EPiC Series in Engineering
Volume 3, 2018, Pages 207–216
HIC 2018. 13th International Conference on Hydroinformatics

Flood forecasting with uncertainty using a fully automated flood model chain: A case study for the City of Kulmbach

Md Nazmul Azim Beg[1], Jorge Leandro[2], Punit Bhola[2], Iris Konnerth[2], Kanwal Amin[2], Florian Köck[3], Rita F. Carvalho[1] and Markus Disse[2]
[1] Marine and Environmental Sciences Centre, Department of Civil Engineering, University of Coimbra, Portugal
[2] Chair of Hydrology and River Basin Management, Technical University of Munich, Germany
[3] Bayerisches Landesamt für Umwelt, Germany
jorge.leandro@tum.de

Abstract

Real time flood forecasting can help authorities in providing reliable warnings to the public. This process is, however, non-deterministic such that uncertainty sources need to be accounted before issuing forecasts. In the FloodEvac project, we have developed a tool which takes as inputs rainfall forecasts and links a hydrological with a hydraulic model for producing flood forecasts. The tool is able to handle calibration/validation of the hydrological model (LARSIM) and produces real-time flood forecast with associated uncertainty of flood discharges and flood extents. In this case study, we focus on the linkage with the hydrological model and on the real-time discharge forecasts generated.

1 Introduction

Forecasting of flood events is a non-deterministic process in which uncertainty stems from different sources (Deletic et al., 2012). Disregarding its non-deterministic nature leads to disregarding events with lower probability (Leandro, Leitão, & de Lima, 2013) which may already trigger warnings to specific sensitive areas. In the FloodEvac project, we develop a real-time flood forecasting tool which is able to forecast flood discharges and flood extents with the inclusion of uncertainties. The tool is developed within the FloodEvac project funded by the *Bundesministerium für Bildung und Forschung* (BMBF, FKZ 13N13196 (TUM)).

G. La Loggia, G. Freni, V. Puleo and M. De Marchis (eds.), HIC 2018 (EPiC Series in Engineering, vol. 3), pp. 207–216

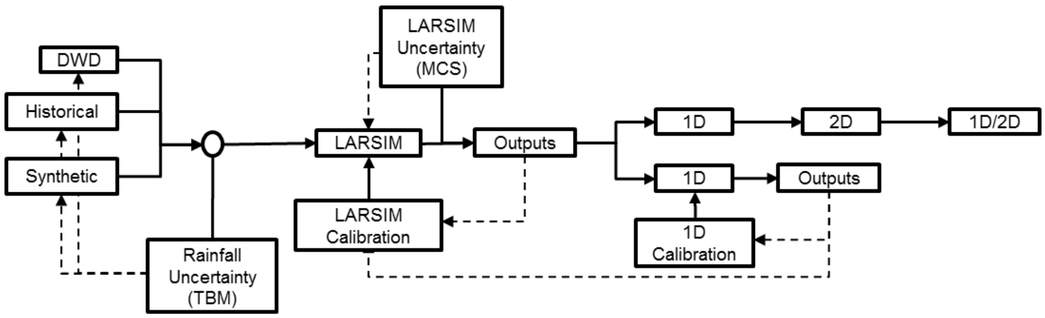

Figure 1: FloodEvac Tool and Model chain

2 Methodologies: The FloodEvac Tool

The FloodEvac tool allows the simulation of the rainfall-runoff process while including uncertainty from different sources (**Figure 1**). The tool can be run in simulation or forecast modus. The former is suitable for reproducing specific flood events or for the simulation of long time series (e.g. yearly). The latter is suitable for real-time flood forecasting. The model chain includes a rainfall uncertainty module, an uncertainty and calibration module for the hydrological model, and a link to several hydraulic models. The three modules will now be shortly explained.

In the rainfall module, rainfall data can be introduced in three different ways: using 1) observed/forecast rainfall from German Meteorological Services (Deutscher Wetterdienst, DWD), 2) generated rainfall based on historical data, or 3) generated rainfall based on synthetic data. Uncertainty can then be added to catchment rainfall based on the sequential conditional simulation (Seo, Kim, & Singh, 2014).

LARSIM (Large Area Runoff Simulation Model) is the hydrological conceptual model used in the tool. It is suitable for the simulation of rainfall-runoff in large catchments. The soil module consists of three storages: upper, middle and lower soil storage which contribute to the discharge components modelled as a linear storage system (**Figure 2** left panel). It includes 34 parameters which allow modelling of different processes such as direct discharge, interflow and groundwater flow (please see (Haag, A, M, & Bremicker, 2016) for a complete description of the parameters). The tool also includes a calibration module for LARSIM, based on the Shuffled Complex Evolution Algorithm (SCE-UA) (Deubler, 2017), which is widely used in hydrology for model calibration (Seong, Her, & Benham, 2015). In this module, it is possible to define the calibration and validation time windows, as well as the calibration parameters and ranges.

The final hydraulic module includes the linkage to the Hydro_AS-2D, a 2D fully dynamic model with unstructured grid, from which the flood inundation extents are generated. Other possible linkages include MIKE 11, MIKE Urban and HEC-RAS 2D.

In forecast modus, the FloodEvac tool generates ensembles by sampling from LARSIM parameters using a beta or a normal probability distribution function. The beta function can produce skewed shapes of the distribution function, and hence take into account asymmetric uncertainty parameter intervals around the calibrated parameter set when generating the ensembles.

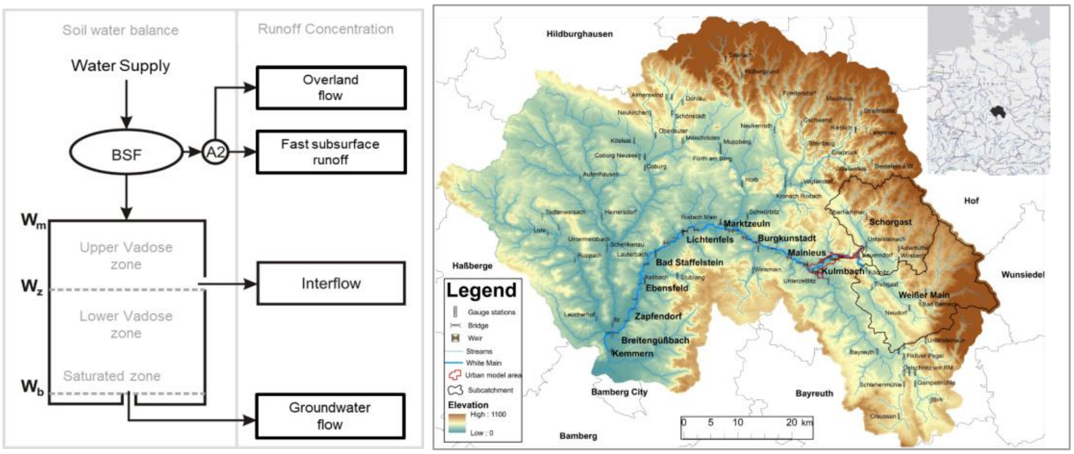

Figure 2: LARSIM water balance model (left) and Case study map of Upper-Main Catchment (right)

3 Case Study

The FloodEvac tool was applied to the catchment area of the Upper.Main located in Bavaria with a total area of 4244 km^2 (**Figure 2** right panel). In this work, we are hindcasting the flood discharge of December 2012-January 2013 to exemplify the forecasting tool. The Hydro-AS 2D model will be implemented at a later stage to use the forecast discharge and predict possible inundation. For the sake of simplicity, the actual DWD measured rainfall is used here as a forecast data. The considered flood event had two peaks on 24 and 28 December respectively. The discharge was recorded at the Ködnitz gauge situated just upstream of the city of Kulmbach. The two peak discharges were recorded as 64.2 m^3/s and 58.4 m^3/s on 24 and 28 December respectively. The uncertainty estimation in the forecast was done at two stages: a) rainfall (input) uncertainty and b) parameter uncertainty.

3.1 Generation of rainfall uncertainty

For the forecast, more than 50 points gauge rainfall station data were available at the catchment area, each having an hourly time step. The rainfall uncertainty module checks observed or forecast rainfall data at these stations and distribute the data within the whole catchment area considering sequential conditional geospatial simulation.

In this method, the variable is considered as normally distributed and continuous. However, using geospatial simulation in case of precipitation has some challenges as due to zero precipitation (no rainfall) at some stations, the distribution is not normal but positively skewed. Moreover, the spatial distribution of rainfall is not constant but varies temporarily. To adapt these issues, a suitable mixed distribution is considered here using two variants. The discrete part of the distribution is empirically recorded via the proportion of zeros in the total sample, and the continuous part is mapped on the three-parametric gamma distribution as well as on the basis of a nonparametric nuclear density distribution considering gammaMix and kdeMix. The whole geostatistical simulation is implemented using two different R-packages, namely gstat (Gomez-Hernandez & Journal, 1993; Pebesma, 2004) and RandomFields (Schalather, Malinowski, Menck, Oesting, & Strokorb, 2015).

The LARSIM model takes input from distributed rainfall data for the catchment at a spatial resolution of 1 km x 1 km. At this stage, 10 rainfall simulations were considered to estimate the uncertainty of the spatial rainfall distribution (**Figure 3**).

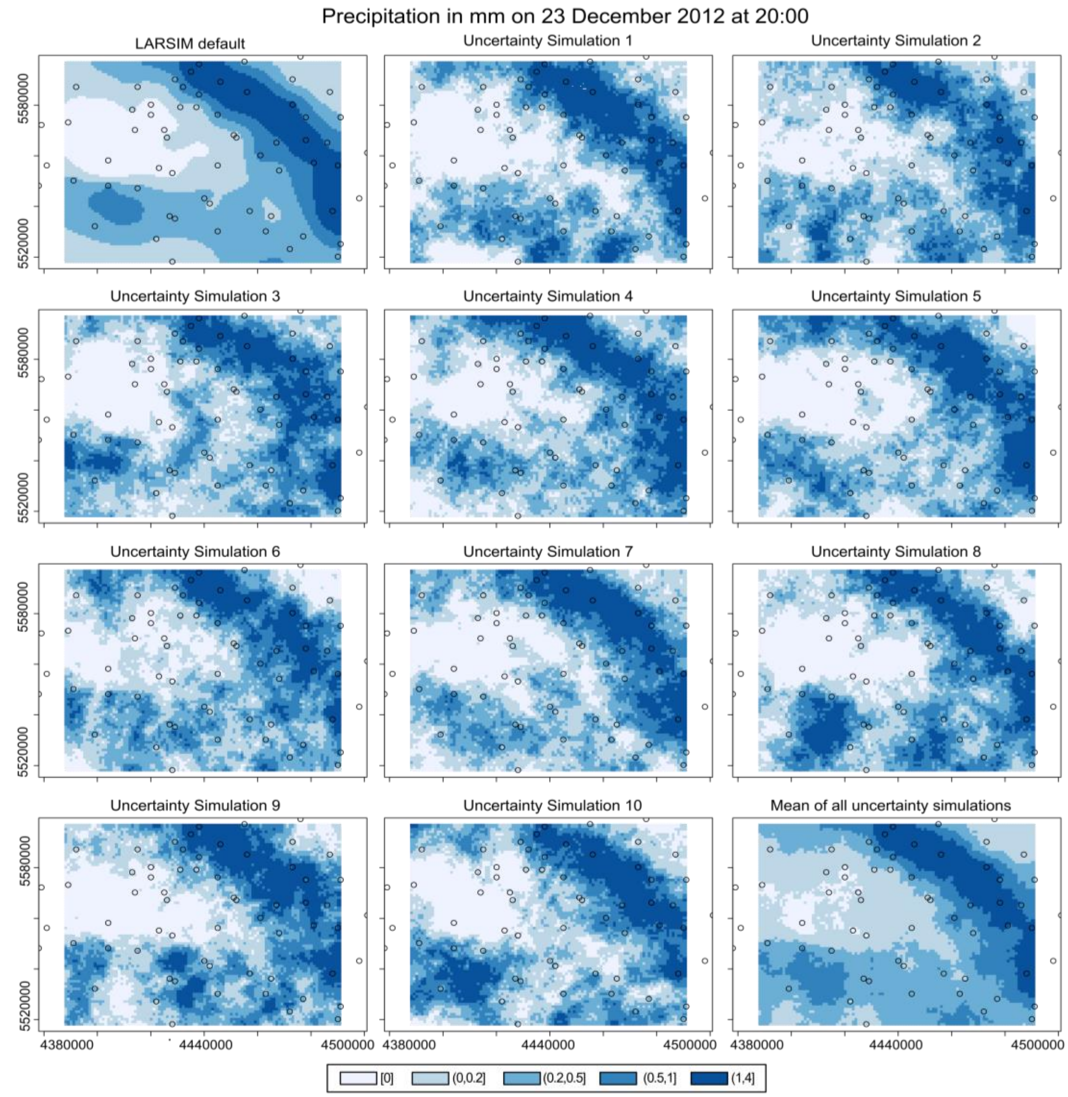

Figure 3: Uncertainty quantification of rainfall distribution

3.2 Generation of parameter uncertainty

A sensitivity analysis was performed in LARSIM to derive the most sensitive parameters of the model regarding flood discharge at the upstream gages of Kulmbach (Ködnitz and Kauerndorf). Eight parameters were identified as most sensitive such as *EQD*: the index for lateral drainage to the lower soil storage, *beta*: the shape parameter of the soil-moisture – saturated-areas function, followed by *BSF*, *EQB*, *EQI*, *EQD2* and *Dmax* (Härder, 2017). In this work, only these parameters were considered for uncertainty analysis out of the 34 parameters available in the LARSIM model. In a next step, the original model (Haag et al., 2016) currently in use by the Flood Forecast Centre at Bayerisches Landesamt für Umwelt (Bavarian Water Authorities) - LfU Bayern was compared with the one obtained using the automatic calibrated model. Since similar results were obtained between the original and the automated calibrated model, the original model was kept unchanged. The Upper.Main catchment area has 81 sub-catchments. Each sub-catchment has its own set of calibrated model parameters. Different sets of parameters were generated for each sub-catchments and applied accordingly for the forecast runs. The ensembles were generated using Monte Carlo method. The parameters and intervals were selected based on the sensitivity analysis of each sub-catchment. **Figure 4** shows the probability distribution curve for the 8 selected parameters of a sub-catchment in the model area.

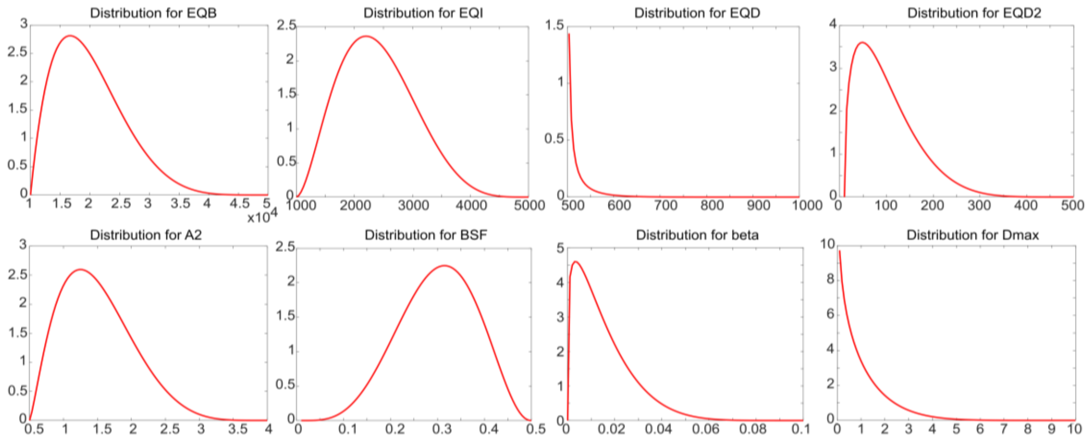

Figure 4: Probability distribution curve of the eight considered parameters at one sub-catchment

3.3 Forecast of flood discharge

Before the hindcasting process can start, a warm-up period is run. The model is run using observed precipitation and temperature data for one year of warm-up period until 49 hours before the forecast initialisation time. At this stage, the model uses the previously calibrated parameters. The model results are saved in an '*initial state file*' at the end of the one-year warm-up simulation time. As such, we assure that the internal model states condition of the basins are as close as possible to the real conditions. Later, this initialization state is used to simulate each forecast ensemble run. However, in case of starting the simulation from an '*initial state file*', it is recommended to start the flood forecasting for at least 49 hours ahead of the initial time. Therefore, each forecast ensemble simulation was run for 63 hours; the first 49 hours of simulation results are deleted and the last 12 hours of forecast data are stored. In this process, the model collects 49 hours of observed hourly rainfall followed by 12 hours of forecast rainfall data. These 63 hours of rainfall data are passed through the rainfall distribution uncertainty module and 10 different rainfall uncertainty dataset are prepared. Later, 50 different parameter sets are produced using parameter uncertainty module. These 50 parameter uncertainty sets are combined with the 10 rainfall uncertainty cases linking one rainfall uncertainty scenario with every five parameter uncertainty sets in a sequential order, making 50 sets of hydrological models for the Upper.Main catchment. These 50 models are run and results of discharge datasets are saved. The total run time for completing this whole process was around 25 minutes in a three core desktop in parallel mode.

As in this work, the actual measured rainfall data is used, which has a temporal resolution of one hour, the whole process was repeated at every one-hour interval. In the next forecast process, the model simulates new 50 cycles of 63 hours simulation, of which the first 48 hours are a repetition of previously done simulation; the 49th-hour data uses the latest available observed data and the last 12 hours data uses the new weather forecast. At this stage, the model uses the same parameter set which was generated at the first stage. **Figure 5** shows the results of all the simulations for the flood discharge forecast at Ködnitz gauge point.

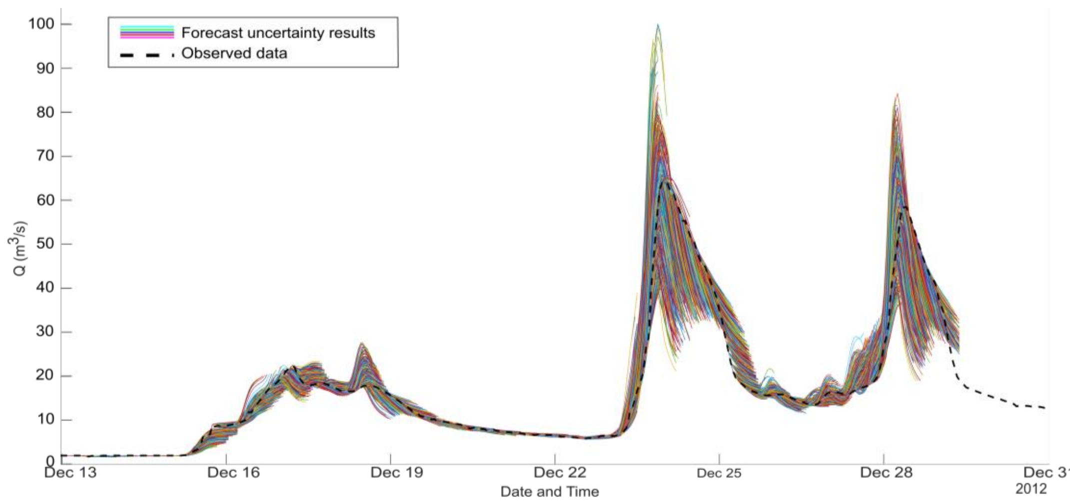

Figure 5: All forecast uncertainty results

4 Uncertainty analysis in the model results

The quality of the forecast data is assessed from each hourly lead of the forecast (**Figure 6**). It can be seen that the inconsistency between forecast and observed data increases with increasing forecast lead time. It is apparent that the model is very good for forecasting flood up to 4 hours in advance. The difference between forecast and observed data at 5 or 6 hours lead time is comparably better than forecast data with lead time of 7 to 12 hours. The deviation of simulated data from the observed is consistent at 9 to 12 hours lead time, which indicates that after a certain lead time, the error in forecast becomes stable and stops increasing.

In assessing the uncertainty of the basin response to spatial and temporal rainfall distribution, confidence intervals are calculated. According to the procedure described above, each temporal result was forecast 12 times using 50 ensembles of uncertainty runs. In this way, each temporal result is predicted 600 times, which are used for the statistical assessments for each time step.

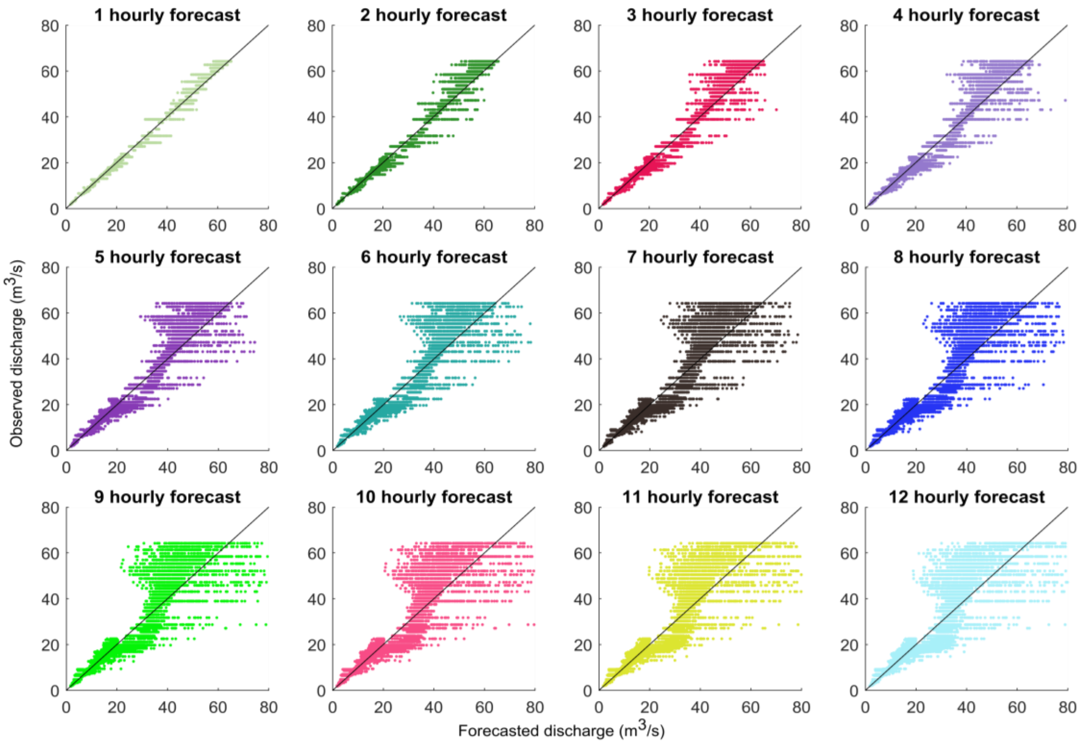

Figure 6: Scatter plots of observed vs simulated data for each hourly forecast leads

A confidence interval chart is shown for all the simulated forecast data (**Figure 7**). We calculate the confidence intervals of the whole forecast data in two parts: considering a forecast lead time up to 6 hours and 12 hours. Comparing with the observed flood discharge of the area, it can be seen that the model can forecast the rising limb of the flood peak fairly well. Both of the rising limbs of the observed flood discharge lie within the 50 percentile of the simulated results at both 6 hours and 12 hours lead forecasts. The peak discharge predicted in the model simulation is slightly earlier than the actual flood peak time. However, the uncertainty interval of the model peak is found moderately high. The falling limb of the flood discharge is found within the 98 percent confidence interval of the simulated discharges. The uncertainty is considerably lower at 6 hours lead forecast than that of 12 hours lead. However, all these analyses represent only one flood event. Analysing more events in a similar way would give a better representation of the uncertainty in the model results.

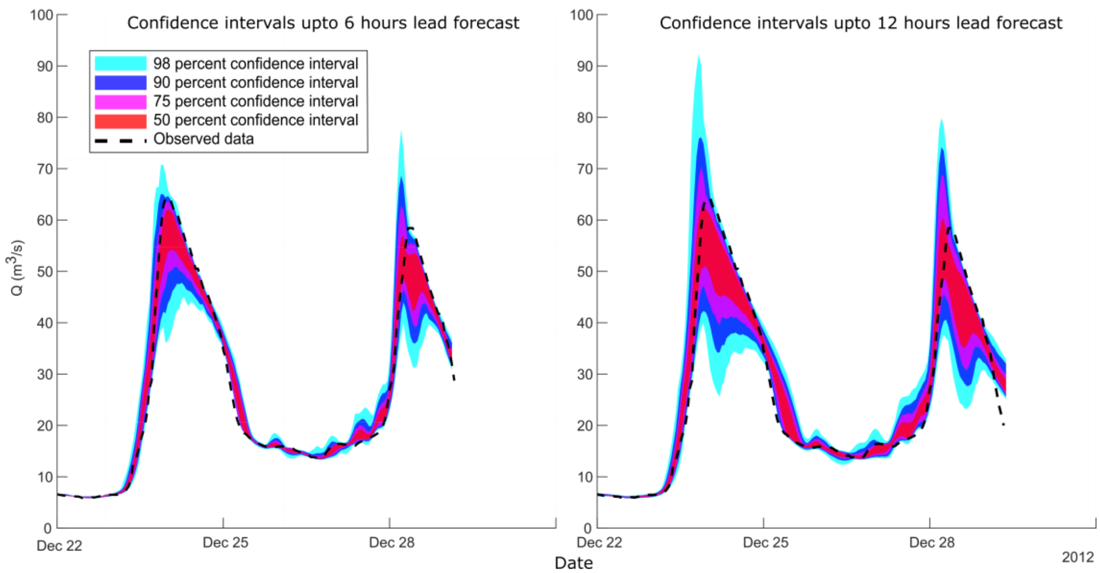

Figure 7: Confidence intervals in flood forecasting

5 Towards reducing uncertainty in the forecast

This work was intended to check the effectiveness of the FloodEvac tool in real-time flood forecasting. The rainfall forecast is available at every hour. Considering this fact, it was intended to obtain forecast results within a considerable fraction of an hour. Due to available computational resources, this forecast was done using 50 uncertainty runs only; which took around 25 minutes. The forecast quality might have been improved if more uncertainty runs were used. In this section, we propose another option to produce similar uncertainty bands of forecast within a shorter time period.

In this new proposed method, a pair of uncertainty ensembles is chosen within some predefined discharge intervals. Each possible pairs of uncertainty results are investigated to choose one pair of ensemble results that contains the maximum numbers of observation points within them. Later, the user can consider using the parameters used in preparing those two ensembles and regenerate new sets of normal and/or beta distribution curves and therefore predict new sets of hydrological model parameters. As the model receives observed discharge data at the time of forecast initialization, the parameter ranges can be defined considering that discharge observed at that time.

In this work, we divide the observed discharge into seven segments with an interval of 10 m^3/s starting from 0 discharge. One pair of ensemble is chosen for each segment based on the observed discharge at the forecast initiation time considering that the pair bounds the maximum numbers of observed discharge data. The observed discharge along with the results of these seven ensemble pairs are plotted in **Figure 8**.

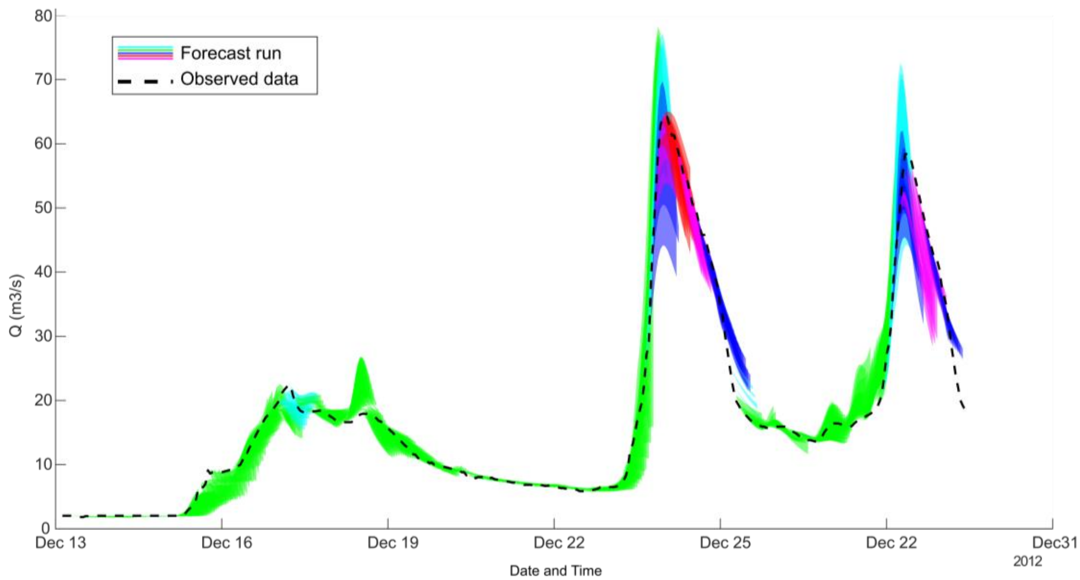

Figure 8: Observed discharge with updated forecasts

In this proposed forecasting method, only one pair of uncertainty simulation run is required to predict the flood forecast within a given flood range. However, this method requires prior knowledge of the hydrological response in the catchment area before actual forecast, by running a large number of uncertainty based hindcast simulations. Moreover, in this work, the proposed methodology is analysed and applied to only a single event. It has to be compared, tested and applied to more flood events in order to find the best pair of parameter sets for different flood magnitudes. This will be done in the future works and the corresponding results will be compared.

6 Discussion/Conclusions

In this case study, the FloodEvac tool was able to produce real-time forecasts with uncertainty for the investigated flood event. It was possible to consider forecasts with a lower probability which may trigger warnings earlier to sensitive areas. The total number of ensembles which can be generated in forecasting mode is limited by the available computation time. In any case, a larger number of ensembles could be used depending on of available cores computation power.

References

Deletic, A., Dotto, C. B. S., McCarthy, D. T., Kleidorfer, M., Freni, G., Mannina, G., … Tait, S. (2012). Assessing uncertainties in urban drainage models. *Physics and Chemistry of the Earth*, *42–44*, 3–10. https://doi.org/10.1016/j.pce.2011.04.007

Deubler, M. (2017). *Der Shuffled-Complex-Evolution-Algorithmus (SCE-UA) - Anwendung für die Kalibrierung des Wasserhaushaltsmodells LARSIM am Beispiel des Einzugsgebiets Oberer Main.* Technische Universität München.

Gomez-Hernandez, J. J., & Journal, A. G. (1993). Joint Sequential Simulation of MultiGaussian

Fields. In A. Soares (Ed.), *Quantitative Geology and Geostatistics vol 5* (pp. 85–94). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-011-1739-5_8

Haag, I. M., A, J., M, S., & Bremicker. (2016). Guideline for the Calibration of LARSIM Water Balance Models for operational Application in Flood Forecasting; LARSIM-Entwicklergemeinschaft - Hochwasserzentralen. *LUBW, BLfU, LfU RP, HLNUG, BAFU*.

Härder, M. (2017). *Sensitivitätsanalyse des Wasserhaushaltsmodells LARSIM für das Einzugsgebiet des Oberen Mains,*. Technische Universität München.

Leandro, J., Leitão, J. P., & de Lima, J. L. M. P. (2013). Quantifying the uncertainty in the Soil Conservation Service flood hydrographs: a case study in the Azores Islands. *Journal of Flood Risk Management, Outstanding Paper Award Supported by The JBA Trust – Commended Paper*, *6*(3), 279–288.

Pebesma, E. J. (2004). Multivariable geostatistics in S: The gstat package. *Computers and Geosciences*, *30*(7), 683–691. https://doi.org/10.1016/j.cageo.2004.03.012

Schalather, M., Malinowski, A., Menck, P. J., Oesting, M., & Strokorb, K. (2015). Analysis, Simulation and Prediction of Multivariate Random Fields with Package RandomFields. *Journal of Statistical Software*, *63*(8), 1–25. https://doi.org/10.18637/jss.v063.i08

Seo, Y., Kim, S., & Singh, V. P. (2014). Assessment of Uncertainty in the Spatial Distribution of Rainfall Using Geostochastic Simulation. *Journal of Hydrologic Engineering*, *19*(5), 978–992. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000882

Seong, C., Her, Y., & Benham, B. L. (2015). Automatic Calibration Tool for Hydrologic Simulation Program-FORTRAN Using a Shuffled Complex Evolution Algorithm, 503–527. https://doi.org/10.3390/w7020503