

Interactive comment on “Regional ensemble forecast for early warning system over small Apennine catchments on Central Italy” by Rossella Ferretti et al.

Anonymous Referee #1

Received and published: 26 June 2019

This manuscript presents a regional-scale flood forecasting system based on the combination of ensemble meteorological predictions with hydrological modeling. The hydro-meteorological chain is based on the dynamical downscaling of short-term (i.e., 3 days) GFS predictions using WRF-ARW and offline CHyM hydrological model simulations. The forecasting system is tested for an extreme flood event that hit central part of Italy on November 2017. Flood warnings are issued using an index based on the ratio between the flow discharge and the hydraulic radius.

Establishing and testing an advanced flood warning system is a subject of interest for the broad hydro-meteorological community and certainly suitable for HESS readership.

C1

Furthermore, water agencies operationally dealing with flood predictions should greatly benefit from well-designed and technically advanced studies aimed at investigating the skills and/or deficiencies of hydro-meteorological prediction chains. Unfortunately, this work does not offer new and useful insights, lacking novelty and a rigorous evaluation approach that could eventually turn into valuable information for both the scientific and operational hydro-meteorological communities. In light of this overall assessment, I would consider this manuscript not suitable for publication in HESS. I highlight below a series of more specific comments.

1. The manuscript fails to identify an outstanding research question associated to the development and testing of probabilistic flood forecasting. In my opinion, the novelty of the work cannot rely on the simple combination of downscaled probabilistic meteorological forecasts with hydrological simulations. Moreover, after reading the manuscript it is not clear what's the message the authors want to convey when they generically discuss about the pros and cons of deterministic vs probabilistic forecasting approach. Nowadays the combined use of both is an established practice implemented by many hydro-meteorological centers. That is, the “complementarity” should be substantiated with ad-hoc results and not just advocated. In addition, saying that the probabilistic approach allows for longer forecast periods is absolutely misleading. Finally, in order to put this work in the right perspective I would have referred to the recent and innovative efforts behind the development of WRF-Hydro aiming at fully integrated hydro-meteorological forecasts.

2. The manuscript lacks a quantitative approach in the analysis/interpretation of the modeling results. A robust verification framework should be part of a reliable early warning system. Several common verification scores should have been implemented in order to assess the performance of the modeling results; mainly precipitation and streamflow. Here I would have also paid special attention on the quantification of the spatial agreement between simulated and observed precipitation (using radar data), which is key for short-term distributed flood forecasting. The reader is left with unver-

C2

ifiable statements (e.g., “very good response”, “good agreement”, “suggesting a more accurate forecast”, etc.) that do not make things clearer.

3. The content of the manuscript is not well organized (i.e., section order). In general you should present first data and methods (i.e., numerical models), define the skill metrics and/or indices, and at the end you interpret/discuss the results. The language should be improved. This is not just a matter of typos, grammar mistakes, and unclear sentences. For instance, “reliability” and “uncertainty” do not have the same meaning in probabilistic forecasting. I am not sure authors are using the term “spatial coherence” in the right way. Finally, many figures (e.g., Fig. 3, Fig. 7, Fig. 11) and Table 1 are really not necessary. Please note also that the geographical location of the two Italian regions (i.e., Umbria and Abruzzo) is not shown in any map. The same apply for the hydrometric stations.

4. I have some remarks on the model setup and configuration: - The use of 1° GFS forecast is not fully justified in my opinion because other deterministic and probabilistic products (e.g., ECMWF) are available at higher resolution. - Domain definition, grid resolution, and physical parameterizations have a large influence on model results. Did the authors made preliminary tests to check their impacts on the selected events? This is key for a sound modeling strategy, especially from an operational perspective. - The definition of the benchmark configuration is not clearly discussed. If I look at Fig. 1 of Pichelli et al., 2017, the 1km domain (D3) does not fully cover the study area of this work, am I wrong? In the same paper it is mentioned that the operational setup of WRF-CETEMPS is different from the one shown in Fig.1. Further, it seems that the high-resolution setup uses GFS later boundary forcing with different resolutions (i.e., 0.25° instead of 1° resolution) and different physical model parameterizations. Finally, it seems not completely justified to directly nest the 9km WRF into the 1° GFS forecast. I would have expected an intermediate step to reduce lateral boundary effects. In general, these aspects of the work are not clearly explained.

5. I have several remarks concerning the adopted discharge index and the discussion

C3

of the related results: - Authors consider the definition of the BDD index necessary due to the lack of discharge measurements. This contradicts the definition (Eq. 2) of the index itself, which is based on discharge values! - What's the equation used to calculate the hydraulic radius as a function of the drainage area? - The comparison between Fig. 10 and Fig. 9 is not intuitive. - I do not fully agree with the interpretation of Fig. 10. I see a good agreement in the timing even for those stations heavily impacted from hydropower production (i.e., Vomano and Todino). I also think that the mismatch for Pescara River could be due to some error in the observed atmospheric forcing at the local stations. That's why a more careful evaluation of the atmospheric forcing would have provided more useful insights.

6. It is really difficult to follow the discussion around Fig. 12. For instance, authors interpret the results saying that the mismatch between “observed” and “simulated” BDD index is due to precipitation occurring only on a very small area and not capture by the model. What do you mean with “small”? The high-resolution simulations are at 1km! I suspect that you can get the same issue if you go down to 100m resolution. Again, if you do not carefully evaluate the atmospheric simulations it is difficult to provide convincing interpretation of the BDD index. Finally, I would also remark that authors talk about “overestimation” and “underestimation” of the BDD index using as a reference the model results driven with observed (interpolated?) precipitation. I am fine with this as long as you cross-validate local precipitation measurements with other sources of information, e.g., spatially distributed information obtained from radar retrievals.

7. I would expect the same kind of curve when I look at the black (“observed”) lines in Fig. 14 and the red ones in Fig. 10, am I wrong somewhere? One of the main conclusions is that the uncertainty in the BDD index is underestimated if you do not perturb the parameters of the hydrological model. This is intuitive and this is the reason why you should take both (“atmospheric” and “hydrologic”) into account. In my opinion this opportunity was missed in this work.

C4

