
Flood forecasting using medium-range probabilistic weather prediction

B.T. Gouweleeuw¹, J. Thielen¹, G. Franchello¹, A.P.J. De Roo¹ and R. Buizza²

¹NASA/GSFC, Mail Code 614.3, Greenbelt, MD 20770, USA

²European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, UK

Email for corresponding author: bingo@hsb.gsfc.nasa.gov

Abstract

Following the developments in short- and medium-range weather forecasting over the last decade, operational flood forecasting also appears to show a shift from a so-called single solution or 'best guess' deterministic approach towards a probabilistic approach based on ensemble techniques. While this probabilistic approach is now more or less common practice and well established in the meteorological community, operational flood forecasters have only started to look for ways to interpret and mitigate for end-users the prediction products obtained by combining so-called Ensemble Prediction Systems (EPS) of Numerical Weather Prediction (NWP) models with rainfall-runoff models. This paper presents initial results obtained by combining deterministic and EPS hindcasts of the global NWP model of the European Centre for Medium-Range Weather Forecasts (ECMWF) with the large-scale hydrological model LISFLOOD for two historic flood events: the river Meuse flood in January 1995 and the river Odra flood in July 1997. In addition, a possible way to interpret the obtained ensemble based stream flow prediction is proposed.

Keywords: probabilistic weather prediction, flood forecasting, ensemble based stream flow prediction

Introduction

The information provided to end-users in operational flood forecasting is typically deterministic. Prompted by the use of probabilistic weather predictions in river flow forecasting, operational flood forecasters face the challenge to interpret and mitigate to end-users an ensemble-based stream flow prediction rather than a single solution product. Although concepts that provide a framework for probabilistic flood forecasting do exist (Krzysztofowicz, 1999; Kelly and Krzysztofowicz, 2000; Krzysztofowicz and Kelly, 2000), these prototypes have not yet found many applications in practical everyday flood forecasting. Within a framework of Bayesian forecasting systems (BFSs), Krzysztofowicz (2002) proposed a methodology to produce a probabilistic river stage forecast for the short-range term. The system decomposes the total forecasting uncertainty into input uncertainty and hydrological uncertainty. The former source of uncertainty is dominated by the unknown future rainfall; the latter generally aggregates all other uncertainties and

includes model, parameter, estimation and measurement errors. Maranzano and Krzysztofowicz (2004) developed a second BFS to produce a probabilistic stage transition forecast for the short-range term, and found for their test case-study that the influence of the hydrological model on river stage transitions is restricted to 48 hrs, given a 24-hr probabilistic rain forecast input. Beyond this limit the forecast is essentially statistical, estimated from climatic data and the meteorological conditions.

The exercise presented in this paper deviates from the BSF approach described by Maranzano and Krzysztofowicz (2004) in the sense that it addresses only the input uncertainty by simply feeding probabilistic weather forecasts extended beyond the short-range term to a deterministic hydrological model. The hydrological uncertainty associated with this exercise is described by Pappenberger *et al.* (2004). This paper presents the computation of deterministic and ensemble-based river flow predictions by forcing the deterministic weather forecast and the forecasts derived by

the Ensemble Prediction System (EPS) of the global atmospheric model of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Molteni *et al.*, 1996) through the distributed grid-based hydrological model LISFLOOD (De Roo, 1999; De Roo *et al.*, 2000, 2001). This work has been carried out within the European Commission 5th Framework Program (FP5) European Flood Forecasting System (EFFS) project, which aimed to test the feasibility of a prototype methodology for a flood forecasting system on a European scale by combining probabilistic weather forecasts in the short- and medium-range (≤ 10 days) and a large scale rainfall-runoff model (De Roo *et al.*, 2003; Gouweleeuw *et al.*, 2004). Results are presented for two historic flood events for which hindcasts are performed, i.e. the Meuse flood in January 1995 and the Odra flood in July 1997. In addition, a classification scheme is presented, based on probabilities of exceedence of a certain threshold level. These thresholds can reflect so-called 'critical' discharge levels, e.g. an alert level for overtopping of dykes. This would provide very useful probabilistic information to national river authorities in decision making, e.g. to evaluate cost-damage functions (Dutta *et al.*, 2003), in addition to the commonly used deterministic or single solution forecast.

Description of meteorological and hydrological model

ECMWF GLOBAL ATMOSPHERIC MODEL

A detailed description of the ECMWF NWP model can be found in Molteni *et al.* (1996), Persson (2001) and Buizza *et al.* (2001). This section provides a brief overview of the elements relevant to the present study.

The latest version of the ECMWF general circulation model, encoded T_L511L60, provides a deterministic weather forecast of up to 10 days. The encoding relates to the vertical and horizontal resolution of the model. Based on a spherical harmonic expression and truncated at total wave number 511, the horizontal resolution equates to an approximate 40 km grid spacing at mid-latitudes. The atmosphere is divided into 60 layers up to 0.1 hPa (about 64 km). The vertical resolution varies with height and is based on surfaces of constant pressure which follow the Earth's surface in the lower and mid-troposphere. It is highest in the planetary boundary layer and lowest in the stratosphere and lower mesosphere. The EPS is provided by the T_L250L40 model and consists of 50 perturbed and one unperturbed (or control) 10-day weather forecasts. The horizontal grid spacing equates to an approximate 80 km resolution and 40 layers in the vertical up to 10 hPa.

The 50 perturbed forecasts of the ensemble system result

from a perturbation technique, which is based on a mathematical method called singular vector analysis. It attempts to locate the dynamically most unstable regions of the atmosphere by calculating where small initial uncertainties would affect a 48-hour forecast most rapidly (Buizza and Palmer, 1995). From these, 50 alternative forecasts are produced. Additionally, the ESP includes a simulation of random model error due to parameterised physical processes, i.e. so-called stochastic physics (Buizza *et al.*, 1999). Results have previously shown that both stochastic physics and a higher resolution (from T_L159L40, i.e. ~ 120 km resolution to T_L255L40, i.e. ~ 80 km resolution) had a measurable positive impact on precipitation prediction (Mullen and Buizza, 2001). One forecast, the 'control', is run with a non-perturbed analysis, but at the low resolution. All the different initial states are *a priori* assumed to be equally likely, i.e. the unperturbed forecast is not necessarily the most probable.

For the present study only a selection of the model output parameters at surface level is used. These parameters are large scale precipitation, convective precipitation, snowfall, dry air temperature at a height of 2 m and evaporation (deterministic forecast mode) or surface latent heat flux (ensemble forecast mode). These parameters are stored for each grid cell every 6 hours up to 10 days ahead in either mode.

LISFLOOD RAINFALL-RUNOFF MODEL

LISFLOOD is a distributed, raster-based, combined rainfall-runoff and hydrodynamic model embedded in a dynamic GIS environment (De Roo, 1999; De Roo *et al.*, 2000, 2002). It has been developed as a flexible tool for the simulation of hydrological processes and floods, in particular in large European river basins, on a wide range of temporal and spatial scales. Spatial scales may vary from grid sizes of 100 m for small catchments (< 100 km²) to 5 km on a European scale. The temporal model flexibility allows for a so-called time-cascading structure: the water balance of a catchment may be simulated with a daily time step for a (1 year) pre-flood period, and its output then serves as initial condition for a flood simulation with an hourly time step.

In the vertical, the model considers two soil layers and two groundwater layers. The river channel network is derived by combining digital elevation and actual river network data.

Model inputs are meteorological data such as precipitation, temperature and evaporation. If the latter is not available, it may be calculated from other available meteorological data. A height dependent temperature correction is carried out, if the spatial resolution is finer than the resolution of the

meteorological data, which is usually the case. Further, data layers such as topography, land use, LAI, soil type, soil texture, soil depth are required to define the river basin.

The model simulates processes at the soil–canopy–atmosphere interface, e.g. separation of precipitation in snow and rainfall depending on temperature, interception and evapotranspiration. Simple routines for snowmelt and glacier melt are also included. The model considers soil freezing, impeding infiltration if a certain threshold is reached. Once infiltrated, water is transported vertically and horizontally between cells and it percolates from the bottom soil layer to the upper groundwater layer. Eventually all water is transferred to the channel network where it is routed with a either a dynamic or a kinematic wave, dependent on

the presence of backwater effects and the bed slope gradient.

Numerous outputs are available from the system and may include maps of time series of any variable computed in the model, e.g. maps of soil moisture content, predicted flow and stage hydrographs for any point on the drainage network, flood source areas and estimates of groundwater recharge. While LISFLOOD simulates surface processes using physically-based algorithms, the subsurface flow modelling routine is more conceptual. It requires little additional calibration for individual catchments, mainly on the groundwater parameters. Figure 1 shows a flowchart of the model, visualising the main hydrological processes outlined above.

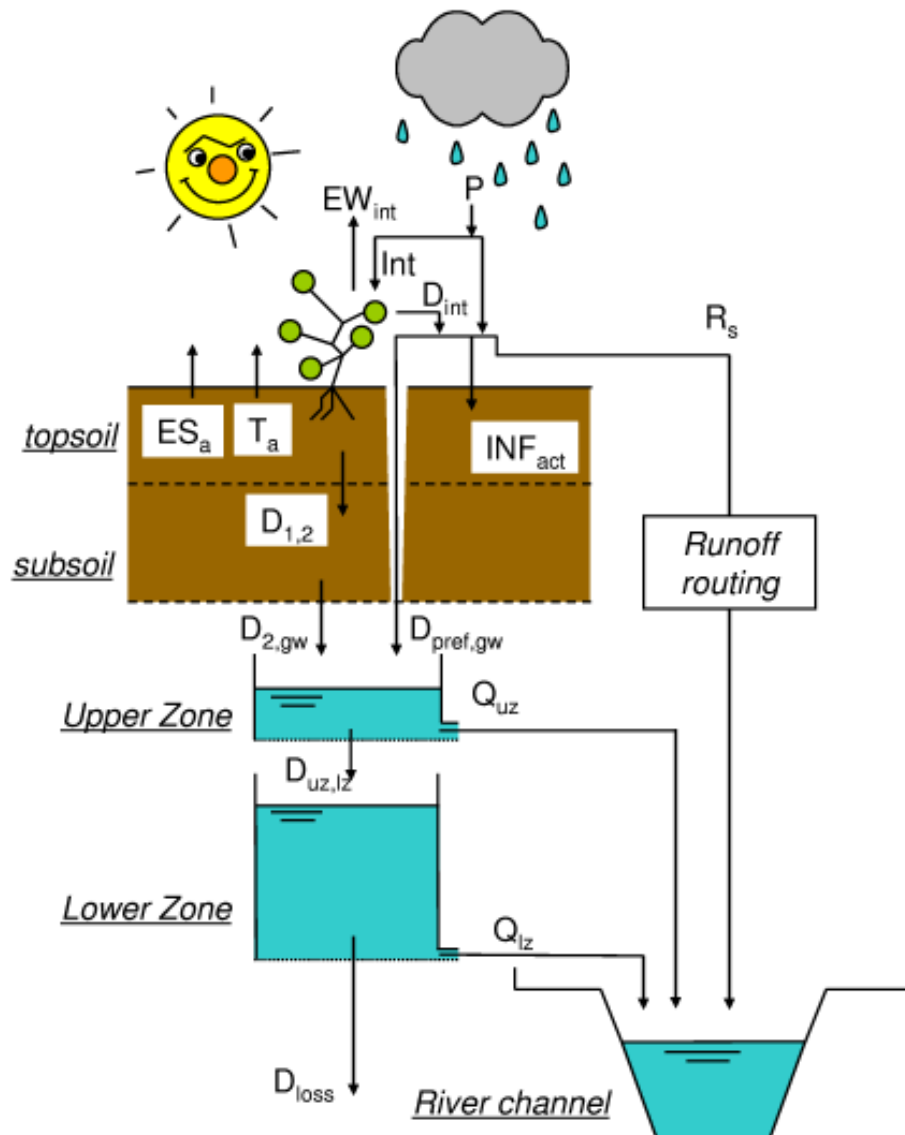


Fig. 1. Flowchart of the LISFLOOD model, showing the key processes included.

The January 1995 Meuse flood

INTRODUCTION

Late January 1995 floods along the rivers Meuse and Rhine were induced by a chain of frontal systems passing over a relatively flat area causing a number of medium intensity rainfall events over several weeks (Sattler and Feddersen, 2003). Snowmelt also contributed to the flood (Ulbrich and Fink, 1995). In the Meuse downstream reaches, a maximum discharge close to $3000 \text{ m}^3 \text{ s}^{-1}$ was recorded, coinciding with an estimated return period of 63 years.

For the River Meuse catchment (Fig. 2), the LISFLOOD model has been run on a 1 km grid for a 21.128 km^2 total area upstream of the Borgharen gauging station. Calibrated on the December 1993 flood for four stations, the model has been initialised for the 1995 flood with the water balance version (i.e. LISFLOOD-WB). These initial conditions were computed with a daily time step for two years using 57 stations of daily rainfall and 33 stations with daily

meteorological variables as model boundary conditions. Hindcast simulations using the flood forecast version (i.e. LISFLOOD-FF) commenced at 12:00 h UTC on January 20, 1995 and were subsequently updated for 10 consecutive days to January 30 1995 for the observed high resolution meteorological data, the deterministic forecast, the ensemble control and each ensemble member.

While a 1-km resolution Digital Elevation Model (DEM) is available for the Meuse river basin, a 75-m DEM is used, together with GIS data on actual rivers, to derive the flow network. A 1-km resolution land use map is based on the CORINE data with sub-grid information available based on 100-m resolution maps for urban and forest coverage. Information on soil depth and soil texture on a 1-km resolution is obtained from the 1:250.000 scale soil data base of the Meuse (source: European soil database, JRC). The saturated hydraulic conductivity and soil water content at saturation is based on the HYPRES database (Wösten *et al.*, 1998). Leaf area index and vegetated soil cover are derived from IRS-WiFS satellite images.

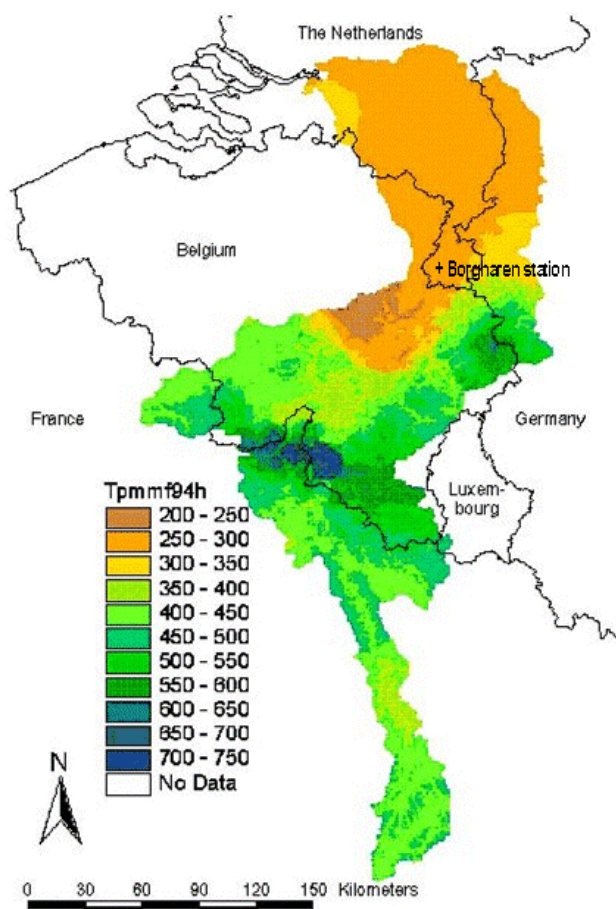


Fig. 2. Map of the river Meuse basin, showing location of Borgharen gauging station together with the accumulated observed precipitation over the period December 1994-January 1995.

RESULTS

Figure 3 illustrates for January 20, seven days before the start of the major flood peak, a complete set of the basin-averaged accumulated precipitation of the deterministic forecast (red curve), the ensemble control run (green) and all ensemble members (thin black) for the Meuse catchment upstream of Borgharen station, located at the outlet of the modelled river basin. For comparison, the accumulated observed precipitation interpolated from a high density rain gauge network (57 stations) is plotted as a blue line. The deterministic rainfall forecast overestimates the observed precipitation from day 3 to 5, after which it starts to underestimate. The ensemble control precipitation forecast is typically lower than the deterministic forecast, caused by its lower model resolution. A few ensemble members exceed the deterministic forecast and even fewer members exceed the observed precipitation. The majority of ensemble members, however, forecast a lower rainfall total.

Figure 4a illustrates for January 20 a complete set of the predicted flow for the deterministic forecast (red curve), the ensemble control run (green) and all ensemble members (thin black) for the Borgharen station. For comparison, the hourly observed discharge is shown as a dark blue bold line, and the bright blue curve shows the LISFLOOD-FF simulation when using high-resolution observed meteorological data. The observed flow rate curve shows a minor peak of about $2000 \text{ m}^3 \text{ s}^{-1}$ at January 24. Then, on January 27, a first major peak of $2800 \text{ m}^3 \text{ s}^{-1}$ is observed, after which the flow remains above $2500 \text{ m}^3 \text{ s}^{-1}$ until

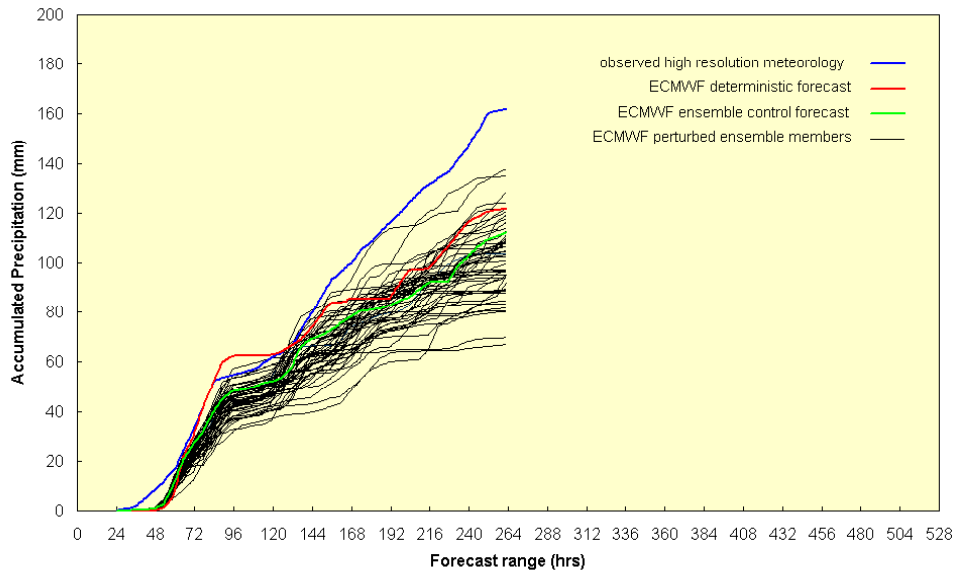


Fig. 3. 10-day accumulated basin-averaged precipitation forecasts of the ECMWF atmospheric global circulation model for January 20 1995 at 12:00 UTC (hour 24) over the Meuse catchment, The Netherlands.

February 1. A second major peak is registered on January 31.

The results show that LISFLOOD-FF is capable of simulating the observed flow at Borgharen when forced with observed meteorological data with an average relative RMSE of 0.10. It should be noted that these results are obtained without updating the flow. In terms of pre-warning, the deterministic forecast ($T_{L511L60}$) on January 20 produces an overestimation up to day 6, caused by the overestimation in the forecasted rainfall from day 3-5 (Fig. 3). The over-estimation in precipitation, however, seems to trigger a relatively large flow over-estimation. In addition, some ensemble members over-estimate the deterministic flow forecast and the simulated flow forced with the high resolution observed meteorology, while hardly any over-estimation is present in the forecasted accumulated rainfall. One possible explanation for this is that an erroneous spreading of the predicted rainfall event over the Meuse river basin could have a significant effect on the streamflow forecast, while this is not reflected in the basin-averaged accumulated precipitation. After day 5 the accumulated deterministic rainfall forecast starts to under-estimate the accumulated observed rainfall resulting in a slightly delayed under-prediction of the forecasted flow from day 6 on. The control ensemble produces a lower rainfall estimation due to its lower model resolution ($T_{L255L40}$) and is therefore more accurate during the first six forecast days in predicting the first minor peak on January 24. The flow forecast forced by the ensemble predictions then follows more or less the same tendency, showing some over-estimation up to day 6 and almost exclusively underestimation after six days.

However, a few of the high extreme ensemble members reach or even exceed the observed discharge, indicating at least the possibility of an extreme event in the medium term forecast range.

The spread in the EPS-predicted flow reaches a maximum on forecast day 3, after which it decreases to remain more or less constant. This is shown in Fig 4b, which presents, for the same forecast day, the cumulative distribution represented by the median and the 0.05 and 0.95 percentile boundaries at each time step. The ensemble spread is then illustrated by the difference between the 0.05 and 0.95 percentile boundaries. The low ensemble spread in the first one to two days is likely due to the fact that the river basin did not have enough time to respond fully to the forecast forcing and is still influenced mainly by relatively low past rainfall. Further, a low spread should imply that the control or ensemble mean forecast will be good, while a larger spread implies a greater forecast uncertainty, possibly indicating unpredictable or unstable atmospheric conditions and an increased probability of an extreme event. As such, the maximum spread in the short forecast range (peak on forecast day 3) could be considered as an indication of the first minor flood peak on January 24. However, the lower spread at a lower flow level in the medium forecast range, i.e. beyond the response time of the catchment, would then suggest a higher probability of an extreme event not happening at the time of the second major peak. Figure 4b further shows that the upper half of the accumulated distribution brackets the increased observed flow from forecast day 3-6 and gives an indication of the first minor peak on January 24. However, the major flow peak starting

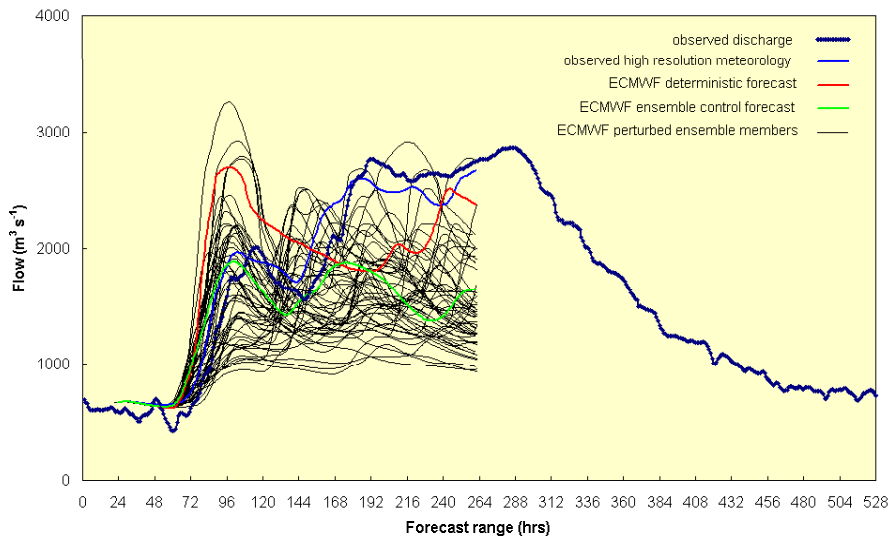


Fig. 4a. 10-day flow forecasts from the LISFLOOD-FF model for January 20 1995 at 12:00 UTC (hour 24) for the Borgharen gauging station on the river Meuse, The Netherlands.

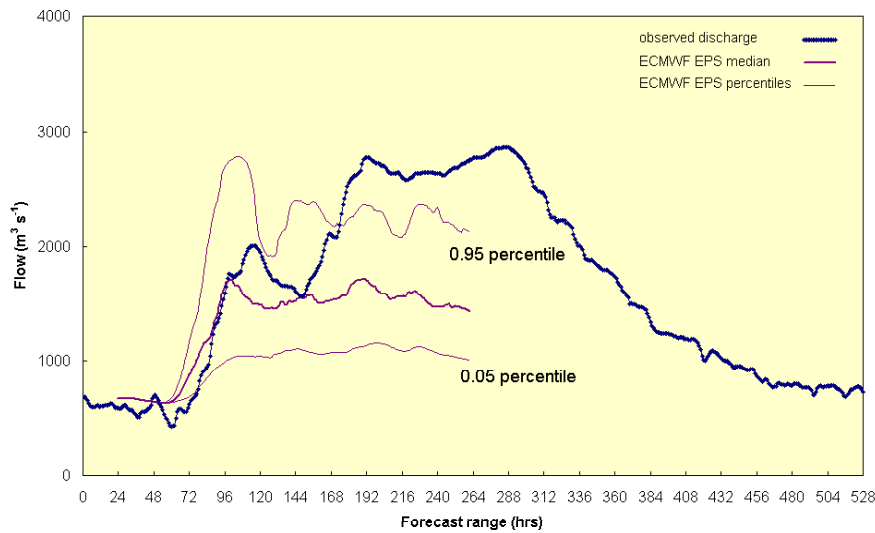


Fig. 4b. Interpretation of the 10-day flow forecast from the LISFLOOD-FF model from 12:00 UTC on January 20 1995 (hour 24) for the Borgharen gauging station on the river Meuse, the Netherlands driven by the ECMWF EPS.

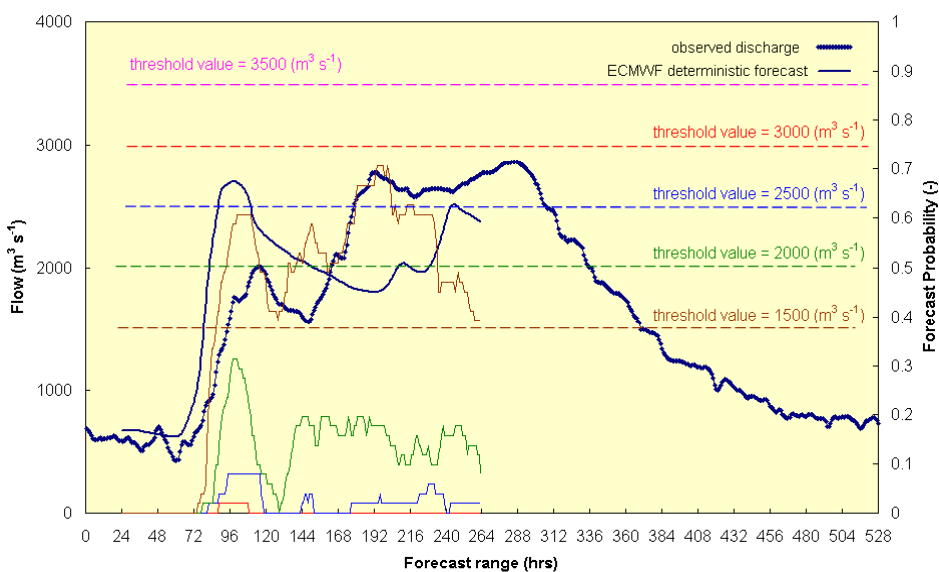


Fig. 4c. Interpretation of the 10-day flow forecast from the LISFLOOD-FF model from 12:00 UTC on January 20 1995 (hour 24) for the Borgharen gauging station on the river Meuse, the Netherlands driven by the ECMWF EPS and deterministic forecast.

from January 25 onwards is not captured by the 0.95 percentile boundary, indicating under-estimation of the ensemble predictions in the medium term 6–10-day forecast range for this particular event. In terms of exceedence levels, the median captures an approximate threshold of $1700 \text{ m}^3 \text{ s}^{-1}$, which is too low compared to the observed flow, certainly in the medium 5–10 forecast range. The exceedence level of $2800 \text{ m}^3 \text{ s}^{-1}$ reached by the 0.95 percentile is accurate, but not in timing and duration.

An alternative way to use thresholds is to compute for each time step the probability that a threshold is exceeded. This probability is simply the proportion of ensemble members exceeding the threshold. Figure 4c shows the forecast probability for several threshold values, where the different colours of the exceedence levels correspond to the plotted forecast probabilities. It shows that from the start of the hindcast issued on January 20, the probability of exceeding the $2000 \text{ m}^3 \text{ s}^{-1}$ threshold reaches 30%. For the medium-term forecast range of 3–10 days, the exceedence probability of this threshold remains about 20%, while also some probability (up to 10%) is given for the $2500 \text{ m}^3 \text{ s}^{-1}$ threshold, i.e. a critical flood risk level in the Meuse river at the Borgharen station. While the deterministic forecast gives an underestimation after six days, the forecast probabilities of exceedence of $2000 \text{ m}^3 \text{ s}^{-1}$ and $2500 \text{ m}^3 \text{ s}^{-1}$ remain more or less at a low but constant level. This is in accordance with Mylne (2003), who reported that ensemble systems are at present able to obtain useful probabilistic skill of severe weather several days ahead, but only with low forecast probabilities. This is mainly due to their inability to resolve small-scale severe weather features on a low model resolution. Despite the low probabilities, the ensemble of flow forecasts provides additional information to the deterministic flood forecast in the medium forecast range. It provides a signal in terms of pre-warning, i.e. four days before the $2000 \text{ m}^3 \text{ s}^{-1}$ exceedence level is reached and six days before the $2500 \text{ m}^3 \text{ s}^{-1}$ threshold is exceeded.

Figures 5a–j show the updates for subsequent days up to January 30 of the cumulative distribution of the ensemble of flow predictions for the Borgharen gauging station. The spread of the ensemble of flow forecasts generally starts to widen beyond the catchment's concentration time after the short 0–3 day forecast range. Conversely, the ensemble spread is not necessarily at its maximum at the end of the 10-day forecast horizon. This is especially the case in the first days during the building up of the flood, which probably relates to underestimation in the rainfall forecast leading to a (too) early indication of the end of the flood. In the wane of the flood the ensemble spread increases towards the end, indicating a possibility of an increased flow. This increased flow was observed to some extent, only much later, on

February 21 at a flow rate of $1206 \text{ m}^3 \text{ s}^{-1}$. On the other hand, the ensemble spread diminishes from forecast day to forecast day towards the end of the flood. This may suggest that in comparing one forecast day with the other the ensemble spread is more indicative of the unpredictability of the meteorological conditions and of unstable atmospheric conditions, which may lead to an increased probability of an extreme event.

In this case study the high extreme ensemble members appear generally to provide the best flow prediction estimate, indicating a general under-prediction by the probabilistic forecasts. Hence, the forecast probability of the observed values or critical exceedence levels is low.

Figures 6a–j show the updates for subsequent days up to January 30 of the forecast probability of exceedence for the Borgharen gauging station. It illustrates that the latest forecast is not necessarily better, although on average it is. For example, the forecast of January 22 is better in the median 3–10 day forecast range, but is worse in the short 0–3 day forecast range. The forecast of January 23 is again more accurate, at least in the short forecast range. Figure 6a–j further confirms that the ensemble flow prediction provides additional information to the deterministic forecast. While the latter may underestimate the observed stream flow, probabilities of 10% and higher may still be computed from the flow ensemble in the medium 3–10 day forecast range, e.g. on January 21 for $2000 \text{ m}^3 \text{ s}^{-1}$ and $2500 \text{ m}^3 \text{ s}^{-1}$. Also, in the case of an accurate deterministic forecast or an overestimation, the flow ensemble provides a measure of confidence or probability, e.g. on January 24.

Overall, the flood wave is captured quite well by the deterministic forecast. Both the minor peak on January 24 and the major peak of January 27–31 are forecasted 3–4 days ahead. The deterministic forecast typically plots in the upper half of the cumulative distribution of the ensemble forecasts, often close to the high extreme (not shown). This confirms that the under-prediction by the probabilistic forecasts is related to its lower model resolution and improved results would be expected at a higher resolution. It is further illustrated that the computed forecast probability is very dependent on the initial flow level, e.g. at high initial discharge rates the forecast probability equals unity for the relatively low threshold values.

The July 1997 Odra flood

INTRODUCTION

The Odra flood of 1997 was triggered by complex large scale dynamic developments and convective episodes. The orography intensified rainfall over the mountainous area in

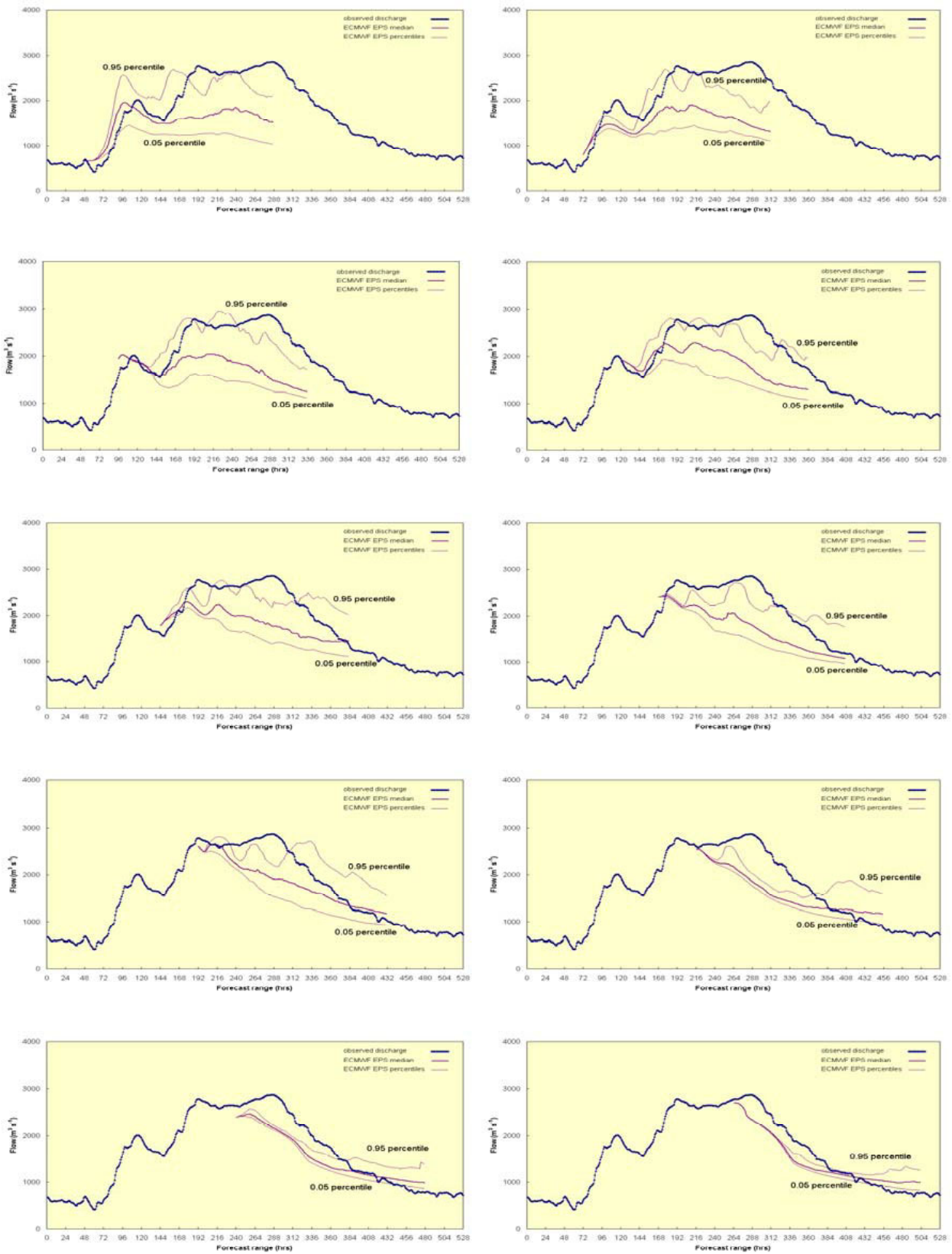


Fig. 5a–h. Subsequent updates of the interpretation of the 10-day flow forecast from the LISFLOOD-FF model for consecutive days from 12:00 UTC on January 21 (hour 48) to 12:00 UTC on January 30 1995 (hour 240) for the Borgharen gauging station on the river Meuse, The Netherlands driven by the ECMWF EPS (left to right).

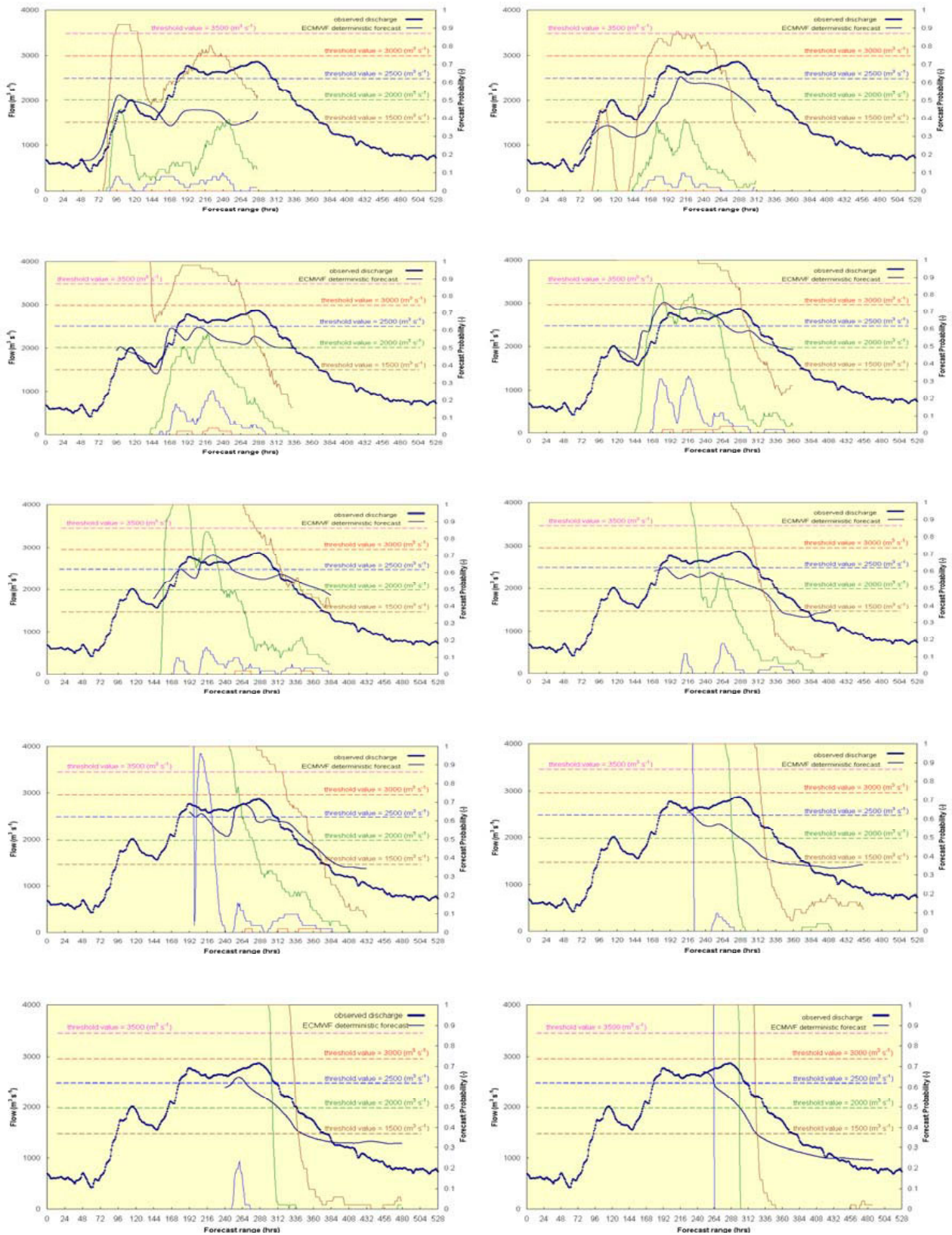


Fig. 6a–h. Subsequent updates of the interpretation of the 10-day flow forecast from the LISFLOOD-FF model for consecutive days from 12:00 UTC on January 21 (hour 48) to 12:00 UTC on January 30 1995 (hour 240) for the Borgharen gauging station on the river Meuse, The Netherlands driven by the ECMWF EPS (left to right).

the SE corner of the basin, with an observed maximum of 600 mm between July 5–10 (Keil *et al.*, 1999). In hydrological terms it was a very rare event, with a return period in the upper Odra river cross-sections in the order of several hundred years or more. Maximum discharge rates of 3260 m³ s⁻¹ were recorded at Miedonia gauging station in SE Poland.

For the upper-Odra catchment (Fig. 7), the LISFLOOD model has been run on a 1 km grid of 59.162 km² total area size upstream of the Warta confluence. Calibrated on the August 1977 and the August 1985 flood for six gauging stations (De Roo and Schmuck, 2002), the model has been initialised for the 1997 flood with LISFLOOD-WB computing with a daily time step for two years using 137 stations with hourly and daily rainfall data and 14 stations with other daily meteorological parameters as model boundary conditions. Hindcast simulations using LISFLOOD-FF commenced at 12:00 h UTC on July 1, 1997, and were updated for eight subsequent days to July 8, 1997, for the observed high resolution meteorological data, the deterministic forecast, the ensemble control and each ensemble member.

For the Odra catchment, a 1-km resolution Digital Elevation Model (DEM) is available. For the derivation of the flow network, however, a 75-m DEM is used, together with GIS data on actual rivers. A 1-km resolution land-use map is based on the CORINE data, with sub-grid information

available based on 100-m resolution maps for urban and forest coverage. Information on soil depth and soil texture on a 1-km resolution is obtained from the 1:250 000 scale soil database of the Odra (source: European soil database, JRC). The saturated hydraulic conductivity and soil water content at saturation are based on the HYPRES database (Wösten *et al.*, 1998). Leaf area index and vegetated soil cover are derived from IRS-WiFS satellite images. Additional information is available in the form of 158 river and floodplain cross-sections. Furthermore, along the German/Polish border, cross-sections were extracted at every river km from a high resolution DEM (sources: IMGW, CHMI, LUAB, Sachsen). Manning roughness coefficients were provided by IMGW and were available for 111 locations. In addition, inflow and outflow data were available for 11 water reservoirs (sources: IMGW and CHMI).

RESULTS

The analysis for the upper Odra river is performed in a similar manner to that for the river Meuse. Figure 8 illustrates for July 1 a complete set of the basin-averaged accumulated precipitation of the deterministic forecast (red curve), the ensemble control run (green) and all ensemble members (thin black) for the Odra catchment up to the Miedonia gauging station (Fig. 7). The upstream area of this station

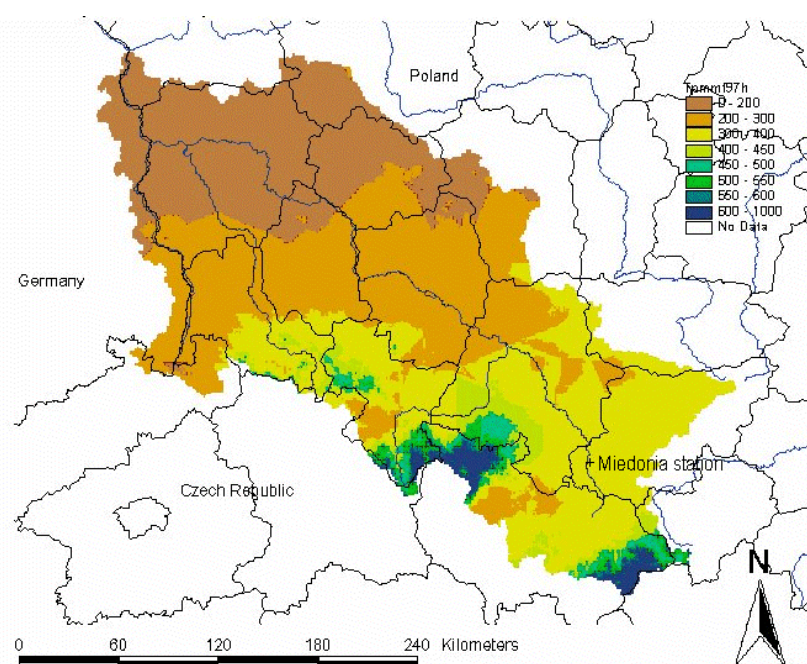


Fig. 7. Map of the river Odra basin, showing location of Miedonia gauging station together with the accumulated observed precipitation for July 1997.

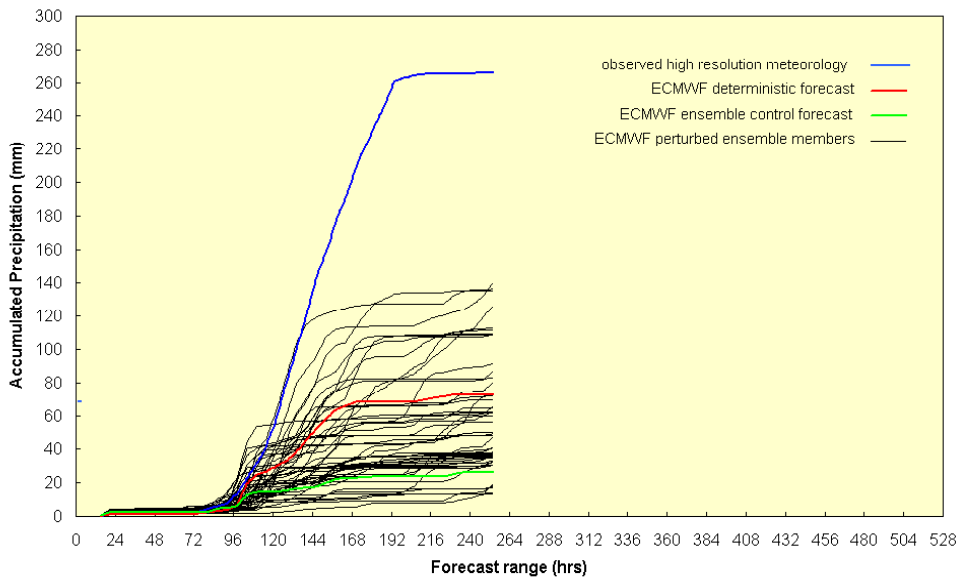


Fig. 8. 10-day accumulated basin-average precipitation forecasts from the ECMWF atmospheric global circulation model for July 1 1997 at 12:00 UTC (hour 12) over the Odra catchment up to Miedonia gauging station, SE Poland.

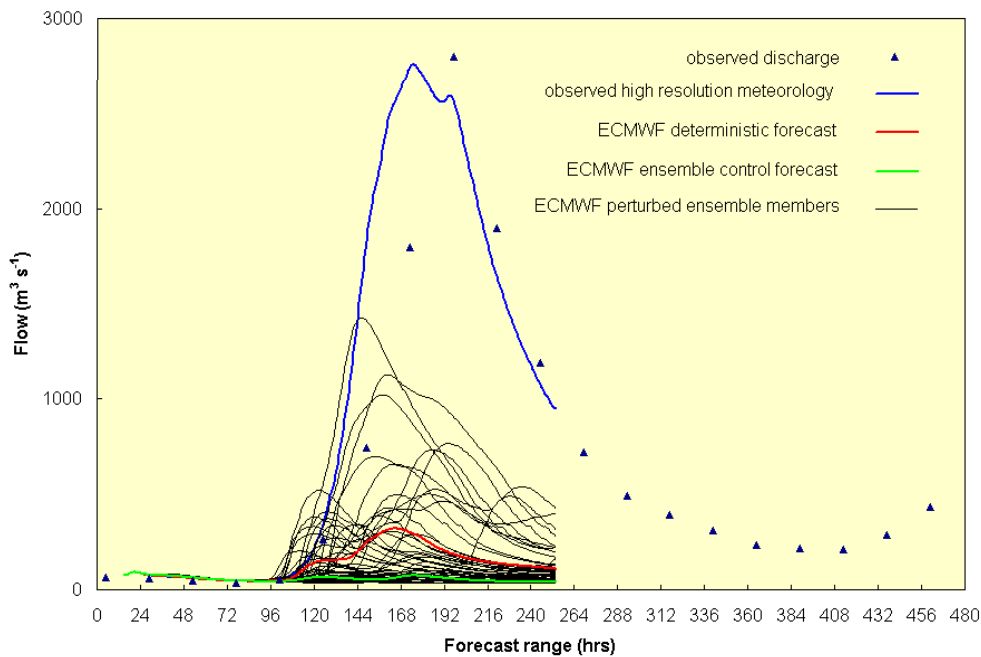


Fig. 9a. 10-day flow forecasts from the LISFLOOD-FF model for July 1 1997 at 12:00 UTC (hour 12) for the Miedonia gauging station on the river Odra, SE Poland.

was hit by heavy rainfall between July 5–8, as shown by the bright blue line that represents the accumulated observed precipitation interpolated from a high density rain gauge network (137 stations). The observed rainfall is largely under-estimated by all forecasts. However, up to day 4, some ensemble members exceed the observed precipitation. As

in the Meuse case study, the ensemble control precipitation forecast is lower than the deterministic forecast, although it plots here in the lower half of the ensemble forecasts.

Figure 9a illustrates for July 1 a complete set of the predicted flow for the deterministic forecast (red curve), the ensemble control run (green) and all ensemble members

(thin black) for the Miedonia gauging station. The observed flow remains level up to July 5, after which a distinct reaction to the rainfall event mentioned above is visible, culminating in a peak discharge of $2800 \text{ m}^3 \text{ s}^{-1}$ on July 9 (shown as dark blue triangles). The thin bright blue curve represents the LISFLOOD-FF simulation when using the high-resolution observed meteorological data. It shows that LISFLOOD-FF is capable of simulating the observed flow at Miedonia, although the peak discharge is simulated a day early. This is probably due to the storage of water in reservoirs during the early phase of the flood, which is not simulated in the

hydrological model. Again, the results are obtained without flow updating.

In terms of pre-warning, the flow forecast reflects the under-estimation observed in the accumulated precipitation forecast. The deterministic flow forecast is much too low, while the control ensemble prediction completely fails to register a signal. The ensemble flow prediction under-estimates the observed flow too, although some members over-estimate the observed flow in the first four days, following the over-prediction in the accumulated rainfall. While the deterministic forecasts provide no information

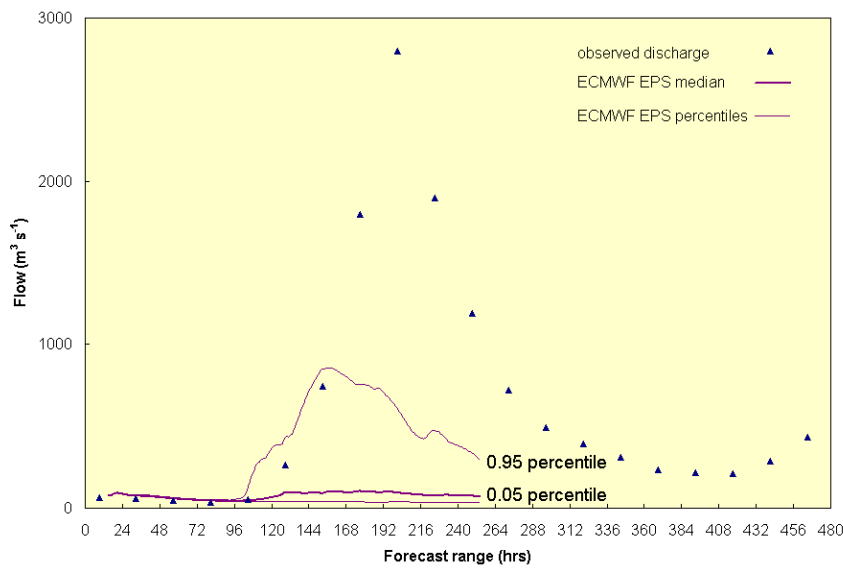


Fig. 9b. Interpretation of the 10-day flow forecast from the LISFLOOD-FF model from 12:00 UTC on July 1 1997 (hour 12) for the Miedonia gauging station on the river Odra, SE Poland driven by the ECMWF EPS.

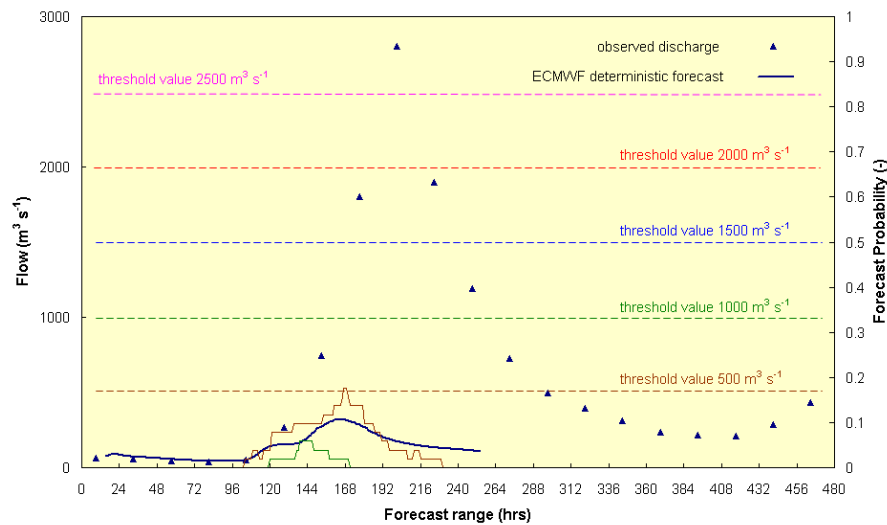


Fig. 9c. Interpretation of the 10-day flow forecast from the LISFLOOD-FF model from 12:00 UTC on July 1 1997 (hour 12) for the Miedonia gauging station on the river Odra, SE Poland, driven by the ECMWF EPS and deterministic forecast.

on the upcoming flood, one may argue that the ensemble spread, and more precisely the high extreme, contains some indication of the observed flow. This is also reflected to some extent in the cumulative distribution (Fig. 9b), although the under-estimation is substantial. Similar to the Meuse flood and taking into account the catchment's response time of 2–3 days for the discharge station

considered, the maximum spread is not at the end of the forecast range, but on day 6. The decreased ensemble spread would suggest a decreased probability of an increased flow after day 6. The early timing of the peak flow and its recession may be partly due to the imperfection of the hydrological model, as it is also observed in the model results forced with the high-resolution observed meteorological

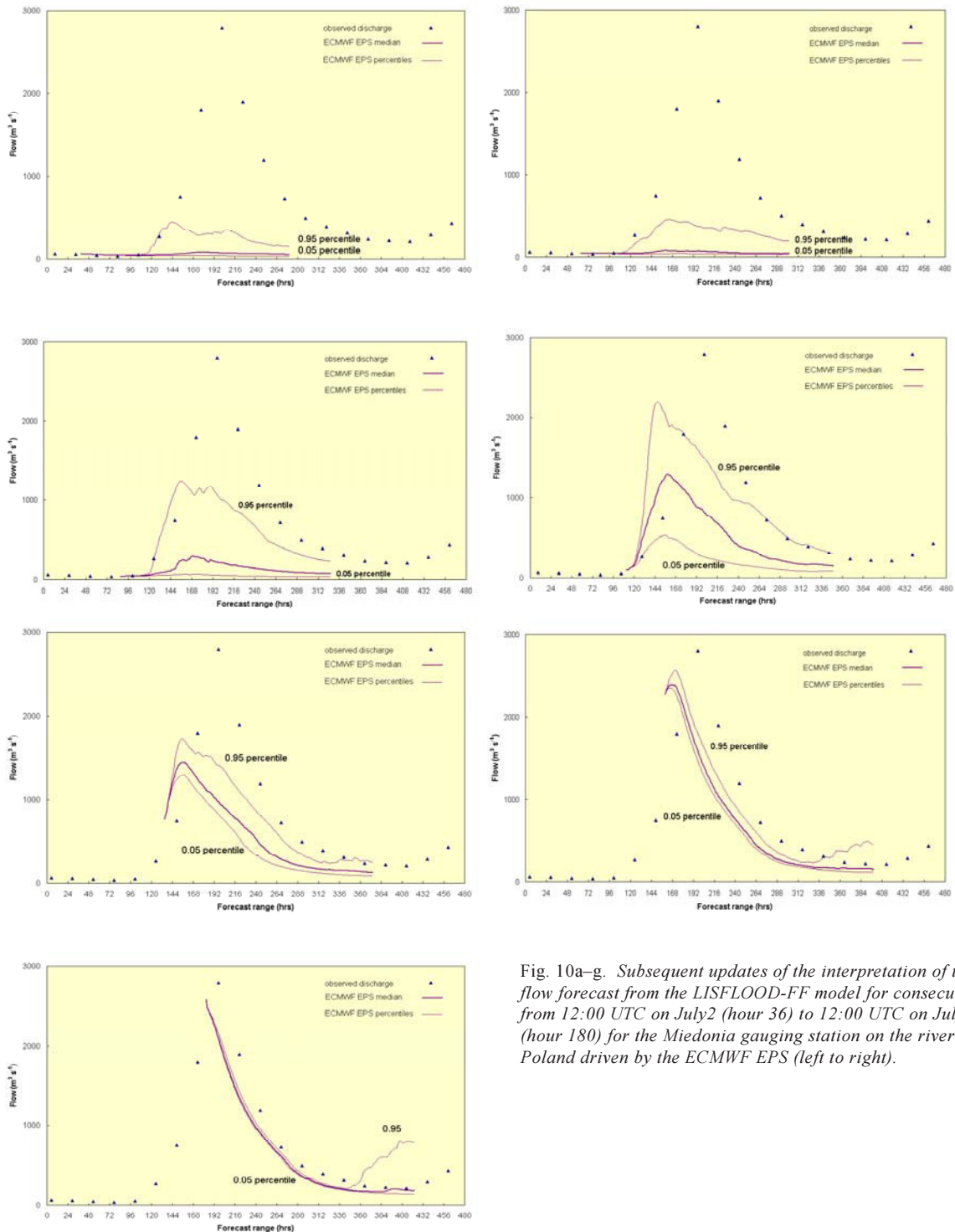


Fig. 10a–g. Subsequent updates of the interpretation of the 10-day flow forecast from the LISFLOOD-FF model for consecutive days from 12:00 UTC on July 2 (hour 36) to 12:00 UTC on July 8 1995 (hour 180) for the Miedonia gauging station on the river Odra, SE Poland driven by the ECMWF EPS (left to right).

data. Taking this effect into account, the forecasted peak flow and its recession is at least a day too early. The general under-prediction is further confirmed by the computed forecast probability (Fig. 9c).

Figures 10a-g show the subsequent updates of the performance of the ensemble of flow forecasts in terms of its cumulative distribution for July 1 to July 8 1997 at

Miedonia station. Similar to the Meuse case, the spread of the ensemble of flow forecasts generally starts to widen after the short 0–3 day forecast range, reflecting the response time for the station considered. Again, the ensemble spread is generally not at its maximum at the end of the 10-day forecast horizon, especially in the first days during the building up of the flood. This may again relate to under-

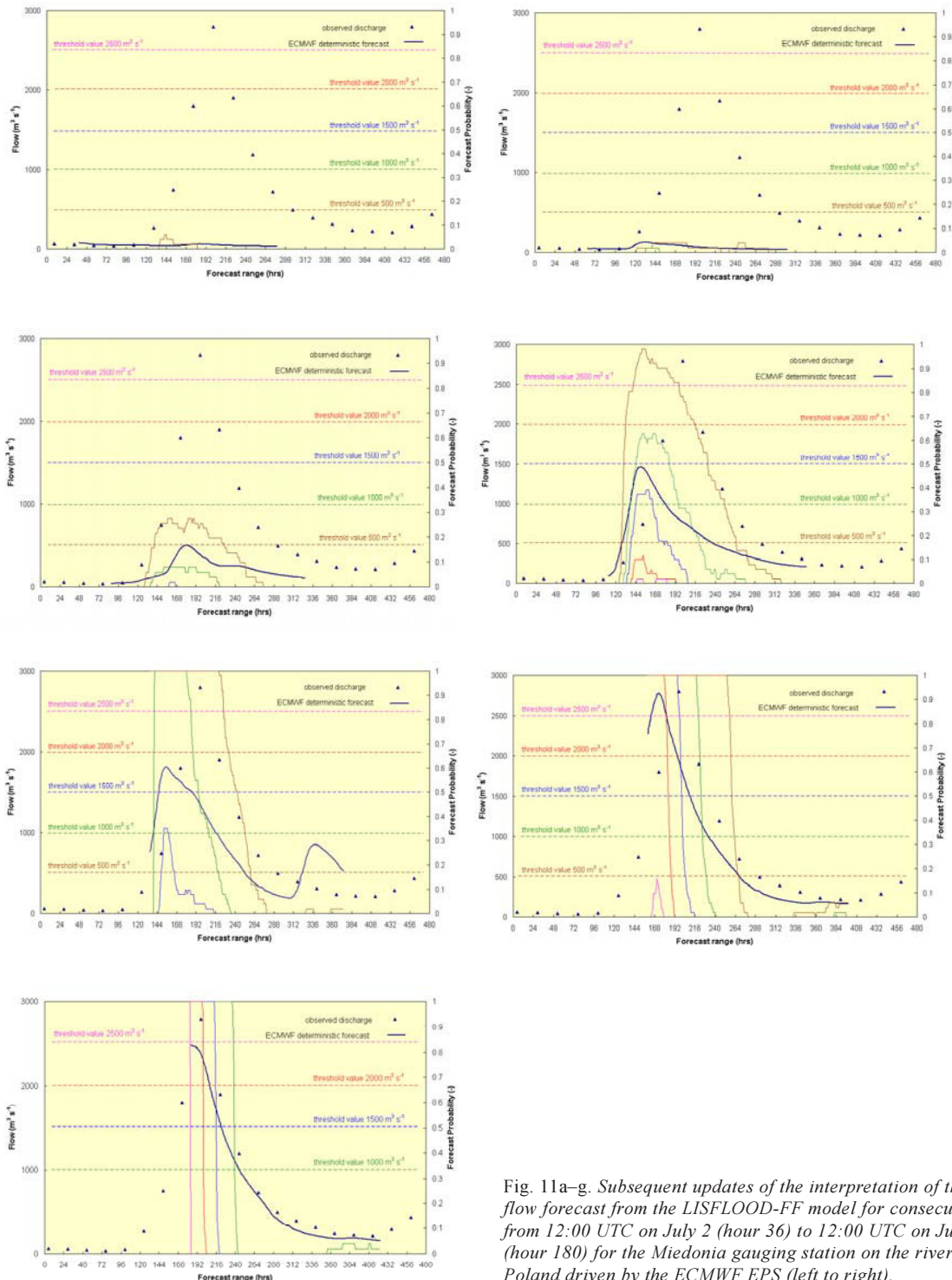


Fig. 11a–g. Subsequent updates of the interpretation of the 10-day flow forecast from the LISFLOOD-FF model for consecutive days from 12:00 UTC on July 2 (hour 36) to 12:00 UTC on July 8 1997 (hour 180) for the Miedonia gauging station on the river Odra, SE Poland driven by the ECMWF EPS (left to right).

estimation in the rainfall forecast and a (too) early indication of the end of the flood. In the wane of the flood the ensemble spread increases towards the end, indicating a possibility of a newly increased flow. This increased flow was also observed to some extent in this case, only again much later on July 22 at a flow rate of $695 \text{ m}^3 \text{ s}^{-1}$. Conversely, the ensemble spread diminishes from forecast day to forecast day towards the end of the flood. As in the Meuse case, this may suggest that in comparing one forecast day with the other the ensemble spread is more indicative of the forecast uncertainty than is a set of ensemble forecasts on one particular day.

In general, the observed flow is largely under-predicted, especially prior to the onset of the major peak flow on July 5. Only at the start and during peak flow do discharge predictions give a somewhat improved indication of the observed flow. Contrary to the Meuse, however, the 0.95 percentile boundary remains well below observed flow rates.

Figures 10a-g compare the performance of the deterministic and probabilistic flow forecast updates for the Odra flood event for July 1 to July 8 1997 at the Miedonia station. It is evident that both forecast modes under-estimate the observed flow, especially prior to July 5. However, as mentioned for the July 1 flow prediction, the deterministic forecast provides only a single under-estimated value. Alternatively, the probabilistic forecast provides a low probability underprediction of a high extreme, which at least indicates the possibility of an extreme event that is relatively close to the observed flow. Hence, the ensemble of flow predictions contains important information, which is not present in a single flow forecast. How this information should translate into decision-making, however, needs to be investigated in more detail, using a number of (historical) case-studies, in particular non-flood events, in order to assess the statistical value.

Discussion and conclusions

Ensemble-based and deterministic flow predictions have been performed and interpreted for the Meuse 1995 flood and the Odra 1997 flood, using a skeleton version of the flood forecasting system developed in the EFFS project. This version basically consists of the EPS and deterministic meteorological hindcast data supplied by the ECMWF global atmospheric model, which are subsequently processed through the deterministic GIS-based distributed hydrological model LISFLOOD. From the perspective of Bayesian forecasting systems (BFSs) (Krzysztofowicz, 2002), this implies that only the meteorological input uncertainty is addressed, while the hydrological uncertainty is ignored. Maranzano and Krzysztofowicz (2004) found

that the influence of the hydrological model on river stage transitions is restricted to 48 hrs, given a 24-hr probabilistic rain forecast. Given the medium-range 10-day probabilistic weather forecast used in this study, the total uncertainty would be expected to increase over this forecast range, if the hydrological uncertainty were included.

The deterministic flow forecasts for the Meuse river showed encouraging results in that the flood event was captured as early as 2–3 days before the onset of the flood. The respective first and second peak flood were both forecast 3–4 days in advance. Further, it has been shown that the ensemble-based stream flow predictions, as interpreted by the probability of exceedence, provides complementary information at a lower resolution to the deterministic flow forecast, both in case of under- and over-prediction, especially in the medium forecast range. The added value of the probabilistic forecast is therefore the combination with the deterministic forecast, rather than its replacement. However, at a higher resolution the performance of the probabilistic forecast is likely to improve with respect to the under-estimation of extreme rainfall events. Run at an identical resolution, it would then replace the deterministic forecast. However, at present this is restricted by computational constraints. Alternatively, one could consider a smaller number of ensembles at a higher resolution but this would reduce the ensemble's statistical significance.

The flow forecasts for the Odra river basin present quite a different picture compared to the Meuse case, in that they only registered a signal for the emerging flood event shortly prior to its actual occurrence. Even after the start of the flood, peak flow was under-estimated. Nonetheless, despite the overall poor performance for this particular case, it was shown that the ensemble of flow forecasts provides additional information to the deterministic forecast, i.e. the indication of the possibility of an extreme event. In case an extreme event actually does occur, this represents vital information *a posteriori*. There is a need for further research assessing the performance of the probabilistic forecasts for more flood events and, more importantly, for non-flood events.

The difference in forecast results of the forecasting system developed for the Meuse and Odra flood event shows that performances may differ from event to event and/or region to region. This confirms the need to consider as many (historical) flood events as possible and to look at non-flood events to gain experience in the perspective of (false) alarm and (non-) flood event ratios and to assess the statistical value. This should preferably be carried out for a continuous period of at least a (hydrological) year.

Acknowledgements

Parts of this research have been funded from the European Commission funded project EFFS (EVG1-CT-1999-00011 EFFS) coordinated by Delft Hydraulics. All EFFS participants are thanked for the fruitful discussions on the work and for data support. Furthermore, the following institutes are thanked for their cooperation in the flood research presented in this paper: Czech Hydro-Meteorological Institute (Prague), Institute Of Meteorology And Water Management (Wroclaw), Regional Water Development Authority (Wroclaw), Landesumweltamt Brandenburg (Potsdam), Sachsisches Landesamt Fuer Umwelt Und Geologie (Dresden), Riza (Arnhem), Bundesanstalt Für Gewässerkunde (Koblenz, Berlin), Rijkswaterstaat Directie Limburg (Maastricht), Dienst Hydrologisch Onderzoek (Brussels), Sathy (Namur) and Diren (Nancy). The help of the JRC-IES staff, in particular R. Hiederer, G. Liberta, S. Peedell, L. Montanarella, B. Jones, A. De Jaeger and G. Schmuck is also acknowledged.

References

- Buizza, R. and Palmer, T.N., 1995. The singular-vector structure of the atmospheric general circulation. *J. Atmos. Sci.*, **52**, 1434–1456.
- Buizza, R., Miller, M. and Palmer, T.N., 1999. Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteorol. Soc.*, **125**, 2887–2908.
- Buizza, R., Richardson, D.S. and Palmer, T.N., 2001. The new 80-km high-resolution ECMWF EPS. *ECMWF Newsletter No. 90*, 2–9.
- Cressie, Noel A.C., 1993. *Statistics for spatial data* [revised edition]. Wiley, New York, 900pp.
- De Roo, A.P.J., 1999. LISFLOOD: A rainfall-runoff model for large river basins to assess the influence of land use changes on flood risk. In: *Ribamod: river basin modelling, management and flood mitigation, Concerted action*, P. Balabanis et al. (Eds.). European Commission, EUR 18287 EN, 349–357.
- De Roo, A.P.J., Wesseling, C.G. and Van Deursen, W.P.A., 2000. Physically-based river basin modelling within a GIS: the LISFLOOD model. *Hydrolog. Process.*, **14**, 1981–1992.
- De Roo, A.P.J., Thielen, J. and Gouweleeuw, B.T., 2002. *LISFLOOD, a distributed water balance, flood simulation and flood inundation model. User manual version 1.0*. European Commission, Special Publications No. I.02.131
- De Roo, A.P.J. and Schmuck, G. K., 2002. ODER-LISFLOOD, *Assessment of the effects of engineering, land use and climate scenarios on flood risk in the Oder catchment*. European Commission, EUR 20276 EN.
- De Roo, A.P.J., Bartholmes, J., Bates, P., Beven, K.J., Bongioannini-Cerlini, P., Gouweleeuw, B., Heise, E., Hils, M., Hollingsworth, A., Holst, B., Horritt, M., Hunter, N., Kwadijk, J., Pappenberger, F., Reggiani, P., Rivin, G., Sattler, K., Sprokkereef, E., Thielen, J., Todini, E. and Van Dijk, M., 2003. Development of a European Flood Forecasting System. *River Basin Manage.*, **1**, 49–59.
- Deutsch, C.V. and Journel, A.G., 1992. *GSLIB: Geostatistical software library and user's guide*. Oxford University Press, New York, USA. 340pp.
- Dutta, D., Herath, S. and Musiak, K., 2003. A mathematical model for flood loss estimation. *J. Hydrol.*, **277**, 24–49.
- Gouweleeuw, B.T., Reggiani, P. and De Roo, A.P.J., 2004. *A European Flood Forecasting System EFFS*. Full report, EUR21208EN, Ispra, Italy
- Keil, C., Volkert, H. and Majewski, D., 1999. The Oder flood in July 1997: Transport routes of precipitable water diagnosed with an operational forecast model. *Geophys. Res. Lett.*, **26**, 235–238.
- Kelly, K.S. and Krzysztofowicz, R., 2000. Precipitation uncertainty processor for probabilistic river stage forecasting. *Water Resour. Res.*, **36**, 2643–2653.
- Krzysztofowicz, R., 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resour. Res.*, **35**, 2739–2750.
- Krzysztofowicz, R., 2002. Bayesian system for probabilistic river stage forecasting. *J. Hydrol.*, **268**, 16–40.
- Krzysztofowicz, R. and Kelly, K.S., 2000. Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resour. Res.*, **36**, 3265–3277.
- Krzysztofowicz, R. and Maranzano, C.J., 2004. Hydrology processor for probabilistic stage transition forecasting. *J. Hydrol.*, **293**, 57–73.
- Lamb, R., Beven, K.J. and Myrabo, S., 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-run off model. *Adv. Water Resour.*, **22**, 305–317.
- Molteni, F., Buizza, R., Palmer, T.N. and Petroliaigis, T., 1996. The ECMWF Ensemble Prediction System: methodology and validation. *Quart. J. Roy. Meteorol. Soc.*, **122**, 73–119.
- Mullen, S. and Buizza, R., 2001. Quantitative precipitation forecasts over the United States by the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **129**, 638–663.
- Mylne, K., 2003. Use of ensemble prediction to improve forecast of hazardous weather. *Geophys. Res. Abstracts*, **5**, 01765.
- Pappenberger, F., Beven, K.J., Hunter, N., Bates, P., Gouweleeuw, B., Thielen, J. and de Roo, A., 2004. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrol. Earth Syst. Sci.*, **9**, 103–115.
- Persson, A., 2001. User Guide to ECMWF forecast products. *ECMWF Meteorological Bulletin M3.2*.
- Sattler, K. and Feddersen, H., 2003. *A European Flood Forecasting System, Study on the treatment of uncertainties in limited area NWP rainfall using different ensemble approaches*. DMI Tech. Rep. Available from the Danish Meteorological Institute (DMI), Copenhagen, Denmark.
- Ulbrich, U. and Fink, A., 1995. The January 1995 flood in Germany: Meteorological versus hydrological causes. *Phys. Chem. Earth*, **20**, 439–444.
- Wösten, J.H.M., Lilly, A., Nemes, A. and Le Bas, C., 1998. *Using existing soil data to derive hydraulic parameters for simulation models in environmental studies and in land use planning*. Winand Staring Centre, Wageningen, The Netherlands Report no. 156. 106pp.