DANISH METEOROLOGICAL INSTITUTE

——— SCIENTIFIC REPORT ———

03-07

An European Flood Forecasting System EFFS

Treatment of uncertainties in the prediction of heavy rainfall using different ensemble approaches with DMI-HIRLAM

Kai Sattler and Henrik Feddersen

COPENHAGEN 2003

Contract no: EVG1-CT-1999-00011
Project co-ordinator: WL|DELFT HYDRAULICS
Internet site: http://effs.wldelft.nl
Workpackage 6
Deliverable no. 6.2

ISSN Nr. 0905-3263 (printed)
ISSN Nr. 1399-1949 (online)
ISBN-Nr. 87-7478-480-3

An European Flood Forecasting System EFFS

Treatment of uncertainties in the prediction of heavy rainfall using different ensemble approaches with DMI-HIRLAM

Kai Sattler and Henrik Feddersen

Danish Meteorological Institute, Lyngbyvey 100, DK-2100 Copenhagen Ø

DMI Scientific report 03-07

ABSTRACT

The EU-supported project *An European Flood Forecasting System* (EFFS) involves many meteorological and hydrological institutions in Europe. Its aim is to provide a prototype of a European flood forecasting system for 4–10 days in advance. Within the work package 6 *Downscaling and Validating Meteorological Forecasts* an investigation on the treatment of uncertainty inherent in the prediction of high-resolution precipitation fields has been performed by the Danish Meteorological Institute (DMI) using the DMI-HIRLAM numerical weather prediction model. The investigation and its results are described in this report.

The study concentrates on the periods around three historical heavy rain events, which caused or at least contributed significantly to a subsequent severe river flooding. The three cases differ in their dynamic structure both in space and in time as well as by the type of dynamic forcing that triggered precipitation. The periods covered by this study comprise at least 10 days for each of the events. The three historical events are: the severe flooding in the Piemonte region in North Italy in November 1994, affecting the rivers Po and Reno, the Rhine/Meuse flooding at the end of January 1995 and the severe Odra flooding in Poland during July 1997.

Two different ensemble approaches are investigated using the high-resolution limted-area model for the atmosphere DMI-HIRLAM. The first ensemble is designed to represent forecast uncertainty arising from the initial atmospheric state uncertainty and at the lateral boundaries of the limited-area model. The second ensemble design addresses model inherent uncertainty through a multi-scheme approach for the parameterization of condensation and convection. A probabilistic verification using Brier Skill score, reliability diagrams, rank histograms and relative operating characteristic is performed against precipitation analysis from the Deutsche Wetterdienst (DWD). It is completed by case studies regarding basin average rainfall and probabilities of precipitation (POP). The verification and the case studies show the capabilities of the ensemble approaches to predict the heavy rain events and provide indication for future improvements.

The DMI-HIRLAM ensemble predictions are also to be applied in the hydrological models within the EFFS project in order to assess water levels probabilistically and also to investigate the usefulness of the high-resolution ensembles compared to the use of the data from the global ensemble prediction system (EPS) of the European Centre for Medium-Range Weather Forecasts (ECMWF).

Contents

1 Introduction

Uncertainty occurs virtually in all fields of science, where predictions are to be carried out. In the very complex system of the atmosphere, uncertainty arises from a wide spectrum of spatio-temporal scales of the atmospheric state variables and parameters. This is firstly due to the impossibility of precisely and completely observing and describing the state of the atmosphere, the ocean and the land surface at a given point of time, and secondly due to the imperfect representation of the atmosphere's dynamics and its energetic cycles and cascades in the model description. Additionally, the equations used to describe the atmospheric behaviour are tightly coupled and of highly non-linear nature, and this leads to a complex cascade of perturbation propagation in time and space, making a simple error tracking impossible. If, for example, the initial uncertainties in the state variables all were normally distributed, a Gaußian probability distribution function (PDF) in state space would describe this circumstance. In the system of the atmosphere, this symetric PDF evolves typically non-linearly towards a multi-modal distribution and finally may disperse completely within the state attractor. As a consequence, the atmosphere exhibits a limited predictability, which turns out to be very inhomogeneous when regarding different synoptic weather situations. These circumstances have been known since decades (Thompson, 1957; Lorenz, 1963).

The most promising and still practical approach to address the complex propagation of uncertainty in numerical weather prediction (NWP) seems to be the use of an adequately designed ensemble prediction system, attempting to forecast the PDF in order to estimate possible future events rather than a single "best guess" forecast (Ehrendorfer, 1997). This approach has been followed for quite some time, and the first operational ensemble prediction systems (EPS) based on global NWP models were established during in the early 1990s (Houtekamer and Derome, 1994; Molteni et al., 1996; Toth and Kalnay, 1993). In order to enhance the ensemble forecast of the early medium-range over a specific area of interest, recent experiments investigated the use of targeting the EPS from the European Centre for Medium-Range Weather Forecasts (ECMWF) onto a certain area, like for example Europe, by utilizing targeted singular vectors for deriving of the initial perturbations (Hersbach et al., 2000; Frogner and Iversen, 2001). These ensemble approaches were still based on the global model from ECMWF.

Some years after the operational establishment of the global EPSs, studies investigating the usefulness of ensemble designs that utilize a limited-area model (LAM) for short-range probabilistic weather prediction were reported (Du et al., 1997; Hamill and Colucci, 1997; Stensrud et al., 1999; Molteni et al., 2001; Frogner and Iversen, 2002). Specific emphasis has been put on quantitative precipitation forecasting (QPF) since then, because the predictability horizon is shorter for precipitation than for more conservative parameters like mean sea level pressure, geopotential height or even upper level temperature.

One of the targets for utilizing the LAM ensembles has been the prediction, or better, the probabilistic assessment for the occurrence of significant rain events, which can lead to severe damage (Du et al., 1997; Bright and Mullen, 2002; Frogner and Iversen, 2002). This target is pursued in this work, too. The study investigates the use of high-resolution LAM ensemble designs, which try to capture the comparatively rare events of strong precipitation despite of a small ensemble size.

The model used for the investigations is the HIgh Resolution Limited-Area Model (HIRLAM) in its version as it is used at the Danish Meteorological Institute (DMI), DMI-HIRLAM. Even though the size of the ensembles had to be limited in order to respect the available computer resources, two different approaches have been investigated. They address two of those uncertainty sources, which are regarded to be among the most significant error sources in quantitative precipitation prediction. These are the initial condition and the parameterization of condensation and convection. A third source of uncertainty arising at the lateral boundary of the LAM is being addressed to the extent, to which it is connected to the atmospheric state description. This is done through the use of a "perturbed" host model in form of a host model EPS.

The report is structured as follows. The second chapter gives a description of the treatment of the mentioned uncertainties arising in the LAM. It includes a description of the member selection for the first ensemble approach as well as a short description of the parameterization schemes utilized in the second ensemble approach. Section 3 outlines the DMI-HIRLAM model and its configuration for the ensemble integrations of this study. A concise description of the three historical heavy rain events that were selected is presented in Section 4. The probabilistic verification against the DWD precipitation analysis follows in Section 5, and results from the case studies are given in Section 6, which also includes a simple evaluation of the ensemble member selection procedure. Finally, Section 7 summarizes the conclusions from the study.

2 Treatment of uncertainties in a LAM

It has become common to address uncertainty in NWP by using an ensemble of forecasts in the hope of capturing the possible future developments of the weather with the modelled PDF as good as possible. Instead of one single assessment of the future state, the ensemble provides a probability distribution for virtually each atmospheric variable, thus including information about the possibility and the extent of deviations from the expectation value, or more precisely, providing indication of a possible multi-modal behaviour.

Several studies have been undertaken in recent years to design LAM ensembles. Fig. 1 outlines a number of possibilities for perturbing a LAM forecast. They are not described in detail here, some of them are quite complex already in their basic concept. The sketch makes, however, clear the complexity of the matter, in that it visualizes the many sources for uncertainty and their connection to each other. One of the early studies regarding different perturbation approaches is probably that of Houtekamer et al. (1996), who investigated a system of initial perturbations including also perturbations of surface parameters. The perturbation of the initial state is the most investigated one, because it is regarded as one of the major processes concerning uncertainty, and several global EPSs have been running operationally since about a decade (Houtekamer and Derome, 1994; Molteni et al., 1996; Toth and Kalnay, 1993). Where regional models are concerned, Misra and Yau (2001) studied e. g. a summer case with a high-resolution model using an ensemble of perturbations based on diabatic initialization in order to address initial uncertainties in the heat sources and sinks.

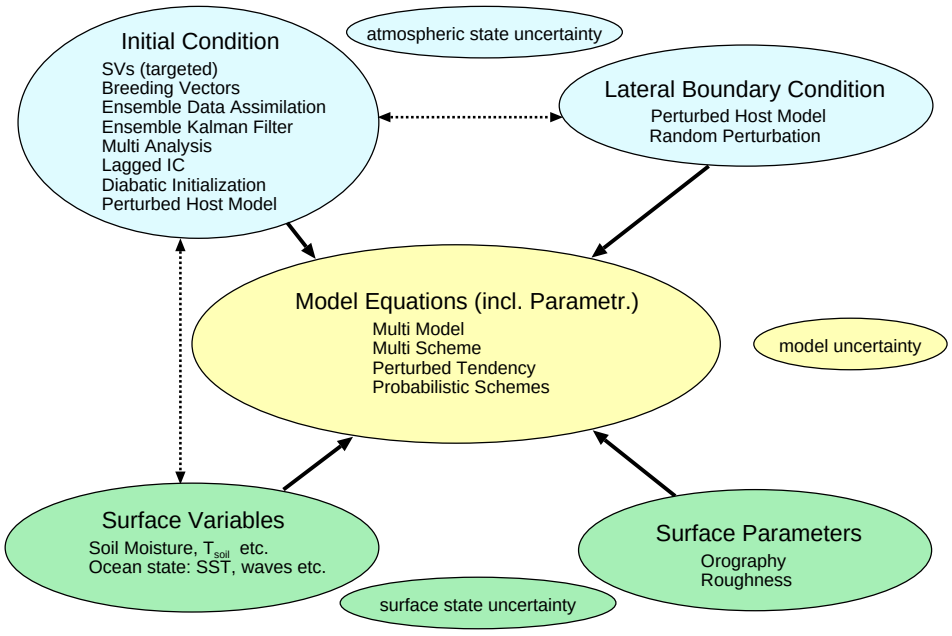

FIG. 1: Perturbation concepts for LAM ensemble prediction.

Another major contributor to the forecast errors are the model equations. Most of the approaches listed in Fig. 1 (centre) have been or are being investigated. An approach of stochastic perturbations of the tendencies within the model parameterizations was established by Buizza et al. (1999). Multi-scheme approaches have been studied recently for example by Stensrud et al. (2000) as well as Bright

and Mullen (2002). It may be expected in the future that "probabilistic parameterization schemes", which try to model a PDF for the parameterized variables will come up. The lateral boundary condition only plays a role in LAMs. Du and Tracton (1999) have pointed out that the LAM domain size has a significant impact on the ensemble spread because of the role of the lateral boundary. Finally, the processes close to the surface become more and more important when increasing the horizontal resolution of the model.

As precipitation is one of those atmospheric parameters, which is related to the prognostic variables both through dynamic as well as thermodynamic processes, there is virtually no atmospheric equation, which does not exhibit an implicit or explicit impact on precipitation. As a consequence, it is sensitive to perturbations of practically all prognostic variables and thus also to the initial and the boundary conditions of the involved equations. This means that all sources of uncertainty in an NWP model can be expected to contribute to some extent to the uncertainty in a precipitation forecast. As the number of sources is vast, it is important to try to address those sources in the first place, which are expected to have a major impact on precipitation uncertainty. The uncertainty in the initial atmospheric condition plays one major role, because it is an inherent property of the atmospheric dynamics and affects already the conservative variables, which precipitation depends on. It is addressed in the first LAM ensemble approach of this work by utilizing a perturbed host model, which also involves the lateral boundaries. Another source, which is regarded to have a major impact on precipitation prediction, is the model treatment of the cloud processes, which directly lead to precipitation. The major part of this treatment happens within the parameterization schemes of condensation and convection, because the model cannot resolve these small scale processes exhaustively. The second LAM ensemble approach tries therefore to address this model uncertainty by a multi-scheme approach. The basic concepts of the two ensemble approaches are described in the following sections.

2.1 Uncertainties in the initial condition

A reason why the behaviour of the atmosphere appears to be chaotic and hard to predict is the impossibility of precisely and completely determining the state of the atmosphere, ocean and land surface globally at a certain point of time. In order to perform a numerical simulation of the future atmospheric development with an NWP model it is, however, necessary to provide the atmospheric state to the model at the time the forecast is to start. Even though the common data assimilation techniques in NWP provide astonishingly good estimates for this initial condition (IC), some uncertainty remains. As the atmospheric behaviour is still very sensitive to even very small changes in the state estimation, two only slightly different state estimations may result in exponentially diverging future states already at an early stage exhibiting a behaviour, which is described as to be chaotic (Lorenz, 1993; Ott, 1993).

This realization led to the design of ensemble prediction systems based on perturbed initial condition (Toth and Kalnay, 1993; Houtekamer and Derome, 1994; Molteni et al., 1996). They are to address the uncertain future development due to the intrinsically aperiodic behaviour of the atmosphere when its state is not precisely known. None of the commonly utilized methods for IC perturbation are utilized directly in the approach of this work, because a LAM not only needs a state description

at its initial point, but frequent additional state descriptions at its lateral boundaries (see below), which is not necessary for global NWP models.

In order to address both initial and boundary condition consistently a perturbed host model is utilized in the LAM ensemble approach. This means that ensemble members from perturbed host model simulations represent the host data for an ensemble of LAM integrations. ECMWF's operational global EPS has been chosen as host model for this purpose. Its ensemble size is, however, too large for currently given computer resources in order to apply a high-resolution LAM like HIRLAM on each member of the ensemble. Therefore a selection of members is undertaken sampling those members, which are believed to be most significant for the prediction. In the case of this work the significance is measured with respect to accumulated precipitation. The resulting HIRLAM selection ensemble (SE) is chosen to be of size five. In addition, the control forecast from the ECMWF-EPS is always considered, because it is based on the unperturbed analyses of the initial atmospheric state, and this is still believed the "best guess" for the atmospheric state resembling the most probable state from a statistical point of view.

Another reason for using a reduced ensemble member set is the assumption that a part of the members from the global EPS do not behave significantly different with respect to the control or other members of the ensembles within the limited area covered by the LAM domain. An approximated selection procedure is expected to be capable of identifying such members implicitly.

The selection of the global ensemble members is performed in the following way. A parameter space of dimension D_p is regarded, where each dimension is described by the 5-day accumulated grid-point precipitation on a geographical target grid covering the area of interest. This area contains the river basins under consideration plus a surrounding area of approximately 1° thickness in longitude and latitude. Thus, D_p is defined by the number of grid points in the target area. The accumulation period starts at the initial time of the global EPS forecast and ends at forecast day 5, and both stratiform and convective precipitation are accumulated.

Let $\{E_m \,|\, m = 0, M\}$ be the global EPS of size M plus the control E_0, and let $\{S_l \,|\, l = 0, L\}$ be the LAM-EPS of size L plus the control S_0. Then the selection is described by

$$
\begin{array}{rcll}
S_0 & = & E_0 & \\
S_1 & = & E_k \quad \wedge & \| E_k - \langle E_M \rangle \| \;\; \leq \;\; \| E_m - \langle E_M \rangle \| \qquad \forall\ m \\
S_2 & = & E_k \quad \wedge & \| E_k - \langle E_M \rangle \| \;\; \geq \;\; \| E_m - \langle E_M \rangle \| \qquad \forall\ m \\
S_{i+1} & = & E_k \quad \wedge & \| E_k - \langle S_{i+1} \rangle \| + \sum_{j=2}^{i} \| S_j - \langle S_{i+1} \rangle \| \;\; \geq \\
 & & & \| E_m - \langle S_l \rangle \| + \sum_{j=2}^{i} \| S_j - \langle S_l \rangle \| \qquad \forall\ m, \quad i = 2, L-1
\end{array}
\tag{1}
$$

The pointed brackets denote the average, where $\langle E_N \rangle$ is the mean of the global ensemble, and $\langle S_{i+1} \rangle$ denotes the mean of the $i+1$ selected members. The norm $\| \ldots \|$ denotes the Euclidean distance, which is used as similarity measure in the selection process. It is determined between two points I_a and I_b in the parameter space,

$$
\| I_a - I_b \| = \sqrt{\sum_{i=0}^{D_p} (r_{a,i} - r_{b,i})^2}, \tag{2}
$$

where $r_{a,i}$ denotes the rainfall at parameter space point a and target grid point i. The SE consists of at least 3 members: the control, the member closest to the global ensemble mean, and the member farthest away from the global ensemble mean. Further members are selected such that the dispersion of the SE is maximized.

2.2 Uncertainties at the lateral boundary

Uncertainties arising at the lateral boundary are twofold. Firstly, a LAM not only needs a state description at its initial point, but frequently additional state descriptions at its lateral boundaries, which is especially important at the inflow boundaries. Otherwise important upstream atmospheric developments outside the LAM domain cannot be captured by the LAM simulation. Frogner and Iversen (2002) have shown this necessity in their LAM ensemble experiments. As the state description is, however, tainted with uncertainty, a similar situation occurs here as for the IC. The dynamical prescription of the lateral boundaries at all times during the integration is therefore addressed together with the IC uncertainty. This is realized by utilizing the perturbed ICs of the host model together with resulting perturbed host model integrations. This also assures that the dynamical development starting from the IC is fully consistent with the dynamic prescription at the lateral boundaries at all times during the LAM forecast. An uncertainty remains, however, due to the gap in time between two boundary descriptions from the host model, which in this case is typically 6 hours.

A second type of uncertainty at the lateral boundaries arises from the approximations used to deal with the boundary effects. In addition to the model intrinsic uncertainty due to the approximated description of the atmospheric equations, which are also present at the boundaries, the treatments to make the boundary as transparent as possible as well as the relaxation scheme to damp the rest of undesired gravity wave reflections may introduce further sources of uncertainty. This type of uncertainty, however, has not been explicitly addressed in this work. However, the size of the domain of the outer model of the chosen nested system has been chosen to be large (Sec. 3.1), which is in accordance with recommendations from Du and Tracton (1999).

2.3 Model uncertainties

Uncertainties arising from the model formulation and discretization are expected to occur nearly everywhere in the model. However, the components of the model contributing to the uncertainty of a certain variable may be very different when regarding different parameters, and some components may exhibit more significance than others. Where precipitation is concerned, it is assumed that one of the largest contributions to the model intrinsic uncertainty arises from the parameterization of convection and condensation. One possibility to address this is to use different parameterization schemes for convection and condensation, and this parameterization ensemble (PE) approach has been followed in this work. An important prerequisite for the ensemble design is, however, that the utilized parameterization schemes perform approximately equally well. This allows for an equal distribution of the assigned probability weights among the members of the ensemble.

The following subsections outline the different schemes for the parameterization of convection and condensation that are available in DMI-HIRLAM and that are utilized in the PE ensemble design. Except for the Tiedtke scheme they where evaluated by Nielsen et al. (1995) both objectively by verification over several periods for mean sea-level pressure, geopotential height, temperature and wind, and subjectively through case studies. They found that the schemes performed comparably well within the DMI-HIRLAM model. A first approximation validation of the performance of the schemes during the heavy rain periods investigated in this work is given in Sec. 2.3.6.

2.3.1 The stratiform condensation scheme (STC)

This scheme has been included in HIRLAM since the model's first version (Källberg, 1989). It determines stratiform condensation and precipitation by a first order adjustment of temperature and humidity. Condensation is defined to occur beyond a critical saturation value smaller or equal to the grid-box saturation specific humidity. The scheme includes the parameterization of stratiform precipitation evaporation proportional to the saturation deficit (Kessler, 1969). The scheme does not account for unresolved convective processes.

2.3.2 The Anthes-Kuo scheme (AKC)

This scheme was proposed by Kuo et al. (1965) and further developed by Kuo et al. (1974) and Anthes (1977). It was implemented into HIRLAM as described by Källberg (1989) and parameterizes convection, that may occur within conditionally unstable layers on the basis of resolved moisture convergence. This affects the moisture change in the layer as well as the release of precipitation, which is dependent on the relative humidity and its vertical distribution within the layer. The scheme describes the temperature and humidity of the emerging clouds using a quasi-static assumption between air parcel and surrounding pressure and separating dry and moist ascent. It includes also cloud temperature changes due to latent heat release. Evaporation of convective precipitation is determined similarly as in the stratiform condensation scheme using the procedure of Kessler (1969). The precipitation rate is, however, adjusted by the fraction of cumulus cloud cover in order to account for sub-grid variability. Freezing and melting processes of precipitation are not included in the scheme.

The parameterization was designed for large scale models and proved to be robust (Pielke, 2001). In its version included in DMI-HIRLAM the scheme can be combined with the stratiform condensation scheme described above in order to represent the condensation and precipitation processes more comprehensively. This was done in the simulations performed with the scheme in this work.

2.3.3 The Sundqvist scheme (SQS)

This scheme treats stratiform and convective condensation regimes and includes a parameterization of microphysical processes (Sundqvist et al., 1989). The criterion for distinguishing the two regimes is based on the buoyancy state of surface air after being lifted to the lifting condensation level (LCL). The regime is assumed convective, if the air is positive buoyant, and stratiform otherwise.

Both treatments include a diagnosis of cloud cover. Anvil representation above a convective cell is realized by a stratiform treatment at the top level of convection, as soon as the temperature at this level falls below -20 °C.

The scheme for microphysical processes introduces cloud water as a prognostic variable including the ice phase (Sundqvist, 1993). Thus, it parameterizes autoconversion, coalescence, the Bergeron-Findeisen mechanism as well as evaporation and melting processes of precipitation. A special treatment of the Bergeron-Findeisen mechanism is included for cirrus clouds.

The convective part of the scheme adopts the scheme of Kuo et al. (1974) and includes a similar scheme like Anthes (1977) for distributing the moisture convergence to condensation and moistening within the convective layer. An additional buoyancy evaluation in thin layers allows for parameterizing shallow convection. Detrained cloud water evaporation is included, too.

The stratiform parameterization of the scheme determines stratiform condensation due to humidity exceeding a saturation threshold, which is based on the diagnosed cloud cover in order to account for the limited resolution of the model. The convergence of available latent heat is divided into condensation within the cloudy parts, and into moistening in the cloud free areas.

The implementation of the Sundqvist scheme into HIRLAM is described by Källén (1996).

2.3.4 The Tiedtke scheme (TDS)

The Tiedtke scheme is a mass-flux type scheme for convection and condensation (Tiedtke, 1989; Tiedtke, 1993). It is capable of distinguishing between shallow, medium and deep convection. In the version adopted in DMI-HIRLAM shallow convection is related to surface evaporation, and to the large-scale moisture convergence otherwise. A description for cumulus downdrafts is included as well.

The treatment of stratiform cloud regimes as well as the parameterization of microphysical processes were adopted from the Sundqvist scheme in the version utilized in DMI-HIRLAM. This includes the possibility for the existence of cirrus clouds and anvils. The representation of microphysical processes comprises parameterizations for autoconversion, coalescence and the Bergeron-Findeisen mechanism.

The Tiedtke scheme was originally developed for large-scale models and implemented into the ECMWF global model for operational use (Miller et al., 1995). From there it was adopted for utilization in the DMI-HIRLAM system.

2.3.5 The STRACO scheme

The STRACO scheme (Soft TRAnsition COndensation) parameterizes moist processes and includes stratiform regimes as well as convection of variable depth and height (Sass et al. (1997), Undén et al. (2002)). It allows for convection and condensation to take place in several vertical layers at different levels at the same time. The number of layers is formally not limited.

The description of convective activity is based on the scheme of Kuo et al. (1974) and includes a modified moisture convergence closure, which also considers the influence of the surface evaporation

flux as well as convective transport of cloud condensate. Stratiform regimes are parameterized by considering an equilibrium in a first order adjustment of the cloud condensate, which is associated with changes in temperature and humidity due to other processes.

A specific feature of the scheme is the inclusion of a relaxed transition between stratiform and convective regimes and vice versa. This transition controls condensation, moistening and latent heat release by considering a scaled depth of the convective layer as well as its buoyancy with respect to a buoyancy threshold. Additional consideration of the mesh size of the model grid in the three mentioned processes allows for a smooth adaptation of the parameterization to horizontal model resolution, i. e. an increased horizontal resolution reduces the effect of the convective parameterization in order to account for the increased part of model resolved convective activity. The smooth transition between stratiform and convective regimes and vice versa is also expressed in the way stratiform and convective cloud cover are referred to the total cloud cover f_{tot}. An equilibrium total cloud fraction f_{eq}, which is either characterized by stratiform or convective cloud cover converts to the new regime in a relaxation process:

$$\frac{\partial f_{tot}}{\partial t} = -\frac{1}{\tau_f}\left(f_{tot} - f_{eq}\right), \tag{3}$$

where the relaxation time for the transition τ_f is set to 900s.

The scheme includes formulations for cloud microphysics, which are very similar to those of the Sundqvist scheme (Sundqvist, 1993), i. e. cloud condensate can be composed of water and ice and processes like autoconversion, coalescence and the Bergeron-Findeisen mechanism are parameterized. A statistical description for the total water content in the cloud using an asymmetric rectangular probability distribution function (PDF) for sub-grid variation is used in the formulation of convective cloud fraction (Redelsperger and Sommeria, 1986; Sass, 2002). Stratiform cloud fraction is similarly formulated using a symmetric rectangular PDF. In addition, a model resolution dependency as originally proposed by Redelsperger and Sommeria (1986) is included in the stratiform case (Sass, 2002), as well as consideration of the variation of the PDF amplitude with height above ground level (Sundqvist et al., 1989).

The STRACO scheme has been implemented in DMI-HIRLAM since the first version of the scheme. It is continuously under development and has been used in the operational NWP at DMI since some years (Sass et al., 2002). The scheme has been used in this work as the standard scheme.

2.3.6 Performance of the parameterization schemes

A simple test on the performance of the parameterization schemes has been performed on basis of a multi-event contingency table for 12h accumulated precipitation. It comprises five categories:

- < 0.2 mm
- $0.2 - 1.0$ mm
- $1.0 - 5.0$ mm
- $5.0 - 10.0$ mm
- > 10.0 mm

An integrated hit rate was determined for these categories by summing over the individual hit rates and relating the sum to the total number of comparisons. This procedure is not very well suited for a quantitative evaluation, because there is a high penalty for those predictions, which come to ly just outside the event category for a hit. For a qualitative estimation of the relative performance between the parameterization schemes, however, this approach is regarded useful in a first approximation.

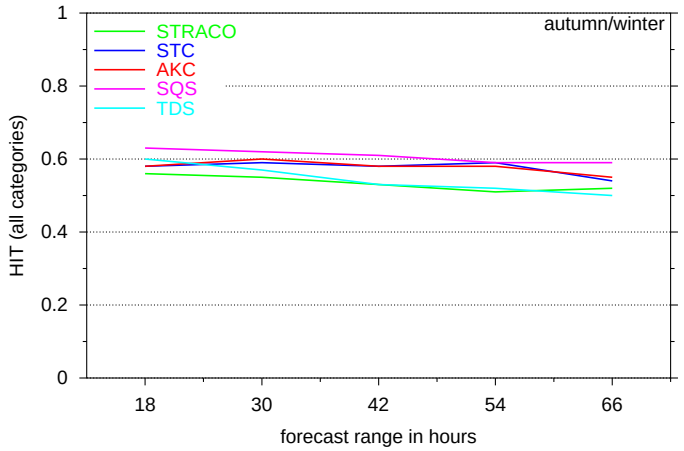

FIG. 2: Integrated hit rates during the autumn/winter periods for 12 hourly accumulated rainfall over the forecast range for the predictions based on the five parameterization schemes of the PE ensemble design.

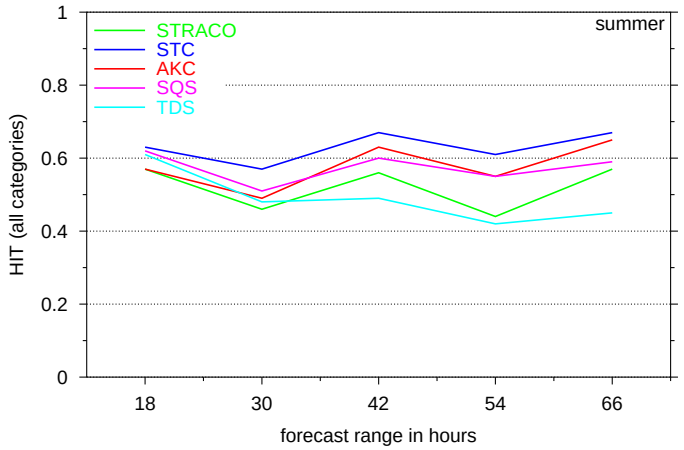

FIG. 3: Integrated hit rates during the summer period for 12 hourly accumulated rainfall over the forecast range for the predictions based on the five parameterization schemes of the PE ensemble design.

The evaluation has been splitted into the periods autumn/winter and a summer period (Figs. 2 and 3). The integrated hit rates for the different parameterization schemes show that the schemes perform with a similar hit rate during the autumn and winter periods. A slight decrease of the hit rates with lead time can be observed. As can be expected for the summer period the integrated hit rates diverge to a larger extent due to the diminished predictability of rainfall connected to convective activity in the summertime, especially occurring during the afternoon, i. e. around the

forecasts ranges 30 and 54 hours of the HIRLAM simulations.

It seems permissible from the figures given above to regard the parameterization schemes as performing comparably well especially in the autumn/winter cases. As a first approximation the parameterization ensemble members have been assigned equal weights in all cases.

3 The NWP model

3.1 General setup and model domain

The HIgh Resolution Limited-Area Model (HIRLAM) is a numerical weather prediction (NWP) system for analysis and simulation of mesoscale dynamical and physical processes in the atmosphere. It has been developed within the international HIRLAM cooperation (Undén et al., 2002) and describes the atmospheric equations of motion, energy and mass conservation in a discrete and approximated way (Källén, 1996). The DMI-HIRLAM has been in operational use at DMI for years. HIRLAM is a hydrostatic grid-point model designed for a high spatial resolution. The atmospheric equations of motion, continuity of mass and conservation for energy and humidity are discretized on an ARAKAWA-C grid with hybrid vertical levels, and Eulerian as well as semi-Lagrangian treatment of advection is possible. HIRLAM includes several parameterization models for the treatment of sub-grid physical processes. Physical parameterizations include several turbulence schemes, different schemes for convection and condensation, a radiation scheme based on Savijärvi (1990) and surface flux parameterization using a drag formulation. Time integration is performed using a semi-implicit scheme, and a split time stepping between the physical parameterizations and the model dynamics is possible in DMI-HIRLAM. This allows for larger time steps in the sub-grid parameterizations with respect to the time step for the model resolved dynamics. The lateral boundaries are treated by a relaxation zone using a scheme, which goes back to Davis (1976). At the upper boundary a rigid top at 10hPa is imposed together with a deep sponge layer below. The DMI-HIRLAM model system is described in more detail by Sass et al. (2002).

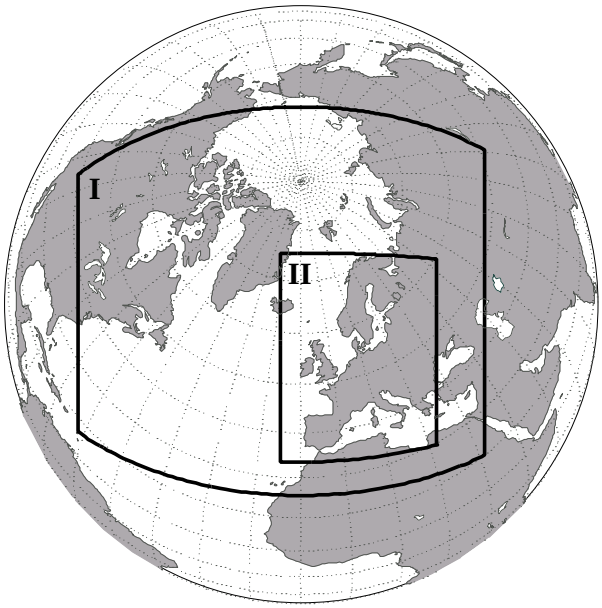

FIG. 4: Regions covered by the two DMI-HIRLAM nested models. The outer model domain (I) is covered by a 0.3 °-grid, while the inner domain (II) is covered by a 0.1 °-grid. The inner model covers all important river catchments of Europe.

The basic configuration of the model for the present study is similar to that used operationally at

DMI, where the choice of the standard physical parameterizations and the advection scheme are concerned. They comprise the STRACO scheme (Sec. 2.3.5), the turbulent kinetic energy (TKE) turbulence scheme of Cuxart et al. (2000), the radiation scheme of Savijärvi (1990), and the simple surface scheme as described in Sass et al. (2002). These components were utilized in the control forecasts and for the members of the SE (Sec. 2.1). The members of the PE were created by replacing the STRACO scheme by one of the parameterization schemes for condensation and convection as described in Sec. 2.3.

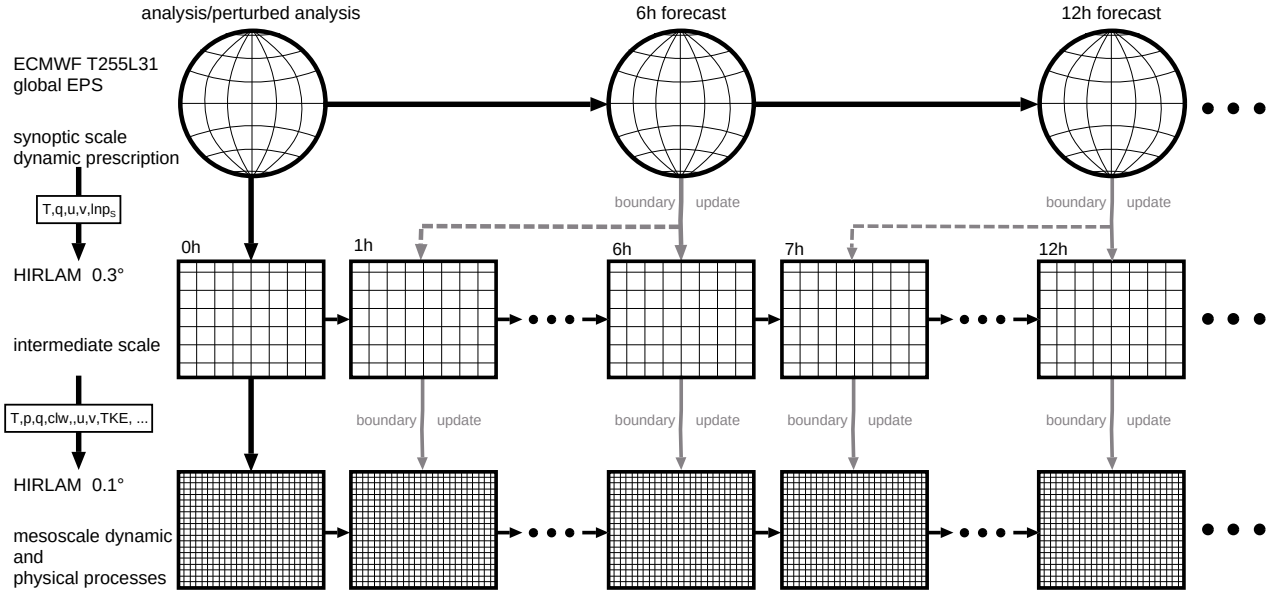

FIG. 5: Sketch of the NWP model chain used for the DMI-HIRLAM ensemble forecasts. The large scale host model is the ECMWF global EPS model T255L31, where boundary data are available to the intermediate model with a 0.3° grid every 6 hours. The high-resolution 0.1°-grid point model is updated every hour at its boundaries by the intermediate model.

The lower boundary model topography is derived from the physiographic data base of the HIRLAM model.

In order to cover most of the European river basins and under consideration of available computational resources, a doubly 1-way nested model setup has been established, which is based on the global model T255L31 of the ECMWF-EPS. The respective domains covered by the two DMI-HIRLAM models are shown in Fig. 4. The grid representation is in rotated latitude-longitude coordinates. A relaxation zone of 15 grid points in width surrounds the domains at the lateral boundaries.

Fig. 5 outlines schematically the model integration process. The first DMI-HIRLAM model is established with 0.3° horizontal grid spacing and uses 31 vertical levels. It is initialized at 12 UTC by the T255L31 analysis data or its perturbation, respectively. The coupling to the host model happens via the vertical model levels in order to preserve as much information about the vertical atmospheric structure as possible. The lateral boundaries are updated by the respective control or perturbed forecast with a 6 hour frequency. An interpolation in time is applied between the boundary data updates in order to assure smooth transitions. The model utilizes semi-Lagrangian

advection. This permits a time step of 450s in the dynamic part of the model, and for the physical parameterizations, 900s have been chosen. The communicated variables from the host model include temperature, specific humidity and wind on the model levels as well as surface pressure (Fig. 5).

The second DMI-HIRLAM model is established on a grid with 0.1° horizontal grid spacing and also with 31 vertical levels, the latter having the same structure as before. It is initialized by the 0.3° DMI-HIRLAM model, and updated hourly at its lateral boundaries by this model. It should be noted that the grid topologies of the two models only differ in horizontal grid spacing. Thus, the boundary fields are treated in a much more consistent way than between the global model and the first DMI-HIRLAM model, where the horizontal grids differ in projection type and also vertical interpolation is involved. The 0.1° DMI-HIRLAM model utilizes the Eulerian advection scheme and runs with about 60s time step for the dynamic part of the model. The time step for the physical parameterizations has been chosen to 540s. The communicated variables between the first and the second DMI-HIRLAM model include all prognostic model variables.

3.2 Ensemble forecast configuration

The integrations over the periods of the three historical heavy rain events were performed on a daily basis. A forecast run was started at 12 UTC each day within the respective period of each event (see Sec. 4). Each forecast produced hourly fields of accumulated large scale and convective precipitation up to a lead time of 72 hours. This means that the forecasts for consecutive days overlap by two days, respectively.

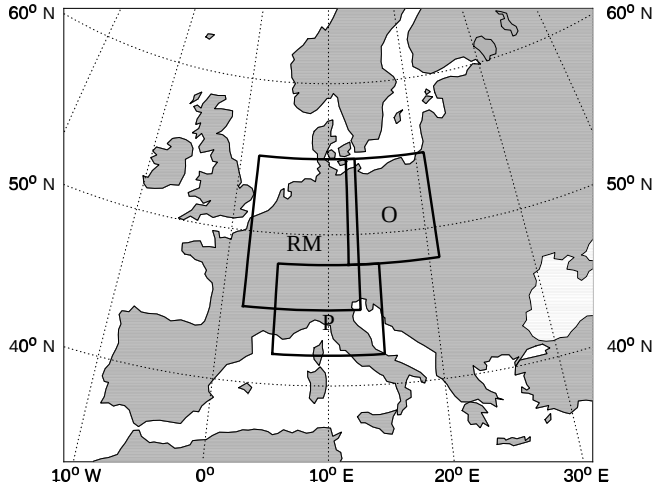

FIG. 6: Target areas of the ensemble member selections for the river basins of Odra (O), Rhine and Meuse (RM) and for the Po (P).

The selection procedure for the SE described in Sec. 2.1 was referred to a specific target area for each of the three events (Fig. 6). The respective area includes the river basin under consideration

including a surrounding frame of about one to two degrees in width. For the Piemonte flood event from November 1994, the area includes the Po river catchment. The area for the Rhine/Meuse flood from January 1995 includes the catchments of both rivers, and the Odra flood from July 1997 covers the complete basin of the Odra. The resolution of the selection area was approximately 0.5° in all three cases.

The members of the PE design (Sec. 2.3) all used the control member from the ECMWF-EPS as host model. As the STRACO scheme (Sec. 2.3.5) was utilized in the integrations of the SE members, the SE control is identical to the PE member using the STRACO scheme. It will be referred to as the control run in the sequel.

The two ensemble designs SE and PE may also be combined. This combined ensemble (CE) includes the control member only once and thus results in a size of 9+1, consisting of the unperturbed control forecast utilizing the STRACO scheme, the 5 perturbed forecasts from the SE, and the 4 forecasts form the PE integrated with the alternative parameterization schemes for condensation and convection.

4 The historical events of heavy rain

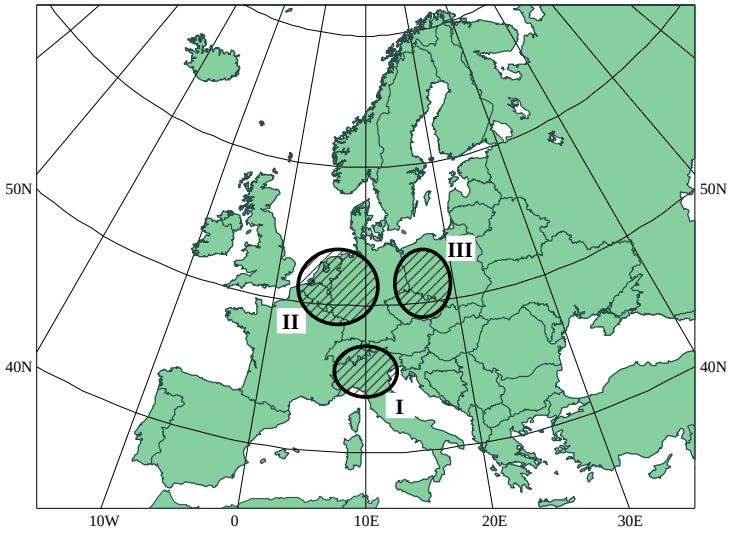

FIG. 7: Location of the three historical precipitation events: Piemonte flood in November 1994 (I), Rhine/Meuse flooding in January 1995 (II) and The Odra flood in July 1997 (III).

The periods on which the ensemble designs were tested comprise periods from within three historical events that caused or contributed to severe river floodings. These events occurred in three different regions in Europe (see Fig. 7), and the meteorological conditions leading to the rainfall were different for the three cases. The cases were also covered by single hindcast simulations using DMI-HIRLAM in a preceding work of Sattler (2002).

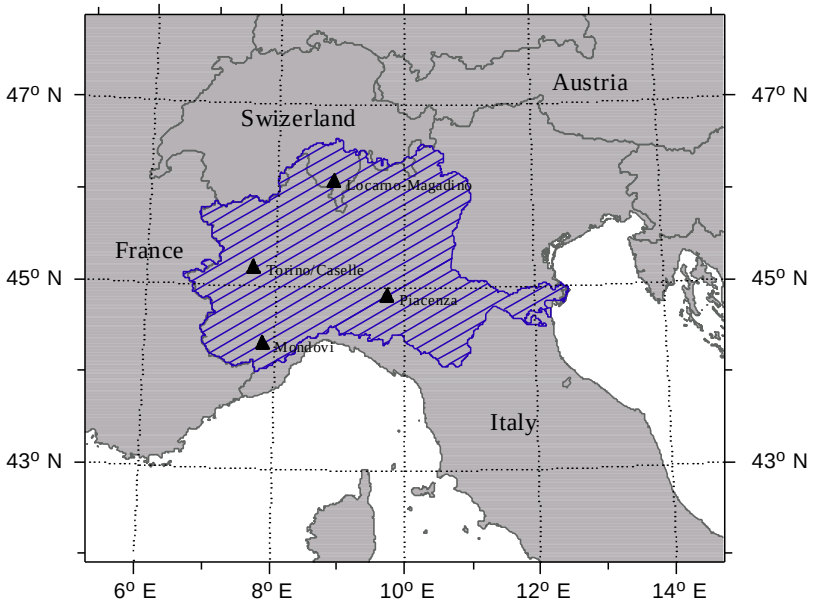

FIG. 8: Location of the four stations, which observed the largest rainfall amounts over the period shown in Fig. 9. The river basin of the Po river is depicted by the hatched area.

The first event is the Piemonte flood that occurred in November 1994. The flooding over the Po river basin was caused mainly by one major rainfall event between the 5^{th} and 6^{th} of November 1994.

It was connected to a small depression developing over the Balearic Islands from a baroclinically unstable trough positioned over the Western Europe. The path of the depression crossed the Alps from south, resulting in an enhancement of precipitation to extreme values by orographic forcing south of the alpine chain. This caused a severe flood over the western part of the Po river basin, the Piemonte region (Buzzi et al., 1995).

The meteorological stations depicted in Fig. 8 were among those registering the largest precipitation amounts. The stations are located in the western part of the Po river basin, the Piemonte region. The precipitation observations from these stations are shown in Fig. 9. The event was neither preceded or succeeded by another heavy rain event, and it resulted locally in more than 250 mm observed precipitation. The period, for which model level data from the ECMWF-EPS were available covers the days up to the event, i. e. from 25$^{\text{th}}$ of October to 3$^{\text{rd}}$ of November 1994. The DMI-HIRLAM ensembles were performed for each day of this period as described in Sec. 3.2.

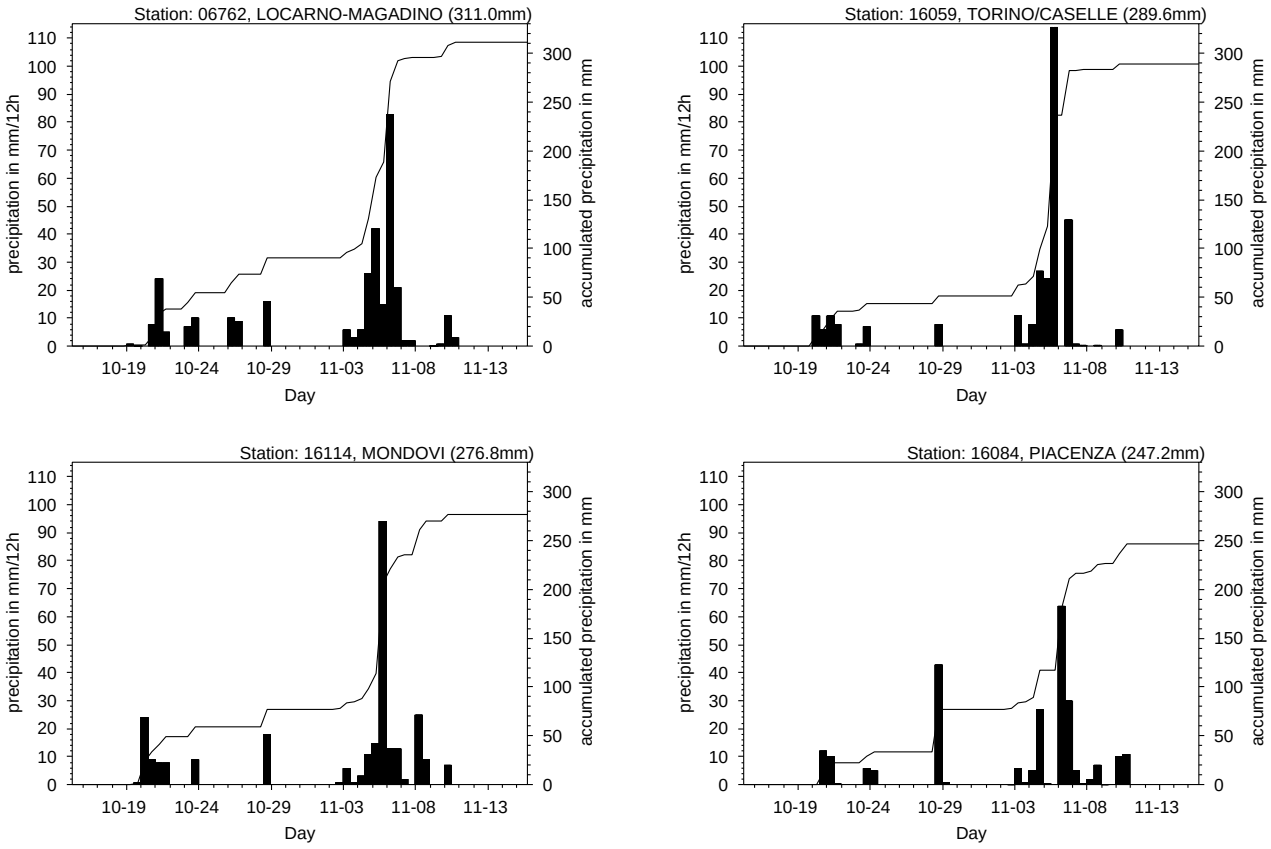

FIG. 9: Observed 12 hourly accumulated precipitation (columns, left ordinate) and pluviographs (curves, right ordinate) of the four synoptic meteorological stations Locarno-Magadino (upper left), Torino/Caselle (upper right), Mondovi (lower left) and Piacenza (lower right). The depicted period comprises the days between 15$^{\text{th}}$ of October to 15$^{\text{th}}$ of November 1994.

The second flooding case occurred in January 1995 and affected the rivers Rhine and Meuse. Contrary to the case in the Piemonte area, the river floods were, besides hydrological causes, linked to a series of precipitation events with a larger spatio-temporal extent than in the Piemonte case (Meijgaard and Jilderda, 1995; Fink et al., 1995). The events occurred over several weeks, each

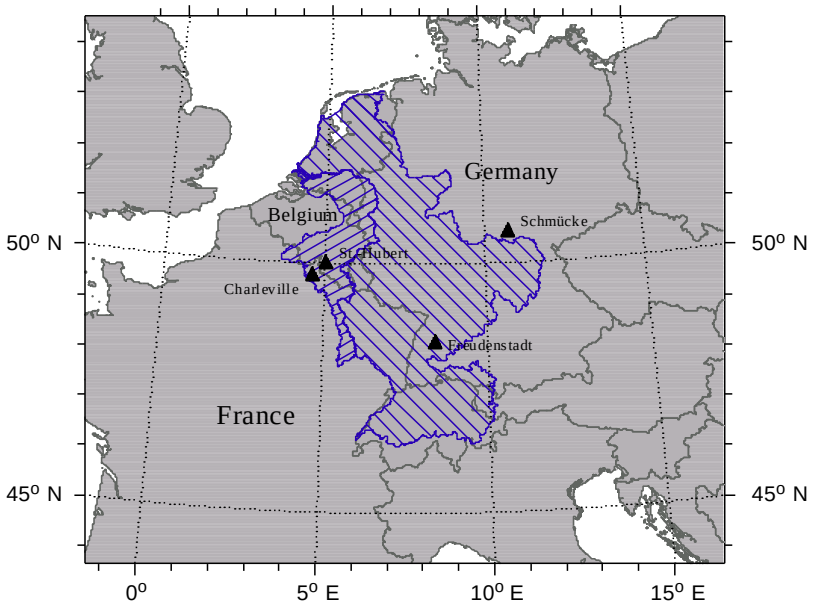

FIG. 10: Location of the four stations, which observed the largest rainfall amounts over the period shown in Fig. 11. The river basin of the rivers Rhine and Meuse are depicted by the hatched areas.

event being of rather medium than of extreme strength. Some of the meteorological observations over the areas of the two river basins (Fig. 10), which observed the largest rainfall amounts over a period of more than five weeks are shown in Fig. 11. As can be seen from Fig. 10, the stations are distributed over a significantly larger area than the most significant observations of the Piemonte case, while at the same time, there are temporal correlations in the observed events, which becomes clear when regarding the columns in Fig. 11. The observations clearly show that several rainfall events of significant but moderate strength occurred quite simultaneously within the shown period ranging from 22$^{\text{nd}}$ of December 1994 to 15$^{\text{th}}$ of February 1995. They all contributed to the floodings of river Meuse and river Rhine at the end of January 1995.

The rainfall events were due to a chain of troughs and frontal systems passing Europe from West, still under development, and persistently transporting humidity from the Atlantic eastwards. Periods with predominance of a north-westerly to westerly flow regime were interrupted by intermediate ridges. Furthermore, hydrological causes like snow melt also played a significant role for the flood to occur (Ulbrich and Fink, 1995).

The period covered by the ensembles is from 20$^{\text{th}}$ to 30$^{\text{th}}$ of January 1995, for which model level data from the ECMWF-EPS was available. This period was dominated by a westerly flow regime and includes three rainfall events with significant precipitation amounts (Fig. 11).

The third rainfall event regarded in this work happened in the beginning of July 1997, when a quasi-stationary depression over Central/Eastern Europe was predominant. The dynamically induced ascent by this depression released significant amounts of precipitation over the area of Southern Poland, South-Eastern Germany and the Czech Republic, lasting several days. In addition to the large scale rainfall dynamically triggered convective episodes occurred (Keil et al., 1999). Even though the orography was not the major rainfall trigger, it played an important role in the rainfall distribution by intensifying precipitation over the mountainous areas.

The first heavy rains occurred between 5$^{\text{th}}$ of July and 10$^{\text{th}}$ of July 1997 and they affected the rivers Odra and Vistula. They initiated the floodings of these rivers during July 1997. The largest rainfall amounts were observed at the meteorological station Lysa Hora, which is elevated at 1327 m MSL and which is located at the SE corner of the Odra river catchment (Fig. 12). It registered about 600 mm precipitation in the mentioned period (Fig. 13, lower left panel). Praded, which is located about 80 km NW of Lysa Hora and also located on a mountain (1512 m MSL), registered almost 500 mm during this period (Fig. 12 and Fig. 13 upper left panel). The other two observations shown in Fig. 13, Cervena and Jelenia Gora, are located in hilly terrain at 753 m and 344 m, respectively. They observed less amounts of precipitation than Lysa Hora and Praded, but still significant amounts. This indicates the non-negligible role the orography still had for the rainfall distribution in this case. A second significant rainfall event occurred after the 15$^{\text{th}}$ of July. It is, however not regarded within this work.

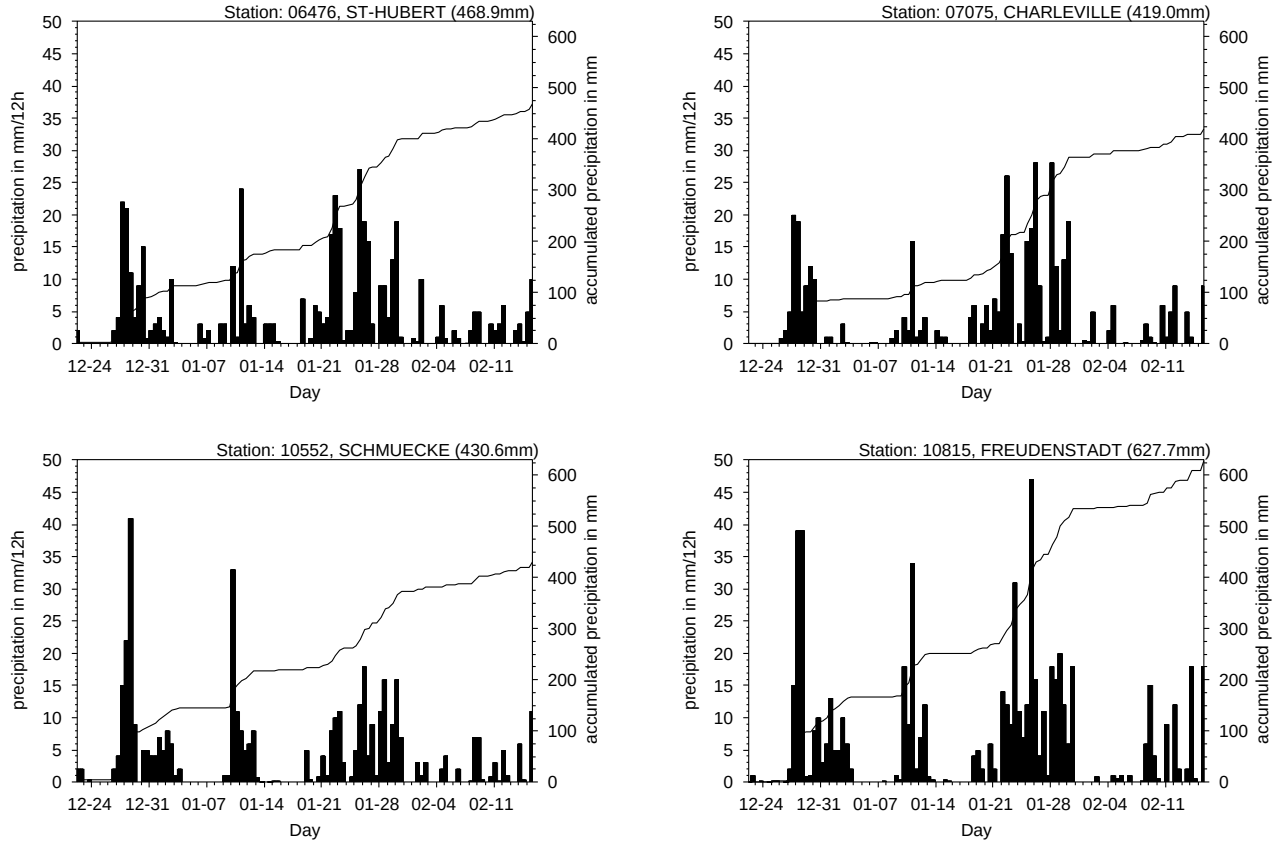

FIG. 11: Observed 12 hourly accumulated precipitation (columns, left ordinate) and pluviographs (curves, right ordinate) of four synoptic meteorological stations. St. Hubert (upper left) and Charleville (upper right) are located within the Meuse river catchment, Schmücke (lower left) and Freudenstadt (lower right) are located in or close to the river Rhine catchment. The depicted period comprises the days between 22$^{\text{nd}}$ of December 1994 to 15$^{\text{th}}$ of February 1995.

The period, for which model level data from the ECMWF-EPS were available, covers the days between 26$^{\text{th}}$ of June and 8$^{\text{th}}$ of July 1997. The DMI-HIRLAM ensemble integrations were performed as outlined in Sec. 3.2.

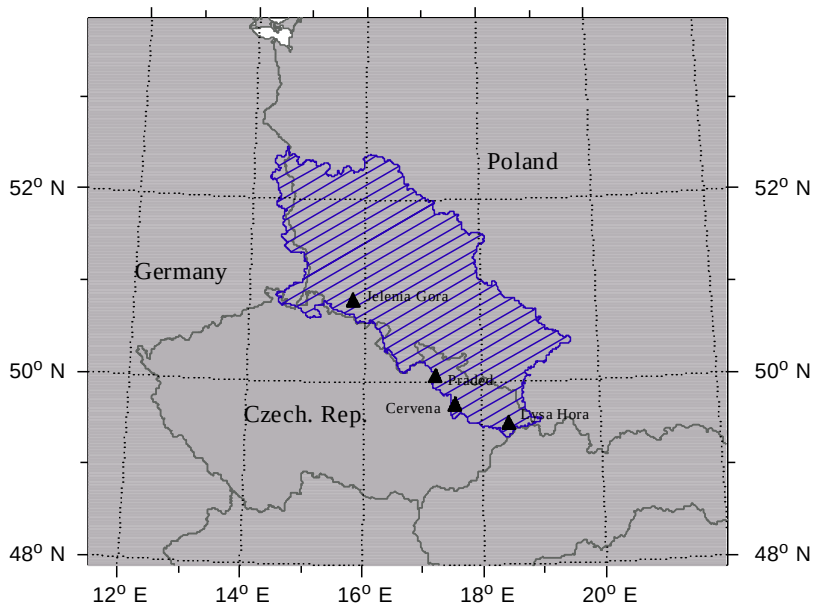

FIG. 12: Location of the four stations, which observed the largest rainfall amounts over the period shown in Fig. 13. The river basin of the Odra river is depicted by the hatched area.

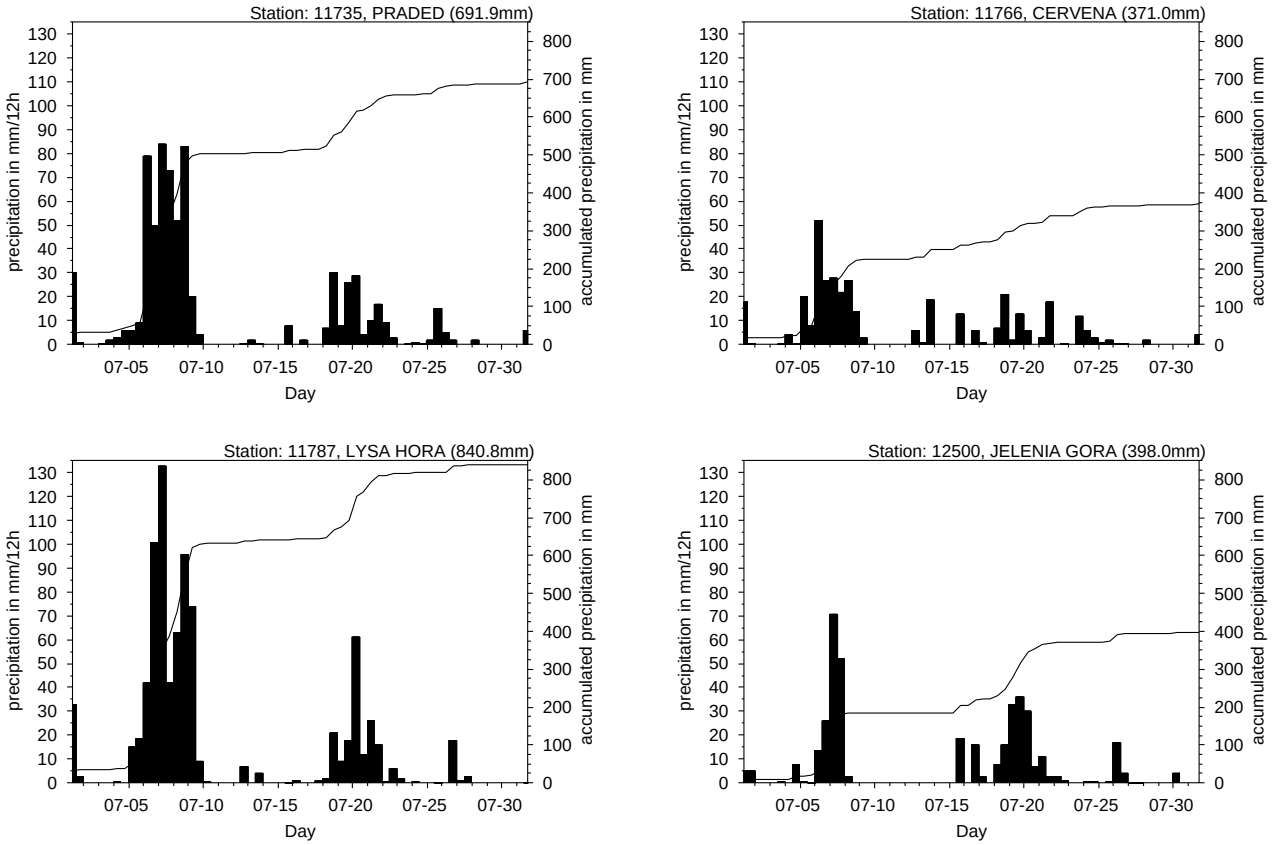

FIG. 13: Observed 12 hourly accumulated precipitation (columns, left ordinate) and pluviographs (curves, right ordinate) of the four synoptic meteorological stations Praded (upper left), Cervena (upper right), Lysa Hora (lower left) and Jelenia Gora (lower right), all located in the Odra river catchment. The depicted period comprises 1st to 31st of July 1997.

5 Verification

The DMI-HIRLAM ensembles are evaluated in their probabilistic performance by common methodologies. This is supplemented by case studies regarding precipitated water over the considered river basin. The verification focuses on precipitation, and a comparison between the ensemble designs is included.

5.1 Methodology

A good probability forecast based on an ensemble of deterministic forecasts should be both reliable and skillful. That is, the statistical distribution of the ensemble should match that of the verifying observations, and the errors of the forecast should be smaller than those of a reference forecast (typically based on climatology).

Several verification measures exist that can be used to evaluate the quality of probability forecasts (Stanski et al., 1989; Wilks, 1995). In the following the DMI-HIRLAM ensembles are verified using rank histograms (a.k.a. Talagrand histograms), reliability diagrams and Brier skill scores and relative operating characteristic (ROC).

A rank histogram is a simple histogram that shows where in the ranked ensemble the verifying observation lies (Strauss and Lanzinger, 1996; Hamill and Colucci, 1997). For a ranked n-member ensemble the verifying observation will fall in one of n+1 intervals defined by the ensemble members, including the two extreme intervals, i. e. less than the smallest ensemble member and larger than the largest ensemble member, respectively. If the ensemble distribution is the same as the distribution of the verifying observations then the rank histogram will be flat. Note that for a flat rank histogram only 100(n-1)/(n+1) percent of the observations will be encompassed by the ensemble. Note also that a flat rank histogram does not imply that the ensemble distribution is the same as the distribution of the verifying observations (but it is a necessary condition).

Unlike the rank histogram the reliability diagram is associated with a forecast of a particular event, e.g. 24h accumulated rainfall exceeding 1 mm. The reliability diagram is a plot of the observed frequency of that event versus the forecast probability (Wilks, 1995). For the perfectly reliable forecast the observed frequency equals the forecast probability. For example, for forecasts that predict a 10% probability for 24h accumulated rainfall exceeding 1 mm we expect to observe more than 1 mm in 10% of the cases.

The Brier skill score (BSS) is also associated with a forecast of a particular event. It is based on the Brier score, which is the mean squared error in probability space where occurrence of the event is associated with an observation probability of 1, and non-occurrence is associated with an observation probability of 0 (Wilks, 1995). Hence, a perfect deterministic forecast has a Brier score of 0. The Brier skill score is defined as the relative improvement over a reference forecast, i.e.

$$\mathrm{BSS} = 1 - \frac{\mathrm{BS}}{\mathrm{BS}_{\mathrm{ref}}} \tag{4}$$

where BS is the Brier score and BS_{ref} is the Brier score of the reference forecast, typically based on the probability forecast one would get using only climatological information.

The maximum BSS is 1, and it is obtained for a perfect deterministic forecast. A BSS of 0 is obtained for a forecast that is no better than the reference forecast, and a negative (unbounded) Brier skill score is obtained for a forecast that is worse than the reference forecast.

In this study the BSS is referred to the sample climatology of the verifying DWD rainfall analyses. It should be kept in mind that this is not a very confident sample, and the resulting Brier skill scores should to be regarded with caution.

The relative operating characteristic (ROC) is a curve that shows hit rates versus false alarm rates for forecasts that are given as a probability of occurrence of a predefined event (Stanski et al., 1989). A forecast counts as a hit if the probability exceeds a certain threshold and a false alarm otherwise. The ROC curve is obtained by counting hits and false alarms for thresholds varying from 0 (yielding a hit rate of 1) to 1 (yielding a false alarm rate of 0). The perfect forecast has a hit rate of 1 and a false alarm rate of 0, but as long as the hit rate is greater than the false alarm rate the forecast is regarded as skillful. A forecast skill score that is often derived from the ROC curve, is the area under the curve. Skillful forecasts have a ROC area between 0.5 and 1, while unskillful forecasts have a ROC area between 0 and 0.5.

The ROC curve probably provides a better measure of skill than Brier skill scores for forecasts that will be used in flood forecasting, as decisions will be made based on forecasts for which the probability for a particular event exceeds a certain threshold. Hit rates and false alarm rates for deterministic forecasts can also be plotted in the same diagram as ROC curves, allowing a simple comparison between deterministic and probabilistic skill.

The rank histograms, reliability diagrams, Brier skill scores and relative operating characteristic that are shown in Sec. 5.2, are based on all grid boxes in a specified region for the available time range in order to get statistically more robust results. Thus, variations in forecast skill and reliability over time and/or space do not show up.

The verifications with the measures mentioned above are performed against data from precipitation analysis of the DWD. It is based on the grid of the German "Lokal-Modell" LM (Doms and Schättler, 1999), and utilizes rain gauge observations from synoptical as well as climatological stations, the latter of which make up a significantly denser network of rainfall observations than the synoptic stations alone do. In the case of the Piemonte flooding of November 1994, DWD included additional rainfall data from the Mesoscale Alpine Project MAP (MAP, 2002) in the analyses. The verification area comprises the region 41 °N–58 °N latitude and 5 °W–19 °E longitude. This is nearly the area covered by the LM-grid.

A slight complication arises in the verification as the forecasts and the verifying DWD analyses exist on different grids. The forecast grid has a horizontal resolution of 0.1 °×0.1 ° and the southern pole of rotation at (0 °, 80 °E), while the analysis grid has a horizontal resolution of 0.0625 °×0.0625 ° and the southern pole of rotation at (32.5 °S, 10 °E). In order not to introduce interpolation errors the rainfall forecasts are aggregated in 3×3 neighbouring grid boxes and the rainfall analyses in 5×5 corresponding grid boxes, and the verification is made using the aggregated values, which for

forecasts as well as analyses cover approximately 30×30 km^2.

The periods, where precipitation analysis data were available coincide with the period, where ECMWF-EPS model level data was available in the two flooding cases of the Rhine/Meuse (January 1995) and the Odra flood (July 1997). In the case of the Piemonte flood, the periods do not overlap exactly. Thus, the verification had to be restricted to a shorter period in this case.

The synoptic rainfall observations refer to 6 UTC and 18 UTC, measuring the accumulated precipitation of the last 12 hours, and climatological stations measure the accumulation over 24 hours, reporting once a day. As a consequence, the precipitation analyses were performed such that they contain the 24 hour accumulation between two consecutive days from 6 UTC to 6 UTC, respectively. The ensemble forecasts, however, start at 12 UTC. The verified forecast periods have therefore been chosen between 18–42 hours and between 42–66 hours in order to be time consistent to the analysis.

5.2 Results

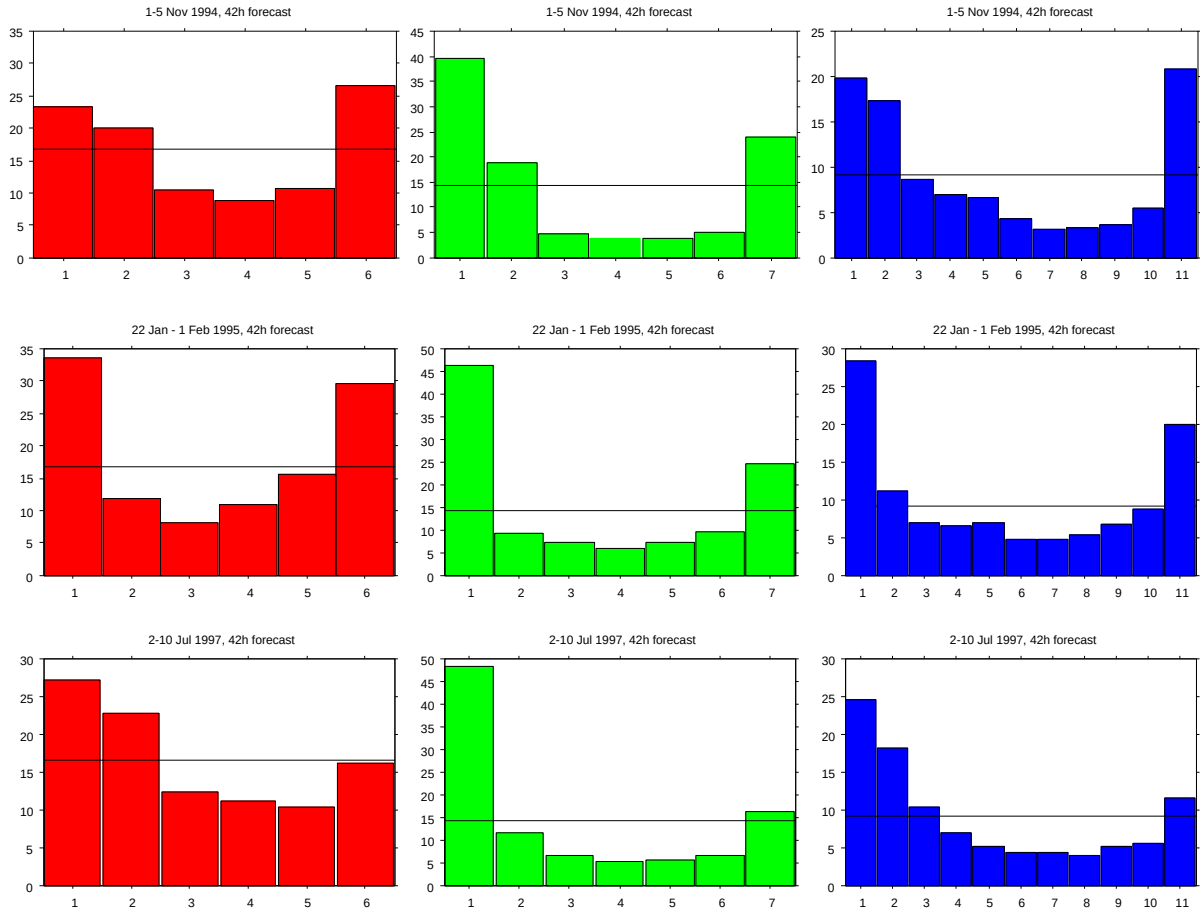

FIG. 14: Ranked histograms for the ensemble designs PE (red), SE (green) and CE (blue) for the three heavy rain events Piemonte 1994 (upper panels), Rhine/Meuse 1995 (panels in the middle) and for the Odra 1997 case (lower panels). The histograms refer to the 18–42h forecast range.

The ranked histograms for the ensemble designs are given in Fig. 14 for the 18–42h forecast range, and Fig. 15 for the 42–66h forecast range. The black horizontal line in the diagrams depicts the flat

distribution, which an ideal ensemble is regarded to exhibit. The figures show that the spread of the ensembles generally does not capture the observations completely. This behaviour is, however, not unusual (Hou et al., 1994; Bright and Mullen, 2002). All figures except the PE histograms for the Piemonte 1994 case consistently show a significant overcrowding of class 1, the class of the smallest rainfall amounts. This is an indication for the LAM to overpredict small amounts of precipitation, as well as to predict small rainfall amounts, where actually no rain was observed.

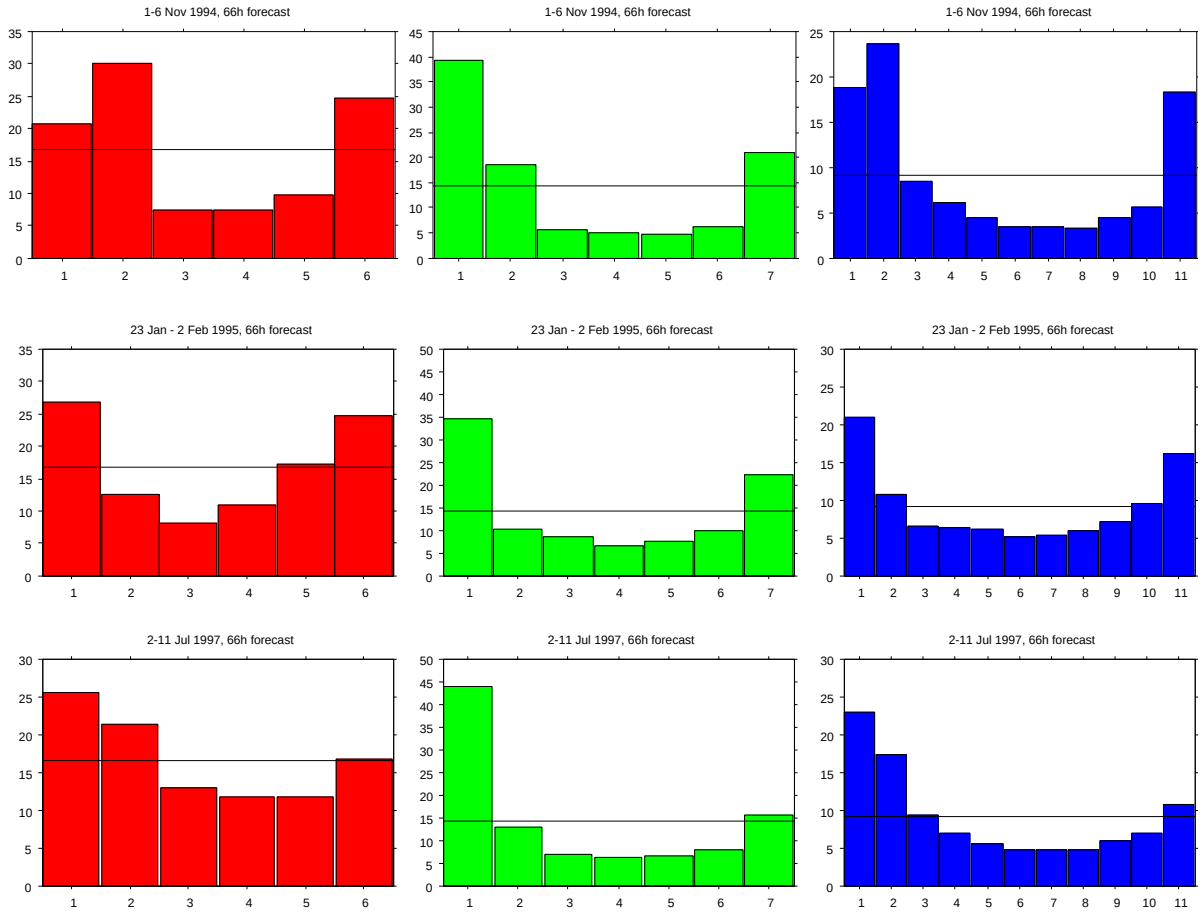

FIG. 15: Ranked histograms for the ensemble designs PE (red), SE (green) and CE (blue) for the three heavy rain events Piemonte 1994 (upper panels), Rhine/Meuse 1995 (panels in the middle) and for the Odra 1997 case (lower panels). The histograms refer to the 42–66h forecast range.

The PE tends to have a larger spread (flatter histogram) in the Piemonte 1994 case and to a lesser extent also in the Odra 1997 case than the SE (Fig. 14). This is due to the fact that the PE members do not exhibit the same tendency like the SE members to overpredict small rainfall or no rain in this event.

There is only little difference between the ranked histograms of the SE for the autumn and the winter historical periods, i. e. the Piemonte 1994 case and the Rhine/Meuse 1995 event (Figs. 14 and 15, green histograms). The SE spread of the Piemonte 1994 ensemble is smaller, which can be recognized by the smaller frequencies in the middle classes 3 to 6. The CE, the combination of the SE and PE, reflects the distribution characteristics of both ensemble designs and shows the best distribution for the Rhine/Meuse 1995 case (Figs. 14 and 15, panels in the middle to the right).

For the period of the Odra 1997 event the frequency of observations in the largest class (Numbers 6, 7 and 11 for the PE, SE and CE respectively) all the ensembles show values close to the frequency of the flat distribution. This indicates that over the short 10 day period of the Odra 1997 event, the frequency of underpredictions of the large rainfall events is not significantly larger than could be expected theoretically from a discrete ensemble.

The reliability diagrams for the three periods are shown in Fig. 16 for the forecast range 18–42h. They refer to the thresholds of exceeding 5 mm and 25 mm of precipitation within 24 hours. At the 5 mm threshold the low frequencies tend to be underestimated by all ensembles in the Piemonte 1994 case and the Rhine/Meuse 1995 case, and the high frequencies tend to be overestimated in these cases (Fig. 16, upper left and middle panel). During the period of the Odra 1997 case, The ensembles overestimate nearly all frequencies (Fig. 16, upper right panel). The same is the case for the longer forecast range 42–66h (Fig. 17, upper panels). The ensemble designs show similar reliability at the 5 mm threshold, the PE design often being slightly better. The combination of the SE and the PE to CE is advantageous for the reliability of the ensemble at the 5 mm threshold in general (blue curves in Figs. 16 and 17, upper panels).

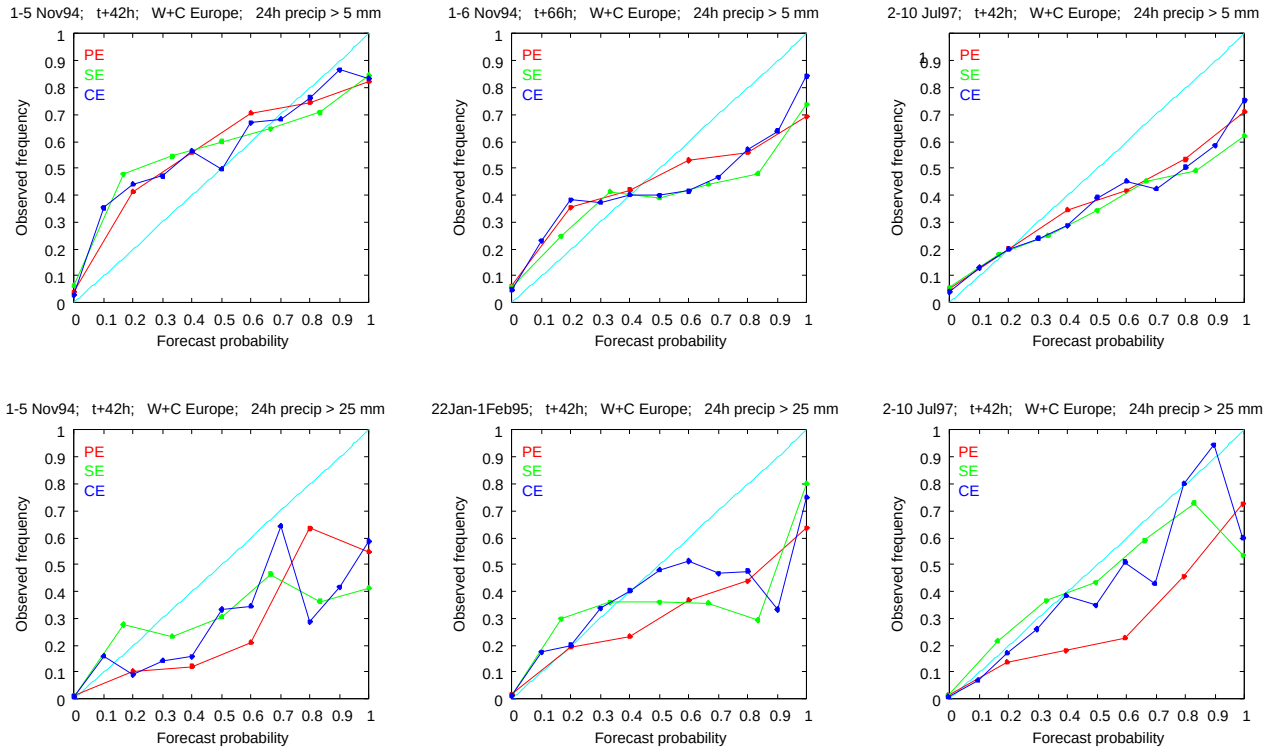

FIG. 16: Reliability diagrams for the ensemble designs PE (red), SE (green) and CE (blue) for the 5 mm threshold (upper panels) and for the 25 mm threshold (lower panels) of 24 hourly rainfall from the forecast range 18-42h. Depicted are the diagrams for the periods from the Piemonte 1994 event (left panels), the Rhine/Meuse 1995 case (panels in the middle) and the Odra 1997 case (right panels).

The reliability of the ensembles to predict the observed frequencies is decreased at the 25 mm threshold (Figs. 16 and 17, lower panels). The differences between the ensemble designs is more pronounced at this threshold, and there are cases, where SE performs very well compared to the PE, e. g. in

the Piemonte 1994 case in the 42-66h range (Fig. 17, lower left panel) and in the Odra 1997 case in the 18-42h range (Fig. 16, lower right panel). There are also cases, where PE is the more reliable ensemble, as in the Odra 1997 case in the 42-66h range (Fig. 17, lower right panel). The CE seems to be advantageous in all these cases also at the 25 mm threshold.

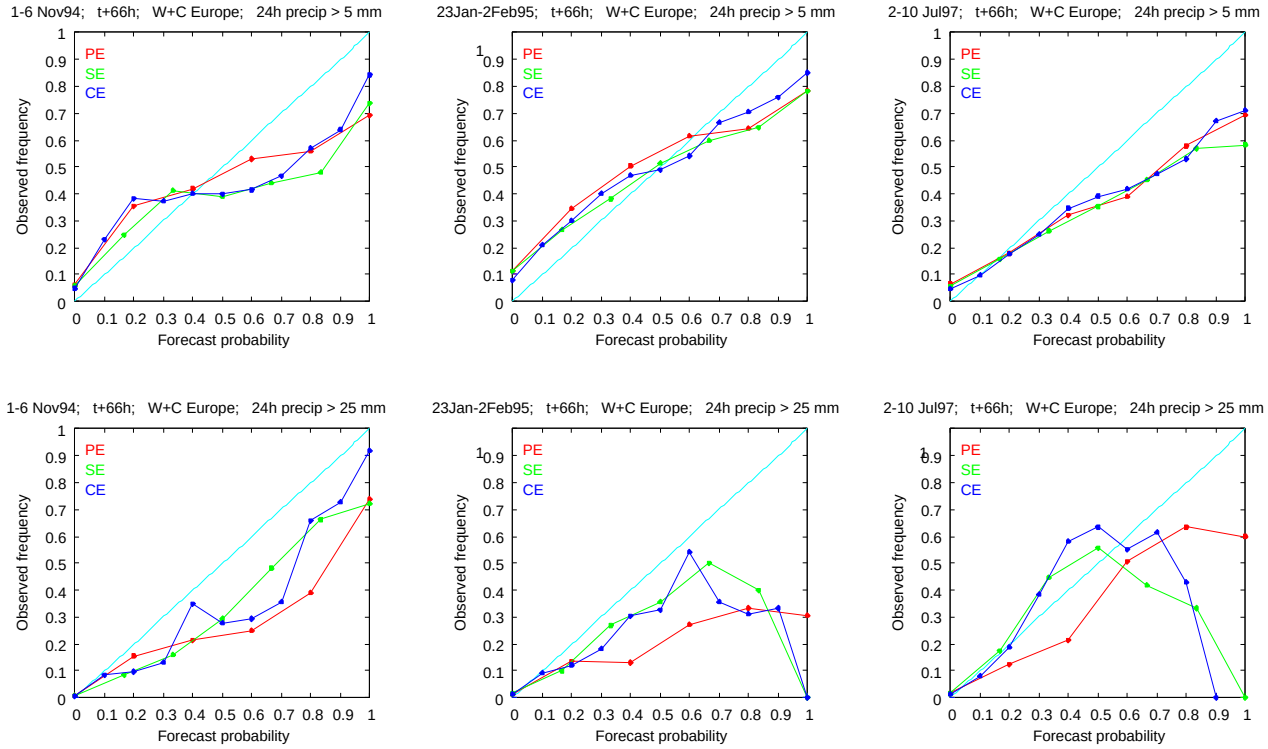

FIG. 17: Reliability diagrams for the ensemble designs PE (red), SE (green) and CE (blue) for the 5 mm threshold (upper panels) and for the 25 mm threshold (lower panels) of 24 hourly rainfall from the forecast range 42-66h. Depicted are the diagrams for the periods from the Piemonte 1994 event (left panels), the Rhine/Meuse 1995 case (panels in the middle) and the Odra 1997 case (right panels).

In the Rhine/Meuse 1995 and the Odra 1997 case, the reliability curves for the 25 mm threshold for the SE and, as a consequence for the CE too, "drop down" at the high forecast probabilities. This is an indicator for the circumstance that when all SE members predict rainfall beyond 25 mm, then no such event was observed, i. e. the SE exhibits a complete false alarm. There can be different reasons for this. The model may overestimate the rainfall amount or the rain area may be misplaced. A third possibility could be errors in the timeliness of the event. All three possibilities could play a role at the same time, of course. In the two cases, choosing the PE only would give the best reliability, as can be seen from the figures. Note also that in the 42–66h range, forecast probabilities of 1 are very rare, and so the reliability diagrams are not very robust when this is the case. In the Rhine/Meuse 1995 case a SE and CE forecast probability of 1 for the 42–66h range was only encountered once, in which case the observed precipitation amount was 21.5 mm. In the Odra 1997 case the SE forecast probability was 1 in only four cases. The corresponding observed precipitation amounts were between 7.3 and 18.4 mm.

The BSS are shown in Fig. 18 for the rainfall thresholds 0.5, 1, 5 , 10 and 25 mm, separately for the

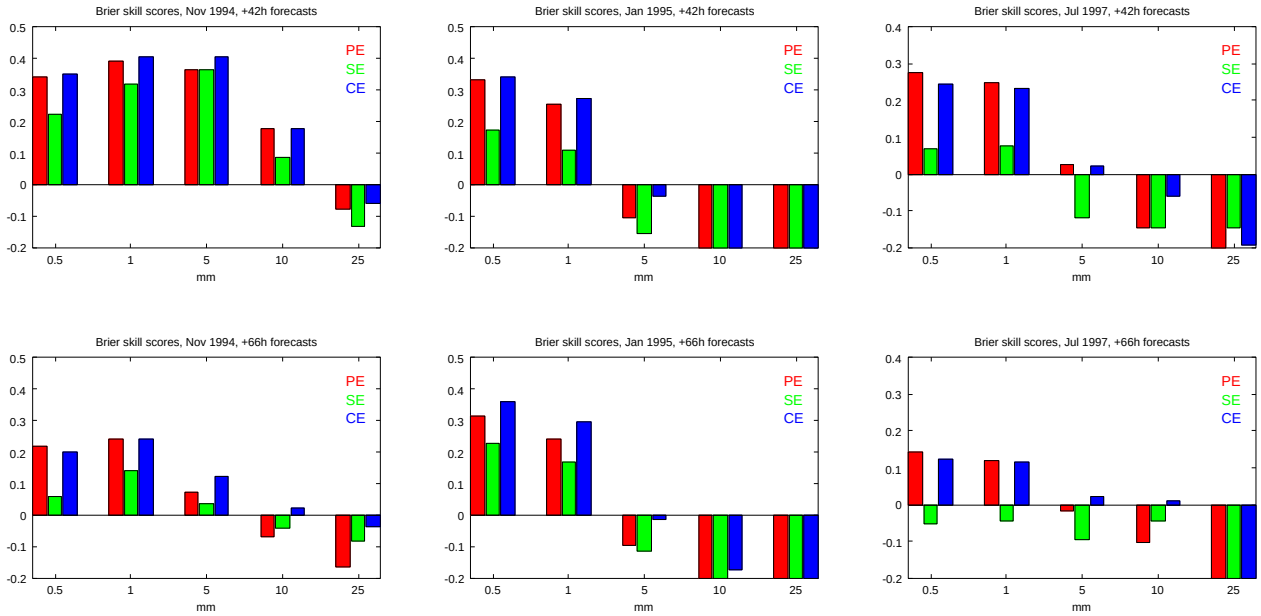

FIG. 18: Brier skill scores for the ensemble designs PE (red), SE (green) and CE (blue) over several thresholds of 24 hourly rainfall, for the 18-42h forecast range (upper panels) and for the 42-66h forecast range (lower panels). Depicted are the diagrams for the periods from the Piemonte 1994 event (left panels), the Rhine/Meuse 1995 case (panels in the middle) and the Odra 1997 case (right panels).

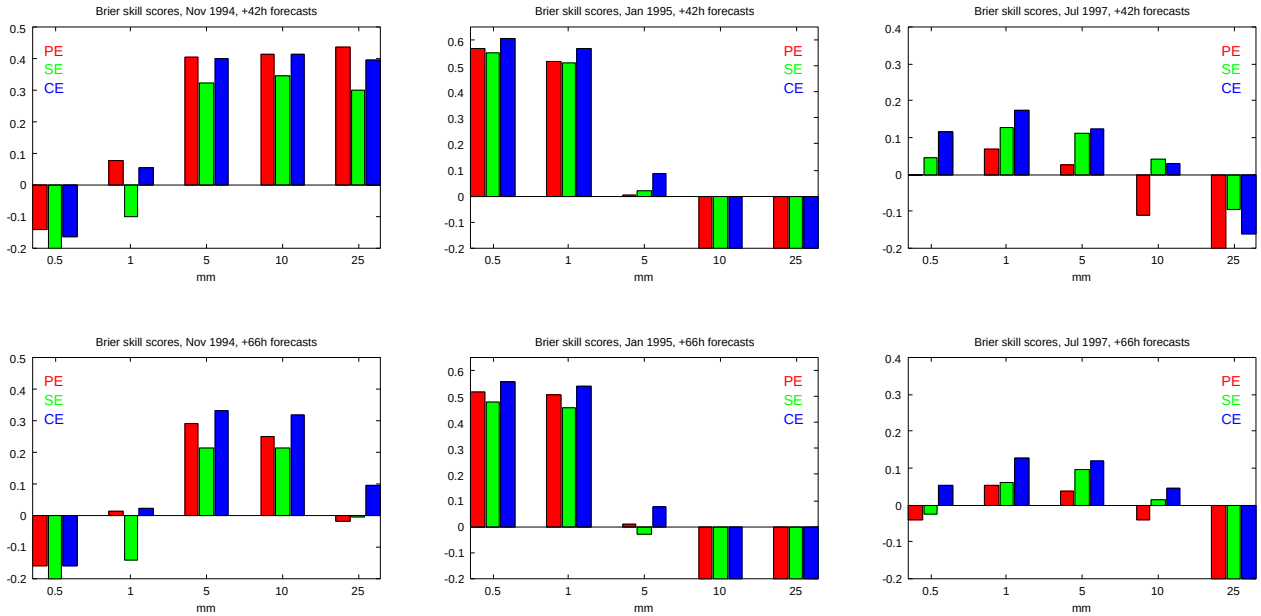

FIG. 19: Brier skill scores based on data only over the respective target areas used to create the SE, for the ensemble designs PE (red), SE (green) and CE (blue) over several thresholds of 24 hourly rainfall, for the 18-42h forecast range (upper panels) and for the 42-66h forecast range (lower panels). Depicted are the diagrams for the periods from the Piemonte 1994 event (left panels), the Rhine/Meuse 1995 case (panels in the middle) and the Odra 1997 case (right panels).

three historical heavy rain events and for the forecast ranges 18–42h and 42–66h. The decrease of the skill score with lead time is obvious in the Piemnonte 1994 case and the Odra 1997 case. This does not happen in the period of the Rhine/Meuse 1995 event. In the Piemnonte 1994 event the ensembles exhibit skill up to the 10 mm threshold at the 18-42h range and up to 5 mm in the longer forecast range. For the other two periods, the BSS becomes negative from the 5 mm threshold onwards. The authors do not believe that this necessarily indicates no skill in the ensembles, because in these two historical events, and especially in the Rhine/Meuse 1995 event, the statistical basis is rather small.

The CE turns out to be a good choice when regarding the BSS, except in the Odra 1997 case, where the SE seems to exhibit a poor skill. However, the comparison between the BSSs of the PE and the SE on basis of the scores from Fig. 18 is not really fair, because the BSSs are calculated over the Central European area (Sec. 5.1). The SE is however targeted for a certain river catchment, an area, which is significantly smaller than the verification area and which is different in the three historical periods that were studied (Sec. 3.2). Therefore, the BSSs were also determined on basis of the respective target area for each heavy rain period. They are shown in Fig. 19. It is illuminating to see how the BSS changes. In the Piemonte 1994 case, there seems to be significant skill even at the 25 mm threshold at 18–42h, the SE cannot outperform the PE, however. The extreme variation of the BSSs with the threshold indicates in this case, that these results should be regarded with much caution, because the statistical basis is even smaller now.

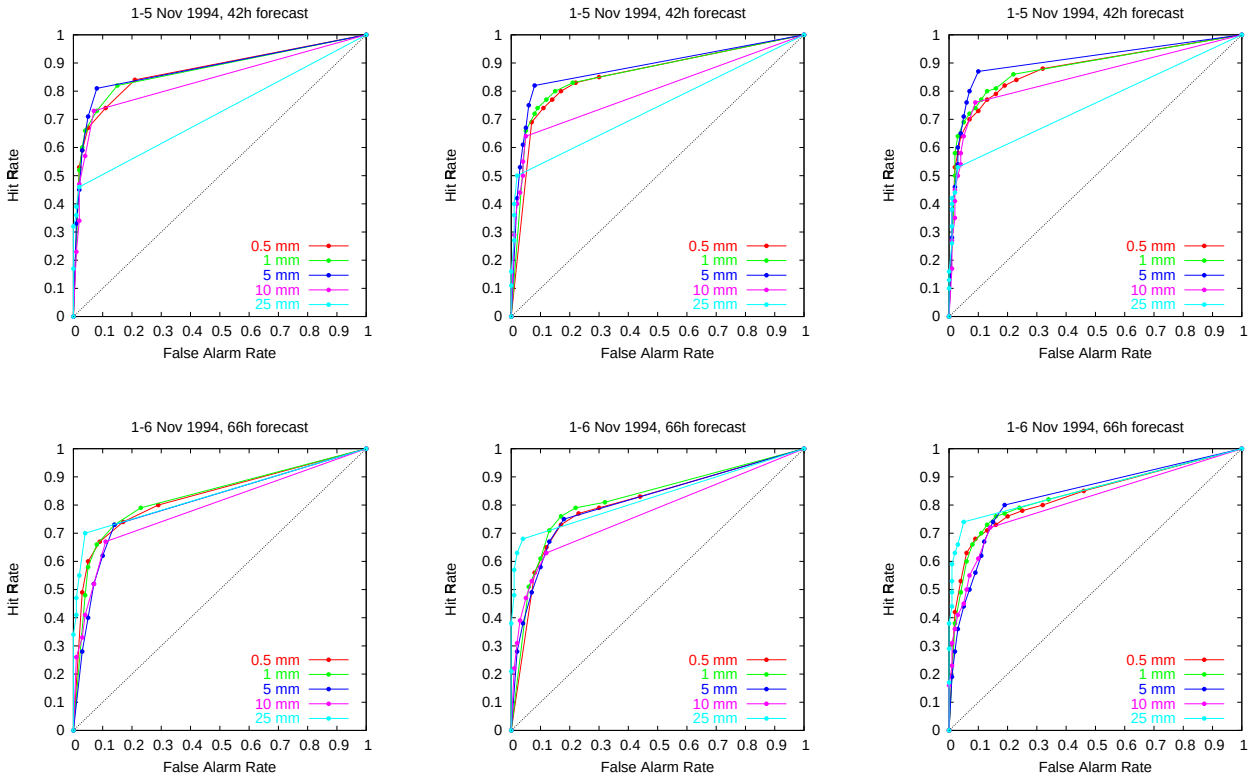

FIG. 20: ROC curves for the ensemble designs PE (left panels), SE (center panels) and CE (right panels) over several thresholds of 24 hourly rainfall within the period of the Piemonte 1994 event, for the 18-42h forecast range (upper panels) and for the 42-66h forecast range (lower panels).

In the Rhine/Meuse 1995 case the PE and SE show comparable BSSs in both forecast ranges, exhibiting high values, and the CE turns out to be the most skillful ensemble. The sharp drop in the BSS at the higher thresholds again is assumed to be due to the small statistical basis for these thresholds. Contrary to the Piemonte 1994 and the Rhine/Meuse 1995 case, where the PE still can be regarded as more skillful than the SE, the situation is vice versa in the Odra 1997 case (Fig. 19, right panels). Here the SE shows more skill than the PE over the target area of the Odra river basin. The BSSs for this event are, however, generally smaller than in the other to events. This indicates the difficulties the HIRLAM simulations had in the Odra 1997 event, which is regarded in more detail in Sec. 6.1.

It is worth noting that also when regarding the BSS over the respective target areas of the river basins the CE often outperforms both the SE and the PE, indicating that a combination of the SE and the PE is advantageous with respect to the ensemble forecast skill.

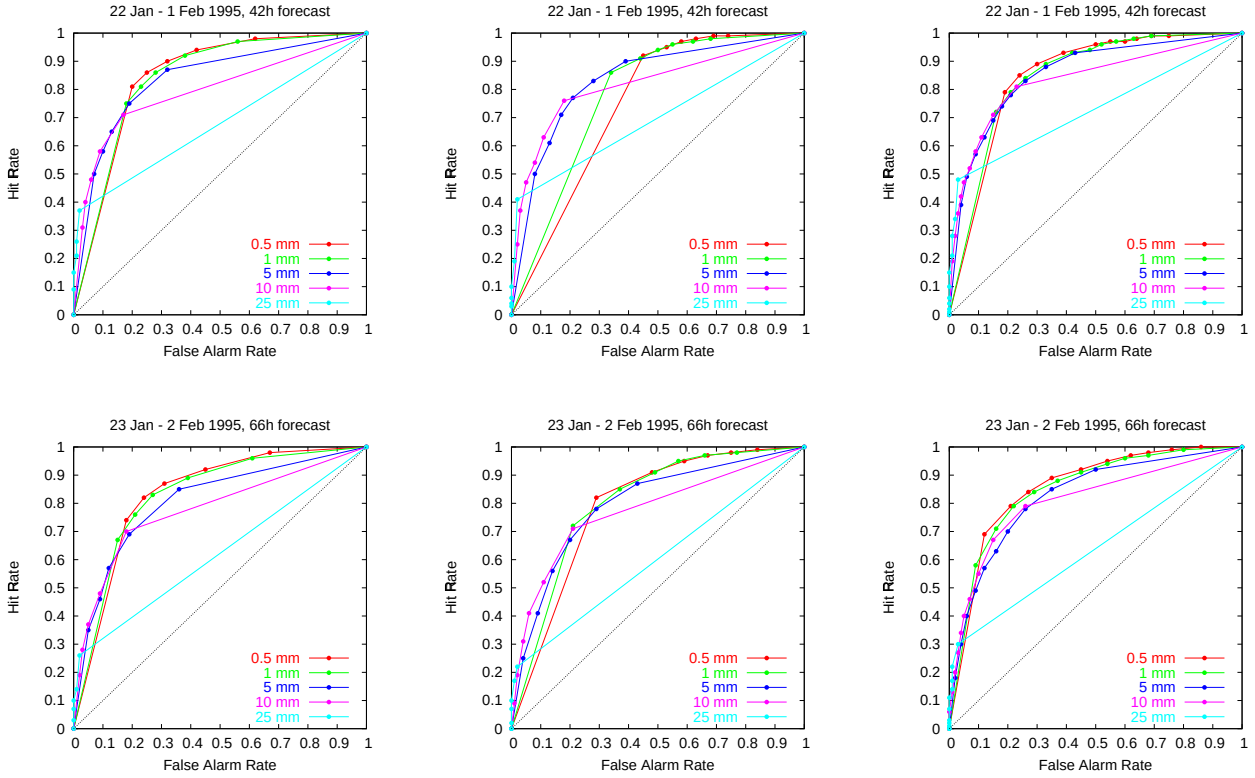

FIG. 21: ROC curves for the ensemble designs PE (left panels), SE (center panels) and CE (right panels) over several thresholds of 24 hourly rainfall within the period of the Rhine/Meuse 1995 event, for the 18-42h forecast range (upper panels) and for the 42-66h forecast range (lower panels).

The ROC for the Piemonte 1994 case is shown in Fig. 20, again for the thresholds 0.5, 1, 5 , 10 and 25 mm of rainfall within 24 hours for the two investigated forecast ranges. The curves reveal a very good portion of hit rates, which decreases with lead time (compare upper panels with the lower panels in Fig. 20). Also with lead time the false alarm rate increases, as can be expected. Again the PE shows slightly better hit rates than the SE in the 18-42h range, especially for the 10 mm threshold. At the 42-66h range PE and SE perform comparably well, and the CE reveals a slight

enhancement in the ROC curves.

The ROC curves for the Rhine/Meuse 1995 event are shown in Fig.21 and they reveal significantly more false alarms than in the first event, which can be easily seen by the curves being shifted towards the right. At the small thresholds the SE exhbits more false alarms and slighly more hits than the PE. Beyond the 10 mm threshold the hit rates reduce significantly, but also the false alarm rates, which still leaves the ensembles skillful. The CE seems to show again a slight improvement in the ROC curves with respect to the other two ensembles.

Fig.22 finally shows the ROC curves for the Odra 1997 event period. The curves for the small and medium thresholds are comparable with the ROC curves for the Rhine/Meuse 1995 event for all ensemble designs. The hit rates at the 25 mm threshold are, however, significantly better than in the former case.

A comparison of the ROC curves from the CE with the ROC curves from the SE and the PE again shows, as was also the case for the BSS, that the combination of the ensemble designs SE and PE improves the ensemble prediction in general.

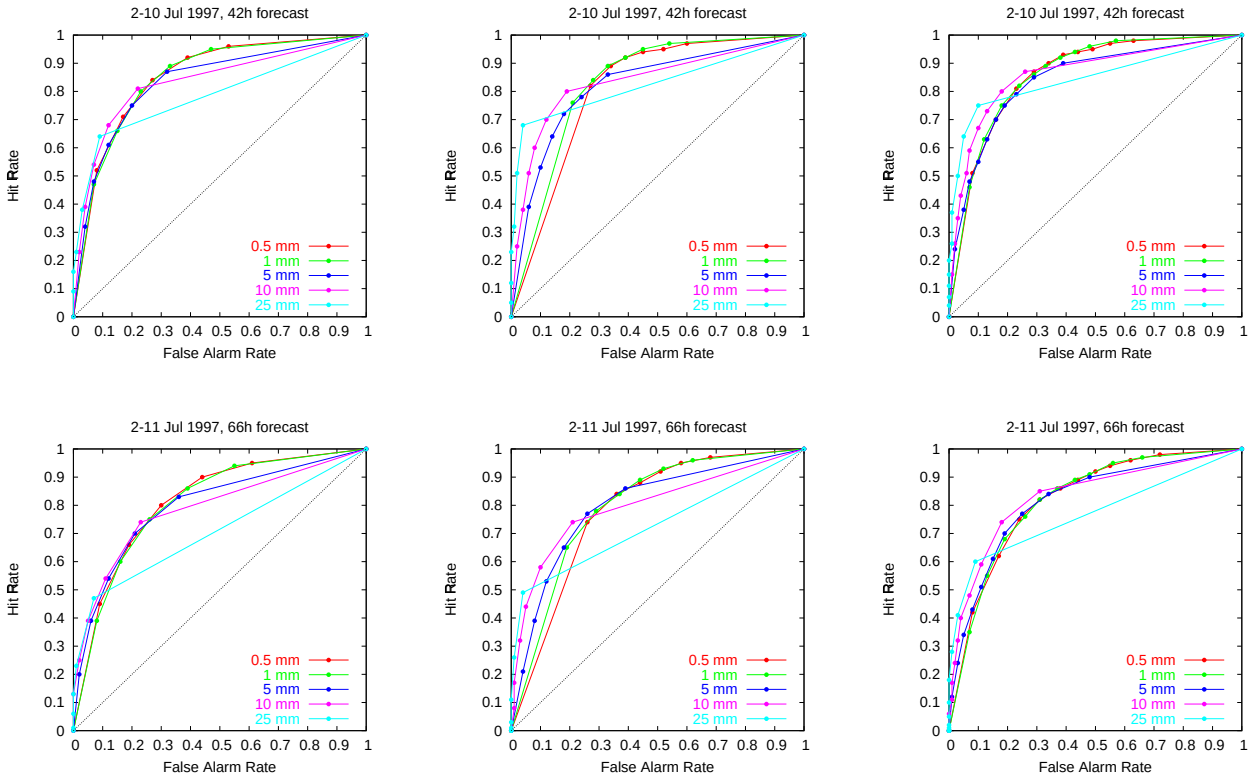

FIG. 22: ROC curves for the ensemble designs PE (left panels), SE (center panels) and CE (right panels) over several thresholds of 24 hourly rainfall within the period of the Odra 1997 event, for the 18-42h forecast range (upper panels) and for the 42-66h forecast range (lower panels).

6 Case studies

In addition to the verification from the previous section it has been of interest to investigate the performance of the ensemble designs during the single days of the forecast periods. This can give an indication on the usefulness of the different ensemble approaches in the very different meteorological situations around the three heavy rain events. The next section is dedicated to this investigation. It has been done in the light of using the ensemble forecasts in the context of river catchment hydrology. A second section will take a closer look at the performance of the selection procedure used to create the SE.

6.1 River basin rainfall

As for hydrological applications and in the light of flood forecasting the basin related rainfall amounts are especially of interest, the following case studies have been performed on basis of basin averaged rainfall amounts. Areal averaged precipitation amounts were determined from the rainfall fields of the ensemble forecasts as well as from the DWD precipitation analysis by calculating areally weighted averages. This accounts for the different grid topologies of the DMI-HIRLAM grid and the LM-grid, on which the precipitation analysis is based.

Fig. 23 shows the daily accumulated 24 hour rainfall over the simulated period of the Piemonte event for the forecast range 18–42h and 42–66h, respectively. The dark colours in the ensemble columns (blue, green and red) depict the minimum rainfall given by any of the ensemble members. The light colours analogously depict the maximum rainfall, and the medium colours indicate the ensemble mean. It should be noted that the period, where the model level data from the ECMWF-EPS were available, is different from the period, for which the rainfall analysis was available. The first ends at 4^{th} November and the latter starts at 1^{st} November. Therefore, the rainfall forecasted by the ensembles around the 28^{th} October are not a false alarm, but occurred actually, as can be seen when looking at the in situ observations shown in Fig. 9. All four stations observed rainfall, and three of them amounts beyond 10 mm.

During the days leading to the heavy rain event of the Piemonte flood the 18–42h forecasts underestimate precipitation around the 4^{th} November (Fig. 23, upper panel). The 42–66h forecasts for the same day perform better (Fig. 23, lower panel), and both the SE and the PE capture the observed rainfall amounts. Up to the 4^{th} of November, the SE and the PE perform mainly with a similar ensemble spread when regarding minimum and maximum, and the ensemble averages correspond nearly. The spread of the CE is always largest, because it includes the extreme members from both SE and PE.

In the 42–66h forecast for the 5^{th} November the SE spread and the PE spread differ significantly, the latter indicating clearly higher probabilities for the extreme event. This can be related to the fact that the PE members are based on the control data from the host model ensemble. Thus the synoptic prescription is the same for those members. As the heavy rain of this event is significantly enhanced by the effect of the orography of the Alps, there is a clear and strong forcing for condensation processes, inducing a strong reaction by all convection/condensation schemes. This can be regarded

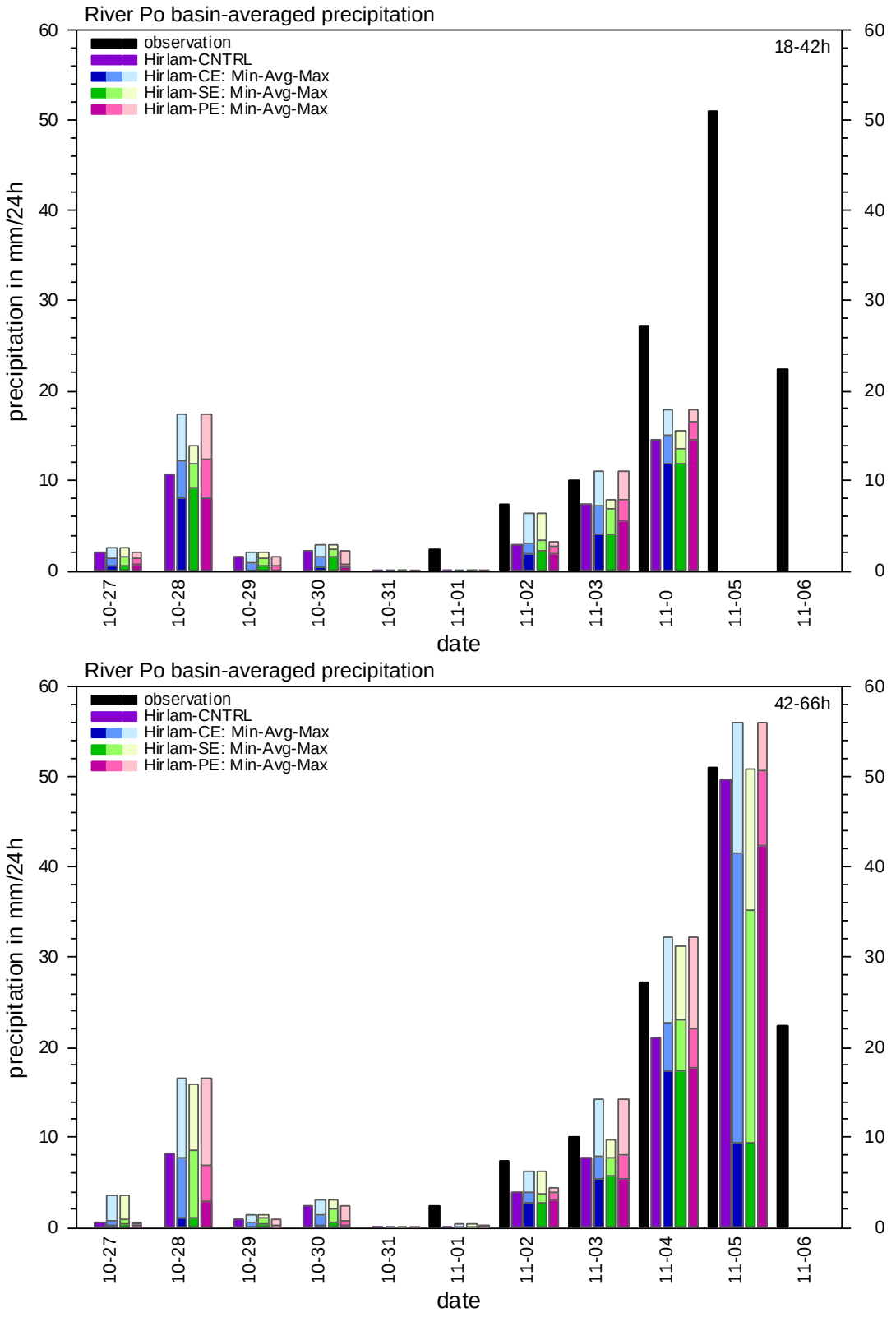

FIG. 23: 24-hourly accumulated daily precipitation between 18 and 42 hours (upper panel) and between 42 and 66 hours (lower panel) over the Po river basin during the November 1994 period of the HIRLAM ensemble runs. The columns show the DWD rainfall analysis of observations (black), the ensemble control (violet), and the ensembles CE, SE and PE (see legend).

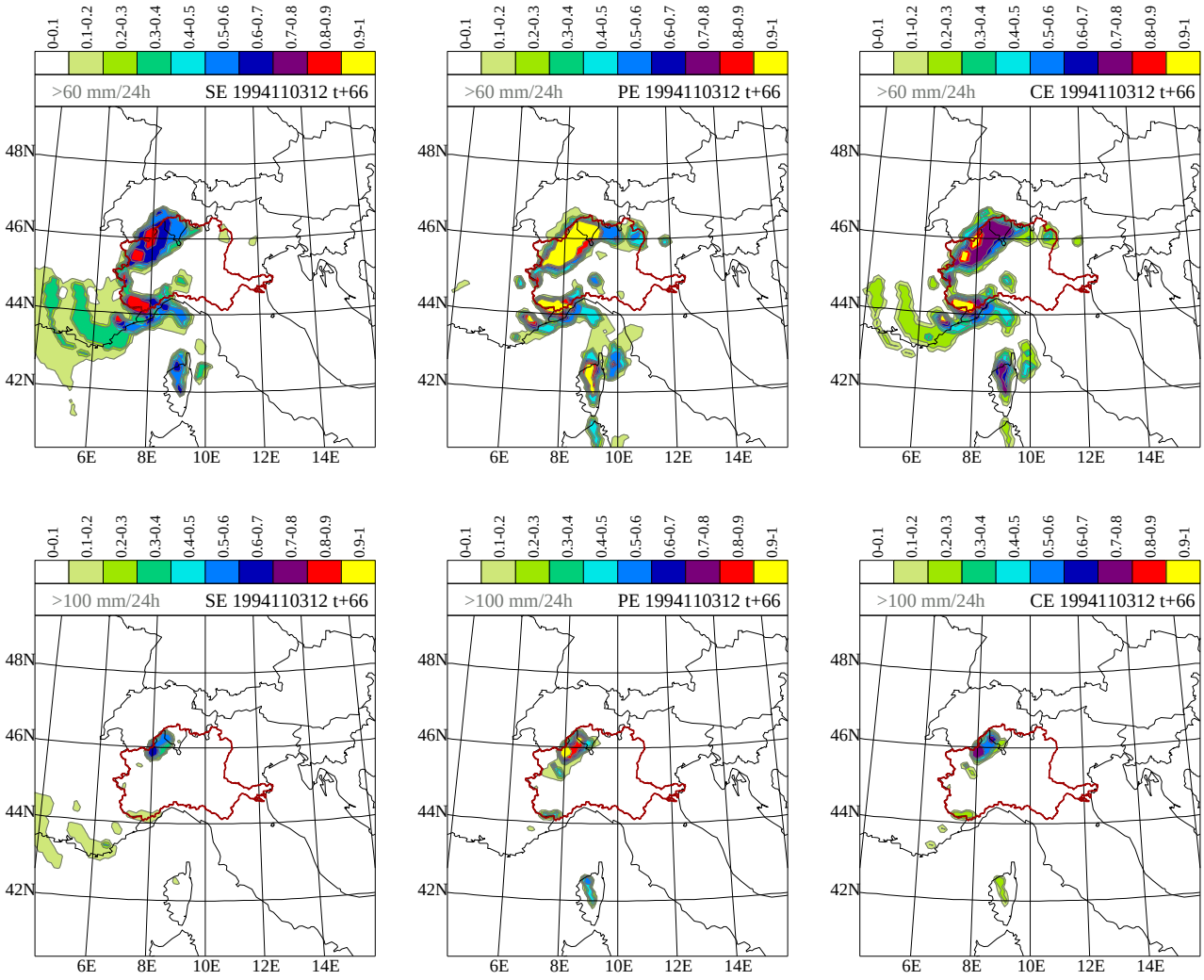

FIG. 24: Probability of precipitation over the Piemonte region for 24 hourly rainfall between 5th November 1994, 6:00 UTC and 6th November 1994, 6:00 UTC. Shown are the probabilities for exceeding a threshold of 60 mm (upper panels) and for exceeding 100 mm (lower panels), as forecasted by the ensembles SE (left panels), PE (panels in the middle) and CE (right panels). The river basin of the Po river is depicted in brown.

as one of the reasons for the reduced PE spread. Even the ensemble mean from the PE therefore practically reaches the observed rainfall average amounts of about 51 mm (Fig. 23, lower panel, medium red column).

The SE performs, however, not bad in the 42–66h forecast range. At the 5th November the control forecast (Fig. 23, lower panel, violet column) performs very well and one of the perturbed SE members also shows rainfall amounts beyond 50 mm (Fig. 23, lower panel, light green column). As a consequence the SE would indicate the heavy rain with a probability of 33% for rainfall above 50 mm (2 out of the 6 members) on average over the Po river basin.

It is also of interest in this case to see how the ensembles distribute rainfall probabilities in space over the river catchment. The spatial distribution over the Po river basin of the probabilities of precipitation (POP) exceeding a threshold of 60 mm and 100 mm is therefore shown in Fig. 24 for the ensembles SE, PE and CE. The underlying ensemble forecasts are the 42–66h predictions

from the integrations started at 3rd November 12:00 UTC. The forecast period corresponds with the period referred to by the columns at day '11-05' in Fig. 23, lower panel. The POP figures clearly indicate the significant signal of high probabilities given by the PE over the NW part of the Po river catchment. The location of the signals is very good when comparing with the locations of the most significant observed rainfall amounts as indicated by the DWD precipitation analysis for the same period (Fig. 25). In the SE the signal is weaker, but its location is very good. This becomes especially clear when regarding the 100 mm threshold probabilities. Both the SE and the PE seem to perform well in this case, and the probabilities given by the combination ensemble CE seem to be even more precise (Fig. 24, panels to the right).

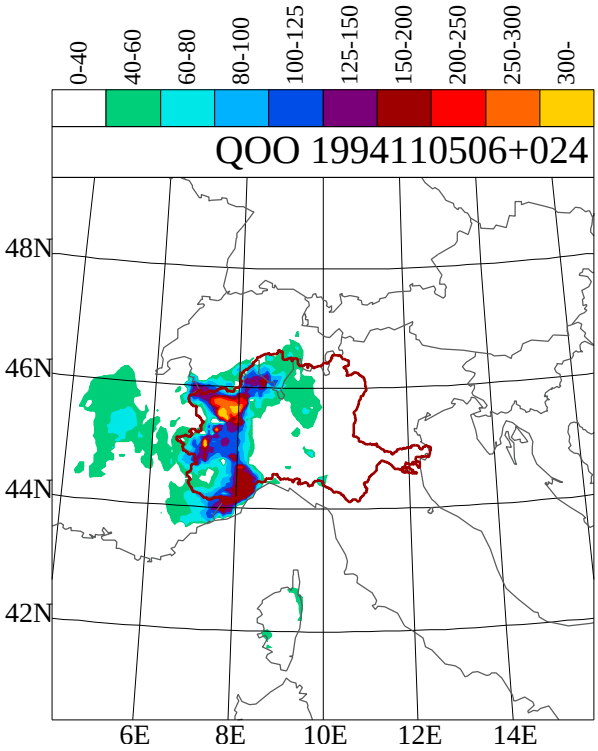

FIG. 25: DWD analysis of the observed 24 hourly precipitation in mm over the Piemonte region between 5th November 1994, 6:00 UTC and 6th November 1994, 6:00 UTC. The river basin of the Po river is depicted in brown.

It should be noted that the HIRLAM predictions underestimate the extremes of rainfall. This negative bias in the rainfall forecasts is illustrated indirectly by the chosen thresholds in Fig. 24, which are far below the more than 300 mm of rainfall that show up in the DWD precipitation analysis. It becomes clear that the probabilities provided by the HIRLAM ensembles for the threshold 300 mm are zero, and thus a valuable probabilistic forecast on basis of the shown probability maps would be problematic. If the probability distribution given through the ensemble is regarded relative to the probability distribution based on a climatology of the underlying model, however, the ensemble may provide qualitative information about a possible significant rainfall event. This is to say that if the model predicts rainfall amounts beyond 100 mm , then this may be regarded as an indication for extreme rainfall, even though the real amounts may be higher. As an example, a forecaster would probably consider to give out a warning, if the model would show significantly larger rainfall amounts over a certain region than usually, provided that the model is reliable.

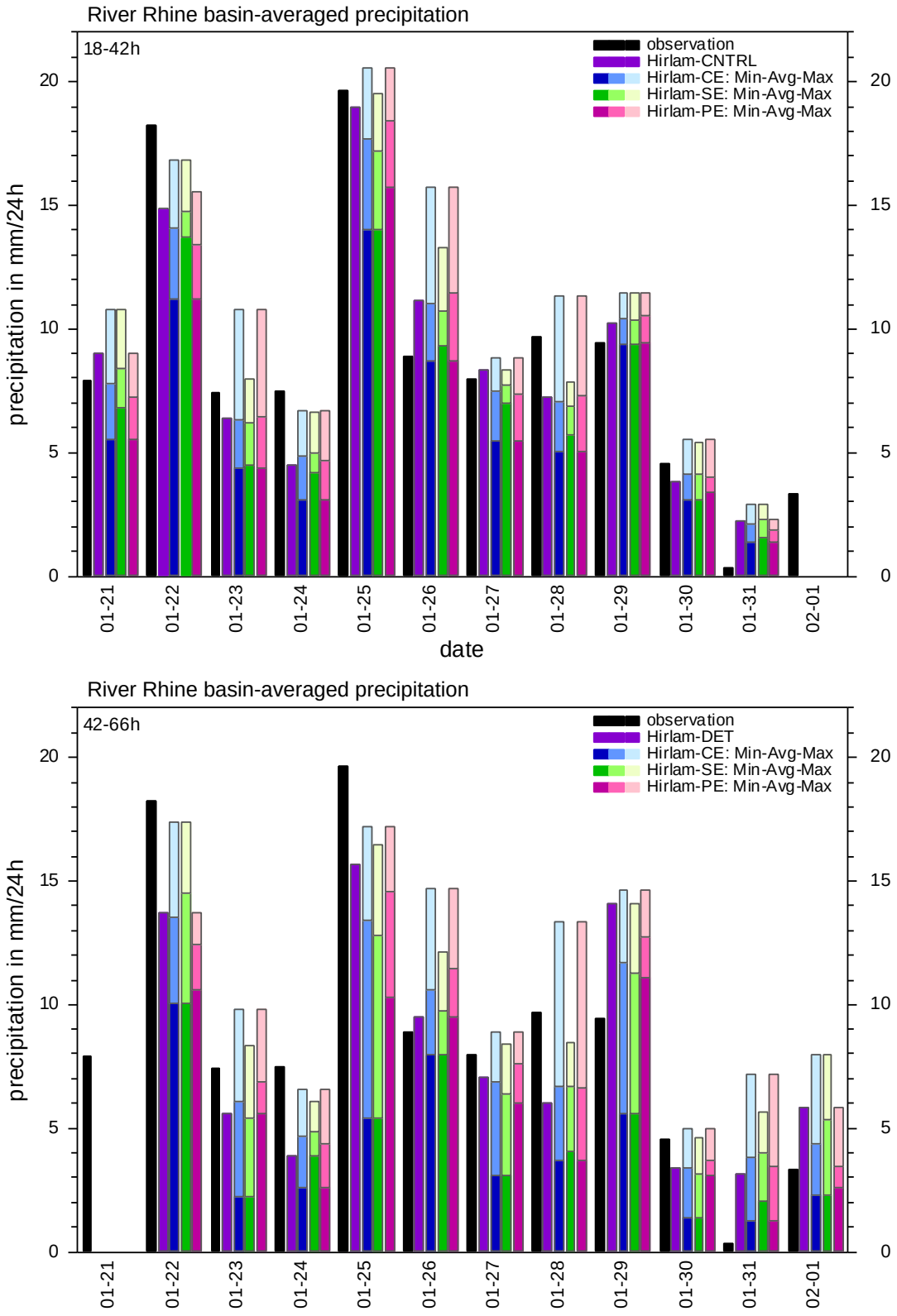

FIG. 26: 24-hourly accumulated daily precipitation between 18 and 42 hours (upper panel) and between 42 and 66 hours (lower panel) over the river Rhine basin during the January 1995 period of the HIRLAM ensemble runs. The columns show the DWD rainfall analysis of observations (black), the ensemble control (violet), and the ensembles CE, SE and PE (see legend).

For the second historical event, Fig. 26 shows the average 24 hourly precipitation over the river Rhine basin during the period covered by the DMI-HIRLAM ensemble simulations. A diagram for the 18–42h forecast range as well as for the 42–66h forecast range are given. Similar diagrams are shown in Fig. 27 for the average rainfall over the Meuse river basin. At the 18–42h forecast range the control forecast captures the synoptic development very well and gives good estimates for the river Rhine basin average precipitation (Fig. 26, violet columns). The control forecasts over the much smaller Meuse basin do not exhibit the same quality (Fig. 27, violet columns), which can be an indicator for forecast errors in the rain area. The ensembles do only slightly enhance the quality of the rainfall prediction over the river basins at the 18–42h forecast range.

At the 42–66h forecast range the quality of the control forecasts has decreased. There seems to be, however, only marginal improvement in the predictions over the river Rhine catchment through the ensemble designs (Fig. 26). Over the basin of the river Meuse, on the other hand, the ensembles seem to become more useful. Especially during the two days showing the maximum observed average precipitation amounts over the Meuse basin, the days '01-22' and '01-29', at least one member of either the SE (day '01-22') or the PE (day '01-29') predict the observed amounts significantly better than the control (Fig. 27, lower panel). The same is the case during the days '01-27' and '01-28'. Utilizing the CE would also in this case be of advantage during most of the days from the period.

The results from the third heavy rain event are different from the first two, in that the control forecasts performed rather poorly during the whole integration period. Fig. 28 shows, analogously as in the figures above, the catchment averaged daily rainfall amounts over the Odra river basin. In this case the spread of the ensembles is large during the days with the large rainfall amounts, and the members from the CE predicting maximum precipitation amounts are significantly closer to the analyzed values than the control forecast during most of the days. The PE in the 18–42h range issues a false alarm at day '07-01', where most of the parameterization schemes over-react on the underlying dynamic conditions (Fig. 28, upper panel). The SE behaves more realistic at this day. During the days with the extreme rainfall all the ensembles tend to indicate the start of the rainfall culmination over the Odra river basin earlier than observed when regarding the 18–42h forecast range (Fig. 28, upper panel). The culmination of the rainfall seems to exhibit a flatter course in the forecasts than the analysis data shows. Besides possible errors in the timeliness of the heavy rain event and the rain volume, spatial displacements of the dynamically most active regions in the model simulations are to be considered as possible causes for this.

The columns for day '07-05' in the diagram for the 18–42h range and the columns at day '07-06' in the 42–66h range refer to the ensemble integrations started at the 4^{th} of July 12:00 UTC. These integrations clearly indicate the earlier start of the heavy rain event and its broader maximum when compared to the DWD rainfall analysis. However, the most extreme member of the CE, which arises from the PE, nearly reaches the observed values of the basin average precipitation (light red and light blue colours in Fig. 28). The ensemble forecasts launched one day later, namely at the 5^{th} of July 12:00 UTC, occur at day '07-06' in the upper diagram of Fig. 28, and at day '07-07' in the lower diagram of the same figure. It is interesting to note that the SE exhibits better estimates for the rainfall development than the PE in this integration, especially at the longer forecast range 42–66h. This may indicate that the atmospheric state estimation of the control was not the "best guess"

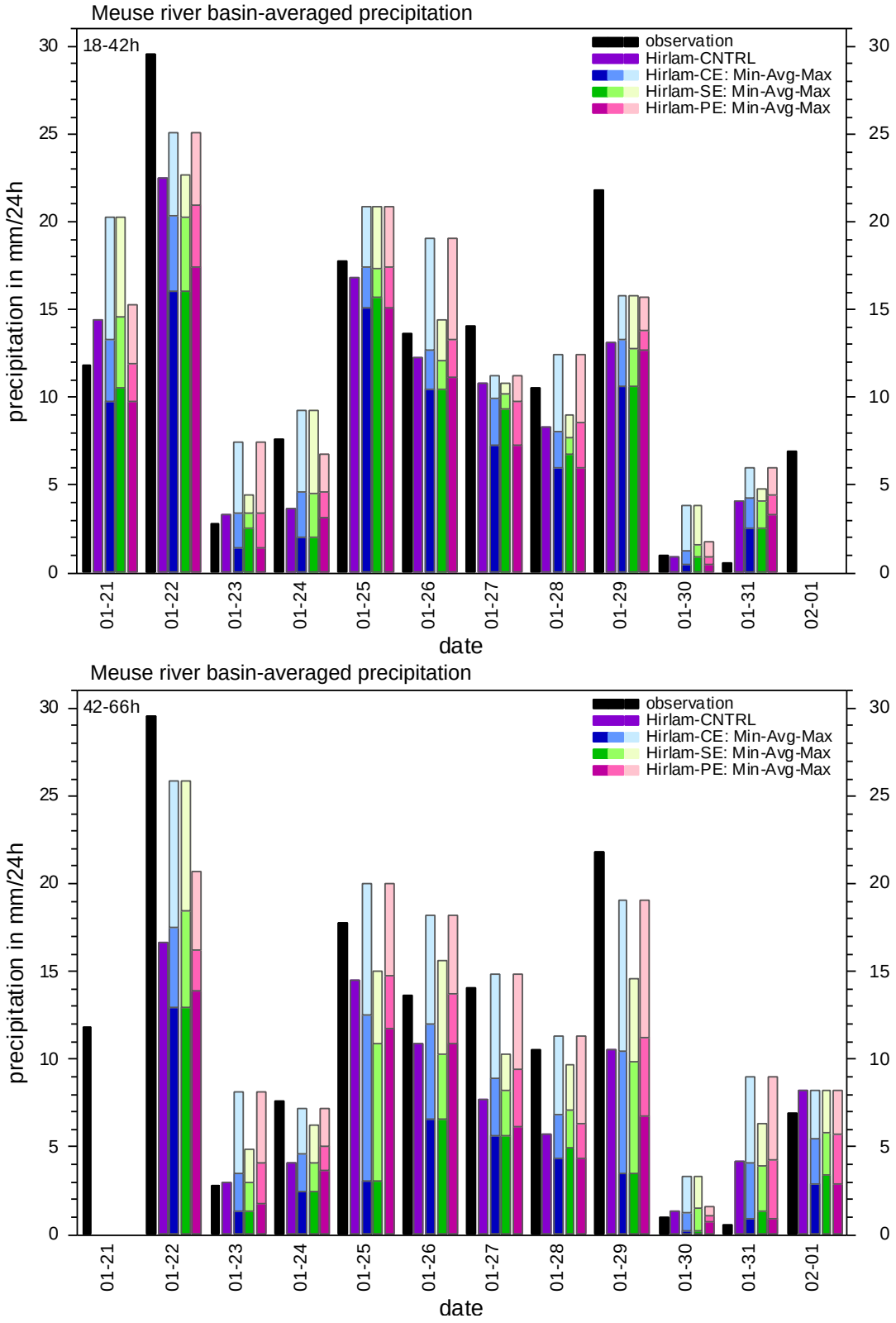

FIG. 27: 24-hourly accumulated daily precipitation between 18 and 42 hours (upper panel) and between 42 and 66 hours (lower panel) over the river Meuse basin during the January 1995 period of the HIRLAM ensemble runs. The columns show the DWD rainfall analysis of observations (black), the ensemble control (violet), and the ensembles CE, SE and PE (see legend).

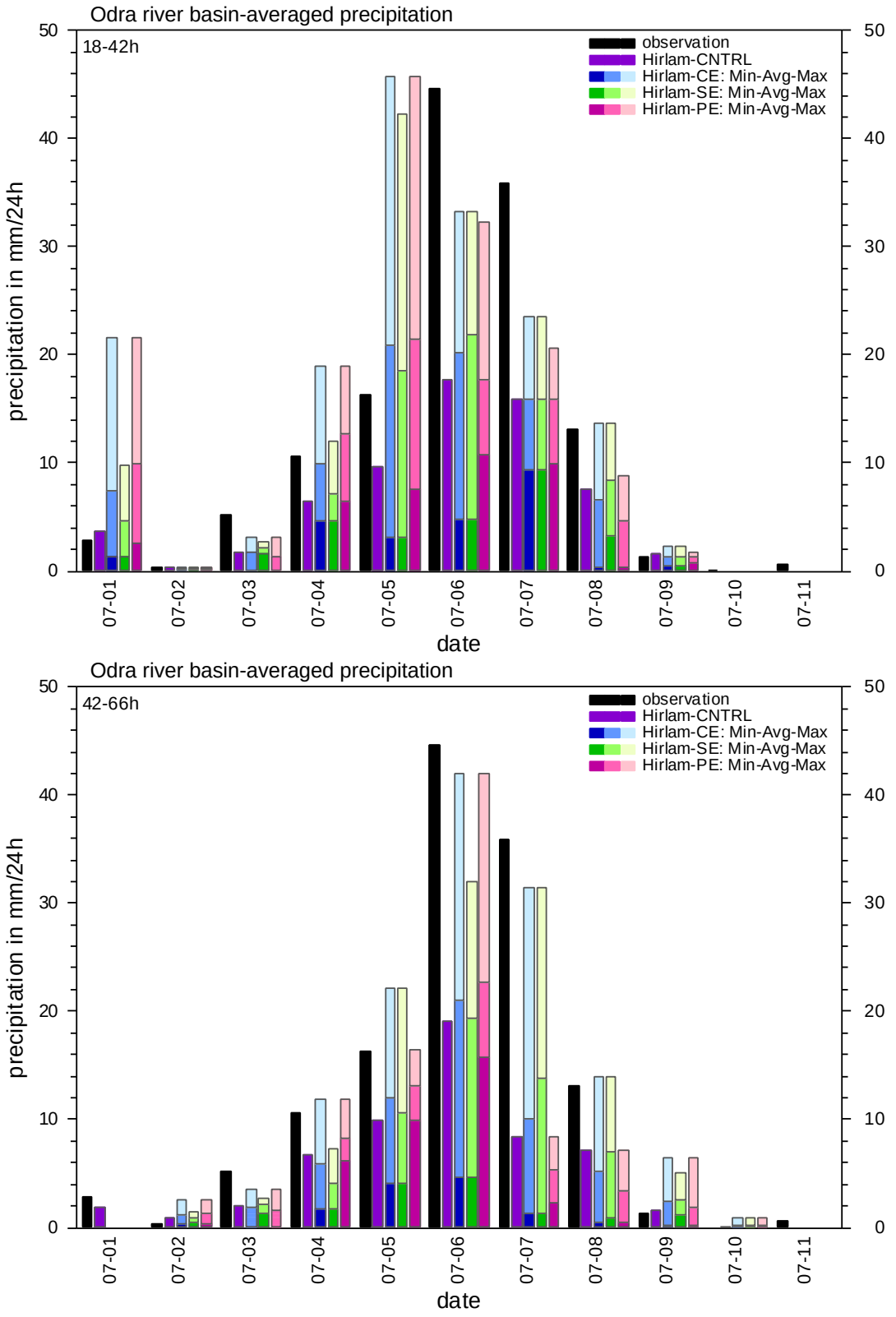

FIG. 28: 24-hourly accumulated daily precipitation between 18 and 42 hours (upper panel) and between 42 and 66 hours (lower panel) over the river Odra basin during the July 1997 period of the HIRLAM ensemble runs. The columns show the DWD rainfall analysis of observations (black), the ensemble control (violet), and the ensembles CE, SE and PE (see legend).

for this day. It is consistent with the poor performance of the control forecast (violet columns). Addressing the uncertainty in the atmospheric state at the initial time and at the lateral boundaries of the LAM seems to be of predominant importance in this forecast. As the PE is solely based on the control of the host model, it cannot account for different dynamical developments arising from uncertainties in the atmospheric state prescription. After all parameterization schemes tended to release more precipitation at day '07-06' than the control forecast (Fig. 28, upper panel, violet versus red columns), the rainfall over the Odra river basin stops rather early during day '07-07' in all members of the PE (Fig. 28, lower panel).

The spatial structure of the ensemble forecasts differs, too. This can be recognized when regarding the POP charts. As an accumulation period for rainfall over 24 hours does not show the structures very clearly, an accumulation over 48 hours was chosen, which comprises the forecast range 18–66h.

Fig. 29 shows the POP for the 40 mm/48h rainfall threshold from the ensemble integrations started at the 4th of July 12:00 UTC, valid for the lead time 5th July to 7th July 6:00 UTC, respectively. For this period the respective DWD precipitation analysis are shown in Fig. 32, upper panels. They show clearly the two locations with the major rainfall, namely around Lysa Hora in the SE corner of the Odra basin and around Praded several degrees NW of Lysa Hora (see also Fig. 12). Even though POP maps are not quantitatively comparable with analysis of rainfall observations, similarities in the field structures are obvious. They arise from the fact that the contour of the 50%-quantile of the POP field referring to a certain rainfall threshold corresponds with the contour for that rainfall threshold in the precipitation field of the ensemble median.

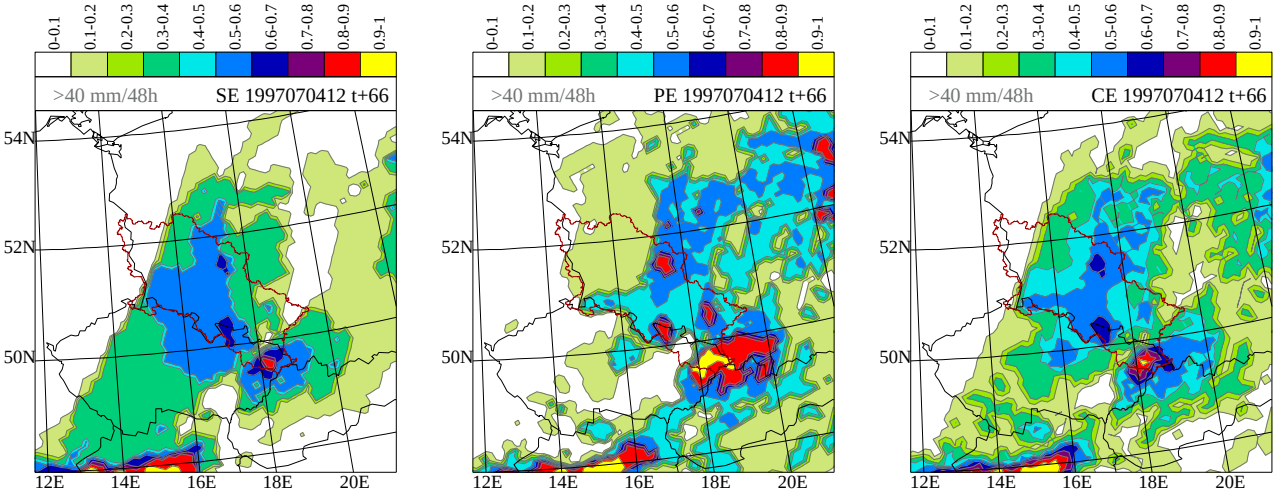

FIG. 29: Probability of precipitation exceeding a rainfall threshold of 40 mm/48h over the region of the Odra river basin (brown) between 5th July 1997, 6:00 UTC and 7th July 1997, 6:00 UTC, as forecasted by the ensembles SE (left panel), PE (panel in the middle) and CE (right panel).

The POP from the SE indicates a high probability (> 80%) for Lysa Hora, where over Praded it is weaker. In the POP from the PE the signals are stronger over the two locations, and there are additional locations exhibiting high probabilities for exceeding the rainfall threshold. The probability structures between the POP form the SE and the POP from the PE are quite different in this case. It is interesting to note that the structure of the POP from the CE exhibits most similarity with

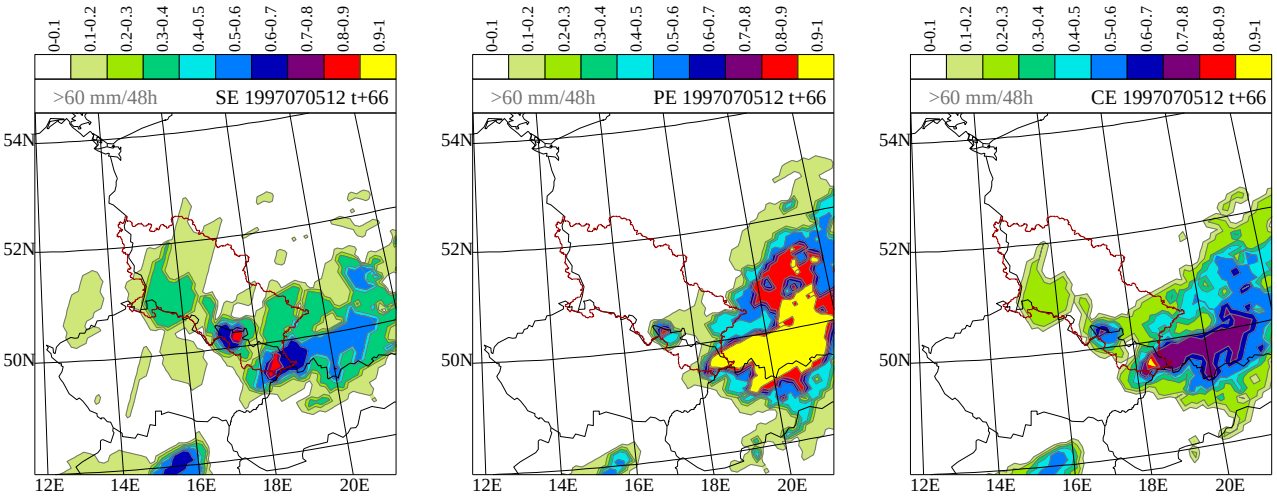

FIG. 30: Probability of precipitation exceeding a rainfall threshold of 60 mm/48h over the region of the Odra river basin (brown) between 6th July 1997, 6:00 UTC and 8th July 1997, 6:00 UTC, as forecasted by the ensembles SE (left panel), PE (panel in the middle) and CE (right panel).

the rainfall structures of the analyzed rainfall fields. Both the SE and the PE seem to add value to the POP.

The POP maps from the ensemble forecasts started at the 5th of July 12:00 UTC, and valid for the time between 6th July and 8th July, are given in Fig. 30, now for the 60 mm/48h rainfall threshold in order to better detect the probability signals. The PE shows again larger probabilities than the SE, and it places the major part of the signal outside the river catchment. In the SE, only some members exhibit this tendency, leaving the highest probabilities to the locations around Lysa Hora and Praded. The POP field structures may be compared to the respective analyzed rainfall fields shown in Fig. 32, the upper right and lower left panel. It becomes immediately clear that the PE

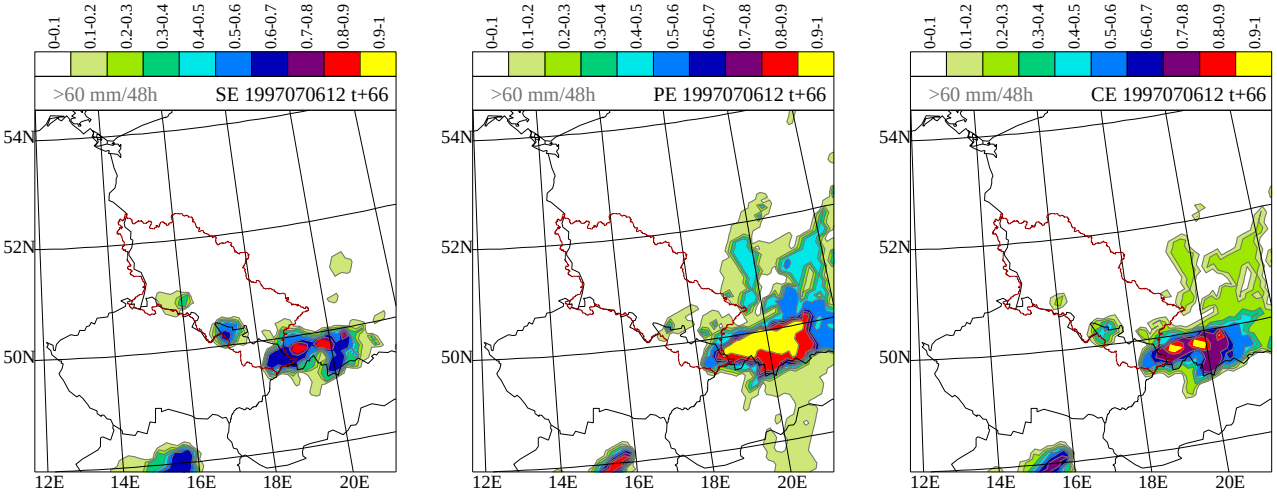

FIG. 31: Probability of precipitation exceeding a rainfall threshold of 60 mm/48h over the region of the Odra river basin (brown) between 7th July 1997, 6:00 UTC and 9th July 1997, 6:00 UTC, as forecasted by the ensembles SE (left panel), PE (panel in the middle) and CE (right panel).

estimation of the POP location over-emphasizes the the area east of the Odra river catchment. The SE seems to capture the situation more realistically. A tendency towards estimating the rain area east of the Odra basin is, however, also visible. Actually, this development happened first about one day later, as can be seen from the precipitation analysis for the next day (Fig. 32, lower right panel).

The ensemble forecasts started another day later and valid for the time between 7th July and 9th July, are shown in Fig. 31 for the 60 mm/48h rainfall threshold. The effect mentioned above can be recognized in these ensembles, too. Only the SE shows a medium probability signal of about 60% over Praded still, where the DWD precipitation analysis for the 7th of July (Fig. 32, lower left panel) reveals heavy rain still over this location. In the CE the signal over Praded has become even weaker because of the influence from the PE ensemble (Fig. 31, right panel).

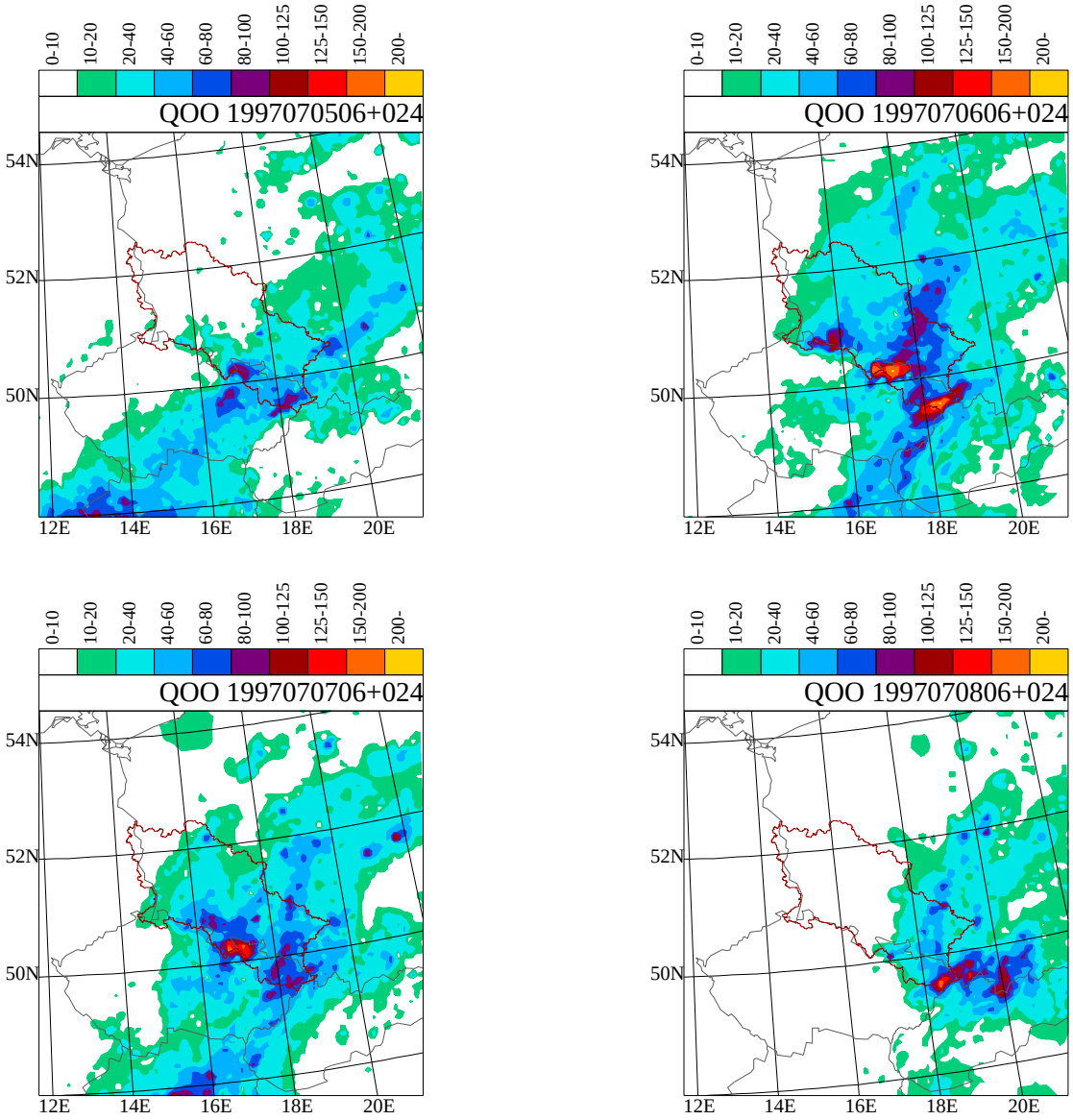

FIG. 32: DWD analysis of the observed 24 hourly precipitation in mm over the region of the Odra river basin (brown) between the 5th and 6th July (upper left), the 6th and 7th July (upper right), the 7th and 8th July (lower left) and the 8th and 9th July (lower right), 6:00 UTC respectively.

6.2 Selection performance

In order to get an impression of how well the selection procedure for the SE (Sec. 2.1) works, a DMI-HIRLAM ensemble including all the 50+1 members was integrated. As this necessitates an enormous amount of computer resources, one day from the period of the Rhine/Meuse flood of January 1995 was selected only. The selected day was the 21st January, with the forecast starting at 12 UTC. The integration time was 72 hours ahead as in the SE, i. e. it ended at 12 UTC the 24th January. Within this period, there was a significant rainfall event of about 35 mm of precipitation on average over the Meuse river basin, and more than 20 mm average precipitation over the basin of the river Rhine. This happened within 24 hours roughly during the 22nd January.

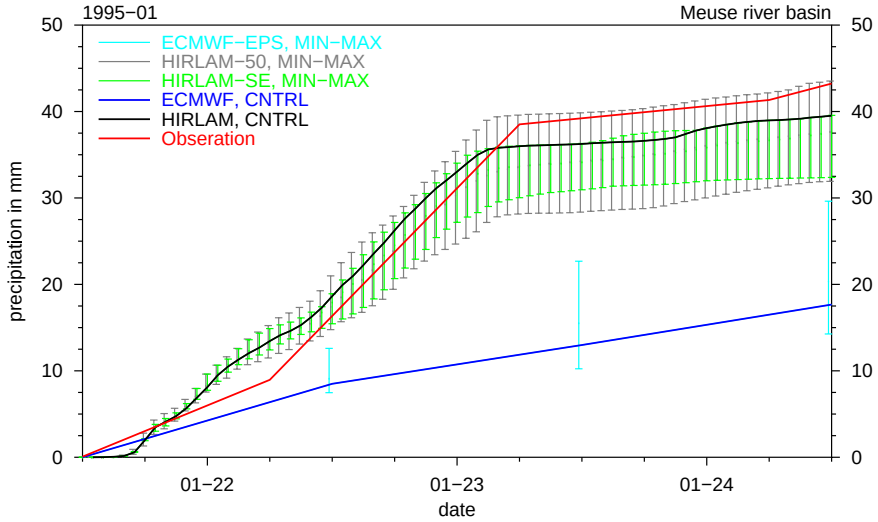

FIG. 33: Rainfall development between 21st and 24th January 1995, averaged over the Meuse river basin. The observed values, as analyzed by the DWD precipitation analyses are plotted in red. The shown EPS prediction from the global ECMWF model (50+1 members) includes the control run (dark blue) and the spread between the members exhibiting the most extreme rainfall (light blue). The 50+1 members ensemble from DMI-HIRLAM is depicted in black for the control and in grey for the spread between the extreme members. The respective DMI-HIRLAM 5+1 members spread between the extremes from the SE is shown in green.

Figs. 33 and 34 show the results for the mentioned period for three ensembles. The first is the global ECMWF-EPS, the second one is the 50+1 members DMI-HIRLAM ensemble (HIRLAM-51), which used all the members from the ECMWF-EPS as host respectively, and the third ensemble is the HIRLAM-SE including 5+1 members. Note that the control for the two DMI-HIRLAM ensembles is the same. The figures clearly show the positive effect of utilizing a higher spatial resolution, which is more pronounced over the smaller Meuse basin (Fig. 33). For the river Rhine basin, the ECMWF-EPS already captures nearly the observed precipitation, at least with the extreme members (Fig. 34). The DMI-HIRLAM control reveals to perform very well, almost forecasting the observed values. When regarding the HIRLAM-51 ensemble, the observed rainfall amounts are captured, and over the river Rhine basin, even the HIRLAM-SE captures the analyzed values. Where the Meuse catchment is concerned, the HIRLAM-SE does not fully reach the observed amounts, only HIRLAM-51 does.

Over both river basins, the spread between the most extreme members is larger in the HIRLAM-

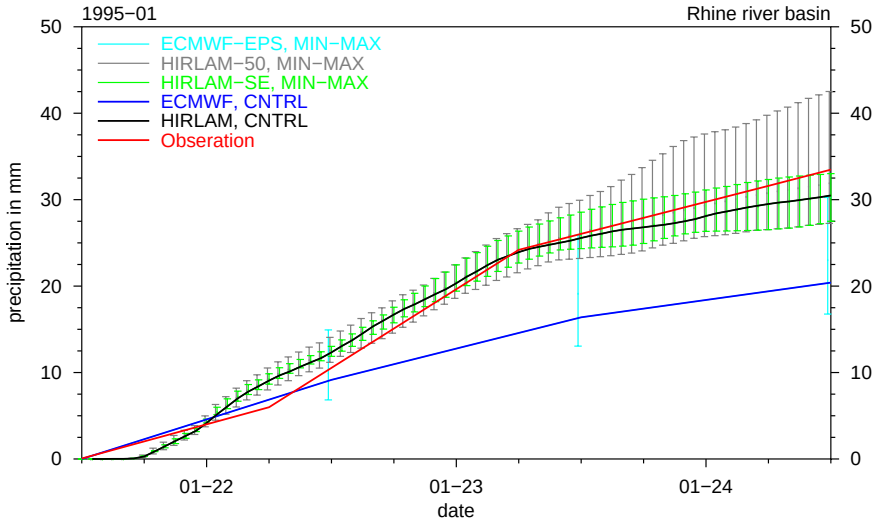

FIG. 34: Rainfall development between 21st and 24th January 1995, averaged over the basin of river Rhine. The observed values, as analyzed by the DWD precipitation analyses are plotted in red. The shown EPS prediction from the global ECMWF model (50+1 members) includes the control run (dark blue) and the spread between the members exhibiting the most extreme rainfall (light blue). The 50+1 members ensemble from DMI-HIRLAM is depicted in black for the control and in grey for the spread between the extreme members. The respective DMI-HIRLAM 5+1 members spread between the extremes from the SE is shown in green.

51 ensemble than in HIRLAM-SE. Why did the selection procedure not choose those members, which would give the largest rainfall amounts? There are three possible reasons for this. First, the target area includes a surrounding around the river basins, which is about 1 degree broad. The spatial extent may be too large and reduce the impact of the catchment related precipitation on the selection procedure. A second possibility could be the temporal linearization over five days, where precipitation is accumulated (Sec. 2.1). Rainfall causing processes that occur first after day three may influence the selection negatively with respect to the three day integration period. A third possibility could be the fact that the selection cannot account for scales in the rainfall structures, which the global EPS model does not resolve. As HIRLAM models a wider range of processes towards small scales than the global model, effects arising from this enhancement probably account for the behaviour of HIRLAM.

In order to investigate this in more detail, several member selections were performed and compared. The first two selections were based on the target area as described in Sec. 3.2, i. e. including both the basin from Meuse and from river Rhine. The accumulation time for the first selection was five days (Sec. 2.1), and for the second selection three days. A third and fourth selection was performed over three and five days respectively, now especially for the Meuse basin with the target area only including the Meuse basin without the large area of the basin of river Rhine. The results from these selection procedures are summarized in Table 1, which lists the 3-day accumulated precipitation averages from the ECMWF-EPS and from the HIRLAM-51 ensemble over the two river basins for all 50+1 members, sorted after rainfall amount. The columns with the bullets denote the members, which were selected by the different procedures. Sorting the members after rainfall amount makes it easier to asses the representativeness of a selection.

It should be noted that the selection algorithm bases its choice on the data from the ECMWF-

EPS only, because only this information is available a priori. Therefore, the columns under the ECMWF-EPS give a picture of the selection quality for the different settings, and the columns under HIRLAM-51 show, how the selection is represented through the HIRLAM model. It can be immediately recognized from Table 1 that the selection via 5-day accumulation over the large area (columns '5' under ECMWF-EPS in Table 1) represents virtually an optimal spread, because the member exhibiting the least (or almost the least in the case of the river Rhine basin) and the largest rainfall accumulation in the ECMWF-EPS model, are included in the selection. The selection based on 3-day accumulation on the other hand exhibits a smaller spread, seen from the ECMWF-EPS. This picture changes in the HIRLAM-51 ensemble. The behaviour of HIRLAM is very inhomogeneous between the members. For example, member number 023, which showed the maximum precipitation amount within the ECMWF-EPS, turns out to be ranked at rather low values in HIRLAM-51. On the other hand, the control forecasts 000, which exhibited small rainfall amounts within the ECMWF-EPS, is ranked much higher in HIRLAM-51. This is a clear indication for the existence and involvement of nonlinear processes at scales, which are not resolved in the ECMWF-EPS model, but which are resolved in HIRLAM. As these processes are not known a priori, it is impossible to account for them during the selection procedure. The selection ensemble based on the 3-day accumulation would have performed better over the basin of river Rhine than the one based on 5-day accumulation, but this may be different at another day. It cannot be regarded as a clear indication that the linearization in time would account for the deficiencies.

Concerning the Meuse catchment, the selections '5M' and '3M' (see Table 1) have been performed additionally in order to investigate the influence of the target area size. These two selections were targeted on a small area, that only contains the Meuse river catchment without the large river Rhine basin. As the leftmost columns under ECMWF-EPS in Table 1 show, the 5-day selection of the smaller area '5M' is very similar to the 5-day selection '5' over the large area with both river catchments included. They only differ by one member (022 in '5M' with 16.9 mm of rainfall instead of 021 in '5' with 21.7 mm). Due to the nonlinear integration in HIRLAM, however, this difference turns into the opposite in the HIRLAM selection ensemble, where member 022 shows 39.8 mm of rainfall, whereas 021 gave 38.5 mm.

The selection over the 3-day period over the Meuse basin (columns '3M' in Table 1) distributes very well with respect to the ECMWF-EPS ranked members, and also with respect to the ranked HIRLAM members. For the 3-day based selections over the large area the distribution of the Meuse basin rainfall (columns marked '3' on the left hand side of Table 1) shows a smaller spread than '3M'. Whether a targeting over the reduced area and using a shorter temporal window in general would be advantageous cannot be concluded just from this case, however. The results at hand can only serve as an indication and should be confirmed by more investigations.

TAB. 1: List of 3-day accumulated rainfall amounts in mm as predicted by the ECMWF-EPS and the HIRLAM-51 ensemble, averaged over the river basin of the Meuse and river Rhine, respectively. The lists are sorted from the member with the smallest rainfall amounts to those with the largest values, respectively. The bullets indicate the selected members when using an accumulation period of 5 days (5), 3 days (3), as well as using an area reduced to the Meuse basin, also for 5 day (5M) and for 3 day (3M) accumulation.

Meuse												Rhine							
ECMWF-EPS						HIRLAM-51						ECMWF-EPS				HIRLAM-51			
NR	mm	5	3	5M	3M	NR	mm	5	3	5M	3M	NR	mm	5	3	NR	mm	5	3
032	14.3	•		•	•	046	31.9					029	16.8			033	27.3		
029	14.8					019	32.0					032	17.6	•		044	27.5		
014	14.9					031	32.0					017	17.7			039	27.5	•	•
004	15.3					032	32.4	•		•	•	025	17.8		•	037	27.5		
007	15.3					036	32.8					046	17.8			011	27.8		
046	15.4					048	33.8					038	18.3			042	28.0		
006	15.7					033	34.1					014	18.5			036	28.4		
010	15.9					028	34.5					010	19.3			028	28.4		
027	16.3					030	34.7					027	19.4			018	28.5	•	•
017	16.5					008	34.7					007	19.6			024	28.6		
022	16.9			•		009	35.0		•			019	20.0			030	29.0		
036	17.3					025	35.4		•			000	20.4	•	•	002	29.1		
034	17.4					042	35.5					048	20.8			026	29.5		
043	17.4					044	35.6					034	21.2			031	29.6		
038	17.5					002	35.7					001	21.2			045	29.6		
000	17.7	•	•	•	•	023	36.5	•		•		050	21.2			000	30.5	•	•
050	18.2					015	36.5					043	21.3			009	30.7		•
025	18.3		•			038	36.5					049	21.5			049	30.7		
049	18.7					029	37.0					042	21.6			013	30.9		
019	18.7					049	37.1					024	21.6			019	30.9		
020	19.1					024	37.1					006	22.0			005	31.0		
026	19.5					045	37.1					021	22.3	•		016	31.0		
048	19.5					011	37.2					036	22.5			015	31.1		
024	19.7					017	37.2					026	22.5			023	31.1	•	
015	19.8					006	37.4					040	22.6		•	008	31.1		
001	20.1					005	37.4				•	020	22.8			010	31.1		
018	20.4	•	•	•	•	003	37.5					022	22.8			047	31.4		
011	20.7					035	37.5					009	22.9		•	007	31.4		
047	20.9					004	37.7					015	23.2			022	31.4		
039	21.2	•	•	•		021	38.5	•				004	23.2			050	31.4		
009	21.4		•			034	38.5					011	23.3			003	32.1		
002	21.4					013	38.6					003	24.1			021	32.6	•	
044	21.6					014	39.1					016	24.2			048	32.9		
021	21.7	•				026	39.2					044	24.3			032	33.0	•	
016	21.8					018	39.4	•	•	•	•	018	24.5	•	•	006	33.5		
033	22.1					012	39.5					013	24.9			046	33.8		
013	22.6					000	39.5	•	•	•	•	008	25.3			034	33.9		
030	22.9					039	39.6	•	•	•		039	25.4	•	•	038	34.0		
003	23.4					022	39.8			•		031	26.0			012	34.0		
042	23.7					007	39.9					033	26.1			004	34.5		
012	23.7					041	40.1					041	26.8			041	35.0		
040	23.8		•		•	016	40.8					030	26.8			029	35.0		
041	23.8					050	40.9					047	26.9			020	35.1		
008	24.0					037	41.2				•	012	27.0			001	35.2		
037	24.1				•	040	41.8		•		•	037	27.2			014	35.3		
031	24.6					010	42.0					005	27.8			043	35.7		
045	26.2					020	42.2					002	28.0			035	36.6		
005	26.3				•	001	42.2					028	28.4			017	36.6		
028	28.7					047	43.3					045	29.3			025	37.8		•
035	29.1					043	43.4					035	29.3			027	38.4		
023	29.6	•		•		027	43.5					023	30.3	•		040	42.5		•

7 Discussion and Conclusions

This report investigated two different ensemble approaches for the prediction of extreme rain events using the high-resolution limited-area model DMI-HIRLAM. The first ensemble approach (selection ensemble, SE) is based on a perturbed host model, which is represented by the ECMWF-EPS. A selection from this 50+1-member global EPS targeted on accumulated precipitation over the area of interest identifies 5+1 members, which are then integrated in DMI-HIRLAM. This ensemble is to address uncertainty in the initial condition as well as at the lateral boundary. The second ensemble approach (parameterization ensemble, PE) tries to address a portion of the model inherent uncertainty responsible for forecast errors in the precipitation field by utilizing different parameterization schemes for condensation and convection. The combination of the two ensemble approaches (combined ensemble, CE) was investigated, too.

Three periods of about 10 days duration around historical heavy rain events that caused or contributed to disastrous river flooding in Europe were studied using the ensembles. They differ significantly in the meteorological synoptic situation that led to the rainfall events. While the first event (Piemonte 1994) occurred during autumn and was caused under a baroclinically unstable regime, where orographic forcing played a major role, the second event (Rhine/Meuse 1995) occurred during winter and was characterized by a chain of frontal systems passing over a rather flat region compared to the orography in the first event. The third event (Odra 1997) was characterized by large scale dynamic developments and convective episodes during summertime over an orographically structured region.

A verification based on common procedures for probabilistic forecasts was applied. It comprises Brier skill score (BSS), reliability diagram, ranked probability diagram (Talagrand diagram) and relative operating characteristic (ROC). The verifying data consists of a precipitation analysis of rain gauge observations. It was performed at the Deutsche Wetterdienst (DWD) for short periods around the three historical heavy rain events.

The probabilistic verification shows that the ensemble designs have skill in BSS up to a threshold of 5 mm/24h over the area of Central Europe. In the case of Piemonte 1994 the skill goes beyond this threshold, which can be regarded to be due to the clear orographic forcing. The relation between hit rates and false alarm rates with respect to probability thresholds, which is given through the ROC curves, shows good skills for all ensemble designs. The best skill shows the CE in most cases, indicating that combining the different ensemble approaches is advantageous.

The reliability diagrams show that the ensembles tend to overestimate the observed frequencies of exceedance for a certain rainfall threshold. The reliability especially at the smaller rainfall thresholds appears, however, to be quite good, when taking into account the small HIRLAM ensemble sizes. The PE ensemble often shows a slightly better reliability than the SE, and the CE turns out to improve the reliability further. The PE design also shows a slightly larger ensemble spread in the ranked histograms than the SE. This occurs in the classes of small rainfall amounts, where the PE members have less tendency to overpredict small amounts of rain or no rain than the model with the standard parameterization.

Case studies were performed in addition to the probabilistic verification. They were based on the

average rainfall over the considered river basin, which gives an indication on the impact of the rainfall on the river basin and which can be useful for flood risk assessment over small river basins. The studies clearly show how different the ensembles perform in different weather situations. In cases where orographic forcing plays a major role, the development of condensation and convection is rather well defined, and the PE tends to exhibit less spread than the SE. The latter includes the possibility for deviations in the synoptic situation connected to less or a modified orographic forcing in some of the ensemble members.

In the cases, where a chain of cyclones of modest depth pass over flat or slightly hilly terrain during the winter, forcing from the surface is weak and both the large scale synoptic flow and the frontal dynamics affect the development of condensation and rainfall in a more sensitive way than in the case of strong orographic forcing. In this case both the SE and the PE exhibit comparable spread, and days can be found where either of them seems superior to the other one in capturing the actual development respectively. Combining the ensembles in the CE is generally advantageous.

The most difficult to predict case has been the summer event, where a large scale ascent in a quasistationary depression was connected to dynamically triggered convection. This means that uncertainties in the initial condition and connected to that the boundary conditions as well as in the model play a decisive role. The CE reveals to be the best ensemble, which becomes especially clear in the POP maps.

The selection procedure for the SE was evaluated over one 3-day forecast period against the respective 50+1-member DMI-HIRLAM ensemble integrating all members of the ECMWF-EPS. The ensemble spread between the most extreme members of the 5+1-member selection reveals to be similar to the global 50+1-member ensemble, when regarded with respect to the rainfall from the global model. When referred to the rainfall amounts that the LAM simulates, however, the 5+1-member selection ensemble can show reduced spread relative to the respective 50+1-member ensemble. This reveals some weaknesses of the selection procedure with respect to the absolute extremes, where the linearized consideration of precipitation accumulation over a 5-day period seems to have more influence on this than the approximation of the target area. The largest impact, however, seems to come from nonlinear processes resolved in the HIRLAM integration, which can enhance or damp precipitation developments locally. This changes the ranking order of the members in the ensemble, i. e. the member exhibiting the largest rainfall amounts in the global model forecast can end up as a member only showing modest rainfall amounts in the LAM forecast. The same can happen the other way around. These nonlinear effects are not known a priori and therefore they cannot be accounted for by the selection procedure used in this study. These results should be regarded with caution and taken as indicative, however, because they are only based on a single forecast. A systematic investigation over forecasts covering a longer period is necessary in order to get results with greater confidence.

The present study did not try to deal with mesoscale features included in the initial state, as Frogner and Iversen (2002) did in order to improve the ensemble skill especially during the early forecast range, because the high-resolution data assimilation of HIRLAM was not included. This should, however, be addressed when trying to optimize the SE ensemble approach. It may help to overcome the weaknesses of the selection procedure mentioned above.

The POP maps have potential to provide valuable information about locations exhibiting a risk for significant precipitation. Especially using the CE when regarding POP distributions over the river basin of interest seems to give some useful indications about locations with a risk for significant precipitation, where the SE and the PE either only show a weak signal or a spatial spread of the probabilities over a larger area. The probability maps created on basis of the ensemble designs in this work seem to be difficult to use in practice, however, because a model bias in the rainfall prediction can disturb the probability distribution. A qualitative estimation may be conceived. This would then necessitate experience about the performance of the ensemble forecasts, analogously to a forecasters evaluation of a significant behaviour of a single deterministic forecast model based on the forecasters experience with the model. To what extent a correction of the bias can improve the probability maps towards a more quantitatively probabilistic rainfall prediction has not been part of the scope of this work.

The POP maps shown for the Odra 1997 case reveal another delicate situation that can occur, namely that the uncertainty about the location of an area showing high probabilities makes a risk assessment difficult when the area is close to the catchment boundary. Regarding a small buffer zone around the actual river catchment and including spatial correlations of the rainfall fields like e. g. the predominant flow over the area of interest can probably be helpful. If the average rainfall over the buffer zone shows significant values, or if the area covered by a predefined probability quantile in the POP map increases significantly, then it can be expected that a reliable risk assessment still may be difficult.

With respect to an application of the ensemble designs for hydrological purposes like flood warning or even flood forecasting, a typical usage would be to perform an ensemble integration with a hydrological model. As long as the rainfall forecasts are biased, this will also become apparent in the water level forecasts. A correction of the rainfall field bias is therefore regarded as a necessary requirement. Future improvements in the NWP model will certainly diminish this error, but it cannot be expected to disappear. An appropriate treatment of the rainfall fields on basis of model related statistics should be envisaged.

A weakness in using LAM ensemble forecasts of small size is that they cannot be expected to represent the PDF of a given variable completely. Especially in cases, where the PDF exhibits multi-modal character with a complex structure, the representation can only be crude, and some modes may even not be represented at all. The probabilistic "resolution", if one may call it this way, is too small for this. A smoothing of the PDF as Wilks (2002) proposes can be expected to improve the representation of the PDF. A useful way how to exploit this in the light of a hydrological application still would have to be found then.

It is open to what extent an increase in the ensemble size of rainfall forecasts based on the same model may help to improve situations, where the ensemble forecast completely fails to predict the occurrence of an event. It may still happen that all ensemble members consistently underpredict the rainfall. An ensemble based on multiple models is probably more powerful in this respect, because the member predictions are less correlated. The use of a multi-scheme ensemble design as used in the PE design of this work can also help to diminish the correlations of the errors between the ensemble members, because the model has changed in some parts. Using this approach on more

model components can be expected to give further improvement. Bright and Mullen (2002) went even further and combined the initial state perturbation with the multi-scheme approach. Their results indicate that such an ensemble design can be superior to a pure multi-scheme approach or an ensemble purely based on initial condition perturbation. Such an approach could probably have enhanced the ensemble predictions further over the Odra river basin in the period of the Odra 1997 heavy rain event.

There are many possibilities to present the forecasts from NWP ensembles, and the choice of the methodology clearly depends on the purpose and the application the forecast is aiming at. Where the assessment of potential heavy rains that may lead to river floodings is concerned, rainfall forecasts have to be related to the river basin under consideration. A first assessment outside a hydrological model application can therefore regard the basin averaged rainfall, as was done within this work. In order to detect regions within the river catchment, where there is potential risk for heavy rain, a refinement onto sub-basins of the river catchment could be considered. Probability maps of precipitation may be a help, too. Relating them to the model climatology can at the same time be helpful for a risk assessment (Lalaurette, 2002). The differences in the spatio-temporal structure of heavy rain events due to differences in large scale and mesoscale dynamic processes as well as orographic influence do not seem to allow a static approach, which e. g. only utilized 24 hourly accumulated rainfall. It seems to be necessary to regard different accumulation periods as well as a series of thresholds in order to detect regions of potential risk for the occurrence of significant precipitation.

Acknowledgments

The Dipartimento di Scienze della Terra e Geologico Ambientali provided the river catchment data of the Po basin, and Riza, WL Delft Hydraulics and the Joint Research Centre (JRC) in Ispra made the river catchment data for the rivers Rhine and Meuse available. We are also grateful to Dr. Roberto Buizza (ECMWF), who kindly supplied the model level data from the ensemble integrations with the ECMWF-EPS, and to the Deutsche Wetterdienst (DWD), who provided the precipitation analysis data. Bjarne Amstrup (DMI) performed the multi-event contingency table calculations for evaluating the different parameterization schemes, which we gratefully acknowledge. Acknowledgement is also made for the use of ECMWF's computing and archive facilities in this research.

References

Anthes, R.A., 1977: A Cumulus Parameterization Scheme Utilizing a One-Dimensional Cloud Model. *Mon. Wea. Rev.*, **105**, 270–286.

Bright, D.R., Mullen, S.L., 2002: Short-Range Ensemble Forecasts of Precipitation during the Southwest Monsoon. *Wea. and Forecasting*, **17**, 5, 1080–1100.

Buizza, R., Miller, M., Palmer, T.N., 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.

Buzzi, A., Cacciamani, C., Paccagnella, T., Patruno, P., Tartaglione, N, 1995: Preliminary meteorological analysis of the Piedmont flood of November 1994. *MAP Newsletter*, **2**, 2–6.

Cuxart, J., Bougeault, P., and Redelsperger, j.-L., 2000: Turbulence closure for a non-hydrostatic model. *Q. J. R. Meteor. Soc.*, **126**, 1–30.

Davis, H.C., 1976: A lateral boundary formulation for multi-level prediction models. *Quart. J. Roy. Meteor. Soc.*, **102**, 405–418.

Doms, G., Schättler, U., 1999: The non-hydrostatic limited-area model LM (Lokal-Modell) of DWD. Part I: Scientific documentation. Available from Deutscher Wetterdienst, Geschäftsbereich Forschung und Entwicklung, Postfach 100465, 63004 Offenbach, Germany.

Du, J., Mullen, S.L., Sanders, F., 1997: Short-Range Ensemble Forecasting of Quantitative Precipitation. *Mon. Wea. Rev.*, **125**, 2427–2459.

Du, J., Tracton, M.S., 1999: Impact of lateral boundary conditions on regional-model ensemble prediction. *Research activities in atmospheric and oceanic modelling*, (edited by H. Ritchie), Report **28**, CAS/JSC Working Group Numerical Experimentation (WGNE), WMO/TD-No. 942, 6.7–6.8.

Ehrendorfer, M., 1997: Predicting the uncertainty of numerical weather forecasts: A review. *Meteorol. Z.*, **6**, 147–183.

Fink, A., Ulbrich, U., Engel, H., 1995: Aspects of the January 1995 flood in Germany. *Weather*, **51-2**, 34–39.

Frogner, I.-L., Iversen, T., 2001: Targeted ensemble prediction for northern Europe and parts of the north Atlantic Ocean. *Tellus*, **53A**, 35–55.

Frogner, I.-L., Iversen, T., 2002: High-resolution limited-area ensemble predictions based on low-resolution targeted singular vectors. *Quart. J. Roy. Meteor. Soc.*, **128**, 1321–1341.

Hamill, T.M., Colucci, S.J., 1997: Evaluation of Eta-RSM short-range ensemble forecasts. *Mon. Wea. Rev.*, **125**, 1312–1327.

Hamill, T.M., Colucci, S.J., 1998: Evaluation of Eta-RSM ensemble probabilistic precipitation forecasts. *Mon. Wea. Rev.*, **126**, 711–724.

Hersbach, H., Mureau, R., Opsteegh, J.D., Barkmeijer, J., 2000: A Short-Range to Early-Medium-Range Ensemble Prediction Sastem for the European Area. *Mon. Wea. Rev.*, **128**, 3501–3519.

Hou, D., Kalnay, E., Droegemeier, K.K., 2001: Objective Verification of the SAMEX'98 Ensemble Forecasts. *Mon. Wea. Rev.*, **129**, 73–91.

Houtekamer, P.L., Derome, J., 1994: Prediction Experiments with Two-Member Ensembles. *Mon. Wea. Rev.*, **122**, 2179–2191.

Houtekamer, P.L., Lefaivre, L., Derome, J., Ritchie, H., Mitchell, H.L., 1996: A System Simulation Approach to Ensemble Prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.

Källén, E. (ed.), 1996: *HIRLAM Documentation Manual system 2.5.* Available from the Swedish Meteorological and Hydrological Institute (SMHI) Norrköping, Sweden.

Keil, C., Volkert, H., Majewski, D., 1999: The Oder flood in July 1997: Transport routes of precipitable water diagnosed with an operational forecast model. *Geophys. Res. Lett.*, **26**, 2, 235–238.

Kessler, E., 1969: On the distribution and continuity of water substance in atmospheric circulations. *Meteor. Monogr.*, **32**, 84 pp.

Kuo, H.L., 1965: On Formation and Intensification of Tropical Cyclones Through Latent Heat Release by Cumulus Convection. *J. Atmos. Sci.*, **22**, 40–63.

Kuo, H.L., 1974: Further Studies of the Parameterization of the Influence of Cumulus Convection on Large-Scale Flow. *J. Atmos. Sci.*, **31**, 1232–1240.

Källberg, P., 1989: *HIRLAM Forecast Model Level 1 Documentation Manual.* 77 pp. Available from the Swedish Meteorological and Hydrological Institute (SMHI) Norrköping, Sweden.

Lalaurette, F., 2002: Early Detection of Abnormal Weather Using a Probabilistic Extreme Forecast Index. *ECMWF Tech. Memor.*, **373**, 27 pp.

Lorenz, E., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **42**, 433–471.

Lorenz, E. N.; 1993: The Essence of Chaos. University of Washington Press, 1, Seattle, 227 pp.

MAP, 2002: The MAP Data Centre. MeteoSwiss, Krähbühlstr. 58, CH-8044 Zürich.

Marsigli, C., Montani, A.,Nerozzi, F., Paccagnella, T., Tibaldi, S., Molteni, F., 2001: A strategy for high-resolution ensemble prediction. II: Limited-area experiments in four Alpine flood events. *Quart. J. Roy. Meteor. Soc.*, **127**, 2095–2115.

Meijgaard, E. van and Jilderda, R., 1995: The Meuse flood in January 1995. *Weather*, **51-2**, 39–45.

Miller, M., Hortal, M., Jakob, C., 1995: A major operational forecast model change. *ECMWF Newsletter*, **70**, 2–8.

Misra, V., Yau, M.K., 2001: An ensemble strategy for high-resolution regional model forecasts. *Meteor. Atmos. Phys.*, **78**, 61–74.

Molteni, F., Buizza, R., Palmer, T.N., Petroliagis, T., 1996: The ECMWF ensemble prediction system: methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119.

Molteni, F., Buizza, R.,Marsigli, C., Montani, A., Nerozzi, F., Paccagnella, T., 2001: A strategy for high-resolution ensemble prediction. I: Definition of representative members and global-model experiments. *Quart. J. Roy. Meteor. Soc.*, **127**, 2069–2094.

Nielsen, N.W., Amstrup, B., Jørgensen, J.U., 1998: HIRLAM 2.5 parallel tests at DMI: Sensitivity to type of schemes for turbulence, moist processes and advection. *DMI Sci. Rep.*, **98-01**, Available from the Swedish Meteorological and Hydrological Institute (SMHI) Norrköping, Sweden.

Ott, E., 1993: Chaos in Dynamical Systems. Cambridge University Press, 397 pp.

Pielke, R.A.; 2001: Mesoscale Meteorological Modeling. Academic Press, 676 pp.

Redelsperger, J.L., Sommeria, G., 1986: Three-Dimensional Simulation of a Convective Storm: Sensitivity Studies on Subgrid Parameterization and Spatial Resolution. *J. Atmos. Sci.*, **43**, 2619–2635.

Sarvijärvi, H., 1990: Fast Radiation parameterization schemes for Mesoscale and Short-Range Forecast Models. *J. Appl. Meteor.*, **29**, 437–447.

Sass, B.H., 1997: Reduction of numerical noise connected to the parameterization of cloud and condensation processes in the HIRLAM model. *HIRLAM Newsletter*, **29**, 37–45. Available from the Swedish Meteorological and Hydrological Institute (SMHI) Norrköping, Sweden.

Sass, B.H., Nielsen, N.W., Jørgensen, J.U., Amstrup, B., Kmit, M., Mogensen, K.S., 2002: The operational DMI-HIRLAM system 2002-version. *DMI Tech. Rep.*, **02-05**. Available from the Danish Meteorological Institute (DMI) Copenhagen, Denmark.

Sass, B.H. 2002: A research version of the STRACO cloud scheme. *DMI Tech. Rep.*, **02-10**. Available from the Danish Meteorological Institute (DMI) Copenhagen, Denmark.

Sattler, K.; 2002. Precipitation hindcasts of historical flood events. *DMI Sci. Rep.*, **02-13**, Available from the Danish Meteorological Institute (DMI) Copenhagen, Denmark.

Stanski, H.R., Wilson, L.J., Burrows, R., 1989: Survey of common verification methods in meteorology. World Weather Watch Tech. Rep. 8, WMO/TD No. 358, WMO, Geneva, Switzerland, 114 pp.

Stensrud, D.J., Brooks, H.E., Du, J., Tracton, S., Rogers, E., 1999: Using ensembles for short-range forecasting. *Mon. Wea. Rev.*, **127**, 433–446.

Stensrud, D.J., Bao, J.-W., Warner, T.T., 2000: Using initial condition and model physics perturbations in short-range ensembles. *Mon. Wea. Rev.*, **128**, 2077–2107.

Strauss, B. and Lanzinger, A., 1996: Validation of the ECMWF ensemble prediction system. *Proceedings of ECMWF seminar on predictability*, **Vol. II**, 4-8 September 1995, 157–166.

Sundqvist, H., Berge, E., Kristjánsson, J.E., 1989: Condensation and Cloud Parameterization Studies with a Mesoscale Numerical Weather Prediction Model. *Mon. Wea. Rev.*, **117**, 1641–1657.

Sundqvist, H., 1993: Inclusion of Ice Phase of Hydrometeors in Cloud Parameterization for Mesoscale and Large Scale Models. *Beitr. Phys. Atmos.*, **66**, 137–147.

Thompson, P.D., 1957: Uncertainty of Initial State as a Factor in the Predictability of Large Scale Atmospheric Flow Patterns. *Tellus*, **4**, 275–295.

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

Tiedtke, M., 1993: Representation of Clouds in Large-Scale Models. *Mon. Wea. Rev.*, **121**, 3040–3061.

Toth, Z. and Kalnay, E., 1993: Ensemble Forecasting at NMC: The Generation of Perturbations. *Bull. Am. Meteor. Soc.*, **74**, 2317–2330.

Ulbrich, U. and Fink, A., 1995: The January 1995 Flood in Germany: Meteorological Versus Hydrological Causes. *Phys. Chem. Earth*, **20**, 439–444.

Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Ereola, K., Fortelius, C., Garçia-Moya, J.A., Jones, C., Lenderlink, G., McDonald, A., McGrath, R., Navascues, B., Nielsen, N.W., Ødegaard, V., Rodriguez, E., Rummukainen, M., Rõõm, R., Sattler, K., Sass, B.H., Savijärvi, H., Schreur, B.W., Sigg, R., The, H., Tijm, S., 2002. *HIRLAM-5 scientific documentation.* Available from the Swedish Meteorological and Hydrological Institute (SMHI) Norrköping, Sweden.

Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences, Chapter 7.4.* Academic Press, San Diego. 467pp.

Wilks, D.S., 2002: Smoothing forecast ensembles with fitted probability distributions. *Quart. J. Roy. Meteor. Soc.*, **128**, 2821–2836.

DANISH METEOROLOGICAL INSTITUTE

Scientific Reports

Scientific reports from the Danish Meteorological Institute cover a variety of geophysical fields, i.e. meteorology (including climatology), oceanography, subjects on air and sea pollution, geomagnetism, solar-terrestrial physics, and physics of the middle and upper atmosphere.

Reports in the series within the last five years:

No. 99-1
Henrik Feddersen: Project on prediction of climate variations on seasonel to interannual timescales (PROVOST) EU contract ENVA4-CT95-0109: DMI contribution to the final report: Statistical analysis and post-processing of uncoupled PROVOST simulations

No. 99-2
Wilhelm May: A time-slice experiment with the ECHAM4 A-GCM at high resolution: the experimental design and the assessment of climate change as compared to a greenhouse gas experiment with ECHAM4/OPYC at low resolution

No. 99-3
Niels Larsen et al.: European stratospheric monitoring stations in the Artic II: CEC Environment and Climate Programme Contract ENV4-CT95-0136. DMI Contributions to the project

No. 99-4
Alexander Baklanov: Parameterisation of the deposition processes and radioactive decay: a review and some preliminary results with the DERMA model

No. 99-5
Mette Dahl Mortensen: Non-linear high resolution inversion of radio occultation data

No. 99-6
Stig Syndergaard: Retrieval analysis and methodologies in atmospheric limb sounding using the GNSS radio occultation technique

No. 99-7
Jun She, Jacob Woge Nielsen: Operational wave forecasts over the Baltic and North Sea

No. 99-8
Henrik Feddersen: Monthly temperature forecasts for Denmark - statistical or dynamical?

No. 99-9
P. Thejll, K. Lassen: Solar forcing of the Northern hemisphere air temperature: new data

No. 99-10
Torben Stockflet Jørgensen, Aksel Walløe Hansen: Comment on "Variation of cosmic ray flux and global coverage - a missing link in solar-climate relationships" by Henrik Svensmark and Eigil Friis-Christensen

No. 99-11
Mette Dahl Meincke: Inversion methods for atmospheric profiling with GPS occultations

No. 99-12
Hans-Henrik Benzon; Laust Olsen; Per Høeg: Simulations of current density measurements with a Faraday Current Meter and a magnetometer

No. 00-01
Per Høeg; G. Leppelmeier: ACE - Atmosphere Climate Experiment

No. 00-02
Per Høeg: FACE-IT: Field-Aligned Current Experiment in the Ionosphere and Thermosphere

No. 00-03
Allan Gross: Surface ozone and tropospheric chemistry with applications to regional air quality modeling. PhD thesis

No. 00-04
Henrik Vedel: Conversion of WGS84 geometric heights to NWP model HIRLAM geopotential heights

No. 00-05
Jérôme Chenevez: Advection experiments with DMI-Hirlam-Tracer

No. 00-06
Niels Larsen: Polar stratospheric clouds microphysical and optical models

No. 00-07
Alix Rasmussen: "Uncertainty of meteorological parameters from DMI-HIRLAM"

No. 00-08
A.L. Morozova: Solar activity and Earth's weather. Effect of the forced atmospheric transparency changes on the troposphere temperature profile studied with atmospheric models

No. 00-09
Niels Larsen, Bjørn M. Knudsen, Michael Gauss, Giovanni Pitari: Effects from high-speed civil traffic aircraft emissions on polar stratospheric clouds

No. 00-10
Søren Andersen: Evaluation of SSM/I sea ice algorithms for use in the SAF on ocean and sea ice, July 2000

No. 00-11
Claus Petersen, Niels Woetmann Nielsen: Diagnosis of visibility in DMI-HIRLAM

No. 00-12
Erik Buch: A monograph on the physical oceanography of the Greenland waters

No. 00-13
M. Steffensen: Stability indices as indicators of lightning and thunder

No. 00-14
Bjarne Amstrup, Kristian S. Mogensen, Xiang-Yu Huang: Use of GPS observations in an optimum interpolation based data assimilation system

No. 00-15
Mads Hvid Nielsen: Dynamisk beskrivelse og hydrografisk klassifikation af den jyske kyststrøm

No. 00-16
Kristian S. Mogensen, Jess U. Jørgensen, Bjarne Amstrup, Xiaohua Yang and Xiang-Yu Huang: Towards an operational implementation of HIRLAM 3D-VAR at DMI

No. 00-17
Sattler, Kai; Huang, Xiang-Yu: Structure function characteristics for 2 meter temperature and relative humidity in different horizontal resolutions

No. 00-18
Niels Larsen, Ib Steen Mikkelsen, Bjørn M. Knudsen m.fl.: In-situ analysis of aerosols and gases in the polar stratosphere. A contribution to THESEO. Environment and climate research programme. Contract no. ENV4-CT97-0523. Final report

No. 00-19
Amstrup, Bjarne: EUCOS observing system experiments with the DMI HIRLAM optimum interpolation analysis and forecasting system

No. 01-01
V.O. Papitashvili, L.I. Gromova, V.A. Popov and O. Rasmussen: Northern polar cap magnetic activity index PCN: Effective area, universal time, seasonal, and solar cycle variations

No. 01-02
M.E. Gorbunov: Radioholographic methods for processing radio occultation data in multipath regions

No. 01-03
Niels Woetmann Nielsen; Claus Petersen: Calculation of wind gusts in DMI-HIRLAM

No. 01-04
Vladimir Penenko; Alexander Baklanov: Methods of sensitivity theory and inverse modeling for estimation of source parameter and risk/vulnerability areas

No. 01-05
Sergej Zilitinkevich; Alexander Baklanov; Jutta Rost; Ann-Sofi Smedman, Vasiliy Lykosov and Pierluigi Calanca: Diagnostic and prognostic equations for the depth of the stably stratified Ekman boundary layer

No. 01-06
Bjarne Amstrup: Impact of ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-Var analysis and forecasting system

No. 01-07
Sergej Zilitinkevich; Alexander Baklanov: Calculation of the height of stable boundary layers in operational models

No. 01-08
Vibeke Huess: Sea level variations in the North Sea – from tide gauges, altimetry and modelling

No. 01-09
Alexander Baklanov and Alexander Mahura: Atmospheric transport pathways, vulnerability and possible accidental consequences from nuclear risk sites: methodology for probabilistic atmospheric studies

No. 02-01
Bent Hansen Sass and Claus Petersen: Short range atmospheric forecasts using a nudging procedure to combine analyses of cloud and precipitation with a numerical forecast model

No. 02-02
Erik Buch: Present oceanographic conditions in Greenland waters

No. 02-03
Bjørn M. Knudsen, Signe B. Andersen and Allan Gross: Contribution of the Danish Meteorological Institute to the final report of SAMMOA. CEC contract EVK2-1999-00315: Spring-to.-autumn measurements and modelling of ozone and active species

No. 02-04
Nicolai Kliem: Numerical ocean and sea ice modelling: the area around Cape Farewell (Ph.D. thesis)

No. 02-05
Niels Woetmann Nielsen: The structure and dynamics of the atmospheric boundary layer

No. 02-06
Arne Skov Jensen, Hans-Henrik Benzon and Martin S. Lohmann: A new high resolution method for processing radio occultation data

No. 02-07
Per Høeg and Gottfried Kirchengast: ACE+: Atmosphere and Climate Explorer

No. 02-08
Rashpal Gill: SAR surface cover classification using distribution matching

No. 02-09
Kai Sattler, Jun She, Bent Hansen Sass, Leif Laursen, Lars Landberg, Morten Nielsen og Henning S. Christensen: Enhanced description of the wind climate in Denmark for determination of wind resources: final report for 1363/00-0020: Supported by the Danish Energy Authority

No. 02-10
Michael E. Gorbunov and Kent B. Lauritsen: Canonical transform methods for radio occultation data

No. 02-11
Kent B. Lauritsen and Martin S. Lohmann: Unfolding of radio occultation multipath behavior using phase models

No. 02-12
Rashpal Gill: SAR ice classification using fuzzy screening method

No. 02-13
Kai Sattler: Precipitation hindcasts of historical flood events

No. 02-14
Tina Christensen: Energetic electron precipitation studied by atmospheric x-rays

No. 02-15
Alexander Mahura and Alexander Baklanov: Probablistic analysis of atmospheric transport patterns from nuclear risk sites in Euro-Arctic Region

No. 02-16
A. Baklanov, A. Mahura, J.H. Sørensen, O. Rigina, R. Bergman: Methodology for risk analysis based on atmospheric dispersion modelling from nuclear risk sites

No. 02-17
A. Mahura, A. Baklanov, J.H. Sørensen, F. Parker, F. Novikov K. Brown, K. Compton: Probabilistic analysis of atmospheric transport and deposition patterns from nuclear risk sites in russian far east

No. 03-01
Hans-Henrik Benzon, Alan Steen Nielsen, Laust Olsen: An atmospheric wave optics propagator, theory and applications

No. 03-02
A.S. Jensen, M.S. Lohmann, H.-H. Benzon and A.S. Nielsen: Geometrical optics phase matching of radio occultation signals

No. 03-03
Bjarne Amstrup, Niels Woetmann Nielsen and Bent Hansen Sass: DMI-HIRLAM parallel tests with upstream and centered difference advection of the moisture variables for a summer and winter period in 2002

No. 03-04
Alexander Mahura, Dan Jaffe and Joyce Harris: Identification of sources and long term trends for pollutants in the Arctic using isentropic trajectory analysis

No. 03-05
Jakob Grove-Rasmussen: Atmospheric Water Vapour Detection using Satellite GPS Profiling

No. 03-06
Bjarne Amstrup: Impact of NOAA16 and NOAA17 ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-VAR analysis and forecasting system - January and February 2003

No. 03-07
Kai Sattler and Henrik Feddersen: An European Flood Forecasting System EFFS. Treatment of uncertainties in the prediction of heavy rainfall using different ensemble approaches with DMI-HIRLAM

No. 03-08
Peter Thejll and Torben Schmith: Limitations on regression analysis due to serially correlated residuals: Application to climate reconstruction from proxies

No. 03-09
Peter Stauning, Hermann Lühr, Pascale Ultré-Guérard, John LaBrecque, Michael Purucker, Fritz Primdahl, John L. Jørgensen, Freddy Christiansen, Per Høeg, Kent B. Lauritsen: OIST-4 Proceedings. 4'th Oersted International Science Team Conference. Copenhagen 23-27 September 2002

No. 03-10
Niels Woetmann Nielsen: A note on the sea surface momentum roughness length.

No. 03-11
Niels Woetmann Nielsen: Quasigeostrophic interpretation of extratropical cyclogenesis