**Natural Hazards
and Earth
System Sciences**

Dealing with uncertainty: an analysis of the severe weather events over Italy in 2006

L. Molini[1], A. Parodi[1], and F. Siccardi[1,2]

[1]CIMA Research Foundation, Savona, Italy
[2]Department of Communication Computer and Systems, University of Genova, Genova, Italy

Abstract. Forecast verification is a long-standing issue of the whole meteorologists' community. A common definition of a truly satisfying prediction skill has not been achieved so far. Even the definition of "event", due to its spatio-temporal discontinuity, is highly affected by uncertainty.

Moreover, decision-making demands numerical weather prediction modellers to provide information about the "inner" uncertainty, i.e. the degree of uncertainty related to the choice of a specific setting of the model (microphysics, turbulence scheme, convective closure, etc.).

Most European Mediterranean countries, due to dense development, steep coastal orography and short hydrological response time of the drainage basins, have to deal very frequently with flash floods and sudden shallow land sliding impacting on urban areas. Civil protection organizations are in place to issue early warnings in order to allow local authorities and population to take precautionary measures. To do so in Mediterranean catchments, hydrologists are required to use numerical rainfall predictions in place of rainfall observations on large European catchments.

Estimating the measure of uncertainty is for this reason crucial.

The goal of this work is to propose an objective evaluation of the performance of the currently operational weather prediction model COSMO-I7 over quite a long time period and to check forecast verification at different space-time scales by the comparison of predictions with observations.

Due to large investments in the last years, in fact, Italy has built up one of the most dense hourly-reporting network of rain gauges. The network has a mean space density of about $1/100\,km^2$, very similar to the horizontal resolution of currently operating limited area models.

An objective procedure to identify and compare the extreme events of precipitation has been applied to the full set of rainfall observations and over the severe events forecast by COSMO-I7 and announced in official warnings by Italian Civil Protection Department.

The procedure allows to classify rainfall events as long-lived and spatially distributed or as having a shorter duration and a minor spatial extent. We show that long-lived events are less affected by overall uncertainty than short-lived ones, yet the inner uncertainty of the event affects both.

1 Introduction

European Mediterranean countries are densely developed. The coastal orography is steep and the hydrological response time quite short. Extreme rainfall events frequently produce flash floods (Ferraris et al., 2002) and sudden shallow land sliding in urban areas (Wieczorek and Guzzetti, 2000; Siccardi et al., 2002). The interval between rainfall and the impact on developed areas is in the order of hours, shorter than the time needed for the population to take precautionary measures (UNISDR, 2008). Hydrologists are required to make use of rainfall predictions in order to timely estimate the effects on the ground.

Italy has a body of official rules and technical tools for the operational prediction of impending risk scenarios. Predictions and observations are regularly archived in order to judge the efficiency of the system and to possibly improve it. Predictions are made at the central and/or regional level and inform about the severity of the incoming event (Italy Official Gazette, 2004).

The definition of the level of criticality of an event, on which possible warnings are based, is summarized in Fig. 1. The predictive chain is first devoted to defining the criticality of the physical scenario stemming from a numerically predicted severe weather condition. An operational probabilistic
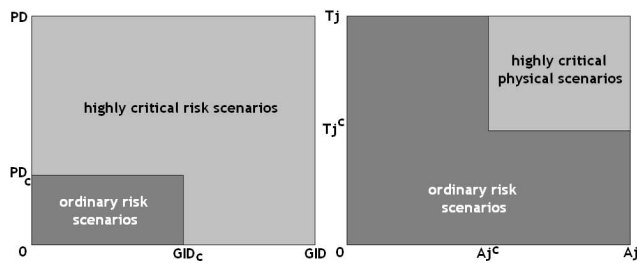
Fig. 1. A synthetic depiction of the different levels of criticality used by the Italian Civil Protection Department that defines physical and risk scenarios.

chain is being implemented to generate, over all the hydrometric sections of interest, a sufficiently large ensemble of outputs (Siccardi et al., 2005). The predictive chain includes the available ensemble of the numerical weather predictions and the rainfall/runoff forecasts on rivers crossing the target area; just recently the real-time observed degree of saturation of the slopes has been included (Caparrini, 2007).

The outcomes of the physical predicting chain are couples of T_j, A_j, where T_j is a generalized measure of the predicted water depth averaged on a target region and A_j is the generalized measure of the extent of the corresponding inundation. Thus, the description in probability of different ensembles of couples A_j, T_j is enabled, defining events that appropriately partition the sample space of physical scenarios (Fig. 1, right panel: highly critical physical scenarios $A_j > A_j^c \cap T_j > T_j^c$).

When the area of highly critical physical scenarios in the right panel of Fig. 1 is populated, the procedure continues assessing the risk scenarios (Fig. 1, left panel: highly critical risk scenarios $PD > PD_c \cup GID > GID_c$, where PD is a generalized measure of risk for persons in the target region and GID is a generalized measure of risk of damage to goods and infrastructure in the same target region). The estimation of the risk is based not only on the level of exposure and vulnerability of the developed flood-prone and landslides-prone areas, but also on the archive of the past hundred years of disastrous events that have affected the country (AVI-CNR, 2004). Then, warnings are issued when the highly critical risk scenarios area in the left panel of Fig. 1 is populated.

Many attempts are going on in most European Mediterranean countries to build up a sound operational hydrometeorological forecasting chain. The European Flood Alert System (EFAS) provides early warnings, especially focused on cross-border river basins (van der Knijff et al., 2008; Ramos et al., 2007). On both scales, large European rivers and small/medium size Mediterranean catchments, the major concern is how large is the uncertainty on the simulation of each single physical process, and how to propagate it through the process cascade.

Assessing objectively forecast quality and reliability of quantitative precipitation forecast is a mandatory task not only for operational services but mainly for civil protection

organizations. It appears also that the predictive skill of numerical models depends on specific meteorological patterns (Casati et al., 2008).

The aim of this project is to assess to which extent COSMO-I7, the Italian operational limited area model, is able to provide a correct modelling of the main kinematic and thermodynamic ingredients typical for intense precipitation processes in complex orography regions (Elementi et al., 2005; Milelli et al., 2008) by characterizing the uncertainty affecting the first step of the predictive procedure in order to issue warnings.

The present paper reports the first findings for some severe hydrometeorological events that occurred over the Italian territory and is structured as follows: Sect. 2 gives an overview of the tools and the datasets used to perform the analysis, together with a diagram summarizing the performances of the national predictions and warnings for the year 2006. Section 3 describes the analytical methodology used to assess forecast verification benchmarking on predictions of severe events. Predictions are intended as forecasts made by Italian Civil Protection Department, which procedures make use of either objective and subjective tools (see Footnote 3). A representative study case is described in detail. Contrary, Sect. 4 describes the analytical methodology used to assess forecast verification benchmarking on ex-post observations of severe events. Section 5 discusses the results.

2 Observational and predictive tools

This study is primarily based on the comparison of two different kinds of data: rainfall observations and rainfall prediction by a limited area model. The first one is the hourly-accumulated rainfall observed by the national rain-gauge network. Due to large investments funded by the Italian Government in the last decades, Italy has one of the world's most dense raingauge network. Figure 2 shows the reporting network made up of more than 1200 WMO-standard operating stations. The network's spatial density varies between $1/50\,km^2$ in the area of north-central Italy and $1/200\,km^2$ in Po Valley, Puglia and Sicily, with a mean coverage of about $1/100\,km^2$, which is close to the horizontal resolution of COSMO-I7 ($\approx 1/50\,km^2$). Each station reports by radio link the accumulated rainfall depth, or liquid equivalent of snow, over short intervals. For the purposes of this project, the hourly archived depths are used; continuous observations are provided by more than 90% of the rain-gauges. The network is designed so that it fills the archive in deferred time, when for some reasons the radio link failed but the measure does exists.

The rainfall predictions are provided by COSMO-I7. COSMO-I7 is the Italian version of the COSMO-Model (Doms and Schättler, 1998), a non-hydrostatic and fully compressible numerical weather prediction model developed at DWD (Deutscher Wetterdienst – German National
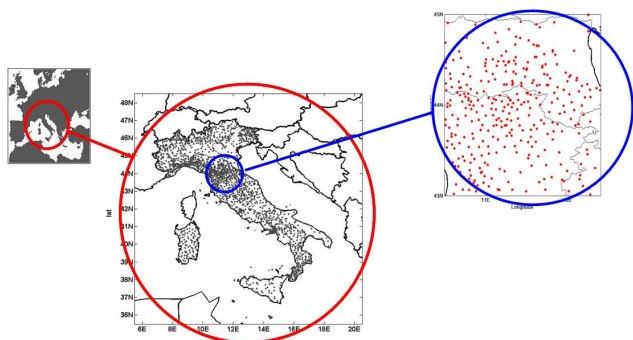
Fig. 2. The Italian real-time reporting rain-gauge network showing its density all over the national territory, and in particular in central Italy, Tuscany, where the network reaches its higher density.

Fig. 3. The Venn diagram summarizes the findings for year 2006, both benchmarking on observations (right panel) and predictions (left panel).

Weather Service). COSMO-I7 is operated by the Hydro-meteorological Service of the Emilia-Romagna Region and distributed by the Italian Air Force Meteorological Service. The COSMO-Model offers a wide range of turbulence, surface and microphysical schemes. For a more comprehensive description of the model, the reader is referred to Steppeler et al. (2003).

The primitive hydro-thermodynamic equations are used for describing compressible non-hydrostatic flow in a moist atmosphere, without any scale approximation. The model uses hybrid terrain-following coordinates, while the vertical resolution may be varied from 50 m near the surface up to several hundred meters in higher altitudes.

COSMO-I7 runs at its operational horizontal resolution of 7 km using 50 vertical levels. It uses the Mellor-Yamada PBL parameterization, which is a 2.5-order local closure scheme, the Tiedtke parameterization for convection and a bulk microphysics parameterization that includes water vapour, cloud water, rain and snow.

As it is reported below, sometimes incoherence emerges between rainfall pattern observations and predictions: in such cases the Advanced Microwave Scanning Radiometer (AMSR-E)[1] water vapour columnar content data are referred to in order to get possible explanations. To familiarize the reader with the goal of this work, a summary of the performance of the Italian civil protection system is given in Fig. 3 and Table 1.

Going deeper into the morphology of the events, it turns out that the success differs according to the severity of the event[2], which is the reason why a thorough comparison of

predictions and observations was undertaken. The present paper reports the first findings for the year 2006.

3 A forecast-driven analysis

The first part of the study concerns the events predicted by Italian Civil Protection Department. The characterization of a so-called "predicted severe event" is contained within the official alert bulletins[3]. The bulletins are issued before the beginning of each event expected to be potentially severe and are stored[4].

Each warning informs about the spatial extent of a predicted extreme event by listing the areas about to be impacted, its starting time and duration.

In this project, for each forecast event, the 00:00 UTC COSMO-I7 run nearest to the event's beginning has been taken into account to compute the estimated amount of rainfall over the predicted duration.

3.1 Forecast verification criteria: morphology, spectral analysis, thermodynamic and kinematic features, basin-scale verification

The corresponding observed rain depths measured by the gauges network in the same region are the terms of comparison for the forecast-driven analysis.

For each forecast, the observed event was chosen in order to have the duration prescribed by the bulletin and minimizing the differences to the prediction in terms of total rainfall

---

[1] AMSR-E is a microwave radiometer launched aboard NASA's Aqua spacecraft in collaboration with the Japan Aerospace Exploration Agency (JAXA). It measures over the oceans certain geophysical parameters, including sea surface temperature (SST), wind speed, atmospheric water vapour, cloud water, and rainfall rate with a resolution of 25 km and revisiting time of about 12 h.

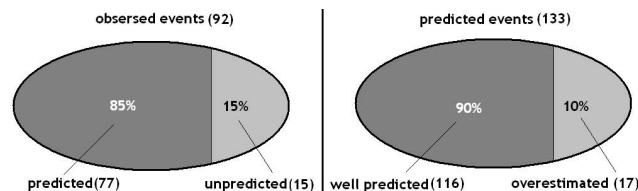[2] For a comprehensive definition of event, see Sect. 3.1

[3] Avviso di Avverse Condizioni Meteorologiche, i.e. Severe Weather Warning. The decision-maker usually chooses to issue a warning on the bases of a consolidated procedure, which involves different features that can be considered either "objective" and "subjective". In the first category, data from both deterministic and probabilistic models or satellite and ground-based observations are recounted. The subjective aspect of prediction encompasses, for instance, the personal degree of familiarity of the forecasting officer with the territory on which the severe precipitation will impact, the knowledge of the particular characteristics of local climatology and the awareness of both antecedent moisture conditions and soils saturation.

[4] CIMA Foundation's archive.

Table 1. **(a)** Contingency score for 2006 ordinary risk scenarios, **(b)** contingency score for 2006 high risk scenarios.

	Operationally predicted events	Unpredicted events	Total observed events
(a) Ordinary Risk Scenarios			
Observed events	77	15	92
No event observed	22		
Total opearationally predicted events	99		
(b) High Risk Scenarios			
Observed events	17	10	27
No event observed	17		
Total opearationally predicted events	34		

amounts. To avoid a too strict constraint for both prediction and observation to be contemporary, the observed event has been searched in a timeslot starting twelve hours before the predicted start and ending twelve hours following the predicted end.

The forecast verification was undertaken through the application of different techniques: morphology, spectral analysis, thermodynamic and kinematic features and basin-scale verification.

Though qualitative, a first rating of the reliability of the prediction is given by a purely visual agreement between the 3-D maps of observed and forecast total rainfall amounts. The comparison of the total rainfall depths can be a useful approach to highlight the general structure of the precipitation fields, and in some cases can provide immediate information about the possible differences in terms of spatial shift and severity.

The spectral analysis was based on the application of the Fourier transform to the space distribution of the total rainfall amounts of each event, both predicted and observed. By transforming the signal to the frequency domain representation, a twofold aim was pursued. The first one was to compare the amplitude of both signals, so as to verify the forecast skill in reproducing whether correctly or not the severity of the observed event. The second goal was to identify the leading scales of the process, i.e., to identify a spatial scale L beyond which the contribution to signal total variance is negligible. This means that if the power spectrum is drawn in a log-log plot, to single out where the downward gradient is maximum, and depending on the extension of the leading scale, each event can be classified as mainly distributed or localized.

The kinematical and thermodynamic study has been performed by comparing on one side the water vapour and wind fields at 500 hPa, 750 hPa, and 850 hPa predicted by

COSMO-I7 and on the other, the water vapour columnar content observed by AMSR-E. This part of the work was devoted, in fact, to gaining a deeper insight in the main triggering factors leading to severe weather, whether localized or not, paying particular attention to the motion of humid air masses over the central Mediterranean Sea in terms of origin, direction, velocity and interaction with coastal and non-coastal mountainous systems. By doing this, an overall coherence of prediction and observation at the global scale of event is analyzed.

However, the area impacted by the event is drained by many catchments of different size: for this reason, ground effects have been investigated with scatterplots of the total amounts of rain, both predicted and observed, averaged over the surface of the basins having areas greater or equal to 200 km^2.

This approach allows to consider the skill of the forecast from a hydrological viewpoint at a basin-scale level for all the catchments hit, totally or partially, by the precipitation. Only raingauges are well-distributed on the ground whilst only a few hydrometers equipped with flow rating curve are currently operating.

3.2 Study case: the 14-15-16 September 2006 event over north-western Italy

Severe weather conditions were reported by the bulletin issued by Italian Civil Protection Department on 13 September, at 15:15 LT, which warned about a low pressure system lasting 36 h centred over the north-central part of the Mediterranean Sea. The event was forecast to produce heavy rainfall over north-western Italy since the early morning of 14 September. The critical time-window was successively extended by another 24 h in the next bulletin for a total duration of a 48 to 60 h. As stated in the previous section, the predicted rainfall depths were provided by the operational COSMO-I7 run initialized on 14 September at 00:00 UTC with a predicted maximum slightly exceeding 250 mm/60 h. The observational dataset was measured by more than 200 rain-gauge stations. Figure 4 shows the good agreement between the spatial distribution of predicted and observed rainfall that affected mainly the western Alps region.

The similarity between the predicted and observed patterns is quite apparent, especially in the north-western alpine sectors of the map, even if a certain degree of overestimation is present. A qualitative comparison between numerically-modelled and satellite-observed humidity fields confirms the good skill of the prediction. Figure 5 displays the relative humidity and the wind direction and intensity forecast by COSMO-I7 at three different heights for 14 September, 15:00 UTC; these are typical of the atmospheric circulation predicted by the model for those two-and-half days, when changes were not substantial. The main features are an almost constant high degree of saturation (around 100%) for
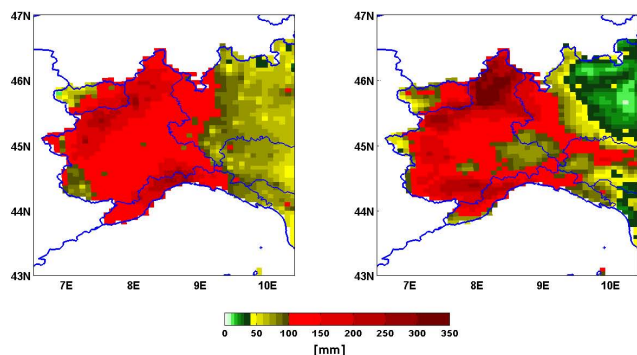
Fig. 4. The rainfall depth of the 60-h lasting September event in northern Italy. The left panel shows the 60-h accumulated rainfall predicted by COSMO-I7, the right one displays the accumulated rainfall observed by the raingauge network.

the whole atmospheric column and the direction of the wind, which was in the direction of the maximum slope of the main coastal orographic systems (the Apennines). This configuration is usually proven (Buzzi and Foschini, 2000; Rotunno and Ferretti, 2001) to produce large amounts of rainfall triggered by the fast orographic lifting of humid air, as was indeed observed in this case. In fact, AMSR-E detected a columnar water content prior to the event, shown in Fig. 6, that is consistent with the high quantity of atmospheric water required to supply heavy rains. Dynamical and thermodynamical factors suitable for triggering orographic lifting of humid air are strongly present in both prediction and observations.

The columnar content of water vapour retrieved by AMSR-E is compared with the one provided by COSMO-I7 (Fig. 7). To that end, AMSR-E data have been interpolated over the computational grid of the model over sea within the 200×200 km box containing the area from which the humid air has been advected (yellow-filled area, Fig. 7, upper right corner). Results show a fairly constant trend for the observational data, with a slow decrease beginning at the second part of the event (whose duration is represented in green). The trend of the model data is quite different, but anyway the water content predicted by COSMO-I7 was enough to trigger heavy rainfall.

The spectral analysis reported in Fig. 8 indicates an almost complete overlapping of the predicted and observed spectra. The smoothness of the rainfall fields revealed by the morphologic inspection of total precipitation amounts (see Fig. 4) is confirmed by the constant slope of the power spectra. This implies that no leading scale is found for the process, since its variance decreases proportionally as the spatial scale becomes finer. Moreover, the amplitudes of both spectra are very similar, demonstrating the event to have occurred with the predicted severity. Consequently, the issuing of the alert bulletins was justified. The spatial shift between the observed and forecast precipitation centre of mass, barely reaching a
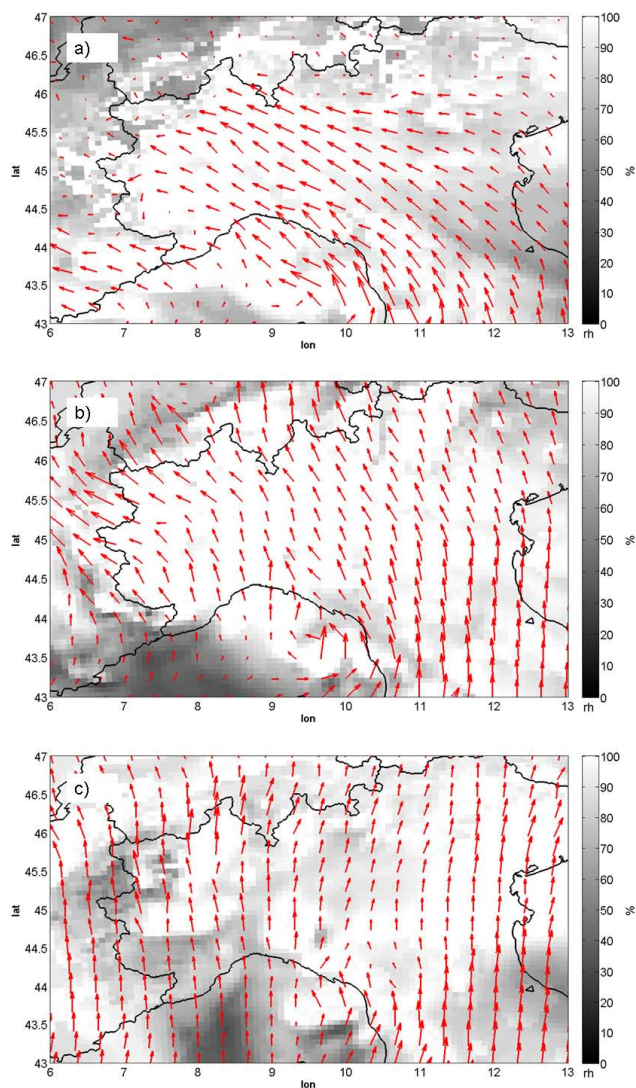

Fig. 5. The wind and the relative humidity predicted by 14 September 00:00 UTC COSMO-I7 run at the height of 850, 700, and 500 hPa (respectively, panel **a**, **b**, and **c**).

10-km extent (very close to the maximum model horizontal resolution), can be considered negligible for this case (not shown).

Figure 9 compares observed and predicted rainfall. The region impacted by the event contains approximately 50 watersheds having an area major or equal to 200 km^2. Generally speaking, rainfall was overestimated. For each basin, the diagram displays the observed total amount of rainfall averaged on basin's area (on the x-axis) and the corresponding prediction (on y-axis). The scatterpoints in the top part of the figure. represent the catchments, not necessarily adjacent to each other, reaching 150/200 mm in prediction and 100/150 mm in observation. From a civil protection point perspective, such overestimation is not a drawback, so the prediction can be assumed true. For the basins (not
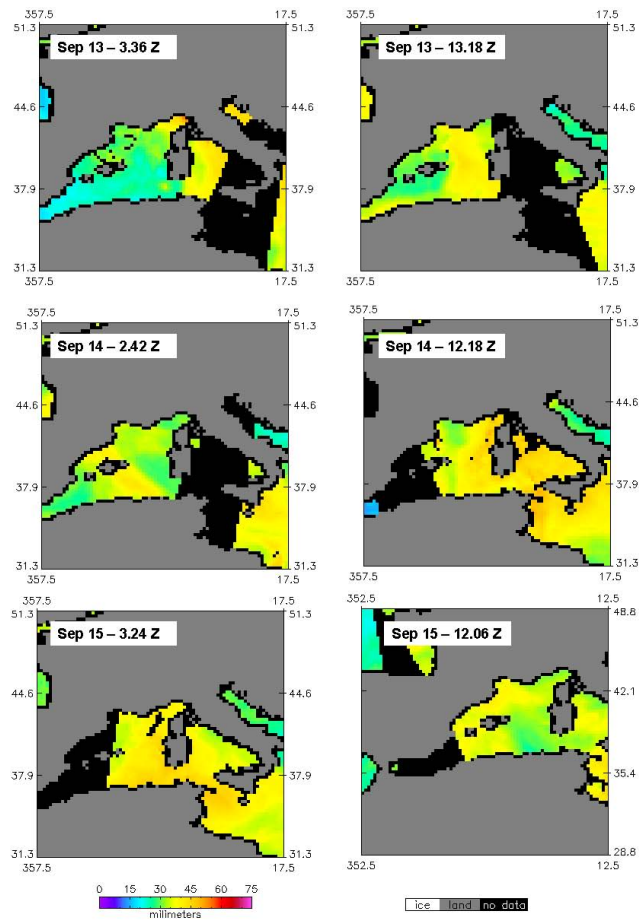
Fig. 6. The columnar content of water retrieved by AMSR-E in 6 different passes dating up to 1 day before the event (from http://www.remss.com/).

Fig. 7. Time trend for the columnar content of water vapour: observations (in blue) and prediction (red).

Fig. 8. Fourier spectral decomposition of the predicted and observed total rain depths. The spectral behaviour of the fields (the amplitude, computed in mm^2 versus the wavelength in km) are pretty akin denoting no leading scale.

contiguous) on the left side, total rainfall was forecast between 100 and 150 mm, whereas observation is largely under 50 mm; hence, those cases might be considered false alarms. Actually, this cluster of watersheds spreads randomly across the area hit by heavy precipitation and these basins border on other basins for which the alert was true. The last small group of scatterpoints, under the bisector, shows low amounts of rainfall, both predicted and observed; underestimation is quite remarkable in only one case. The random spread of basins for which the alert is true or false is a feature not known a priori, since it depends on the small-scale distribution of rainfall.

For this reason, the decision to extend the alert to the whole area allowed to encompass even the small fraction of rivers for which predicted rainfall was less than the actually observed; this means that even the population, living along or near under-predicted rivers, were warned. Thus the forecast skill can be considered satisfying from a civil protection viewpoint. Contrary, from a modelling point of view, such a misprediction from catchment to catchment is a measure of the inherent uncertainty of the predictive chain.
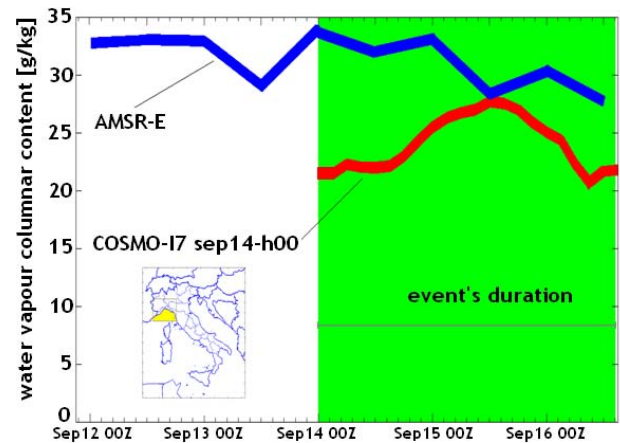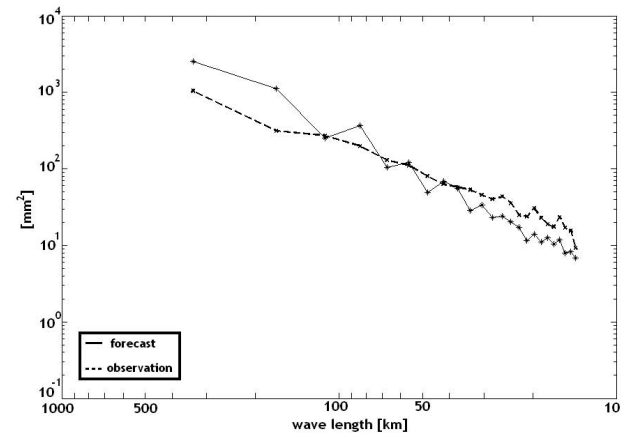
4 An observation-driven analysis

In this section we change the viewpoint of the analysis as mentioned at the end of the introduction: the forecast verification is driven by ex-post observations of severe events. The observation-driven analysis of the extreme rainfall events in Italy has been organized in two distinct phases: the identification and the selection of the periods of intense precipitation.

The first phase identifies the beginning of a precipitation period as the hour in which at least one raingauge over the Italian territory records an hourly rainfall depth higher or equal to 2 mm. A precipitation period ends when the same threshold is not exceeded in any sensor for at least 6 h. For instance, this results in the identification of 134 periods of precipitation for the whole year 2006. Additional criteria are needed in order to isolate a subset of periods really significant
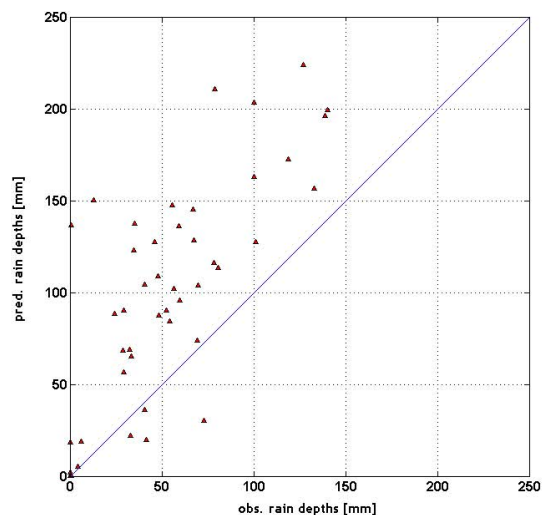
Fig. 9. In this panel the total rainfall depths (computed in millimetres) averaged over larger-than-200 km^2 catchments are presented. The y-axis reports the predicted rainfall while the x-axis shows the observations.
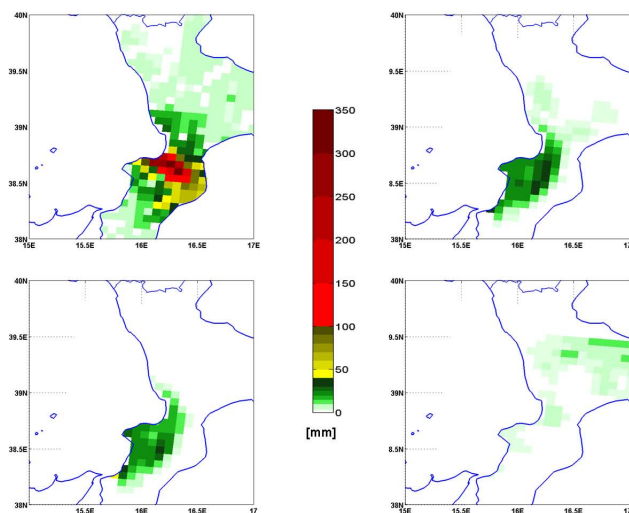

Fig. 10. In top left panel, the total rainfall measured by raingauges between 09:00 and 18:00 UTC fallen over south-western Calabria. Moving clockwise from the upper left, the rainfall predicted by RUN-0, RUN-24, and RUN-48 runs, respectively.

from a hydrometeorological viewpoint. A threshold has been applied in order to single out the period of intense precipitation containing at least one hourly rainfall depth at a given rain-gauge higher than 50 mm (pivot rain-gauge). This value corresponds to a local return period of about 50 years and to a nondimensional quantile equal to 2, as derived from the regional frequency analysis of annual rainfall maxima (for duration of 1 to 24 h) over the Italian territory based on the Two Components Extreme Value Distribution (TCEV): that quantile value discriminates the transition between ordinary and extraordinary annual rainfall maxima (Boni et al., 2006, 2008). By direct inspection, the areal extent has been characterized by considering the neighbouring stations records.

The number of severe events were 19 for the year 2006. They are characterized by a relevant spatial extension and duration of the order of 12–60 h. The precipitation event is a process represented through hourly rainfall depths occurring over a number of points: the spatial and temporal precipitation process may allow null values in some cells, provided that cumulated rainfall depths overall increase until the end of the event itself.

4.1 Events classification and analysis criteria

The analysis of intense precipitation periods allowed selecting a subset of 19 precipitation events containing at least one hourly rainfall depth higher than 50 mm and lasting more than 8 h for the year 2006.

Of these events, 10 occurred during the fall season and were characterized by long duration (higher than 12 h) and relevant spatial extension (larger than 2500 km^2), while the remaining events occurred mainly during the summer season and were short in time and very localized.

To avoid too strict a constraint for observation and prediction to be contemporary, the forecast event has been searched in a 24-h wide time-window centred on the observed starting time.

The procedure was applied to three different initial times of COSMO-I7 runs: the run that started the same day (RUN-0) of the observed precipitation event, the day before (RUN-24) and two days before (RUN-48).

The comparison between the observed precipitations and the predicted ones was undertaken according to the same logic of the forecast-driven study, that is, by means of the analysis of the morphology of the event, spectral analysis, thermodynamic and kinematic features, basin-scale verification.

4.2 Case studies: the 3 July 2006 event over Vibo Valentia

This event, localized and short-lived, occurred on 3 July 2006 and hit the town of Vibo Valentia, in the South-Western part of Calabria. Precipitation began at 09:00 UTC and lasted, in total, 9 h, forcing a maximum rainfall depth of about 350 mm, observed over two small river basins in the area. The most severe rainfall lasted in 3 h (11:00 UTC–14:00 UTC) and was observed by the 30 hourly reporting rain-gauges of the region. A few of them, contiguous, observed an hourly maximum of 250 mm in one rain-gauge. The strongly convective triggering of precipitation was the main feature of this event (Chiaravalloti and Gabriele, 2008), which actually caused flooding of minor rivers, localized landslides damaging highways and 4 casualties.

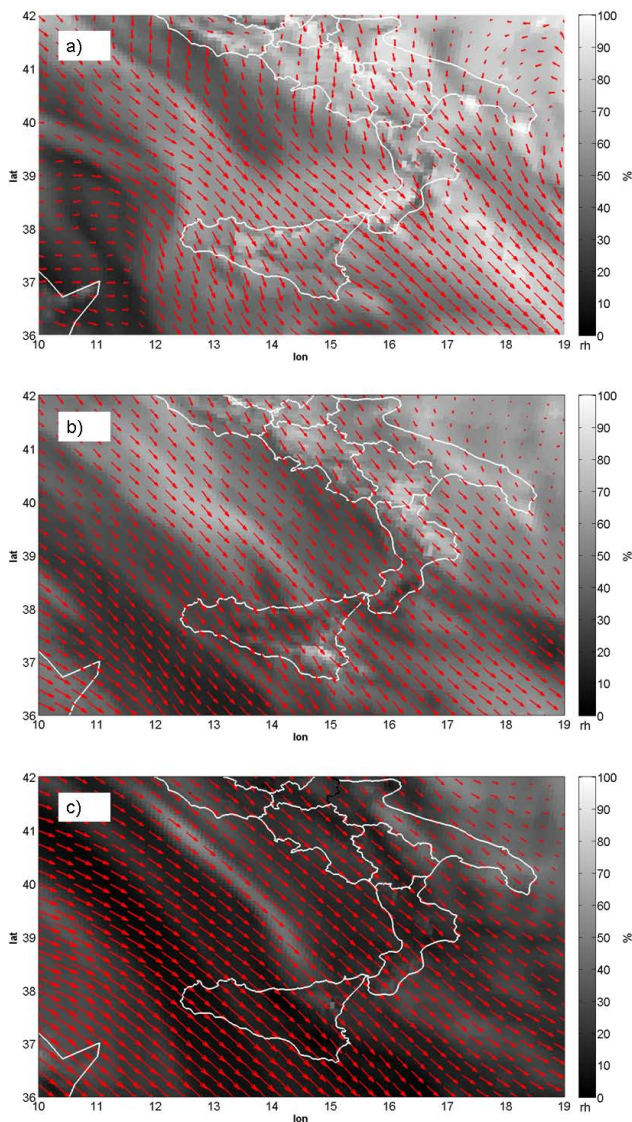No warning bulletin was issued by Italian Civil Protection Department.

Fig. 12. The columnar content of water vapour retrieved by AMSR-E in 6 different passes dating up to 2 days before the event (from http://www.remss.com/).

Fig. 11. Relative humidity and wind by RUN-0 at different elevations (850, 700, and 500 hPa, in panel **a**, **b**, and **c**, respectively).

In Fig. 10, the total rainfall measured by raingauges is shown in the top left panel, while RUN-0 h, RUN-24 and RUN-48 predictions are also displayed clockwisely. None of those runs was able to model correctly the severity of this event. Only the localization was predicted.

Water vapour kinematics and thermodynamics predicted by COSMO-I7 (RUN-0) is presented in the four panels of Fig. 11. The three-hour time-window starting at 12:00 UTC was chosen because it synthesizes well the overall prevailing dynamic configuration of the forecast during this event.

The crosschecking humidity data reveals a very low degree of saturation, hardly exceeding 50%, at 500, 700, and 850 hPa. That is the reason why no heavy precipitation was forecast.
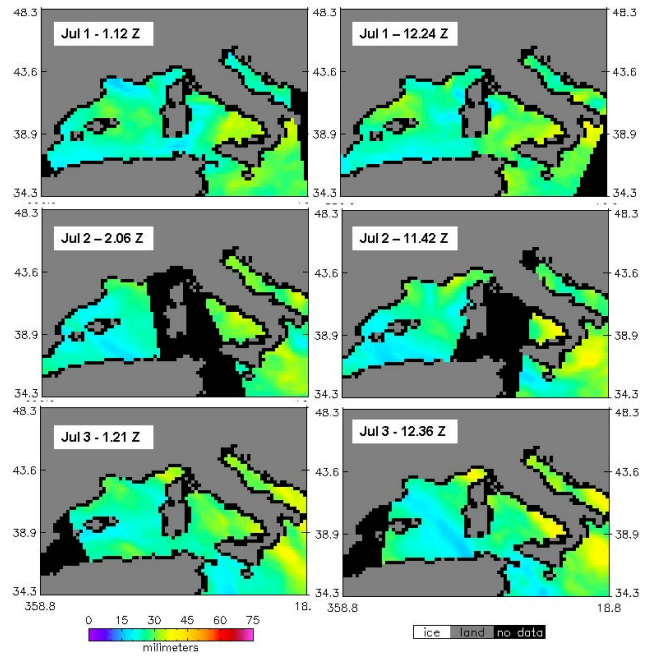
The predicted water vapour amount in the atmosphere was not enough to trigger rainfall due to orographic effect, in spite of the fact that the wind from 1500 to 5000 m was nearly orthogonal to the Apennine divide and shearless. In this case, a major discrepancy between observation and prediction does exist. AMSR-E water content was checked in order to understand the reasons of that discrepancy. The data reveal the presence of a band (150–200 km) of very humid air, denoting a water vapour columnar content of more than 40 mm, arranged consistently with the main wind direction in correspondence with the southern Tyrrhenian Sea (Fig. 12).

In order to gain a deeper understanding of the event's thermodynamic structure, the temporal trends of columnar content of water vapour, both observed and forecast, have been compared in Fig. 13. As the previous case study, the variables were averaged only over sea pixels. COSMO-I7 prediction of atmospheric water vapour content is dramatically underestimated, also considering all three model runs. RUN-0, for instance, predicts less 10 g/kg of water vapour columnar content while AMSR-E observations overcome 30 g/kg.

The observation-driven analysis has been further substantiated by the spectral analysis of observed and predicted data (Fig. 14).

From the viewpoint of ground effects, the performance of each of the 3 runs compared with observations is rather poor (Fig. 15). The worst prediction errors occur for basins hit by the most severe part of the event, where, against
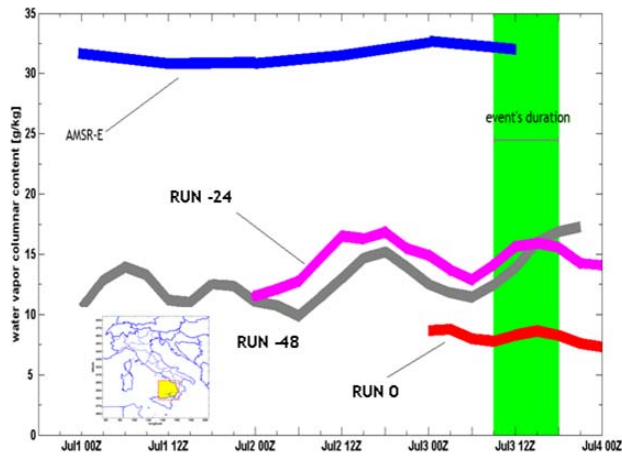
Fig. 13. Comparison of time series for water vapour columnar content. AMSR-E observations (2 daily passes) are represented in blue (aside is reported the local time of each pass). Dark grey, purple and red lines refer to COSMO-I7 runs. The domain used to average the columnar water vapour content for both AMSR-E and COSMO-I7 data is filled in yellow.

Fig. 14. Power spectra of the amplitudes for observed rainfall and for the three runs (dotted and dashed lines): all the predicted data underestimate the severity of the 3 July event by at least one order of magnitude.

measurements over 75 mm (averaged over the basin area), the total forecast rainfall barely reached 25 mm. Such a poor prediction eventually led to the decision not to warn local population, causing damages to infrastructures, casualties and, finally, legal actions brought against Italian Civil Protection Department.

Consequently, a quite good spatial location of precipitation fields cannot be traded off against quantitatively very poor prediction that heavily underestimates the observed rainfall, on the average, as underscored by the outcome of the basin-scale analysis. Due to the small scale of the whole event, in this case overall and inner uncertainties coincide both from a general and a local standpoint and the forecast has to be considered missed.

5 Discussion

In the previous sections, two of the over one-hundred analyzed cases have been fully described. One of them (Sect. 3) was a characteristic early fall meso-α scale event started by the orographic lifting over Alps-Apennine divide. The second one (Sect.4), also typical, was due to a large-scale transport that conveyed humidity to a small-scale orography-triggered process.

The whole set of the 2006 events is summarized in Table 1a–b from the point of view of predictive ability of COSMO I7. Out of the total observed events, extracted from the archive benchmarking, ninety-two belong to the set of ordinary risk scenarios, in the sense that slope saturation and river hydrographs did not reach extreme values. Twenty-seven belong to the high risk scenarios, in the sense that shallow landslides and/or river bank overtopping were observed.
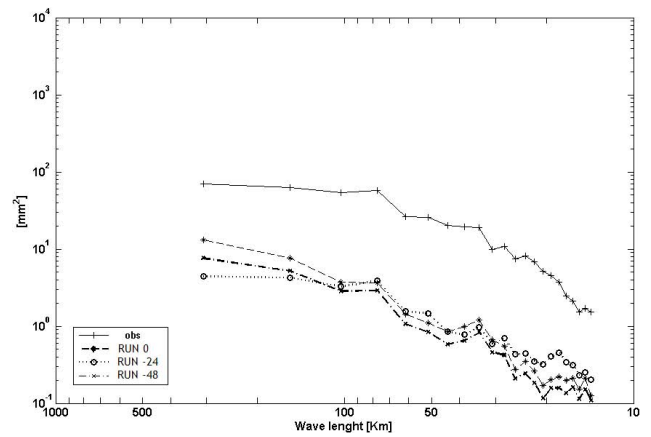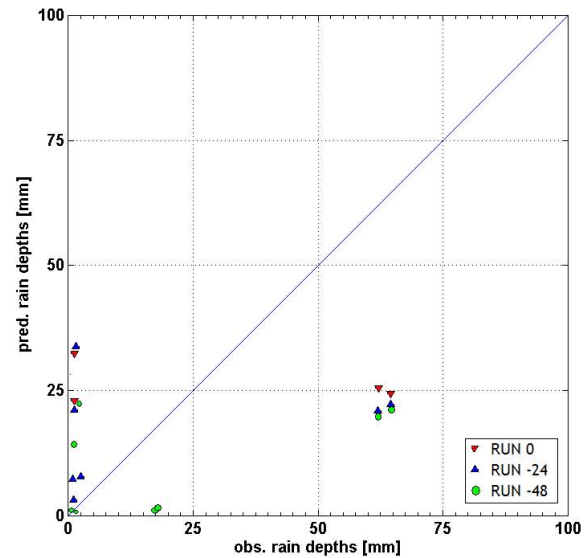

Fig. 15. The basin scale analysis display the very similar performance of the three model runs. The observed rainfall averaged over the basins hit by the event (x-axis) is compared with the predicted one plotted on y-axis: the largest underestimation occurred for basins that inundated Vibo Valentia.

Out of the 92 observed events with ordinary criticality, 15 occurred without prediction from the operational system and 77 were correctly forecast. Inspecting column 2 of Table 1a, it is also evident that 99 events were operationally predicted and 77 were eventually observed, whilst in 22 cases no event took place. It is noteworthy that none of the ordinary scenarios predicted exceeded its severity. It comes clear from the previous findings that the operational system did not underestimate predicted events. The false alarm rate is about 20% whilst the misses are about 15% and so the hit rate is 80%.
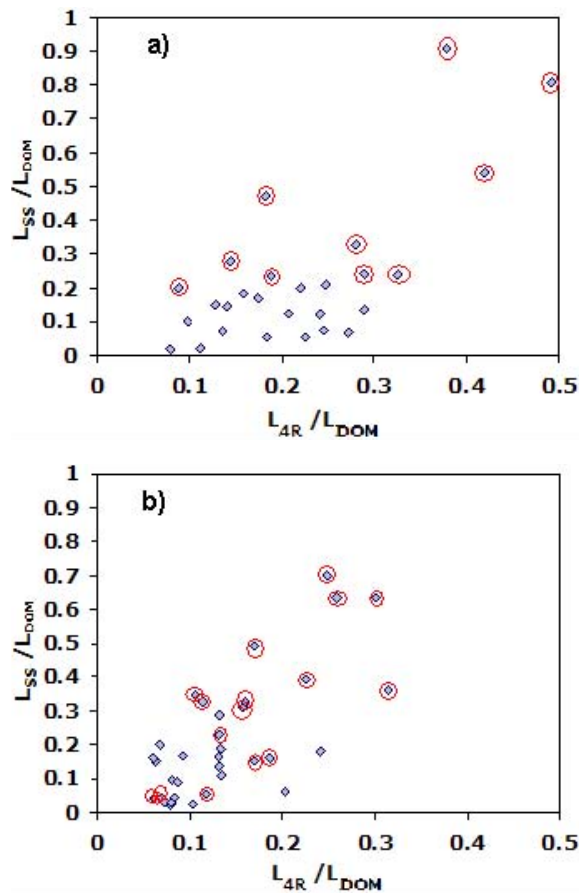
Fig. 16. In both panels, the scatterplots regard the ratio of spatial shift scale (L_{SS}), to the domain scale (L_{DOM}) on y-axis and the ratio of the "physical resolution" (L_{4R}) to L_{DOM} for high risk scenarios reported in Table 1b. Observation-driven results are displayed in panel **a**, whereas panel **b** shows the outcomes of the forecast-driven analysis. Red circles highlight severe events that were not correctly forecast (panel a) and predictions that were not verified by observations (panel b).

The system behaves differently for the high risk scenarios. Out of the total population of 34 severe events, the system overestimated 17 of them and correctly hit the remaining 17. Out of the total high criticality observed events, extracted from the archive benchmarking on observations, 17 were correctly forecast and the missed events were as high as 10. On the whole, the operational system for high risk scenarios has a false alarm rate about 50%, a rate of misses of about 35% and a hit rate of about 65%.

It was observed that events producing ordinary risk scenarios are normally induced by meso-α systems, in which the cyclonic circulation due to the cyclogenesis in the lee of the Alps, which is well described by the classical paper

by Buzzi and Tibaldi (1978), produces extended precipitation triggered by orographic lifting. For the events producing high risk scenarios, only part of them has the previous meteorological structure which generates distributed events of long duration, usually over 24 h. A fair part of the events are meso-β/meso-γ in structure, with duration less than 24 h.

Table 1b makes evident that the predictive ability of the operational COSMO I-7 system for such events is quite limited: 17 of the 27 observed events were meso-α and 10 meso-β/meso-γ. The system hit 16 out of 17 of the meso-α events, but hit only 1 of the meso-β/meso-γ.

The previous discussion makes evident that the scale lengths of rainfall events are one of the key features of the processes we investigated and the interdependencies of their ratios can characterize their predictability. In order to summarize our findings, let us assume that the space scale of the event is represented by L_{DOM}, defined as the width of the square domain associated with each rainfall event, L_{SS} represents the amplitude of the spatial shift between observed and predicted rainfall patterns, measured as the distance between the centre of mass of the rainfall volume, and finally L_{4R} is four times the horizontal resolution of the model (7×4=28 km), in some sense representing the "physical" resolution.

The inspection of Fig. 16 makes immediately evident that the localization of the events is most frequently hit than the severity. The ratio between the spatial shift to the domain scale (L_{SS}/L_{DOM}), plotted on y-axis, while x-axis reports the ratio between the "physical" resolution of the model to the domain scale (L_{4R}/L_{DOM}). Panel a shows results for observation-driven high risk events and panel b displays the same results for forecast-driven ones. Red circles highlight severe events that were not correctly forecast (panel a) and predictions that were not verified by observations (panel b).

On both panels most part of the events shows a ratio between the spatial shift L_{SS} and size L_{DOM} less than 0.3. Only a few outliers exceed L_{SS}/L_{DOM}=0.5 and only a few show L_{SS}=0. Such a consideration means that for most of the events at least 70% of the rainfall volume was observed in an area contained within the predicted one.

The space predictive ability is a quite complex concept: in each predicted event the alarm is false in some part of the domain, is missed in another part but it is true in most part of it. Fortunately from the observations for the year 2006, the missed alarm part and the false alarm part of the area associated with the event are less than 30%. Ultimately, this factor allows citizens to perceive as true an alert warning about landslides or floods not affecting them directly but the nearby areas.

Hence, the ratio of L_{SS} and L_{DOM} on the y-axis gives information on the range of model mislocation as function of the magnitude of the area on which the rain fell and, in the same way, the x-axis ratio (L_{4R}/L_{DOM}) fixes the magnitude of area impacted by the storm as a function of the reliable scale of the model. So, moving leftwards along the

x-axis means dealing with larger storms, while if the origin is approached along the y-axis means having small space shift and a definitely correctly placed prediction.

This kind of analysis has been split between observation-driven and forecast-driven (for both high criticality risk and ordinary risk events) case studies.

If the event duration is considered, a quite apparent distinction of the model performances arises whether precipitation is long lasting or not.

In case of long-lived events the meteorological model provides a satisfactory representation of the low-level and mid-tropospheric horizontal wind and of its interaction with the orography. The wind at the low-level is consistently aligned along the direction of maximum slope and this guarantees one of the key ingredients for orography-induced intense rainfall processes.

The same is true for the spatio-temporal evolution of the water vapour field, which is correctly modelled both in its impinging on orography in the boundary layer region and in its advection by the large-scale wind in the mid and high-tropospheric region. This result confirms that the uncertainty in the physical modelling of intense precipitation events is mainly associated with small-scale processes O(10 km), while the processes at larger scales, O(100 km), are generally reliably modelled. When we move our attention to short duration precipitation events the results are certainly more contradictory. On one hand, the model still shows in the majority of cases a correct representation of kinematics over the orography, on the other hand however, a too much dry moisture initialization arises as one of the main reasons for the misprediction of such class of events.

6   Conclusions

The performed analyses succeeded in spotting the major features causing COSMO-I7 QPF failures and also the good forecast skill shown for a well-defined set of events. Firstly, a poor modelling when mesoscale orographic lifting of humid air processes are addressed. Usually, this leads to large underestimation of up-wind side rain depths and to dramatic underestimation in the leeside. Secondly, an inadequate rate of turbulent mixing within the planetary boundary layer (PBL), induced by the so-called water vapour "channelling", is often found as narrow mountainous valleys could not be modelled satisfactorily. In these sectors, an artificial saturation takes place, usually joined to a strongly stable atmosphere, inhibiting the accurate modelling of latent and sensible heat in the PBL. This orographic drawback is mainly due to common horizontal resolution of non-hydrostatic mesoscale models (∼10 km). The last point concerns a lacking humidity advection in the PBL at synoptic scale. The aftermath is the incapability of triggering deep convection with a sufficient humidity rate. Indeed, this kind of miscalculation is not actually a mesoscale model inaccuracy, but mostly a global

circulation mismodelling, the results of which provide the boundary conditions of the high resolution model.

This feature is common to almost all operational mesoscale models (WRF, MM5, Arome, etc.). A widely shared future research agenda aimes to reach a finer resolution as a means of improving limited area models QPF (Davis et al., 2004; Deng Stauffer, 2006; Kain et al., 2008; Petch, 2006). Nevertheless, moving towards finer scales by itself, as shown by Bryan et al. (2003), may not grant a proportional upgrading in terms of prediction ability. In fact, although the orographic representation will surely progress, thus eliminating the requirement for parameterizing deep humid convection processes, this should be coupled to the assimilation of compatible-scale observations, whether direct or remote.

A final remark must be spent on Table 1a–b. Both of them have been filled with numbers derived from a rigorous analysis based on physical, microphysical, dynamical and kinematical aspects of rainfall processes and predictions. Nevertheless, the evaluation of each prediction is under some aspects still subjective and above all synthetic. The subjectivity relies, for instance, on the weight given to outliers in Figs. 9 and 15. Is it affordable to define as overestimated (underestimated) the event of 14 September (3 July) and to disregard that at the same time some basins of the domain experienced severe rainfall, or even flood, although these were not predicted (and vice versa)? In other words, attention must be paid to which scale refers, stressing again the possibility of locally unpredicted events (and almost unpredictable, for their characteristic dimension lies under, or slightly over, the model resolution), even if a forecast is satisfying "on the average".

For all these motivations and to go beyond a 1-year-only record, an analogous analysis on 2007 and 2008 databases will be soon available.

References

AVI-CNR, Italian Research Council-Group for Natural Disasters Mitigation: Vulnerable areas of Italy project final report, Rome, 2004.

Bryan, G. H., Wyngaard, J. C., and Fritsch, J. M.: Resolution requirements for the simulation of deep moist convection, Mon. Weather Rev., 131, 2394–2416, 2003.

Boni, G., Parodi, A., and Rudari, R.: Extreme rainfall events: learning from rain gauge time series, J. Hydrol., 327, 304–314, 2006

Boni, G., Parodi, A., and Siccardi, F.: A New Parsimonious Methodology of Mapping the Spatial Variability of Annual Maximum Rainfall in Mountainous Environments, J. Hydrometeorol., 9, 492–506, 2008.

Buzzi, A. and Tibaldi, A.: Cyclogenesis in the lee of the Alps: A case study, Q. J. Roy. Meteor. Soc., 104, 271–287, 1978.

Buzzi, A. and Foschini, L.: Mesoscale meteorological features associated with heavy precipitation in the southern Alpine region, Meteorol. Atmos. Phys., 72, 131–146, 2000.

Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocernich, M., Damrath, U., Ebert, E. E., Brown, B. G., and Mason, S.: Forecast verification: current status and future directions, Meteorol. Appl., 15–1, 3–18, 2008.

Chiaravalloti, F. and Gabriele, S.: Vibo Valentia flood and MSG rainfall evaluation, Atmos. Res., 93(1–3), 286–294, 2009.

Davis, C.: The bow echo and MCV Experiment (BAMEX): Observations and opportunities. B. Amer. Meteorol. Soc., 85, 1075–1093, 2004.

Deng, A. and Stauffer, D. R.: On improving 4-km mesoscale model simulations, J. Appl. Meteorol. Clim., 45, 361–381, 2006.

Elementi, M., Marsigli, C., and Paccagnella, T.: High resolution forecast of heavy precipitation with Lokal Modell: analysis of two case studies in the Alpine area, Nat. Hazards Earth Syst. Sci., 5, 593–602, 2005,
http://www.nat-hazards-earth-syst-sci.net/5/593/2005/.

Ferraris, L., Rudari, R., and Siccardi, F.: The Uncertainty in the Prediction of Flash Floods in the Northern Mediterranean Environment, J. Hydrometeorol., 3, 714–727, 2002.

Kain, J. S., Weiss, S. J., Bright, D. R., Baldwin, M. E., Levit, J. J., Carbin, G. W., Schwartz, C. S., Weisman, M. L., Droegemeier, K. K., Weber, D. B., and Thomas, K. W.: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP, Weather Forecast., 23, 931–952, 2008.

Milelli, M., Oberto, E., and Parodi, A.: Sensitivity experiments of a severe rainfall event in north-western Italy: 17 August 2006, Advances in Science and Research, 2, 133–138, 2008.

Petch, J. C.: Sensitivity studies of developing convection in a cloud-resolving model, Q. J. Roy. Meteorol. Soc., 132, 345–358, 2006.

President of the Council of Ministers Directive, 2/27/2004 Italy Official Gazette, Rome, 2004

Rotunno, R. and Ferretti, R. : Mechanisms of intense Alpine rainfall, J. Atmos. Sci., 58, 1732–1749, 2001.

Siccardi, F., Guzzetti, F., Cardinali, M., Reichenbach, P., Cipolla, F., Sebastiani, C., Galli, M., and Salvati, P.: Preliminary analysis of landslides triggered by the 23 November Event in Western Liguria, in: Proceedings 3rd Plinius Conference on Mediterranean Storms, Cagliari, Italy, ISBN 88-8080-031-0, 383–392, 2002.

Ramos, M. H., Bartholmes, J., and Thielen-del Pozo, J.: Development of decision support products based on ensemble weather forecasts in the European Flood Alert System, Atmos. Sci. Lett., 8, 113–119, 2007.

Siccardi, F., Boni, G., Ferraris, L., and Rudari, R.: A Hydrometeorological Approach for Probabilistic Flood Forecast, J. Geophys. Res., 110, D05101, doi:10.1029/2004JD005314, 2005.

Steppeler, J., Hess, R., Doms, G., Schattler, U., and Bonaventura, L.: Review of numerical methods for non hydrostatic weather prediction models, Meteorol, Atmos. Phys., 82, 287–301, 2003.

van der Knijff, J. M., Younis, J., and De Roo, A. P. J.: LISFLOOD: a GIS-based distributed flood and water balance model for European catchments, Int. J. Geogr. Inf. Sci., in press, 2008.

Wieczorek, G. W. and Guzzetti, F.: A review of rainfall thresholds for triggering landslides, in: Proceedings 2.nd Plinius Conference on Mediterranean Storms, GN-DCI pub. n. 2012, 407–414, 2000.

UNISDR Towards National Resilience: Good practices of National Platforms for Disaster Risk Reduction, Rome, 2008.