BOREAL ENVIRONMENT RESEARCH 7: 219–223 ISSN 1239-6095
Helsinki 2 October 2002

Cluster analysis results of regional climate model simulations in the PIDCAP period

Martin Kücken, Friedrich-Wilhelm Gerstengarbe and Peter C. Werner

Potsdam Institute for Climate Impact Research, Telegrafenberg, P.O. Box 601203, D-14412 Potsdam, Germany

Kücken, M., Gerstengarbe, F.-W. & Werner, P.C. 2002: Cluster analysis results of regional climate model simulations in the PIDCAP period. — *Boreal Env. Res.* 7: 219–223. ISSN 1239-6095

On the basis of an extended cluster analysis algorithm, a new validation method is presented for the quality control of climate model simulation results. To test the method, a comparison of four climate model simulation runs was carried out, comprising on the one hand results of the non-hydrostatic model (LM) of the German Weather Service (DWD) and on the other hand results of a hydrostatic model (REMO) of the Max-Planck-Institute for Meteorology, Hamburg (MPI). The results of the cluster analysis show that the method is suitable to describe the differences between reference and simulation data in space and time.

Introduction

The local model LM is used by the German Weather Service (DWD) for weather forecasts for Germany. Version 2.1 of the LM was implemented on the IBM RS/6000 SP computer system of the Potsdam Institute for Climate Impact Research (PIK) to carry out climate simulation experiments.

The first simulation runs were started using the actual forecast area of the DWD in Europe, with the purpose of studying the long-term behaviour of the model in space and time. In the case of the Baltic Sea region, experiments are primarily conducted with a 0.5° horizontal resolution and 20 vertical levels for present-day climate conditions. For the simulation experiments, DWD analysis of the PIDCAP (Pilot Study for Intensive Data Collection and Analysis of Precipitation in the Baltic Sea water catchment region) project period from August to October 1995 were used as initial and boundary conditions. Furthermore, three-month climate simulation runs were performed with the higher horizontal resolutions of 0.25° and 0.125°.

These studies were motivated on the one hand by the need to provide reliable models for long-term climate simulation runs and on the other hand by the more fundamental interest in the underlying physical processes. One main point of research is the development and theoretical foundation of new statistical approaches and their application to practical model validation problems.

The validation of the simulation results is based on comparison with operational analysis data of the DWD with a 0.5° horizontal resolution. For this task, it is significant that for the PIDCAP period the analysis data quality is good thanks to the large amount of observations (Isemer 1996) and in this sense these data represent the structures of a climate regime as exactly as possible and are well suited as reference data.

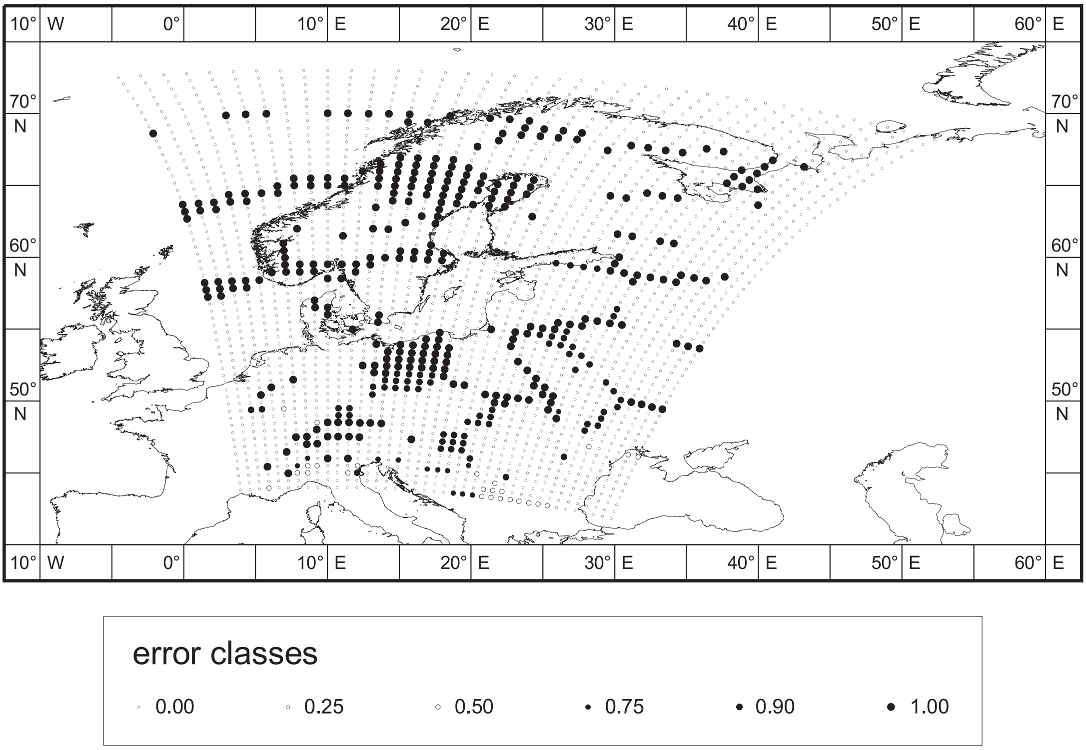

Fig. 1. Error classes of a three-month LM simulation.

A multi-variate statistical technique (cluster analysis) is used to validate different model variables in different model runs simultaneously. Detailed studies of the differences help to identify any weaknesses in the models.

To compare LM simulation runs with other models, simulation results of the hydrostatic local grid-point model REMO (Jacob *et al.* 1997, Jacob *et al.* 2001) for the same model area and the same simulation period were chosen. The investigated REMO simulation run employs a physical parameterization scheme developed by the DWD (Majewski 1991. The same scheme was further refined for the LM (Doms *et al*. 1999).

This paper will describe the cluster analysis validation approach and will give examples on how the cluster analysis method have been applied succesfully.

Validation method

Model validation plays an important role in the development of climate models. Examples of validation methods are simple visual comparison, the use of generalizing measurement figures (e.g. root mean square error: RMSE) and the application of fully developed statistical methods. The latter often require relatively amounts of high calculation time and so they are rarely used in practice. Nevertheless, the evaluation of complex correlations can only be accomplished by using such complex methods. In order to make progress, a validation method was developed which is based on multi-variate pattern recognition and provides measurement figures for quality evaluation that can easily be interpreted, and whose spatial distribution and temporal development serve to derive information on possible error sources (Gerstengarbe *et al.* 1997, Gerstengarbe *et al.* 1999).

The question that should be answered by the method is: How well does a model represent defined complex structures of a climate regime? As a first step, the validation goals were defined by selecting interesting model parameters. Near-surface parameters were used to validate the results of the long runs from the viewpoint of

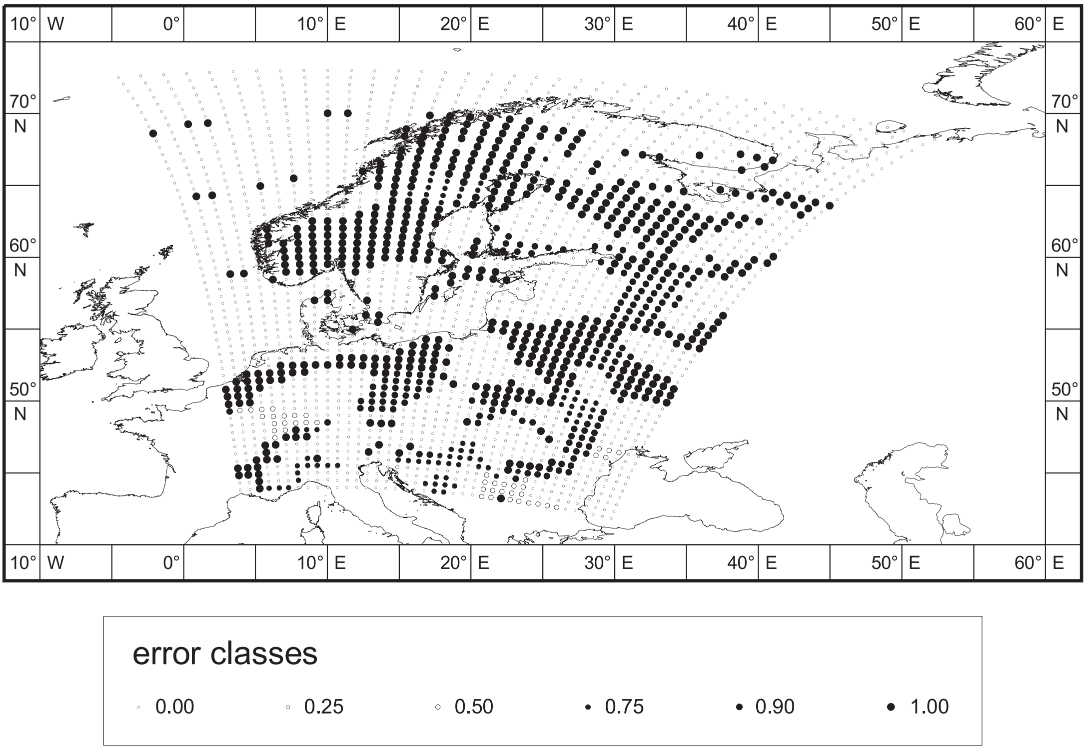

Fig. 2. Error classes of a three-month REMO simulation.

a climatologist. The cluster analysis itself was carried out with weekly averaged mean values and variances of both 2-m dew point temperature (TD2M) and 2-m temperature (T2M). This approach leads to a parameter number of 4 parameters for a cluster analysis for a single week (mean value and variance both for TD2M and T2M) and of 52 parameters for a period of 13 weeks (4 parameter per week multiplied by 13). Then, based on the selected model parameters, the reference grid points of the validation area are grouped into clusters by means of the cluster analysis algorithm of Gerstengarbe *et al.* (1997). The number of clusters depends on the heterogenity of the validation area.

After that it is only necessary to use the euclidian distance to relate the parameter combination of the modelled data set in accordance to the cluster number of the same gridpoint within the reference data set. Each cluster of the reference data and the simulation data represents its own parameter space. In case of a non-agreement of the cluster assignment between the simulation and the reference data set the value of the simulation error is inversely proportional to the distance of these clusters. The value zero indicates that the clusters of these reference and simulation grid points are identical, the numerical value one indicates that they are completely different. However, values between zero and one can also serve as a deviation criterion. Generally, these numerical values can be considered as error classes and simply indicate the shift of model outputs compared to the model patterns identified in the reference data set.

In the following the effectiveness of the method is demonstrated by comparing the quality of different regional climate simulation runs.

Simulation results

The simulation carried out with the LM for a three-month simulation period (Fig. 1) and the corresponding REMO simulation run (Fig. 2) are based on the same model region and the same simulation period. The most important difference between the models is the use of non-hydrostatic

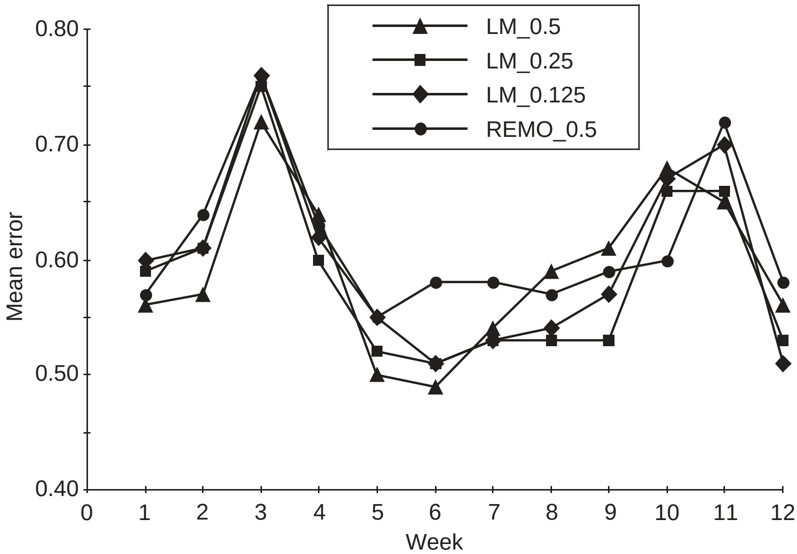

Fig. 3. Temporal development of error classes for a three-month REMO and three three-month LM simulations.

prognostic model equations for LM but hydrostatic prognostic model equations for REMO. Because of this, even the numerical calculation of the integration scheme is different, in particular there is a strong difference between the time integration schemes of REMO and LM.

There are remarkable differences between the two scenarios. In the ideal case (where the grid points of simulation and analysis represent the same cluster) only the error class zero will occur. The area mean value of the grid point error classes for the REMO simulation of 0.34 decreases in the LM simulation to a value of 0.16. The model REMO which is based on DWD physics is identical to the DWD operational weather forecast model of 1990. Hence, a great improvement can be seen in the decrease of the error classes of the new DWD weather prediction model version LM. But especially in the case of mountain barriers or transitions from land to sea, the model LM still needs some improvement. This can be seen in some areas in the fronthills of the Alps.

The investigation of the model's behaviour over time is another approach used to check the quality of the model. The question of whether and for how long a model is able to simulate a sequence of well-known real meteorological conditions realistically has the same importance as the question of whether the models show the same persistence characteristics for the modelling area. Figure 3 shows the time series of the averaged error classes over the entire area for each week in a period of three months both for the LM with different grid resolutions and for REMO. For these comparisions LM simulation runs with a 0.5° horizontal grid resolution (LM_0.5), with a 0.25° horizontal grid resolution (LM_0.25) and with a 0.125° horizontal grid resolution (LM_0.125) were investigated, further a REMO run with 0.5° horizontal grid resolution (REMO_0.5). Again the analysis data are used as the reference field for both models. On the basis of these time series, the temporal dynamics of the individual model runs can be identified. However, the error classes are much larger here compared to the error classes for the entire period with 52 parameters (seen in Figs. 1 and 2) due to the sharper cluster separation which is a consequence the lower number of only four parameters per week.

All investigated simulation runs show a similar run time behaviour. One reason for this effect could be the increasing dominance of the similiar physical parameterization in relation to the different model dynamics. Remarkable is on the one hand the similar temporal behaviour of the LM and REMO but on the other hand the similiarity of the LM runs with different grid resolutions. From the statistical point of view, an interpretation of trends on the basis of only 12 time series intervals is not sufficient. The temporal fluctuations of the error classes that can be observed in all model variants could be caused by different weather conditions.

Conclusions

It has been shown that the suggested cluster method is able to evaluate the quality of the spatial and temporal simulation of the climate conditions. The error classes do not significantly increase for the investigated model versions, except for some accidental fluctuations. The model results do not 'drift off'.

Differences in the simulations can easily be worked out with the presented method. For a three-month prediction period, the new LM of the DWD is able to reproduce the basic weather patterns in space and time. The error analysis of the deviations over Central Europe is difficult since a clear reason for these deviations was not found in a first assessment. An intensive analysis of the model behaviour and the model structure would be necessary for this (this is not part of these investigations).

Acknowledgement: The authors wish to thank Daniela Jacob (MPI) for her permission to use REMO results for this report.

References

Doms G., Schättler U. 1999. The non-hydrostatic limited-area model LM (Local Model) of DWD Part I: Scientific documentation. *Technical Report DWD*.

Gerstengarbe F.-W., Werner P.C. 1997. A method to estimate the statistical confidence of cluster separation. *Theor. Appl. Climatol.* 57: 103–110.

Gerstengarbe F.-W., Werner P.C., Fraedrich K. 1999. Applying non-hierarchical cluster analysis algorithms to climate classification: Some problems and their resolution. *Theor. Appl. Climatol.* 64: 143–150.

Isemer H.J. 1996. Weather patterns and selected precipitation records in the PIDCAP period, August to November 1995. *GKSS external report GKSS* 96/E/55: 1–92.

Jacob D., Podzun R. 1997. Sensitivity studies with the Regional Climate Model REMO. *Meteorol. Atmos. Phys.* 63: 119–129.

Jacob D., Van den Hurk B. J. J. M., Andrae U., Elgered G., Fortelius G., Graham L.P., Jackson S.D., Karstens. U., Köpken C., Lindau R., Podzun R., Rockel B., Rubel F., Sass B.H., Smith R.N.B., Yang X. 2001. A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period. *Meteorol. Atmos. Phys.* 77: 19–43.

Majewski D. 1991. The Europa-Modell(EM) of the Deutscher Wetterdienst. *ECMWF seminar on numerical methods in atmosphere models* 2: 147–191.

Received 23 January 2002, accepted 5 July 2002