Responses to Reviewers
Manuscript hess-2021-229: "Applying Non-Parametric Bayesian Network to estimate monthly maximum river discharge: potential and challenges" Elisa Ragno et al.

Dear Editor,

Thank you very much for considering our manuscript for publication and for taking the time to review our work in detail. Below the answers to the comments provided by the reviewers in blue.

Best,

E. Ragno, on behalf of M. Hrachowitz, and O. Morales-Nápoles

Reviewer 1:

In the study, the authors mainly investigated the performance of the Non-Parametric Bayesian Network for the estimation of monthly maximum river discharge, and also discussed its challenges, with a case study in the 240 catchments in USA. Overall, the paper was rewritten well, and many details were clearly explained. However, there two main issues that should be clarified to further improve the quality of the paper before its submission.

First, the authors briefly explained the motivation of this study as: "very little attention has so far been given to explicitly representing the interdependence between inflow and outflow via probability functions". However, it is not clear enough, as there have been many methods used for describing the relationship among variables through probability functions. The key issue should be further explained very clearly to clarify the potential novelty of this study in Introduction.

We would like to thank the Reviewer for taking the time to review our work in detail. Non-parametric Bayesian Networks are tools for defining the joint distribution function of a set of variables. This joint distribution may be used to generate river discharge samples (or samples of any other variable in the model if required). The advantage of such a method is that the joint probability distribution is determined by defining the dependence between pair of variables. Such a non-parametric joint probability distribution is then more flexible compared to a theoretical parametric multivariate distribution because the dependence between variables is not fixed by the theoretical parametric model, but it depends on how the variables (nodes of the network) are connected to each other (arcs and parenting order). The dependence between pairs of variables can be determined based on prior knowledge of the underlying system dynamics, so it can be case dependent. As the Reviewer mentioned, other studies have implemented such methods, and references are provided in the manuscript – but applications in hydrology remain scarce. Here, we further explore the potential of such methods in providing estimates of river discharge by defining the joint distribution between environmental variables that are used for physical-based hydrological models and which are considered drivers of the discharge generation process. Then, we use the joint probability distribution to derive discharge via conditional probability, rather than characterizing the joint occurrence of the modelled variables, which is the most common implementation of multivariate distribution function.

In the revised manuscript, we have modified the *Introduction* to better highlight the novelty of the study and the main objective. Please refer to Lines 44-79 of the revised manuscript.

Second, there lacks "comparison discussion" between the NPBN-based results here and the previous studies in the study area. There have been so many studies in these catchments and others in USA. Without comparison, the advantages and challenges of the NPBN model cannot be easily understood. Thus I suggest adding some comparison contents to prove the advantages of the NPBN.

We agree with the Reviewer's comment that the lack of comparison of the NPBN-based results and previous studies makes difficult to appreciate the advantages and challenges of the proposed model. We would like to add also that the aim of this study was to provide a comprehensive analysis of the suitability of Non-Parametric Bayesian Networks (NPBNs) to derive river discharge via conditional probability given its several advantages in terms of the characteristics of NPBNs compared to other process-agnostic models.

Following the Reviewer suggestions, in the revise manuscript, we included the performances of other methods used for river discharge generation found in the literature

> Lines 353-368 – *"In a recent study, Ren et al. (2020) investigated the performances of regression models based on a variety of filter-based feature selection methods to estimate average monthly river discharge in three catchments from the CAMELS data set. The results obtained in terms of NSE ranged from ~0.6 to ~0.8, values similar to average (mean) performance of the network SN-1 (NSE ~0.596) here investigated for maximum river discharge. Kratzert et al. (2019) used CAMELS data set to evaluate the performances of hydrological models. They investigated the performances of Long Short Term Memory (LSTM) network to estimate daily river discharge in 530 catchments and included also, among other models, the performances of the Sacramento Soil Moisture Accounting (SAC-SMA) conceptual model. For the sake of discussion, we look at the performances of LSTM network without catchments attributes and SAC-SMA from Kratzert et al. (2019) for a subset of catchments also analysed in this study. Kratzert et al. (2019) results are available at https://github.com/kratzert/lstm_for_pub. The LSTM network without catchments attributes and the SAC-SMA conceptual model for daily river discharge have an average (mean) performance of NSE ~ 0.603 and 0.598 respectively. SN-1 network, here investigated, for monthly maximum river discharge has an average (mean) performance of NSE ~0.596. In general, NSEs obtained for simulations on a daily temporal scale tend to be lower than the ones on a monthly temporal scale due to the higher amount of observations over a common fixed period of time (Moriasi et al., 2007). However, other studies suggest that for both daily and monthly model simulations a satisfactory performance is given when 0.37<NSE<0.75 (Van Liew et al., 2007). "*

> Lines 514-518 – Conclusion: "*The NPBNs individually trained to specific catchments showed potential to reproduce monthly maximum river discharge in a wide range of environments with an average NSE of 0.59 (predictive models), while in the literature the performance of regression models for average monthly river discharge were NSE ~0.6 to ~0.8 (Ren et al., 2020), and the performances for daily river discharge were NSE of ~0.603 and ~0.598 for LSTM network and SAC-SMA model respectively (Kratzert et al., 2019) "*

Reviewer #2

This manuscript explores the application of a Non-Parametric Bayesian Network to estimate monthly maximum river discharge and its potentiality and challenges. The topic is important and would be of great interest to the readers of this field and falls within the scope of HESS. The paper has many grammatical errors and needs lots of editing. I did some of them in the abstract section, but it is not the role of reviewers to edit the full manuscript. The paper is not well organized. This reviewer wants to re-review the article after consideration given to the comments listed below.

We first would like to thank the Reviewer for taking the time to review this work in detail. Below, we addressed the comments received.

1. The authors ought to re-write the abstract so that it briefly presents the problem at hand, objectives of the study, methods used to achieve the objectives in a logical order before presenting a summary of major results and conclusions drawn from the study.

We are a bit surprised by this comment as the information mentioned is already included in the abstract, following the same order mentioned by the reviewer (the lines refer to the first version of the manuscript, version reviewed by the Reviewer):

Knowledge gap: Unclear whether Non-Parametric Bayesian Networks (NPBNs) are suitable tools to predict river discharge, as there are only very few studies using this method in hydrology (Line 2)

Objective: Explore here the potential of NPBNs to reproduce catchment-scale hydrological dynamics. (Lines 2-3)

Methods: 3 different Nonparametric Bayesian Networks (Unsaturated Network (UN-1) and Saturated Network (SN-1) with only hydro-meteorological variables and trained on one catchment; Saturated Network with hydro-meteorological variables and catchment properties (SN-C) and trained on all the catchments. (Lines 4-8)

Following the Reviewer suggestion, we have modified the abstract so that the knowledge gap, objective, and methods are more clearly highlighted in the order suggested. Please, refer to Lines 1-13 of the revised manuscript.

2. The introduction of the manuscript was very poorly written. The reason why you carried out this study does not seem to justify a publication. Try to highlight the regional or national significance of this study, especially since a lot of similar work has been done.

The main objective of this study is to explore and test the potential of Non-Parametric Bayesian Network (NPBN) to reproduce river discharge given its several potential advantages, e.g., the uncertainty quantification is embedded in the model, all the variables can be inferred via conditioning on the remaining variables, knowledge on the relationship between variables can be imposed a priori, information from different catchments can contribute to improve inference, and the computational time is limited. Hence, the significance of the study lies in the appraisal of this specific method rather than in a comparison of regional/national patterns of streamflow. The selection of the study basins, as dictated by the necessity of having a consistent and complete dataset of large number of catchments from diverse climate, then served as actual means to test the method using a large sample of study basins characterized by different environmental conditions. However, the main objective of this paper remains to test NPBNs for their suitability as tools/methods to estimate river discharge. We appreciate the comment of the Reviewer. We have clarify the objective of the manuscript as follows:

Lines 74-79: *Starting from these premises, the main objective of this study is to further explore and test the suitability of NPBNs as a tool to reproduce catchment-scale hydrological dynamics and to explore challenges involved when inferring monthly maximum discharge. More specifically, long-term data from 240 river catchments across the United States from the Catchment Attributes and Meteorology for Large-sample Studies (CAMELS, Newman et al. (2015) and Addor et al. (2017)) data set will be used as actual means to test the utility of NPBNs as descriptive models and to evaluate them as predictive models for monthly maximum river discharge considering the catchments individually and in group, to explore catchment similarity.*

3. The authors should also discuss other algorithms previously used by other researchers to predict the monthly maximum river discharge in the introduction section and explain why only NPBN were chosen for this study?

   As discussed in our response to the previous comment, the main objective of this study is to explore the potential of Non-Parametric Bayesian Network to reproduce river discharge. This was an a-priori decision based on the potential advantages of this method. In the introduction, we provided an overview of methods used in hydrology for generation of river discharge values and we divided these methods into process-based models and process-agnostic models.

   Highlight the key points of the paper, the innovative part of your work. What differentiates it from other works? Why should the journal publish it?

   We thank the Reviewer for this comment.
   The key point of the paper is to investigate the applicability of a fully probabilistic process-agnostic approach to predict river discharge generation. Given its several potential advantages, such as the uncertainty quantification embedded in the model (see line 68-73 of revised manuscript for details), we decided to test whether this type of probabilistic model, frequently used in other disciplines for risk and reliability assessment, could be implemented also for generating samples of river discharge. While investigating its suitability for river discharge characterization, we identified some benefits (e.g., embedded uncertainty quantification) and challenges (e.g., Gaussian assumption for bivariate dependence) and we reported them in the Discussion section to incentivize further studies.

   In the revised manuscript, we have addressed this comment in the *Introduction*. Please refer to Lines 44-79 in the revised manuscript.

4. The quality of the figures should be improved.

   Following the suggestion of the reviewer, we have improved the quality of the figures to increase their readability. At the same time, it was not very clear to us which aspects of the figures need improvement, e.g., resolution, colour codes, legends, content presented.

   We modified the bottom panels in Figure 5 to better show the difference in the performances of the networks UN-1 and SN-1. The histograms represent the ratio between the simulated and the observed quantiles for different quantiles value, i.e., 0.05, 0.50, and 0.95. The light grey histograms refer to the network UN-1 and the dark grey histograms refer to the network SN-1

   We also modified Figure 6 and now the colorbar of each map to improve readability.

5. Make your conclusion more clear and simplified. Highlight your important result or findings.

   We agree with the Reviewer's comment and we have modified the Discussion session to improve it readability. We have divided it in subsections to highlight the key points of discussion. Please refers to the section *"Discussion and Challenges"* , section 6, Lines 430-505

   Moreover, in the revised version of the manuscript we have also included the average performance of other river discharge methods found in the literature.

   Lines 353-368 – *"In a recent study, Ren et al. (2020) investigated the performances of regression models based on a variety of filter-based feature selection methods to estimate average monthly river discharge in three catchments from the CAMELS data set. The results obtained in terms of NSE ranged from ~0.6 to ~0.8, values similar to average (mean) performance of the network SN-1 (NSE ~0.596) here investigated for maximum river discharge. Kratzert et al. (2019) used CAMELS data set to evaluate the performances of hydrological models. They investigated the performances of Long Short Term Memory (LSTM) network to estimate daily river discharge in 530 catchments and included also, among other models, the performances of the Sacramento Soil Moisture Accounting (SAC-SMA) conceptual model. For the sake of discussion, we look at the performances of LSTM network without catchments attributes and SAC-SMA from Kratzert et al. (2019) for a subset of catchments also analysed in this study. Kratzert et al. (2019) results are available at https://github.com/kratzert/lstm_for_pub. The LSTM network without catchments attributes and the SAC-SMA conceptual model for daily river discharge have an average (mean) performance of NSE ~ 0.603 and 0.598 respectively. SN-1 network, here investigated, for monthly maximum river discharge has an average (mean) performance of NSE ~0.596. In general, NSEs obtained for simulations on a daily temporal scale tend to be lower than the ones on a monthly temporal scale due to the higher amount of observations over a common fixed period of time (Moriasi et al., 2007). However, other studies suggest that for both daily and monthly model simulations a satisfactory performance is given when 0.37<NSE<0.75 (Van Liew et al., 2007). "*

   Lines 514-518 – Conclusion: "*The NPBNs individually trained to specific catchments showed potential to reproduce monthly maximum river discharge in a wide range of environments with an average NSE of 0.59 (predictive models), while in the literature the performance of regression models for average monthly river discharge were NSE ~0.6 to ~0.8 (Ren et al., 2020), and the performances for daily river discharge were NSE of ~0.603 and ~0.598 for LSTM network and SAC-SMA model respectively (Kratzert et al., 2019) "*

6. Remove unnecessary "the" from the manuscript.

   We thank the reviewer for the suggestion. We have done a thorough grammar check to minimize the amount of placed articles in the revised manuscript.

   Specific comments:

We thank the reviewer for the specific comments. Below a summary of the changes made following the reviewer's suggestions.

Abstract:

In line 4 authors wrote UN and SN networks, and then in line 8, they said UN and SN models. Is there any difference between these two? If not, then please use only one for uniformity in the manuscript.P1 L2: "However, few hydrological applications can be found in the literature." This sentence doesn't fit after the prior sentence.

In this study, the Bayesian networks represent the numerical model used to determine the joint probability distribution of the hydro-meteorological variables and catchment characteristics, and then infer from it river discharge. We will replace models with networks to avoid confusion. Thanks for the suggestion.

L2 has been modified as:

Lines 2-3*: "Despite their several advantages, such as the embedded uncertainty quantification and the limited computational time required for the inference process, NPBNs' applications in hydrological studies are still scarce."*

P1 L2: Change "We therefore" to We, therefore," The change has been implemented.

P1 L4: Write the full form of CAMELS first before using its abbreviation. We have implemented the suggested comment.

Lines 5-8: *"Long-term data from 240 river catchments with contrasting climates across the United States from the Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) data set will be used as actual means to test the utility of NPBNs as descriptive models and to evaluate them as predictive models for monthly maximum river discharge."*

P1 L4: Change "one saturated" to "one saturated network" The change has been implemented

P1 L6: What is SN-C? SN-C is the name given to the network that estimates river discharge using also information from the catchments' attributes. We will clarify this in the revised version. We have revised the manuscript as follows:

Lines 11-13: *"Then, we analyse the performance of a saturated network (hereafter SN-C), consisting of the network SN-1 and including physical catchments attributes, to model a group of catchments and infer monthly maximum river discharge in ungauged basins based on the similarity of the attributes.*

P1 L6: Delete "but additionally" The change has implemented. Please, see the response to the comment P1 L6

P1 L8: Change "the attributes similarity" to "the similarity of the attributes" The change has implemented. Please, see the response to the comment P1 L6

P1 L10: Use "," after "analysed" Changed

P1 L14: Use "," after "catchments" Changed

P1 L15: Remove “,” before “once” Changed

P1 L15: Remove “,” after “discharge” Changed

P1 L16: Change “Despite these advantages, the result also suggest that there are considerable challenges in defining a suitable NPBN, in particular for predictions in ungauged basins.” to “Despite these advantages, the result also suggests considerable challenges in defining a suitable NPBN, particularly for predictions in ungauged basins.” The change has been implemented as follows:

Lines 22-23 *“[…]the results also suggest considerable challenges in defining a suitable NPBN, particularly for predictions in ungauged basins”*

Please make these changes in the abstract section and revise the whole manuscript for other English grammar and typing errors.

We appreciate the Reviewer’s suggestion and will pay extra attention in the revision of the manuscript to avoid grammar and typing errors.