Short report of the COSMO workshop on Verification of Lokal Model. Bologna, 14-15 February 2000

Authors: All the responsibles of the WPs in the R&D area Verification.

Verification of Lokal Model (LM) is an important R&D area within the COSMO Consortium: the verification activities allow to understand the capability of LM in reproducing correctly the mesoscale features of upper air and surface meteorological parameters.
Several topics have been discussed during the workshop and can be summarised in the following points:

- ✓ During the first part of the workshop the verification scores of the different LM forecasted parameters and the activities performed so far have been shown. Seventeen WPs are present in the R&D area Verification and most of them are in a good status of advancement. Some other started just some weeks ago and some other will start this spring. In Appendix the different WPs and the main results obtained so far are described.

- ✓ After the first part, a detailed discussion of the techniques and methods used to achieve the LM verification has followed. We believed to be necessary to define a minimum common agreement for the verification of some surface parameters (Precipitation, 2 m. Temperature, cloud coverage). This common agreement will allow us to compare results of the verification performed in the different geographical areas. During the meeting we have also stressed the importance to evaluate skill scores (against persistence forecast, for example) and not only "absolute" indices for the different parameters simulated by the model. It has been decided that each participant goes on in doing the activities already in course and, additionally, some common verification practices should be achieved according to some agreed common rules:
    1. As a common criteria we decided to make the verification of model output against observation for different class of elevation. The procedure should be the following: for each station and for all the parameter, we have to select the nearest grid-point to the OBS station. If height difference is < 100 m we chose the nearest grid point, else we select from the four surrounding grid points the one with the minimum height difference;
    2. For temperature, we see to be necessary to make a correction from grid point elevation to station point elevation by using an "appropriate" profile that should be defined according to the different local PBL structure. We discussed in some detail the problem of the vertical profile to use for the vertical reduction. Probably the best choice could be to use the LM temperature profile forecast in order not to create artificial forecast error. *Question to the COSMO scientific community: could it be possible to introduce this vertical reduction to the true station elevation into the LM post-processing code ?*. As verification scores we agreed to evaluate (at minimum) RMS, BIAS, Standard Deviation and number of useful forecasts using different thresholds (2 K). Furthermore, an useful skill score for temperature is: $1-[(RMS(LM)/RMS(persistence)]^2$;
    3. For cloud cover verification against surface observation it should be considered also all grid points in a radius of around 30 Km. The common thresholdes for cloudness are: $n<3/8$, $3/8\leq n\leq 7/8$; $n\geq 7/8$ (see EWGLAM Newsletter 26, 111-114);
    4. As regard precipitation common thresholds have been defined for 6 hours accumulated precipitation: 0.1mm (or a different threshold to discriminate yes/no precipitation according to different measuring methods); 2mm; 10 mm/6 hours. Starting from contingence tables, evaluated for different thresholds, we proposed to evaluate the

following (at a minimum) indices (Wilks, 1996): Bias, False Alarm rate, Probability of detection. In order to evaluate the skill of LM, similar indices should be evaluated also for persistence forecast and must be compared with those evaluated for LM. For persistence we intend that it must be used the same interval in the day time.

- ✓ An important discussion has followed about the best way to exchange verification results on a regular basis and to transfer these results to the COSMO WEB page. We decided that at least every season each participant to the WG should produce a short report where the main results of LM verification are described for the occurred season. This short report must be written in *html* format and the graphs, figures etc. in *postscript* format. The coordinator of the R&D area Verification will keep the contacts with the COSMO scientific coordinator in order to define the best way to transfer these reports into a reserved area of the COSMO web pages.

- ✓ We did a discussion about the difference between the standard methods used so far to verify the accuracy of forecast (Hit rates, Bias etc…) and some "new" methods to investigate the realism of very high resolution models (pattern correlation, probability distributions etc..). We decided to continue the discussion and eventually propose some new activities to be done starting from the next COSMO meeting. We have to collect more experience about that. Ideas between people involved in our WG will be circulated.

- ✓ We also do not see now the need to define a single score to assess the global forecast quality. The main reason is the impossibility to reduce the state of the atmosphere in a single parameter. Furthermore this single parameter can not give an indication about the exact "location" of the main deficiencies of the LM model and this single number would be also of low value for the end users of the COSMO consortium.

- ✓ Finally the group stressed the importance to achieve a stratification for "weather regimes" in order to discriminate the scores of LM forecast for each regime. These weather regimes can be different for each region and differently defined using subjective or objective classification criteria.

APPENDIX

WP 5.1/5.2 : RECENT RESULTS OF LOKAL MODELL VERIFICATION IN EMILIA-ROMAGNA AND MARCHE REGIONS, ITALY.

Carlo Cacciamani [1], Laura Mannozzi [2] and Laura Sandri [1]

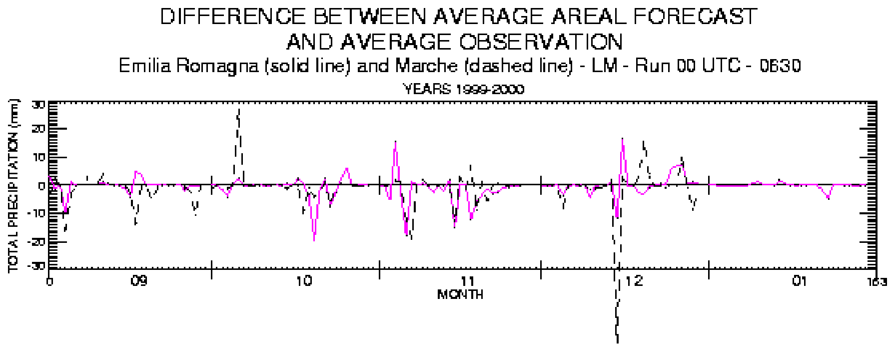
(1) ARPA SMR Emilia Romagna, Italy

(2) OSGM Marche, Italy

The goal of this study, performed in collaboration between ARPA SMR and OSGM, is to assess the quality of Lokal Modell (LM) quantitative precipitation forecast (QPF) over the Emilia Romagna and Marche regions in Italy, from February 1999 to January 2000. In this short report, we will discuss the results obtained during the last 5 months of verification (from September 1999 to January 2000). The verification has been carried out by means of the synoptic and local stations, in both regions. In order to measure the quality of the QPF, we selected two different verification methods, meant to explore different aspects of the model forecast.

The first and less strict method, called *areal* method, is applied to obtain important insights regarding the integral quantities of the simulation. It is sensitive to errors in the magnitude of the precipitating structures and not in their location, provided that both the observed and forecast structures fall into the verification area. The areal verification is carried out by computing the spatial average over all the model grid points that fall into the verification area, and comparing it (daily and monthly) with the observed one.

Regarding the daily comparison, we found out that during the last 5 months of verification the model generally underestimated the total precipitation over both regions, especially during November 1999. The behaviour of LM was generally better in Emilia Romagna than in Marche, where also in September 1999 LM underestimated the rainfall amount. In the Marche region, on Dec. 15th, 1999, a heavy precipitation event was recorded (over 70mm/day observed on average over the whole area) which was seriously underpredicted (less than 20mm/day forecast on average over the whole area).



It is interesting to analyse the behaviour of LM during the so-called Christmas storms occurred in Europe on the 26th and 28th of December 1999. During these events, in Emilia Romagna there was almost no observed precipitation but strong winds, and LM correctly forecast low rain amounts (less than 10mm/day on the 26th and 2mm/day on the 28th on average over the whole area). In Marche almost no precipitation was recorded on the 1st event, but 10mm/day on average fell during the 2nd one. The LM forecast was quite wrong, because the simulation yielded 12mm/day on the 26th and less than 2mm/day on the 28th, always on average over the whole area of Marche.

As concerns the monthly comparison of spatial averages, we computed Mean and Root mean Square Errors (ME and RMSE). For 24-hour cumulated precipitation forecasts, in both regions the bias is very different from the high and positive one we observed during spring and summer 1999. In fact, in Emilia Romagna the ME is close to zero (although slightly negative in November 1999) and in Marche is negative. This confirms the above discussed considerations regarding the daily comparison. We also checked the first 6 and 12 hours into forecast ME, and found out that the spin-up problem is much less significant in autumn and winter than in summer.

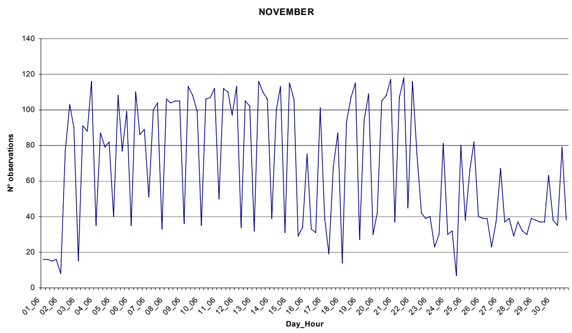
We also compared the daily maximum observed value with the forecast one, regardless of the correspondence between the station and the grid point where they were respectively recorded. This was done to measure the realism of the LM forecast, that is how realistic are the precipitation maxima produced by LM. The results show a more realistic behaviour over the Marche region, where the forecast maxima do not exceed significantly (like they do over Emilia Romagna) the observed ones.

The second method we use is called *punctual*. It is a stricter method if compared to the areal one, because it is sensitive to errors both in the magnitude and in the location of precipitating structures. The punctual verification is performed by comparing the daily value recorded at each station with the one forecast at its closest model grid point. The comparison for each station at different precipitation thresholds (we used 1mm/day, 5mm/day, 10mm/day, 20mm/day, 50mm/day) produces a 6x6 contingency table, which is then reduced, for a given threshold, to a 2x2 yes(event above threshold)/no(event below threshold) contingency table. Thus, we obtained monthly tables at different thresholds, from which bias and some accuracy measures, such as the Hit Rate (HR), the Hit Rate Rain (HRR), the False Alarm Rate (FAR), the Threat Score (TS) and the Heidke Skill Score (HSS) were computed. Due to the scarce number of high rainfall rates events in these Italian regions, we cannot give results for thresholds higher than 10mm/day. However, the bias for both regions is about one for the 1 to 10mm/day thresholds during the autumn and winter months, being less than 1 in November 1999. During these months, there is also a much lower FAR, and a much higher TS (up to 0.55-0.6 at 1mm, up to 0.5 at 10mm) and HSS (up to 0.6 for 1 and 10mm) in both regions, compared to the spring and especially the summer scores. The most remarkable feature of our punctual verification, apparent for both regions at all the significant thresholds (1 to 10mm/day) is a systematic improvement of bias, FAR,TS and HSS during the whole year of LM QPF verification, apart from the summer overprediction problems outlined in the September 1999 COSMO meeting in Bologna.

WP 5.3: Work done at IMS: VERIFICATION OF "LOKAL MODELL"
Massimo Ferri, Massimo Bonavita and Claudio Reina

During the MAP project it has been done a lot of measures on the metheorologicals parameters.
In the north of Italy many stations (about 250) made hourly observations between October and November.
Because the observation zone is inside the deutsche model domain the precipitation measures have been used as verification parameters of "Lokal Modell".
The first step was to create a database of observations chronologically organised. The time steps chosen are of three and six hours.

The second step was to realise a statistical study on the frequency of observations. From this study we have noted that the available data were less then 50% of nominal stations related to the threehourly observations and less concerning the cumulated precipitation in six hours. The availability of last observations presents a diurnal cycle.



Before and after the central part of observations (October-November) the data available have reduced their number by a factor two.
We have extracted the cumulated precipitation field from the "Lokal Modell" GRIB. We have organised this datasheet to have the cumulated precipitation every three and six hours vs. initial time of forecast and time step.
The two homogeneous databases of observations and cumulated precipitation are used to produce some scores concerning the "Lokal Modell".
About the GRID points we decided, as a common criteria, to make the verification of model output against observations for different class of elevation. The procedure is the following: for each station if the height difference between the nearest grid-point and the observation station is less then 100 m we select this nearest grid-point else we select the grid-point with the minor height difference among the four surrounding.
The skill scores elaborated now are ME, MAE and RMSE. Later we will do the contingence tables and we will create the following index: Bias, False Alarm Rate and Probability of Detection.
The final step of this work will be the subdivision of North Italy region into six climatological zones to calculate again the previous index.
We hope to have the final results, concerning the "Lokal Modell" verification from the month of May to July and from September to December, before June 2000.

Working Package 5: WORK DONE AT SMI (October 1999 - February 2000)

In early Summer 2000 the preoperational phase of the LM on the NEC SX-5 at CSCS/Manno (TI) will begin. At that time all the Working Packages described below will be performed with this swiss LM-configuration.

WP 5.4: High resolution verification of daily cycle over Switzerland
The results of LM (and SM) for the last months (till January 2000) of 2m-temperature, 10m-wind, precipitation (hourly intervals) and for cloud cover (3-hourly intervals) have been shown at the workshop. In addition to the results presented at the COSMO meeting in September 1999, the following points are of main interest:
(1) the 2m-temperature negative bias is very pronounced in wintertime. In January 2000 it is of the order of 3K for gridpoints < 800m and even ~ 8K (!) for gridpoints > 1500 m.
(2) since November 1999 there is no more a spin-up in precipitation for the first hours of forecast. But the results of precipitation still must be interpreted as a mean for an area of ~ 14x14 km: the variability of total monthly precipitation for neighbouring grid points remains very large.
(3) by comparing LM with SM results, some changes in the behaviour for 2m-temperature, 10m-wind, precipitation and cloud cover are noticed since December 1999. These are possibly due to the change for the SM of initial and lateral boundary conditions from EM to GME (see below 'Other SMI-activities presented at the workshop').
The next steps in this Working Package are:
- the verification of Td2m will start this Spring.
- with begin of the preoperational phase of the LM, the verification will be performed on both LM-configurations (DWD and SMI).

WP 5.5: Low PBL verification with masts of swiss nuclear powerplants
This work has not yet begun.

WP 5.6: Daily verification of LM/SM cloudiness with the Meteosat VIS-Channel
This verification is operational with the SM since February 1, 2000. It will begin for the LM with the preoperational phase in early Summer 2000.

WP 5.13: Verification of the vertical profiles at TEMPs stations
This verification will begin with the preoperational phase of the LM in early Summer 2000.

WP 5.14: Verification of precipitation with the swiss raingauge data
The programs for this verification with the SM (on an seasonal basis) will be adapted to the LM-grid in the next months. The great variability in the LM for precipitation over the Alpine area will be investigated.

WP 5.15: Verification of SM(/LM) surface precipitation with the Swiss radar composite network
A systematic evaluation using simple statistics, including pattern correlation (as presented at the DWD/SMI Meeting at Bad Säckingen, 4-6.11.97) is now underway. It is being developed in the SM framework but will be applied to the LM after the start of its preoperational phase.

WP 5.16: Verification of MAP-SOP data with SM, LM and MC2
There was no commitment for this Package but there is a strong interest for a common verification of these three models with the verification package
of WP 5.4. The first step - the extraction of MC2 GRIB-data to the format for WP 5.4 - has to be done outside SMI by students involved with MAP- activities.

Other SMI-activities presented at the Workshop
Results with the verification package of WP 5.4 of a one month-parallel runs of SM with initial and lateral boundary conditions from (a) EM + initial surface fields from DM [abbreviated SM@EM] and (b) GME [abbreviated SM@GME] have been presented.
In spite of the fact that these results are not from the LM, they are of interest to the COSMO-community, because they show the dependency to initial and lateral boundary conditions. The main results are:
- 2m-temperature with SM@GME is lower than with SM@EM (i.e. increased negative bias). For 00 UTC-forecasts and gridpoints < 800m: SM@GME gives an increased negative bias of ~ 0.5 K for all the forecast length
- less precipitation with SM@GME (SM@GME gives ~ 20% too less precipitation and SM@EM ~ 15% too much)
- positive bias in 10m wind speed is a little bit reduced with SM@GME
- total cloud cover of SM@GME ~ 0.5 octa less than SM@EM (but SM@EM has already a negative bias).

WP 5.2 and 5.11: Lokal Modell Verification at CMIRL (Meteo-Hydrological Centre of Liguria Region)

At CMIRL in Genoa we daily receive grib formatted files from DWD for the two runs of the model (00 UTC and 12 UTC), containing the total precipitation and the high-medium-low and total cloud cover over the verification domain; these files are visualized using GrADS.

The work of verification of precipitation is carried out in similar way and in cooperation with the SMR of Emilia Romagna Region in Bologna.
The verification is performed over the Liguria Region, in two different ways: one is carried out by considering spatial averages over all the model grid points that fall into the area (areal verification), and the other one by considering the forecast values at each model grid point that is closest to a station (punctual verification).
For 6h and 24h cumulated precipitation and for different forecast time (00-06, 06-12, 12-18, 18-24, 24-30, 30-36, 36-42, 42-48; 00-24, 24-48, 06-30) contingency tables and statistical parameters are computed.
The thresholds for the contingency tables are 0, 1, 5, 10, 20 and 50 mm, and the statistical indexes are Bias, Heidke Skill Score, Hit Rate Rain, Hit Rate, False Alarm Rate and Threat Score.
These indexes and others statistical quantities, like spatial average, standard deviation, maximum value, mean error and percentual error, are displayed using Xmgr.

The work of verification of cloudiness is carried in cooperation with CNR-ISAO (Bologna, Italy) and with OGS (Macerata, Italy).
By means of High Resolution Images in IR and VIS METEOSAT channels, daily available at CMIRL in Genoa, a cloud classification scheme for low, middle and high clouds will be operationally available at CMIRL. The procedure takes into account spectral proprieties of the METEOSAT data as well as statistical proprieties of the clouds detected into the verification area (42N-47N, 6W-16W). By means of comparison with results of the classificator, verification of LM cludiness will be then performed.

WP 5.7: Verification of LM-forecasts for surface weather elements state, problems and outlook
Ulrich Damrath; Deutscher Wetterdienst

*Verification system

-weather parameters:
cloud cover, wind speed, wind direction wind vector, 2m temperature, 2m dew point(difference), 2m extreme tempreatures, 6h sums of precipitation, 24h sums of precipitation

-scores:
hit rate for cloud cover, mean error, mean absolute error, root mean square error, standard deviation vor other elements, true skill statistics for precipitation
conditional bias for cloud cover, wind speed 2m temperature and mslp

-regions
different regions in Europe (especially Germany and Switzerland), different height of stations

-special investigation concerning precipitation

*Results

-comparison of LM with DM
advantages of LM especially concerning wind, nearly same quality converning temperature and dew point,
bias of cloud cover strongly dependent on cloud cover itself, bias of temperature depends on wind direction (negative bias in cases of southerly winds)
horizontal distirbution of errors: highest errors over the Alps, lowest over the northern part of Germany

*Problems

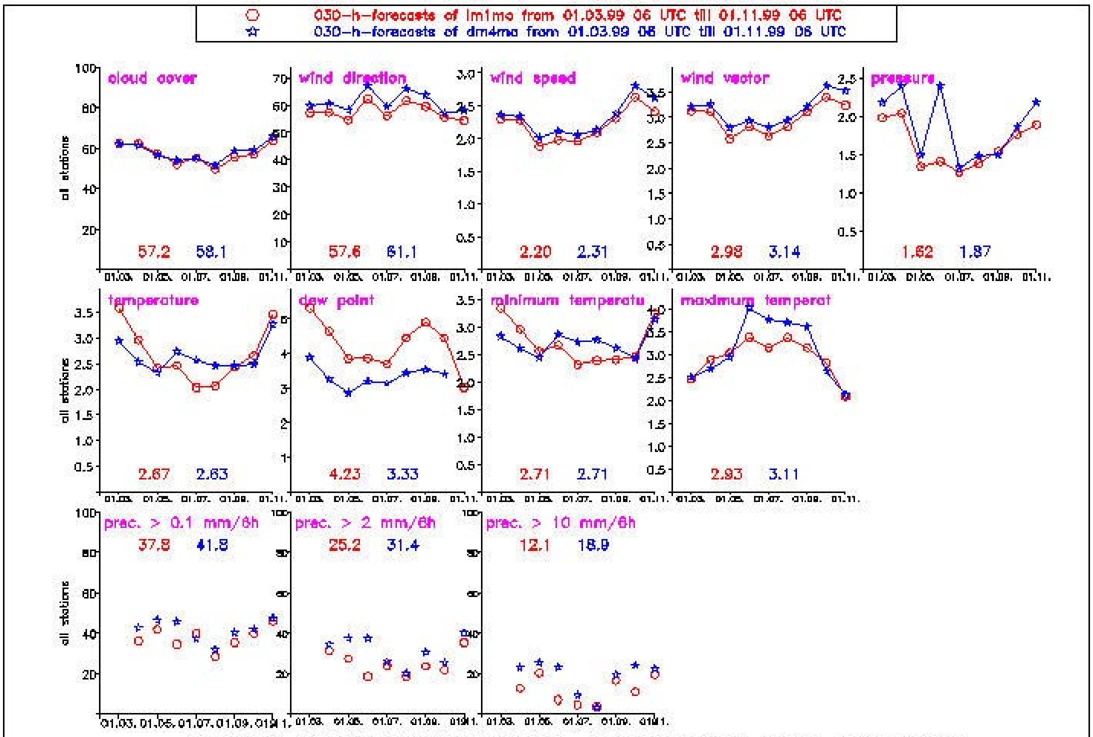
-precipitation over regions with complex orography
underestimation of precipation amount in deep valleys, overestimation on the windward side of mountains
to dry summer especially over the southern part of Germany and over the Alps due to nudging problems

-verification of small scale processes
two different approaches to incorporate small scale processes (B.Ritter and U.Damrath)
but: horizontal varability of forecasted precipitation is higher than those of observations!

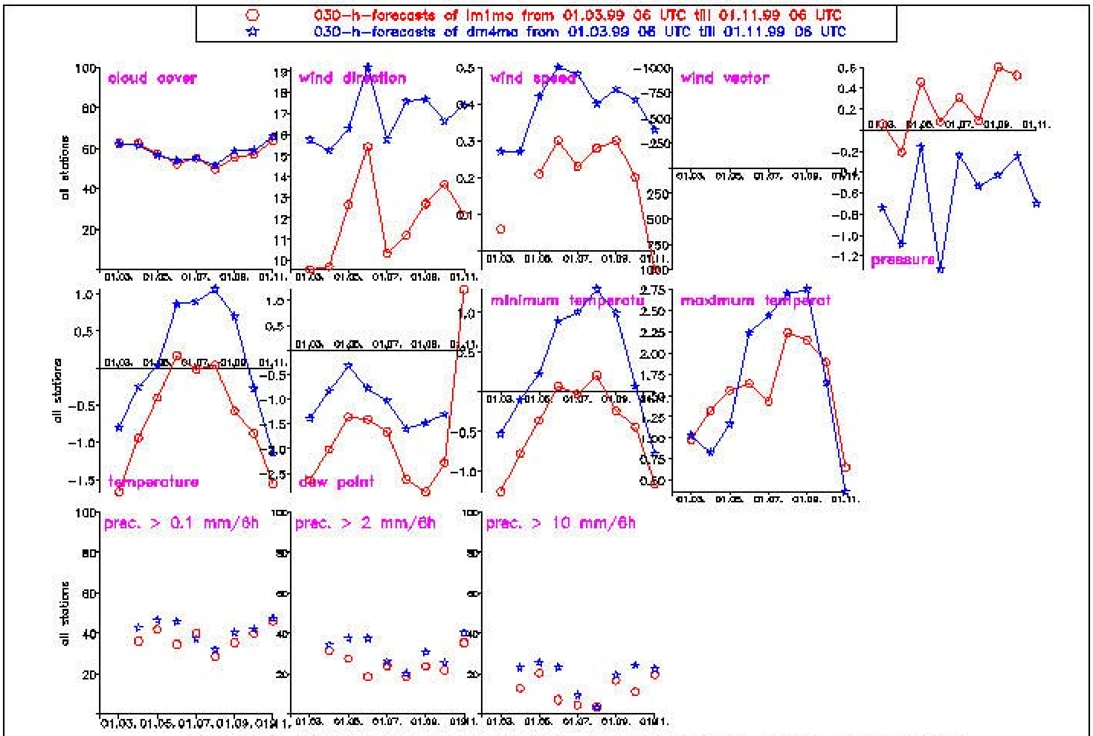
*Plans

-verification against radiosonde data
-verification against observations at towers

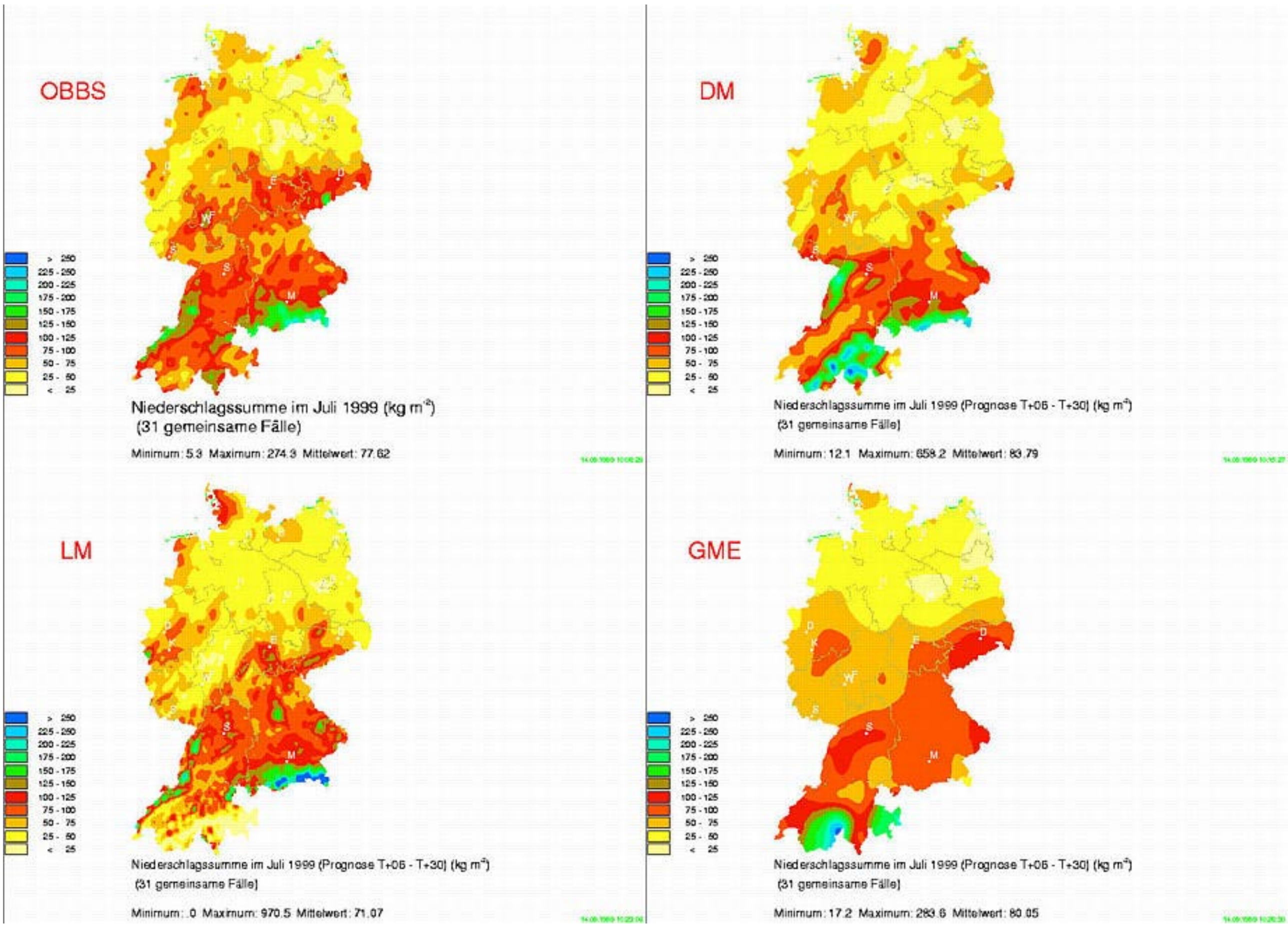
Graphics
Comparison of LM and DM (rmse of forecasts for surface weather elements)
Comparison of LM and DM (bias of forecasts for surface weather elements)
Monthly sum of precipitation for July 1999
Single case of precipitation over the Alps

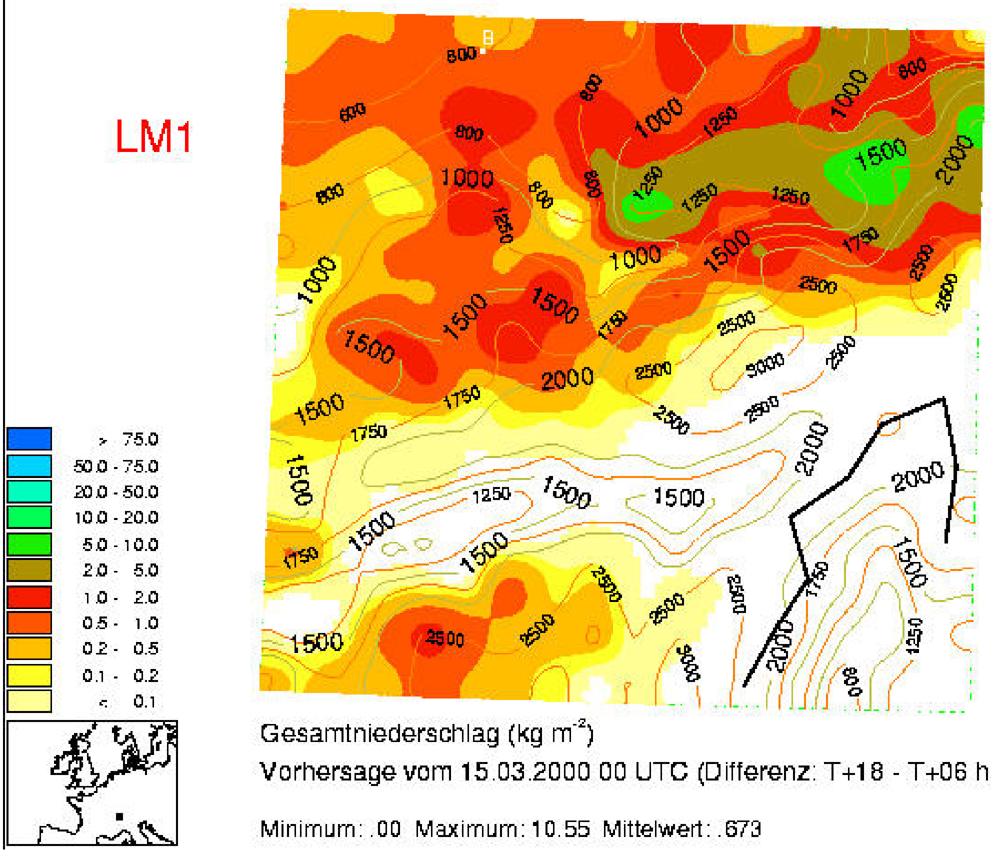


Results of verification of forecasts for local weather elements at surface weather stations
tss for precipitation, percent correct for cloud cover, RMSE for other elements
Mean values over all forecasttimes as numbers
Observations from 01.03.99 till 01.11.99



Results of verification of forecasts for local weather elements at surface weather stations
tss for precipitation, percent correct for cloud cover, mean error for other elements
Observations from 01.03.99 till 01.11.99


OBBS
> 250
225 - 250
200 - 225
175 - 200
150 - 175
125 - 150
100 - 125
75 - 100
50 - 75
25 - 50
< 25
Niederschlagssumme im Juli 1999 (kg m^{-2})
(31 gemeinsame Fälle)
Minimum: 5.3 Maximum: 274.3 Mittelwert: 77.62
DM
Niederschlagssumme im Juli 1999 (Prognose T+06 - T+30) (kg m^{-2})
(31 gemeinsame Fälle)
Minimum: 12.1 Maximum: 658.2 Mittelwert: 83.79
LM
Niederschlagssumme im Juli 1999 (Prognose T+06 - T+30) (kg m^{-2})
(31 gemeinsame Fälle)
Minimum: .0 Maximum: 970.5 Mittelwert: 71.07
GME
Niederschlagssumme im Juli 1999 (Prognose T+06 - T+30) (kg m^{-2})
(31 gemeinsame Fälle)
Minimum: 17.2 Maximum: 283.6 Mittelwert: 80.05



Gesamtniederschlag (kg m^{-2})
Vorhersage vom 15.03.2000 00 UTC (Differenz: T+18 - T+06 h)

Minimum: .00 Maximum: 10.55 Mittelwert: .673

WP 5.12: VERIFICATION OF VERTICAL PROFILES OF WATER AND ICE CONTENT USING A RADIATIVE TRANSFER MODEL ON SELECTED CASE STUDIES

L. Mannozzi and Marco Lazzeri (OGSM, Regione Marche); R.Rizzi and R.Amorati (ADGB-Bologna); C.Cacciamani and T. Paccagnella (ARPA-SMR)

The aim of this work is to verify the vertical profiles of some variables forecast by Lokal Modell (LM) using an indirect technique based on the comparison between the simulated radiances and the METEOSAT measured ones.

The LM fields we need to carry out the verification are the following:
-Temperature at the model levels;
-Mixing ratio at the model levels;
-Cloud fraction at the model levels;
-Cloud ice content at the model levels (*not yet available at DWD*);
-Cloud water content at the model levels;
-The value of pressure on the model levels;
-The surface parameters: pressure, skin temperature, 2mt temperature, 2mt dew point temperature and Land sea mask.
These parameters are given for each LM grid point in the chosen domain and at three different forecast time: 6, 12, 18 hours after the run time of the model.

The computation of simulated radiances at each Field of View (FOV) of the METEOSAT satellite requires both time and space interpolation of the LM fields. Two kinds of tools are necessary: an algorithm to compute the FOV geometry and timing, given the satellite orbital parameters and an interpolation algorithm to generate the horizontal grid and the vertical profiles of the meteorological variables to be used in the computation of the radiances. In this way, starting from the LM fields, the horizontal interpolation of each three dimensional meteorological parameter to the latitude and longitude of each FOV follows. Finally a vertical interpolation/extrapolation to the pressure levels is required.

These interpolated profiles are used as input for a radiative transfer model. In general, a radiative transfer model solves the basic radiative transfer equations of the whole atmosphere taking into account the absorption of the gases and the scattering by the aerosols and clouds. Some different numerical methods to resolve the radiative equations exists. The ADGB group performed a comparison between a fast radiative transfer model (RTTOV5) and a full scattering radiative transfer scheme, where the gaseous and cloud optical properties are computed at high spectral resolution by a line by line code. The results are satisfactory and they encourage us to use the fast model RTTOV5 in order to have an operative system with a short computation time.

The Radiative Transfer for TIROS Operational Vertical sounder (RTTOV5) is operationally used at the European Centre for Medium-Range Weather Forecasts (ECMWF) for radiance assimilation and it is implemented to calculate radiance in NOAA and METEOSAT instruments spectral channels. An absorption-type radiation scheme is incorporated in the model to take into account the cloud presence. It is a sub-grid cloud overlap scheme leading to the computation of some columns with different vertical structures. In this scheme the radiative properties of the clouds are defined by the absorption coefficient. A parameterisation of the absorption coefficients has been developed at the ADGB group. However this parameterisation was developed for spherical particles, so at the moment the ADGB group and L. Mannozzi are testing a new parameterisation for non-spherical particles.
The RTTOV5 run produces the simulated radiances at each METEOSAT FOV and these ones can be immediately compared with the satellite data.