**Natural Hazards
and Earth
System Sciences**

Verification of operational Quantitative Discharge Forecast (QDF) for a regional warning system – the AMPHORE case studies in the upper Po River

D. Rabuffetti[1], **G. Ravazzani**[2], **C. Corbari**[2], **and M. Mancini**[2]

[1]ARPA Piemonte-Forecast and Survey area, Turin, Italy
[2]DIIAR-Politecnico di Milano, Milan, Italy

Abstract. In recent years, the interest in the prediction and prevention of natural hazards related to hydrometeorological events has grown due to the increased frequency of extreme rainstorms. Several research projects have been developed to test hydrometeorological models for real-time flood forecasting. However, flood forecasting systems are still not widespread in operational context. Real-world examples are mainly dedicated to the use of flood routing model, best suited for large river basins. For small basins, it is necessary to take advantage of the lag time between the onset of a rainstorm and the beginning of the hydrograph rise, with the use of rainfall-runoff transformation models. Nevertheless, when the lag time is very short, a rainfall predictor is required, as a result, meteorological models are often coupled with hydrological simulation. While this chaining allows floods to be forecasted on small catchments with response times ranging from 6 to 12 h it, however, causes new problems for the reliability of Quantitative Precipitation Forecasts (QPF) and also creates additional accuracy problems for space and time scales.

The aim of this work is to evaluate the degree to which uncertain QPF affects the reliability of the whole hydrometeorological alert system for small catchments. For this purpose, a distributed hydrological model (FEST-WB) was developed and analysed in operational setting experiments, i.e. the hydrological model was forced with rain observation until the time of forecast and with the QPF for the successive period, as is usual in real-time procedures. Analysis focuses on the AMPHORE case studies in Piemonte in November 2002.

Correspondence to: G. Ravazzani
(giovanni.ravazzani@polimi.it)

1 Introduction

In the field of natural weather-related calamities, flooding is the worst hazard, causing loss of life and significant damage to property (Carpenter et al., 1999). In 2005, 171 large flood events were reported worldwide (Brakenridge et al., 2005) with damages of more than 82 thousand million U.S. dollars.

In the Upper Po River basin in Italy, located directly south of the Alps, several flood events have occurred in the very last years, for example, September 1993, November 1994, November 1996, June and October 2000 and November 2002. On the basis of historical data that goes back to 1800, this region has been hit by severe meteorological events, on average, once every two years.

Two strategies are required to reduce flood losses, structural protection measures (such as flood retention basins for the reduction of flood hazard), and real time flood forecasting systems that reduce flood risk by issuing warnings (with the complementary strategy of the education of the public on the appropriate response to warnings).

As far as the operational implementation of real time flood forecasting systems for Civil Protection is concerned, the numerical models main requirements are: reliability for an accurate flood forecasting and computational speed to maximize the warning lead-time. Nevertheless, when the lag time is very short, to extend the lead-time between warning and flood event, a rainfall predictor is required (Lardet and Obled, 1994; Brath et al., 1988). Starting from the pioneering work of Georgakakos and Hudlow (1984), meteorological models are often coupled with hydrological simulations (Jasper et al., 2002; Montaldo et al., 2002; Ludwig et al., 2003). While this chaining allows floods to be forecasted on small catchments with response times ranging from 6 to 12 h it, however, causes new problems for the reliability of Quantitative

Table 1. Main characteristics of the catchments involved in the analysis.

Cross-section	River	Drained Area [km^2]	Lag time [h]	Warning code 2 [m^3/s]	Warning code 3 [m^3/s]
Castelnuovo	Belbo	422	15.1	300	500
Carignano	Po	3976	18	700	1300
Cassine	Bormida	1521	23.2	750	1300
Serravalle	Scrivia	619	10	800	1500
Palestro	Sesia	2587	18.8	1900	3400
Farigliano	Tanaro	1508	14.8	750	1500
Candoglia	Toce	1531	9	1150	2100

Precipitation Forecasts (QPF) and also creates additional accuracy problems for space and time scales.

In this work a distributed hydrological model (FEST-WB) was developed to co-operate with the existing warning system of Regione Piemonte (Rabuffetti and Barbero, 2005), in Italy, to predict floods in small catchments. The use of FEST-WB is analysed in an operational setting. As it is usual in real-time procedures, the hydrological model is forced with rain observation until the time of forecast and with the QPF for the successive period. Analysis focuses on two case studies in Piemonte, i.e. the ITALIA1 (14–18 November 2002) and the ITALIA2 (24–28 November 2002) test cases. The objective of analysing these is to evaluate the degree to which uncertain QPF affects the model's capability of producing reliable QDF. Also, the accuracy of the flood forecasting system is investigated assuming a lack of meteorological data.

This study was part of the AMPHORE (Application des Methodologies de Previsions Hydrometeorologiques Orientees aux Risques Environnementaux) European Union research project whose objective regards the forecasting and the prevention of natural hazards, with particular reference to risks coming from severe hydro-meteorological events. The main goal was the optimisation of the existing warning systems in the Western Mediterranean Basin (Medocc area) by the implementation and the usage of probabilistic and statistical methods in which different scenarios are embedded together to improve the Quantitative Precipitation Forecast (QPF).

2 The upper Po river basin

The subject area is the upper Po River basin and covers 38 000 km^2. This is predominantly an alpine region that is bounded on three sides by mountain chains covering 73% of its territory. While the north-east part of the basin is located in Switzerland, most of it is in the north-west of Italy. Piemonte, located in the Padana Plain, is the principal Italian region in the area while other Italian regions include Valle D'Aosta, Liguria and Lombardia.

The Civil Protection (Table 1) adopts two levels of warning discharge thresholds: moderate flood (Code 2) and extreme flood (Code 3). A Code 2 flood refers to a peak discharge that does not produce flooding but can seriously damage structures in the river; a Code 3 flood addresses a peak discharge that produces flooding. Code 1 identifies normal floods characterized by no significant effects on the territory. The discharge thresholds are defined on the basis of the capacity of each cross-section under consideration and its river branch. They are evaluated by means of off-line hydraulic analysis and, when available, on the basis of historical floods data.

2.1 Physiographic basin characterization

Available digital cartographic data include: the Digital Elevation Model (DEM) available in raster format at 100 m×100 m resolution; CORINE land cover maps (CEC, 1994, EEA, 2000) for the Italian part updated in the year 2000 available in vector format and CORINE maps for the Switzerland part updated in the year 1990 available in raster format (spatial resolution of 250 m×250 m); pedologic characteristics for soils available in vector format. From the available basic thematic layers, basin parameters required for the application of the hydrological model have been derived at a spatial resolution of 1000 m×1000 m. These include: Curve Number (Soil Conservation Service, 1986), flow direction, slope and aspect, residual and saturated soil moisture, albedo, pore size distribution index, saturated hydraulic conductivity, wilting point, field capacity and soil depth.

2.2 Hydrologic and meteorological data

For this study, meteorological and hydrologic ground measured data were collected by the telemetric monitoring system of the Regione Piemonte, Regione Lombardia and Switzerland. Data of rainfall, air temperature, incident short wave solar radiation and air relative humidity are available from 1 January 2000 to 31 December 2004 at hourly or sub-hourly time steps.

Hydrometric observations at 30 min time step are available at 40 locations from 1 January 2000 to 31 December 2003. In the present analysis, only medium-sized basins (Table 1) with areas ranging from 422 to 3976 km^2 are considered
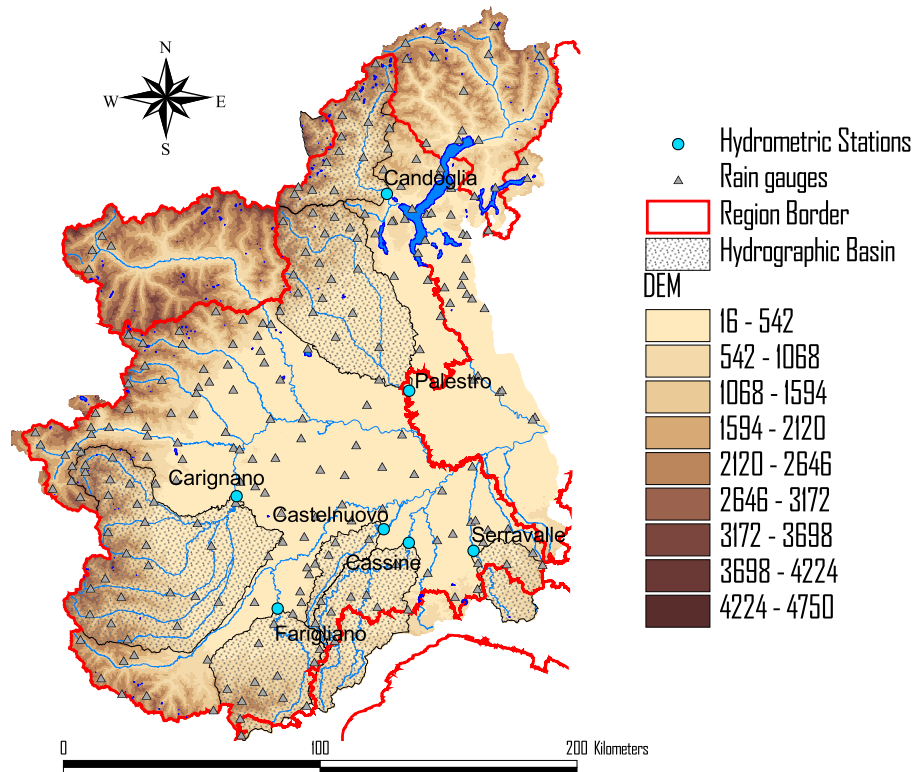
Fig. 1. The upper Po River basin showing locations of the rain gauges and hydrometric stations and related watersheds.

(Fig. 1) as these were the most affected by the events used in this study. Locations characterized by lack of meteorological observations have been intentionally excluded from analysis.

2.3 The case-study flood events

The analysis focuses on the test cases selected in the AM-PHORE project, i.e. the 14–18 November 2002 (ITALIA1) and 24-28 November 2002 (ITALIA2). As already stated, these mainly affected small and medium-sized catchments and locations around their confluences with the main rivers. On the whole, these events did not produce extreme floods and, in terms of damage, effects were negligible.

During the ITALIA1 event, Code 2 warning was detected in four locations, namely, Cassine, Palestro, Farigliano and Candoglia. In the ITALIA2 event, only at Serravalle did discharge exceed Code 2 warning threshold. Discharge data for Carignano are not available for the ITALIA2 event. No any Code 3 warning was detected during the two flood events. These events can, therefore, be considered as lower thresholds above which floods will become very dangerous. It is in moderate events such as these that it is most important to understand the capability of the warning system.

Available QPFs for ITALIA1 and ITALIA2 events are from ECMWF (spatial resolution of 40 km×50 km, time step of 6 h) and from two different sets of LAMI (resolution of

7 km×7 km, time step of 6 h) obtained by changing the initial and boundary condition of the meteorological model: S1 set is obtained by nesting LAMI with the ECMWF runs and, similarly, S2 with GME runs (Rabuffetti and Milelli, 2005).

3 The hydrological model

A distributed hydrological water-balance model (FEST-WB) was developed by starting from the event-based models FEST98 and FEST04 (Mancini, 1990; Montaldo et al., 2007). FEST-WB computes the main processes of the hydrological cycle: evapotranspiration, infiltration, surface runoff, flow routing, subsurface flow and snow dynamics. The computation domain is discretized with a mesh of regular square cells, within which water fluxes are calculated.

The model needs spatially distributed meteorological forcings. The observed data at ground stations are interpolated to a regular grid using the inverse distance weighting technique.

Soil moisture, θ, evolution for the generic cell at position i, j, is described by the water balance equation:

$$\frac{\partial \theta_{i,j}}{\partial t} = \frac{1}{Z_{i,j}} \left(P_{i,j} - R_{i,j} - D_{i,j} - ET_{i,j} \right) \tag{1}$$

where P is the precipitation rate, R is runoff flux, D is drainage flux, ET is evapotranspiration rate and Z is the soil depth.

Table 2. Statistics of peak discharge and flood volume relative errors before and after calibration. Negative bias denotes model underestimation.

Relative error	Before calibration		After calibration	
	Mean	Standard deviation	Mean	Standard deviation
Peak discharge	0.30	1.06	0.07	0.65
Flood volume	−0.01	0.70	-0.009	0.46

Runoff is computed according to a modified SCS-CN method extended for continuous simulation (Ravazzani et al., 2007) where the potential maximum retention, S, is updated at the beginning of storm as a linear function of the degree of saturation, ε.

$$S = S_1 \cdot (1 - \varepsilon) \tag{2}$$

where S_1 is the maximum value of S when the soil is dry (AMC 1).

The actual evapotranspiration, ET, is computed as a fraction of the potential rate tuned by the beta function that, in turn, depends on soil moisture content (Montaldo et al., 2003). Potential evapotranspiration is computed with a radiation-based equation (Priestley and Taylor, 1972).

The surface and subsurface flow routing is based on the Muskingum-Cunge method in its non-linear form with the time variable celerity (Montaldo et al., 2007).

The snow model includes the snow melt and the snow accumulation dynamics. The partitioning of total precipitation, P, in liquid, P_l, and solid, P_s, phase is a function of air temperature, T_a (Tarboton et al., 1994)

$$P_l = \alpha_P P \tag{3a}$$

$$P_s = (1 - \alpha_P)P \tag{3b}$$

where α_P is computed by:

$$\begin{cases} \alpha_P = 0 & \text{if } T_a \leq T_{\text{inf}} \\ \alpha_P = 1 & \text{if } T_a \geq T_{\text{sup}} \\ \alpha_P = \frac{T_a - T_{\text{inf}}}{T_{\text{sup}} - T_{\text{inf}}} & \text{if } T_{\text{inf}} \leq T_a \leq T_{\text{sup}} \end{cases} \tag{4}$$

where T_{inf} and T_{sup} are calibrating parameters.

The snow melt simulation is based on the degree day concept (Martinec et al., 1960). The melt rate in m/s, M_s, is proportional to the difference between air temperature and a predefined threshold temperature, T_b (Salandin et al., 2004):

$$M_s = C_m(T_a - T_b) \tag{5}$$

where C_m is an empirical coefficient depending on meteorological conditions and geographic location.

3.1 Model calibration

The model was subjected to a rigorous process of calibration and validation by comparison of simulated and observed discharge. Thanks to discharge time series availability, eleven locations could be chosen for this study. About 60 flood events from 2000 were used for calibration, while about 80 additional flood events from 2001–2003 were exploited for verification. Simulated hydrographs were extracted from the hourly time step continuous simulation for the period 2000–2003. The number of events that exceeded Code 2 and Code 3 warning threshold was, respectively, 23 and 7. The calibration and validation activity took into account only flood events in those locations where no significant debris blockages at river structures were reported. This to exclude the additional risk, namely, failure of the debris blockage and the release of impounded water that may result in extreme flooding from a modest rainstorm, not addressed in the present analysis. Flood events in those locations where widespread inundation was observed were excluded from calibration and validation process as well, as the hydrologic distributed model cannot simulate this.

The main calibration activity, based on the "trial and error" approach, was primarily focused on flood volumes, which are strongly dependent on the infiltration process. The parameters subjected to calibration were the soil hydraulic conductivity, regulating saturation dynamics and, undirectly, "Hortonian" infiltration excess, and soil depth, regulating "Dunnian" saturation excess (soil porosity is generally reliably taken from soil maps). To this aim, observed discharge hydrograph was subdivided into its superficial and deep flow components (Pilgrim and Cordery, 1993). The deep soil hydraulic conductivity, responsible for hypodermic flow velocity, was calibrated as well. The deep soil conductivity was checked comparing observed and modelled hypodermic flow and taking into account that an increase of deep soil conductivity implies an increase of hypodermic flow.

Furthermore, the parameters for the simulation of the snow dynamic were calibrated using snow cover maps derived from satellite images (Corbari et al., 2007; Rabuffetti et al., 2006).

The calibration effectiveness was evaluated with the change in peak discharge, $\varepsilon_{q\,\text{max,sim}}$, and flood volume, $\varepsilon_{\text{vol,sim}}$, relative error (Table 2).

$$\varepsilon_{q\,\text{max,sim}} = \frac{Q_{\text{sim}}^{\text{max}} - Q_{\text{obs}}^{\text{max}}}{Q_{\text{obs}}^{\text{max}}} \tag{6}$$

$$\varepsilon_{\text{vol,sim}} = \frac{V_{\text{sim}} - V_{\text{obs}}}{V_{\text{obs}}} \tag{7}$$

where $Q_{\text{sim}}^{\text{max}}$ is the simulated peak flow (m³/s); $Q_{\text{obs}}^{\text{max}}$ is the observed peak flow (m³/s); V_{sim} and V_{obs} are the volume (m³) of the discharge hydrograph, respectively, simulated and observed.
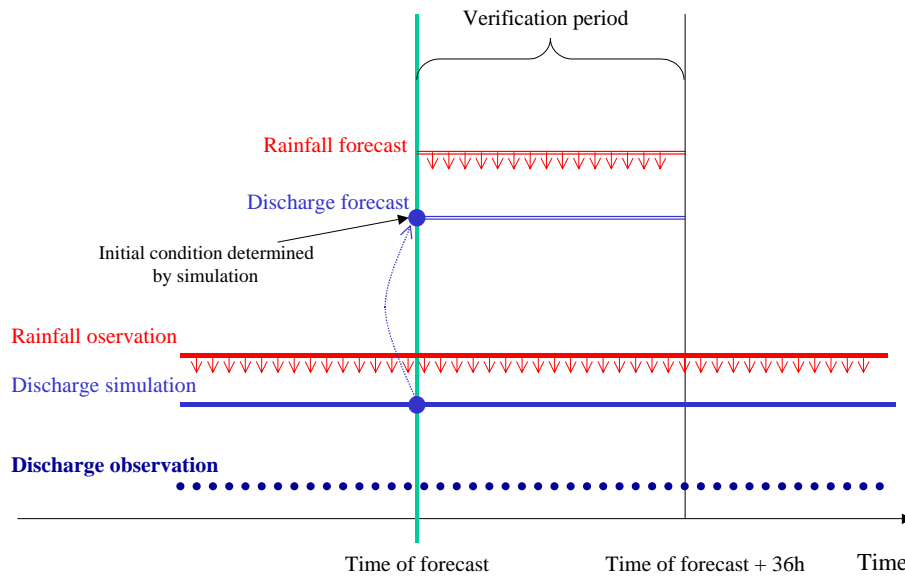
Fig. 2. Sketch of the verification process.

4 Coupling strategy

The runs were initialised at: 12:00 UTC of 14 November 2002, 12:00 UTC of 15 November 2002, 00:00 UTC of 24 November 2002, 00:00 UTC of 25 November 2002. The hydrological forecasts have a horizon of 36 h. Rainfall forecasted fields were downscaled to hydrological model spatial resolution using the simple nearest neighbour resampling method. Rainfall intensity was kept constant for each computational time step of the meteorological models (6 h).

Discharge forecasts at hourly time step, were obtained by employing meteorological forecasts in the hydrological model and soil moisture initial conditions were obtained from an off-line simulation of the whole period, 2000–2003, by using the observed data at ground stations (Fig. 2).

In the verification period, the assumption that evapotranspiration fluxes are negligible during flood events occurring in November meant that the hydrologic model was only forced with forecasted rainfall.

5 Uncertainty propagation from QPF to QDF and warning system

5.1 Verification procedure description

The primary need for the analysis of the ITALIA1 and the ITALIA2 test cases was to evaluate how much QPF uncertainty affects model capability in producing reliable QDF.

A comparison was also made where the QDF is obtained when QPFs are not available and the forecasted precipitation is taken as null. This offers the opportunity of understanding the real adjoined value in the use of QPF in flood forecasting; of course it is not a general result as it is specific to the catchments and the events considered.

To synthesize hydrological model behaviour, a relative error was calculated for each flood forecast either in terms of peak discharge, ε_{QDF}, (Eq. 8) or in terms of flood volume, ε_V, (Eq. 9).

$$\varepsilon_{\text{QDF}} = \frac{Q_{\text{sim,for}}^{\max} - Q_{\text{obs}}^{\max}}{Q_{\text{obs}}^{\max}} \tag{8}$$

$$\varepsilon_V = \frac{V_{\text{sim,for}} - V_{\text{obs}}}{V_{\text{obs}}} \tag{9}$$

where $Q_{\text{sim,for}}^{\max}$ is the simulated or forecasted peak flow (m³/s), $V_{\text{sim,for}}$ is the volume (m³) of the discharge hydrograph simulated or forecasted. In addition, to account for the geomorphologic differences of each studied catchment, the advance time index, a normalized index that takes into account the advance of forecasts versus lag time, is defined (Eq. 10) (Rabuffetti and Milelli, 2005).

$$i_a = \frac{T(Q_{\text{obs}}^{\max}) - \text{TOF}}{T_c} \tag{10}$$

where $T(Q_{\text{obs}}^{\max})$ is the time, in hours, of the observed peak discharge; TOF (hour) is the time of forecast; T_c (hour) is the lag time of the basin obtained on the base of the SCS method (USDA, 1986) (Table 1). A sketch is given in Fig. 3 for Candoglia on the Toce River. Normalization is needed to compare errors for the different catchments and different flood events.

Table 3. Observed, simulated and forecasted peak discharge (m³/s) for the ITALIA1 case study (with 14 November and 15 November runs of S1, S2 and ECMWF).

Cross section	Q max Observed	Q max Simulated	S1 14/11 12:00	S1 15/11 12:00	S2 14/11 12:00	S2 15/11 12:00	ECMWF 14/11 12:00	ECMWF 15/11 12:00
Castelnuovo	220.76	178	4.78	120.86	38.99	99.54	31.80	98.04
Carignano	262.90	259	61.08	174.58	69.96	178.32	122.93	285.50
Cassine	806.10	895	332.70	652.58	158.35	653.57	125.10	666.18
Serravalle	364.09	285	34.93	128.88	48.49	128.88	28.73	128.88
Palestro	2862.20	2980	2568.99	3021.52	1590.80	3731.11	929.38	1504.79
Farigliano	1371.57	1713	565.74	1535.42	688.94	1535.42	648.36	1658.22
Candoglia	1347.70	1870	971.41	1729.51	559.20	1547.34	447.85	981.86

Table 4. Observed, simulated and forecasted peak discharge (m³/s) for the ITALIA2 case study (with 24 November and 25 November runs of S1, S2 and ECMWF). Discharge measurements for Carignano are not available.

Cross section	Q max Observed	Q max Simulated	S1 24/11 00:00	S1 25/11 00:00	S2 24/11 00:00	S2 25/11 00:00	ECMWF 24/11 12:00	ECMWF 25/11 12:00
Castelnuovo	124.94	89	89.73	42.83	37.14	96.94	69.50	74.53
Cassine	748.40	720	1422.68	216.84	187.32	762.53	203.25	320.40
Serravalle	935.85	1536	80.97	583.85	196.22	583.85	360.14	670.41
Palestro	1112.20	1100	2707.31	664.47	909.91	1615.73	803.38	1250.02
Farigliano	514.21	557	432.34	312.43	229.05	479.26	246.96	436.34
Candoglia	550.30	679	512.49	184.05	597.73	486.07	397.41	633.04

5.2 QDF results

5.2.1 Using QPF

Tables 3 and 4 show the results obtained in terms of peak discharge forecasts for the locations involved in the two case studies for the available different meteorological runs; observed and simulated flood peaks are also presented.

Comparison in terms of peak discharge and flood volume relative errors is presented, respectively, in Tables 5 and 6, for the two case studies and for the four meteorological forcings (observed and forecasted with S1, S2 and ECMWF). Mean, standard deviation and the mean of the absolute values of errors (absolute mean), are shown. For the computation of the statistical indices, hydrographs produced by the different meteorological model initializations were considered as part of a unique set of data. Peak discharge and flood volume forecast errors are both generally negative thus showing an increased bias to underestimation increased with respect to the simulation cases, this is consistent with the clear underestimation of QPFs displayed in Table 7. Nevertheless, the mean QDF relative errors are only a little lower than the mean QPF relative errors while the standard deviation is higher. This shows that the integrating effect of the catchment converting rainfall to runoff has a minimal role in smoothing uncertainties.

The higher spatial resolution of the LAMI meteorological model do not seem to result in a significant reduction of forecast errors: in the ITALIA2 event, the mean rainfall errors for S1 and S2 are greater than those for ECMWF model (Table 7) as for the flood volume underestimation (Table 6). Probably the required spatial resolution capable of resolving the spatial detail of the complex topography of the area is higher than the ones used in this study.

Figure 4 presents the dependency of the errors on the advance time index for a deeper analysis focused on operational use. A slight trend of the errors to decrease as far as the forecast horizon can be detected and this is of course due to the decrease of QPF errors with the forecast horizon (Fig. 5). This means that a new initialization of the meteorological model is necessary to improve significantly the forecasted rainfall, as much as approaching the flood event.

For each of the analysed locations, QPF errors were calculated by comparing them with the mean basin cumulated rain on the contributing catchment; the advance index is the same as used for QDF in each section.

5.2.2 Using zero rain

A further experiment was carried out at the end of the analysis to simulate the case of unavailable QPF. The objective was to quantify the adjoint value of a complete
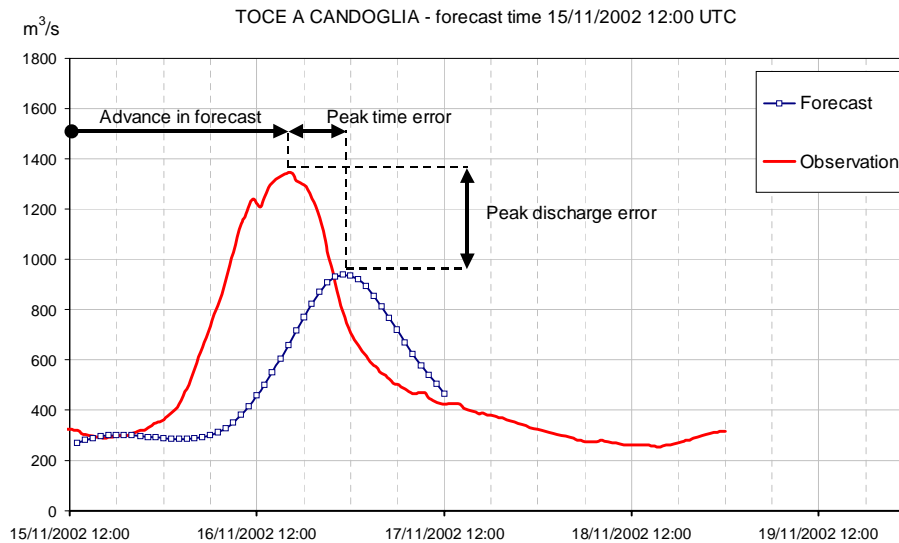
Fig. 3. Sketch comparing the forecasted and observed hydrographs.

Table 5. Statistics of peak discharge relative errors for the two case studies, ITALIA1 and ITALIA2, and for the four meteorological forcings, observed and forecasted with S1, S2 and ECMWF. Negative bias in the mean denotes model underestimation.

| Forcing | ITALIA1 | | | ITALIA2 | | |
	Mean	Absolute mean	Standard deviation	Mean	Absolute mean	Standard deviation
Simulation	0.05	0.17	0.20	0.1	0.21	0.28
S1	−0.38	0.45	0.37	−0.19	0.58	0.66
S2	−0.42	0.50	0.37	−0.26	0.36	0.36
ECMWF	−0.48	0.53	0.33	−0.33	0.38	0.26

hydrometeorological chain with respect to the simple hydrological forecast. In this situation, the forecasted rain was always taken equal to 0.

Hydrological forecasts were run every six hours. Figure 6 shows the results obtained for Palestro during the ITALIA1 event and Cassine during the ITALIA2 event. To provide a comparison, result of the simulation forced with the entire available rainfall observations (best available forecast) is also reported; their performance is presented in Fig. 7. The underestimation is of course high when the advance time is high but, as the advance time decreases, the performance is very similar to that obtained with the whole hydro-meteorological chain. This behaviour is dependent on the time between the onset of the rainstorm and the basin response, and hence is tight linked to the basin lag-time.

5.3 Errors in rainfall versus errors in discharge

A comparison between meteorological and hydrological performance is evaluated in this section.

Mean and standard deviations of the errors in mean basin rainfall classified by advance time index are summarised in

Fig. 8. We recall that negative errors denote underestimation of the model with respect to observations. As expected, the general trend can be explained by an increase of mean error with the growth in the advance time index. Mean errors vary from a value less than −20% (when the advance time index is in the range 1–2) to about −70% (when the advance time index is in the range 5–6). Standard deviation follows a similar trend, except when the advance time index is in the range 5–6; here standard deviation is lower.

Mean and standard deviations of the errors in peak discharge prediction classified by advance time index are summarised in Fig. 9. Results of simulation forced with gauge data (best available forecast) are reported for comparison. In opposition to the mean basin rainfall behaviour, as a general trend, mean forecast errors slightly decrease with the reduction of advance time index but increase when the advance time is comparable with the lag-time of the catchment (i.e. advance time index equal to 1). The standard deviation of errors does not decrease significantly with the advance time index. For advance time lower than 2 times the basin lag-time, QDF performance in the case of unavailable QPF is similar to that obtained with the other models. The performance

www.nat-hazards-earth-syst-sci.net/8/161/2008/

Nat. Hazards Earth Syst. Sci., 8, 161–173, 2008

Table 6. Statistics of flood volume relative errors for the two case studies, ITALIA1 and ITALIA2, and for the four meteorological forcings, observed and forecasted with S1, S2 and ECMWF. Negative bias in the mean denotes model underestimation.

Forcing	ITALIA1			ITALIA2		
	Mean	Absolute mean	Standard deviation	Mean	Absolute mean	Standard deviation
Observations	−0.01	0.20	0.25	0.06	0.32	0.40
S1	−0.37	0.43	0.38	−0.45	0.46	0.33
S2	−0.42	0.50	0.38	−0.42	0.47	0.33
ECMWF	−0.45	0.51	0.34	−0.29	0.45	0.46
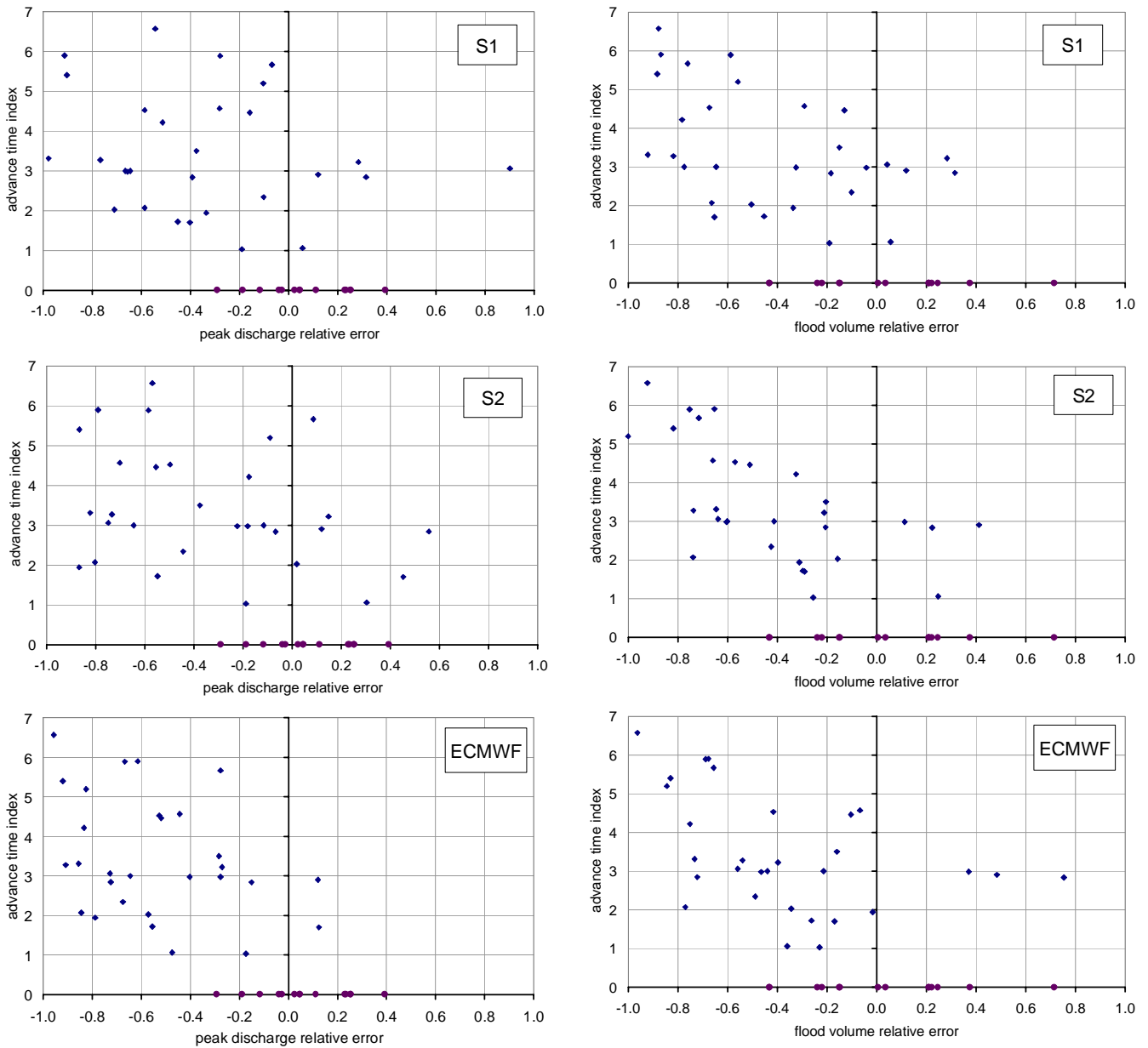
Fig. 4. Peak discharge and flood volume forecast errors versus advance time index for the three different forcing QPF. Simulation relative errors are displayed for advance time index equal to 0. Negative errors denote model underestimation.

Table 7. Statistics of mean basin cumulated rainfall relative errors for the two case studies.

| | ITALIA1 | | | ITALIA2 | | |
	Mean	Absolute mean	Standard deviation	Mean	Absolute mean	Standard deviation
S1	−0.43	0.44	0.32	−0.51	0.51	0.22
S2	−0.44	0.46	0.28	−0.45	0.45	0.30
ECMWF	−0.44	0.44	0.23	−0.31	0.35	0.28

of the hydrological model coupled to observed meteorological data, this can be considered the best available forecast, is quite good (mean relative error less than 10%). Comparing Figs. 8 and 9, it can be noted that mean QDF errors are lower than QPF's while the standard deviations are higher.

Probably, an explanation of this can be found in the fact that, along with the errors in cumulated rainfall volume, there is also the problem with errors in the timing of forecasted rainfall that particularly affects the results for the short forecast advance. Updating the hydrological forecast to reduce the forecast advance proves useful in reducing bias but seems to enhance the uncertainties in QPF and in the hydrological models.

5.4 Impacts on the warning system

The results in terms of warning system performance are evaluated and compared with those obtained in simulation context (best available forecast).

A perfect forecast of runoff peak, volume and timing would be the ideal scenario for a warning system. Fortunately, however, effective emergency procedures can be undertaken with less than a perfect forecast. Due to the operational-oriented application, it is important to understand the capability of the model to catch hazardous events. To this end, despite the possible errors in the simulated peak discharge, it is more important for the simulation to provide a good estimate of hazard linked to the flood.

Using the simple classification of flood hazard based on peak discharge thresholds and adopted by the Civil Protection (Table 1) and following the meteorological practice (Mason and Graham, 1999), it is possible to build up a contingency table comparing the observed and the simulated or forecasted hazard. The contingency tables indicate the quality of a forecast system by considering its ability to correctly anticipate the occurrence or non-occurrence of predefined events that are expressed in binary terms. In this work, the contingency table is a matrix composed of 3 rows by 3 columns (i.e. 9 possible outcomes), as there are 3 possible warning codes (Fig. 10).

To account for the fact that errors are not equally important in flood management, the following definitions have been adopted:
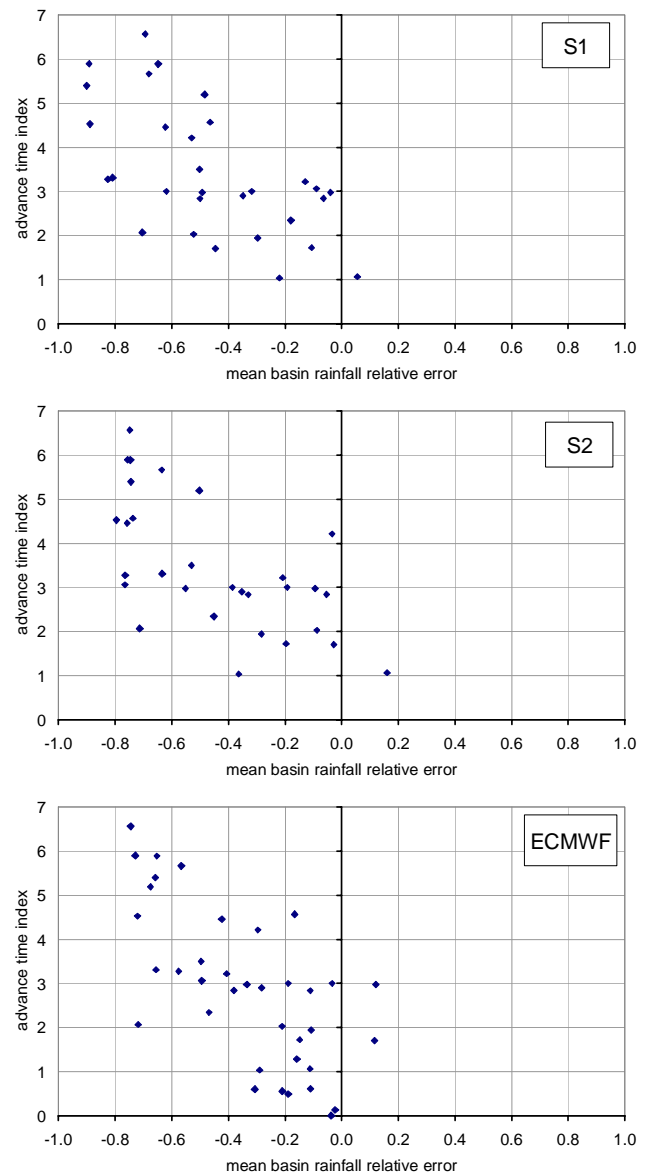

Fig. 5. QPF errors versus advance time index for the three meteorological models. Negative errors denote model underestimation.

– Heavy missed alarms (when code 3 occurred and code 1 is simulated or forecasted). These are very important mistakes.
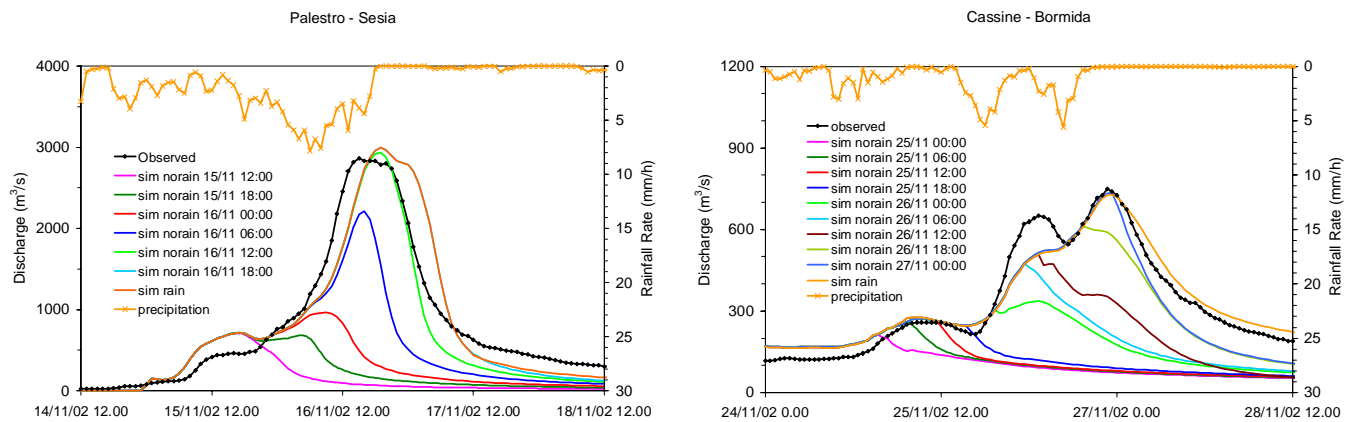
Fig. 6. Results of the simulations conducted assuming a lack of QPFs for the Palestro location during ITALIA1 event (left) and the Cassine location during the ITALIA2 event (right). Result of the simulation forced with the entire available rainfall observations (best available forecast) is reported for comparison.
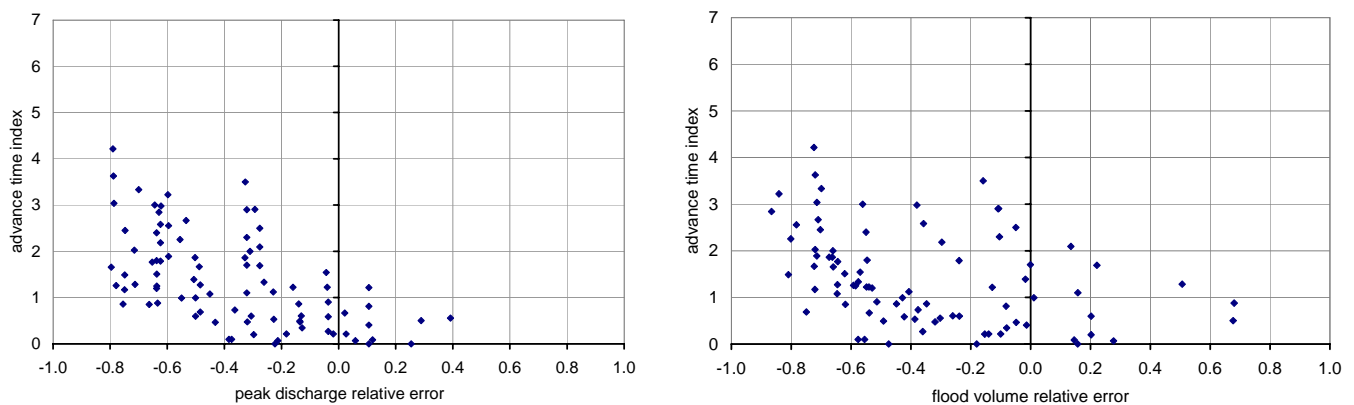

Fig. 7. Peak discharge and flood volume forecast errors versus advance time index for the case of unavailable QPF. Negative errors denote model underestimation.

– Missed alarms (when either code 3 occurred and code 2 is simulated or forecasted, or code 2 occurred and code 1 is simulated or forecasted). These are important mistakes.

– False alarms (when either code 2 occurred and code 3 is simulated or forecasted, or code 1 occurred and code 2 is simulated or forecasted). These are minor mistakes.

– Heavy false alarms (when code 1 occurred and code 3 is simulated or forecasted). These are important mistakes.

– A hit (when there is a perfect agreement between the code that occurred and the code of the the forecasted or simulated peak).

The correct performance index: CPI (defined as the ratio between the number of hits and the total number of outcomes) is chosen as a measure of the overall performance of the system (Jolliffe and Stephenson, 2003; Ravazzani et al., 2007).

Simulations driven with observed forcings show that the performance of the hydrologic model is not affected by errors in precipitation forecast. Only 2 false alarms (minor mistakes) from a total of 13 outcomes were detected (CPI=84,6%). Note that when coupled with a meteorological model, there are 26 outcomes due to the fact that two data sets of forecasted rainfall are available for every event, one for each run of the meteorological model. When coupled with the S1 data, the warning system detected 2 false alarms, 1 heavy false alarm and 6 missed alarms (CPI=65,4%). Performance was about the same when coupled with the S2 data, 3 false alarms and 7 missed alarms (CPI=61,5%). When coupled to the ECMWF data, the warning system detected 1 false alarm and 9 missed alarms (CPI=61,5%).
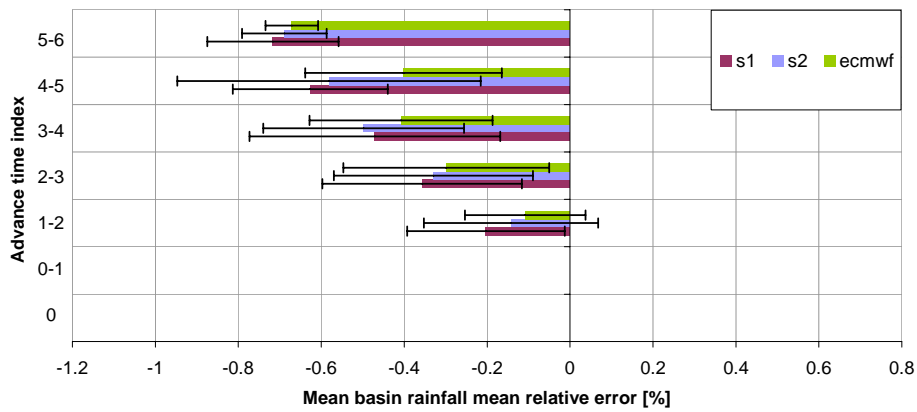
Fig. 8. Histogram comparing the mean basin rainfall mean (and standard deviation with bars) relative errors for the different meteorological models.
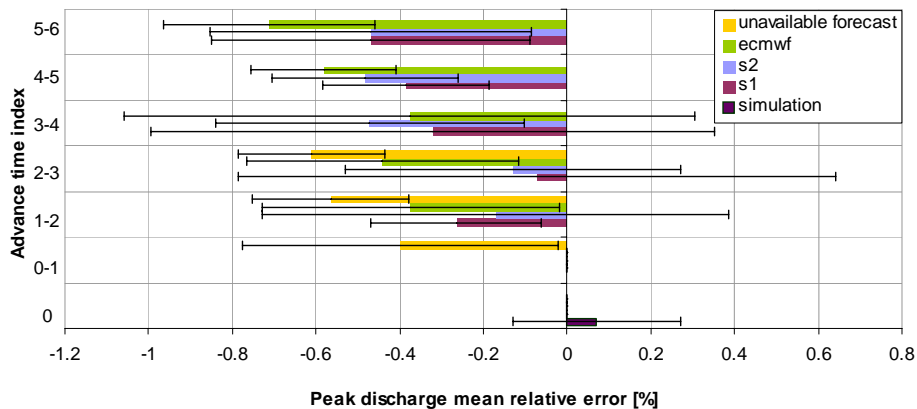

Fig. 9. Histogram comparing the peak discharge mean (and standard deviation with bars) relative errors for the different meteorological models.

6 Conclusions

This paper evaluated a forecasting operational system based on a distributed hydrological model. The FEST-WB model was analysed in operational setting experiments, coupling the hydrological model to the output of two different meteorological models: two parameter sets (S1 and S2) drove the local area model LAMI and the general circulation model ECMWF.

When forced by meteorological ground observed data (best available forecast experiment), the FEST-WB performed quite well in simulating discharge hydrographs, showing a slight mean overestimation in peak discharge (the mean relative error is less than 10%).

When coupled to the meteorological models, there was a heavy degradation in the performance of the hydrological model. Mean errors in QDF slightly decrease with the reduction of advance time index but increase when the advance time is comparable with the lag-time of the catchment (i.e. advance time index equal to 1). The standard deviation

of QDF errors, however, does not decrease significantly with the reduction of forecast advance. In the comparison of meteorological and hydrological performance, it is shown that QDF errors are lower than QPF's, while standard deviations are higher.

The aforementioned behaviours may be related to the fact that, beyond the errors in cumulated rainfall volume, the timing of forecasted rainfall provides an additional source of error and this particularly affects the results for short forecast advance. Updating the hydrological forecast, in order to reduce the forecast advance, proves to be useful at reducing bias but can result in enhancing the uncertainties in QPF and in the hydrological models. This behaviour highlights the problem that the chains of hydro-meteorological models are affected by the large variance of the QDF errors so that the results are not reliable for decision makers to take the most effective decisions under uncertainty (Todini, 1999).

For alpine watersheds with complex topography, the higher spatial resolution of the LAMI meteorological model did not seem to result in a significant reduction of forecast

SIM

Observed code \ Simulated code	1	2	3
1	8	0	0
2	0	3	2
3	0	0	0

S1

Observed code \ Forecasted code	1	2	3
1	14	1	1
2	6	3	1
3	0	0	0

S2

Observed code \ Forecasted code	1	2	3
1	15	1	0
2	7	1	2
3	0	0	0

ECMWF

Observed code \ Forecasted code	1	2	3
1	16	0	0
2	9	0	1
3	0	0	0

Fig. 10. Warning system performance in terms of contingency table with respect to simulation with observed data (SIM) and forecasted with S1, S2 and ECMWF meteorological model. Green cells denote perfect agreement, yellow cells minor mistakes, orange cells important mistakes and red cells very important mistakes.

errors: a meteorological model with higher spatial resolution is probably needed to reduce forecast errors. Experiments using zero rain shows that, for advance time lower than two times the basin lag-time, QDF performance is similar to that obtained with the other meteorological models.

Errors in discharge simulation are, generally, not directly related to the reliability of the warning system, which is linked to the capability of the hydro-meteorological chain to estimate the flood hazard. The analysis in this paper, based on contingency tables, shows that, despite the high errors detected in peak discharge analysis, the warning system can reach a reliability of 65.4% (measured with the correct performance index) in conjunction with LAMI-S1 model and 61.5% in conjunction with LAMI-S2 or ECMWF model. However, degradation from the best available forecast experiment (reliability = 84.6%) is remarkable.

It is important to underline that these conclusions refer to the analysis of the selected case studies, so they cannot be directly taken as universally true. Of course, as pointed out, an assessment of the general performance of the warning system needs an extensive long period to validate its operational forecasts.

Finally, it needs to be stated that the problems associated with the on-line evaluation of forecast uncertainty, which is one of the most relevant aspects of real time flood forecasting, is still not considered in this work. This represents a strict requirement for future developmental work.

References

Brakenridge, G. R., Anderson, E., and Caquard, S.: Global Register of Major Flood Events. Dartmouth Flood Observatory, Hanover, USA, digital media, http://www.dartmouth.edu/~floods/Archives/2005sum.htm, 2005.

Brath, A., Burlando, P., and Rosso, R.: Sensitivity analysis of real time flood forecasting to on-line rainfall predictions, in: Natural Disasters in European Mediterranean Countries, edited by: Siccardi, F. and Bras, R. L., 469–488, 1998.

Carpenter, T. M., Sperfllage, J. A., Georgakakos, K. P., Sweeney, T., and Fread, D. L.: National threshold runoff estimation utilizing GIS in support of operational flash flood warning systems, J. Hydrol., 224, 21–44, 1999.

CEC: CORINE land cover. Technical guide, Commission of the European Communities, Luxembourg, 1994.

Corbari, C., Martinelli, J., Ravazzani, G., and Mancini, M.: Snow satellite images for calibration of snow dynamic in a continuous distributed hydrological model, Hydrol. Earth Syst. Sci. Discuss., 4, 3979–4004, 2007,
http://www.hydrol-earth-syst-sci-discuss.net/4/3979/2007/.

EEA: CORINE land cover technical guide – Addendum 2000, European Environment Agency, 2000.

Georgakakos, K. P. and Hudlow, M. D.: Quantitative precipitation forecast techniques for use in hydrologic forecasting, B. Am. Meteor. Soc., 65(11), 1186–1200, 1984.

Jasper, K., Gurtz, J., and Lang, H.: Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model, J. Hydrol., 267, 40–52, 2002.

Jolliffe, I. T. and Stephenson, D. B.: Forecast Verification. A Practitioner's Guide in Atmospheric Science, Wiley and Sons Ltd, 240 pp, 2003.

Lardet, P. and Obled, C.: Real-time flood forecasting using a stochastic rainfall generator, J. Hydrol., 162, 391–408, 1994.

Ludwig, R., Taschner, S., and Mauser, W.: Modelling floods in the Ammer catchment: limitations and challenges with a coupled meteo-hydrological model approach, Hydrol. Earth Syst. Sci., 7, 833–847, 2003,
http://www.hydrol-earth-syst-sci.net/7/833/2003/.

Mancini, M.: La modellazione distribuita della risposta idrologica: effetti della variabilità spaziale e della scala di rappresentazione del fenomeno dell'assorbimento, PhD thesis, Politecnico di Milano, Milan, Italy, 1990 (in Italian).

Martinec, J.: The degree-day factor for snowmelt runoff forecasting, proceedings of general assembly of Helsinki commission on surface waters, IAHS Publ. 51, 1960.

Mason, S. J. and Graham, N. E.: Conditional Probabilities, Relative Operating Characteristics, and Relative Operating Levels, Weather Forecast., 14(5), 713–725, 1999.

Montaldo, N., Ravazzani, G., and Mancini, M.: On the prediction of the Toce alpine basin floods with distributed hydrologic models, Hydrol. Processes, 21, 608–621, 2007.

Montaldo, N., Toninelli, V., Albertson J. D., Mancini, M., and Troch P. A.: The Effect of Background Hydrometeorological Conditions on the Sensitivity of Evapotranspiration to Model Parameters: Analysis with Measurements from an Italian Alpine Catchment, Hydrol. Earth Syst. Sci., 7, 848–861, 2003,
http://www.hydrol-earth-syst-sci.net/7/848/2003/.

Montaldo, N., Toninelli, V., Mancini, M., and Rosso, R.: Coupling Limited Area Models with Distributed Hydrologic Models for Flood Forecasting: the Toce Basin Case Study, IAHS publ. no. 274, 229–236, 2002.

Pilgrim, D. H. and Cordery, I.: Flood runoff, in: Handbook of hydrology, edited by: Maidment, D. R., McGraw-Hill, Chap. 9, 1993.

Priestley, C. H. B. and Taylor, R. G.: On the assessment of surface heat flux and evaporation using large scale parameters. Mon. Weather Rev., 100, 81–92, 1972.

Rabuffetti, D. and Barbero, S.: Operational hydro-meteorological warning and real-time flood forecasting: the Piemonte Region case study, Hydrol. Earth Syst. Sci., 9, 457–466, 2005,
http://www.hydrol-earth-syst-sci.net/9/457/2005/.

Rabuffetti, D. and Milelli, M.: The hydro-meteorological chain in Piemonte region, NorthWestern Italy – analysis of the HYDROPTIMET test cases, Nat. Hazards Earth Syst. Sci., 5, 845–852, 2005,
http://www.nat-hazards-earth-syst-sci.net/5/845/2005/.

Rabuffetti, D., Salandin, A., and Cremonini, R.: Hydrological modelling of snow cover in the large upper Po river basin: winter 2004 results and validation with snow cover estimation from satellite, Geo-Environment and Landscape Evolution II, ISBN: 1-84564-168-X, 293–302, 2006.

Ravazzani, G., Mancini, M., Giudici, I., and Amadio, P.: Effects of soil moisture parameterization on a real-time flood forecasting system based on rainfall thresholds, in: Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management, edited by: Boegh, E., Kunstmann, H., Wagener, T., Hall, A., Bastidas, L., Franks, S., Gupta, H., Rosbjerg, D., and Schaake, J., IAHS Publ. 313, 407–416, 2007.

Salandin, A., Rabuffetti, B., Barbero, S., Cordola, M., Volontè, G., and Mancini, M.: Il lago effimero sul ghiacciaio del Belvedere: monitoraggio e simulazione numerica del fenomeno finalizzata alla previsione e gestione dell'emergenza, Neve e Valanghe, 51, 58–65, 2004.

Soil Conservation Service: Urban hydrology for small watershed, Tech. Rel. no. 55. U.S. Department of Agriculture, Washington D.C., 1986.

Tarboton, D. G., Chowdhury, T. G., and Jackson, T. H.: A Spatially Distributed Energy Balance Snowmelt Model, Utah Water Research Laboratory, 1994.

Todini, E.: Using phase-space modelling for inferring forecasting uncertainty in non-linear stochastic decision schemes, J. Hydroinform., 01.2, 75–82, 1999.

USDA – U.S. Department of Agriculture: Soil Conservation Service, National Engineering Handbook, section 4, Hydrology, Rev. ed., U.S. Department of Agriculture, Washington D.C., USA, 1986.