Institute for
Environment and
Sustainability

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

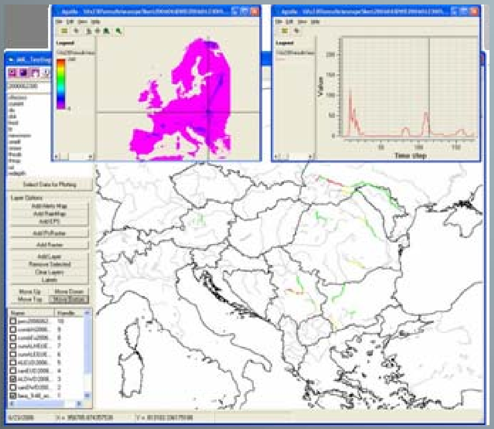
The benefit of probabilistic flood forecasting on European scale

-

Results of the European Flood Alert System for 2005/2006

Jutta Thielen (Editor)

with contributions from
Bartholmes J. , Ramos M.H., Thielen J., and et al.



Institute for Environment and Sustainability

2007

EUR 22560 EN

The mission of the Institute for Environment and Sustainability is to provide scientific and technical support to the European Union's policies for protecting the environment and the EU Strategy for Sustainable Development.

European Commission
Directorate-General Joint Research Centre
Institute for Environment and Sustainability

Contact information:
Address: Jutta Thielen, JRC/IES, TP261, via E. Fermi, 1, 21020 Ispra (VA), Italy
E-mail: jutta.thielen@jrc.it
Tel.: +39 0332 785455
Fax: +39 0332 786653

http://ies.jrc.cec.eu.int
http://www.jrc.cec.eu.int

EUR 22560 EN
ISSN 1018-5593
Luxembourg: Office for Official Publications of the European Communities

Printed in Country

Acknowledgements

The editor and authors would like to thank all those that have given support to the work presented in this report. In particular the floods team of the Weather Driven Natural Hazards Action, the IT support, and the valuable GIS support. Further the authors would like to thank P. Bougeault from the European Centre for Medium-Range Weather Forecasting for critical questioning of the results which has given more depth to the analysis. Finally all data providers should be thanked at this place, which are the AgriFish unit of the JRC for providing observed meteorological data, the Deutsche Wetterdienst for providing observed and forecasted data and the European Centre for Medium Range Forecasting for providing deterministic and probabilistic weather forecasts. In particular I would like to thank their data support unit for swift solutions to data problems.

Table of Contents

Chapter 1

Introduction to the European Flood Alert System development

Jutta Thielen and Ad de Roo
European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, TP261, 21020 Ispra (Va), Italy

1.1 Introduction to EFAS and its aims

The European Flood Alert System (EFAS) is a research project that aims at improving preparedness for oncoming flood events by:

- informing the authorities in the Member States of the possibility of a flood to happen *before* the local systems capture the event with their own monitoring and forecasting systems;
- providing catchment based information that gives downstream authorities an overview of the current and forecasted flood situation also in upstream countries. This does not only include the immediate upstream country but also any further potentially useful upstream information.

EFAS provides medium-term (3 to 10 days) flood forecasts, based on different meteorological inputs, as complementary information to the typical short-term (less than 48 hours) forecasts performed by national centres. The challenge of EFAS is to combine all hydrographs, calculated from different medium-range weather forecasts, into one early flood warning information that is useful for local flood forecasting centres. EFAS information is disseminated to the EFAS users in the form of EFAS Information Reports, which summarise the situation in a concise way with spatial and temporal information.

The establishment of an EFAS network was initiated in August 2004, with a formal agreement between the JRC EFAS activity and ECMWF (meteorological data provider). In the agreement, about 30 hydrological services were identified as potential partner organisations. They were contacted in January 2005 with a Memorandum of Understanding (MoU) which identified the rules of collaboration for the EFAS research activity. The first MoU was signed in May 2005. Once EFAS could demonstrate its potential benefit during the summer floods in 2005, more authorities have joined the network, which comprises currently 22 authorities. Figure 1.1 shows the status of the river basin areas covered by EFAS MoU agreements on 26th June 2006.

At the moment, the crucial question for further development of EFAS is: does EFAS information actually help the receiving authorities to be better prepared for an oncoming flood event? To answer this question it is first necessary to discuss the potential benefits and limitations of EFAS forecasts for increasing preparedness for flood events.

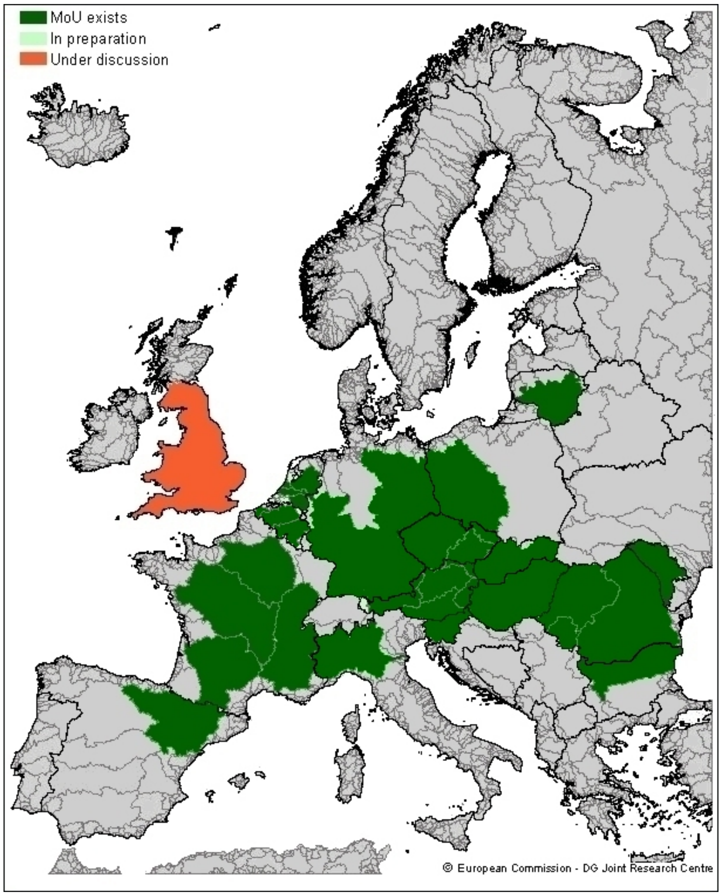


Figure 1.1: Overview of trans-national river basin areas for which an EFAS partner organisation exists on 30th November 2006.

1.2 Possible benefits and adverse effects of EFAS forecasts

Globally, the local flood forecasting centres will not activate an emergency procedure only based on the early flood alerts forecasted by EFAS. For this specific action, the national authorities take a decision based mainly on their local information, which is more accurate than EFAS information due to the use of higher resolution and locally calibrated models, as well as specific expert knowledge. EFAS forecasts are supposed to be used as a *pre-alert* to allow the receiving authorities to be aware of the possibility of a flood to take place. In other words, with EFAS forecasts in hands, local forecasters can already play through a number of different scenarii “what to do if” and, as the event approaches and its location and magnitude become more certain, national authorities can act more quickly and accurately, increasing the economic value of their forecasts. In fact, cost-benefit discussions typically revolve around matrixes such as the one in Table 1.1.

Table 1.1: Typical contingence table for assessing the quality and value of a forecast.

	Forecasted event yes	**Forecasted event no**
Observed event yes	Positive hit	Missed event
Observed event no	False alarm	Correct rejection

Considering that EFAS medium-range forecasts are designed for better preparedness for oncoming flood events, the benefits or adverse effects of the individual members of the matrix in Table 1.1 can be qualitatively assessed as indicated in the four distinct scenarii detailed below :

a) A ***positive hit*** by EFAS is ***always beneficial***.
   a. Receiving authorities are aware of a possible flood event to happen before the local systems capture the event. Different scenarii for location, severity and counter-acting actions can be explored. As the event draws nearer and becomes more defined within the local forecasting system, the authorities can act more quickly and with more confidence.
   b. If the local flood forecasting system has also already captured the event, there is more confidence in the forecasts and again action can be taken faster.
   c. If there is conflict between the – in this case correct – EFAS information and the – in this case wrong – local information, the receiving authorities will be more cautious with the interpretation of their own results and at least will not be taken entirely by surprise if the flood event takes place after all.

b) A ***missed event*** by EFAS ***should not have but may have adverse effects***.
   a. The receiving authorities have no early notice of the potential danger of a flood to happen and are not prepared in advance. However, they have their own forecasting systems in place and therefore act according to “normal” mode. There is no adverse effect.
   b. However, in case the receiving authorities are already used to receiving early EFAS information reports there could be a conflicting situation if the local systems capture the event and EFAS has not/is not reporting the event. In this case the receiving authorities may not right away trust their local forecasting system at first signal. This may incur some delay and thus have adverse effects on the preparedness.

c) A ***false alarm*** by EFAS ***has little* adverse effects**.
   a. If EFAS predicts a flood more than 3 days in advance the authorities may be on stand-by for action but the closer the potential event draws the clearer it should become that the event is not going

to take place and therefore adverse effects do not occur or are very little.
   b. Local flood forecasting results may be discussed and checked in more detail. This may take some of the forecasters' time but does not have adverse effects on the preparedness.

d) ***Correct rejections*** are ***always beneficial***
   a. Correct rejections are always beneficial because receiving authorities are correctly not alerted and do not need to act. There is agreement and therefore confidence in the local forecasts. Since this will be the situation in the majority of the cases the impact of correct rejections is of course very small.

In summary, it can be concluded that in the majority of the cases the additional information provided by EFAS is beneficial. This is particularly true for a correct early forecast, in which case the benefit would be greatest. The impact of false alarms is, comparatively, small. This is under the assumption, of course, that the number of false alerts is not over-proportionally high and within the probability range: flood forecasts with a 20% probability should happen 1 out of 5 forecasts! The only case where EFAS information could have adverse effects on preparedness is a totally "missed" event if this incurs mistrusting the local forecasts. Since EFAS, however, by definition, covers the early warning range while the local forecasting systems pre-dominantly cover the short-range, there is little chance for this to happen.

1.3 Objectives and content of this report

The objective of this report is to assess EFAS performance and verify EFAS results for the 2005/2006 period, starting on summer 2005, under three specific perspectives:

i. **Enduser perception on the usefulness of EFAS information reports (Chapter 3)**
The impact of EFAS information reports in the local operational flood forecasting centres is analysed with respect to hit/false alarms, perception of the enduser concerning the usefulness of the information and also the clarity of its presentation in EFAS reports, as well as its effective dissemination. The analysis is based on feedback questionnaires, which are systematically sent out after the last EFAS Information Report for an event, on discussions during technical EFAS workshops (the first one was held in January 2006), as well as on any other personal or informal email communications. Results for 2005/2006 are presented in Chapter 3.

ii. **In-depth case study analyses of selected events and statistical evaluation of flood-prone periods (Chapter 4)**
Case studies in EFAS serve to understand the potential capabilities of EFAS results for early warning for different types of flood events. They are also used to identify interpretation and decision rules for individual forecasts (deterministic and EPS forecasts), as well as for combined forecasts (using the ensemble of all EFAS forecasts in a complementary way). The case studies in Chapter 4 focus on the

Danube catchment which experienced repeatedly different types of flooding events in summer 2005 and spring 2006.

iii. **Quantitative analyses of different verification (skill) scores on an European scale (Chapter 5)**
The aim is to evaluate the skill of EFAS hydrologic forecasts under different aspects of deterministic, probabilistic and combined forecasting. Statistical methods are used and scores are computed to assess the performance of EFAS forecasts. Analyses of hit and false alarm rates are performed and the Brier score and the Brier skill score are calculated. Results are given in Chapter 5.

This study aims at analysing the usefulness of EFAS results not under just one specific aspect, but in its totality of using all the available input data (currently, deterministic DWD and ECWMF weather forecasts and ECMWF-EPS probabilistic forecasts) for issuing a combined early flood alert.

This report summarises briefly important aspects of EFAS set-up, methodologies and visualisation tools in Chapter 2 and presents in-depth analyses of EFAS results in Chapters 3 to 5, as indicated above. In Chapter 6, general conclusions are drawn.

Chapter 2

EFAS set-up, methodology and visualisation

Jutta Thielen[1], Johann van der Knijff[1], Giovanni Franchello[2], Milan Kalas[3],
Maria-Helena Ramos[1,4] and Jens Bartholmes[1]

[1]European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, TP261, 21020 Ispra (Va), Italy
[2]European Commission, DG Joint Research Centre, Institute for the Protection and Security of the Citizen, TP 670, 21020 Ispra (VA), Italy
[3]Detached from Department of Land and Water Resources Management, Faculty of Civil Engineering, Slovak University of Technology, Slovakia
[4] Cemagref - Groupement de Lyon ; 3 bis quai Chauveau CP 220, 69336 LYON cedex 09 FRANCE

2.1 LISFLOOD: the rainfall-runoff model for EFAS

EFAS is based on the rainfall-runoff model LISFLOOD, a grid-based catchment model that has been developed to simulate floods in large European river basin (de Roo et al, 2000, van der Knijff, 2004). The typical size of one grid cell is 1 by 1-km, although the model can be run at both much finer and coarser resolutions if needed. It is used in EFAS both for the calculation of the initial conditions using observed meteorological data as input and for the flood forecasting using the different meteorological weather forecasts as input. A schematic view of the physical process simulated is given in Figure 2.1.

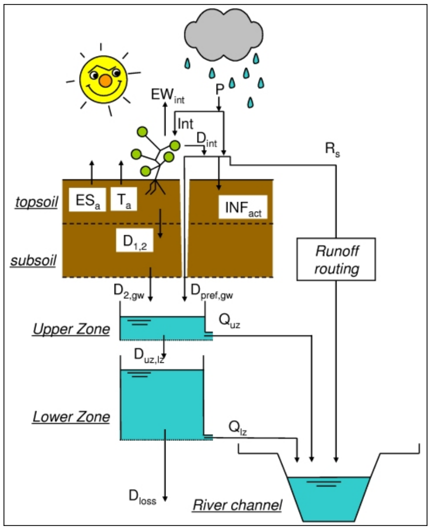


Figure 2.1: Schematic view of the physical processes simulated in LISFLOOD model.

One of the difficulties in the numerical description of hydrological processes is that soil types and textures range from porous grounds to solid rock or clay and are highly variable in space. Consequently, hydrological processes cannot be totally physically described and there will always be elements that need to be parameterized and calibrated on observed discharges.

For EFAS, the need of observed discharges constitutes a challenge: discharge data have been collected for two pilot basins, the Danube and the Elbe (Wachter *et al.* 2004, Gierk et al., 2004), but for most large European rivers discharge data are not yet available. Two important projects have been launched in 2006 to collect these data (ETN-R and EU-FLOOD-GIS) but the data is not expected before 2-3 years time. In the meantime, the development of the prototype needs to be based on what is currently available for calibration, validation and forecast verification: discharge data at a reduced number of stations, information from EFAS partners and information from the European Media Monitoring (EMM) on observed events (http://emm.jrc.it).

However, it should be noted that even with relatively long time series, the calibration of a hydrological model is not a trivial task. The hydrologic environment of river catchments and also the river beds undergo constant modifications through sedimentation, erosion and man-made interventions. These changes can sometimes be very abrupt, as for instance after a flood event, and have consequences for the observed hydrographs (introduction of trends or jumps in the time series). In these cases long-term reference values, climatologic values and statistical distributions are more difficult to derive of non-stationary time series.

2.2. EFAS set-up

The prototype development of EFAS focuses on a robust forecasting system that incorporates a minimum of different medium-range weather forecasts to produce flood forecasts across Europe that can be distributed and analysed for their usefulness and impact. The concept of the development is to start from a simple set-up that can be improved step-by-step. The first prototype will then be accompanied with a set of suggested improvements for the second phase of the project. At present, for example, improvements already identified for the next step of development are the geographical bias correction of the meteorological forecasts and the assimilation of real-time observed discharges, soil moisture and/or snow extent information.

Following the set-up which was successfully tested for the competitive European Flood Forecasting System project (Gouweleeuw et al, 2004, Gouweleeuw et al., 2005), for EFAS the LISFLOOD model has been set-up for the whole of Europe with a grid-spacing of 5 km. Input data consist of static data on one hand and dynamic data on the other hand. Static data sets for example are those describing the river basins, e.g. soil data, landuse data, river network data, which are available on European scale.

The flood forecasts, however, are essentially driven by the dynamic meteorological data inputs. An overview of these data is given in Table 2.1. Essentially they consist of:

i) observed meteorological data from the JRC MARS-STAT database (http://agrifish.jrc.it/marsstat/datadistribution), necessary to determine the initial conditions,

ii) medium-range weather forecast data from ECMWF and the DWD for the deterministic flood forecasts and

iii) medium-range Ensemble Prediction System (EPS) also provided by the European Centre for Medium-Range Weather Forecasts (Molteni *et al.*, 1996, Buizza *et al.*, 2003) All weather forecasts are run through the EFAS system individually, i.e. each EPS member is an input to the model just like a deterministic forecast. The ensemble of streamflow series obtained are analysed and interpreted at the end of the simulation. Combined results from the deterministic and the EPS weather forecasts are explored before flood forecasts and eventual warnings are issued.

The observed meteorological data, which are used for climatological studies and for the calculation of the initial conditions in EFAS, are derived from the JRC MARS-STAT data base. These data play a key role for EFAS because they do not only determine the initial conditions but also the (model) climatology. They JRC MARS data base comprises data from an irregular network of synoptic weather stations. The data are available from 1975 onwards. The time resolution is 1 day and covers a 24-hour period from 6 am to 6 am. The weather stations are not homogeneously distributed (Figure 2.2, left) and, when assigned to a 50x50 km grid, not every cell contains a weather station (Figure 2.2, right). Good coverage is found in Central Europe and Great Britain, while the coverage of the weather stations in Eastern Europe has shown a significant improvement after 1995, in particularly in East Europe.

Table 2.1: EFAS input data in the period 2005/2006

	DWD	**ECMWF-deterministic**	**ECMWF-EPS**	**Observed meteorological data JRC-MARS**
Temporal resolution	Staggered, 1h (1-3 days); 3h (4-7 days)	Staggered 3h (1-3 days); 6h (4-10 days)	6h (1-10 days)	Daily
Spatial resolution	Staggered, 7 km (1- 3 days) & 40km	TL511L60 (~40km)	TL255L40 (~80km)*	Gridded 50x50km^2
Times provided	12:00; 00:00	12:00; 00:00	12:00; 00:00	Irregular, typically at 23:00
Input fields	1 (P, T, E)	1 (P, T, E)	50+1 (P, T, E)	P, T, E0, ES0, ET0
Bias removal	None	None	None	None

Where: P=total precipitation, T-temperature, E=evaporation, E0=potential evaporation over water, ES0=potential evaporation over bare soil and ET0=potential evapotranspiration

* Since autumn 2005 ECMWF is producing the EPS at a higher spatial resolution. For technical reasons they are at present still resampled for EFAS on 80km

Table 2.2: EFAS simulation set-up for EFAS version 3.1

	DWD	**ECMWF-deterministic**	**ECMWF-EPS**
Temporal resolution	1h	1h	24 h
Spatial resolution	5km	5 km	5 km
Downscaling	Dynamical downscaling Lokalmodell for the first 2 days (2005) and 3 days (2006).	No	No

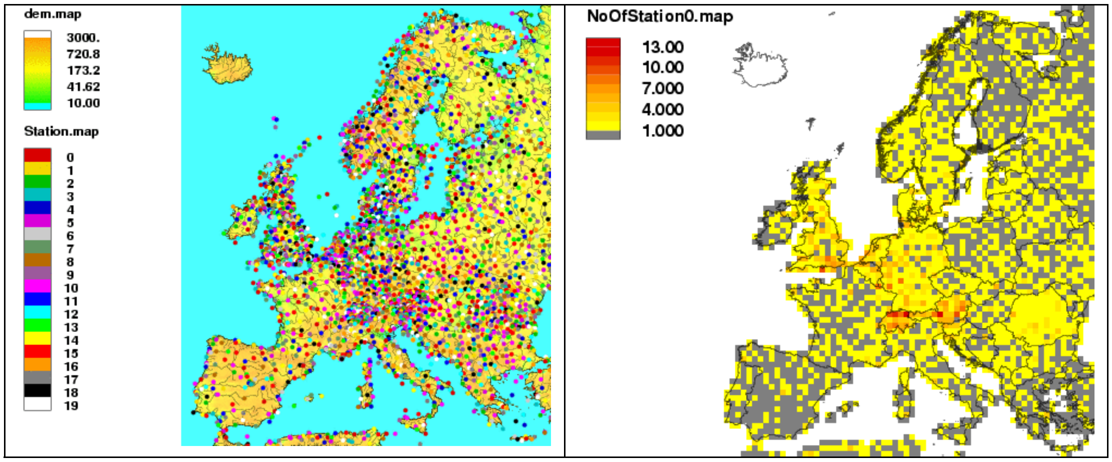


Figure 2.2: The irregular network of weather stations from the JRC MARS-STAT data base (left) and number of stations per 50x50 km grid cell (right).

2.2.1 Principle of threshold exceedance

With the current limitations in the meteorological input data to drive the model and the scarcity of data to calibrate and verify the model results, it is unlikely that the model will perform perfectly well in the quantitative prediction of discharge at this stage of the project. The goal of EFAS is, however, to produce early flood warning and *not* quantitatively correct discharges. The high degree of uncertainty in the meteorological rainfall forecasts at the medium-range together with the parameter and calibration uncertainty in the present model set-up would make a quantitative forecast meaningless. Therefore the principle of threshold exceedance is being applied. Instead of quantitative discharge or height information the flood warning procedure is based on a dichotomous information of yes/no that a flood threshold has been exceeded where the thresholds have been derived from the model simulations themselves. There are two reasons for this in a model-consistent approach:

First, LISFLOOD uses a regular grid structure and EFAS results are being evaluated for the entire grid and not just for selected points. Thus EFAS calculates if a critical threshold is being exceeded at every grid point of the model. Therefore critical values need to be determined at *every* model grid point. Observations of critical values are, however, only available at selected gauging stations.

Second, EFAS is set-up on a 5-km grid for the whole of Europe and simulates currently the river basins essentially as natural catchments. This means that information on steering rules for lakes, reservoirs, polders or any other measures are not included in the set-up. Consequently the simulated discharges are likely to differ from observed ones. They cannot thus be compared against observed thresholds, defined from observed discharge time series. They need to be compared against thresholds derived within the same modelling environment and the same parameter set.

To tackle these problems a model consistent approach has been chosen for EFAS (see also Figure 2.3):

1) A long term water balance simulation based on observed meteorological data (mostly synoptic network and other available meteorological stations) was calculated with the LISFLOOD model for a period of 14 years (1990-2004).
2) The discharges from the long-term time series were ranked from the highest to the lowest value. This was done for every river pixel across Europe in a 5-km grid space.
3) From the ranked discharges, four specific discharge values at given rank positions were selected as thresholds (alert levels).

This approach has the advantage that any systematic over- or under-prediction of the model is compensated. If, because of the model set-up as a natural catchment or because of its non-optimised parameterisation, the model tends to overestimate discharges at a given location this will be reflected in the thresholds as well as in the simulations. In this way, in EFAS forecasts, the exceedances of a threshold become more informative than the magnitude of the simulated discharges and the actual values of the thresholds.

In EFAS, currently four different thresholds are defined (Table 2.3). At present, the highest discharge simulated during the 14-year water balance simulation (gives the EFAS Severe threshold. The EFAS High Threshold has been selected as the Q-value for which only 1% of all discharges were higher during the simulation period. It has been seen that in many situations this threshold corresponds roughly to a 1-2 year return period flood, which corresponds to typical bankful conditions in a natural catchment. The drawback of this approach is the relatively short time span from which the thresholds have been calculated. The reason for this is that the quality of the observed meteorological data is much poorer in the years before 1994 due to the fewer stations. Taking into consideration too many years with poor data quality would have impact on the threshold. On the other hand the thresholds depend very much on the hydrological regime over the last 14 years. If the basin experienced rather low flow periods during this time, the thresholds would be comparatively low.

Definitions of the alert levels are given in Table 2.3. Figure 2.4 shows graphically how the thresholds relate to the real geometry and water level conditions in the river. It illustrates also the uncertainties of a qualitative interpretation (obviously, the better EFAS is calibrated and validated, the smaller the uncertainty range will be).

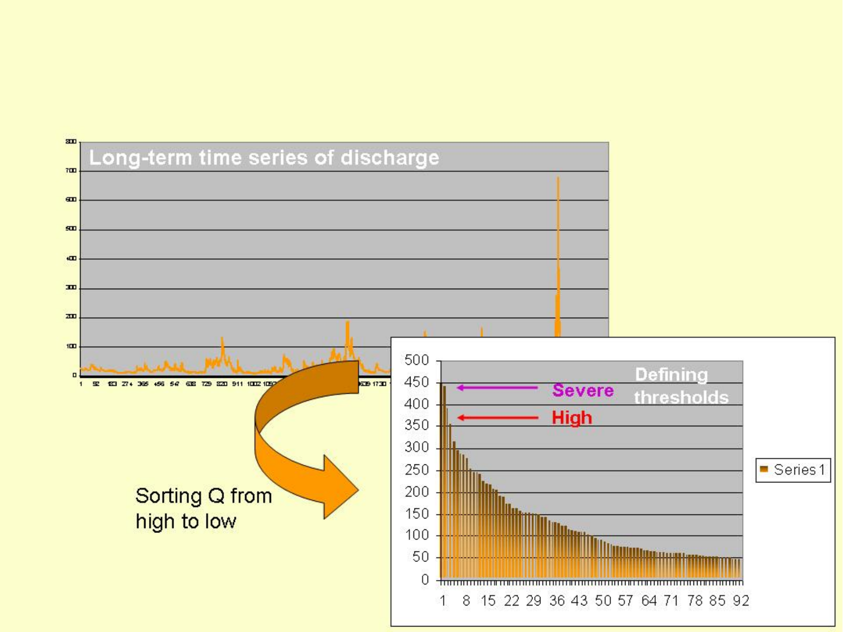


Figure 2.3: Illustration of EFAS threshold determination.

Table 2.3: EFAS levels

EFAS Level	Color	Description
S (Severe)	🟪	very high possibility of flooding, potentially severe flooding expecting
H (High)	🟥	high possibility of flooding, bankful conditions or higher expected
M (Medium)	🟨	water levels high but no flooding expected
L (Low)	🟩	water levels higher than normal but no flooding expected

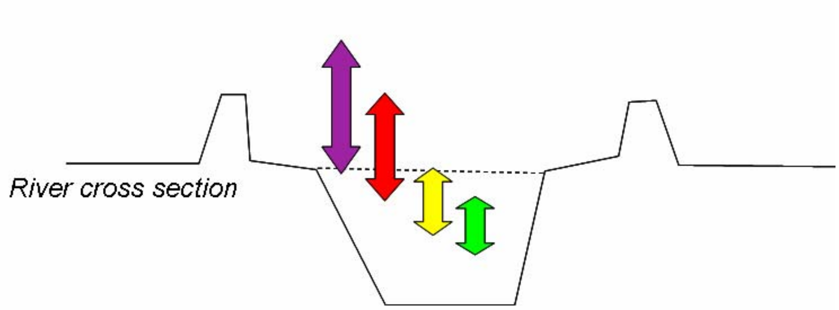


Figure 2.4: Graphical illustration of how EFAS thresholds compare indicatively to real river geometry and water level conditions.

An important assumption of the threshold exceedance principle is that the thresholds derived with observed rainfall data from the JRC-MARS STAT database can also be applied to the flood forecasts driven with meteorological weather prediction data. In other words, the forecasted rainfall fields from the DWD and ECMWF are comparable to the observed rainfall fields from the JRC-MARS STAT data base.

Some analysis has been performed to verify if this assumption holds. Figure 2.5 shows, for example, a comparison between the 00:00 and 12:00 forecasted precipitation of the DWD and ECMWF and the observed precipitation of the JRC-MARS STAT data base. First, the ECMWF and DWD forecasted data are aggregated to 24h – per each day leading time – and then they are compared to the MARS-STAT data. Results from the analysis of the years 2004 and 2005 show that ECMWF overpredicts about 10% with respect to the MARS-STAT, while DWD estimates are within a range of -10% to +5 %. Looking to the mean *RelRmse* (relative mean square error), ECMWF is performing slightly better than DWD: the mean *RelRmse* is higher than 1 after 2-3 days for DWD and after 4 days for ECMWF.

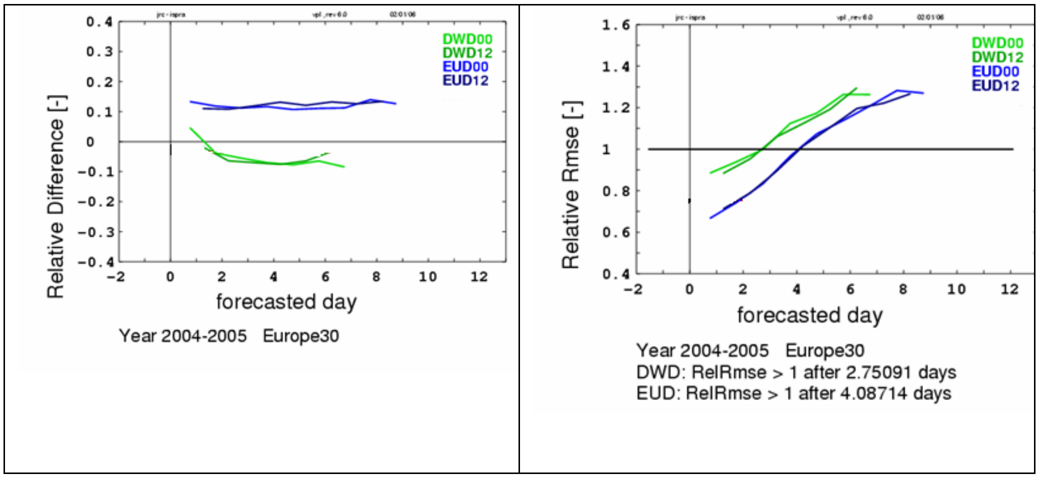


Figure 2.5: Comparison of mean forecasted and observed precipitation values for the whole of Europe with observed precipitation derived from the JRC MARS-STAT database for the period December 2004 to November 2005. Relative difference (left) and mean root square error (right) are shown.

It is also important to quantify the spatial bias over Europe for the different lead times of the forecasting ranges. A typical example is the seasonal precipitation bias in mountainous areas such as the Alps. Figure 2.6 illustrates the spatial distribution of the relative difference between the forecasted and the observed precipitation data. As could be expected, the differences are comparatively smaller at the first 24 hours of forecasting range and larger at the 4th day of forecast. The ECWMF forecasts have a tendency to overestimate, while DWD forecasts have a tendency to underestimate as compared to the JRC MARS STAT data.

Considering the uncertainty in spatial precipitation fields derived from available point measurements – an uncertainty of 30-50% has been calculated - it is assumed that the assumption of comparable rainfall fields from observed and forecasted rainfall fields is satisfactory for the demands on the EFAS prototype.

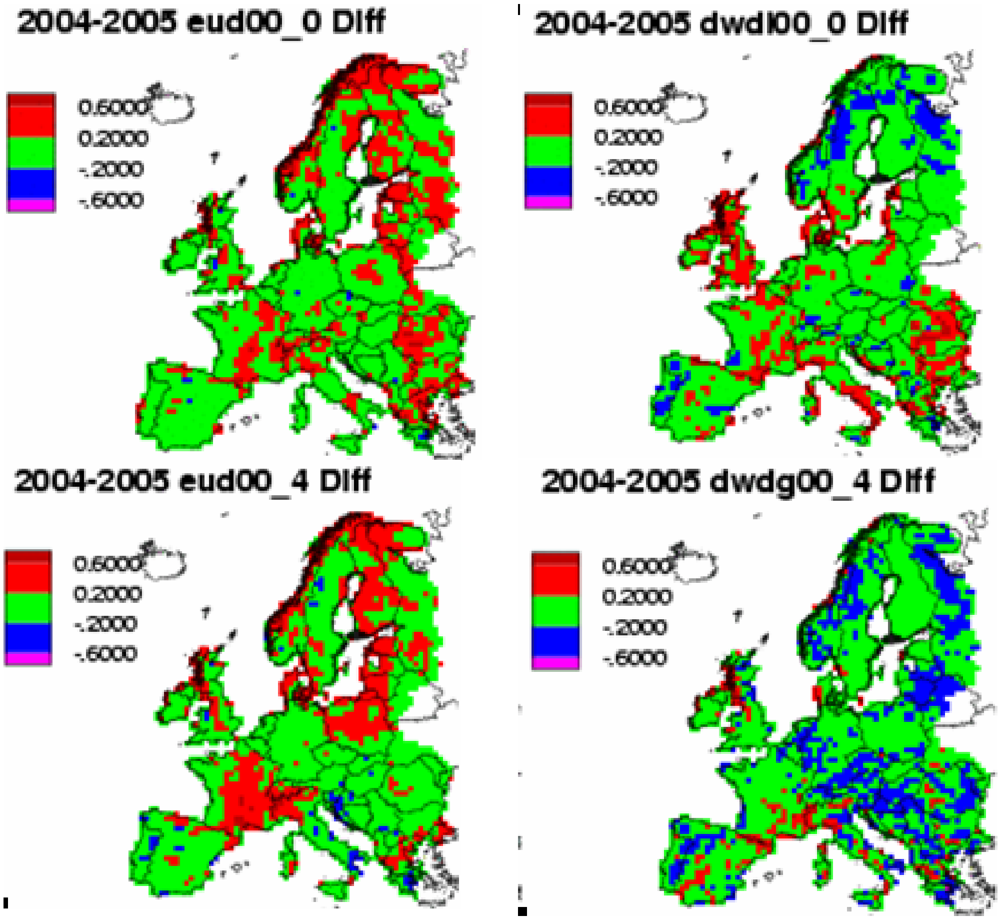
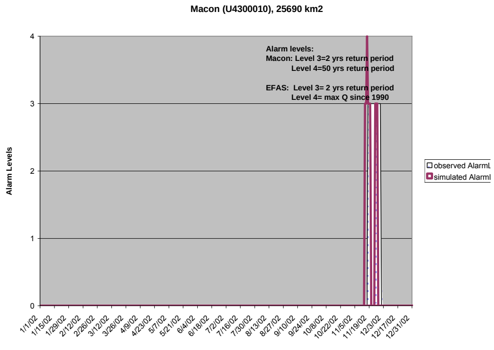


Figure 2.6: Spatial distribution of relative difference between forecasted and observed precipitation for the deterministic ECWMF (left) and the DWD (right) forecasts for the first 24 hours (top) and for a 4-day leadtime (bottom): in red, relative overestimation and in blue, underestimation of precipitation during the years 2004-2005. For the DWD forecasts, the data are derived from the LokalModell for the first 24 h (Day 0) and from the GlobalModell for Day 4.

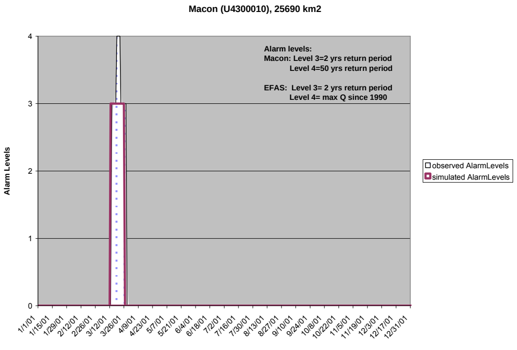
2.2.2. A case study illustrating the threshold exceedance principle

In EFAS the forecasted discharges calculated with the different weather forecasts are compared against the four threshold levels: Severe, High, Medium and Low. The assessment of EFAS hits and false alarms then reduces to the analysis of two dichotomous time series: observed yes/no, forecasted threshold exceeded yes/no.

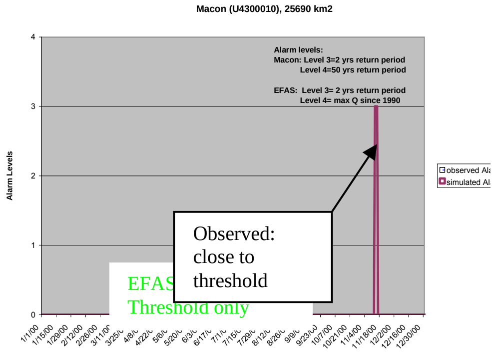
Figure 2.7 shows the time series of discharge exceedances for two comparable thresholds for a gauging station in Macon (station U4300010) in the Sôane River. The station has an upstream area of 25,690 km^2. The chosen thresholds are the observed 1-year return period discharge, which compares with the EFAS High threshold, and the observed 10-year return period discharge, which can be compared with the EFAS Severe threshold. The analysis of this station is interesting as it represents an example of a totally non-calibrated river basin in the EFAS system. The graphs shown illustrate the capacity of the system to simulate flooding using *observed* meteorological data as input, when comparing the simulations against model internal thresholds. In this case the same type of data inputs is used to calculate the thresholds and to calculate the threshold exceedances. The quality of the input data for a particular catchment depends largely on the density of stations inside the area. This density is not very high for France (see Figure 2.2). Also, the density of gauging stations is not static in time. Over the years more and more stations have been incorporated to the network. In particular, since 1994, the quality of the database has much improved.



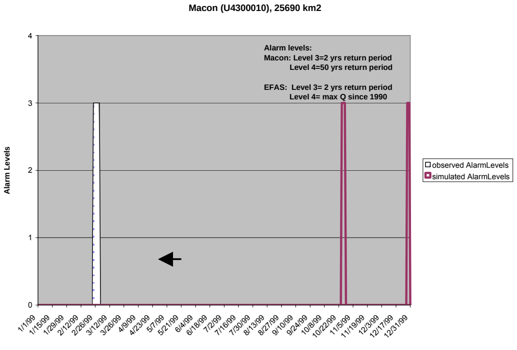
Analysis 2002 – EFAS Hit



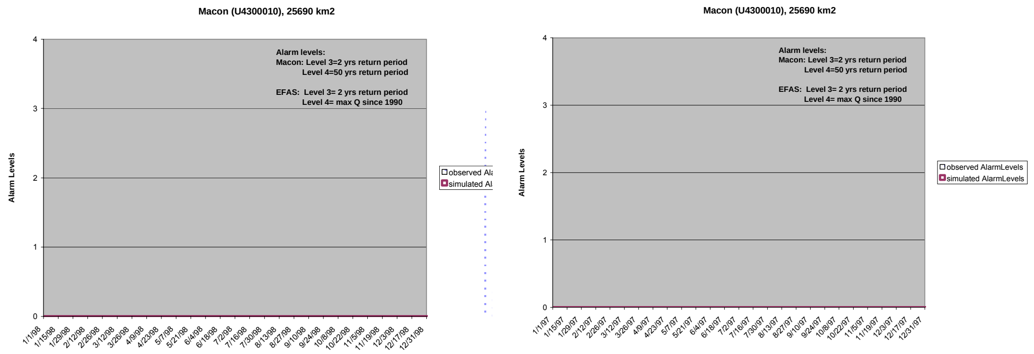
Analysis 2001 – EFAS Hit



Analysis 2000 – 1 EFAS false Alarm



Analysis 1999: 1 Missed Event, 2 False Alarms



Analysis: 1998: only correct rejections

Analysis : 1997: only correct rejections

Figure 2.7: Comparison of threshold exceedances between observations (white area) and EFAS simulations (red outline) based on observed meteorological data for the years 2002 (tl), 2001 (tr), 2000 (ml), 1999 (mr), 1998 (bl) and 1997 (br).

The example in Figure 2.7 shows that:

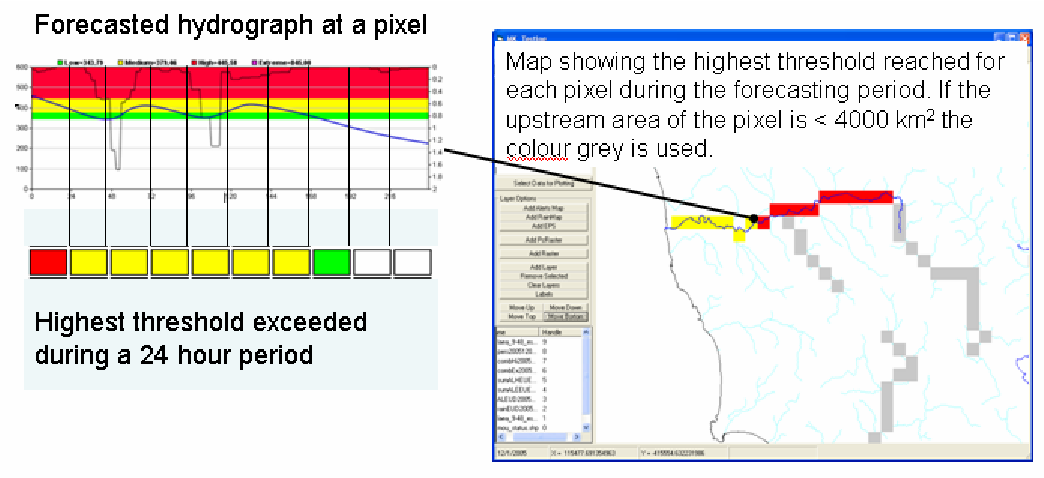
- EFAS correctly simulates threshold exceedances in 2002 and 2001. In both cases the timing of the floods is well simulated. The severity of the event is overestimated in 2002 and underestimated in 2001. However, the results are in general very good.
- In the year 2000, EFAS simulations are exceeding the EFAS high alert threshold by approximately 300 m^3/s. The observed discharges are also high, but do not exceed the threshold (in fact, it would be necessary more 400 m^3/s to reach the alarm level). Thus, EFAS effectively overestimated this event.
- In 1999, the situation is different: the observed threshold exceedances match in EFAS against the exceedances of the lowest EFAS threshold only (not shown here). In October/November, the observed discharges are just 100 m^3/s below the observed threshold, while EFAS is just slightly exceeding its thresholds. This illustrates the problem of defining absolute values of thresholds, without considering uncertainties. In this case, a more fuzzy approach in defining the thresholds might be helpful.
- The years 1997 and 1998 are two years in which floods were neither forecasted nor simulated.

Figure 2.7 indicates a result that has been confirmed by experience with some other discharge gauging stations: in the present version, the model has difficulties to correctly simulate snowmelt-related floods, but it performs quite well for floods taking place in the summer months. This problem is already being addressed through improvement of the model code and parameterisation.

2.2.3 Visualising combined EFAS information

EFAS results are principally summarised as daily information. If during any 24 hours of simulations a threshold is exceeded for that day the river pixel is flagged visually with the colour of the highest threshold exceeded during the 24-hour period. With this approach the information presented is reduced, but also is reduced the uncertainty in the timing of an event. Figure 2.8 illustrates the calculation and visualisation of EFAS results for the forecasts based on the deterministic weather forecasts. In analogy, threshold exceedance box plots are also produced for the EFAS results based on EPS. In this case the 51 different forecasts are summarised by numbers. This is illustrated in Figure 2.9. The highest number exceeding the EFAS High threshold (HAL) or EFAS Severe threshold (SAL) during a 24-hour period is counted and presented in the corresponding lead time box.

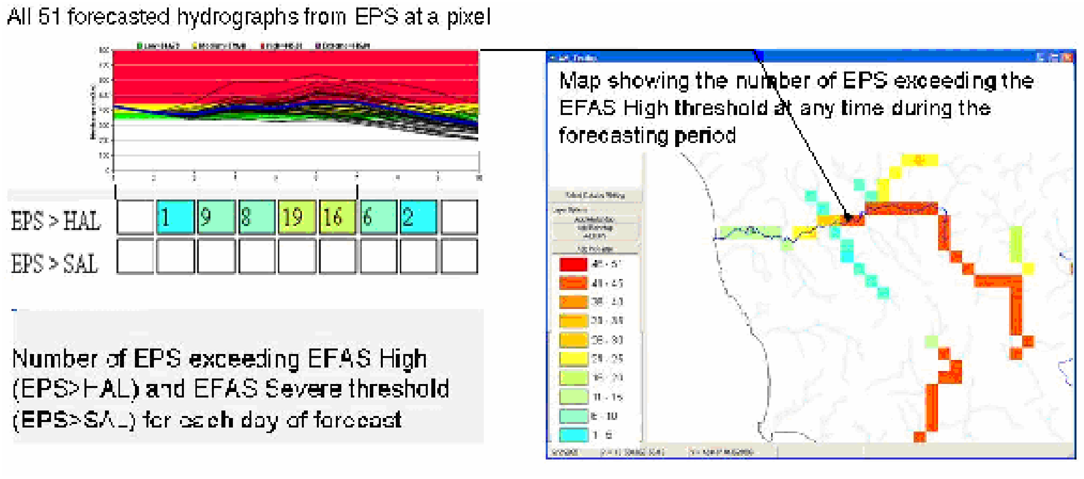
Figure 2.10 shows an example of EFAS forecast diagrams for a given forecast date and a given point in a river. In this example, simulations are based on weather forecasts issued on 10th July 2005 at 12:00. In this case the information of all forecasts is visualised in a concise and easy to understand way.


Forecasted hydrograph at a pixel

Highest threshold exceeded during a 24 hour period

Map showing the highest threshold reached for each pixel during the forecasting period. If the upstream area of the pixel is < 4000 km² the colour grey is used.

Figure 2.8: Summarising deterministic hydrograph information into threshold exceedance box plots and overview maps, each box representing a 24 hour time span.

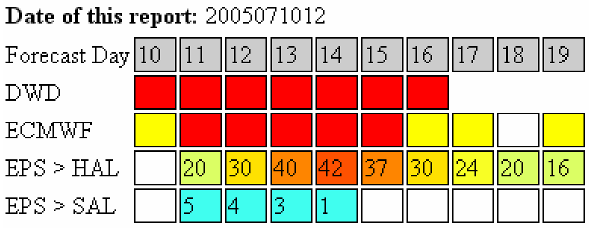

All 51 forecasted hydrographs from EPS at a pixel

Map showing the number of EPS exceeding the EFAS High threshold at any time during the forecasting period

EPS > HAL		1	9	8	19	16	6	2	
EPS > SAL									

Number of EPS exceeding EFAS High (EPS>HAL) and EFAS Severe threshold (EPS>SAL) for each day of forecast

Figure 2.9: Summarising flood EPS forecasted hydrograph information into threshold exceedance box plots and overview maps.

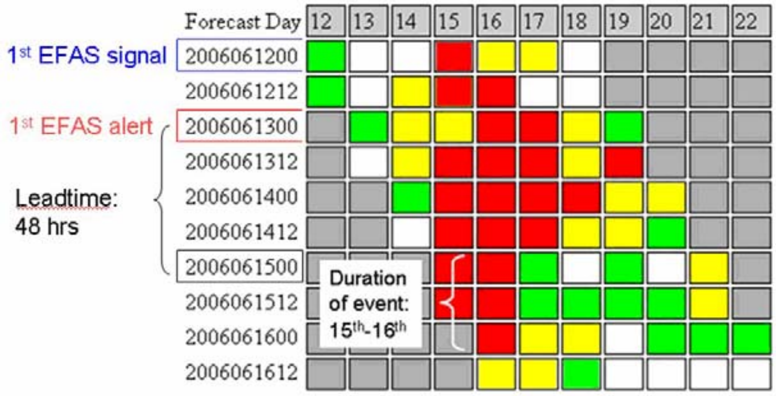

Date of this report: 2005071012

Forecast Day	10	11	12	13	14	15	16	17	18	19
DWD	🟥	🟥	🟥	🟥	🟥	🟥	🟥			
ECMWF	🟨	🟥	🟥	🟥	🟥	🟥	🟨	🟨	⬜	🟨
EPS > HAL	⬜	20	30	40	42	37	30	24	20	16
EPS > SAL	⬜	5	4	3	1	⬜	⬜	⬜	⬜	⬜

Figure 2.10: Example of EFAS Forecast Diagrams for the 10th July 2005 (12:00) at a point in a river: alert levels exceeded in the next 7 days based on DWD forecasts; alert levels exceeded in the next 10 days based on ECMWF forecasts; number of EPS-based simulations above EFAS High alert levels (EPS>HAL) and above EFAS Severe alert level (EPS>SAL) for each day of lead time (forecast range of 10 days).

When, instead of considering just one forecast date, EFAS forecast diagrams are built for an entire month, it is possible to visualize backwards what EFAS forecasted on each day and to detect the possible flood situations forecasted by the system on an event basis (Figure 2.11). This representation is particularly useful to detect persistence and consistency in the forecasts, for post-event analyses and case study analyses. They are also helpful in understanding the system's forecasts retrospectively and in investigating the way to better define decision rules to be applied in "real-time" (i.e., when a flood situation is being forecasted by EFAS) in order to improve the global performance of the system.

The example given in Figure 2.11 shows a series of DWD forecasts (7 days lead time). Based on simulations with observed meteorological input data, EFAS high alert levels were exceeded for two days within the period reported in the diagram: from the 15th to the 16th. The first time EFAS shows a signal of possible flooding (high levels) for the 15th-16th period is on the 12th, 4 days before the event. Since 2005 and up to now, any EFAS action is conditioned on *persistence*: only if a signal appears at least in 3 consecutive deterministic 12-hour forecasts the event is considered a “forecasted flood event”. This strategy aims to reduce false alarms and increase confidence in a forecast. The definition of the most suitable and effective decision rules, including for forecasts based on probabilistic EPS weather forecasts, is part of the ongoing research activities in EFAS.



Forecast Day	12	13	14	15	16	17	18	19	20	21	22
2006061200	🟩	⬜	⬜	🟥	🟨	🟨	⬜	▨	▨	▨	▨
2006061212	🟩	⬜	🟨	🟥	🟥	⬜	⬜	▨	▨	▨	▨
2006061300	▨	🟩	🟨	🟨	🟥	🟥	🟨	🟩	▨	▨	▨
2006061312	▨	⬜	🟨	🟥	🟥	🟥	🟨	🟥	▨	▨	▨
2006061400	▨	▨	🟩	🟥	🟥	🟥	🟥	🟨	🟨	▨	▨
2006061412	▨	▨	⬜	🟥	🟥	🟥	🟨	🟨	🟩	▨	▨
2006061500				🟥	🟥	🟩	⬜	🟩	⬜	🟨	▨
2006061512				🟥	🟥	🟩	🟩	🟩	🟩	🟨	▨
2006061600				▨	🟥	🟨	🟨	⬜	🟩	🟩	🟩
2006061612	▨	▨	▨	▨	🟨	🟨	🟩	⬜	⬜	⬜	⬜

Figure 2.11: Example of persistence diagram of EFAS forecasts based on DWD weather forecasts.

The use of non ambiguous terminology is always difficult when dealing with interdisciplinary subjects. Therefore, it is useful to clarify hereafter the following terminology used in EFAS:

Consistency: EFAS forecasts are considered consistent if flood forecasts based on two or more weather forecasts (DWD deterministic and ECMWF deterministic or probabilistic weather forecasts) show the same threshold exceedances.

Persistence: EFAS forecasts are persistent if from one forecast date to another, the same threshold levels are reached by the forecasted discharges. There are different types of persistence, e.g. persistence in DWD based forecasts only, persistence in ECMWF-based forecasts only or persistence in mixed forecasts. In the latter case, for instance, a river pixel may be simulated to exceed the EFAS high threshold based on DWD data today, while having exceeded the threshold in the previous forecast with ECWMF data.

References

Buizza, R., Richardson D.S., and Palmer T.N. (2003) Benefits of increased resolution on the ECMWF ensemble system and comparison with poor-man’s ensembles. Q.J.R. Meteorol. Soc., 129, 1269-1288

De Roo, A., Wesseling C., van Deursen, W. (2000) Physically-based river basin modelling within a GIS: the LISFLOOD model; Hydrological Processes, 14, 1081-1992

Gierk, M., Younis J., Szabo J., Kalas M., Bodis K., and van der Knijff J. EFAS – Status of data collection for the Ebe River basin and initial results of hydrological model calibration for the german Elbe on 1 km – from *Abstracts of the 2nd*

European Flood Alert System (EFAS) workshop, Ispra, 10-12 November 2004, European Commission, Joint Research Centre, S.P.I. 04.187,pp 55-58

Gouweleeuw, B., Reggiani P., de Roo A.P. (eds) 2004. A European Flood Forecasting System EFFS, Full Report, European Commission, EUR 21208EN

Gouweleeuw B.T., Thielen, J., Franchello G., de Roo APJ., Buizza R. (2005) Flood forecasting using medium-range probabilistic weather prediction, Hydrology and Earth System Sciences, 9(4), 365-380

Molteni , F., Buizza R., Palmer T.N. and Petroliagis TT. (1996) The new ECMWF ensemble prediction system: methodology and validation Q.J.R. Meteorol. Soc., 122, 73-119

Van der Knijff J., de Roo, A., Franchello G. (2004) The Lisflood model; from *Abstracts of the 2nd European Flood Alert System (EFAS) workshop, Ispra, 10-12 November 2004, European Commission, Joint Research Centre, S.P.I. 04.187, pp 35-39*

Wachter K., Kalas, M., Szabo J., Younis J., Niemeyer S., Gierk M., Bodis, K., de Roo, A. (2004) EFAS Pilotbasins – Description of the Danube river basin and status of data collection; from *Abstracts of the 2nd European Flood Alert System (EFAS) workshop, Ispra, 10-12 November 2004, European Commission, Joint Research Centre, S.P.I. 04.187, pp 40-43*

Chapter 3

Enduser perception and feedback of early flood warnings provided by EFAS

Jutta Thielen[1] and Maria-Helena Ramos[1,2]
[1]*European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, TP261, 21020 Ispra (Va), Italy*
[24] *Cemagref - Groupement de Lyon ; 3 bis quai Chauveau CP 220, 69336 LYON cedex 09 FRANCE*

3.1 Dissemination process of EFAS results

An important measure for the success of EFAS is the usefulness and actual benefit it represents for the receiving operational partner organisations. When analysing the potential usefulness of EFAS information one has obviously to bear in mind that although already being operated in pre-operational mode, EFAS is a research project and not an operational service.

Since 2005 EFAS information is disseminated in the following ways:

1) if EFAS simulates potential flooding more than 48 hours in advance and for an upstream area larger than 30000 km^2 in an area covered by an MoU, an EFAS information report is sent to the partner organisation.
2) If EFAS simulates potential flooding in an area covered by an MoU, but the flooding does not qualify within the MoU agreement, e.g. the flooding is for upstream areas smaller than 30000 km^2 or within the 48-hour range, then an informal email may be sent to the authorities informing them of the fact that EFAS is simulating a potential flood hazard and explaining why it does not qualify for a full report. This option was introduced on request of some receiving authorities and accepted during the 1st technical EFAS meeting in January 2006.
3) EFAS bulletin are issued typically bi-monthly, summarising the most important EFAS events during a two- to three-month period and reporting on system development issues.
4) Technical EFAS meetings are held once a year to discuss EFAS related issues with regard to system development, system performance, and future developments
5) Specific workshops or expert meetings may be held, e.g. on using EPS in flood forecasting in November 2005.
6) Presentations during scientific or technical conferences and meetings are encouraged.

The potential usefulness of EFAS information is monitored throughout the project, and again by using different means:

1) at the end of each flood event for which EFAS reports were issued, a feedback questionnaire is sent out to the authorities to inquire about different aspects of the provided information, e.g. questions if the stated information was correct, the information was presented clearly, how the information was used (Annex 1).
2) Free-format emails and other feedbacks are collected and individually analysed.

3) Personal communications are collected and documented, and during workshops or seminars, receiving authorities are encouraged to report on EFAS success rates and on how the information is locally being used.

In this chapter the feedback to EFAS Information reports for external flood alerts from July 2005 to June 2006 is summarised. The analysis is based on the evaluation of the returned feedback questionnaires and on other communications.

3.1 Summary of external flood alerts for June 2005-June 2006

The years 2005 and 2006 were flood-prone years particularly for the Danube catchment. The first EFAS information reports were sent out to Hungary and Slovakia at the beginning of July 2005 (Figure 3.1). During 2005, the EFAS team reported for 10 external flood events and a total of 50 reports were sent out during these events to authorities in Germany (Danube), Austria (Danube), Hungary (Danube), Slovakia (Danube), Bulgaria (Danube and others) and Italy (Po). If other MoU's were already signed at the beginning of the year, for example with Bulgaria or Romania, then more EFAS information reports would surely had been sent out during the year, since other European countries and river catchments were also affected by floods.

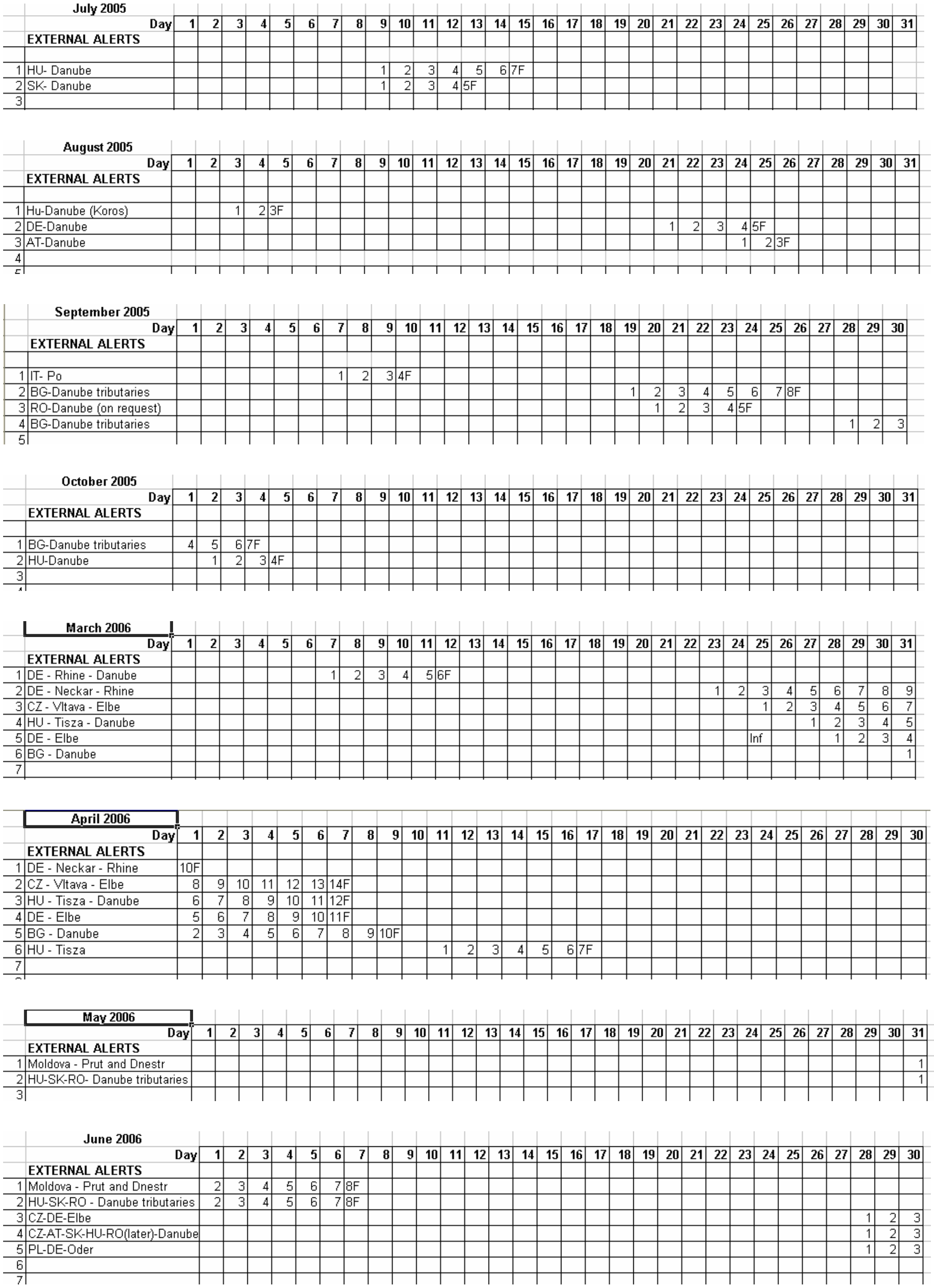
In 2006 record high flooding took place again in the Danube and Elbe river basins. Also other rivers such as the Rhine, the Odra, the Rhone and the Seine rivers had high waters and minor flooding in some upstream areas. From March to June, EFAS reported to 10 different organisations and sent out a total of 95 EFAS Information Reports to Germany (Elbe and Rhine), Slovakia (Danube), the Czech Republic (Elbe and Danube), Hungary (Danube and Tisza), Bulgaria (Danube) and Moldova (Prut and Dnestr).

Figure 3.1 gives an overview of the different full EFAS information reports that were sent since the first report was sent out in July 2005. During November 2005-February 2006 no external EFAS information Reports were sent out, although several internal EFAS alerts were active in areas where no MoU was established. In special cases, e.g. for floods in Austria and Romania, a special request was made to ECMWF in order to be able to report to the authorities. In these cases the flood crisis was already going on and the receiving partner organisations were aware of the EFAS research project and its features through prior workshops.

3.2 Summary of feedback to EFAS information reports from partner organisations

After the last EFAS information report, the authorities receive a feedback questionnaire (Annex 1). While verbal communications typically include subjective elements, the feedback questionnaires are the most quantitative measure for assessing the potential impact of EFAS information for the receiving authorities. However, some authorities prefer to communicate per email or personal communications, e.g. communications to the detached national experts or the EFAS team. This is particularly true in the wake of a flood crisis when the authorities have little time for additional work.

In the next sections the analysis of 7 feedback questionnaires are summarised in tables and the equivalent feedback through other means described. Feedback and quotations are kept anonymous.



July 2005

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	EXTERNAL ALERTS																															
1	HU- Danube									1	2	3	4	5	6	7F																
2	SK- Danube									1	2	3	4	5F																		
3																																

August 2005

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	EXTERNAL ALERTS																															
1	Hu-Danube (Koros)			1	2	3F																										
2	DE-Danube																					1	2	3	4	5F						
3	AT-Danube																								1	2	3F					
4																																

September 2005

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	EXTERNAL ALERTS																														
1	IT- Po							1	2	3	4F																				
2	BG-Danube tributaries																			1	2	3	4	5	6	7	8F				
3	RO-Danube (on request)																				1	2	3	4	5F						
4	BG-Danube tributaries																												1	2	3
5																															

October 2005

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	EXTERNAL ALERTS																															
1	BG-Danube tributaries	4	5	6	7F																											
2	HU-Danube		1	2	3	4F																										
3																																

March 2006

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	EXTERNAL ALERTS																															
1	DE - Rhine - Danube							1	2	3	4	5	6F																			
2	DE - Neckar - Rhine																							1	2	3	4	5	6	7	8	9
3	CZ - Vltava - Elbe																									1	2	3	4	5	6	7
4	HU - Tisza - Danube																											1	2	3	4	5
5	DE - Elbe																									Inf			1	2	3	4
6	BG - Danube																															1
7																																

April 2006

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	EXTERNAL ALERTS																														
1	DE - Neckar - Rhine	10F																													
2	CZ - Vltava - Elbe	8	9	10	11	12	13	14F																							
3	HU - Tisza - Danube	6	7	8	9	10	11	12F																							
4	DE - Elbe	5	6	7	8	9	10	11F																							
5	BG - Danube	2	3	4	5	6	7	8	9	10F																					
6	HU - Tisza											1	2	3	4	5	6	7F													
7																															

May 2006

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	EXTERNAL ALERTS																															
1	Moldova - Prut and Dnestr																															1
2	HU-SK-RO- Danube tributaries																															1
3																																

June 2006

	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	EXTERNAL ALERTS																														
1	Moldova - Prut and Dnestr	2	3	4	5	6	7	8F																							
2	HU-SK-RO - Danube tributaries	2	3	4	5	6	7	8F																							
3	CZ-DE-Elbe																												1	2	3
4	CZ-AT-SK-HU-RO(later)-Danube																												1	2	3
5	PL-DE-Oder																												1	2	3
6																															
7																															

Figure 3.1: Summary tables of external EFAS information reports from July 2005 to June 2006 (from November 2005 to Feb 2006 no EFAS Information Reports were sent out)

3.2.1. Correctness of EFAS forecasts

Table 3.1: Correctness of EFAS forecasts

Question	**yes**	**no**	**Don't Know**
For the river basins flagged in the EFAS reports flooding was observed?	6	2[*]	
For the river basins flagged in the EFAS reports the river level reached bankful conditions somewhere in the river basin?	5	1	
River levels were mainly:			
High and critical	1		
High but not critical	5		
Only moderately increased or normal	2		
Regarding timing, EFAS forecasts were			
Mainly correct in time	7		
Mainly too late			
Mainly too early	1		

[*] The numbers in table 3.1 do not necessarily add up to 7 reports because sometimes more than 1 river basin was mentioned in 1 report but addressed to the same authority. In this case, occasionally, two values were reported.

3.2.2 Usefulness and impact of EFAS Information Reports

Table 3.2: Usefulness and impact of EFAS information reports

Question	**yes**	**no**	**DN**
Do you find EFAS information reports useful	7		
Flood ensemble prediction system information is given in the form of maps counting the number of ensemble forecasts generating discharges exceeding critical flood level thresholds.	7		
Do you find this information useful?			
Were the EFAS reports used in some way by the flood forecasting team?	6	1 *(the reports arrived too late because of technical problems with the receiving mail server)*	
Did the EFAS reports effectively help you?	6	1 *(no earlier information than*	

from local sources)[*]

[*] with the new rules of dissemination the information (send reports to all authorities within the catchment and not only when the area of MoU is affected) this authority would have been informed 5 days earlier.

Table 3.3: Type of use of EFAS information reports

Question: How were the EFAS information reports used?	**Category[*]**		
Answers	**Early warning**	**Additional information**	**Decision making**
It was the first warning that focused our attention to the Drava river	X	X	
The reports were useful for the estimation of peak discharge		X	
We are using the reports as indication (the local 48h forecasts from Meteo-France Aladin and DWD HRM are used in a quantitative respect)	X	X	
We get an overview of the situation in the whole catchment, e.g. which tributaries are affected		X	
it is good to know which general development is predicted by EFAS	X	X	
EFAS reports are used to present the hydrological situation in the near days to institutes responsible for flood protection	X	X	X
EFAS reports were used as orientation information	X	X	
EFAS reports were used as support to create statement of development of flood situation		X	X

[*] *The category was estimated a posteriori and not ticked by the partner organisations*

Summary of other feedback (examples)
(See also minutes of 1st technical EFAS enduser meeting in Annex 2)

- "We knew already that something was going to happen. We used the EFAS information to know what was going on in the neighbouring countries."
- "We received the EFAS information reports when flood alerts had already been issued. If we could have received them a day or two earlier they would have been useful" [*note: this authority had not signed the MoU and received EFAS information reports only after the JRC had requested from ECWMF an exception from the agreement. The reports were sent out when the event was already ongoing*.]
- "Our service uses EFAS information as qualitative and indicative information. They are the only way to get information from other

weather services. For the next 48 hours we have our own quantitative information from the Aladin model".

- [*the EFAS team had informed a partner that EFAS was predicting potential flooding in the major tributaries, located in another country, to a river within their authority, but that the initial conditions in that river did not seem to correspond to reality – the answer was:*] "Any information is useful and welcome."
- "Thanks to your reports we have re-discussed and revised the week-end schedule for forecasters on duty."
- "We are only at the beginning of this project. We will have to gain experience with the system and evaluate its benefits in operational terms with time. So far, we found the additional information useful for us."
- "Thank you very much for the support during the flood crisis with the EFAS information reports. In particular, they were helpful when discussing with Politicians" (*translation*).
- "Leadtime gain >=1 day. Very useful for operational forecast. Correct forecast of 2nd wave."

With no doubt, all organisations that filled out the questionnaires found EFAS information useful. This is also reflected in the list of other comments. For many organisations, the EFAS report is the first quantitative indication of the possibility of flooding. EFAS information helps as orientation and preparation and serves typically as additional information. According to the 1st technical EFAS meeting, the EFAS information reports are used mostly to stimulate discussion within the teams and to check their own forecasts, often done more carefully and in more detail.

It can be concluded from the above results, and taking also into account discussions with the EFAS partners during the 1st technical EFAS workshop and on other occasions, that the use of the EFAS information reports depends largely on the leadtime:

a) first reports usually have a long lead-time and lead to a raised awareness that a flood might take place. The local forecasts have not captured the event yet and EFAS reports lead to discussions, raised awareness and preparedness. Time tables and weekend shifts may be revised.
b) If the reports are confirmed in subsequent days, the reports are used to roughly localise the area and to intensify local flood forecasting analysis. EFAS is used more and more as additional information and to "extrapolate" the situation beyond the short-term local forecasts – e.g. to know if a second flood wave is to be expected or not, if the flooding will spread also to other areas, etc.
c) When the events are ongoing, the EFAS information reports are used as additional information that can otherwise not be received, e.g. rainfall information from another weather service or possible discharge contributions from tributaries in other countries. It can also assist in the decision making, probably mostly when it is in line with the local forecasts. Otherwise, EFAS information is used to assess the tendency of the flood situation for the days beyond the ranges of the local flood forecasting systems.

3.2.3 Trading off early alerts for false alarms

Weather forecasts become increasingly uncertain with increasing leadtime. Thus any flood simulations based on these weather forecasts have a higher probability of having false alarms or missed events the longer the leadtime. It is the challenge of EFAS to provide meaningful early warning and reduce the number of false alarms (Table 3.4).

Table 3.4: EFAS strategy to reduce false alarms

Question	**yes**	**no**	**Don't know**
The first EFAS report is produced when high or severe flood threshold levels have been exceeded at least for 3 consecutive EFAS forecasts. This can decrease the false alarm rates but also shortens the leadtime. Would you like to receive the reports earlier although this may mean an increase in false alarm rates?	6	1	
At present, the first report is sent when at least one of our meteorological forecasts (DWD and ECMWF) results in high severe flood levels. Would you prefer to receive the first EFAS report only when **both** meteorological forecasts indicate them?	1	6	

Summary of other feedback (examples)

- "We would like to know as soon as possible. It makes us aware of the possibility of a flood and we monitor our own systems more closely. If nothing happens, no damage is done."
- "We rather receive a few false alarms too many than miss a true alert. Obviously, if alerts come every day, this is not useful anymore and the alert-rules would have to be revised."
- "the earlier the warning the better"
- "We [the receiving partner] would like to receive all the information to get experience with the system's performance. The decision if it is a true alert or a false alarm should be made by the receiving partner, not by the EFAS team."

The question on trade-off between false alarms and early warning is quite a complex issue. The information summarised in Table 3.4 clearly shows that the receiving authorities would trade off a higher false alarm rates to the benefit of earlier warning. This has also repeatedly been confirmed in personal communications on the subject (see examples in the summary of quotations above) and has been discussed during the 1st technical EFAS meeting on the 23rd January 2006 (minutes in Annex 2). This result may be surprising when thinking in terms of the "cry wolf"-effect: if a system issues too many false warnings the warnings become meaningless. In fact, the second point in the summary list above is reflecting this. However, it is also clear that the receiving local authorities understand the EFAS information as a possibility to become earlier aware of the possibility of a flood to happen: it does not force them to take decisions and therefore does not have adverse effects (see discussion in the introduction of this report).

The last item in the summary of other feedback reflects the view of several partner organisations. They would like to receive the EFAS information unfiltered, daily, and then decide themselves how to do deal with them. One of the main arguments is that only by receiving daily the EFAS information they will get sufficient experience with the performance, bias, weaknesses and strengths of the system. This also implies that the partner organisations would not like to have the constraint on 30000 km^2 upstream area, as is currently imposed by the Memorandum of Understanding. This upstream area threshold was imposed by ECWMF to avoid a high number of false alarm rates for small upstream areas and based only on criteria related to the meteorological modelling framework (model grid size, time step). But again, the discussion with the receiving EFAS partner organisations showed clearly that most EFAS partners do not share this concern: in case of the probability of a flood most of them would like to have EFAS information also for smaller upstream areas – with the argument that floods are typically produced by the confluence of smaller tributaries and that it is important to know where the source of the water comes from.

3.2.4. Using Ensemble Prediction System information

Basing flood forecasting on EPS is relatively new in operational flood forecasting and few organisations currently deal with the topic. The JRC organised a workshop in November 2005 to explore together with flood forecasting experts from the Member States the usefulness of EPS information implemented in EFAS for operational flood forecasting and decision making, as well as their perception of uncertainty in flood forecasting. The workshop's concept was to have a small group of flood forecasters from different river basins working through a number of case-studies, each one representing a potential flood situation as forecasted by EFAS. On the first day, the participants worked in groups on each case-study. The second day was targeted mostly to plenary discussions on the use of meteorological EPS for ensemble flood forecasting. The results of this workshop are described in a EUR report by Thielen *et al.* (2005), and can be downloaded from the EFAS webpage (http://efas.jrc.it). Results on the use of EPS in EFAS were also shown in several conferences so far (Thielen *et al.*, 2005; Thielen *et al.*, 2006a, 2006b; Bartholmes *et al.*, 2006; Ramos *et al.*, 2006a, 2006b, 2006c; Watcher *et al.*, 2006) and are the subject of publications currently under preparation on quantifying and communicating uncertainty in flood forecasting.

The feedback questionnaires to EFAS information reports also contain questions on EPS. They are summarized in Table 3.5.

Table 3.5: Using EPS in flood forecasting

Question	**Yes**	**no**	**Don't Know**
Flood ensemble prediction system information is given in the form of maps counting the number of Ensemble Forecasts (EPS) generating discharges exceeding critical flood level thresholds. Do you find this information useful?	7		
Would you like to have more information on Flood Ensemble Prediction System, i.e., more EPS results as, for instance, the spread of the forecasted hydrographs, statistical information for each leadtime, or others?	3	4	
If Yes, which information?	EPS spread		

Summary of other feedback (examples from *the workshop on EPS*)

- "The most important information for us was the EPS. Because it was very comfortable to say 'yes we are sure because the most of the EPS forecasts say the same information' (…). We are not afraid to say this".
- "It was helpful to have the EPS to make the decision [but] not always".
- "If the EPS shows the same direction like the other forecast, it makes us more confident".

One of the most important results of the workshop on EPS was that the participants felt that more background knowledge and experience was needed to understand and work with probabilistic flood forecasting.

Overall the discussions on EPS show that the receiving authorities are very open to new ways and methods to assess and describe flood hazard in a probabilistic way, but are still not sure on how to use the information. It is clear that there is a need to visualise the EPS results in a clear and simple way. Workshops, such as the one organised by the JRC in November 2005, was welcomed as a training possibility.

3.2.5 Quality aspects of the reports and its dissemination

EFAS information reports have undergone some changes since the first layout that was proposed to the ECMWF council and that was finally sent to the partner organisations. Obviously such reports need to be adapted to the needs of the partners.

Table 3.6: Editorial aspects of EFAS Information Reports

Question	**Yes**	**no**	**Don't Know**
Is EFAS information clearly stated	6	1	
Is all information necessary	7		
Would you suggest improvements of the EFAS reports	1	6	

Table 3.7: Timeliness and dissemination of reports

Timeliness and dissemination of Report			
You consider 16:00 a good time to receive forecasts?	Yes: 1	Too late: 3	Does not matter: 3
The forecasters receive EFAS reports	Directly: 5	Via technical contact: 2	
The forecasters receive EFAS reports	At the time it was sent: 5	Within the following 24 h: 1	Later than 24 h: 1

Examples of other feedback

Editorial aspects

- "EFAS maps are too small"
- "It is difficult to find an orientation in the maps without the river network"
- "The templates are too repetitive – there are too many captions stating the same"
- "The river basin should be mentioned on the first page"
- "The overview map is not necessary"
- "The overview map is very useful"
- "The colour schemes are confusing"

Timeliness and dissemination

- "We received the reports only from Nr. 5 onwards, but then our own services were already alarmed. We would like to receive the EFAS information reports from the very beginning"

In some cases suggestions of the partner organisations can be immediately incorporated into the EFAS information reports, e.g. following comments such as "*the maps are too small*", "*there are too many captions*" or "*this particular colour is difficult to see*", etc. In other cases the consent of all partners was sought, e.g. during the 1st technical EFAS meeting in 2006. For example, some partners find the overview map in the report useful, others find it confusing (see points 5 and 6 on summary above). Most changes were incorporated in July and August 2005 following the first EFAS information reports but since lately there are only minor changes requested.

One issue was discussed during the 1st technical EFAS meeting for which, so far, no solution has being proposed in the EFAS information reports: EFAS information reports are combining a lot of different information for which only a limited colour palette is available and the printing of the reports have to be taken into account as well. One example is the map showing the number of EPS above high and severe EFAS threshold. At present the colour scheme ranges from blue (low number) to red (high number). Some partners found this confusing because in fact, what should logically be shown are just different intensities of red for the numbers of EPS above high alert threshold (flagged as red in deterministic forecasts) and different intensities of purple for the numbers of EPS above the severe threshold (flagged as purple in deterministic forecasts). However, this colour scheme is difficult to visualise on a map and often prints very badly. The JRC is still looking into a solution for this, but in the meantime the maps are shown on different pages to avoid misunderstandings.

It can be concluded from the feedback questionnaires, as well as from verbal feedbacks, that in general the EFAS partners are satisfied with the way EFAS information is provided.

With regard to the dissemination of the reports, it can be concluded that most EFAS information reports are usually directly received by the flood forecasters on duty. Only in a few cases – and often this is the case when EFAS information reports are received for the first time – the reports are distributed to the technical contact point first and then forwarded to the forecasting team. In case the contact point is not on duty or absent, this would cause a delay. Typically this is then changed by requesting EFAS information reports to be sent to operational addresses.

Two main problems were identified regarding the timing of dissemination:

First, at present, the latest EFAS forecasts based on the midnight weather forecasts are ready at around 13:00 local time for analysis (the 00:00 DWD weather forecast data are received at 7:30 local time, the 00:00 ECMWF deterministic weather forecasts at around 09:15 and the latest EPS data files at around 10:45). With the present computing environment it takes about 2 hours to process all the incoming data, to run the model and to produce the analysis maps. Although improvements to the EFAS user interface allows now a relatively fast analysis of the flood situation, detailed analysis during a flood event, decision making, compilation of the reports and cross checking still takes on average around 30-45 minutes per report. On working days several people can work in parallel but during the weekend typically only 1 or 2 persons are on duty. Thus if 5-6 reports have to be produced, as was the case during the widespread flooding in summer 2005 and again spring 2006, then the last reports are usually only sent around 16:00-17:00 (local time at the JRC). If the dissemination of EFAS information can be automated, e.g. through a specific and password-protected web page, then this could represent a gain of 4-6 hours for the 00:00 forecast. More importantly, in this case, also the 12:00 forecasts could be processed automatically and put available for the receiving authorities – thus giving a gain in lead time of at least 12 hours.

Second, concerning the last point on the summary list above: as long as only few partner organisations had signed the MoU, the EFAS information reports were not sent immediately to all organisations belonging to a river basin, but it was only sent to those partner organisations for which EFAS simulated the exceedance of a high alert. In other words, downstream organisations would only get the reports if previous upstream flooding was spreading downstream. Since most of the floods that took place in 2005 were rather local floods, this did not cause any problems. For the snowmelt floods in 2006, however, a revision of the dissemination strategy would have been beneficial: the downstream Elbe authorities could have received, in fact, the 1st EFAS information report that was sent to the Czech Republic authorities 5 days earlier than the time they actually received a first report. In this case the EFAS information report would have come earlier than they had received the first indication of flooding. Since now most authorities in the major catchments have signed the MoU, the dissemination strategy was revised and EFAS information reports are now sent immediately to the downstream partners as well.

3.3. Conclusions on enduser perception

Feedback from partner organisations on the EFAS information reports is generally very positive. It appears that the hit rate of EFAS for forecasting flood events and/or bankful conditions is considerably higher than its false alarm rate. The receiving partner organisations are glad with the EFAS information reports as well as with the quality of the reports. The EFAS information is used actively by most organisations as additional information for orientation and occasionally even for the decision making process. The use of the reports seems to depend largely on the leadtime. As pre-warning information EFAS reports mainly serve to increase the preparedness at the local water authority level. As soon as the event takes place, it serves as additional information, as outlook for second flood waves and also as support to the decision making process and discussion with civil protection authorities (and politicians). Partner organisations are very keen to gain more experience with EFAS information reports.

Improvements are needed in the dissemination of the EFAS information reports. Earlier dissemination by sending all reports directly also to all downstream authorities (even if their administrative borders are not yet affected) could increase the preparedness for flood events in particular in the downstream areas. This strategy has already been adopted in the latest flood event. Further, more automation in the production of the EFAS information reports and putting daily information on a web-restricted webpage will probably gain valuable time in the future.

Chapter 4

EFAS-EPS forecasts: in-depth case study analyses and statistical evaluation of summer 2005 and spring 2006 flood forecasts in the Danube catchment

Maria-Helena Ramos[1,2]
[1]*European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, TP261, 21020 Ispra (Va), Italy*
[2] *Cemagref - Groupement de Lyon ; 3 bis quai Chauveau CP 220, 69336 LYON cedex 09 FRANCE*

4.1. Introduction

This study essentially deals with the analysis and interpretation of case studies with a view to defining rules to an objective and statistical evaluation of EFAS-EPS forecasts. It investigates EFAS-EPS forecasts on a flood-event basis and the ability of the system to provide early warning, comparatively and/or complementarily to EFAS forecasts based on deterministic weather forecasts. Post-event analyses are important to understand the potential capabilities of EFAS results to forecast different types of flood events in different geographical conditions (river basins). They allow to identify interpretation and decision rules for individual forecasts (deterministic and EPS forecasts), as well as for combined forecasts (using the ensemble of all EFAS forecasts in a complementary way). This study focuses on the Danube river basin, the European catchment most affected by floods in 2005-2006. The analyses consider mostly the weather input from ECMWF (weather forecasts generated by the deterministic model and the Ensemble Prediction System) and, for simplicity, sequences of 12:00 forecasts. Flood events as forecasted by EFAS are illustrated and the statistical evaluation of hits, misses and false alarms is assessed on a seasonal basis (summer 2005 and spring 2006). The study also presents a first evaluation of gain in preparedness (early forecasting) from EPS-based forecasts comparatively to the deterministic ECMWF-based forecasts.

4.2. Overview of the Danube floods in 2005

In the year 2005 the Danube river basin faced a high number of devastating floods, causing several deaths, a thousand of evacuations and heavy economic losses. The main causes were related to snow melting in mountainous areas, heavy rainfall events and continuous precipitation leading to saturated soils. The 2005 Flood Archive of the Dartmouth Flood Observatory[1] reported seven major flood events or high river level occurrences in the Danube River Basin. We combined this register with additional information from national authorities and from the JRC-EMM European Media Monitoring[2] to produce a summary overview of the most important flood events in the Danube and its tributaries in 2005 (Figure 4.1). The most severe events were observed in Austria and Germany on 21st-26th

---

[1] http://www.dartmouth.edu/~floods/Archives/.
[2] http://emm.jrc.it

August (peak discharges with estimated recurrence interval greater than 100 years – De Roo *et al.*, 2006) and in Romania and Bulgaria, where widespread floods and historic high levels occurred due to heavy and recurrent rainfall during July-August-September 2005. The Danubian events provided several case-studies to investigate the use of EPS in EFAS and the additional value of EPS-based forecasts to increase preparedness in flood forecasting (Ramos *et al.*, 2006a, 2006b).

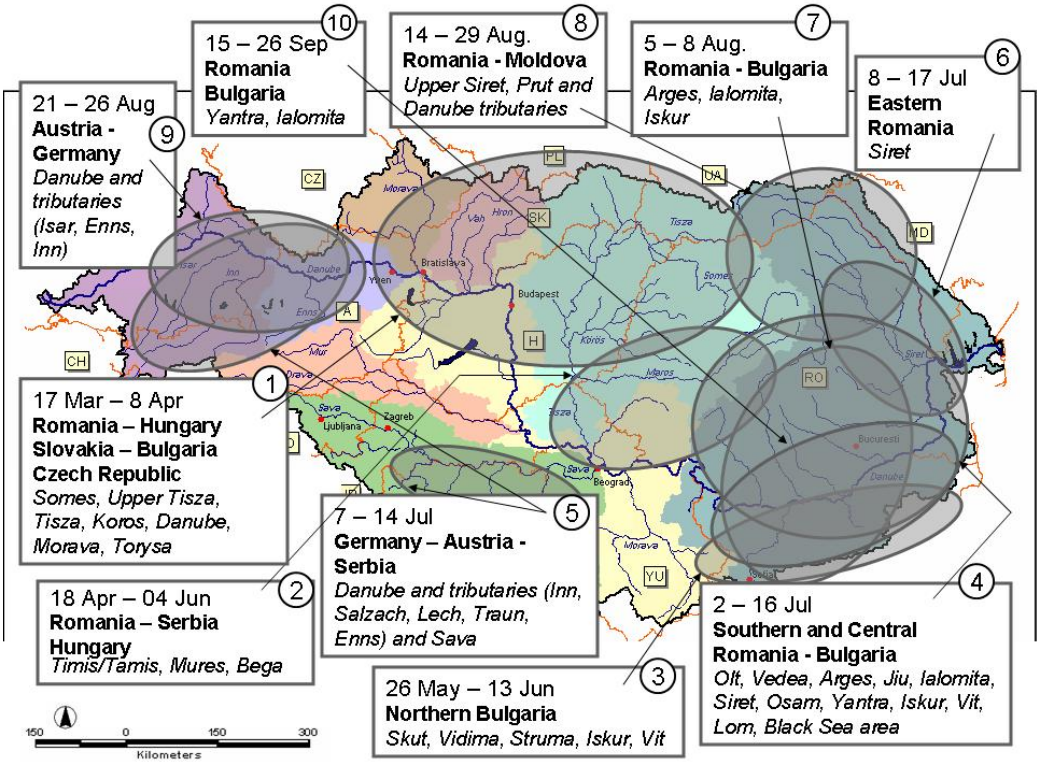


Figure 4.1: Summary of major flood events or high water level occurrences in the Danube River Basin during 2005 (based on the Flood Archive of the Dartmouth Flood Observatory, on information from national authorities and on the JRC-EMM European Media Monitoring).

4.3. EFAS results for the Danube river basin in July-Aug. 2005

4.3.1 Overview maps

Figure 4.2 shows an overview of the number of days EFAS simulations based on observed meteorological data exceeded the EFAS high alert level during July and August 2005. In comparison to the observed situation illustrated in Figure 4.1, it can be seen that the system captured well the areas of flooding during these months.

It can also be interesting to compare these "proxy-results" used as a good assessment of observed high water levels against EFAS-EPS forecasts. Figure 4.3 presents an overview of the number of early EFAS forecasts (forecast lead time of 5-days) based on ECMWF-EPS weather forecasts (issued at 12:00) showing more than 50% of simulations with discharges above EFAS High alert level for the months of July and August. In practice, for every day of forecast, we check if the number of EPS-based simulations is greater than 25 (~ 50% of 51 EPS simulations) for the 5-day lead time. The criteria chosen are therefore, in general, more adapted for capturing important flood events in river catchments associated with relatively great upstream areas.

It is clear from Figure 4.3 that the ensemble based system identified well the flooded areas during the two months investigated. When comparing the maps in Figures 4.1 and 4.2 to those in Figure 4.3, one can clearly see a correspondence between the areas where observed floods were reported and the areas where more than 50% of EPS-based simulations forecast the possibility of a flooding with a 5-day lead time. The areas affected by floods in Romania and Bulgaria show up clearly in the overview maps: for July (Figure 4.3a), the Danube tributaries in Bulgaria and the Siret tributaries in Romania show the highest number of occurrences (possibly corresponding to the events 4 and 6 highlighted in Figure 4.1), while for August (Figure 4.3b), Danube tributaries of Southern Romania and Upper Siret can be distinguished (possibly corresponding to the events 7 and 8 in Figure 4.1). Other areas show also some early EPS signal, like tributaries in Austria for the July map.

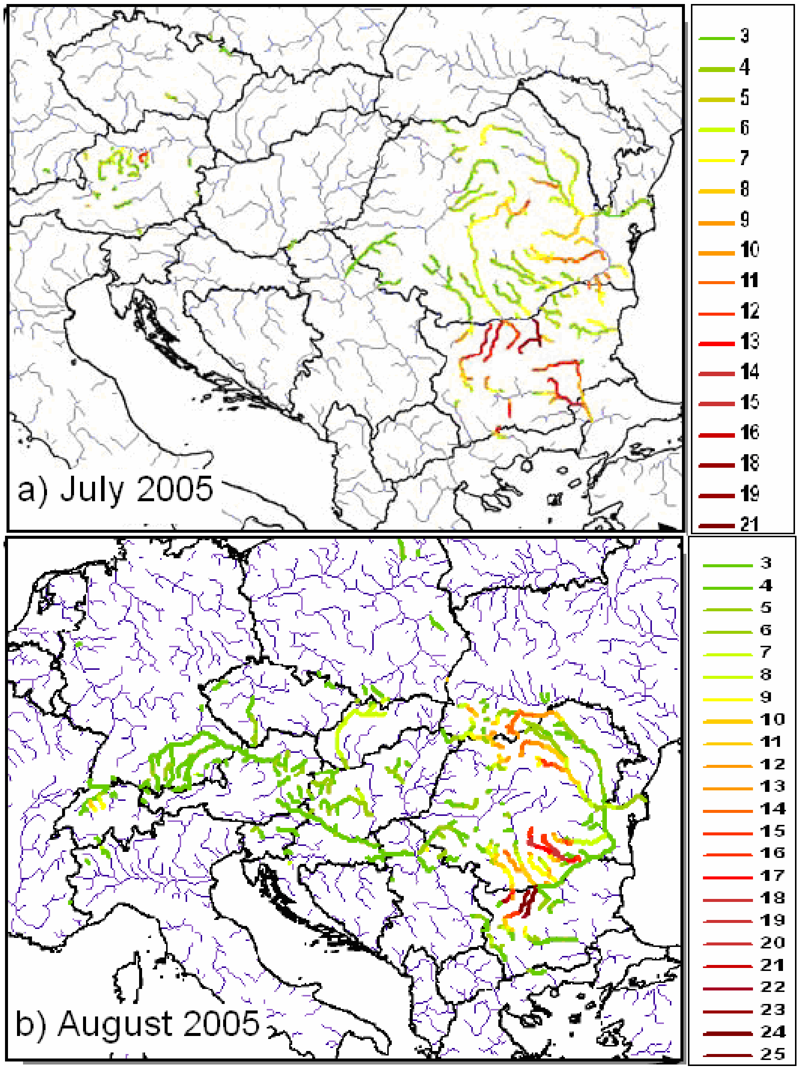


Figure 4.2: Number of days EFAS simulations based on observed meteorological data exceed the EFAS High alert level during a) July and b) August 2005.

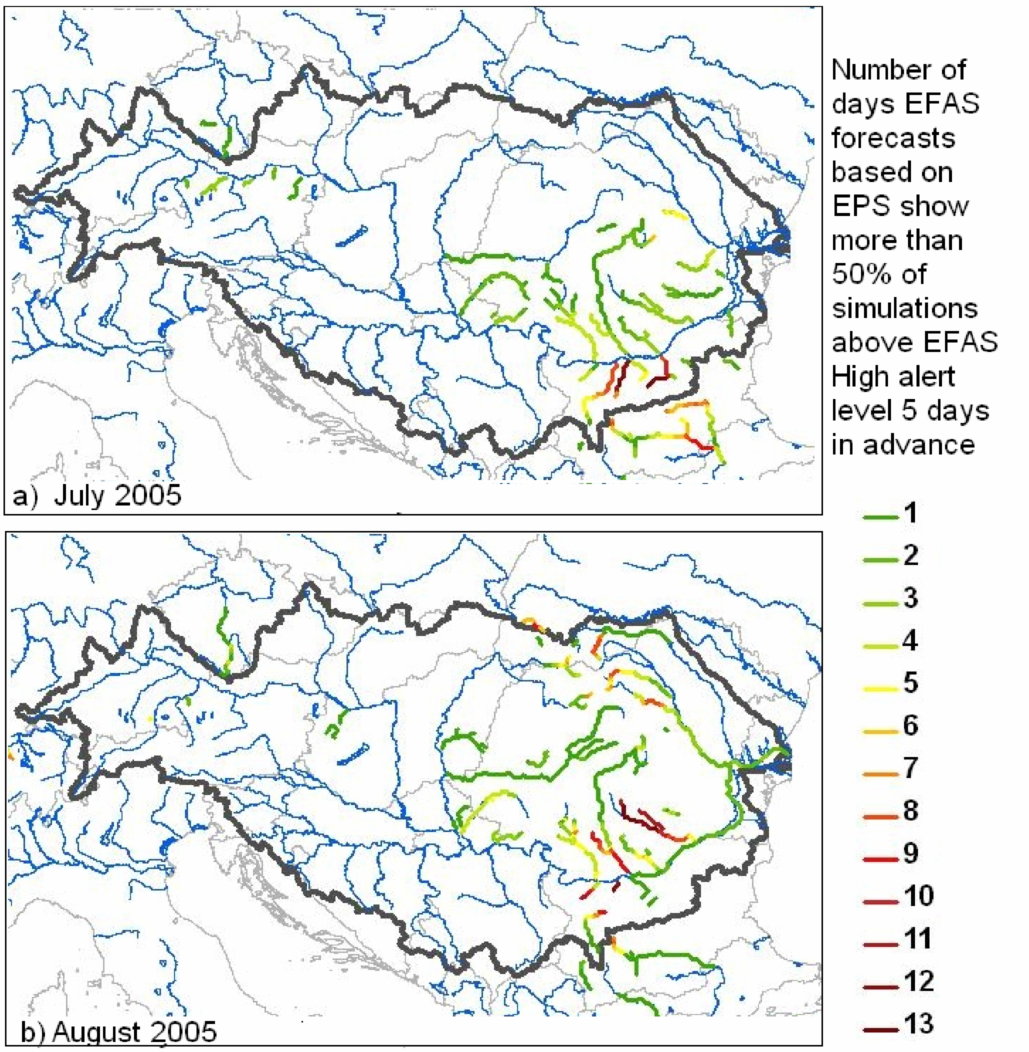
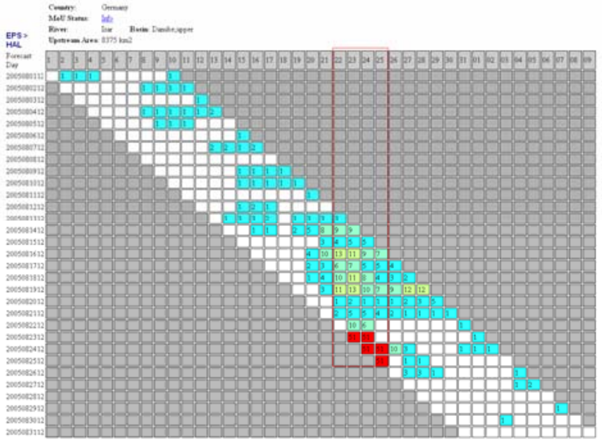


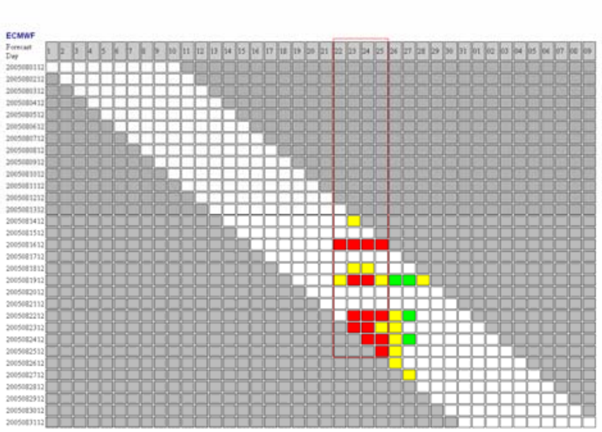
Figure 4.3: Number of days EFAS forecasts based on ECMWF-EPS weather forecasts show more than 50% of EPS-based simulations with discharges above EFAS High alert level for a lead-time of 5 days and a) July 2005 and b) August 2005.

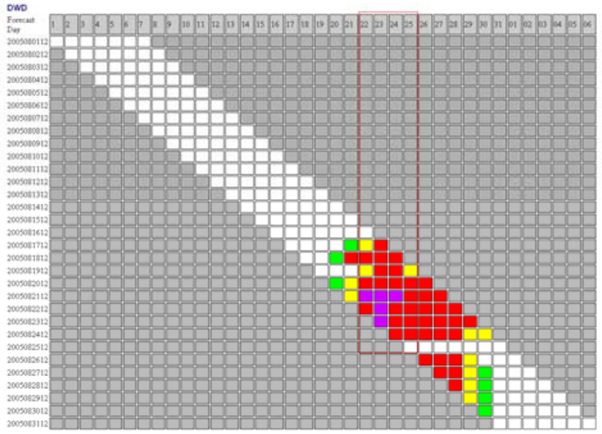
4.3.2 EFAS temporal diagrams at different river locations

In order to illustrate EFAS forecasts at different points in rivers and for hydrological events occurring on July and August 2005, we built, for each month, historical EFAS forecast diagrams at selected locations (Figures 4.4 to 4.7). These diagrams highlight some important features of EFAS forecast procedure and results, as commented in each figure presented.

Location 1 – Isar river in Germany (8375 km^2)







Example: Intermittent forecasts

One observed and forecasted event from 22nd-25th August 2005.

This example shows the difficulty in dealing with non-persistent forecasts. Although the ECMWF based forecasts indicate potential flooding as early as the 16th July, thus a 7-day leadtime, the signal is not persistent over the following forecasts. The combination of the deterministic with the probabilistic information gives more confidence in the signal, especially when the EFAS high thresholds are again exceeded on the 19th July (lead time of 4 days).

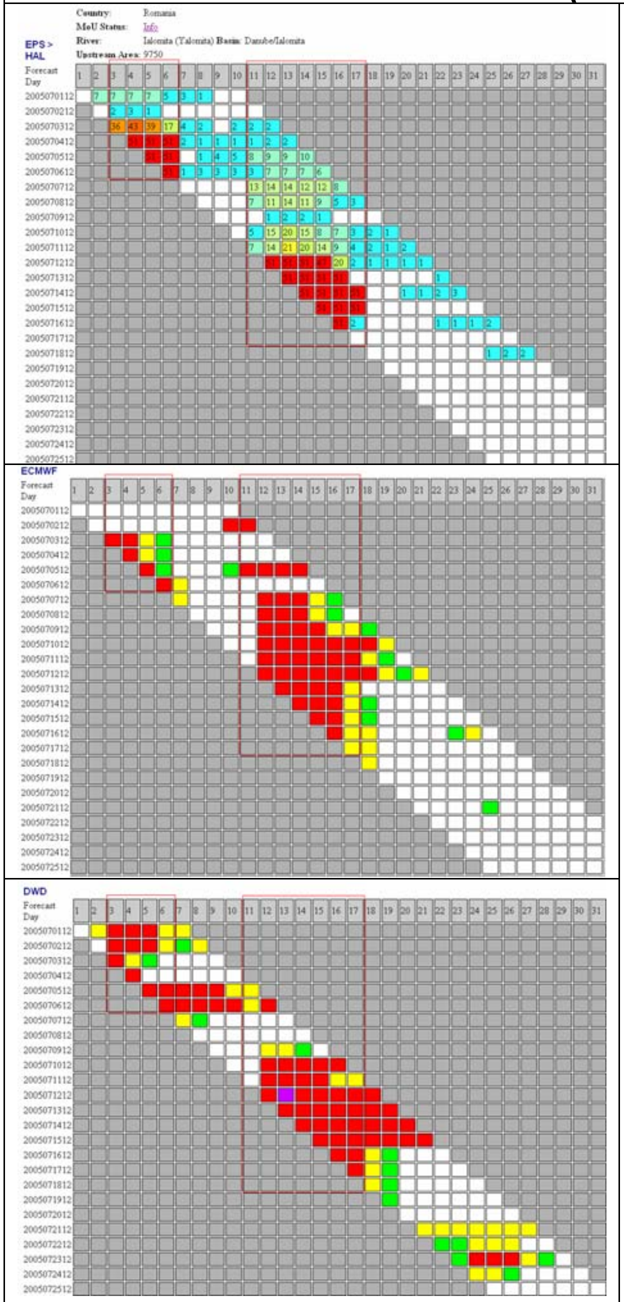
Using the EFAS procedure to base decisions on persistence in both deterministic forecasts (ECMWF and DWD) and using the EPS forecasts as additional information, the flood event in this case can be forecasted with a lead time of 5 days.

This example also shows a number of occurrences of 1 to 3 EPS simulations above high alert levels that are not associated with observed high levels (and also not associated with deterministic-based forecasts). Following the current EFAS decision rules on flood forecasting, these occurrences would not constitute a forecasted event – thus not counting as "false alerts".

The importance of defining a "forecasted event" before counting hits and misses in EFAS forecast verification is here highlighted.

Figure 4.4: History of alert levels forecasted by EFAS at Isar river in Germany (8375 km^2) for August 2005. Rows go from 01st to 31st August 2005 and correspond to EFAS simulations based on weather forecasts issued at 12:00. Columns correspond to the dates for which the forecast applies (each box is a 24-hour lead time). Red rectangle: dates for which high flood levels are simulated by EFAS with observed meteorological data.

Location 2 – Ialomita river in Romania (9750 km^2)



Example: Persistent and consistent forecasts

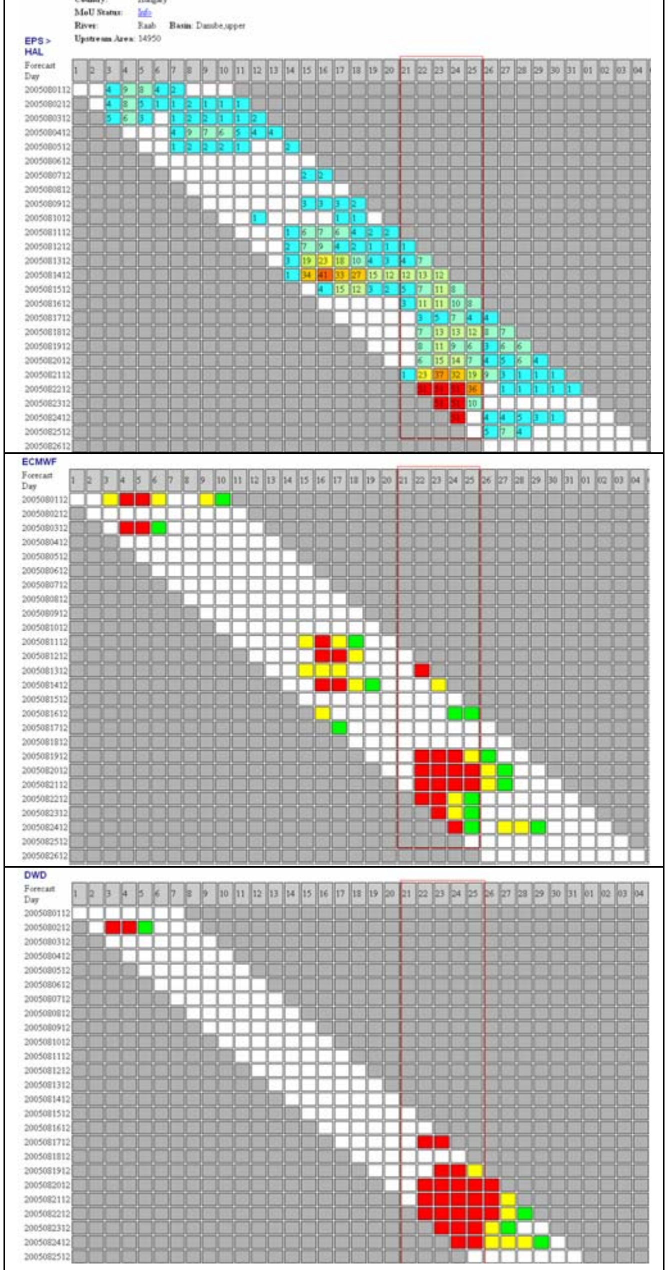
Two observed and forecasted events from 3rd-6th July and 11-17th July 2005.

For the first event, the condition of having at least 2 consecutive forecasts before issuing an alert and both deterministic-based forecasts reduces significantly the effective lead time for EFAS alerts. Lead times within the 48-hour range are not qualified for an early warning in EFAS. In this case, the combination of EPS-based forecasts and deterministic-based forecasts is crucial for the early forecasting of the flood event.

For the second event, deterministic ECMWF-based simulations show a non-persistent signal of high levels on the 02nd and 5th July. Only from the 7th July onwards the signal becomes persistent. For the DWD-based forecasts, persistence is only achieved from 10th July onwards. In this case EPS information can again contribute to earlier preparedness. The signal in EPS-based forecasts starts already on the forecast of the 03rd July, although it shows only 1 to 2 forecasts above EFAS high levels. In simulations from the 05th July onwards, however, the results show more than 5 simulations (out of 51) above EFAS high level.

Figure 4.5: History of alert levels forecasted by EFAS at Ialomita river in Romania (9750 km^2) for July 2005. Rows go from 01st to 25th July 2005 and correspond to EFAS simulations based on weather forecasts issued at 12:00. Columns correspond to the dates for which the forecast applies (each box is a 24-hour lead time). Red rectangle: dates for which high flood levels are simulated by EFAS with observed meteorological data.

Location 3 – Raab river in Hungary (14950 km^2)



Example: Hits and false alarms

One forecasted but not observed event from 16th-17th August and one observed and forecasted event from 22nd-24th August.

One can show in this example a false alert forecasted by EPS-based simulations in combination with the simulations based on the deterministic ECMWF weather forecasts. If one considers that more than 5 simulations above high levels in EPS-based simulations during two consecutive forecasts is enough to obtain persistence and define a potential flood event, one would forecast an event for the 15th-16th August. The warning would be issued on the 13th, with a signal already since the 11th and with the support of two consecutive high levels forecasted on the simulations based on ECMWF deterministic weather forecasts. With the current decision rules implemented in EFAS this would not, however, let to an alert because there is no support from the deterministic forecasts based on DWD weather forecasts and no persistence on the ECMWF forecasts from 13th onwards.

Here, one can also see a situation where EPS-based forecasts merge into ECMWF-based forecasts as the forecast date approaches the date of the forecasted event: on 15th, for 16th -17th, the EPS signal is weaker and from 16th onwards it disappears, while ECMWF-based forecasts only show low to medium levels forecasted.

Figure 4.6: History of alert levels forecasted by EFAS at Raab river in Hungary (14950 km^2) for August 2005. Rows go from 01st to 31st August 2005 and correspond to EFAS simulations based on weather forecasts issued at 12:00. Columns correspond to the dates for which the forecast applies (each box is a 24-hour lead time). Red rectangle: dates for which high flood levels are simulated by EFAS with observed meteorological data.

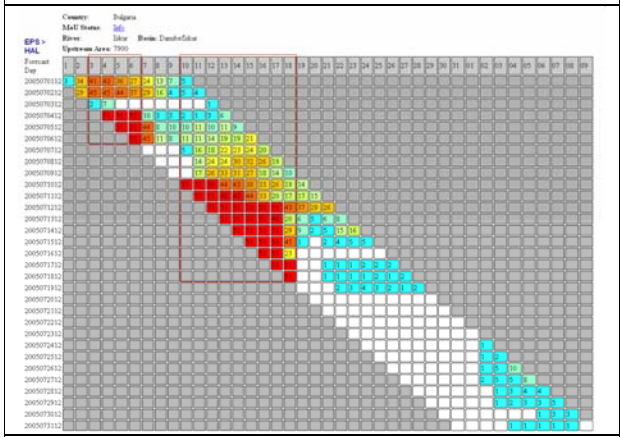
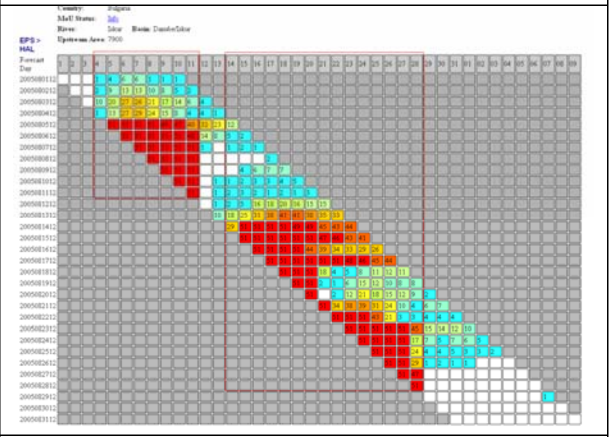
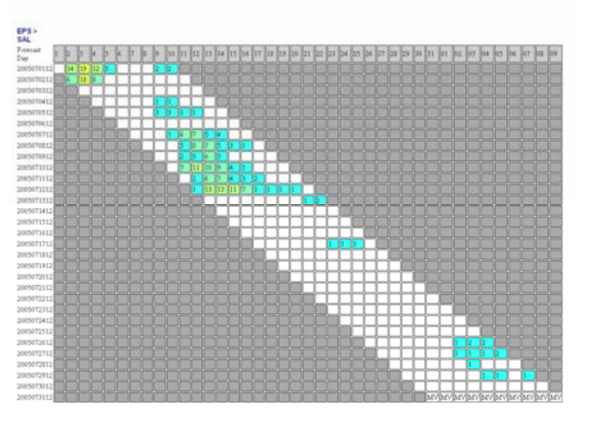
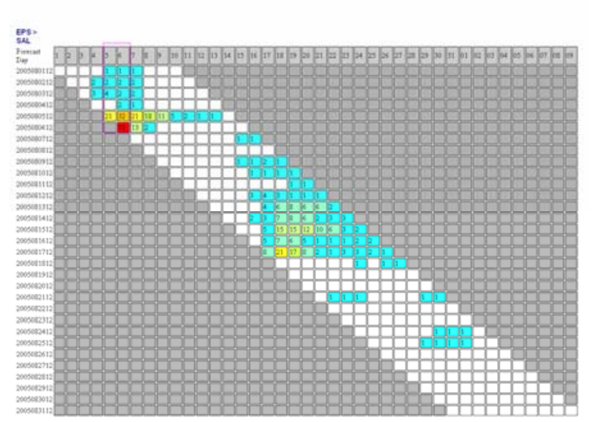
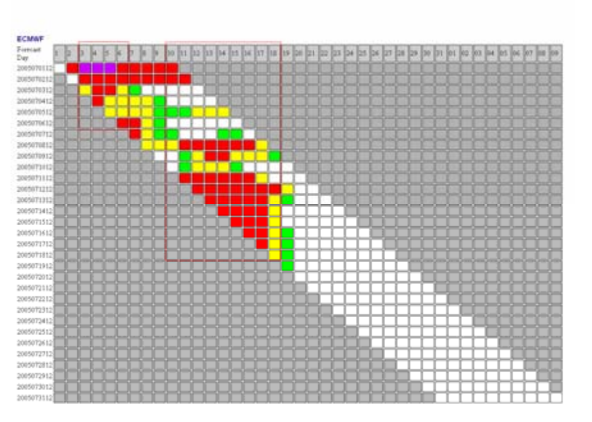
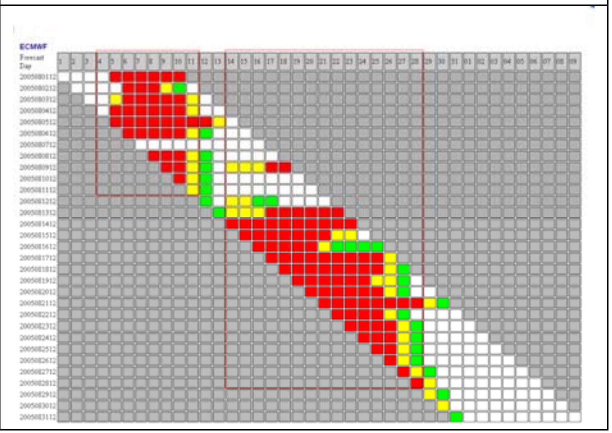
Location 4 – Iskur river in Bulgaria (7900 km^2)	
JULY 2005	***AUGUST 2005***
	
	
	

Figure 4.7: History of alert levels forecasted by EFAS at Iskur river in Bulgaria (7900 km^2) for the months of July (left) and August (right) 2005. Rows go from 01st to 31st and correspond to EFAS simulations based on weather forecasts issued at 12:00. Columns correspond to the dates for which the forecast applies (each box is a 24-hour lead time). Red and pink rectangle: flood event as simulated by EFAS with observed meteorological data (exceedances of EFAS High and Severe levels respectively).

Figure 4.7 illustrates the exceptional flood prone period of July-August 2005 for the Iskur River in Bulgaria in EFAS forecasts. Feedback from the river authorities in Bulgaria confirmed that during 5th-8th of August serious floods took place in the Iskur catchment, whereas during the other periods the river discharges were high and reaching bankful conditions (corresponding approximately to EFAS High alert levels) but that flooding did not occur.

From the historical EFAS forecast diagrams presented (Figures 4.4 to 4.7), it is possible to capture several features of EFAS forecasts such as:

- **Event definition:** forecasted events of potential flooding can be detected in the diagrams. In cases when forecasts are clearly persistent they can be easily defined. At other situations, a signal of forecasted alert levels can be visualized, but an event can not be well-defined. So far, an event in EFAS consists of grouped high or severe alert levels and not on isolated exceedances of these levels.

- **Persistence in forecasts:** a number of criteria can be used to establish that a situation has been forecasted persistently. EFAS alerts are only useful if delivered to users with more than 2 days of lead time. Therefore waiting too long to issue an alert would reduce the preparedness and consequently the usefulness of the information. This means that it is more acceptable to have a number of false alerts than to issue alerts too late that will not give significant additional information to the local forecasters. Persistence in EPS-based forecasts is not only related to continuous exceedances of EFAS high levels (as it is for the deterministic case), but also on having a steady number of EPS members above the alert threshold. So far, there is not a rule established in EFAS for persistence in EPS-based forecasts. Ongoing studies are investigating the subject.

- **Misses and complementary forecasts in EFAS:** one of the advantages of the system is that it deals with different meteorological forecasts as input. This helps the EFAS forecaster to gain confidence on whether send an alert to the users or not. In some cases an event can be missed by one of the simulations, but EFAS forecasts can still be based on the other two. This complementary aspect of the system can help in preventing EFAS from missing events, especially in situations where the forecasts are not very steady from one day to another.

It is the challenge of EFAS research studies to extract ***decision rules*** from these features to better define alert situations that produce meaningful early flood alerts, with a small number of false alarms, and reduce the number of missed events as much as possible. The case-study approach presented in this chapter is useful for the identification of possible best rules. Based on the identified features, a more quantitative approach, as described in the next section, can then be applied to verify the proposed rules in an objective way and against a large data sample on European scale and over long time periods.

4.4. EFAS-EPS forecast verification: statistical analysis

In this section, a statistical analysis of the flood events forecasted over the summer-autumn period from July to October 2005 at a number of locations in the Danube catchment is presented, as well as a first application to the spring 2006 period (March-April). The methodology adopted is explained in detail. The analysis focuses on the comparative analysis of gain in lead time (preparedness) when alert rules are applied on both EFAS forecasts based on EPS and ECMWF deterministic weather forecasts.

4.4.1 Definitions and methodology

For a large number of locations, a typical contingence table of hits, misses and false alarms is calculated. Each forecast date is a step and, therefore, an entry in the table. The total number of entries in the table is then equal to the product: number of forecast dates x number of locations x number of lead times comprised in the analysis. Only lead times 2 to 9 are taken into account *and* without distinction: the performance of the system is supposed to be of the same quality for all lead times. Lead time = 1 (first 24 hours of the forecasting range) is not considered because at this lead time the influence of the initial conditions is still very important. Lead time = 10 (last 24 hours of the forecasting range) is not taken into account because at this time step it is not possible to check for persistence. However, EFAS forecasts at lead time 10 days are considered when checking persistence at the next forecast at lead time 9 days.

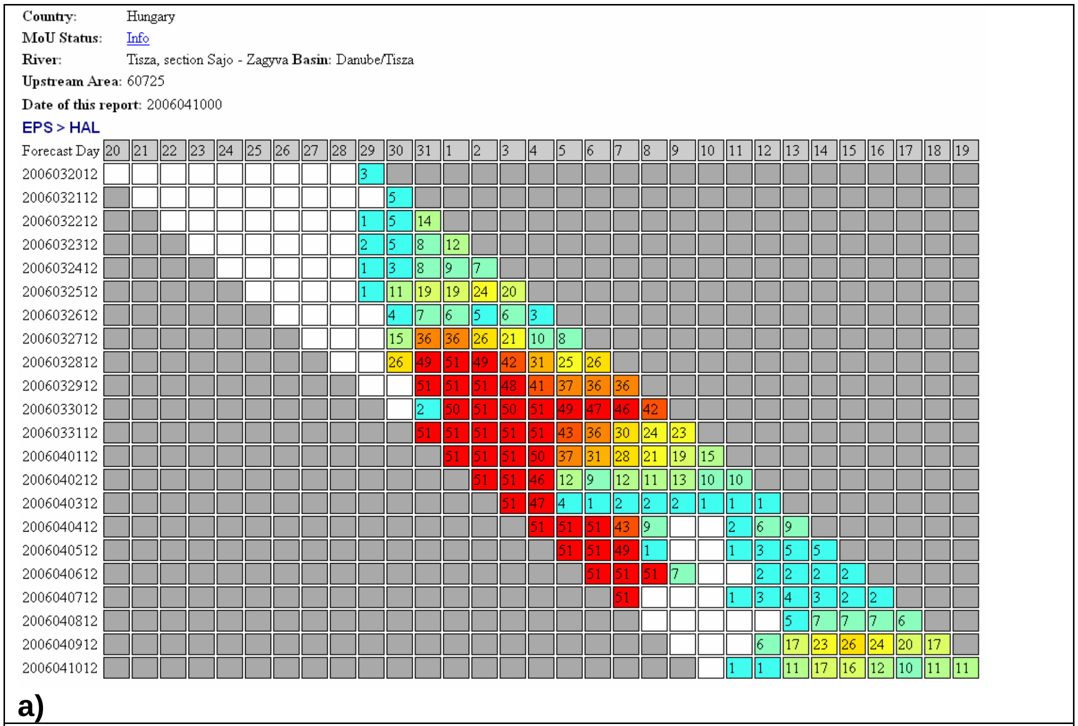
For the definition of hits, misses and false alarms in the contingence table, we considered:

- An ***observed event*** is defined as YES (*or NO*) when discharges simulated with observed meteorological data as input exceed (*or do not exceed*) EFAS high alert levels.

- A ***forecasted event*** is defined as YES (*or NO*) when:

  - Deterministic ECMWF (hereafter named EUD): forecasted discharges exceed *(or do not exceed)* EFAS high alert levels in two consecutive simulations based on weather forecasts issued at 12:00.
  - Probabilistic ECMWF-EPS (hereafter named EUE): forecasted discharges exceed (*or do not exceed*) EFAS high alert levels in two consecutive simulations based on weather forecasts issued at 12:00, both with N_{th} or more EPS simulations above EFAS high levels (N_{th} ranges from 1 to 50). We note that a contingence table can thus be computed for each N_{th} considered in the analysis.

Figure 4.8 illustrates, for a single case, the methodology described above. The example computes hits, misses, false alarms and correct rejection for N_{th} = 5 (i.e., at least 5 EPS-based simulations show discharges above EFAS High alert level).

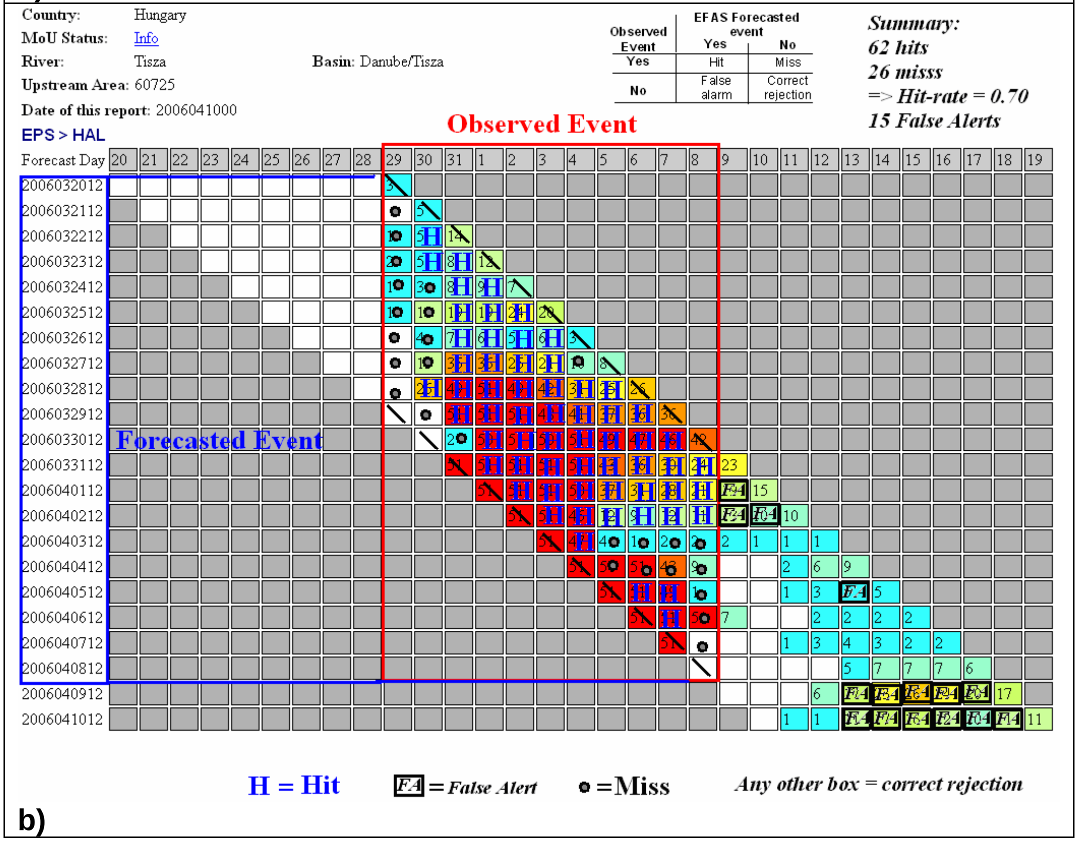


Country: Hungary
MoU Status: Info
River: Tisza, section Sajo - Zagyva **Basin:** Danube/Tisza
Upstream Area: 60725
Date of this report: 2006041000

EPS > HAL

Forecast Day	20	21	22	23	24	25	26	27	28	29	30	31	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2006032012										3																					
2006032112											5																				
2006032212										1	5	14																			
2006032312										2	5	8	12																		
2006032412										1	3	8	9	7																	
2006032512										1	11	19	19	24	20																
2006032612											4	7	6	5	6	3															
2006032712											15	36	36	26	21	10	8														
2006032812											26	49	51	49	42	31	25	26													
2006032912												51	51	51	48	41	37	36	36												
2006033012												2	50	51	50	51	49	47	46	42											
2006033112												51	51	51	51	51	43	36	30	24	23										
2006040112													51	51	51	50	37	31	28	21	19	15									
2006040212														51	51	46	12	9	12	11	13	10	10								
2006040312															51	47	4	1	2	2	2	1	1	1							
2006040412																51	51	51	43	9			2	6	9						
2006040512																	51	51	49	1			1	3	5	5					
2006040612																		51	51	51	7			2	2	2	2				
2006040712																			51				1	3	4	3	2	2			
2006040812																									5	7	7	7	6		
2006040912																								6	17	23	26	24	20	17	
2006041012																							1	1	11	17	16	12	10	11	11

a)



Country: Hungary
MoU Status: Info
River: Tisza **Basin:** Danube/Tisza
Upstream Area: 60725
Date of this report: 2006041000

EPS > HAL

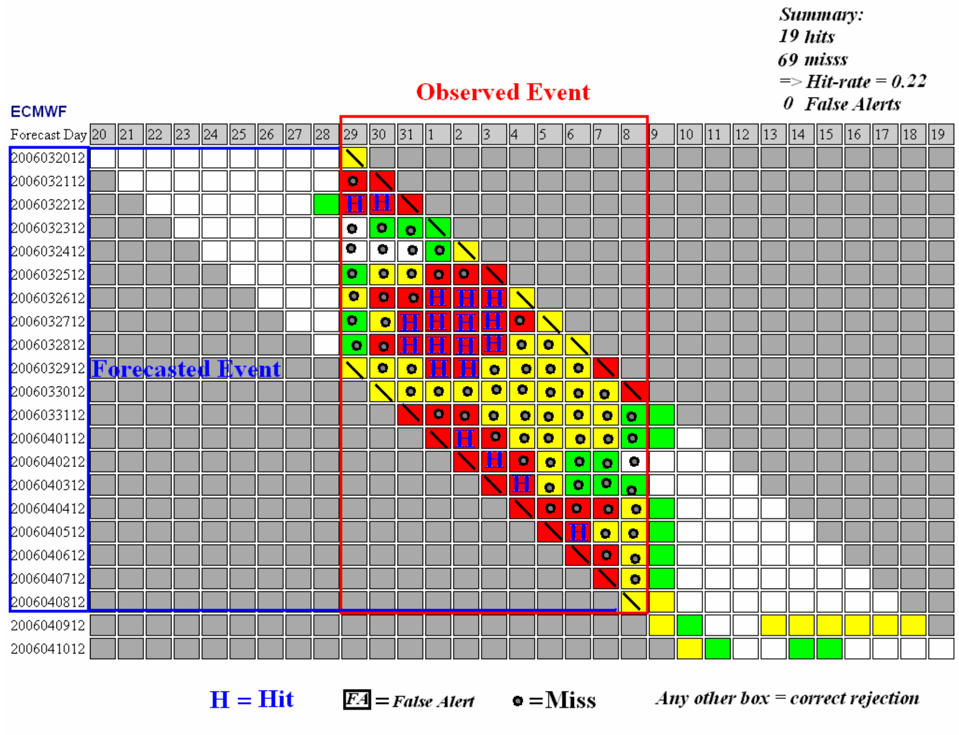
Observed Event	EFAS Forecasted event Yes	EFAS Forecasted event No
Yes	Hit	Miss
No	False alarm	Correct rejection

Summary:
62 hits
26 misss
=> Hit-rate = 0.70
15 False Alerts

Forecast Day	20	21	22	23	24	25	26	27	28	29	30	31	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2006032012										╲																					
2006032112										●	╲																				
2006032212										●	H	╲																			
2006032312										●	H	H	╲																		
2006032412										●	●	H	H	╲																	
2006032512										●	●	H	H	H	╲																
2006032612										●	●	H	H	H	H	╲															
2006032712										●	●	H	H	H	H	●	╲														
2006032812										●	H	H	H	H	H	H	H	╲													
2006032912										╲	●	H	H	H	H	H	H	H	╲												
2006033012											╲	●	H	H	H	H	H	H	H	╲											
2006033112												╲	H	H	H	H	H	H	H	H	23										
2006040112													╲	H	H	H	H	H	H	H	FA	15									
2006040212														╲	H	H	H	H	H	H	FA	FA	10								
2006040312															╲	H	●	●	●	●	2	1	1	1							
2006040412																╲	●	●	●	●			2	6	9						
2006040512																	╲	H	H	●			1	3	FA	5					
2006040612																		╲	H	●	7			2	2	2	2				
2006040712																			╲	●			1	3	4	3	2	2			
2006040812																				╲					5	7	7	7	6		
2006040912																								6	FA	FA	FA	FA	FA	17	
2006041012																							1	1	FA	FA	FA	FA	FA	FA	11

H = Hit **FA** = *False Alert* ● = Miss *Any other box = correct rejection*

b)



c)

Figure 4.8: Illustration of the methodology adopted: a) EFAS diagram of forecasted alert levels from 20th March to 10th April 2006 for EPS-based EFAS forecasts; b) indication of hits, misses and false alerts on the criteria that a forecast event needs at least 5 EPS-based simulations on two consecutive simulations above EFAS high alert levels; c) EFAS diagram of forecasted alert levels for EFAS forecasts based on deterministic ECMWF weather forecasts and indication of hits, misses and false alerts on the criteria that a forecast event needs at least two consecutive simulations above EFAS high alert levels.

For the evaluation of gain in lead times from different forecasts, we have the following definitions:

- ***Preparedness*** is defined as the lead time associated with the first signal in EFAS of a forecasted event (after persistence as defined above).

- **Gain/loss** (Δ prep) is *only* calculated when we have observed events: the comparison is not relevant for EFAS when flooding is not observed. It is evaluated by taking the difference between the EUE preparedness and the EUD preparedness.

$$\Delta\ \text{prep} = \text{prep}_{\text{EUE}} - \text{prep}_{\text{EUD}} \quad ; \quad -10 \le \Delta\ \text{prep} \le +10$$

- If there is a “Miss”, the preparedness (prep_{EUE} or prep_{EUD}) is set to 0 (i.e., no preparedness). If both, preparedness are equal to zero (a “miss” in EUE and also in EUD), Δ prep is not considered.
- If there is a “Hit”, the preparedness (prep_{EUE} or prep_{EUD}) is calculated *for each given observed event date* as the lead time associated with the first signal of an event in the forecasted event.

- If (Δ prep < 0) => *loss* for EUE: a “hit” was forecasted earlier by ECMWF deterministic OR a hit was forecasted by ECMWF, while EPS simulations indicated a “miss” ($prep_{EUE}$=0).
- If (Δ prep > 0) => *gain* for EUE: a hit was forecasted earlier by EPS simulations OR a hit was forecasted by EPS simulations, while ECMWF deterministic simulations indicated a “miss” ($prep_{EUD}$=0).
- If (Δ prep = 0) => EUE and EUD show the same preparedness: a hit was forecasted by both simulations with the same time in advance.

Table 4.1 illustrates for the example shown in Figure 4.8 the gain/loss evaluation.

Table 4.1: Gain/Loss table extracted from the example in Figure 4.8.

Date of observed event	29	30	31	1	2	3	4	5	6	7	8
$prep_{EUE}$	0	10	10	10	10	10	9	10	10	10	10
$prep_{EUD}$	9	10	6	8	9	10	3	0	3	0	0
Δ prep: Gain(+)/Loss(-)	-9	0	+4	+2	+1	0	+6	+10	+7	+10	+10

When a long time series of forecasts is considered, a graph with the variability of the number of occurrences of hits, misses and false alarms according to the criteria of “EPS threshold” (N_{th}) used to define a forecasted event can be drawn. Figure 4.9 shows the typical distribution one can obtain for forecasts based on the deterministic weather forecasts (EUD) and on the EPS weather forecasts (EUE):

- If 1 EPS member (N_{th} = 1) above the EFAS high alert level is sufficient to flag the location as “under flood alert”, then one can expect a high number of false alarms, a much lower number of hits, and an even lower number of missed events. If, on the contrary, 50 EPS out of 51 need to exceed the EFAS high alert level in order to flag the location as “under flood alert”, then there will be a low number of false alarms and a higher number of hits. However, the number of missed events will be even higher, since a flood event may take place already with a lesser number of EPS simulations exceeding the EFAS high alert level.
- For the analysis of forecasts based on the deterministic weather forecasts, the corresponding curves of false alarms, hits and misses are given as simple lines (constants) in these diagrams, since they are not dependent of the “EPS threshold” (N_{th}).

From these curves, three thresholds can be obtained:

- ***t1***: corresponds to the EPS-threshold N_{th} at which *eue* hits become more important than *eue* false alerts;
- ***t2***: corresponds to the EPS-threshold N_{th} at which *eue* misses become more important than *eue* hits;
- ***t3*:** corresponds to the EPS-threshold N_{th} at which *eue* misses become more important than *eud* misses and, consequently, *eue* hits become less important than *eud* hits. We note that, since the hit-rate is defined as the proportion of hits to the total number of observed events [Hit-rate=Hits/(Hits+Misses)], this threshold corresponds to the moment

when *eue* hit-rate becomes smaller than *eud* hit-rate. In a sense, EFAS EPS-based forecasts statistically loose their additional value and the system could run only with the deterministic-based forecasts with, statistically, better performance.

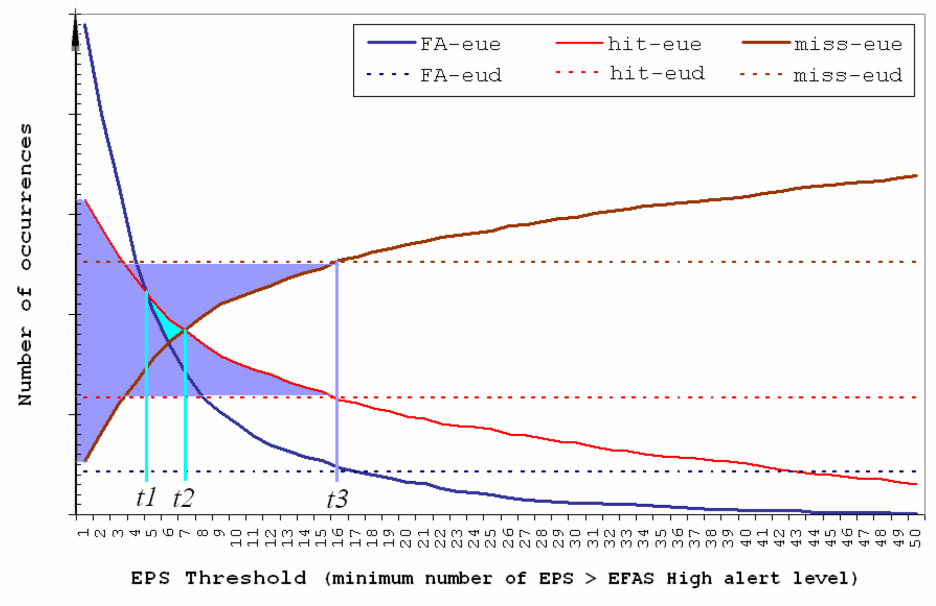


Figure 4.9: Typical Hit, Miss and False Alert curves for deterministic-based forecasts (eud) and probabilistic-based forecasts (eue).

4.4.2 Application to EFAS forecasts in the Danube catchment for July-Oct. 2005

The analysis described in the previous section was performed over 70 representative points selected throughout the Danube river basin (Figure 4.10). The selected points are associated with a wide range of upstream areas, from 1,000 km^2 to 660,000 km^2. About 50% of the points have upstream areas less than 12,000 km^2 and 85% less than 40,000 km^2.

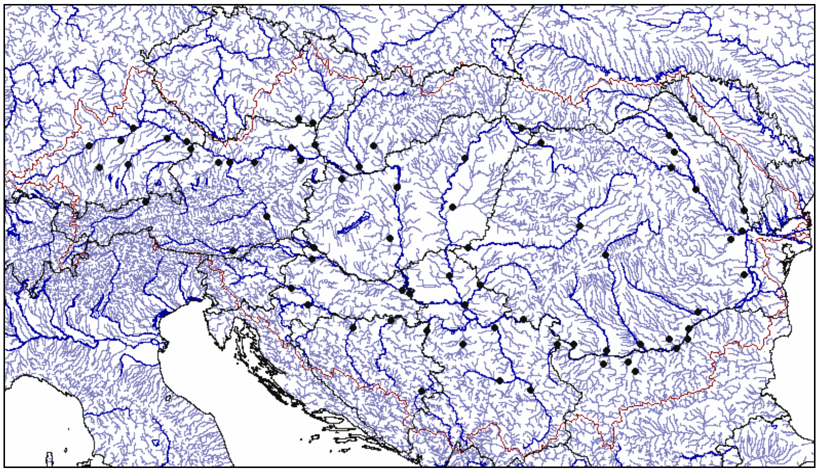


Figure 4.10: Map of 70 locations in the Danube River and its main tributaries.

The analysis was performed for the months of July to October 2005. Figures 4.11 and 4.12 show the curves obtained for hits, misses and false alerts for the months of July and August 2005. The derived EPS-thresholds *t1*, *t2* and *t3* for the whole four-month period analysed are shown in Figure 4.13.

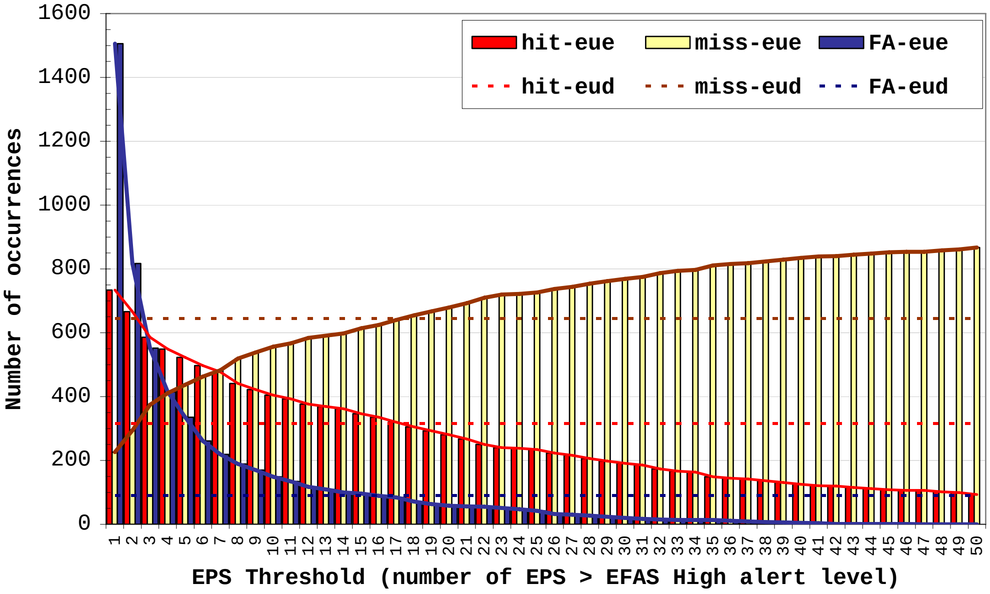


Figure 4.11: Hit, Miss and False Alert curves for deterministic-based forecasts (eud) and probabilistic-based forecasts (eue) for EFAS forecasts in July 2005.

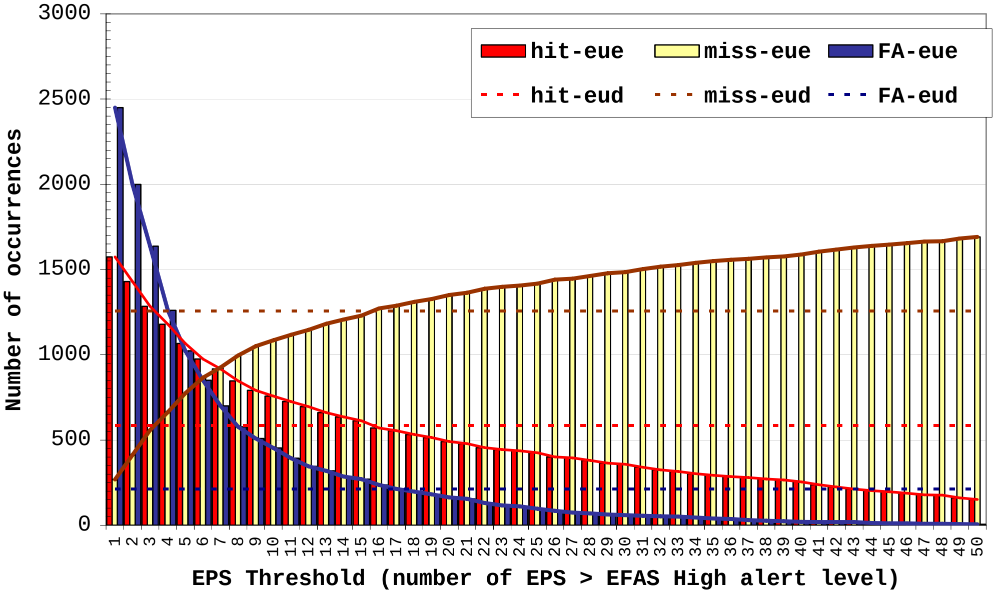


Figure 4.12: Hit, Miss and False alert curves for deterministic-based forecasts (eud) and probabilistic-based forecasts (eue) for EFAS forecasts in August 2005.

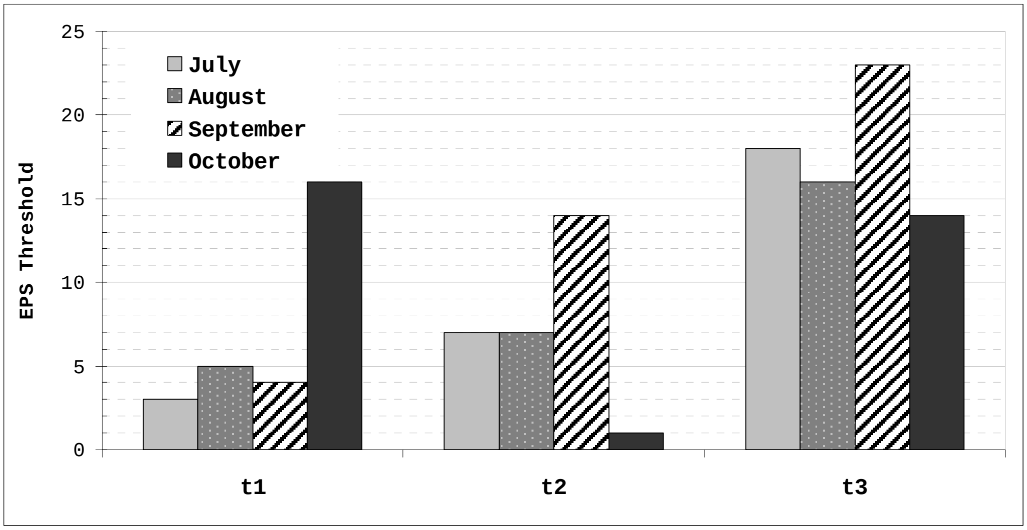


Figure 4.13: Values of t1, t2 and t3 parameters for the months July to October 2005.

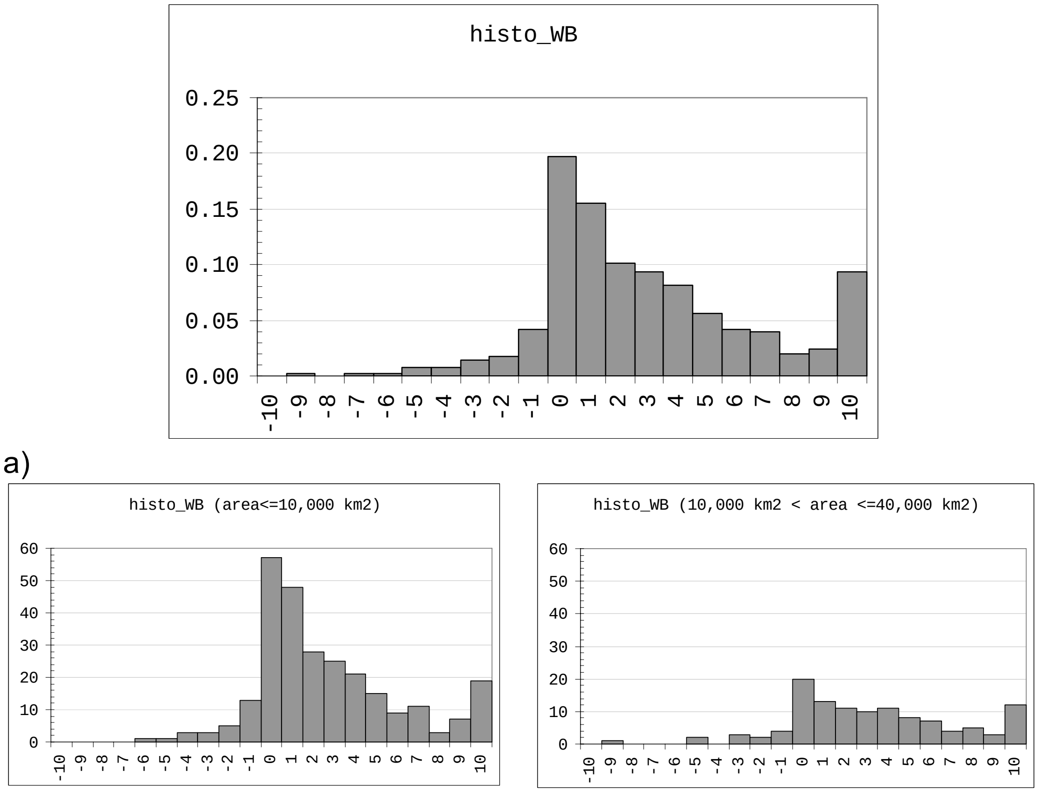
The results show that:

- the most flood prone months of July and August follow the typical expected curves of hit, misses and false alerts as a function of EPS criteria to define a forecasted event. The distribution curves for October (not shown) differ considerably, which is also reflected in the different *t1* and *t2* thresholds obtained for this month (Figure 4.13). This difference is possibly due to a smaller number of occurrences in October 2005 since this was not a flood prone month for the Danube catchment.
- Based on July-September data, if one waits for having more than 5 EPS above EFAS high levels, the *eue* hits become more important than the *eue* false alerts (*t1*). However, when waiting for more than 7-10 EPS, the *eue* misses become more important than the *eue* hits (*t2*). A threshold around 5-10 EPS simulations above high alert levels seems therefore to be a good comprise between hits, misses and false alerts in EPS-based forecasts.
- Also based on July-September data, the EPS threshold around 15-20 EPS simulations appears to be the one above which the *eue* hit-rate becomes smaller than the *eud* hit-rate (*t3*). It shows therefore that if the system waits to have more than 15-20 EPS above high alert levels to issue a warning, the EPS-based simulations partially and statistically loose their additional value in increasing the successful early forecasting of flood events.

The additional value of EPS-based flood forecasts to increased preparedness was investigated by statistically evaluating the gain in preparedness (in days of lead time) of the *EUE* hits comparatively to the preparedness one can get from the analyses of *EUD* hits. Since EFAS is a system for medium-range forecasts, only lead times greater than 3 days were considered. Again, the methodology applied was the one described in section 4.3.1 and illustrated in the example given at Table 4.1.

The gain in preparedness was computed for the period from July to October 2005 and for different “EPS-Threshold”, N_{th} (number of EPS above EFAS High alert level). The results obtained for $N_{th} = 5$ – around *t1* and *t2* values – and $N_{th} = 20$ – around *t3* value – are presented in Figures 4.14 and 4.15 respectively.

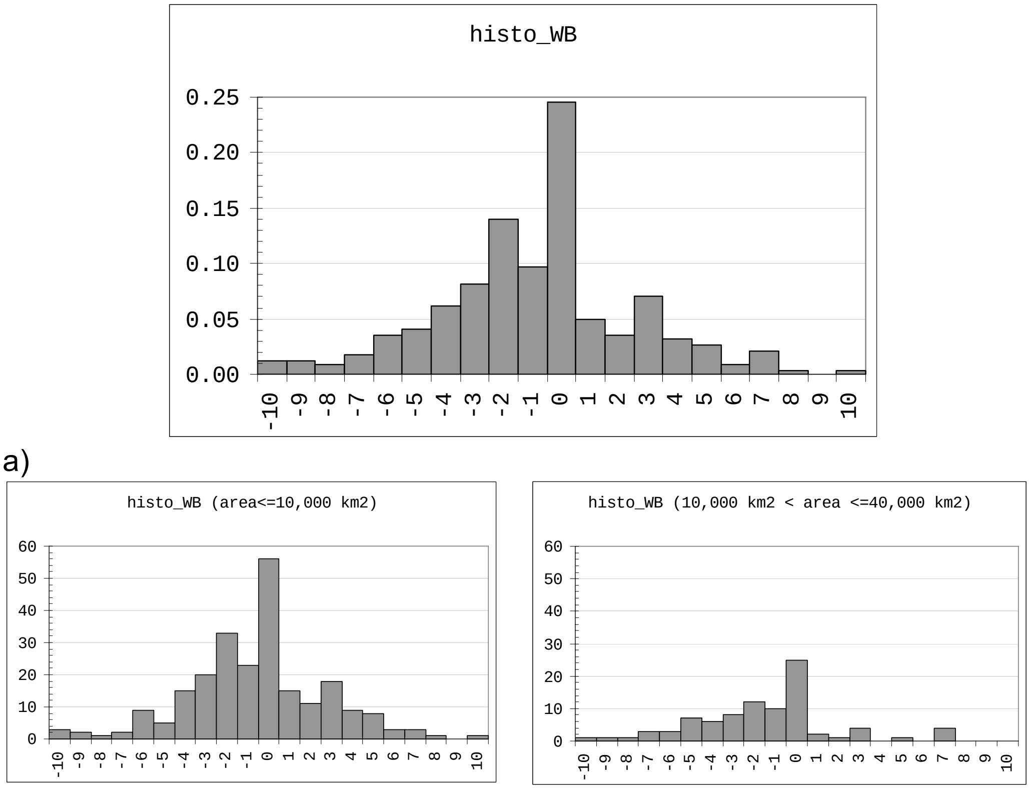
In order to investigate the influence of the upstream catchment area in the EPS gain/loss of preparedness, stratified samples of points with upstream areas smaller than 10,000 km^2 (a total of 30 points and an average area of 5,300 km^2) and between 10,000 km^2 and 40,000 km^2 (a total of 30 points and an average area of 19,000 km^2) were constituted. Results for the ensemble of all cases at the 70 locations and for each stratified sample are also shown in the figures.



a)

b)

Figure 4.14: Gain (+) or Loss (-) in preparedness when comparing EFAS simulations based on EPS to those based on ECMWF deterministic forecasts for EPS-Threshold $N_{th} = 5$ (minimum number of EPS above EFAS High alert levels) for July to October 2005 in the Danube catchment: a) relative frequency considering all 70 points, and b) number of occurrences for points in small upstream areas (left) and larger upstream areas (right). X-axes in days of lead time.



a)

b)

Fig. 4.15 Gain (+) or Loss (-) in preparedness when comparing EFAS simulations based on EPS to those based on ECMWF deterministic forecasts for EPS-Threshold N_{th} = 20 (minimum number of EPS above EFAS High alert levels) for July to October 2005 in the Danube catchment: a) relative frequency considering all 70 points, and b) number of occurrences for points in small upstream areas (left) and larger upstream areas (right). X-axes in days of lead time.

The results show that:

- there is an important gain in lead time when considering at least 5 EPS simulations above EFAS high levels to define a forecasted event (Figure 4.14): the area under the positive part of the histogram is greater than the area under the negative part. The histogram of gain is less skewed for bigger areas and shows a more uniform positive gain. One can also see that for the same number of locations in each upstream area stratification (30 points at each sample), the number of occurrences (observed flood events forecasted by EFAS with more than 3 days in advance) for smaller areas is significantly higher.
- The general behaviour of the histogram of gain when considering at least 20 simulations above EFAS high levels to define a forecasted event opposes to the previous case (Figure 4.15): the area under the negative part of histogram is greater than the area under the positive part. For this EPS-threshold, the gain in lead time that EFAS can get from EPS-based simulations is less important. Most frequently there is no gain in preparedness compared to the deterministic-based forecasts.

We should note that the analyses presented here do not take into account false alerts: from figures 4.11 and 4.12, for instance, one can see that at the EPS-threshold = 5, false alerts in EPS-based simulations are much bigger than false alerts in ECMWD deterministic-based simulations. If we only consider the moment when the number of false alerts is equal in both simulations, the gain in preparedness of EUE forecasts over EUD forecasts will be lesser than the one showed in Figure 4.14.

4.4.3 Application to EFAS forecasts in the Danube catchment for March-April 2006

The results presented in section 4.3.2 showed some statistical properties of EFAS forecasts based on deterministic and EPS weather forecasts for July to October 2005 in the Danube catchment. Floods occurring during these summer-autumn months are basically driven by heavy convective rainfall. The weather systems are heavily influenced either by the Alps or the Carpathian mountains and, typically, catchments with smaller upstream areas are mostly affected by floods. In comparison, floods that take place in spring are mainly driven by snowmelt, often combined with rainfall. In this case, catchments with bigger upstream areas can be affected by serious floods due to the cumulative effect of high water levels coming from upstream areas and the concomitance of river flows from different tributaries.

In spring 2006, severe floods were particularly observed in the downstream parts of the rivers Elbe and Danube. In order to compare EFAS forecast performance in spring time to the one in the summer-autumn 2005 period, the analysis described in section 4.3.1 was applied to the months of March and April 2006. Figures 4.16 and 4.17 show the curves obtained for hits, misses and false alerts for these months. The derived EPS-thresholds *t1*, *t2* and *t3* for this two-month period analysed are shown in Figure 4.18. The gain in preparedness for different "EPS-Threshold" (N_{th}) was also computed. The results obtained for $N_{th} = 5$ (Figure 4.19) and $N_{th} = 30$ (Figure 4.20) are presented for all 70 locations and for stratified samples of points with upstream areas smaller than 10,000 km^2 and between 10,000 km^2 and 40,000 km^2.

The main results from this preliminary comparative analysis indicate that:

- as for July-September 2005, a threshold around 5-10 EPS simulations above high alert levels seems to be a good comprise between hits, misses and false alerts in EPS-based forecasts.
- The EPS threshold for which the *eue* hit-rate becomes smaller than the *eud* hit-rate is around 20-35 EPS, thus, greater than that obtained for July-September 2005 (around 15-20 EPS). It shows maybe that the EPS simulations are more "robust" for the spring period: the EPS-based simulations partially loose their additional value in increasing the successful forecasting of flood events only if the system waits to have more than 30-35 EPS above high alert levels to issue a warning.
- As for July-September 2005, there is also an important gain in preparedness when considering at least five 5 simulations above EFAS high levels to define a forecasted event (Figure 4.19): the area under

the positive part of the histogram is greater than the area under the negative part. However, the histogram of gain is more *U-shaped*, with an important peak at gain in +10 days.

- Contrary to the 2005 results, the contribution of bigger areas is at least as important as the one from smaller areas: the total histogram has its form more dominated by the occurrences in bigger areas (sample of areas between 10,000 km^2 and 40,000 km^2).
- The histogram of gain when considering at least 20 EPS simulations above EFAS high levels (not shown) is less skewed to the left (negative gains) than that obtained for the 2005 period. Only when considering around at least 30 EPS simulations above EFAS high, a more *negative skew* shape is achieved in the histogram of gain (Figure 4.20). These results are linked to the threshold *t3* of 20-35 EPS commented above.

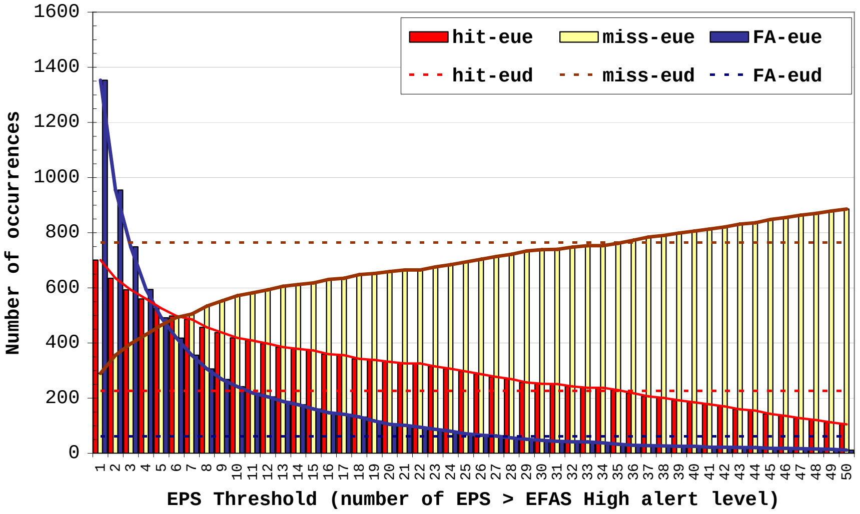


Figure 4.16: Hit, Miss and False Alert curves for deterministic-based forecasts (eud) and probabilistic-based forecasts (eue) for EFAS forecasts in March 2006.

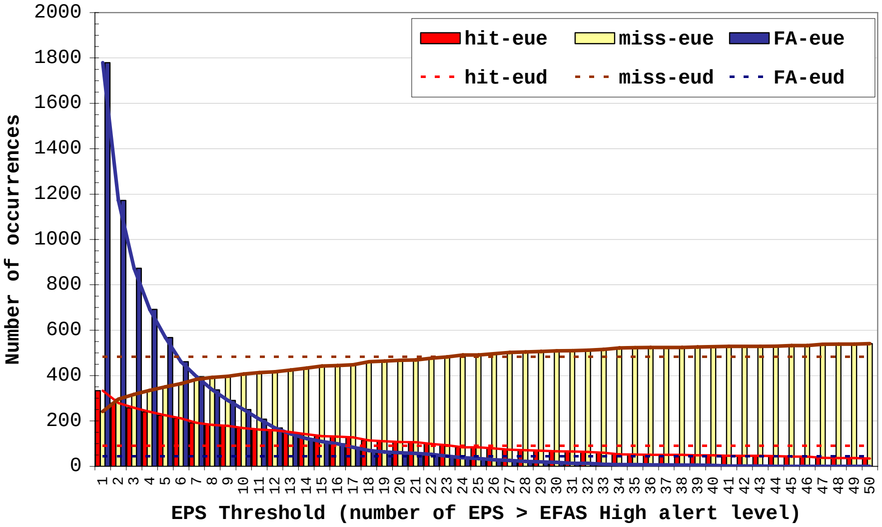


Figure 4.17: Hit, Miss and False Alert curves for deterministic-based forecasts (eud) and probabilistic-based forecasts (eue) for EFAS forecasts in April 2006.

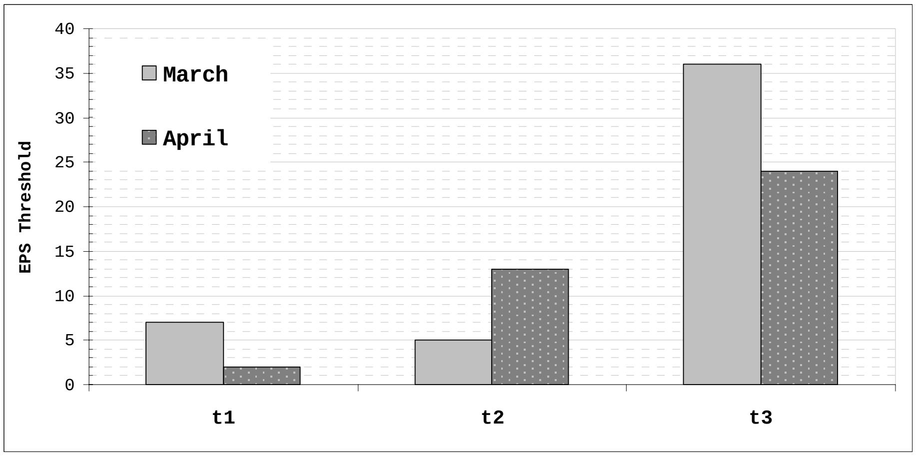
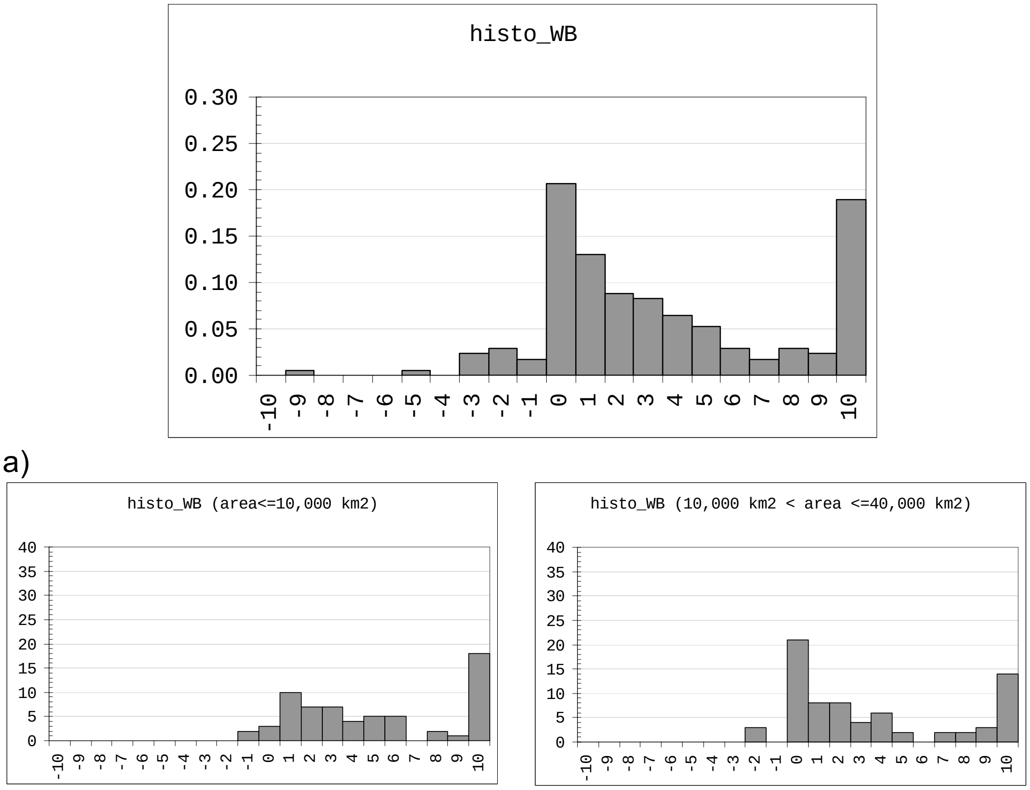


Figure 4.18: Values of t1, t2 and t3 parameters for the months March and April 2005.



a)

b)

Figure 4.19: Gain (+) or Loss (-) in preparedness when comparing EFAS simulations based on EPS to those based on ECMWF deterministic forecasts for EPS-Threshold N_{th} = 5 (number of EPS above EFAS High alert levels) for March to April 2006 in the Danube catchment: a) relative frequency considering all 70 points, and b) number of occurrences for points in small upstream areas (left) and larger upstream areas (right). X-axes in days of lead time.

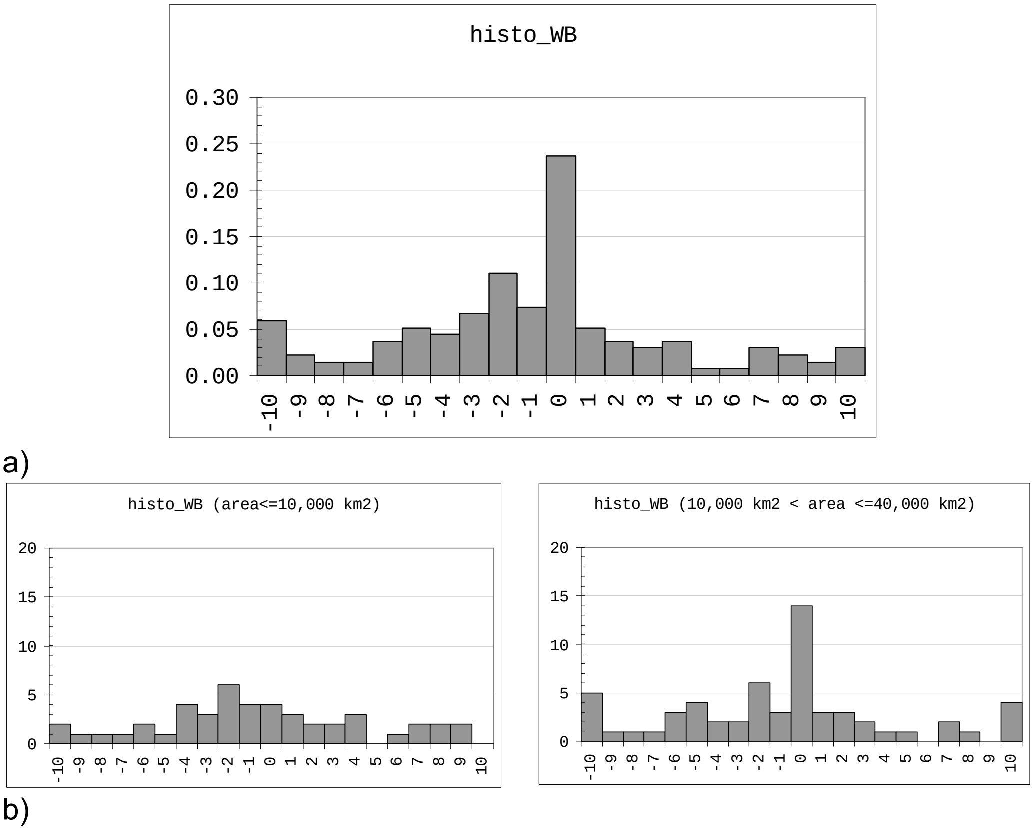


***Figure 4.20: Gain (+) or Loss (-) in preparedness when comparing EFAS simulations based on EPS to those based on ECMWF deterministic forecasts for EPS-Threshold N_{th} = 30 (number of EPS above EFAS High alert levels) for March to April 2006 in the Danube catchment: a) relative frequency considering all 70 points, and b) number of occurrences for points in small upstream areas (left) and larger upstream areas (right). X-axes in days of lead time*.**

4.5. Conclusions

The main conclusions of this study can be summarized in the following points:

- The in-depth analysis of case-studies proved to be very useful to understanding the behaviour of EFAS forecasts under different hydro-meteorological conditions.
- The importance of defining a criterion of persistence in forecasts of consecutive days, together with a criterion of consistence between different forecasts at the same forecast date, for EFAS to flag a possible flooding event was highlighted.
- The analyses of case-studies were also essential to the definition of appropriate rules for an objective assessment of EFAS performance over long time periods and on European scale.

- The additional value of EFAS forecasts based on ECMWF EPS weather forecasts for increased preparedness was demonstrated: the evaluation of the gain in preparedness comparatively to the use of deterministic-based forecasts showed an important gain in preparedness when considering at least 5 to 20 EPS simulations above EFAS high alert levels to flag a forecasted event.
- The comparative analysis of summer-autumn and spring periods (July-October 2005 and March-April 2006, respectively) allowed to a first investigation of seasonal effects in EFAS forecast performance: preliminary results indicate that forecasts at locations associated with smaller upstream catchment areas (less than 10,000 km²) dominate EFAS performance during the summer-autumn period (floods caused mainly by heavy rainfalls), while those at bigger areas (between 10,000 km² and 40,000 km²) shape the results for the spring time (floods caused mainly by snowmelt).

References

Ramos, M.H., Bartholmes, J., Thielen, J., Kalas, M. & de Roo, A. (2006*a)* The additional value of ensemble weather forecasts to flood forecasting: first results on EFAS forecasts for the Danube river basin. *In: Proceedings of the 23rd Conf. Danubian Countries on the Hydrological Forecasting and Hydrological Bases of Water Management*, Belgrade, Republic of Serbia, 28-31.08.06, CD-ROM, 6p.

Ramos, M.H., Thielen, J., Bartholmes, J. & de Roo, A. (2006b*)* The use of meteorological ensemble prediction in the European flood alert system. In: *Geophysical Research Abstracts*, EGU, Vienna, 2006.

Chapter 5

Quantitative analyses of EFAS forecasts using different verification (skill) scores

Jens Bartholmes[1], Jutta Thielen[1], Maria-Helena Ramos[1,2]

[1] *European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, TP261, 21020 Ispra (Va), Italy*
[2] *Cemagref - Groupement de Lyon ; 3 bis quai Chauveau CP 220, 69336 LYON cedex 09 FRANCE*

In this chapter results obtained from statistical analyses of EFAS forecast data for the whole of Europe on a 5x5km^2 grid are shown. The analyses are based on an 11-month period starting in July 2005. The quantitative approach assesses the performance of EFAS forecasts with regard to hit and false alarm rates, leadtime and upstream area. The improvement of forecast performance regarding the use of persistence criteria is analysed. Furthermore, potential gain in leadtime of EFAS-EPS forecasts over EFAS deterministic forecasts is investigated.

Only representative examples of the results are shown.

5.1 Methodology

5.1.1 Specifications of analysis

- The period analyzed spans 11 consecutive months from 1st July 2005 to 31st May 2006.
- Only the 12:00 forecasts are used for the analysis.
- Only leadtimes of 3-10 days are analysed
- The data are the same that were used for the pre-operational EFAS forecasts (and not hind-cast data).
- The analysis is done for every 5x5 km^2 pixel on the grid for the whole of Europe. Excluded are only river basins with a total size of less than 2000 km^2.
- 11 months of data are available and the EPS have been analysed in bins distributed as
  [1-2 3-5 6-10 11-15 16-20 21-25 26-30 31-35 36-40 41-45 46-50 51].
- The EFAS forecasts are compared against a proxy, which is the simulated discharge obtained with observed meteorological input data.

5.1.2 Definition of persistence

Persistence for this study is defined as

- **Persistence in deterministic forecasts**: forecasts at time t and at time t-1 have discharges exceeding HAL in the same pixel. In this case the difference between t and t-1 is 12 hours. This is a persistence criterion that is already used in the semi-operational EFAS.
- **Persistence in EPS forecasts**: forecasts at time t and time t-1 have at least 5 EPS exceeding HAL in the same pixel. In this case the difference between t and t-1 is 24 hours as only the 12:00 EPS forecast were run over the whole analyzed period.

5.1.3 Definition of hits, misses and false alarms

For *deterministic forecasts*, we have:

- **PH**: a positive hit means that both forecast and proxy have exceeded the EFAS high alert threshold
- **FA:** a false alarm means that the exceedance of the EFAS high alert threshold was forecasted but not simulated in the proxy
- **ME**: a missed event means that the proxy exceeded the EFAS high threshold but it was not forecasted.
- **CR**: correct rejections, meaning that both forecast and proxy are not exceeding the EFAS high alert, are not considered in this analysis.

In analogy, the definitions for the *EPS forecasts* are as follows:

- **nPH:** the number n of EPS forecasts out of 51 members that exceed the EFAS High Alert threshold (HAL) if the proxy has also a discharge exceeding the HAL.
- **nFA:** the number of n EPS forecasts out of 51 members that are exceeding the EFAS high alert thresholds when the proxy does not exceed the HAL.
- **nME**: no EPS member forecasted the exceedance of the EFAS high alert threshold but the proxy discharge does exceed the HAL.
- **nCR**: both proxy and forecast do not exceed HAL. This part is not considered in this study.

5.2. Probability distribution function of EPS

An important question with regard to EFAS EPS concerns the probability distribution:

“*What is the probability of a flood event to happen if 10 or 15 out of 51 EPS members exceed the EFAS high alert threshold?*”

To answer this question PH, FA and ME were counted and relative frequencies have been calculated for the 11-month period and the different EPS bins.

Figure 5.1 shows the occurrences of hits and false alarms for leadtime 4 days, per EPS bin and for two classes of upstream areas. The first class is defined as 4000 < ups < 30.000 km^2 and the second as ups > 30.000 km^2. For easier visualisation the number of occurrences has been normalised with the maximum number of occurrences.

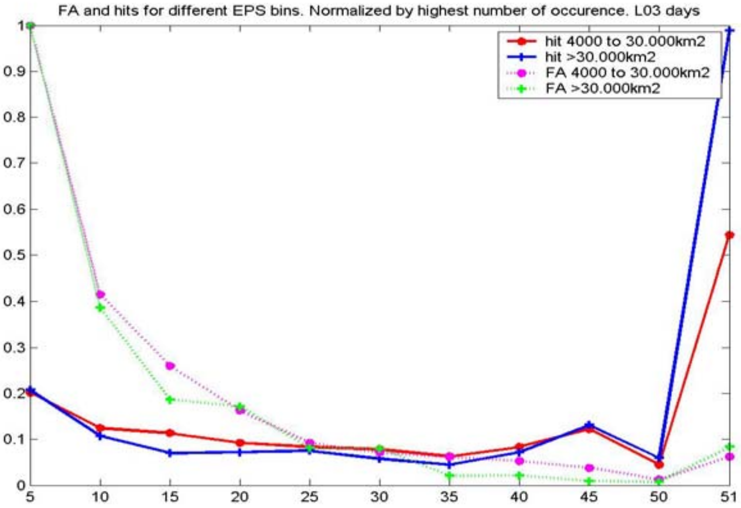


***Figure 5.1** : Normalized number of nPH (solid line) and nFA (dotted lines) on the y-axis plotted against different EPS bins (x-axis) for a lead time of 4 days (L03)[3]. Results are shown for two different classes of upstream area.*

Figure 5.1 shows a number of interesting features. Regarding upstream area it is obvious that the behaviour of the curves for the two different upstream areas is very similar and a clear distinction in occurrence of hits and false alarm rates for the two upstream area (ups) classes “4000-30.000 km^2 and > 30.000 km^2 can not be made.

As a general tendency one can see that the occurrences of false alarms (FA) are decreasing with increasing number of EPS: the more EPS exceed the EFAS high alert threshold (HAL), the more likely the event is going to happen. The fewer EPS are exceeding the EFAS HAL, the higher the false alarm rate. At around 25-30 EPS the corresponding FA and PH curves are intersecting both for smaller and larger ups classes, indicating that at this point the ratio of false alerts to hits is around 1.
If the probability distribution function was linear (see Eq 5.1) and 25 EPS out of 51 represented 50% probability for an event to happen, this result was obvious. However, looking at the EPS bins it becomes clear that the distribution of EFAS-EPS forecast results is slightly different from the expected one.

$$\text{Prob. EPS}_{\text{linear}} \text{ in \%} = \frac{1}{51}\left(nEPS\right) \qquad \text{Eq. 5.1}$$

In figure 5.2 these results are shown as a cumulative distribution function of *Hit ratio$_{EPS}$ (see Eq.5.2)* for a leadtime of 4 days (L03)[1].

$$Hit\ ratio_{EPS} = 100 * nPH/(nFA+nPH)\ [in\ \%] \qquad \text{Eq. 5.2}$$

---

[3] *The naming convention used for the analysis and shown in the diagrams is such that a leadtime of 0 means today. A leadtime of 4 days is then named as L03.*

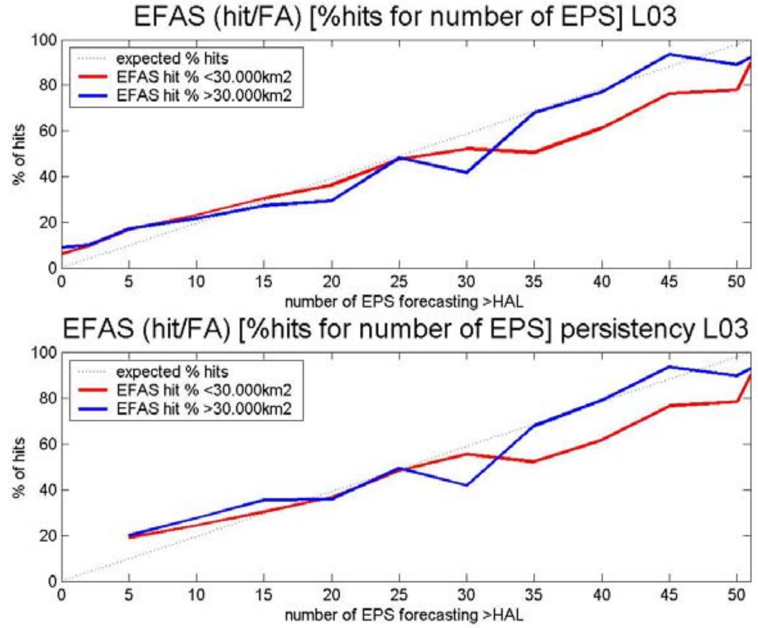


Figure 5.2 : Percentage of hits defined as nPH/(nFA+nPH) shown on the y-axis for different EPS bins shown on the x-axis for a leadtime of 4 days (L03). Results for upstream areas less than 30.000 km^2 are shown in red and larger than 30.000 km^{2}are shown in blue. The top diagram shows unconditioned results and the bottom diagram results conditioned on persistency of at least 5 EPS.

A linear probability distribution function as defined in (Eq 5.1) is represented by a straight diagonal line, plotted as grey dotted line in Figure 5.2.

Results for all lead times show, however, that there is a bias in lower EPS bins towards higher probabilities and in higher EPS bins towards lower probabilities. The probability of a hit with 5 EPS is of the order of 20% for a leadtime of 4 days, whereas 30 EPS have a lower probability than 60%. From 35 EPS onwards and for large upstream areas, the distribution follows more or less the linear probability distribution for the large upstream areas. For the smaller upstream areas there is a lower probability associated with higher EPS numbers compared to the linear distribution. A similar tendency can be found for other leadtimes as well.

Figure 5.2. illustrates also the effect of persistence on the results. Conditioning the analysis on the persistence of at least 5 EPS, the probability to have a hit is raised particularly in the lower bins (bottom diagram compared to top diagram): without persistence 5 EPS out of 51 represent a 18% probability for the event to happen, whereas with persistence it is 22%, thus a gain of almost 5% probability.

To analyse this phenomenon in more detail the *Relative Hit ratio 2* $_{EPS}$ was chosen:

$$\textit{Relative Hit ratio}_{EPS}\text{: Ratio} = (nPH/nFA)_{EFAS} / (nPH/nFA)_{linear} \quad \text{Eq. 5.3}$$

In figure 5.3 this score (Eq. 5.3) is plotted against the EPS bins for leadtime 4 days (L03). The grey dotted reference line at 1.0 marks the points where the single ratios $(nPH/nFA)_{EFAS}$ and $(nPH/nFA)_{linear}$ would be the same, i.e. the forecasted distribution follows equation 5.1.

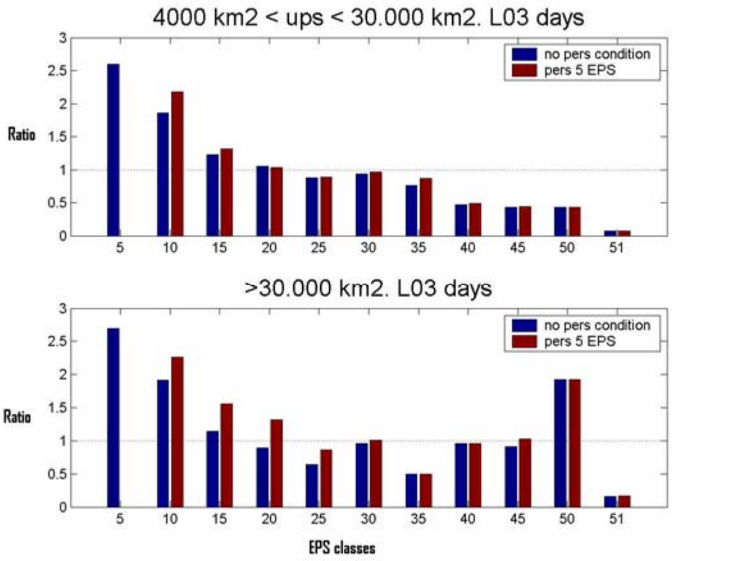


Figure 5.3: Relative Hit ratio $(nPH/nFA)_{EFAS}$ / $(nPH/nFA)_{linear}$ on the y-axis for different EPS bins shown on the x-axis. Blue bars are unconditioned, red bars are conditioned on persistence of > 5 EPS.

The *Relative Hit ratio* can be interpreted as:

> 1: the forecasts have a higher probability to get a hit (for the respective EPS bin) than the linear reference distribution would let expect.

< 1: the forecasts have a lower probability to get a hit (for the respective EPS bin) than the linear reference distribution would let expect.

The figure shows three important results:

1) Conditioning the EFAS forecasts on persistence mostly raises significantly the probability to get a hit. The improvement is particularly important in the lower bin classes and up to 25 EPS whereas it has little effect in larger bin classes. This result is also consistent for other leadtimes.
2) Up to 15 EPS there is clearly more skill than would be expected if the distribution was linear (Eq 5.1). This is also true for other leadtimes..
3) For smaller upstream areas the ratio as compared to the linear distribution decreases steadily with increasing EPS bins. This is different for the larger upstream areas, where the distribution of the ratio on the bin classes is u-shaped: after a local minimum around 25-35 EPS, the ration increases again.

5.3. EFAS skill as a function of upstream area classes

So far the analysis has not shown a clear dependence of the EFAS results on upstream area (distinction was only made between > and < 30.000 km^2). This is analysed further in this section. As before, hits and false alarms (nPH and nFA) have been evaluated for every pixel and then summed up over classes of upstream areas. Again, this analysis has been done unconditioned and conditioned on persistence. In this section both results for the deterministic as well as EPS results are shown.

5.3.1. EFAS probabilistic forecasts

Figures 5.4a/b show *Relative Hit ratio*$_{EPS}$ (Eq. 5.3) but only for EPS bin 10:[8-12] and EPS bin 25:[23-27] [4] for different upstream areas ranging from 500 km^2 to larger than 30.000 km^2. The results are shown for 2 leadtimes 4 days (L03) and 5 days (L04). Again the reference ratio from the linear distribution is shown as line at 1.0.

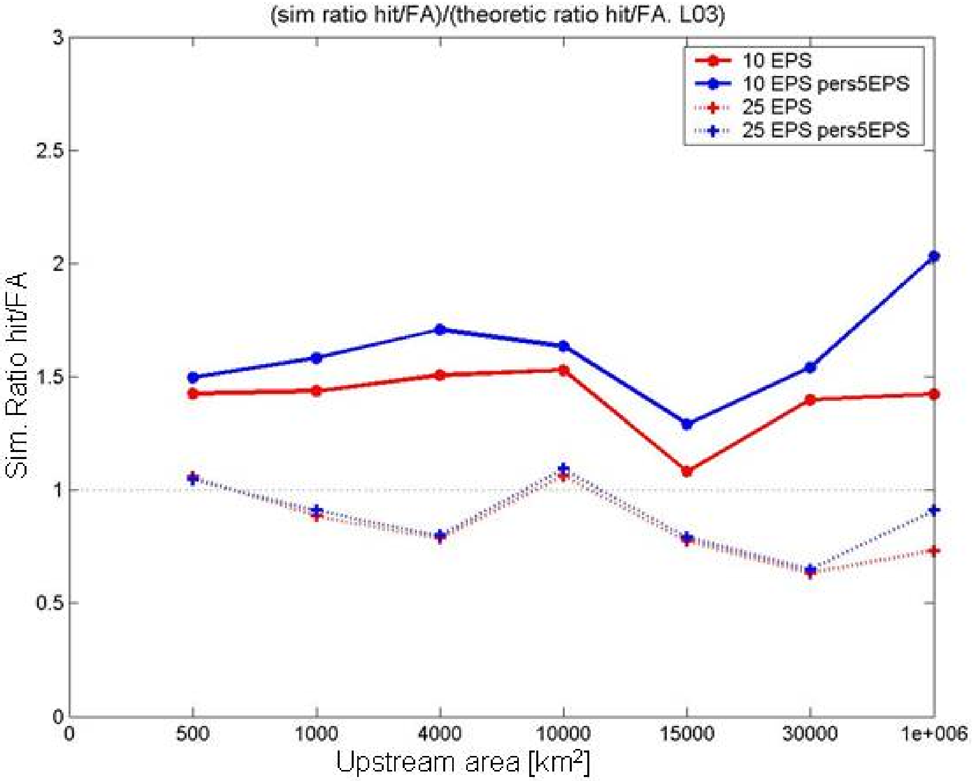


Figure 5.4a: Relative Hit ratio (nPH/nFA)$_{EFAS}$ / (nPH/nFA)$_{linear}$ for EPS bins 10 and 25 plotted over upstream area for leadtime 4 days. The full lines show the results for the 10 EPS bin and the dotted the results for the 25 EPS bin. The blue lines are with conditioning on persistency, the red without.

---

[4] *These bins differ slightly from the ones used in the previous part. 2 bins that had exactly 10 and 25 EPS members as mean were chosen for this visualization.*

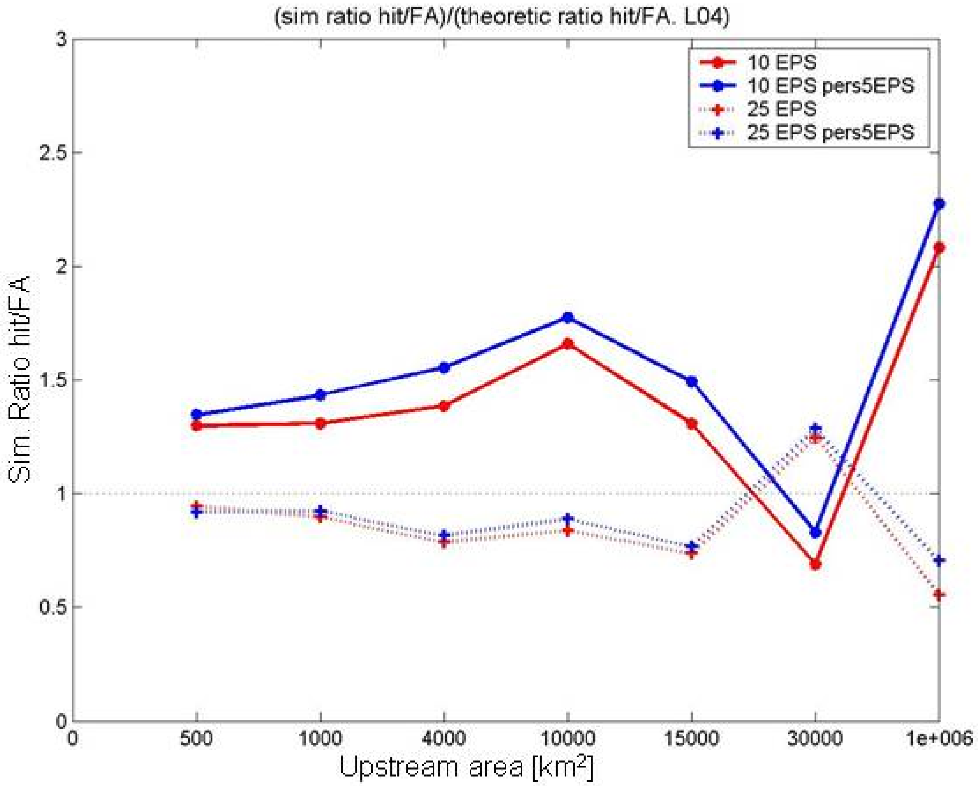


Figure 5.4b: Relative Hit ratio $(nPH/nFA)_{EFAS}$ / $(nPH/nFA)_{linear}$ for EPS bins 10 and 25 plotted over upstream area for leadtime 5 days. The full lines show the results for the 10 EPS bin and the dotted the results for the 25 EPS bin. The blue lines are with conditioning on persistency, the red without.

Figure 5.4 shows clearly that:

- a clear cut-off line for the ratio of hit to false alarms depending on upstream area cannot be identified. When looking at the ensemble of all results the best skill can even be found in the smaller upstream areas. Results, not shown in this report, indicate that only for very small upstream areas of much less than 500 km^2 a tendency can be identified.

- For the 10 EPS bin the ration of hit to false alarms is mostly larger than the reference values, whereas for the EPS bin of 25 it varies around 1. This confirms the results in previous findings that 25 EPS roughly correspond to a 50% probability, whereas a smaller number of EPS has a higher probability than a linear distribution.

Similar results can be derived from the analysis of the Brier Skill Score which is shown in Section 5.4.

5.3.2. EFAS deterministic forecasts

Figures 5.5a/b shows a comparison between the performance of EFAS forecasts based on ECMWF data (EUD) and DWD data (DWD). Shown in figure 5.5a is the number of pixels of hits or false alarms (on the y-axis) as a function of upstream area (x-axis). The upper graph shows results for which persistence was not taken into consideration while the lower graph shows results for forecasts that were conditioned on persistence.

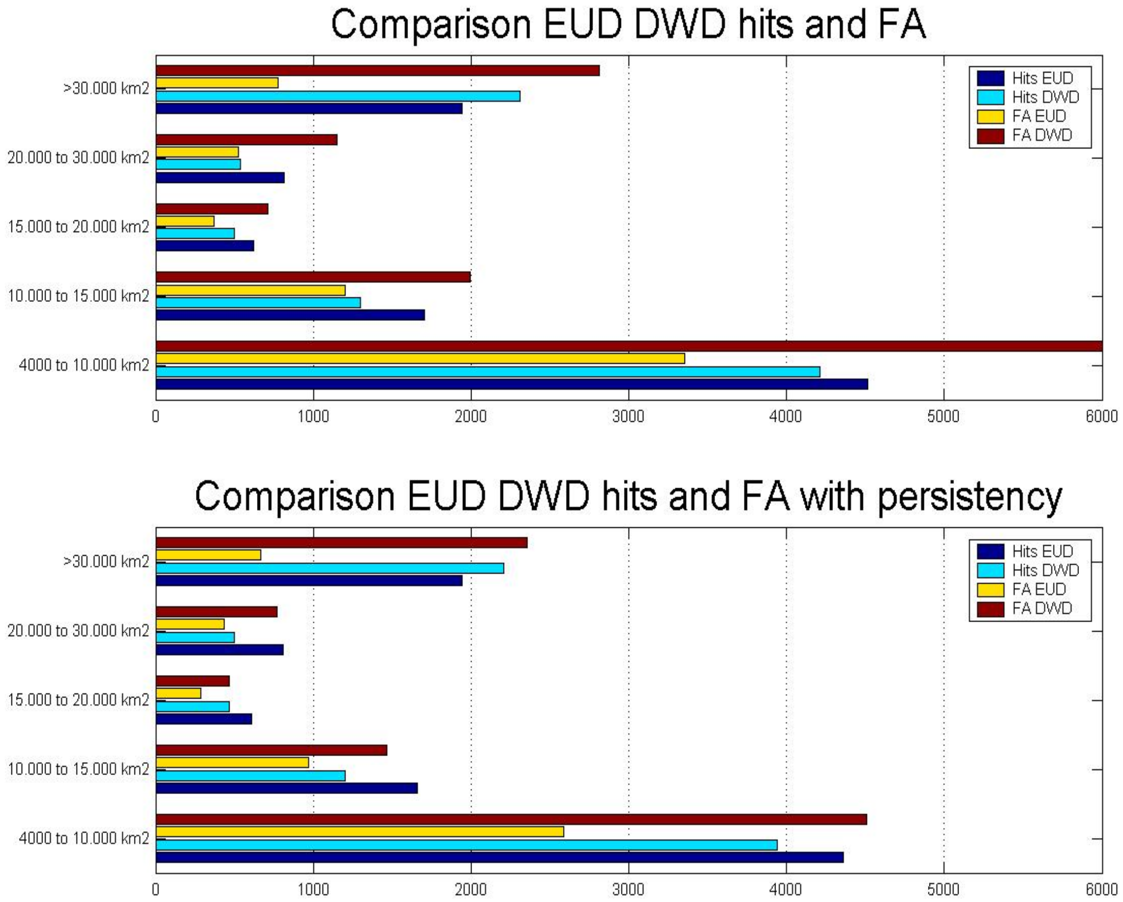


Figure 5.5a Occurrence of hits and false alarms (x-axis) for ECWMF deterministic forecasts (EUD) and DWD forecasts (DWD) as a function of upstream area (y-axis) for a leadtime of 4 days. Results are shown unconditioned (top) and conditioned on persistence (bottom).

Figure 5.5a shows an example for leadtime 4 days and figure 5.5b summarizes the results for the deterministic EFAS forecast in one diagram where hits/FA were plotted over all leadtimes.

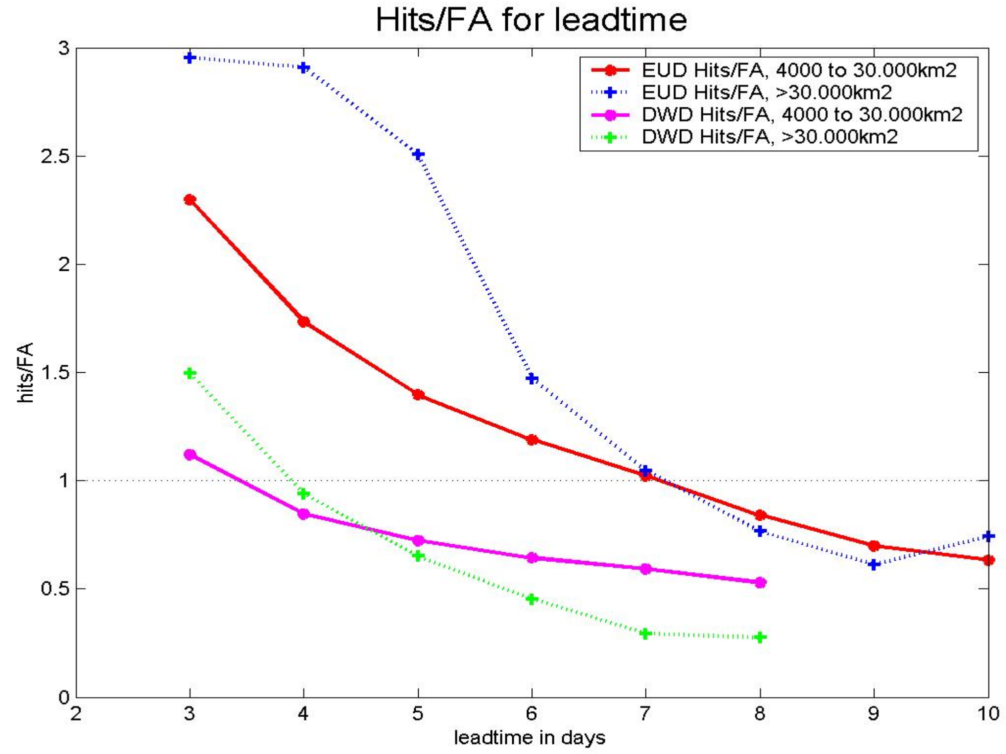


Figure 5.5b : Hits/FA for EFAS deterministic (ECMWF [EUD] and DWD) forecasts for 2 different upstream area classes plotted over lead time.

Overall, the results show

- Also for deterministic forecasts the condition of persistence reduces considerably the number of false alarms (see figure 5.5a) without significantly increasing the number of missed events.
- The number of hits with ECMWF based forecasts is higher than with DWD based forecasts and the number of false alarms is lower respectively. In detail, and taking also into account also the results from other leadtimes, one can see in figure 5.5a/b for example that

  a. for pixels with upstream areas of 4000 to 10.000 km^2: Up to 6 days of lead time the EFAS ECMWF deterministic forecast (conditioned on persistence) have more hits than false alerts, while for the EFAS DWD deterministic forecast this is only the case for lead time 3 days.
  b. for pixels with upstream areas of 10.000 to 30.000 km^2: Up to 5 days of lead time the EFAS ECMWF deterministic forecast (conditioned on persistence) have more hits than false alerts. For the EFAS DWD deterministic forecast this is not the case for any lead time
  c. for pixels with upstream areas of > 30.000 km^2: Up to 6 days of lead time the EFAS ECMWF deterministic forecast (conditioned on persistence) have more hits than false alerts. For the EFAS DWD deterministic forecast this is only the case for lead time 3 days.

5.4. Evaluation of Brier skill score

The Brier score and the Brier skill score (Wilks 1995) are used to assess the skill in probabilistic forecasts. The Brier score is a measure of mean-square error of probability forecasts for a binary (yes/no) event. The Brier score (BS_f) of the forecast t is defined as

$$BS_f = \frac{1}{N}\sum_{1}^{N}(p-o)^2 \qquad \text{Eq. 5.4}$$

where *p* refers to the probability with which an event is forecasted and *o* to the binary value of the observation (*o* = 1 if event observed and *o* = 0 if not observed). N is the total number of forecast dates.
The Brier score indicates a perfect forecast if BS equals 0. Unfortunatelly, it is difficult to interpret in absolute terms and hence the Brier skill score (BSS) is being computed. In the BSS the Brier score of the forecast is compared against the Brier score of a reference climatology (BS_{clim}).

$$BSS = 1 - \frac{BS_f}{BS_{c\lim}} \qquad \text{Eq. 5.5}$$

Thus the interpretation of the skill is strictly related to the reference: i.e. is the current forecast better or worse than assuming climatology. A BSS of 0 indicates that the forecast is not better than the reference forecast, while BSS values > 0 indicate that the forecast is better than the reference. Hamill et al. (2005) have shown that the interpretation of BSS is very sensitive to the choice of the reference climatology.

In the case of EFAS the choice of suitable climatology is not trivial, because strictly speaking there is no climatology to compare the results to. EFAS forecasts run with EPS only since a 12-13 months period. Using the simulations with observed meteorological data as reference introduces also a bias into the analysis because the data sets are not consistent. Observed data can, in any case not be used. It has therefore been decided to follow a suggestion of Legg and Mylne (2004) who propose to confront the Brier score of rare events with a forecast that always forecasts “no event”. In this case the climatic Brier score becomes:

$$BS_{c\lim} = \frac{1}{N}\sum_{1}^{N}(0.0-o)^2 \qquad \text{Eq 5.6}$$

The EFAS high alert threshold corresponds roughly to a 1-2 year return period. If this justifies the “no event” criteria can be debated – but as could be any other climatology.

It should further be noted that for this analysis Brier scores of correct rejections are not included, i.e. if nothing happened during the whole period and the forecast never indicated any probability (=> BS = 0) for an event, then this was not counted.

The following diagram shows the relative frequency (y-axis) of Brier skill scores (x-axis) for different upstream areas at a leadtime of 4 days not conditioned (top) and conditioned on persistence on 5 EPS (bottom).

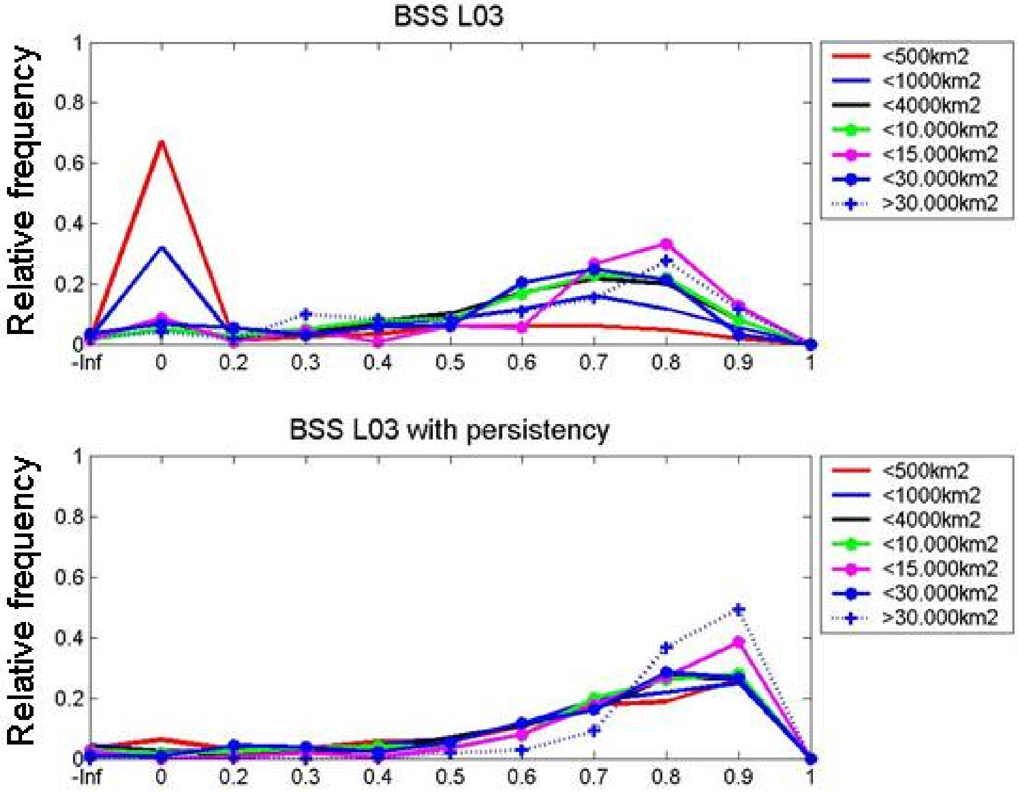


Figure 5.6: Relative frequency (y-axis) of Brier skill score (x-axis) distribution of EFAS EPS forecasts for different upstream areas for a leadtime of 4 days

The most striking feature of Figure 5.6 is the benefit of introducing persistence in EFAS. In this case the high number of BSS around zero in small upstream areas is totally eliminated (in this case of leadtime 4 days up to 4000 km^2) and the whole distribution of BSS is shifted to higher skills.

Figure 5.7 shows that also for larger leadtimes, in this case leadtime 10 days, the forecasts remain surprisingly skilful, in particular for the small river basins when considering persistence.

Obviously the results shown above need to be interpreted in view of the reference climatology: EFAS results are more skilful than assuming that nothing happens. Since 2005 was a particularly flood prone year, the high BSS is perhaps not surprising.

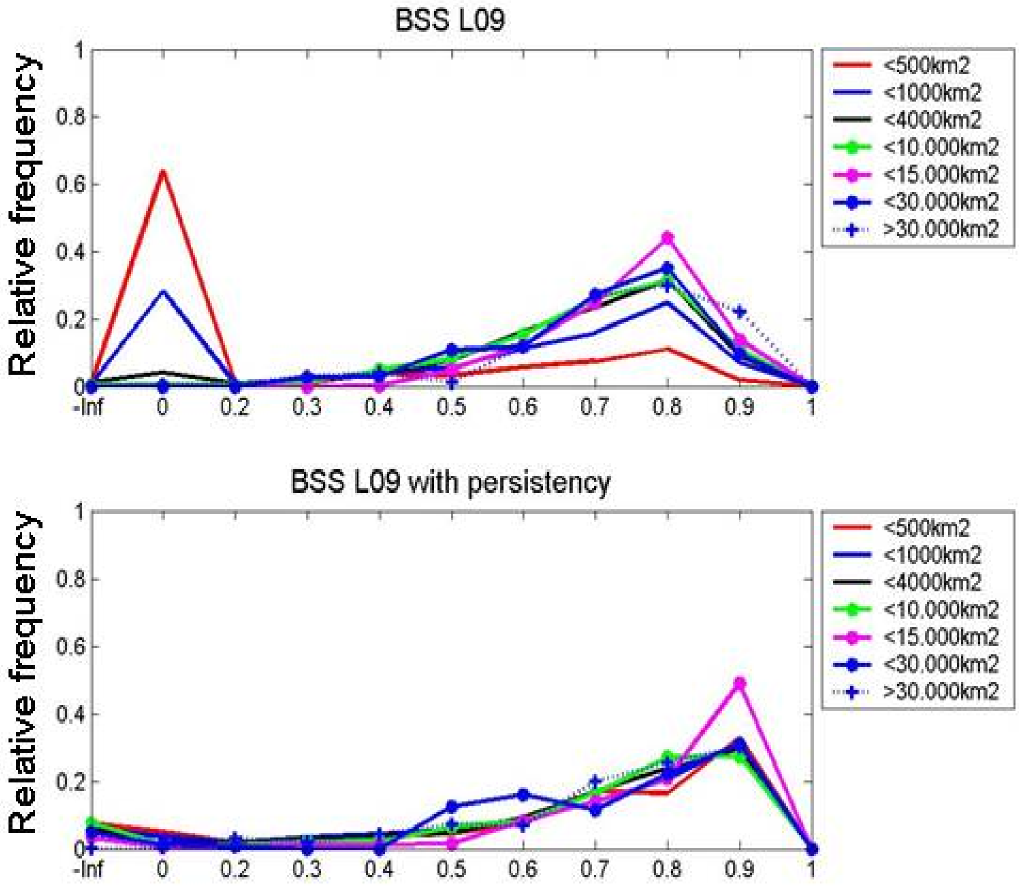


Figure 5.7: Relative frequency (y-axis) of Brier skill score (x-axis) distribution of EFAS EPS forecasts for different upstream areas for a leadtime of 10 days

Taking into consideration another climatology the picture changes. When assuming a constant probability of 1% of an event to happen the BSS drops drastically. The positive influence of using the persistence criterion is also visible for this climatology, however, it is not enough to prevent negative BSS values. The most positive changes are to be observed for pixels with upstream areas larger than 30.000 km^2. Results for a 1% climatology are shown in Figures 5.8. These results are supporting the findings of Hamill et al. (2005) regarding the interpretation of BSS and its sensitiveness to the choice of the reference climatology.

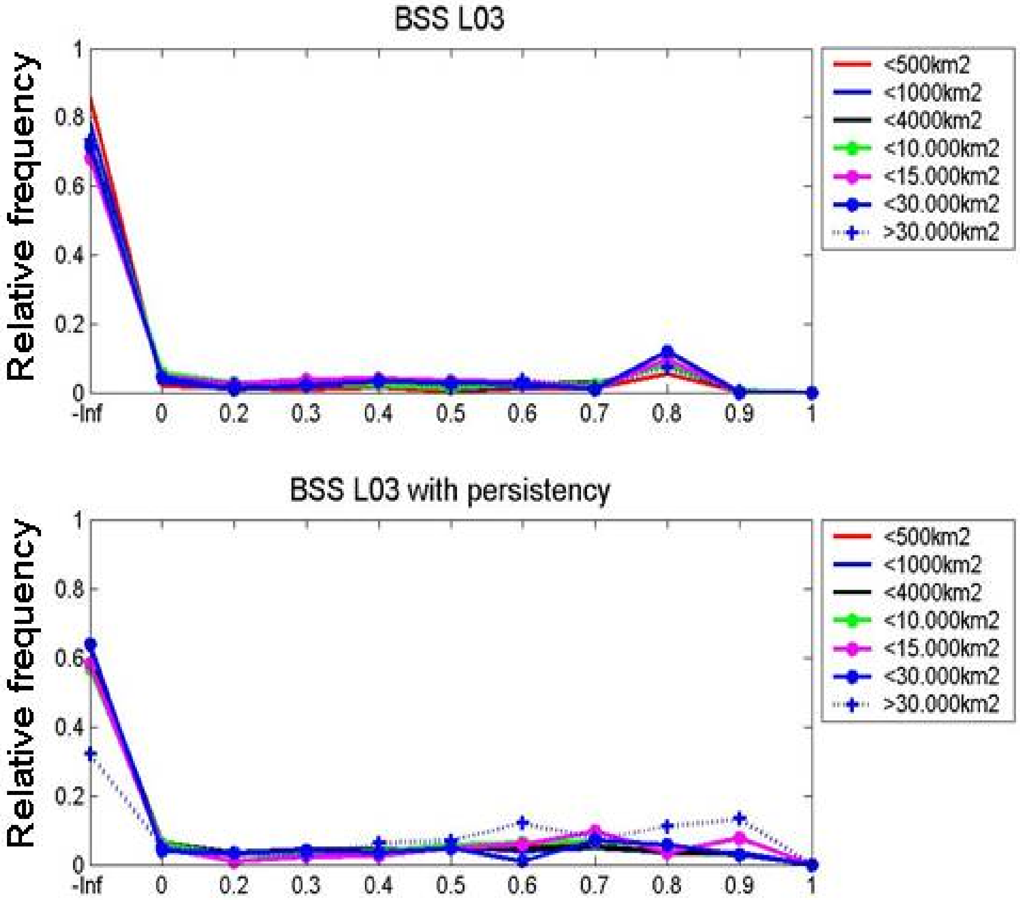
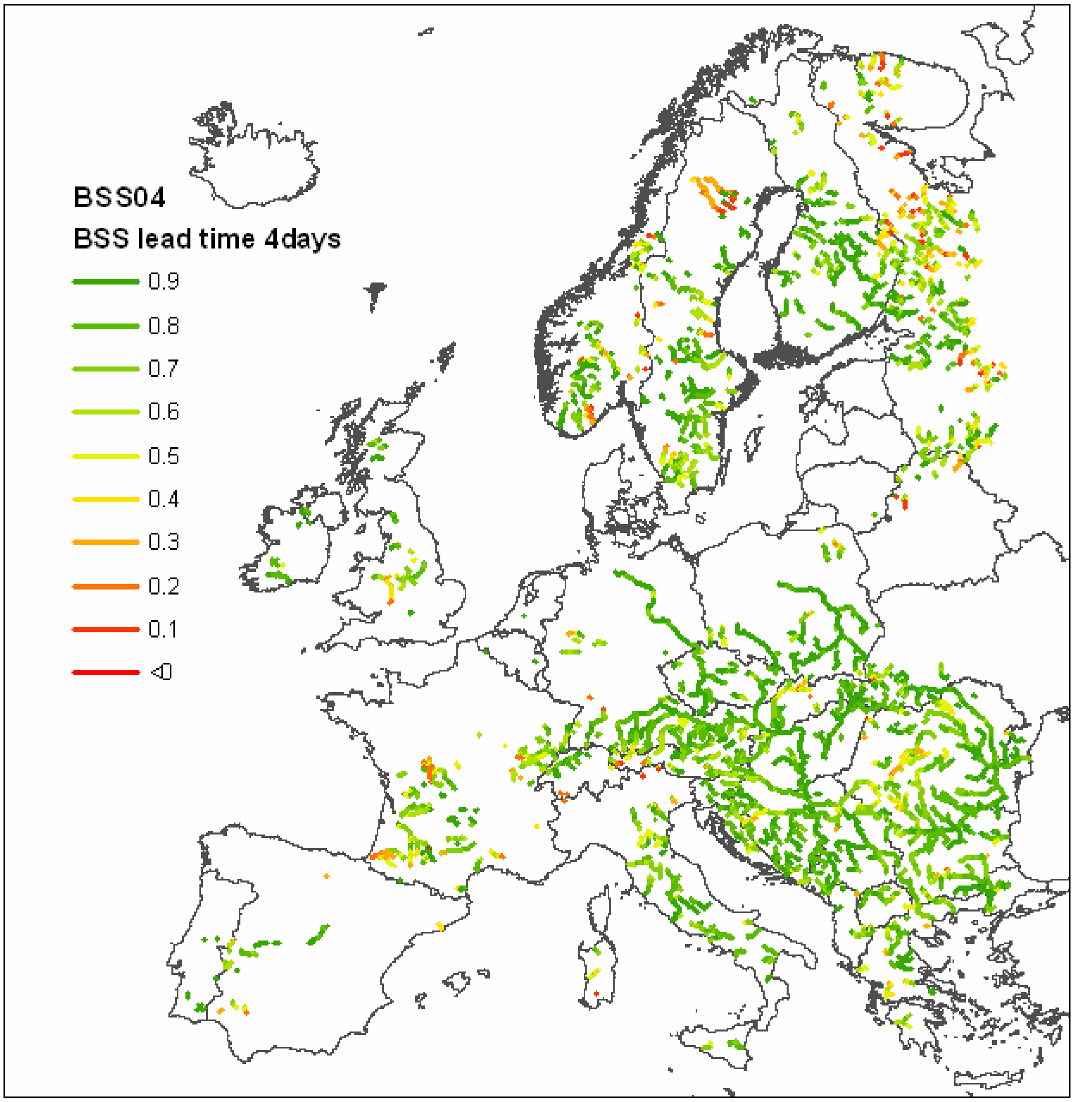


Figure 5.8: Relative frequency (y-axis) of Brier skill score (x-axis) distribution of EFAS EPS forecasts for different upstream areas for a leadtime of 4 days. For this elaboration a climatology assuming a constant probability of 1% of an event to happen was applied.

More research is also needed to investigate the impact of dependency in hydrological analysis – results in downstream river pixels are not independent from upstream river pixels. Further research to clarify this issue is ongoing. Another limitation of this approach may be the limited amount of data. Ideally the BSS should be calculated over longer timeseries and is limited here to a period of 11 months only. Within these 11 months a number of very distinctly different flood events took place – summer floods and snow melt winter floods. From post event analysis it is clear that EFAS is not performing well for the snowmelt floods in terms of timing and quantity. Particularly for the snowmelt floods in spring 2006, a part of the problem consisted in the poor observed meteorological data. In how far this introduces a bias when assessing the forecasted data with the observed data is yet to be investigated.

Interesting for EFAS applications is also the spatial distribution of the forecast skill. Figure 5.9a an shows a map of Europe with BSS for leadtime 4 days and figure 5.9b shows a map of Europe with BSS for leadtime 10 days.



.

Figure 5.9a: Spatial distribution of BSS conditioned on persistency (5 EPS) for lead time 4 days

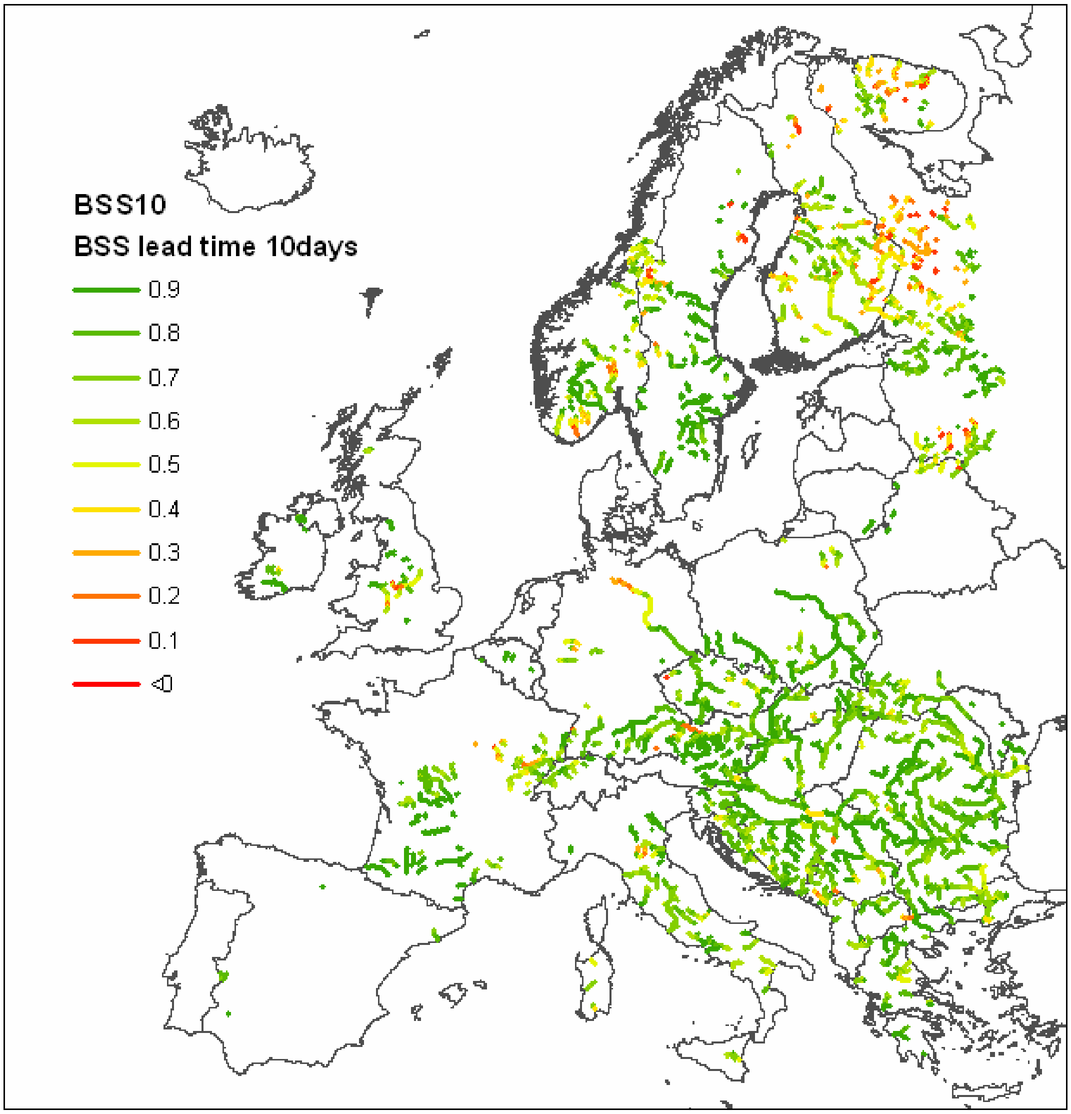


Figure 5.9b: Spatial distribution of BSS conditioned on persistency (5EPS) for lead time 10 days

The results of figures 5.9a/b show that the BSS is only defined for those areas where floods took place or were forecasted (with at least 5 EPS) during the reporting period. It also shows that in large areas where no floods took place also no floods were forecasted (with more than 5 EPS). Furthermore, after 10 days of leadtime there is still skill in the forecasts.

5.5. EFAS results as a function of leadtime

For EFAS forecasters it is important to know at what rate the skill of the forecasts decreases with increasing leadtime. This section analyses these results for the probabilistic and the deterministic results. Again, the analysis focuses on two classes of upstream areas, with 4000km^2 < ups< 30.000 km^2 and 30.000 km^2 < ups.

5.5.1 Probabilistic EFAS

The mean of the BSS is shown as function of leadtime and for the two upstream area classes and only conditioned on persistence of 5 EPS.

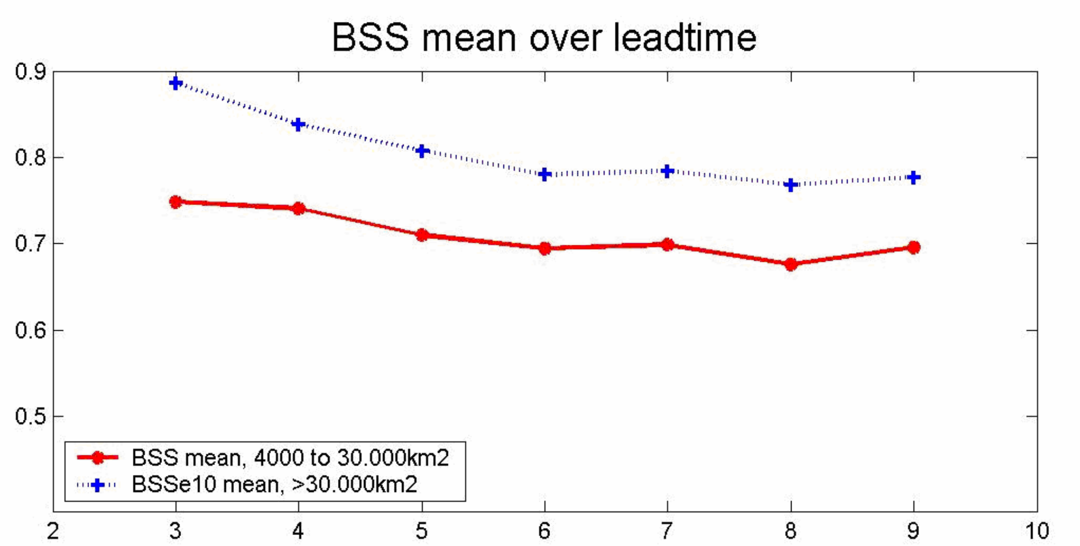


Figure 5.10 : Mean BSS (y-axis) as a function of leadtime (x axis) and upstream area. Upstream areas of 4000 km^2 < ups < 30.000 km^2 is shown in red and ups < 30000 km^2 shown in blue.

Figure 5.10 clearly illustrates a number of interesting results.

- EFAS EPS based on persistence and using the no event criteria as climatology is skilful even at very long leadtimes
- The skill decreases (almost) steadily with increasing leadtime
- The skill is slightly lower for the smaller upstream areas than for the larger upstream areas. However, it still remains skilful for both classes.

In figure 5.11 the absolute numbers of hits, false alerts and misses (conditioned on persistency) are plotted over leadtime and are separated for different upstream areas. Here hits are not classified into separate EPS bins. Every forecast with at last 5 EPS forecasting >HAL and proxy >HAL is counted as hit. As well, every forecast with at least 5 EPS forecasting >HAL and proxy <HAL is counted as FA.

The number of hits and misses decreases slightly with increasing lead time while the number of false alert increases almost linearly with increasing lead time. This does not mean that the meteorological EPS are over-predicting precipitation amounts at higher leadtimes, but that the spatial spread increases with leadtime. The same number of EPS members predicting discharges > HAL at short leadtimes will be concentrated on fewer pixels, thus resulting in less pixels with a FA (also due to the definition used here: no distinction in FA as long as ≥ 5 EPS members > HAL).

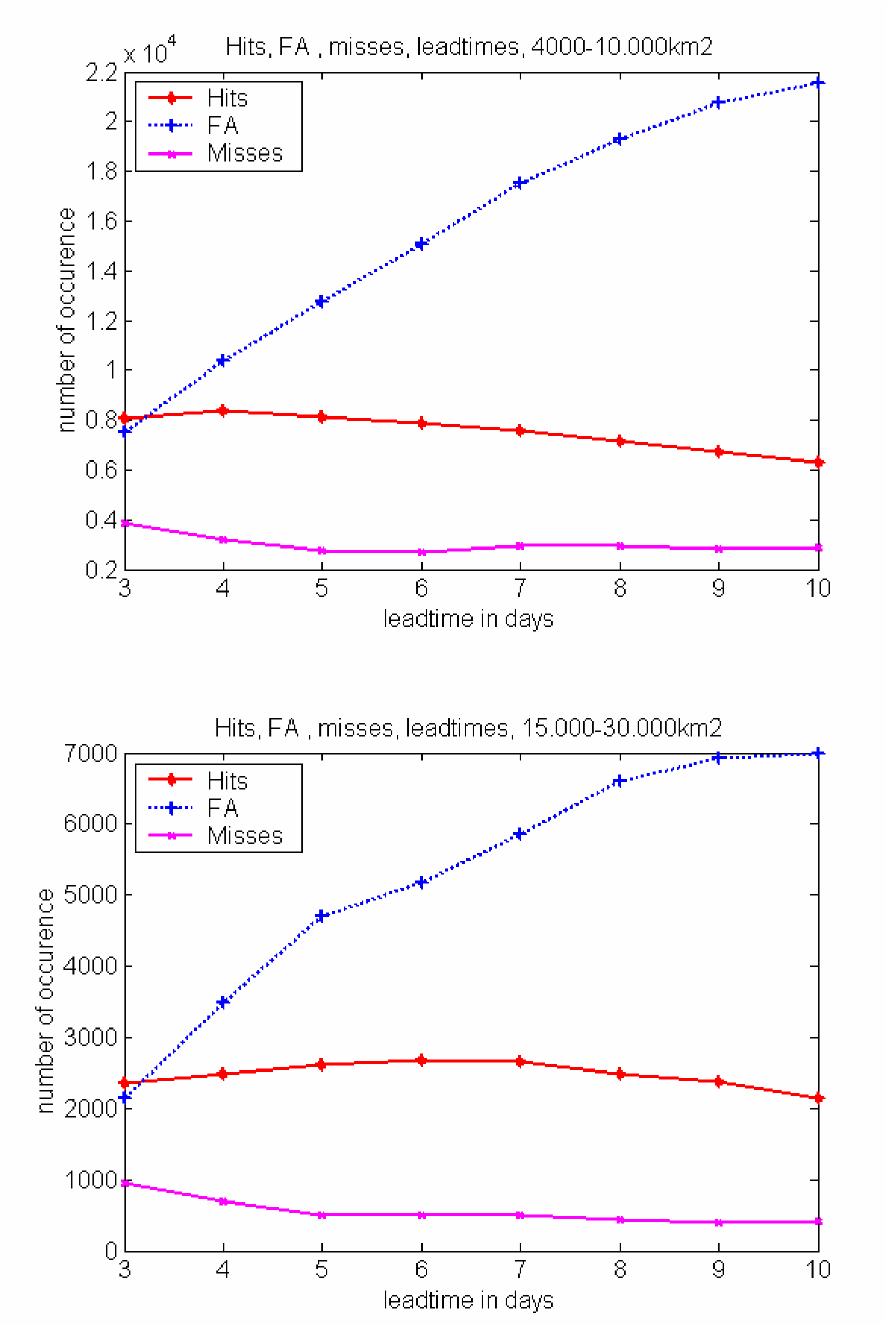


Figure 5.11a : Absolute numbers of hits, false alerts and misses are plotted over leadtime. Top : 4000 to 10.000 km^2, bottom 15.000 to 30.000 km^2

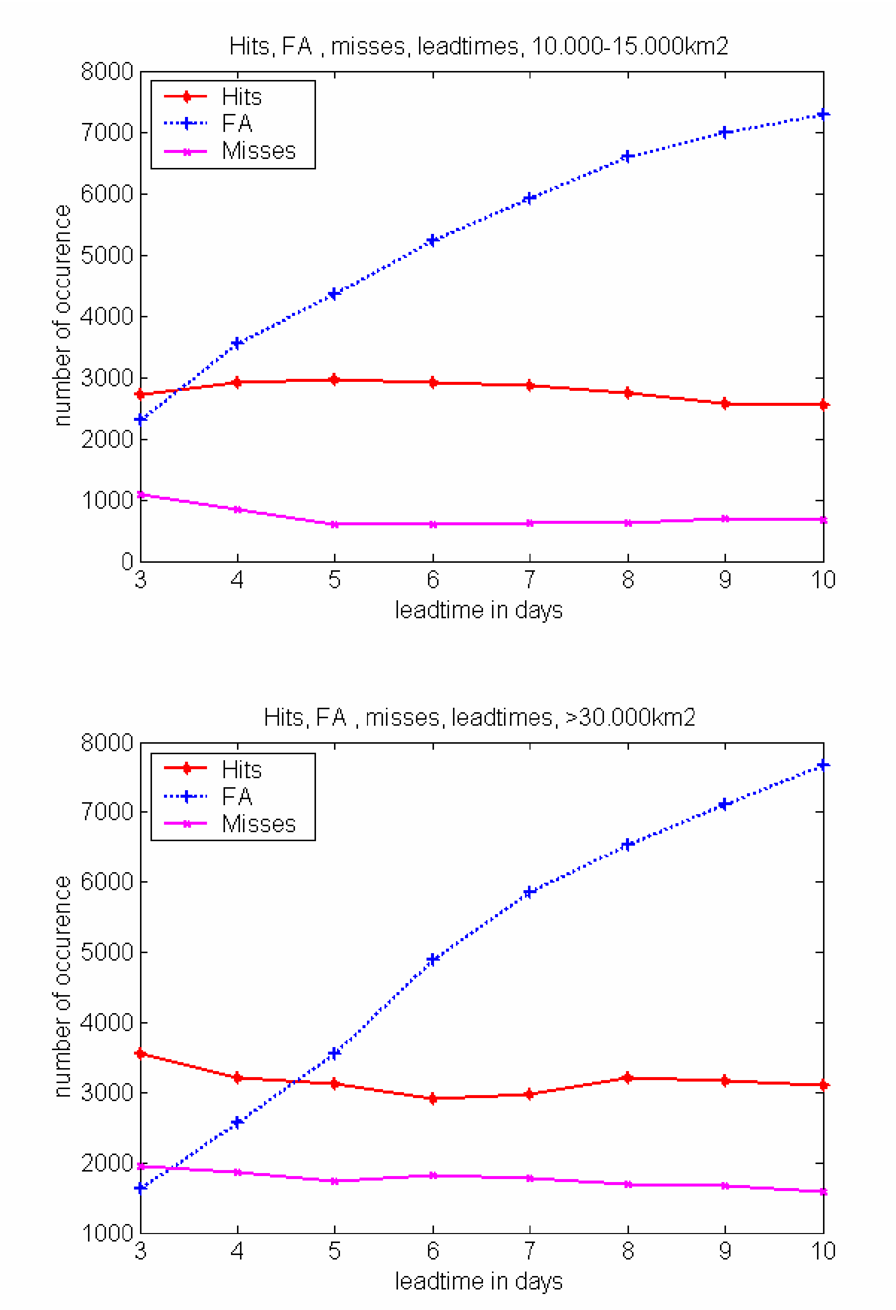


Figure 5.11b : Absolute numbers of hits, false alerts and misses are plotted over leadtime. Top : 10.000 to 15.000 km^2, bottom > 30.000 km^2

5.5.2 Deterministic EFAS

Figure 5.12 shows the skill in leadtime for the deterministic EFAS results again for the two classes of upstream area and conditioned on persistence.

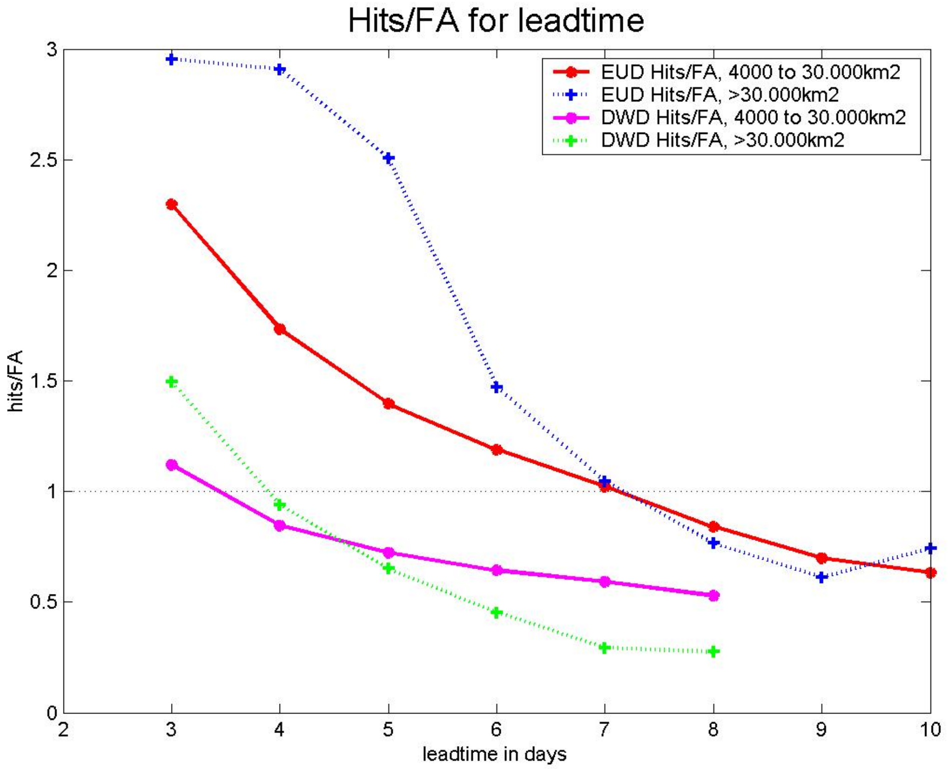


Figure 5.12: Ratio of hits over false alarm as a function of leadtime for EFAS forecasts based on ECMWF deterministic runs (EUD) and DWD deterministic runs (DWD).

The results show clearly the skill of EFAS 12:00 forecasts based on both weather forecasts decreases with increasing leadtime, but that ECMWF seems to have a better performance than DWD. Based on DWD forecasts the skill for small upstream areas is lower than for larger upstream areas. This is somewhat unexpected since the DWD has a higher resolution forecast for the first 3 days and one could expect a higher skill particularly during the shorter leadtimes. Surprisingly, at larger leadtimes the skill for smaller upstream areas seems to decrease less than for higher upstream areas. This might be connected to the grade of influence of observed or forecasted meteo data. For small catchments the influence of observed meteo data decreases earlier than for large catchments. So for small catchments at 3 days leadtime the influence of observed meteo data is already small (thus the skill is already smaller) and the decrease in hydrological forecasting skill beyond 3 days leadtime is mainly linked to the quality of the forecasted meteo data. For large catchments the initially faster decrease in skill becomes similar to the one for small catchments at higher leadtimes when also for large catchments the influence of forecasted meteodata is dominant.

5.7 Gain in leadtime

Having analysed the forecasting skill regarding leadtime and upstream area it is also of interest to compare the different forecast types regarding potential leadtime gains.

Leadtime gains can only be calculated for events that happened, i.e. in the proxy the pixel registered a discharge > HAL. Therefore also FA cannot be considered here and thus the EPS threshold (here 5) is influencing the gain analysis. If the EPS threshold is set high there will be less gain than for low thresholds: Waiting for higher probability means coming closer to the event and having less leadtime gain.

Please note that this lead time gain analysis is based on an approach that takes into account only persistent (and not intermittent) forecasts. This choice was made due to restrictions in the used GIS macro language PCRaster.

The Δ in leadtime was computed as follows:

If the proxy for a pixel has a discharge >HAL

- and at least 5 EPS member with a leadtime corresponding to the date of the proxy also forecasted a discharge >HAL then this pixel “gains” 1 (Δ+1)
- and the respective deterministic forecast with a leadtime corresponding to the date of the proxy also forecasted a discharge >HAL then this pixel “loses” 1 (Δ-1)

This means that if both EPS and deterministic forecasts indicate discharges >HAL for this pixel at a certain data the Δ remains 0, is +1 if only EPS indicates discharges >HAL and -1 if only the deterministic forecast indicates a discharge >HAL.

This is repeated over all lead times and the final Δt is reported in the following graphic distinguishing between several upstream area classes.

On the y-axis the number of pixels in the class is reported and on the x-axis the Δ in lead time. The blue bars show the results for the ECMWF EPS forecasts when compared to the ECMWF deterministic forecasts, while the red bars show the results for the ECMWF EPS forecasts when compared to the DWD deterministic forecasts.

The zero Δ class is left vacant as a result of the above explained procedure.

Days 1 and 2 are not considered in this analysis (that is why DWD has max Δ 6 days and ECMWF det. max Δ 8 days).

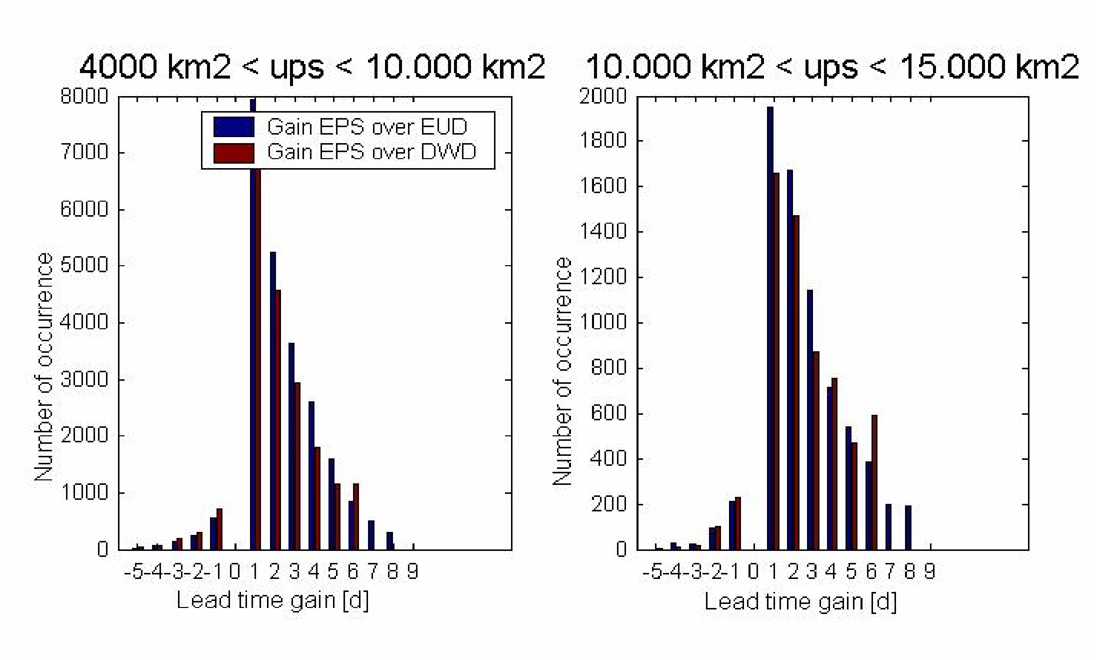


Figure 5.13a: Lead time gain of EFAS-EPS over EFAS deterministic forecasts

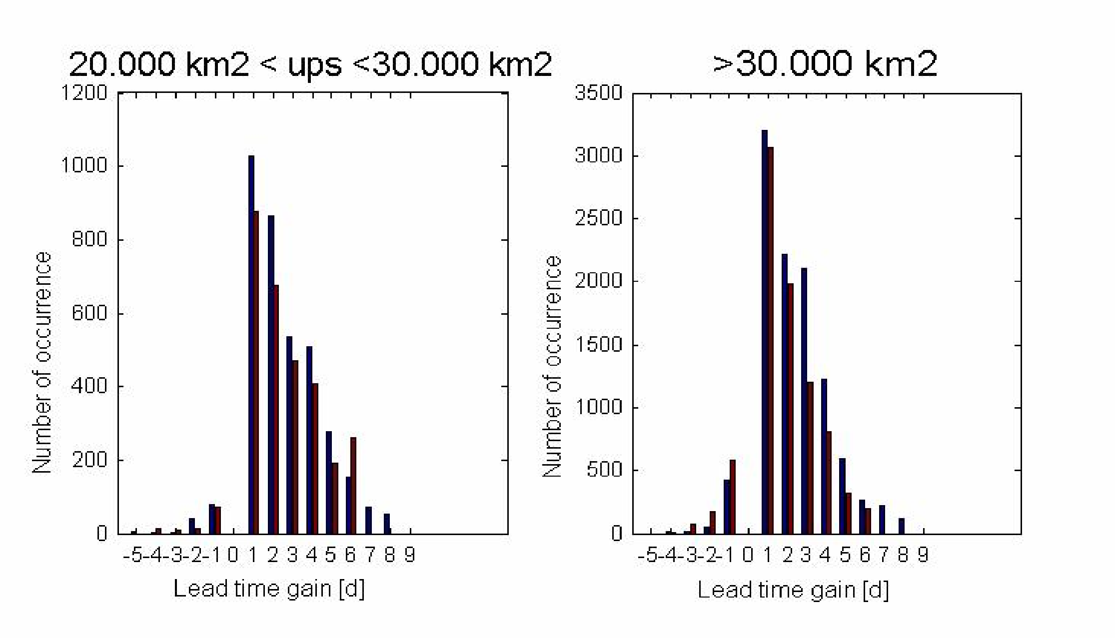


Figure 5.13b: Lead time gain of EFAS-EPS over EFAS deterministic forecasts

A similar analysis was done comparing the ECMWF deterministic forecast to the DWD deterministic forecast. In figures 5.14a/b positive Δ indicate that the ECMWF deterministic forecast had a lead time gain over the DWD deterministic forecast and negative Δ indicate that the DWD deterministic forecast had a lead time gain over the ECMWF deterministic forecast.

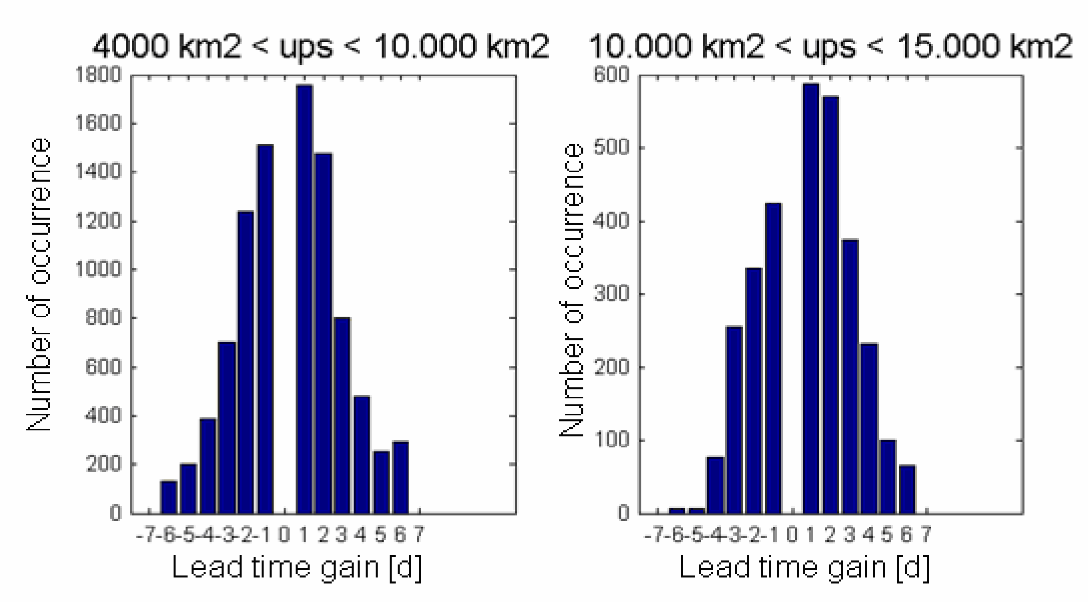


Figure 5.14a: Lead time gain of EFAS-ECMWF deterministic over EFAS DWD deterministic forecasts for different upstream area classes

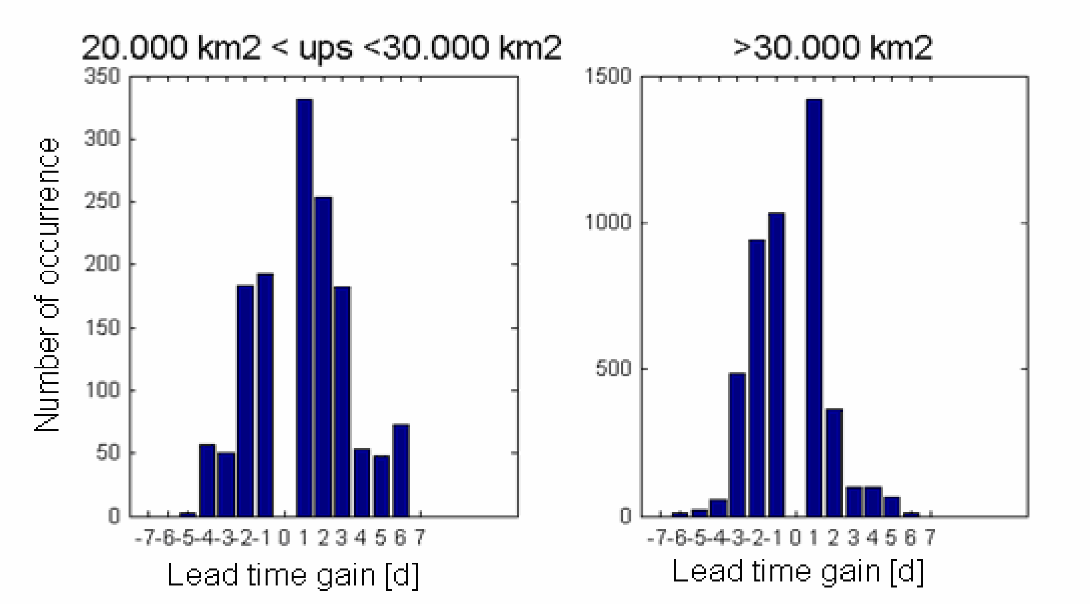


Figure 5.14b: Lead time gain of EFAS-ECMWF deterministic over EFAS DWD deterministic forecasts for different upstream area classes

The distribution can be approximated as Gaussian with mean 0 indicating no substantial leadtime gain of EFAS-ECMWF deterministic over EFAS-DWD deterministic forecasts. There is no evidence in any of the upstream area classes that one deterministic (DWD or ECMWF) forecast would allow for a longer leadtime than the other deterministic forecast.

5.7 Conclusions

The statistical analysis, pixel by pixel, of 11 months of EFAS-forecast data revealed most of all the usefulness of conditioning EFAS forecasts on a persistence criterion. If the persistence is based on a minimum number of EPS or on the deterministic forecast it always results in a considerable improvement of skill and reduction of false alarm rates. Other criteria can be envisaged and will be investigated in the future (i.e. a combination of EPS based and deterministic criteria could be envisaged).

If persistence is not taken into account, the forecasts for small upstream areas below 4000 km^2 reveal little to no skill. Conditioned on persistence, however, also the small scale areas reveal as much skill as the larger areas.

The probability distribution of EFAS-EPS members exceeding the EFAS high alert threshold is not a linear distribution like in equation 5.1. From the current analysis it can be estimated that the probability of an event to happen for the lower EPS bins up to 20 is higher than expected from the linear probability distribution.

In terms of leadtime, there seems to be a large potential of gaining lead time when using EFAS-EPS over the deterministic forecasts.
The comparison between the performance of EFAS based on different deterministic weather forecasts shows that EFAS ECMWF deterministic forecasts performed better as for all upstream areas there were more hits

than False Alarms up to 5 days of lead time while this was not the case for the EFAS–DWD forecast.

References

Wilks, D. S., 1995: Statistical Methods in the Atmospheric Sciences. Cambridge Press, 467 pp.

Hamill, T.M., and J. Juras, 2005: Common forecast verification metrics can overestimate skill. Mon. Wea. Rev., submitted

Legg, T. P., and Mylne, K. R., 2004: Early Warnings of Severe Weather from Ensemble Forecast Information. Wea. Forecasting, 19, 891-906.

ANNEX 1: Example of Feedback questionnaire

FEEDBACK QUESTIONNAIRE

for

EFAS INFORMATION REPORTS

Basin:
From:
To:

Question	Yes	No	Don’t know
1. Is the information in the EFAS reports stated clearly? Comments:			
2. Is all the information in the EFAS reports necessary? If No, what information is not useful?			
3. Would you like to have more information? If Yes, what kind of information?			
4.☐o you find the EFAS reports useful? If Yes, the EFAS reports were useful as early warning information additional information to local forecasts overview information from neighbouring countries help in decision making other If No, what could be improved?			

5. The first EFAS report is produced when high or severe flood threshold levels have been exceeded at least for 3 consecutive EFAS forecasts. This can decrease the false alarm rates but also shortens the leadtime. Would you like to receive the reports earlier although this may mean an increase in false alarm rates? Comments:			
6. At present, the first report is sent when at least one of our meteorological forecasts (ECMWF or DWD) results in high or severe flood levels. Would you prefer to receive the first EFAS report only when **both** meteorological forecasts indicate them? Comments:			
7. Flood ensemble prediction system information is given in the form of maps counting the number of Ensemble Forecasts (EPS) generating discharges exceeding critical flood level thresholds. Do you find this information useful? Comments:			
8. Would you like to have more information on Flood Ensemble Prediction System, i.e., more EPS results as, for instance, the spread of the forecasted hydrographs, statistical information for each leadtime, or others? If Yes, which information?			
9. ☐ EFAS reports were sent to the technical contact e-mail specified in the Memorandum of Understanding. This e-mail address is: a single physical person. an e-mail list addressing to a group of forecasters. other : ___________________________________			
10. The EFAS reports reached the forecasters on duty a.			

☐directly.

via the technical contact specified in the Memorandum of Understanding.

other : ______________________________

b.

at the time it was sent.

in the following 24 hours.

more than 24 hours after it was sent.

11. EFAS is a system under research and development, not an operational system. In the same way, the EFAS team is composed of researchers and not operational forecasters. Nevertheless, the EFAS forecasts are produced on a daily basis including weekends.
External reports are issued only once a day when both meteorological forecasts (00:00 and 12:00) are available and not before 16:00.

You would prefer receiving the EFAS reports
☐ in the morning, based on the 12:00 weather forecast from the previous day
☐ in the afternoon, based on the 00:00 weather forecast from the current day
☐ morning *and* afternoon, two reports based on 12:00 and 00:00 weather forecasts

☐ in its current format (pdf-File sent by email)
☐ via a web-interface that can be accessed any time

Comments:

POST-ANALYSIS

Please answer these questions once the period of the forecasted event is over

1. ☐ For the river basin flagged in the EFAS reports

a. Flooding was observed ?

Yes.

No.

b. If flooding was not observed, did river levels reach bankful conditions somewhere in the river basin?

Yes. Where? ____

No.

c. River levels were mainly:

high and critical.

high but not critical.

only moderately increased or normal.

d. If flooding was observed, indicate the areas where it was observed and, if possible, give some summary qualitative and/or quantitative information about the event:

2. ☐ Regarding to timing, EFAS forecasts were:

mainly correctly in time.

mainly too late.

mainly too early.

3. Were the EFAS reports used in some way by the flood forecasting team?

☐ Yes.

☐ ☐ No.

If Yes, how?

- Basis for discussion in the forecasting team
- Arrange working schedule
- Increase local flood forecasting frequency
- Other

If No, why not?

Did the EFAS reports effectively help you?

Yes.

No.

If Yes, how? If No, why not?

4. ☐ Would you suggest improvements to the EFAS reports?

Yes.

No.

If Yes, indicate your suggestions:

Any other comments?

THE EFAS TEAM THANKS YOU FOR GIVING FEEDBACK WHICH IS VERY VALUABLE FOR THE DEVELOPMENT AND TESTING OF EFAS!

PLEASE RETURN THIS QUESTIONNAIRE AS SOON AS POSSIBLE:

- by e-mail: efas@jrc.it

or

- by fax: Jutta Thielen
  NHLMU/IES/WDNH
  INT+39 + 0332 786653

ANNEX 2: Summary of 1st EFAS enduser meeting

JRC WDNH Action *nr 4336*

Floods & other weather driven natural hazards Action

Title and Description	*Minutes of the 1st 2006 EFAS-MoU technical Meeting*
Creator	*J. Bartholmes (JB)*
Date	*2006-01-23*
Publisher	JRC-Institute for Environment and Sustainability, Ispra
Contributors	Jutta Thielen
Rights	
Identifier	*EFAS-technical-meeting_minutes_20060123.pdf*
Coverage	

Location: JRC *Bld.36, Room 3*	Date of Meeting:*23-01-2006*	Duration: *09.00 – 16.00pm*
Recorder of Minutes: JB	Name: JB	
List of Attendees	EFAS related staff : • Ad de Roo (AdR) • Davide Muraro • Helena Ramos (HR) • Jalal Younis • Janos Szabo • Jens Bartholmes (JB) • Jutta Thielen (JT) • Karl Wachter (KW) • Meike Gierk (MG) • Milan Kalas • Mauro del Medico • Simone Gentilini	External participants : *See attached participants list with organization and contact details*

Attendance:

Representatives of 14 MoU partner organisation including Czech Hydrometeorological Institute (CZ), National Institute of Meteorology and Hydrology (BG), Vituki (HU), SAIH-Ebro (ES) , Landesamt fuer Umwelt, Wasserwirtschaft und Gewerbeaufsicht Rhein (DE), Slovak Hydrometeorological Institute (SK), Environment Agency (UK), Landesumweltamt Brandenburg (DE), Agenzia per la protezione dell'Ambiente e per I Servizi Tecnici (IT), Vilnius University (LT), ARPA servizio Idrometeo (IT), Agencija republike Slovenije (SI), Flanders Hydraulics research (BE), Ministry for ecology and sustainable development- SCHAPI (FR) participated.

The following organisations could not send representatives at this particular date: Institute of Meteorlogy and Water Management Wroclaw Branch (PL), RIZA (NL), Bundesanstalt fuer Gewaesserkunde (DE), the Landesanstalt fuer Umweltschutz BadenWuerttemberg (DE), Federal Ministry of Agriculture, Forestry, Environment and Water Management (AT).

In two cases there was no response to the invitation: Saxon flood centre (DE) and the State Hydrometeorological Service (SHS) of Moldova (MD).

Objectives of meeting:

The JRC invited to the 1st EFAS MoU technical meeting all EFAS partners that signed the Memorandum of Understanding (MoU), have a MoU in preparation, or are currently negotiating a MoU with the JRC. The scope was to:

- present to the MoU partners the current state of the EFAS system
- inform about the related research
- get their feedback on operational issues
- discuss the future improvements for EFAS directly with them
- inform about meteorological data issues
- inform about MoU related issues
- facilitate in general network building and experience exchange

Meeting results:

The meeting was divided into 2 parts. In the morning a series of EFAS related presentations and guided discussions took place. In the afternoon the focus was on feedback and experience exchange with the MoU partners and open discussions.

The EFAS team presentations dealt with:

- the scope of EFAS and its general political context
- The status quo of the MoU up to date
    - 14 signed

  - o 3 in progress
  - o 5 awaiting agreement from ECMWF
  - o In average it took 2-3 months to complete the process but in some cases there were big delays due to late response or bureaucratic hold-up
- • Technical background of EFAS
  - o Discussion on EFAS critical thresholds and their statistical and data background
    - Concern that 14 years as base for a statistical calculation of thresholds are not enough but all that is available right now. Longer data time series needed.
    - ECMWF reanalysis data (lower resolution that the now used (JRC-MARS data) will be evaluated for extending the threshold calculation to a base of 40 years.
- • EFAS performance measured in hit and false alarms over the whole year 2005
  - o 10 external EFAS information reports send out to MoU partners (specific feedback see also below => EFAS Information Report (EIR) feedback section)
  - o 27 Internal EFAS alerts (situation under surveillance but no external EIR) in comparison to hydrological simulations with observed meteorological data
  - o Verification difficulties due to lack of information on minor flooding events
- • Presentation of results from the 1st EFAS workshop on the use of EPS in flood forecasting, held 21-22nd November 2005 with selected EFAS partners.
- • Communication issues:
  - o Problems sometimes reaching partners, e.g. Italy did not receive the EFAS Information reports on time because of email problems
  - o E.g. some MoU partners never responded to the invitation of the meeting
  - o In one case the email server address had changed and the partner could not be contacted.
  - o Some have technical problems, e.g. APAT, and then should provide an alternative contact address
  - o MoU partners should inform EFAS of any change of address (email address, server, etc)
- • Highlighting of the issues that have to be discussed with ECMWF:
  - o How to react with pending MoU's, e.g. case of Slovenia. General understanding that full EFAS reports can only be send when the agreement has been signed – but what about informal emails? This should be iterated with ECMWF.
  - o Information email only in case MoU exist but area too small?
  - o Daily information on a password protected web?
  - o Sometimes the pending of MoU requests can be very long – e.g. the case of Slovenia (pending since June) and UK ENV Agency, where conflicting information is received from the Agency and the ECMWF council. To be iterated with ECMWF
- • Presentation of the EU-FLOOD-GIS and IDA-GRDC initiative for data

collection within the framework of EFAS. Announcement that the National Authorities will be approached here in the near future.

- Announced dates:
  - 1st March, next ECMWF meeting. On the agenda of this meeting, amongst others, also catchment size, informal emails, daily exchange of information with partners
  - Next technical EFAS meeting again end of January/beginning of February 2007

Summary of feedback, related discussions and ad hoc presentations:
(The .ppt presentations are annexed at the end)

JT gave an overview presentation of the feedbacks received so far to the EFAS information reports (see annexed presentation). In total 5 feedback questionnaires were returned and 3 informal emails. Highlighted should be the following results:

- Form, presentation and contents of the EFAS information reports appear to be appropriate. Suggestions on colour schemes, image size have been given and partially already taken into account in the reporting.
- The feedbacks are inconclusive regarding information on EPS. Half of the responses indicate that more information on EPS is desired, the other half is content with the present information
- Generally the MoU partners would like to have the EFAS reports as early as possible even if this may mean a higher false alarm rate
- EFAS information reports that were received on time have been used actively in the local flood forecasting team.

Discussion: what to do if EFAS simulates flooding in small upstream areas only and not in the larger rivers qualifying for the 30000 km2. In this case an email without EFAS Information Report (EIR) informing the partners that a possibility of flooding in small upstream areas is possible would be appreciated.

EFAS Information Report (EIR) feedback from the MoU partners

Gabor Balint (Vituki) received EIR on three different occasions:

- **07.2005** *Danube well forecasted and EIR useful. Raba, Drava medium high as forecasted.*
- **08.2005** *Mures: historical high observed, forecasted: severe. Leadtime gain >=1 day. Very useful for operational forecast. Correct forecast of 2nd wave.*
- **10.2005** *Drava medium high. First EFAS forecasts overestimated, but with each forecast the predicted severity decreased. Timing ok.*

Gabriela Babiakova (SMHI):

- **07.2005** *Danube 4days lead time with well forecasted time to peak.*

3rd national alert level reached.

AdR for BfG:

- **08.2005** *If EPS had been available operationally during South-Germany flood, 1 day more of EFAS lead time could have been gained. Small Bavarian and Swiss rivers were forecasted well.*

Mojca Susnik (Slovenia):

- **10.2005** *Slovenia was contacted by EFAS that there was a possibility of flooding in the Drava, and if ECMWF should be approached to ask for permission to send a full EFAS information report. Mrs Susnik said that although the permission for a full EFAS information report was never given, the email brought their attention back to the Drava, which was not their focus of priority because of the expected flooding in the Sava. She said that this was very useful for them.*

Catchment size discussion

Jan Kubat (CHMI)
30.000 km2 catchment EFAS restriction should be re-discussed with ECMWF. Inclusion in EFAS of smaller catchments would be appreciated. 10 day old EFAS forecast on password restricted web page.

Gabor Balint (VITUKI)
Area 30.000 km2 ok but just referring to meteo even not catchment size would be a better definition.

General agreement on this point.

Jutta Thielen, Ad de Roo(JRC)
30.000 km2 issue has to be renegotiated in ECMWF council.

EIR cover period

Gabriela Babiakova:
Useful to receive EFAS Information Report (EIR) for at least 2 days after the flood peak is over.

JT proposes to split EIR in two if 2 consecutive flood peaks are forecasted.

Caroline Wittwer:
Better no splitting of EIR for 2 consecutive peaks as hydrological influence of 1st peak is high (IC for next peak). Continued EIR until 2nd peak is over.

Fixed point forecasts

Shall fixed predefined points (GRDC ?) be integrated into EFAS Information Report (EIR) ?

Discussion on this topic postponed.

EFAS Colour scale

Stefano Tibaldi:
If the necessarily different colour scales of EPS and deterministic forecasts are display close to each other it is visually confusing if similar colours are used, e.g. red. EFAS Information Report (EIR) strategy of putting EPS and deterministic forecasts on separate pages already helps.

Jan Kubat:
Check colour coding of other natural phenomena alerts.

Caroline Wittwer:
Colour connected to different risk interpretation. Distinction between risk and hazard could be discussed in the future.

The EFAS colour scale was already examined closely during the past months and adapted to needs of end users. Differences in monitor/print appearances are taken into account. This also applies to EPS counting colour code. No further explicit change was demanded during the meeting.

Layouts of EFAS products

General agreement that the present layouts of EFAS Information Reports, EFAS feedback questionnaires, and bulletins are ok.

Information exchange between MoU partners and JRC

Gintautas Stankunavicius:
Can EFAS give access to a ftp folder per MoU partner for data exchange?

Jutta Thielen
JRC is planning to upgrade the EFAS webpage in a dynamic way to include possibilities of discussion forum and specific folders per MoU partner. Daily data could also be put there on password protected sites to give the MoU partners restricted use.

Jutta Thielen
Question if there is a need for onsite EFAS training similar to the EPS workshop held in November 2005 in Ispra. This possibility has been generally perceived as a good idea by the partners and should perhaps be envisaged in the future. One could also image that JRC staff goes to the MS organisations to present EFAS to the local forecasting staff. The aspect of "training" on EFAS products was a general issue. Providing daily EFAS results for a certain period per catchment might be a possibility to give the local forecasting centres experience with the EFAS products.

EFAS issues raised during discussions:

Jan Kubat:
EFAS Information Report (EIR) receipt return message can be helpful for sent EIR.

Eram Artinyan:
Precipitation and snowmelt info together with EIR could be helpful.

Patricia Pérez:
Ebro flood 02.2003 EFAS rerun possible?

Caroline Wittwer:
Would it be possible to have access to a password protected webpage with daily EFAS forecasts even in non flooding situations to get expert knowledge on the products.

This is already planed for Vituki and Czech Hydrometeorological Institute as a test phase.

Jan Kubat:
Would be good if every partner could send mail to EFAS if flood was observed but no EIR received.

Closure of open discussion with applause.

****

Extended AGENDA of the meeting

09:00-09:15 Welcome (AdR)
- Political background EFAS
- ECMWF-JRC agreement

09:15-09:30 Introduction of participants
09:30-10:00 Presentation on EFAS in 2005 (JT)
- MoU negotiations
- Number of MoU's established
- Problems in establishing MoU's
- Number of external reports send out + feedback
- EFAS products: reports, Bulletins
- Research on EPS & EFAS workshop on the use of EPS
- guidelines
- aim of this meeting

10:00-10:30 Technical presentation of EFAS (JB)
- system set-up

- EFAS interface
- EMM
- Overall results 2005

10:30-10:45 Coffee Break

10:45-11:15 EPS research results in EFAS (HR)
- case study
- 1st FEPS workshop

11:15-11:30 Summary Danube catchment on 1 km (KW)
11:30-11:45 Summary Elbe catchment on 1km (MG)

11:45-12:30 general discussion

12:30-14:00 LUNCH

14:00-14:15 overview of feedback from EIR (JT)
14:15-15:00 experience from WA having received EFAS information report (EIR)
- Gabor Balint summarizes Vituki's experience with EIR
- Gabriela Babiakova summarizes SMHI's experience with EIR
- AdR reports on BFG feedback regarding August flood in South Germany
- Gintautas Stankunavicius EIR related presentation regarding Lithuania

15:00-17:00 Specific discussions on
- EIR (style, colour schemes, etc)
- Monthly bulletins
- Training documents or training data sets
- Feedback questionnaire
- Information flow between JRC and organisations
- Information flow within the receiving organisations
- Improve feedback: from receiving organisation necessary – perhaps per web in the future??
- Webpage- more interactively in the future including online surveys for research, EFAS products, …
- Password protected sites for EFAS with full information? Decision on 1-2 test partner
- Frequency of EFAS technical meetings
- AOB

European Commission

EUR 22560 EN – DG Joint Research Centre, Institute for the Environment and Sustainability
Title: The benefit of probabilistic flood forecasting on European scale
Authors: Bartholmes J. , Ramos M.H., Thielen J., and et al.
Luxembourg: Office for Official Publications of the European Communities
2007 – 99 pp. – 21 x 29.7 cm
EUR - Scientific and Technical Research series; ISSN 1018-5593

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

Publications Office
Publications.eu.int