic profiles are determined, containing different constraints. On one hand, *AWATOS Correlation* includes inter-voxel constraints between all neighboring voxels. Compared to a double-difference observation, the constraints are down-weighted by a factor of 50^2 (regularization factor = $\frac{1}{2500}$). In addition, an a priori wet refractivity of zero is assigned to the uppermost layer (8.000-15.000 m) with a regularization factor of $\frac{1}{900}$. On the other hand, *AWATOS ANETZ* contains the same constraints as *AWATOS Correlation* and in addition one a priori wet refractivity value for each voxel lower than 2000 m. These refractivity values are based on a collocation and interpolation procedure (COMEDIE software package, Troller et. al., 2000) using the operationally available meteorological ground station data (ANETZ, MeteoSwiss).

The least-squares adjustment of the tomographic equation system yields the matrices of the cofactors of the unknown parameters from the inverted normal matrices (e.g. Mikhail, 1976). The comparison of the two cofactor matrices of the two solutions reveals the impact on quality of the two different types of solution. The a priori calculation takes no measurements into account. It assesses the geometry and the weighting of the measurements, only. The solution *AWATOS Correlation* shows a slight increase of the precision with height. Regarding the solution *AWATOS ANETZ*, a significant improvement of the square root of cofactors is visible in the first 2.000 m height but also in the higher voxels (Fig. 3).

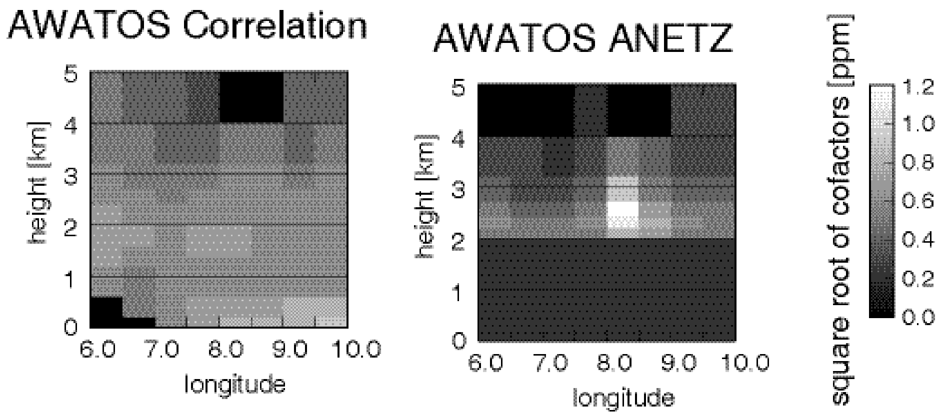


Figure 3: Cross section of the square root of cofactors at latitude 46.75°. The two top layers (5.000 - 15.000 m height) are not shown in the figures. (a) represents the situation of the solution *AWATOS Correlation* and (b) *AWATOS ANETZ*. It is seen, that if a priori refractivity is introduced, the precision increases significantly.

In Fig. 4, two examples of tomographic profiles, aLMo data and radiosonde profiles are displayed. Furthermore, the profile obtained with COMEDIE is plotted. The lowermost 2.000 m of the latter profile are used to constrain the solution *AWATOS ANETZ*. The comparison of the profiles confirms the conclusions made of the matrices of cofactors: The degree of agreement of the individual profiles is varying. *AWATOS ANETZ* fits always more accurately to aLMo than *AWATOS Correlation*. The 22 radiosonde launches during the investigation week at station Payerne have been used to perform a comparison to the tomographic profiles and the aLMo model (Table 1). The conclusion of Fig. 4 can be verified, the mean rms of *AWATOS ANETZ* is half as large as of *AWATOS Correlation*. Fig. 5 shows the mean rms of the tomographic profiles compard to the radiosondes as function of the height. The impact of the a priori refractivity in the first 2000 m height is clearly visible.

The comparison of aLMo to the radiosonde profiles is also documented in Table 1. The agreement is within 2.6 ppm (refractivity units) at station Payerne, where radiosondes are available to compare the aLMo model.

As radiosondes are available every six or twelf hours and at the station Payerne only, the aLMo data with an hourly resolution have been used to compare the tomographic solutions over entire Switzerland. Thus, a total of 3024 profiles are compared to aLMo for the one-week measurements (Fig. 6). In Fig. 4, two profiles with different levels of agreement, mainly depending on the atmosphere's actual state are shown. This behavior can be confirmed by

Table 1: Statistical analysis of the tomographic profiles and aLMo compared to the radiosondes. The analysis contains 22 profiles. The accuracy of the two tomographic solutions varies significantly. The mean rms of AWATOS Correlation is approximately double as the coresponding value of AWATOS ANETZ. Furthermore, the latter has no significant mean offset.

	AWATOS Correlation	AWATOS ANETZ	aLMo[1]
mean offset	-6.3	-1.4	0.5
mean rms	12.7	5.1	2.6
mean σ	11.0	4.9	2.6

[1] aLMo data for the lowermost voxel (200 - 600 m height) are not available.

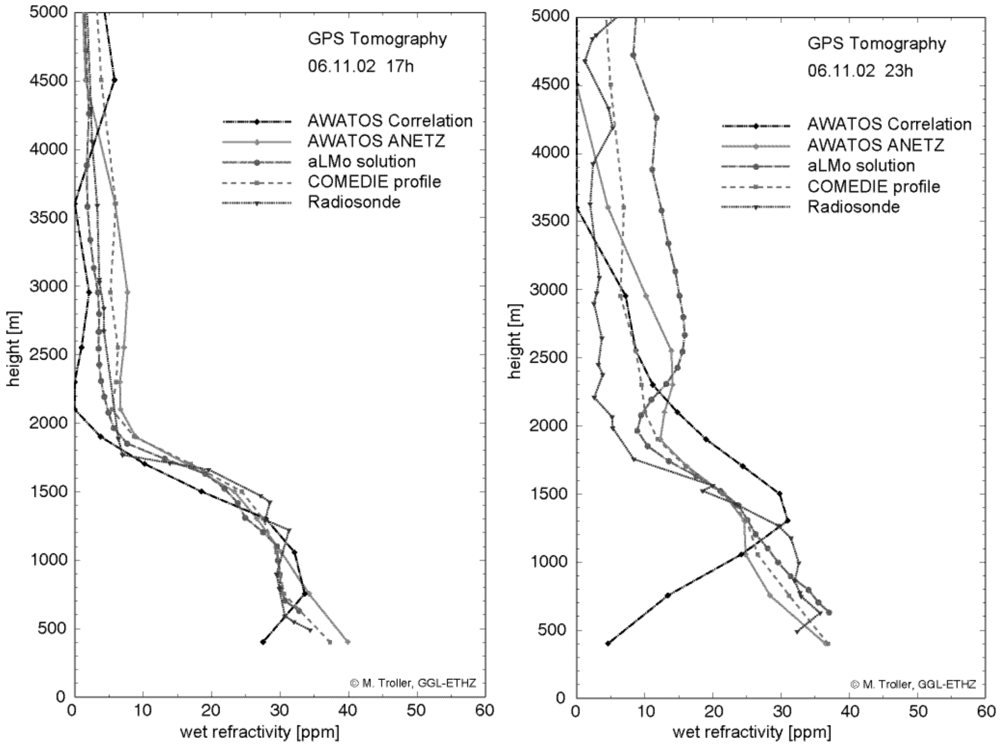


Figure 4: Wet refractivity profiles of the two tomographic solutions, the radiosonde ascending at the station Payerne (MeteoSwiss), data of the numerical weather model aLMo (MeteoSwiss) and the COMEDIE profile. (a) shows a profile with a steep refractivity gradient at 1.800 m height. All solutions match accurately together on the order of 10 ppm. (b) shows a situation with an inhomogeneous decrease of refractivity with height. *AWATOS Correlation* underestimates the wet refractivity in the lowest 1.000 m height. However, also the other tomographic solution as well as the radiosonde show differences to the aLMo solution.

inspecting the time series analysis of Fig. 6. Periods with large rms on November 3, 6 and 9, coincide with atmospheric situations rapidly changing from heavy rain to dry periods and contrariwise. During these time periods, the differences between GPS-tomography and aLMo are larger. This may mainly be caused by the temporal averaging during each data assimilation period, thus neglecting gradients in time for the assimilation interval.

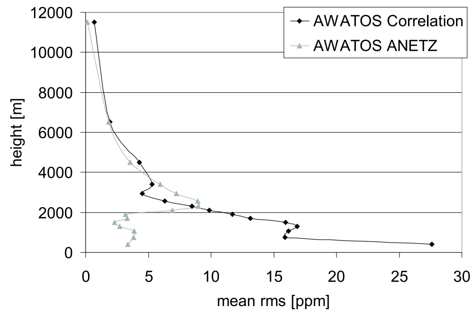


Figure 5: Mean rms of the two tomographic solutions compared to the radiosondes as a function of the height. The plot shows the comparison of the 22 radiosonde launches during the investigation week at the station Payerne. Generally, the rms is decreasing with increasing height. The impact of the a priori profiles on the first 2000 m height from COMEDIE is clearly visible in the solution AWATOS ANETZ.

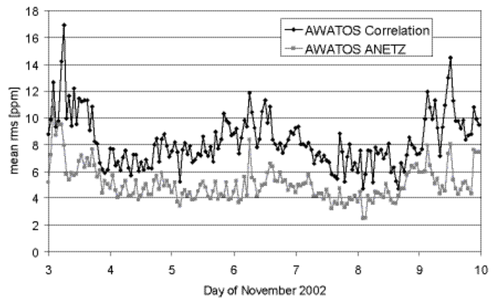


Figure 6: Time series of the mean of all profiles compared to aLMo. Overall, *AWATOS ANETZ* has a smaller rms than *AWATOS Correlation*. Three jumps are visible on November 3, 6 and 9.

5 Conclusions

The tomographic approach was successfully performed to process GPS data for the determination of 4 dimensional water vapor data.

Regularization methods are necessary to achieve a stable tomographic solution. Inter-voxel constraints and a priori refractivity information allow usually to achieve a high accuracy. An overall rms of about 5-8 *ppm* (refractivity units) or 0.8-1.3 $\frac{g}{m^3}$ absolute humidity compared to aLMo has been reached, depending on the tomographic solution. Using a temporal resolution of one hour, the accuracy remains approximately stable during the whole week, at least for the solution *AWATOS ANETZ*. Only during rapidly changing atmospheric situations, the rms is slightly decreasing. Further investigations showed (Troller, 2004) that the accuracy achieved at station Payerne, can be expected in whole Switzerland.

References

Beutler, G., Bock, H., Brockmann, E., Dach, R., Fridez, P., Gurtner, W., Hugentobler, U., Ineichen, D., Johnson, J., Meindl, M., Mervart, L., Rothacher, M., Schaer, S., Springer, T., and Weber, R., 2001. Bernese GPS software version 4.2. Edited by U. Hugentobler, S. Schaer and P. Fridez. Astronomical Institute, University of Berne.

Bevis, M., Businger, S., Herring, T., Rocken, C., Anthes, R., and Ware, R., 1992. GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *J. of Geophysical Research*, 97(D14), pp. 15787-15801.

Elgered, G. (1993). *Atmospheric Remote Sensing by Microwave Radiometry*, chapter Tropospheric Radio-Path Delay from Ground-based Microwave Radiometry, pp. 215–258. John Wiley and Sons, Inc.

Flores, A., Ruffini, G., and Rius, A., 2000. 4d tropospheric tomography using GPS slant wet delays. *Ann. Geophysicae*, 18, pp. 223-234.

Hirahara, K. 2000. Local GPS tropospheric tomography. *Earth Planets Space*, 52, pp. 935-939.

Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J., 2001. GPS Theory and Practice. Springer, Wien, New York.

Kruse, L., 2001. Spatial and temporal distribution of atmospheric water vapor using space geodetic techniques. Geodätisch-geophysikalische Arbeiten in der Schweiz, Volume 61, Schweizerische Geodätische Kommision.

Mendes, V.B., 1999. Modeling the neutral-atmosphere propagation delay in radiometric space techniques. PhD dissertation, University of New Brunswick.

Mikhail, E.M., 1976. Observations and least squares. University Press of America.

Saastamoinen, J., 1972. Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites. In *Henriksen, S.W.; Mancini, A.; Chovitz, B.H. (ed.): The use of artificial satellites for geodesy, Geophys. Monogr. Ser., AGU*, pp. 15:247–251.

SMA, 1985. Charakteristiken der ANETZ-Daten. In *Beiträge zum ANETZ-Daten Kolloquium vom 17. April 1985 in Zürich.*, p. 56. Schweizerische Meteorologische Anstalt Zürich.

Troller, M., 2004. GPS based Determination of the Integrated and Spatially Distributed Water Vapor in the Troposphere. Diss ETH No. 15513.

Troller, M., Bürki, B., Cocard, M., Geiger, A., and Kahle, H.-G., 2002. 3d refractivity field from GPS double difference tomography. *Geophys. Res. Let.*, 29(24), pp. 2149–2152.

Troller, M., Cocard, M., and Geiger, A., 2000. Modellierung 4 dimensionaler Refraktionsfelder zur Berechnung von Weglängen-Korrekturen bei Satellitenmessungen. In *Simulation raumbezogener Prozesse: Methoden und Anwendungen*, 9, *IfGI prints.*

Improved Methods for Snow-Cloud Separation Using Multi-temporal Meteosat-8 SEVIRI Imagery

Martijn de Ruyter de Wildt

Swiss Federal Institute of Technology, Institute of Geodesy and Photogrammetry, Zürich, Switzerland
and
MeteoSwiss, Zürich, Switzerland

Abstract

Meteosat-8 is the first geostationary satellite that possesses channels at all bandwidths that are of use for snow mapping. It therefore offers new possibilities for multi-temporal snow mapping, as well as for snow mapping in short time intervals, which is for example required for numerical weather prediction models. The spectral capabilities of Meteosat-8 allow an optimal spectral separation of clouds and surface snow cover, whereas the high temporal frequency introduces temporal information that might be used in the classification process. In this work we describe an algorithm that uses a new spectral feature and temporal for snow mapping and cloud detection.

1 Introduction

Snow cover influences several processes that occur at or near the earth's surface. It affects the exchange of energy and moisture between the surface and the atmosphere and is an important aspect of the hydrological cycle. Furthermore, snow cover extent is an indicator of climatic change and affects many human activities. Near real-time information about the surface snow cover is therefore important for studies and applications in many disciplines. This is particularly the case for Numerical Weather Prediction (NWP) models, which are initialised several times per day and require the latest information about the state of the atmosphere and the surface, including snow cover. A valuable tool for detecting snow cover is remote sensing, because it allows us to monitor large areas of the earth at regular time intervals.

A regularly encountered problem in remote sensing of snow is the sometimes similar spectral appearance of snow and clouds. In general, clouds have a similar reflectance as snow, and when they also have the same brightness temperature and phase (i.e. ice clouds), it can be difficult to distinguish them from snow with spectral information alone. Some authors have therefore used the spatial context of satellite pixels to detect clouds, but these methods are based on spatial inhomogeneity and are mainly suited for differentiating between cloud types and for detecting clouds over homogeneous surfaces. To differentiate between clouds and natural land surfaces, which often are quite inhomogeneous themselves, such methods are less suitable.

Another type of contextual information that can be used to classify satellite images is of temporal nature. Image classification that uses temporal information is generally referred to as (digital) change detection. In the literature, this notion generally refers to changes at the earth's surface, but there seems no reason why change detection methods might not be used for detecting clouds. Mostly, change detection involves two images of the same scene acquired at different dates, but in a number of applications temporal series of images are used. To the latter category, which clearly offers more opportunities for detecting change as it uses more images, belong temporal trajectory analysis and temporal compositing. In temporal trajectory analysis, the temporal trajectory of a pixel is compared with a predefined trajectory, whereas in temporal compositing a composite is made from a series of individual images by retaining those pixels that satisfy a certain criterion. Trajectory analysis of high-frequency images is used for background estimation in video surveillance and photogrammetry. There, the task is to detect and/or remove objects that temporarily obscure the background, which is in fact comparable to detecting moving clouds over a static surface. However, background estimation requires the obscuring object to be present in only a few of a series of images, whereas clouds often cover large areas and single pixels can be cloud-covered during large parts of an observation period.

In remote sensing, change detection is generally used to study processes that occur at rather long time scales of months to years. This is the case for change detection over land, as well as over the oceans (e.g. sea ice). As a consequence, in most of these studies polar orbiting sensors have been used, as these have repeat times of hours to weeks and offer a wide range of spatial and spectral resolutions. In contrast to land surfaces and oceans, clouds often display a dynamic behaviour at time scales of minutes to hours, and only geostationary satellites have a frequency that is high enough to monitor this behaviour. However, unlike many polar orbiting sensors, geostationary platforms did until recently not possess all spectral channels that are required for optimal spectral separation of snow and clouds.

In 2002, the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), launched the first of a new series of geostationary satellites, called Meteosat Second Generation (MSG). This new satellite, Meteosat-8 (MSG-1) bridges the gap between polar orbiting sensors with good spectral resolution and geostationary sensors with high temporal frequency. It thus offers an unprecedented dataset of spectral and temporal information, which can be used to detect clouds over cold regions and to map surface snow. Here we describe an algorithm that uses temporal trajectory analysis in conjunction with pixel-based spectral classification to detect clouds and to map surface snow cover. This algorithm is intended for delivering real-time snow cover data to the operational mesoscale NWP model of MeteoSwiss, the Alpine Model (aLMo).

2 Data

Meteosat-8 is currently situated at 3.4° western longitude at an altitude of 36.000 km. It carries the Spinning Enhanced Visible and Infrared Imager (SEVIRI), which has improved spectral, spatial and temporal resolution with respect to its predecessors on board of the previous Meteosat satelites. It continuously monitors the entire earth disk with a frequency of 15 minutes. SEVIRI has twelve spectral channels, most of which measure radiation from the surface. Only the water vapour absorption channels, at 6.25 and 7.35 μm, contain no information about the surface at all. The ozone absorption channel (9.66 μm) measures radiation from the troposphere and the surface and is also sensitive to ozone concentration in the lower stratosphere. The CO_2 absorption channel (13.4 μm) mainly measures radiation from the troposphere with only a small contribution from the surface. Some CO_2 is also detected by the 3.9 μm channel, which slightly overlaps with one of the CO_2 absorption bands. Channel 12 is a high resolution visible (hrv) channel, which has a spatial resolution of 1 km at the sub-satellite point. All of the other channels have a spatial resolution of 3 km at the sub-satellite point. The region of interest in this study is the model domain of aLMo, which corresponds to western and central Europe. The spatial resolution of Meteosat-8 over this area is 1.5 to 2 km for the hrv channel and 5 to 6 km for the other channels.

For testing and validation we selected a three-day period in March 2004. During the selected period, March 8$^{\text{th}}$ till March 10$^{\text{th}}$, all mountainous regions of Europe were covered with snow, as well as large parts of Central, Eastern and Northern Europe. The weather was variable, with low pressure activity over Central Europe and the Mediterranean. Clouds, some of them containing ice particles, bare land and snow covered land were therefore all well represented over Europe, making this period suitable for testing a snow mapping algorithm.

3 Pre-processing

The Meteosat-8 data that we use are provided in Level 1.5 Native Format. These consist of raw satellite counts, which need to be calibrated and converted into reflectances (r) and brightness temperatures (BT). We also apply a sea mask to the data, which is based on the SRTM30 Digital Elevation Model (DEM) of NASA and the United States Geological Survey. This global DEM has a horizontal resolution of 1 km, and we resampled it to the Meteosat-8 grid with bicubic interpolation. Because of the large viewing angles of Meteosat-8 over Europe, it is furthermore necessary to ortho-rectify the data.

3.1 Correction for atmospheric effects and anisotropy

Further pre-processing of the measured reflectances involves correcting for the influence of the atmospheric and for anisotropy of reflection at the surface. The atmospheric influence depends on the state of the atmosphere and on the solar and satellite viewing angles, and can be described with Radiative Transfer Models (RTMs). However, state-of-the-art RTMs are not very reliable for solar zenith angles over 70°, whereas snow cover is mainly present during the winter season, when the sun remains rather low above the horizon. Also, RTMs require the atmospheric aerosol content and atmospheric profiles of water vapour, ozone and CO_2, which are generally not available. Water vapour can be provided by NWP models, but always with some degree of uncertainty. We therefore only apply one correction for all angular effects, including those caused by atmospheric radiative transfer and by anisotropic reflection at the surface. For this correction, a semi-empirical model that describes bi-directional surface reflectance is used. The model has five coefficients that can be related to the Normalised Differential Vegetation Index (NDVI). From Meteosat-8 SEVIRI data, the NDVI can be computed for the low resolution channels. For the hrv channel, we use the downscaled low-resolution NDVI values.

The temporal behaviour of the reflectances at two cloud-free locations, illustrating the angular effects, is shown in Fig. 1. The first location was snow-free, whereas the second was partially snow-covered. At both locations the reflectances display a peak in the early afternoon, which corresponds to the hot spot in the BRDF. Also, the angular effects increase with wavelength at both locations. This is most clearly the case over the snow-free location, whereas it is less obvious over snow-cover. Over snow the 1.6 μm reflectance is much lower than the other reflectances, so that the absolute effect of anisotropy is not very large. The relative effect, however, is largest for the 1.6 μm reflectance, and this increase with wavelength corresponds with the observations of other authors.

For each of the solar channels (channel 1, 2, 3 and 12), the coefficients c_i were derived by tuning the bi-directional reflectance model to the cloud-free pixels in the available satellite images. Because the true value of r (the actual hemispherical reflectance of the surface) is unknown, r is set equal to 1 in the tuning procedure. The resulting BRDF's can then be used to bring all observed reflectances to a reference viewing and illumination geometry. Over vegetated surfaces without snow cover the model performs well, but for pixels that contain

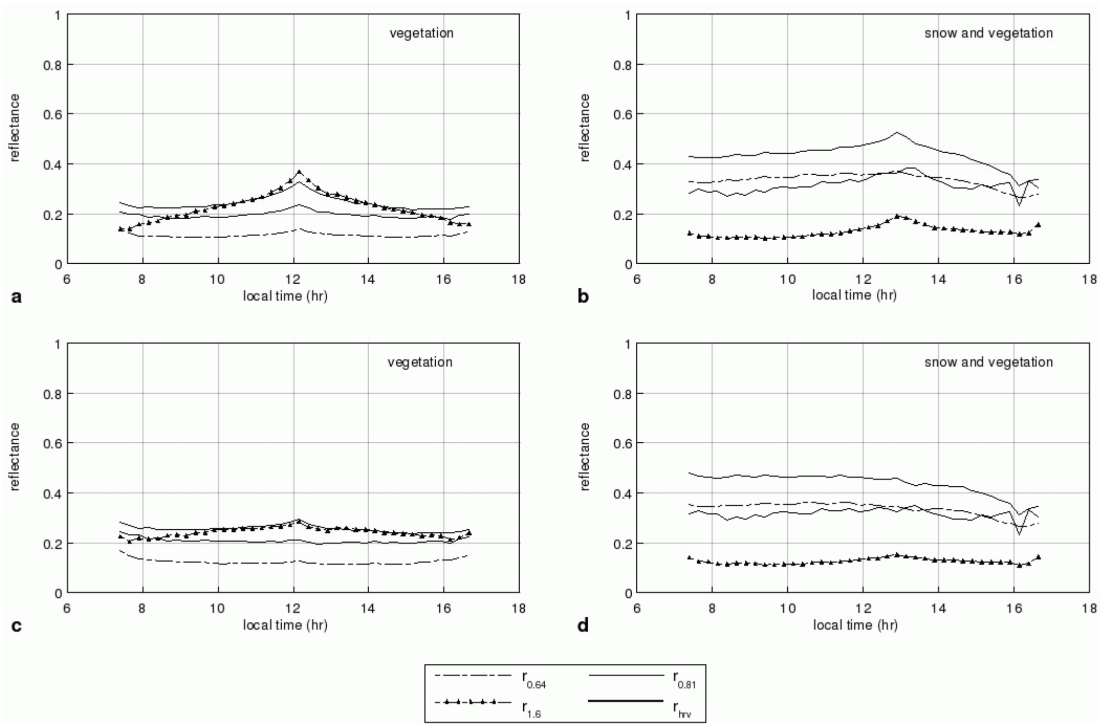


Figure 1: Reflectances as a function of local time at two cloud-free locations on March 10$^{\text{th}}$, 2004. Only data with corresponding solar zenith angles below 75° are shown. Temporal profiles are shown for a vegetated location (a) and for a vegetated location that was partially covered with snow (b). Temporal profiles that have been corrected for angular effects are shown in c and d, respectively.

snow cover, mostly mixed with vegetation in our images, it could not be adequately tuned. We therefore determined the tuning coefficients only for pixels without snow, which display a smooth dependence of the coefficients on NDVI. These dependencies can be described by simple polynomial or exponential functions and we let these functions approach zero for very low NDVI (corresponding to pixels with snow). This approach means that over mixed pixels, we only apply BRDF's for vegetation and not for snow. At the locations for which the time series are shown in Fig. 1a and b, the BRDF's that we use remove a large part of the temporal variation (Fig. 1c and d). The most variation remains at the partly snow covered location (Fig. 1d), which may be caused by anisotropic reflectance of the snow. It could also be due to melting of snow and an increasingly lower snow fraction during the afternoon, as suggested by the temporal behaviour of the reflectances. If the snow fraction would remain constant throughout the day and the temporal behaviour was only caused by angular effects, all reflectance channels would follow a similar pattern. Here, however, the visual reflectances are constant in the morning and decrease during the afternoon, whereas the near infrared reflectance is slightly increased during the afternoon. Both effects are well explained by a lower snow fraction.

4 Classification

4.1 Temporal features

Examples of the temporal behaviour of clouds are shown in Fig. 2. The first location (Fig. 2a and b) was covered with ice clouds in the morning, as indicated by the low 1.6 μm reflectance,

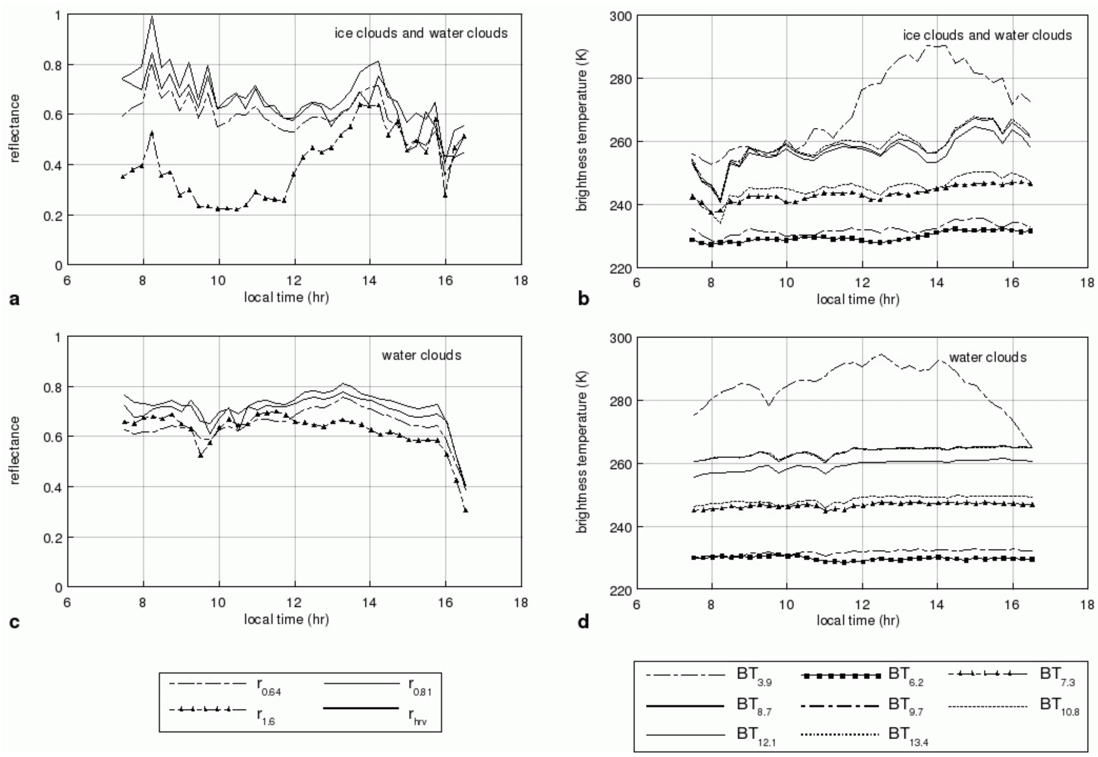


Figure 2: Reflectances and brightness temperatures as a function of local time at two cloudy locations on March 10$^{\text{th}}$, 2004. Only data with corresponding solar zenith angles below 75° are shown. Temporal profiles are shown for a location that was covered with ice clouds (a, b) and for a location that was covered with ice and water clouds (c, d).

and in the afternoon with water clouds. There is considerable temporal variability in most of the spectral channels. The second location (Fig. 2c and d) was also covered with clouds, but here the cloud cover consisted entirely of water clouds that were fairly homogeneous in space and time. Consequently, the overall temporal variability is lower. At both locations, the infrared absorption channels, in which information from the surface and the lower atmosphere has been (partly) filtered out, display less variation than the other infrared channels. As a measure of temporal variability we use the standard deviation in time. We found that it is also useful to take same temporal information from the eight surrounding pixels into account by averaging the standard deviation in time over each block of nine pixels. Although this classifier considers the eight surrounding pixels, it does not quantify spatial variability, and can therefore be regarded as a quasi three-dimensional classifier. To illustrate the usefulness of the temporal standard deviation, scatter plots of this classifier against the near-infrared reflectance are shown in Fig. 3. Pixels that represent water clouds, which have a high near-infrared reflectance, appear on the right hand sides of these plots. Ice clouds and snow have low infrared reflectances (on the left), whereas mixed clouds and snow-free surfaces display intermediate values. The cluster in the bottom left corners could be identified as surface snow, and the cluster next to it as snow-free surface. These pixels display a low temporal variability. Mixed clouds and ice clouds generally display larger temporal variabilities, whereas water clouds, represented by the cluster on the right-hand side of both plots, display low variabilities. This is especially the case for the brightness temperature (Fig. 3b).

We found the optimal number of time steps for separating clouds from surface snow by calculating a divergence parameter, which indicates the ability of a feature to separate two classes. When the temporal standard deviation at one pixel is used, for most channels the

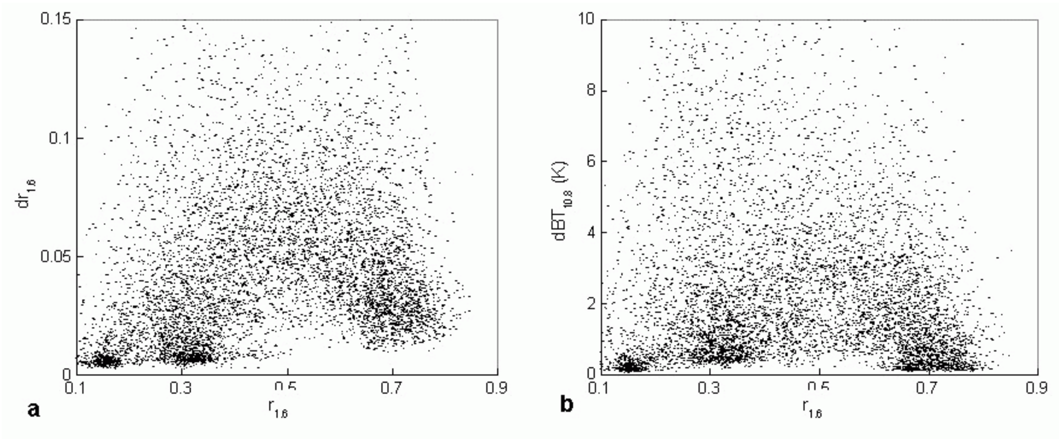


Figure 3: Scatter plots of the quasi 3-D temporal standard deviation against the near-infrared reflectance on March 10^{th}, 2004, 12:12. Displayed are scatter plots for the 1.6 μm reflectance (a) and for the 10.8 μm brightness temperature (b).

divergence is largest when 7 successive images are used. Shorter time series obviously include too little temporal information. On the other hand, the temporal variability of pixels that are cloud-free in the current image and clouded several time steps earlier or later (or vice versa), will increase for longer time series. Such pixels will be classified as clouds when the temporal variability at the beginning and/or at the end of the time series is large. When the temporal standard deviation of the eight surrounding pixels is also taken into account, the divergence improves and only 5 successive image are needed for the best results.

Table 1: The twelve channels of the SEVIRI instrument on board of Meteosat-8.

Channel	Spectral band μm			Description
	centre	min.	max.	
1	0.635	0.56	0.71	visual
2	0.81	0.74	0.88	visual
3	1.64	1.50	1.78	near infrared
4	3.90	3.48	4.36	solar + terrestrial infrared
5	6.25	5.35	7.15	infrared (water vapour absorption)
6	7.35	6.85	7.85	infrared (water vapour absorption)
7	8.70	8.30	9.10	infrared
8	9.66	9.38	9.94	infrared (ozone absorption)
9	10.80	9.80	11.80	infrared
10	12.00	11.00	13.00	infrared
11	13.40	12.40	14.40	infrared (CO_2 absorption)
12	0.75	0.60	0.90	high resolution visual broadband

We found higher divergences for the reflectance channels than for the infra red channels, which is caused by the often static behaviour of water clouds (Fig. 3b). When we omit all cloudy pixels with a near infrared reflectances above 0.5, we find comparable divergences for all channels (Table 1). The highest divergence is found for the reflectance in the hrv channel, which detects the most detailed information. The other three solar channels also display high divergences whereas the infrared channels display somewhat lower divergences. As expected, the lowest values are found for the water vapour absorption channels at 6.2 and 7.3 μm. These channels measure mid-atmospheric water content, and changes in this quantity can occur independently of cloudiness at lower levels. These channels are therefore not suitable for detecting clouds. Significantly larger divergences are found for the other absorption channels at 9.7 and 13.4 μm, which detect significant amounts of information from low atmospheric levels. Thus, we use all channels except the water vapour absorption channels for temporal detection of clouds.

4.2 Spectral features

SEVIRI has channels in more spectral bands that can be used for cloud detection and snow mapping than any other currently available sensor, apart from MODIS on board of NASA's Terra and Aqua satellites. An existing cloud mask for the MODIS snow product uses $r_{0.64}$, $r_{1.6}$, $BT_{3.9} - BT_{10.8}$, $BT_{13.4}$ and the Normalised Differential Snow Index (NDSI), which equals $(r_{0.64} - r_{1.6})/(r_{0.64} + r_{1.6})$. In addition to these features we also use the 3.9 - 13.4 μm thermal difference. The latter feature often reveals water clouds, but it does not detect optically thick ice clouds, as illustrated by Fig. 4a. Water clouds have a large $BT_{3.9} - BT_{10.8}$ and appear bright, whereas unclouded regions, which have a small or even negative $BT_{3.9} - BT_{10.8}$, appear dark. Clouds with a high ice content, some of which are indicated in Fig. 4a, appear as dark as or somewhat brighter than snow-covered areas. A similar picture arises when we compute $BT_{3.9} - BT_{13.4}$ (Fig. 4b), but now many ice clouds tend to be darker than snow.

This difference can be attributed to CO_2 absorption, which occurs in the 3.9 and 13.4 μm channels. It reduces the amount of observed radiation, leading to lower observed brightness temperatures. The effect is far larger at 13.4 μm than at 3.9 μm and consequently, $BT_{3.9} - BT_{13.4}$ is always strongly positive. No CO_2 absorption takes place at 10.8 μm so that $BT_{10.8}$ is much higher than $BT_{13.4}$ and $BT_{3.9} - BT_{10.8}$ always smaller than $BT_{3.9} - BT_{13.4}$. Furthermore, more CO_2 absorption takes place when the atmospheric path length is longer, so that in general it has a stronger cooling effect over the surface than over clouds. The difference between $BT_{3.9} - BT_{10.8}$ and $BT_{3.9} - BT_{13.4}$ is therefore smallest for high altitude pixels. A scatter plot of the two brightness temperature differences (Fig. 5) clearly shows two bands of pixels, one corresponding to ice clouds and one corresponding to surface pixels. The use of both features should thus improve the separation of ice clouds and snow, which can be visualised by computing the ration between them. A plot of this ratio (Fig. 4c) clearly reveals many ice clouds that are not detectable with each separate brightness temperature difference (Fig. 4a and 4b).

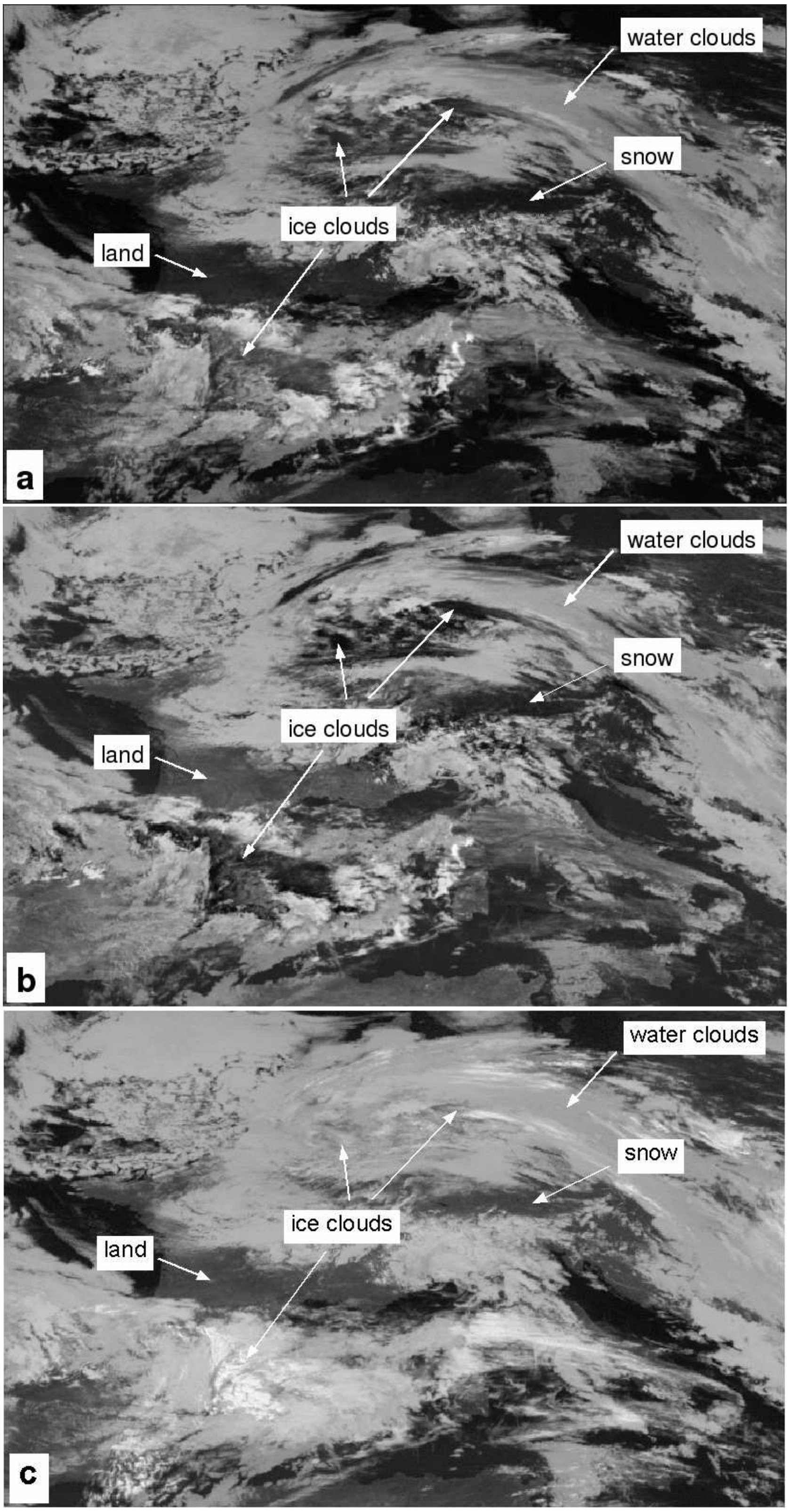


Figure 4: Normalised Meteosat-8 brightness temperature differences over the study area on March 10$^{\mathrm{th}}$, 2004, 12:12 UTC. Shown are $BT_{3.9} - BT_{10.8}$ (a), $BT_{3.9} - BT_{13.4}$ (b) and $(BT_{3.9} - BT_{10.8} - 5)/(BT_{3.9} - BT_{13.4})$ (c). This scene could be visually classified by making use of the multi-spectral and multi-temporal information that is available.

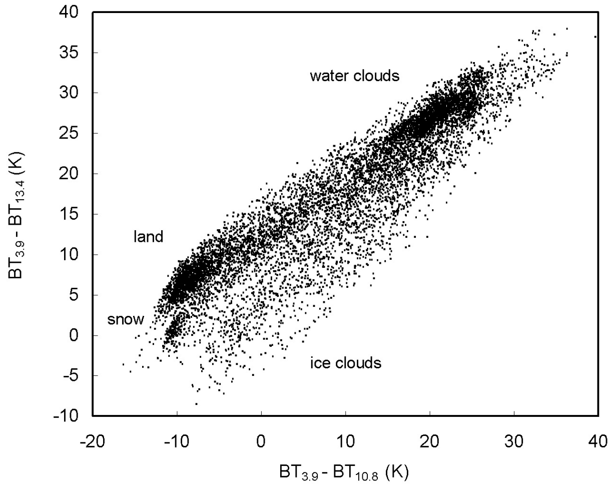


Figure 5: Scatter plot of $BT_{3.9} - BT_{13.4}$ against $BT_{3.9} - BT_{10.8}$ for the same scene as shown in Fig. 3.

4.3 Classification method

In remote sensing of snow cover, often threshold based classification trees are used. The spectral properties of clouds and snow are well known, which makes it straight-forward to choose threshold tests and to set values for the thresholds. This classification method generally gives good results and is easy to implement. Here we choose another standard classification method, namely maximum likelihood classification. With this method each pixel is assigned to the class for which the conditional probability of the pixel is highest. The advantage of this method is that it can adequately classify pixel distributions like the one shown in Fig. 5. Also, maximum likelihood classification gives probabilities in stead of rigid values (e.g. snow or cloud), which can be used for assigning quality flags to the pixels.

When we assume that the features are normally distributed, the conditional Probability Density Functions (PDFs) are given by the multi-variate normal distribution. For each images class, this distribution is described be the mean feature values and the feature covariance matrix. We chose four classes to which pixels can be assigned: snow-free land, snow, ice clouds and water clouds. Although we are not interested here in distinguishing between different cloud types, we do make the division between ice clouds and water clouds in order to improve the separation of clear and cloudy pixels. Clouds containing ice particles differ in appearance from water clouds in several ways: $r_{1.6}$, $BT_{3.9} - BT_{10.8}$ and $BT_{3.9} - BT_{13.4}$ are lower (see Fig. 3 and 5) when ice particles are present. There is also a difference in mean temporal variability between ice clouds and water clouds (Fig. 3). A simple threshold of for $r_{1.6}$ is used for differentiating between training areas for ice clouds and water clouds.

In order to determine for all classes the means of the features and the covariance matrices, we first classified all images with a simple threshold-based classification using only the spectral features. The values of the thresholds were chosen such that the best classification results were obtained, as could be subjectively judge by visual inspection. For the threshold-based classification we use a simple scheme that includes the ratio between $BT_{3.9} - BT_{10.8}$ and $BT_{3.9} - BT_{13.4}$, which we found very suitable for detecting clouds (Fig. 4c). Pixels that are not classified as cloudy in this way, are checked for the presence of snow by a second suite of tests.

The threshold-based classification missed some clouds that were misinterpreted as snow, but the overall quality was judged acceptable. The classification results of all images were then used as training areas for computing the means and covariances for all features and classes. Then, the images were classified again, now with the maximum likelihood method. In the new results the misclassified clouds were no longer present and these results were therefor used for a final determination of the means and covariances.

5 Results

For analysing the performance of the algorithm, we focus on March 10^{th}, 2004, 12:12 UTC. An RGB image for this image which we found very useful for visual inspection is shown in Fig. 6a. Most image classes are clearly discernible from each other in this RGB combination. Only the colour ranges of snow (red) and ice clouds (red/pink) slightly overlap. For visual discrimination between these two classes one can view animated time-series of this RGB combination, which visualise both the spectral and the temporal component. On the project web-site (`www.photogrammetry.ethz.ch/research/snow/index.html`) examples of such animated time series are available. Before we discuss the classification result of this image, it is worthwhile to have a look at the conditional probabilities. Three of these conditional probabilities can be combined into one RGB image, and this is shown for snow-free surfaces, snow and water clouds in Fig. 6b. The three classes very clearly emerge in different colour groups, and even snow and ice clouds are now clearly distinguishable from each other. Pixels that have comparable conditional probabilities for all three classes appear in grey tones. In this RGB combination, this is the case for water surface and for some ice clouds, which both appear in white.

The result for the full maximum likelihood classification, i.e. including both spectral and temporal features, is shown in Fig. 7a. There are no false positives, i.e. no snow is detected where it is not present. However, a few false negatives occur: these are transparent and/or sub-pixel clouds over snow that are missed. When a binary snow map is requested the latter aspect is an advantage, as more snow is detected, but when fractional snow cover is to be derived pixels should be completely cloud-free. The influence of each type of information, spectral and temporal, upon the classification result can be investigated by using only one of these types of information for the classification. In Fig. 7b the result is shown for the case when only spectral information is used. Now, more snow is detected in some places and less snow in other. The reason for this is that some clouds display very low temporal variability, which lowers the conditional probability for clouds. These clouds may therefor not be detected when temporal information is used, whereas they are detected when no temporal information is used. For clouds with high temporal variability the opposite may be true. When only temporal information is used to mask clouds with the maximum likelihood method, we found that many clouds are missed. The cause of this poor performance is that whereas all temporal features display low values over cloud-free pixels, the opposite is not necessarily true over cloudy pixels. Slightly better results are obtained when we obtain a temporal cloud mask in each single channel and then stack all single channel cloud masks. A more substantial improvement is obtained when in each single channel the decision rule is changed in favour of the conditional probability for clouds. For example, when pixels are masked as cloud when the conditional probability of the temporal classifier is twice as high for clouds as for cloud-free surfaces, we obtain the temporal cloud mask that is shown in Fig. 7c. This cloud mask can be used to check all pixels that were classified as snow in the spectral classification (Fig. 7b) for high temporal variability. As a result (Fig. 7d) many mixed pixels near cloud edged that were previously classified as snow, are now classified as cloud.

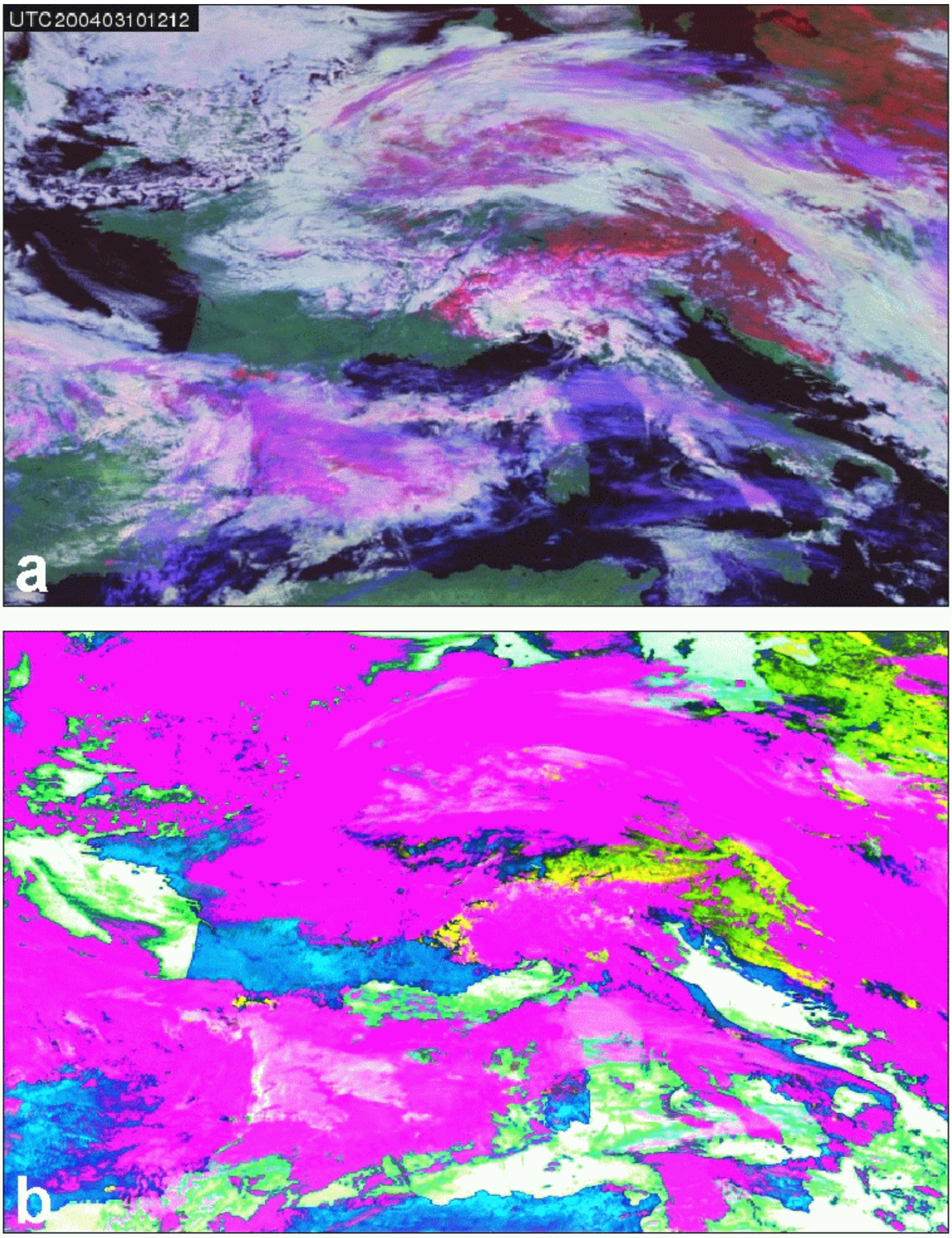


Figure 6: Meteosat-8 RGB images of central Europe, acquired on March 10^{th}, 2004, 12:12 UTC. (a) combination of $r_{0.64}$ (red), $r_{1.6}$ (green) and $(BT_{3.9} - BT_{10.8})/(BT_{3.9} - BT_{13.4})$ (blue). Snow-free surfaces are green, snow is read, water clouds are white, optically thin ice clouds (cirrus) tend to be purple and optically thick ice clouds are pink or red. (b) combination of the conditional probabilities for snow-free surfaces (red), snow (green) and water clouds (blue). Here, snow-free surfaces appear blue, snow appears green or yellow and clouds appear pink. Note that water appears black in (a) and white in (b).

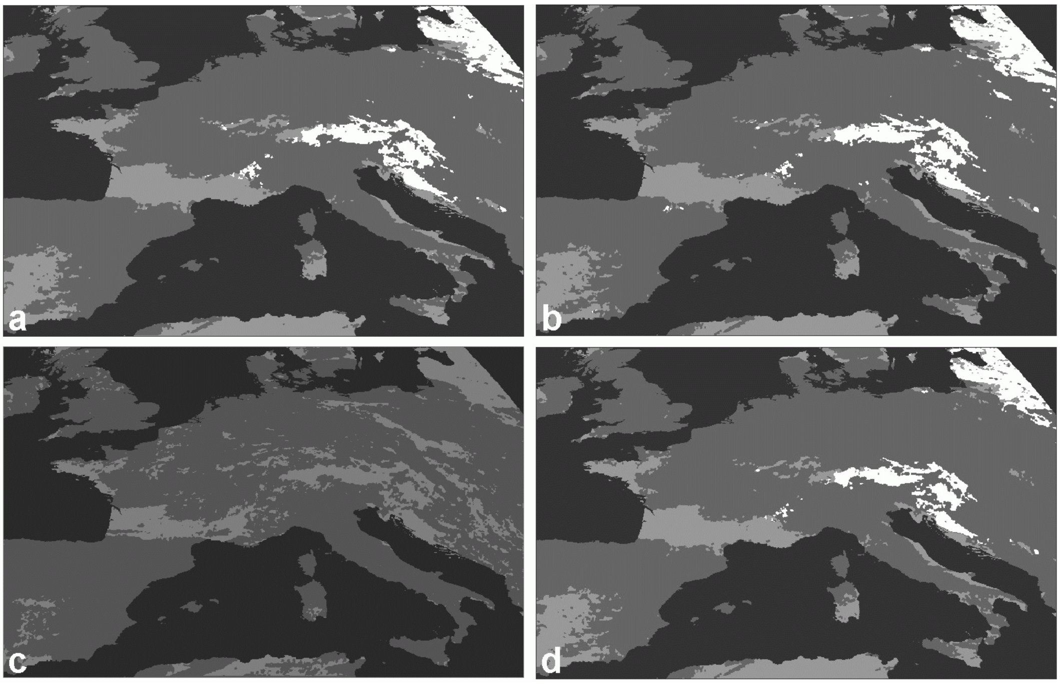


Figure 7: Classification results for March 10^{th}, 2004. (a) maximum likelihood classification with spectral and temporal features; (b) maximum likelihood with only spectral features; (c) temporal cloud mask, obtained from stacking all single channel cloud masks; (d) as (b), but now the temporal cloud mask of plot (c) is used to mask snowy pixels with high temporal variability.

6 Conclusions

With the spectral features that we use for maximum likelihood classification, all pixels that were classified as snow actually contained snow, as judged by visual inspection. Many of these pixels were of mixed type, representing both snow and snow-free land, and sometimes also transparent and/or sub-pixel clouds, mainly near cloud edges. This type of classification thus produces a liberal snow map, in the sense that it detects the highest amount of pixel where snow is to some extent present. No false positives and only some false negatives are present in this snow map. When temporal information is used to filter out pixels with high temporal variability, false negatives are removed and the snow map becomes more conservative (more mixed snow/cloud pixels are classified as clouds). The liberal snow map could be of use when one wants to obtain a binary snow map, i.e. a snow map that simply indicates whether snow is present or not. When a fractional snow map is required, the conservative snow map is more appropriate, because then the detected snow pixels are more likely to be cloud free and to contain only contributions from surface classes.

Apart from improving the detection of clouds during day-light, we anticipate that temporal information can also be used for cloud detection during the night. Of course, the solar channels can not be used during the night, but the temporal variability will still be measurable in the infrared channels. For the purpose of detecting surface snow cover however, which is not possible during the night, this is of no importance.

Sensitivity of the LHN Scheme to Non-Rain Echoes

Daniel Leuenberger[1], Andrea Rossa[2]

[1] *MeteoSwiss, Krähbühlstrasse 58, 8044 Zürich, Switzerland*
[2] *Centro Meteorologico di Teolo, ARPA Veneto, Via Marconi 55, 35037 Teolo, Italy*

1 Introduction

Radar-derived quantitative precipitation estimates (QPE) are becoming an increasingly important element in high-resolution numerical weather prediction (NWP). As such they complement conventional data like surface or upper-air observations. COSMO has chosen to use the Latent Heat Nudging (LHN) method (Jones and Macpherson, 1997; Leuenberger and Rossa, 2003; Klink and Stephan, 2005) to assimilate radar-derived QPE.

In a recent study Leuenberger (2005) systematically investigated the performance of the LHN scheme at meso-γ scale, both in an idealised setup and in the context of real cases. He found that LHN has considerable potential at the convective scale in that for an idealised supercell it successfully initialised the storm in a perfect environment and - to a lesser extent - in non-perfect environments in which low-level humidity or wind fields were altered. For the real case convective systems, a supercell and a squall line case, LHN was able to capture the salient features of the storms. Persistence of the assimilated systems in the subsequent free forecasts appeared to depend much on the instability of the environment into which the observed systems were forced.

Unlike conventional observations, radar data exhibit a highly variable quality, in that they are affected by a number of factors that limit their accuracy in estimating precipitation at the surface. In the context of assimilating radar-derived QPE in high-resolution NWP models this poses two salient questions, i.e. how is a specific assimilation scheme affected by errors in the observations, and how can such variable quality be accounted for?

This paper addresses the first question and presents a sensitivity study of the Latent Heat Nudging scheme to gross errors in the radar data, notably non-rain echoes. These include ground clutter returns and spurious signals due to anomalous propagation of the radar beam. Consideration is given to the dynamical response of the model to the continuous forcing of idealised and real signals during assimilation time, and to the performance of free forecasts started from the LHN analyses.

2 Methodology

2.1 Model and assimilation scheme

All simulations are conducted with the LM (Version 3.1) in an idealized mode. The parametrisation for grid-scale precipitation accounts for four categories of water (water vapour, cloud water, rain and snow), the mass fractions of rain water (q_r) and snow (q_s) are treated diagnostically. Vertical subgrid turbulence and the surface flux formulation are switched on, whereas cumulus parametrization, radiation and soil processes are switched off. The LHN scheme used in this study is described in Leuenberger and Rossa (2003).

All model integrations were uniformly performed on a 50 × 50 gridpoint domain. In the vertical, a stretched grid is employed composed of 60 levels and separated by 67 m near the ground and 2000 m near the model top at 23500 m. Above 11000 m a Rayleigh damping layer is used to absorb vertically propagating waves. In order to damp grid-scale noise, fourth-order numerical diffusion is applied. All simulations are integrated to 2 hours.

2.2 Setup of the sensitivity experiments

The basic atmospheric environment for the sensitivity experiments was chosen following Weisman and Klemp (1982) for the study of splitting supercell storms (Fig. 1a). They used a conditionally unstable thermodynamical profile and a moist, well mixed boundary layer with constant water vapour mixing ratio r with a reference value of $r = 12$ g/kg, yielding a lifting condensation level of ∼1500 m, a level of free convection of ∼1900 m, a level of neutral buoyancy of ∼10000 m and a Convective Available Potential Energy (CAPE) of ∼1200 J/kg. As a simplification for the present study the environmental wind was set to zero for most experiments. For selected experiments, the wind profile was set to exhibit a vertical shear of 20 m/s over the lowest 4000 m and constant wind aloft with no variations of the wind direction with height ($V = 0$). The lateral boundaries are relaxed towards the initial state throughout the whole simulation.

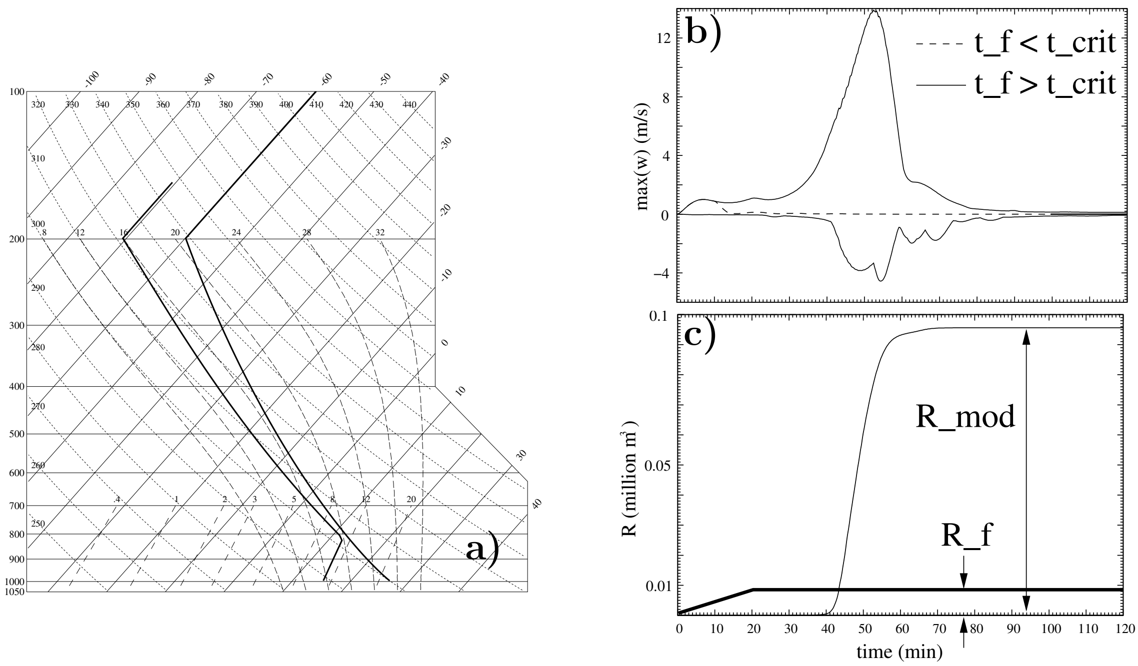


Figure 1: Panel a) shows the reference sounding following Weisman and Klemp (1982) used for the clutter experiments. The instability was varied by varying the boundary layer humidity. Here the profile for a maximum mixing ratio of 11 g/kg (resulting in a CAPE of 800 J/kg) is displayed. Panel b) displays the time evolution of the maximum up- and downdraft (m/s) during an individual assimilation and subsequent forecast experiment. The solid (dashed) line denotes an experiment in which the convective instability was (not) released (i.e. the forcing time is larger (smaller) than the critical forcing time). Panel c) shows the corresponding cumulated forcing (*R_f*) and resulting model precipitation (*R_mod*) (mm) for an experiment in which the convective instability is released with a corresponding ratio of roughly 10.

Ground clutter can be considered as one of the most important source of non-rain echoes, most of which is eliminated by appropriate clutter filters. However, in order to minimise eliminating real rain echoes, MeteoSwiss' clutter filter, for instance, leaves some 2% of the non-rain echoes in the data (Germann and Joss, 2004). This residual clutter often manifests as small-scale, quasi-static, medium to high intensity signals. On the basis of this, ground clutter is modeled for the purposes of this study as isolated, one-pixel signals of varying intensities I_f. These signals are assimilated during forcing times of various length (t_f). The boundary layer humidity was varied to obtain environments of various degrees of instability (see Tab. 1). An NWP model's numerical diffusion scheme is designed to act strongly upon one-pixel signals so that results obtained by these experiments can be taken as lower limits of the respective impact. I.e. in reality, and for larger non-rain echo areas, the impacts are expected to be more pronounced than what results from these experiments.

A large number of of experiments, i.e. 125, have been conducted, varying the clutter intensity I_f = 2, 5, 10, 20, 30, 50, 60 mm/h, the forcing times t_f from 2 min up to 2 hours. The varying degree of instability with the boundary layer moisture content results in differing lifting condensation levels and levels of free convection. Often, these levels are lower for environments of larger instabilities.

r_{max} (g/kg)	CAPE (J/kg)	LCL (m)	LFC (m)
10	400	1630	2840
11	800	1450	2400
12	1200	1290	2040
13	1700	1130	1730
14	2200	990	1470

Table 1: Specifications of the environments used for the numerical experiments. Values include the maximum water vapour mixing ratio in the boundary layer, CAPE, lifting condensation level and level of free convection.

3 Results

3.1 Description of a single experiment

Figure 1b,c) summarises the outcome of two individual experiments both conducted in a 1700 J/kg CAPE atmosphere, with a 10 mm/h clutter forced during 20 min (solid lines in panel b) and 6 min (dashed line). The model response to the applied forcing is depicted in terms of maximum up- and downdraft (panel b) and total accumulated precipitation (panel c). The larger forcing causes an updraft which reaches a strength of 1 m/s after about 7 min. Continuing the forcing out to 20 min does not increase the vertical velocity, but keeps it at this level even though a slight modulation is visible, indicating an interaction between the forcing and the developing model dynamics. However, after having switched off the forcing air parcels seem to have reached the level of free convection sometime between 20 and 30 min. Once this level is reached the instability present in the basic state is released, exhibiting values of the vertical velocity up to 13 m/s at $t = 52$ min. Substantial rain falls out beginning at $t = 41$ min. and stops when the system relaxes at $t = 60$ min. Total rainfall accumulates close to $0.1 \cdot 10^6$ m^3 whereas the precipitation equivalent of the forcing amounts to $0.01 \cdot 10^6$ m^3, i.e. the error given by the non-rain echo has been amplified by the assimilation scheme by a factor of close to 10. Note, that in this experiment downdrafts form of several m/s in response to the convectively driven updraft.

The second experiment in Fig. 1b,c), on the other hand, is an example in which the initial erroneous forcing is not sufficient to lift air parcels to their level of free convection. Hence the initially triggered vertical velocity gradually decreases without producing any precipitation and downdrafts.

3.2 Evaluation of the idealised clutter experiments

The impact of a particular non-rain echo on the assimilation and the subsequent forecast in an individual experiment can be measured as the ratio of the resulting model-produced precipitation R_mod and the precipitation equivalent of the total forcing R_f calculated from the product of the forcing time t_f and the clutter amplitude I_f (Fig. 1c). If the ratio R_mod/R_f is zero or much smaller than one, the effect of the spurious signal on the assimilation is negligible. If, however, the ratio is larger than one, the assimilation has amplified the error in the radar data. In the former case, the conditional instability present in the environment was not released by the applied forcing, i.e. it is too small to lift an air parcel to reach the level of free convection. For the latter, however, this level is eventually reached and the instability released. Consequently, the model-produced rain can be much larger than the forcing equivalent. The instability is accompanied by significant values of vertical velocity of the order of several tens of m/s as illustrated in Fig. 1b).

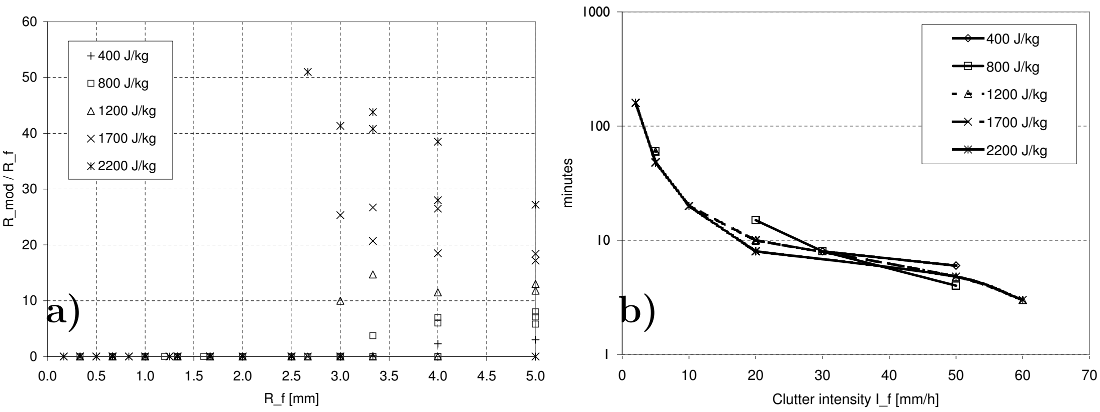


Figure 2: Panel a): Sensitivity of the Latent Heat Nudging scheme to ground clutter-like non-rain echoes (refer to Section 3.2). The symbols denote individual experiments in which a single-pixel forcing was applied during various periods of time in atmospheres of various instabilities. The x-axis denotes the total applied forcing (i.e. the product of intensity of the echo times the time over which it is applied), the y-axis the ratio of the resulting total model precipitation and the total forcing. Note, that even for relatively moderate instabilities amplification of the signal (i.e. ratio larger than 1) takes place after modest forcing. Panel b): Minimum time needed for the Latent Heat Nudging scheme to amplify a ground clutter-like non-rain echo for various instabilities (denoted by different symbols (refer to Section 3.2). The x-axis represents the clutter amplitude (mm/h), while the y-axis the critical forcing time (min) in logarithmic scale. Note, that for the higher clutter amplitudes it takes only a few minutes of forcing for the amplification to take place.

The ensemble of one-pixel clutter experiments with zero wind is evaluated in terms of the resulting amplification factors which are summarised in Fig. 2a). It becomes evident that even medium-intensity ground clutter signals can be dramatically amplified in unstable environments and, therefore, hamper the precipitation assimilation substantially. This occurs when the forcing induced by the spurious echo is sufficient to lift air parcels to their level of free convection. For instance, a clutter signal of 20 mm/h nudged during 10 min into an environment with a CAPE of 1200 J/kg, is amplified by a factor of 15, while in a 400 J/kg environment a 50 mm/h clutter amplifies by a factor of 3 after 8 minutes, i.e. even for relatively moderate instabilities amplification of the signal (ratio larger than 1) takes place after modest total forcing. Accompanying updrafts easily reach values between 10 and 20 m/s. This scatter plot suggests the existence of a threshold forcing, for the present configuration at R_f between 2.5 and 3 mm, above which air parcels do reach their level of free convection, and the instability is released. However, there are cases with a larger total forcing, e.g. 10 mm resulting from a combination of small clutter intensity and long forcing time, in which the level of free convection is not attained. This may partly be due to an interference of the LHN forcing with the model dynamics (such as numerical diffusion), when the convective system starts to develop in the model. Investigation of this is beyond the scope of the present study.

A slightly different way of representation is given in Fig. 2b), in which the critical forcing time t_crit, i.e. the minimum time for a given amplitude to reach amplification, is depicted for several degrees of instability. Again, for unstable environments even very small amplitude signals are amplified given sufficient forcing time. For high-amplitude signals dramatic error amplification is almost immediate, i.e. takes place after as little as a few minutes. For smaller values of CAPE and clutter amplitude, however, the assimilation scheme is able to dampen the error.

In the light of these results and given that real ground clutter amplitudes often reach, or even exceed, such amplitudes, a thorough clutter elimination in convectively unstable situations seems to be fundamental.

3.3 Real case example

In order to illustrate what can happen in real cases of clutter, Swiss Radar Network (SRN) data for a non-rain day are assimilated for a six hour period into the experimental setup of this study using the reference profile with CAPE=800 J/kg (Fig. 1a). The simulation was performed with the setting used by LR05, i.e. a model domain of 361 $\times$ 333 horizontal grid points, with a mesh size of 2.2 km and 45 vertical levels. The six hour accumulation of the resulting model precipitation (Fig. 3b) exhibits dramatic amplification of the original clutter signals. It is evident that regions of large coherent clutter amplify to larger intensities than pixel-sized signals, as the former are less dampened by the model's numerical diffusion scheme that acts primarily on the structures with sizes of the order of the gridlength. The problem is somewhat mitigated if the SRN data are run through a Shapiro type observation filter with length 4 (Shapiro 1975) (Fig. 3c). In addition, the presence of appreciable wind causes the precipitation resulting from the clutter assimilation to be exported to neighbouring regions, in which new convection can be triggered (Fig. 3d).

3.4 Anomalous propagation conditions

The signal resulting from anomalous propagation of the radar beam is another important source of non-rain echoes (e.g. Koistinen et al., 2004). In contrast to regular ground clutter, anomalous propagation clutter can be more coherent in space but more intermittent in time.

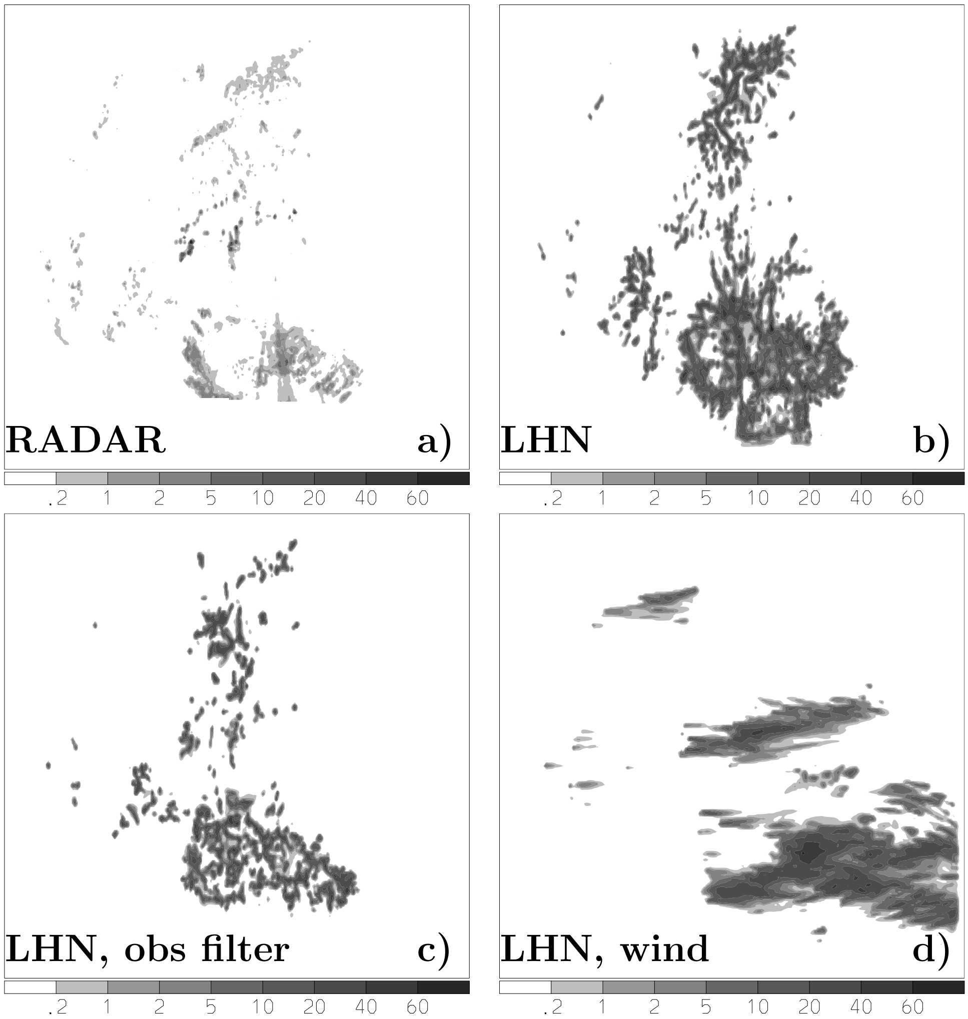


Figure 3: Examples for the assimilation of real clutter in a convectively unstable situation. Panel a) shows a six-hour accumulation (mm/h) of the Swiss Radar Network for a non-rain day, and panel b) displays the model precipitation resulting from a continuous forcing of the clutter by LHN. Panel c) is as b) except that the observed clutter is filtered. Finally, panel d) is as b) except that wind is added in the basic state. The domain has a size of $730 \times 800\,\text{km}$.

In Switzerland, conditions conducive to anomalous propagation are characterized by very stable conditions and often occur in concomittance with low-stratus, in which often very dry upper-level air tops the planetary boundary layer and the thermal inversion, thus giving rise to strong vertical refractivity gradients. Consequently, assimilating such clutter signals does not usually result in precipitation amplification, due to the absence of convective instability and sufficient moisture. However, the LHN forces the model in trying to match the model precipitation with the input signal. As the model does not produce precipitation the forcing is continued and may yield significant vertical circulations throughout the troposphere. Values for up-downdrafts can reach $7\,\text{m/s}$ and $-2\,\text{m/s}$, respectively (Fig. 4a). This spurious circulation may distort the dynamical fields locally and interact adversely with the mesoscale flow. In particular, substantial vertical mixing of the local model atmosphere can take place. Fig. 4b) illustrates how more humid air of the boundary layer is generously mixed into the dryer air of the free troposphere. This effect would be undesired, for instance, in the context of air pollution modelling, as critical pollution episodes are usually tied to strong inversions. However, this issue was not pursued further in this study.

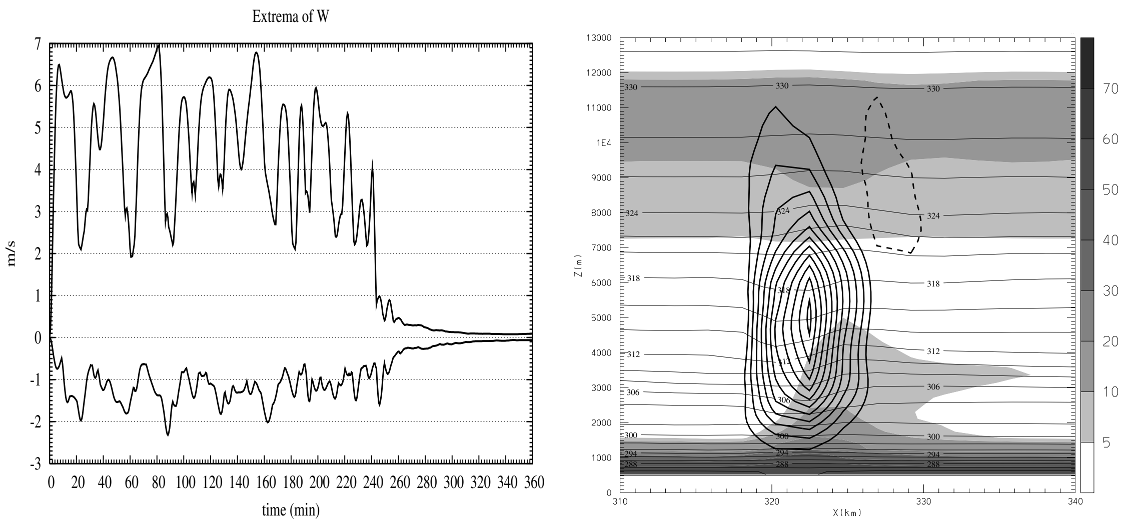


Figure 4: Panel a): Time evolution of the maximum up- and downdraft (m/s) during the assimilation experiment in conditions conducive to anomalous propagation of the radar beam. Positive values denote updrafts, negative values downdrafts. The forcing is applied during 4 hours. Panel b): Vertical cross-section through maximum updraft of the anaprop experiment at 2h into simulation time. Displayed are RH (in %, shaded), potential temperature (2K contour interval, thin lines) and vertical velocity (0.5 m/s contour interval, bold lines, solid updrafts, dashed downdrafts).

4 Summary and discussion

In this study the sensitivity of the Latent Heat Nudging (LHN) scheme to non-rain echoes was investigated by means of idealised experimentation with synthetic and real radar data. It constitutes one part of an effort to judge the LHN's aptness as an efficient and economic scheme for operational high-resolution rainfall assimilation. The main findings of this study are:

- non-rain, or clutter, echoes as small as one pixel can trigger the release of convective instabilities when forced by the LHN scheme;
- the resulting precipitation can be large compared to the original signal, i.e. factors 3 up to 50 have been found for moderate to high value of CAPE;
- the response of the model atmosphere to the forcing is very quick, i.e. on the time scale of convection (less than ten minutes for strong forcing to a couple of hours for moderate forcing);
- large, coherent areas of non-rain echoes pose a more stringent problem;
- filtering the input data can significantly mitigate the problem;
- non-rain echoes resulting from anomalous propagation of the radar beam in a low-stratus case over Switzerland, by virtue of the usually stable and dry conditions associated, are not conducive to error amplification. However, a strong spurious vertical circulation, along with undesired mixing, may be induced and adversely impact the mesoscale circulation.

A limitation of the study is that the impact of non-rain echoes is investigated in isolation. The negative impact found may be less dramatic in situations, when clutter is embedded in real precipitation echoes. Furthermore the LHN scheme used in this study does not contain the modifications of Klink and Stephan (2005), i.e. it is not compatible with the prognostic treatment of precipitation. However, given the highly idealised nature of the sensitivity experiments we expect the findings of this study to apply, at least qualitatively, also for the new version of the LHN scheme for the following reasons: Firstly the 1×1 clutter pixel is constant in time and space, secondly there is no horizontal wind in most of the experiments and thirdly the forcing time of the LHN scheme would not be much different if the vertically integrated precipitation flux would be taken as model reference precipitation.

Characterisation of the radar data quality for use in atmospheric data assimilation schemes is very important. It has been shown that errors can, in certain circumstances, be dramatically amplified and cause the QPF to deteriorate. Quality characterisation of radar data could, therefore, include at the pixel level some sort of probability for the signal to be rain in terms of a static clutter map of zero probability of rain and a dynamic estimate of varying amplitude. An assimilation scheme like LHN can include such information into a quality or weighting function as proposed by Jones and MacPherson (1977) and Leuenberger and Rossa (2003). It is conceivable to make the assimilation of pixels with non-zero probabilities for being spurious conditional on the prevailing atmospheric conditions. For instance, a pixel with a 50% probability of being real rain would be assimilated in stable to neutral environments, while rejected in highly unstable situations. The dialog between radar data producer and users is absolutely vital in this context.

References

Germann, U., and J. Joss, 2004: Operational measurement of precipitation in mountainous terrain. *Weather Radar. Principles and Advanced Applications*, P. Meischner, Ed., Springer Verlag, pp. 52–75.

Jones, C. D., and B. Macpherson, 1997: A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model. *Meteorol. Appl.*, 4, 269–277.

Klink, S., and K. Stephan, 2005: Latent heat nudging and prognostic precipitation, *COSMO Newsletter*, No. 5, available from http://www.cosmo-model.org.

Koistinen, J., D. B. Michelson, H. Hohti, and M. Peura, 2004: Operational measurement of precipitation in cold climates. *Weather Radar. Principles and Advanced Applications*, P. Meischner, Ed., Springer Verlag, pp. 78–114.

Leuenberger, D., 2005: High-resolution radar rainfall assimilation: Exploratory studies with latent heat nudging, Ph.D. thesis, Swiss Federal Institute of Technology, Zürich, Switzerland.

Leuenberger, D., and A. Rossa, 2003: Assimilation of radar information in aLMo, *COSMO Newsletter*, No. 3, available from http://www.cosmo-model.org.

Shapiro, R., 1975: Linear filtering. *Math. Comp.*, 29, 1094–1097.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, 110, 504–520.

Revised Latent Heat Nudging to cope with Prognostic Precipitation

Christoph Schraff, Klaus Stephan, and Stefan Klink

Deutscher Wetterdient, Offenbach am Main, Germany

1 Introduction

The prognostic treatment of precipitation (Gassmann 2002, Baldauf and Schulz 2004) used operationally in the LM tends to decorrelate the surface precipitation rate from the vertically integrated latent heat release and thereby violate the basic assumption of the Latent Heat Nudging (LHN) approach. This, and resulting problems have been shown by Klink and Stephan (2005), and they also suggested possible adaptations to the LHN scheme. More recent experiments have allowed to better specify preferable choices and parameters for at least some of the adaptations. Here, the specifications for the most important ones are briefly described, and results of recent experiments are shown. In the concluding remarks, the current status is summarized, and some remaining problems with LHN are outlined.

2 Major revisions to the LHN scheme

At horizontal model resolutions of 3 km or less, the prognostic treatment of precipitation allows the model to distinguish between updrafts and downdrafts inside deep convective systems. Compared to using the diagnostic precipitation scheme, it modifies both the 3-D spatial structure and the timing of the latent heating with respect to surface precipitation. Two revisions address spatial aspects and a third one an important temporal issue:

- In updraft regions at the leading edge of convective cells, very high values of latent heat release ΔT_{LHmo} occur often where precipitation rates RR_{mo} are low. Considering that

  $$\Delta T_{LHN} = (\alpha - 1) \cdot \Delta T_{LHmo} \qquad , \qquad \alpha = \frac{RR_{obs}}{RR_{mo}}$$

  high values of the scaling factor α and of the latent heat nudging temperature increments ΔT_{LHN} often occur. To mitigate this, the upper limit for α is reduced to 2 and the lower limit increased accordingly to 0.5 . In addition, the linear scaling $(\alpha - 1)$ is replaced by a logarithmic scaling $\ln(\alpha)$ in order to unbias the scheme in terms of adding or taking away absolute amounts of heat energy. The effective upper and lower scaling limits are then 1.7 and 0.3 respectively. This adaptation reduces the simulated precipitation amounts during the LHN.

- In downdraft regions further upstream in convective cells, high precipitation rates occur often where latent heating is weak or even negative in most vertical layers. In order to avoid negative LHN temperature increments and cooling where the precipitation rate should be increased (and vice versa), only the vertical model layers with positive simulated latent heating are used to compute and insert the LHN increments. These layers coincide very roughly with the cloudy (saturated) layers. This modification tends to render the increments more coherent and the scheme more efficient.

- Precipitation produced by the prognostic scheme will take some time to reach the ground where it is compared to the radar-derived surface precipitation rate. Thus, the conventional LHN scheme can notice only with some temporal delay when it has already initiated precipitation aloft, and it will continue to add (or take away) heat energy for some time when it is not required any more.
  Therefore, an immediate information on the precipitation rate already initialised is required, i.e. a sort of undelayed 'reference precipitation' RR_{ref} which is used merely to replace the delayed prognostic precipitation RR_{mo} in the computation of the scaling factor α. Deploying the diagnostically calculated precipitation rate (by an additional call to the diagnostic precipitation scheme without any feedback on other model variables) is found to be prone to problems since the diagnostic and prognstic schemes are not consistent with each other. A better choice is found to be the vertically averaged precipitation flux, defined as follows:

$$RR_{ref} \;=\; \frac{\sum_{k_{top}}^{ke} RR_{flux}^{k} \cdot \Delta h^{k}}{\sum_{k_{top}}^{ke} \Delta h^{k}} \qquad , \qquad RR_{flux}^{k} \;=\; \sum_{x} (q_x^k \cdot \rho^k \cdot v_{sed,x}^k)$$

  where q_x is the mass fraction and $v_{sed,x}$ the sedimentation velocity of precipitate x (rain, snow, or graupel), ρ is the density of dry air, Δh the model layer thickness, k

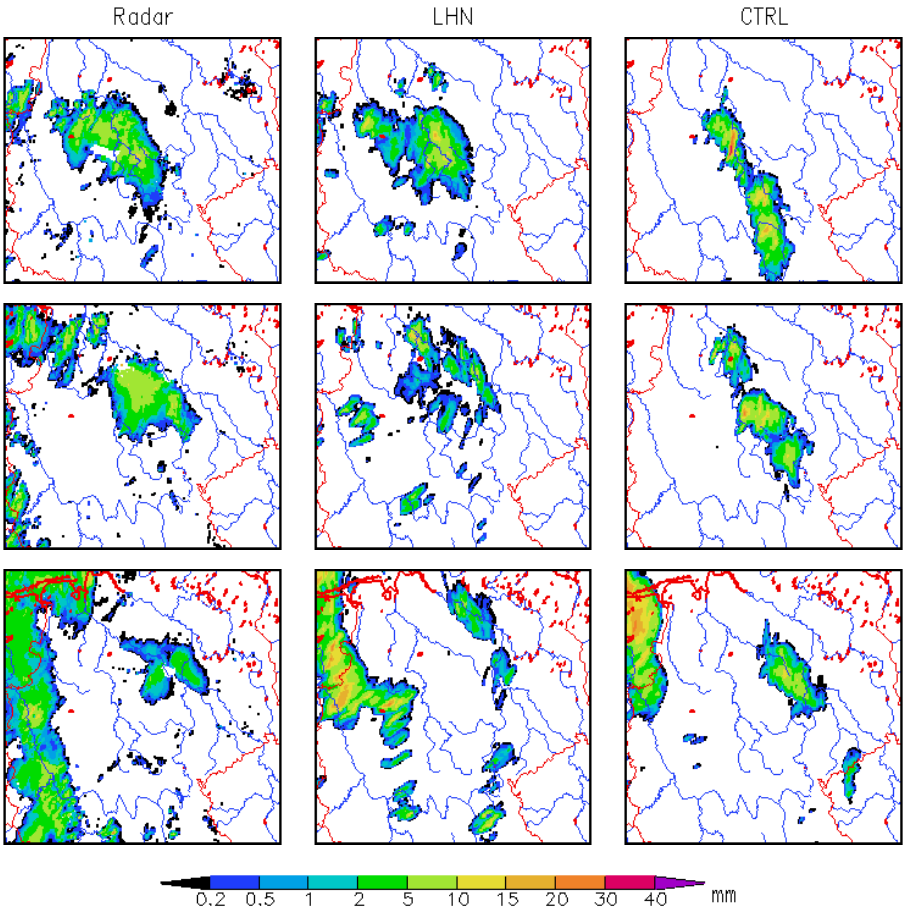


Figure 1: Hourly precipitation over northern and central Germany for LMK forecasts starting at 17 July 2004, 15 UTC. Left column: radar-derived surface precipitation; middle: LMK free forecast from the assimilation cycle with LHN; right: control LMK forecast without LHN. Upper row: 0-h forecast valid for 15 UTC; middle: 2-h forecast for 17 UTC; lower row: 4-h forecast for 19 UTC.

the model level index, ke the index of the lowest model level, and k_{top} the uppermost layer in the grid point column with $|RR^k_{flux}| > 0.01\,\mathrm{mm/h}$.
This type of reference precipitation is compatible with the prognostic model precipitation since both quantities are produced by the same scheme. Note, however, that the averaged flux is a mixture of undelayed and 'fully delayed' information and therefore does only mitigate rather than eliminate the temporal delay problem.

3 Results for an 11-day case study

The above mentioned revisions have been tested for an 11-day convective summer period from 7 to 18 July 2004. An assimilation cycle and 3 daily forecast runs from 00, 12, and 18 UTC have been carried out with the LMK configurations for the general model setup (with Bott advection for humidity and condensate). Note that during the first 3 hours of the forecast runs, the assimilation including LHN was still switched on (unintendedly) so that the free forecasts started in fact at 03, 15, and 21 UTC. In addition to the major revisions, several minor modifications have been implemented in the LHN scheme (e.g. at the grid point search), and the LHN configuration in the experiments also included the following features:

- use of radar observations from the so-called precipitation scan every 5 minutes, and application of a blacklist to reject suspicious radar pixels (e.g. near wind power plants)
- limitation of LHN to grid points with $RR_{obs} > 0.1\,\mathrm{mm/h}$ or $RR_{mo} > 0.1\,\mathrm{mm/h}$

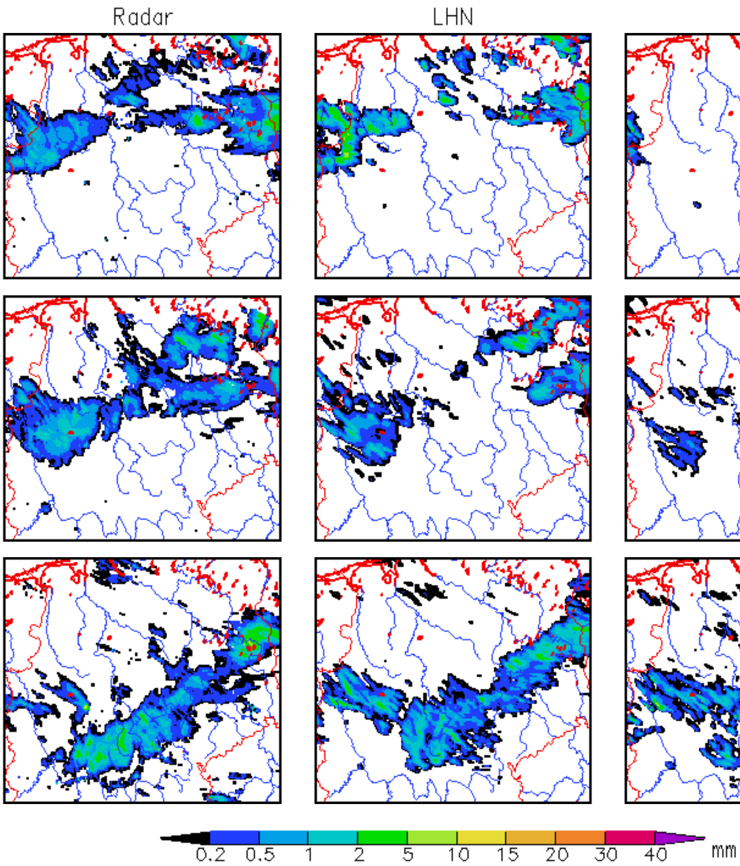


Figure 2: As Fig. 1, but for LMK forecasts starting at 12 July 2004, 3 UTC. Upper row: 0-h forecast valid for 3 UTC; middle row: 2-h forecast for 5 UTC; lower row: 7-h forecast for 10 UTC.

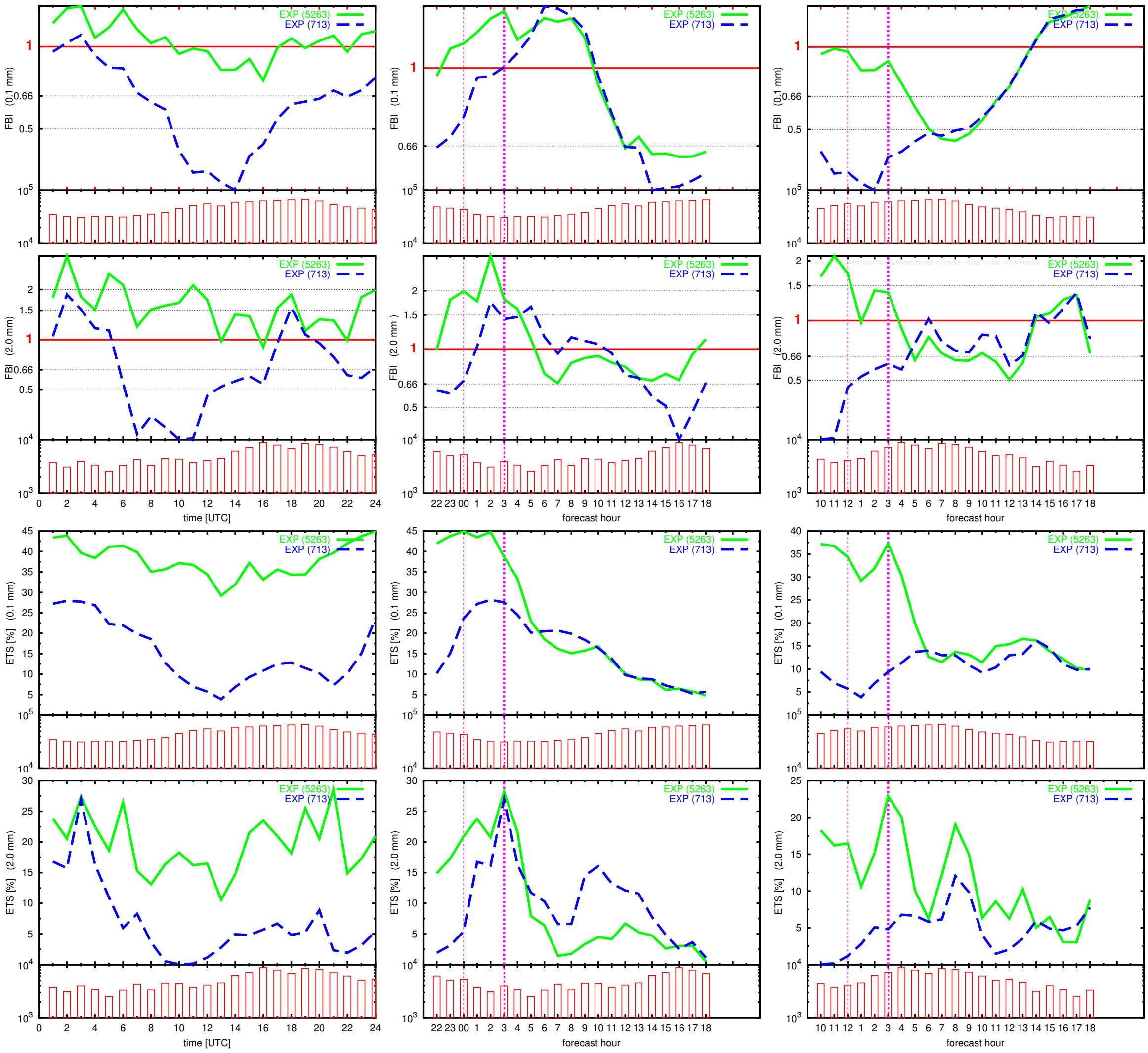


Figure 3: Scores of hourly precipitation (LMK versus radar) as a function of time for a 10-day period from 8 to 17 July 2004. Upper two rows of panels: Frequency Bias (FBI) for a threshold of 0.1 mm and 2.0 mm , respectively; lower two rows: Equitable Threat Score (ETS) for 0.1 mm and 2.0 mm . Left column of panels: assimilation cycle as a function of daytime; middle panels: 0-UTC forecast runs as a function of forecast time (free forecasts starting only at 3 UTC, indicated by the thick pink vertical lines); right panels: 12-UTC forecast runs (free forecasts from 15 UTC). Within each panel: green solid line: LHN experiment; blue dotted line: control experiment without LHN; red columns in lower part: total number of grid points with observed precipitation larger than threshold.

- search for nearby profiles of latent heat release, if both RR_{mo} and the latent heating are 'too small'; use of an idealised 'climatological' profile in case of unsuccessful search
- adjustment of specific humidity (by preserving relative humidity, and by nudging towards saturation at cloud-free model grid points with observed precipitation)

The LHN experiment is evaluated in comparison to a control experiment without LHN. Plots of surface precipitation fields (see e.g. Figs. 1, 2) reveal that during the assimilation, LHN greatly improves the match to the observed rain patterns. In the forecasts however, the improvement is usually reduced very rapidly. In Figure 1, the squall line tends to break up erroneously within two hours and then rearrange in an elongated broken north-south band, so that it is even degraded compared to the 4-hour control forecast. On the other hand, a

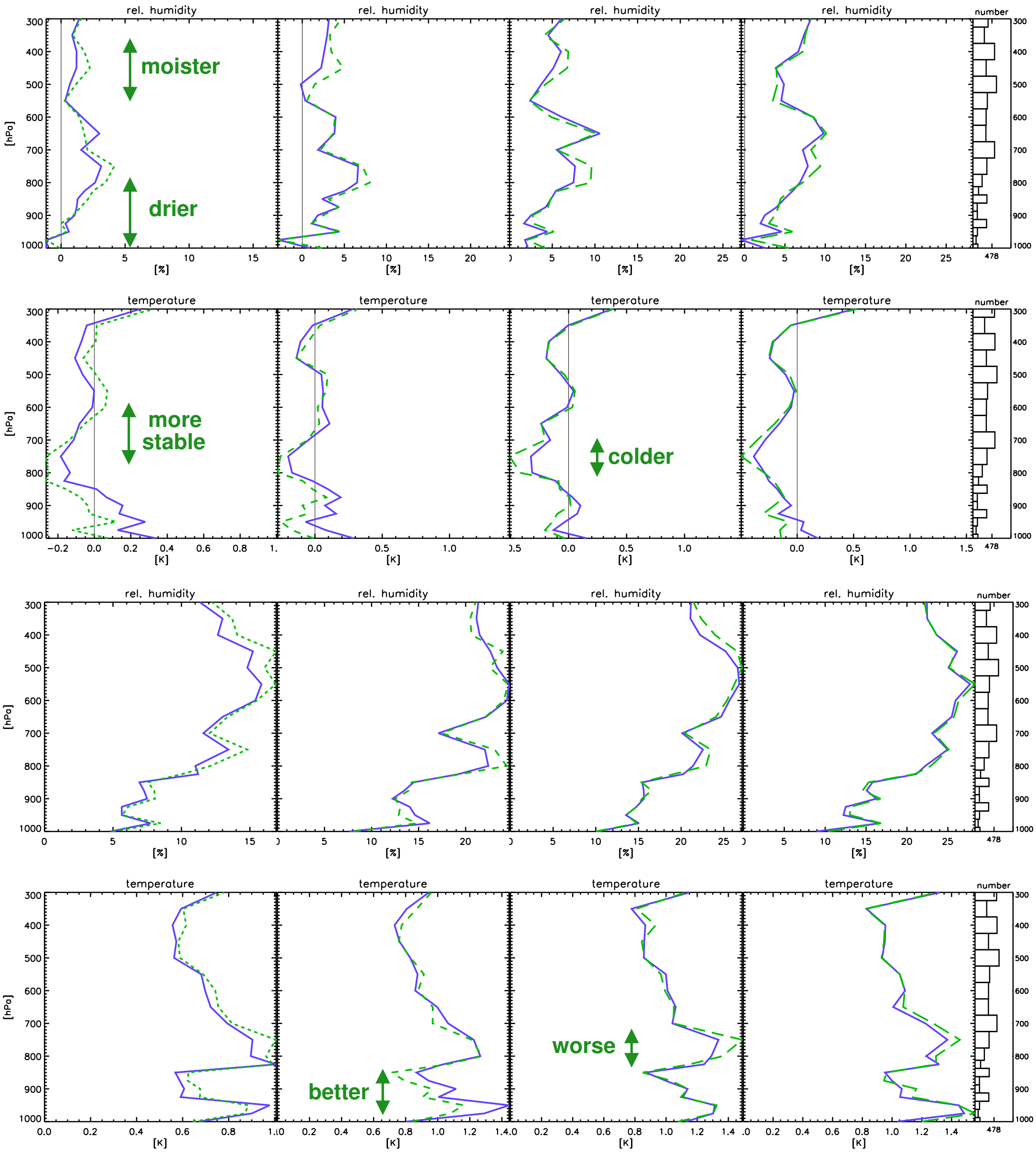


Figure 4: Upper-air verification against German radiosondes for an 11-day period from 8 to 18 July 2004. Panel rows from top down: bias for relative humidity, bias for temperature, rmse for relative humidity, rmse for temperature. Panel columns from left to right: 0-h, 3-h, 9-h, resp. 15-h free forecasts. Green dashed lines: LHN experiment; blue solid lines: control experiment without LHN.

better indication is given of the rain in southwestern Germany. Figure 2 shows a favourable case, where a significant benefit from LHN prevails for 7 hours in the forecast.

Figure 3 shows statistical scores for the whole period. The frequency bias (FBI) indicates that during the assimilation, precipitation is greatly underestimated at daytime without LHN, and it is increased significantly by LHN. While the areal extent (low threshold) is matched very well with LHN, rain amounts are overestimated (by about 50 % for the 2-mm threshold), but less strongly than in previous experiments that used the old LHN scheme. Moreover, the equitable threat scores (ETS) confirm that LHN greatly improves the location of the precipitation patterns. In the forecasts, however, the benefit from LHN decreases rapidly within 2−3 hours. After this, the impact on ETS is neutral for the 18-UTC forecast runs (not shown), remains slightly positive for the 12-UTC runs, and becomes even moderately negative for the 0-UTC runs. Whether this result given by the ETS reflects a real degradation or is an effect of the double penalty problem inherent to the ETS still needs to be evaluated.

The upper-air verification against radiosonde data (Figure 4) indicates that LHN modifies the vertical stratification in the troposphere significantly. In the analyses, it cools and, in terms of absolute humidity, dries the lower troposphere (below 750 hPa resp. 850 hPa) and heats and moistens the upper troposphere (above 600 hPa). As a result, the stability is increased considerably between 750 hPa and 600 hPa . This may be due to an enhanced triggering of convection (reflected by the higher precipitation rates), which acts to reduce atmospheric instability. In terms of rms error, the fit of the analyses to the assimilated temperature and humidity radiosonde observations is decreased. However, the overall impact on the forecasts is very close to neutral (for temperature, humidity, and wind).

4 Concluding Remarks

Several adaptions to the LHN algorithm have been developed to mitigate the problems of LHN related to the prognostic treatment of precipitation. Most importantly, a vertically averaged precipitation flux is used as a 'reference precipitation' instead of the real model precipitation for comparison to the observed precipitation. The revised LHN scheme has been tested for an 11-day convective summer period. During the assimilation, the simulated rain patterns agree well with radar observations, and the overestimation of precipitation is reduced significantly compared to previous LHN versions. In the forecasts, the impact on precipitation decreases rapidly, similarly to past experiments using a diagnostic treatment of precipitation. With respect to other forecast parameters, the overall impact is nearly neutral, e.g. in terms of rmse against radiosonde observations.

Thus, the problems related to prognostic precipitation appear to be mitigated to a satisfactory degree. However, the scheme needs still to be tested for stratiform precipitation, and at least two important shortcomings remain. Firstly, this is the rapid decrease of benefit in the forecasts, and secondly, there are indications that the LHN forcing is too strong:

- The surface pressure fields indicate that strong gravity waves are induced during the assimilation. While local pressure disturbances of $2-3$ hPa can be realistic for convective systems, they sometimes exceed 5 hPa in LHN simulations. The fact that these perturbations do also occur before precipitation is triggered by LHN indicates that they are not primarily linked to the problems related to prognostic precipitation.
- In comparison to surface observations, the convective outflow appears to be too strong (Leuenberger, personal communication).
- LHN tends to stabilise the mid troposphere too much.
- LHN leads to significant cooling and drying of the planetary boundary layer (PBL). This is likely to contribute to the rapid decrease of impact in the precipitation forecasts.

Hence, there is still a need to improve and better balance the scheme. One line of thought is to modify the vertical distribution of LHN increments and add (or take away) more energy and humidity at lower levels and less further above. Unless this is found to require larger LHN increments altogether, it may reduce the effects on the stratification and possibly even extend slightly the period of positive impact on predicted precipitation as a result of increased PBL humidity and decreased stability. To pave the way for developing modifications or new methods with significantly longer forecast impact, a better understanding is needed of how the model itself produces convection, which conditions (such as low-level moisture convergence) it needs, and consequently what kind of observational information and forcing it should be given at which scale.

References

Baldauf, M., J.-P. Schulz, 2004: Prognostic precipitation in the Lokal Modell (LM) of DWD. *COSMO Newsletter*, No. 4, 177 - 180.

Gassmann, A., 2002: 3D-transport of precipitation. *COSMO Newsletter*, No. 2, 113 - 117.

Klink, S., K. Stephan, 2005: Latent heat nudging and prognostic precipitation. *COSMO Newsletter*, No. 5, 124 - 131.

Revised quality control for radiosonde humidity

Christoph Schraff

Deutscher Wetterdient, Offenbach am Main, Germany

1 Introduction

In July 2004, several soundings with strongly overestimated humidity values were issued by the radiosonde station Stuttgart. While the OI-based GME analysis scheme rejected at least some of these data, the operational nudging scheme of LM failed to detect these errors and assimilated the humidity observations. The impact on LM forecasts was rather limited and did not attract attention. In the assimilation cycle, however, it caused spurious rain of up to 200 mm in 24 hours (31 July − 1 August 2004, 6 UTC).

In the threshold quality control (QC) of the nudging, which is similar to a first-guess check of 3-D analysis schemes, it is assumed that the assimilating model run provides a fair estimate of the true atmospheric state. Hence, a relative humidity observation U_o can be considered erroneous and is rejected if it deviates from the model value U_b by more than a threshold value U_{thr} (i.e. if $|U_o - U_b| > U_{thr}$). In the operational LM, the threshold $U_{thr}(t_o)$ valid at observation time t_o is 70 % whereas for GME, it had been reduced to about 28 % for most data. In addition, the LM threshold is further enhanced with increasing distance to the observation time (by error up to 100 %). This leads to hardly any data ever being rejected.

Therefore, a modification to the quality control of upper-air humidity is required which is able to reject at least some of the bad data, including those leading to the 200 mm of spurious rain on 1 August 2004, but which at the same time accepts most of the good data. Note that near strong inversions in wintertime low stratus periods, good observations often deviate from the model values by (far) more than 50 % relative humidity since such inversions tend to be simulated poorly by the model. In order to meet these requirements to a satisfactory degree, it is found not to be sufficient to only modify the existing QC steps, i.e. the threshold QC for individual humidity observations and the multi-level check. A spatial consistency check for integrated water vapour is added for this reason.

2 Revisions to the quality control

Revised, stability-dependent thresholds in the QC for individual observations

In a first trial, the QC thresholds and multi-level check of the GME OI analysis were adopted. This, however, resulted in far too many observations being rejected (about 40 %) and initiated also a revision the GME QC by increasing the thresholds from about 28 % to 44 % for most data. These new thresholds are also used in the revised QC for LM:

$$U_{thr\,(1,3)}\,(t_o) \quad = \quad \min\left[\left({\sigma_0}^2 + {\sigma_b}^2\right)^{1/2} \ , \ 2\ \sigma_b\right] \cdot c_{flag\,(1,3)}$$

where the observation error $\sigma_0 = 10\,\%$ ($15\,\%$ for $T_o < 233\,\mathrm{K}$, $20\,\%$ for $U_o < 20\,\%$), the background error $\sigma_b = 10\,\%$ ($15\,\%$ south of $30\,\mathrm{N}$), and the constant $c_{flag\,(3)} = 3.1$ ($c_{flag\,(1)} = 1.8$ for flag 1 as used in the multi-level check).

In strongly stable situations and in particular at inversions, model errors are known to be increased often. In the revised QC for LM, the assumed background error σ_b is therefore enhanced by 2 terms selectively for those humidity observations at which the observed lapse rate β to the next humidity observation further above or below is $\beta > \beta_{crit} = -0.0065\,\mathrm{K/m}$:

$$\begin{aligned}
\sigma_b &\rightarrow \sigma_b \cdot (1 + f_{stable} + f_{invers}) \\
f_{stable} &= 1/4 \cdot (1 - \min(\beta, 0) / \beta_{crit}) \cdot (1 + c_s), \qquad c_s = \Delta_\beta T / (1 + \Delta_\beta T) \\
f_{invers} &= 1/5 \cdot \max(\Delta T, 0) \cdot (1 + c_i), \qquad c_i = \min(2, \beta / 0.05)
\end{aligned}$$

$\Delta T = T_k - T_{k-1}$, where T_k and T_{k-1} are the temperature observations at the humidity observation level k respectively at the next level $k-1$ further below. $\Delta_\beta T = T_k^\beta - T_{k-1}$, where T_k^β is T_k extrapolated to level $k-1$ with the lapse rate β_{crit} . Both terms f_{stable} and f_{invers} increase with increasing stability and with increasing thickness of the stable layer (given by the two successive humidity observation levels).

Finally, an upper limit of $70\,\%$ is imposed to the threshold $U_{thr}\,(t_o)$ at observation time. With increasing distance to the observation time, the threshold is enhanced linearly to a maximum of $77\,\%$ (with the temporal weight function used currently in the nudging).

Multi-level check

The revised multi-level check is analogous to that of the GME OI analysis (but not equivalent, due to the different, stability-dependent thresholds for flag ≥ 1 in the first guess check):

- Analysis layers are defined equal to the standard layers except below $700\,\mathrm{hPa}$, where the thickness of the analysis layers is reduced to $50\,\mathrm{hPa}$ and below $800\,\mathrm{hPa}$ to $25\,\mathrm{hPa}$.
- Criterion: If 4 or all consecutive standard layers contain humidity observations with flag ≥ 1, then these standard layers are set to 'rejected'. Each analysis layer within those rejected standard layers is set to 'rejected' if it contains observations with flag ≥ 1. All observations within these rejected analysis layers are rejected.

Spatial consistency check of integrated water vapour (IWV)

A spatial consistency check of integrated water vapour has been developed to detect a general bias in a radiosonde humidity sounding. As a first step, observation increments of IWV are derived from radiosonde humidity profiles and optionally also from ground-based GPS zenith path delay data. At the location of each IWV 'observation' Q_k , an IWV 'analysis increment' ΔQ_k^{ai} is then computed using only the neighbouring observations $Q_{j \neq k}$:

$$\Delta Q_k^{ai} = \frac{\sum_{j \neq k} w_{kj}^2 \cdot \frac{Q^{sat}(\mathbf{x}_k, t)}{Q^{sat}(\mathbf{x}_j, t)} \cdot (Q_j - Q(\mathbf{x}_j, t))}{\max\left(\sum_{j \neq k} w_{kj}^2 \, , \, 1\right)}$$

Here, $Q^{sat}(\mathbf{x}_k, t)$ is the IWV derived from the model temperature profile at the observation location assuming saturation. The Q^{sat} term scales the observation increment, mainly in order to account for differences in orographic height. The weight w_{kj} consists of a horizontal weight (equal to that used for the nudging of radiosonde humidity data at 850 hPa respectively for GPS data), and of a temporal weight (given by a linear function of time within $\pm 2\,h$ respectively $\pm 1\,h$ from the observation time).

The spatial consistency check of IWV is a revised first guess check, in which the model background is corrected by the above 'analysis increment' in order to obtain a better estimate of truth. The complete humidity profile of the sounding k is rejected if

$$\left| \; Q_k \;-\; \left(Q(\mathbf{x}_k, t) \;+\; \Delta Q^{ai}_k \right) \; \right| \;\; > \;\; Q^{ai}_{thr\,k}$$

This check corresponds to a first guess check of IWV if there are no neighbouring observations influencing the observation location $\mathbf{x}_k$. This usually applies approximately if GPS data are not used. The basic threshold $Q_{thr\,k}(t_o)$ depends on temperature and is set to (in [mm]):

$$Q_{thr\,k}(t_o) \;\;=\;\; \left(\, 1 \;+\; 0.15 \cdot Q^{sat}(\mathbf{x}_k, t) \right)$$

In the presence of many neighbouring IWV observations, however, the check addresses the spatial consistency between them. The more observations are used for the 'analysis increment', the more accurate the estimate of truth, and the smaller the threshold $Q^{ai}_{thr\,k}$ should be set. On the other hand, the larger the 'analysis increment' and hence the disagreement between model and observations, the more uncertain the estimate of truth, and the larger the threshold should be. Therefore, the following correction is applied to $Q_{thr\,k}$:

$$Q^{ai}_{thr\,k}(t_o) \quad = \quad Q_{thr\,k}(t_o) \cdot \left(\, 1 \;-\; 0.2 \cdot \min \left(0.2 \cdot \textstyle\sum_j w_{kj}^2 \;,\; 1 \right) \right) \;\; + \;\; \Delta Q^{ai}_k$$

3 Results

The revised QC for radiosonde humidity has been tested for 14 days in July 2004 and a 5-day wintertime low stratus period from 9 to 13 February 2003. The humidity profiles of radisonde Stuttgart, that lead to the strong spurious rain of 1 August 2004, are rejected successfully by the IWV check (not shown).

Figure 1 illustrates the negative effects if the QC is too strict. Subjective evaluation does not give any indication for errors in the radiosonde observations within the domain shown. Accepting all humidity data with the operational QC renders a fairly good analysis of low cloud for 13 February 0 UTC. Rejecting many data when using the small thresholds of the old GME OI version strongly degrades the analysis in the region around the Lyon and Payerne radiosonde stations. With preliminary stability-dependent but still too small thresholds (Figure 1, lower left), more data are accepted again, and most of the cloud around Lyon comes back. In this analysis, however, low cloud is missing in a large area around the Paris sounding, because the thresholds are still so small as to reject the moist data below the inversion, but large enough now to accept all the dry data above it. In the final revised QC, the introction of the IWV check allows to further relax the thresholds in the first guess and multi-level checks, so that for the case shown, almost all the relevant observations are accepted, and the low cloud analysis is as good as the original one.

In the whole 5-day low stratus period, the new QC rejects 4 % of the humidity profiles (completely, or partially from the top down to at least 700 hPa) in the multi-level check and another 1 % in the IWV check. Very few data are rejected additionally by the first guess check. About 80 % of the rejected profiles are relatively close (within 50 grid points) to the lateral boundaries of the LM domain. Some of them are rejected erroneously when the model fields are far too moist above the inversion, after this moisture has been advected from the lateral boundaries (see e.g. Figure 2). The latter reflect the GME OI analysis which often grossly overestimates moisture above inversions.

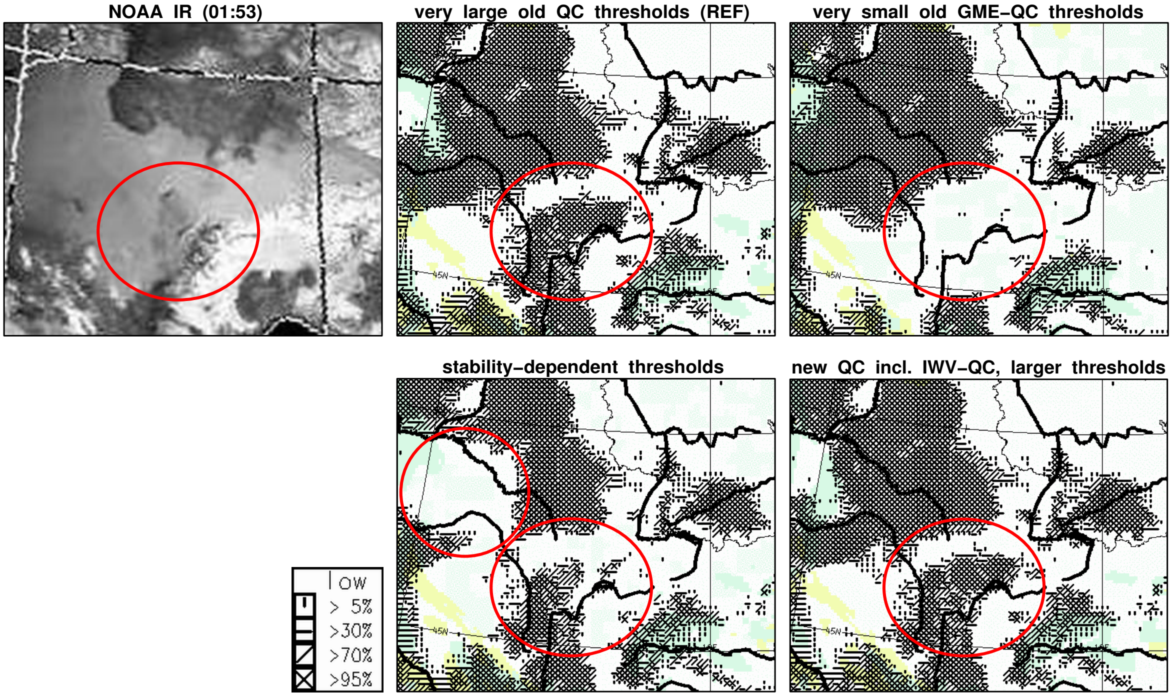


Figure 1: Cloud cover for southwestern Germany, northeastern France, and parts of the Alps on 13 February 2003, 00 UTC. Upper left panel: NOAA IR image (at 01:53 UTC); upper middle: reference LM analysis with old QC; upper right: LM analysis with QC thresholds as for operational GME; lower left: LM analysis with preliminary stability-dependent QC; lower right: LM analysis with new QC. In LM analyses panels, low cloud cover is displayed in black patterns as by the legend, middle and high cloud cover > 50 % in green and yellow shading. Red circles indicate areas of main interest.

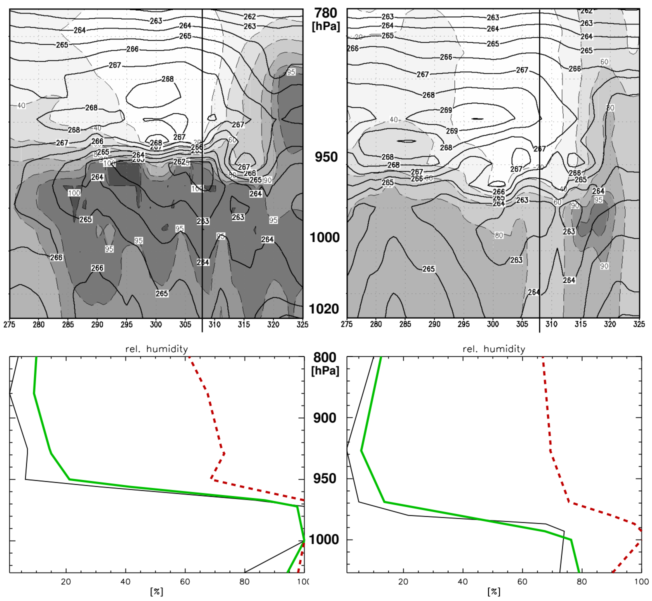
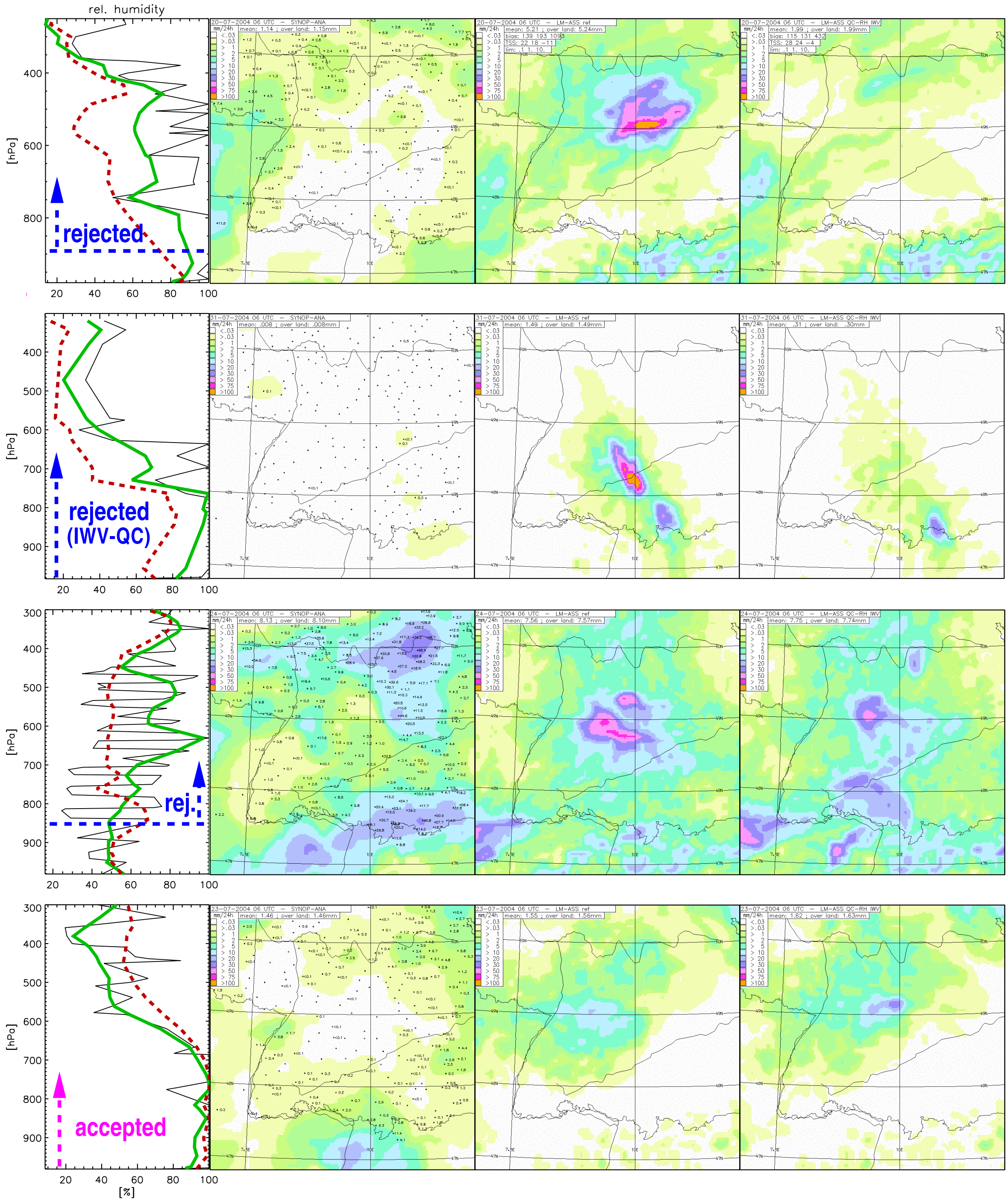


Figure 2: Upper panels: west-east vertical cross sections through the lowest 15 model levels (the numbers between the panels indicate the approximate pressure) and 50 horizontal LM grid points, which include the location of the Warsaw radiosonde station (indicated by the black vertical lines) and the eastern boundary of the LM domain (right boundary of panels). Solid black contours: temperature [K] of reference LM analyses with the old QC; grey shading and thin dashed contours (for 20, 40, 60, 80, 90, 95, 100 %): relative humidity. Lower panels: vertical profiles of relative humdity at Warsaw for 12 February 2003. Solid black line: observation; thick solid green line: reference LM analysis with old QC; red dashed line: LM analysis with new QC. Left panels: 00 UTC; right panels: 12 UTC.

In the 14-day period in summer 2004, significant spurious rain has occured with the old QC in 4 cases (shown in Figure 3, apart from the 1 August case that was used to tune the new QC). In the two most severe cases with spurious rain exceeding $100\,\mathrm{mm}$, the revised QC is able to remove that rain (almost) completely, in the second case by means of the new IWV check. In the third case, it rejects the data of an erroneous profile only above $850\,\mathrm{hPa}$. This does not eliminate the spurious rain but reduces its area and maximum amount (from 75 to $50\,\mathrm{mm}$) and also tends to improve the rain patterns in the environs. It is the forth and least severe case only, where the revised QC completely fails to reject the erroneous data. In the whole period, $2\,\%$ of the profiles are (at least partially) rejected by the multi-level check and $1\,\%$ by the IWV check. $35\,\%$ of the rejected profiles are Stuttgart soundings.

4 Concluding Remarks

The quality control (QC) for radiosonde humidity has been revised. This includes a significant reduction of the general threshold in the 'first guess' check. Yet, a new stability-dependent enhancement to it allows to account for large model errors and observation increments near inversions. Furthermore, a spatial consistency check for integrated water vapour (IWV) derived from radiosonde humidity and optionally from GPS-derived zenith path delay has been developed. This check uses model-derived IWV as background information and is equivalent to a first guess check of IWV in the absence of neighbouring observations. The revised QC rejects about $2-5\,\%$ of the humidity data, including most of the erroneous data from the radiosonde Stuttgart in July 2004, but it accepts most of the correct data near strong wintertime inversions. It is planned to become operational in LME at the beginning of 2006.



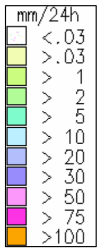


Figure 3: Left panels: Vertical profiles of relative humidity at Stuttgart on (from top to bottom) 19 July 2004 23 UTC, 30 July 23 UTC, 23 July 17 UTC, 22 July 23 UTC; solid black lines: observation; green thick solid lines: reference LM analysis with old QC; red thick dashed lines: LM analysis with new QC; horizontal blue dashed lines: approx. level above which all humidity data are rejected by the new QC.
Other panels: 24-hour sum of precipitation in southwestern Germany valid at (from top to bottom) 20, 31, 24, resp. 23 July, 06 UTC; panels from left to right: analysis from synop observations; reference LM analysis with old QC; LM analysis with new QC.

Implementation of the 3D-Turbulence Metric Terms in LMK

Michael Baldauf

Deutscher Wetterdienst, Offenbach am Main, Germany

1 Introduction

At the Deutscher Wetterdienst the numerical weather forecast model LMK (LM-Kürzestfrist) is currently under development. It is based on the LM (Lokal-Modell) and will be used for very short range forecasts (up to 18 hours) and with a resolution on the meso-γ-scale (about 2.8 km). The development tasks cover the areas of data assimilation, numerics (e.g. Förstner et al., 2005), physical parameterisations (e.g. Theunert and Seifert (2005), and Reinhardt and Seifert (2005)), and new verification approaches (e.g. Lenz and Damrath, 2005).

In Baldauf (2005) the formulation of the turbulent fluxes and flux divergences in terrain following coordinates and for spherical base vectors was derived. The result for the scalar flux divergence is

$$\rho\frac{\partial s}{\partial t} = \underbrace{-\frac{1}{r\cos\phi}\frac{\partial H^{*1}}{\partial\lambda}}_{(a)}\underbrace{-\frac{J_\lambda}{\sqrt{G}}\frac{1}{r\cos\phi}\frac{\partial H^{*1}}{\partial\zeta}}_{(b)}\underbrace{-\frac{1}{r}\frac{\partial H^{*2}}{\partial\phi}}_{(c)}\underbrace{-\frac{J_\phi}{\sqrt{G}}\frac{1}{r}\frac{\partial H^{*2}}{\partial\zeta}}_{(d)}\underbrace{+\frac{1}{\sqrt{G}}\frac{\partial H^{*3}}{\partial\zeta}}_{(e)}$$
$$\underbrace{-\frac{2}{r}H^{*3}}_{(f)}\underbrace{+\frac{\tan\phi}{r}H^{*2}}_{(g)}, \qquad (1)$$

and for scalar fluxes:

$$H^{*1} = -\rho K_s\frac{1}{r\cos\phi}\left(\frac{\partial s}{\partial\lambda}+\frac{J_\lambda}{\sqrt{G}}\frac{\partial s}{\partial\zeta}\right), \qquad (2)$$
$$H^{*2} = -\rho K_s\frac{1}{r}\left(\frac{\partial s}{\partial\phi}+\frac{J_\phi}{\sqrt{G}}\frac{\partial s}{\partial\zeta}\right), \qquad (3)$$
$$H^{*3} = +\rho K_s\frac{1}{\sqrt{G}}\frac{\partial s}{\partial\zeta}. \qquad (4)$$

(Analogous expressions for the 'vectorial' diffusion of u, v and w). The terms (a), (c) and (e) in equation (1) describe the cartesian components of the flux divergence and are already contained in the 3D-turbulence scheme of the Litfass-LM (Herzog et al., 2002), which was implemented into the LMK code (Förstner et al., 2004). The metric terms (b) and (d) describe corrections due to the terrain following coordinate. As shown in Baldauf (2005), the terms (f) and (g) are due to the earth curvature and can be neglected with good approximation.

In this article, the implementation of the metric terms (b) and (d), (and also their counterparts in the scalar and vectorial fluxes and vectorial flux divergences) is described. Results of a test procedure are shown and in a first real case study the importance of the 3D-turbulence for meso-γ-models is examined.

2 Implementation and test of the metric terms

In subroutine `explicit_horizontal_diffusion`, the horizontal cartesian terms (a) and (c) are discretised explicitely whereas the vertical cartesian term (e) is in implicit form to consider the fact that the stability criterion can be violated in the case of small vertical level distance or large diffusion coefficients. The metric terms (b) and (d) are some sort of horizontal correction due to the terrain following coordinate system and therefore the first attempt for their discretisation is also in explicit form.

In Fig. 1 the positions of different sorts of variables in the staggered grid are shown. The scalar variables and the diagonal terms of the momentum fluxes are sitting at the center of the grid (an exception of this rule is the scalar variable 'turbulent kinetic energy', which is defined at the w-positions). The velocity variables and the fluxes of scalar variables have positions due to the Arakawa-C/Lorenz grid. The non-diagonal momentum fluxes are again staggered relative to their velocity components. A positive side effect of the Arakawa-C/Lorenz grid is that this sort of staggering simplifies the discretisation of the cartesian components of fluxes and flux divergences: simple centered differences of these terms are easily possible. In contrast to this, for the non-cartesian metric components (like the terms (b) and (d) above) additional averages to other positions in the staggered grid are needed. Surprisingly these additional calculations are not very costly: in the real case study presented below, the complete 3D-turbulence scheme needs about 8.5 % of the total calculation time. The calculation of the metric correction terms alone needs about 3.5 % of the total calculation time of the LMK run (activation of the metric terms will be possible by setting the new namelist-parameter `l3dturb_metr=.TRUE.` in one of the subsequent LM versions).

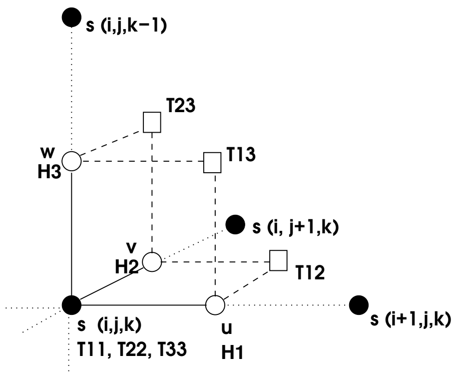


Figure 1: Positions of scalars s, vectors v^i or H^i and 2. rank tensor components T^{ij} in the staggered grid.

For the upper and lower boundaries ($k = 1$ and $k = ke$) all the centered differences were replaced by one-sided differences (for fluxes and flux divergences). In an idealised diffusion test, where the diffusion cloud arrives at the boundaries, no artefacts at the boundaries were observed.

The correct discretisation of the 3D-turbulence and especially the metric terms was tested with an isotropic diffusion problem. It is well known that the diffusion equation (3D) for a tracer ϕ

$$\frac{\partial \phi}{\partial t} = K \Delta \phi \tag{5}$$

with constant diffusion coefficient $K = const.$ and a Gaussian initial distribution

$$\phi(x, y, z, t = 0) = 1 \cdot e^{-r^2/a^2} \tag{6}$$

has also a Gaussian solution

$$\phi(x,y,z,t) = \frac{\Phi_0}{\sqrt{4\pi K(t+t_0)}^3}\, e^{-\frac{r^2}{4K(t+t_0)}}, \qquad r := \sqrt{x^2+y^2+z^2} \tag{7}$$

with

$$t_0 = \frac{a^2}{4K}, \qquad \frac{\Phi_0}{\sqrt{4\pi K t_0}^3} = 1 \tag{8}$$

For this test a resolution of $\Delta x = \Delta y = \Delta z = 50$ m was used[1] with $60 \times 60 \times 60$ gridpoints. For $K = 100$ m^2/s and a good resolution of the initial distribution ($a = 250$ m was chosen) it follows $t_0 = 156.25$ sec. The stability criterion for explicit horizontal discretisation needs a time step of at most $[2K\,((\Delta x)^{-2} + (\Delta y)^{-2})]^{-1} \approx 6$ sec. Here a time step of 3 sec. was chosen.

To test the metric terms an orography with a sinus shape in both horizontal directions $h(x,y) = h_0/2\;(1 + \sin(k_x x)\cdot\sin(k_y y))$, $h_0 = 250$ m, $k_x = 2\pi/(12\Delta x)$, $k_y = 2\pi/(12\Delta y)$, was chosen and shifted in a way that the initial distribution and the orography have no common reflection symmetry planes or points. This orography is plotted in the horizontal cross-section in Fig. 3 (right). Of course, the orography should not at all influence the diffusion process, if it is away enough from the tracer. But it induces a distortion of the numerical grid and this grid distortion has to be corrected by the metric terms.

Three tests were performed:

- case 1: vertical (i.e. 1D) diffusion,
- case 2: 3D-diffusion, without metric,
- case 3: 3D-diffusion, with metric.

In each of the cross-section plots in Figrs. 2 to 4, the analytic solution (7) is plotted with thin, black lines, whereas the numerical solution is plotted with thick coloured lines.

Case 1: Due to the missing horizontal diffusion terms, the 1D-diffusion scheme diffuses anisotropically only in vertical direction as shown in a vertical (x-z-)cross-section in Fig. 2.

Case 2: In the 3D-diffusion scheme without metric terms a horizontal diffusion takes place but the distorted grid also distorts the solution. Comparing this with Case 1, one could be astonished that the 1D-turbulence has no problems with orography. This is indeed the case, because each column of the 1D turbulence scheme works correctly. The false influence of the distorted grid in Case 2 comes purely from the horizontal diffusion terms.

Case 3: In contrast to Case 2 the 3D-diffusion scheme with metric terms is able to reproduce almost correctly the analytic solution. Figure 4 shows vertical and horizontal (i.e. x-z-, y-z-, x-y-) cross-sections and also the temporal behaviour of the numerical solution. As can be seen, the metric terms are able to correct the deformation of the purely cartesian diffusion almost completely. Any mistake in the discretisation would lead to a deviation at least in one spatial direction. Therefore this test should prove the correct discretisation of each single term.

It should be remarked here, that this test is designed only for the scalar variables. It does not work exactly for the velocity components ('vectorial' diffusion). Nevertheless a similar test for the u-component performed rather well, too.

[1] The reason for such a fine resolution is that LM uses a base state with a variation of the density with height. Therefore also the diffusion coefficient (which is multiplied by the density) would no longer be constant. To reduce this artefact in this special test the model area was chosen to have a relatively small vertical extension. Otherwise more parts of LM would have had to be reprogrammed for this test.

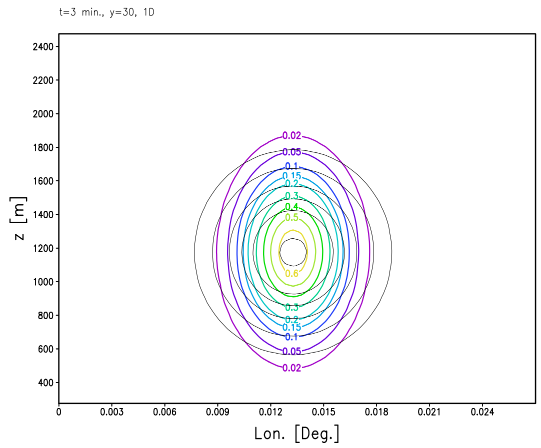


Figure 2: Case 1: 1D diffusion; vertical (x-z-)cross-section after 3 Min., analytic solution (thin, black lines) and numeric solution (thick, coloured lines).

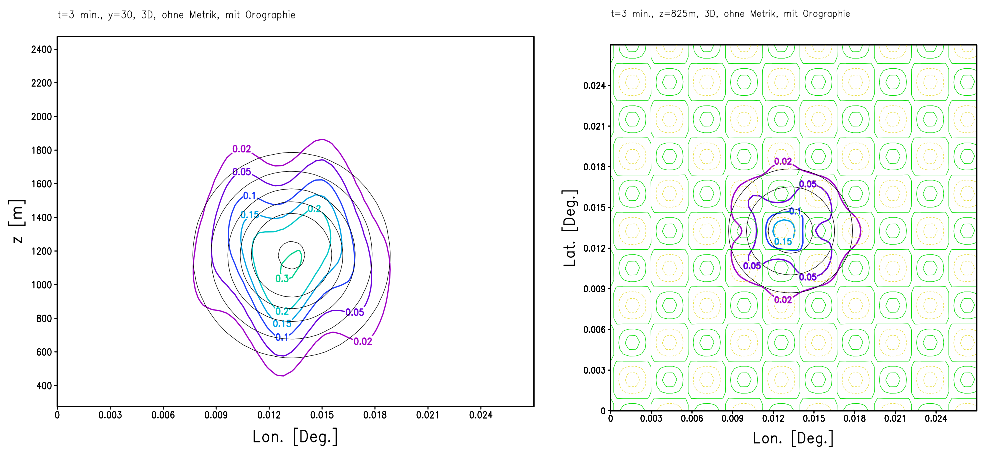


Figure 3: Case 2: 3D diffusion without metric terms; vertical (x-z-)cross-section (left), horizontal (x-y-)cross-section (right).

3 A real case study

To inspect the importance of 3D-turbulence and especially the metric terms a first real case study was performed. The simulation was started at the 12. August 2004, 12 UTC, and lasted 18 h. This was a strong convective situation with the development of a squall line. Again the three cases (1) only 1D (vertical) turbulence, (2) 3D-diffusion without metric terms, and (3) 3D-diffusion with metric terms, were carried out. The simulations were performed with the 'standard' LMK horizontal resolution of 2.8 km and a time step of 30 sec. As in the idealized simulations, the explicit discretisation of the metric terms did not generate any stability problems.

The precipitation sum after 18 h simulation time is plotted in Fig. 5 (upper part). The left plot shows the whole simulation area, whereas the right plot zooms into the part of Southern Germany with a wider distribution of mid-level mountains. This area was chosen, because during the simulation time the precipitation event travelled completely over this area and lied outside of it at the end of the simulation. Therefore in the following difference plots, we can be sure, that the differences do not alter because of new precipitation events.

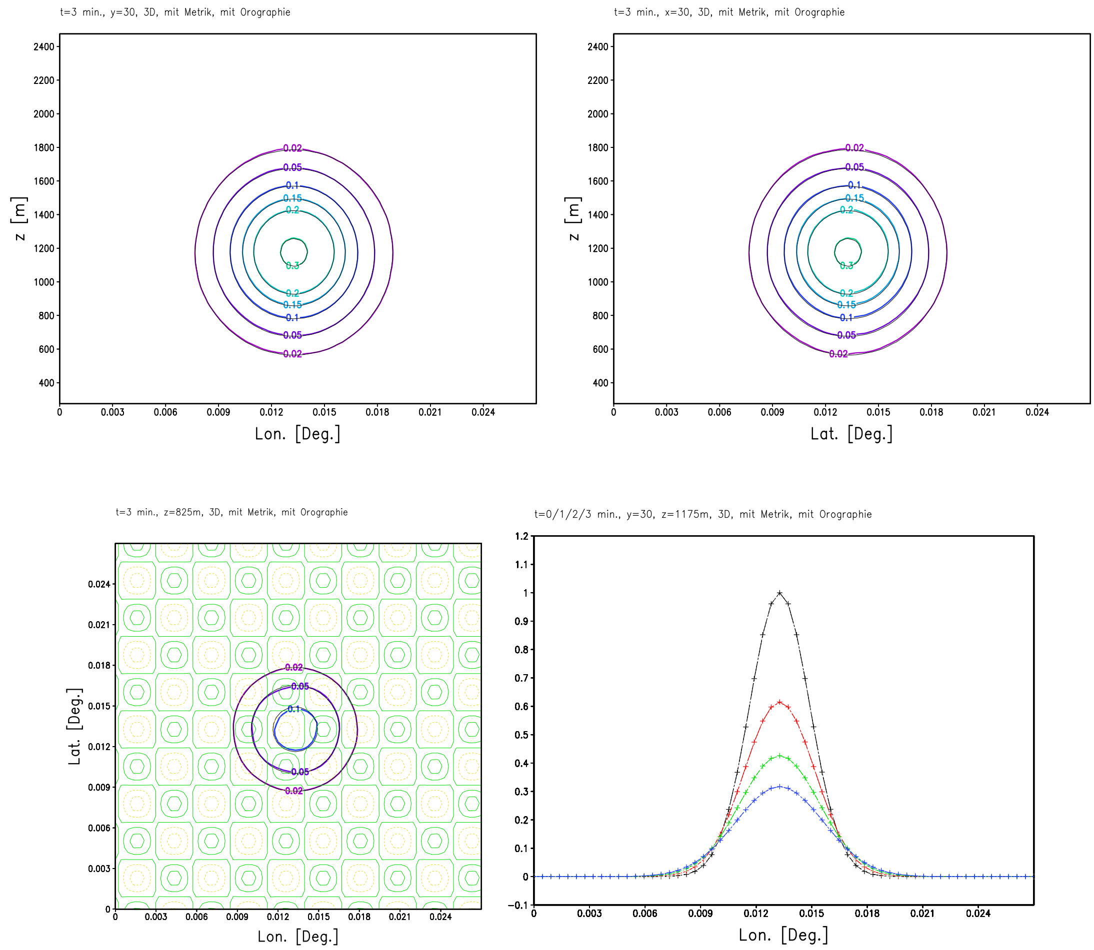


Figure 4: Case 3: 3D-diffusion with metric terms; x-z-cross-section (upper left), y-z-cross-section (upper right), x-y-cross-section (lower left), x-cross-section for different times (lower right).

Figure 5 (below, left) shows a difference plot between Case 3 (complete 3D-turbulence) and Case 1 (only vertical turbulence). The differences can reach maximum and minimum values nearly at the same order as the precipitation sum itself (see Fig. 5). But a look to the mean value shows, that the 3D-turbulence has almost no influence to the total amount of precipitation. This seems to be reasonable, as turbulence occurs mostly in the boundary layer, whereas the most part of precipitation is generated above. But the transport of precipitation (especially with the prognostic precipitation scheme, which is unconditionally necessary at this resolution) is heavily influenced by the boundary layer flow. The 18h-precipitation sum is a marker of all these integrated flow changes due to 3D-turbulence.

The effect of the metric terms themselves is shown in Fig. 5 (below, right), the difference between Case 3 (with metric terms) and Case 2 (without metric terms). The maxima and minima are slightly smaller, but obviously they cannot be neglected in comparison to Case 2 (this was already theoretically derived in Baldauf, 2005). The mean value of the difference 'Case 1 - Case 3' is even smaller as in the difference 'Case 2 - Case 3'. If one accepts the statement, that 3D-turbulence does not alter the total amount of precipitation, then the metric terms obviously have a positive impact on the conservation of total precipitation amount, too.

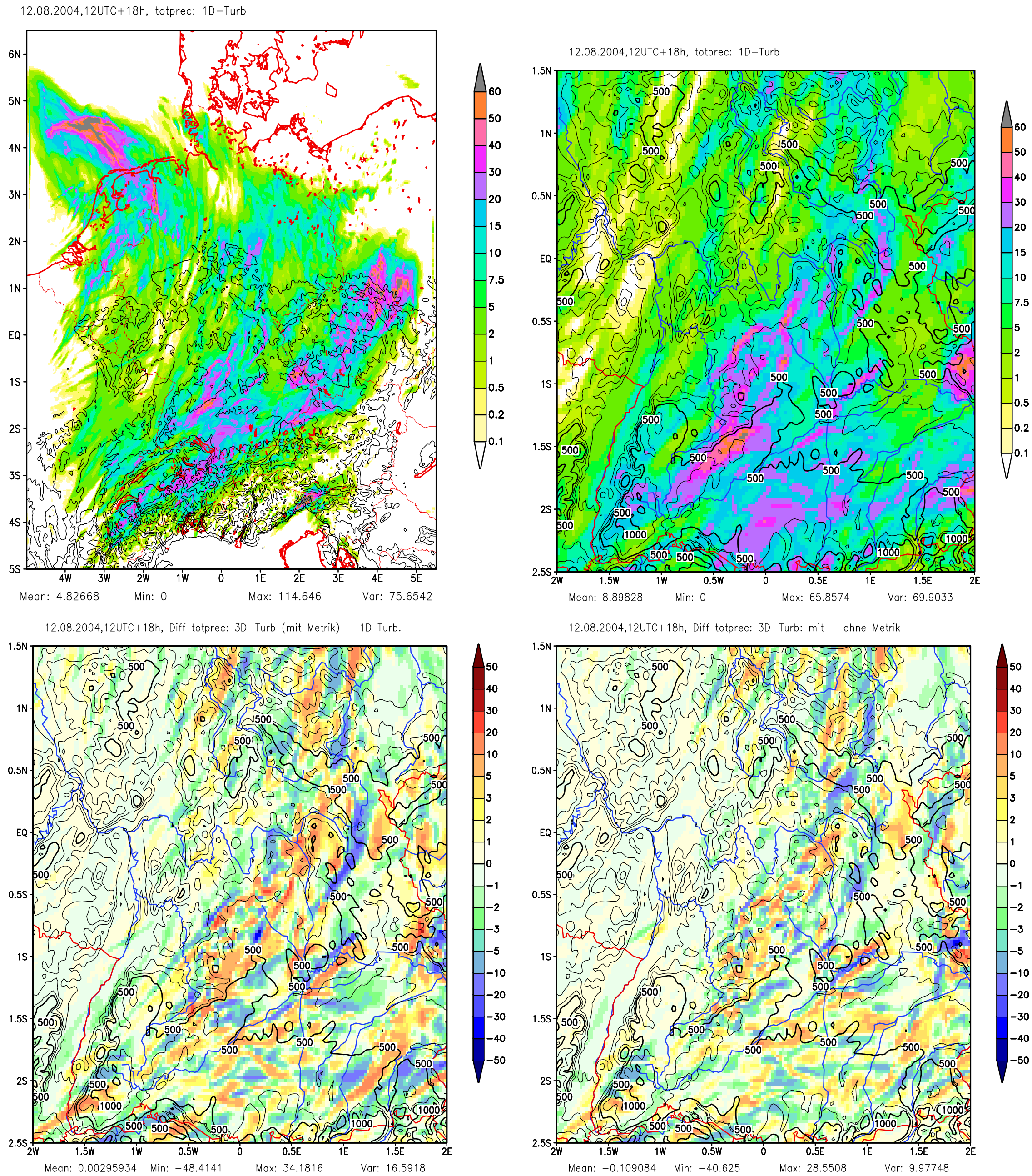


Figure 5: Above: Total precipitation sum over 18h for the 12.08.2004, 12 UTC run with the currently used 1D turbulence. Left: total simulation area, right: zooming into southern Germany.
Below: Differences of the 18h precipitation sum. Left: between 3D-turb. with metric terms and 1D-turb., right: between 3D-turb. with and without metric terms.

A remark should be made about the diffusion coefficient. In the Litfass-LM-turbulence scheme (Herzog et al., 2002) the ratio between horizontal and vertical diffusion coefficient was stretched with the grid length ratio $\Delta x/\Delta z$ with the consequence that near the bottom K_{hor} is about one order of magnitude bigger than K_{vert}, whereas above the boundary layer, K_{hor} is much smaller than K_{vert}. This stretching designed for LES-simulations seems not to be adequate in a statistical turbulence model. Here, as a first guess, the diffusion coefficient was assumed to be isotropic $K_{hor} \equiv K_{vert}$ (e.g. Klemp and Wilhelmson, 1978). This choice has the further advantage to circumvent the problem that the distinction between horizontal and vertical diffusion coefficients becomes much more difficult in steep terrain and should be set on a more profound theoretical basis.

4 Summary and outlook

The correct implementation of the 3D-diffusion terms (at least for the scalar variables) for terrain following coordinates into the LMK was shown with the analytically known case of isotropic Gaussian diffusion. The stabilitiy of the explicit discretisation of the metric terms was demonstrated in this idealised test and also in the real case study.

A certain influence of the 3D-turbulence on the precipitation pattern was found from the one real case study. The 3D-turbulence seems not to alter the total amount of precipitation but can shift the precipitation areas up to 30-40 km. This would be a non-neglectable effect if one has e.g. hydrological applications in mind. Of course further work has to be done to inspect if this influence can be seen also in other weather situations.

References

M. Baldauf, 2005: The Coordinate Transformations of the 3-dimensional Turbulent Diffusion in LMK, *COSMO-Newsletter*, No. 5, 132-140

Förstner, J., M. Baldauf and A. Seifert, 2005: Courant Number Independent Advection of the Moisture Quantities for the LMK, *COSMO-Newsletter*, No. 6

Förstner, J. and H.-J. Herzog and G. Vogel, 2004: Implementation of a 3D-Turbulence Parameterisation for the Very Short Range Forecast Model LMK, *WGNE Blue Book*, `http://www.cmc.ec.gc.ca/rpn/wgne/`

H.-J. Herzog and G. Vogel and U. Schubert, 2002: LLM - a nonhydrostatic model applied to high-resolving simulations of turbulent fluxes over heterogeneous terrain, *Theor. Appl. Climatol.*, 73, 67-86

J. B. Klemp and R. B. Wilhelmson, 1978: The simulation of three-dimensional convective storm dynamics, *J. Atmos. Sci.*, 35, 1070-1096

Lenz, C.-J. and U. Damrath, 2005: First Results on Verification of LMK Test Runs basing on Synop Data, *COSMO-Newsletter*, No. 6

Reinhardt, T. and A. Seifert, 2005: A Three-Category Ice Scheme for LMK, *COSMO-Newsletter*, No. 6

Theunert, F. and A. Seifert, 2005: Simulation Studies of Shallow Convection with the Convection-Resolving Version of the DWD Lokal-Modell *COSMO-Newsletter*, No. 6

Courant Number Independent Advection of the Moisture Quantities for the LMK

Jochen Förstner, Michael Baldauf and Axel Seifert

Deutscher Wetterdienst, Offenbach am Main, Germany

1 Introduction

The main part of the development of the two-timelevel dynamical core for the very short-range forecast model LMK was presented in the *COSMO Newsletter No. 4* (Förstner and Doms 2004; Doms and Förstner 2004). It follows closely the ideas of Wicker and Skamarock (2002) for the WRF model. The two most noteworthy differences are an option to use a truely third order Runge-Kutta scheme in the LMK, which is known in the literature (Liu et.al. 1994) as total variation diminishing (TVD) and, maybe more important, the formulation of the basic model equations. While in the WRF model the conservation form is used, in the LMK, which is based on the original LM code, we use the advection form. Idealized tests of the new dynamical core show good results (Baldauf et.al. 2005) and the TVD variant is used as default setting. While the integration of the dynamics is costly, we are able to use a bigger time step (30 s at 2.8 km) which leads in the end to a better efficiency than with the old Leapfrog core. Although the wind speed is restricted to 111 m/s, we have to deal with Courant numbers bigger than one — for the vertical this is the case anyway.

In the beginning of the LMK project stable integration of the prognostic equations for the moisture quantities (q_x) was realized by doing two sequential Eulerian advection steps with half the time step of the dynamics to ensure Courant numbers less than one. For the advection a positive definite but rather diffusive flux correction method with a monotonized-centered limiter was used.
To be able to do the advection of q_x in a single step, different methods for Courant number independent (CRI) transport have been implemented. Semi-Lagrange advection of the precipitating quantities was realized for the operational Leapfrog version. To retain positive definiteness (PD) of the scheme only tri-linear interpolation is used. While this leads to extreme numerical diffusion this might not be a problem for the precipitating fields. But for water vapor, cloud water and cloud ice this is not sufficient. Therefore in the combination with the Runge-Kutta core tri-cubic interpolation is used for these quantities, but the higher order is payed with the loss of PD.

It is also possible to construct CRI versions Eulerian schemes in flux-form and the way to do this is described. As an alternative to the semi-Lagrangian transport, in this paper we present a CRI version of Bott's advection scheme which has the advantage of being PD.

With respect to the quantitative precipitation forecast a new way to couple microphysics and dynamics, which is described in the latest WRF documentation (Skamarock et.al. 2005a), was recently implemented for the LMK and seems to be quite promising so far. This new variant is discussed briefly and first results are shown below.

2 Splitting, Conservation and Courant Number Independent Eulerian Advection

Skamarock (2005b) summarizes and presents ideas which combine several desirable properties of transport schemes used up to day. He concentrates on Eulerian schemes in flux-form and addresses splitting errors, mass conservation, Courant number restriction and positive definiteness or even monotonicity. To reduce the splitting error associated with multidimensionality we use the Strang-splitting technique

$$q_x^{t+\Delta t} = (I + A_z)(I + A_y)(I + A_x)\; q_x^t \tag{1}$$

$$q_x^{t+2\Delta t} = (I + A_x)(I + A_y)(I + A_z)\; q_x^{t+\Delta t} \tag{2}$$

as a good and cheap compromise, where the sequence of the Marchuk-splitting of the different directions is reversed for each new time step. Here I is the identity operator and the A's are the advection operators in the different spatial directions.

Since the model equations of LM/LMK are formulated in advection form

$$\frac{\partial q_x}{\partial t} + \vec{v} \cdot \nabla q_x = \ldots, \tag{3}$$

we have to deal with the problem that the schemes in mind compute the divergence of the flux. Different options to do this are implemented. The first ("Flux Form - DIV") is to subtract the missing term which includes the 3D wind divergence

$$\frac{\partial q_x}{\partial t} + \nabla \cdot (\vec{v} q_x) - q_x \nabla \cdot \vec{v} = \ldots, \tag{4}$$

the second is to switch to the "Conservation Form"

$$\frac{\partial \rho q_x}{\partial t} + \nabla \cdot (\rho \vec{v} q_x) = \ldots \tag{5}$$

for the moisture transport by multiplying the specific values with the air density:

$$\rho_x^{(n)} = \rho^{(n)} q_x^{(n)}. \tag{6}$$

Afterwards the resulting moisture densities are transported

$$\begin{aligned} \rho_x^{(*)} &= \rho_x^{(n)} \;\; + \Delta t\; A_x\big(\rho_x^{(n)}\big) \\ \rho_x^{(**)} &= \rho_x^{(*)} \;\; + \Delta t\; A_y\big(\rho_x^{(*)}\big) \\ \rho_x^{(n_{adv})} &= \rho_x^{(**)} + \Delta t\; A_z\big(\rho_x^{(**)}\big) \end{aligned} \tag{7}$$

and in the end we have to transform them back to the specific values

$$q_x^{(n+1)} = \frac{\rho_x^{(n_{adv})}}{\rho^{(n_{adv})}}. \tag{8}$$

To do this we need the updated $\rho^{(n_{adv})}$ and two different variants for its calculation are implemented in the LMK. Either to use the following diagnostic relation based on the equation of state:

$$\rho^{(n_{adv})} = \frac{p_0 + p^{*(n+1)}}{R_d T^{(n+1)}} - \left(\left(\frac{R_v}{R_d} - 1 \right) \rho_v^{(n_{adv})} - \rho_c^{(n_{adv})} - \rho_s^{(n_{adv})} \right) \tag{9}$$

where we use the updated $\rho^{(n_{adv})}$ in the remaining calculations of the time step. Or to do a prognostic calculation of the continuity equation

$$\begin{aligned} \rho^{(*)} &= \rho^{(n)} \;\; + \Delta t\; A_x\big(\rho^{(n)}\big) \\ \rho^{(**)} &= \rho^{(*)} \;\; + \Delta t\; A_y\big(\rho^{(*)}\big) \\ \rho^{(n_{adv})} &= \rho^{(**)} + \Delta t\; A_z\big(\rho^{(**)}\big) \end{aligned} \tag{10}$$

where the air density is advected in the same way as the moisture densities (7). In this case $\rho^{(n_{adv})}$ is used for (8) only, afterwards we continue to use $\rho^{(n)}$.

Our results so far showed that the prognostic variant (10) is to be preferred and all simulations shown below for which the conservation form was used are done in this way.

The basic idea to remove the Courant number restriction of advection schemes which use forward-in-time differencing is described, e.g. by Lin and Rood (1996). During the calculation, the fluxes at the cell interfaces are split into their integer (12) and fractional (13) part

$$Cr_{j+\frac{1}{2}} = u_{j+\frac{1}{2}} \frac{\Delta t}{\Delta x} \tag{11}$$

$$K_{j+\frac{1}{2}} = \mathrm{INT}(Cr_{j+\frac{1}{2}}) \tag{12}$$

$$Cr'_{j+\frac{1}{2}} = \mathrm{MOD}(Cr_{j+\frac{1}{2}}, K_{j+\frac{1}{2}}) \tag{13}$$

and only the fractional part is treated by a standard flux-form advection scheme which is restricted to $Cr \leq 1$:

$$f'_{j+\frac{1}{2}} = f'(Cr'_{j+\frac{1}{2}}, j - K_{j+\frac{1}{2}}) \tag{14}$$

$$\frac{\Delta t}{\Delta x} F_{j+\frac{1}{2}} = \begin{cases} \sum_{k=1}^{K_{j+\frac{1}{2}}} \rho_{x\,j-k+1}, & K_{j+\frac{1}{2}} \geq 1 \\ 0, & K_{j+\frac{1}{2}} = 0 \\ \sum_{k=-1}^{K_{j+\frac{1}{2}}} \rho_{x\,j-k}, & K_{j+\frac{1}{2}} \leq -1. \end{cases} \tag{15}$$

The total flux at a cell boundary is then given as the sum of the integer flux $F_{j+\frac{1}{2}}$ (15) and the fractional flux $f'_{j+\frac{1}{2}}$ (14).

3 CRI Version of Bott's Positive Definite Advection Scheme

Several Eulerian flux-form advection schemes are now implemented in a CRI way in the LMK, i.e. for the calculation of the fractional flux (14): First a van Leer-type scheme with a monotonized centered flux limiter (van Leer 1977), second the PPM scheme (with no flux limitation) used in Skamarock (2005b) — while the former is even monotone, it is also rather diffusive, whereas the latter is indeed less diffusive, but lacks the property of being positive definite. Therefore, as a third scheme with low numerical diffusion, we concentrate on the positive definite version of Bott's (1989a, 1989b) integrated flux form method.

The Bott scheme is realized in two variants with either second or fourth order polynomials and the procedure starts with the calculation of the integrated (fractional) fluxes given by

$$I^{+}_{j+\frac{1}{2}} = \sum_{k=0}^{l=2/4} \frac{\alpha_{j-K_{j+\frac{1}{2}},k}}{(k+1)2^{k+1}} \left[1 - \left(1 - 2Cr'^{+}_{j+\frac{1}{2}} \right)^{k+1} \right] \tag{16}$$

$$I^{-}_{j+\frac{1}{2}} = \sum_{k=0}^{l=2/4} \frac{\alpha_{j+1-K_{j+\frac{1}{2}},k}}{(k+1)2^{k+1}} \left[1 - \left(1 + 2Cr'^{-}_{j+\frac{1}{2}} \right)^{k+1} \right] (-1)^{k} \tag{17}$$

$$I_j = \sum_{k=0}^{l=2/4} \frac{\alpha_{j-K_{j+\frac{1}{2}},k}}{(k+1)2^{k+1}} \Big[(-1)^k + 1\Big] \tag{18}$$

where the $\alpha_{j,k}$ are the mentioned polynomials of order l which are listed in Bott (1989b, Table 1). Equation (16) is the flux for positive $Cr\,'$, (17) is the flux for negative $Cr\,'$ and (18) is the flux corresponding to $Cr = 1.0$ which is used for normalization.

Applying the following constraints for positive definiteness of the scheme

$$i^+_{j+\frac{1}{2}} = \max\Big(0, I^+_{j+\frac{1}{2}}\Big) \tag{19}$$

$$i^-_{j+\frac{1}{2}} = \max\Big(0, I^-_{j+\frac{1}{2}}\Big) \tag{20}$$

$$i_j = \max\Big(I_j, i^+_{j+\frac{1}{2}} + i^-_{j-\frac{1}{2}} + \epsilon\Big) \tag{21}$$

the fractional flux at a cell boundary is finally given by

$$f\,'_{j+\frac{1}{2}} = \frac{\Delta x}{\Delta t}\left[\frac{i^+_{j+\frac{1}{2}}}{i_j}\psi_{j-K_{j+\frac{1}{2}}} - \frac{i^-_{j+\frac{1}{2}}}{i_{j+1}}\psi_{j+1-K_{j+\frac{1}{2}}}\right]. \tag{22}$$

When we compare the CPU time for a real case simulation, where advection is calculated for all six prognostic moisture variables q_v, q_c, q_i, q_r, q_s and q_g, while the fourth order version is slightly more expensive, the second order one is comparable to the semi-Lagrangian transport in this respect. All in all with the CRI version of the Bott scheme we get a high order positive definite scheme at a reasonable computational cost.

4 Tri-cubic Semi-Lagrange Advection

A Semi-Lagrange (SL) scheme for the advection equation (3) is usually performed in two steps: first in estimating the point, from where a fluid particle started at the beginning of the time step, i.e. in calculating a backtrajectory and second in interpolating properties at the starting point from the neighboring grid points (Staniforth and Coté 1991).

The backtrajectory of 2^{nd} order is calculated as described in Baldauf and Schulz (2004). The result is a shift vector of the starting point relative to the actual grid point. The integer part of this shift vector delivers the grid position of the interpolation polynomial and the fractional part delivers the interpolation weights (for consistency with the tri-linear SL-routine these weights lie between -1 and 0). Therefore the following interpolation takes place in the transformed (or index) space.

For the tri-cubic interpolation a polynomial $p(x, y, z)$ is searched with the property $p(i, j, k) = q_{i,j,k}$, where i, j, $k = -2, -1, 0, 1$, and $q_{i,j,k}$ is the value at the appropriate grid point. This problem can be reduced to the estimation of polynomials $P_k(x)$ of only one variable x with the property

$$P_i(j) = \begin{cases} 1 & : \quad i = j \\ 0 & : \quad i \neq j \end{cases}, \qquad i, j = -2, -1, 0, 1, \tag{23}$$

which is fulfilled by

$$P_{-2}(x) \;=\; \frac{1}{6}\,(x+1)\;x\;(x-1), \tag{24}$$

$$P_{-1}(x) \;=\; -\frac{1}{2}\,(x+2)\;x\;(x-1), \tag{25}$$

$$P_0(x) \;=\; \frac{1}{2}\,(x+2)\;(x+1)\;(x-1), \tag{26}$$

$$P_1(x) \;=\; -\frac{1}{6}\,(x+2)\,(x+1)\,x. \tag{27}$$

The polynomial p can then be constructed by

$$p(x,y,z) = \sum_{i,j,k=-2}^{1} P_i(x)\;P_j(y)\;P_k(z)\;q_{i,j,k}, \tag{28}$$

where x, y, z are the interpolation weights. The sum (28) runs over the neighboring 64 grid points of the backtrajectory starting point. These are much more points than are needed to construct a polynomial of only 3^{rd} order in 3 dimensions; for this task only 20 grid points would be necessary. But the high symmetry of equation (28) allows a quick way for calculating the sum. Therefore this SL variant is rather efficient and nevertheless possesses rather good transport properties. Another advantage of this SL-method is, that it calculates the interpolation in three dimensions in one step, therefore no splitting error occurs. However a disadvantage is, that the method can produce negative undershoots for a positive definite field. Clipping of these negative values can destroy the rather good conservation properties of this SL-method. To reduce sharp edges, which can produce stronger undershoots, a smoothing filter is applied before clipping.

5 New Physics-Dynamics-Coupling

A special procedure to couple microphysics and dynamics is described in the latest WRF documentation (Skamarock et.al. 2005a). A similar numerical treatment of latent heat is now implemented in the Runge-Kutta core of the LMK. In both models the microphysic parameterization is calculated in one Eulerian time step after the Runge-Kutta integration of the dynamical core and the Marchuk-splitting method is used here to finally update the fields in a balanced way and complete the time step.

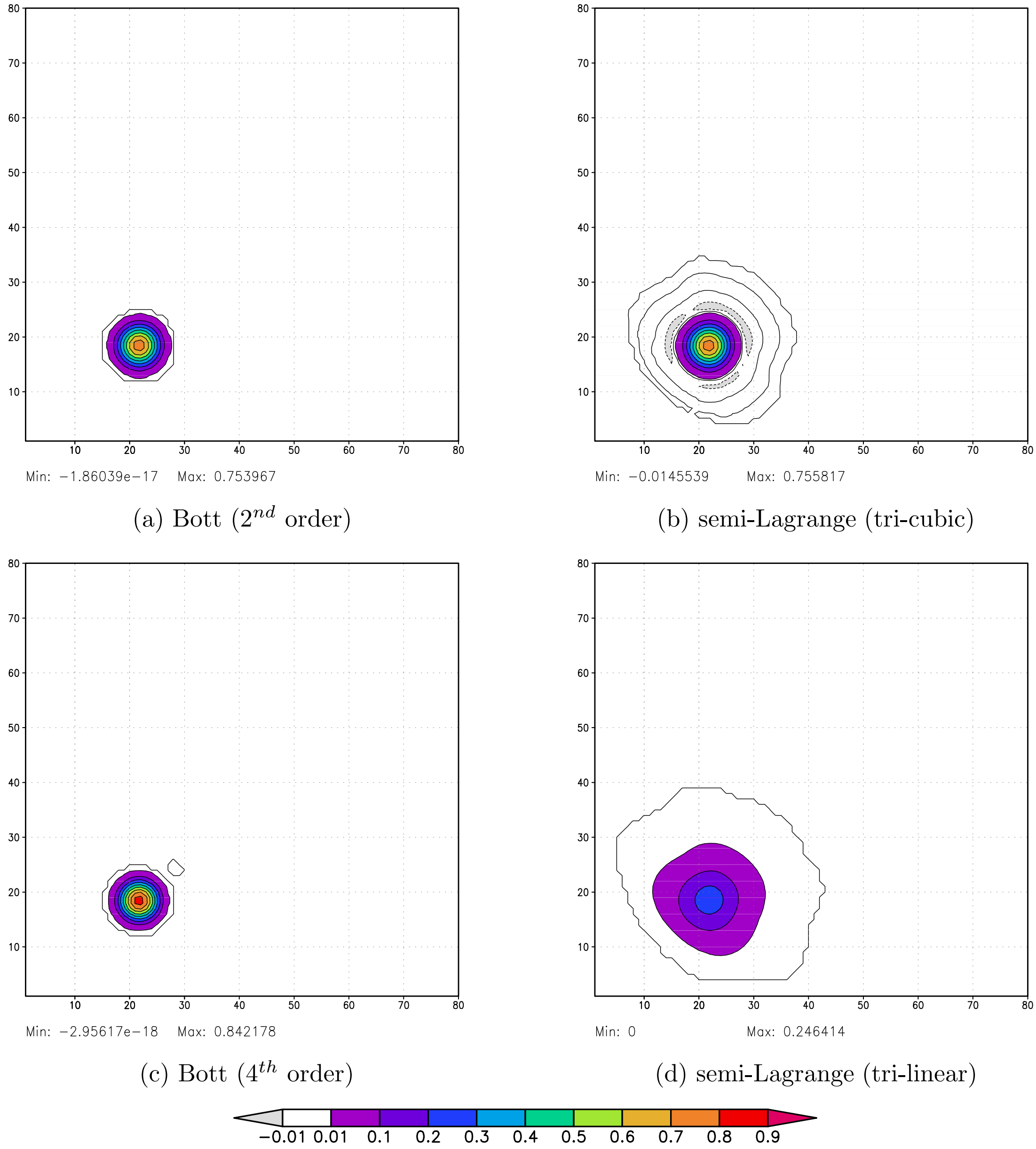
The diabatic heating term in the prognostic equation for temperature Q_T in the LM(K) includes the physical tendencies due to radiation, convection, turbulent mixing and latent heat conversion in the microphysics (e.g. the saturation adjustment). While the former tendencies are integrated within the acoustic steps the last temperature tendency has not been part of this integration.

The new variant now uses the tendency due to latent heat conversion of the previous time step as an estimate which enters the integration of the fast waves in the same way than the other physical tendencies. After the Runge-Kutta update, this tendency estimate is subtracted again and the time step is completed as before.

6 Idealized Advection Tests

To verify the correct implementation and for comparison of the different schemes idealized advection tests were carried out with the LMK.
As a first and also very common test, the solid body rotation of a tracer cone was simulated (Fig. 1). In the case shown, the angular velocity was rather high, leading to Courant Numbers well above one.
All of the four schemes perform equally well, at least in two aspects: First the Courant number independence is fulfilled, since in each case we get a stable integration. Second the circular shape of the cone is more or less retained, i.e. errors due to the directional splitting method used for the Bott scheme are negligible. For a further investigation of this last property idealized tests utilizing a deformational flow field are planned.

The Bott scheme using 4^{th} order polynomials (Fig. 1(c)) performs best. It is positive definite (up to an ε) and the maximum value is retained well, that is the scheme shows only weak numerical diffusion. For the 2^{nd} order Bott scheme (Fig. 1(a)) we get a similar result with an only slightly smaller maximum, which is almost equal to the one we get for semi-Lagrangian variant using tri-cubic interpolation (Fig. 1(b)). But the higher order interpolation in this case leads to negative undershoots, i.e. the loss of positive definiteness of the scheme. While this does not happen when tri-linear interpolation (Fig. 1(d)) is used, the numerical diffusion of this scheme is extremely large and perhaps comparable to a 1^{st} order upwind scheme.



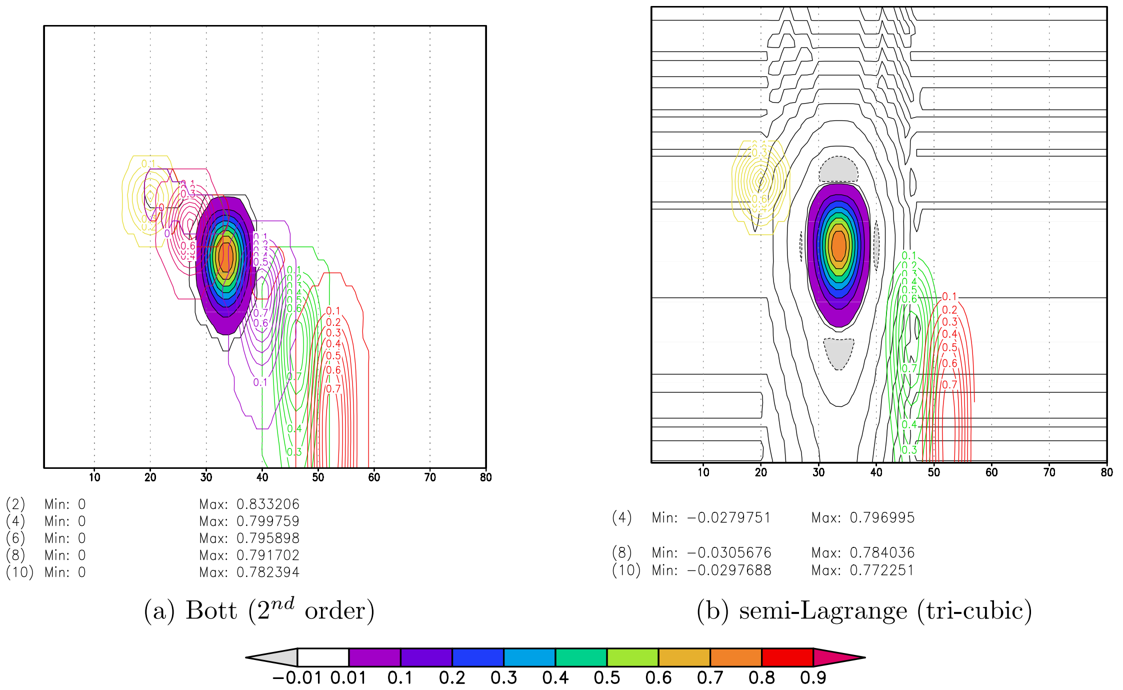
(a) Bott (2^{nd} order) (b) semi-Lagrange (tri-cubic)

(c) Bott (4^{th} order) (d) semi-Lagrange (tri-linear)

Figure 1: Solid body rotation of a tracer cone with an initial maximum of 1.0. Results after 80 time steps equivalent to one revolution. The Courant Number in the domain is close to a value of three near the lateral boundaries and approximately a value of 2.2 at the center of the cone.

From now on we will concentrate on two schemes, namely the Bott scheme formulated with the 2^{nd} order polynomials and the semi-Lagrangian scheme using the tri-cubic interpolation to calculate the values at the starting point of the backtrajectory.

In a second test the tracer cone is advected in x- and z-direction (Fig. 2). The 35 layers used are stretched in the vertical as it was done in the once operational 7 km version of LM. This is a good test of the metrics in such a configuration.
The statements made for the solid body rotation test apply here in an analogous way. The increasingly ellipsoidal shape as the tracer cone is advected downwards is a drawing artefact. A real problem with the directional splitting would give rise to a tilted structure.

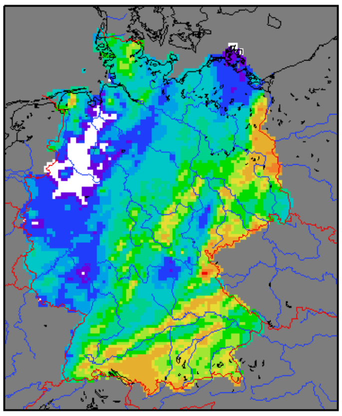


(a) Bott (2^{nd} order) (b) semi-Lagrange (tri-cubic)

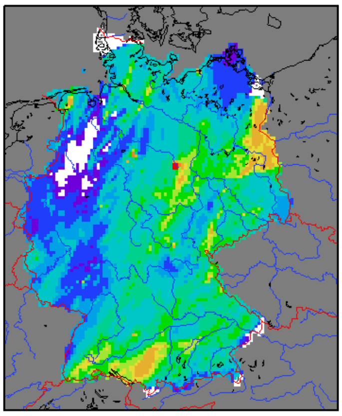
Figure 2: xz-Advection with $u = 150$m/s and $w = -10$m/s of a tracer cone with an initial maximum of 1.0. Simulation times in minutes are given in parentheses. A vertically stretched grid is used which leads to values of $Cr_z = w\frac{\Delta t}{\Delta z}$ exceeding one near the bottom boundary. (The vertically stretched grid is not taken into account when directly plotting the model layers in GrADS.)

7 Real Case Studies

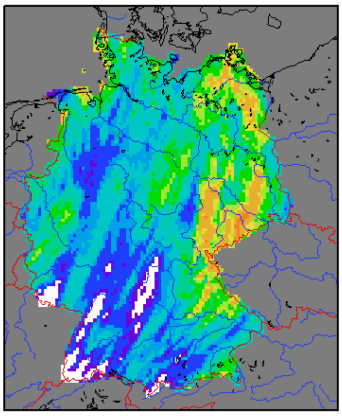
In real case studies we experience a big sensitivity of the model with respect to the different choices for moisture advection and the way in which we couple dynamics and the micro-physic parameterization. As two examples Fig. 3 and Fig. 4 show a comparison of the 24 h precipitation over Germany for different advection schemes and observations. The data for each plot is aggregated to the LM grid with a resolution of 7 km.
The results differ significantly in the mean, their maximum and the precipitation pattern itself – especially when we look at the Southern border of Germany. The 24 h precipitation shows us to a certain degree the "integrated" performance of the model.
But the problems start when we look at the given observational data (Figs. 3(a) and 3(b), respectively 4(a) and 4(b)). The uncorrected radar data show obvious erroneous high values, e.g. in the vicinity of the radar sites, but very low values for the overall precipitation mean. Therefore we take the quality controlled data of the high-density rain-gauge observation network as the "truth".



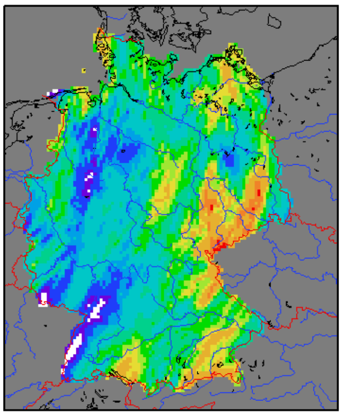
(a) High-Density Rain-Gauge Observation (= ”truth”)
Mean: 12.62 mm   Max.: 90.16 mm



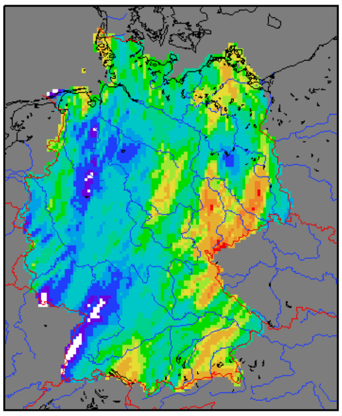
(b) Radar Observation (uncorrected)
Mean: 9.51 mm   Max.: 120.47 mm



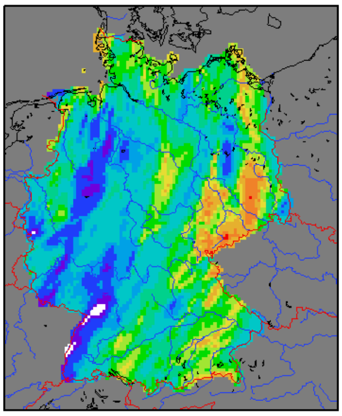
(c) Old Physics-Dynamics-Coupling: Bott (2^{nd} order) in Flux Form - DIV
Mean: 10.51 mm   Max.: 69.79 mm



(d) Old Physics-Dynamics-Coupling: semi-Lagrange
Mean: 11.84 mm   Max.: 87.48 mm

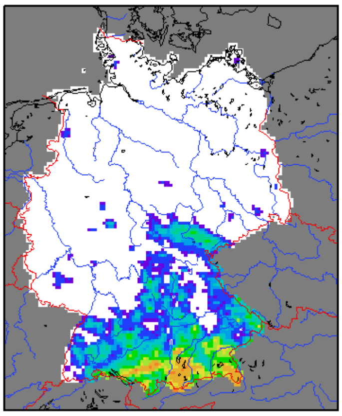


(e) New Physics-Dynamics-Coupling: Bott (2^{nd} order) in Conservation Form
Mean: 14.62 mm   Max.: 94.56 mm

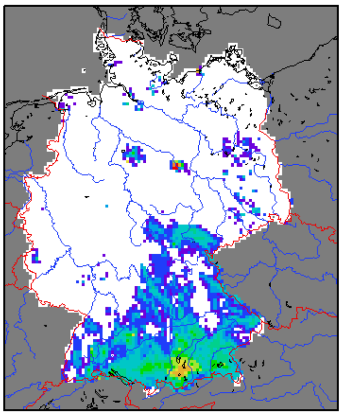


(f) New Physics-Dynamics-Coupling: semi-Lagrange
Mean: 13.19 mm   Max.: 88.48 mm

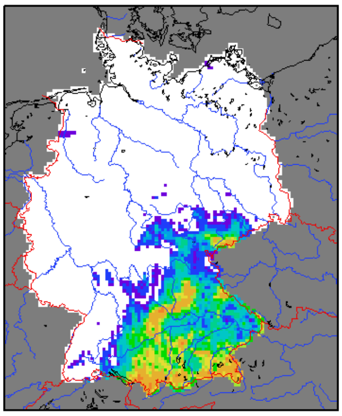
Figure 3: Precipitation 07/08/2004 – 6:00 UTC + 24 h (Figures by Thorsten Reinhardt).



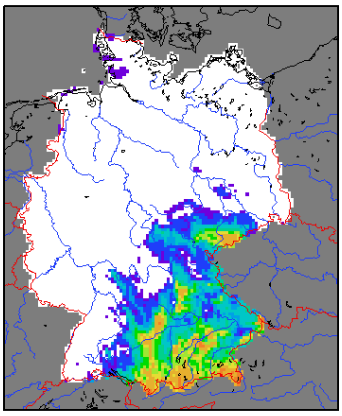
(a) High-Density Rain-Gauge Observation (= ”truth”)
Mean: 2.54 mm Max.: 63.60 mm



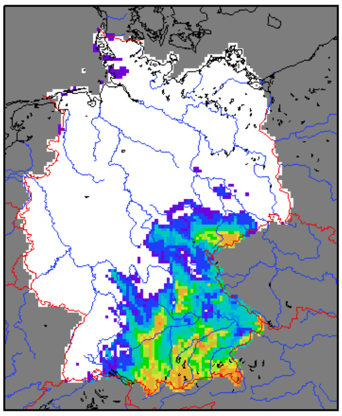
(b) Radar Observation (uncorrected)
Mean: 1.63 mm Max.: 139.28 mm



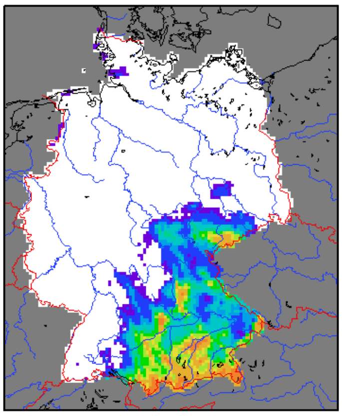
(c) Old Physics-Dynamics-Coupling: Bott (2^{nd} order) in Flux Form - DIV
Mean: 3.72 mm Max.: 61.04 mm



(d) Old Physics-Dynamics-Coupling: semi-Lagrange
Mean: 4.30 mm Max.: 138.36 mm



(e) New Physics-Dynamics-Coupling: Bott (2^{nd} order) in Conservation Form
Mean: 3.54 mm Max.: 77.23 mm



(f) New Physics-Dynamics-Coupling: semi-Lagrange
Mean: 3.77 mm Max.: 74.92 mm

Figure 4: Precipitation 07/24/2004 – 6:00 UTC + 24 h (Figures by Thorsten Reinhardt).

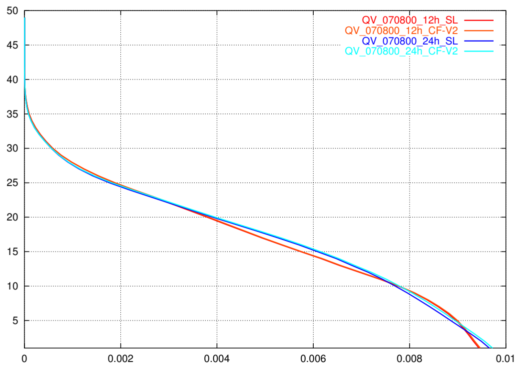
For both cases the two plots in the middle row show results obtained with the old coupling, whereas the plots in the bottom row were obtained using the new coupling.

For the semi-Lagrangian runs this is the only change, but in the runs with Bott's Eulerian scheme a second change was made, namely a switch from the non-conservative flux-form (4) to the conservation form (5).

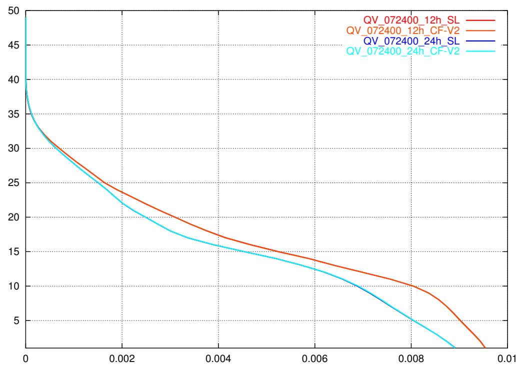
Most noteworthy when we look at the results for the 07/08/2004 (Fig. 3) is the difference between Fig. 3(c) and Fig. 3(e) at the Southern border of Germany. In the run using the non-conservation form we get too little precipitation compared to the rain-gauge observation. This is probably due to the fact, that the term $-q_x\nabla \cdot \vec{v}$ in (4) which is only calculated in 2^{nd} order has a great potential to deteriorate the conservation of the specific moisture quantities. This problem will be addressed in more detail in the next section.
For the two different semi-Lagrangian runs (Fig. 3(d) and Fig. 3(f)) it is not obvious which variant of coupling gives us the better result — at least in this case.

When we look at the results for the 24/07/2004 (Fig. 4) it is somehow the other way round. Now the results for the Bott scheme (Fig. 4(c) and Fig. 4(e)) are more similar to each other, while the runs with the semi-Lagrange scheme (Fig. 4(d) and Fig. 4(f)) differ significantly in their maximum values of precipitation near and in the alpine region.

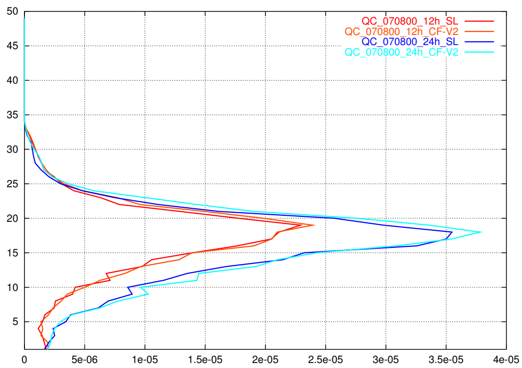
But for both cases the newer variants (in the bottom row) are at least as good or remarkably better than the old ones. And, what is important, the results for the Bott scheme and the semi-Lagrange scheme are very similar now.



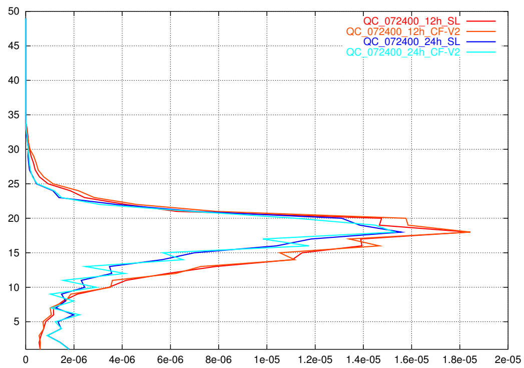
(a) q_v – 07/08/2004 00 UTC Run



(b) q_v – 07/24/2004 00 UTC Run



(c) q_c – 07/08/2004 00 UTC Run



(d) q_c – 07/24/2004 00 UTC Run

Figure 5: Mean vertical profiles of specific water vapor and cloud water [kg/kg] for forecast times of 12 (red) and 24 (blue) hours.
lighter colors: Bott (2^{nd} order) in Conservation Form / darker colors: semi-Lagrange

This can also be seen in Fig. 5, where the domain averaged vertical profiles of water vapor and cloud water are plotted for both new variants and both dates (forecast times of 12 and 24 hours). Especially the profiles of water vapor lie almost on top of each other.
In the profiles of cloud water we notice a small scale zigzag-structure, most notably in the second case. The reason for this structure is not clear yet, but one explanation might be the rather poor quality of the only 2^{nd} order implicit vertical advection scheme for the wind components, pressure perturbation and temperature in the dynamical core. This low order centered scheme produces rather large undershoots and is also not free of numerical dispersion. Here, higher order schemes might have a big potential to further improve the model.

8 Tracer Experiment

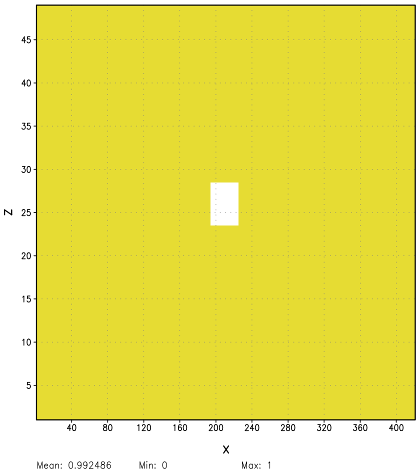
The reasons for the differences we have seen, have to be investigated and discussed further, but they are at least partly a result of the varying numerical diffusivity and (missing) positive definiteness / mass conservation properties of the schemes.

Therefore a further experiment has been performed. In this experiment a tracer is transported in the flow field of the real case study for the date 07/08/2004 (00 UTC run) discussed in the previous section.
The tracer was initialized with a value equal to one, with the exception of a small cuboidal area in the middle of the domain (Fig. 6(a)). In this way we get a structure which is characterized by sharp gradients in an otherwise homogeneous field. In addition a zero-gradient lateral boundary condition was used for the tracer.
As long as the tracer structure remains inside the domain, the integral tracer mass should be conserved.

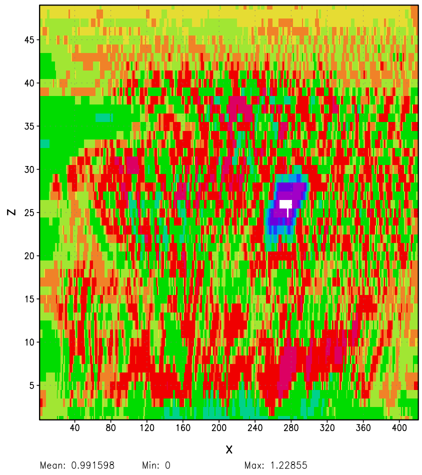
Figs. 6(b) to 6(d) show the tracer field after three hours of simulation time for different variants of advection. In each case, all moisture quantities were transported in the same way as the tracer, with the exception of q_r, q_s and q_g in the semi-Lagrangian case, for which the default tri-linear interpolation was used, i.e. the flow field differs from run to run for this reason. For all runs the new coupling of physics and dynamics was switched on and a clipping of negative tracer values was performed. This clipping acts as an artificial source if the overall advection scheme is not positive definite, and this is the case for all runs, with the only exception being the Bott scheme in conservation form (Fig. 6(c)) for which we get identical results for the runs with and without (not shown) clipping.

The result for the Bott scheme used in the non-conservative form given by Eq. (4) shows a tracer field with a lot of small scale disturbances over the whole domain. This pattern is due to the term including the 3D wind divergence, and a look at the divergence field (not shown here) shows similar small scale features especially pronounced over orographically structured terrain.
It is clear from this experiment, that it is probably not a good idea to use this form of advection, and it explains the deficiencies we have noticed in the precipitation pattern in the last section.

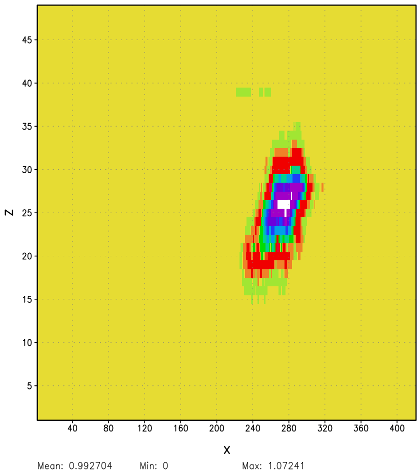
It was already stated that the conservative Bott scheme and the semi-Lagrangian scheme in the new form produce quite similar results. This is also true for the tracer field (Fig. 6(c) and 6(d)). A closer look shows some minor artifacts around level 39 in the plot for the Bott scheme. This is approximately the level where the Rayleigh damping layer begins, but a real explanation is still missing. It is also worth to notice, that in the homogeneous parts of the tracer field, the interpolation step in the semi-Lagrange scheme is quite trivial. On the other



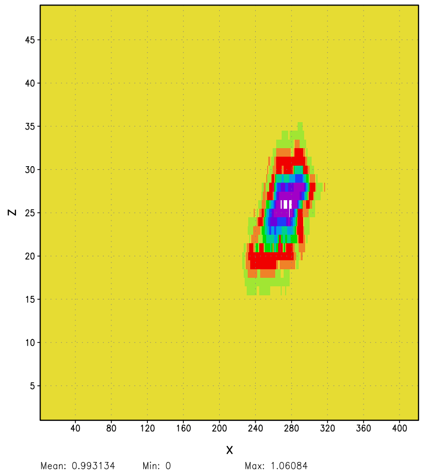
(a) Initialized tracer distribution at y-index $j = 230$ (middle of domain).



(b) Tracer distribution after 3 h at y-index $j = 260$ – Bott (2^{nd} order) in Flux Form - DIV with Clipping of Negative Values.



(c) as (b) – Bott (2^{nd} order) in Conservation Form.



(d) as (b) – semi-Lagrange (tri-cubic) with Clipping of Negative Values.

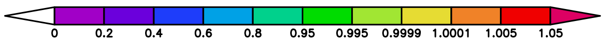


Figure 6: Advection of a tracer in a real case flow field (07/08/2004 00 UTC run). Shown are vertical cross sections of the tracer field. (The vertically stretched grid is not taken into account when directly plotting the model layers in GrADS.)

hand in the Bott scheme in conservation form, the mass specific tracer field is multiplied with the air density, and afterwards the resulting exponentially distributed tracer density has to be transported.

In Fig. 7 the normalized volume integral of the tracer field is given as time series over the first 360 steps. The red curve corresponds to the run given in Fig. 6(c), and the blue one corresponds to Fig. 6(d). In addition, shown in green is the result for a semi-Lagrangian run without the clipping of negative values. As stated in Sec. 4, the semi-Lagrange scheme with tri-cubic interpolation, for itself, has good conservation properties, and it follows closely the line of the conservative Bott scheme. But the clipping of undershoots — which arise near sharp gradients, to circumvent negative values, acts as an artificial source. As soon as this sharp gradients are diffused to a certain extend — in approximately the first 100 time steps

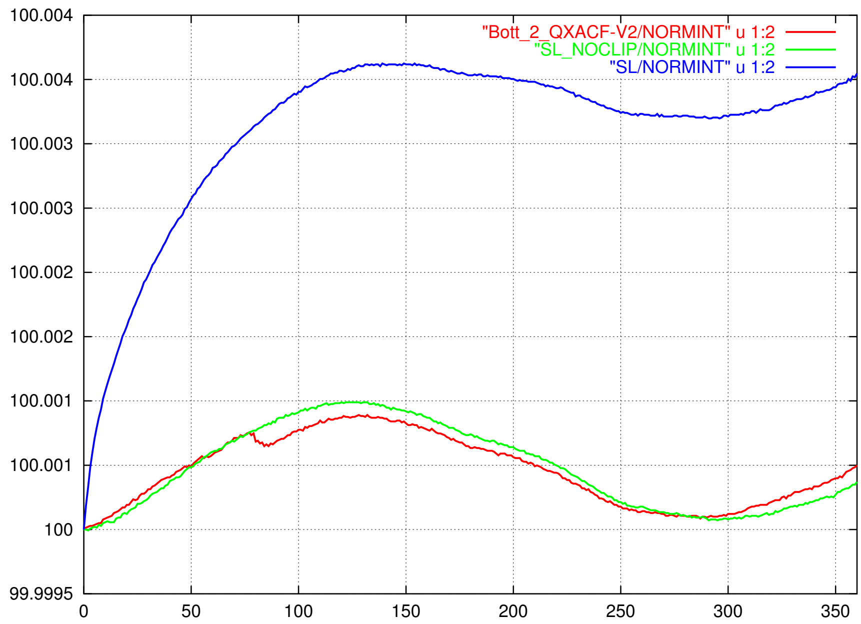


Figure 7: Time series of the normalized volume integral of the tracer field for the first three hours (360 time steps). The initialized value corresponds to 100 percent.
red: Bott (2^{nd} order) in Conservation Form / blue: semi-Lagrange (tri-cubic) with Clipping of Negative Values / green: semi-Lagrange (tri-cubic) without Clipping of Negative Values.

in this case, the curves for the runs with and without clipping stay more or less parallel.

9 Conclusions

The two different advection schemes, namely the semi-Lagrangian one and the Eulerian Bott scheme in conservation form, have been developed independently from each other. And the fact that they produce very similar results, at least in the cases investigated so far, makes us quite confident, with respect to their (finally?) correct implementation into the LMK. The most recent changes are not part of the official version 3.16 yet, but will come in the next release.
Both variants have their pros and cons — being positive definite or not, being a real 3D or only a split 1D scheme, as well as the question, if all quantities should be transported in the same way, or, as it is done in the semi-Lagrangian variant, to make a distinction between the precipitating and non-precipitating quantities.
Further tests, comparisons and verifications, especially for runs over longer periods, are on their way and will help us to decide, which one of the two schemes is to be preferred — if we see significant differences then.

10 References

Baldauf, M. and J.-P. Schulz (2004). Prognostic Precipitation in the Lokal Modell (LM) of DWD. *COSMO-Newsletter*, No. 4, 177–180. Deutscher Wetterdienst, Offenbach a. M., Germany. availabe at `http://www.cosmo-model.org`.

Baldauf, M., J. Förstner, and P. Prohl (2005). Idealized Tests of the Very Short-Range Forecast Model LMK. *WGNE Blue Book*, `http://www.cmc.ec.gc.ca/rpn/wgne/`.

Bott, A. (1989a). A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes. *Mon. Wea. Rev.*, 117, 1006–1015.

Bott, A. (1989b). Reply. *Mon. Wea. Rev.*, 117, 2633–2636.

Doms, G. and J. Förstner (2004). Development of a kilometer-scale NWP-system: LMK. *COSMO Newsletter*, No. 4, 159–167. Deutscher Wetterdienst, Offenbach a. M., Germany. availabe at `http://www.cosmo-model.org`.

Förstner, J. and G. Doms (2004). Runge-Kutta Time Integration and High-Order Spatial Discretization of Advection – A New Dynamical Core for the LMK. *COSMO Newsletter*, No. 4, 168–176. Deutscher Wetterdienst, Offenbach a. M., Germany. availabe at `http://www.cosmo-model.org`.

Lin, S. J. and R. B. Rood (1996). Multidimensional flux-form semi-Lagrangian transport schemes. *Mon. Wea. Rev.*, 124, 2046–2070.

Liu, X.-D., S. Osher and T. Chan (1994). Weighted essentially non-oscillatory schemes. *J. Comput. Phys.*, 115, 200–212.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers (2005a). A Description of the Advanced Research WRF Version 2. *NCAR Technical Note*, `http://www.wrf-model.org/wrfadmin/docs/arw_v2.pdf`. Boulder, Colorado, USA: Mesoscale and Microscale Meteorology Division NCAR.

Skamarock, W. C. (2005b). Positive-definite and montonic limiters for unrestricted-timestep transport schemes. accepted for *Mon. Wea. Rev.*,

Staniforth, A. and J. Coté (1991). Semi-Lagrangian integration schemes for atmospheric models — A review. *Mon. Wea. Rev.*, 119, 2206–2223.

van Leer, B. (1977). Towards the ultimate conservative difference scheme IV. A new approach to numerical convection. *J. Comp. Phys.*, 23, 276–299.

Wicker, L. J. and W. C. Skamarock (2002). Time-splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, 130, 2088–2097.

Sensitivity Experiments with the Runge Kutta Time Integration Scheme

Lucio Torrisi

CNMCA, via Pratica di Mare 45, 00040 Pomezia (RM), Italy

1 Introduction

The new dynamical core developed in LM (Förstner and Doms, 2004) is based on a TVD variant of 3rd-order Runge Kutta time integration scheme (RK) using a 5th-order spatial discretization of advection. The RK dynamical core should be more accurate than the standard Leap-Frog / 2nd-order advection scheme (LF) and it will be used for very detailed short range forecasts. The RK core needs to be tested and evaluated intensely, before it can be operationally implemented. This work aims to objectively evaluate the current version of the RK dynamics (LM 3.16+) through a comparison to the LF core and some sensitivity experiments in the period 24-28 March 2005. The LM configuration used in these experiments is shown in Tab. 1. The prognostic TKE turbulent scheme, the prognostic precipitation scheme, the new option of the upper level Rayleigh damping and a 72 s time step (40 s for LF runs) are used in the RK runs. IFS fields were used as initial and boundary conditions. The LM forecast fields were objectively evaluated through comparisons with radiosonde and conventional surface observations. Mean error (ME or bias) and root mean square error (RMSE) vertical profiles are computed for temperature and wind. Surface parameters, such as two meter temperature (2T), two meter dew point (2TD), ten meter wind speed (10U), mean sea level pressure (MSLP) and 6h accumulated total precipitation larger than 2 mm (6TP_2) are verified for Synop stations satisfying the COSMO WG5 specification ($|H_s - H_n| < 100m$), where H_s is the station height and H_n is the height of nearest land grid point). About 3500 forecast-observation pairs are used to calculate the ME and RMSE of the surface variables forecast. They are considered enough to make statistical comparison between different configurations of LM.

Table 1: **LM Configuration (Version 3.16+)**

Domain Size	465 × 385 grid points (EuroLM)
Horizontal Grid Spacing	0.0625° ($\sim 7\ km$)
Number of Layers	35
Time Step and Integration Scheme	72 s (RK) — 40 s (LF)
Forecast Range	24 h
Initial Time of Model Runs	00 UTC
Lateral Boundary and Initial Conditions	from IFS
L.B.C. update frequency	3 h
Orography:	filtered (`eps`=0.1)
Turbulence parameterization	prognostic TKE

2 LF and RK runs with the prognostic TKE turbulence scheme

Temperature and wind ME and RMSE vertical profiles for T+24 h forecast are shown in Fig. 1 for RK (red lines) and LF (blue lines) runs. The RK temperature RMSE is slightly

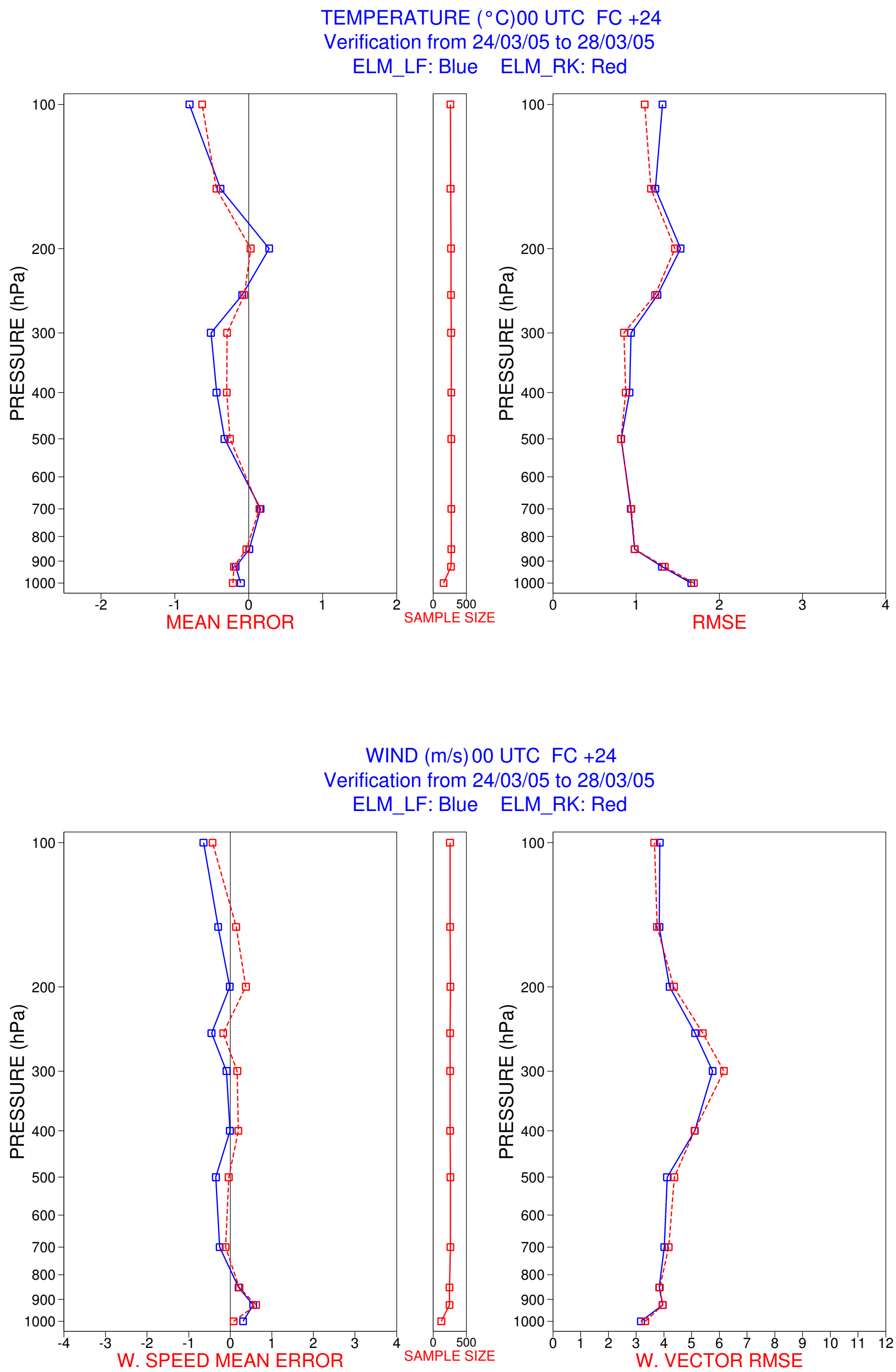


Figure 1: LF and RK runs with prognostic TKE turbulence: temperature mean error and root mean square error vertical profiles of 24 h forecast; wind speed mean error and wind vector root mean square error vertical profiles of 24 h forecast. (2 files: *024*lf-rk.ps)

smaller than the LF one above 500 hPa, while the RK wind vector RMSE is larger than the LF one at almost all levels. The interpretation of these results has to be done with caution, because of the small number of observation-forecast pairs used. 2T, 2TD, 10U, MSLP bias and RMSE (and standard deviation - STDV for MSLP) as a function of the forecast time are shown in the Fig. 2 for RK and LF runs. Frequency bias index (FBI) and threat score (TS) for 6TP_2 are also computed (Fig. 2). A very slight worsening is found in RK forecasts for 2TD, 10U and 6TP_2 (no significant difference in 2T). The RK MSLP bias is larger than the LF one leading to a worsening in MSLP forecast skill of the RK dynamics. This large difference in the MSLP bias is the most important result of the LF-RK comparison.

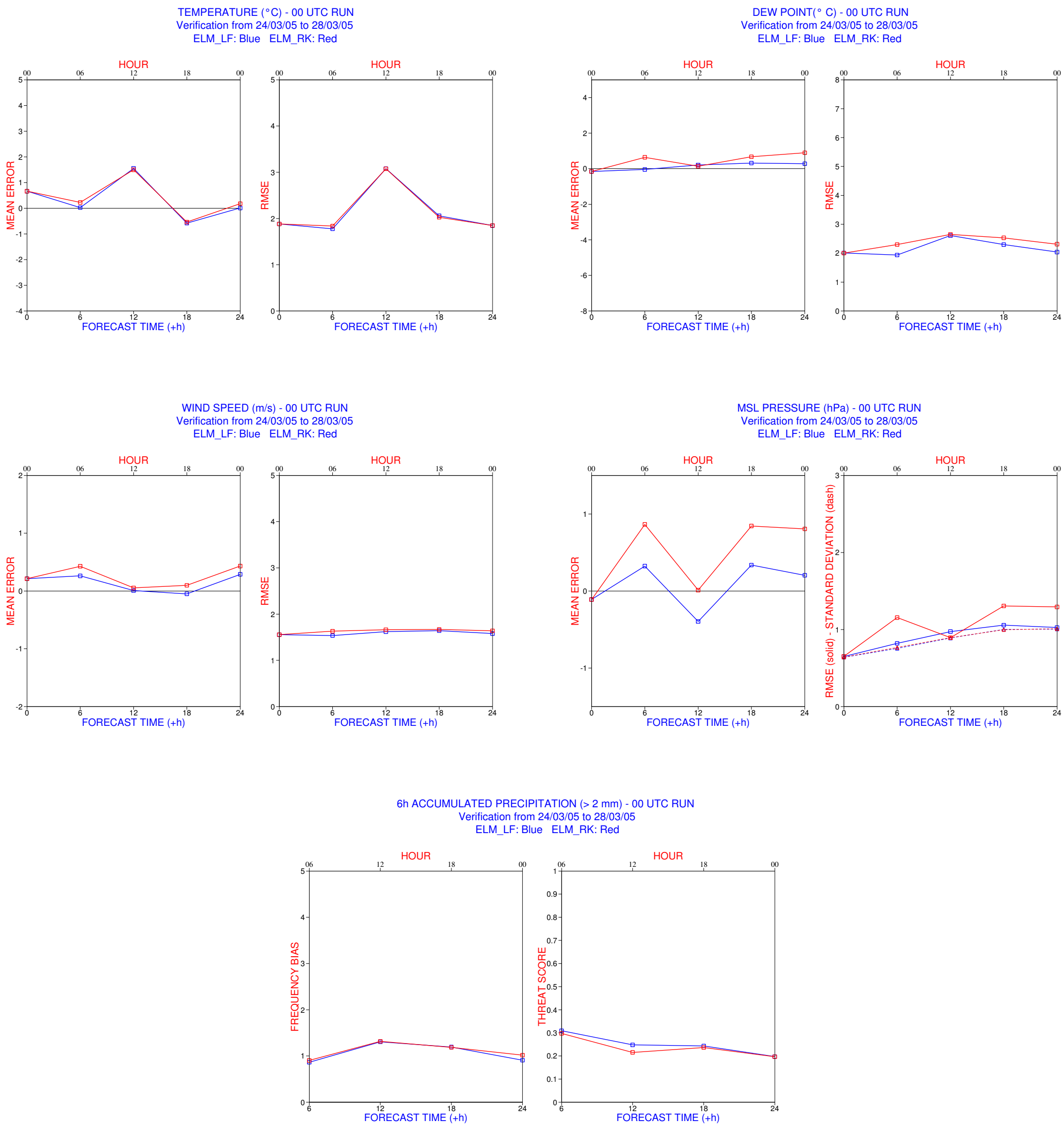


Figure 2: LF and RK runs with prognostic TKE turbulence: mean error and root mean square error of two meter temperature, two meter dew point, ten meter wind speed and mean sea level pressure (also standard deviation for MSLP); frequency bias index and threat score of 6h accumulated total precipitation larger than 2 mm.

3 Sensitivity experiments

The sensitivity of RK dynamics to the integration time step, the interval between two calls of (convection, turbulence) parameterization, the turbulence scheme, the domain size and moisture variables advection scheme is also investigated. Verification plots are shown if significant differences are found.

3.1 Integration time step

RK and LF runs use different time steps (72 and 40 s respectively) that determine a different time interval between two calls of physics. In this experiment the RK runs are performed with the same time step of the LF one, in order to exclude the difference in the parameterizations calling frequency as a possible cause of the LF-RK difference in the forecast skill (MSLP deficiency in RK). Mean sea level pressure ME, RMSE and STDV as a function of the forecast step are shown in Fig. 3 for RK runs with 40 s (red lines) and 72 s (blue lines) time step. RK forecasts with 40 s time step have a slightly smaller MSLP bias than that of RK forecasts with 72 s time step (except for 12h forecast). This result seems to be due to the higher accuracy associated with the smaller time step.

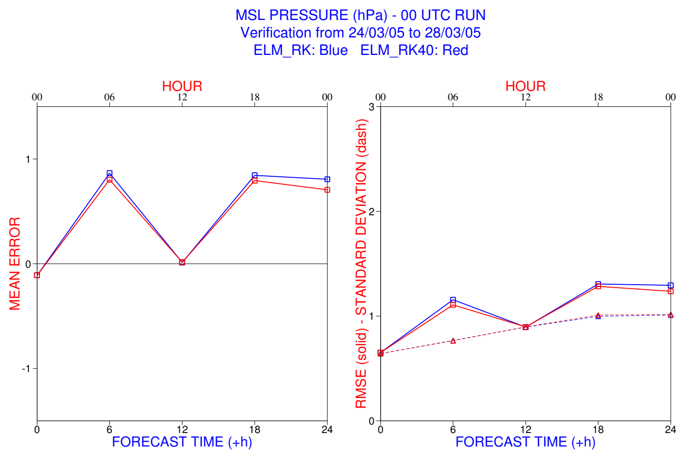


Figure 3: Mean error, root mean square error and standard deviation of mean sea level pressure for RK and 40s RK runs.

3.2 Parameterizations calling frequency

To totally exclude the influence of the interval between two calls of parameterization schemes on the MSLP deficiency in RK runs, other two experiments are useful. One experiment is to decrease the convection calling frequency nincconv from 10 (default for previous experiments) to 5. The other experiment is to increase the interval between two calls of the prognostic TKE turbulence scheme `ninctura` from 1 (default for previous experiments) to 2. In both experiments no significant difference is found from the reference RK runs.

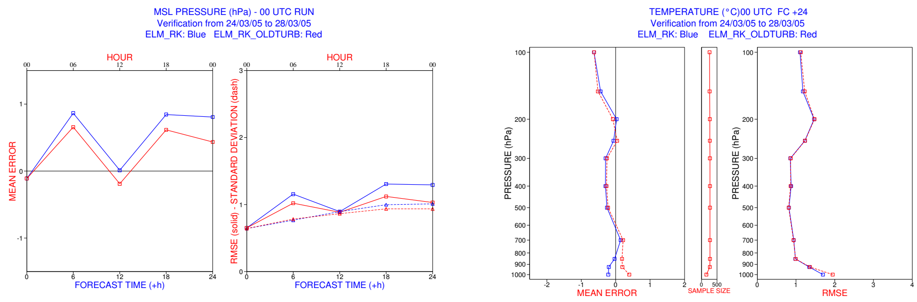


Figure 4: RK runs with diagnostic and prognostic TKE turbulence: mean error, root mean square error and standard deviation of mean sea level pressure; temperature mean error and root mean square error vertical profiles of 24 h forecast.

3.3 Turbulence scheme

The sensitivity of the RK core to the two turbulent schemes implemented in the LM, the prognostic (`itype_turb=3, imode_turb=1, itype_tran=2`) and the diagnostic (`itype_turb=1, imode_turb=0, itype_tran=1`) TKE, is also investigated. Mean sea level pressure ME, RMSE and STDV as a function of the forecast step are shown in Fig. 4 for RK forecasts with diagnostic (red lines) and prognostic (blue lines) TKE turbulence scheme. Temperature ME and RMSE vertical profiles for 24 h forecast are also computed (Fig. 4). RK dynamics with the old turbulence scheme performs better for MSLP forecast (smaller bias and standard deviation) than RK dynamics with the prognostic TKE turbulence parameterization. On the other hand, 2T and 2TD of RK runs with the prognostic TKE turbulence scheme seem to have a slightly better skill (not shown). The reduction of the positive MSLP bias seems to be related to the low level positive temperature bias of RK with old turbulence scheme. The different temperature bias behaviour may be due to the different surface layer formulation associated with each turbulence scheme. The TKE prognostic scheme seems to be one of the possible candidates to justify the MSLP deficiency in RK runs, but a large positive MSLP bias is still present in RK forecasts with the old turbulence scheme. This result is an indication that more work is needed to tune the turbulence parameterization schemes (surface layer, exchange coefficients, etc.) and to improve the dynamics and physics coupling.

3.4 Domain size

The impact of the domain size on LM forecast was evaluated in Torrisi (2005) using the LF core. A similar experiment (much shorter period) is also performed for the RK dynamics using EuroLM and LAMI (smaller) domain. Mean sea level pressure ME, RMSE and STDV as a function of the forecast step are shown in Fig. 5 for LAMI (red lines) and EuroLM (blue lines) domain. The enlargement of the domain size has a positive impact (smaller bias and standard deviation) on the MSLP forecast, as found for LF in Torrisi (2005). This could be related to the improvement of the intrinsic variability of the numerical model associated with the enlargement of the domain, since the boundary conditions fields are slightly affecting the forecast.

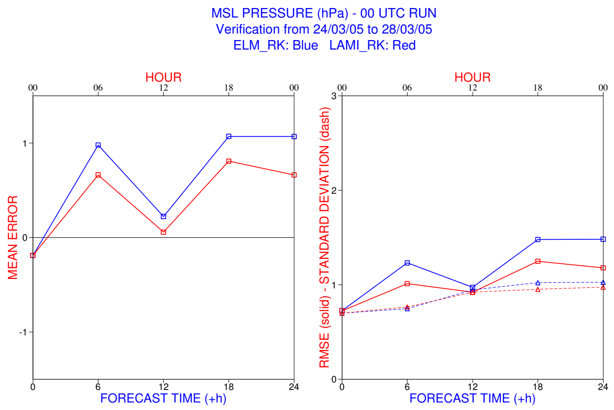


Figure 5: Mean error, root mean square error and standard deviation of mean sea level pressure for LAMI and EuroLM domain.

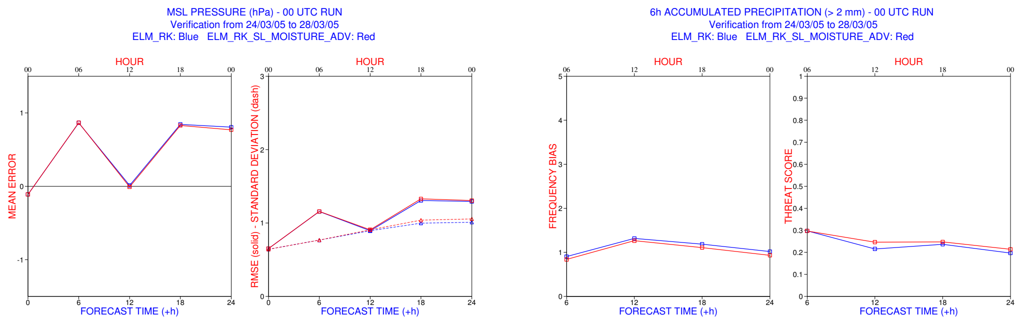
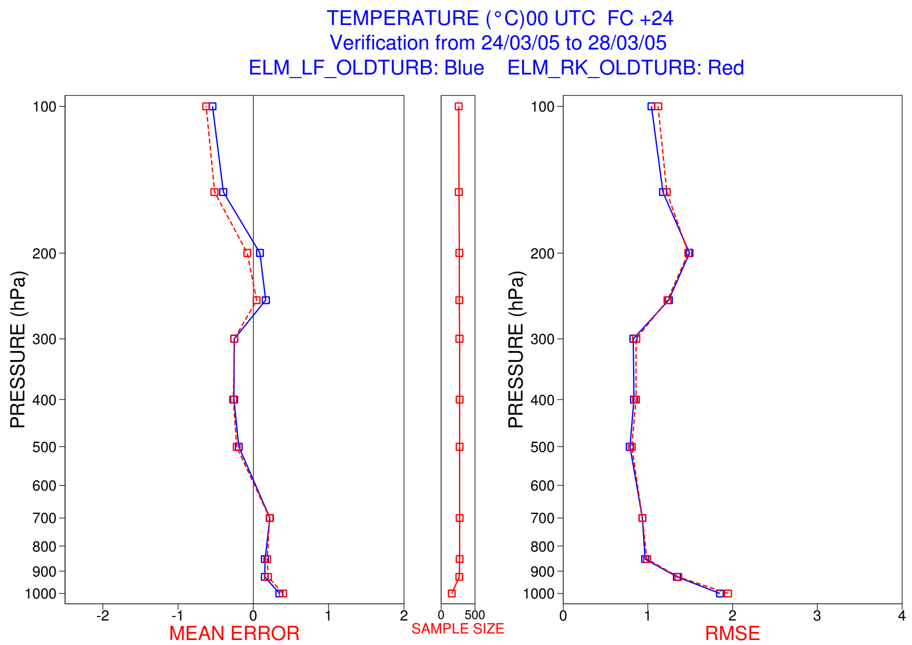


Figure 6: RK runs with eulerian and semi-lagrangian formulation of the moisture variables advection: mean error, root mean square error and standard deviation of mean sea level pressure; frequency bias index and threat score of 6h accumulated total precipitation larger than 2 mm.

3.5 Moisture variables transport scheme

All the RK experiments were performed using the eulerian formulation of the moisture variables advection scheme. In this experiment the semi-lagrangian formulation (SL) is compared to the eulerian one (EU). MSLP ME, RMSE and STDV as a function of the forecast step are shown in Fig. 6 for the SL (red lines) and the EU (blue lines) runs. 6TP_2 FBI and TS are also computed (Fig. 6). The semi-lagrangian moisture variables advection scheme does not show any significant difference in MSLP forecast compared to the eulerian version (slightly larger standard deviation after 18h forecast balanced by a slightly smaller bias), but SL forecasts seem to have a slightly better skill for 6h accumulated precipitation (slightly larger TS) and 2m dew point (not shown).



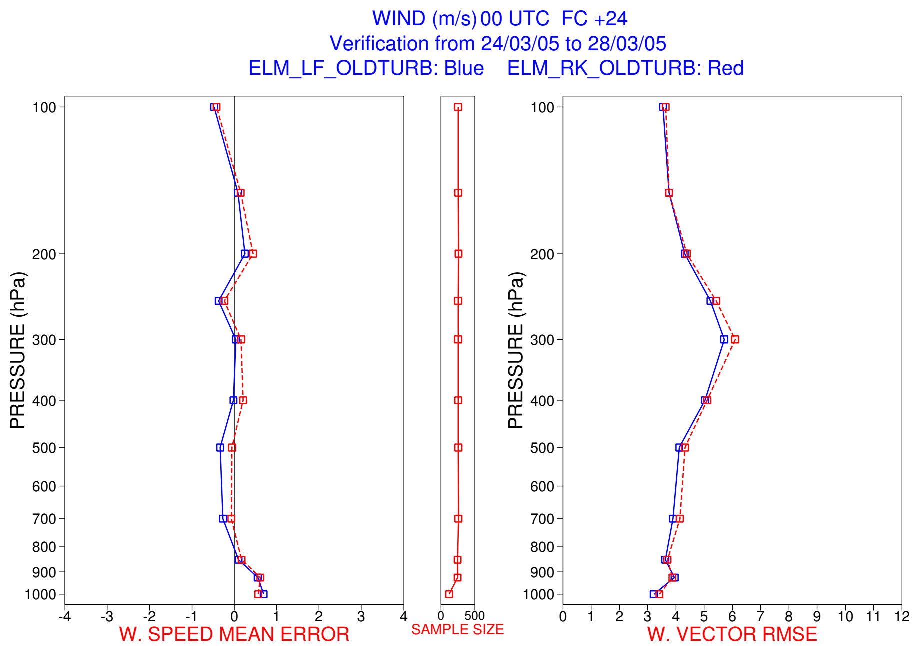


Figure 7: LF and RK runs with diagnostic TKE turbulence: temperature mean error and root mean square error vertical profiles of 24 h forecasts; wind speed mean error and wind vector root mean square error vertical profiles of 24 h forecast.

4 LF and RK with the diagnostic TKE turbulence scheme

Temperature and wind ME and RMSE vertical profiles for 24 h forecast are shown in Fig. 7 for the RK (red lines) and LF (blue lines) runs. The RK wind vector RMSE is larger than the LF one at almost all levels, while the RK and LF temperature RMSE have no significant differences. 2T, 2TD, 10U, MSLP bias and RMSE (also STDV for MSLP) as a function of the forecast time are represented in the Fig. 8 for RK (red lines) and LF (blue lines) runs. 6TP_2 FBI and TS are also computed (Fig. 8). A very slight improvement is found in RK forecast skill for 2T (18h and 24h forecast) and 2TD (6h and 24h forecast), while a slight worsening is found for 6TP_2 (18h and 24h forecast). A positive MSLP bias (except for 12h forecast) is found in RK forecasts, but the RK-LF MSLP bias difference found using the prognostic TKE turbulence scheme, even if shifted and slightly reduced, is still present. On the other hand, RK forecasts have a slightly smaller MSLP standard deviation than LF ones.

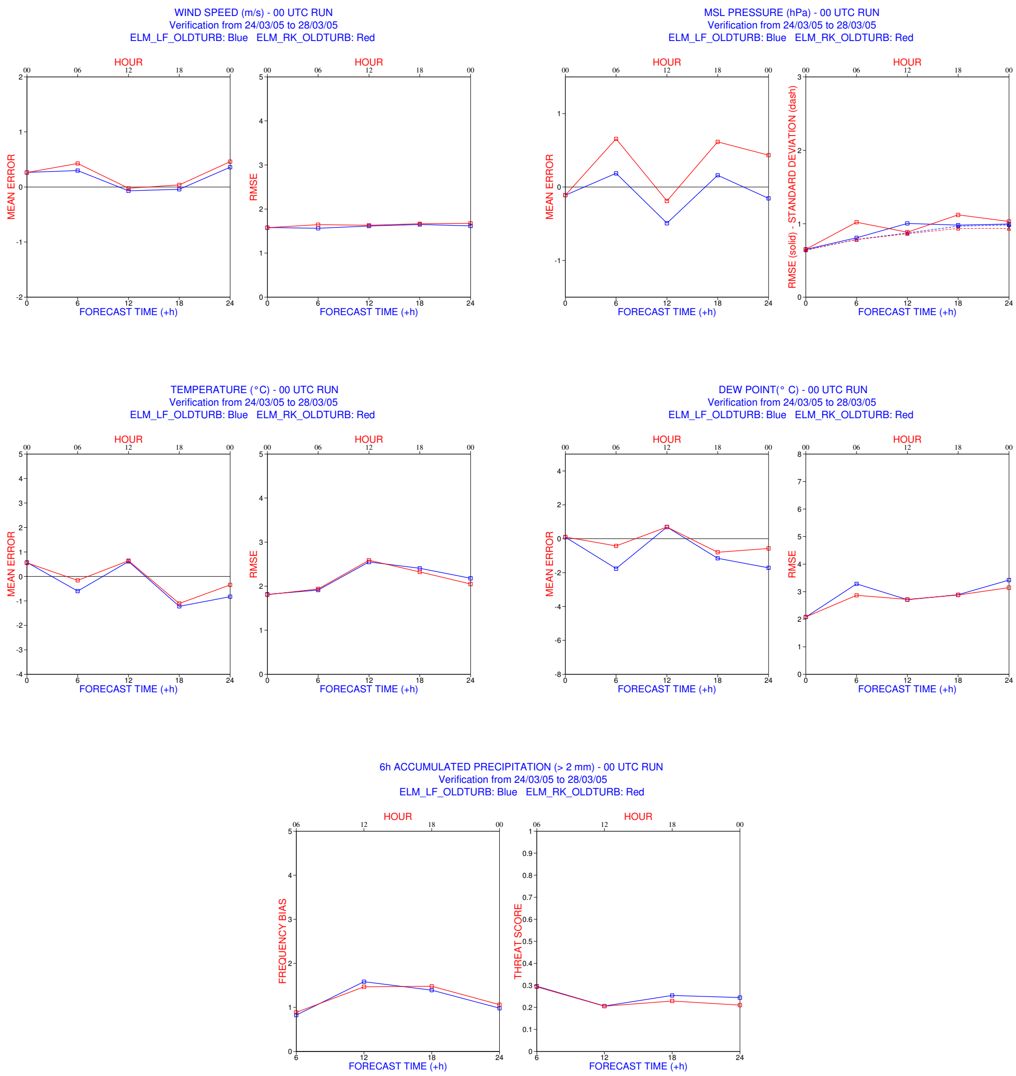


Figure 8: LF and RK runs with diagnostic TKE turbulence: mean error and root mean square error of two meter temperature, two meter dew point, ten meter wind speed, mean sea level pressure (also standard deviation for MSLP); frequency bias index and threat score of 6h accumulated total precipitation larger than 2 mm.

5 Summary and conclusions

The comparison of LF and RK dynamical cores was performed for a 5 days period using the EuroLM configuration. Statistical verification results showed that RK performance for surface variables was slightly better than LF one. A large positive MSLP bias was typical of the RK runs. Some sensitivity studies were performed on RK to determine the cause of the RK-LF differences. RK did not show any sensitive to the calls of the prognostic TKE turbulence and convection schemes. An improvement in the MSLP forecast skill was obtained using the diagnostic TKE turbulent scheme, but a positive bias (slightly reduced) was found again. The domain size sensitivity experiment showed similar results to those found in a previous work with LF core. The moisture variables transport experiment showed that semi-lagrangian version has a slightly better skill for precipitation than the eulerian one. Longer periods of investigations in different seasons are necessary to substantiate that the MSLP forecast deficiency found in RK runs is a real problem, but there are indications that more work is needed to tune the turbulence parameterization schemes and to improve the dynamics and physics coupling in the RK core.

References

Förstner, J. and G. Doms, 2004: Runge-Kutta time integration and high-order spatial discretization of advection: a new dynamical core for LM. *COSMO Newsletter*, No. 4, 168-176.

Torrisi, L., 2005: Impact of the domain size on LM forecast. *COSMO Newsletter*, No. 5, 145-148.

Application of the Z-Coordinate Version vs. the Terrain Following Version of LM Nonhydrostatic Model over Greece

E. Avgoustoglou, T. Tzeferi, V. Tirli and I. Papageorgiou

Hellenic National Meteorological Service, El. Venizelou 14, Hellinikon, Athens 16777, Greece

1 Introduction

Local numerical weather prediction models using terrain following coordinates are sensitive to numerical errors induced by the singular behavior of the coordinate transformation Jacobian around steep orographic slopes (Sundqvist, 1976; Gallus and Klemp, 2000). This misfunction may produce artificial circulations that destroy clouds in the vicinity of the mountains leading to significant limitations regarding local weather forecasting.

In order for this pathology to be properly treated, a new version of the LM (subsequently denoted as "LM_Z") has been developed that explicitly uses the Z-coordinate representation in the model numerical scheme (Steppeler et.al., 2002; Bitzer and Steppeler, 2002; and others). The results for tests, mainly involving idealized cases, have been successful, leaving space towards its operational validation.

2 Test Case Justification and Results

The geographical domain of Greece may be considered an excellent candidate towards the relative evaluation of LM_Z against its terrain following coordinates operational version (LM_TF); since the area is characterized by the equipartitioned land-sea interchange combined with a complex orography as well as a large number of mountainous islands. In this particular test case, we investigated the weather development during the three day period of the 6^{th}, 7^{th}, and 8^{th} of March 2005. As it can be seen from the satellite pictures as well as the synoptic analysis (Fig. 1), on March 6, a deep low pressure system over East Balkans associated with frontal activity extended to East Aegean was moving East/Northeast. This activity was followed on March 7 and 8 by a moderate frontal development over South Italy moving East combined with a strong Southwestern wind field in the middle troposphere. The North to South orographic structure of mainland Greece was expected to effect cloud formation and precipitation in a way that might demonstrate differences between LM_TF and LM_Z. In Figs. 2, 3, 4, we show the relative forecasted low, medium and total cloud cover for LM_TF and LM_Z respectively. We used boundary conditions from the Global Model of the German Meteorological Service (DWD) with analysis of 00 UTC for every date under consideration. The cloud cover forecasted by LM_Z conforms more with the satellite pictures of Fig. 1. This looks consistent with Fig. 5 where the 12-hour forecasted accumulated precipitation in LM_Z is overall downgraded and less dispersed in reference to LM_TF, particularly over the sea surface. Regarding observation, the measured values of the 12-hour accumulated precipitation over the local meteorological stations were compared to the forecasted values of the nearest grid point. By summing these values, it was found that the total forecasted precipitation for LM_Z was closer to the total precipitation measured (Table 1). In Fig. 6, we depict with "R" the positions of the meteorological stations where the observed value for the precipitation was closer to LM_TF and with "Z" when this value was closer to LM_Z. The bullet sign corresponds to stations where precipitation was neither observed nor predicted by any version of LM. Within this context, it may be seen again that the forecasted values from LM_Z are relatively closer to observation.

		March 06	March 07	March 08
Observed:	Total	10.52	152.00	78.01
	Average	0.18	2.82	1.37
LM_TF:	Total	145.99	304.56	266.06
	Average	2.52	5.64	4.67
LM_Z:	Total	36.87	167.82	155.57
	Average	0.64	3.11	2.73

Table 1: Total and average observed and forecasted precipitation height (mm)

Conclusions

For the test case under consideration, LM_Z forecast shows relative preponderance over LM both for cloud coverage and 12-hour accumulated precipitation. However, more systematic investigation is necessary in the direction of further validating LM_Z for real weather situations in connection of further development of the code both in the direction of numerics as well as that of physics.

References

H. Sundqvist, 1976: On vertical interpolation and truncation in connection with use of sigma system models.*Atmosphere*, 14, 37-52.

W. Gallus and J. Klemp, 2000: Behavior of flow over steep orography. *Mon. Wea. Rev.*, 128, 1153-1164.

J. Steppeler, H.-W. Bitzer, M. Minotte and L. Bonaventura, 2002: Nonhydrostatic Atmospheric Modeling using a z-Coordinate Representation. *Mon. Wea. Rev.*, 130, 2143-2149.

H.-W. Bitzer and J. Steppeler, 2002: A description of the Z-Coordinate Dynamical Core of Lm. *COSMO Technical Report*, No. 6, Deutscher Wetterdienst (DWD), Offenbach.

J. Steppeler and H.-W. Bitzer , 2002: The Z-Coordinate Version of the LM. *COSMO Newsletter*, No. 2, 111-112.

J. Steppeler and H.-W. Bitzer and U. Schättler, 2003: New Developments Concerning the Z-Coordinate Version of the LM. *COSMO Newsletter*, No. 3, 177-178.

J. Steppeler, S. Janjic, H.-W. Bitzer, P. Prohl and U. Schättler, 2004: The Z-Coordinate LM. *COSMO Newsletter*, No. 4, 152-154.

J. Steppeler, H.-W. Bitzer, S. Janjic, U. Schättler, P. Prohl, J. Parfiniewicz, U. Damrathh and E. Avgoustoglou 2005: A Z-Coordinate Version of the Nonhydrostatic Model LM. *COSMO Newsletter*, No. 5, 149-150.

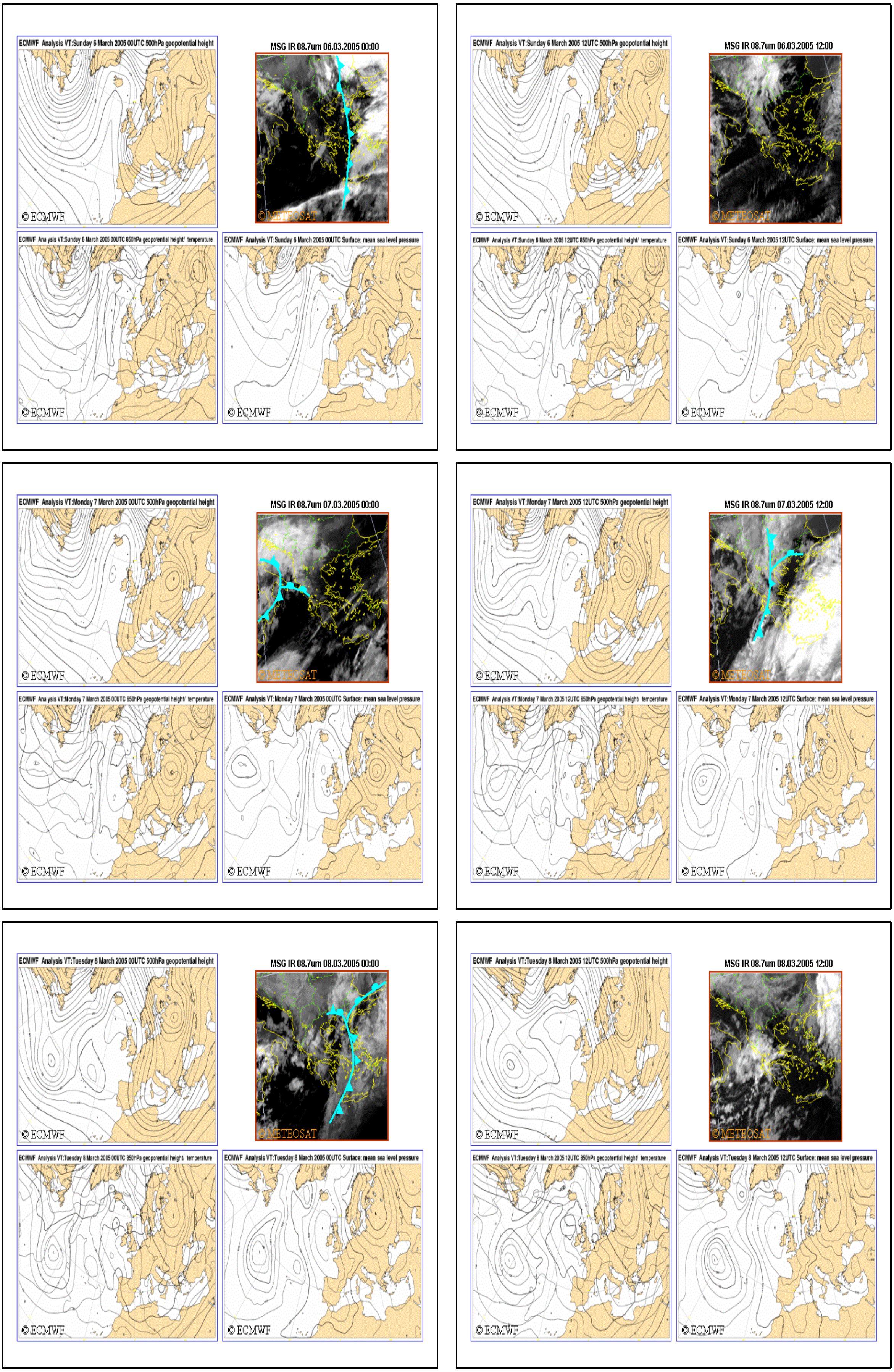


Figure 1: Satelite pictures and analysis charts for 6, 7 and 8 of March 2005.

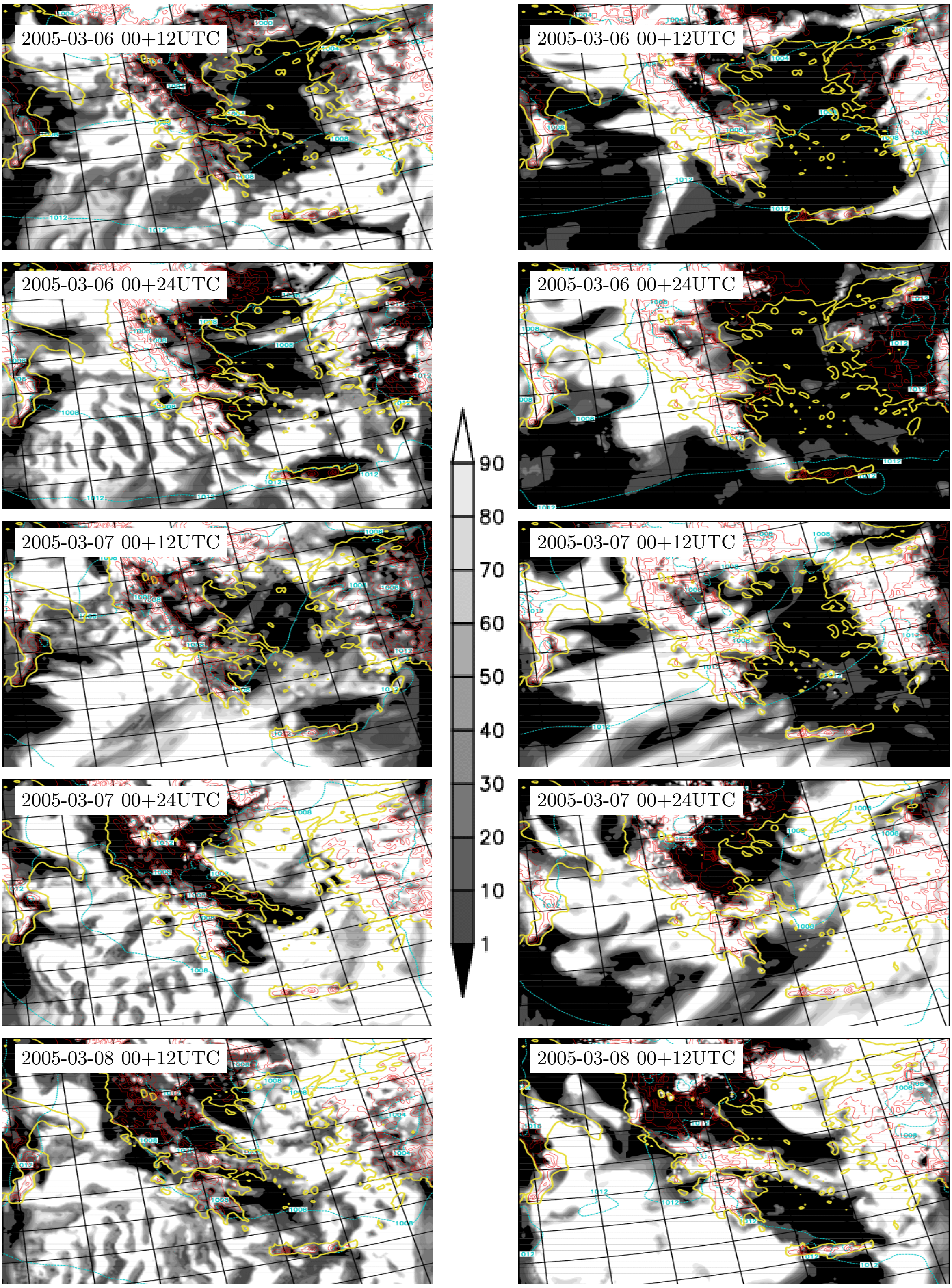


Figure 2: Low cloud cover forecast (%) and PMSL (HPa) from LM_TF (left column) and LM_Z (right column).

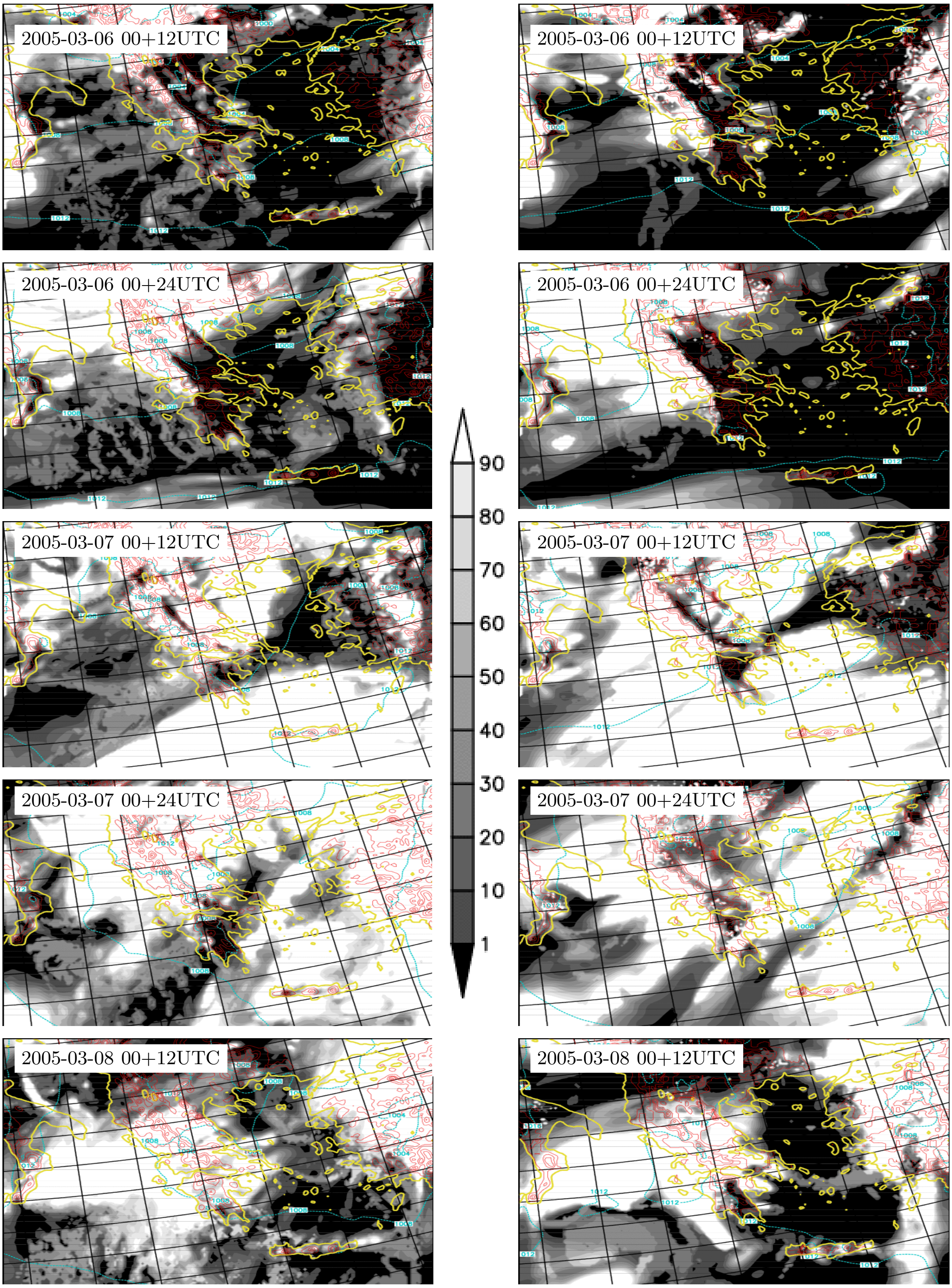


Figure 3: Medium cloud cover forecast (%) and PMSL (HPa) from LM_TF (left column) and LM_Z (right column).

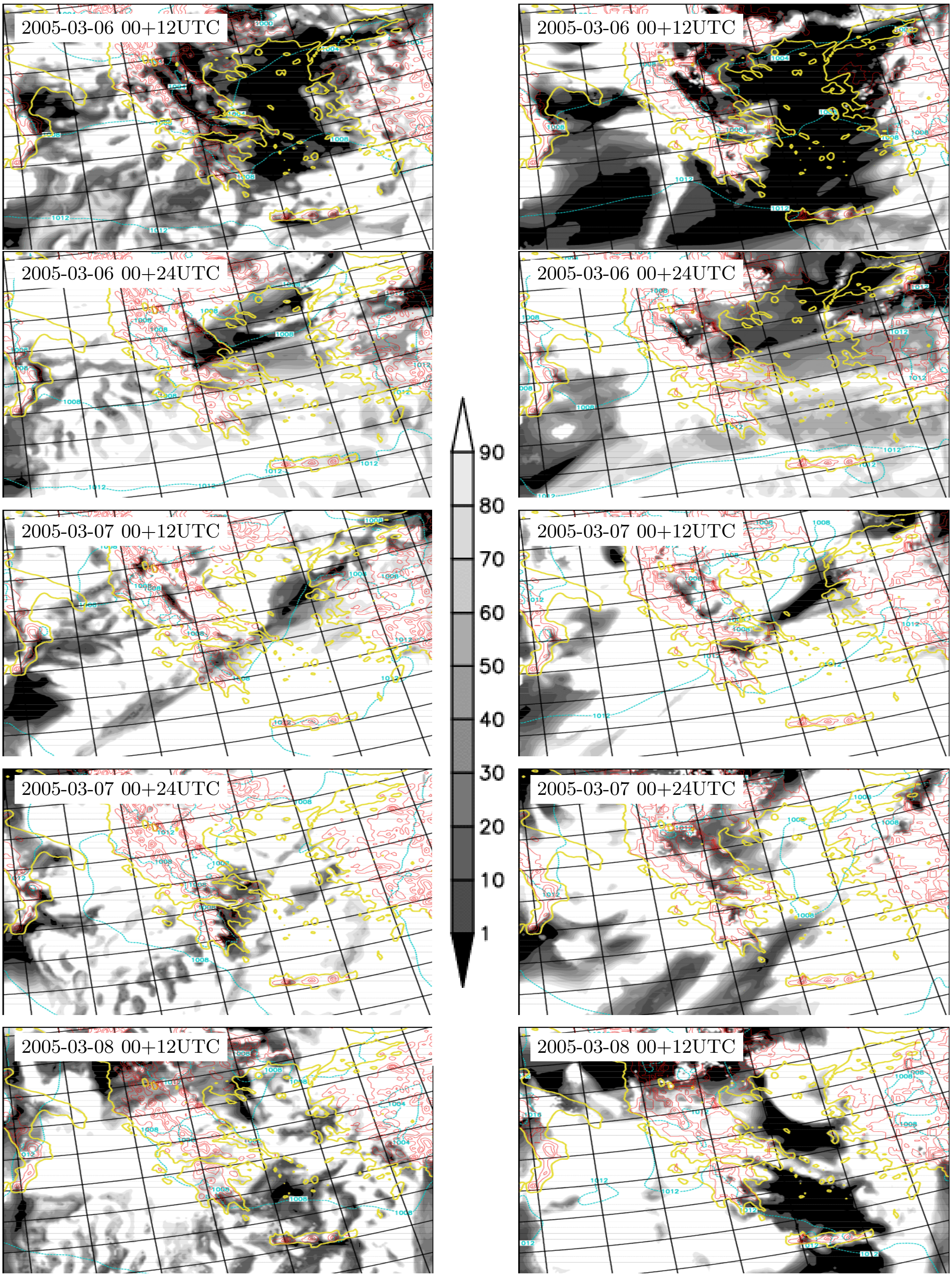


Figure 4: Total cloud cover forecast (%) and PMSL (HPa) from LM_TF (left column) and LM_Z (right column).

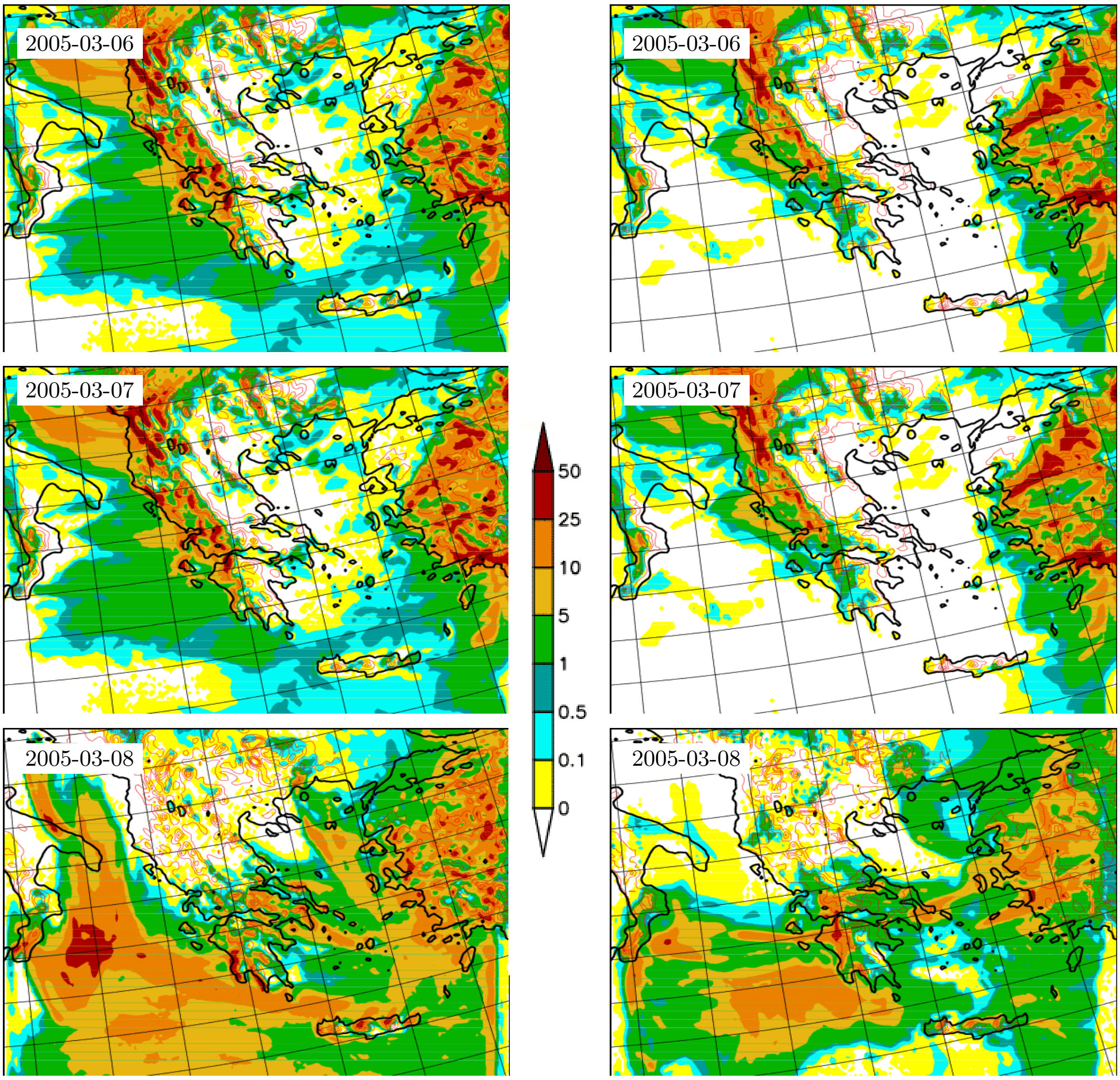


Figure 5: 12-hour accumulated precipitation (mm) from LM_TF (left column) and LM_Z (right column).

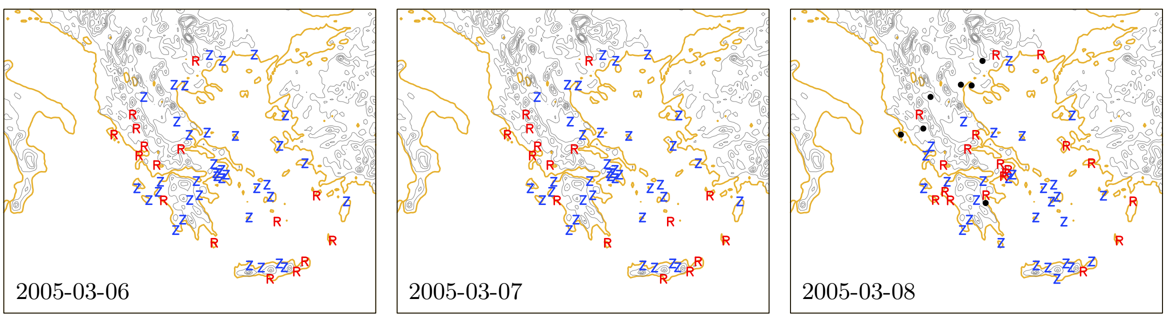


Figure 6: Represantation of the positions of the meteorological stations in reference to the relation of measured against the 12-hour accumulated precipitation; "R" and "Z" stand for stations where LM_TF and LM_Z forecast was closer to observation respectively. The "bullet" sign stands for stations where no precipitation was observed or forecasted by any version of LM.

6.3 Working Group 3: Physical Aspects

The main responsibility of this working group is to develop new physics packages for future operational applications and to improve existing parameterisation schemes. The WG is coordinated by Marco Arpagaus (MeteoSwiss). The work packages of this group are splitted into various sub-themes, such as planetary boundary layer, microphysics, clouds, convection, radiation, and soil processes. The detailed annual status reports for each work package can be obtained from within the member area of the COSMO web-site. Short summaries of some selected topics of the last COSMO period are given below:

Planetary boundary layer: Work continued on the new turbulence scheme based on a prognostic treatment of turbulent kinetic energy (TKE) as well as on the new surface transfer scheme. The main effort went into setting up a 1-d version of the model to be able to make thorough validation studies. Writing an extended documentation of the new scheme was another priority of last years work. — A technical report on parts of this work package ("Evaluation of Empirical Parameters of the New LM Surface-Layer Parameterisation Scheme") can be obtained on the COSMO web-site at `http://www.cosmo-model.org/cosmoPublic/technicalReports.htm`.

Microphysics: A three-category ice scheme has been developed and implemented into the LM. After extensive testing, the scheme is now the default option of the test-suites for the 2.8 km LM version 'LMK' running at DWD.

Clouds: First tests with the statistical cloud scheme (which is part of the new turbulence scheme) have been done and showed encouraging results. More tests and systematic studies are certainly needed, and the work will be continued within the framework of the new priority project *Towards a Unified Turbulence Shallow Convection Scheme (UTCS)* (see below).

Convection: The high resolution versions of LM (e.g. LMK, aLMo2) treat deep convection explicitly. However, to prevent the boundary layer from becoming too wet (the transport across the boundary layer top due to vertical mixing alone seems to be insufficient), a shallow convection scheme is needed also for grid-spacings of 2.8 km or 2.2 km. Tests with a stripped-down version of the Tiedtke cumulus parametrisation scheme showed satisfactory results, and is being used for the test-suites of LMK and aLMo2. However, a physically more appealing description of the moisture transport in and across the boundary layer top is part of the new priority project *Towards a Unified Turbulence Shallow Convection Scheme (UTCS)* (see below).

Radiation: First attempts towards a (poor-man's) three-dimensional radiation scheme have been made by applying correction factors to the solar and thermal radiation budgets at the surface due to shadowing, terrain inclination and orientation, reduced sky-view, etc.

Soil processes: The new multi-layer version of the soil model TERRA, which includes freezing and melting of soil layers and a revised formulation of the snow model, has finally been put into operational service at DWD in autumn 2005 (LME). A technical report describing the changes to TERRA is available on the COSMO website at `http://www.cosmo-model.org/cosmoPublic/technicalReports.htm` ("The Multi-Layer Version of the DWD Soil Model TERRA_LM").

According to the new organisational structure of COSMO, introduced at the last COSMO meeting in Zurich (September 2005), there are now priority projects in addition to the work packages already known from earlier years. The two priority projects associated to WG3 (but explicitly not restricting their focus to physical aspects only) and starting as of March 2006 and December 2005, respectively, are *Towards a Unified Turbulence Shallow Convection Scheme (UTCS)* and *Tackle deficiencies in quantitative precipitation forecasts*.

Towards a Unified Turbulence Shallow Convection Scheme (UTCS):

Representation of shallow convection and boundary-layer turbulence in numerical models of atmospheric circulation is one of the key unresolved issues that slows down progress in numerical weather prediction. The goal of this project is to make a step forward in this area. The project is aimed at (i) parameterising boundary-layer turbulence and shallow non-precipitating convection in a unified framework, and (ii) achieving a better coupling between turbulence, convection and radiation. Boundary-layer turbulence and shallow convection will be treated in a unified second-order closure framework. Apart from the transport equation for the sub-grid scale turbulence kinetic energy (TKE), the new scheme will carry at least one transport equation for the sub-grid scale variance of scalar quantities (potential temperature, total water). The second-order equations will be closed through the use of a number of advanced formulations, where the key point is the non-local parameterisation of the third-order turbulence moments.

Tackle deficiencies in quantitative precipitation forecasts:

This project aims at looking into the LM deficiencies concerning precipitation by running sensitivity experiments on a series of well chosen cases which have verified very poorly. If successful, the outcome of these sensitivity experiments will be a more effective set of LM namelist or model parameters for quantitative precipitation forecasting, or a clear idea of what parts of the model need to be reformulated and improved most urgently to obtain better quantitative precipitation forecasts.

The working plan for the next COSMO period includes — additionally to the priority projects just described — further ongoing or new work related to the parameterisation schemes dealing with the planetary boundary layer, microphysics, clouds, convection, radiation, and the soil processes. Examples for such work packages are the revision of the surface transfer scheme to improve the daily cycle of 2m temperature, a detailed comparison of the (fast) radiation scheme with an elaborate line-by-line code, and the implementation of the new lake model FLake.

Preliminary Results on Comparison of LM Radiation Code to the LbL Model RTX

A. Bozzo

University of Bologna, ADGB, Bologna, Italy

1 Introduction and Model Description

The LM fast radiation scheme (Ritter and Geleyn, 1992) has been compared with the LbL-multiple scattering model RTX, that has been improved and extended during the last 5 years at ADGB (Atmospheric Dynamic Group – Dep. of Physics, University of Bologna). Motivation of the comparison is testing the accuracy of LM radiation code using RTX computations as benchmark. Main interest is checking the optical properties of gases and clouds, as modeled in LM.

RTX spectroscopic database is based on the modern HITRAN2000 and extends from far infrared (FIR) to 0.4 μm. We are working in order to extend the gaseous optical database up to the UV bands. Single scattering properties of water clouds are computed with exact Mie theory calculations. Optical properties of ice clouds are based on the most advanced parameterizations for aggregate and columns, made available to ADGB in a long standing cooperation with the Met Office. Adding and doubling method provide exact radiative transfer calculations in multiple scattering environments.

The LM radiative transfer code (named GRAALS, General Radiative Algorithm Adapted to Linear-type Solutions) is a fast delta-two stream radiation code with treatment of partial cloudiness via maximum-random overlap method. The optical properties of clouds and gases are parameterized over 8 wide spectral intervals from visible (VIS, 0.2 μm) to far infrared (FIR, 104.5 μm). Parameterizations by Slingo-Schrecker (1982) and Rockel (1991) is adopted for the computation of water and ice clouds single scattering properties respectively. The spectroscopic gaseous properties are based on the old AFGL '84 database.

In this work we will present the very first comparisons between the two models in some simple atmospheric situations. The basic settings of both the models concern a Lambertian surface and the gaseous absorption of the major atmospheric gases as CO_2, H_2O, O_3, N_2O, O_2, CO, CH_4 for both models. RTX consider also trace gases as CFCs (F11,F12,CCl4), NO, NO_2, SO_2, N_2. Monocromatic RTX fluxes are grouped in 7 of the 8 LM-GRAALS bands (the LM VIS range 0.2 μm - 0.7 μm is not taken in account due to the (for now) limited RTX solar spectroscopic database).

2 Comparisons

The first comparison scenario is a clear Standard tropical atmosphere (McClatchey et al, 1971). In Fig. 1 the thermal IR net fluxes show good agreement in terms of the shapes of the curves, but a bias affect the absolute values. In Fig. 2 the IR net flux is splitted in the contributions of the 5 thermal IR LM-GRAALS bands (i.e. contributions of the major gases). All curves show discrepancies between RTX and LM-GRAALS. In particular, the CO_2 bands show considerable higher LM-GRAALS net fluxes throughout the whole profile.

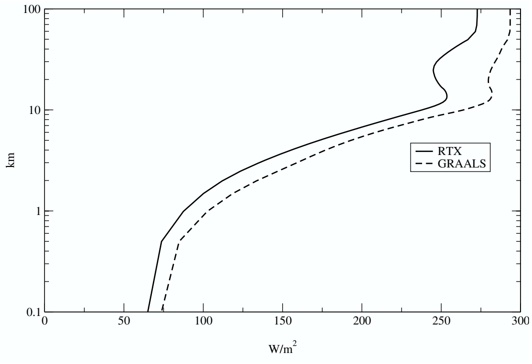


Figure 1: Thermal-IR (104-4.6 μm) Net Flux; (STD Trop. Atm. Profile – Sun zenith angle = 0°)

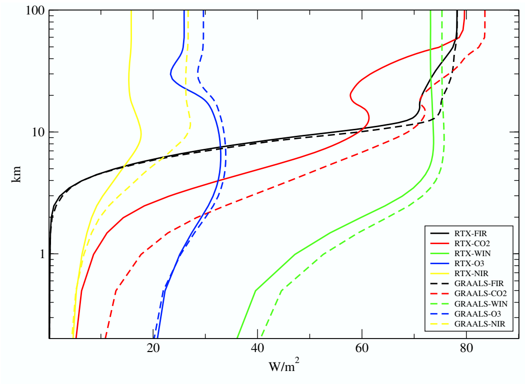


Figure 2: Thermal-IR (104-4.6 μm) Spectral Net Fluxes; (STD Trop. Atm. Profile – Sun zenith angle = 0°)

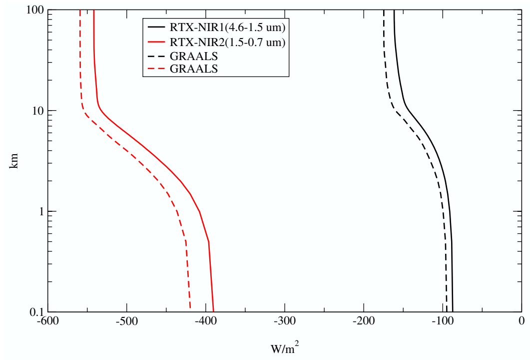


Figure 3: Solar-IR Spectral Net Fluxes; (STD Trop. Atm. Profile – Sun zenith angle = 0°)

Moreover the net fluxes diverge in the O_3 band in the layer where the maximum absorption is located, i.e. in the first stratospheric layers. The large discrepancy in the near infrared band (NIR) is due to the fact that the sun irradiance is distributed in LM_GRAALS only in the 3 solar bands (i.e. from 0.2 μm to 4.5 μm) instead to be treated as a continue spectral distribution as RTX do: the integral of the down-welling solar irradiance in the spectral range covered by the NIR LM-GRAALS band (4.5-8 μm) gives approximately $8W/m^2$, explaining the flux difference of the two models at TOA in this band. The bad agreement of the net atmospheric fluxes in the other bands are most probably due to the approximations used by LM-GRAALS in the computations of gaseous optical properties, although the different spectroscopic database used by the models can be an important other source of discrepancy.

In the 2 analysed solar bands (Fig. 3), we can again see a good agreement in the net fluxes profile but a bias in the absolute values. Table 1 shows the percentage difference at surface and at top of atmosphere between the two model's up- and down-welling fluxes. Greatest values can be found in the solar bands and in thermal IR TOA outgoing fluxes.

	TOA Fluxes		Surface Fluxes	
Up Flux	-12.3	-4.6	12.9	0.8
Down Flux	-4.4		-8.8	0.7

Table 1: TOA and Surface Net Fluxes RTX-GRAALS (%)

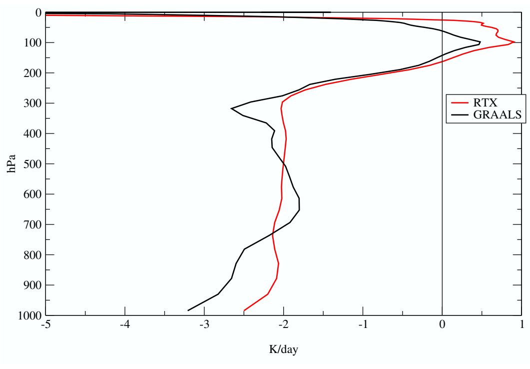


Figure 4: Thermal-IR (104-4.6 μm) Heating Rate; (STD Trop. Atm. Profile)

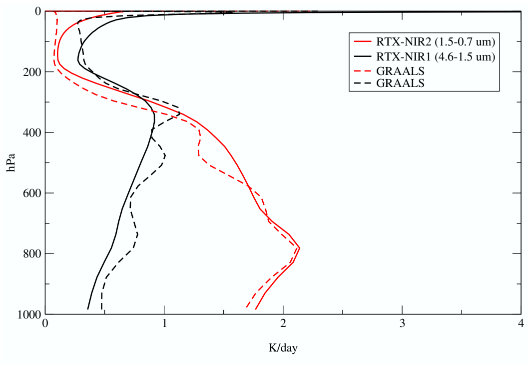


Figure 5: Solar-IR (4.5-0.7 μm) Heating Rate; (STD Trop. Atm. Profile)

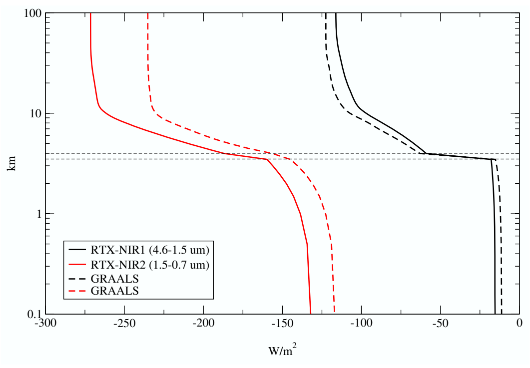


Figure 6: Solar-IR Spectral Net Flux; (Overcast medium-level water cloud: 500m thick; LWC=0.129 g/m^3, re = 5.2 μm; Sun zenith angle = 0°)

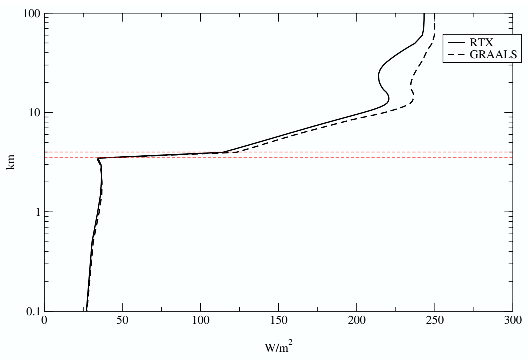


Figure 7: Thermal-IR (104-4.6 μm) Net Flux; (Overcast medium-level water cloud: 500m thick; LWC=0.129 g/m^3, re = 5.2 μm)

The thermal IR heating rate profiles (Fig. 4) show some differences, again due probably to the approximations adopted in LM-GRAALS and to the different spectroscopic database of the two models. We think that upgrading the LM gaseous optical properties can help to improve the computations. In the solar bands, the heating rates (Fig. 5) show good agreement between the models. Nevertheless, work is still in progress in order to extend the comparison at the whole solar spectral range.

In the second scenario, various cloud types have been introduced in the standard tropical profile. Three overcast, homogeneus cloud layers have been generated:

- low-level thick water cloud: 850-715 hPa (1500m thick); LWC=0.5 g/m^3, re = 9.3 μm
- medium-level water cloud: 672-633 hPa (500m thick); LWC=0.129 g/m^3, re = 5.2 μm
- high-level cloud: 266-247 hPa (500m thick); LWC=0.013 g/m^3, re = 4.1 μm

For the medium level clouds, solar net fluxes (Fig. 6, for the low cloud and for different zenith angles a similar situation holds) show the same features emphasized in the clear-sky case; in the thermal IR spectrum (Fig. 7), we can see a fairly good agreement for the net fluxes below the cloud layer, but still disagreement above, as noted in the clear sky case.

Hence it appears that the thermal IR heating rates are in better agreement in cloudy profile than in the clear one (Figs. 8, 9), the differences in the higher layers being unchanged with

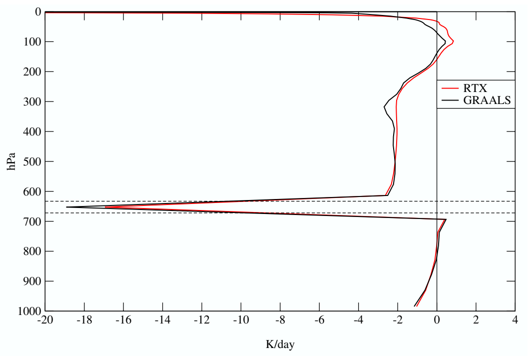


Figure 8: Thermal-IR (104-4.6 μm) Heating Rate; (Overcast medium-level water cloud: 500m thick; LWC=0.129 g/m^3, re = 5.2 μm)

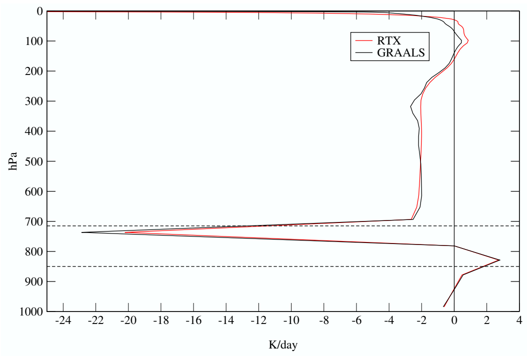


Figure 9: Thermal-IR (104-4.6 μm) Heating Rate; (Overcast low-level water cloud: 1500m thick; LWC=0.5 g/m^3, re = 9.33 μm)

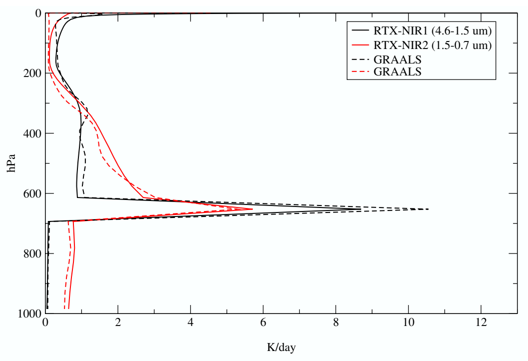


Figure 10: Solar-IR Spectral Heating Rate; (Overcast medium-level water cloud: 500m thick; LWC=0.129 g/m^3, re = 5.2 μm)

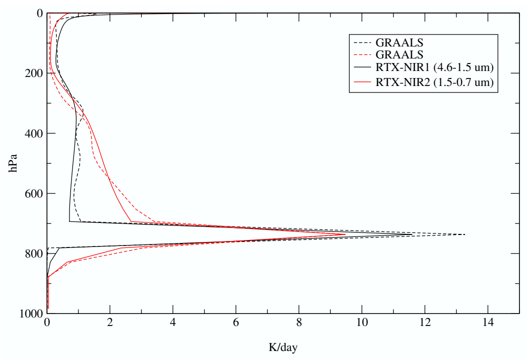


Figure 11: Solar-IR Spectral Heating Rate; (Overcast low-level water cloud: 1500m thick; LWC=0.5 g/m^3, re = 9.33 μm)

respect to the clear sky case. Agreement between the two models can be seen also in the LM solar bands (Figs. 10, 11), the maximum difference between models is noted near the top of the cloud layers, where the maximum heating rate is located. The difference appear to be less evident at shorter wavelengths probably linked to slightly different layer absorption properties, computed approximately by LM_GRAALS and following the exact solution of Mie theory in multiple scattering layers by RTX.

The high, shallow cloud layer has been treated, in first approximation, as water cloud, in spite of the high altitude. It can be seen (Fig. 12) that the thermal IR heating rates show a mean difference of 0.2 K/day to 0.5 K/day below the cloud layer (as in the clear sky case) and a maximum of 1 K/day at the bottom of the cloud layer. The comparison in the solar bands (Fig. 13) is worse as RTX shows strongest heating rate inside the cloud layer. It seems that differences between the models are enhanced by small IWP values.

3 Conclusions and further work

These very first comparisons are showing a reasonably good agreement of the two models in some simple clear and cloudy atmospheric profiles. Some discrepancies in the clear sky profile are linked with the parameterizations adopted by LM-GRAALS and probably also with the old spectroscopic database adopted by the fast radiation code. Actually it appears that the water cloud model can gain fairly good results; nevertheless deeper investigations

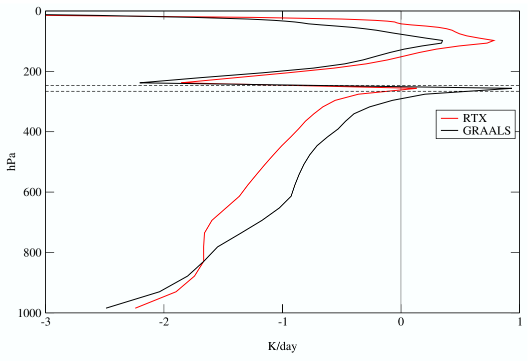


Figure 12: Thermal-IR (104-4.6 μm) Heating Rate; (Overcast high-level water cloud: 500m thick; LWC=0.013 g/m^3, re = 4.1 μm)

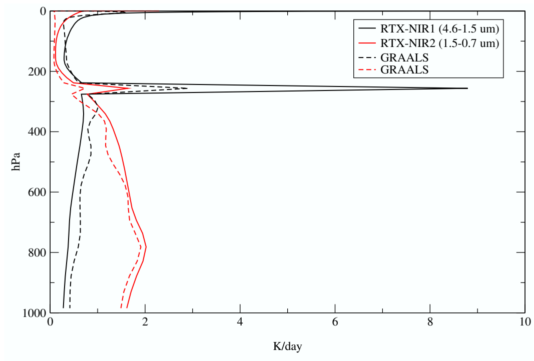


Figure 13: Solar-IR Spectral Heating Rate; (Overcast high-level water cloud: 500m thick; LWC=0.013 g/m^3, re = 4.1 μm)

will be necessary, mostly to clarify the shallow cloud layers optical properties.

This preliminary work is the starting point for further, systematic analysis which will concern the following points:

1. Compute the spectral coefficients for gaseous absorption using the latest release of HITRAN. We are aware of work being done by F. Geleyn on this issue and do not intend to duplicate his effort;

2. Test and eventually improve the parameterization of optical properties of ice clouds, using the most advanced parameterisations for aggregate, columns and possibly hexagonal plates made available to us by the Met Office;

3. Compare the treatment of aerosols in GRAALS and RTX using possibly the most recent version of OPAC (Optical Properties of Aerosols and Clouds) (Hess, Koepke, and Schult, 1998).

References

Slingo, J.M., 1989: A GCM parameterization for the shortwave radiative properties of water clouds. *J. Atmos. Sci.*, 46, pp. 1419-1427.

Slingo, J.M., and Schrecker H. M., 1982: On the shortwave radiative properties of stratiform water clouds. *Quart. J. Roy. Meteor. Soc.*, 108, pp. 407-426.

Stephens, G.L. 1979: Optical properties of eight water cloud types. *CSIRO, Division of Atmospheric Physics*, Technical Paper No. 36, 35 pp.

Ritter B. and Geleyn J.-F. (1992), A comprehensive Radiation Scheme for numerical weather pretiction models with potential applications in climate simulations, Monthly weather review, 120, 303-325

Zdunkowski, W.G., and Korb G.J. (1985), Numerische Methoden zur Loesung der Strahlungsuebertragungsgleichung, PROMET 2/3, 1985, 26-39

Implementation of the Statistical Cloud Scheme Option: Preliminary Tests

E. Avgoustoglou, T. Tzeferi and I. Papageorgiou

Hellenic National Meteorological Service, Athens, Greece

1 Introduction

In general, cloud models are implemented in numerical weather prediction with the main assumption that the computational grid volume is either entirely saturated or entirely unsaturated. From a closer physics standpoint, this conjecture might be insufficient since substantial portions of grid volumes contain saturated air near boundaries. Additionally, cumulus interiors may contain unsaturated air due to lateral merging of adjacent clouds and entrainment while the initial stage of cloud growth might be treated incorrectly because no latent heat is released until an entire grid volume is saturated. The complexity of these processes leads to the investigation for possible improvement to the cloud model of LM by deriving dependencies of mean cloud fraction upon humidity statistics by assuming Gaussian quasi-conservative properties (Refs. Betts (1973), Sommeria and Deardorff (1977), Mellor (1977), Mellor and Yamada (1982), Raschendorfer (2005)).

Currently, there are two different schemes for sub-grid scale cloudiness in the LM. A simple one based on relative humidity, and another one, called a statistical cloud scheme, which depends on the statistical properties of the saturation deficit within the turbulence scheme. The first scheme is currently used to feed the radiation scheme with cloud information, and has hence been tuned for this purpose. In addition, its results are used as model output as well. The statistical scheme on the other hand is currently used only within the moist turbulence scheme. As the statistical scheme is more sophisticated, it would be of considerable importance to use only this scheme, both for turbulence and radiation as well as for model output.

2 Analysis

For the implementation of the statistical cloud scheme we follow the works of Sommeria and Deardorf (1977) as well as Mellor (1977) where the subgrid low cloud fraction R and mean liquid water content $\overline{q}_l$ are estimated as

$$R = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(q_w - q_s) G dq_w d\theta_l \tag{1}$$

and

$$\overline{q}_l = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (q_w - q_s) H(q_w - q_s) G dq_w d\theta_l \tag{2}$$

where q_w and q_s correspond to the total-water and saturation specific humidities respectively, θ_l is the liquid water potential temperature, H stands for the Heaviside function

$$H = \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases} \tag{3}$$

and G is a bivariate normal function

$$G = \frac{1}{2\pi\sigma_{\theta_l}\sigma_{q_w}(1-r^2)^{\frac{1}{2}}} exp\left[\frac{-1}{1-r^2}\left(\frac{\theta_l'^2}{2\sigma_{\theta_l}^2} - r\frac{\theta_l' q_w'}{\sigma_{\theta_l}\sigma_{q_w}} + \frac{q_w'^2}{2\sigma_{q_w}^2}\right)\right] \tag{4}$$

with primed quantities defined as $x' \equiv x - \overline{x}$ and the correlation factor $r = \overline{\theta_l' q_w'}/(\sigma_{\theta_l}\sigma_{q_w})$.

By assuming a linear approximation for q_s around the value $\overline{q}_{sl} = q_s(\overline{\theta}_l, \overline{p})$ and with the help of Clasius-Clapeyron equation the expressions for R and $\overline{q}_l$ become

$$R \approx \frac{1}{2}\left[1 + erf\left(\frac{Q}{\sqrt{2}}\right)\right] \tag{5}$$

$$\overline{q}_l \approx \frac{1}{1+\beta\overline{q}_{sl}}\left[RQ + \frac{exp\left(\frac{-Q^2}{2}\right)}{\sqrt{2\pi}}\right] \tag{6}$$

where

$$Q = \frac{\overline{q}_w - \overline{q}_{sl}}{\sigma}, \qquad \sigma = (\overline{q_w'^2} + \overline{q_{sl}'^2} - 2\overline{q_w' q_{sl}'})^{\frac{1}{2}}, \qquad \beta = 0.622\frac{L^2}{R_d c_p T_l^2} \tag{7}$$

with T_l standing for the liquid water temperature, L is the latent heat for vaporization, R_d is the gas constant for dry air and c_p is the specific heat at constant pressure.

Sommeria and Deardorff (1977), further approximated R through the linear part of an empirical curve that they drew for R by using an ensemble of 400 bivariate normal distributions

$$R \approx 0.5(1 + \frac{Q}{1.6}), \qquad 0 \le R \le 1 \tag{8}$$

In the statistical cloud scheme implemented in the LM, the low cloud cover is parameterized through a similar relation

$$R \approx A(1 + \frac{Q}{B}), \qquad 0 \le R \le 1 \tag{9}$$

The parameters A (cloud cover at saturation) and B (critical value of the saturation deficit) are denoted as $zlc0$ and zq_crit and are tunable in the physical parameterization of LM code.

3 Preliminary Tests

Our investigation regarding the implementation of the statistical cloud scheme for low clouds is focused on the weather situation over Greece on January 2 2005. The analysis charts at 500 and 850 Hpa at 06 UTC (right upper and middle graphs of Fig. 1) show an anticyclonic circulation over the western parts of Greece leading to westerly moving warmer air masses streaming over existing colder ones. In the eastern part of the country, a fainting dynamic activity in reference to a passing “trough” retains a weak northern wind current. The vertical correspondence of the anticyclonic circulation over the western parts of Greece is concluded from the surface analysis chart (lower right graph of Fig. 1). The resulting weak surface winds favor the low cloud development as it is shown on the corresponding infrared satellite picture (Fig. 1).

We present the low and total cloud coverage forecasts at 06 UTC of January 2 2005 from LM Version 3.15 with a 7 kilometer horizontal grid of 35 vertical levels, integration time step of 30 seconds and initial conditions from the DWD global model based on 12 UTC analysis of January 1 2005 (Figs. 2, 3). The rest of the configuration of the LM is described in previous COSMO Newsletters (Avgoustoglou; 2003). The test-runs were performed at the

Test Number	Cloud-Ice	*icldm_rad*	*zq_crit*	*zlc0*	*ztkhmin*	*ztkmmin*
1	T	4	4.0	0.5	0	0
2	T	4	4.0	0.5	1	1
3	F	4	4.0	0.5	1	1
4	F	2	4.0	0.5	1	1
5	F	2	6.0	0.5	1	1
6	F	2	2.0	0.5	1	1
7	F	2	6.0	0.8	1	1
8	F	2	2.0	0.8	1	1
9	F	2	6.0	0.2	1	1
10	F	2	2.0	0.2	1	1
11	F	2	4.0	0.8	1	1
12	F	2	4.0	0.2	1	1

Table 1: Combinations of the parameters modified for the considered test runs.

IBM Supercomputing System of the European Center of Medium-Range Weather Forecasts (ECMWF). In addition to the physical parameters modified in order to show the sensitivity of the statistical cloud scheme (i.e *zlc0* and *zq_crit*), we also performed tests on the cloud-ice scheme impact (Doms, 2002; Doms, 2004) as well as the absolute minimum value diffusion coefficients for heat and momentum (i.e. *ztkhmin*, *ztkmmin*) (Heise, 2005). The parameter choice for the test-runs are presented on Table 1.

4 Results and Conclusions

From the first and second test runs (first and second rows of graphs in Fig. 2), we see that the choice of the minimum value of diffusion coefficients does not effect the low and total cloud cover. The same holds for the inclusion or not of the cloud-ice scheme regarding the simple scheme for cloudiness based on relative humidity (i.e. second and third tests runs referring to the second and third rows of graphs in Fig. 2 respectively). Therefore, the weather situation under study may be considered as particularly suitable regarding isolating the sensitivity of the statistical low cloud cover scheme.

Test runs four to six (Fig. 2) and seven to twelve (Fig. 3) show the response of the low cloud cover to the implementation of the statistical scheme with the parameters *zq_crit* and *zlc0* ranging from 0.8 to 0.2 the former and from 2 to 6 the latter. We tentatively infer that the low cloud cover from the simple scheme for cloudiness (test runs 1-3) as well as test runs 4, 5, 7 and 11 from the statistical cloud scheme give comparable results to the satellite picture. For the test runs where either of the parameters *zq_crit* and *zlc0* take their smallest value (i.e 2 and 0.2 respectively), cloud cover is considerably reduced.

The results, although versatile, look realistic leaving space for further investigation. However, due to the more general impact the statistical cloud scheme may have to the physics of the model, further understanding and testing must be considered before its operational implementation.

References

Betts, A.K., 1973: Non-precipitating cumulus convection and its parameterization. *Quart J. R. Met. Soc.*, 99, pp. 178-196.

Sommeria G. and J. W. Deadorff, 1977: Subgrid-Scale Condensation in Models of Nonprecip-

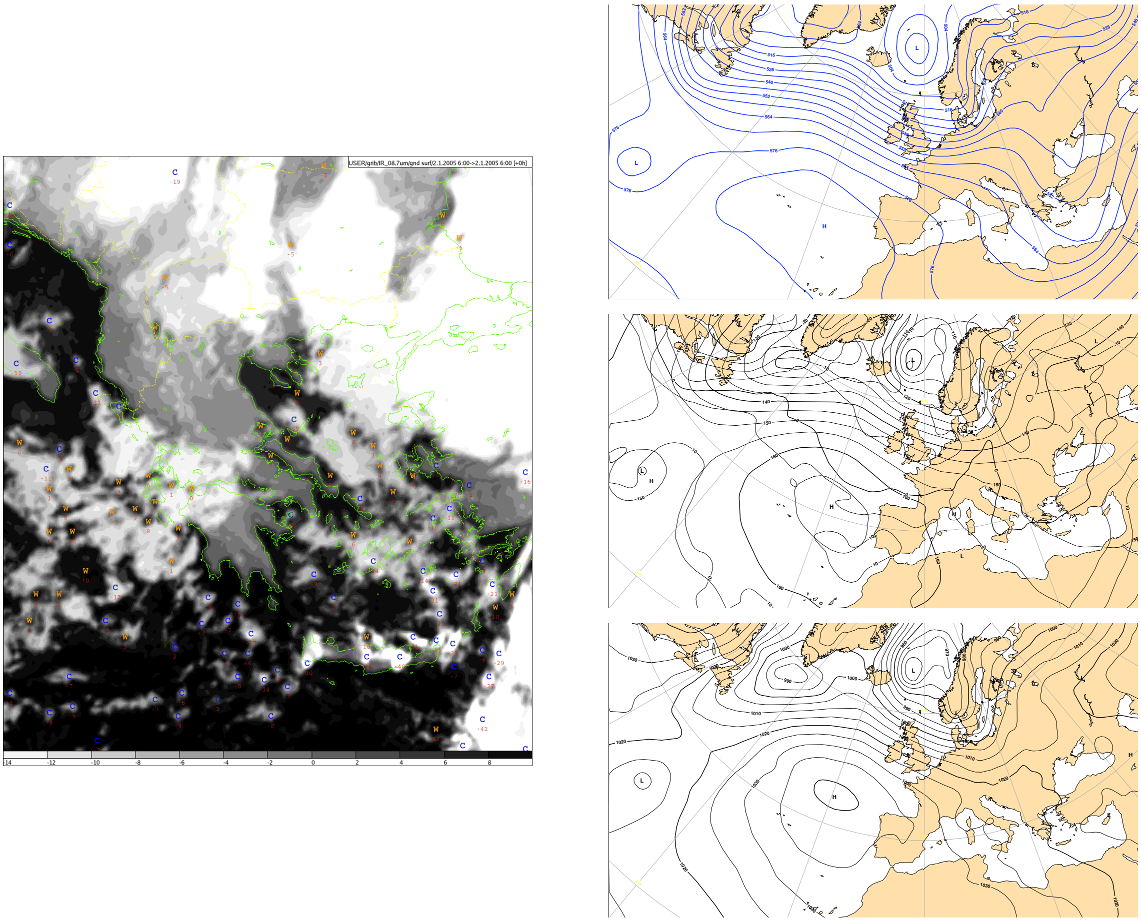


Figure 1: Left: Infrared satellite image, ©METEOSAT. Upper right: Analysis of the geopotential height (meters) for 500 HPa, ©ECMWF. Middle right: Analysis of the geopotential height (meters) and temperature (degrees Celcius) for 850 HPa, ©ECMWF. Lower right: Mean sea level pressure analysis (HPa), ©ECMWF. Date: 2005-01-02:06UTC.

itating clouds. *J. Atmos. Sci.*, 34, pp. 344-355.

Mellor, G.L., 1977: The Gaussian Cloud Model Relations. *J. Atmos. Sci.*, 34, pp. 356-358.

Mellor, G. and T. Yamada, 1982: Development of a Turbulence Closure Model for Geophysical Fluid Problems. *Rev. Geophys. Space Phys.*, 20, pp. 851-875.

Raschendorfer, M., 2005: A new TKE-based Scheme for Vertical Diffusion and Surface-layer Transfer. Deutscher Wetterdienst (DWD), Offenbach and private communication.

Avgoustoglou, E., 2003: Operational Applications of LM at HNMS. *COSMO Newsletter*, No. 3, pp. 28-29.

Doms, G., 2002: The LM cloud ice scheme. *COSMO Newsletter*, No. 2, pp. 128-136.

Doms, G., D. Majewski, A. Müller and B. Ritter, 2004: Recent Changes to the Cloud-Ice scheme. *COSMO Newsletter*, No. 4, pp. 181-188.

Heise, E., 2005: Validation of boundary layer clouds: Test results with the minimum vertical diffusion coefficient set equal to zero in LM. Deutscher Wetterdienst (DWD) and private communication.

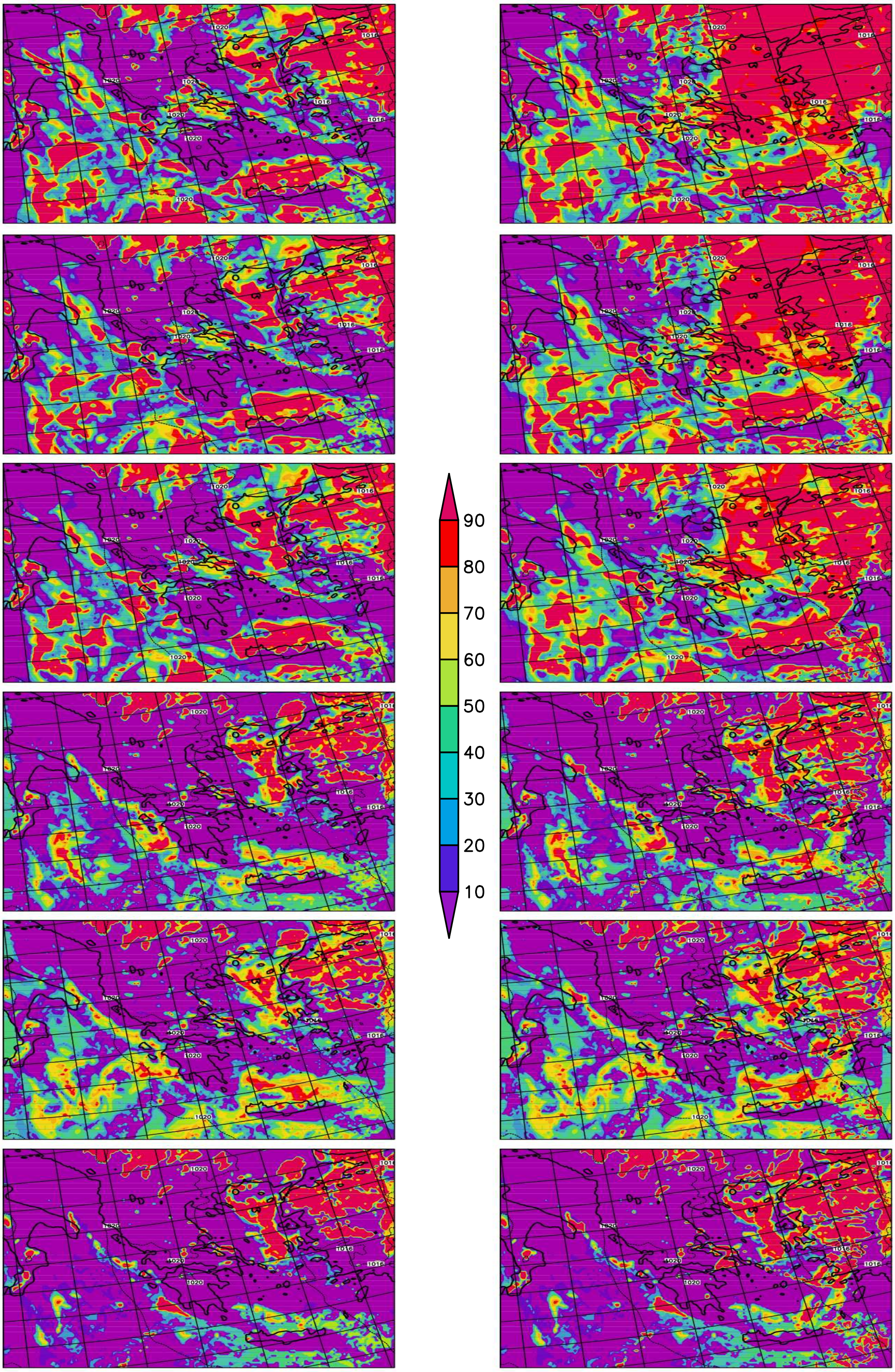


Figure 2: Low (left column) and Total (right column) Cloud Cover forecast (%) and PMSL (HPa), Test Runs No. 1 to No. 6 (from top to bottom) for 2005-01-02:06UTC.

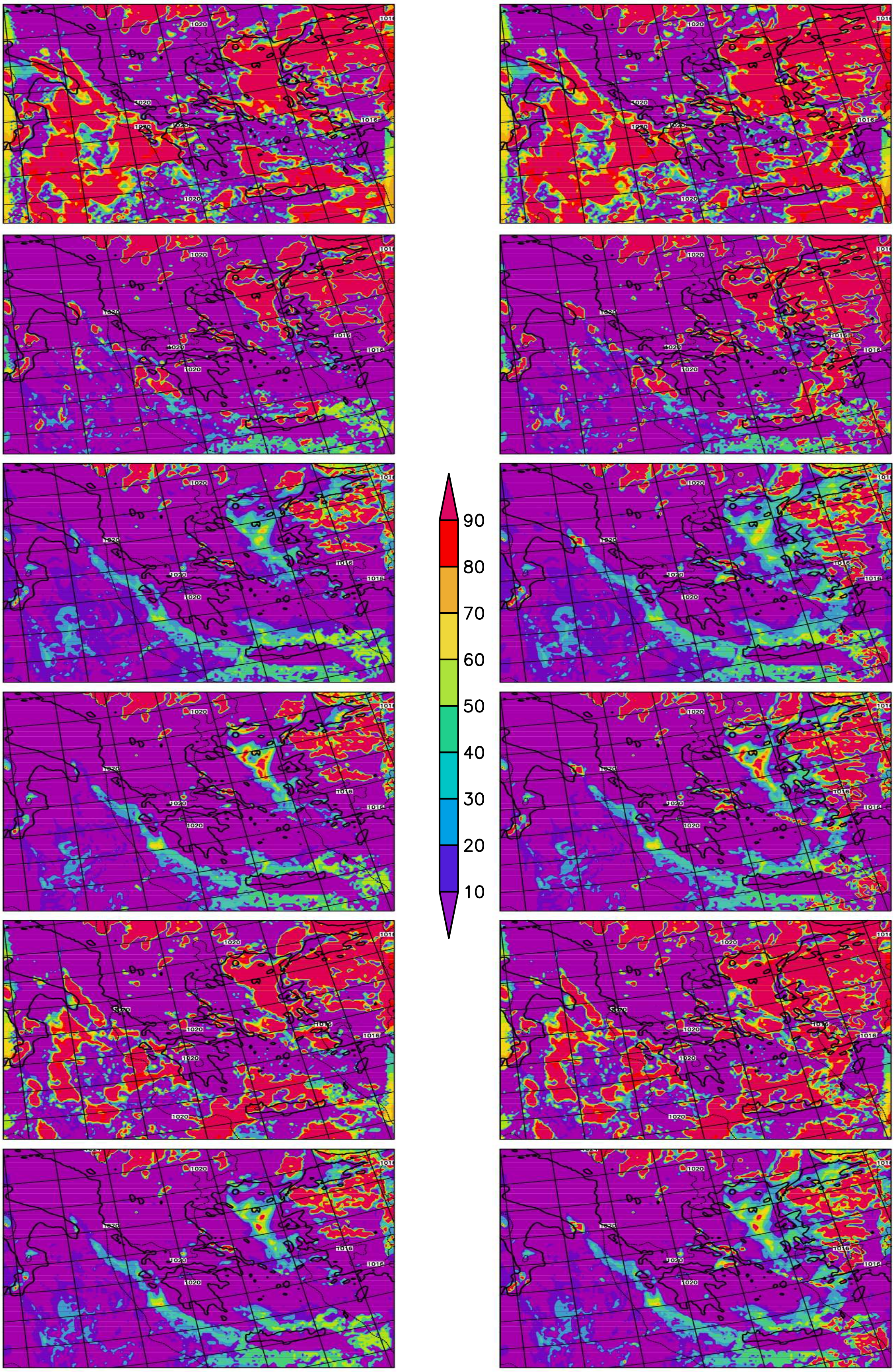


Figure 3: Low (left column) and Total (right column) Cloud Cover forecast (%) and PMSL (HPa), Test Runs No. 7 to No. 12 (from top to bottom) for 2005-01-02:06UTC.

Validation of Boundary Layer Clouds: Test Results with the Minimum Vertical Diffusion Coefficient set equal to Zero in LM (Interim Report on Work Package 3.5.1.)

E. Heise

Deutscher Wetterdienst, Offenbach a.M., Germany

1 Introduction

There is a long lasting problem of LM to rapidly dissolve low stratus/stratocumulus over land in late autumn and in winter. This results in extremely high errors of near surface temperatures. Some earlier tests performed by D. Mironov suggested the prescribed minimum value of the vertical diffusion coefficient to be the reason for this behaviour. The minimum value of $1\,m/s^2$ was introduced because of problems over the North Sea. In the early stage of the operational LM the North Sea appeared to be completely covered by low level clouds most of the time. The introduction of the minimum vertical diffusion coefficient solved this problem. Meanwhile it became obvious that the evaporation rate from water surfaces was too high in LM, contributing to an overdevelopment of depressions over water. A considerable reduction of the evaporation rate by increasing the laminar resistance for scalar fluxes over water by a factor of 10 was introduced operationally in April 2004. This reduced the problem of overdevelopment. At the same time this action should reduce an overprediction of low clouds by reducing the available moisture in the lower layers of the model atmosphere. Therefore it seems to be logically to set back the minimum value of the vertical diffusion coefficient to zero. In a long parallel experiment the consequences of this change are tested. The following questions have to be answered by evaluating the simulations:

1. Is the tendency of LM to dissolve too rapidly low level clouds over land reduced to a tolerable amount?
2. Is the temperature prediction improved?
3. Does LM simulate again a more or less completely cloud covered North Sea?
4. Are there detrimental effects to other results?

A positive outcome of the experiment would result in the answer 'Yes' for the first two questions and in 'No' for the next two questions.

2 Parallel experiments

Two parallel experiments were conducted for the period 01 October 2004 to 31 December 2004. The reference run uses LM Version 3.15 without changes [this is the first version with the so-called aerosol-bug corrected]. In the experiment there is only one change compared to the reference run: the minimum vertical diffusion coefficient is set to zero. Objective verification of the simulation results will be used to assess the overall quality of both the reference run and of the experiment. In addition to the objective verification, a visual inspection of the cloud cover over the North Sea is conducted in order to exclude (hopefully) the occurrence of unrealistic overcast situations.

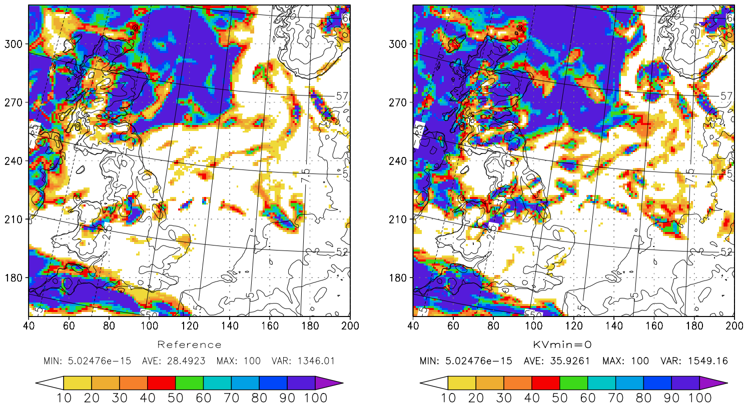


Figure 1: Distribution of low clouds (%) for 10 October 2004, 12 UTC. Left part: reference run, right part: experiment.

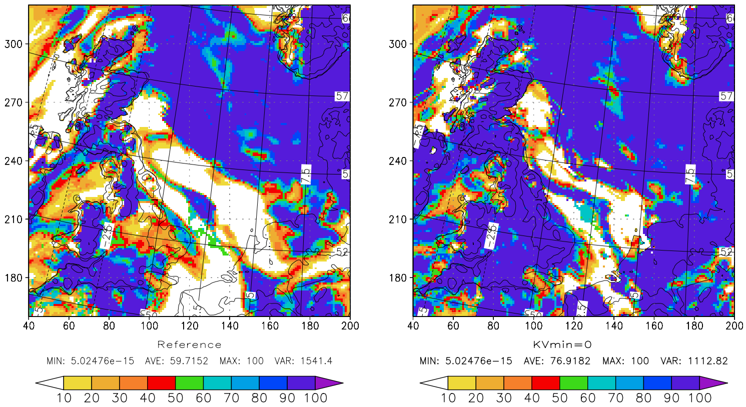


Figure 2: Distribution of low clouds (%) for 30 October 2004, 00 UTC, 24 hours prediction time. Left part: reference run, right part: experiment.

October 2004

In general October is not the month where extremely large errors in the prediction of low stratus clouds are anticipated. Although the domination of convection is not as present as in the summer months, solar insolation is still strong enough to quickly dissolve low level stratus clouds developing during night in stable high pressure systems. But sea surface temperatures are rather warm and could perhaps cause excessive low cloud cover over the large water areas.

As a first example Fig. 1 shows the distribution of low clouds over the North Sea and over adjacent areas for 10 October 2004 after 12 hours prediction time. Both over land and over sea cloud cover is larger in the experiment with the minimum vertical diffusion coefficient set to zero than in the reference run. But there is no indication of an excessive increase over sea. Satellite pictures (not shown here) indicate an overcast situation with low clouds over the northern part of the North Sea with the exception of the area close to the Norwegian coast. The model results are in reasonably good agreement with the satellite pictures.

A second example is presented in Fig. 2. A large increase in the mean cloud cover in the experiment compared to the reference run is ovious (from 59.7 % to 76.3 %). But the increase is rather similar both over the sea and over land areas. Especially the cloudless area off the coast of the British Isles is conserved in the experiment, although it is a bit reduced in size. This area can also be seen in the respective satellite picture (not shown here).

October 2004	**Reference**	**Experiment**
frequency bias		
total cloud cover (0-2/8)	1.26	1.04
total cloud cover (7-8/8)	1.08	1.22
low level cloud cover (0-2/8)	1.09	0.92
low level cloud cover (7-8/8)	0.91	1.23
percent correct		
total cloud cover	60.9	62.1
low level cloud cover	60.5	58.9
mid level cloud cover	60.4	58.4
high level cloud cover	55.9	55.5
2m temperature	73.9	75.1
2m dew point	69.9	71.5
root mean square error (K)		
2m temperature	1.93	1.89
2m dew point	2.36	2.27
true skill statistics		
precipitation $>$ 0.1 mm/6h	49.2	46.3
precipitation $>$ 2.0 mm/6h	38.5	38.5
precipitation $>$ 10.0 mm/6h	22.0	23.2
equitable threat score		
gusts $>$ 12 m/s	25.86	27.52
gusts $>$ 15 m/s	24.12	23.56
gusts $>$ 20 m/s	15.98	12.12
gusts $>$ 25 m/s	3.00	3.80

Table 1: Mean values of verification scores for October 2004 for the reference run and for the experiment. The verification is performed for all observing stations available in the model area.

The objective verification provides mean values over all forecast times for statistical measures. Some of them are shown in Table 1. The frequency bias for small ($\leq$ 2/8) and large ($\geq$ 7/8) values of cloud amount shows the anticipated result of reducing low and increasing high cloud amounts. But in total this effect seems to be too large in this month. This is also reflected in the percent correct values, which show a minor improvement for total cloud cover only. According to the percent correct values and to the root mean square error, the temperatur and dew point predictions improved a little. There is a mixed signal for precipitation, where the prediction of high values is slightly improved. But the success for precipitation yes/no ($>$ 0.1 mm/6h) is decreased considerably. The bias of 2m temperature and dew point (not shown in the table) is on average reduced, especially the warm temperature bias at noon is reduced from 0.78 K to 0.48 K. But the cold bias at 6.00 p.m. increased from -0.20 K to -0.38 K. The positive bias of dew point prediction at noon is reduced from 1.55 K to 1.25 K, and at 6.00 p.m. from 0.45 to 0.20 K. There are only small changes of temperature and dew point biases for the other verification times. There are moderate and not expected changes in the verification of gusts. On average the gusts verification for the experiment is worse compared to the reference run, especially for higher values of the gusts.

November 2004

In November there was especially one case where the synoptic meteorologists vehemently complained about the cloud cover forecast. This was the situation of November 11. The

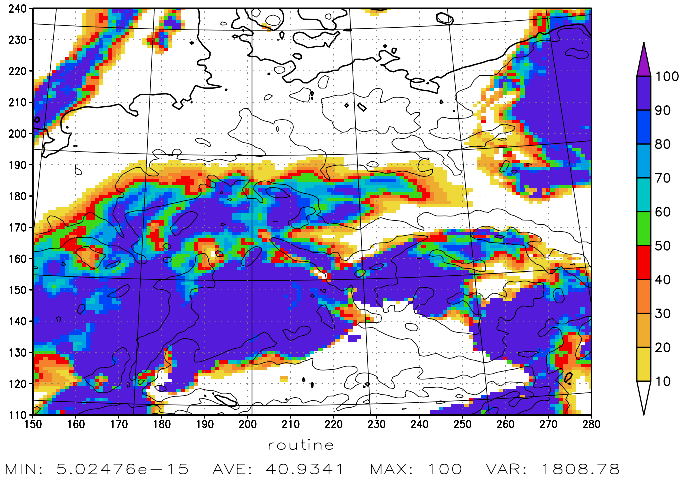


Figure 3: Distribution of low clouds (%) in the routine run of 10 November 2004, 12 UTC, 24 hours prediction time. Validation time 11 November 2004, 12 UTC.

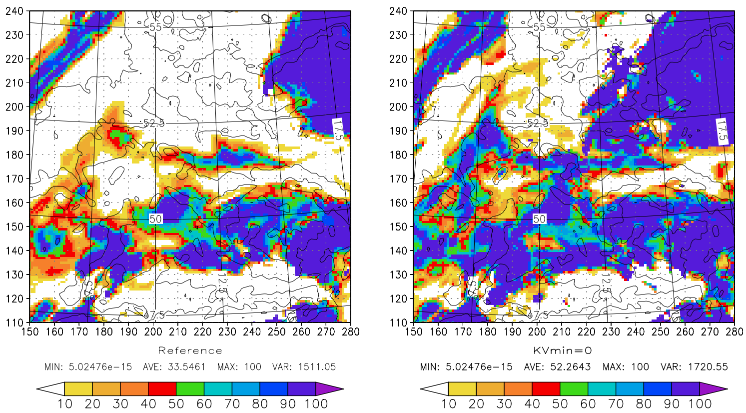


Figure 4: Distribution of low clouds (%) for 11 November 2004, 12 UTC, 12 hours prediction time. Left part: reference run, right part: minimum vertical diffusion coefficient set to zero.

central and the southern parts of Germany were under the very weak influence of a cyclone over Sardinia, whereas there was a ridge of comparably high pressure over northern Germany. In general the surface pressure gradient was small over Germany. According to the satellite pictures, a mainly cloudless band was stretching from the area of Aachen (ca. 51^{o}N, 6^{o}E) to the isle of Gotland. Even in this band occasionally low clouds were present. Also the foothills of the Alps and a small band parallel to the Ore Mountains were mainly free of clouds. The other parts of Germany as well as the Czech Republic were covered by low clouds.
The operational LM-forecast of 12 UTC on November 10 in Fig. 3 (this was the forecast which led to the complaints) shows large areas free of clouds in northern Germany, whereas southern Germany and the Czech Republic are covered with clouds. As in the observations, the foothills of the Alps and the band parallel to the Ore Mountains are cloudless, but the north-south extension of these areas is too large, especially in southeastern Bavaria.
Because in the parallel experiments forecasts were run only at 00 UTC (over a period of 24 hours), a direct comparison with the operational forecast is not possible. Therefore, we have to rely on the 00 UTC forecast of November 11. Fig. 4 shows a comparison of the prediction of low clouds for the reference run and for the experiment. Apparently, the reference run (left hand part of Fig. 4) underestimates the cloud cover even more than the operational forcast

November 2004	**Reference**	**Experiment**
frequency bias		
total cloud cover (0-2/8)	1.47	0.95
total cloud cover (7-8/8)	0.94	1.12
low level cloud cover (0-2/8)	1.22	0.87
low level cloud cover (7-8/8)	0.83	1.10
percent correct		
total cloud cover	63.9	67.9
low level cloud cover	55.6	59.0
mid level cloud cover	61.3	58.0
high level cloud cover	65.7	64.4
2m temperature	70.8	71.7
2m dew point	78.8	79.4
root mean square error (K)		
2m temperature	2.17	2.15
2m dew point	2.43	2.42
true skill statistics		
precip > 0.1 mm/6h	53.4	43.2
precip > 2.0 mm/6h	56.0	56.6
precip > 10.0 mm/6h	47.7	49.6
equitable threat score		
gusts > 12 m/s	34.52	37.86
gusts > 15 m/s	32.46	33.28
gusts > 20 m/s	27.26	21.66
gusts > 25 m/s	14.68	14.00

Table 2: Mean values of verification scores for November 2004 for the reference run and the experiment. The verification is performed for all observing stations available in the model area.

in Fig. 3. But the forecast is significantly improved in the experiment (right hand part of Fig. 4). One can just make out the mainly cloudless band from Aachen to the isle of Gotland (this island is not shown in the figure). The extension of the cloudless bands adjacent to the Alps and to the Ore Mountains is reduced, there are some clouds in northwestern Germany, and eastern Germany is mainly covered by clouds.
The objective verification scores are shown in Table 2. In the reference run a significant overprediction of low cloud cover values of both total and low level cloud cover is diagnosed by the frequency bias. And simultaneously high cloud cover values are underpredicted. The situation changes in the experiment, on average the values of the frequency bias are now closer to one, with the exception of high values of total cloud cover. The percent correct values of total cloud cover and low level cloud cover increase in the experiment compared to the reference run, but mid and high level cloud prediction is better in the reference run. There are only negligible changes in the scores for 2m temperature and dew point. As for the October verification there is a mixed signal for precipitation verification with a small improvement for high values but a large worsening of the prediction of low precipitation rates. There is some improvement of the verification of low gust values, but a significant worsening for high values.

December 2004

From 6 to 09 December 2004 there was a period of significant underestimation of low cloud

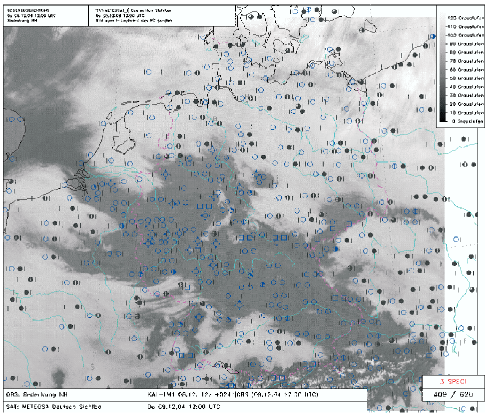


Figure 5: Satellite distribution of cloud cover for 09 December 2004, 12 UTC. Clear sky is shown by dark areas. The station symbols show surface observations (to the left of the vertical bar) and LM-predictions (to the right of the vertical bar).

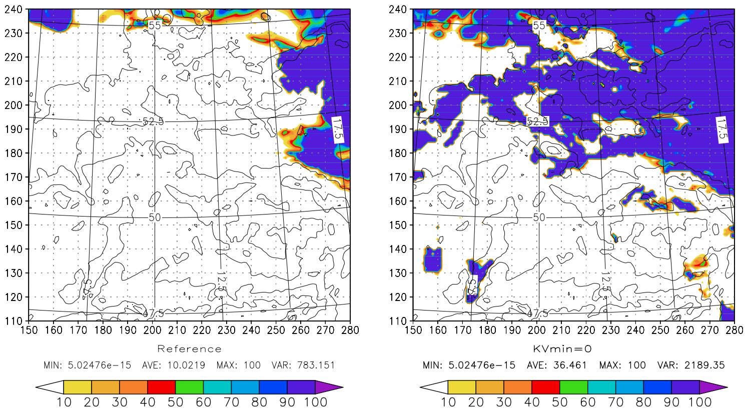


Figure 6: Distribution of low clouds (%) for 09 December 2004, 00 UTC, 12 hours prediction time. Left part: reference run, right part: experiment.

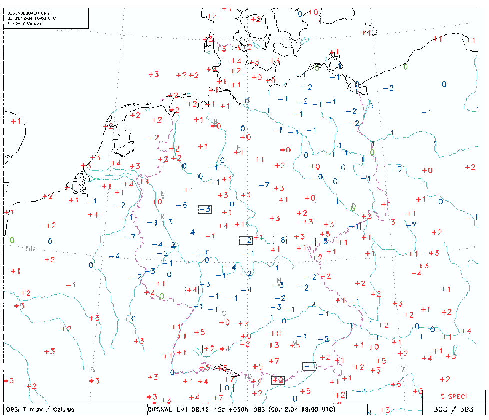


Figure 7: Difference of predicted minus observed 2m maximum temperatures on 09 December 2004.

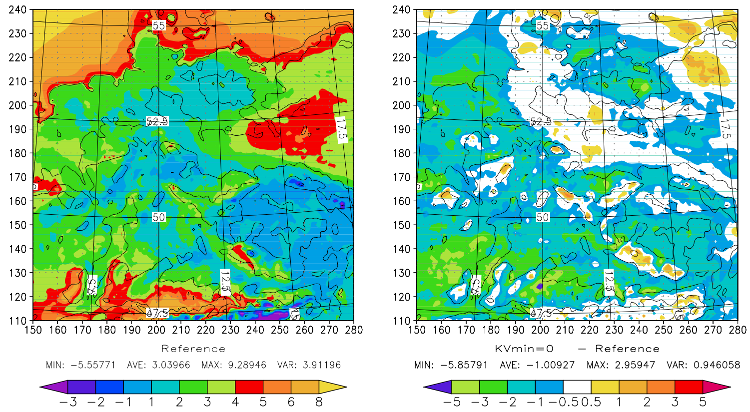


Figure 8: Predicted maximum 2m temperatures in the reference run (left part), and difference experiment minus reference run (right part) on 09 December 2004.

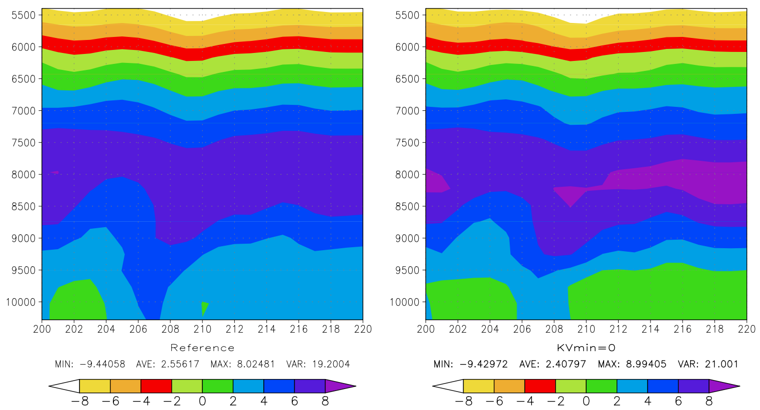


Figure 9: Cross section of temperature in the lowest 20 layers of LM on 09 December 2004. The cross section runs in zonal direction through the northern part of the Thuringian Forest on model row 163 from model column 200 to 220.

cover in LM over Germany. The case of 09 December is used here for demonstration. The satellite picture in Fig. 5 shows large parts of Germany and the surroundig area cloud covered with the exception of a northwest/southeast oriented band in central and southern Germany, which is mainly cloudfree. Looking at the reference run in Fig. 6, only part of Poland is cloud covered in the LM simulation. In the experiment the simulation is improved, northern Germany is mainly covered by clouds, but still southern Germany and northeastern France are nearly completely free of clouds. In the operational model, which is nearly identical to the reference run, the large errors in cloud cover led to large errors in the predicted 2m maximum temperatures (Fig. 7). In southern Germany predicted temperatures are too warm by up to 7 K (synoptic station Altenstadt), whereas in the areas, which are cloudfree in the observations, predicted temperatures are too low.

On average, the 2m temperature maxima are lower in the experiment by 1 K (Fig. 8). The largest temperature reduction occurs in southwestern Germany, but there is also a reduction of temperatures in northeastern Germany, where the operationally predicted temperatures are too cold already. A significant warming takes place at mountain tops of the Thuringian Forest, the Ore Mountains and the Bavarian Forest, where the operational forecast at some stations showed significantly too cold temperatures. No change in the cloud cover prediction occured here. The reason for this significant temperature increase is explained in Fig. 9. In the reference run the strong temperature inversion in the lower atmosphere is smoothed considerably compared to the experiment. Clearly this is due to the prescribed minimum vertical diffusion coefficient in the reference run.

December 2004	**Reference**	**Experiment**
frequency bias		
total cloud cover (0-2/8)	1.64	1.12
total cloud cover (7-8/8)	0.82	1.01
low level cloud cover (0-2/8)	1.31	0.93
low level cloud cover (7-8/8)	0.75	1.05
percent correct		
total cloud cover	63.0	71.0
low level cloud cover	58.7	65.4
mid level cloud cover	68.6	67.2
high level cloud cover	77.7	77.0
2m temperature	63.9	64.2
2m dew point	76.4	78.0
root mean square error (K)		
2m temperature	2.67	2.70
2m dew point	2.89	2.85
true skill statistics		
precip > 0.1 mm/6h	60.7	47.0
precip > 2.0 mm/6h	63.9	63.9
precip > 10.0 mm/6h	44.6	42.7
equitable threat score		
gusts > 12 m/s	34.18	35.54
gusts > 15 m/s	29.76	29.84
gusts > 20 m/s	20.18	16.30
gusts > 25 m/s	8.68	4.17

Table 3: Mean values of verification scores for December 2004 for the reference run and for the experiment. The verification is performed for all observing stations available in the model area.

The results of objective verification are presented in Table 3. According to frequency bias and percent correct the cloud cover forecast was improved significantly in the experiment compared to the reference run. On average there is no significant change of temperature and dew point forecasts. Again, there is a large degradation of precipitation forecast, and also a degradation of the prediction of high gust values.

3 Summary and discussion

The LM tends to rapidly dissolve low stratus in late autumn and in winter. This leads to cloudfree situations in the model in contrast to overcast situations in reality and to associated temperature errors. This problem was suspected to be caused by the introduction of a minimum vertical diffusion coefficient of 1 m/s^2 some years ago. The reason for using the minimum vertical diffusion coefficient were problems with cloud cover over water areas. Here cloud cover tended to be much too high. As last year evaporation over water surfaces was reduced considerably, the minimum vertical diffusion coefficient might turn out to be dispensable.

A couple of single test cases with the minimum vertical diffusion coefficient set to zero showed improvements in the cloud cover prediction. Therefore, the period 1 October 2004 to 31 December 2004 was simulated with a reference run and an experiment. The aim of the experiment was to quantify the effect of this reduction on the simulation of low clouds in autumn and winter.

The objective verification showed a mixed signal. On the one hand cloud cover prediction was improved, but on the other hand there were detrimental effects on precipitation and gusts. The questions at the end of the introduction gain the following answers:

1. Is the tendency of LM to dissolve too rapidly low level clouds over land reduced to a tolerable amount? **Yes**

2. Is the temperature prediction improved? **There was only a marginal improvement**

3. Does LM simulate again a more or less completely cloud covered North Sea? **No evaluation has been done so far**

4. Are there detrimental effects to other results? **Yes**

These results cannot be considered as satisfying. A rigorous investigation of the results is necessary to find out the reasons for the failure of the experiment.

Improved Diagnosis of Convective and Turbulent Gusts: Test Results of new Gust Parameterizations (Interim Report on Work Package 3.10.2.)

E. Heise

Deutscher Wetterdienst, Offenbach a.M., Germany

1 Introduction

In the former operational version of LM, gusts were diagnosed by increasing the wind in the lowest model layer depending on stability. In general forecasters complained about seriously underestimated gusts in convective situations and overestimated gusts in non-convective situations. In a first attempt to improve the gust parameterisation, convective gusts were considered in addition to the operational gust determination. The kinetic energy of convective downdrafts was computed depending on the vertical integral of their negative buoyancy. The square root of kinetic energy, multiplied by a tuning parameter to account for the horizontal distribution of the downward directed kinetic energy, was considered to be the convective gust. The results were not completely satisfactory as convective gusts still seemed to be underestimated. In a new approach described here, the water loading effect was included in the parameterisation of convective gusts. Additionally, a parameterisation of gusts after Brasseur (2001) was introduced to improve the results in non-convective situations.

2 Theory

In the following the parameterisations of convective and turbulent gusts are shortly described.

Convective gusts

The parameterisation of convective gusts is based on Nakamura et al. (1996), who proposed to use the downdraft kinetic energy produced by the negative buoyancy and the direct transport of momentum from higher layers to the surface:

$$V_{gust,con0} = \sqrt{\alpha \int_0^H 2g \left(\frac{\Delta\theta}{\theta} + \gamma q_r \right) dz + \beta V(H)^2}\ , \tag{1}$$

where H is the height of the downdraft generation, $\Delta\theta$ is the difference of the potential temperatures of downdraft and environment, and q_r is the mixing ratio of precipitation. The factor α accounts for the distribution of kinetic energy to different directions.

First operational tests showed significant problems with the direct momentum transport term. In cases of rather light but penetrative convection connected to the polar front, this term is able to transport momentum from the layer of the jet-stream down to the surface, producing unrealistically high gusts. Therefore the last term in (1) is neglected in the following ($\beta = 0$). Operationally also the effect of water loading is neglected ($\gamma = 0$) and we use $\alpha = 0.2$. This configuration is used in the reference simulations. In the test simulations we increase the tuning parameter for the horizontal distribution of kinetic energy to $\alpha = 1/\pi$ and we include the water loading effect by $\gamma = 1$.

An additional problem of the parameterisation of convective gusts is the case of small or even vanishing (due to evaporation below cloud base) convective precipitation. Although also in reality convective gusts may occur without convective precipitation, this situation seemed to be overpredicted both with respect to the frequency of occurence and with respect to the windspeed of the gusts. Therefore, the gusts parameterised by (1) with the parameters as given above are suppressed, if the convective precipitation rate at the surface falls below $0.015\,mm/h$. This threshold value is chosen in accordance with the interpretation of model results, where showers or thunderstorms are diagnosed only if this value is exceeded. Therefore, the final convective gust in the test simulation is determined by

$$\begin{array}{rcll} V_{gust,con} = & V_{gust,con0} & \text{if} & rr_{con} > 0.015mm/h \\ V_{gust,con} = & 0.0 & \text{if} & rr_{con} \leq 0.015mm/h \quad , \end{array} \tag{2}$$

where rr_{con} is the convective precipitation rate. This reduction in case of low convective precipitation is not used in the reference experiment, here $V_{gust,con} = V_{gust,con0}$.

Turbulent gusts

The present operational method for the determination of turbulent gusts relies on the windspeed in the lowest model level and a stability dependent increase of the wind speed:

$$V_{gust,turb} = |V_{ke}| + 7.2 \cdot u_* \quad , \tag{3}$$

where V_{ke} is the wind speed in the lowest model level, and u_* is the friction velocity.

The formulation (3) makes the gust determination dependent on the height of the lowest model level. Additionally, the formulation is not really physically based but rather a product of a limited amount of tuning. Therefore it seemed to be appropriate to change to the method proposed by Brasseur (2001). He based his approach on the consideration of turbulent and buoyant energies. Brasseur's (2001) basic assumption is that turbulent motions are able to transport momentum downward from a height z_p to the surface as long as the energy of large turbulent eddies averaged from the surface to the height z_p is larger than the buoyant energy between the surface and z_p:

$$\frac{1}{z_p}\int_0^{z_p} 0.5\; q^2(z)dz \geq \int_0^{z_p} g\frac{\Delta\theta_v(z)}{\theta_v(z)}dz \tag{4}$$

Here $0.5\; q^2$ is the turbulent kinetic energy (in J/kg), $\Delta\theta_v$ is the difference of the virtual potential temperatures between the environment and a rising parcel. The integration starts from the surface, which is assumed at $z = 0$, and is continued as long as the inequation holds. Then the largest value of the grid-scale momentum $V(z)$ between $z = 0$ and $z = z_p$ is assumed to be the maximum gust at the surface:

$$V_{gust,turb0} = Max[V(z), z = 0...z_p] \tag{5}$$

Clearly, this approach is more or less independent of the layer structure of the model. But tests revealed that situations exist with significantly too low diagnosed gusts using this method. To overcome this problem, a contribution of the turbulent kinetic energy at the surface $E_{z=0}$ is added to yield the final value of the turbulent gust:

$$V_{gust,turb} = \sqrt{V_{gust,turb0}^2 + 0.5q_{z=0}^2} \tag{6}$$

In convective situations this parameterisation might experience the same problems as were noted in the first attempt to parameterise convective gusts ($\beta \neq 0$, see above). In such

	turbulent	**convective**	
reference run	eq. (3)	eq. (1)	$\alpha = 0.2$ $\beta = 0.0$ $\gamma = 0.0$
experiment	eqs. (4)-(6)	eqs. (1),(2)	$\alpha = 1/\pi$ $\beta = 0.0$ $\gamma = 1.0$

Table 1: Setup of operational and modified runs

situations the difference in virtual temperature might become negative over a large height interval, making the inequation being valid up to very large heights. Therefore, this parameterisation is strictly confined to turbulent gusts of the lower troposphere by defining an upper limit for the integration: $z_{p,max} = 2000m$. Tests in non convective situations with an increased value of this limit (4000 m) did not show any change in the resulting gusts.

Equations and parameters used in the reference runs and in the experiments

Table 1 compiles the equations and parameters used for the reference runs and for the experiments.

3 Single test cases

Predominantly convective situations

In the following examples it would be somewhat difficult to see the effect of the convective gusts, because normally also turbulent gusts play a role. Therefore, additional runs have been made with the turbulent gusts switched off, in order to see the changes in the convective gusts more clearly. The respective results are shown here.

July 10, 2002

One of the most important test cases for convective gusts is the Berlin storm of 10 July 2002, where in the late afternoon and in the evening a cold front and a prefrontal mesoscale convective complex (MCC) caused violent gusts of up to 42 m/s in Berlin and in the surrounding area (Gatzen, 2004). Gusts of this force were not predicted by the operational models. A convergence line, which in reality was transformed into the MCC, was simulated far to the east of the cold front. There was no convective rain predicted at the convergence line, but moderate gusts were diagnosed here. The cold front and the associated convective rain were simulated rather well, but also only moderate gusts were diagnosed here. This was in accordance with the observations, which showed the intensity of the cold front to weaken rapidly and the MCC becoming the main feature. In fact, this situation turned out to be a very difficult task for model prediction. One reason might have been the very special synoptic situation of a derecho-development (Gatzen, 2004) in connection with the MCC. Nearly all of the most violent gusts ($>$ 34 m/s) were observed in the region of the derecho.

From the viewpoint of gust diagnosis the most serious problem appeared to be the determination of the level of free sink. In terms of convection parameterisation here the downdrafts originate and the integration in (1) ends. The level of free sink was diagnosed at rather low altitudes in some areas of heavy convective rain in the region of the cold front. This led to much too low values of the convective gusts. This problem could not yet be solved.

The distribution of convective precipitation from 9 UTC to 12 UTC is shown in Fig. 1. The approaching cold front can be seen over northern Germany, whereas the convergence line

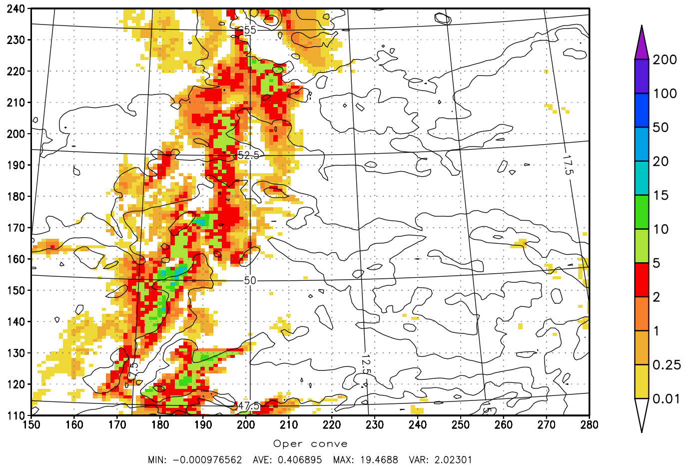


Figure 1: Distribution of convective precipitation [mm] from 09 UTC to 12 UTC on 10 July 2002

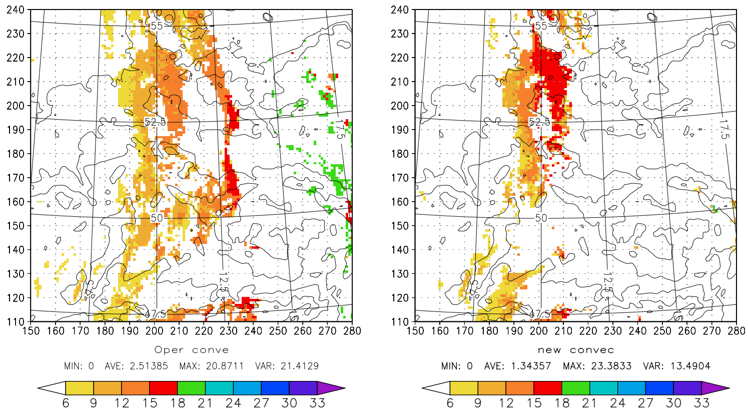


Figure 2: Distribution of the maximum convective gusts [m/s] from 09 UTC to 12 UTC on 10 July 2002. Left hand part: operational version of the gust diagnosis, right hand part: modified diagnosis

does not show up in the precipitation distribution. But the distribution of convective gusts in the operational version clearly shows the convergence line (left part of Fig. 2). In the modified version (right part of Fig. 2) all gusts in areas without convective precipitation are suppressed. In the region of the cold front, the gusts are higher than before because of the consideration of the effect of water loading and because of the new value for α. These effects are more pronounced in the evening, see Fig. 3. Maximum gusts are higher bei nearly 5 m/s, although still far from reaching the observed values. It should be mentioned that the area averaged value of the convective gusts is lower in the modified version of gust diagnosis than in the operational version because gusts connected to low precipitation rates are suppressed.

August 12, 2004

In the afternoon of August 12 a very narrow low pressure trough crossed the southwestern and southern parts of Germany with rather high convective precipitation rates and strong gusts. The observed precipitation for 06 UTC to 12 UTC (Fig. 4) shows a band of high precipitation stretching from Switzerland/northeastern France to the Netherlands, which is clearly separated from the dry area to the east. The LM simulation (Fig. 5) is partly successful. The trough is well positioned, but LM erroneously predicts moderate convective

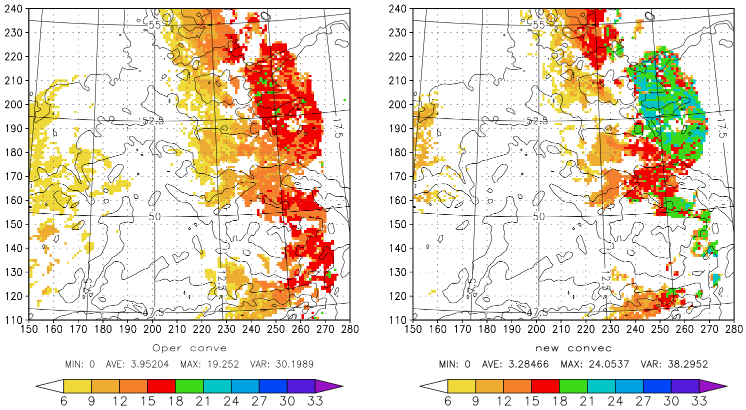


Figure 3: Distribution of the maximum convective gusts [m/s] from 18 UTC to 21 UTC on 10 July 2002. Left hand part: operational version of the gust diagnosis, right hand part: modified diagnosis.

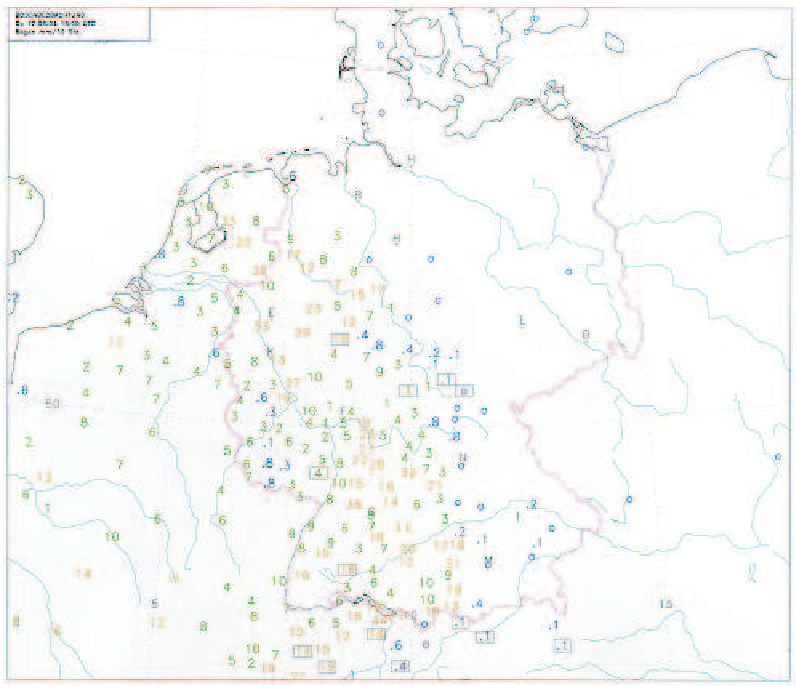


Figure 4: Distribution of observed precipitation [mm/12h] from 06 UTC to 18 UTC on 12 August 2004.

activity in Northeastern Germany, the Czech Republic and in Slovakia. In southwestern and southern Germany the precipitation amount predicted, with values up to some 50 mm/12h, is slightly larger than observed. In contrast, in the Benelux countries predicted precipitation amounts are somewhat too low.

The distribution of observed gusts (Fig. 6) for the period 15 to 18 UTC shows highest values in the easternmost part of the trough. Here values up to 70 knots are reported. Values in the western part of the trough are lower with the exception of reports from high level stations (indicated by a square around the value). The general structure is very well simulated by LM (Fig. 7). But in the operational version (left hand part of the figure) the predicted gusts are much too low (note the dimension m/s in Fig. 7 instead of knots in Fig. 6). The new version

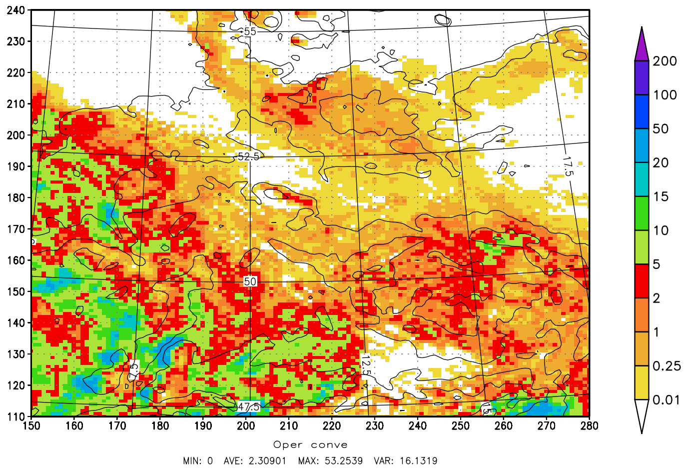


Figure 5: Distribution of predicted precipitation [mm/12h] from 06 UTC to 18 UTC on 12 August 2004.

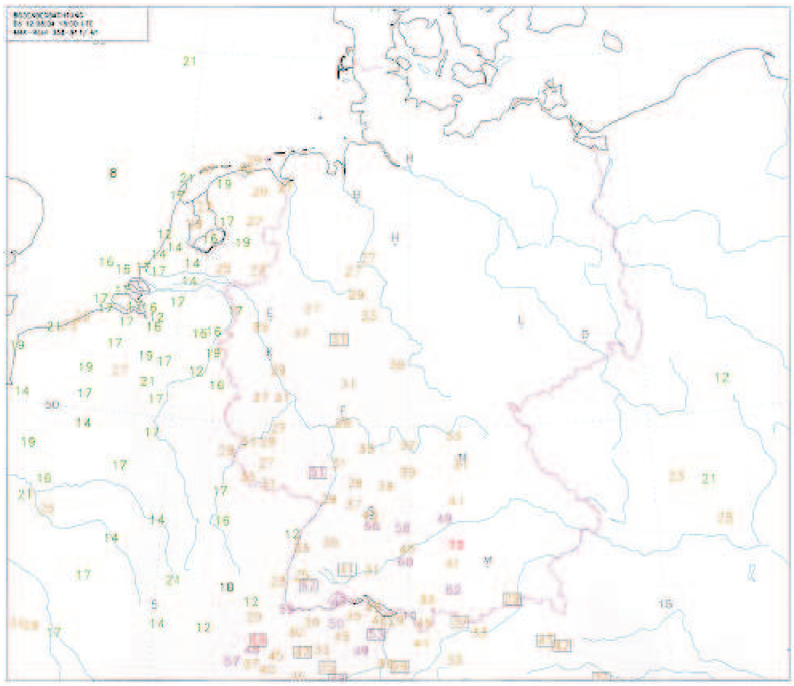


Figure 6: Distribution of observed gusts [knots] from 15 UTC to 18 UTC on 12 August 2004.

of the gust parameterisation (right hand part of Fig. 7) shows a significant improvement, although still the values are too low. Because of the erroneous prediction of convection east of the trough region also gusts are predicted here. And inevitably these gusts are higher in the modified version of the gust parameterisation. Nevertheless, as in the case of 10 June 2002 the area averaged value is lower for the modified version.

A situation with low values of gusts: April 07, 2004

This day was simulated in order to test the behaviour of both the new gust parameterisations in a situation of no significant gusts. There was no need for warning of gusts. On April 07, 2004, widespread showers occured over Germany. In central and in southern Germany gusts

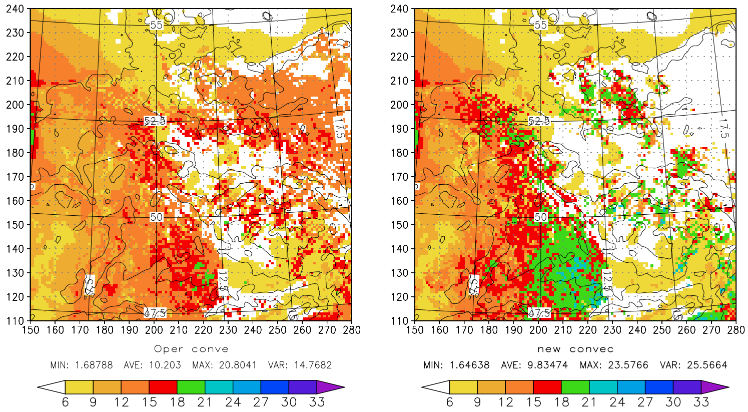


Figure 7: Distribution of the maximum convective gusts [m/s] from 15 UTC to 18 UTC on 12 August 2004. Left hand part: operational version of the gust diagnosis, right hand part: modified version of the gusts diagnosis.

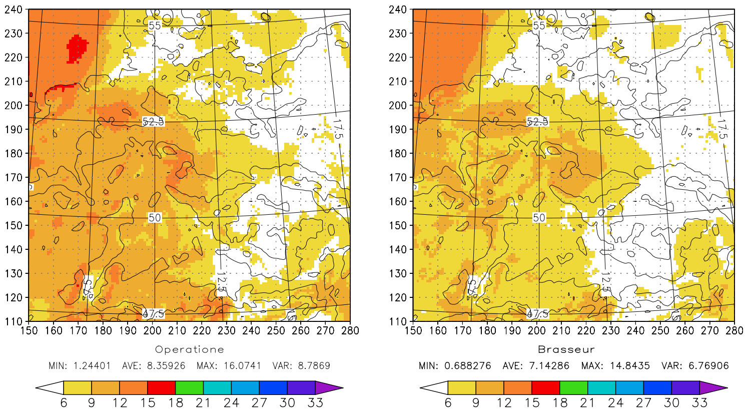


Figure 8: Distribution of predicted gusts [m/s] from 06 UTC to 09 UTC on 07 April 2004. Left hand part: Operational version of gust diagnosis, right hand part: modified diagnosis of convective and turbulent gusts.

of up to 12 m/s were reported. Occasionally somewhat higher values occured at higher level stations. Only over the German Bay values up to 15 m/s were observed. In this case the maximum of convective and turbulent gusts is shown. The operational prediction was rather successful with perhaps slightly too high values (Fig. 8, left hand part). The modified parameterisations (Fig. 8, right hand part; both the convective and the turbulent parts use the new versions of the gusts parameterisation) show a small reduction of the values, which might even better fit to the observations.

Predominantly turbulent situations

In this section some cases with turbulence dominated gusts are investigated. In all these cases convection does not play a significant role. Therefore, no separation between the different origins of the gusts is made, the maximum gusts are shown.

May 13, 2004

This situation is characterised by very low gusts over most of the LM-area. Exceptions are the Rhone valley and the Golfe du Lion. In these regions a strong Mistral is blowing. Gusts of 27 m/s are reported in Orange in the Rhone valley and 30 m/s at Cape Bear (southeast of Perpignan). The results of the simulations are shown in Fig. 9 for the whole model area,

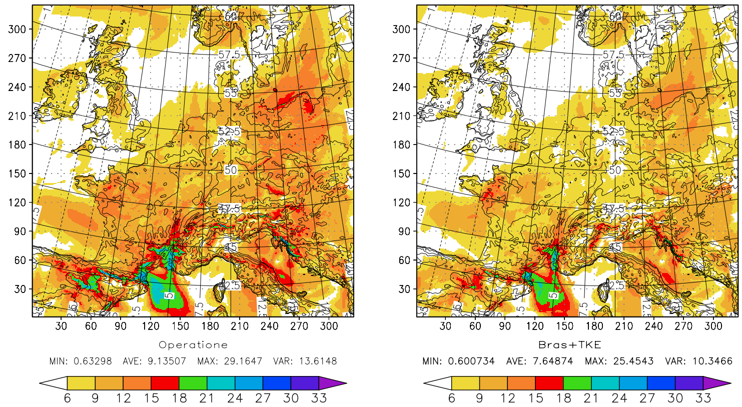


Figure 9: Distribution of predicted gusts [m/s] from 09 UTC to 12 UTC on 13 May 2004. Left hand part: Operational version of gust diagnosis, right hand part: modified diagnosis of gusts after Brasseur.

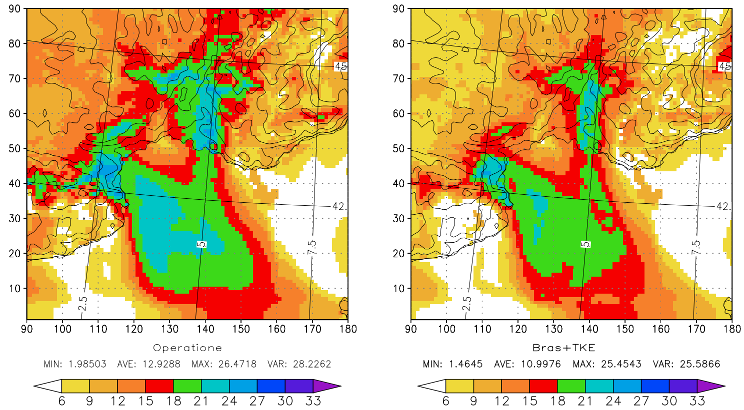


Figure 10: Distribution of predicted gusts [m/s] from 09 UTC to 12 UTC on 13 May 2004 for the Rhone valley and the Golfe du Lion. Left hand part: Operational version of gust diagnosis, right hand part: modified diagnosis of gusts after Brasseur.

and in Fig. 10 for the region of the Rhone valley and the Golfe du Lion. Observations are not shown here. Over most of the LM-area the operational version overestimates the gusts. Either no gusts are reported, or the reported values are around 8 to 10 m/s. Only at the coast of the Baltic Sea and in Brittany values of up to 12 m/s are reported. Here the values are captured quite well by the predictions of the operational model. The Brasseur gusts are lower on average by 1.5 m/s. These lower values better fit to the observations than the higher values in the operational model. E.g., in the region of the Thuringian Forest, the Ore Mountains, the Bavarian Forest, the Czech Republic and Slovakia no gusts are reported, but the operational model simulates values around 15 m/s. Only at the coast of the Baltic Sea the Brasseur gusts might be slightly too low. In the region of the Rhone valley the Brasseur method restricts the highest gusts to the inner part of the valley. This seems to better reproduce the few observations available. But the highest values observed here are not met by both the model predictions.

Nov 18/19, 2004

Two intensive depressions crossed Germany in November 2004. The first one on November 18, 2004 caused gale-force winds in northern Germany, whereas the second one on November 19, 2004 mainly hit southern Germany. For the first situation Fig. 11 shows the distribution

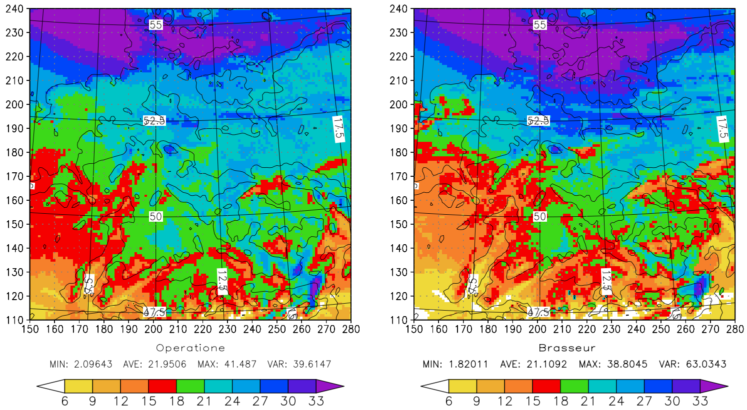


Figure 11: Distribution of predicted gusts [m/s] from 03 UTC to 06 UTC on 18 November 2004 for Germany. Left hand part: Operational version of gust diagnosis, right hand part: modified diagnosis of gusts after Brasseur.

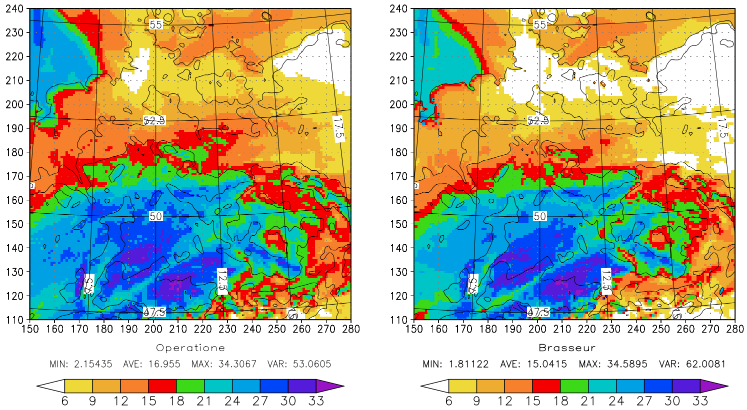


Figure 12: Distribution of predicted gusts [m/s] from 03 UTC to 06 UTC on 19 November 2004 for Germany. Left hand part: Operational version of gust diagnosis, right hand part: modified diagnosis of gusts after Brasseur.

of gusts in Germany. Compared to the operational version the area of high wind speeds is somewhat more extended using the Brasseur version. But the maximum and the average values are slightly reduced. Especially in southern Germany the gusts are less pronounced with the Brasseur version. The somewhat lower values in northern Germany are closer to the observations. The results are similar for the second case on 19 November (Fig. 12). There is a slight reduction of both the average and the maximum values. A more drastic reduction with the Brasseur method occurs in the region of Corsica (Fig. 13). The maximum value is reduced from 46 to 36.5 m/s. The latter value seems to be more likely, but no observations are available to verify the results of one or the other method.

4 Parallel experiments

The parameterisations as described above have to be tested in parallel experiments in order to judge their overall performance with respect to the operational verification. The period 01 July 2004 to 31 August 2004 was chosen to represent mainly convective gusts, whereas for turbulent gusts the period was 01 November 2004 to 31 December 2004. For the summer period the operational determination of turbulent gusts is used in combination with the

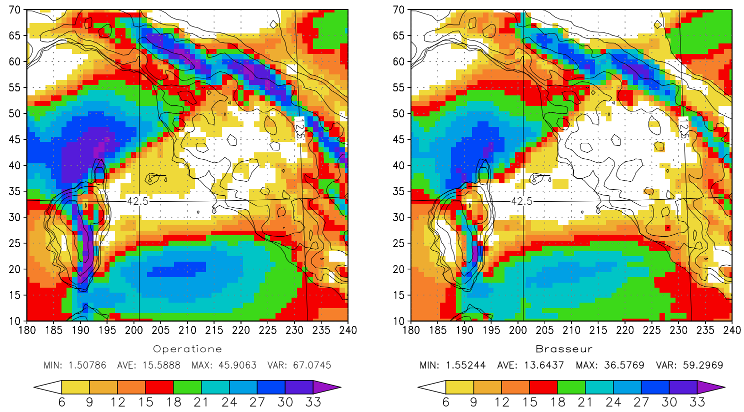


Figure 13: Distribution of predicted gusts [m/s] from 03 UTC to 06 UTC on 19 November 2004 for the regions of Corsica and part of Italy. Left hand part: Operational version of gust diagnosis, right hand part: modified diagnosis of gusts after Brasseur.

		convective gusts	**turbulent gusts**
summer	**reference runs**	operational	operational
	experiments	modified	operational
winter	**reference runs**	operational	operational
	experiments	operational	modified

Table 2: Determination of gusts in the reference runs and in the experiments

modified version of convective gusts. For the autumn/winter period the operational diagnosis of convective gusts is combined with the modified version for turbulent gusts. Table 2 shows the usage of the different variants of gust diagnosis according to Table 1.

For each of the periods a reference run was performed. Reference run and experiment for the turbulent gusts use LM Version 3.15, and the reference run and the experiment for the convective gusts are based on LM Version 3.16.

Convective gusts

At present no objective verification results are available, because reference run and experiment were not identical in all results but gusts. Therefore both runs had to be repeated.

Turbulent gusts

Reference run and experiment were finished successfully, including the operational verification of the results. Table 3 shows mean verification results for the three months for all significant gusts (> 12 m/s) and for severe gusts (> 20 m/s). The evaluation is for all stations. The statistical measures are: equitable threat score (ETS), frequency bias (FBI), probability of detection (POD), and false alarm rate (FAR). Optimum values are 100 (%) for ETS and POD, 1 for FBI, and 0 (%) for FAR. If the change from the reference run to the experiment is larger than 5 %, the numbers are color-coded: green for improvement and red for deterioration.

In general the results in Table 3 are dissapointing. Just counting the coloured numbers: there are only 7 in green but 13 in red. Looking at all gusts (> 12 m/s) the decrease of the absolute number of simulated gusts compared to observed gusts (FBI) and the reduction of the false alarm rate are positive. But there is a significant decrease in the probability of

	ETS		FBI		POD		FAR	
	ref	exp	ref	exp	ref	exp	ref	exp
> 12 m/s								
October	25.9	24.7	1.7	1.2	66.1	53.8	59.6	55.0
November	34.5	34.5	1.6	1.4	77.7	70.2	51.1	47.4
December	34.2	33.6	1.7	1.3	84.6	74.1	50.4	47.4
> 20 m/s								
October	16.0	9.5	0.7	0.4	25.7	13.3	63.4	67.8
November	27.3	23.3	1.1	1.3	46.9	64.7	58.4	65.4
December	20.2	18.0	1.2	1.4	38.3	38.6	64.8	70.6

Table 3: Verification of the results of the reference run (ref) and of the experiment using the modified diagnosis of turbulent gusts (exp). See text for details.

		> 12 m/s		**> 20 m/s**	
		ref	exp	ref	exp
ETS	stable	32	31	22	17
	unstable	31	34	28	21
FBI	stable	1.9	1.6	1.2	1.3
	unstable	1.8	1.4	1.1	1.0
POD	stable	80	71	41	37
	unstable	82	71	46	41
FAR	stable	56	52	63	70
	unstable	53	45	57	59

Table 4: Verification of the results of the reference run (ref) and of the experiment (exp), for stable and unstable conditions. See text for details.

detection of gusts. A major positive outcome is the considerable increase in the probability of detection for severe gusts in November and the slight improvement in December. But for severe gusts also the false alarm ratio increases considerably in all three months. Also, for severe gusts the equitable threat score is reduced to much lower values in the experiment compared to the reference run.

The operational verification provides estimates of the diurnal cycle of the different scores. Here some indication on a dependence on stability appeared. Therefore, all scores were determined separately i) for 12 UTC (on average unstable stratification) and ii) for 06 and 18 UTC (on average stable stratification). The results for the three months were averaged for this evaluation. Looking at the statistical measures depending on stability in Table 4, it can be observed that indeed with only one exception the scores are better in unstable than in stable situations. But with respect to this behaviour there is no significant difference between reference run and experiment.

5 Continuation of the workpackage during the next one year phase

It is suggested to continue the work during the next phase. The work will comprise the following aspects.

- Evaluation of the results of the reruns of reference run and experiment using the modified diagnosis of convective gusts.

- If necessary, retuning of the parameter α in (1) and conducting a shorter experiment.
- In depth evaluation of the results of reference run and experiment using the Brasseur-method for turbulent gusts.
- Retuning of the Brassuer-method if possible.
- Rerun of reference run and experiment for turbulent gusts, evaluation of the results.

References

Brasseur, O., 2001: Development and application of a physical approach to estimating wind gusts. *Mon. Weath. Rev.*, 129, pp. 5-25

Gatzen, Chr., 2004: A Derecho in Europe: Berlin, 10 July 2002. *Weather and Forecasting*, 19, pp. 639-645

Nakamura, K., R. Kershaw, and N. Gait, 1996: Prediction of near-surface gusts by deep convection. *Meteorol. Appl.*, 3, pp. 157-167

A Three-Category Ice Scheme for LMK

T. Reinhardt and A. Seifert

Deutscher Wetterdienst, Offenbach, Germany

1 Introduction

The present microphysics scheme operational in LM at 7 km mesh size is dedicated to precipitation formation in stratiform clouds. Ice hydrometeors occuring in deep convective clouds – such as graupel and hail – are neglected. In order to simulate such clouds explicitly – as we intend to do with LM's high-resolution version LMK (Doms and Förstner, 2004) –, ice particles with larger fall velocities than snow must be included to allow for a reasonable physical description of precipitation formation. Therefore the present cloud ice scheme has been extended to include graupel as a third ice category. The scheme considers the mixing ratios of cloud water, cloud ice, rain, snow, and graupel as prognostic condensate categories.

2 Method

For the graupel particles, an exponential size distribution is assumed:

$$f_{\mathrm{g}}(D_{\mathrm{g}}) \;=\; N_0^{\mathrm{g}} \exp(-\lambda_{\mathrm{g}} D_{\mathrm{g}})$$

with $N_0^{\mathrm{g}} = 4 \times 10^6$ m^{-4} (Rutledge and Hobbs, 1984), D_{g}: diameter of graupel particle. The properties of single graupel particles in the form of power laws are taken from Heymsfield and Kajikawa (1986) for their (low density, $\rho_{\mathrm{g}} \approx 0.2\,\mathrm{g/cm^3}$) lump graupel: For the mass-size relation, it is assumed: $m_{\mathrm{g}} \;=\; a_m^{\mathrm{g}}\, D_{\mathrm{g}}^{3.1}$ with a_m^{g} =169.6; and for the terminal fall velocity depending on size: $v_T^{\mathrm{gp}}(D_{\mathrm{g}}) \;=\; v_0^{\mathrm{g}} D_{\mathrm{g}}^{0.89}$ with $v_0^{\mathrm{g}} = 442.0$ (all in the corresponding SI units).

Graupel is initiated from freezing of raindrops and from conversion of snow to graupel due to riming. The expression for the conversion rate for snow being converted to graupel due to riming follows from the consideration that a particle is converted from the snow category into the graupel category if the volume of the frozen ice from collected cloud water reaches a certain percentage (here: $\sim 12\%$) of the enveloping sphere associated with the snow particle's maximum diameter, see e.g. Seifert (2002). This process is active if a cloud water threshold of 0.2 g/kg is exceeded. Water vapor deposition, sublimation, melting, and collection of cloud droplets and cloud ice crystals is parameterized for graupel in a way analogous to snow. In contrast to the present scheme, for the (Kessler-type) autoconversion from cloud water to rain water, a cloud water threshold is applied (currently 0.2 g/kg). Figure 1 shows the microphysical processes considered in the parameterization scheme.

3 Results

Idealized 2-d Warm Bubble

Figure 2 compares the simulated hydrometeor distribution of a warm bubble, after 72 min, calculated with standard LM microphysics and graupel microphysics.

The graupel scheme simulates mostly graupel instead of snow, besides the upper part of the cloud. There is also more cloud water and more cloud ice in the simulation with graupel, due

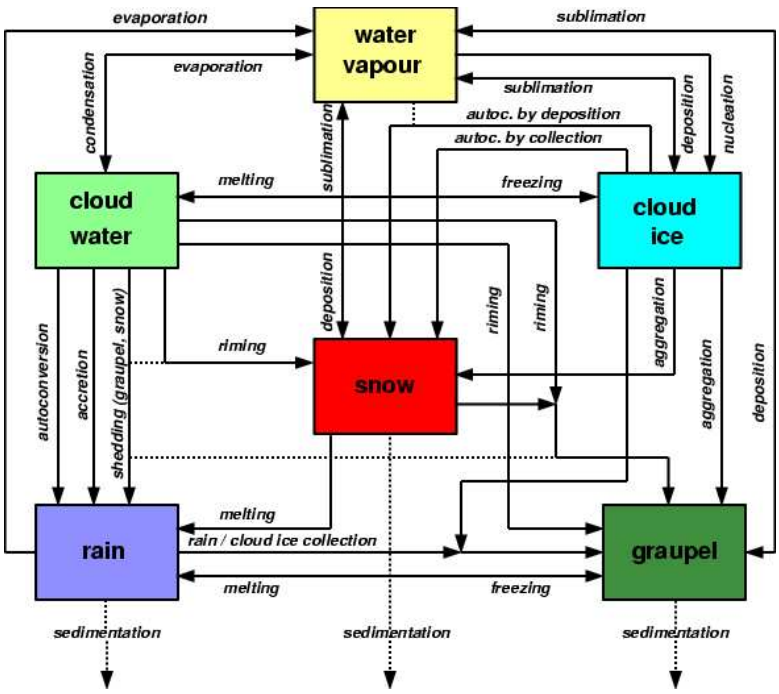


Figure 1: Cloud microphysical processes considered in the graupel scheme.

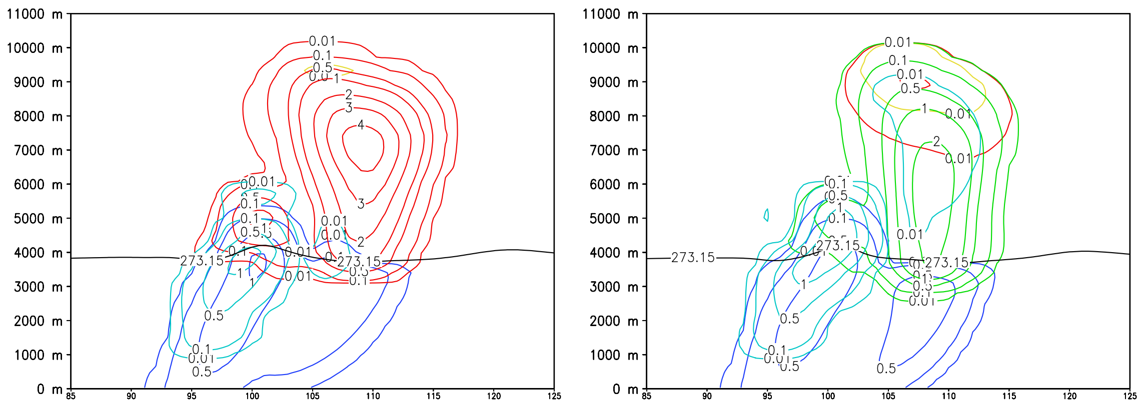


Figure 2: Hydrometeor distribution (mixing ratios in g/kg) of 2-d warm bubble, after 72 min. Mesh size: 2 km. Left: simulation with standard LM microphysics. Right: simulation with graupel scheme. Yellow: cloud ice, red: snow, green: graupel, light blue: cloud water, dark blue: rain.

to less efficient riming and deposition growth of graupel compared to snow, and due to the autoconversion threshold of cloud-water-to-rain autoconversion introduced in the graupel scheme. The zone with precipitating ice (snow, graupel, resp.) reaches further downward in the simulation including graupel, due to higher sedimentation velocity of graupel compared to snow. These features can be assessed as improvements.

Single Cases With LMK

Figure 3 shows west-east cross-sections of hydrometeor distributions for two LMK cases: A stratiform snowfall event from March 2004 and a spring/summer convective event from May 2004.

On the one hand, in the stratiform snowfall event most precipitation ice is simulated as snow, with about only 10 percent graupel. On the other hand, in the convective event, most

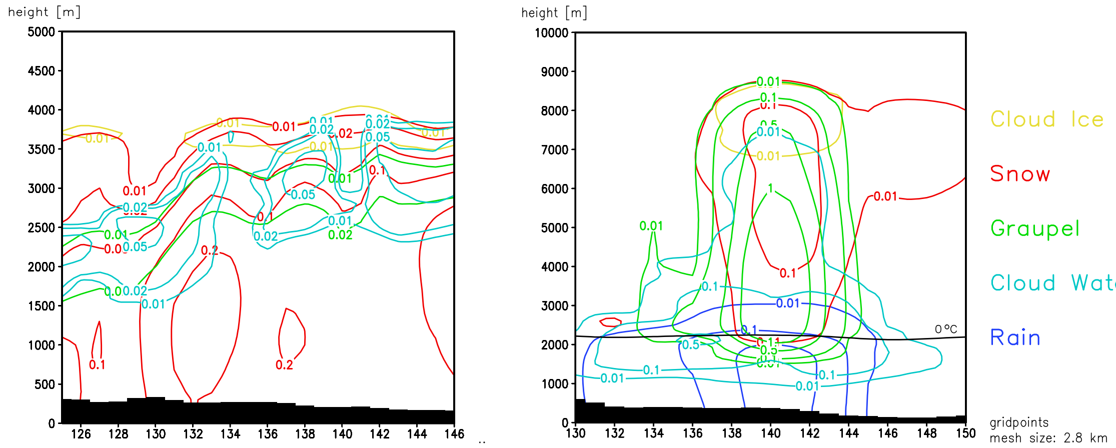


Figure 3: West-east cross-sections of hydrometeor distributions (mixing ratios in g/kg) for two cases simulated with LMK (mesh size: 2.8 km). Left: stratiform snowfall (2004-03-09 00 UTC + 08 h), isolines: 0.01, 0.02, 0.05, 0.1, 0.2. Right: convective cell (2004-05-11 00 UTC + 13 h), isolines: 0.01, 0.1, 0.5, 1. Color code as in Fig. 2.

precipitation ice is simulated as graupel, with snow occuring mostly in the upper part (and also in an anvil-like part) of the cloud. These seem to be reasonable results. Therefore, it can be concluded that the scheme simulates graupel principally in a plausible way.

Testsuite July to September 2004

A three-month comparison (two 18-h forecasts daily, starting 00 UTC and 12 UTC, for July to September 2004) of LMK results computed with the new scheme shows a small (5 %) decrease in total precipitation compared to the present microphysics scheme. Generally, standard verification scores (against synoptic observations) were not affected significantly. The positive frequency bias for small (0.1–2 mm/h) precipitation events was slightly reduced which might be caused by the introduction of the threshold for cloud water autoconversion. It can be concluded that the scheme behaves well also for a large series of forecasts, but improvements in forecast skill could not be found yet from the verification carried out up to now. The graupel scheme is now default for ongoing LMK testsuites.

4 Sensitivity to Graupel Particle Properties

Gilmore et al. (2004) carried out sensitivity tests with respect to the assumed properties of the graupel/hail category within a bulk (one-moment) microphysics parameterization. They used an idealized convective environment for their model setup (1 km mesh size, 30 m/s and 50 m/s wind speed with veering wind shear, supercell development, similar to Weisman and Klemp, 1984) and varied the intercept parameter N_0^g of the graupel particle size distribution and the graupel particle density ρ_g. Decreasing N_0^g as well as increasing ρ_g each changes the bulk properties of the particle ensemble towards more hail-like properties, e.g. faster sedimentation and less rapid melting. In general, more precipitation accumulated at ground was found in the cases with the graupel/hail category weighted towards large hail. These sensitivities could be found also with LMK simulations in a similar idealized 3-d convective setup (2.8 km mesh size, unidirectional wind shear only, wind speed 25 m/s, symmetric storm splitting, similar to Weisman and Klemp, 1982). Surface precipitation (both mean and maximum) tends to be higher with the graupel category weighted towards hail-like properties, i.e. smaller intercept parameter and larger particle density, see Table 1. A large sensitivity

N_0^{g}	ρ_{g}	TotP	TotG	MaxP	MaxG
4×10^4	≈ 0.2	36.17	0.1069	23.03	0.0001
4×10^5	≈ 0.2	27.55	0.0000	16.91	0.0000
4×10^6	≈ 0.2	14.80	0.0000	10.91	0.0000
4×10^4	0.4	35.51	0.1883	22.79	0.5017
4×10^5	0.4	32.02	0.0000	19.07	0.0000
4×10^6	0.4	25.79	0.0000	16.05	0.0000
4×10^4	0.9	32.82	3.4673	25.56	5.8234
4×10^5	0.9	35.01	0.0000	21.27	0.0000
no graupel	–	4.13	–	4.26	–

Table 1: Comparison of surface precipitation for simulations with different assumed intercept parameter N_0^{g} (in m^{-4}) and graupel particle density ρ_{g} (in $\mathrm{g/cm^3}$). Accumulated mass on ground (total precipitation (TotP) and graupel (TotG) in Tg and maximum total precipitation (MaxP) and maximum graupel precipitation (MaxG) in mm. All after 2 hours. For $\rho_{\mathrm{g}} = 0.4\,\mathrm{g/cm^3}$ and $\rho_{\mathrm{g}} = 0.9\,\mathrm{g/cm^3}$, the velocity-size relationship is taken from Lin et al. (1983).

to N_0^{g} is found in the $\rho_{\mathrm{g}} \approx 0.2\,\mathrm{g/cm^3}$ and $\rho_{\mathrm{g}} = 0.4\,\mathrm{g/cm^3}$ cases: With N_0^{g} decreasing from $4 \times 10^6\,\mathrm{m}^{-4}$ to $4 \times 10^4\,\mathrm{m}^{-4}$ total surface precipitation increases by 144 % and 73 %, resp. Higher mass-weighted sedimentation velocity of the graupel particle ensemble (i.e. smaller N_0^{g}) makes the particles less susceptible to horizontal advection (and subsequent evaporation outside the storm) and can therefore lead to more surface precipitation. As to be expected, with decreasing N_0^{g} and increasing ρ_{g} more unmelted graupel/hail can reach the ground. Much less surface precipitation (compared to all simulations including graupel) is found in the no-graupel (= standard LM microphysics) case confirming the need of a faster-than-snow falling ice species when simulating severe convection.

Less sensitivity is found in two simulations of real (convective) weather situations: A pre-frontal squall-line case (July 18, 2004) and a case with less organized, more isolated convection in a situation with weak large-scale gradients (August 7, 2004), see Tab. 2 and 3 and Fig. 4. In contrast to the idealized warm-bubble setup, in the August 07 case area-mean precipitation tends to decrease when moving from light-graupel to hail-like particle properties in the graupel/hail category, while in the July 18 case one might see the same but very much damped tendency as in the idealized setup. As in the idealized setup, in both real cases maximum precipitation is lower in the no-graupel simulation than in any of the simulations including graupel. From the precipitation patterns shown in Fig. 4 it can be inferred that in the August 07 case simulated precipitation becomes less widespread (i.e. areas receiving precipitation becoming smaller without maxima being reduced) when moving from the no-graupel over the low-density-graupel to the high-density-graupel/hail case which might be due to the effect of ice precipitation becoming less subject to horizontal advection when sedimenting faster. In the July 18 case (no figure shown), this feature is not seen.

That the sensitivity to the assumed properties of the graupel category is smaller in simulations of real convective cases compared to the idealized setup may be attributed (i) to more (negative) feedbacks being active in longer integration time and on a larger domain and (ii) to graupel being overall less important in simulations of real weather events (since there are always also more stratiform and snow-dominated areas) compared to the idealized simulations where much more graupel than snow is simulated. Tab. 2 and 3 show also that in simulations weighted towards large hail (all N_0^{g}=$4 \times 10^4\,\mathrm{m}^{-4}$ simulations; the more the higher ρ_{g}) explicit simulation of hail occurence at the ground is possible.

$N_0^{\rm g}$	$\rho_{\rm g}$	MeanP	MeanG	MaxP	MaxG
4×10^4	≈ 0.2	0.3627	0.0004	64.05	0.46
4×10^5	≈ 0.2	0.4461	0.0000	57.64	0.00
4×10^6	≈ 0.2	0.4548	0.0000	47.71	0.00
4×10^4	0.4	0.3136	0.0002	49.55	1.07
4×10^5	0.4	0.4023	0.0000	58.61	0.00
4×10^6	0.4	0.4846	0.0000	57.92	0.00
4×10^4	0.9	0.3109	0.0061	73.71	10.15
4×10^5	0.9	0.3129	0.0000	56.61	0.05
no graupel	–	0.4190	–	41.96	–

Table 2: As Tab. 1, but for simulated 23-hour precipitation sum of LMK forecasts started at August 07, 2004 00 UTC. Numbers are valid for the area shown in Fig. 4 (total domain larger than domain shown). MeanP and MeanG stand for mean total precipitation and mean graupel precipitation (in mm), resp.

$N_0^{\rm g}$	$\rho_{\rm g}$	MeanP	MeanG ($\times 10^3$)	MaxP	MaxG
4×10^4	≈ 0.2	4.384	0.222	91.37	1.65
4×10^5	≈ 0.2	4.318	0.0	94.46	0.0
4×10^6	≈ 0.2	4.183	0.0	81.20	0.0
4×10^4	0.4	4.369	2.098	98.34	3.15
4×10^5	0.4	4.341	0.0	83.31	0.0
4×10^6	0.4	4.341	0.0	83.71	0.0
4×10^4	0.9	4.276	20.906	98.39	9.60
4×10^5	0.9	4.334	0.047	86.73	0.0
no graupel	–	4.154	–	79.41	–

Table 3: As Tab. 2, but for simulated 23-hour precipitation sum of LMK forecasts started 2004-07-24 00 UTC.

5 Outlook

It is under consideration to change the bulk properties of the graupel category in such way that more hail-like particles instead of low-density graupel particles are represented. This would allow for an explicit simulation of surface hail occurence. Then it would be more consistent to take into account also wet growth of the hailstones which is neglected currently. On the other hand, medium- and low-density graupel would then be represented less accurately. A compromise might be to make $N_0^{\rm g}$ dependent on the graupel/hail mixing ratio, i.e. let $N_0^{\rm g}$ decrease when $q_{\rm g}$ increases.

References

Doms, G. and J. Förstner, 2004: Development of a kilometer-scale NWP-system: LMK. *COSMO Newsletter*, No. 4, pp. 159–167.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Mon. Wea. Rev.*, 132, pp. 2610–2627.

Heymsfield, A. J. and M. Kajikawa, 1986: An improved approach to calculating terminal velocities of plate-like crystals and graupel. *J. Atmos. Sci.*, 44, pp. 1088–1099.

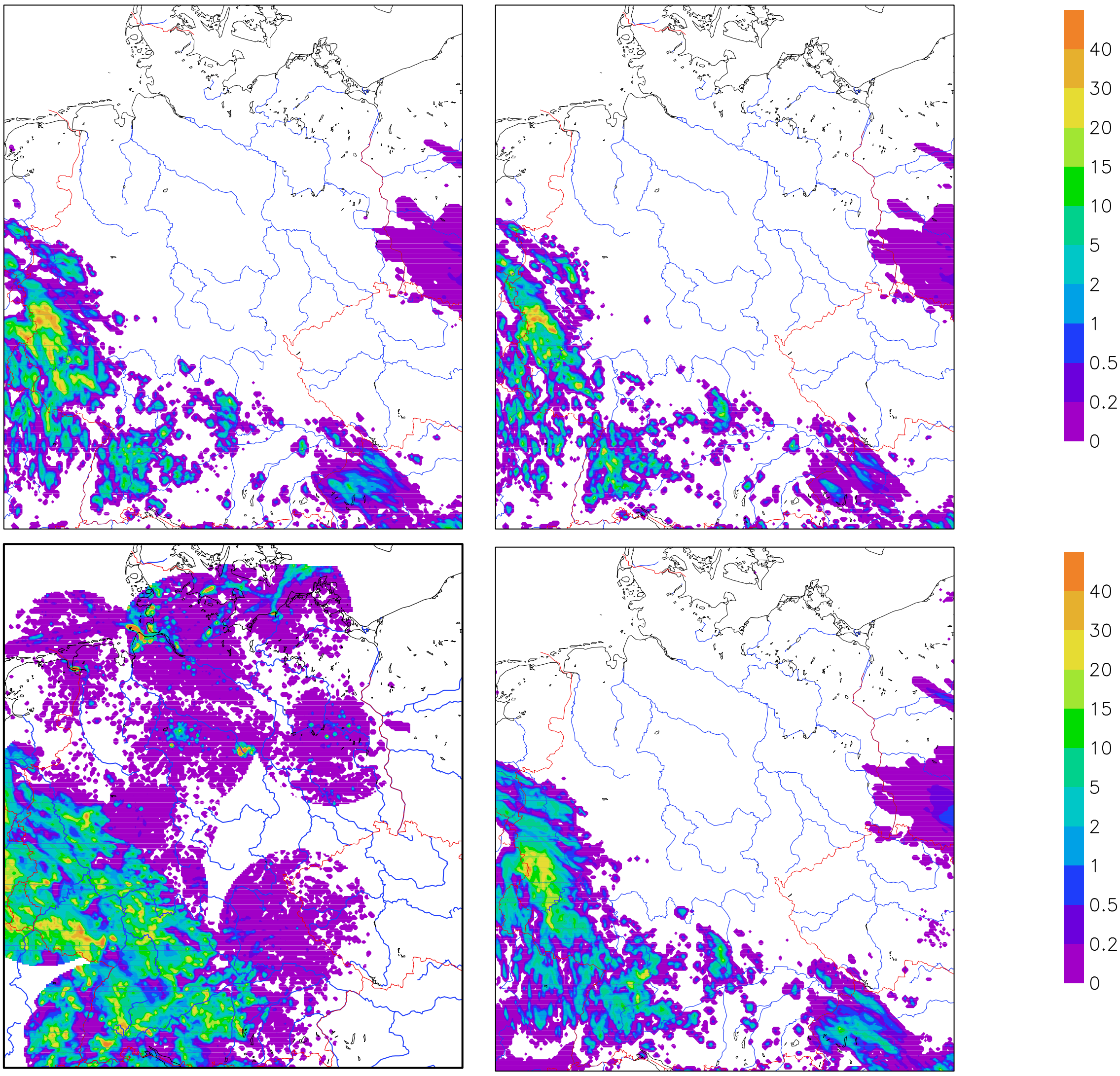


Figure 4: Simulated 23-hour precipitation sums (in mm) of LMK forecasts started 2004-08-07 00 UTC. Upper left: N_0^{g}=4 × $10^6\,\mathrm{m}^{-4}$, $\rho_{\mathrm{g}} \approx 0.2\,\mathrm{g/cm}^3$; upper right: N_0^{g}=4 × $10^4\,\mathrm{m}^{-4}$, $\rho_{\mathrm{g}} = 0.9\,\mathrm{g/cm}^3$; lower right: no graupel (standard LM microphysics). Lower left: corresponding radar-derived precipitation observation.

Lin, Y.-L., R. D. Farley and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, 22, pp. 1065–1092.

Rutledge, S. A. and P. V. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: Diagnostic modelling study of precipitation development in narrow cold-frontal rainbands. *J. Atmos. Sci.*, 41, pp. 2949–2972.

Seifert, A., 2002: *Parametrisierung wolkenmikrophysikalischer Prozesse und Simulation konvektiver Mischwolken*, Dissertation, Department of Physics, Universität Karlsruhe, 232 pp.

Weisman, M. L. and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, 110, pp. 504–520.

Weisman, M. L. and J. B. Klemp, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, 112, pp. 2479–2498.

Simulation Studies of Shallow Convection with the Convection-Resolving Version of the DWD Lokal-Modell

F. Theunert[1], A. Seifert[2]

[1] *Amt für Geo-Informationswesen der Bundeswehr, Traben-Trarbach, Germany*
[2] *Deutscher Wetterdienst, Offenbach, Germany*

1 Introduction

The next generation of mesoscale NWP models will - at least to some extent - resolve deep convection, e.g. squall lines. Most of the modeling systems currently under development aim at a grid resolution of 2-4 km. At DWD the first convection-resolving version of LM, called Lokal-Modell-Kürzestfrist[2] (LMK), will have a horizontal resolution of approximately 2.8 km. Other examples of convection-resolving NWP models are the WRF model, which was successfully applied at 4 km resolution to resolve squall lines and mesoscale convective systems in the continental U.S., the AROME project of MeteoFrance and the high-resolution version of the UK MetOffice Unified Model.

Although all these models may describe deep moist convection explicitly by the model equation system, shallow convection can currently only be considered as a sub-scale process. Therefore the important impact of shallow convection on the vertical transport of energy and water vapor has to be included by applying a special parameterization scheme. For LMK a simple shallow convection scheme based on the cumulus parameterization of Tiedtke (1989) was suggested by Doms and Förstner (2004).

In the following we will give an overview of the Tiedtke-Doms shallow convection scheme, present some case studies and verification as well as some ideas for potential improvements.

2 The Tiedtke-Doms shallow convection scheme

The Tiedtke scheme, which is operationally applied in the 7 km version of LM to parameterize cumulus convection, distinguishes between 3 cloud types: deep convection, mid-level convection and shallow convection. It is therefore rather straightforward to reduce the scheme to shallow convection only as suggested by Doms and Förstner (2004). Here we will just summarize the basic assumptions of this approach:

1. The momentum fluxes are neglected, only temperature and moisture are affected directly by shallow convection.
2. Shallow convection is non-precipitating, i.e. rain formation is neglected completely and no evaporation of rain below cloud base occurs.
3. Shallow convection does not induce convective downdrafts.
4. The moisture convergence mass flux closure is applied (Eq. 19 of Tiedtke, 1989)
5. Organized entrainment is neglected. For turbulent entrainment/detrainment $\epsilon_u = \delta_u = 3 \times 10^{-4}\ \mathrm{m}^{-1}$ is used as in Tiedtke (1989).

[2] Kürzestfrist (German) = shortest-range

a) without shallow conv. scheme b) with shallow conv. scheme

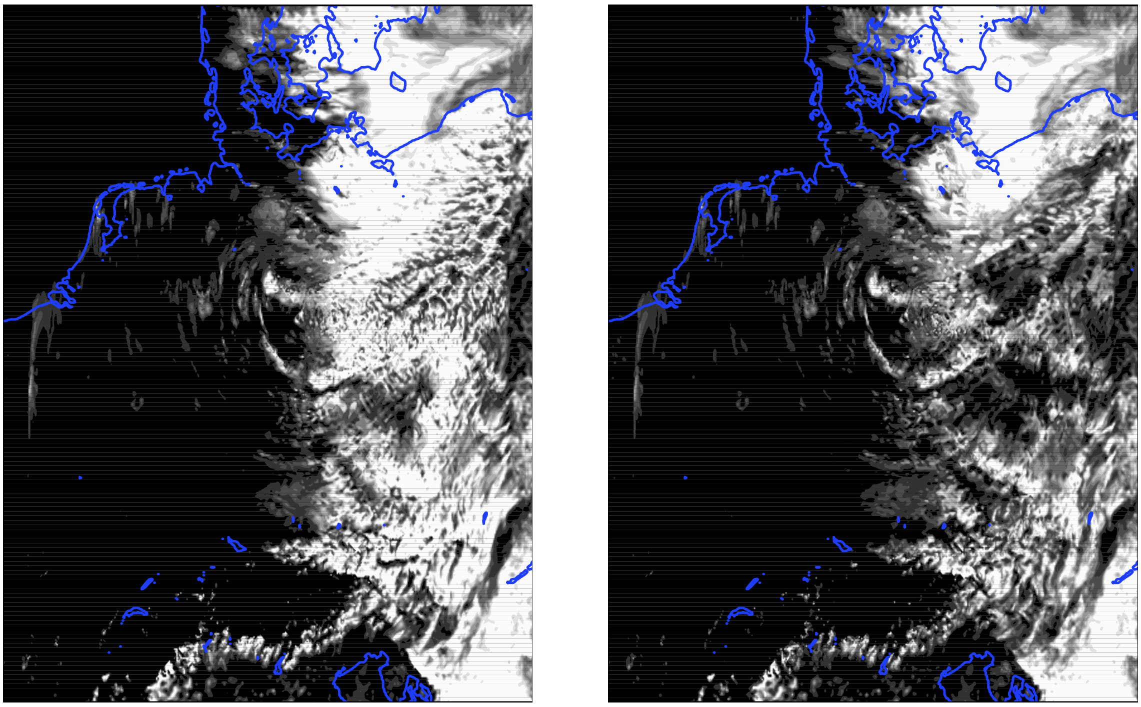


Figure 1: Simulated low-level cloud cover 29.05.2004 at 12 UTC in units of 1/8

6. Exactly the same triggering and parcel ascent calculation is used as in the full Tiedtke scheme. As in the operational Tiedtke scheme of LM horizontally averaged values of the vertical velocity and the moisture convergence are used.

7. Shallow convection is limited to a cloud depth of 250 hPa. For deeper clouds the scheme is simply turned off.

The last assumption replaces the moisture convergence threshold which distinguishes shallow convection from deep convection in the original Tiedtke scheme.

The scheme is implemented in LM 3.16 and can be turned on by setting `lconv=.true.`, `ltiedtke=.false.`, `lshallow=.true.`.

3 Case study 29 May 2004

This was a typical summertime high-pressure situation with a surface high centered over northern Germany. During daytime shallow convection developed especially over eastern Germany. This day was chosen to be able to use measurements of the Lindenberg observatory located about 100 km southeast of Berlin.

Fig. 1 shows the low-level cloud cover at 12 UTC. Compared to the 27 Feb. 2004 case shown by Doms and Förstner (2004), the reduction in cloud cover due to the shallow convection scheme is weaker, but still significant. In the simulation without shallow convection scheme the model tries to represent the cloud-topped convective boundary layer explicitly leading to an overestimation of cloud cover (note that grid-scale clouds are always counted as 100% cloud cover). Using the shallow convection scheme, the small cumuli are described as being of sub-grid scale and the associated cloud cover is drastically reduced. This can lead to a significant change in the radiation budget and the 2m-temperature. Without the shallow

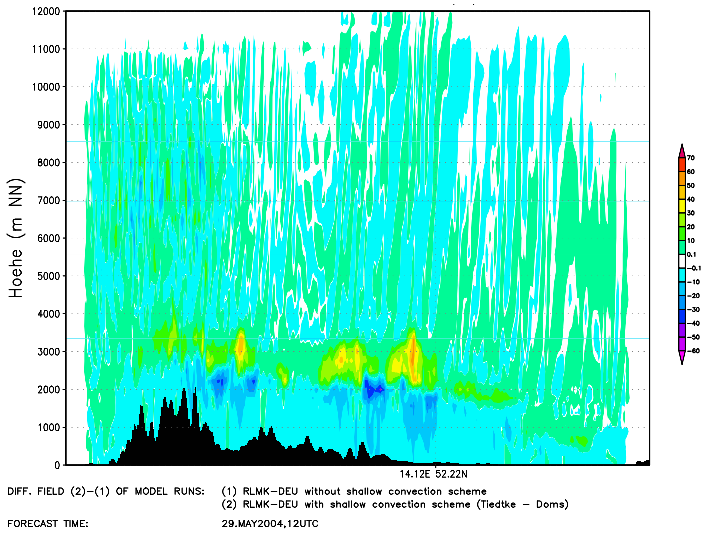


Figure 2: Vertical cross section of the difference in rel. humidity with and without shallow conv. scheme, RH_{with} - $RH_{w/o}$, in %. The cross section is oriented South-North at 14.12^{o}E.

convection parameterization the model has also deficiencies in representing the moisture fluxes due to the vertical motions within the convective boundary layer. Fig. 2 shows a vertical cross section of the difference in relative humidity between the simulation with and without shallow convection scheme. Obviously, the parameterization provides an efficient vertical transport of moisture out of the boundary layer. The rel. humidity in the PBL is reduced by about 10 % or more and the moisture is detrained in the free atmosphere. A more detailed evaluation of this process is possible by comparison with measured vertical profiles at the Lindenberg observatory located at 14.12^{o}E 52.22^{o}N.

Fig. 3 shows profiles of temperature, rel. humidity and water vapor mixing ratio for 12 UTC and 18 UTC. While the temperature profiles of both simulations matches the observations very well including the location of the PBL height, the rel. humidity and vapor mixing ratio within the PBL are overestimated, especially by the simulation without shallow convection scheme. Applying the Tiedtke-Doms scheme results in a significant reduction of the rel. humidity with the maximum being reduced from 100 %, i.e. grid-scale cloud, to 85 %. Although the rel. hum. within the PBL is still overestimated, as is the vapor mixing ratio, this seems to be an improvement compared to the simulation without shallow convection scheme. The moisture is deposited above the PBL in a layer between 2-4 km AGL making this layer much moister than observed in this case. For 18 UTC the observations show an increase in temperature and PBL height as well as a decrease of the moisture within the PBL compared to 12 UTC, both features are not reproduced by either one of the simulations. Obviously the model has some deficiencies here which may be related to treatment of soil moisture or the turbulence scheme. Since these problems are larger than the impact of the shallow convection scheme, the question arises whether the improvement by using the shallow convection scheme, also at 12 UTC, is maybe due to wrong reasons, i.e. that another process might be causing the overestimation of PBL moisture.

4 Case study 14 May 2004

The 14 May 2004 was another day with a typical summertime high-pressure situation and shallow convection was observed during the afternoon and evening over most of Germany.

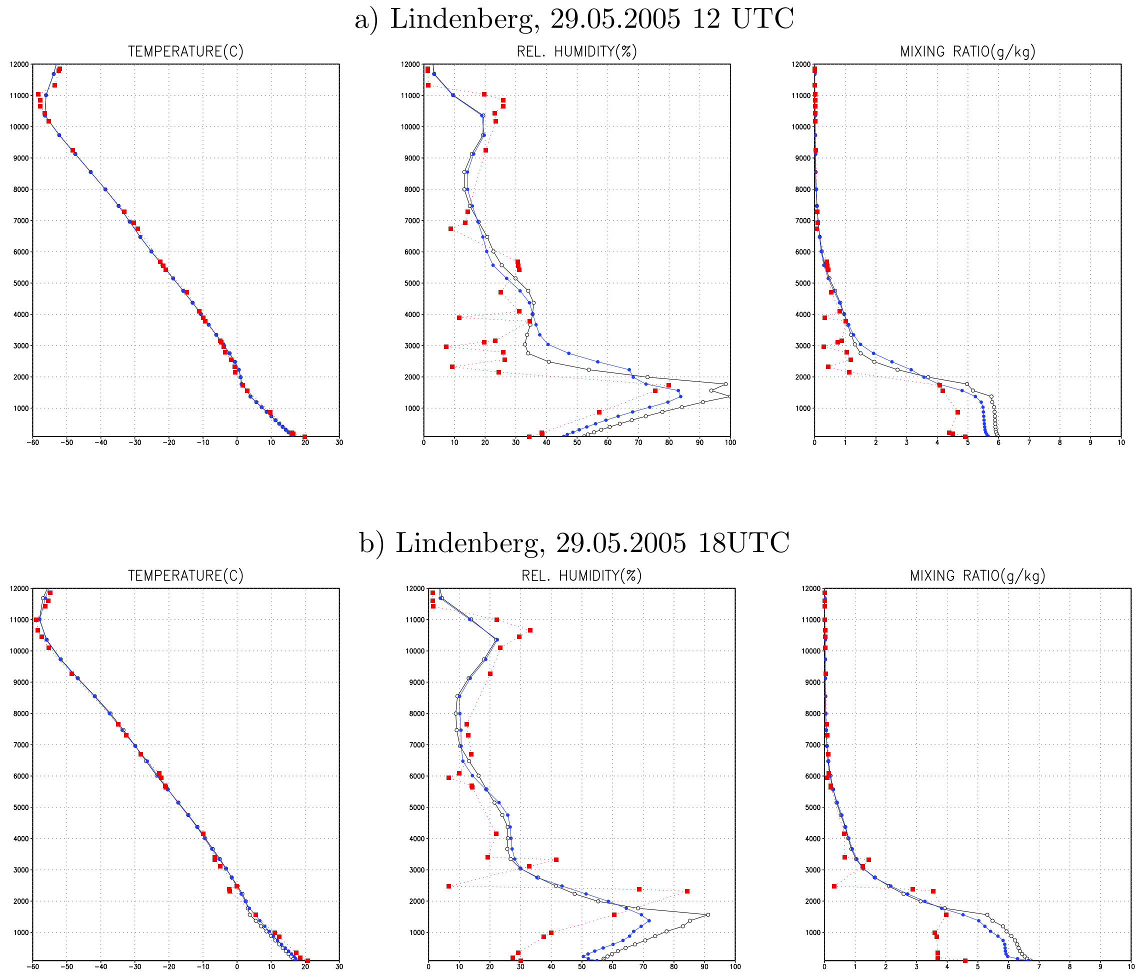


Figure 3: Vertical profiles of temperature, rel. humidity and water vapor mixing ratio at Lindenberg. Radiosonde measurements (red), LMK w/o shallow convection (black) and with shallow convection (blue)

Fig. 4 shows low-level cloud-cover at 18 UTC. In this case the difference between the two simulations is even more pronounced, without shallow convection the model predicts overcast conditions at 18 UTC while the simulation using the parameterization shows a greatly reduced cloud cover. The boundary problems without shallow convection scheme, especially in the SW-corner, result from the difference compared to the driving 7 km model that uses the full Tiedtke scheme and predicts a cloud cover quite similar to the high-resolution run with shallow convection scheme (see Doms and Förstner (2004) for a comparison of the 7 km vs 2.8 km model).

Fig. 5 shows the vertical profiles at Lindenberg. In this case the temperature profile is predicted reasonably well by both simulations, although the inversion at 2000 m AGL is more pronounced in the observations. The profiles of rel. humidity and vapor mixing ratio show that, in this case and at this specific grid point, the vertical transport of moisture by the shallow convection scheme leads to a growth of the PBL itself. By deposition of moisture right on top of the PBL the shallow convection increases the PBL height and also the maximum rel. humidity matching the observation much better than in the simulation without shallow convection scheme. Note that for this grid point the simulation with shallow convection predicts a higher cloud cover compared to the simulation without the parameterization.

Overall this case shows nicely that the LMK shallow convection scheme is necessary to ensure

a) without shallow conv. scheme

b) with shallow conv. scheme

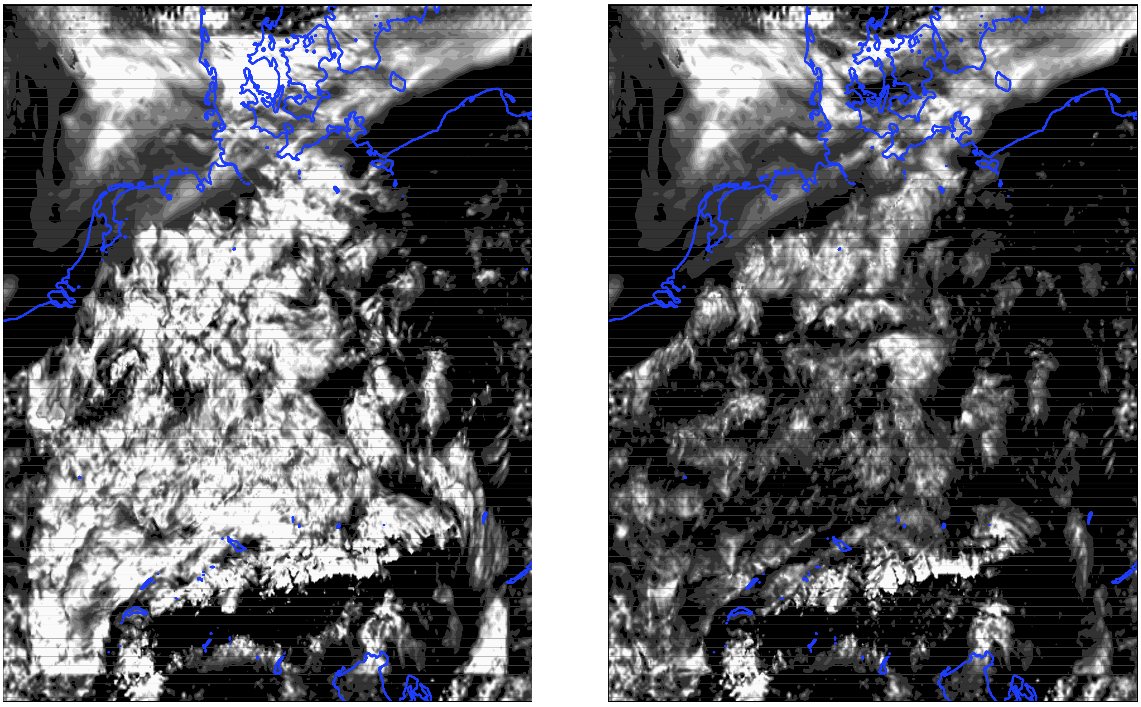


Figure 4: Simulated low-level cloud cover 14.05.2004 at 18 UTC in units of 1/8

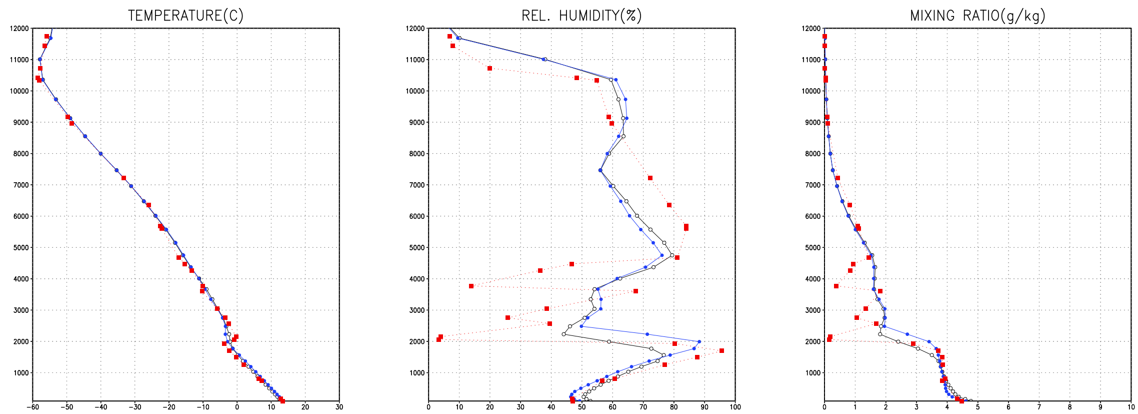


Figure 5: Vertical profiles of temperature, rel. humidity and water vapor mixing ratio at Lindenberg on 14 May 2004 18 UTC. Radiosonde measurements (red), LMK w/o shallow convection (black) and with shallow convection (blue).

consistency with the driving model and that the parameterized vertical transport of moisture may cause an increase of the PBL height at some grid points. Although the interaction of the shallow convection scheme and the PBL scheme has to be investigated more thoroughly, this PBL growth by shallow convection looks quite reasonable and might also be able to remove inversion layers in some cases, triggering resolved deep convection.

5 The w_*-closure

As already mentioned by Tiedtke (1989) and many others, shallow convection is mostly controlled by the sub-cloud layer turbulence. Therefore it is quite plausible to use a mass flux closure which is solely based on PBL turbulence parameters instead of the moisture convergence closure. In a high resolution NWP model this would implicitly assume that mesoscale convergence zones always trigger (resolved) deep convection.

For example, Neggers et al. (2004) suggest to use a mass flux closure proportional to the free convective velocity scale w_*:

$$M_u = a\,w_* = a\left(\frac{g z_b}{\theta_v^0}\,\overline{(w'\theta_v')_s}\right)^{\frac{1}{3}} \tag{1}$$

Here M_u ist the cloud base (specific) mass flux, z_b the PBL height and θ the potential temperature, i.e. $\overline{(w'\theta_v')_s}$ is the surface heat flux. Within the shallow convection scheme we assumed the z_b equals the cloud base height estimated by the parcel ascent.

Following Grant (2001) one may simply set $a = 0.03$ and this will be used in the following simulations, but note that Neggers et al. (2004) argue that a is in fact a function of the cloud fraction and should be parameterized using a statistical cloud scheme.

To compare the moisture convergence closure and the w_*-closure we have chosen the 20 May 2004. Fig. 6 shows low-level cloud cover at 6 UTC and 12 UTC. In the early morning hours we only see some differences over the Baltic sea where shallow convection might have been active in reality, too. Later on, the differences become larger and for this case the w_*-closure is even more efficient in reducing the low-level cloud cover. Fig. 7 shows time-height cross sections of measured radar reflectivity and simulated cloud cover at Lindenberg. The observations, which include also the estimated cloud base by a ceilometer, show an almost ideal development of a cloud-topped free convective boundary layer. Note that even a cloud radar cannot detect small, optically thin cumuli which were probably present during morning and noon, later on the clouds became thicker and even drizzle was formed in the late evening. The simulation using the moisture convergence closure reproduces the development of the PBL height very nicely and the amount of cloud cover, although not directly comparable to radar reflectivity, looks like a reasonable representation of the atmospheric conditions. Applying the w_*-closure leads to a more rapid development of the PBL height in the morning hours and more low-level cloud cover is observed earlier during the day. From this point of view, the moisture convergence closure seems to match the conditions somewhat better, although we have to keep in mind that this is only a single grid point.

6 Conclusions

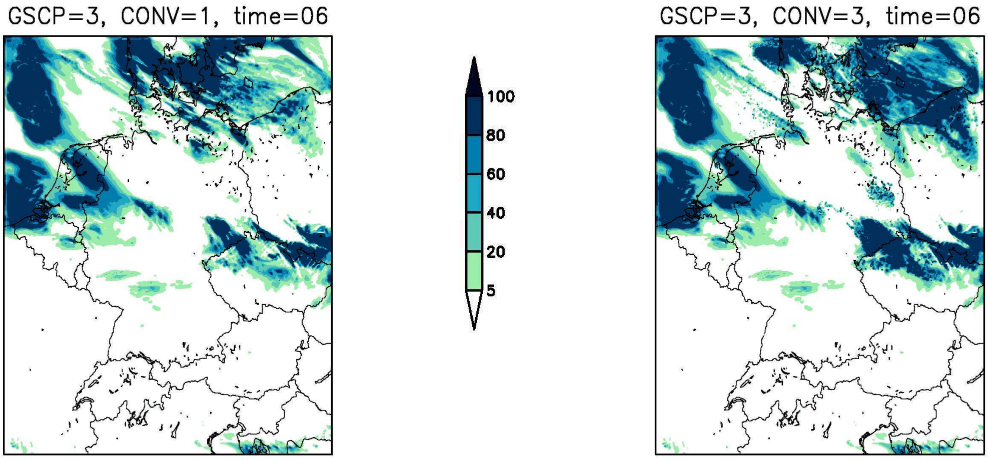
Our sensitivity studies showed that using a simple shallow convection parameterization within the convection-resolving LMK improves directly the forecasts of low level cloud cover. Basically, this confirms the earlier results presented by Doms and Förstner (2004). The shallow convection is also necessary to ensure consistency with the driving larger-scale model which, at least at DWD, uses the Tiedtke parameterization.

The w_*-closure, although interesting, did not show any advantages in our case studies. Future work in this area will probably focus on the development of a generalized PBL scheme which includes shallow convection.

Further testing is needed for cases with (resolved) deep convection. Since the shallow convection scheme reduces the PBL moisture, it also reduces the convective available potential energy CAPE. Therefore we expect the deep convection to be slightly weaker in simulations using the shallow convection scheme.

a) q_v-convergence closure, 6 UTC

b) w_*-closure, 6 UTC



c) q_v-convergence closure, 12 UTC

d) w_*-closure, 12 UTC

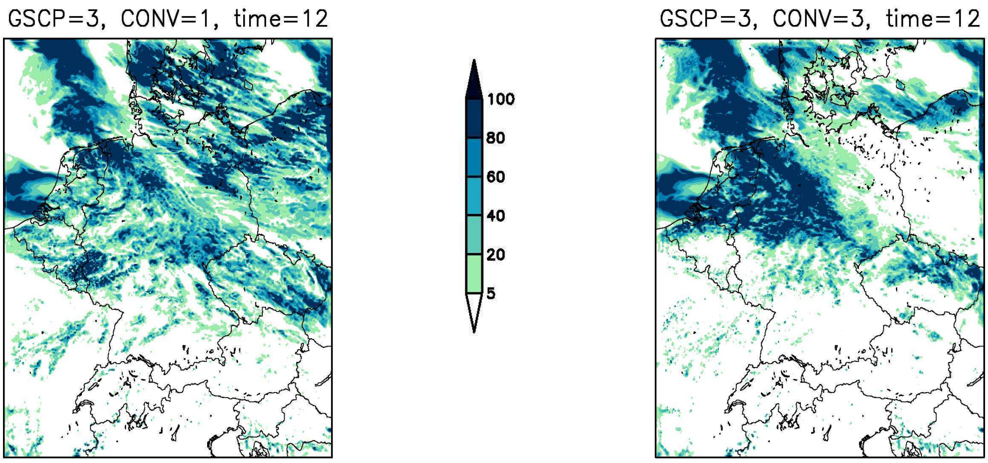


Figure 6: Simulated low-level cloud cover on 20.05.2004 in % (Note that the colors are reversed compared to Fig. 1)

References

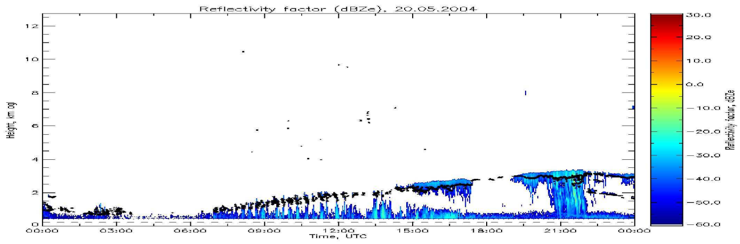
Doms G. and J. Förstner, 2004: Development of a kilometer-scale NWP System: LMK. *COSMO Newsletter*, No. 4, pp. 159-167.

Grant, A., 2001: Cloud-base fluxes in the cumulus-capped boundary layer. *Quart. J. Roy. Met. Soc.*, 126, pp. 1913-1936.

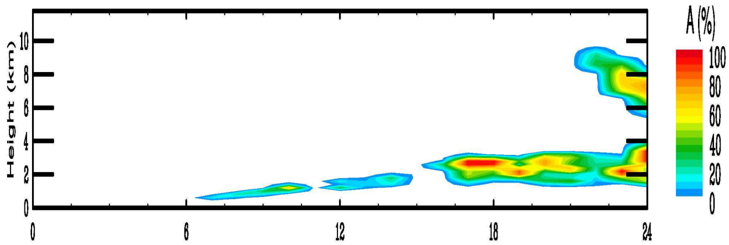
Neggers, R., A. Siebesma, G. Lenderik and A. Holtslag, 2004: An Evaluation of Mass Flux Closures for Diurnal Cycles of Shallow Convection. *Mon. Wea. Rev.*, 132, pp. 2525-2538.

Tiedtke, M., 1989: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. *Mon. Wea. Rev.*, 117, pp. 1779-1800.

a) radar reflectivity and ceilometer cloud base



b) cloud fraction, q_v-convergence closure



c) cloud fraction, w_* closure

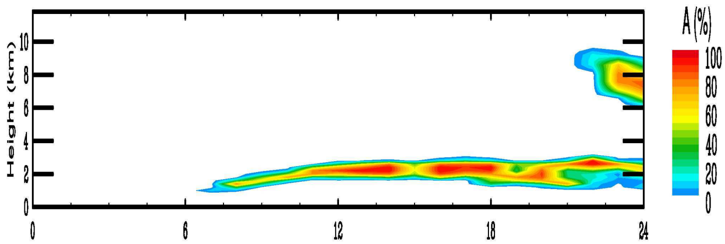


Figure 7: Time-Height cross sections of (a) observed radar reflectivity in dBZ (35 GHz cloud radar) and ceilometer cloud base height (black dots) as well as (b,c) simulated cloud cover in % at Lindenberg.

The COSMO-LEPS Suite at ECMWF: Present Status and Developments

A. Montani, C. Marsigli, T. Paccagnella

ARPA–SIM, Bologna, Italy

1 Introduction

COSMO–LEPS is the Limited–area Ensemble Prediction System that has been developed within the COSMO consortium. It aims to improve the short–to–medium range predictability of localised and intense weather events (e.g. heavy rainfall, intense winds, strong temperature anomalies). Its present set–up comprises 10 Lokal–Modell integrations, nested on selected members of ECMWF EPS global ensemble. The main features of the COSMO–LEPS system are summarised in Fig. 1 and are described in greater details in Montani et al. (2003a and 2003b) and in Marsigli et al. (2005).

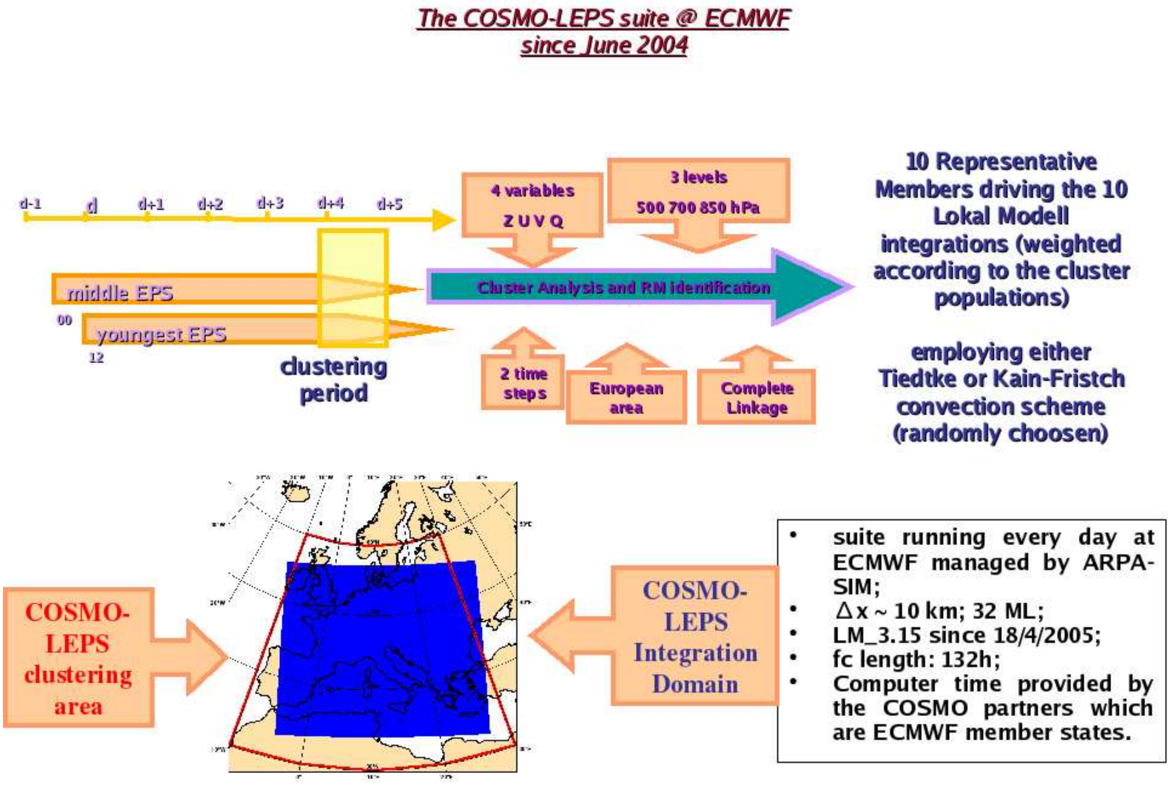


Figure 1: Main features of COSMO–LEPS application running at ECMWF.

More specifically, the products generated by the operational set–up can be summarised as follows:

- core products: ten perturbed LM runs (taking initial and 3–hourly boundary conditions from 10 selected EPS members) to generate probabilistic output (start at 12UTC; forecast length: 132 hours);

- additional products:
  - one *deterministic* run (taking initial and 6–hourly boundary conditions from the high–resolution deterministic ECMWF forecast) to assess the relative merits between deterministic and probabilistic approach (start at 12UTC; forecast length: 132 hours);
  - one *proxy* run (taking initial and 3–hourly boundary conditions from ECMWF analyses) to "downscale" ECMWF information (start at 00UTC; forecast length: 36 hours)

2 Dissemination

The products generated by the operational COSMO–LEPS suite are disseminated to the National and Regional Weather Services of COSMO and, for a number of case studies, also to Hungary. Fig. 2 illustrates some of the products which are sent on a daily basis to the

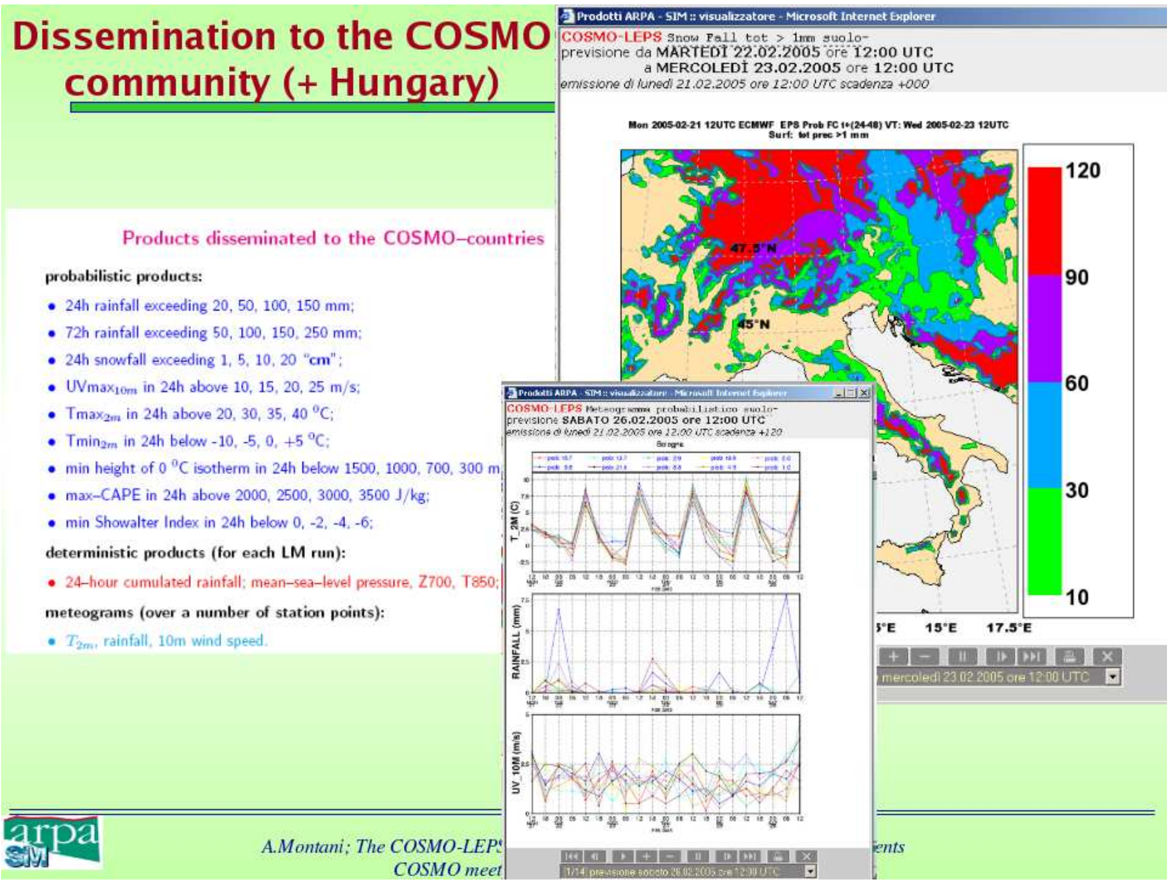


Figure 2: Examples of the COSMO–LEPS products disseminated to the COSMO community.

weather services, including probabilistic output (probability of exceeding of a threshold for a certain variable), deterministic output (the fields predicted by each individual COSMO-LEPS integration) and meteograms over station points (in terms of 2–metre temperature, 10–metre wind speed and rainfall).

As for the timing of delivery (see Fig. 3), COSMO-LEPS application is triggered on ECMWF supercomputers at about 20.30 UTC of day *D* and output files are disseminated to the operational weather services already after 5 hours, at about 1 UTC of day *D+1*. Since forecasts have a range of 132 hours (that is up to the end of day *D+5*), COSMO–LEPS products turn out to have a long range of validity and utility for forecasters.

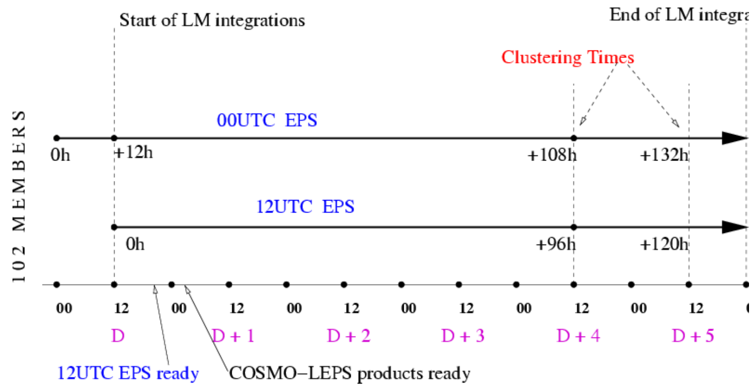


Figure 3: Timing of dissemination of COSMO–LEPS products.

3 Recent changes

Archive at ECMWF

During 2005, the archiving of COSMO–LEPS products on ECMWF's Meteorological Archival and Retrieval System (MARS) was implemented. This important step, achieved thanks to the assistance and collaboration of ECMWF staff, will make the retrieval and use of COSMO–LEPS products simpler and more user–friendly.

From 1 July 2005, the following products are archived on MARS and can be retrieved ("class=co" needs to be specified so as to obtain them):

- deterministic forecast (from fc+0h to fc+132 every 3h):
  - pressure level: geopotential height, relative humidity and temperature at 500, 700 and 850 hPa;
  - surface: albedo, low cloud–cover, medium cloud–cover, total cloud–cover, short–wave radiation flux, CAPE, height of 0^oC isotherm, height of snow–fall limit, mean–sea–level pressure, 2–metre temperature, 2–metre dew–point temperature, minimum of 2–metre temperature, maximum of 2–metre temperature, zonal 10–metre wind, meridional 10–metre wind, maximum of 10–metre wind speed, large-scale rainfall, convective rain, large-scale snowfall, total precipitation.
- ensemble prediction system
  - 10 perturbed forecasts (from fc+0h to fc+132 every 3h):
    - pressure level: geopotential height, relative humidity and temperature at 500, 700 and 850 hPa;
    - surface: albedo, low cloud–cover, medium cloud–cover, total cloud–cover, short–wave radiation flux, CAPE, height of 0^oC isotherm, height of snow–fall limit, mean–sea–level pressure, 2–metre temperature, 2–metre dew–point temperature, minimum of 2–metre temperature, maximum of 2–metre temperature, zonal 10–metre wind, meridional 10–metre wind, maximum of 10–metre wind speed, large-scale rainfall, convective rain, large-scale snowfall, total precipitation.

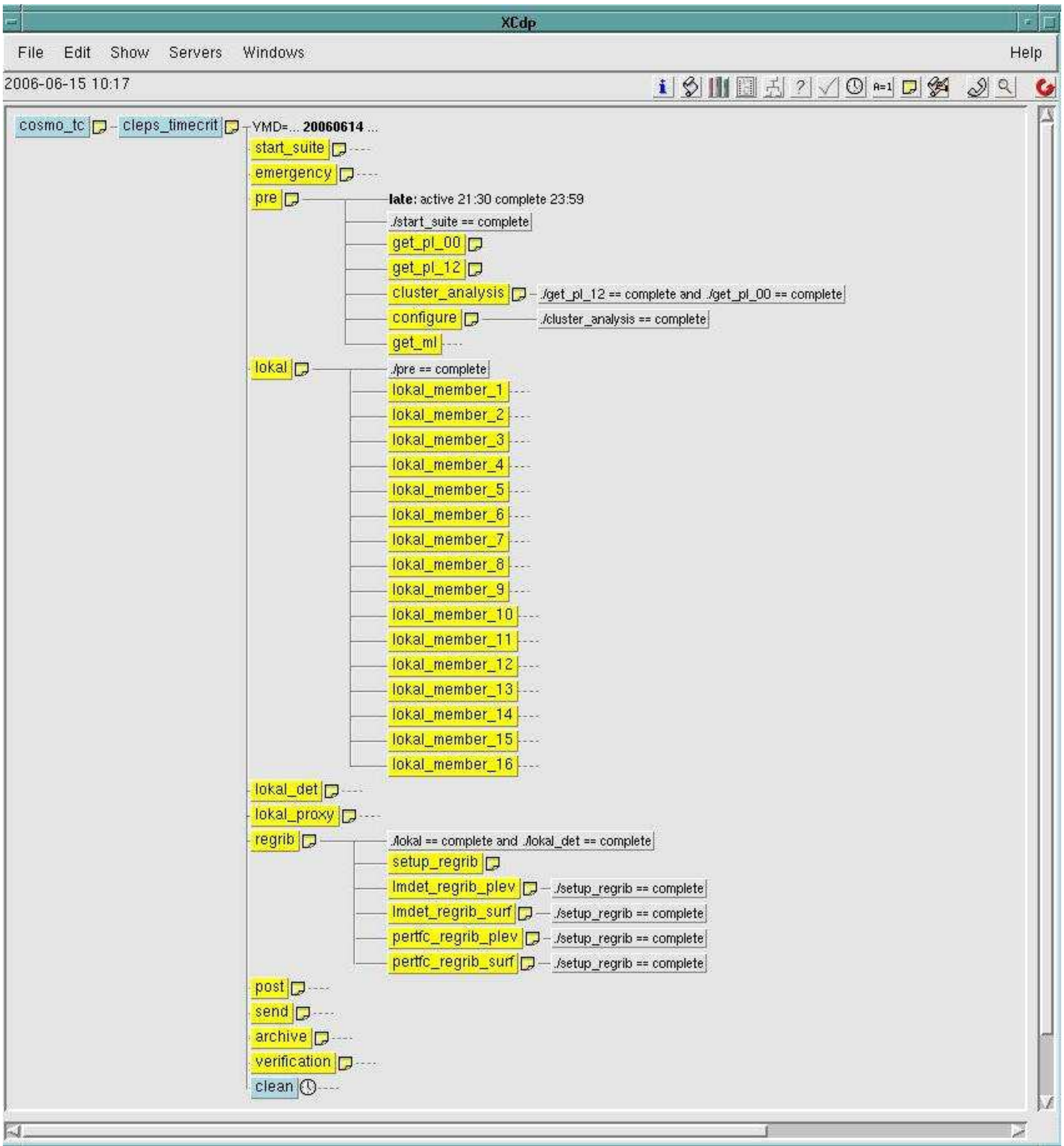


Figure 4: Configuration of COSMO–LEPS time–critical suite at ECMWF (since February 2006 with 16 members).

- Forecast probability (various intervals and thresholds):
  - surface: CAPE, height of 0^{o}C isotherm, minimum of 2–metre temperature, maximum of 2–metre temperature, maximum of 10–metre wind speed, total precipitation, total snowfall, minimum of showalter index.
- Clustering information (population, clustering variables and intervals, . . .).

In the present configuration, about 2.2 GB/day are produced by COSMO–LEPS and archived on MARS.

COSMO–LEPS as a time–critical application

At the end of November 2005, COSMO–LEPS has become a "Member–state time–critical application" at ECMWF. This implies that COSMO–LEPS jobs are given higher priority on ECMWF supercomputers and dedicated file systems are used so as to speed up the application. In addition to this, ECMWF operators (on duty 365 days a year) monitor the suite (shown in Fig. 4) and can take actions in case of problems, if the intervention from ARPA–SIM is not possible. For this purposes, a number of man–pages were developed so as to instruct operators about the corrective actions to be taken in case of failures and COSMO–LEPS application had to be re–organised so as to follow a number of ECMWF requirements.

The implementation of COSMO–LEPS time–critical application has insured in the last months a faster response in case of problems and a safer delivery of products.

4 Future plans

In the near future, the following activities are planned:

- produce and archive on MARS a number of so–called "derived probability products", like the ensemble mean and the ensemble standard deviation, so as to assess the features of COSMO–LEPS system in terms of spread–skill correlation;
- (beginning of 2006), increase the population of COSMO–LEPS system from 10 to 16 members and to increase the vertical resolution from 32 to 40 levels.
- produce and disseminate synthetic satellite and radar pictures in ensemble mode on an operational basis.

References

Marsigli, C., Boccanera, F., Montani, A. and Paccagnella, T., 2005. The COSMO–LEPS ensemble system: validation of the methodology and verification. *Nonlin. Proc. Geophys.*, **12**, pp. 527–536.

Montani, A., Marsigli, C., Nerozzi, F., Paccagnella, T., Tibaldi, S. and Buizza R., 2003a. The Soverato flood in Southern Italy: performance of global and limited–area ensemble forecasts *Nonlin. Proc. Geophys.*, **10**, pp. 261–274.

Montani, A., Capaldo, M., Cesari, D., Marsigli, C., Modigliani, U., Nerozzi, F., Paccagnella, T., Patruno, P. and Tibaldi, S., 2003b. Operational limited–area ensemble forecasts based on the Lokal Modell. *ECMWF Newsletter Summer 2003*, **98**, pp. 2–7.

Verification of the COSMO–LEPS new Suite in Terms of Precipitation Distribution

C. Marsigli, A. Montani, T. Paccagnella

ARPA–SIM, Bologna, Italy

1 Introduction

An objective verification of the mesoscale ensemble system COSMO–LEPS is carried out. This is a 10 member ensemble based on the Limited-Area Model *Lokal Modell* (LM), running daily at ECMWF since November 2002 using the resources shared by the COSMO countries. The system is based on the ECMWF EPS, which provides the perturbations to the limited-area model runs through the initial and boundary conditions. Mesoscale perturbations are also added by letting the different LM run to randomly choose the scheme to be used for the parameterisation of the deep convection (Tiedtke or Kain–Fritsch).
The system has recently been updated to the current configuration (June 2004) and the verification of the new suite is presented here.
The COSMO–LEPS system transfers the EPS probabilistic approach to the scales where a better representation of mesoscale–related processes permits to forecast surface parameters with a greater detail. Then, the verification of the system focuses on features where the impact of the high–resolution is dominant. Precipitation is chosen as the verification parameter, trying to evaluate the ability of the system in forecasting the detailed structure of this field by considering a number of the parameter of the precipitation distribution in the verification process.
The probabilistic verification tools considered here are: Relative Operating Characteristic (ROC Curves, Mason and Graham, 1999), Brier Score and Brier Skill Score (Stanski et al., 1989, Wilks, 1995), Cost–Loss Analysis (Richardson et al., 2000) and Percentage of Outliers (Talagrand et al., 1997, Buizza, 1997). For a brief description of the indices the reader is referred to Marsigli et al. (2004). Though the computation of these scores is rather simple, their interpretation is not straightforward, different indices describing different features of the forecast system. In addition to this, the relationship between these scores is not a linear one. Therefore, a global evaluation of the forecast system should rely on a set of indices. In this report, for brevity reasons, only results in terms of the Brier Skill Score, ROC area and Percentage of Outliers will be presented.
After a brief description of the system (Section 2), the verification methodology is presented (Section 3), followed by the results (Section 4).

2 The COSMO–LEPS operational system

The limited–area ensemble prediction system COSMO–LEPS has been running operationally at ECMWF since November 2002. The suite is run and maintained remotely by ARPA–SIM, with support given by ECMWF, and the necessary Billing Units are made available by the ECMWF COSMO countries.
The system has recently been updated (June 2004) and now 10 runs of the non–hydrostatic limited–area model (LM) are available every day, nested on 10 selected members (the so–called Representative Members, or RMs) of two consecutive 12–hour lagged ECMWF global ensembles. The 10 selected members are representative of 10 clusters, built by grouping all

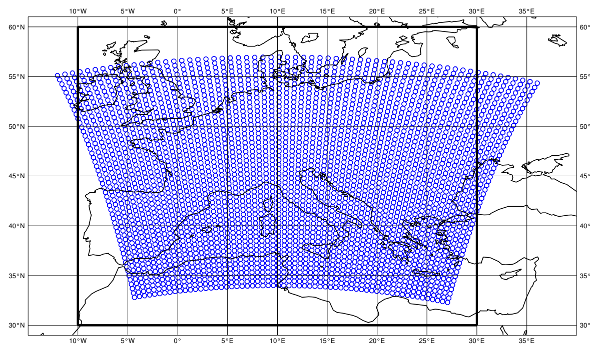


Figure 1: COSMO–LEPS operational domain (small circles) and clustering area (big rectangle).

the global ensemble members on the basis of their similarity in terms of upper–air fields. Mesoscale perturbations are also added by letting the different LM runs randomly choose the scheme to be used for the parameterisation of the deep convection (Tiedtke or Kain–Fritsch). A description of the old suite and the motivation for the suite update are described in Marsigli et al. (2005).
The limited–area ensemble forecasts range up to 120 hours and are integrated over a domain covering all the countries involved in COSMO (Fig. 1). The model version is 3.9 (prognostic precipitation and cloud ice scheme have been activated), the horizontal resolution is about 10 km and 32 vertical layers are used. LM–based probabilistic products covering a "short to medium–range" (48–120 hours) are disseminated to the weather services involved in COSMO.

3 Verification methodology

Verification is performed in terms of daily precipitation, cumulated from 06 to 06 UTC. Precipitation observations are available on the very dense COSMO station network (over 4500 stations, see Fig. 2) covering Germany, Switzerland, Poland and part of Italy.

For verification purposes, the verification area is covered by 12 $\times$ 11 boxes of 1.5 $\times$ 1.5 degrees (approximately equal to 150 km $\times$ 150 km) and a pair of *representative* observed and forecast values is individuated for each box. This approach is followed to permit the comparison between a punctual value (the observation) and an areal value (the forecast). Several observed and several forecast values fall in each box and a comparison between the distribution of the observed and the distribution of the forecast values is attempted. A number of statistical properties of the two distributions are computed over each box, thus allowing a comparison between forecast and observed values which are representative of different features of the precipitation distribution. In this work, the average, the median (50th percentile), the 90th percentile and the maximum are computed.

COSMO–LEPS performances are compared with those of the ECMWF EPS, both the operational full–size 51–member EPS and the reduced–size 10–member EPS made up by the 10 selected Representative Members. The comparison of COSMO–LEPS with this 10–RM–EPS permits to quantify the impact of the high–resolution alone, irrespective of the different number of members of the two operational systems. The need to compare the 10–km COSMO–LEPS with the 80–km EPS determines the choice of boxes as big as 1.5 $\times$ 1.5 degrees, in order to have enough EPS grid points within a box to make possible the computation of the statistical properties. The three systems will be referred to as:

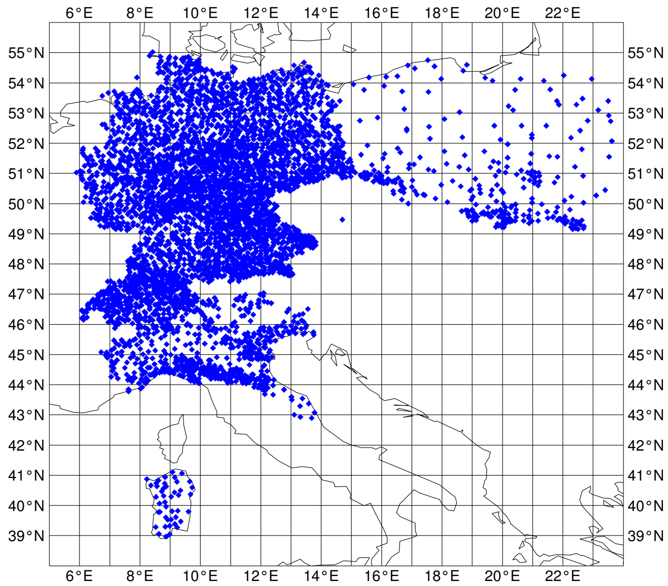


Figure 2: COSMO network of stations were observed precipitation is available. Precipitation data are cumulated over 24 hours from 06 to 06 UTC.

- `cleps`: COSMO-LEPS, 10 members, 10 km hor. res.
- `epsrm`: reduced EPS made up by the RMs, 10 members, 80 km hor. res.
- `eps51`: operational EPS starting at 12 UTC, 51 members, 80 km hor. res.

The period considered for verification is Autumn (September, October and November) 2004. Verification has been performed in terms of 24–hour cumulated precipitation (from 06 to 06 UTC).

4 Results

Results are presented in terms of a set of indicators: Brier Skill Score (BSS), ROC area, Talagrand diagrams and Perecntage of Outliers. For the Brier Skill Score (Stanski et al., 1989) a higher value corresponds to a better results and the zero level indicates the limit of usefulness of the forecasting system. The ROC area (Mason and Graham, 1999) can take values in the range [0,1], the higher the better, and the no–skill level is 0.5. The Talagrand diagram (Talagrand et al., 1999) is obtained by counting how many times the truth falls in each of the bins that are obtained by putting the forecast values in increasing order. An U–shape of the diagram indicates that the ensemble is underdispersive, while a dome shape indicates that is overdispersive. The best shape is the uniform distribution. The Percentage of Outliers (Talagrand et al., 1999, Buizza, 1997) indicates the percentage of times the truth falls outside from the range of the forecast values, so it sums up the informations coming form the two extreme bins of the Talagrand diagrams.

Average values

Results from the verification in terms of average values exceeding 10mm/24h over the 1.5×1.5 degree boxes are presented in Fig. 3.

In terms of BSS (left panel), `cleps` is performing worse than `epsrm`, while the two systems are comparable in terms of ROC area (right panel). This indicates that as regards the total

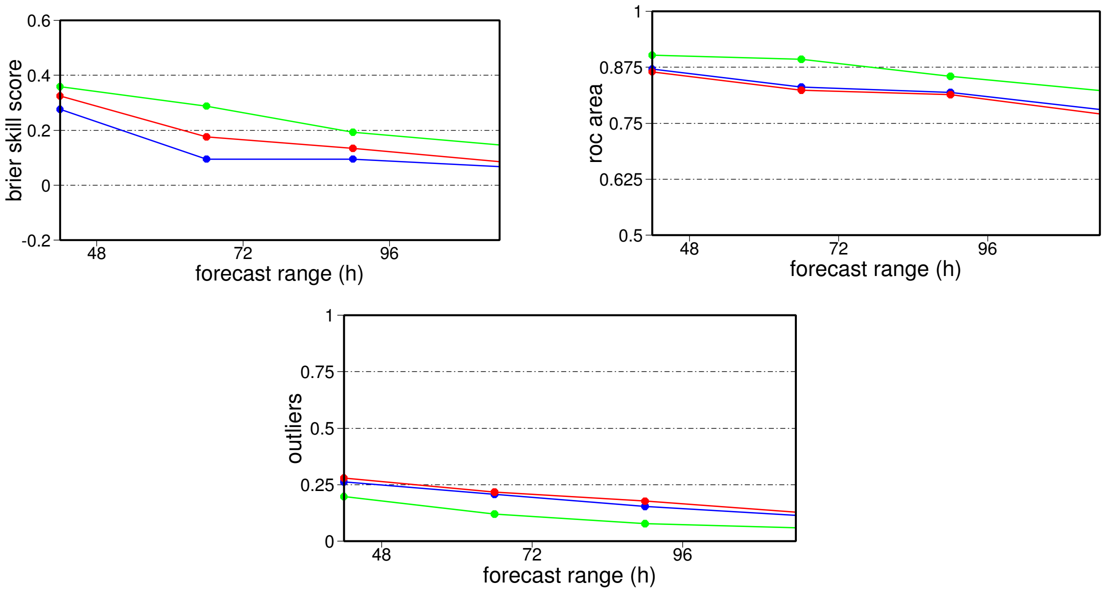


Figure 3: Brier Skill Score (top left panel) ROC area (top right panel) and Percentage of Outliers (bottom panel) as a function of the forecast range (in hours) relative to the 24–hour cumulated precipitation forecasts by COSMO–LEPS (`cleps`, blue line), by the 10–RM EPS (`epsrm`, red line) and by the operational 51–member EPS (`eps51`, green line) for the 10mm precipitation thresholds. For each box, the mean of the forecast values is compared with the mean of the observed forecast values.

amount of precipitation falling within a box, the information provided by the global ensemble is enough for this threshold and even better in terms of reliability. Even in terms of outliers, the high–resolution system is not lowering the percentage (bottom panel).

The use of the operational EPS (`eps51`) seems the best solution for this quantity over an area as big as 150 $\times$ 150 km^2. It is not possible to repeat the verification by decreasing the dimension of the area because of the low resolution of the EPS.

In order to give an idea of what the spread of the considered ensemble is, the Talagrand diagrams are also presented (Fig. 4).

The marked U–shape of the diagram relative to the `eps51` system is evident (bottom panel), indicating underdispersion, even though the total number of outliers is smaller than for the oher two systems. It is also evident that the truth is more frequently in the upper tail of the distribution, both for `eps51` and for `epsrm` (upper right panel), while `cleps` provide a more uniform distribution (upper left panel).

Median (50th percentile) values

A comparison in terms of the median values has also been carried out but the results are comparable to those obtained for the mean values, so they are not shown.

90th percentile values

In the COSMO–LEPS verification, we are mainly interested in the tail of the precipitation distribution over an area, where the lower resolution system is supposed to show some deficiencies and the use of the mesoscale model would permit more realistic precipitation structures with higher peak values in case of intense precipitation. For this purpose, a verification in terms of the 90th percentile of the precipitation distribution has also been performed

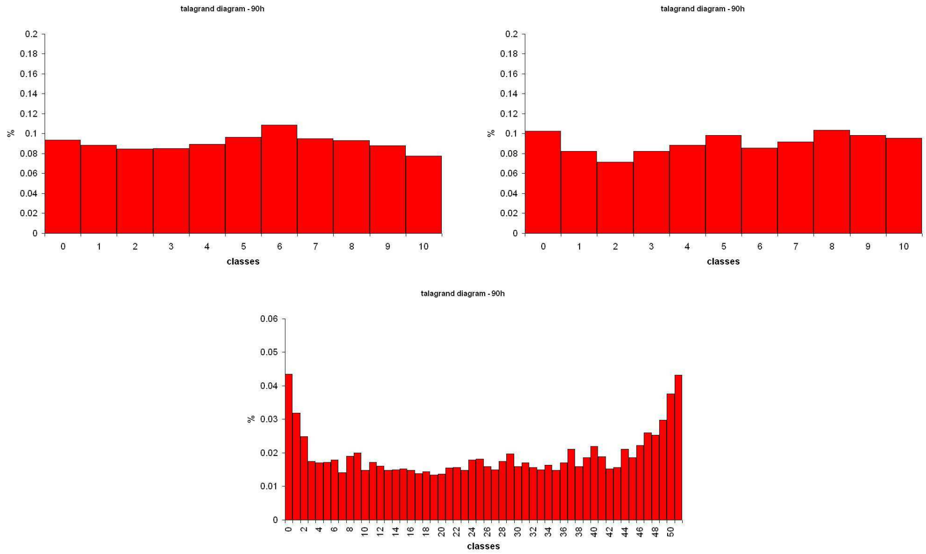


Figure 4: Talagrand diagrams relative to `cleps` (top left panel) `epsrm` (top right panel) and `eps51` (bottom panel) computed for average values within boxes at the 90 hour forecast range. The scale of the y–axis is different for the `eps51` system.

(Fig. 5) for the threshold 20mm/24h.

In terms of BSS (left panel) the three systems exhibit similar performances, with `eps51` still showing slightly higher values, while in terms of ROC area (right panel) the `cleps` system overperforms `epsrm` by a large amount and, to a lesser extent, also `eps51`. The percentage of outliers (bottom panel) produced by `cleps` is lower than that of `epsrm`, while `eps51` has the lowest value but a direct comparison with the other two is not fair due to the very different number of ensemble members. Form this kind of verification it appears that, when considering the tail of the precipitation distribution over a box, the high resolution of the COSMO–LEPS system plays an important role, which is not rewarded when considering average values.

The Talagrand diagrams are also presented (Fig. 6).

For this parameter, `cleps` exhibits a tendency to forecast too high values (top left panle), the upper tail of the distribution being slightly less populated, while the pronounced U–shape provided by `eps51` is still evident (bototm panel), with a tendency to predict too lower values, which is even more evident in `epsrm` (top right panel).

Maximum values

The capability of the COSMO–LEPS system to signal the possibility of the occurrence of very large precipitation over an area can be quantified by repeating the verification in terms of maximum values over a box. Results are shown in Fig. 7 for the threshold 50mm/24h.

In terms of both scores (BSS in the left panel and ROC area in the right panel) `cleps` is overperforming the lower resolution systems, showing a positive skill in forecasting the occurrence of heavy precipitation over an area. The percentage of outliers (bottom panel) is also shown.

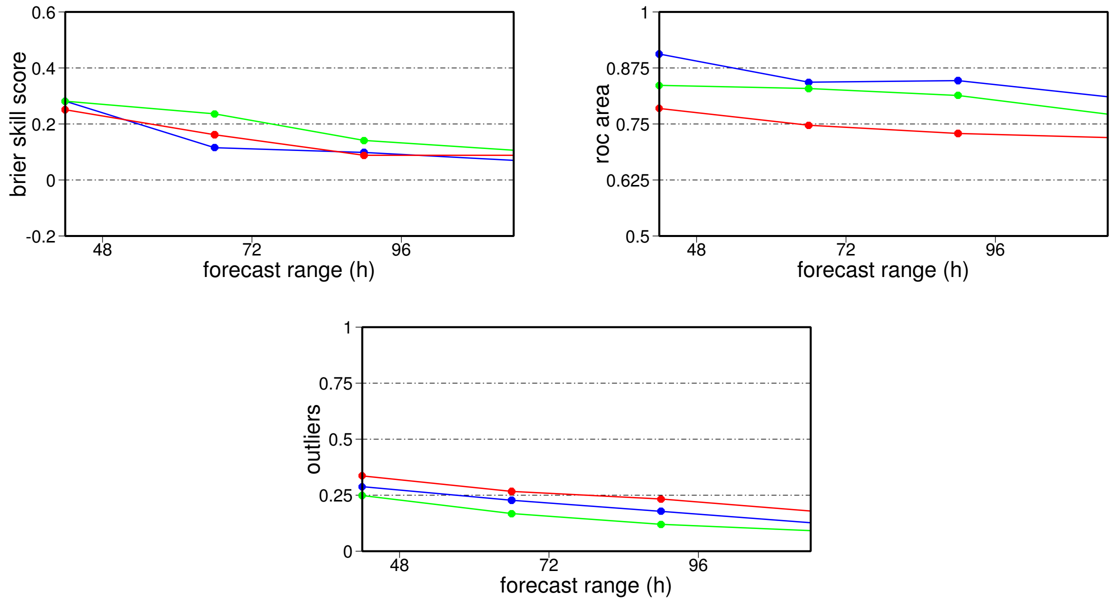


Figure 5: Brier Skill Score (top left panel) ROC area (top right panel) and Percentage of Outliers (bottom panel) as a function of the forecast range (in hours) relative to the 24–hour cumulated precipitation forecasts by COSMO–LEPS (`cleps`, blue line), by the 10–RM EPS (`epsrm`, red line) and by the operational 51–member EPS (`eps51`, green line) for the 20mm precipitation thresholds. For each box, 90^{th} percentile of the forecast values is compared with the 90^{th} percentile of the observed forecast values.

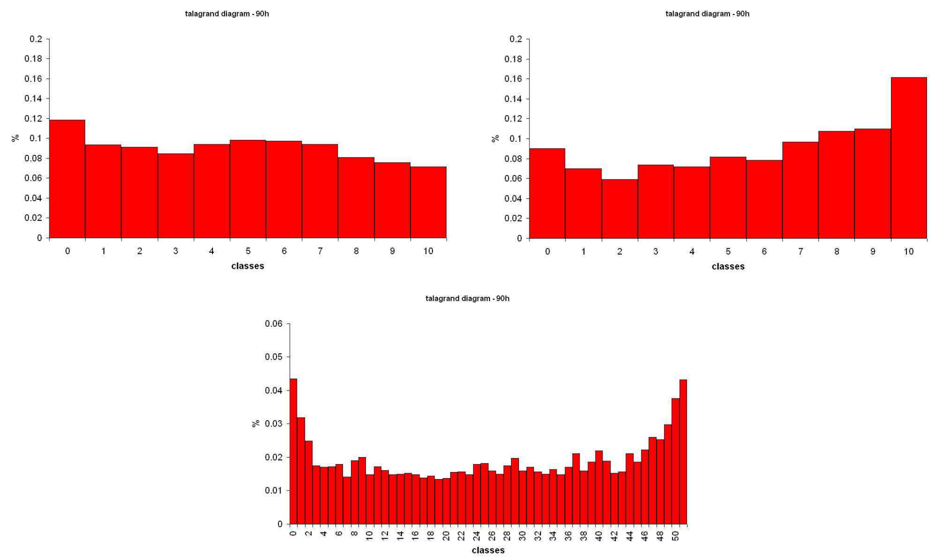


Figure 6: Talagrand diagrams relative to `cleps` (top left panel) `epsrm` (top right panel) and `eps51` (bottom panel) computed for $90^{t}h$ percentile values within boxes at the 90 hour forecast range. The scale of the y–axis is different for the `eps51` system.

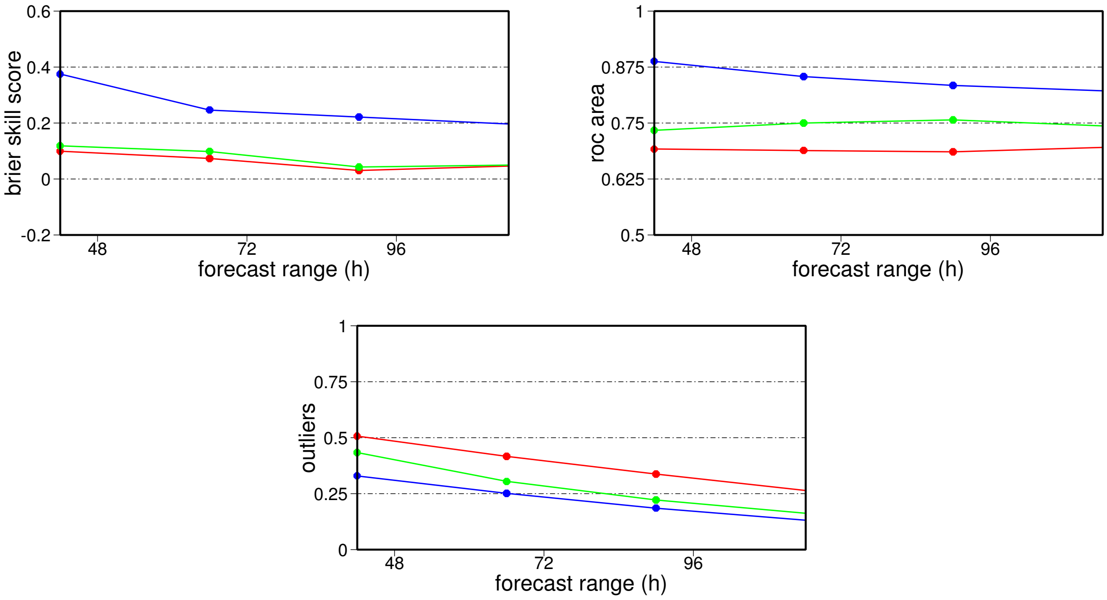


Figure 7: Brier Skill Score (top left panel) ROC area (top right panel) and Percentage of Outliers (bottom panel) as a function of the forecast range (in hours) relative to the 24–hour cumulated precipitation forecasts by COSMO–LEPS (`cleps`, blue line), by the 10–RM EPS (`epsrm`, red line) and by the operational 51–member EPS (`eps51`, green line) for the 20mm precipitation thresholds. For each box, the maximum of the forecast values is compared with the maximim of the observed forecast values.

5 Conclusions

An objective verification of the COSMO–LEPS system in terms of precipitation distribution has been shown. Results indicate that the COSMO–LEPS system is useful for the forecast of intense precipitation over an area, allowing a good description of the tail of the precipitation distribution ($90^t h$ percentile) and permitting to capture the occurrence of high precipitation values (maximum). As regards average values over an area, the best performance is obtained by the operational full-size EPS.

A direct comparison between the old (5–member 3–EPS) suite and the new (10–member 2–EPS) suite is not possible, because they were never run in parallel on a common period. Results obtained with the old suite for Autumn 2003 (Marsigli et al., 2005) are not comparable with the results here presented for Autumn 2004. Nevertheless, more or less the same conclusions can be drawn from the results obtained in 2003 and in 2004.

References

Buizza R., Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system, 1997. *Mon. Wea. Rev.*, 125, pp. 99–119.

Marsigli, C., Montani, A., Nerozzi, F., Paccagnella, T., Tibaldi, S., Molteni, F., Buizza, R., A strategy for High–Resolution Ensemble Prediction. Part II: Limited–area experiments in four Alpine flood events, 2001. *Quart. J. Roy. Meteor. Soc.*, 127, pp. 2095–2115.

Marsigli, C., Boccanera, F., Montani, A., Nerozzi, F., Paccagnella, T., COSMO–LEPS verification: first results, 2004. *COSMO Newsletter*, No. 4, pp. 125–135.

Marsigli, C., Boccanera, F., Montani, A. and Paccagnella, T., The COSMO–LEPS ensemble system: validation of the methodology and verification, 2005. *Nonlinear Processes in Geophysics*, 12, pp. 527–536.

Mason, S. J. and Graham, N. E., Conditional probabilities, relative operating characteristics and relative operating levels, 1999. *Wea. and Forecasting*, 14, pp. 713–725.

Molteni, F., Buizza, R., Marsigli, C., Montani, A., Nerozzi, F. and Paccagnella, T., 2001. A strategy for High–Resolution Ensemble Prediction. Part I: Definition of Representative Members and Global Model Experiments. *Quart. J. Roy. Meteor. Soc.*, 127, pp. 2069–2094.

Montani, A., Marsigli, C., Nerozzi, F., Paccagnella, T. and Buizza, R., 2001. Performance of ARPA–SMR Limited–area Ensemble Prediction System: two flood cases. *Nonlinear Processes in Geophysics*, 8, pp. 387–399.

Montani, A., Marsigli, C., Nerozzi, F., Paccagnella, T., Tibaldi, S. and Buizza R., 2003a. The Soverato flood in Southern Italy: performance of global and limited–area ensmeble forecasts *Nonlinear Processes in Geophysics*, 10, pp. 261–274.

Montani, A., Capaldo, M., Cesari, D., Marsigli, C., Modigliani, U., Nerozzi, F., Paccagnella, T., Patruno, P. and Tibaldi, S., Operational limited–area ensmeble forecasts based on the Lokal Modell, 2003b. *ECMWF Newsletter Summer 2003*, 98, pp. 2–7.

Richardson, D. S., Skill and relative economic value of the ECWMF ensemble prediction system, 2000. *Quart. J. Roy. Meteor. Soc.*, 126, pp. 649–667.

Stanski, H. R., Wilson, L. J. and Burrows, W. R., Survey of common verification methods in meteorology, 1989. *WMO World Weather Watch Tech. Rep.*, 8, pp. 144.

Talagrand, O., Vautard, R. and Strauss, B, Evaluation of probabilistic prediction systems, 1999. *Proceedings of the ECMWF workshop on predictability*, 20–22 October 1997, Reading, UK, pp. 372.

Wilks, D. S., Statistical methods in the atmospheric sciences, 1995. Academic Press, New York, 467 pp..

COSMO-LEPS Forecasts for the August 2005 Floods in Switzerland

André Walser

MeteoSwiss, Krähbühlstrasse 58, 8044 Zürich, Switzerland

1 Introduction

In August 2005, heavy precipitation for three days caused tremendous floods in Switzerland and in adjacent neighborhood countries. The precipitation started on 20 August and was most intense on the northern slopes of the Alpine ridge. The according weather type corresponded to the well known Vb situation with a low pressure system over Italy that transported warm and very moist air at the eastern edge of the Alps to the northern side where the air impinged in the northern Alpine slopes. Fig. 1 illustrates the synoptic situation for August 22 taken from the assimilation cycle of the Swiss Alpine Modell (aLMo; see COSMO Newsletter, No. 5, Section 4.5). This report documents the COSMO-LEPS forecasts for this extreme precipitation event and compares them with observations.

2 Observations

At MeteoSwiss, a high-resolution precipitation analysis has been derived for this event (Fig. 2). The analysis considers measurements of about 400 Swiss rain gauges and a high-resolution precipitation climatology. From 20 August 06 UTC until 23 August 06 UTC, more than 100 mm precipitation occurred in a large area from the western Alps to the north eastern Alpine foreland. In central Switzerland more than 150 mm were observed and at some locations even more than 300 mm during these 72 hours.

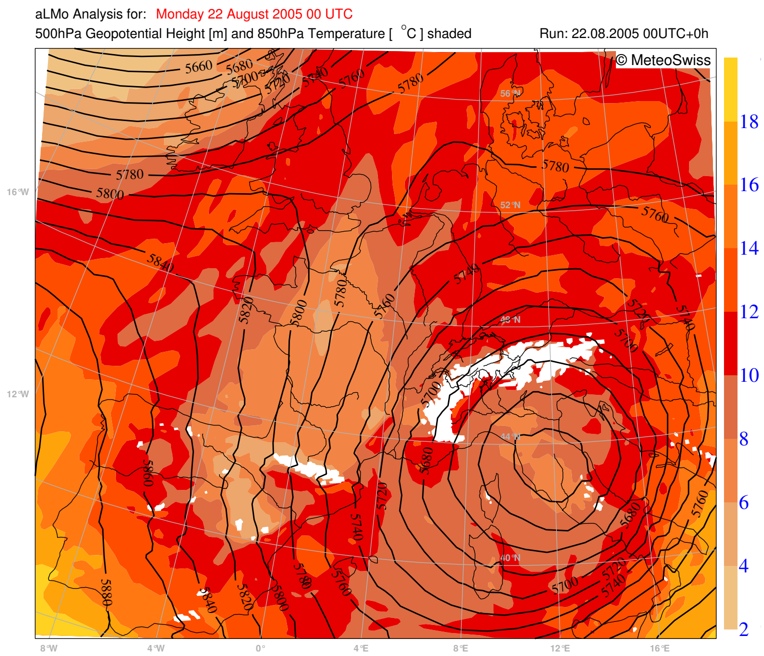


Figure 1: Overview of the synoptic situation for 22 August 2005 00 UTC from the aLMo analysis.

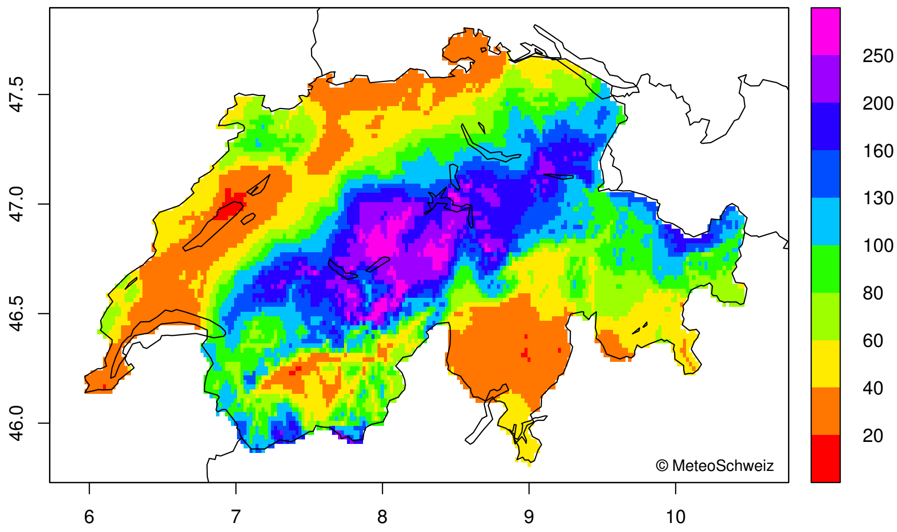


Figure 2: Analyzed accumulated precipitation from 20 Aug 2005 06 UTC to 23 Aug 2005 06 UTC (courtesy C. Frei, MeteoSwiss).

3 COSMO-LEPS forecasts

The limited-area ensemble prediction system COSMO-LEPS (Montani et al., 2003; Marsigli et al, 2005) computes daily probabilistic high-resolution weather forecasts for central and southern Europe for a lead-time of 132 hours.

Figure 3 shows the COSMO-LEPS forecast from 18 August 2005 12 UTC for the 72-h precipitation sum between 20 August 06 UTC and 23 August 06 UTC (lead-time 42h-114h). The four panels show the probabilities to exceed the corresponding thresholds 50, 100, 150 and 250 mm, respectively. This medium-range forecast reveals probabilities up to 60% for precipitation sums higher than 100 mm for the northern Alpine slopes and for the western part of Ticino as well as very low probabilities to exceed 150 mm for large parts of Switzerland, highest in the western part of Ticino. Overall, COSMO-LEPS predicted the event already with a lead-time of almost five days, but with higher probabilities rather for the southern Alpine region than for the northern Alpine area.

The forecast of the following day for the same forecast period is depicted in Fig. 4. The panels show for the entire northern Alpine slopes high probabilities (up to 80%) for accumulated precipitation over 100 mm and high probabilities (over 60%) for precipitation above 150 mm particularly in the Bernese and central Alps. In addition, a scenario with accumulated precipitation over 250 mm for some locations in these regions is predicted with probabilities of 20-30%. The two panels in Fig. 5 show the precipitation analysis with a contour corresponding to two of the precipitation thresholds used for the probability maps, namely for 100 mm and 250 mm, respectively. The panel for the former threshold demonstrates a high correlation of region with precipitation higher than 100 mm with those region showing high predicted probabilities (cf. Fig. 4). Further, the locations with observed precipitation above 250 mm are in those region where the predicted probabilities were highest for this threshold.

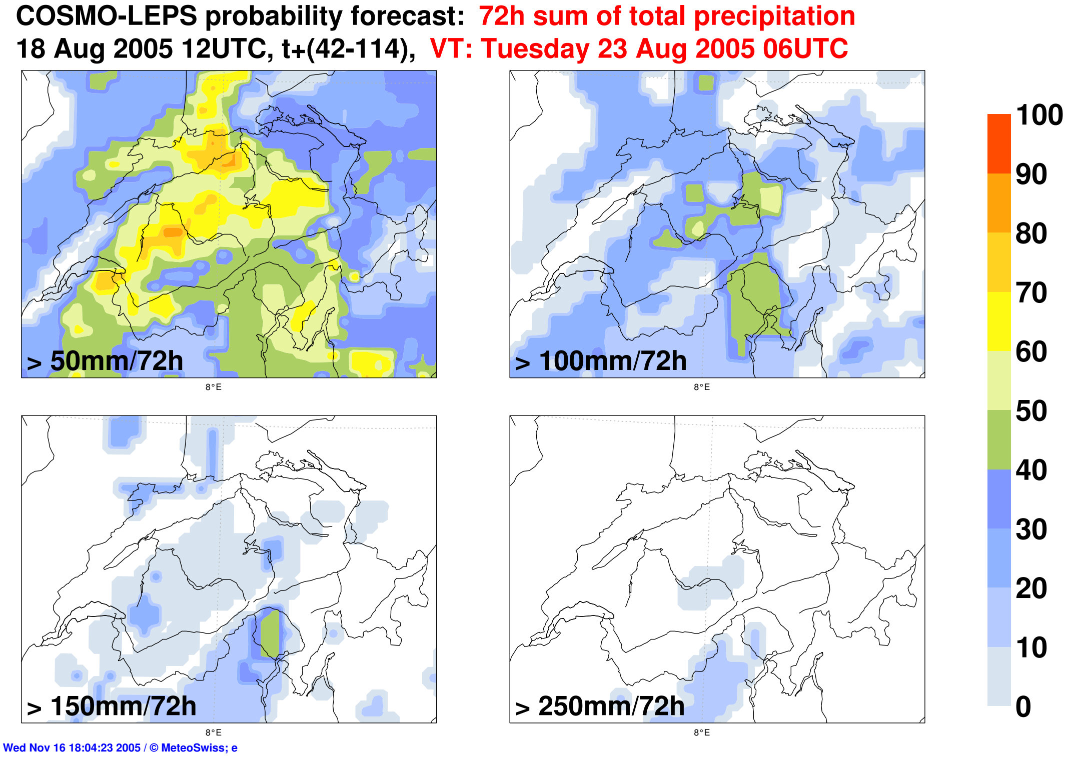


Figure 3: COSMO-LEPS forecast from 18 Aug 2005 12 UTC for 72-h precipitation sum. The panels show the probabilities to exceed the given thresholds 50, 100, 150 and 250 mm/72h, respectively, for the period 20 Aug 06 UTC to 23 Aug 06 UTC.

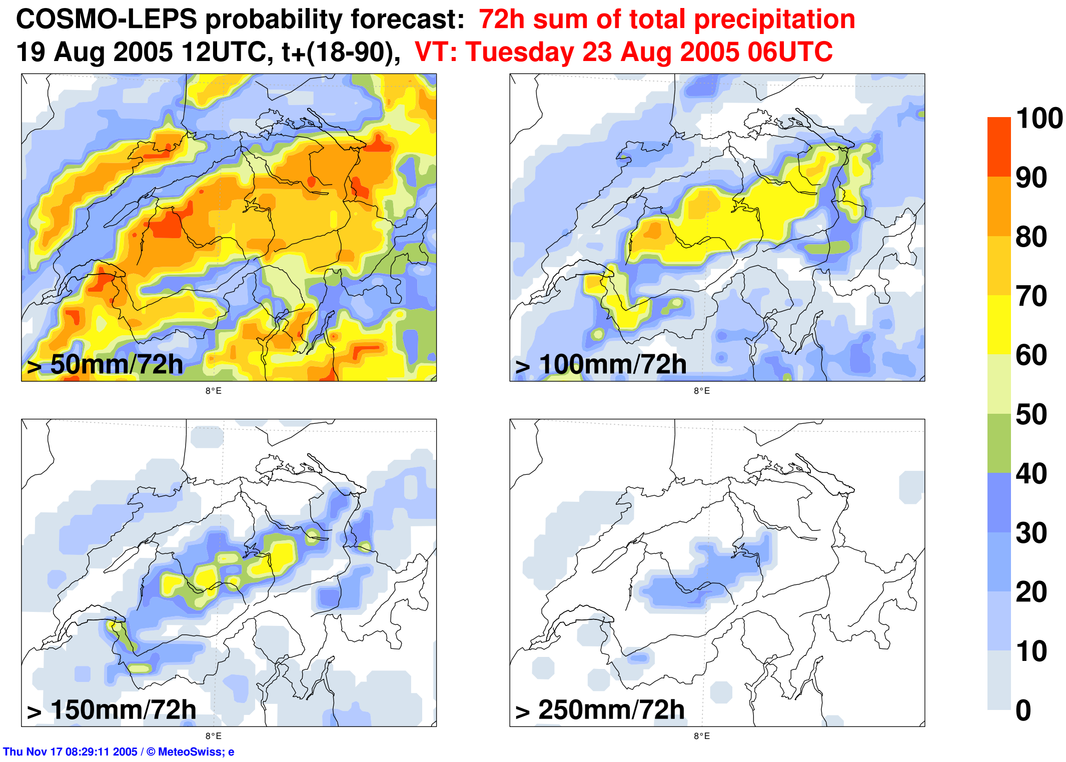


Figure 4: Same as Fig. 3, but for COSMO-LEPS forecast from 19 August 2005 12 UTC.

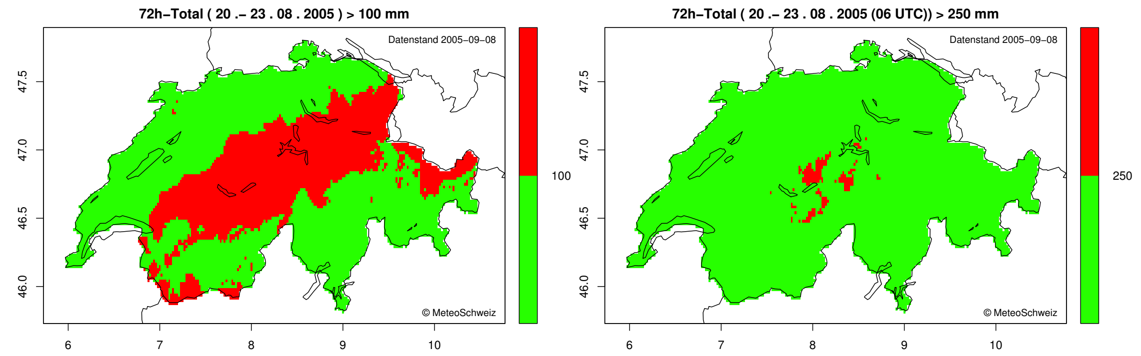


Figure 5: Same as Fig. 2, but (left) with a 100 mm and (right) with a 250 mm contour, respectively.

4 Conclusion

In summary, COSMO-LEPS forecasts provided very appropriate warnings for the extreme precipitation event in August 2005. In particular the forecast initialized at 19 August 2005 12 UTC predicted high probabilities for large precipitation amounts for most of the regions hit by the event without giving obvious false alerts for other regions.

References

Marsigli, C., F. Boccanera, A. Montani, and T. Paccagnella, 2005: The COSMO-LEPS mesoscale ensemble system: validation of the methodology and verification. *Nonlinear Processes in Geophysics.*, 12, pp. 527–536.

Montani, A., M. Capaldo, D. Cesari, C. Marsigli, U. Modigliani, F. Nerozzi, T. Paccagnella, P. Patruno, and S. Tibaldi., 2003: Operational limited-area ensemble forecasts based on the Lokal Modell, ECMWF Newsletter No. 98.

Use of Multimodel SuperEnsemble Technique for Complex Orography Weather Forecast

Massimo Milelli[1], Daniele Cane[1,2]

[1] *ARPA Piemonte*

[2] *Organizing Committee for the XX Winter Olympic Games Torino 2006*

1 Introduction

The Multimodel SuperEnsemble technique (see Krishnamurti et al, 1999 and 2000 for instance) is a powerful post-processing method able to reduce direct model output errors. Several model outputs are put together with adequate weights to obtain a combined estimation of meteorological parameters. Weights are calculated by square error minimization in a so-called training period. In a previous paper (Cane and Milelli, 2005), we applied the Multimodel technique on the operational 00 UTC runs of Local Area Model Italy (LAMI) by UGM, ARPA-SIM, ARPA Piemonte (`nud00`), Lokal Modell (LME) by Deutscher Wetterdienst (`lkd00`) and aLpine Model (aLMo) by MeteoSwiss (`alm00`). This was one of the first implementations of Multimodel technique on limited-area models (in this case of 0.0625° resolution) and we obtained a strong improvement in temperature forecasts in Piedmont region. In this work we extend the application of temperature and precipitation to larger periods and we introduce the method to the calculation of humidity, wind speed and precipitation.

2 Multimodel Theory

As suggested by the name, the Multimodel SuperEnsemble method requires several model outputs, which are weighted with an adequate set of weights calculated during the so-called training period. The simple ensemble methods with biased (Eq. 1) or bias-corrected (Eq. 2) data respectively, are given by

$$S = \overline{O} + \frac{1}{N}\sum_{i=1}^{N}(F_i - \overline{F_i}) \tag{1}$$

and

$$S = \overline{O} + \frac{1}{N}\sum_{i=1}^{N}(F_i - \overline{O}) \tag{2}$$

The conventional superensemble forecast constructed with bias-corrected data is given by

$$S = \overline{O} + \sum_{i=1}^{N} a_i(F_i - \overline{O}) \tag{3}$$

where N is the number of models, F_i is the i^{th} forecast by the model, $\overline{F_i}$ and $\overline{O}$ are the mean forecasts and the mean observation during the training period T.

The calculation of the parameters a_i is given by the minimization of the mean square deviation

$$G = \sum_{k=1}^{T}(S_k - O_k)^2 \tag{4}$$

by derivation $\left(\frac{\partial G}{\partial a_i} = 0\right)$ we obtain a set of N equations, where N is the number of models involved $(i, j = 1, N)$:

$$\left(\sum_{k=1}^{T} \left(F_{i_k} - \overline{F_i}\right)\left(F_{j_k} - \overline{F_j}\right)\right) \cdot (a_i) = \left(\sum_{k=1}^{T} \left(F_{j_k} - \overline{F_j}\right)\left(O_k - \overline{O}\right)\right) \tag{5}$$

We then solve these equations using the Gauss-Jordan method (see Press et al., 1992).

3 Results

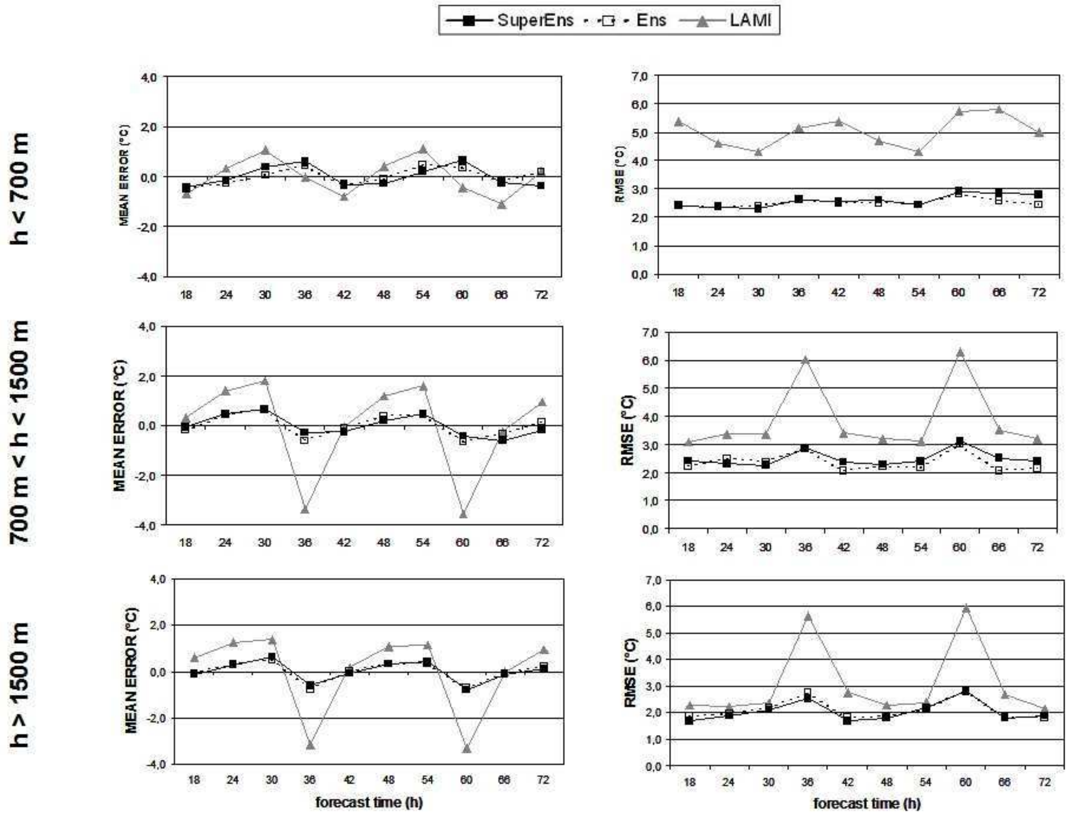


Figure 1: Mean temperature error (left) and RMSE (right) for Superensemble output (black continuous line), Ensemble output (black dotted line) and LAMI output (grey continuous line); low-lying stations (upper panels), middle-mountain stations (middle panels) and high mountain stations (lower panels).

The Piedmont region is monitored by ARPA Piemonte with a very-dense automatic weather station network. We used the data from this non-GTS network for the calculation of the weights in the training period and for validation purposes. In order to obtain more readable graphs, we do not report all the model outputs, but only the operational one (LAMI 00 UTC run). In order to compare with the unbiased values of SuperEnsemble and Ensemble, all the direct model output forecasts here shown are bias-corrected, with the exception of precipitation forecasts since we do not expect to have a systematical error in this case.

Temperature

Stations are grouped by height: 53 low-lying stations (h < 700 m), 34 middle-mountain stations (700 m < h < 1500 m) and 15 high-mountain stations (h > 1500 m). The training period is 90 days (dynamical) and the forecast is on March 2005. We used a bilinear interpolation in the horizontal direction and a linear interpolation (with the geopotential) in the vertical one. In Fig. 1 the BIAS and the RMSE are shown, according to the station elevation. It has to be pointed out the strong systematic error of the direct model outputs, reaching a bias of the order of 4 C, with significant increase around noon (+36 hr and +60 hr forecast time). Multimodel SuperEnsemble substantially eliminates the bias, and the RMSEs also are lower than direct model outputs' ones, with values around 2 C. Moreover we observe a constant performance for all the forecast times.

Relative Humidity

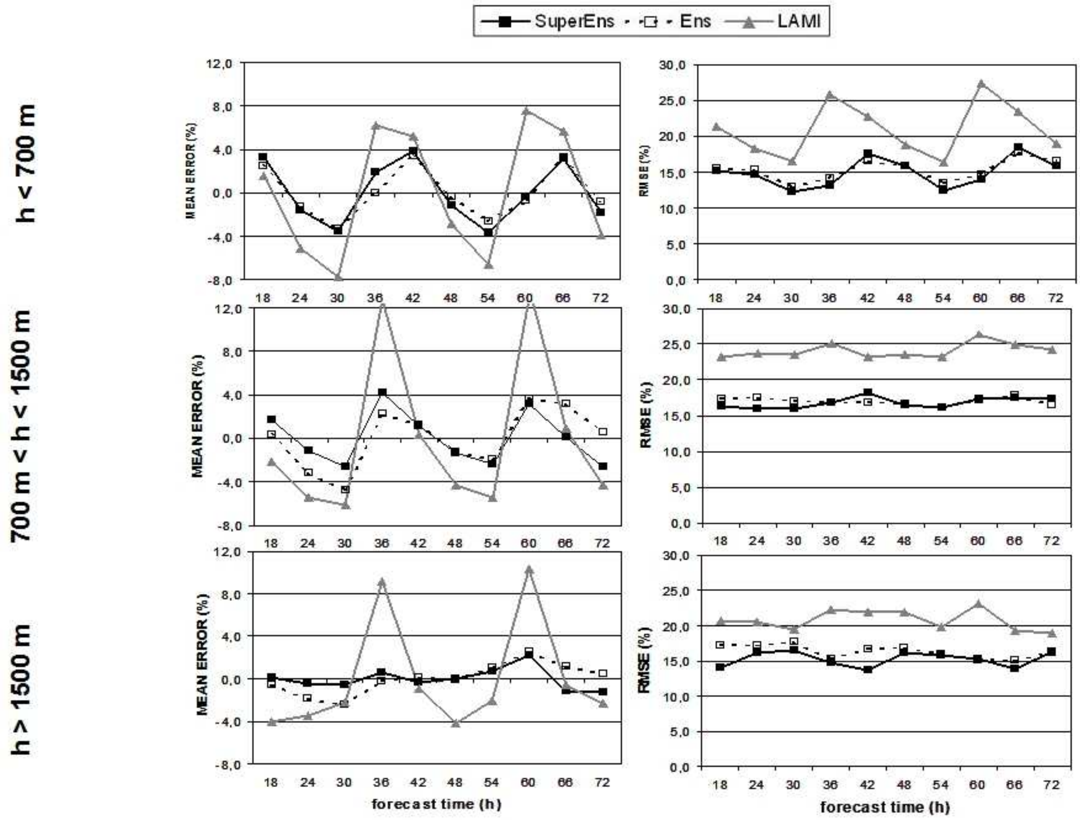


Figure 2: Mean relative humidity error (left) and RMSE (right) for Superensemble output (black continuous line), Ensemble output (black dotted line) and LAMI output (grey continuous line); low-lying stations (upper panels), middle-mountain stations (middle panels) and high mountain stations (lower panels).

The stations are grouped as before and the training period, the forecast time and the interpolation methods are the same used for the temperature forecast. Relative humidity (Fig. 2) shows strong systematic error of the direct model outputs, as temperature does, with high biases and RMSEs. Also in this case the errors are strongly dependent from the forecast time. SuperEnsemble practically eliminates bias, especially for higher elevation stations, with slightly better performances by SuperEnsemble. We also obtained a good RMSE reduction. Both biases and RMSEs are very stable with respect to the forecast time. It has to be highlighted that relative humidity, due to its non-gaussian error distribution, does not satisfy Kalman filter hypothesis. In fact Kalman filter post-processing does not improve significantly relative humidity forecasts. Multimodel SuperEnsemble, on the other hand, does not assume any hypothesis and is suitable to be applied to every meteorological parameter.

Wind Intensity

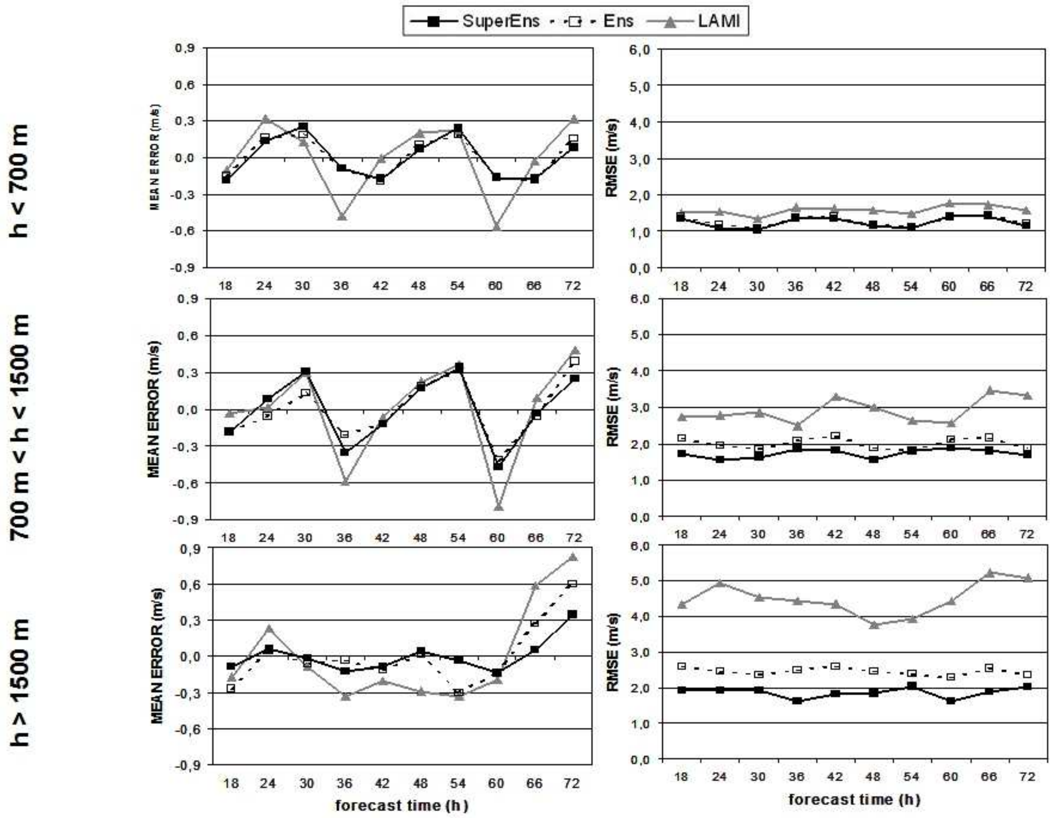


Figure 3: Mean wind intensity error (left) and RMSE (right) for Superensemble output (black continuous line), Ensemble output (black dotted line) and LAMI output (grey continuous line); low-lying stations (upper panels), middle-mountain stations (middle panels) and high mountain stations (lower panels).

Due to model data availability, for this parameter only the ECMWF IFS and the Italian LAMI (00 UTC and 12 UTC operational runs) were used. Stations are grouped in the same groups as for temperature and the training period and the forecast time are the same used for the temperature forecast but here we used the model grid point nearest to the observation. Direct model outputs (Fig. 3) show again strong, forecast time dependent errors. Multimodel permits a strong improvement both in biases and RMSEs, very stable with respect to the forecast time. In this case there is room for improvements: in fact in this work we used model outputs on pressure level, due to data availability, but it would be interesting to check the performance with the model level fields.

Precipitation

Precipitation cannot be easily interpolated to station location without introducing huge errors. For this reason we grouped the same stations we used before in 11 warning areas defined for the regional Civil Protection warning system (see Cane and Milelli, 2005). For each warning area we calculated the 6-hour average and maximum precipitation values. We extracted the same precipitation from the models, calculating the average and maximum values of the grid points covering each warning area. The same method is used operationally for standard precipitation verification at ARPA Piedmont. For further details see Milelli et al., 2003. The training period is 180 days (dynamical). We applied Multimodel Ensemble and SuperEnsemble technique on the average and maximum values, considering as forecast the period July 2004 - March 2005, in order to achieve a good statistics with at least 40

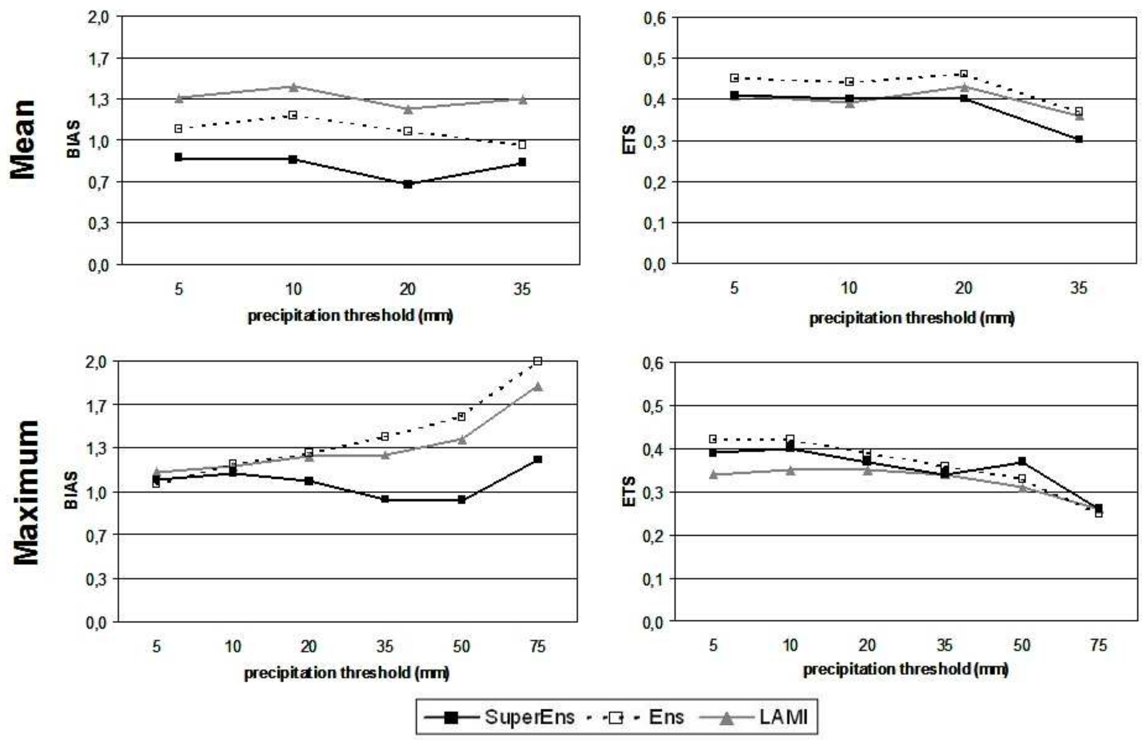


Figure 4: Mean values (upper panels) and maximum values (lower panels) of precipitation in 24h (from +12 to +36) for Superensemble output (black continuous line), Ensemble output (black dotted line)and LAMI output (grey continuous line); BIAS (left panels) and ETS (right panels).

events for each precipitation threshold. We compared the models and Multimodel results by Normalized Bias and Equitable Threat Score (ETS) (see for instance Wilks, 1995). In Piedmont the models usually overestimate average precipitation, as we can see by the BIAS values higher than 1 (Fig. 4 and Fig. 5). Multimodel SuperEnsemble gives a good BIAS reduction. The best improvement is obtained in the spatio-temporal localization of the precipitation events, as described by ETS, for which it shows the highest values. Moreover Multimodel performances are very stable with respect to forecast time, with almost the same BIAS and ETS values for 12-36 UTC and 36-60 UTC forecasts.

4 Conclusions and future perspectives

The Multimodel SuperEnsemble technique has been applied on limited-area and global model in a complex orography alpine region and verified against a large number of weather stations for several weather parameters. For each of them the Multimodel results show good error improvements with respect to the direct model outputs, providing a new powerful post-processing tool. In particular, SuperEnsemble is always superior to Ensemble, except for mean precipitation over warning areas and for ETS in general. The possible future implementations of this technique can be here summarized:

- Extension to other areas and/or variables (observation ⇒ ECMWF analysis):
  - Geopotential
  - MSLP
  - Tracking of cyclones (original purpose, see Krishnamurti et al, 1999 and 2000).

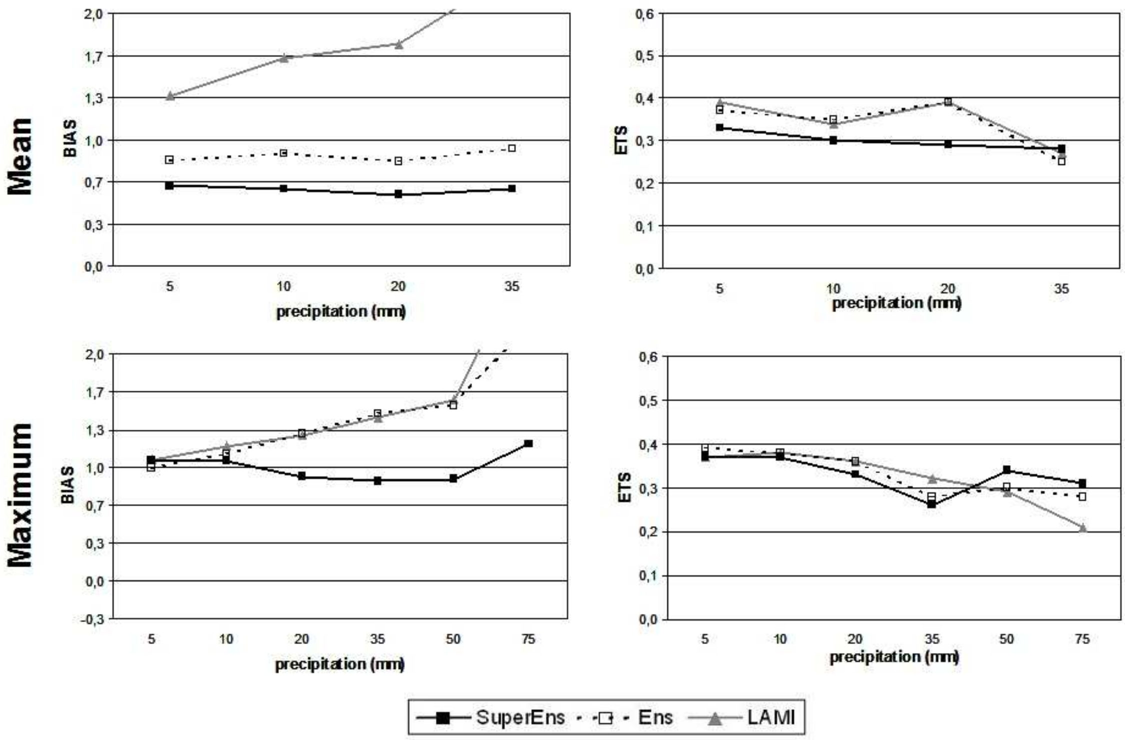


Figure 5: Mean values (upper panels) and maximum values (lower panels) of precipitation in 24h (from +36 to +60) for Superensemble output (black continuous line), Ensemble output (black dotted line)and LAMI output (grey continuous line); BIAS (left panels) and ETS (right panels).

- Study of a spread interval in the forecast of any variable by the introduction of the MultiModel for maximum, mean and minimum values over predefined areas (analogous to precipitation)
- Application to vertical profiles

Moreover, in the framework of the Interreg IIIB-Medocc project Amphore the Multimodel technique will be applied on the Italian LM, Aladin (from MeteoFrance), MM5 (from the University of Balearic Islands), Bolam (from ARPA Liguria) and ECMWF global model for the prediction of 2m temperature and total precipitation.

Acknowledgements

We wish to thank the Deutscher Wetterdienst and MeteoSwiss for providing the model outputs for this research work.

References

Krishnamurti, T. N. et al., *Science*, 285, 1548-1550, 1999

Krishnamurti, T. N. et al., *J. Climate*, 13, 4196-4216, 2000

Cane, D. and Milelli, M., *COSMO Newsletter*, No. 5, 2005

Press, et al., Numerical Recipes in Fortran, Cambridge University Press, 1992

Milelli, M., Oberto, E., Bertolotto, P., Pelosini, R., Proc. of the ECAM/EMS Annual Meeting, Rome, Italy, September, 2003

Wilks, D., Academic Press, 1995

Interpretation of the New High Resolution Model LMK

Heike Hoffmann, Volker Renner and Susanne Theis

Deutscher Wetterdienst, P.O. Box 100465, 63004 Offenbach a. M., Germany

1 Overview

In the framework of the so-called *Aktionsprogramm 2003* at DWD a new high resolution model for the very short range is being developed that is based on the non-hydrostatic limited area model LM. It is called LM Kürzestfrist (LMK) and will have a mesh size of 2 to 3 km which allows a direct calculation of phenomena on the meso-γ-scale. This is especially important for warning purposes. However, there is a problem with the short predictability limit of small scale features. When increasing model resolution the deterministic forecast of a single grid point is often not reliable and the random forecast error can be large. Therefore, the direct model output (DMO) of the LMK has to be interpreted statistically.
Within this project postprocessing methods for the weather parameters of the LMK are being developed. The three aims of the project are:

- to transform the direct model output for point forecasts (smoothed fields),
- to derive probability information for given thresholds and warning events,
- to develop a new weather interpretation for the LMK.

To suppress essentially unpredictable small scale structures, a simple spatial 5×5 averaging will be applied, followed by a re-calibration of the distribution of the smoothed field to the distribution of the direct model output for some of the elements (e.g. precipitation, wind gusts).
Moreover, exceedance probabilities for certain thresholds will be derived, especially with respect to the occurence of severe weather. In a first step these exceedance probabilities will be derived from individual LMK-forecasts by applying the Neighbourhood Method (NM) by Theis et al., 2003. This method was originally developed for the postprocessing of the weather parameters of the LM. The NM transforms the model output into a probabilistic forecast at a given grid point by assuming that points in a spatiotemporal neighbourhood constitute a sample of the forecast at the central grid point.
For this purpose also parameters from the weather interpretation (thunderstorms, fog, etc.) are required. The general task of the weather interpretation is to derive elements that are not calculated by the model directly and to translate the model information into the WMO weather code. The weather interpretation is operational for the LM (Renner 2002). Transfering the method to the LMK is not trivial, because the LMK is a convection resolving model. Therefore, model output parameters originating from the LM convection scheme are no longer available and have to be replaced by other promising parameters like the new LMK graupel, maximum radar reflectivity and maximum vertical velocity for instance.
In the following sections, verification results for the deterministic and probabilistic products as well as a short introduction to the weather interpretation are given.

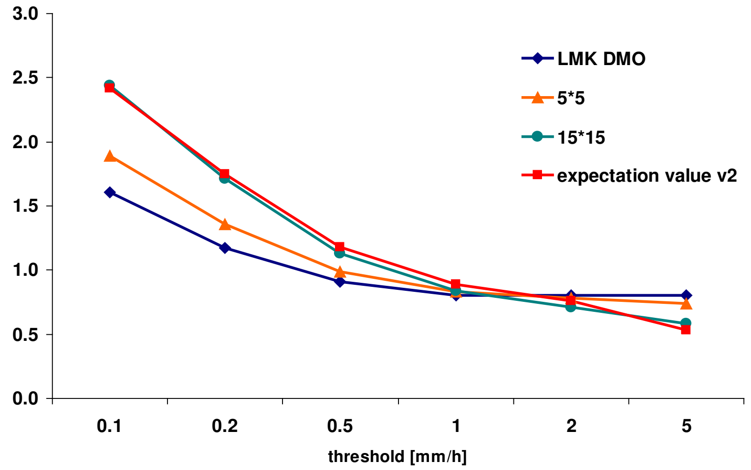


Figure 1: Frequency Bias (FBI) for different precipitation thresholds for the LMK DMO, for 5×5 and 15×15 spatial averaging and the expectation value of the NM version 2 in July 2004, 00 UTC forecasts.

2 Smoothed fields for deterministic point forecasts

High-resolution numerical weather predictions include noticeable stochastic elements even in the short range. In order to suppress essentially unpredictable small scale structures different smoothing techniques have been compared. Figure 1 shows the Frequency Biases (FBI) for the DMO, for simple spatial averaging over quadratic grid boxes of different size as well as the results for the expectation value of Version 2 (see chapter 3) of the Neighbourhood Method. Already for the DMO there is an overestimation of low precipitation amounts (FBI > 1) and an underestimation of high precipitation amounts (FBI < 1). This effect is additionally strengthened by the averaging. In extreme cases the FBI for the highest threshold of 5 mm/h can decrease to zero.
Figure 2 shows the Heidke Skill Scores for the same period. Here only for the highest threshold a degradation of the score due to the averaging is visible. There is however no overall improvement by using the NM instead of simple spatial averaging. Therefore we decided to apply the 5×5 averaging. As it is shown by the FBI, smoothing changes the distribution of the original field. We get more grid points with low precipitation and extreme values are more or less smoothed away. To revoke this change of the distribution a re-calibration algorithm was implemented that reconstitutes the distribution of the DMO also in the smoothed field. With this re-calibration method the obtained frequency bias is nearly the same as for the DMO (see Fig. 3). At this stage we only reconstitute the FBI of the DMO and abandon a re-calibration towards the real 'climate', because the LMK model is still in its development phase.

3 Probabilistic forecasts for weather warnings

A main goal of the LMK-project is the development of a model-based NWP system for very short range forecasts of severe weather especially related to deep moist convection and interactions with fine-scale topography.
Our forecasters have prepared a list of warning criteria that consists of threshold values. Whenever a certain threshold is exceeded, a warning is issued. Our aim is the derivation of

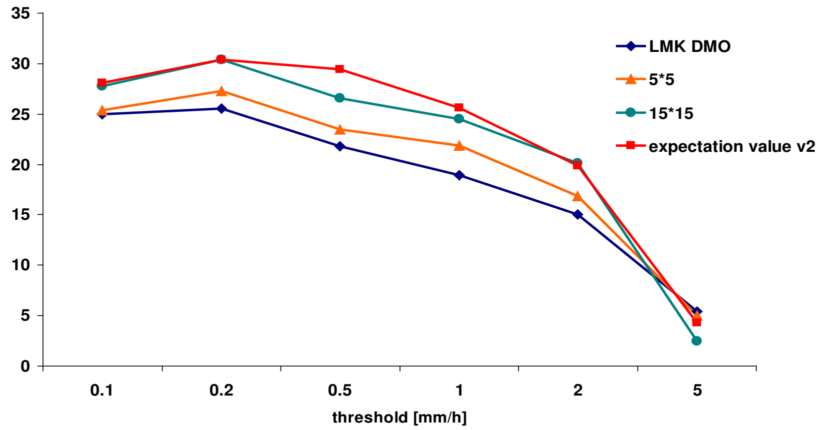


Figure 2: Heidke Skill Score (HSS) for different precipitation thresholds for the LMK DMO, for 5×5 and 15×15 spatial averaging and the expectation value of the NM in July 2004, 00 UTC forecasts.

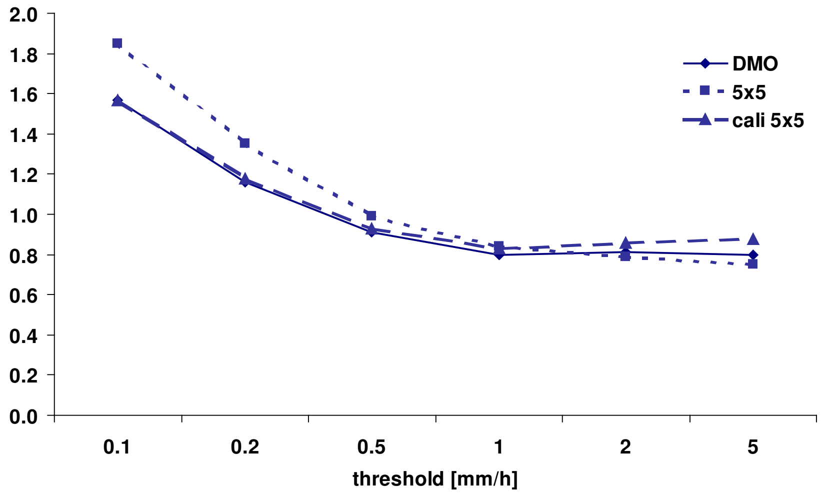


Figure 3: Frequency Bias (FBI) for different precipitation thresholds for the LMK DMO, for the 5×5 spatial averaging and the 5×5 smoothed and calibrated fields in July 2004, 00 UTC forecasts

probabilities for the exceedance of these thresholds.
We plan a two step approach to transform the deterministic LMK forecasts into probabilistic forecasts. In a first step we apply the Neighbourhood Method that uses the information from a spatiotemporal neighbourhood of a single model forecast. In a second step we use the information from the LMK forecasts that will be started every three hours - the so-called lagged average forecast ensemble (LAF).
Current work focusses on the neighbourhood approach. The method has originally been developed for the LM and is now applied to the LMK. To be able to compare the results obtained with the Neighbourhood Method for the LM and for the LMK, the forecasts of a two week period in January 2004 were processed with the NM with similar parameter settings. Figure 4 shows the Equitable Threat Score (ETS) of the 00 UTC forecasts. The dashed lines represent the results for the LM and the solid lines represent the LMK results. In addition to the direct model output, the expectation value and the median are drawn. The results are very similar with slightly better results for the LMK. Due to the critical issue

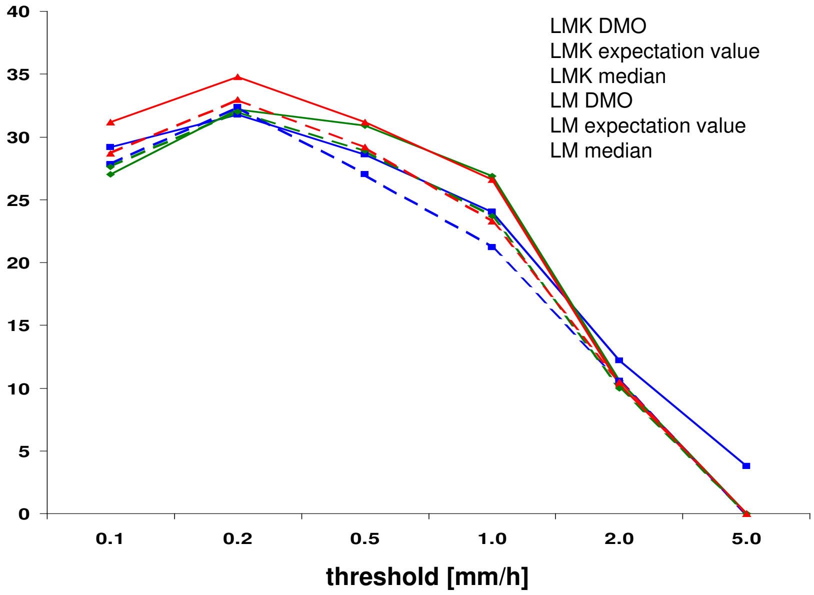


Figure 4: Equitable Threat Score (ETS) for different precipitation thresholds for the LM (dashed lines) and the LMK (solid lines) in January 2004, 00 UTC forecasts.

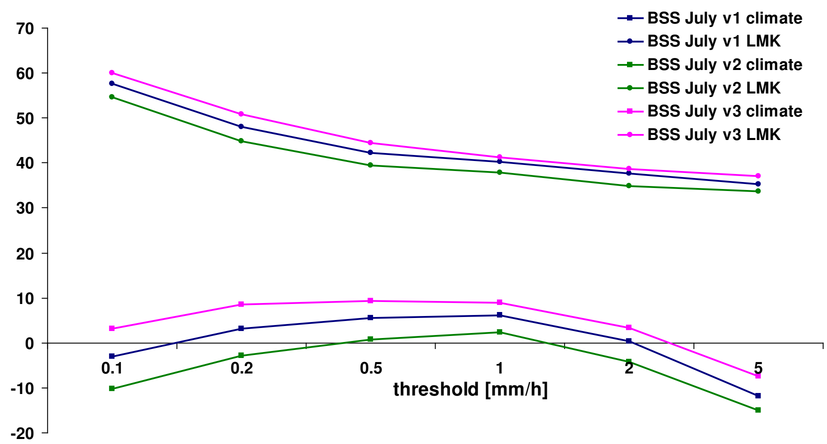


Figure 5: Brier Skill Scores (BSS) for different precipitation thresholds for different versions of the Neigbourhood Method, with 'climate' as reference forecast (lower lines) and with the LMK DMO as reference forecast (upper lines) in July 2004, 00 UTC forecasts.

of the double penalty for higher resolving models, the NM is expected to improve the LMK-DMO to larger extent than the LM-DMO. A higher resolution model predicts an observed small scale structure more realistically but often misplaced in space and time. The model is penalized twice, once for missing the actual feature and again for forecasting it where it does not occur. The NM aims at an alleviation of this problem.

The Brier Skill Score (see Fig. 5) is calculated for several versions of the Neighbourhood Method. The versions differ in the size of the temporal and spatial neighbourhood. The parameters of the used versions were:
vers_01: 3 time levels, radius of 10 grid increments (=28 km)
vers_02: 3 time levels, radius of 5 grid increments (=14 km)
vers_03: 3 time levels, radius of 15 grid increments (=42 km)

The BSS is calculated once with the 'climate' of the respective period as reference forecast and once with the DMO as a reference. However, the 'climate' is not known, so that we have to estimate it from the short verification period itself. In comparison with the real 'climate', better Brier Skill Scores would be achieved, because the estimated 'climate' reference contains too much specific information about the period. We also calculated the BSS with the DMO as reference forecast which leads to a significant improvement with BSS in the range of 0.4 to 0.6.

4 Weather Interpretation

Table 1: WMO code of weather

ww	Description
38	Drifting snow, slight or moderate
39	Drifting snow, heavy
45	Fog
48	Fog, desspositing rime
50	Drizzle
56	Drizzle, freezing
60	Rain, sligth
63	Rain, moderate
65	Rain, heavy
66	Rain, freezing, slight
67	Rain, freezing, moderate or heavy
70	Snowfall, slight
73	Snowfall, moderate
75	Snowfall heavy
80	Rain shower(s), slight
81	Rain shower(s), moderate or heavy
82	Rain shower(s), violent
85	Snow shower(s), slight
86	Snow shower(s), moderate or heavy
95	Thunderstorm, slight or moderate
96	Thunderstorm, strong, wiht hail or gusts > 18 m/s or precipitation > 10 mm/h
99	Thunderstorm, heavy, with hail > 2 cm or gusts > 29 m/s or precipitation > 25 mm/h

The interpretation of model forecasts aims at the derivation of quantities that are not calculated by the model itself. The model information is translated into the WMO code of weather. The derived elements are given in Table 1. For the LMK there are a few more elements than for the LM, resulting from the list of warning criteria. The ww codes for strong and heavy thunderstorms (96 and 99) are not exactly the same as those from the WMO list but include the warning criteria for hail, gusts and precipitation. In the case of the LMK, which is a convection resolving model, the parameters produced by the convective parametrization scheme (convective precipitation, height of convective clouds, temperature at their upper limit) from the LM are missing and should be replaced by other promising parameters like graupel, radar reflectivity or maximum vertical wind velocity.
The weather interpretation is ready for pre-operational tests, but not tuned yet. Tuning is

set aside for the moment, because there is still a problem with the precipitation amount from the LMK in convective cases which is much too low compared to observations.

5 Conclusions and outlook

This study develops new methods to post-process the DMO of the LMK. Three aims are pursued: smoothing, probabilities and weather interpretation.
In terms of smoothing, a simple averaging over a 5×5 domain will be applied followed by a re-calibration of the distributions of the smoothed fields towards the distribution of the original field. Simple averaging atains roughly the same improvement as the NM with the same number of points in the neighbourhood.
In terms of probabilities, the BSS shows a significant improvement of the post-processed fields when compared to the deterministic DMO. The quality of the products improves with increasing size of the spatiotemporal neighbourhood. Only small improvements are achieved when compared to a 'climatological' forecast. However, this might be due to the 'climatology' which potentially leads to an unfair comparison.
In terms of weather interpretation, an operative version has been set up for the LMK.
Future work will deal with the fine-tuning of the weather interpretation and the derivation of probabilities for the full range of the above mentioned list of warning events (precipitation, gusts, thunderstorms, fog, etc.).

References

Renner, V., 2002: Interpretation. *Promet*, Jahrgang 28, Heft 1/2, Die neue Modellkette des DWD, pp 1-7.

Theis, S., Hense, A., Damrath U. and V. Renner, 2002: Ensemble Prediction and Statistical Postprocessing of Weather Parameters for the LM. *COSMO Newsletter*, No. 2, pp 152-161.

Theis, S., Hense, A., Damrath U. and V. Renner, 2003: Deriving Probabilistic Precipitation Forecasts from LM Simulations. *COSMO Newsletter*, No. 3, pp 209-220.

Theis, S.E., Hense, A. and U. Damrath, 2005: Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. *Meteorological Applications*, Vol. 12, No. 3, pp 257-268.

Simple Kalman Filter - A "Smoking Gun" of Shortages of Models?

Andrzej Mazur

Institute of Meteorology and Water Management, Centre of Numerical Weather Forecasts, 61 Podleśna str., PL-01673 Warsaw, Poland

1 Summary

The COSMO-LM model is currently running operationally at the Centre for Development of Numerical Weather Forecasts at IMWM, producing 72-hour forecasts of meteorological fields such as wind, precipitation intensity, cloud cover etc. Additionally it provides data for point locations (e.g., meteorological stations) in the form of meteograms. SHAWrt is a Simultaneous Heat And Water model (road temperature) dedicated for road temperature calculations for road maintenance during winter. Input data are model forecasts (temperature at 2m agl., wind at 10m agl, relative humidity, cloud cover, precipitation, i.e. rain and/or snow), vertical profile (contents of basic materials), site description (height, terrain configuration etc.) and time and date. Basic analysis of "raw" model results showed that they differ from point measurements. So, an application of additional procedure seemed to be necessary. As the beginning, simple Kalman filtering (Adaptive Regression method) was suggested. It seems to work quite good as far as "continuous" meteorological parameters, like temperature, wind speed or air pressure, are concerned.

2 Problem

Every forecast (even numerical forecast) comes with an error. Especially it can be seen when we are talking about point forecast (for instance, at meteorological stations). In this point location a quality of forecast may be easily verified. As an example, simple comparison between observed and predicted maximum and minimum temperatures for Warsaw station is shown in Figs. 1 and 2.

How can we handle this error? Kalman filtering seems to be an appropriate method (among others, of course). It can be used both for direct model results and for processed ones relatively easily. A basic scheme of filter (so called Adaptive Regression Method) is shown below.

$$
\begin{aligned}
y_k^f &= h_k^T b_{k-1}^a \\
P_k^f &= P_{k-1}^a + Q_{k-1} \\
e_k &= y_k^o y_k^f \\
w_k &= h_k^T P_k^f h_k + r_k \\
k_k &= P_k^f h_k w_k^{-1} \\
b_k^a &= b_{k-1}^T + k_k e_k \\
P_k^a &= -k_k w_k k_k^T + P_k^f
\end{aligned}
$$

where:
y - measurement vector
b - multiple regression coefficients (time dependent)
h - predictors – model forecast values
Q - error covariance
r - observational error
P - forecast covariance
e - forecast error
w - temporary scalar
k - Kalman gain

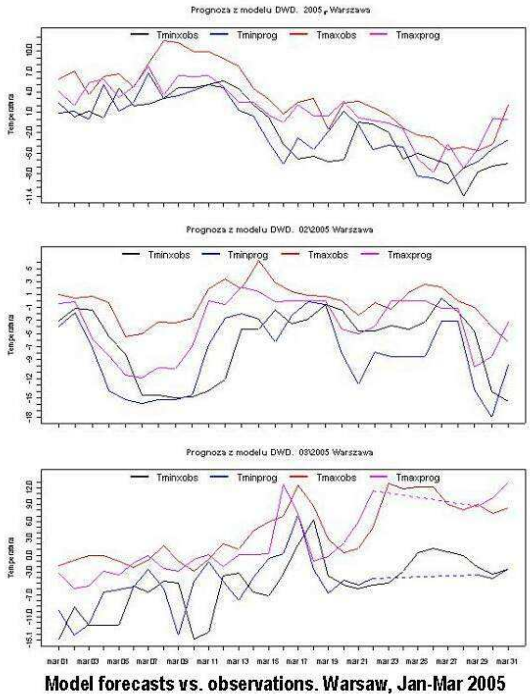


Figure 1: Model forecasts vs. observations (maximum and minimum temperature, Warsaw, Jan-Mar 2005).

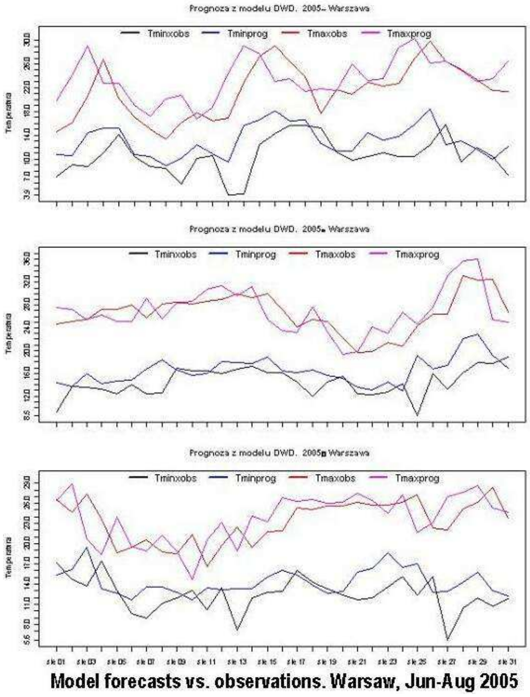


Figure 2: Model forecasts vs. observations (maximum and minimum temperature, Warsaw, Jun-Aug 2005).

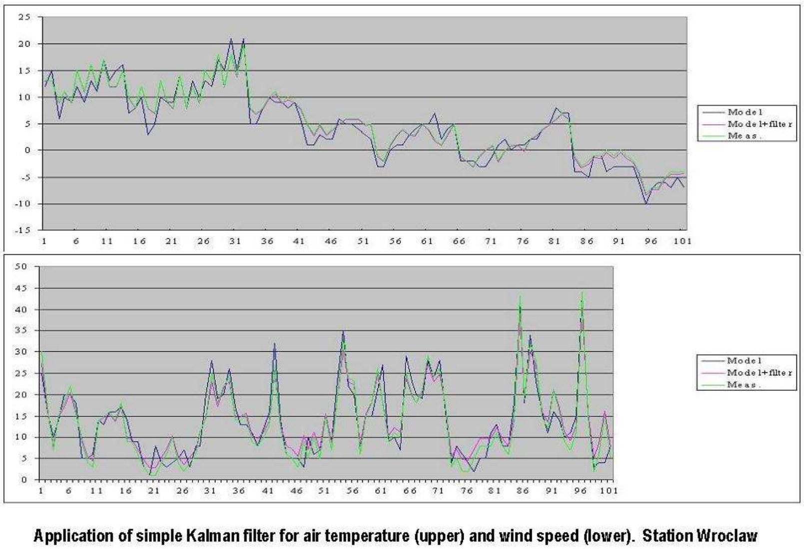
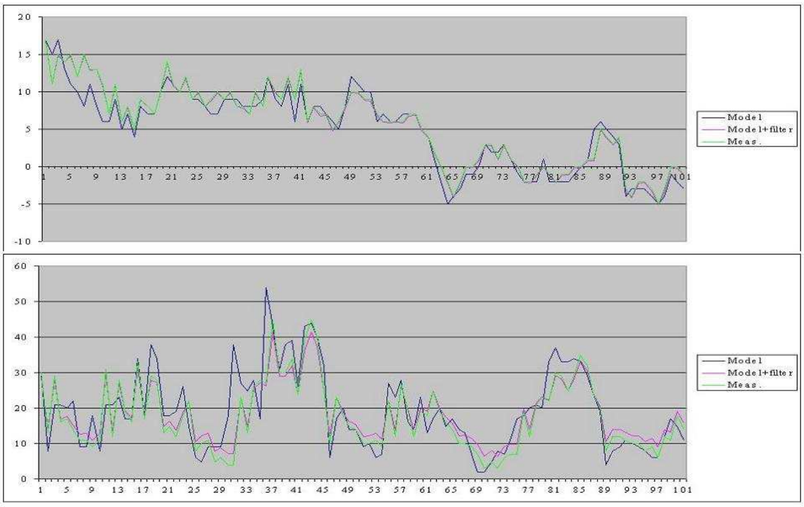


Figure 3: An adaptation of simple Kalman filter for air temperature and wind speed. Wroclaw, 2005.

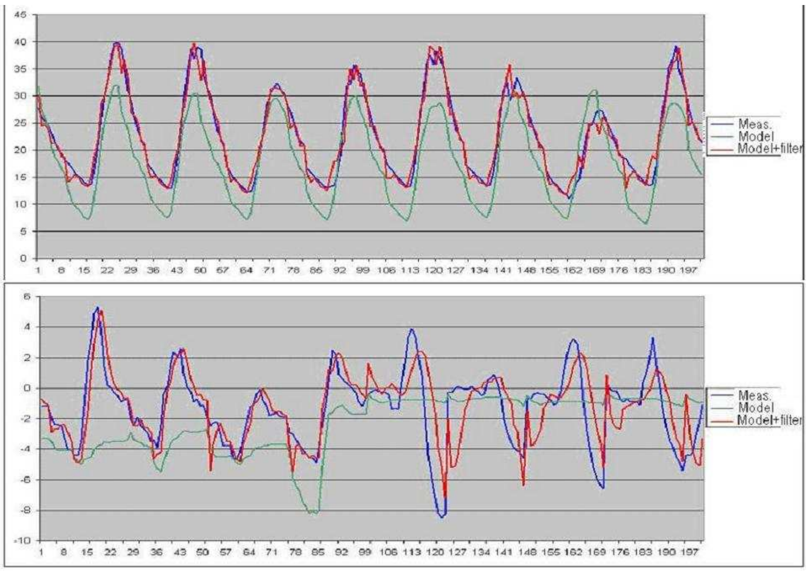
3 Results

In Figs. 3 and 4 results of this kind of filtering approach is presented. Figure 5 in turn shows the utilisation of Kalman filter for road temperature calculations (model SHAWrt) as an example of filtering approach to model results processed by other application. Interesting situation appear during winter season, while un-filtered model results did not take into account an appearance of snow cover (removed shortly afterwards by maintenance services).



Application of simple Kalman filter for air temperature (upper) and wind speed (lower). Station Warsaw.

Figure 4: An adaptation of simple Kalman filter for air temperature and wind speed. Warsaw, 2005.



Application of Kalman filter for road temperature assessment during summer (upper) and winter (lower).

Figure 5: Application of simple Kalman filter for road temperature assessment during summer and winter period.

This snow packet "worked" as a blanket keeping temperature more or less constant (green line in the figure). In reality, after removal of snow, the temperature of the road changed in a wide range (blue line). Filtered results were significantly closer to real observed ones.

4 Discussion and conclusions

Application of filter for "raw" (direct) model results have some characteristic features. First of all, it seems to work quite good as far as "continuous" meteorological parameters, like temperature, wind speed or air pressure, are concerned. Moreover, results seem to depend on differences between observations and "raw" results (i.e., BEFORE filter is applied). In other words, the greater difference - the better result. Other parameters, like precipitation, should be studied in a similar way. They might require different approach due to their different "nature". In both cases, careful selection of predictors is strongly advised. The method - even in this simple approach - can "detect" not only any factor "aside" of the model, but also systematic errors in results.

References

Ott, E., Hunt, B., Szunyogh, I., Zimi, A.V., Kostelich, E.J., Corazza, M., Kalnay, E., Patil, D.J. and Yorke, J.A.: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, 56A (2004).

Sorenson, H.W.: Least-squares estimation: from Gauss to Kalman. *IEEE Spectrum 7*, 1970.

Kalnay, E.: Numerical weather forecasting and predictability. Cambridge University Press, 2001.

5 Working Group on Verification and Case Studies

1 Group Activities

The activities of this group focus on an administrative point of view, in order to have an objective measure of how well LM forecasts are performing, and on a scientific one, in order to have a detailed assessment of the strengths and weaknesses of the model. Thus at the moment the main activities of the working group deal with the following issues:

- Verification of operational model forecast.
- Verification with feedback on the physics parameterization (verification of new predefined LM versions on predefined set of test cases).
- The development of new verification methods and diagnostic tools.
- LM case studies and collection.

The Working Group was coordinated from September 2004 to September 2005 by Patrizio Emiliani (CNMCA-Italy), and now by Adriano Raspanti (CNMCA-Italy). The main activities for the period Oct 2004 - Sep 2005 covered the following points:

- Operational verification of surface parameters, using Synop stations and also regional high resolution networks. Results are summarised in verification reports which are distributed on quarterly basis on the COSMO web site.
- Operational verification of upper-air parameters, using TEMP stations, with results summarised again in quarterly reports distributed on the COSMO web site.
- Exchange of LM maps (24hrs cumulated precipitation, and MSLP), of each operational LM running, on the COSMO web site.
- High resolution verification of precipitation, using available high resolution dense non-GTS surface data. Consolidation of a common high resolution data set of non-GTS daily precipitation.
- Daily cloudiness verification at 12 UTC with the Meteosat VIS channel.
- Verification of integrated water vapour content using GPS data.
- Validation of near-surface boundary layer processes and radiation budget from operational weather prediction runs (LM, GME, aLMo) at selected observatory measurement sites with different land surface properties.
- Weather regime type verification of vertical profiles and precipitation using Radar composite network.
- Verification of precipitation forecast using radar composite network.
- Realization of a common Verification Package.
- Verification of runoff over river basins.

A WG3/WG5 workshop was held on March 2005 in Langen. Besides usual presentations of recent developments and results and the status of the common verification package at ECMWF, the main topic of this joint workshop was dedicated to the problem concerning the needs of so-called *Conditional Verification*. In particular, it has been clarified that this is the best verification technique that can be used for the peculiar needs of WG3, in order to find and solve problems or to optimize the variables parameterization. For this purpose WG3 provided, around mid 2005, a list of criteria to use for conditional verification and a draft of this new project was also discussed at the second WG Coordinator meeting held in Bologna (September 2005).

Finally, during the COSMO general meeting 2005 held in Zürich, the new WG5 Priority Project *Conditional Verification Tool* was presented to the COSMO community and approved by the Steering Committee.

Further, the plan for 2005-2006 includes verification on high resolution verification for precipitation, weather regime verification of surface and upper-air data, verification of near surface boundary layer processes (operational at MeteoSwiss), complete delivery of the Common Verification Suite package and a WG 5 workshop on March 2006 in Langen, pointed on Conditional Verification Tool.

2 Results and Methods of Model Verification

The operational verification results for the LM forecasts at various COSMO meteorological centres, both for near-surface and upper-air parameters, are summarized in this section. More detailed verification results are presented on a quarterly basis at the COSMO web-site.

Also included are contributions related to the development and test of new methods of model verification, including the use of high-resolution non-GTS data and weather situation-dependent verification using radar composite data. Most of the papers are write-ups from the COSMO annual meeting 2005 in Zürich.

Of course, thanks to all of you who provided contributions for the present issue of the Newsletter. The numbering of equations and figures in this section refers to each paper.

Before continuing with the contributions, it can be useful to summarize briefly some conclusions on model deficiencies from the recent verification results as well as from diagnostic evaluations and from case studies.

Model Deficiencies

From the verification results for 2004-2005, we can summarize some basic problems:

- The cloud cover cycle is not well reproduced, generally the high cloud cover is overestimated, with a different behaviour in summer depending on the area and the thresholds; it is better in winter.
- The mean daily cycle of precipitation is not well represented, generally overestimated, with a signal increase with terrain height.
- During summer the model shows a strong phase shift of precipitation maxima in the daily cycle (maxima occur too early in the forecasts), probably this is linked to the fact that also the diurnal cycle of 2m-temperature still shows too rapid increase during the early morning and too rapid decrease in the afternoon

- Also low precipitation amounts appear to be overestimated by the model. Over regions with complex orography (especially over the Alps and Appenini), the precipitation patterns are still not very satisfactory and show great precipitation amounts upwind and an underestimation of rainfall downwind (reported also by forecasters in their subjective evaluations).
- During evening and night-time, the 2m-temperature has a quite large cold bias, especially during winter, while in summer it shows a low absolute accuracy (strong positive mae) around midday.
- The diurnal cycle phase of the 2m-dewpoint-temperature is relatively well captured by the model even if the diurnal wave amplitude is not so good, it is clear an overestimation.
- 10-m winds generally appear to be underestimated on mountain stations and overestimated, especially during the night, for low stations as well as there is a constant overestimation of wind gusts (also the small ones).
- The temperature vertical profiles (as verified with TEMP soundings) show a cold bias in the boundary layer during summer season and in the whole atmosphere during winter. Small positive above 500 hPa in summer.

There are some interesting first results in verifications of LMK test runs:

- It shows an enhanced forecast accuracy for gusts probably due to finer grid resolution.
- For precipitation the results of the smaller scale model lead to partly better verification results than the LM routine forecast.
- In general it leads to higher accuracy for some parameters, but other parameters may be affected negatively.

During the last COSMO meeting, held in Zürich in September 2005, the new concept of *Priority Projects* has been introduced. In this frame, WG5 people have proposed a sort of extension of the Common Verification Suite into a new Priority Project named *Conditional Verification Tool*. This will be the main topic of discussion and the main work duty of WG5 for the next 2-3 years.

Verification of aLMo in the Year 2005

M. Arpagaus, P. Kaufmann, G. de Morsier, D. Ruffieux, F. Schubiger, A. Walser and E. Zala

MeteoSwiss, Zürich, Switzerland

1 Operational Verification

1.1 Verification with European SYNOP (WP 5.1.1 / 5.1.7) [Pirmin Kaufmann]

The operational verification of the Alpine Model (aLMo) has been extended to include the visualization of categorical scores. A number of categorical scores are calculated for total cloud cover (CLCT) and precipitation (TOT_PREC) as reported by Kaufmann (2005). Here, these newly added plots, which are now available back to 2001 on the COSMO verification pages, are presented. While the cloud cover verification is done with 6-hourly SYNOPs like the other parameters, the precipitation is currently verified with 12-hourly sums. The model lead time used for the verification is 42 hours (i.e. the 30 h to 42 h sum) and 66 hours (i.e. the 54 h to 66 h sum). The following results are valid for 42 hours lead time.

The thresholds 30% (2.5 octa) and 80% (6.5 octa) are used for the categorical statistics of total cloud cover. In summer, the occurrence of over 30% total cloud cover is generally underestimated (not shown). The underestimation is considerably higher towards the south of the model domain. In winter, the agreement is much better. Only over the Alps, there is a clear positive frequency bias with up to 60% overestimation of the occurrences of partial cloud coverage.

The frequency bias is considerably higher with the threshold of 80% cloud coverage. In summer, there is usually an overestimation over the Mediterranean region and an underestimation over central Europe (Fig. 1). The variability of the bias from station to station is relatively high even over flat parts of the domain, which is an indication of a possibly too large station-to-station variability in the observations.

The summer 2003 was exceptional with the prevailing sunny and hot weather. The frequency of occurrence of over 80% cloud coverage is overestimated by aLMo by a factor that reaches up to 5 at some stations, especially over the Mediterranean. This large factor is however at least partially due to the very low observed frequency.

In winter, the frequency bias of the 80% cloud coverage threshold is again larger than for the 30% threshold, but smaller than for the 80% threshold in summer. The spatial structure is somewhat similar as in summer, with underestimation over central Europe and overestimation over the Mediterranean. Winter 2004/2005 is shown as a typical example (Fig. 1). A particularly well forecasted season was the winter 2003/2004.

The categorical verification for the 12-hourly precipitations sums uses thresholds of 0.1, 1 and 10 mm. Other thresholds (2, 5, 20, 30, 50 mm) are also calculated but currently not visualized. The frequency bias for the occurrence of precipitation (threshold 0.1 mm) is greater than one at most stations and for all seasons. The spatial structure and extent however vary. In summer, the frequency bias is around 1.5 over the northern half of the domain, but much larger over the regions south of the Alps and the Pyrenees (Fig. 2). In winter, the overestimation is considerably larger over the northern half of the model

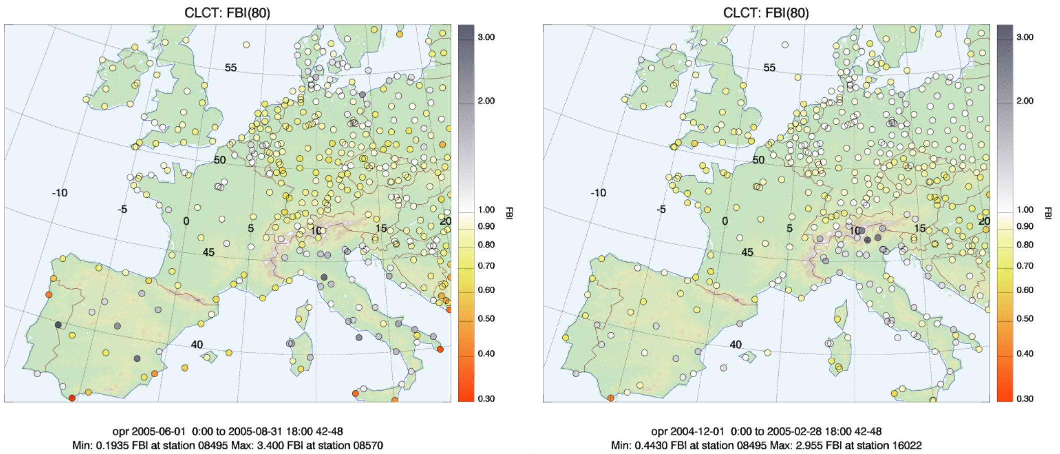


Figure 1: Frequency bias of cloud coverage above 80% for (a) summer 2005 and (b) winter 2004/05, lead times 42 h and 48 h.

domain (Fig. 2). Especially the regions with complex terrain show large factors (up to 5) of overestimation.

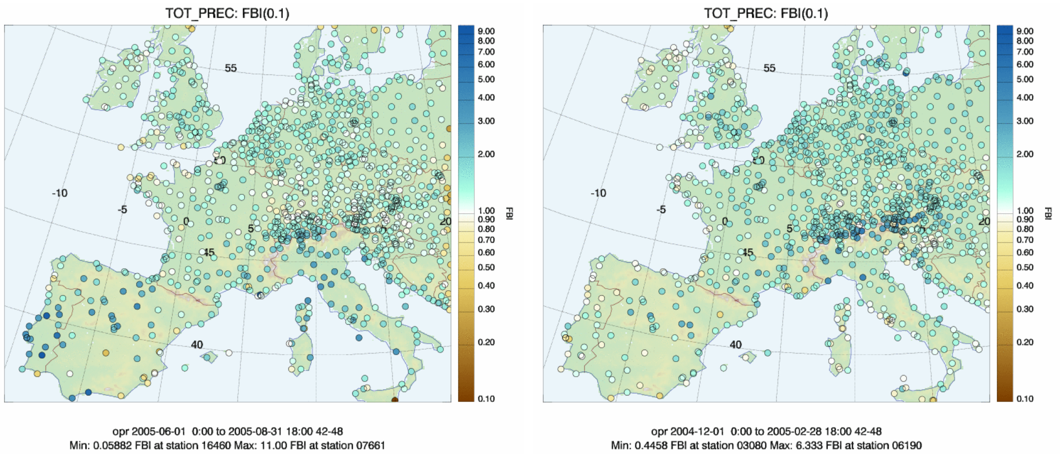


Figure 2: Frequency bias of 12 hourly precipitation sums for (a) summer 2005 and b) winter 2004/05, lead times 30 42 h.

The large model bias leads to a considerable false alarm rate around 40-50% north of the Alps. It is even higher south of the Alps and over the northern half of the Iberian Peninsula, where the observed frequency is very low. The observed frequency is equally low over the southern half of the Iberian Peninsula, but despite this, the false alarm rate here is lower than over the northern half of the peninsula and at the same level as over northern Europe.

1.2 Verification of daily cycle over Switzerland (WP 5.1.1 / 5.1.7) [Francis Schubiger]

No new developments in this package have been done this year and the main results of the verification of aLMo and LM inclusive Winter 2004/05 have been published by Schubiger (2005). The most important feature that should be investigated (by WG3) concerns the cloud cover in case of convection. Results for summer over the Alps suggest that the cloud amount in convective situations is too low (Fig. 3). While the observed cloud coverage shows a clear diurnal cycle (lower panel) in accordance with the observed precipitation (upper panel), the diurnal variation is absent in the model cloud coverage.

The results for precipitation are summarized in Table 1 with the scores for the frequency bias of the five seasons (from Summer 2004 to Summer 2005) for the thresholds 0.1, 2, 10 and 30 mm/6h for aLMo and LM. It shows an overestimation for low amounts (0.1 mm/6h) of 20-40% in summer and up to 60-80% in winter: this overestimation is most pronounced in the Prealps (altitude range 800-1500m). The high amounts (10 mm/6h) show a tendency towards higher values since begin of 2004 (each season of 2004 compared to the corresponding season in 2005): the underestimation of 15-25% in Summer 2004 changed into a slight overestimation of 5-15% in Summer 2005.

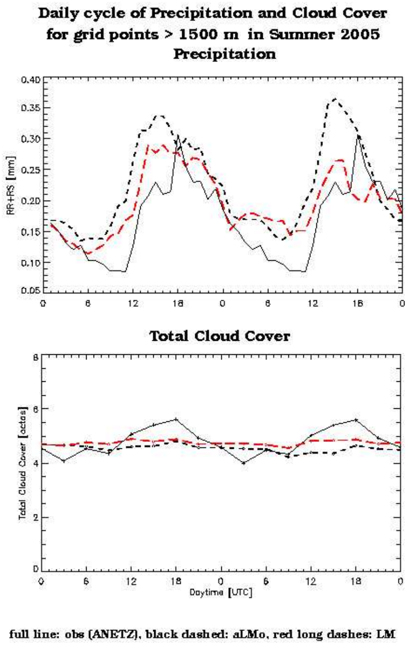


Figure 3: Verification of the daily cycle of precipitation (upper part) and total cloud cover (lower part) for grid points > 1500m over Switzerland in Summer 2005. Observations (ANETZ): full line black; aLMo: black dashed; LM: red long dashed.

1.3 Verification of the vertical profiles at TEMP stations (WP 5.2.2 / 5.2.6) [André Walser]

In Winter 2004/2005 the mean error for temperature shows a clear cold bias of up to 0.6 K between 1000 hPa and 300 hPa, which is increasing with forecast time. Above the tropopause height, an unprecedented cold bias of up to 2 K at 50 hPa can be observed, also increasing with forecast time (Fig. 4). Although this cold bias in the stratosphere is known from other seasons, its magnitude is considerably larger than previously seen. This negative temperature bias is at least partly due to the IFS boundary conditions. ECMWF confirmed to see a larger negative bias in the IFS forecasts for the northern hemisphere than a year or two before (but it was worse earlier). It is a model generated drift and as such sensitive to many aspects of the model formulation including physics, numerics and vertical resolution.

2 Verification studies

2.1 Weather situation dependent verification of upper-air data (WP 5.3.3) [Marco Arpagaus]

A weather situation dependent verification of the vertical structure of the Alpine Model (aLMo) based on the Schüepp classification (Schüepp 1979) is performed over the full dataset since the beginning of the climatic year 2004. The most interesting results are:

- Significant differences between classes high, flat, and low for temperature (1000 - 700 hPa) and relative humidity (900 - 700 hPa)
- Distinct differences between advective classes, especially west and east, for wind direction and wind speed.

Table 1: Frequency bias (%) of predicted precipitation over Switzerland. For all 6h-sums from +6h until +48h of all 00 UTC- and 12 UTC-forecasts, compared to 69 ANETZ stations. The LM and aLMo precipitation is the mean over 5 grid points. For the high amounts (10 and 30 mm/6h) the percentage of occurrences (%) is given. The columns give the values for the seasons since Summer 2004.

Threshold	Summer 04	Autumn 04	Winter 04/05	Spring 05	Summer 05
0.1 mm/6h					
aLMo	120	135	162	136	136
LM	121	145	183	143	129
2 mm/6h					
aLMo	101	109	162	122	113
LM	97	112	219	136	102
10 mm/6h					
aLMo	84	106	91	113	118
LM	76	94	231	140	102
30 mm/6h					
aLMo	114	268	0	51	216
LM	82	166	7	85	154

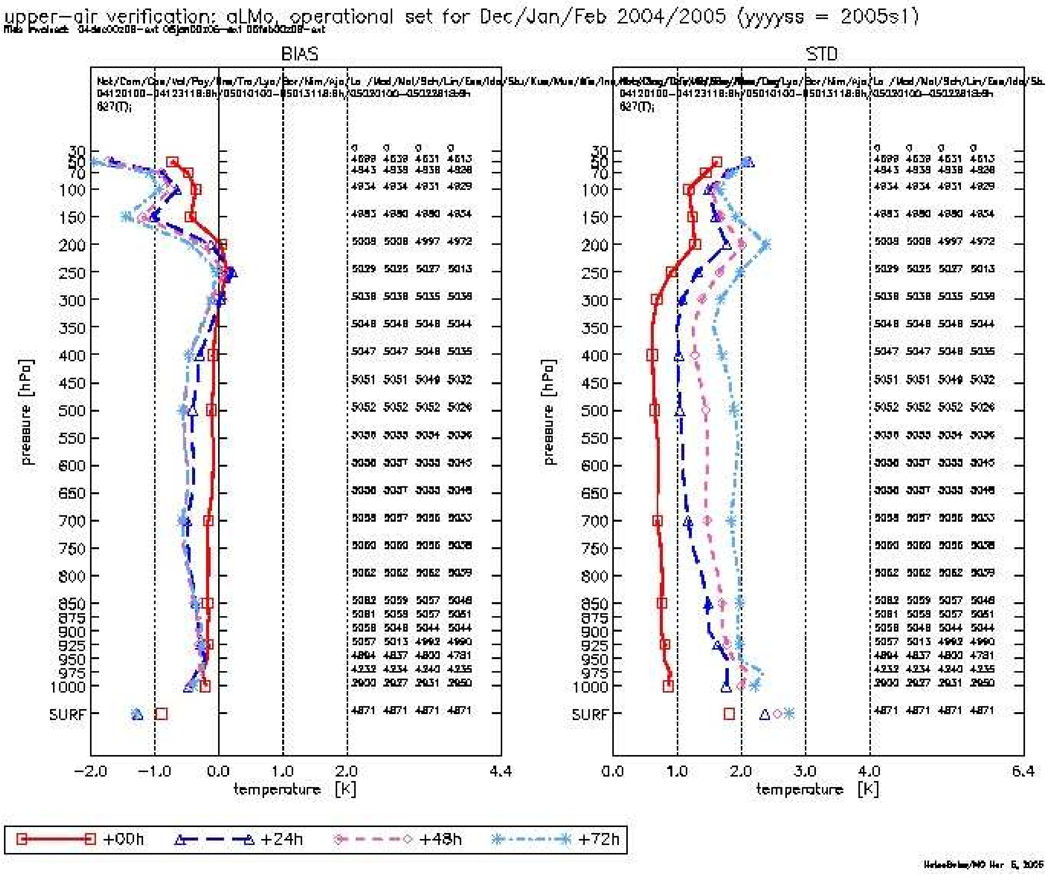


Figure 4: Vertical profile of temperature bias (left) and standard deviation (right) for winter 2004/05.

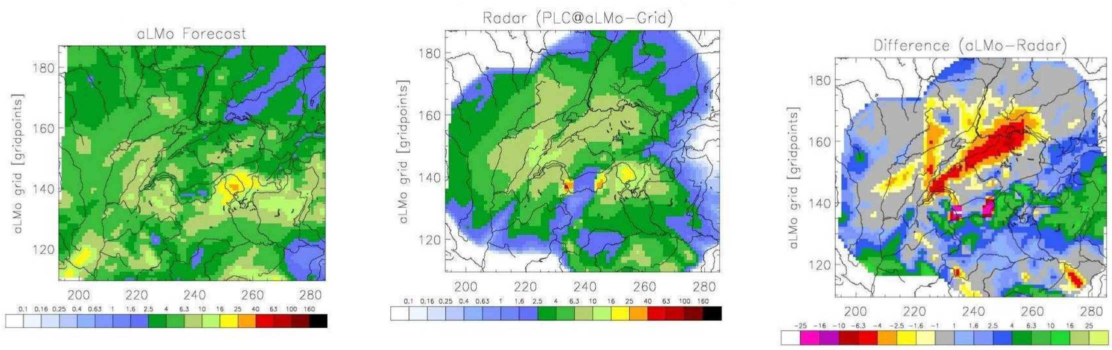


Figure 5: Precipitation sum for the southwest weather situation, (a) aLMo forecast, (b) Swiss radar composite, (c) model bias.

Large regional differences can be observed (e.g., wind speed for all stations and for Alpine stations only, respectively), as can many other interesting features.

2.2 Verification of aLMo precipitation forecast using radar composite network (WP 5.4.2) [Emanuele Zala]

A weather situation-dependent verification of aLMo precipitation based on Swiss radar composite data was performed over the climatic year 2004. Two weather classification were used: the Schüep classification (Schüep 1979), which is used daily by MeteoSwiss forecast office, and a simple experimental classification based mainly on 500 hPa winds and surface pressure distribution over the alpine region. Both classifications deliver similar results, the signal however is clearer in the experimental classification.

Main results:

- significant differences of aLMo QPF for different weather classes
- confirmation of the aLMo QPF overestimation over the relief, but only in situations with advection (Fig. 5)
- significant underestimation of precipitations over Swiss plateau, specifically in SW regimes (Fig. 5)
- generally good performance in situations with weak advection.

2.3 Verification with wind profilers [Dominique Ruffieux]

Wind speed and wind direction are verified on a daily basis for Payerne with wind profiler and radio soundings (Fig. 6). For Kloten, profiles have been verified with a wind profiler from autumn 2004 to Spring 2005.

3 Verification of different test suites

3.1 Hourly boundary conditions [Guy de Morsier]

During the Olympic and Paralympic Games of summer 2004 (from August 4 to September 30), ECMWF increased the update frequency of the LBCs from 3 to 1 hour. The largest impact of this higher frequency was found in the upper-air verification of the geopotential field (Fig. 7). Although some impact was also present in the temperature, wind speed and slightly in the wind direction after a forecast time of 24h, there was no signal in the humidity

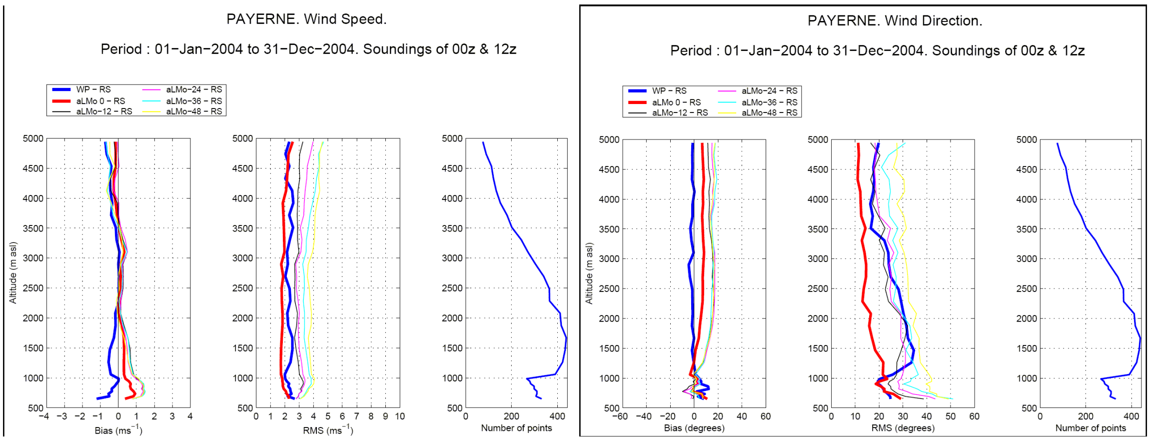


Figure 6: Verification against profiler data of (a) wind speed and (b) wind direction.

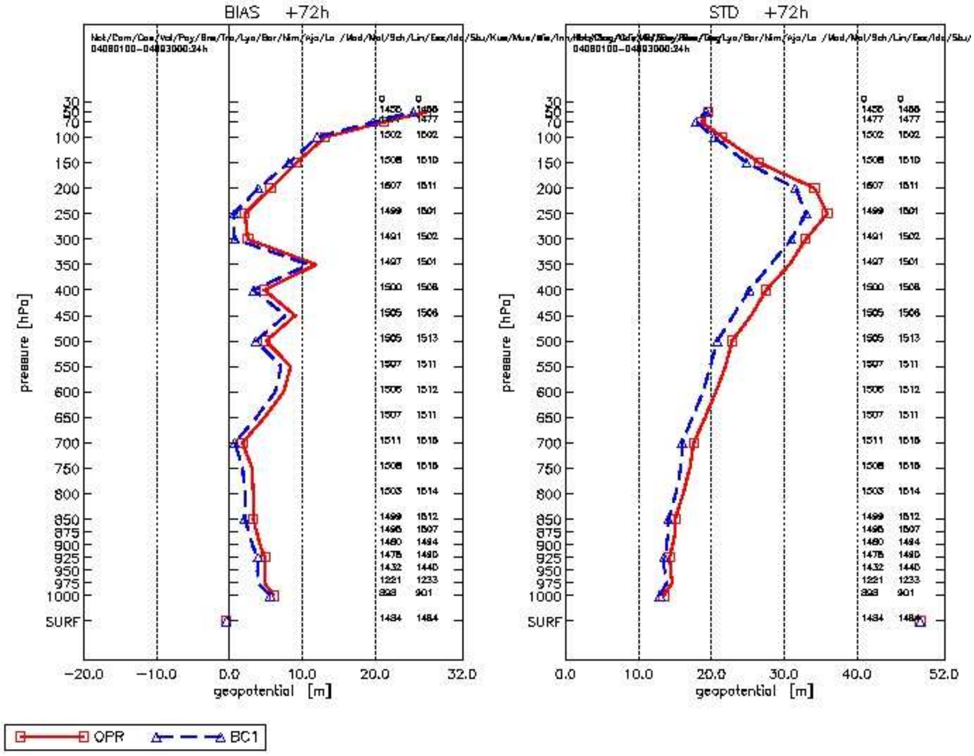


Figure 7: Vertical profile of geopotential height bias (left) and standard deviation (right) for winter 2004/5.

fields. The success of this first test suite with hourly LBCs was quite remarkable for a summer period.

Last winter (from January 5 to February 7, 2005) a second test phase was started at ECMWF. MeteoSwiss took this opportunity to run another parallel test suite of aLMo with hourly LBCs and to compare it with the production suite. This time the positive impact on the geopotential field could only be found for 6 TEMP stations west of the Alps and only for the standard deviation of the errors. This result was somewhat disappointing, as we expected the most positive impact of the hourly BC during strong winter synoptic developments. A problem could be the inadequacy of our local verification based on low frequency observations (SYNOPs each 6h and TEMPs each 12h).

A case study verification for the strongest storm during this winter period was then performed: It was the storm called Erwin (January 8, 2005) which caused a lot of damage in the UK, Denmark, Germany and Sweden. The forecasts valid at January 8, 2005 12UTC have been verified with SYNOPs and TEMPs. It was the time of the maximum development of the storm on the northern part of the domain. As opposed to our expectation, the results showed a greater standard deviation of the errors in the vertical structure of the geopotential and in the surface pressure for the forecasts with the hourly BC as compared to those with 3 hourly BCs. This suggests that a problem could also be the inadequacy of our relaxation scheme to cope with such high temporal non-linear BCs. The forecast with hourly BCs exhibits a greater bias over the northern part of the domain, although in the region with more than 12 hPa differences between forecast and analysis, the isolines in the hourly BC forecast are smoother and thus more realistic.

3.2 New LM-versions [Guy de Morsier]

In 2005 the following aLMo test suites have been verified with SYNOPs, hourly data from ANETZ and TEMPs (the verification has been documented on internal web pages that could be made available on request:

- LM version 3.15 with/without Rayleigh damping in the upper layers from 8-22/3/2005
- multilevel soil model (assimilation for March / April 2005; forecasts from 7-20/4/2005)
- lowest model at 10m and new vertical layering from 7-20/4/2005
- SLEVE (Smooth LEvel VErtical coordinates) for two periods (1-23/8/2004 and 7-20/4/2005)
- new divergence damping from A. Gassmann for two periods (1-23/8/2004 and 7-20/4/2005).

References

Kaufmann, P., 2005: Verification of aLMo with SYNOP and GPS over Europe. *COSMO Newsletter*, No. 5, pp. 113–117. available from http://www.cosmo-model.org.

Schubiger, F., 2005: High resolution verification of daily cycle over Switzerland. *COSMO Newsletter*, No. 5, pp. 88–93. available from http://www.cosmo-model.org.

Schüepp, M., 1979: Witterungsklimatologie. Beiheft zu den Annalen der Schweizerischen Meteorologischen Anstalt (Jahrgang 1978), MeteoSwiss, Zürich

Verification of aLMo Precipitation using a Coupled Hydro-Meteorological Modeling approach for the Alpine Tributaries of the Rhine

M. Verbunt[1], J. Gurtz[1], M. Zappa[2], P. Kaufmann[3]

[1]*Swiss Federal Institute of Technology (ETH), Zürich,*
[2]*Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf,*
[3]*MeteoSwiss, Zürich*

1 Introduction

The Alpine region is often struck by devastating flood events like the recent flooding in summer 2005. Due to its topography, this region is further vulnerable to secondary effects like flooding, landslides and erosion, which endanger environment, inhabitants and industries. A timely forecast of the events could help mitigate some of the possible consequences of extreme flooding. Recent developments in coupling hydrologic and atmospheric models show that these coupled approaches have, despite current limitations, a great potential in flood forecasting and impact assessments (e.g. Benoit et al., 2003). Especially in the case of runoff prediction in alpine catchments, the use of surface observations for flood forecasting is limited due to the short response times to precipitation events, which requires precipitation data to be available ahead of time.

The coupled atmospheric-hydrologic applications offer the possibility of verification that will be very important in the improvement of technologies associated with atmospheric models and water resources management. Jasper and Kaufmann (2003) used a coupled atmospheric-hydrologic system as a validation tool for Swiss Model (SM) forecasts in southern Switzerland. The present paper validates precipitation forecasts of the Alpine Model (aLMo) over a period of 2 years (2001-2002) in the upper stretch of the Rhine river basin.

2 Results

PREVAH is a hydrological model with a fine spatial resolution, including the simulation of glacier- and snow melt and the retention of lakes. It has proven its abilities in simulating hydrological processes in several Swiss catchments (Gurtz et al., 1999; Zappa et al., 2003). It has been applied to the Rhine basin down to the gauge Rheinfelden (34,550 km^2) with a spatial resolution of 500 by 500 m^2 using an hourly time-step. The spatial discretization of the hydrological PREVAH model (Gurz et al. 1999) relies on the aggregation of gridded spatial information into so-called hydrologic response units (Ross et al. 1979). The model contains modules to calculate evapotranspiration, snow- and glacier melt, and soil moisture.

The hydrological model has been calibrated with the use of hourly ground observations (precipitation, air temperature, wind speed, global radiation, sunshine duration and vapor pressure) during the period 1997 - 1998 after being initialized during 1996. A set of 656 stations (52 with hourly resolution, 56 which measure twice a day, and 548 rain gauges with daily resolution) were used. The calibration has been carried out separately for each of the 23 sub-catchments. Because of the presence of lakes in the investigated catchments, it was necessary to include a function which represents the retarding and flattening of flood peaks by lakes. Results show (Verbunt et al. 2006) that the model is capable to reproduce the relevant hydrological processes in the investigated catchments and that the model properly captures the extreme runoff peaks both during the calibration (1997 - 1998) and validation period (1999 - 2002).

The catchments in the upper part of the Rhine basin show a clear annual runoff cycle, caused by snow and glacier melt in spring and summer. Discharges at gauges in the lower part of the river basin are more influenced by lakes. The runoff in the Thur catchment in the northeast of Switzerland is mainly precipitation dominated, showing very strong fluctuations in runoff.

Precipitation forecasts of the atmospheric model are interpolated to the hydrological grid using a bilinear interpolation between the nodes. The verification used +19 h to +42 h aLMo forecasts for the period 2001 – 2002.

For the modeled runoff, the use of precipitation forecasts considerably increases the False Alarm Rate, while the Critical Success Index decreases. Consequences of errors in the precipitation forecasts are most pronounced for higher thresholds, while the coupled modeling system performs better for smaller precipitation events.

Fig. 1 shows the bias in annual runoff, calculated by subtracting the observed annual runoff from the simulated annual runoff amount for each investigated catchment. Although these deviations may vary over the years and show a high spatial heterogeneity, some general tendencies can be noticed.

The annual runoff in the upper Aare catchment is always overestimated, caused by an overestimation of precipitation by the NWP model in this region. This overestimation is especially evident in summer, due to too strong convection in the NWP model in mountainous areas. The model further overestimates precipitation on the northwest facing slopes. In contrast to the upper Aare region, runoff in the easternmost watershed of the Aare and the uppermost Rhine catchment is generally underestimated. This can partly be explained by the lack of advection of falling precipitation in the NWP model, which causes the precipitation in these areas to be underestimated, and which has recently been included in the aLMo with the prognostic precipitation scheme.

The heterogeneity in the runoff bias for each catchment (Fig. 1) causes the accumulated bias (not shown) to decrease in catchments further downstream. Deviations caused in upstream catchments however influence runoff volumes further downstream and are still noticeable at the lowest Rhine river gauge at Rheinfelden.

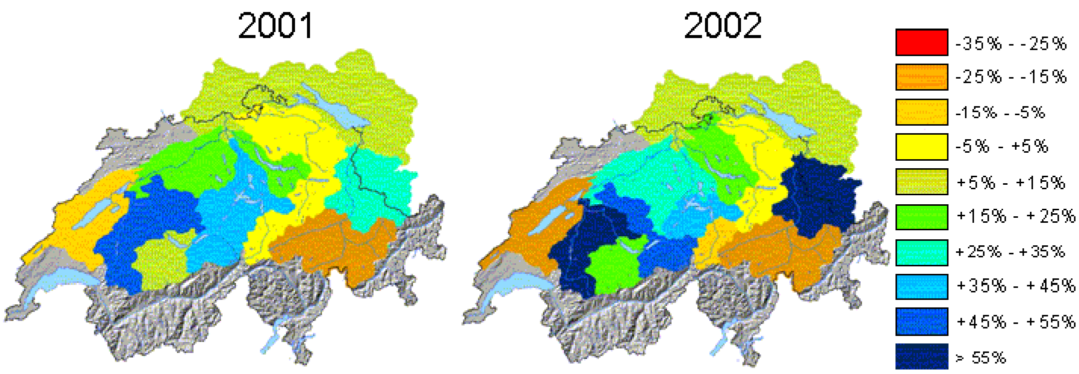


Figure 1: Annual bias of runoff simulated based on aLMo precipitation relative to observed runoff for the years 2001 (left) and 2002 (right).

3 Conclusions

These results agree with the findings of Kaufmann et al. (2003), who concluded for the SM that precipitation amount in the mountains above 1500 m is considerably overestimated in summer due to the too strong convection. This is still the case for the aLMo. It leads to clear overestimation of runoff peaks, especially during the summer season. These results also agree with those from Jasper and Kaufmann (2003), who showed that SM forecasts lead to high false alarm rates in runoff forecasts due to an overestimation of precipitation peaks in the Ticino-Verzasca-Maggia basin. Further downstream, the consequences of errors in the precipitation forecasts are still considerable but reduced mainly because of

- the retention capacity of larger lakes and
- by an underestimation of precipitation in the lower situated catchments.

The spatially distributed verification of coupled atmospheric-hydrologic model systems enables the detection of inaccuracies in numerical weather prediction models and is important to meet the ever increasing demands in operational forecasting.

References

Benoit R., N. Kouwen, W. Yu, S. Chamberland and P. Pellerin, 2003: Hydrometeorological aspects of the real-time ultrafinescale forecast support during the Special Observing Period of the MAP. *Hydrology and Earth System Sciences*, 7(6), pp. 877–889

Gurtz, J., A. Baltensweiler, and H. Lang, 1999: Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountainous basins. *Hydrological Processes*, 13, pp. 2751-2768

Jasper, K., and P. Kaufmann, 2003: Coupled runoff simulations as validation tools for atmospheric models at the regional scale. *Quarterly Journal of the Royal Meteorological Society*, 129(588), pp. 673–693

Kaufmann, P., F. Schubiger, and P. Binder, 2003: Precipitation forecasting by a mesoscale numerical weather prediction (NWP) model: eight years of experience. *Hydrology and Earth System Sciences*, 7(6), pp. 812–832

Ross B.B., D.N. Contractor, and V.O. Shanholtz, 1979: A Finite-Element Model of Overland and Channel Flow for assessing the hydrologic Impact of Land-Use Change. *Journal of Hydrology*, 41, pp. 11–30

Verbunt, M., M. Zappa, J. Gurtz, and P. Kaufmann, 2006: Verification of a coupled NWP-hydrological modeling approach for different catchments in the Rhine basin. *J. Hydrol.*, in press.

Zappa, M., F. Pos, U. Strasser, P. Warmerdam, and J. Gurtz, 2003: Seasonal water balance of an Alpine catchment as evaluated by different methods for spatially distributed snowmelt modeling. *Nordic hydrology*, 34(3), pp. 179-202

First Results on Verification of LMK Test Runs Basing on SYNOP Data

C.-J. Lenz, U. Damrath

Deutscher Wetterdienst, Offenbach am Main, Germany

1 Introduction

Within DWD's *Aktionsprogramm 2003* a main focus is on the development, the implementation, and the evaluation of a high–resolution model system for the very short forecast range of 3 to 24 hours basing on the non–hydrostatic regional model LM. The mesh size of this model system, which is called *Lokal Modell Kürzestfrist* (LMK), will be about 2 to 3 kilometers horizontally and 50 model layers in vertical direction. This numerical resolution in space and time allows a direct calculation of phenomena of the meso–γ–scale as larger convective cells and enables DWD a more precise forecast of strong thunderstorms, and hence an earlier warning of the authorities, the customers, and the population potentially affected by severe weather.

Considering mainly January and two summer months in the year 2004 different LMK test series have been performed with respect to numerical aspects, boundary values and relaxation methods, introduction of prognostic graupel calculation, as well as data assimilation. Accompanying the LMK test series a standard verification of the results following Nurmi (2003) is done permanently basing on hourly SYNOP data. Due to the progress in instrumentation and measuring methods concerning area distributed data about cloud coverage and precipitation rates, new possibilities for verification of weather forecast results arise. In the following an overview will be given about present verification results of the LMK as well as on future plans.

2 Model configuration

The model domain used for the LMK test runs extends about 10.5 degrees in longitudinal and 11.5 degrees in latitudinal direction with the center of the model domain near Offenbach/Main at 10 degrees E and 50 degrees N (see Fig. 1). The horizontal grid length used is 2.8 km and the number of grid points count 421 in longitudinal and 461 in latitudinal direction. In vertical direction the model domain is divided into 50 layers with a height of the lowest model half layer of about 44 m and the lowest main level about 22 m above ground. Initialization and hourly boundary values are taken from the operational LM runs.

In the following two tables (Table 1 and Table 2) the different LMK test suites performed for January 2004 and July-September 2004 and their different configurations are listed. The first experiment 673 represents a LM run done with the usual features of the operational LM in January 2004 except the differences in spatial and temporal resolution. The basic features for the summer test suites are similar to those mentioned for experiment 683 except the use of an explicitly formulated cosine function for the lateral boundary relaxation.

Table 1: Different LMK test suites and their configurations. Part 1.

experiment	time integration scheme, time step	horizontal advection scheme	calculation of precipitation	thermodynamics	relaxation at lateral boundaries
673	leapfrog $\Delta t = 16$ s	centered 2. order	diagnostic	advection of $T = T_0 + T*$	*tanh* function, implicit, as to LM
683	TVD–RK 3. order $\Delta t = 30$ s	upwind 5. order	prognostic	advection of $T = T_0 + T*$	*tanh* function, as in LM, precipitation diagnostic
687	TVD–RK 3. order $\Delta t = 30$ s	upwind 5. order	prognostic	advection of $T* = T - T_0$	*cos* function, explicit, precipitation diagnostic
688	TVD–RK 3. order $\Delta t = 30$ s	upwind 5. order	prognostic	advection of $T* = T - T_0$	*cos* function, explicit, precipitation diagnostic, relaxation of q_v, q_c, and q_i

Table 2: Different LMK test suites and their configurations. Part 2.

experiment	shallow convection parametr.	lateral boundary values for pressure	precip. scheme, nr.of classes	data assimilation	advection of q_x
689	no	interpolation	5 classes	no	Euler–forward
696	yes	vertical integration	5 classes	no	Euler–forward
698	yes	vertical integration	including graupel 6 classes	no	Euler–forward
701	yes	vertical integration	5 classes	yes	Euler–forward
709	yes	vertical integration	including graupel 6 classes	yes	Semi–Lagrange
713	yes	vertical integration	including graupel 6 classes	yes	Bott–2–Scheme

3 Results on the Verification Using SYNOP Data

The temporal development of the True Skill Statistics (TSS) graphs shown in Fig. 2 to Fig. 4 has been calculated using hourly SYNOP data from stations situated within the LMK – and hence in the LM – model domain. For the LM the assignment of the SYNOP stations to the model grid has been done via the nearest grid point approach. In the case of the LMK the same approach has been used, but additionally the surrounding 8 LMK grid points have been taken for an arithmetic averaging of the model results on a grid size comparable to the operational LM.

The TSS graphs for all precipitation rates shown in Fig. 2 show two groups of lines: the first group is represented by the LM routine run (black line) and LMK experiment 673 (red line), whereas the other lines form a second group with different characteristics. In

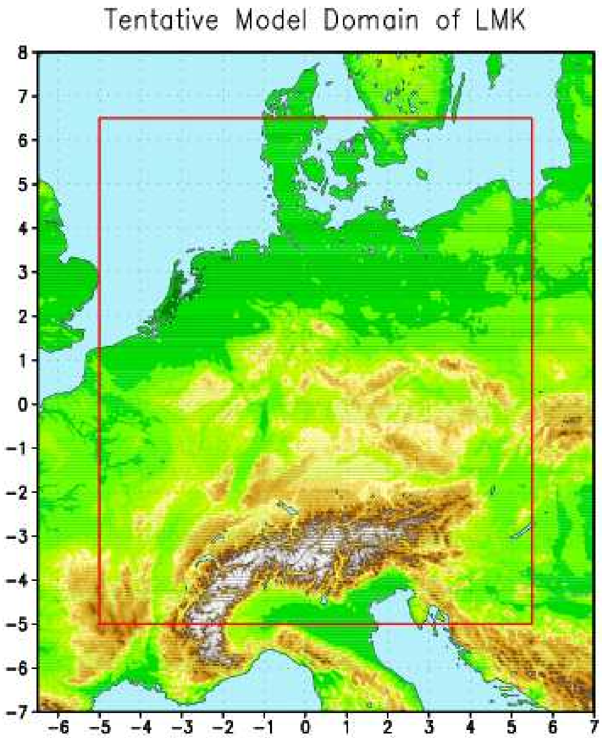


Figure 1: Integration area of the LMK model, topographical height [m].

the operational LM run as well as in LMK experiment 673 the precipitation is calculated diagnostically, whereas precipitation is computed via a prognostic equation in the other LM experiments shown. The TSS values for the diagnostic precipitation calculation are rather constant during the considered simulation time. In the prognostic case a strong spin–up effect can be seen in the TSS graphs during the first 2 to 6 hours of simulation. During daytime the TSS increases considerably to higher values than in the model runs with a diagnostic precipitation computation. This effect can be seen mostly remarkable in the first figure valid for precipitation rates > 0.1 mm/h.

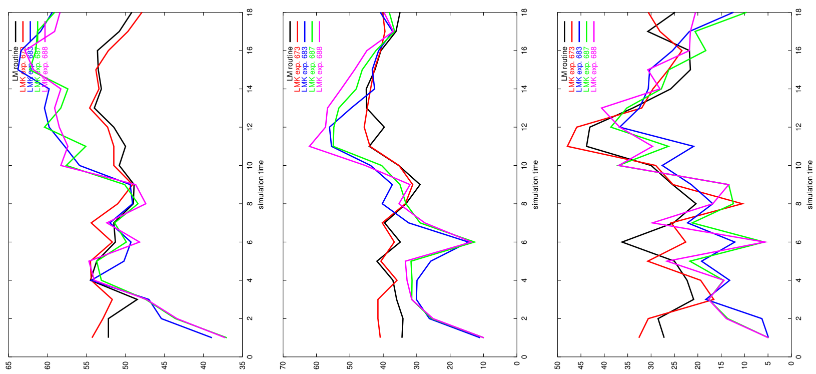


Figure 2: True Skill Statistics (TSS) for precipitation rates > 0.1 mm/h (top left), > 2 mm/h (top right) and > 10 mm/h (bottom). LM routine and different LMK test runs for January 2004, simulation start: 00 UTC.

Generally, the TSS for the July runs show much lower values as in the January simulations for the operational LM run as well as for all different LMK test runs (Fig. 3). The spin–up effect mentioned in the context of Fig. 2 can be seen again, but is significantly reduced in the

LMK test cases performed using data assimilation (LMK experiments 709 and 713). The precipitation calculated during these experiments results in much higher TSS values during the mainly convection governed afternoon hours considering medium and high precipitation rates. The TSS values decrease in all cases strongly during nighttime.

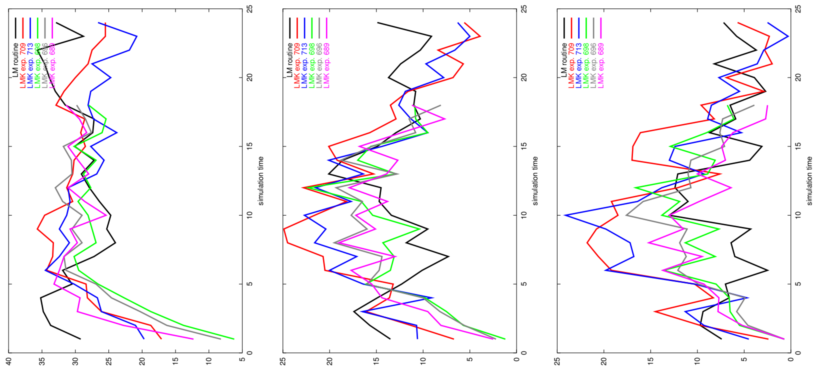


Figure 3: True Skill Statistics (TSS) for precipitation rates > 0.1 mm/h (top left), > 2 mm/h (top right) and > 10 mm/h (bottom). LM routine and different LMK test runs for July 2004, simulation start: 12 UTC.

Besides the LMK experiments described above, an additional LMK test run using latent heat nudging during the first 3 hours of simulation has been performed for the time period from 07 to 19 July 2004. For all considered precipitation rate classes the TSS resulting from the latent heat nudging shows a very strong increase during these nudging period (Fig. 4). The TSS values decrease after switching off the latent heat nudging for several hours until the typical TSS valued of non–nudged LMK experiments is approached. After 6 to 9 simulation hours the effect of the latent heat nudging seems to be very small and the TSS is rather similar to unnudged LMK runs.

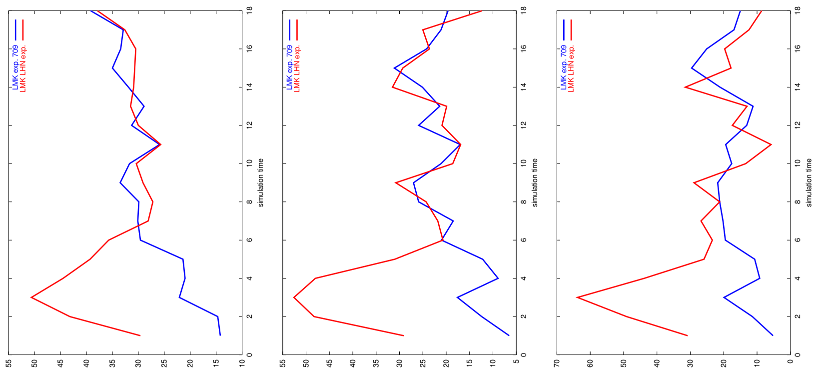


Figure 4: True Skill Statistics (TSS) for precipitation rates > 0.1 mm/h (top left), > 2 mm/h (top right) and > 10 mm/h (bottom). LMK experiment 709 and LMK experiment using latent heat nudging for 07th to 19th July 2004, simulation start: 12 UTC.

4 Future Plans

The verification of LMK test suites will be continued during the further development of LMK. The main focus of the verification will be on precipitation, but extends to other prognostic and diagnostic variables computed by LMK, e.g. temperature, pressure, cloudiness, wind vector and gusts etc. To allow a deeper investigation of the processes interacting in the model the development of a tool for a simple conditional verification of LMK results is ongoing. The conditional verification will include dependencies of two or more model variables, but no temporal or time–delayed interaction of variables. Results of the LMK model are used to calculate synthetic RADAR images for the RADAR sites of Germany according to the RADAR Simulation Model (RSM) of Haase (1998). Preliminary results of the application of the RSM on LMK output can be seen in Fig. 5: Entering the considered domain from the southwestern corner a simulated squall line crosses the domain of the RADAR site Hannover to northeastern direction within several hours. These synthetic RADAR images calculated from model results together with measured RADAR pictures will be taken for verification of precipitation patterns using pattern recognition methods as e.g. described by Ebert and McBride (2000).

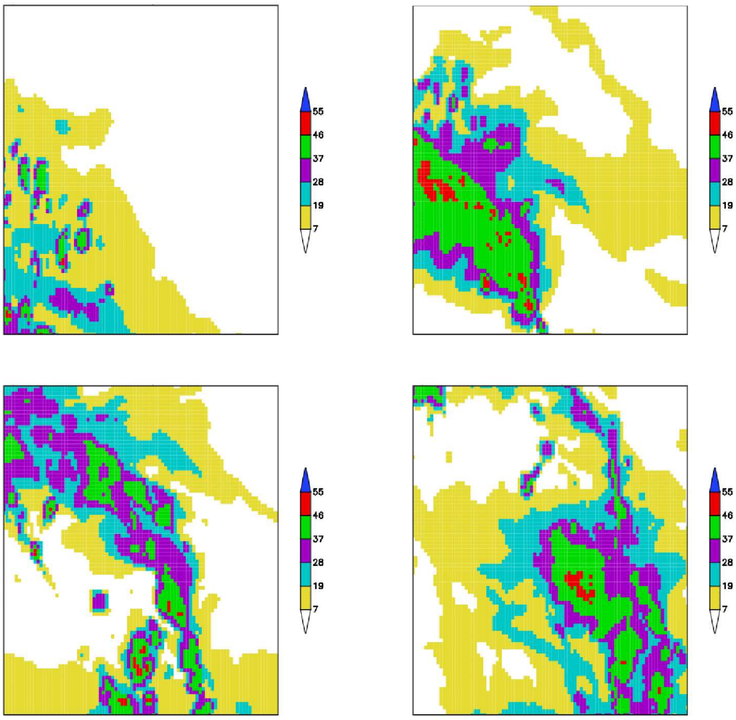


Figure 5: Results of the RADAR Simulation Model (RSM) for the LMK experiment 698 for 12th August 2004, simulation start: 12 UTC. Results after 4 h (top left), 6 h (top right), 8 h (bottom left), and 10 h (bottom right) of simulation time.

References

Ebert, E. E.; McBride, J. L., 2000: Verification of precipitation in weather systems: determination of systematic errors. *J. Hydrology*, 239, pp. 179–202.

Haase, G., 1998: Simulation von Radarmessungen mit Daten des Lokalmodells. Diploma Thesis, Meteorological Institute, University of Bonn, Germany, 84 pp.

Nurmi, P., 2003: Recommendations on the verification of local weather forecasts (at ECMWF member states). Consultancy Report, ECMWF Operational Department, http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/Rec_FIN_Oct.pdf.

Latest Results in the Precipitation Verification over Northern Italy

E. Oberto, M. Turco, P. Bertolotto

ARPA Piemonte, Torino, Italy

1 Introduction

The purpose of this work is to summarize the main results on precipitation verification during last year. More specifically, in order to highlight the different behaviour and performance of the three model versions (aLMo, LAMI, LM-DWD), we consider a common domain, composed by 47 meteo-hydrological basins of Northern Italy (about 1023 stations), as shown in Fig. 1. In addition we take into account also the Piedmont region, with high density, spatial homogeneity and high operational percentage of the observational network, for a special focus about LAMI study.

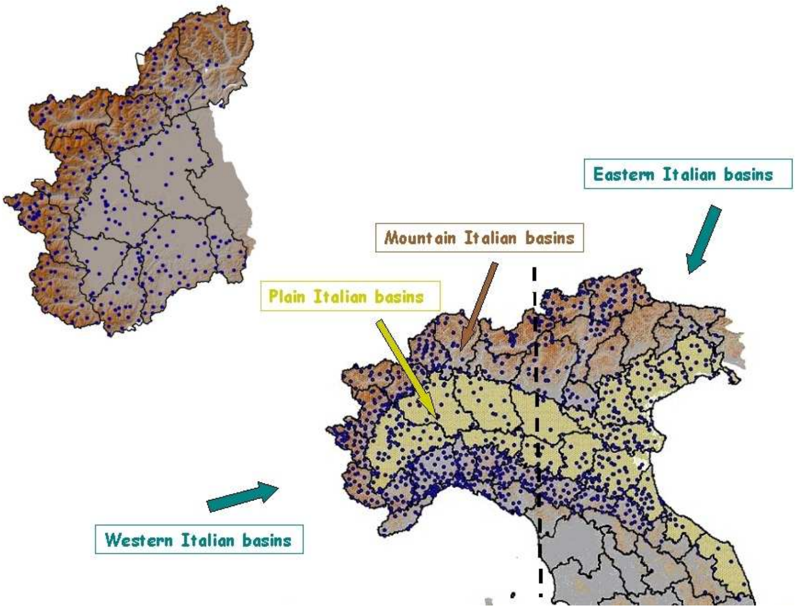


Figure 1: Right: Northern Italy raingauges and basins (30 mountainous areas in brown, 17 flat areas in yellow). Left: zoom with North-West warning areas and raingauges distribution.

2 The main purpose

In the fist part of this work we present the latest results concerning the precipitation verification over Northern Italy, pointing out the model error trend during the last seasons. We verify both runs (00UTC and 12UTC) over the latest six seasons (from December 2003 to May 2005) for each model version: we calculate skills and scores considering 24h cumulated precipitation averaged over basins for several increasing precipitation thresholds for +24h and +48h forecast time. Furthermore, in order to check if there is an error linked to the orography or to the synoptic, we subdivide the 47 meteo-hydrological basins into two big subsets:

mountain and plain areas, and western and eastern areas. We estimate therefore seasonal skills and scores considering 24h cumulated precipitation averaged over these different kinds of basins for several increasing precipitation thresholds, for +24h and +48h forecast time for both runs.

In the second part, we suggest a comparison over Piedmont between the standard version of LAMI and LAMI with the prognostic precipitation scheme, using a eyeball verification approach linked to the statistical categorical indices. In particular, we present seasonal observed/forecasted precipitation maps considering 24h cumulated precipitation every day for the first and the second 24h forecast time (D+1 and D+2), in order to obtain a visual comparison of the two different versions; in addition we also carry out the statical verification averaging 24h forecasted and observed rain over 13 warning areas (11 in Piedmont, Val d'Aosta and Ticino).

3 General results

The following three figures (Fig. 2, Fig. 3, Fig. 4) show the seasonal errors at a fixed threshold of 10 mm/24h over the whole domain, represented by the 47 basins, and besides, over mountain/plain basins (30 and 17 warning areas respectively) and over western/eastern basins (21 and 26 warning areas respectively). In this study we consider 00UTC run for each model version. The main remarks are:

- As we note in Fig. 2 there is a general increasing in overestimation trend during the latest seasons, with the greatest BIAS in JJA 2004 and DJF 2005.
- The role of soil moisture analysis during summer seasons seems to be decisive in term of QPF (better BIAS index for LM-DWD in JJA 2004 as seen in Fig. 2) but not in term of capability to localize and predict accurately the precipitation pattern (ETS very low as seen in Fig. 2).
- There are large differences in term of BIAS between mountainous and plain areas (see Fig. 3) and we obtain the greatest overestimation over the mountain.
- Concerning the error sensitivity with respect to western or eastern areas subdivision (see Fig. 4), aLMo shows a general increasing in QPF, but much greater overestimation on Western areas; LM-DWD shows a large worsening during the latest seasons but it has anyway a great overestimation over western areas; LAMI has similar behaviour with respect to the other version, but not for DJF 2005, in fact there is a large overestimation over the East where we have the greatest majority of precipitation cases (110 cases over the East and 40 over the West respectively).
- We find again a strong overestimation during the latest winter time probably due to the introduction of prognostic cloud-ice scheme (open problem). But how much does an halved statistic affect the results interpretation? How much more difficult is to estimate quantitatively the precipitation pattern during a particularly dry winter with respect to a normal winter? In fact, during the latest winter only few events occurred: about 150 events with precipitation $>$ 10mm/24h on average, in comparison with about 400 during DJF 2004.
- There is a remarkable LAMI performance during DJF 2005 (Fig. 3 and Fig. 4): only LAMI runs without prognostic precipitation scheme, so that the role of the prognostic precipitation scheme has been investigated during a very dry season (see results over Piedmont).

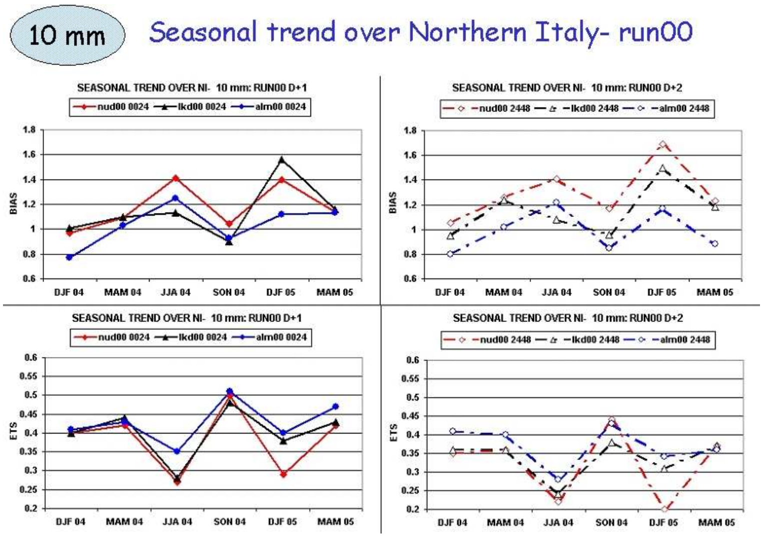


Figure 2: Seasonal indices over whole domain for each model version (00UTC run). On the left +24h forecast time, on the right +48h forecast time.

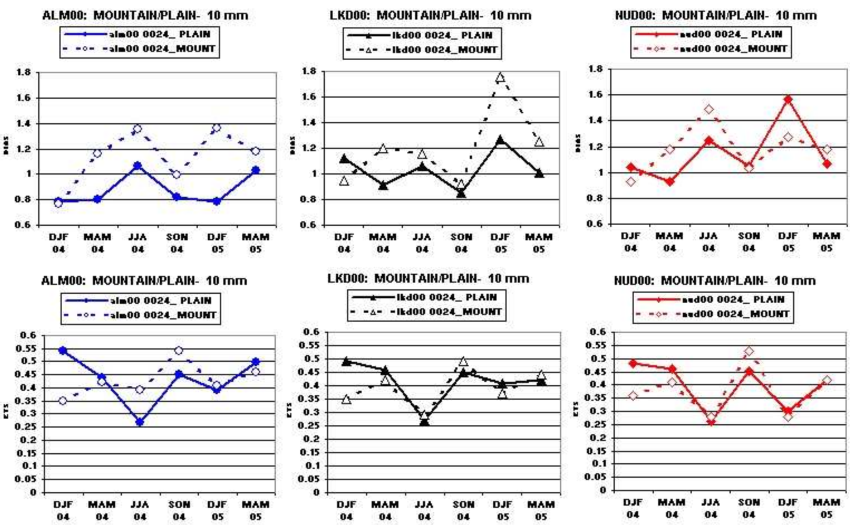


Figure 3: Seasonal indices over mountain/plain areas separately for each model version.

Finally we present a comparison between the standard version of LAMI and LAMI with the prognostic precipitation scheme, using Piedmont (11 basins), Val d'Aosta and Ticino on the period August 2004 - July 2005. In this case BIAS and ETS are obtained by averaging the 24h precipitation over the 13 mentioned warning areas: the error bars indicates 2.5th and 97.5th percentiles of resampled distribution, applied to the reference model (see Turco, 2005). BIAS and ETS calculated over the whole study period are shown in Fig. 5: there is a statistically significant reduction of the error for high thresholds that is evident especially for the wet seasons. In fact if we analyze separately the seasonal cumulated maps in Fig. 6, we note a different behaviour during the last dry winter time: for LAMI with prognostic precipitation we do not find the precipitation amount reduction as occurred in the other seasons and the precipitation pattern seems to be better described by the standard version of LAMI. This effect could be strictly linked to a weather type dependent verification: in fact during the latest winter a strong dry northwestern flux was the predominant type of weather and that could have caused an incorrect pattern prediction up/down the Alps for

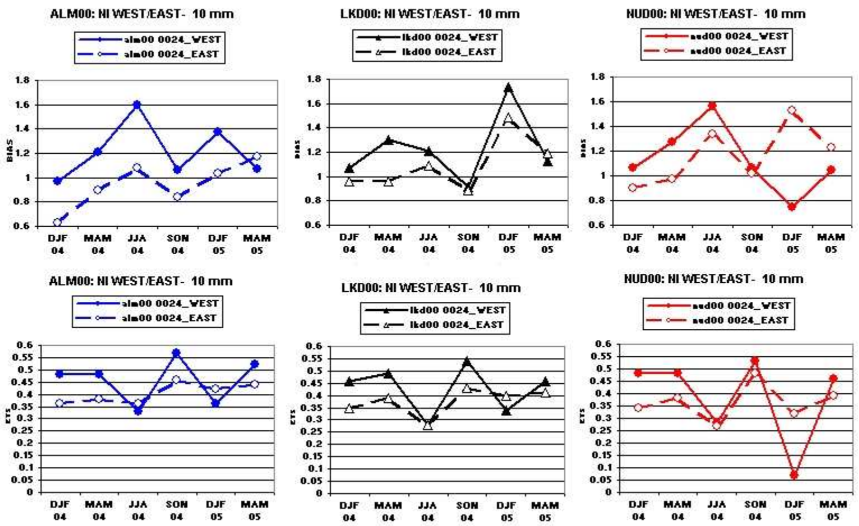


Figure 4: Seasonal indices over western/eastern areas separately for each model version.

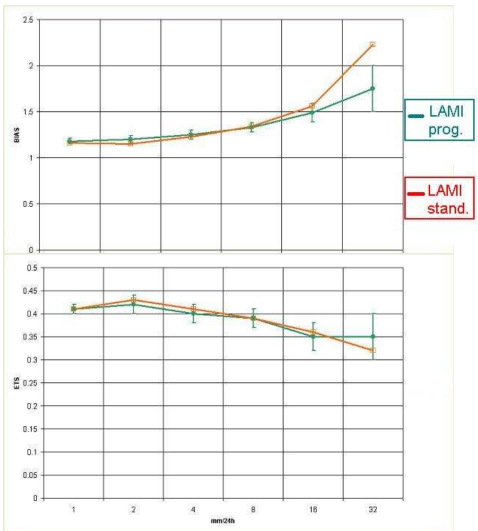


Figure 5: BIAS and ETS (Aug 2004 - Jul 2005) over Piedmont, Val d'Aosta and Ticino: standard version of LAMI versus LAMI with prognostic precipitation.

the prognostic precipitation version; in that case LAMI standard version performs better probably because it does not feel the effects of particle drift through the alpine obstacle.

References

Turco, M., Oberto, E. and Bertolotto, P., *COSMO Newsletter*, No. 5, 2005

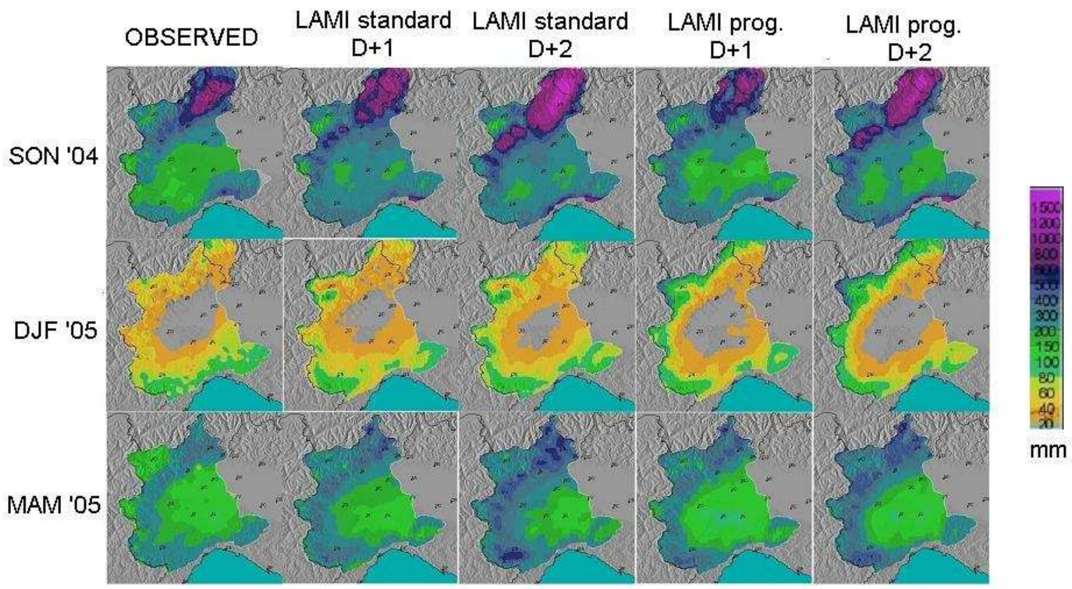


Figure 6: Seasonal maps: standard version of LAMI versus LAMI with prognostic precipitation, D+1 and D+2. For the last season (MAM 2005), the Ticino data have not been included because too few stations were present.

Verification of LAMI at SYNOP Stations

P. Emiliani, A. Raspanti

Ufficio Generale per la Meteorologia (UGM), Roma, Italy

1 Introduction

A synthesis of LAMI (the Italian version of LM) verification results for winter, spring and summer 2005 is presented, along with some seasonal comparison from 2003 to 2005.

In this paper only the following surface parameters are analysed: 2m Temperature (2m T), 2m Dew Point Temperature(2m TD), 10m Wind Speed (10m WS), Mean sea Level Pressure (MSLP) and precipitation (PP). Further information concerning verification of upper-air parameters can be found on the COSMO web internet site. The verification concerns the 3.9 LAMI reference version.

These five parameters are not explicit model variables but they are computed through some internal post-processing which may introduce extra errors. Nevertheless, since the internal post-processing is generally based on some diagnostic balance among the model variables, which is derived from physical constraints, it is still possible to have some important information about problems in the formulation and in the configuration of the model itself.

The observations forming the control data set were collected on 3-hourly basis from synoptic Italian network, including 91 manned stations and distributed over the Italian area. Stations were divided in three classes according to geographical location; mountain stations (> 700m), valley stations or inner lowland stations and coastal stations. Stations subdivision in different classes has been choosen in order to check systematic errors related with different geographical and surface conditions.

This approach can give two type of results: information about models ability in reproducing correct surface processes through a correct climatology in different geographical areas and indication of possible error sources through error comparison in different areas. For this reason, the results obtained in the verification of daily cycle for 2m T, 2m TD, 10m WS, MSLP and for categorical rainfall verification are presented.

2 Daily Cycle

In order to verify the diurnal behaviour of the model, the couples observation-forecast were stratified according to the hour of the day (3-hourly frequency), the season of the year and the forecast range (day 1 and day 2). Synchronous and co-located couples observation-forecast independently from the station position then form each sample. In such way systematic errors due to inconsistency in the surface representation of the model (inconsistency in the terrain elevation and in the percentage of the surface covered by water are the main error sources over Italy) are somewhat dumped and the signal of daily and seasonal oscillation is retained. For each of the obtained samples the mean error (ME, forecast-obs) and mean absolute error (MAE) were computed.

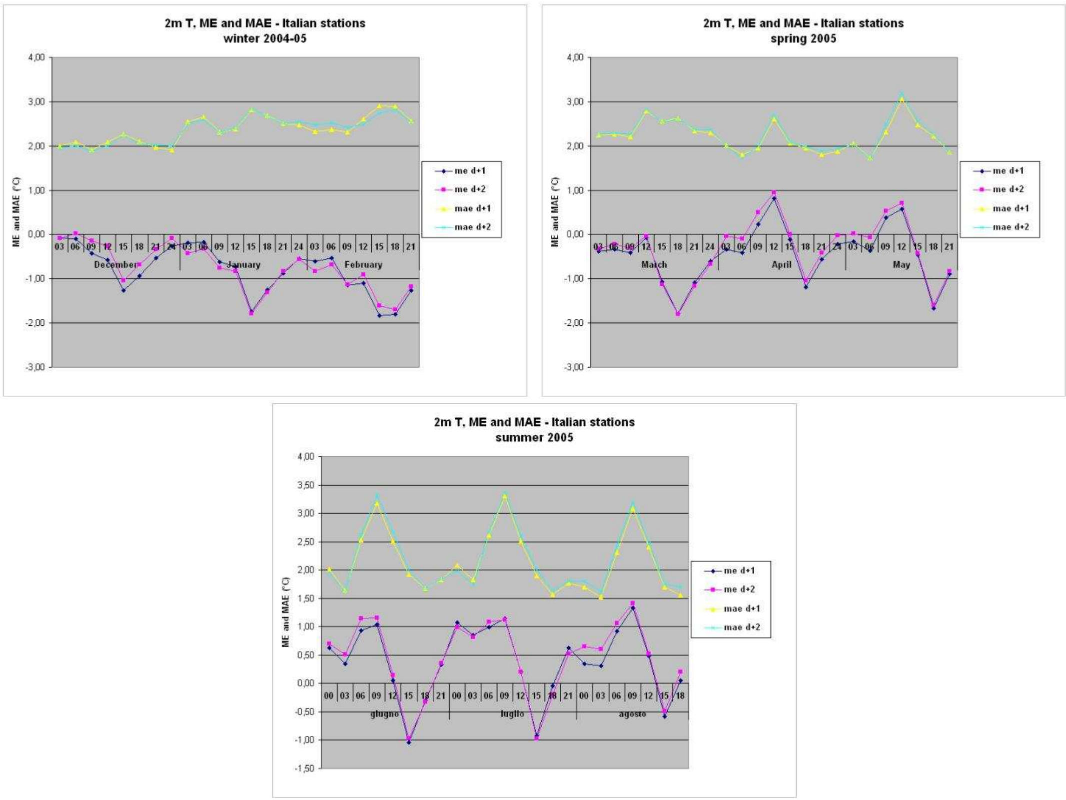


Figure 1: 2 m temperature forecast error for all italian stations.

3 2m Temperature

Figure 1 shows the behaviour of 2m-Temperature forecast error for all the set of italian stations. A clear diurnal cycle is present for all months especially from April to August when it becomes more evident. About the error pattern the figures show a strong cold bias in winter and already from April an increase in ME and MAE (low absolute accuracy) to reach the maxima in July always around 09-12 UTC, with a smaller but clear secondary peak around sunset, maybe a signal of an early warming.

Figures 2 and 3, resp., the seasonal 2m-Temperature for coastal and valley stations, show, of course, the same behaviour of the previous graphs, with some more interesting characteristics. For example for coastal stations the model seems to be colder during afternoon, while for valley stations this seems to disappear (in summer) or to be less evident (in spring). Again for valley stations MAE seems to be always higher (lower accuracy) and the bias in summer shows us a model warmer for almost all the day.

4 2m Dew Point Temperature

A diurnal cycle is also present in ME curves of 2m Dew Point Temperature, see Fig. 4 for all Italian stations. In general, from the ME or bias point of view, it can be said, the model has a better behaviour compared with temperature; in fact the mean is around zero except for the colder months (winter plus march) when a positive bias is evident (the model is too humid). About the absolute accuracy (MAE) the value remains relatively high, while a diurnal cycle is less evident.

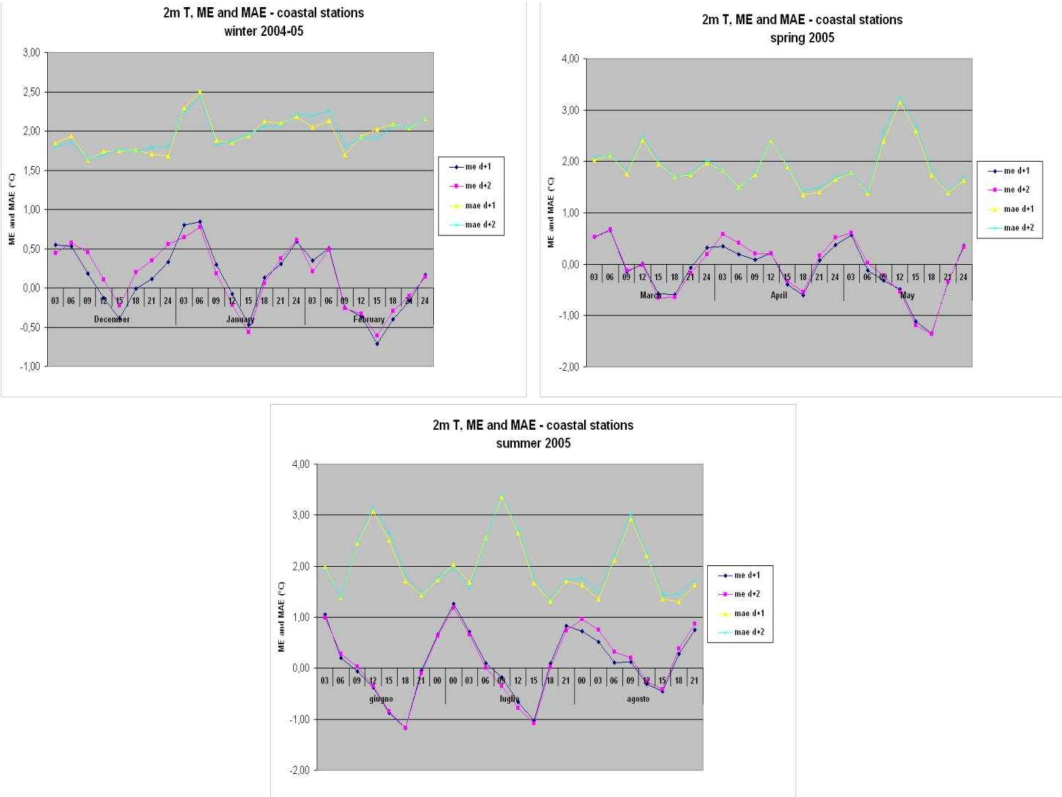


Figure 2: 2 m temperature forecast error for coastal stations.

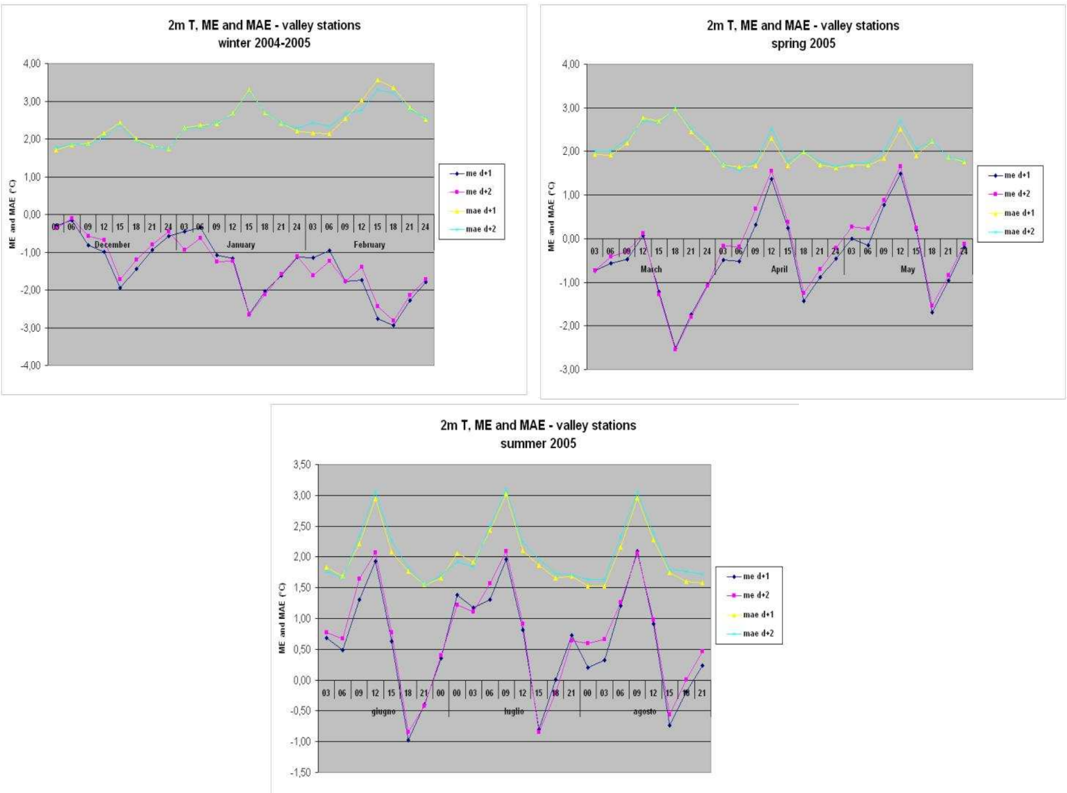


Figure 3: 2 m temperature forecast error for valley stations.

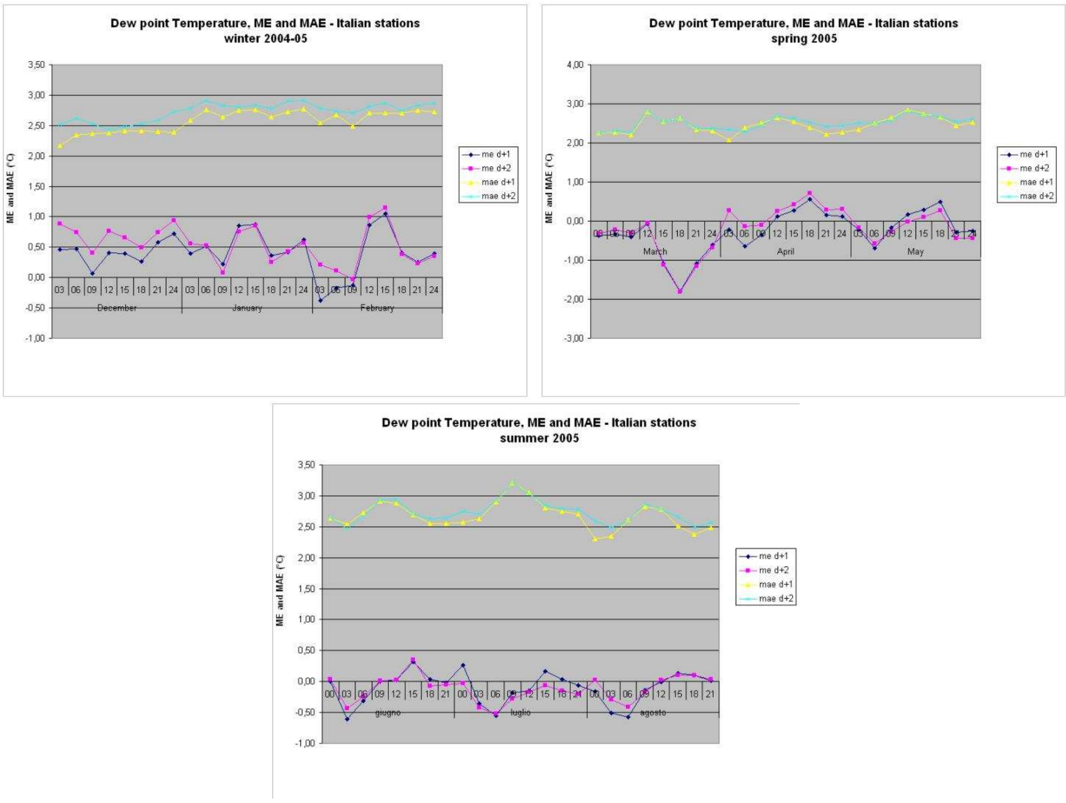


Figure 4: 2 m dew point temperature forecast error for all italian stations.

5 10m wind speed

In Fig. 5 the curves relative to mean error and mean absolute error of 10m wind speed for all Italian stations, are shown. Even if the amplitude is small, a diurnal cycle is present in ME curves. An overestimation of wind speed, positive bias, occurs especially during the cold months for valley and coastal stations, when dynamical circulation is dominant. It is interesting to point the attention to low ME and MEA values in summer months for coastal stations: it could be interpreted as a good model interpretation of local breeze circulation.

6 Mean Sea level Pressure

Figs. 6 show MSLP mean error and mean absolute error for 2005 seasons all over Italy. Mean error curves does not show a clear diurnal cycle, also there is a good phase agreement between ME d+1 curve and ME d+2 curve. MAE curves shows how the mean sea level pressure is less affected by local circulations or by Model physics and is dominated by atmosphere dynamic; in fact, MAE increases quasi-linearly in function of forecast range (for each month, d+1 curve starts with +03 hrs and stops with +24 hrs while d+2 curve starts with +27 hrs to +48hrs forecast range) with a degradation in MAE values during the winter (characterized by strong atmospheric motions). Besides, in summer, there is a clear and strong negative bias for d+2 (a loss of mass?).

7 Precipitation

The results for 2005 seasons are summarized in Figs. 7, 8 and 9, where FBIAS, TS and POD-FAR scores are presented, respectively, for all Italian stations stratified for 12h cumu-

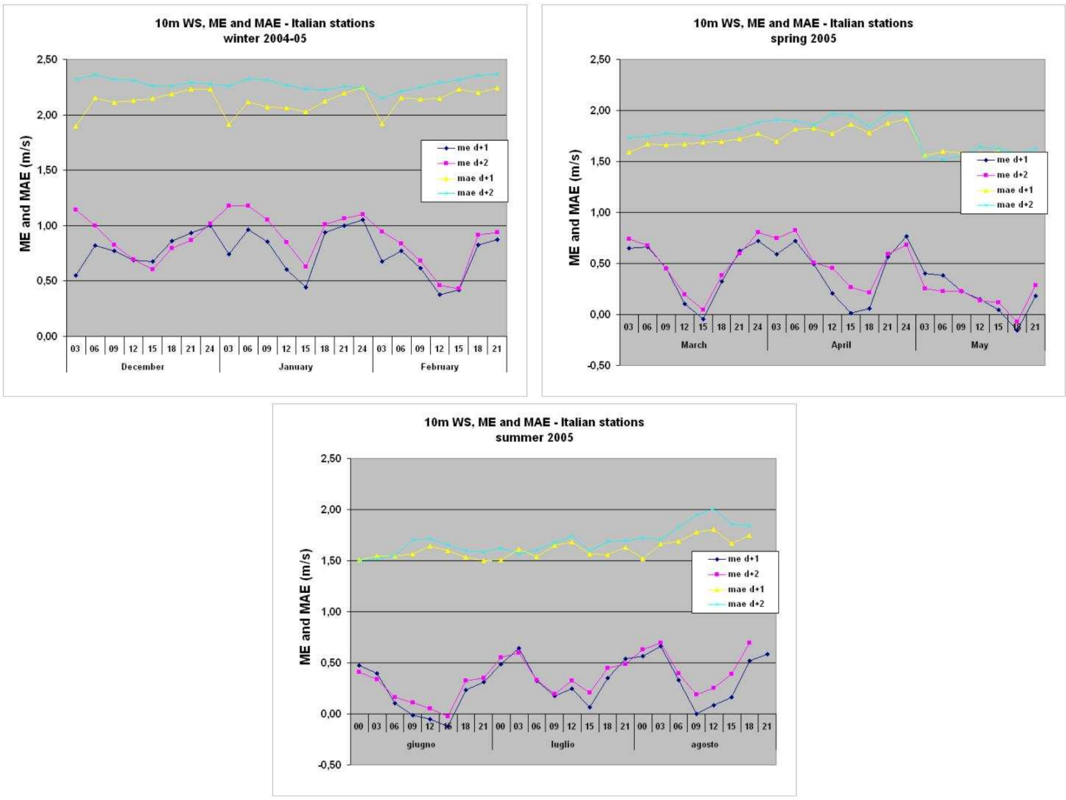


Figure 5: 10 m wind speed forecast error for all italian stations.

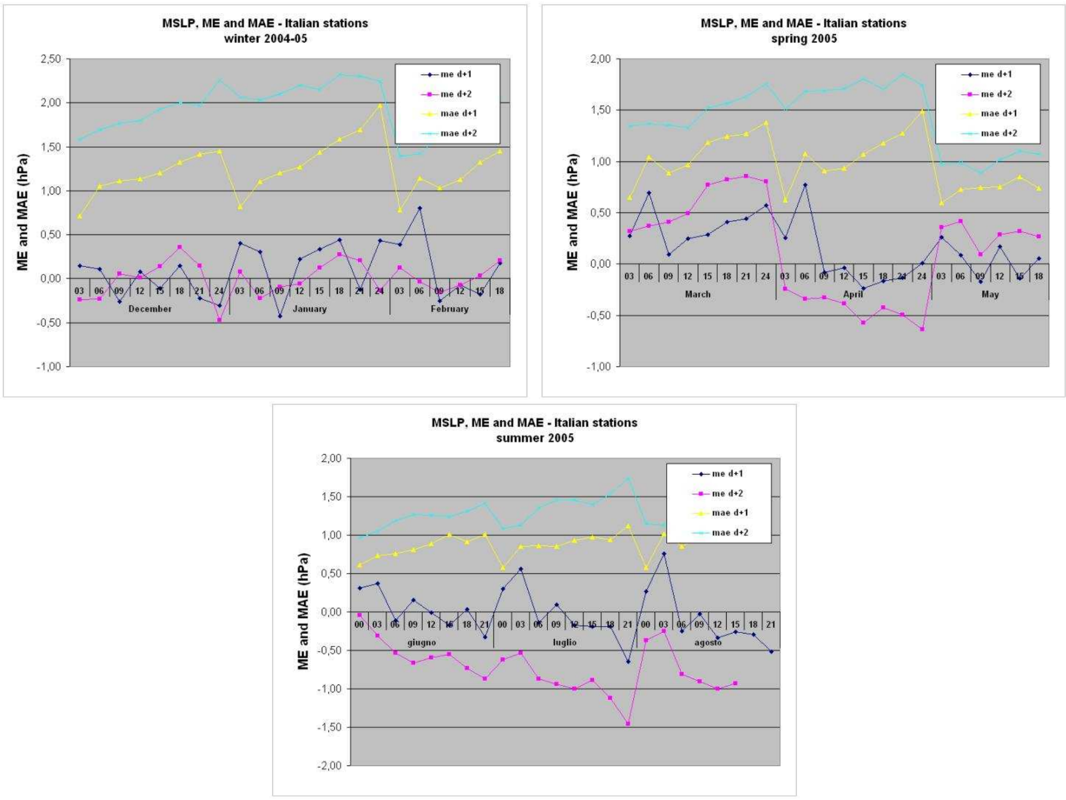


Figure 6: MSLP forecast error for all italian stations.

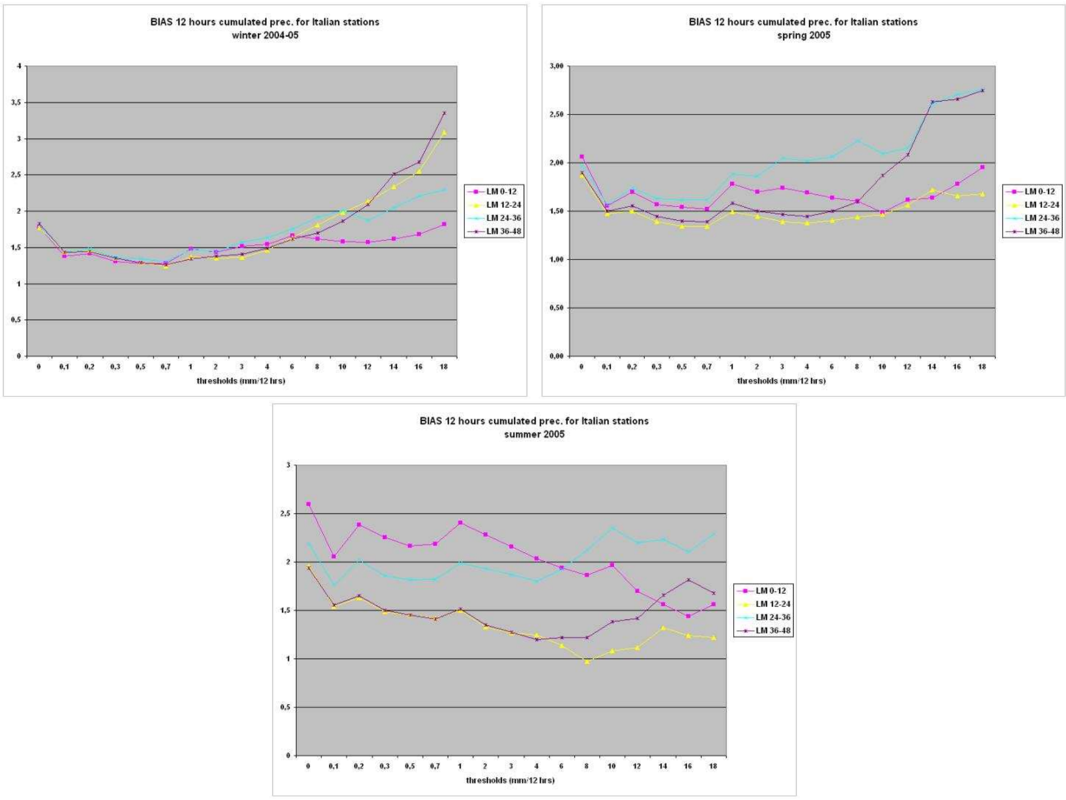


Figure 7: BIAS for 12-h accumulated precipitation for all italian stations

lated precipitations, without any morphological or regional stratifications (for details about stratified precipitation scores see COSMO web site).

Fig. 7 (FBIAS) shows in winter (and less in spring) a better model performance, probably due to the type of precipitations (mainly large scale vs. convective ones) up to 4mm/12h, while in summer the shift in convective daily precipitations (model anticipates the occurrence) can be seen as a clear link with the same kind of signal in 2m-Temperature; in fact there are clear larger FBIAS scores for the morning ranges (00-12 and 24-36). In spring the signals are more complicated and, probably, a mix of the previous two.

Threat Score plots for 12 hours cumulated rainfall, reported in Fig 8, show that model performances decrease with the season and only 12-24 range remains on a acceptable values, also in summer.

LAMI Probability of Detection and False Alarm Ratio plots, Fig. 9, give useful information to understand the threshold range where forecast can be used with high benefit, that is the plot area where POD-FAR. The transition threshold is around 5-6 mm/12h for winter for all forecast ranges, and decrease with the season, becoming really low and confused in summer. Again this is probably due to the peculiar nature of precipitation (convective) and to the early morning shift in temperature.

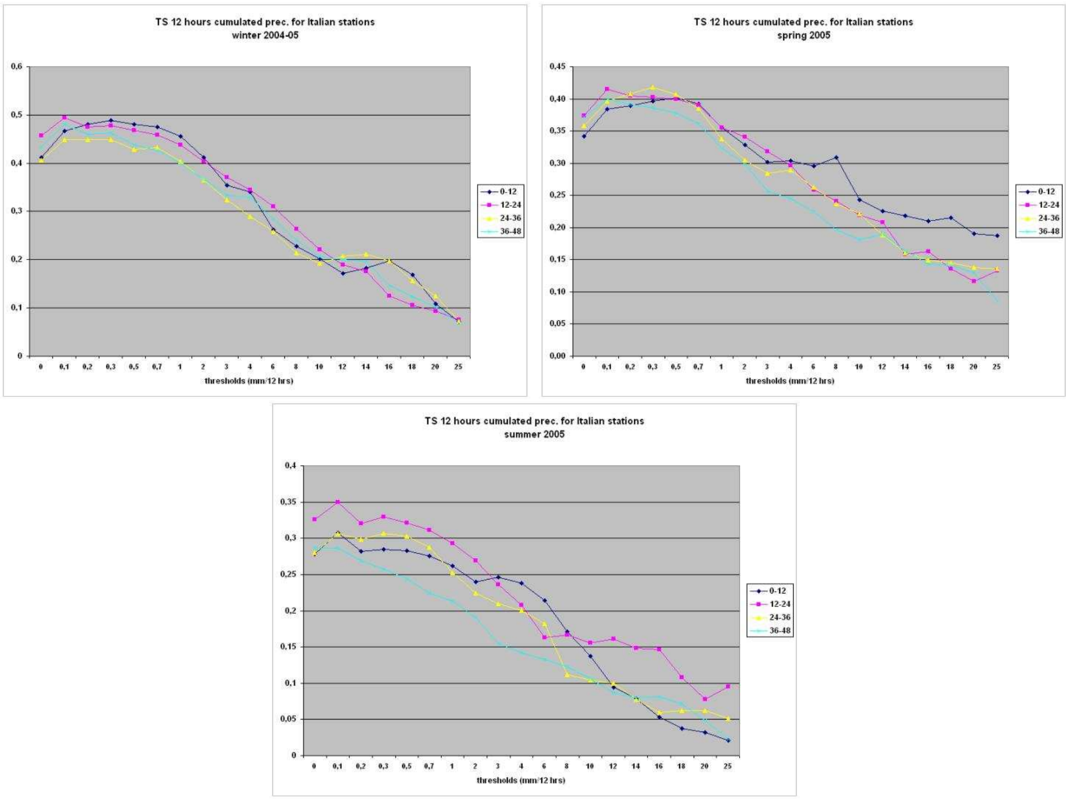


Figure 8: TS for 12-h accumulated precipitation for all italian stations

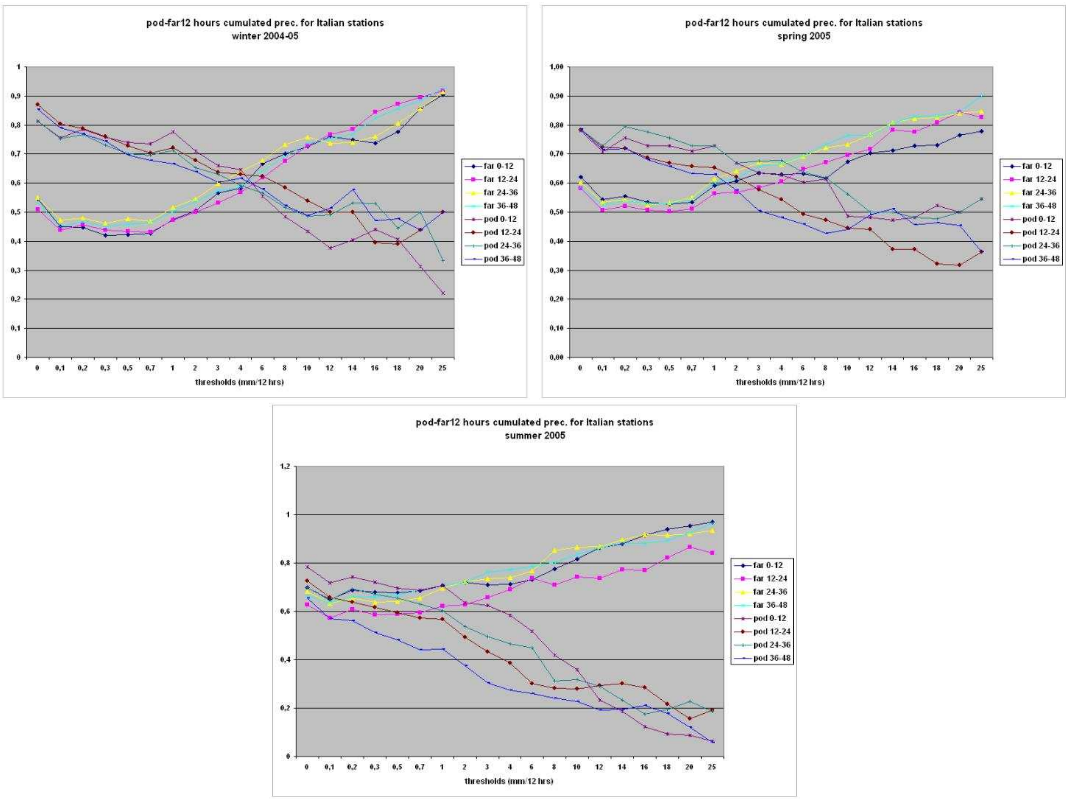


Figure 9: POD-FAR for 12-h accumulated precipitation for all italian stations

Precipitation Verification (Overestimation): A Common View of the Behaviour of the LM, aLMo and LAMI

Francis Schubiger[1], Pirmin Kaufmann[1], Uli Damrath[2] and Elena Oberto[3]

[1]*MeteoSwiss, Zürich,* [2]*Deutscher Wetterdienst, Offenbach am Main, Germany,* [3]*ARPA Piemonte, Torino, Italy*

1 Introduction

The operational verification of precipitation over Germany and especially over Switzerland showed an increased overestimation during the last two winters (2003/04 and 2004/05) as compared to earlier winters (Sec. 2). The differences of mean monthly precipitation (24h-sums of LM +6 to +30h) to rain gauges in Germany and Switzerland showed a strong increase of the positive bias since the end of 2003 (C. Schraff, personal communication).

These two verification results were the starting point of this examination. The aim is to give some hints and information to the WG3-scientists for identifying the causes of possible precipitation overestimation in winter by looking more in detail into the verification of precipitation over the full domains of the operational model versions running in COSMO.

The relation between monthly sums of forecasted (by LM, aLMo and LAMI) and observed precipitation (SYNOP and rain gauges network) has been examined since January 2000 (Sec. 3 and 4). The bias (from monthly or seasonal sums) shows a different signal from season to season, region to region but also from year to year (for the same region). One part comes from the precipitation variability, another part could arise from model changes (namely the cloud ice scheme). The overestimation in winter seems more concentrated for the last two/three winters (and partly over the mountainous regions).

Further studies are necessary to identify possible causes of precipitation overestimation (Sec. 5): the problem is that precipitation is a very hard parameter not only in prediction but also in verification! Some few cases can "offset" a mean statistics based on biases. A first step was an examination of scatter plots of daily precipitation (Sec. 6): it shows that the overestimation is not only a problem of some isolated cases, but is visible on almost all days. The next steps (Sec. 7) will be studied in 2006 within the COSMO priority project "Tackle deficiencies in quantitative precipitation forecasts".

2 Verification of aLMo and LM with the hourly observations from the automatic network of MeteoSwiss during the four last winters and for the test chains at DWD and MeteoSwiss with the cloud ice scheme.

Fig. 1 shows the mean daily cycle of precipitation during the last winters (a) 2002/03, (b) 2003/04, (c) 2004/05 and (d) for a testsuite at MeteoSwiss with the cloud ice scheme in Spring 2004.

During winter 2002/03 LM and aLMo gave a similar overestimation: LM 25%, aLMo 30%. The slightly stronger overestimation in aLMo comes from the high amounts (10 mm/6h). The results from the earlier winter 2001/02 are similar (not shown): overestimation in both models, of the order of 20% in aLMo and 25% in LM. During these two winters, both models were driven by GME and run without cloud ice scheme.

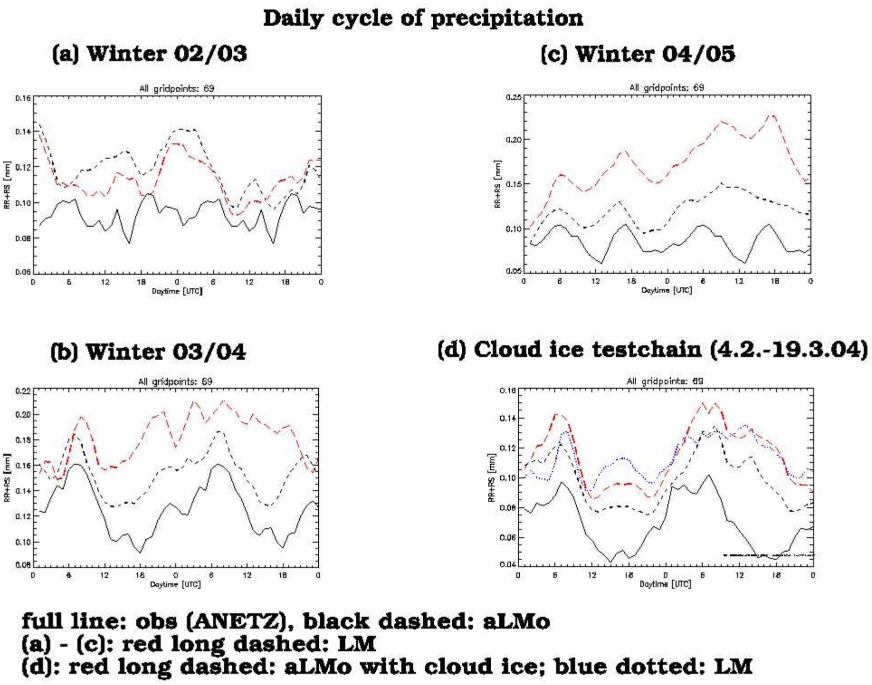


Figure 1: Verification of the daily cycle of precipitation for all 69 grid points corresponding to an ANETZ-station during the winters (a) 2002/03, (b) 2003/04, (c) 2004/05 and (d) for a test suite from 4/2-19/3/2004 with the cloud ice scheme. Observations (ANETZ): full line black. In (a)-(c): aLMo: black dashed line; LM: red long dashes. In (d): aLMo without cloud ice: black dashed line, aLMo with cloud ice: red long dashes, LM (also with cloud ice scheme; here just for comparison): blue dotted line.

During the winter 2003/04 LM run with cloud ice scheme, but aLMo still without the cloud ice scheme (see COSMO Newsletter No. 4, pp. 181-188 "Recent changes to the cloud-ice scheme"). aLMo shows an overestimation of ∽20% and LM ∽40%. aLMo captures better the mean daily cycle, possibly due to the lateral boundary conditions (LBC) from ECMWF (IFS-frames since 16.09.03), whereas LM takes the LBC from the GME (DWD). The frequency bias shows an overestimation for low amounts (0.1 mm/6h) in both models, of the order of 35-40% for grid points < 800m and 55-65% for grid points > 800 m. The high amounts (10 mm/6h) are underestimated with aLMo (∽10%) and overestimated with LM (∽35%) but in the range 800-1500 m both models show an overestimation, aLMo of 30% and LM of almost 100%.

A possible candidate for this stronger overestimation in LM as compared to aLMo could be the cloud ice scheme. DWD tested the cloud ice scheme in a test suite of LM during May 2003 (with an almost neutral impact) and with a revised version in September 2003 (with noticeable improvements) [see COSMO Newsletter No. 4, p. 187]. MeteoSwiss tested the cloud ice in a test suite from 4 February to 19 March 2004 (Fig. 1): the results over Switzerland showed an increase of ∽15% in precipitation (but already the operational version gave an overestimation of ∽20%). The frequency bias shows an overestimation for low amounts (0.1 mm/6h) in both versions, of the order of ∽70-80%. The high amounts (10 mm/6h) are much more frequent in the cloud ice version: aLMo without cloud ice scheme gives an underestimation of ∽20% and aLMo with cloud ice an overestimation of ∽30%.

So the increased precipitation amount in LM as compared to aLMo during winter 2003/04 could come from the cloud ice scheme already operational in LM, but not yet in aLMo, where it was introduced in May 2004. During the year 2004 the prognostic precipitation scheme has been introduced operationally in LM (26/04/2004) and in aLMo (15/11/2004), so in winter 2004/05 both models run with the cloud ice scheme and with prognostic precipitation scheme. But LM run since 15.12.2004 with a bug correction in the numerical treatment of the prognostic precipitation: this correction gives a higher precipitation amount of $\backsim$10-15% in the area mean. In aLMo this correction has been introduced only in July 2005.

During the winter 2004/05 LM gives about 35% more precipitation than aLMo over a domain of 57x39 grid points covering Switzerland, but already aLMo gives $\backsim$20% too much precipitation as compared to all 69 ANETZ stations. In the first forecast hours the difference between both models is less pronounced. The frequency bias shows an overestimation for low amounts (0.1 mm/6h) in both models, for grid points $<$ 800m of $\backsim$40% in aLMo and 65% in LM, for grid points $>$ 800 m of $\backsim$70-90% in aLMo and even 20% more in LM. The high amounts (10 mm/6h) are slightly underestimated with aLMo ($\backsim$10%) but strongly overestimated with LM ($\backsim$230%). During this winter 2004/05 LM run with a bug correction in the prognostic precipitation scheme giving $\backsim$10-15% more precipitation, so about a third of the higher precipitation amount in LM is explained by this fact.

Summarized we can say: especially during the last two winters LM and aLMo gave a strong overestimation in precipitation over the mountainous regions. A possible cause could be the introduction of the prognostic cloud ice scheme that was not enough tested during wintertime conditions.

3 Verification of LM with rain gauges over Germany and Switzerland and of aLMo with all European SYNOPS.

Fig. 2 shows the monthly evolution from January 2000 to December 2004 of the 24h precipitation sums from +6h to +30h (blue surface) and the mean bias (red line) of LM as compared to the 3300 rain gauge stations in Germany. It shows the increased positive bias during the last two winters. This increase comes mainly from the mountainous regions in the southern part of Germany (not shown). The verification with the 450 rain gauges over Switzerland (not shown) gave already an overestimation in winter 2001/02 when LM and aLMo run without the cloud ice scheme.

Fig. 3 shows the winter season precipitation bias of 12 hourly precipitation sums at all European SYNOP stations over the full domain of aLMo (from Portugal in South-West to Poland in North-East). The mean bias shows a decrease in the overestimation during the last winters. Fig. 4 shows the mean error of 12 hourly precipitation during Winter 2004/05 and confirms that the overestimation is mainly concentrated over the Alpine region. The behaviour over the Alpine region discussed in Sec. 2 (increase of the bias) is just reverse over the major part of Europe outside the mountainous regions!

4 Verification of LM, aLMo and LAMI with raingauges over Piedmont.

Fig. 5 show the seasonal cumulated precipitation maps from the three operational models LAMI, LM and aLMo and the rain gauge network over the Piedmont region. In particular winter 2003/04 (DJF'04) and winter 2004/05 (DJF'05) are remarkable both in term of wet/dry seasons and in term of different model versions performance. There is a general overestimation during both winters: in winter 2003/04 aLMo, that run without prognostic

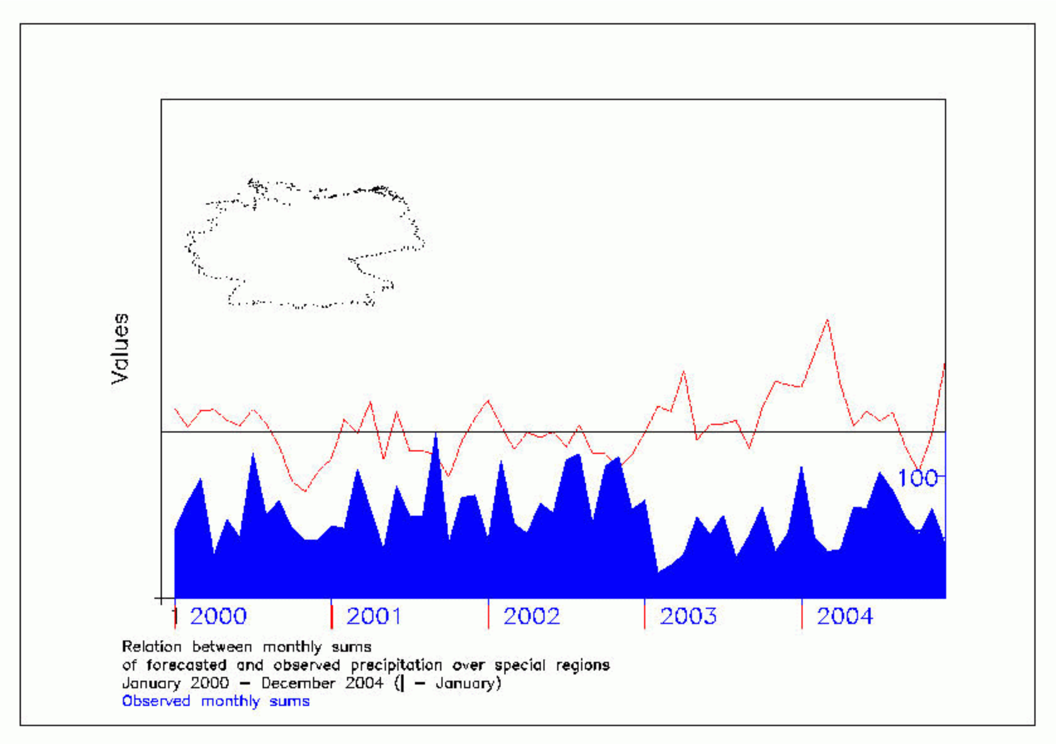


Figure 2: Monthly evolution from January 2000 to December 2004 of the 24h precipitation sums from +6h to +30h (blue surface) and the mean bias (red line) of LM as compared to the 3300 rain gauges stations in Germany.

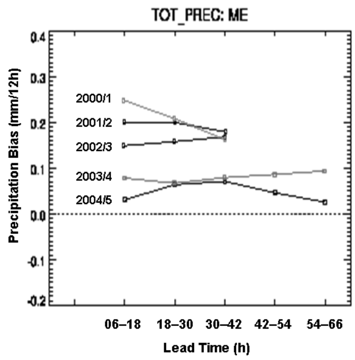


Figure 3: Precipitation bias for the winter seasons of 12 hourly precipitation sums at European SYNOP stations of all 00 UTC and 12 UTC aLMo-forecasts. Results are plotted for the different lead times.

cloud-ice scheme, performs slightly better with respect to the other ones (with cloud ice scheme): this is also visible on the frequency bias results. In winter 2004/05 all three models run with prognostic cloud ice scheme and LM and aLMo also with prognostic precipitation scheme (but not LAMI): in that last winter LAMI seems to overestimate less than aLMo and LM, but this winter was unfortunately very dry in Piedmont, so the results are difficult to interpret.

5 Preliminary conclusions.

A possible cause for the precipitation overestimation is the cloud ice scheme, but other causes (masked by the precipitation variability) are also possible, because

- already before the introduction of cloud ice scheme we had seasons with precipitation

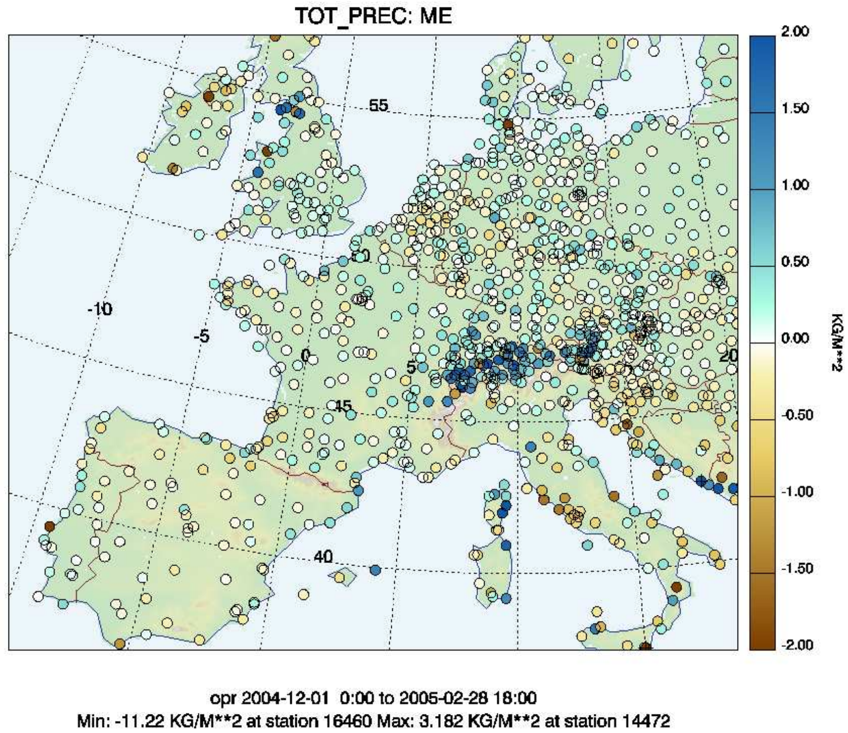


Figure 4: Mean error of 12 hourly precipitation sums from all 00 UTC and 12 UTC aLMo-forecasts (range +30h to +42h).

overestimation;

- we have great regions (especially in Southern Europe) with a decrease of the (seasonal) biases during the last winters.

Clearly, more work is necessary to find causes of the precipitation overestimation in mountainous regions.

6 Scatter plots of daily precipitations sums over Germany and Piedmont

Scatter plots of daily precipitation amounts can help to see if this overestimation is due to (some few) cases where the model gives (much) precipitation that is not observed, or whether overestimation of precipitation is a problem in all (or most) cases. Monthly scatter plots of the daily sums of LM vs rain gauges in Germany for three months are shown in Fig. 6-Fig. 8: August 2002 with the flooding events in Central Europe, December 2002 when LM run without cloud ice scheme and December 2004 when LM run with cloud ice scheme. The overestimation is not only a problem of some cases (i.e. days). In 2002 the observed strong precipitation (more than 10 mm in the areal mean) were underestimated. This behaviour is just reversed in December 2004, where almost all days show an overestimation. The last two winters (2003/04 and 2004/05) gave systematically overestimated daily sums whereas during the summer months there is no systematic over/underestimation (not shown).

Fig. 9 shows similar scatter plots over the Piedmont region for the two winters 2003/04 and 2004/05 for the three models aLMo, LM, and LAMI. Also here it is evident that the biases in the seasonal means is not a problem of isolated days. Figure 10 gives the results of the daily scatter plots of January 2005: during this month the overestimation of the three models is

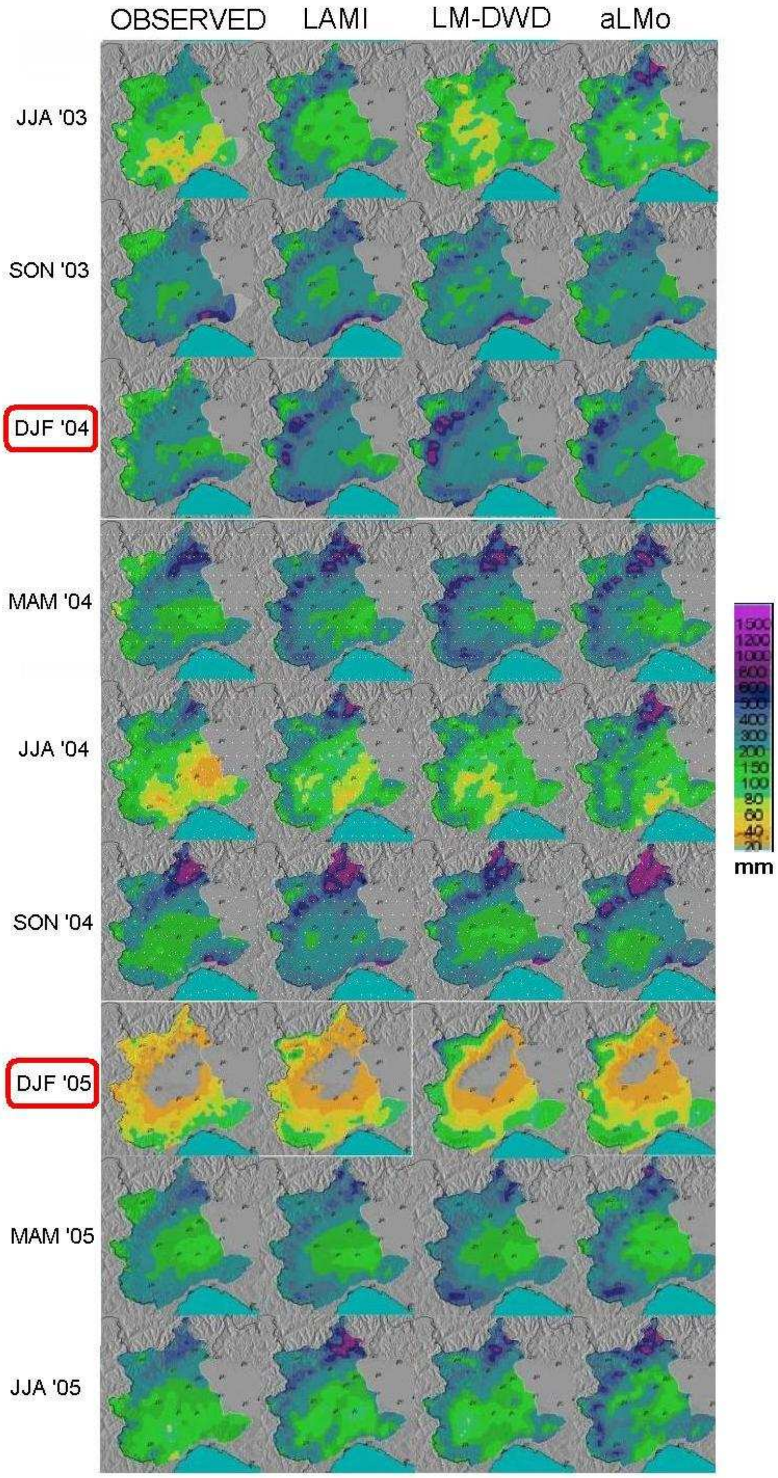


Figure 5: Seasonal cumulated precipitation maps from the 00 UTC operational forecasts (24h sums from +0h to +24h) of LAMI, LM and aLMo and observed from the rain gauges over the Piedmont region.

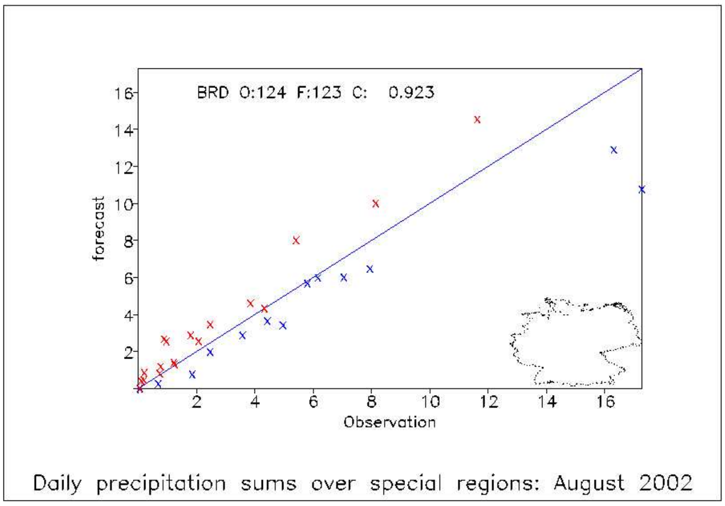


Figure 6: Monthly scatter plots of the daily sums of LM vs rain gauges in Germany for August 2002.

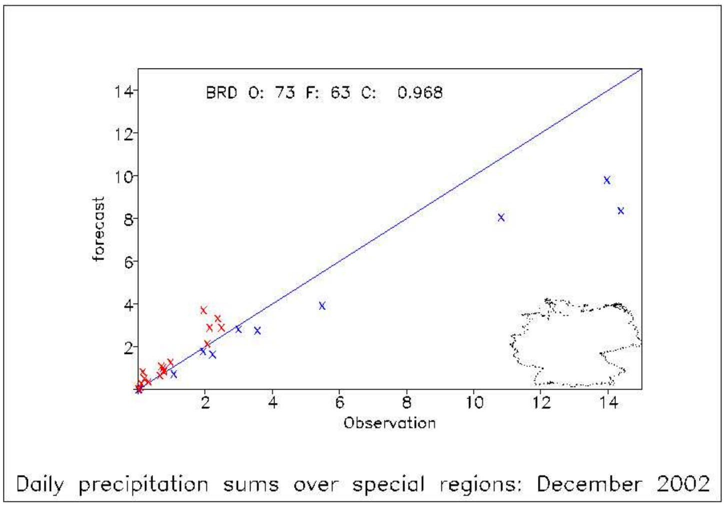


Figure 7: As in Fig. 6 but for December 2002.

very high for all days (Fig. 10 shows LM, but the results for LAMI and aLMo are similar). The overestimation is much higher in January 2005 than on the other two winter months: during this month the precipitation events were mainly due to a cold air outbreak and two days of strong northwesterly flow, i.e. they are connected to synoptic situations with cold air. It could be a hint that the overestimation is more pronounced during very cold precipitation events. This must be further investigated with a more systematic weather-type dependant verification of these wintertime precipitations.

7 Outlook.

Further systematic scatter plots of daily precipitation sums (also from radar areal mean estimates) can help to better identify the forecast failures. Alternatively simple conditional verifications by discriminating between events of different (observed) vertical stability (i.e., unstable/convective vs stable/stratiform) or with parameters such as 'convective/stratiform

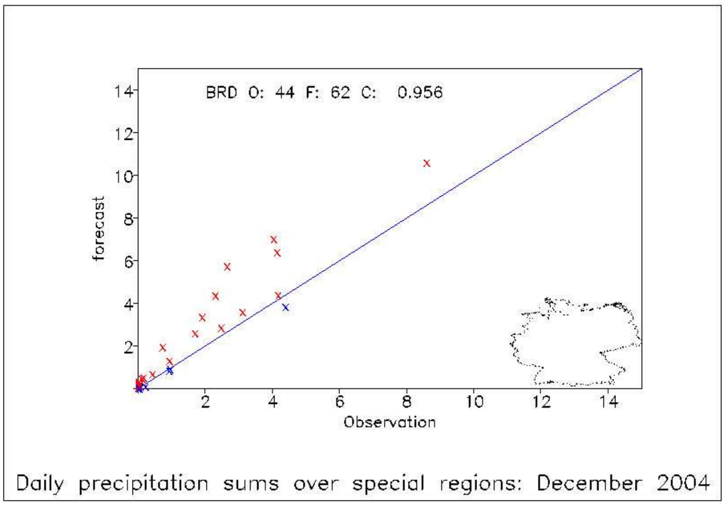


Figure 8: As in Fig. 6 but for December 2004.

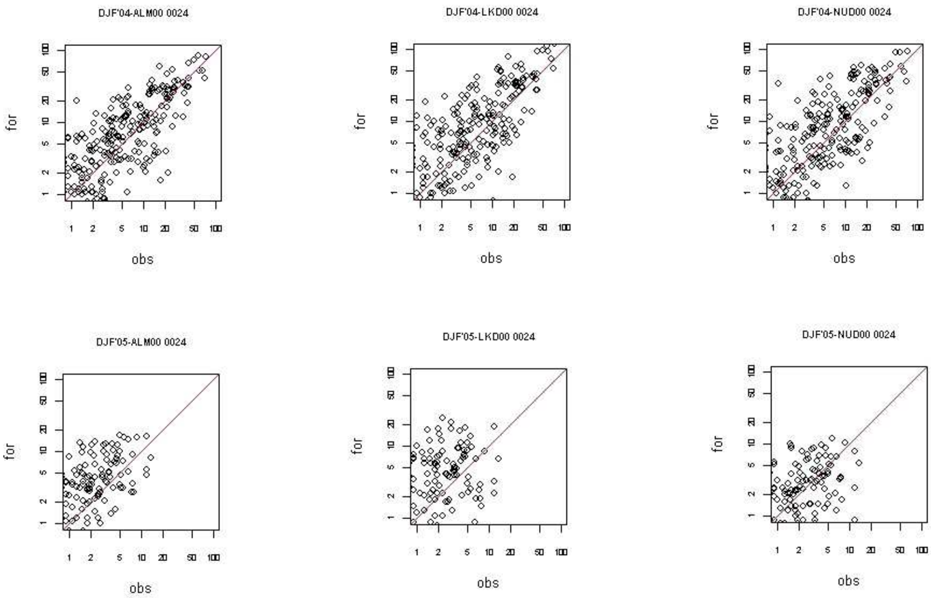


Figure 9: Scatter plots over the Piedmont region for the two winters 2003/04 (upper part) and 2004/05 (lower part) for the three models aLMo (left), LM (middle) and LAMI (right).

precipitation amounts in the model' would help to attribute the problem to specific parts of the physics parameterizations or other model properties.

The COSMO priority project "Tackle deficiencies in quantitative precipitation forecasts" will pursuit this study: a first task will be a consolidated report of forecast failures and verification findings.

References

COSMO Newsletter, No. 4, 2004.

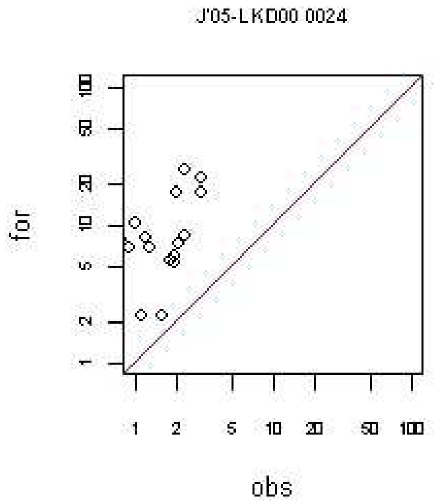


Figure 10: Scatter plots over the Piedmont region for LM in January 2005.

Verification of COSMO-LM in Poland

Katarzyna Starosta, Joanna Linkowska and Andrzej Mazur

Institute of Meteorology and Water Management
Centre of Numerical Weather Forecasts
61 Podleśna str., PL-01673 Warsaw, Poland

1 Introduction

The results of the verification of COSMO-LM in Poland from January 2005 to June 2005 are shown below. In our research we verified the following parameters: the surface meteorological parameters, the 24-hour precipitation amounts and the upper-air parameters.

2 Results

2.1 Verification of surface parameters using 56 Polish SYNOP stations

The mesoscale LM model was verified daily (operational verification) and monthly. For the fields generated by the model the following parameters were extracted: the temperature at 2 m above ground level, the dew point temperature at 2 m a.g.l, the air pressure at sea level, the wind speed at 10 m a.g.l. and cloud cover. For operational verification we compared the present data with 6 earlier forecasts for the same hour. For monthly verification, the mean error (ME) and the root mean square error (RMSE) were calculated using a 12 forecast range (every 6 hours) for a 72 hour forecast starting at 00 UTC. The error estimators were calculated by stations throughout Poland.

2.1.1 The temperature at 2 m above ground level.

We observed the diurnal and monthly cycle of the RMSE and ME. (Fig. 1) The diurnal cycle in the spring and summer (March, June) was bigger than in the winter and achieved a maximum at 12 UTC and a minimum at 6 UTC. The biggest amplitude of RMSE and ME occurred in March. The overestimation of temperature was observed in the summer and the underestimation of temperature was observed in the winter. (Fig. 2). In Fig. 3 is shown the distribution of ME for all synoptic stations (the two highest mountain stations are not shown). In this case, the temperature was overpredicted for almost all stations. The mean errors on each station are different and depend on the location of the station to the nearest grid points and the altitude of the station. (Fig. 3)

2.1.2 The dew point temperature at 2 m a.g.l.

The diurnal cycle of RMSE only occurred in May and June. In January and April we observed the growth of the RMSE with forecast time. In January, the ME was negative regardless of the diurnal cycle. In February and March, the ME was negative at night and positive during the day, with the maximum at 12 UTC. From April to June, the ME was positive. The ME achieved a maximum at 18 UTC and a minimum at 6 UTC in May and June. The value of the diurnal amplitude of the ME was small for almost all months, except February and March. The mean error for each station is different (Fig. 4 - Fig. 6).

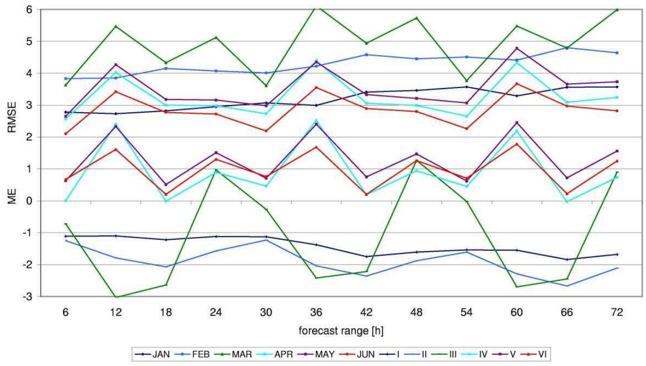


Figure 1: RMSE, ME, Temperature 2m [°C], January - June 2005.

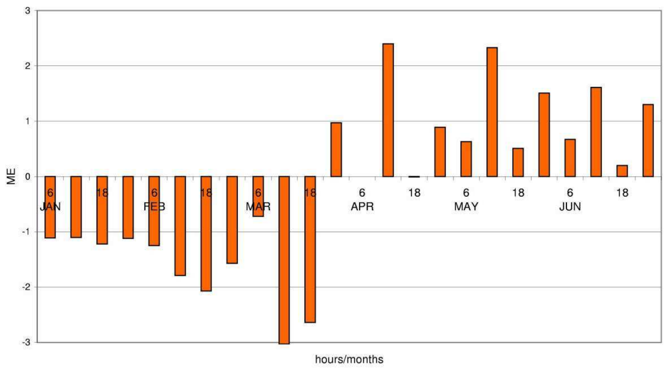


Figure 2: ME, 1st day, Temperature 2m [°C], January - June 2005.

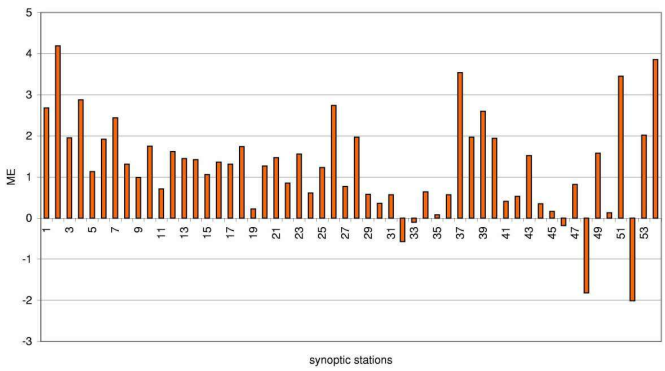


Figure 3: ME, Temperature 2m [°C], 36h-forecast, June 2005.

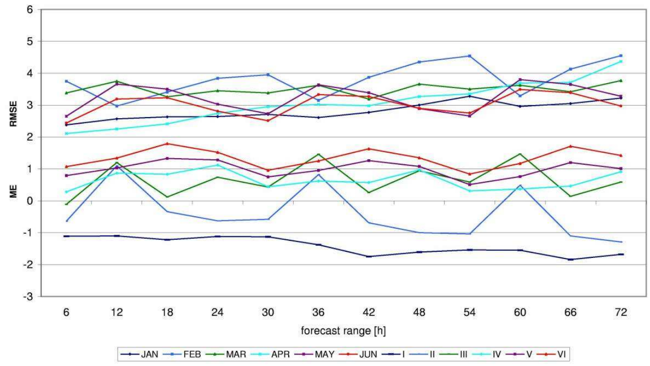


Figure 4: RMSE, ME, Dew point temperature 2m [°C], January - June 2005.

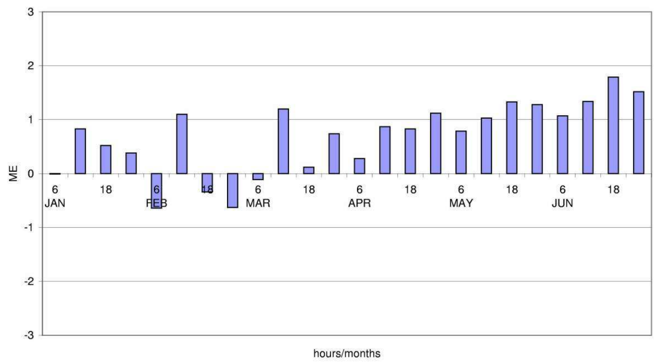


Figure 5: ME, 1st day, Dew point temperature 2m [°C], January - June 2005.

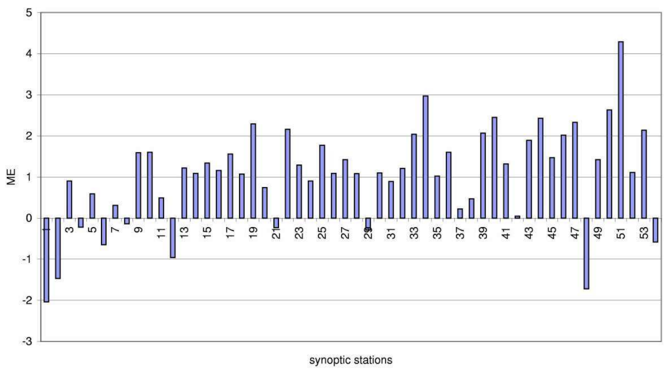


Figure 6: ME, Dew point temperature 2m [°C], 36h-forecast, June 2005.

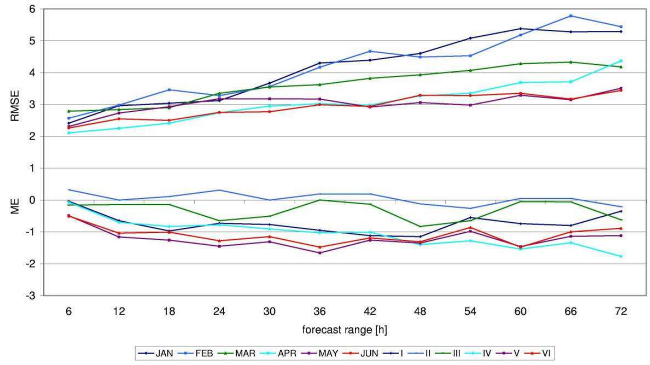


Figure 7: RMSE, ME, Pressure [hPa], January - June 2005.

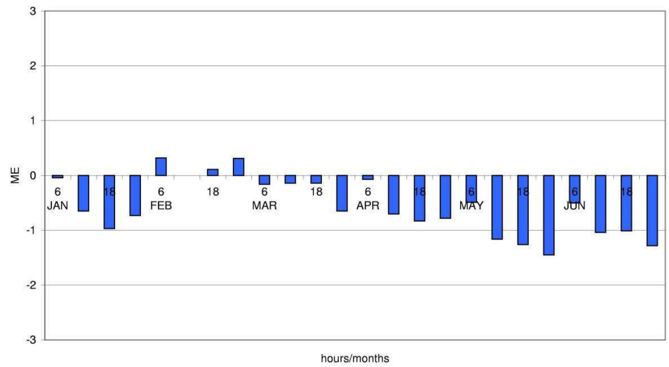


Figure 8: ME, 1st day, Pressure [hPa], January - June 2005.

2.1.3 The air pressure at sea level

The RMSE clearly increased with forecast time. The growth was smaller in the summer than in the winter. The ME in January was closed to 0 hPa. During the winter the range of ME is from -1.0 hPa to 0 hPa and during the summer the range is from -1.5 hPa to 1.0 hPa (Fig. 7 - Fig. 9).

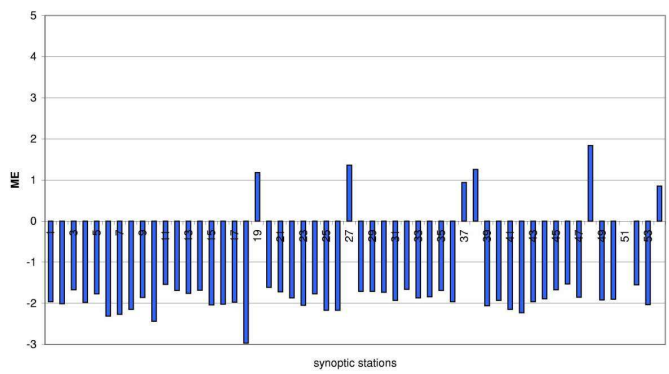


Figure 9: ME, Pressure [hPa], 36h-forecast, June 2005.

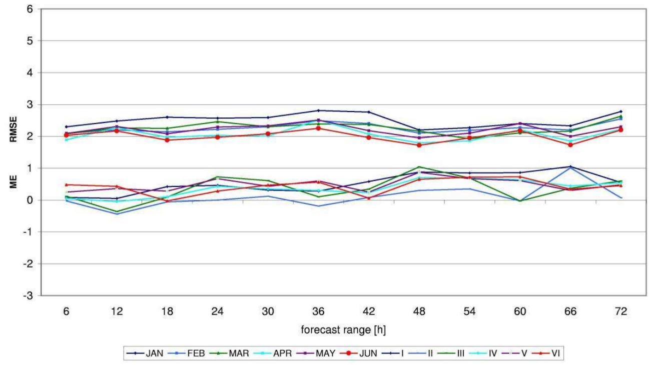


Figure 10: RMSE, ME, Wind speed [m/s], January - June 2005.

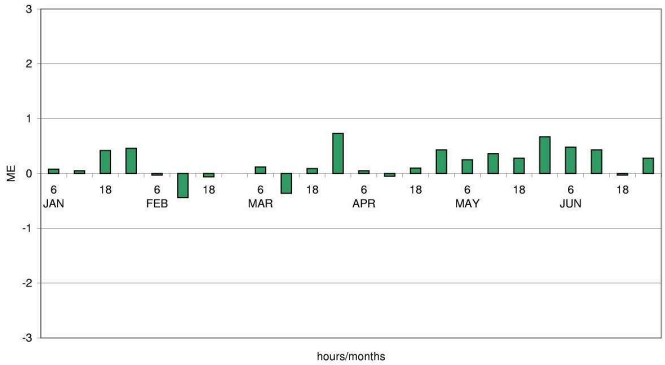


Figure 11: ME, 1st day, Wind speed [m/s], January - June 2005.

2.1.4 The wind speed at 10 m a.g.l.

The mean error is almost positive during the whole period. The ME increased with the forecast time. For 1st day and 2nd day is about 0.5 m/s and for 3rd day is 1 m/s. The RMSE is bigger in the winter than in the summer. We observed the diurnal cycle of RMSE from April to June (Fig. 10 - Fig. 12).

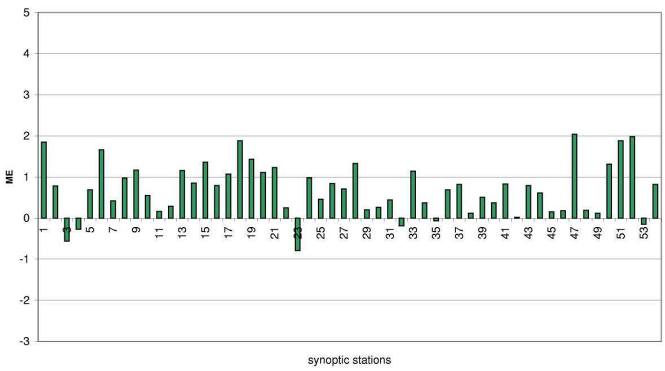


Figure 12: ME, Wind speed [m/s], 36h-forecast, June 2005.

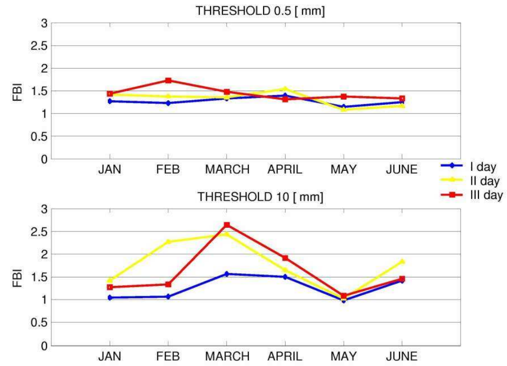


Figure 13: Frequency bias index, January -June 2005.

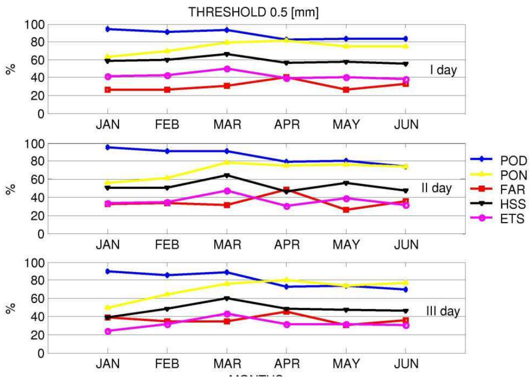


Figure 14: Indices for 24h accumulated precipitation, January -June 2005, threshold 0.5 [mm].

2.2 Verification of 24-hour precipitation amounts using 308 rain gauge stations

For the calculations we interpolated the gridded forecast values on the station points where observations are available. The interpolation of the forecast values on the station points was performed by averaging the values on the four nearest grid points. For this purpose we used the bilinear interpolation. We verified the 24-hour precipitation amounts using 7 indices from the contingency table for the 3 day forecast range (1st day, 2nd day, 3rd day). For verification of the precipitation thresholds 0.5, 1, 2.5, 5, 10, 15, 20, 25, 30 mm were used. For each threshold the following scores were calculated: frequency bias index (FBI), probability of detection of event (POD), false alarm rate (FAR), true skill statistic (TSS), Heidke skill score (HSS) and equitable skill score (ETS) (Fig. 13 - Fig. 16). The figures shown the overestimation of precipitation amount. This overestimation was bigger in 2nd day and 3rd day than in 1st day. Especially it was apparent for threshold 10.0 mm. The forecast of precipitation amount was better from January to March than in next months. For threshold 0.5 mm in 1st day of forecast the POD index was always bigger than 80 % and the FAR index was smaller than 40 %. Those indices did not deteriorate significantly in subsequent days of the forecast. For higher precipitation the FAR index significantly increased and exceeded the POD index.

2.3 Verification of upper-air parameters using 3 TEMP stations

The quality of 72- hour mesoscale forecast of DWD model for Poland was estimated through comparison of forecast results with upper-air soundings, carried out at three Polish stations, located in Leba, Legionowo and Wroclaw. The results of COSMO-LM were compared to

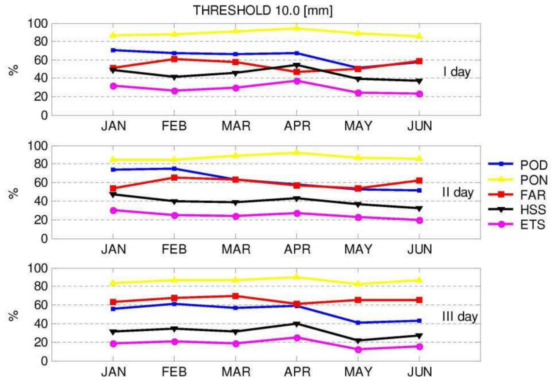


Figure 15: Indices for 24h accumulated precipitation, January -June 2005, threshold 10 [mm].

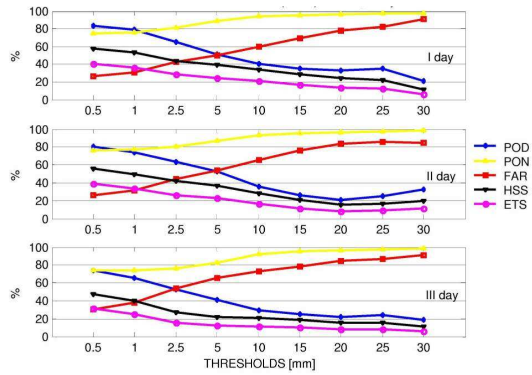


Figure 16: Indices for 24h accumulated precipitation, May 2005.

actual values, observed at the stations. Following meteorological elements (at standard pressure levels 1000, 850, 700, 500, 400, 300, 250 and 200 hPa) were concerned:

- Air temperature;
- Relative humidity;
- Height of a pressure level (air pressure);
- Wind speed.

Verification was carried out for 0, 12, 24, 36, 48, 60 and 72 hour forecast. Mean Error and Root Mean Square Error were calculated as basic scores (Fig. 17 - Fig. 19).

3 Conclusions

3.1 The 2 m temperature

- A monthly and seasonal variation for the scores of temperature is observed.
- The mean error is negative in the winter and positive in the spring and summer.
- In the summer we observed a large diurnal amplitude of mean error and amplitude of RMSE with maximum value during a day.

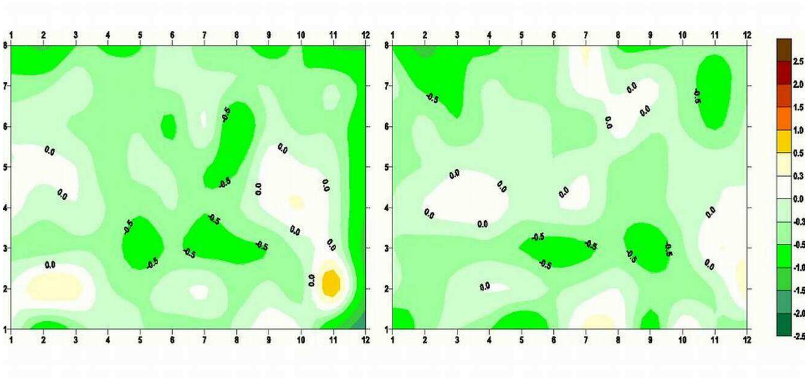


Figure 17: Mean error (observed-predicted) for temperature, soundings in Leba, 2003 (left chart) and 2004 (right chart).

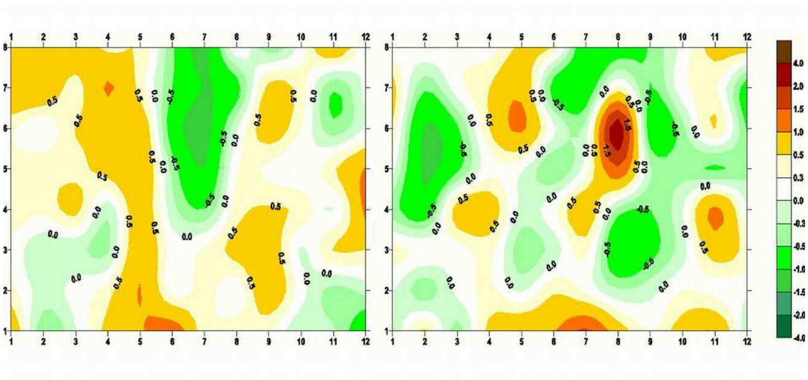


Figure 18: Mean error (observed-predicted) for windspeed, soundings in Leba, 2003 (left chart) and 2004 (right chart).

3.2 The dew point temperature at 2 m a.g.l.

- The monthly variation of mean error is observed. The bias is negative in January, positive in the summer and diurnal amplitude in the spring.
- The RMSE increased with the forecast time.

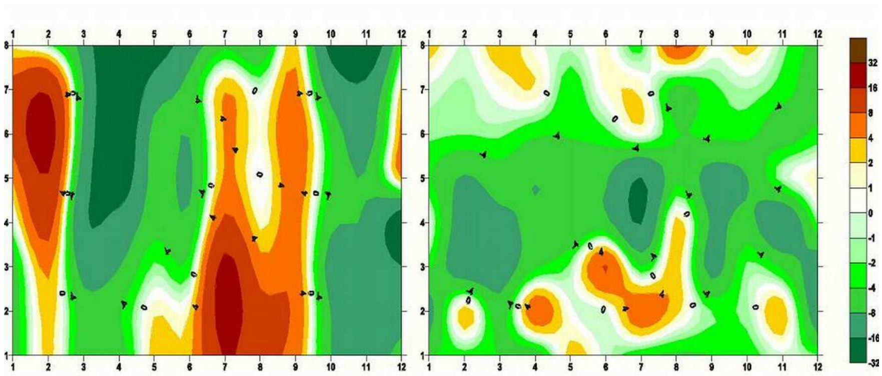


Figure 19: Mean error (observed-predicted) for relative humidity, soundings in Leba, 2003 (left chart) and 2004 (right chart).

3.3 The air pressure at sea level

- The RMSE increased with the forecast time.
- The error is smaller in the summer and higher in the winter.
- The ME is quite smooth (about zero in the winter and negative in the summer).

3.4 The wind speed at 10 m a.g.l.

- The ME is mostly positive and increases with the forecast time.
- The RMSE is quite smooth, bigger in the winter than the summer (with daily amplitude in the summer).

3.5 24-hour precipitation amounts

- The model overestimates the amount of precipitation.
- The precipitation forecast quality in the period from January to June in 2005 are better than in 2004.
- The plots of the indices for the 24-h accumulated precipitation for the threshold 0.5 mm in 2005 show smaller variability than in 2004.
- The quality of the forecast does not deteriorate significantly in subsequent days of the forecast.

3.6 Upper-air parameters

- Forecasts seem to be good as far as temperature and wind speed are concerned (ME about 0 and 1 m/s, respectively).
- Model still seems to be ”too wet” (ME about 30%, in extreme case relative error ∽50%).
- However, this looks like it improved, comparing results for 2003 with these for 2004, probably due to changes in (some) parameterizations.
- The quality of forecast, which is naturally expected tendency, decreases monotonously with time, especially for relative humidity. For other parameters this tendency is not so clearly seen.

References

Wilks,D.S., 1995 Statistical methods in the atmospheric sciences. Academic Press, New York

Application and verification of ECMWF products in Member States and Co-operating States. Report 1998.

Jolliffe I.T, Stephenson D.B, 2003, Forecast Verification. A Practitioner's Guide in Atmospheric Science, Wiley, London.

The new Lokal-Modell LME of the German Weather Service

J.-P. Schulz

Deutscher Wetterdienst, Offenbach, Germany

1 Introduction

In order to fulfill new requirements of both external and internal customers, for instance in aviation, sea traffic or air pollution modelling, the German Weather Service (DWD) decided to expand the model domain of the operational limited area model, the Lokal-Modell (LM, Doms and Schättler 2002, Steppeler et al. 2003, Schulz 2005). The new version has successfully been introduced in the operational numerical weather prediction system of DWD on 28 September 2005.

2 The model LME

The former version covered basically Central Europe, including Germany and its neighbouring countries. The new version covers almost entire Europe and therefore got the name LM Europe (LME). The integration domain of LME is shown in Fig. 1.

The number of grid points per layer is enhanced from 325 $\times$ 325 to 665 $\times$ 657, while the mesh size is kept unchanged at 7 km $\times$ 7 km. The number of vertical layers is increased from 35 to 40. The additional layers are mainly located in the lower troposphere, the height of the lowest layer is reduced from 33 m to 10 m. This is in accordance with the new 40-km version of the driving global model GME which started operation at DWD in September 2004. The poles of the rotated LME coordinate system are different from the LM system. The LME system is rotated in a way that the equator is located within the center of the model domain. This has the advantage that the grid cells have a similar size and shape throughout the entire domain or, in other words, the divergence of the longitude rows is minimal. The main non-technical model change is the introduction of a new multi-layer soil model, the same that was incorporated into GME in 2004.

3 Results

The introduction of LME at DWD was done in several steps. First of all, two experiments were set up at ECMWF in 2004, namely LME and LM, running daily forecasts driven by GME. Here, the influence of the domain size or the distance between the boundaries and the region of interest, respectively, can be tested. It turns out that in most weather situations there is very little influence. But, there are sporadic cases where for example the development of a cyclone evolves significantly differently. The results of an objective verification show some advantage for LME forecasts for precipitation and gusts and some disadvantage for mean sea level pressure.

In January 2005 a full LME data assimilation cycle has been set up in an operational parallel suite at DWD. This parallel suite also includes two 78h-forecasts (00 and 12 UTC) per day. Hence, LME could be tested in operational mode against LM and GME during spring and summer 2005. All postprocessing procedures had to be adjusted. Preliminary verification showed similar results as the experiments at ECMWF.

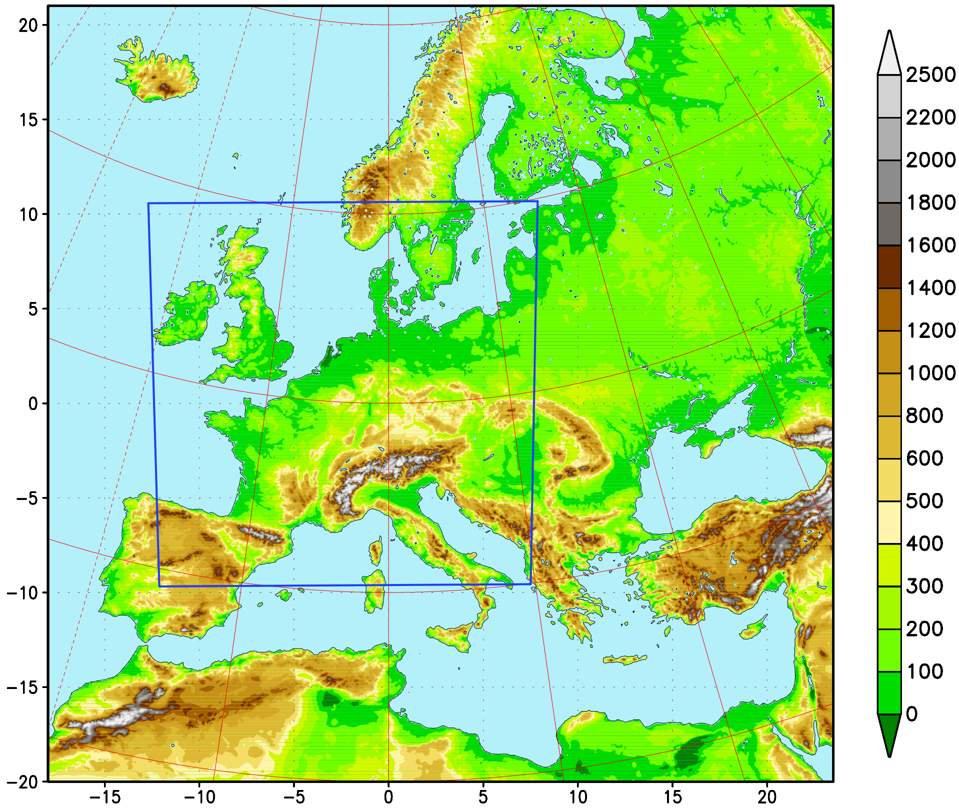


Figure 1: Model domain of LME. Topographical height (m) for land fractions > 50% (for the operationally used filtered orography). The frame in the figure depicts the integration domain of the former LM.

More detailed comparison revealed that the evaporation over sea in LME is up to 30% higher than in GME. Furthermore, precipitation in LME tends to show a systematic positive trend during the forecasts, even on a monthly mean basis, while precipitation in observations and also in GME is balanced. This behaviour indicates that evaporation over sea in LME is likely to be overestimated. Some sensitivity tests were carried out at DWD and a parameter tuning led to a LME version with reduced evaporation over sea. Preliminary verification of this version showed some improvement in the simulated moisture budget and also the mean sea level pressure.

A quantity of particular importance is the simlated soil moisture. First of all, it is a component of the new multi-layer soil model of LME and therefore needs to be monitored with great care. This has certainly been done already during the development of the model, but due to the very long memory of the soil with respect to e. g. temperature and soil moisture content, it is hard to run it long enough to ensure reasonable initial states for all its variables. Secondly, it is affected by the variational soil moisture analysis (SMA) scheme. This external analysis scheme is part of the data assimilation scheme of LME and is run once per day, at 00 UTC. It adjusts the soil moisture in an indirect way by minimizing the model bias of the near-surface air temperature. It has been switched on in LME in mid May 2005.

Figure 2 shows the simulated soil moisture of the third layer of the multi-layer soil model of LME compared to in-situ measurements from January to November 2005. Generally, the simulated soil moisture resembles the observations very well during most of the annual cycle of 2005, in particular during the first half of the year. After about day 205 LME becomes a bit drier than the observation, but the curves tend to converge again later in the year. But

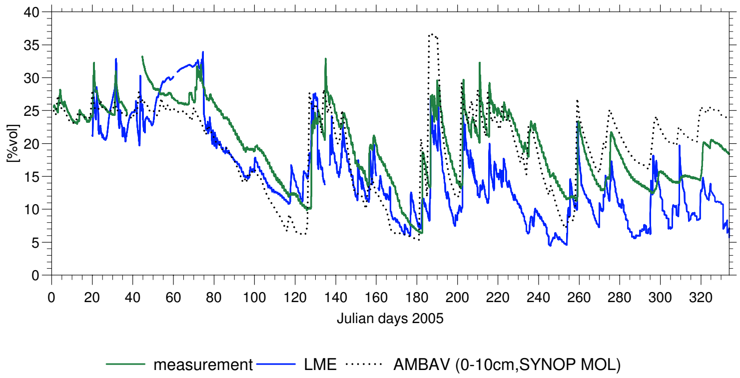


Figure 2: Soil moisture simulated by LME from January to November 2005 compared to measurements at the site Falkenberg (Meteorological Observatory Lindenberg, DWD). Shown is the soil moisture of the third layer (depth: 3–9 cm) of the new multi-layer soil model and a corresponding measurement. The third curve depicts results of AMBAV, which is a land surface scheme used at DWD for agricultural applications. It is more complex than the LME land surface scheme. For this study, it has been run off-line, forced by atmospheric conditions from the Meteorological Observatory Lindenberg which is close to the soil moisture measurement site. This figure was provided by G. Vogel, DWD.

during the entire period the tendencies of soil moisture variations are very similar between model and observation which is a good sign for the soil model performance. A likely reason for the difference between model and observation is that there was too little precipitation in the model at the beginning of the second half of the year. A hint for this is that the second model, AMBAV, run off-line with atmospheric forcing is much better able to follow the observed soil moisture evolution during this period.

4 Conclusions

In order to fulfill the requirements of several customers the German Weather Service (DWD) decided to expand the model domain of its operational limited area model, the Lokal-Modell (LM). The new LME, covering almost entire Europe, has successfully been introduced in the operational numerical weather prediction system of DWD on 28 September 2005. Current verification results look reasonable, further subjective and objective verification is carried out.

References

Doms, G. and U. Schättler, 2002: A description of the nonhydrostatic regional model LM. Part I: Dynamics and Numerics. *Deutscher Wetterdienst*, Offenbach, 134 pp. (Available at: www.cosmo-model.org).

Schulz, J.-P., 2005: Introducing the Lokal-Modell LME at the German Weather Service. *COSMO Newsletter*, No. 5, pp. 158–159. (Available at: www.cosmo-model.org).

Steppeler, J., G. Doms, U. Schättler, H.W. Bitzer, A. Gassmann, U. Damrath, G. Gregoric, 2003: Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteor. Atm. Phys.*, 82, pp. 75–96.

Flow Across the Antarctic Plateau

Ulrike Wacker[1], Hinnerk Ries[1], Ulrich Schättler[2]
[1]*Alfred-Wegener-Institut für Polar– und Meeresforschung, Bremerhaven*
[2]*Deutscher Wetterdienst, Offenbach am Main, Germany*

1 Introduction

Precipitation is an important parameter to control the growth and decay of the Antarctic ice shield. Due to lack of adequate measuring facilities, only few reliable observation data on precipitation sums are available. At present, precipitation climatology for Antarctica is only interpreted from annual ice accumulation observations (King and Turner, 1997). Therefore, precipitation simulation for Antarctic conditions with numerical weather prediction models (see e.g. Bromwich et al. (2001) and Powers et al. (2003) for the US American Antarctic Mesoscale Prediction System (AMPS)) and with climate models (e.g. van Lipzig et al., 2002; Bromwich et al., 2004) has gained much attention during the last years.

A striking feature in Antarctic precipitation climatology is the sharp inland decrease in precipitation amount from the order of 500 mm water equivalent to less than 50 mm on the central Antarctic plateau. The precipitation distribution in the coastal regions is strongly influenced by topography, namely the escarpment of the plateau with more than 2000 m height difference.

In this contribution, the influence of the plateau and its precipice on the distribution of clouds and precipitation will be investigated. To that end we perform and evaluate simulations with LM of DWD for Antarctic conditions. A weather situation is selected in which the air flow approaches the plateau nearly perpendicular. The effect of the excited gravity waves on the precipitation distribution will be investigated with the help of two-dimensional simulations of an idealized flow over a plateau.

2 Model Setup

The LM version V3.15 is implemented for the Antarctic region. For the case studies presented the domain is centered around the Neumayer Station (70°39'S, 8°15'W) with 321 × 201 gridpoints of about 7 km mesh size. The atmosphere is divided into 35 vertical layers of unequal thicknesses. The simulation is started on 10 January, 2002 0000 UTC and run for 48 hours. The initial and boundary data are obtained by interpolating ECMWF analysis and forecast data onto the LM grid.

The simulation of the idealized flow is performed with the two-dimensional version of LM (here: V2.19) for a channel of 22000 km length in x-direction. This length is sufficient to find in the interior a reasonably large area where gravity waves spread without any serious boundary effects. The topography is prescribed in form of a cosine-shaped slope towards the plateau of 2000 m height. As in the three-dimensional simulations, a 7 km mesh size and 35 vertical layers are chosen. The model run is started from a prescribed field and run for 10 days to reach a quasi-steady state; this state will be interpreted later. At the inflow boundary, the values are kept constant; at the outflow boundary, we assume a vanishing gradient in x-direction.

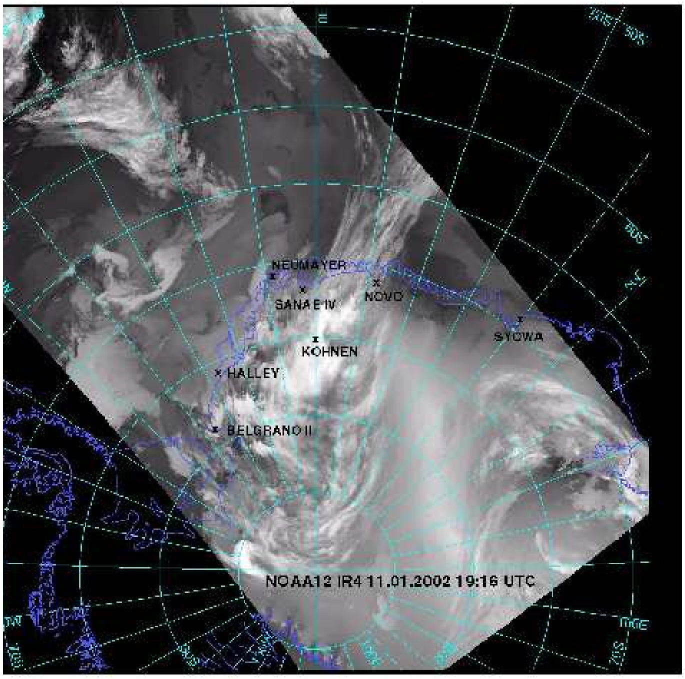


Figure 1: IR satellite image valid at 11 January, 2002 1916 UTC.

In both types of simulation we use the *cloud ice scheme* for parameterization of microphysics in grid scale clouds, as described in Doms et.al. (2005). The precipitation concentrations are calculated either according to the column equilibrium approach or by solving the full prognostic budget equation (Baldauf and Schulz, 2004).

3 Situation from January, 2002

The weather situation from 10 to 12 January, 2002, in the larger surrounding of Neumayer Station was characterized by a huge frontal system moving from the Drake passage in south-eastern direction toward the Antarctic coast and penetrating into Dronning Maud Land (DML); see the cloud bands in the satellite image valid for 11 January, 2002 1916 UTC given in Fig. 1.

The horizontal distribution of mean sea level pressure, given in Figure 2 (left) for simulation valid at 11 January, 2002 0000 UTC, shows the low pressure system with center over the Weddell Sea. The turning of the near surface wind direction to the left indicates the position of a cold front/occluded front. East of the front the flow at about 15°E over the ocean is nearly perpendicular to the plateau; close to the coast and inland the flow is deflected due to the mountains. As a result of this flow field, (see the horizontal distribution of relative humidity in Fig. 2 (right)) moist air is advected aloft by the northerly flow in the warm sector.

The meteorological observatory of the Neumayer Station provides 3-hourly routine synoptic observations. In Fig. 3 the observed vertical profiles of temperature and of relative humidity are given together with a time series of simulated profiles taken at the grid point nearest to the station. The thick green lines mark the simulated profiles for the time closest to the launch of the radiosonde.

The temperature profiles agree reasonbly well in the free atmosphere and show a stable stratification. In the lower troposphere, the simulated profiles evolve towards nearly isothermal

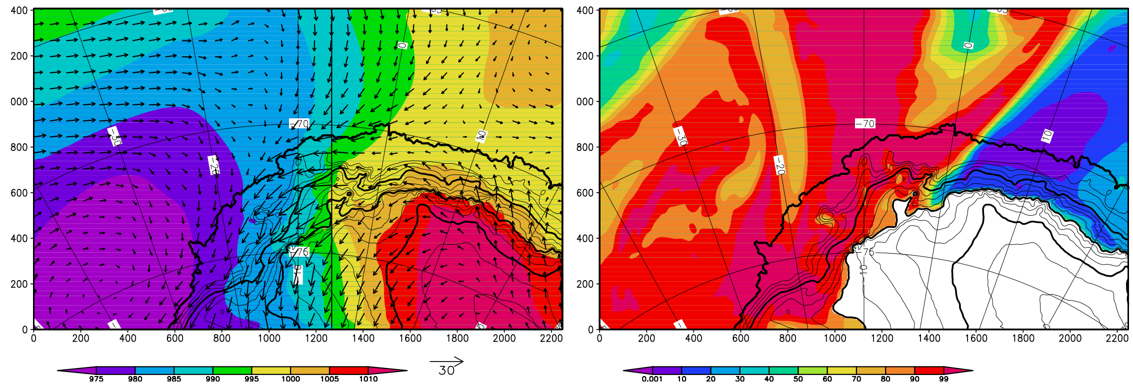


Figure 2: Horizontal distribution of simulated mean sea level pressure (in hPa) and near surface horizontal wind field (left) and of relative humidity (in %) at 2000 m above sea level (right). No relative humidity data are plotted for surface heights larger than 2000 m. Data follow from 24h simulation valid at 11 January, 2002 0000 UTC. Black lines give topographical heights in 250 m intervals starting at 0 m. Black vertical line at ca 1260 km in left figure marks position of vertical cross section for Fig. 5.

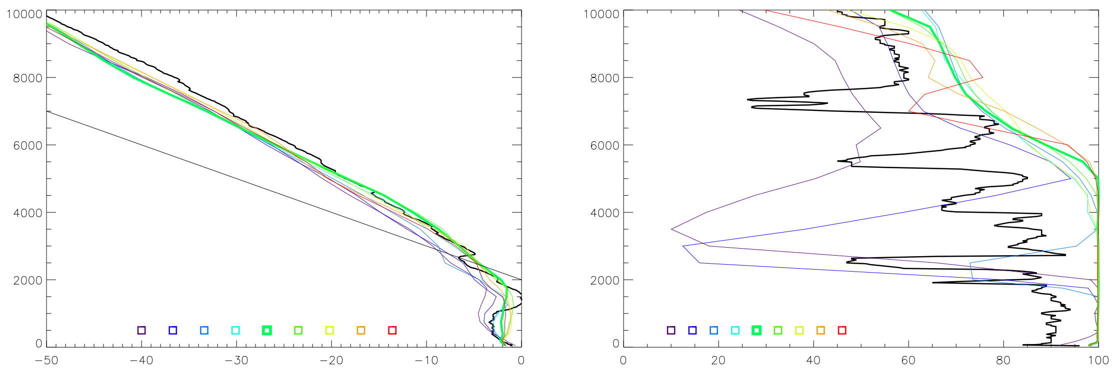


Figure 3: Vertical profiles of temperature (in °C; left) and of relative humidity (in %; right). Thick black lines give results from radiosonde launched at Neumayer Station on 10 January, 2002 1100 UTC. Coloured lines give simulated profiles from the model run started at 10 January, 2002, 0000 UTC for the grid point closest to Neumayer Station; data are valid from 00 h (purple) to 24 h (red) simulation time in three hourly intervals. Thin black straight line gives a dry adiabatic temperature profile.

ones; however, they do not show the observed inversion between 1000 to 1500 m height. The observed profile of relative humidity shows values of about 80 to 90 % in the lower and mid troposphere and a decrease above. In between, dry layers are observed with changes of up to 40 % over a distance less than 100 m. It is expected that such variations cannot be simulated with LM let alone due to the larger depth of the vertical layers. Although the simulations start with a dry mid troposphere, a water saturated lower and mid troposphere develops, probably due to the persistent advection of moist air and only few ice formation.

Figure 4 (left) shows the horizontal distribution of the simulated 6 hours precipitation sums. The distribution reflects the forcing due to the synoptic situation and topography. The latter is well seen in the precipitation enhancement on the upslope side of the plateau. The peak values with precipitation sums above 30 mm may be overestimated, since the annual mean precipitation sums amount here to only several hundreds of mm. From synoptic observations at the Neumayer Station, continuous moderate snow fall was reported on 10

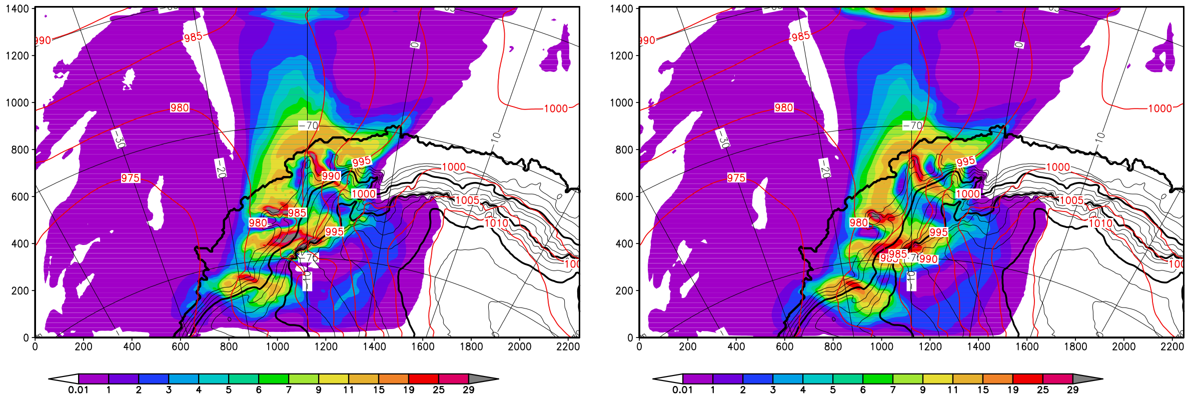


Figure 4: Horizontal distribution of simulated 6 hours precipitation sums valid for the 6 hours interval 10 January, 1800 UTC to 11 January, 0000 UTC. Left: Precipitation fluxes are calculated using the column equilibrium approach. Right: Precipitation concentrations are calculated using the full precipitation mass budget equation. Red lines denote mean sea level pressure (in hPa) for 11 January, 2002 0000 UTC.

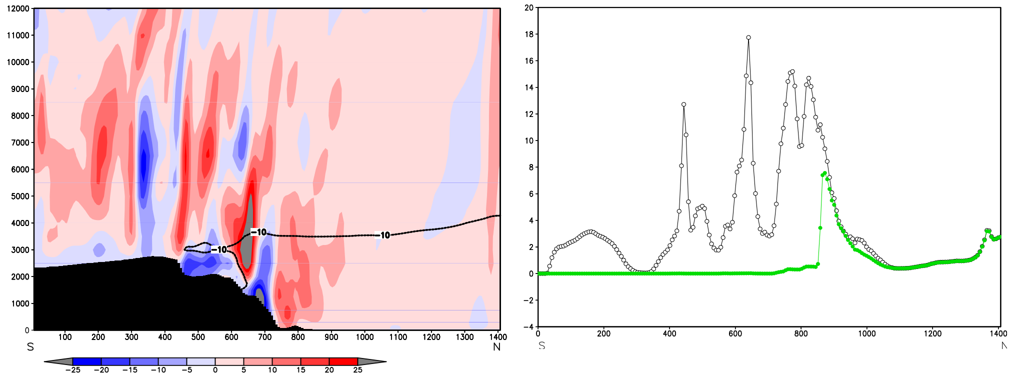


Figure 5: Left: Vertical cross section for vertical velocity w (in cm s^{-1}) in north-south direction along the line marked in Fig. 2 (left). Thick line gives $-$ 10°C isotherm. Simulated data are valid for 11 January, 2002 0000 UTC. x-axis gives distance from southern domain boundary in km. Right: Surface precipitation sum (in mm) along the cross-section for period as in Fig. 4. Green line indicates rain, black line indicates rain plus snow.

and 11 January, 2002 from 10 January, 2002 1500 UTC to 12 January, 2002 0900 UTC. No precipitation sums are registered. In the region on the plateau, which does not show strong topographic variations and which is yet unaffected by the approaching front, remarkable horizontal variations in precipitation are found.

The precipitation variations in north–south direction are inspected by looking at a cross section of vertical velocity and of precipitation (Fig. 5) along the line marked in Fig. 2. Certainly the w-field is strongly affected by topography. Over land the near surface wind is frequently directed downward along the sloping surface and thus has negative vertical component in those regions. Nevertheless, a wavy pattern can be recognized, whereby the axes of the upwind and downwind cells are tilted with height slightly towards the north, that is in upstream direction. This distribution suggests the presence of gravity waves.

Due to the topographically induced rising of the comparatively warm and moist air, the simulated cloud water and cloud ice cells are in close connection with the upwind pattern (not shown), with maximum concentrations at the downstream side of the upwind cells. The simulated 6 hour sum of surface precipitation, given in Fig. 5 (right), shows three successive peaks, each around 15 mm. In the cold air over the slope and the plateau all simulated precipitation falls as ice as expected, while over the ocean most of the falling precipitation reaches the surface as rain.

4. Gravity Waves in an Idealized Flow Field

Apart from the major mechanisms influencing the distribution of precipitation, that are the approaching front and the topographical details, we suggest that the velocity field and the precipitation pattern over the continent (as seen in Figs. 4 and 5 (right)) are also influenced by the occurrence gravity waves. To find out the pure gravity wave phenomena for a flow over a plateau, the results from two-dimensional simulations with idealized initial and topographic conditions are now presented. A topography is prescribed in a form to resemble the Antarctic plateau in DML. The surface height increases from zero to 2000 m height over a distance of approximately 500 km. The following case study is calculated, however, for conditions of 45°N.

The vertical cross section of the field of vertical velocity is given in Fig. 6 (left). A strong rising of air occurs immediately above the slope; further downstream one finds the bands of rising and sinking air with axes tilted upstream with height. Even though we consider a plateau, the cells are similar to the pattern found for mountain gravity waves excited by a single hill or a periodic sequel of hills, as assumed in typical idealized gravity wave studies (e.g. Smith, 1979). The w-fields in Fig. 5 (left) and in Fig. 6 (left) show a similar overall spatial pattern. They differ, however, insofar as that in the idealized case (i) the vertical velocities are generally weaker than in the case shown in Fig. 5), (ii) the cells above the plateau are much weaker than the cell over the slope, and (iii) the wave length, being of the order of 1000 km, is larger than in the case from Section 3. The differences may be attributed, at least partly, to the deviations in the topography, and certainly also to the deviations in the atmospheric general situation. Moreover, in the southern part of the model domain on the left hand side in Fig. 5 (left) the intensity of the wave seems to be damped; this, however, may also be due to boundary effects.

The precipitation pattern in the idealized simulation is shown in Fig. 6 (right). The highest precipitation sum is found above the escarpment, related to the strong upwind cell. Downstream, across the plateau, the intensity rapidly decreases below to about 1 mm, but due to the up-/downwind pattern, weak peaks are perceived.

5. Discussion

This study has reported on results from the first-time application of LM for Antarctic conditions. Bearing in mind that no specific modifications were introduced, the overall quality of the LM-simulations is satisfying.

The study reveals first informations on the order of magnitude and the horizontal distribution of precipitation on the Antarctic continent. The horizontal variability reflects strongly the synoptic and topographic forcing. On the plateau, where these forcings become weak, we do not find a monotonous decrease in precipitation. From the distribution of vertical velocity in combination with results of two-dimensional simulations we suggest the presence of topographically induced gravity waves. Weather situations, in which the northerly flow is nearly perpendicular to the plateau, are characteristic for the generation of gravity waves. The re-

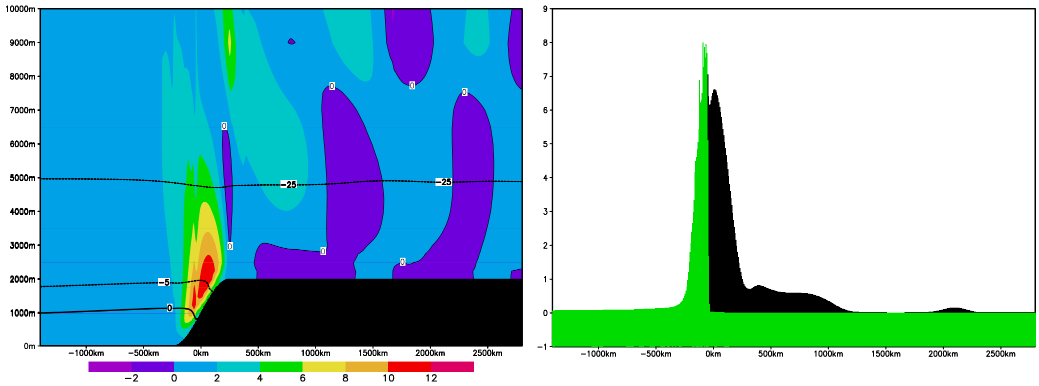


Figure 6: Left: Vertical cross section for vertical velocity w (in cm s^{-1}) for the idealized 2-dimensional quasi-steady state flow field. Horizontal flow is from left to right. x-axis gives distance from the center of the slope in km. Thick black line marks isotherms (in °C). Right: Surface precipitation sum (in mm within 6 hours) along the cross-section. Green shading marks rain, black shading marks rain plus snow.

lated upwind and downwind cells provoke a corresponding surface precipitation pattern with several precipitation peak values above the plateau. Although the peaks are strongly damped with increasing distance from the escarpment, they nevertheless inhibit a monotonous decrease in precipitation amount inland the Antarctic continent. This result from the idealized case study confirms the interpretation, that despite of the plateau-shaped topography, we should not expect a monotonously decreasing precipitation distribution downstream over the plateau in synoptic situations when the air flows perpendicular towards the plateau.

In the simulations discussed up to now, the precipitation flux is calculated by using the so-called column equilibrium approach. Fig. 4 (right) now shows the horizontal distribution of 6 hour precipitation sums, as follow if the full precipitation mass balance equation is solved. Both precipitation distributions given in Fig. 4 are very similar, since precipitation is mainly forced by the synoptic situation and topography. When using the prognostic treatment, however, the peak values of precipitation sums are damped and shifted downstream. With regard to the profile shown in Fig. 5 (right), the two peaks over the plateau are reduced to less than 15 mm, and all three maxima are shifted about 20 km downstream.

Certainly, several deficits in the simulations are obvious, although lacking the possibility of a proper verification. (i) The precipitation sums over the Antarctic plateau seem to be overestimated. (ii) Huge precipitation sums are found at the northern inflow boundary, see Fig. 4, in particular when using the solution of the full budget equation for precipitation mass; their origin should be looked for in the model physics of LM and of ECMWF-model, which provides the boundary data. (iii) Cloud ice is found mostly at levels where temperature drops below -25°C (not shown here) in the 3-dimensional as well as in the 2-dimensional simulations. This is in agreement with the LM results presented by Doms et al. (2004). It is speculated that the confine of cloud ice to such a threshold temperature may be related to an assumption in the cloud ice parameterization scheme: namely that cloud ice nucleation requires water saturation for $T > -25$°C, and only for $T \leq -25$°C ice saturation is sufficient for the initiation of cloud ice. This hypothesis is supported by the simulated vertical profile of relative humidity e.g. for 1200 UTC (see Fig. 3), which indicates a mostly water saturated troposphere for temperatures above about -20°C. Since the cloud microphysics parameterization used in the ECMWF model simulates clouds at temperature below 0°C already at water subsaturation, the prescription of ECMWF-model boundary data may cause a too dry model atmosphere for the temperature range $-25°C < T < 0°C$ with regard to cloud ice formation by LM.

Acknowledgement

The research was supported by the Deutsche Forschungsgemeinschaft (grant WA1334/4 and BI816/1).

References

Baldauf, M. and J.-P. Schulz, 2004: Prognostic precipitation in the Lokal Modell (LM) of DWD. *COSMO Newsletter*, No. 4, pp. 177-180. availabe at `http://www.cosmo-model.org`.

Bromwich, D.H., A.J. Monaghan, J.J. Powers, J.J. Cassano, H. Wei, Y. Kuo, and A. Pellegrini, 2003: Antarctic Mesoscale Prediction System (AMPS): A case study from the 2000/2001 field season. *Mon. Wea. Rev.*, 131, 412-434.

Bromwich, D.H., Z.Guo, L.Bai, and A. Chen, 2004: Modeled antarctic precipitation. Part I: Spatial and temporal variability. *J. Climate.*, 17, 427-447.

Doms, G., J. Förstner, E. Heise, H.-J. Herzog, M. Raschendorfer, R. Schrodin, T. Reinhardt, and G. Vogel, 2005: *A Description of the Nonhydrostatic Regional Model LM, Part II: Physical Parameterizations.* DWD, GB Forschung und Entwicklung, 2005.

Doms, G., D. Majewski, A. Müller, and B. Ritter, 2004: Recent changes to the cloud-ice scheme. *COSMO Newsletter*, No. 4, pp 177-180. availabe at `http://www.cosmo-model.org`

King, J.C. and J. Turner, 1997: *Antarctic Meteorology and Climatology.* Cambridge: University Press, 409 pp.

van Lipzig, N.P.M., E. v. Meijgaard, and J. Oerlemans, 2002: The spatial and temporal variability of the surface mass balance in Antarctica: Results from a regional climate model. *Int. J. Climatol.*, 22, pp. 1197-1217.

Powers, J.J., A.J. Monaghan, A.M. Cayette, D.H. Bromwich, Y.-H. Kuo, and K.W. Manning, 2003: Real-time mesoscale modeling over Antarctica. *Bull.Amer.Meteor.Soc.*, 84, pp. 1533-1545.

Smith, R.B., 1979: The influence of mountains on the atmosphere. *Advances in Geophys., Academic Press*, Vol.21, pp. 87-230.

An Objective Quality Measure based on a Pattern Recognition Technique to validate Regional Ensemble Forecasts

Christian Keil, George Craig

Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82234 Wessling, Germany

1 Introduction

The overall objective is an investigation into the predictability of convective storms, with the aim of determining the relative importance of several different sources of uncertainty. Such knowledge is needed for the design of an ensemble forecasting system, and eventually of an appropriate observing network. The work consists of ensemble simulations with differing sources of variability and development of objective methods of assessing accuracy of individual forecasts. Errors in regional forecasts often take the form of phase errors, where a forecasted weather system is displaced in space or time. For such errors, a direct measure of the displacement is likely to be more valuable than traditional measures, such as RMS error. A displacement measure is developed with a view to using it to explore the relative importance of various sources of uncertainty in regional ensemble forecasts.

2 The Regional Ensemble System

In general, a limited area model is influenced by the following sources of uncertainty:

(1) boundary conditions (uncertainty in the synoptic and meso-scale environment provided by a global model),

(2) initial conditions (uncertainty due to structures not seen by the observing system or due to limited resolution) and

(3) physical parameterizations (uncertainty resulting from the model formulation of convection, cloud microphysical, planetary boundary layer, or other processes).

The most obvious way to account for boundary condition uncertainty is to use a set of boundary conditions generated by a global ensemble forecasting system. Following the limited-area ensemble prediction system COSMO-LEPS methodology (Molteni et al., 2001; Marsigli et al., 2001), the high-resolution non-hydrostatic Lokal-Modell (LM) (Steppeler et al., 2003; Doms and Schättler, 2002) is nested on selected members of the global ECMWF EPS. It was found in this system that most of the variability in the 51 member ECMWF EPS for a region centered on the European Alps can be retained by as few as ten members (Marsigli et al., 2005). Here, 51 ECMWF EPS T255L40 ensemble members (Init. 2002070712 +72h fc) were down-scaled by a cluster analysis into 10 classes, for which LM experiments (Version 3.12) with 7 km horizontal resolution were conducted.

Secondly, using the new LM module LMSynSat (available from version 3.12 onwards) allows the production of synthetic satellite imagery. The synthetic satellite imagery is generated using the fast radiative transfer model for TIROS Operational Vertical Sounder (RTTOV-7), that allows fast simulation of brightness temperatures for various satellite radiometers

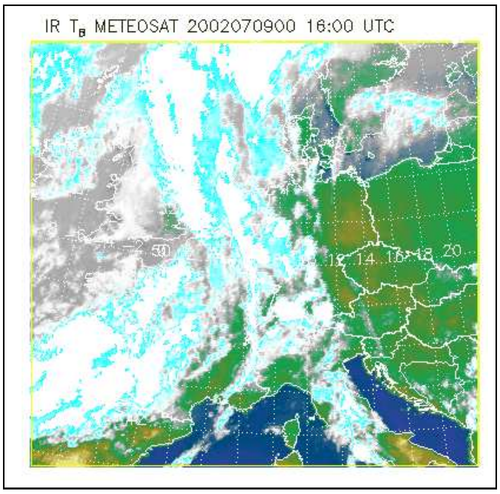


Figure 1: Observed Meteosat 7 IR imagery at 16:00 UTC 9 July 2002.

(e.g. Meteosat 7 MVIRI and Meteosat 8 SEVIRI). The input variables provided by LM are atmospheric profiles of temperature, specific humidity, cloud properties (cloud cover, cloud liquid water, cloud ice), specific content of snow and rain, and surface properties (skin temperature, temperature and specific humidity at 2m, land-sea mask). The output variables are clear and cloudy-sky radiance and brightness temperatures in IR and WV of Meteosat 7 and 8 channels of Meteosat 8. Sensitivity studies showed that a more realistic representation of clouds in LM can be achieved using the prognostic precipitation scheme (incl. precipitating snow) and a modified critical ice-mixing ratio (Keil et al., 2005).

Thirdly, using the model-forecast and the observed satellite image a field of displacement vectors is computed which 'morphs' the simulated image into a best match of the observed image. The magnitude of the mean displacement vector and the quality of the final match measured by the correlation give objective measures of the quality of the forecast.

After implementation, a case study observed during the VERTIKATOR field campaign (Vertical Transport and Orography, Lugauer and co-authors (2003)) in the northern Alpine forelands on 9 July 2002 has been examined.

3 Results and discussion

Ahead of an eastward propagating cold front, pre-frontal convection developed in the northern Alpine region in the afternoon of 9 July 2002. The cloud signature of a convective cell that has been initiated two hours before in the northern Alps is clearly visible in the IR image across southern Bavaria at 16:00 UTC (Fig. 1). The elongated cloud band across eastern France marks the cold front.

Model-forecast synthetic IR images of each representative member of the 10 clusters, the ensemble mean, and its spread are displayed for a subdomain in Fig. 2. While most of the clusters capture the synoptic scale cloud pattern (outside the subdomain, not shown) there are large differences in the pre-frontal convection and the position of the cold front at 16:00 UTC. Visual intercomparison of the observed and synthetic IR images gives a first, subjective ranking of the realism of the different clusters. The mean ranking based on a subjective evaluation by 8 scientists is given in Table 1. The top scoring of clusters 2, 10 and 4, reproducing the convection and the corresponding cloud signature in this region, can be

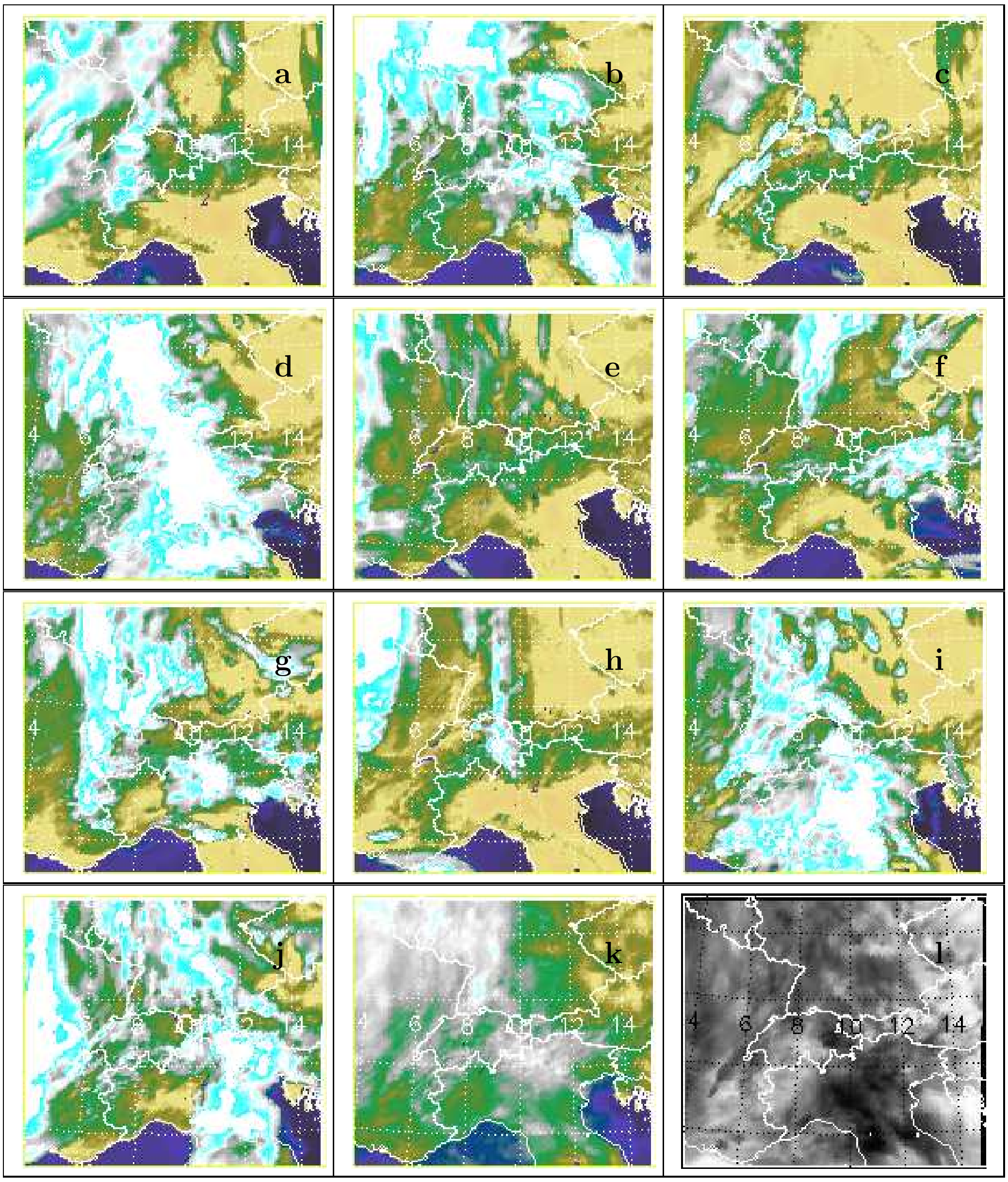


Figure 2: Forecast IR synthetic satellite imagery of the representative members of each cluster on 9 July 2002 16:00 UTC (LM +52h forecast range): (a-j) the individual ensemble members ranging from 1 to 10, (k) the ensemble mean and (l) ensemble spread (dark colored areas denote large spread).

confirmed in Fig. 2b,j,d. Likewise the low scoring of cluster 3 is evidently due to the severe underprediction of clouds in that area. The rank correlation between the scientists is quite high (0.85) confirming a good agreement among themselves and pointing towards a clear ranking of the clusters.

Next, the Pyramidal Image Matching technique (Mannstein, personnel communication) is applied to weight the different clusters according to their correspondence to the observed satellite imagery. In essence, differently coarse-grained pixel elements are compared. Starting at the largest scale (one pixel element containing 8×8 LM grid cells), a displacement vector field that minimizes the total squared error in brightness temperature within the range of ± 2 pixel elements, that is within a range of about 250 km, is computed. Subsequently, this image processing is done at successively finer scales (pyramidal). Finally a displacement

Table 1: Rankings of the 10 clusters according to (i) subjective eyeball evaluation of 8 scientists, (ii) the cluster population (number of members per cluster) and (iii) the objectively calculated forecast quality index.

Rank	1	2	3	4	5	6	7	8	9	10	rank corr.
subjective	2	10	4	7	9	1	5	6	8	3	0.85
population	3	2	4	1	5	7	6	8	9	10	-0.29
FQI	7	2	9	4	10	1	8	6	3	5	0.77

vector for every pixel is obtained from the sum over all scales.

However, the mean vector length of the displacement vector field contains no one-to-one information of the forecast quality. For instance, imaging a forecast showing no (or very few) cloud features at all (subjectively a forecast failure, e.g. cluster 3) would result in a mean displacement vector equal (or close to) zero. On the other hand, a perfect forecast would result to zero as well. Thus, a quality measure is constructed containing different measures of quality: the objectively computed mean displacement *displ*, the ratio of forecast and observed cloud occurrence CC_{LM}/CC_{Sat} (below a threshold brightness temperature), and the spatial correlation of observed and forecast-matched cloud structures *corr*. This measure is the normalized forecast quality index FQI, attaining zero for a *perfect* forecast:

$$FQI = 0.33 * [displ + (1 - CC_{LM}/CC_{Sat})_{+} + (1 - corr)].$$

Application of the Pyramidal Image Matcher on the cloud pattern at 16:00 UTC allows an objective ranking of each cluster that is shown in Table 1, too. Comparison of the subjective ranking and the one obtained from the object-oriented algorithm shows that the image matching provides a reasonable error measure for phase errors: the subjectively top-scored clusters are within the top five of the objective technique, while the lowest scored cluster agree reasonably well, too. The rank correlation between the average subjective and the objective ranking attains 0.77, confirming the consistent results of both rankings. In contrast, the ranking based on cluster population (number of members per cluster) shows no correlation with the other rankings (-0.29).

At 16:00 UTC, i.e. after +52 h forecast time, the variance of clouds is not only confined to the pre-frontal and frontal regions. Instead the ensemble spread shown in Fig. 2l shows a considerable variance of brightness temperatures in large areas of the domain. Due to the long forecast range, there is considerable noise in the forecast. The persistence of skill of individual members can be assessed by computing the rank correlation with different lead times. For this episode the persistence of skill is about 12 hours owing to a change of weather regime in the region. This new synoptic scale weather system moving into the region from 00:00 UTC onwards developed to the violent *Berlin* storm on 10 July 2002 (Gatzen, 2004).

4 Summary

The Regional Ensemble System currently developed at DLR consists of the following main components: (i) the COSMO-LEPS system, (ii) forward operators to generate synthetic satellite imagery based on model fields, and (iii) a pattern recognition algorithm to measure the quality objectively.

Validating the ensemble output of different episodes of pre-frontal summertime convection in Bavaria using the novel forecast quality measure FQI leads to the following conclusions:

- Pyramidal image matching provides a plausible measure of forecast error, which is consistent with subjective rankings.
- COSMO-LEPS cluster populations are a poor indicator of local skill.
- Persistence of skill is about 12 hours owing to change of weather regime in region.

In future, additional case studies from the Schwarzwald (moderate orography) and the southern UK will be simulated, to explore the performance of the system in different predictability regimes. Next to satellite observations, radar data will be utilized to validate the ensemble output using the radar forward operator developed by Pfeifer et al. (2004). An opportunity has arisen to implement a stochastic convective parameterisation in the LM (Craig et al., 2005), and if initial tests are successful, this will be used to compare its contribution to ensemble spread to that of the EPS boundary conditions.

Acknowledgements

Meteosat data are copyrighted by EUMETSAT. The authors thank Andrea Montani (ARPA-SIM) for his guidance to run COSMO-LEPS.

References

Craig, G.C., B.G. Cohen and R.S. Plant, 2005: Statistical mechanics and stochastic convective parameterisation. In: Workshop on Representation of sub-grid processes using stochastic dynamic models, ECMWF, Reading, England.

Doms, G. and U. Schättler, 2002: A Description of the Nonhydrostatic Regional Model LM. Deutscher Wetterdienst, Offenbach am Main.

Gatzen, C., 2004: A Derecho in Europe: Berlin, 10 July 2002, *Weather and Forecasting*, 19, pp. 639–645.

Keil, C. A. Tafferner, and T. Reinhardt, 2005: Synthetic satellite imagery in the Lokal-Modell, *Atmos. Res.*, accepted.

Lugauer, M. and co-authors, 2003: An overview of the VERTIKATOR project and results of Alpine Pumping, In: *Proceed. Int. Conf. Alpine Meteorol.*, Brig, Switzerland, pp. 129–132.

Marsigli, C., F. Boccanera, A. Montani, and T. Paccagnella, 2005: The COSMO-LEPS mesoscale ensemble system: validation of the methodology and verification, *Nonlinear Processes in Geophysics*, 12, pp. 527-536.

Marsigli, C., A. Montani, F. Nerozzi, T. Paccagnella, S. Tibaldi, F. Molteni, and R. Buizza, 2001: A strategy for high-resolution ensemble prediction. Part II: limited-area experiments in four Alpine flood events, *Quart. J. Roy. Meteor. Soc.*, 127, pp. 2095–2115.

Molteni, F. R. Buizza, C. Marsigli, A. Montani, F. Nerozzi, and T. Paccagnella, 2001: A strategy for high-resolution ensemble prediction. Part I: definition of representative members and global-model experiments, *Quart. J. Roy. Meteor. Soc.*, 127, pp. 2069–2094.

Pfeifer, M., G. Craig, M. Hagen and C. Keil, 2004: A polarimetric radar forward operator, In: *Proceed. Third European Conference on Radar in Meteorology and Hydrology (ERAD)*, Visby, Sweden, pp. 494–498.

Steppeler, J., G. Doms, U. Schättler, H.W. Bitzer, A. Gassmann, U. Damrath and G. Gregoric, 2003: Meso-gamma scale forecasts using the nonhydrostatic model LM, *Meteorol. Atmos. Phys.*, 82, pp. 75-96.

CLM - The Climate Version of LM: Brief Description and Long-Term Applications

U. Böhm[1], M. Kücken[1], W. Ahrens[2], A. Block[2], D. Hauffe[1], K. Keuler[2], B. Rockel[3], A. Will[2]

[1] *Potsdam Institut für Klimafolgenforschung, P.O. Box 60 12 03, 14412 Potsdam, Germany*
[2] *Brandenburgische Technische Universität Cottbus, Burger Chaussee 2, 03044 Cottbus, Germany*
[3] *GKSS Forschungszentrum Geesthacht, Max-Planck-Straße 1, 21502 Geesthacht, Germany*

1 Introduction

In 2001, LM was chosen as the basis for a new regional climate model called CLM. Although at this time numerous regional climate models already existed, there were at least two reasons for this decision. First, the non-hydrostatic formulation of the dynamical equations in LM without any scale assumptions made it eligible for use at horizontal grid resolutions of about 20 km and below, coming closer to the spatial scales requested by modelers of regional climate impacts. Secondly, the continuous development of LM at the DWD allows improvements in major LM versions to be adopted in the climate version ensuring that the central core is not a *frozen* one, but that of an up-to-date *living* forecast model. Meanwhile, we have succeeded in implementing extensions for long-term simulations into LM and in applying and evaluating the resulting CLM for multi-year climate reconstructions over Europe. Vice versa, these *climate mode* extensions will be made available in the forecast model as a new feature of LM 4.1.

In Section 2, we give a brief description of the major extensions implemented in LM 3.1 to obtain the CLM version 2.0. In Section 3, we present the CLM 2.0 setup for a 15-year simulation over Europe, give a short introduction into the developed validation strategy and assess the model's performance compared to analyses of the ECMWF (European Centre for Medium-Range Weather Forecasts), to observations and to results from simulations with other regional climate models under comparable conditions for reference. Finally, we summarize our findings, refer to the CLM as community model of the German regional climate modeling community and give an outlook on further model applications and recently started regional scenario constructions.

2 Brief Description of CLM Extensions

To enable LM for long-term simulations, the CLM community implemented several new features, some of them also of practical importance to the classical LM users. Therefore, all of them will be implemented into the upcoming LM version 4.1. In the CLM, we followed the general approach of adding switches to activate/deactivate the implemented extensions. In line with the general LM philosophy of modularity, all main climate extensions can be used by appropriate setting of the corresponding switches.

On climatological time scales, such model formulations as e.g. that describing the vegetation state of the soil cannot be assumed to be constant anymore as in the forecast mode. Therefore, we enabled the model to use not only initial values but also dynamic boundary data for `plcov`, `lai`, `rootdp`, `w_cl`, `t_cl` over land, for `t_s` and `qv_s` over sea, and for `vio3` and `hmo3` for the entire model area. To activate this feature, we implemented the logical switch lbdclim into the namelist group GRIBIN. Also, the switch lbdclim controls the output of variables with time range indicators 3 and 4 (mean and sum over the forecast time, respectively) which are reset after each output interval in order to avoid the calculation of output values as small differences of large numbers with limited numerical accuracy, as for instance for precipitation after 150 years model integration when estimating daily totals. There are intentions to extend the functionality of lbdclim to that of a *main climate mode switch*.

For climate change scenario simulations, the CO_2 concentration can be specified (constant at 330 ppm or an increase following the A1B and the B1 SRES scenario between 1950 and 2100, either for CO_2 only or composite CO_2).

We implemented a scale-selective type of relaxation, the spectral nudging. It can be activated and configured by setting parameters which we included in namelist group DYNCTL. In module organize_dynamics, a list is defined to indicate what set of boundary data are to be nudged. All subroutines performing the spectral nudging (initialization, spectral decomposition, nudging, spectral re-composition) are concentrated in module src_spectral_nudging.

For describing the thermal and hydrological processes in a deep soil, we use in the CLM Version 2.0 a modified DWD beta version of the multi-layer soil model. Different to this version, we apply this integration scheme in any case for both parts of the model (terra1_multlay and terra2_multlay). Having a thin first soil layer of about 1 cm thickness and a longer time step than 90 s, numerical oscillations of the surface temperature were encountered. To avoid such oscillations, a switch was implemented to allow restriction of the maximum change of t_so per time step to a user-defined value. Furthermore, the vertical soil moisture diffusion is restricted to layers above a certain depth, and a modified runoff computation is implemented. In namelist group PHYCTL, an additional parameter is introduced to externally control the lowermost depth for these hydrologically active soil layers. For the cases that there is snow cover, or there is no snow cover but t_snow is less than 0°C, the interception store water content is added to the snow store for consistency.

We also added parameters and routines that allow to write and read restart files and in this way to continue any model run.

Because of its portability, its self-describing character and because there is no risk of accuracy loss by data packing, we additionally implemented the NetCDF (Network Common Data Form) format for model input and output. In IOCTL, either NetCDF (following the CF, Climate and Forecast conventions) or GRIB1 can be selected individually for both the format of the initial and boundary data input files and that of the model output files. To handle NetCDF input and output, additional subroutines have been incorporated into the module io_utilities. Among them are all routines for opening/closing of NetCDF files, for reading and writing global definitions, variable-related attributes and data itself, but also for checking the input and output records.

Furthermore, several additional output variables were made available. In module data_fields, we defined equivalents for already existing model variables at 2m and 10m height indicated by the suffix _av with a time range indicator 3. They represent averages over the output interval (t_2m_av, td_2m_av, u_10m_av, v_10m_av) for use in namelist group GRIBOUT. Another 30 or so new variables have been defined for model output based on user demand. In addition to the cases already implemented in LM, the parameter ytunit in GRIBOUT is used to indicate a further 4-element date format (ytunit='d') including the month number in the file name convention (resulting date string: {yyyy}{mm}{dd}{hh}) set in subroutine make_fn.

In addition, further technical extensions are implemented. For example, in RUNCTL, the use

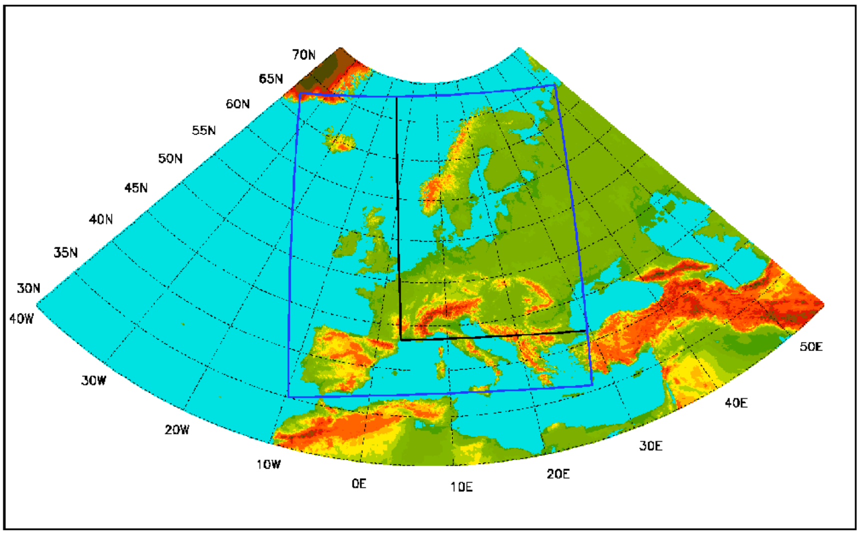


Figure 1: Extended BALTEX model region (blue borderlines) for ERA15-driven long-term runs over Europe. The original BALTEX model region is indicated by the black rectangle.

of a Gregorian or a climatological year can be specified.

3 Long-Term Climate Reconstruction over Europe

In a first long-term application, we used the CLM to reconstruct the climatic conditions over Europe for a 15-year period. For this purpose, we generated initial and lateral boundary conditions from ECMWF re-analysis data (ERA15). We set up the model for the so-called extended BALTEX region (Baltic Sea Experiment, `http://w3.gkss.de/baltex/`). The original and the extended BALTEX region which covers Europe almost completely are shown in Fig. 1.

We configured the multi-layer soil model with 10 layers and a total depth of $\sim$15 m. We used the climatological 2 m temperature as the lower boundary condition for the thermal part, which is in good agreement e.g. with observations at the station Potsdam. Initial soil moisture is set to 75% of the soil type depending pore volume in each grid box for all hydrological active layers. Initial and boundary data are generated for the entire ERA15 period 1979-1993. We used a horizontal resolution of 1/6 deg. (lat/lon) and 20 vertical levels in the pressure-based hybrid η-system. The north pole of the geographical grid is located at $\lambda = 170.0°$W and $\varphi = 32.5°$N. In this configuration, the lower left corner of the model area in rotated coordinates is located at $\lambda_r = 17.005°$W and $\varphi_r = 19.996°$S for mass grid points. We used 193×217 grid points which leads to the rotated coordinates for the upper right corner of $\lambda_r = 14.995°$E and $\varphi_r = 16.004°$N. We used a time step of 90 s. Boundary data are provided every 6 hours and relaxed using the Davies (1976) relaxation technique. The model output is stored with the temporal resolution of 6 h, too.

We elaborated a general verification strategy (Böhm et al., 2004) to assess the performance of a regional climate model using uni- and multivariate methods. Here, we concentrate on the model's ability to reproduce the regional climate exemplarily by evaluating the signed difference (bias) between the CLM results and suitable reference data for the three near-surface variables mean sea level pressure (MSLP), 2 m temperature and precipitation. Additionally,

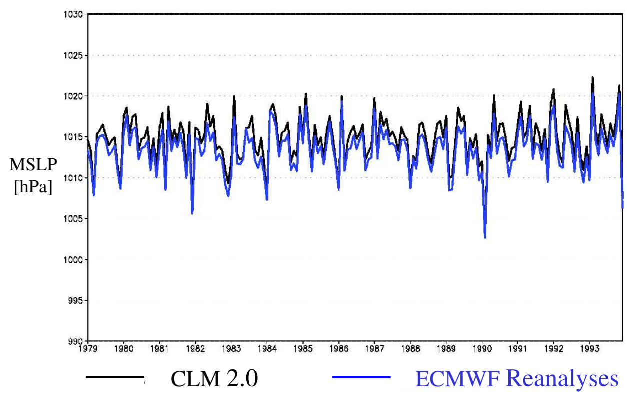


Figure 2: Area average of monthly mean MSLP 1979-1993.

we analyzed the accuracy of two temperature-related extremes – the number of summer days and the number of frost days – in the model results. For reference, we used ERA15 data, gridded observations that are compiled by the Climate Unit of the University of East Anglia (New, Hulme, Jones, 2000) and an observational data set of the German Weather Service. To compare the quality of the CLM results with those of other regional climate models that have been applied using the same forcing data, we supplied them to the QUIRCS project (Quantification of Uncertainties In Regional Climate and climate change Simulations, `http://www.tu-cottbus.de/meteo/Quircs/home.html`).

We analyzed the simulated mean sea level pressure as a representative of a rather large-scale climate variable. Fig. 2 shows that the CLM (black line) reproduces the ERA15 reference data (blue line) almost perfectly on average for the entire model area. Furthermore, there are no indications of a noticeable initial bias due to any cold start problem. The temporal mean of the area averaged MSLP generated by the CLM is, however, about 0.8 hPa higher.

In climate impact research, the 2 m temperature is a key variable. In Fig. 3, the 15-year annual mean is shown for both the CLM (left panel) and the CRU high-resolution gridded reference data set for most of the European land areas (right panel). Please note that the coordinate axes are labeled in rotated geographical coordinates in this figure and also in Figs. 4, 6, 7 and 9.

As a general result, there is evidence for the model's ability to reproduce the spatial patterns as represented in the gridded reference data set. This is especially true for mountainous regions.

However, the model underestimated the 2 m temperatures nearly everywhere over land. This becomes more clearly visible in Fig. 4 where the 15-year mean bias is shown. Whereas over Northern and North-Western Europe the smallest differences can be observed, the largest systematic deviations occur over land areas of high elevation in the south.

The already mentioned QUIRCS project addresses the question of how the identified model inaccuracies of different regional climate models compare. A final report is in preparation, and status reports are available on the above-mentioned home page. Here, we use the QUIRCS model results as an anonymous reference for the CLM. Figure 5 shows the mean annual cycle of the 2 m temperature for CLM 2.0, three other models and two reference data sets, averaged over the CRU region as indicated by the dark green-colored area in Fig. 9. Again, CLM 2.0 (black line) provides colder temperatures than both ERA (red line) and

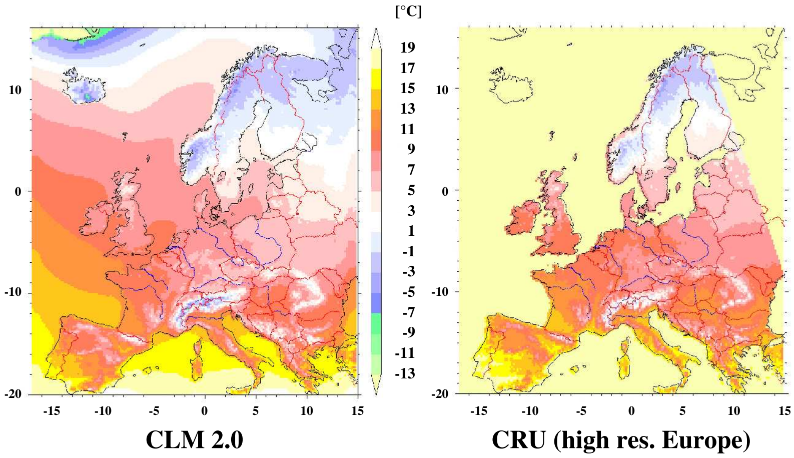


Figure 3: 15-year mean of annual mean 2 m temperature 1979-1993.

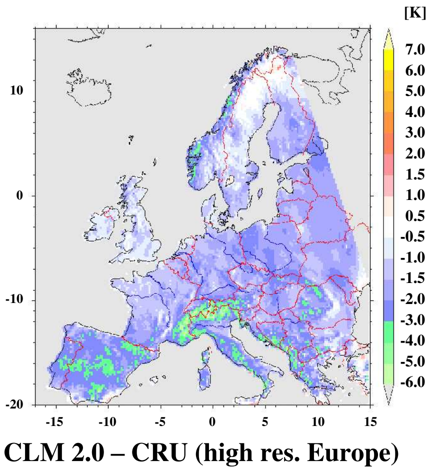


Figure 4: 15-year mean bias CLM – CRU of annual mean 2 m temperature 1979-1993 over the region as shown in Fig. 3, right panel.

CRU reference data (grey line) during summer. During winter, the bias compared to ERA is smaller than the difference between both reference data sets. Two of the three other models in Fig. 5 simulate too high temperatures especially during summer and early autumn with higher absolute differences for July and August than for the CLM. During winter, the other models range between the two reference data sets.

On average over the year, the CLM underestimates the 2 m near-surface temperatures by about 1.75K/ 0.67K compared to the CRU/ERA data set. Bearing in mind the different sign of the bias for two of the three other models during summer and winter compared to the CRU data set, this result indicates that the CLM provides an accuracy which is comparable to that of the other models participating in QUIRCS. The identified CLM bias lies within the range of about 2 K which is a typical order of magnitude for present-day regional climate models

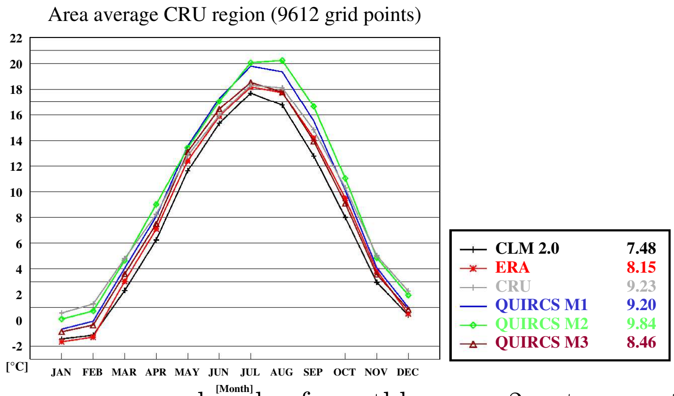


Figure 5: 15-year mean annual cycle of monthly mean 2 m temperature 1979-1993. CLM 2.0, ERA and CRU: as explained in the text. QUIRCS M1, M2 and M3: three other models participating in the QUIRCS project.

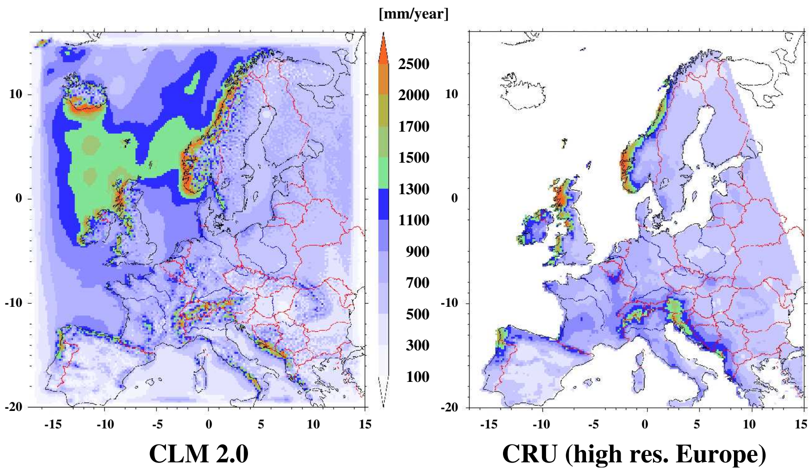


Figure 6: 15-year mean of annual total precipitation 1979-1993.

as concluded in the PRUDENCE project (Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects, see Christensen, 2005). Possible sources of this model inaccuracy are discussed after assessing the model's ability to reproduce the temperature-related extremes at the end of this section.

Although precipitation is a much more inhomogeneous quantity, Fig. 6 shows in the left panel that the model is also able to reproduce the major features of the spatial patterns in the CRU reference data set (right panel), especially over Great Britain, Scandinavia, the Alps, the northern Balkan region and northern Portugal.

Different to the 2 m temperature, however, there is no homogeneous picture for the sign of the bias. In Fig. 7 it becomes clearly visible that a gradient exists from the north - where the model simulates too much precipitation - to the south, which is reproduced as being too dry by the model, with the exception of some mountainous and coastal areas there. From subsequent investigations, there are indications that this north-south gradient of the bias at least partially may be caused by a too restrictive soil temperature change damping (not reported here).

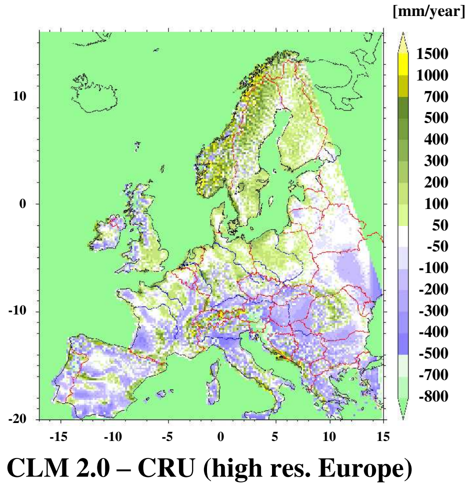


Figure 7: 15-year mean bias CLM-CRU of annual total precipitation 1979-1993 over the region as shown in Fig. 6, right panel.

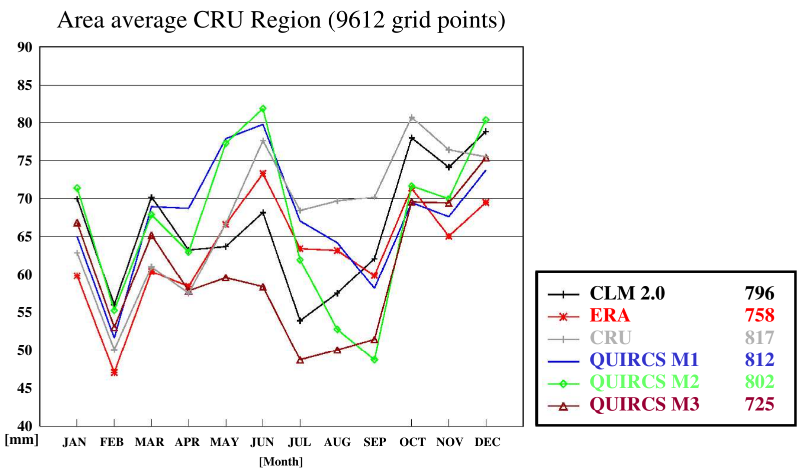


Figure 8: 15-year mean annual cycle of monthly total precipitation 1979-1993. Meaning of acronyms as in Fig. 5.

We also analyzed the relation of CLM's precipitation bias to that of other models participating in the QUIRCS project. In general, the variety between the models is much wider than in the case of the 2 m temperature, as it is obvious in Fig. 8, where the annual cycle of precipitation, averaged over the CRU region is shown in CLM 2.0 results, in the results of three other models involved in QUIRCS, and also in the ERA15 and CRU reference data.

The CLM provides some evidence of an overestimation of precipitation during winter and early spring, whereas the model underestimates precipitation during the rest of the year. Compared to the other models, the CLM is again within their characteristic range of uncertainty. We estimated an annual relative bias of -3% and of +5% in relation to the CRU and ERA data set, respectively. These values represent the lower boundary of errors as estimated for typical present-day regional climate models in the PRUDENCE project (Christensen, 2005).

In Kotlarski et al. (2005), the model intercomparison results for 2 m temperature and precip-

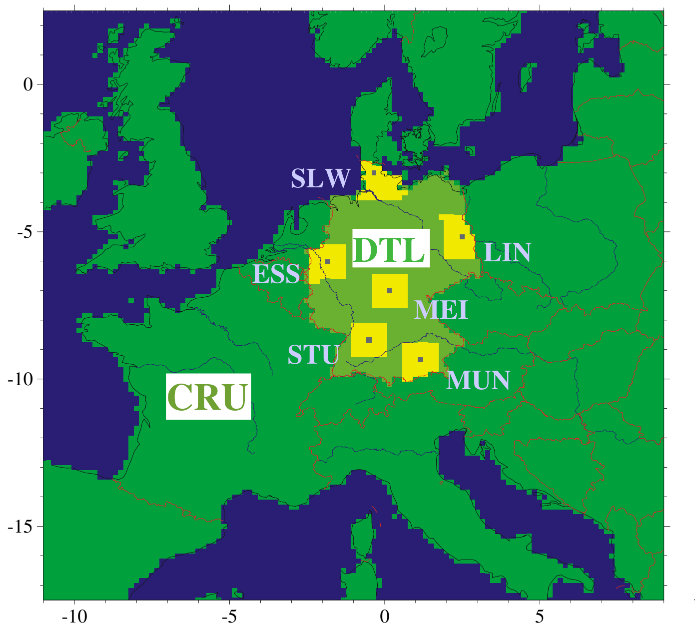


Figure 9: Regions for model verification. CRU: Land area covered commonly by all models and the CRU high-resolution reference data set. DTL: Germany SLW (Schleswig), ESS (Essen), LIN (Lindenberg), MEI (Meiningen), STU (Stuttgart) and MUN (Muenchen): smaller diagnosis regions centered around the indicated cities.

itation are described in more detail, for Germany in particular, and reveal the applicability of the CLM for regional climate simulations at spatial resolution of about 18 km (Kotlarski et al., 2005).

For climate impact research, extremes become more and more important. Especially during transitions from one climate to another one, there are indications that extremes occur more often and their amplitude intensifies.

Therefore, we assessed CLM's ability to reproduce temperature and precipitation extremes. Here, we concentrate on two examples of temperature-related extremes over Germany as a whole and over several smaller regions with specific climatic conditions ranging from north to south as shown in Fig. 9. We calculated the bias between the results of the CLM and a set of station observations from the German Weather Service, and compared it to the bias as estimated within the QUIRCS project for the same three models as listed in Fig. 5.

In Fig. 10, the number of summer days, i.e. the number of days with a daily maximum temperature equal to or higher than 25°C, is shown for the individual models and sub regions. The green circles representing the bias CLM results - DWD data set illustrate that the CLM performs well for DTL. The results for the smaller sub regions, however, give evidence that this outcome is based on averaging effects for larger differences with a slightly increasing tendency from north to south. Compared to other models, the CLM results are located at the lower limit of the range of errors.

The picture is different for the number of frost days with minimum temperatures below 0°C. Figure 11 shows that the CLM overestimates these extremes over almost all investigated regions and ranges at the upper error limit as estimated within the QUIRCS project.

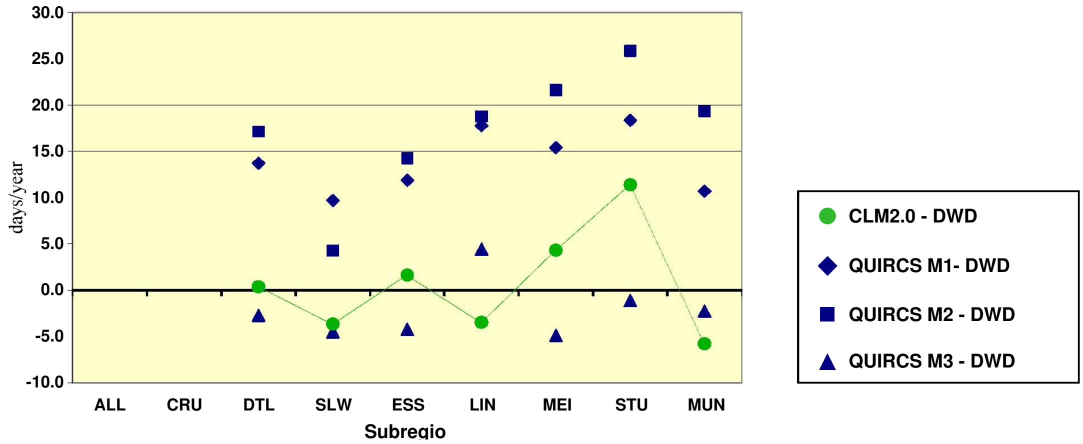


Figure 10: 15-year mean bias for the number of summer days per year for the CLM (green circles) and the models as indicated in Fig. 5, averaged for the regions as shown in Fig. 9.

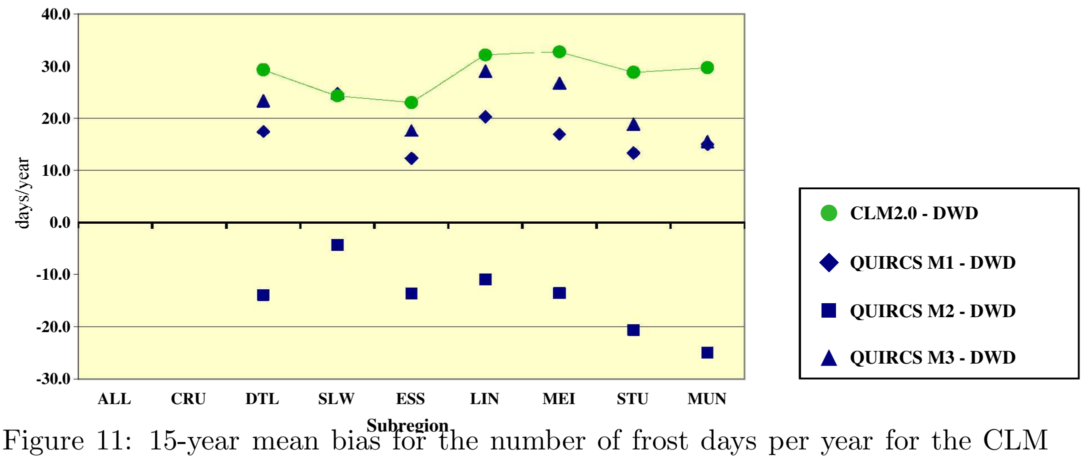


Figure 11: 15-year mean bias for the number of frost days per year for the CLM (green circles) and the models as indicated in Fig. 5, averaged for the regions as shown in Fig. 9.

This result gives evidence that the negative 2 m temperature bias in Fig. 5 compared to the CRU data set over Europe during winter exists also in relation to DWD observations over Germany and that ERA data seem to be too cold.

There are indications that the cold bias of the CLM may be caused at least partially by the strong damping of the soil temperature change of 2K/h in the multi-layer soil model which becomes important during summer. Subsequent analyses (not reported here) revealed, that this configuration leads to an imbalance of the surface energy budget. Therefore, we changed this limit to a less restrictive value (actually 20K/h in CLM 2.4).

Consistently with it, further analysis of the bowen ratio showed the sensible heat flux to be overestimated and the latent heat flux to be underestimated during spring, summer and autumn. This problem may, in turn, be also linked to the representation of root depths in the soil sub model in the CLM. Recent sensitivity experiments with a modified description of root depths provide changes in the bowen ratio in the right direction and are, therefore, encouraging. So far, however, no long-term simulations considering these changes are completed which would allow complementing the verification and intercomparison as described here.

4 Summary and Outlook

We extended the LM version 3.1 to the CLM version 2.0 to enable the model to perform long-term simulations. Most of the extensions are controlled by independent switches ensuring compatibility with the underlying forecast version. The most important new features of the CLM are: the usability of additional dynamic boundary conditions for vegetation and ozone parameter, for the surface temperature and humidity over sea and for the deep-soil lower boundaries, the possibility to specify changing CO_2 concentrations following two climate change scenarios, the spectral nudging technique, the use of an altered type of a modified DWD beta version of the multi-layer soil model, the possibility to continue a simulation from a user-defined model restart point, NetCDF following the CF conventions as an additional input and/or output format, the output of mean and total values for different output intervals instead for the entire forecast period only, the availability of additional output variables and the choice between a Gregorian or a climatological calendar year. They will be made available to the LM users in LM 4.1.

We presented results of the 15-year evaluation run of CLM 2.0 over Europe using ERA15 data as initial and boundary data. We evaluated atmospheric and near-surface variables for this simulation experiment. In this contribution, we discussed the model performance exemplarily for near-surface climate elements representing both mean conditions and extremes. The model simulated too cold temperatures, overestimated the precipitation over Northern Europe, and underestimated it over Southern Europe. The diagnosed bias for both 2 m temperature and precipitation is comparable to that of other present-day regional climate models and could at least partially be attributed to a too restrictive damping of the soil temperature changes in the applied multi-layer soil model, which could be overcome, however, in forthcoming model versions. Therefore, we conclude that CLM 2.0 is appropriate for simulating regional climates. Due to its non-hydrostatic formulation, we see the potential of the model, however, at finer horizontal grid sizes of about 10 km and below.

Based on the evaluation of the CLM results described here and on additional diagnostics of atmospheric variables, the Scientific Advisory Board (WLA – Wissenschaftlicher Lenkungsausschuss) of the German Climate Computing Centre (DKRZ) declared CLM as community model for the German regional climate modeling community. Linked to this declaration, a Community Agreement has been formulated and all interested scientists are invited to join the community and to contribute to the development of the CLM. Until now, scientists from 11 institutions have joined it. In addition, CLM has been selected to perform so-called consortial runs to generate small ensembles of transient regional climate scenarios for the period 1960-2100 at 0.165 deg. lon/lat resolution according to the SRES scenarios A1B and B1 (IPCC, 2001) forced by the global coupled model ECHAM5/MPIOM.

Acknowledgements

We are grateful to the European Centre for Medium-Range Weather Forecasts for providing the ERA data. Furthermore, we thank the Climate Research Unit of the University of East Anglia for providing the gridded monthly data set we used for model verification.

References

Böhm, U., Kücken, M., Hauffe, D., Gerstengarbe, F.-W., Werner, P.C., Flechsig, M., Keuler, K., Block, A., Ahrens, W., Nocke, T., 2004: Reliability of regional climate model simulations of extremes and of long-term climate. In: *Natural Hazards and Earth System Sciences*, 4, pp. 417-431.

Christensen, Jens Hesselbjerg, 2005: *Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects*, Final Report 2001-2004. `http://prudence.dmi.dk/public/publications/PRUDENCE Final report.pdf`

Davies, 1976: A lateral boundary formulation for multi-level prediction models. *QJRMS*, 102, pp. 405-418.

IPCC, 2001: Climate Change 2001: The Scientific Basis. Contribution of the Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (Houghton, J.T. et al., eds), Cambridge and New York 2001, pp. 609-612.

Kotlarski, S., Block, A., Böhm, U., Jacob, D., Keuler, K., Knoche, R., Rechid, D., Walter, A., 2005: Regional Climate Model Simulations as Input for Hydrological Applications: Evaluation of Uncertainties. *Advances in Geosciences*, Vol. 5, pp. 119-125.

New, M., Hulme, M. Jones, Ph., 2000: Representing twentieth century space-time climate variability. II: Development of 1901-1996 monthly grids of terrestrial surface climate. *J. of Climate*, 13, No. 13, pp. 2217-2238.

The Effect of Ice-Phase Microphysics on Tropical Cyclones Simulated by the Lokal Modell

Thomas Frisius, Thomas Hasselbeck, Fritz Herbert

Goethe-Universität Frankfurt a. M., Theoretische Meteorologie

1 Introduction

It is known that the release of latent heat is essential for the growth and the longevity of tropical cyclones. Furthermore, it is necessary for the maintenance of storms that the formed condensate must be removed from the vortex area before it may be evaporated in subsaturated air. Therefore, the formation of precipitation and fallout are important processes in tropical cyclones. In most situations, the ice phase is incorporated in the dynamics of precipitation. Examples are the Bergeron-Findeisen process, autoconversion, or collection of cloud water by snow or graupel (riming). High resolution mesoscale models like the Lokal Modell (LM) take these cloud microphysical processes into account in their cloud parameterization schemes. The current version LM 3.14 simulates five categories of hydrometeors, namely cloud water, rain, cloud ice, snow and graupel. Therefore, the LM is a suitable model for the investigation of the ice phase dynamics in tropical cyclones.

Willoughby et al. (1984) and Lord et al. (1984) simulated the ice phase in tropical cyclones using an axisymmetric nonhydrostatic model with a cloud parameterization scheme for the five categories of hydrometeors mentioned above. They found that the existence of the ice phase forces downdrafts outside of the eyewall. These downdrafts are generated by the withdrawal of latent heat due to melting of falling graupel. Wang (2002) investigated the impact of the ice phase in a threedimensional nonhydrostatic model. He found a slight decrease of intensity due to the existence of ice. Furthermore, spiral bands in the tropical cyclone are forced by downdrafts which are produced by melting of graupel.

In the present study the effect of the ice phase on tropical cyclones is investigated within the LM. It turned out that the prognostic treatment of precipitating cloud constituents is crucial for the formation of spiral bands, especially, when the ice phase is present.

2 Description of experiments

A detailed description of the LM-configuration and initial state of the experiments is given by Frisius (2004). For all simulations the LM Version 3.14 is used in which the effects of spherical geometry are neglected (f-plane geometry). The lower boundary is a sea surface with constant temperature $T_s = 28°\mathrm{C}$. The model domain extends over a length of 1120 km in the zonal and meridional directions. The horizontal distance between two grid points amounts to 6.95 km so that the domain is divided in each horizontal direction into 161 grid points. In vertical direction the 35 levels of the operational LM version (see Steppeler et al. 2003) are adopted. The model uses no parameterization of convection processes since we believe that hurricane structures can be assigned to the meso-β-scale that cannot be represented properly when parameterization of convection is switched on.

Initially, a circular symmetric balanced cyclone is placed at the center of the model domain. The initial cyclone has a central pressure deficit of $\Delta p = 5\mathrm{hPa}$ and a characteristic radius of

r_0 =150 km. The initial temperature distribution only depends upon z and is characterized by a linear decrease with the lapse rate $\gamma = 0.0065$ km^{-1} below the tropopause lying at a height of $H = 10$ km. The initial relative humidity F is uniform and the pressure is calculated from the hydrostatic balance equation. The initial horizontal wind is given in accord with gradient wind balance.

The experiments are divided into two types. One type of experiments contains simulations using the traditional cloud parameterization scheme with diagnostic treatment of precipitation which is based on a column-equilibrium relation for the precipitation fluxes. A second type of experiments refers to simulations with the refined cloud parameterization scheme that is based on a prognostic treatment of the precipitation categories. This new cloud parameterization scheme is able to forecast five categories of hydrometeors (cloud water, rain, cloud ice, snow and graupel) with the possibility to switch off some categories.

Numerical experiments have been performed for

- cloud water only (NOPREC),
- cloud water and rain (WARMRAIN),
- cloud water, rain and snow (1CATICE),
- cloud water, rain, snow and cloud ice (2CATICE),
- all five categories (3CATICE) .

The shortcut in brackets denotes the name of the respective experiment. Note that the LM cannot simulate the 3CATICE experiment with diagnostic precipitation. The model simulates 144 hours of the hurricane development and the integration time step is 20 seconds.

3 Results

Figure 1 shows the evolution of maximum wind velocity at the lowest model level for all experiments. It comes apparent that in every model run the cyclone reaches hurricane intensity. However, there are considerable differences between the simulations with diagnostically and those with prognostically calculated precipitation. The diagnostic precipitation simulations with ice phase reaches category 4 on the Saffir-Simpson scale (up to 60m/s) while the WARMRAIN experiment reveals only a hurricane of category 1 and for short time periods category 2. The situations are reversed for simulations based on the prognostic precipitation scheme. The model-hurricane in the WARMRAIN experiment attains category 3 while in the experiments including the ice-phase the simulated hurricanes are at best category two hurricanes. The NOPREC experiment also exhibits the development of a tropical cyclone but its spinup takes longer. However, this case will not be discussed further since the structures of this simulated cyclone are rather unrealistic (no eye-formation and heap up of cloud water).

To see how these differences can possibly be explained it is useful to take a look at the horizontal distribution of the cloud elements which are shown as snapshots at $t = 108$ hours in Fig. 2. Obviously, the different cloud structures are related to the different intensities of the storms. The simulations 1CATICE and 2CATICE based on diagnostic precipitation only reveal a single closed eyewall with a marked wavenumber two undulation. In contrast the diagnostic WARMRAIN experiment exhibits the evolution of spiral bands and an incomplete eyewall. In all experiments with diagnostic precipitation the precipitation pattern is correlated with the cloud water pattern. This results from the column-equilibrium assumption of the diagnostic precipitation scheme in which the fallout of precipitation takes place at the

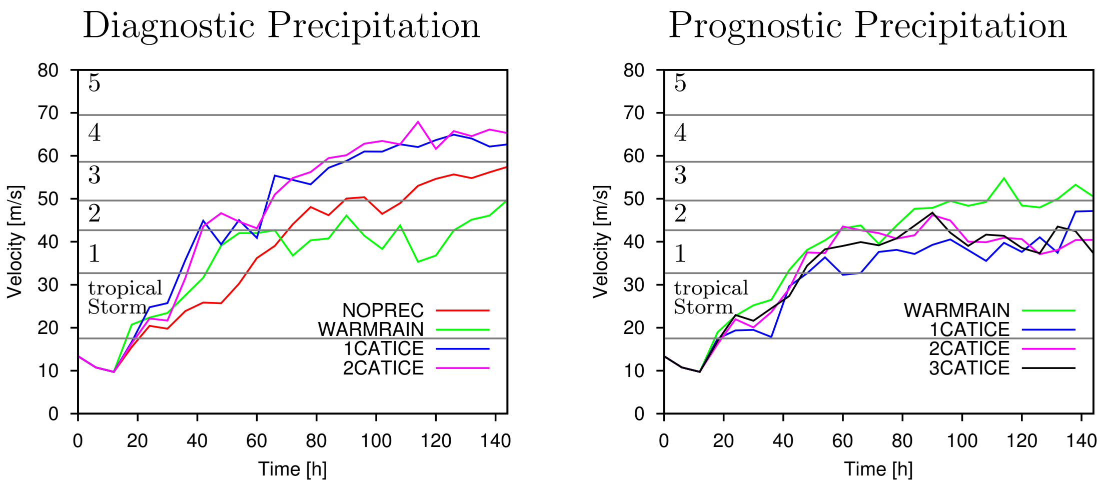


Figure 1: Maximum of wind at the lowest model level as a function of time for the various experiments. The left panel displays results for the cloud scheme with diagnostic precipitation and the right panel the results for the cloud scheme with prognostic precipitation. The thin horizontal lines bounds the areas of the Saffir-Simpson Hurricane scale which is divided into categories 1-5 (number on the left in the figure). The horizontal line at 17.5 m/s marks the minimum velocity of the tropical storm category.

same position where it is produced. The experiments 1CATICE and 2CATICE based on a prognostic precipitation do not show this correlation. The same is true for the 3CATICE experiment (not shown). With the prognostic scheme precipitable water can be advected to other locations. Especially, snow can spread out horizontally over large distance because the fall velocity is relvatively small. This may also be the reason why precipitation is still correlated with cloud water in the prognostic WARMRAIN simulation since slowly falling snow is not present. This can be verified in Fig. 3 where the various azimuthally averaged hydrometeor categories at $t = 108$ are displayed in a cross section for the experiments WARMRAIN, 1CATICE, 2CATICE and 3CATICE based on the prognostic precipitation scheme. Rain forms below the wide snow maximum in the experiments including the icephase while rain only falls close to the eyewall in the WARMRAIN experiment. Note that the amount of graupel is very small relative to the amount of snow. Therefore, graupel does not play an important role for the dynamics of the mature tropical cyclone of the 3CATICE simulation. The same is true for the cloud ice since the experiments 1CATICE and 2CATICE show quite similar results. Possibly cloud ice is only important for the initiation of snow production while the major snow production conversions are rather riming and depositional growth than autoconversion (for details of the ice-phase parameterization see Doms et.al., 2005).

It can be stated that the experiments showing only a single eyewall are associated with a larger storm intensity (maximum wind speed) than in the experiments showing cloud structures outside the eyewall. However, the physical reasons for these differences remain unclear so far. The results of the experiments with prognostic precipitation are in agreement with the findings of Wang (2002). He accentuates the advection of cold air into the boundary layer by downdrafts which limits the intensity of the storm. Figure 4 shows the downdraft velocity at z=1240m together with the equivalent potential temperature field for both 2CATICE experiments. Considerable qualitative differences can be seen. In the simulation with diagnostic precipitation downdrafts occur near the eyewall while with prognostic precipitation downdrafts appear far away from the storm center. Furthermore, the downdraft in the experiment with prognostic precipitation cools the boundary layer equivalent potential temperature up to 40 K below the ambient values. It can be expected that this cooling

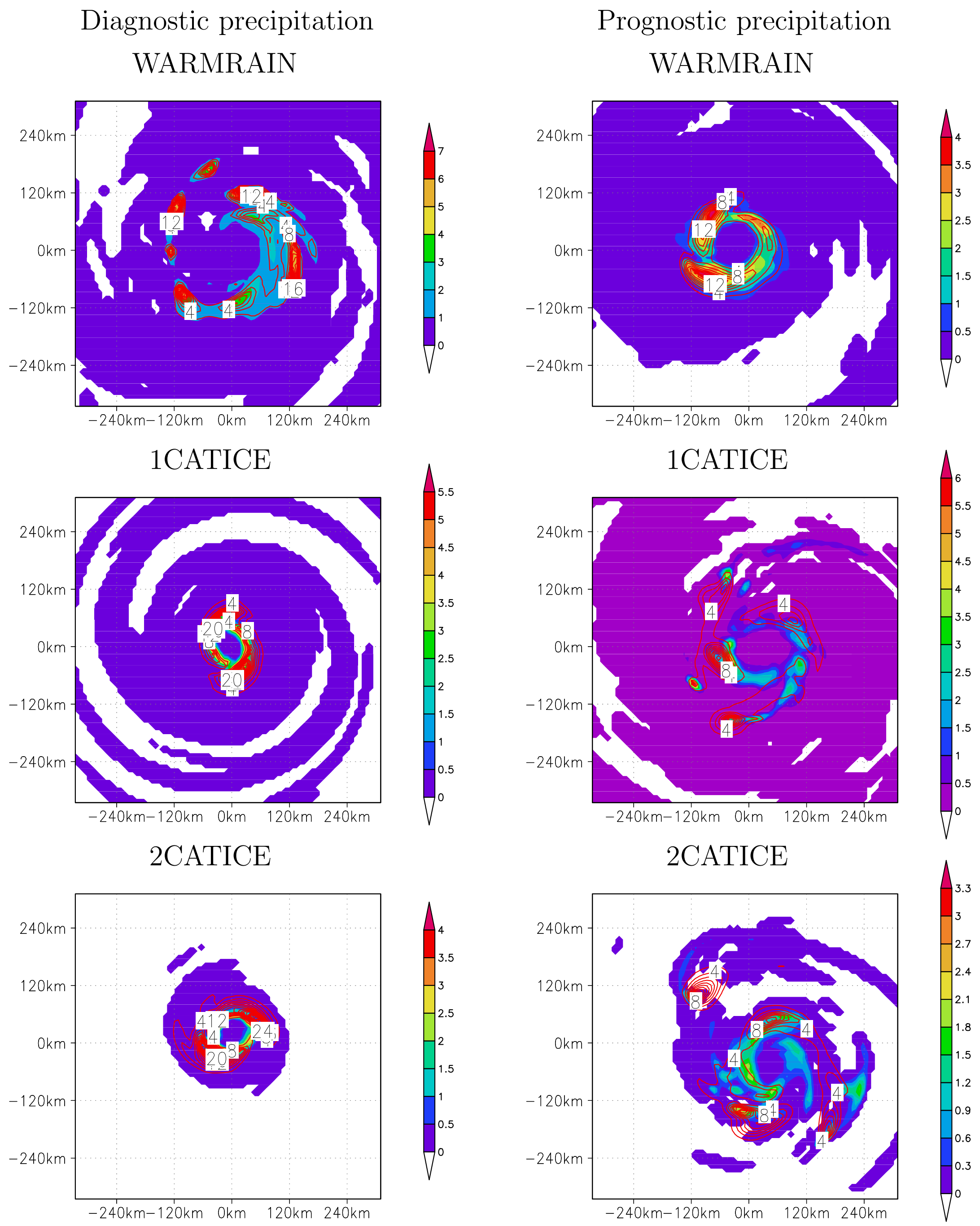


Figure 2: Snap-shots of the vertically integrated cloud water in kg/m^2 (coloured shadings) and vertically integrated precipitable water (red isolines, contour interval 2 kg/m^2) at $t = 108$h. The left (right) panel displays results for the cloud scheme with diagnostic (prognostic) precipitation.

effectively reduces the intensity of the storm since the boundary layer air ascends in the eyewall later on. There, it will release a smaller amount of latent heat than without this downdraft-induced cooling. Most downdrafts are associated with enhanced values of the vertically integrated melting rate (thick white isolines) suggesting that these downdrafts are indeed initiated by melting of snow. The same is true for the rain evaporation rate (not shown). Both processes contribute to a negative buoyancy of the air parcel. In the model experiment based on diagnostic treatment of precipitation no such downdrafts occur since the produced snow cannot be advected away from the eyewall due to the limitations of the

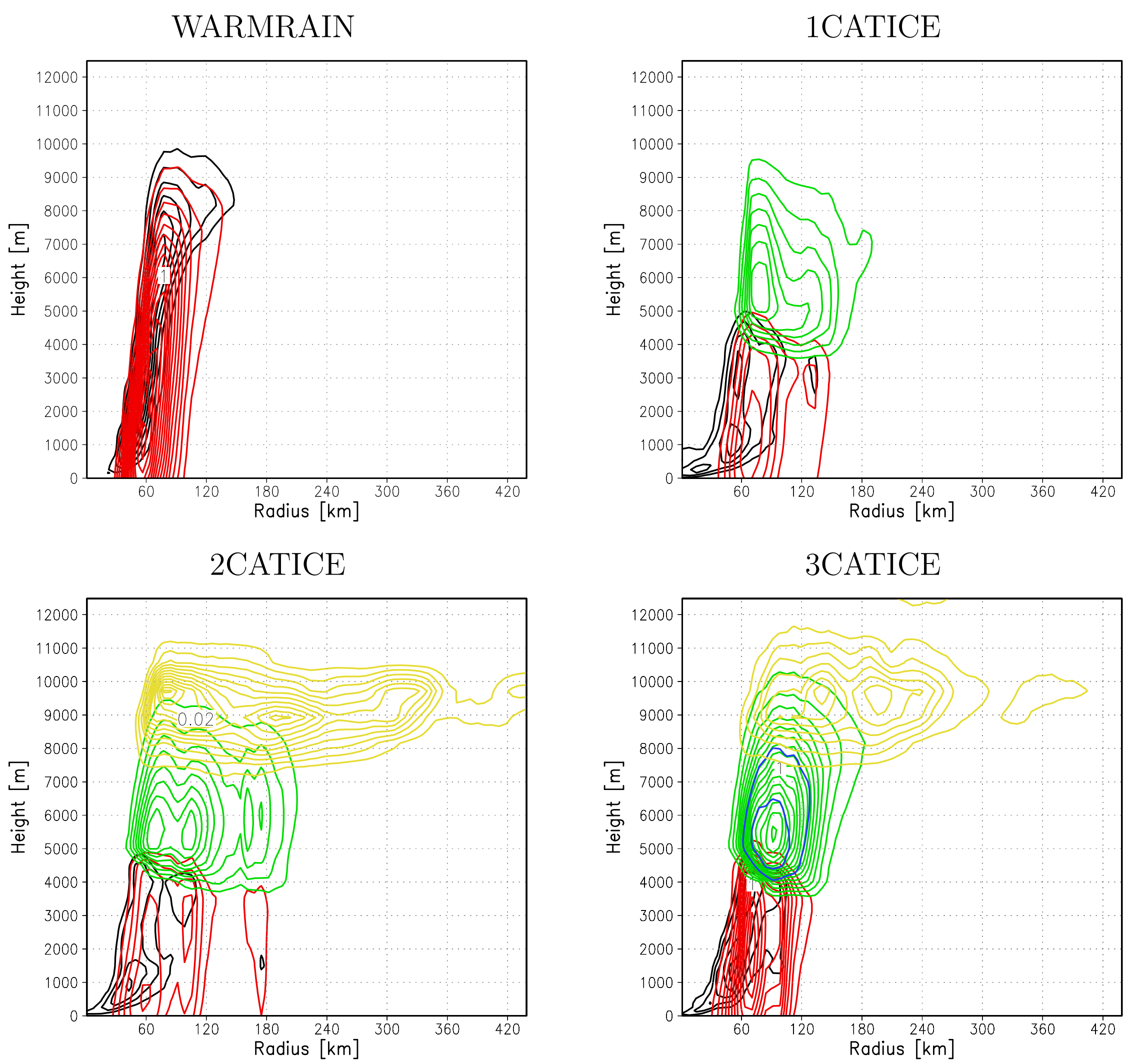


Figure 3: Azimuthal averages at t = 180h of cloud water content (black isolines, contour interval 0.1 g/kg), rain content (red isolines, contour interval 0.1 g/kg), snow content (green isolines, contour interval 0.1g/kg) cloud ice content (yellow isolines, contour interval 0.002g/kg) and graupel content (blue isolines, contour interval 0.1g/kg). Shown are the results of the experiments based on the prognostic treatment of precipitation.

diagnostic scheme.

4 Conclusion

This study reveals that the LM-simulation of tropical cyclones is very sensitive with respect to the details of the microphysical cloud parameterization scheme. The ice phase provides downdrafts due to melting of snow that reduces the intensity of the simulated storm. It seems that the diagnostic precipitation scheme is not capable in producing downdrafts by evaporation of rain and melting of snow. Therefore, no further convective cells are generated outside of the eyewall and no cooling of the boundary layer takes place. This leads to a higher storm intensity. Since cloud structures outside of the eyewall, for instance rainbands, are evident in real tropical cyclones (e. g. Anthes 1982) the cloud parameterization scheme with a prognostic treatment of precipitation is more suitable for tropical cyclone modeling than the old diagnostic scheme.

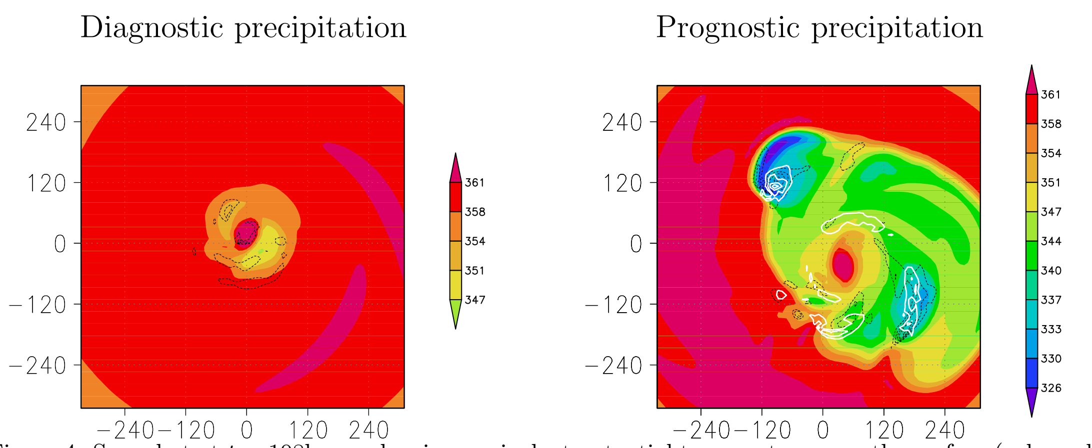


Figure 4: Snapshot at $t = 108$hours showing equivalent potential temperature near the surface (coloured shadings) and downdraft velocity at $z = 1240$m (black dashed isolines, contour interval 0.2m/s). The thick white solid isolines in the right panel display contours of the vertically integrated melting rate. (contour interval $2\text{kg/m}^2\text{/h}$). The left panel displays the 2CATICE experiment based on the diagnostic treatment of precipitation and the right panel the 2CATICE experiment based on the prognostic treatment of precipitation.

Acknowledgments

The authors gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) within the SFB 641 (Die troposphärische Eisphase). We also thank the DWD (Deutscher Wetterdienst) for making the LM available and the use of their high performance computer and the DKRZ (Deutsches Klimarechenzentrum) where parts of the calculations have been performed.

References

Anthes, R. A., 1982: Tropical cyclones - their evolution, structure, and effects. *Meteorol. Monogr.*, 41, 208pp.

Doms, G., J. Förstner, E. Heise, H.-J. Herzog, M. Raschendorfer, R. Schrodin, T. Reinhardt, and G. Vogel, 2005: *A Description of the Nonhydrostatic Regional Model LM, Part II: Physical Parameterizations.* DWD, GB Forschung und Entwicklung, 2005.

Frisius, T., 2004: Numerical simulation of tropical cyclogenesis with the Lokal-Modell. *COSMO Newsletter*, No. 4, pp. 189-196.

Lord, J. S., H. E. Willoughby and J. M. Piotrowicz, 1984: Role of a parameterized ice-phase microphysics in an axisymmetric, nonhydrostatic tropical cyclone model. *J. Atmos. Sci.*, 41, pp. 2836-2848.

Steppeler, J., G. Doms, U. Schättler, H. W. Bitzer, A. Gassmann, U. Damrath, 2003: Mesogamma scale forecasts using the nonhydrostatic model LM. *Meteorol. Atmos. Phys.*, 82, 75-96.

Wang, Y., 2002: An explicit simulation of tropical cyclones with a triply nested movable mesh primitive equation model: TCM3. Part II: Model refinements and sensitivity to cloud microphysics parameterization. *Mon. Wea. Rev.*, 130, pp. 3022-3036.

Willoughby, H. E., H. L. Jin, S.J. Lord, and J.M. Piotrowicz, 1984: Hurricane structure and evolution as simulated by an axisymmetric nonhydrostatic numerical model. *J. Atmos. Sci.*, 41, pp. 1169-1186.

Application of the LM to Investigate the Sharpness of the Extratropical Tropopause

Tamás Szabó, Volkmar Wirth

University of Mainz

1 Introduction

At the University of Mainz the *Lokal Model* (LM) is used to investigate unexplained climatological features of the extratropical tropopause. Previously, Birner et al. (2002) found that observed average profiles of buoyancy frequency squared (N^2) in the extratropics feature a sharp peak just above the tropopause corresponding to a strong thermal inversion in that region (Fig. 1). Their results were based on (1) a large number of high resolution radiosonde measurements and (2) an averaging method that considers the thermal tropopause as a common reference level. The inversion layer was found to have a thickness of 500 – 3000 m and showed characteristic variations with season and latitude (Birner 2005). The reason for the extreme sharpness of the extratropical tropopause must be considered as an unsolved problem. There are several processes which may contribute to this phenomenon, such as balanced dynamics, gravity waves, or radiation at cloud tops. However, the degree to which any of these processes is relevant in a climatological sense is unknown. In our work we concentrate on purely dynamical mechanisms, investigating the impact of synoptic scale dynamics on the tropopause sharpness as well as the potential impact of gravity waves.

2 Theoretical background

A possible explanation for the observed tropopause sharpness was suggested by Wirth (2003), implying balanced dynamics. He found different profiles of N^2 in idealized axisymmetric

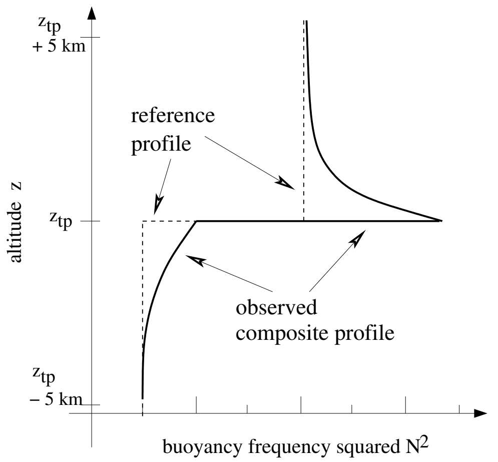


Figure 1: *Schematic representation of profiles of buoyancy frequency squared in the extratropical tropopause region. (Wirth 2004)*

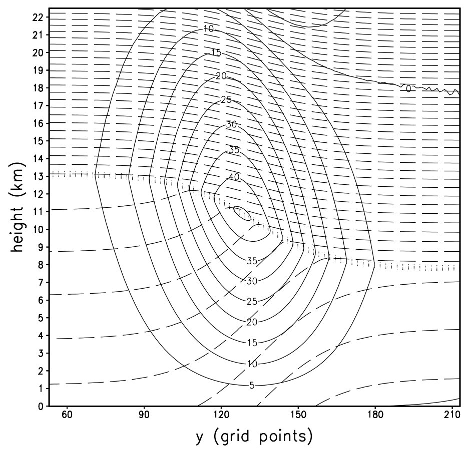


Figure 2: *Initial condition for the LM runs characterized by a jet (the solid lines depict the zonal wind in m/s) and piecewise constant N^2 (dotted). The dashed contours depict potential temperature.*

cyclonic and anticyclonic anomalies, which were produced with the help of a potential vorticity (PV) inversion technique. His anticyclones showed a sharp peak in N^2 just above the tropopause as well as increased values of N^2 below the tropopause compared to the reference profile. In contrast, his cyclones showed a smooth transition of N^2 in the tropopause region between tropospheric and stratospheric values. These features could be related to the specific partitioning of a given PV anomaly into a static stability anomaly and a vorticity anomaly, which differs between cyclonic and anticyclonic anomalies. The average N^2 from a large number of modeled profiles turned out to be qualitatively similar to the observed composite profile, showing in particular the peak just above the tropopause.

In a subsequent paper Wirth (2004) investigated the underlying mechanism. It was shown that the above-mentioned profiles are generated during the anomaly formation through the induced secondary circulation, and that the convergence of the vertical wind plays a major role during this process.

3 Model experiments and results

It is the aim of the current work to examine the proposed mechanism in a more realistic framework by performing simulations of baroclinic wave development. For this purpose the LM was configured in a channel version. Numerical experiments were carried through starting from an initial state (Fig. 2), which consists of a perturbed jet flow with piecewise constant buoyancy frequency in the troposphere ($N_t^2 = 1 \times 10^{-4}\mathrm{s}^{-2}$) and in the stratosphere ($N_s^2 = 4.5 \times 10^{-4}\mathrm{s}^{-2}$). Different model resolutions were tested, the highest of which was 0.3° in the horizontal and 125 m in the vertical direction. The high vertical resolution proved to be necessary to resolve the small scale processes in the tropopause region. In the experiments the atmosphere was dry, and a flat topography was used.

Figure 3 shows the simulation after 7 days of integration. Red and yellow colours indicate a sharp tropopause with large values of N^2 above the local tropopause, while blue and purple colours correspond to a smooth tropopause. As exemplified in this figure, the model simulations generally show a sharp tropopause in ridges and anticyclonic areas, but a smooth tropopause in troughs and cut-off cyclones. The main reason for this proved to be the convergence and divergence of the vertical wind generated during the baroclinic wave development. Figure 4 shows that a sharp tropopause appears downstream of the convergence region in an earlier stage of the wave development. This is in good qualitative agreement with the more idealized studies mentioned above.

Composite profiles of N^2 for the field in Fig. 3 are displayed in Fig. 5. The average over the entire area (red) indicates that by this stage of development the tropopause has overall slightly sharpened in comparison with the specified reference profile (black). Apparently, this is a result of averaging profiles from different locations with locally sharp or smooth tropopauses (cf. the blue and green lines in Fig. 5 and the profiles in Figs. 6 and 7). Again, this is in good qualitative agreement with the earlier studies.

However, there is at least one further mechanism in the LM simulations affecting the tropopause sharpness in addition to balanced dynamics. The single profiles in Figs. 6 and 7 indicate the presence of waves with short vertical wavelength, having a significant influence on the sharpness of the tropopause. In our simulations these waves could be identified as inertia-gravity waves, which are generated along the jet in the tropopause region during the baroclinic wave development (Figs. 8 and 9). This is in agreement with observations from the radiosonde network, which frequently show the presence of large amplitude gravity waves in the tropopause region. We tentatively conclude that these waves can influence the precise

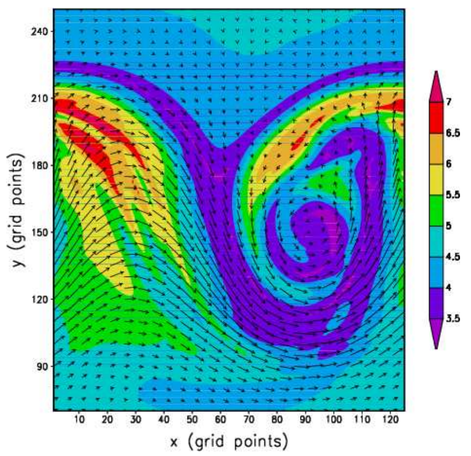


Figure 3: *Quasi-horizontal distribution of N^2 (colours, in 10^{-4}s^{-2}) and wind vectors on a surface which is 750 m above the tropopause after 7 days of integration.*

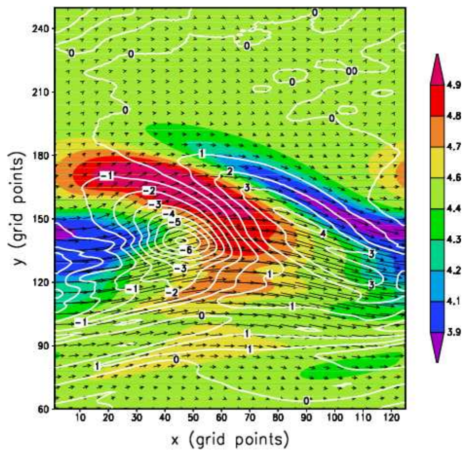


Figure 4: *Quasi-horizontal distribution of N^2 (colors, in 10^{-4}s^{-2}), wind vectors, and vertical divergence (contours, in 10^{-6}s^{-1}) on a surface which is 750 m above the tropopause after 3 days of integration.*

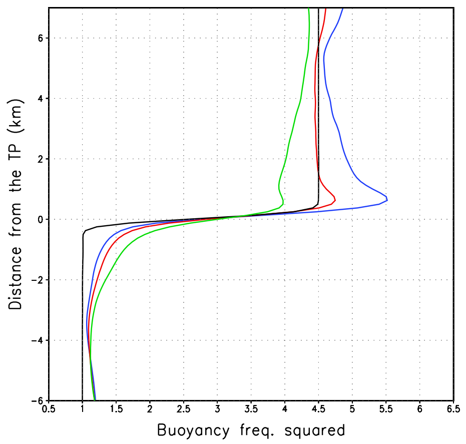


Figure 5: *Composite profiles of N^2 (in 10^{-4}s^{-2}) for the field shown in Fig. 3: average profile (red), average from the 25 percentile of profiles with the largest (blue) and the smallest (green) peaks. For comparison, the black line shows the reference profile.*

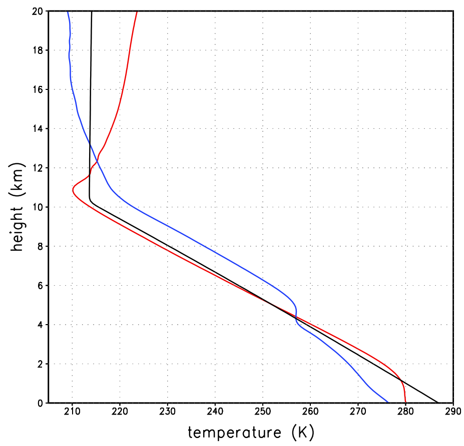


Figure 6: *Two selected temperature profiles from the field shown in Fig. 3: a profile located in the ride (red) and a profile located in the trough (blue). Also shown for comparison is the reference profile (black).*

location of the tropopause in such a way that a strong inversion — being part of the gravity wave — occurs just above the tropopause.

4 Conclusion

We have shown that synoptic scale dynamics during modeled baroclinic wave development plays a significant role for the sharpness of the extratropical tropopause. This corroborates the results from previous, more idealized studies. Key mechanism for the net sharpening is the convergence of the vertical wind during the generation of anticyclonic regions, which is more pronounced than the effect of the divergence of the vertical wind during the formation of cyclonic regions.

At the same time, our simulations indicate that there is more to the tropopause sharpness

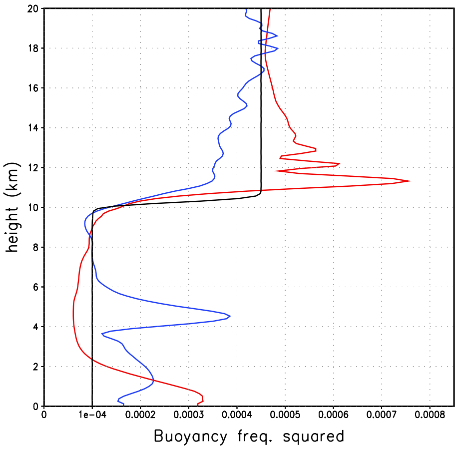


Figure 7: *Profiles of N^2 (in s^{-2}) corresponding to the profiles from Fig. 6.*

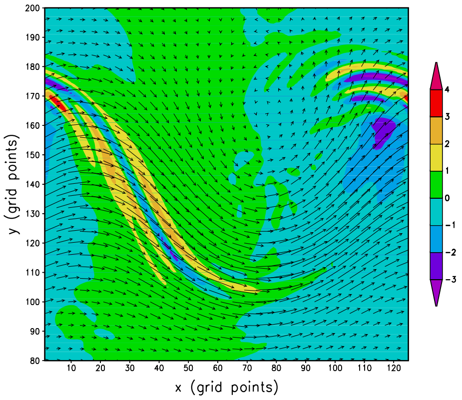


Figure 8: *Divergence of the vertical wind (colours, in $10^{-5}\mathrm{s}^{-1}$) and horizontal wind vectors at 12 km altitude after 6 days of integration.*

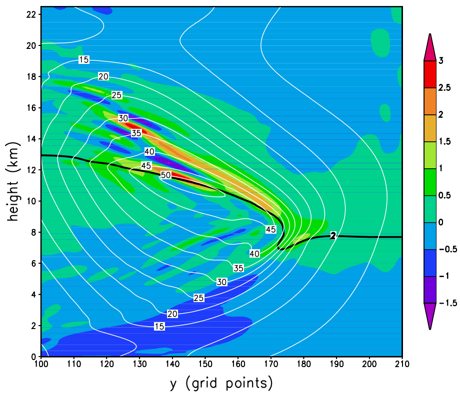


Figure 9: *Vertical cross-section of Fig. 8 at x=30, showing the divergence of the vertical wind (colours, in $10^{-5}\mathrm{s}^{-1}$), wind-speed (white contours), and the dynamical tropopause (defined as PV=2 PVU, black line).*

than just balanced dynamics. It is suggested that large amplitude gravity waves may also play an important role. We are currently investigating whether this can result in tropopause sharpening in a climatological sense.

References

Birner, T., A. Dörnbrack, and U. Schumann, 2002: How sharp is the tropopause at midlatitudes?, *Geophys. Res. Lett.*, 29, 10.1029/2002GL015142

Birner, T., 2005: The fine-scale structure of the extratropical tropopause region. Submitted to *J. Geophys. Res.*

Wirth, V., 2003: Static stability in the extratropical tropopause region, *J. Atmos. Sci.*, 60, pp. 1395-1409.

Wirth, V., 2004: A dynamical mechanism for tropopause sharpening, *Met. Zeitschrift*, Vol. 13, No. 6, pp. 477-484.

Numerical Simulations of the Elbe Flood Case: Sensitivity to Initial and Boundary Data

Günther Zängl[1], Gisela Hartjenstein[1], Günther Doms[2] †, Helmut Frank[2] and Ulrich Schättler[2]

[1]*Meteorologisches Institut der Universität München*
[2]*Deutscher Wetterdienst, Offenbach am Main, Germany*

Abstract

LM simulations have been conducted to examine the reasons for the unsatisfying performance of the operational forecasts for the Elbe flood case in August 2002. We investigated the impact of different initial and boundary conditions, of the cloud microphysical scheme, and of the model resolution. The most important factor turned out to be related to the initial and boundary data. Taking the initial conditions from ECMWF analyses rather than from the operational DWD analyses greatly improves the intensity and spatial structure of the rainfall field. On the other hand, it is comparatively unimportant whether the lateral boundary data are taken from the ECMWF forecast or from a GME forecast initialized with ECMWF analysis data, provided that the horizontal resolution of the GME is similar to that of the ECMWF model. Regarding the cloud microphysical scheme of the LM, it proved to be important to account for the horizontal advection of precipitation particles. Otherwise, the simulated rainfall maximum in the Erzgebirge region is located too far upstream. Refining the model resolution from 7 km to 2.8 km combined with deactivating the convection scheme degrades the model performance because only part of the parameterized (convective) rainfall occurring at 7 km resolution is captured explicitly at 2.8 km resolution.

1 Introduction

In August 2002, the Elbe river catchment area in eastern Germany and the Czech Republic was struck by the heaviest flooding event ever recorded. The largest fraction of the rainfall responsible for the flood fell on 12 and 13 August in connection with a cyclone following the so-called Vb path (van Bebber, 1891), moving from the Adriatic Sea via the eastern edge of the Alps towards Poland. Specifically, the heavy rainfall was caused by a partly occluded warm front on the western and north-western side of the cyclone, which remained essentially stationary over eastern Germany for more than a day although the low-pressure core moved slowly north-eastward. As shown by Ulbrich et al. (2003), the large-scale lifting in the frontal zone was particularly intense because converging surface isobars were combined with pronounced upper-level divergence. Moreover, orographic lifting over the northern slopes of the mountain ranges located in the frontal area led to pronounced local precipitation maxima. The primary one was registered in the eastern part of the Erzgebirge, a mountain range located at the German-Czech border (see Fig. 1). According to the available raingauge measurements, an area of about 25×25 km^2 encountered 36-hour rainfall accumulations in excess of 250 mm, and a peak value of 394 mm was measured at a village named Zinnwald-Georgenfeld near the crest line of the Erzgebirge.

Although the global weather forecasts for 12/13 August 2002 indicated a Vb cyclone track already six days in advance, the operational forecasts of the precipitation field were quite poor. Even the regional-scale Lokal-Modell (LM) forecasts of the German Weather Service (DWD) greatly underestimated the rainfall amounts associated with the cyclone, combined with a mislocation of the rainfall maximum. For example, the LM forecast started at 12 UTC on 11 August placed the primary rainfall maximum ($\sim$ 225 mm) 130 km too far east, affecting the Riesengebirge rather than the Erzgebirge, whereas the secondary rainfall maximum in the Erzgebirge reached only 150 mm (DWD, 2002; see also Fig. 3a). For this rainfall distribution, the Oder river would have been affected by a more severe flooding than the Elbe. In the subsequent forecasts, the rainfall maximum gradually moved westward but still remained significantly east of the observed location. As a consequence, the severe weather warnings issued by the DWD on August 11 were far from indicating a catastrophic flood, predicting only 40–80 mm within 24 hours (DWD, 2002). In the morning of August 12, the rainfall warning was enhanced to 70–120 mm, which would imply a serious flood but is still far below the observed values.

Numerical simulations with the MM5 conducted by Zängl (2004, hereafter Z04) showed a much better agreement with observations although the peak precipitation amounts in the Erzgebirge were still underestimated. Moreover, tests with different model configurations indicated that the model's capability to reproduce the orographic rainfall intensification in the Erzgebirge depends significantly on the model resolution. However, the reasons for the bad performance of the operational forecasts remained unresolved because the simulations presented by Z04 differ in too many aspects from the operational forecasts. This issue will be investigated in the present study, where we use the LM in order to ensure comparability with the operational forecasts. The focus of our sensitivity tests is on the impact of the initial and boundary conditions, but we will also consider the effect of the cloud microphysical parameterization and of the model resolution. The remainder of this note is structured as follows. The setup of the simulations is described in Section 2, followed by a description of the observed precipitation field and the verification methods in Section 3. Section 4 presents the results of the experiments, and a set of conclusions is drawn in Section 5.

2 Setup of the experiments

The sensitivity study reported here was performed with LM version 3.12 (Steppeler et al., 2003) Unless mentioned otherwise, our simulations are based on the operational setup of the LM with a horizontal mesh size of 7 km and 325×325 grid points (Fig. 1). Part of the sensitivity tests are conducted with a mesh size of 2.8 km and 421×461 grid points, corresponding to the domain of the LMK that is planned to become operational (see dashed box in Fig. 1).

The model experiments discussed in section 4 start with the operational forecasts initialized at 12 UTC on August 11 and 00 UTC on August 12 (denoted as OP-1112 and OP-1200, respectively). The OP runs are initialized with the operational nudging-based LM analysis and use GME forecasts as lateral boundary conditions. The GME forecasts are performed with a horizontal mesh size of 60 km and are initialized with the optimal interpolation scheme operationally used for the GME. The LM cloud microphysics scheme does not include the ice phase and does not account for the horizontal advection of precipitation particles (as was operational in 2002). Convection is parameterized with the operational Tiedtke (1989) scheme.

To determine the impact of the LM nudging analysis scheme, the second series of experiments is initialized directly from the GME analysis without applying other changes to the model setup. This series is denoted as GME60-LM-ddhh (ddhh = initialization day and hour). In addition, a re-analysis performed with a more recent GME version that includes ice

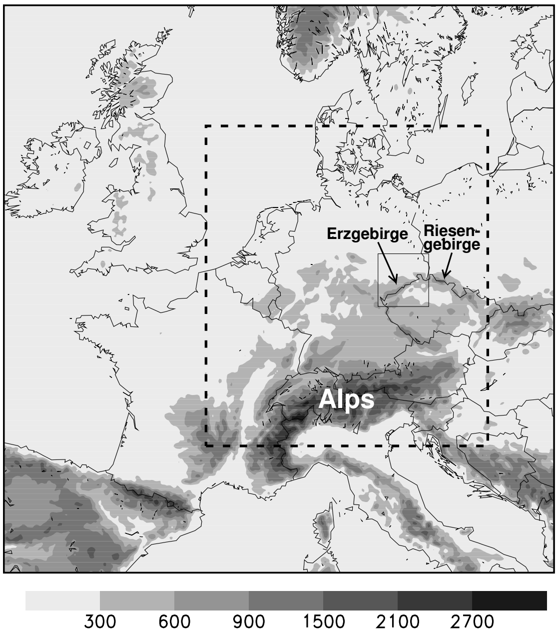


Figure 1: Model topography of the LM domain. The dashed and solid boxes indicate the LMK integration domain and the verification domain, respectively.

microphysics has been tested. For this series of tests, ice microphysics has also been used in the LM. In the simulation name, GME is replaced with GMI when the GME includes ice microphysics.

The remaining experiments are initialized with ECMWF analysis data. Since the ECMWF analyses include cloud ice, the ice phase is also accounted for in the LM microphysics scheme. For the lateral boundary conditions, three different configurations have been tested. The first one, denoted as EC-LM-ddhh, uses ECMWF forecasts corresponding to the respective analysis time. In the other configurations, the lateral boundary conditions are provided by a GME forecast (including ice) initialized with ECMWF analysis data. These GME forecasts have been conducted with a horizontal mesh size of either 60 km or 40 km, corresponding to the operational setup of 2002 and 2004, respectively. The latter series will be denoted as EC-GMIxx-LM-ddhh with xx indicating the GME grid size in km. These boundary data configurations are combined with two versions of the LM microphysics scheme, the more recent one of which accounts for the horizontal advection of precipitation particles. The latter scheme will be referred to as "prognostic precipitation" in the following, abbreviated as LMpp in the simulation acronyms. The sensitivity experiments conducted with the LMK configuration (2.8 km resolution) always include the prognostic precipitation scheme and thus are abbreviated as LMKpp. They differ from the coarser-resolved LM simulations in that the Tiedtke convection scheme is deactivated. However, the shallow convection scheme currently being developed for the LMK is not used because this scheme was not yet available at the time the simulations were conducted. Also, the dynamical core is the same as for the present LM. The LMK grid is one-way nested into the LM, implying that it receives the lateral boundary data from the corresponding LM forecast.

3 Observed precipitation field and verification methods

Since the synoptic evolution of the Elbe flood case is described in some detail in Ulbrich et al. (2003) and Z04, the observed precipitation field in the core precipitation area (see solid box in Fig. 1 for location) is only briefly discussed. Fig. 2 displays the 36-h accumulated

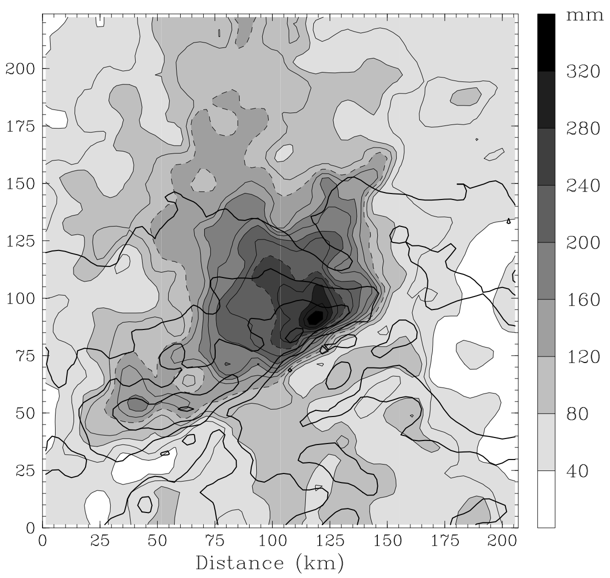


Figure 2: Observed accumulated precipitation (12 August 00 UTC – 13 August 12 UTC) interpolated on the LMK grid. Bold contours indicate the LMK topography (contour interval 200 m), thin contours and shading indicate precipitation at an increment of 20 mm and 40 mm, respectively. Contours for 120 mm and 240 mm are dashed.

precipitation (00 UTC 12 August – 12 UTC 13 August) in this area, roughly corresponding to 50–52°N, 12–15°E. The rainfall field has been constructed by interpolating the measurements from 530 raingauge stations to the LMK grid with a Gaussian weighting method (see Z04 for details). It reveals a wide precipitation band covering most of the verification domain, with precipitation accumulations ranging between 30 and 80 mm near the western and eastern edges and between 100 and 400 mm in the central part. Values exceeding 160 mm are restricted to the northern slope of the Erzgebirge range, indicating that orographic precipitation enhancement played an important role in this case. Analysis of radar data (not shown) reveals that significant precipitation started around 04 UTC on 12 August in the Erzgebirge region, implying that even the simulations started at 00 UTC had several hours to spin up the precipitation field.

The skill scores computed to validate the model results presented in the subsequent section against the precipitation data start with a bilinear interpolation of the simulated values to the locations of the 530 raingauges. This is appropriate in this case because the data density is close to the model resolution (Tustison et al., 2001). The interpolated model output data are then used to compute the relative bias (i.e. the bias normalized by the averaged observed precipitation), the canonical correlation coefficient, the root-mean-square error (rmse) and the mean absolute error (mae). The statistical error measures are summarized in Table 1.

In addition, we computed three standard skill scores based on contingency tables in order to show the dependence of the model performance on the precipitation amount. A contingency table counts the number of simulated and observed data points exceeding a certain threshold value ts, yielding four possible cases:

	$o_i \geq ts$	$o_i < ts$
$s_i \geq ts$	a	b
$s_i < ts$	c	d

Based on this contingency table, the bias score (BS), the false alarm rate (FAR) and the

Simulation	bias	corr	rmse	mae
OP-1112	−45%	0.29	71.7	58.9
OP-1200	−26%	0.22	65.5	48.3
GME60-LM-1112	−28%	−0.03	75.1	59.4
GME60-LM-1200	−18%	0.42	59.7	46.0
GMI60-LM-1112	−23%	0.40	58.6	44.2
GMI60-LM-1200	+6%	0.47	55.6	41.6
EC-LM-1112	−10%	0.65	45.4	32.6
EC-LM-1200	−1%	0.80	35.5	25.4
EC-GMI40-LM-1112	−14%	0.70	43.2	31.9
EC-GMI40-LM-1200	+5%	0.72	40.4	31.2
EC-GMI60-LM-1112	−21%	0.51	54.3	38.9
EC-GMI60-LM-1200	+14%	0.71	46.3	36.1
EC-LMpp-1112	−23%	0.74	43.6	31.8
EC-LMpp-1200	−13%	0.86	31.5	22.1
EC-GMI40-LMpp-1112	−29%	0.78	45.1	34.6
EC-GMI40-LMpp-1200	−12%	0.85	32.5	24.4
EC-GMI60-LMpp-1112	−36%	0.65	55.1	40.7
EC-GMI60-LMpp-1200	−4%	0.78	36.7	27.0
EC-LMKpp-1112	−32%	0.71	50.5	39.2
EC-LMKpp-1200	−23%	0.78	42.3	29.8

Table 1: Relative bias, canonical correlation coefficient (corr), root-mean-square error (rmse) and mean absolute error (mae) for all model experiments discussed in this paper.

equitable threat score (ETS) are defined as

$$\mathrm{BS} = \frac{a+b}{a+c}\,, \qquad \mathrm{FAR} = \frac{b}{a+b}\,, \qquad \mathrm{ETS} = \frac{a-h}{a+c+b-h}\,,$$

where $h = (a+b)(a+c)/(a+b+c+d)$. These quantities are displayed in Fig. 4 for ts ranging from 20 mm to 300 mm at steps of 10 mm.

4 Model results

The results of the operational forecasts OP-1112 and OP-1200 are shown in Fig. 3. Comparing the simulated fields with the observed one (Fig. 2) reveals that the forecasts not only underpredict the maximum precipitation amounts but also fail to capture the spatial structure of the precipitation field. OP-1112 exhibits a local precipitation maximum at approximately the right location, but the main precipitation field is located too far east, and the absolute rainfall maximum ($\sim$ 225 mm) is east of the verification domain in the north-western Riesengebirge (see DWD, 2002, and Hartjenstein et al., 2005). In the later operational forecast OP-1200, the primary rainfall maximum is closer to the observed location, but the overall pattern of the precipitation field differs even more from the observed one than for OP-1112. The statistical measures summarized in Table 1 corroborate that the predicted rainfall fields have a large negative bias and a bad spatial correlation with the observed field, the latter being even worse for OP-1200 than for OP-1112. The negative bias is also evident from the bias score displayed in Fig. 4a, being generally below 1 for the operational forecasts. Rainfall accumulations above 150 mm (OP-1112) and 210 mm (OP-1200) are completely missed, corresponding to a bias score of zero. Despite the large negative bias, both operational forecasts exhibit a substantial false alarm rate (FAR, Fig. 4b), reflecting the misplacement of the simulated rainfall maxima. To summarize, the operational forecasts exhibit effectively no skill, as indicated by an ETS fluctuating around zero (Fig. 4c).

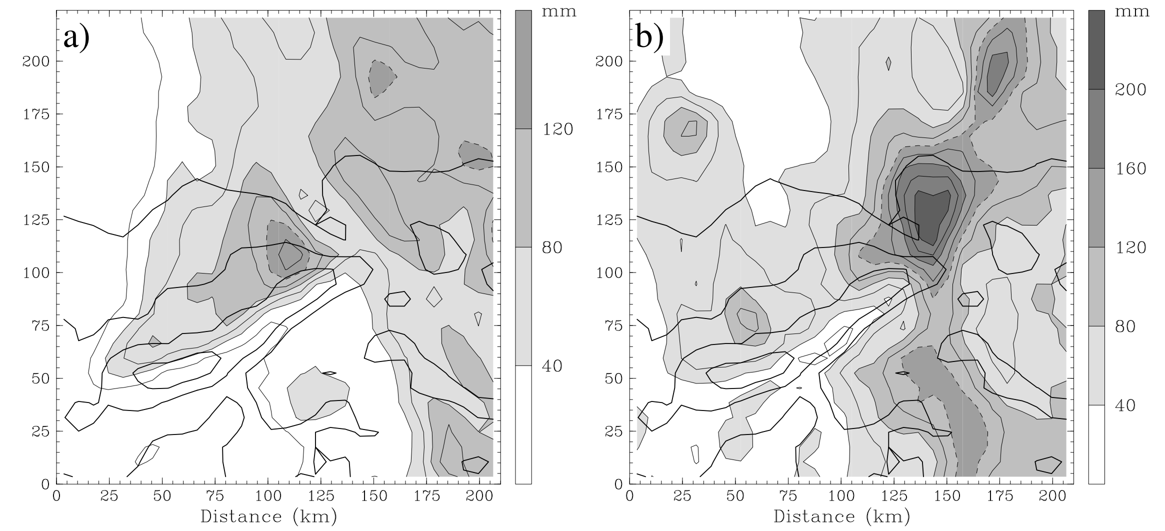


Figure 3: Simulated accumulated precipitation (12 August 00 UTC – 13 August 12 UTC) for experiments (a) OP-1112, (b) OP-1200. Plotting conventions are as in Fig. 2, but the model topography is shown for the LM.

Omitting the LM nudging analysis reduces the negative bias of the operational forecasts (Table 1 and Fig. 4a), but a slight improvement of the forecast quality is found only for the GME60-LM-1200 experiment. In the GME60-LM-1112 run, the mislocation of the precipitation field is even more pronounced than for OP-1112, yielding even larger error measures, a slightly negative correlation coefficient, a greatly increased FAR (Fig. 4b), and a negative ETS (Fig. 4c). Including ice-phase microphysics in the GME and the LM further increases the domain-averaged precipitation (Table 1, GMI60-LM series). Compared to the GME60-LM runs, the error measures indicate a notable improvement for 1112 but no clear tendency for 1200. In both cases, the spatial correlation between the simulated and observed precipitation fields is still unsatisfying.

Substantially better results are obtained when starting from ECMWF analysis data. As evident from Table 1, all the remaining experiments exhibit a higher correlation coefficient and lower error measures than the simulations initialized with the GME analysis. In addition, the lateral boundary data used during the forecast and the microphysical scheme (prognostic vs. diagnostic precipitation) play a significant role.

For the experiments with the diagnostic scheme (no horizontal advection of precipitation particles; 3rd section of Table 1), it can be seen that the forecasts started at 1200 perform better than those started at 1112. This behaviour is most pronounced for the EC-LM series that uses ECMWF forecasts as lateral boundary data. While the EC-LM-1112 run still exhibits a significant negative bias, EC-LM-1200 has almost no bias, and the other validation measures are also better for 1200 than for 1112. When taking the lateral boundary conditions from GME-forecasts based on ECMWF analyses, the negative bias increases for 1112 whereas a positive bias appears for 1200. Both effects also depend on the horizontal resolution of the GME, being larger at 60 km than at 40 km resolution. In addition, the verification results are generally worse for a GME resolution of 60 km than for 40 km. On the other hand, comparing the verification results for the EC-LM and EC-GMI40-LM series does not indicate a systematic signal.

Detailed verification results and accumulated precipitation fields are displayed in Figs. 4d–f and 5 for selected experiments. Fig. 5 readily shows that the spatial structure of the precipitation fields is closer to the observed one (Fig. 2) than for the operational forecasts (Fig. 3). In particular, the precipitation maximum is located in the right region, though being shifted

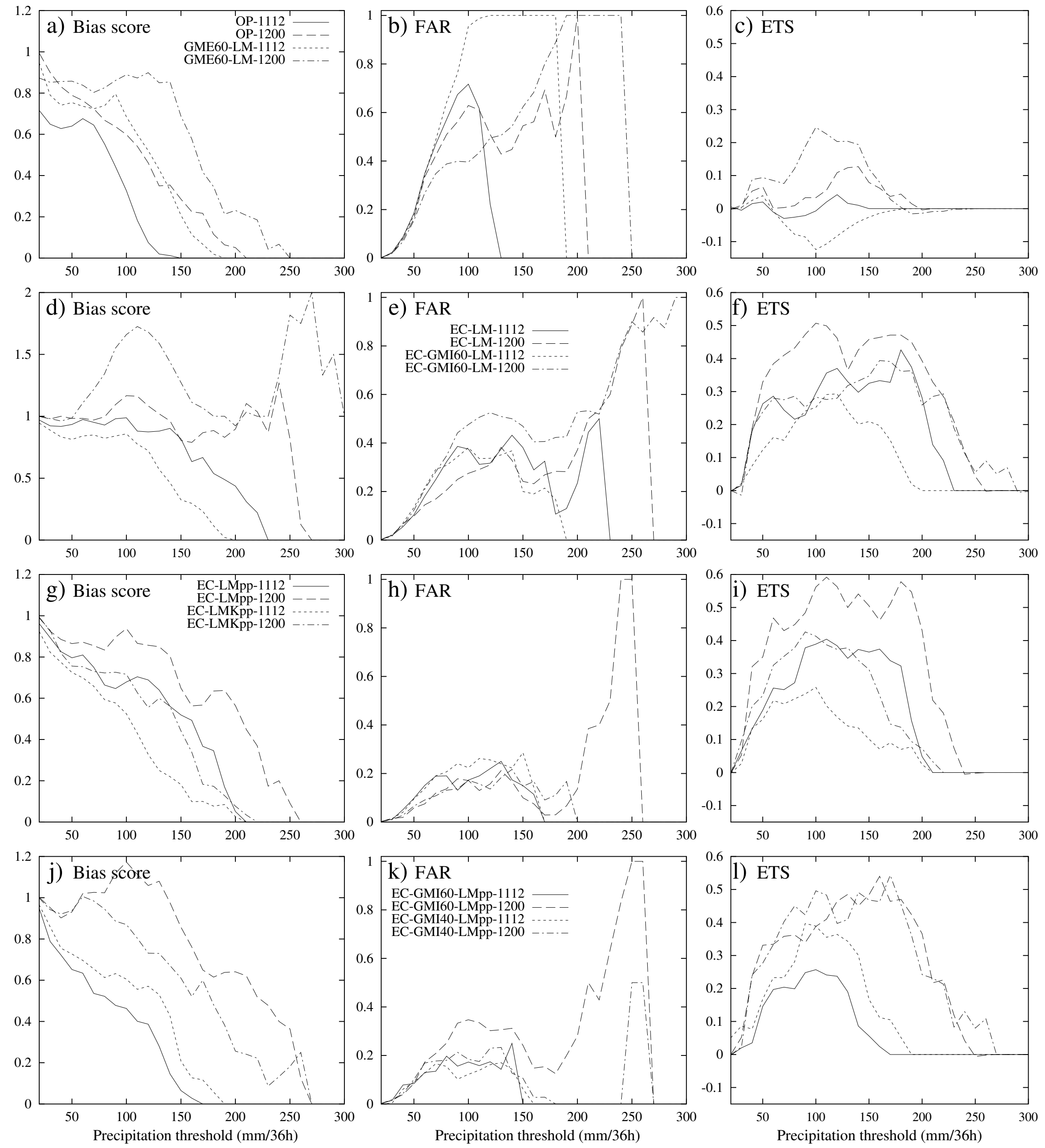


Figure 4: Bias score (left column), false alarm rate (FAR, middle column) and equitable threat score (ETS, right column) for various model experiments (see text for definition). The line keys given in (a), (e), (g) and (k) are valid for all panels of the respective row.

20–30 km upstream compared to the observations. Moreover, Fig. 5 confirms that the simulations started at 1200 produce higher precipitation amounts than those started at 1112. The bias scores shown in Fig. 4d indicate that EC-LM-1112 performs well at precipitation amounts below 150 mm but underestimates higher precipitation amounts, which is in accordance with the visual impression arising from Figs. 2 and 5a. An ETS of 0.3 up to a threshold of 200 mm (Fig. 4f) and a FAR below 0.4 (Fig. 4e) indicate that this forecast has significant skill. The ETS is even higher for the 36-h forecast (EC-LM-1200), combined with a bias score fluctuating around 1 up to a threshold of 250 mm. The high FAR at precipitation thresholds between 200 and 260 mm reflects the fact that the precipitation maximum is located incorrectly. Fig. 4d also shows that taking the lateral boundary conditions from a GME forecast performed at 60 km resolution (EC-GMI60-LM) increases the negative bias for 1112 whereas a substantial positive bias is created for the initialization at 1200. The

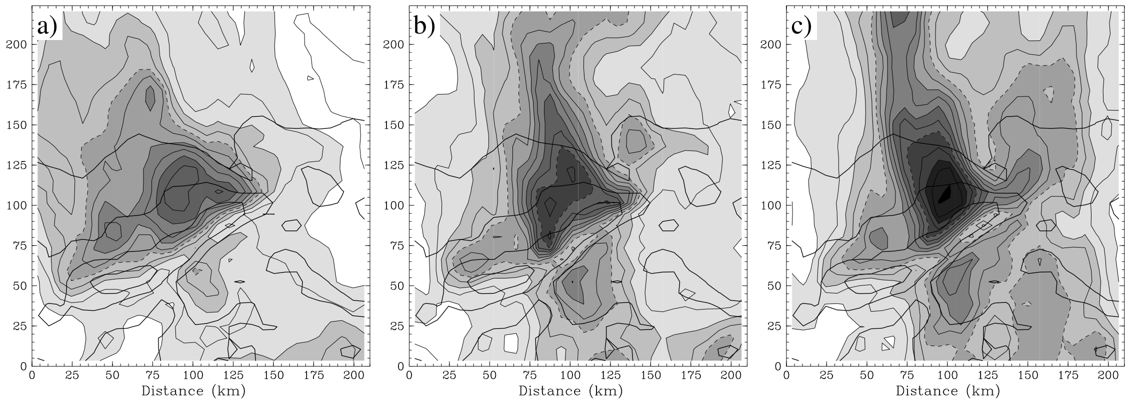


Figure 5: Same as Fig. 3, but for experiments (a) EC-LM-1112, (b) EC-LM-1200, (c) EC-GMI60-LM-1200.

precipitation maximum is still mislocated, leading to a high FAR above 200 mm in the 1200 case. Finally, Fig. 4f confirms that the skill of the EC-GMI60-LM series is not as good as for the EC-LM series.

All six experiments initialized with ECMWF analysis data have been repeated with the prognostic precipitation scheme in the LM (LMpp series). Comparing the verification results summarized in Table 1 with those for the diagnostic scheme indicates that the prognostic scheme significantly decreases the total amount of precipitation. This is mostly related to a bug in the prognostic precipitation scheme of the LM version used for this study, leading to a systematic underestimation of the precipitation reaching the ground. A test with the corrected scheme revealed that about 90% of the domain-averaged difference were related to this bug. However, the spatial distribution of the precipitation remained largely unaffected, so that we decided not to repeat the full suite of sensitivity tests. Table 1 reveals that the pattern correlation between the simulated and observed rainfall fields is generally improved by the prognostic scheme, leading to a substantial decrease of the error measures in the 1200 cases. For the 1112 cases, the increased negative bias tends to balance the improved pattern correlation, so that the error measures do not show up a clear trend. With the corrected precipitation scheme, the error measures would be improved in both cases. The accumulated precipitation fields for the EC-LMpp series (Fig. 6a,b) indicate that the improvement of the pattern correlation is mainly due to a downstream shift of the precipitation maximum over the eastern Erzgebirge, bringing it closer to the observed location. This is also true for the EC-GMI-LMpp experiments (not shown). Correspondingly, Figs. 4h and 4i show that the prognostic scheme reduces the FAR and improves the ETS, particularly for the initialization at 1200. The high peak in the FAR appearing for the EC-LMpp-1200 run is represented by very few data points and thus not quite significant.

Taking the lateral boundary data from GME forecasts initialized with ECMWF analyses has a similar impact as for the experiments with the diagnostic precipitation scheme discussed above. Compared to the simulations driven with ECMWF forecasts, the negative bias increases for the 1112 runs while it decreases for the 1200 runs (Table 1 and Figs. 4g,j). Moreover, the verification results are significantly degraded when the GME forecasts are conducted at 60 km resolution. However, the differences between the EC-LMpp and EC-GMI40-LMpp series are even smaller than for the experiments with the diagnostic scheme (Table 1 and Fig. 4j–l), which is corroborated by a visual inspection of the precipitation fields (not shown). This confirms that the initial data have a much larger impact on the forecast skill than the forecast model used for creating the lateral boundary conditions, provided that the spatial resolution of the outer (global) is not too low.

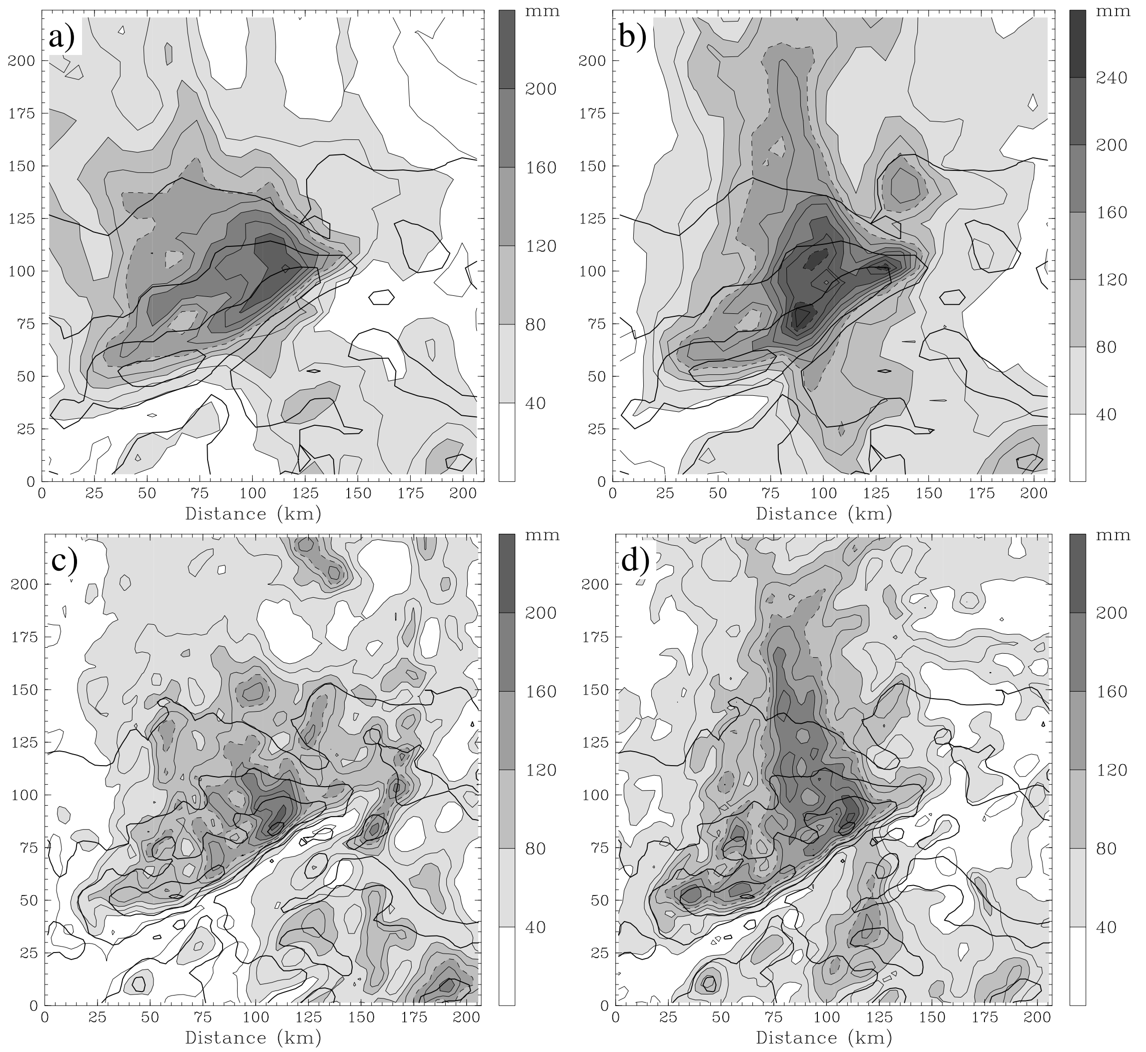


Figure 6: Same as Fig. 3, but for experiments (a) EC-LMpp-1112, (b) EC-LMpp-1200, (c) EC-LMKpp-1112, (d) EC-LMKpp-1200. In (c) and (d), bold lines indicate the LMK topography.

Finally, a look at the results obtained with the LMK configuration (2.8 km mesh size, Fig. 6c,d) reveals that the increased resolution creates much smaller-scale structures in the precipitation fields. This is partly due to a better resolution of the topography and partly due to the fact that the convection scheme is switched off in the LMK experiments. However, a comparison with the corresponding LM runs (Fig. 6a,b) shows that the total amount of precipitation, including the precipitation maxima, is less for the LMK than for the LM. This implies an increased negative bias (Table 1 and Fig. 4g). Moreover, Table 1 and Fig. 4i indicate that the forecast skill of our preliminary LMK version is not as good as that of the LM. A larger number of LMK experiments should be conducted when the development of the model code is completed. In particular, recent tests performed at DWD revealed that a shallow convection parameterization is still necessary at a mesh size of 2.8 km in order to improve the triggering of resolved deep convection.

5. Conclusions

Our sensitivity tests for the Elbe flood case indicate that the limited accuracy of the initial data was the most important reason for the poor quality of the operational forecasts. At least in the present case, the sophisticated 4D-VAR data assimilation performed at ECMWF leads to a substantially better forecast accuracy than the optimal interpolation scheme currently used in the GME. This is in accordance with the notion that forecasts of extreme weather events are particularly sensitive to the accuracy of the initial data (e.g. Wergen and Buchhold, 2002). Thus, the 3D-VAR scheme currently under development for the GME can be regarded as an important step for improving the forecast skill. On the other hand, our experiments in which a GME forecast initialized with ECMWF analysis data was used to provide the lateral boundary conditions for the LM indicate that the global model itself is of comparatively minor importance. It is mainly its horizontal resolution that plays a role. A mesh size of 40 km, as currently used both in the GME and the ECMWF model, yields significantly better results than a mesh size of 60 km (operational in the GME prior to October 2004). In addition, a comparison of the cloud microphysics schemes available in the LM shows that accounting for the downstream advection of precipitation particles is important to get the spatial structure of the precipitation right. The impact is particularly evident over the Erzgebirge, where the rainfall maximum is shifted from the windward slope toward the crest line.

To relate our present results to the MM5 simulations reported in Z04, it has to be mentioned that the MM5 simulations were driven with ECMWF analysis data rather than forecast data and that the MM5 was operated in a multiple-nested configuration with 2–4 domains and a finest resolution of 9 km, 3 km and 1 km, respectively. Thus, the results are not strictly comparable with the LM forecasts. Nevertheless, the verification results of EC-LMpp-1112 are very close to the two-domain MM5 run with a finest resolution of 9 km, and EC-LMpp-1200 performs even better than a corresponding MM5 simulation initialized at 1200 (not reported in Z04). However, the dependence of the model skill on the horizontal resolution is opposite. While the MM5 verification results steadily improve with increasing model resolution (see Z04), the 2.8-km resolution LMK performs not as well as the 7-km LM (implying that it is also inferior to the MM5 at 3 km resolution). The most likely explanation for this puzzling behaviour lies in the different characteristics of the convection parameterizations. The Kain-Fritsch scheme used in the 9-km MM5 domain (Kain and Fritsch, 1993) accounts for less than 3% of the total precipitation in the core precipitation area (Erzgebirge), whereas the Tiedtke scheme of the LM accounts for 20–25% of the total precipitation. In the MM5, the domain-average precipitation then increases with increasing horizontal resolution because resolved embedded convection enhances the precipitation efficiency of the microphysical scheme. This reduces the negative bias and improves the skill scores. In the LM, however, the convection scheme appears to overcompensate the resolution-dependence of explicit precipitation, leading to an opposite behaviour of total precipitation and skill scores. Another contribution probably arises from the lack of an appropriate shallow convection scheme in the LMK, which was not yet available for our tests.

Acknowledgements

The DWD and the CHMI (Czech Hydrometeorological Institute) are acknowledged for providing the raingauge data needed for validating the model results.

References

van Bebber, J. W., 1891: Die Zugstraßen der barometrischen Minima. (The tracks of the barometric minima.) *Meteor. Zeitschr.*, 8, 361–366.

DWD, 2002: Starkniederschläge in Sachsen im August 2002. (Heavy precipitation in Saxony in August 2002.) Internal report, Deutscher Wetterdienst, Offenbach, October 2002, 61 pp.

Hartjenstein, G., G. Zängl, G. Doms, H. Frank, and U. Schättler, 2005: Numerical simulations of the Elbe flood case: Sensitivity to initial and boundary data. Proc. Intern. Conf. on Alpine Meteorology, Zadar, Croatia, *Croat. Meteorol. Journal,* 70, 361–364.

Kain, J. S. and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. *The representation of cumulus convection in numerical models*, K. A. Emanuel and D. J. Raymond, Eds., Amer. Meteor. Soc., 246 pp.

Steppeler, J., G. Doms, U. Schättler, H.-W. Bitzer, A. Gassmann, U. Damrath, and G. Gregoric, 2003: Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteorol. Atmos. Phys.,* 82, 75–96.

Tiedtke, M, 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.,* 117, 1779–1799.

Tustison, B., D. Harris, and E. Foufoula-Georgiou, 2001: Scale issues in verification of precipitation forecasts. *J. Geophys. Res.,* 106, 11775–11784.

Ulbrich, U., T. Brücher, G. C. Leckebusch, A. Krüger, and J. G. Pinto, 2003: The central European floods of August 2002: Part II – Synoptic causes and considerations with respect to climate change. *Weather*, 58 (11), 434–441.

Wergen, W., and M. Buchhold, 2002: Datenassimilation für das Globalmodell GME. (Data assimilation for the global model GME.) *Promet,* 27, 150–155.

Zängl, G., 2004: Numerical simulations of the 12–13 August 2002 flooding event in eastern Germany. *Quart. J. Roy. Met. Soc.,* 130, 1921–1940.

List of COSMO Newsletters and Technical Reports

(available for download from the COSMO Website: www.cosmo-model.org)

COSMO Newsletters

No. 1: February 2001.

No. 2: February 2002.

No. 3: February 2003.

No. 4: February 2004.

No. 5: April 2005.

No. 6: July 2006; Proceedings from the COSMO General Meeting 2005.

COSMO Technical Reports

No. 1: Dmitrii Mironov and Matthias Raschendorfer (2001):
Evaluation of Empirical Parameters of the New LM Surface-Layer Parameterization Scheme. Results from Numerical Experiments Including the Soil Moisture Analysis.

No. 2: Reinhold Schrodin and Erdmann Heise (2001):
The Multi-Layer Version of the DWD Soil Model TERRA_LM.

No. 3: Günther Doms (2001):
A Scheme for Monotonic Numerical Diffusion in the LM.

No. 4: Hans-Joachim Herzog, Ursula Schubert, Gerd Vogel, Adelheid Fiedler and Roswitha Kirchner (2002):
LLM ‾ the High-Resolving Nonhydrostatic Simulation Model in the DWD-Project LITFASS.
Part I: Modelling Technique and Simulation Method.

No. 5: Jean-Marie Bettems (2002):
EUCOS Impact Study Using the Limited-Area Non-Hydrostatic NWP Model in Operational Use at MeteoSwiss.

No. 6: Heinz-Werner Bitzer and Jürgen Steppeler (2004):
Documentation of the Z-Coordinate Dynamical Core of LM.

No. 7: Hans-Joachim Herzog, Almut Gassmann (2005):
Lorenz- and Charney-Phillips vertical grid experimentation using a compressible non-hydrostatic toy-model relevant to the fast-mode part of the 'Lokal-Modell'

No. 8: Chiara Marsigli, Andrea Montani, Tiziana Paccagnella, Davide Sacchetti, André Walser, Marco Arpagaus, Thomas Schumann (2005):
Evaluation of the Performance of the COSMO-LEPS System

No. 9: Erdmann Heise, Bodo Ritter, Reinhold Schrodin (2006):
Operational Implementation of the Multilayer Soil Model