Ensemble flood forecasting considering dominant runoff processes: I. Setup and application to nested basins (Emme, Switzerland)

Authors replies to RC2:

We want to thank the reviewer for his/her assessment of our manuscript. In the following we give our answers to the comments and recommendations that have been raised. Reviewer comments RC are **bold,** our reply AR is in *italic*. Insertions in the revised manuscript MI are underlined.

MAJOR COMMENTS

RC: However, it should be considered that the paper has some conceptual/methodological limitations:

- **The investigation period (of some months) is very limited in order to gain substantial data for further statistical analyses (e.g., skill determination, etc.). However, the authors address and discuss this issue in their text.**
- **The number of investigated catchments is low; conclusions on regional transferability of the suggested methods remains limited due to lacking statistical significance. However, the authors anticipate these shortcomings within the study design (by using MC methods) and briefly discuss the matter.**

AR: We are of course aware, that more basins and longer periods of evaluation are always welcome. NHESS is in this respect a journal that regularly publishes case studies (e.g. Kobayashi et al., 2016; Cane et al., 2013), preliminary assessments (Picciotti et al., 2013) or intercomparison of approaches during limited period of time (e.g. Davoli et al., 2018). Having targeted NHESS as journal for disseminating our experience, this study is designed to evaluate two different model structures, during a representative flood season and in case of nested basins with different area. With this approach we can learn about the quality of the novel approaches from different perspectives at the same time. The transfer of experience to another catchment and climatic region is presented in the companion paper by Horat et al. (2018). As far as the length of the investigation period is concerned, some limitations arise from the use of COSMO-E and COSMO-1. MeteoSwiss decommissioned after several years the antecedent operational NWP COSMO-2 and COSMO-LEPS in 2016. As we want to make our systems operational, it was for us important to focus on a first analysis with the NEW NWPS that we receive and archive in real-time since February 2016.

In the revised manuscript we will better declare our choices concerning selection of basins and investigation period..

RC: For the reviewer, it was a bit hard to get the methodological (?) connection of the reviewed paper to the "accompanying paper" (Horat et. al., 2018), as well as to previous work of the researchers (e.g., Zappa et al., 2011 or Antonetti, 2017). Maybe, a graphical, structured representation of the questions covered in those papers would offer a way to better comprehend the overarching research activities of authors/group/lab and serve the scientific significance of the manuscript.

AR: With thank the reviewer for this suggestion. We first taught to follow his comment and create an image, but finally opted for a table that we will insert in the supplementary material of the revised manuscript. The table is also provided here below.

Paper	*Zappa et al.*	*Addor et al.*	*Liechti et al.*	*Antonetti et al.*	*Antonetti et al.*	*Antonetti et al.*	*Horat et al.*
Year	*2011*	*2011*	*2013*	*2017*	*2018*	*2018*	*2018*
Journal	*At. Research*	*HESS*	*HESS*	*Hydrol. Proc.*	*HESS*	*NHESSD*	*NHESSD*
Target areas							
Verzasca	X		X				X
Sihl		X					
Emme					X	X	
Other			X	X			
Topics							
Forecasting	X	X	X			X	X
Model development				X			
Uncertainty propagation	X		X		X	(X)	(X)
Intercomparison		X	X	(X)	(X)	X	X
Model/module							
PREVAH-HRU	X	X	X				X
RGM-PRO				X	X	X	X
RGM-TRD					X	X	
Rainfall forcing							
Intrepolated gauges	X	X	X		X		X
Combiprecip				X	X	X	X
COSMO-1							
COSMO-2	X	X	X				
COSMO-LEPS	X	X				X	X
COSMO-E						X	X
Weather radar nowcasting	X		X				
Frequency	continuous	continuous	events	events	events	events	events
Period	2007-2010	2007-2009	2007-2010	2005-2016	2005-2016	2016	2016
Analyses							
NSE/KGE	NSE			KGE	KGE	NSE/KGE	
Brier/ROC/FAR/RankHist	(X)	X	X			X	X
MonteCarlo	X		(X)	X	X	X	(X)
Other	SWAE				ANOVA		

Zappa et al. (2011) is our benchmark paper on uncertainty propagation

Addor et al.. (2011) is our reference work on verification of deterministic and ensemble forecasts

Liechti et al. (2013) focuses on flash-flood nowcasting with advanced weather radar products

Antonetti et al. (2017) introduces RGM-PRO

Antonetti et al. (2018, HESS) evaluate structures and configurations of RGM-PRO in the Emme catchment

Antonetti et al. (2018, NHESSD) first apply RGM-PRO in forecasting mode for the Emme catchment and is our first study with COSMO-E/COSMO-1

Horat et al. (2018, NHESSD) applies RGM-PRO in forecasting mode for the Verzasca catchment and compare its quality with our current operational model as forced by COSMO-E/COSMO-1.

SPECIFIC COMMENTS

RC: P1-L1: Add "potential" before the word "risk" >> if there is no vulnerability, a hazard (i.e., heavy rain) would pose no risk.

AR: Will be addressed

RC: P1-L21: What is meant by "satisfying skill" here? Please give a brief comment on that in the manuscript.

AR: Will be addressed

RC: P2-L31: Maybe add a reference to Collier & Fox (2003) who propose a Flash Flood Susceptibility Assessment Procedure (FFSAP) which is quite comparable to the one proposed by Mani et. al (2012).

AR: We will elaborate on FFSAP in the revised manuscript

RC: P2-L32: What is meant here by the word "torrent"?

AR: We replaced "torrent" with "flash floods", it was originally meant "torrential flow"

RC: P3-L2: What is determined "with radar" and why?

AR: We will specify the meaning of "radar", i.e. identification of thresholds according to data from rainfall radar.

RC: P4-Ls27-30: The systematic of "physically-based" and "conceptual" models excludes other approaches (e.g., FFSAPs) which are potentially useful for deriving information on "timing and magnitude" (e.g., see http://www.hochwasserzentrum.sachsen.de/fruehwarnung which is based on a simple FFSAP).

AR: As already states before we will elaborate on FFSAP in the revised manuscript and also present the application for Saxony (Philipp et al., 2016)

RC: P5-L4: Does this mean "assimilation (of transient data)"? If this is not the case, would not it be better to talk of "estimation", rather than "assimilation"?

AR: We will rephrase the sentence: "To what extent does the skill of the FF prediction depend on the use of model structures considering spatially distributed information on runoff processes into a hydrological model?"

RC: P5-L12: What is the "statistical approach"? Better replace with something like "skill assessment procedures. . . "?

AR: Will be addressed

RC: P8-L1: What is meant by ". . . served as fingerprint"?

AR: We used here the terminology introduced by Blöschl et al. (2009). The sentence will be expanded: "With this method, the map of RTs serves as fingerprint since it contains information determining the spatial variability of soil moisture Antonetti (2018)."

RC: P8-L3: "Traditional benchmark version. . . " . . . of what?

AR: "Traditional benchmark version with conventional hydrological runoff generation module"

RC: P9-L2: From our point of view, using continuous measures for skill assessment (e.g., NSE or KGE) should be called "validation". On the other hand, employing event or threshold-based (binary, dichotomous) methods (e.g., AUC or BSS), it is "verification".

AR: We will re-arrange the section, remove the sub-section title and introduce the definition of BSS in order to make the manuscript less dependent on the companion paper.

RC: P11-L5: Are the probabilistic forecasts post-processed? This should be stated if this applies.

AR: no, they are not.

Technical corrections

RC: P5-L8: Please rewrite ". . . model is executed at the runoff gauge" (language).

AR: Will be addressed

RC: P8-L8: Better replace "completed" with "conducted".

AR: Will be addressed

RC: P19-L21: "bericht" should be "Bericht" (capitalized).

AR: Will be addressed

RC: Larger figures/numbers (e.g., "2120") are not separated in the manuscript. Please check, if this is in agreement with the NHESS style guide (e.g., be written as "2,120").

AR: Will be addressed

References:

Addor, N., Jaun, S., Fundel, F., and Zappa, M.: An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): Skill, case studies and scenarios, Hydrology and Earth System Sciences, 15, 2327–2347, doi:10.5194/hess-15-2327-2011, 2011.

Antonetti, M., & Zappa, M. (2018). How can expert knowledge increase the realism of conceptual hydrological models? A case study based on the concept of dominant runoff process in the Swiss Pre-Alps. *Hydrology and Earth System Sciences*, *22*(8), 4425-4447. https://doi.org/10.5194/hess-22-4425-2018

Antonetti, M., Scherrer, S., Kienzler, P. M., Margreth, M., & Zappa, M. (2017). Process-based hydrological modelling: the potential of a bottom-up approach for runoff predictions in ungauged catchments. *Hydrological Processes*, *31*(16), 2902-2920. https://doi.org/10.1002/hyp.11232

Blöschl, G., Reszler, C., and Komma, J.: A spatially distributed flash flood forecasting model, Environmental Modelling and Software, 23, 464–478, doi:10.1016/j.envsoft.2007.06.010, 2008.

Cane, D., Ghigo, S., Rabuffetti, D., and Milelli, M.: Real-time flood forecasting coupling different postprocessing techniques of precipitation forecast ensembles with a distributed hydrological model. The case study of may 2008 flood in western Piemonte, Italy, Nat. Hazards Earth Syst. Sci., 13, 211-220, https://doi.org/10.5194/nhess-13-211-2013, 2013.

Devoli, G., Tiranti, D., Cremonini, R., Sund, M., and Boje, S.: Comparison of landslide forecasting services in Piedmont (Italy) and Norway, illustrated by events in late spring 2013, Nat. Hazards Earth Syst. Sci., 18, 1351-1372, https://doi.org/10.5194/nhess-18-1351-2018, 2018.

Horat, C., Antonetti, M., Liechti, K., Kaufmann, P., and Zappa, M.: Ensemble flood forecasting considering dominant runoff processes: II. Benchmark against a state-of-the-art model-chain (Verzasca, Switzerland), Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2018-119, in review, 2018.

Kobayashi, K., Otsuka, S., Apip, and Saito, K.: Ensemble flood simulation for a small dam catchment in Japan using 10 and 2 km resolution nonhydrostatic model rainfalls, Nat. Hazards Earth Syst. Sci., 16, 1821-1839, https://doi.org/10.5194/nhess-16-1821-2016, 2016.

Liechti, K., Panziera, L., Germann, U., and Zappa, M.: The potential of radar-based ensemble forecasts for flash-flood early warning in the southern Swiss Alps, Hydrology and Earth System Sciences, 17, 3853–3869, doi:doi:10.5194/hess-17-3853-2013, 2013.

Picciotti, E., Marzano, F. S., Anagnostou, E. N., Kalogiros, J., Fessas, Y., Volpi, A., Cazac, V., Pace, R., Cinque, G., Bernardini, L., De Sanctis, K., Di Fabio, S., Montopoli, M., Anagnostou, M. N., Telleschi, A., Dimitriou, E., and Stella, J.: Coupling X-band dual-polarized mini-radars and hydro-meteorological forecast models: the HYDRORAD project, Nat. Hazards Earth Syst. Sci., 13, 1229-1241, https://doi.org/10.5194/nhess-13-1229-2013, 2013.

Philipp, A., Kerl, F., Büttner, U., Metzkes, C., Singer, T., Wagner, M., and Schütze, N.: Small-scale (flash) flood early warning in the light of operational requirements: opportunities and limits with regard to user demands, driving data, and hydrologic modeling techniques, Proc. IAHS, 373, 201-208, https://doi.org/10.5194/piahs-373-201-2016, 2016.

Zappa, M., Jaun, S., Germann, U., Walser, A., and Fundel, F.: Superposition of three sources of uncertainties in operational flood forecasting chains, Atmospheric Research, 100, 246–262, doi:doi:10.1016/j.atmosres.2010.12.005, 2011.