Response to Referee 1 Comments on:

Probabilistic Flood Inundation Mapping through Copula Bayesian Multi-Modelling of Precipitation Products

Francisco Javier Gomez, Keighobad Jafarzadegan, Hamed Moftakhari, and Hamid Moradkhani

We appreciate the constructive comments provided by the reviewer as they have helped enhance the quality of our manuscript. Enclosed, please find the revised manuscript with all the necessary corrections implemented in accordance with the comments and suggestions in tracked and cleaned versions. The following are point-by-point answers, in blue color, to the comments of the first referee:

Referee #1 (RC1):

1. **The role of copula function and the choice for a specific function remains somewhat elusive.** For one, Section 2.2 could benefit from being more specific in the role of copula function in replacing the posterior distribution. How does the copula function relate to the prior distribution? In addition, the section would benefit from an introduction of the concept of a copula function in statistical terms so as to convince the reader that the separation of functional dependence and modeling of marginal distributions is appropriate.

   Response: We appreciate the reviewer's suggestion regarding the introduction of copula functions, which we consider is crucial for enhancing the understanding of the readers. To address this, we have made improvements to the readability of the paragraph that introduces copula functions. Additionally, we have incorporated the statistical concept of copulas within the BMA approach in Section 2.2, as suggested, and referenced the Cop-BMA methodological steps outlined by Madadgar and Moradkhani (2014) and Abbaszadeh et al. (2022a), which were employed in our study. The modified text and equations added reads as follows:

   "Copulas are functions in the unit cube, which can link multi-dimensional distributions to their one-dimensional marginals (Sklar, 1959), they provide a flexible and powerful tool for modelling the dependency structure between variables, regardless of their individual marginal distributions and model dependency. This is particularly valuable in scenarios where the relationships between variables are complex and may not follow a simple linear pattern. Cop-BMA modifies the BMA predictive distribution through relaxing the assumption on parametric posterior distribution $g(y^t|M_i^t, \sigma_i^2)$ replaced with a group of multivariate copula functions. Multiple copula functions have been applied to postprocess hydrological forecasts (Abbaszadeh et al., 2022a; He et al., 2018; Madadgar et al., 2014; Madadgar and Moradkhani, 2014), and are used in this study for the estimation of water surface elevation posterior distribution. Equation 1 is modified to incorporate copula functions replacing the posterior distribution $p(y^t|M_i^t, Y)$ following the procedure from Abbaszadeh et al. (2022a). Supported

by Sklar theorem copulas can express the joint behaviour among correlated variables through their marginal CDFs in equation 3.

$$P(x_1, \dots, x_n) = C[P(x_1, \dots, P(x_n)] = C(u_1, \dots u_n) \qquad (3)$$

where C is the Cumulative Distribution Function (CDF) of the copula and $P(x_i)$ is the marginal distribution of x_i denoted as u_i for the interval [0, 1]. Using the PDF of copula, the joint probability density function of the variables involved can be defined as follows:

$$p(x_1, \dots, x_n) = c(u_1, \dots, u_n) \prod_{i=i}^{n} p(x_i) \qquad (4)$$

The conditional probability distribution of x_1 given x_2 is defined in equation 5:

$$p(x_1 | x_2) = \frac{p(x_1, x_2)}{p(x_2)} \qquad (5)$$

Considering the copula joint probability from equation 4, the equation 5 can be expressed as:

$$p(x_1 | x_2) = \frac{p(x_1, x_2)}{p(x_2)} = \frac{c(u_1, u_2) \cdot p(x_1) \cdot p(x_2)}{p(x_2)} = c(u_1, u_2) \cdot p(x_1) \qquad (6)$$

Since u_1 and u_2 are the observations (y^t) and simulations (M_k^t) respectively, the posterior distribution in equation 1 is replaced with the conditional probability distribution from equation 6 as:

$$p(y^t | M_1^t, M_2^t, \dots, M_k^t, Y) = \sum_{i=1}^{k} w_i \cdot p(y^t | M_i^k, Y) = \sum_{i=1}^{k} w_i \cdot c\left(u_{y^t}, u_{M_i^t}\right) \cdot p(y^t) \qquad (7)$$

Where $c(u_{y^t}, u_{M^t})$ represents the PDF of the copula function. To estimate weight w_i, it is required to maximize the log likelihood function of the vector of parameter $\theta = \{w_i, i = 1, \dots, k\}$ as:

$$l(\theta) = \log\left(\sum_{i=1}^{k} w_i \sum_{t=1}^{T} c\left(u_{y^t}, u_{M_i^t}\right) \cdot p(y^t)\right) \qquad (8)$$

Besides Section 2.2, the choice of copula in Section 4.1, line 329, is not convincing. It should be argued why only three distinct copula functions are evaluated and how that decision to restrict it to these three took place (one could imagine extreme value copulas or the Normal copula to be a candidate as well). For instance, whether there are certain characteristics that are sought in terms of attainable dependence, dimensionality, etc. Also, please detail the fitting process for evaluation on line 330 – it is unclear. The authors could use a copula cross-validation criterion, for instance, to make their choice tractable and not dependent on the rest of the model.

Response: We initially focused on the three most representative Archimedean copulas (Frank, Gumbel, Clayton) within our analysis, as they have been commonly used in previous hydrologic studies (Abbaszadeh et al., 2022; He et al., 2018). In response to the reviewer's comment, we expanded the scope of copulas considered to also include the Elliptic copulas (Gaussian and t-Student), as evaluated in the original COP-BMA paper (Madadgar & Moradkhani, 2014). To ensure robust copula selection, we employed the Akaike Information Criterion (AIC), which serves as a measure of the relative quality of statistical models. While this aspect was not previously mentioned in the article, we have incorporated it into the manuscript in Section 4.1 to provide details regarding our methodology. Furthermore, in accordance with the reviewer's suggestion, we conducted copula cross-validation using the criterion (xv-CIC) proposed by Grønneberg and Hjort (2014) employing a fold $k = 10$, yielding consistent conclusions with the AIC results for copula selection, which is also mentioned. Table S2 in supplementary data is added with results from criterion testing. We have updated the results section in the manuscript and included Table 4 to summarize the copulas selected per QPEs and cluster analyzed. The text reads as follows:

"In this study, five distinct copula functions are evaluated, Gumbel, Clayton, and Frank from the class of Archimedean copulas, and Gaussian and t-Student from Elliptical group. Copulas are constructed and evaluated using the marginals distributions of the observed data and each of the precipitation products modelling results of water surface elevation per cluster as $c(u_y, u_{M_k})$. Fitting and selection process was conducted using Akaike Information Criterion (AIC) and copula cross-validation criterion (xv-CIC) (Grønneberg and Hjort, 2014) using copula package implemented in R (Hofert et al., 2023), where the copula fit with lowest value of AIC and higher xv-CIC was selected. Table 4 shows the selected copulas for the seven QPEs evaluated in HR2D simulations over the three clusters. Calculated values for AIC and xv-CIC are presented in Table S2 in the supplementary material."

Table 4. Summary of copula fitting results per cluster for each precipitation product used in the HEC-RAS 2D model simulations.

Precipitation Product	Transitional cluster	Upper cluster	Coastal cluster
CMORPH	Gumbel	Gumbel	t-Student
Daymet	Gumbel	Gumbel	t-Student
ERA5	Gumbel	Gaussian	t-Student
IMERG	Gumbel	Gumbel	t-Student
NCEP Stage IV	Gumbel	Gumbel	t-Student
MRMS	Gumbel	Gumbel	t-Student
NLDAS	Gumbel	Gumbel	t-Student

2. **Validation:** A substantial part of the manuscript aims to validate the approach by comparing the model prediction with the inundation level at certain stations and comparing these against single data-source predictions. As the COP-BMA nests all other single models, it is no surprise that it outperforms the others by design. A better comparison might be not against the other models but against a non-copula BMA. An alternative would be to omit one data source (likely the best performing rain gauge data) and to make the horserace of the COP-BMA against that single model. This comparison has additional implication as it would be a remote sensing data vs. rain gauge data comparison.

Response: We agree that analyzing QPEs against rain gauges and observations is a more reasonable approach than initially implemented. This approach provides the basis for comparing remote sensing QPE products, their COP-BMA integration and rain gauge data. In response, we have made modifications to the manuscript to reflect this adjustment:

- Rain gauge data has been excluded from the Global Cop-BMA analysis, and this exclusion has been reflected in the updated version of Table 3.
- To summarize the results more effectively and enhance clarity, we have generated a new Figure 6 to compare time-series of average error to observations per cluster. This figure compares errors generated by individual QPEs, Global Cop-BMA, and rain gauge simulations. Discussion of these results has been updated accordingly in the manuscript in Section 4.1. The new Figure 6 and modified text reads as follows:
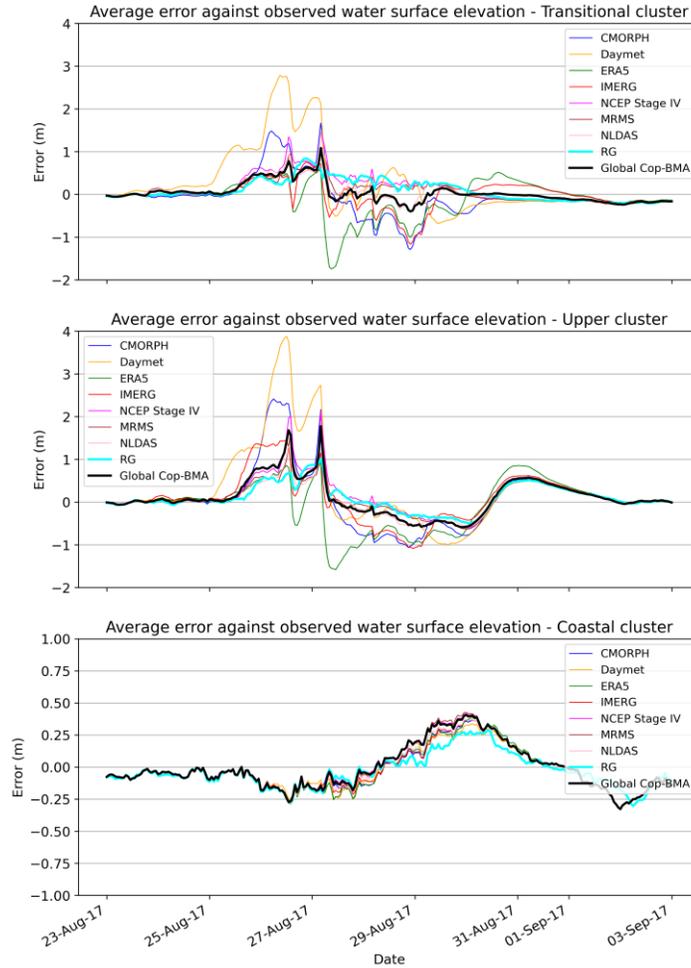
Figure 6. Averaged error time series of validation stations per cluster of simulated water surface elevation (WSE) by the HEC-RAS 2D model using QPE datasets and Global Cop-BMA approach results (black) against observed WSE during Hurricane Harvey.

"The averaged error of simulations using different QPEs against observations from validation stations are shown in Figure 6, featuring rain gauges simulation errors and estimations for Global Cop-BMA approach per cluster. Notably, this method exhibits better results in its responses to different precipitation products and clusters, leading to an enhanced accuracy in water level estimations, particularly during peak periods compared to the range of modelling water surface elevation outputs from the analyzed QPE such as Daymet or ERA5 which exhibit larger averaged errors. This demonstrates the Cop-BMA's capability to generate results that closely correspond to the observed values at the validation stations. It is important to highlight that if all models consistently overestimate or underestimate, Global Cop-BMA may not lead to significant improvement in the result (e.g., NOAA 8770613, USGS 08074710, USGS 08072050 in Figure 4; and Coastal cluster in Figure 6). Despite its advanced weighting mechanism, Global Cop-BMA's effectiveness relies on the diversity and accuracy of the model ensemble. Therefore, while it enhances the integration of diverse model outputs, its capability to improve results may be limited when all models exhibit similar differences compared to the observations at certain sections of the hydrograph".

The modification implies changes to evaluation of new weights distribution per cluster. We updated the weights discussion in Section 4.1 as follows:

"As depicted in Figure 7, the distribution of weights per cluster exhibits greater variability. In the transitional cluster, CMORPH, Daymet, ERA5 and NCEP Stage IV have weights below 0.1, as these four products generated underestimated responses in the hydrographs for most stations within this cluster. Weights center around the precipitation from MRMS, IMERG and NLDAS QPEs. Within the upper cluster, a different weights distribution among the QPEs is observable, with minimal influence from CMORPH, Daymet, MRMS and NLDAS QPEs. A higher difference is observable the three more dominant QPE, where NCEP Stage IV has a weight of 0.663 compared to the 0.144 of IMERG and 0.174 of ERA5 QPEs. For the Coastal cluster, precipitation from Stage IV QPE also holds the greatest weight (0.753) compared to the rest of QPEs which hold weight values below 0.1. Within this cluster, minimal discernible differences exist between QPEs water surface elevation results for the stations, as seen in Figure 4 (NOAA stations 8771013 and 8770613) and in Figure 6."

- Figures 7, 8, and 9 have been updated to reflect the new weights calculation after excluding rain gauge data in the Cop-BMA analysis and incorporating Elliptical copula functions within the fitting process as discussed previously. These updates ensure that the figures accurately represent the latest calculations and findings.

- Section 4.1 was complemented and modified to be consistent with the change suggested.

3. **Implication:** As already indicated in the previous point raised, the manuscript lacks the precise quantification or discussion of the proposed methodology's advantages and applications. For instance, staying with the case of Hurricane Harvey: what are the computational costs and runtime of the model? Could it be employed in risk assessment with rain-on-grid data forecasts? How much better is it compared to a non-COP BMA? While these questions are for illustration only, the discussion in the manuscript could aim to assess the methodology's benefit in a broader context.

   Response: Thank you for your comment. In response, we have made the following revisions to strengthen the implications and broader impacts of our study:

   - Section 5 has been edited to discuss the potential use of alternative precipitation forecast datasets and the application of our framework for operational forecasting purposes. This discussion expands the scope of our study and highlights its relevance for real-world applications and future areas of improvements. We have edited the text in the Discussions and Conclusions section as follows:

      "One advantage of our proposed framework is its flexibility allowing for the use of alternative precipitation products to enhance model simulations. For instance, this framework can be implemented for operational forecasting purposes where the Quantitative Precipitation Estimations (QPEs) utilized in this study can be replaced with Quantitative Precipitation Forecasts (QPF) from numerical weather prediction models such as High-Resolution Rapid Refresh (HRRR), North American Mesoscale Forecast System (NAM), Global Forecast System (GFS), European Centre for Medium-Range Weather Forecasts (ECMWF) among others".

- Computational resources and computational time required for the hydrodynamic modeling task have been addressed in Section 3.1, providing readers with insight into the practical aspects of implementing the proposed methodology. We consider the computational times (~7 hours) to be adequate for the application of the proposed methodology, and future research efforts will focus on utilizing the Linux version of HEC-RAS 2D on High-Performance Computing (HPC) systems. Added text reads as follows:

  "It is worth mentioning that our simulations were performed on a desktop computer with an Intel Core i7-7700 CPU @ 3.60GHz and 32GB RAM memory, averaging about seven hours per simulation for the time window."

- The improvements achieved using the Cop-BMA approach in terms of overall performance have been integrated into Section 4.1. This section provides a detailed analysis of the enhanced performance resulting from the application of the Cop-BMA methodology. The modified text in the manuscript describing the Cop-BMA results reads as follows:

  "Figure 8 provides a comprehensive overview of collective performance metrics of the HR2D model across the seven QPE simulations, rain gauges simulation, and the Global Cop-BMA multi-modelling for the seven QPEs evaluated at 30 validation stations over the 11-day simulation period. In general, the inundation modelling driven by different products consistently exhibits NSE performance with mean values ranging between 0.695 and 0.846. In terms of KGE performance, the interquartile ranges for QPEs display broader ranges, and the medians for Daymet and ERA5 products fall below 0.8 in contrast to other simulations.
  Notably, the Cop-BMA approach exhibits slightly higher performance metrics compared to the QPE products, NSE has an average of 0.858 and its total variability is lower compared to single precipitation products. KGE metric has a similar result with an average value of 0.852. The Averaged RMSE for Cop-BMA is 0.561m which is smaller than all the single QPE except for the rain gauges simulation which is only 3 centimetres lower. The averaged MBE for single QPEs ranged between -0.018 and 0.23m, while the Global Cop-BMA method results in an averaged value of 0.049m. Among individual products, the rain gauge outperforms all spatially distributed precipitation datasets and comes closest to matching the performance of Cop-BMA method. This highlights that reanalysis gridded precipitation products may have higher errors when compared to in-situ rain observations and allows Global Cop-BMA to generate QPE post-processed results that are closer to modelling results with observed precipitation from rain gauges. This methodology could be replicated in areas where measured precipitation is not available and obtain better performance metrics accounting for the uncertainties from this input".

Minor remarks:

- Line 52: please introduce the HEC-RAS 2D when first mentioning it, briefly.
  Response: Description of HEC-RAS 2D model was added in Section 1. Text reads as follows:

"Among various hydrodynamic models, the Hydrologic Engineering Center's River Analysis System (HEC-RAS) developed by the United States Army Corps of Engineers (USACE, 2022). It has the capability to simulate flooding conditions allows in both 1D and 2D."

- Line 83: please discuss the advantages and disadvantages of deterministic vs. probabilistic approach in this setting.
  Response: Description and discussion of deterministic and probabilistic approaches for flood mapping was added to the manuscript along with references from Di Baldassarre et al. (2010); Bates et al. (2004) and Merwade et al. (2008) to the readers interested in the topic. We have included your suggestion in the revised manuscript as follows:

  "On the deterministic front, the numerical results of the HEC-RAS 2D 6.3.1 hydrodynamic model, incorporating RoG, are evaluated to best describe the hydrodynamic behaviour of rivers, coastal and floodplain processes with a computationally affordable model. In parallel, a probabilistic approach is employed to use eight distinct precipitation products as forcing data to the hydrodynamic model to estimate an ensemble of flood extent and water depth in response to this hurricane-induced flood event. The deterministic approach provides a single representation of flood extents and depths based on predefined inputs and parameters, offering a clear understanding of potential inundation scenario evaluated. However, it fails to adequately capture the uncertainty associated with flood modelling, potentially leading to underestimation or overestimation of flood extents in other scenarios considering highly sensitive input parameters, which can impact the accuracy of results (Di Baldassarre et al., 2010; Bates et al., 2004).
  Probabilistic flood inundation mapping incorporates probabilistic techniques to assess and quantify uncertainty, providing a more comprehensive understanding of the range of potential flood outcomes and associated risks. It allows the integration of different datasets and input values, accommodating a wider range of initial and boundary conditions, and improving the robustness of flood predictions (Merwade et al., 2008; Di Baldassarre et al., 2010). Often this approach requires conducting numerous simulations to assess parameter uncertainty, leading to a substantial consumption of computational resources. Consequently, there is a preference for utilizing models that make substantial flow assumptions to conduct these simulations more efficiently and reduce computational costs."

- Line 123: please precisely state which variables are all subject of BMA.
  Response: The water surface elevation was the variable of interest. This information was added.
- Line 133: "In other worlds..."
  Response: Edited to "In summary".
- Line 148: Is the copula function bivariate? Please introduce the concept of a copula function here.
  Response: Copulas evaluated in this study are bivariate. Description of copula functions and formulation within BMA approach were added to the manuscript in Section 2.2 as mentioned in Comment 1.

Additional references included to manuscript:

Bates, P. D., Horritt, M. S., Aronica, G., and Beven, K.: Bayesian updating of flood inundation likelihoods conditioned on flood extent data, Hydrological Processes, 18, 3347–3370, https://doi.org/10.1002/hyp.1499, 2004.

Di Baldassarre, G., Schumann, G., Bates, P. D., Freer, J. E., and Beven, K. J.: Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches, Hydrological Sciences Journal, 55, 364–376, https://doi.org/10.1080/02626661003683389, 2010.

Grønneberg, S. and Hjort, N. L.: The Copula Information Criteria, Scandinavian J Statistics, 41, 436–459, https://doi.org/10.1111/sjos.12042, 2014.

Merwade, V., Olivera, F., Arabi, M., and Edleman, S.: Uncertainty in Flood Inundation Mapping: Current Issues and Future Directions, J. Hydrol. Eng., 13, 608–620, https://doi.org/10.1061/(ASCE)1084-0699(2008)13:7(608), 2008

Sklar, M.: Fonctions de répartition à N dimensions et leurs marges, Annales de l'ISUP, 229–231, 1959.

Thank you again for your constructive comments.

Sincerely,

Francisco J. Gomez, Corresponding Author
Center for Complex Hydrosystems Research (CCHR)
Department of Civil, Construction and Environmental Engineering
The University of Alabama