Reply to the Comments by Reviewer 1:

Rebuttal of the manuscript entitled "Resolving the water budget of a complex carbonate basin in Central Italy with parsimonious modelling solutions"

We thank the Associate Editor and the reviewer for their valuable comments that, hopefully, will bring to a greatly enhanced manuscript. We have carefully read all the comments and provided point-to-point answers along with an indication of possible modifications to the manuscript. **Reviewers' comments** are in boldface. Underlined parts refer to changes in the revised manuscript.

Reviewer's Comments:

I read this manuscript with great interest, knowing the complexity of modeling karst systems and the unique and important role they play in hydrology.

Reviewer's Comment 1: However, in the current reading of the discussion paper, the novelty of the work is not clear beyond the study area. The introduction states that the study demonstrated that "good results can still be obtained by using few experimental data and time series analysis"(L66) but the research experiment itself does not follow a systematic approach that shows how lesser and lesser data still results in similar results. Is the goal of the study to show that less data can still result in good modeling of a karst system or that it is not necessary to fully couple the surface-water and groundwater systems in a karst system (L65)? The study design does not seem to follow a systematic testing for either approach; rather it applies the GEOframe-Newage tools to the study area. It would be more compelling to compare these results to how the system could be modeling with other couplings or with more (or less) data.

Authors' Reply:

We thank the reviewer for the valuable comment. The main novelty of this study is demonstrating that despite the lack of detailed groundwater (GW) information, yet, the modeling can be still conducted by extracting information from the observed hydrological time series within the basin and by relying on flexible modeling solutions.

Specifically, the novelties of our work are:

1. We were able to improve the appropriate goodness of fit indicators of simulated discharges of the karst system by using the obtained information from a new technique of correlation analysis between the rainfall and discharge time series.

2. We have proposed a new methodology based on the estimation of an Empirical Conditional Probability (EPC) which could provide information about the model reliability on top of the classical goodness of fit indicators.

The two previous achievements are methodological and can be applied to the other catchments, even for the non-carbonate ones.

However, the Nera catchment is an interesting study area by itself and is a good representative of the carbonate basins behavior. Therefore, our research experiment is extendable to other carbonate areas across the Mediterranean region. The accurate assessment, that the reviewer asks for, is the dream of any researcher but clearly is not achievable because of the missing appropriate groundwater data sets. In this regard, addressing the challenge of groundwater data scarcity and trying to provide reliable results (in terms of modelled and simulated discharges) is necessary and is the main scope of our study.

Comment 1b: 'Reviewer asked the authors to specify clearly if the study scope is Showing that it is not necessary to fully couple the surface-water and groundwater systems in a karst system'

A: Our contribution does not aim at demonstrating that fully-physical distributed models are not needed in these kinds of complex basins, rather, we are going to show that other effective solutions can be implemented especially when data availability is low. In other words, considering the fact that fully-physical modeling needs more expensive long-term groundwater data for the model calibration and run, we demonstrate here that a solution is still possible with a more parsimonious and fewer data demanding approach (see Fig. 1). The Upper Nera River as the resource of multipurpose water supplies and as a representative of carbonate basins, with the problem of groundwater data shortage derived from climate change, is certainly of interest of many Mediterranean catchments that face similar challenges.

Comment 1c: Reviewer asked the authors to specify clearly if the study scope is Showing that less data can still result in good modeling of a karst system.'

A: We are not going to pursue the demonstration to which extent lesser and lesser data lead to similar results. During this study, we tried to show that in case of a lack of long-term groundwater data, still we are able to represent this complex interactions in a karst catchment in a reasonable way, if we know the response time of the groundwater system as deduced from the precipitation and discharge data and the extension of the external carbonate area. We actually we have also used the field work results of hydrogeologists to verify our hypothesis.

Reviewer's Comment 2: The contribution is made more difficult to understand because the experiment organization in L70-79 appears to read more as results than hypotheses about what the study will test. The lack of clear hypotheses makes it difficult to understand the broader contribution of

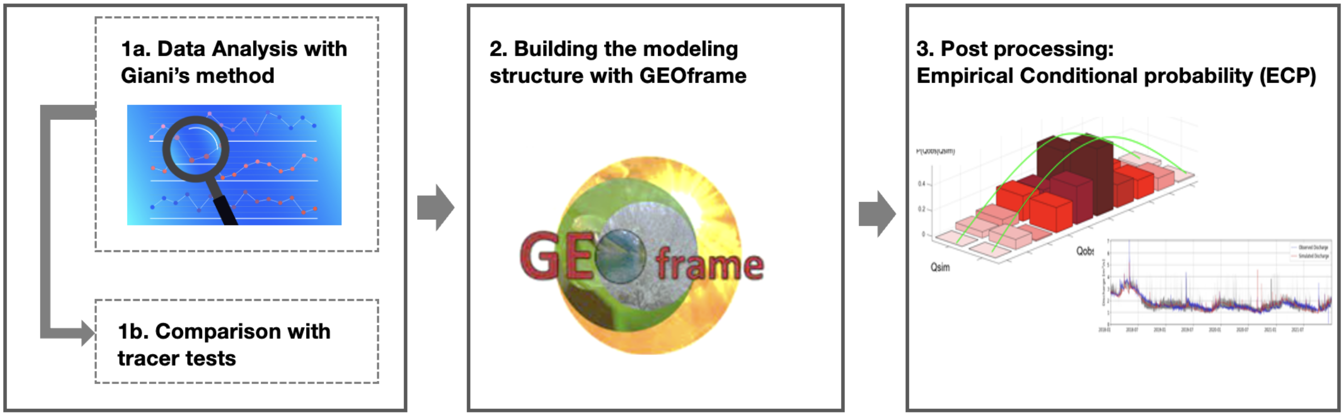

Figure 1. The proposed scheme in case of dealing with carbonate/karst systems without sufficient available groundwater data (the figure is for the reviewer and will not be added to the main text).

this work.

Authors' Reply:

We thank the reviewer for rising this up. We realized that the manuscript needs more work to define better our research hypothesis. Thus the text of those lines will be modified as follows:

"In this study, we are going to explore the following research questions:

1. Is it possible to model the complex carbonate catchment response to precipitation by relying only upon streamflow and precipitation time series? What kind of modelling solution is suitable to do that? And, is a parsimonious modeling solution suitable for that?

2. What is the impact of the external contributing area on streamflow in catchments characterized by fractured carbonate rocks behavior? And, to what extent does this contributing area impact the total streamflow from small headwater catchments to the main outlet?

3. What is the role of the storage for these types of catchments in sustaining streamflow during years of significant precipitation deficit?

We answer these questions on the Nera River basin, one of the main tributaries of the Tiber River (the second largest river in Italy) which contributes to almost 50% of its total discharge. Nera River is characterized by a significant portion of fissured and fractured carbonate rocks feeding the river discharge by releasing a large amount of groundwater into the riverbed by the streambed springs. Thus, this catchment is a good representative of the carbonate catchments to answer the three research questions.

Groundwater data shortage is a problem that is not unique to the Upper Nera River area, and the findings of this study could help inform water management and policy decisions in other carbonate basins as well. By providing a comprehensive analysis of the water cycle in this area, this study could also help identify potential sources of water stress and may suggest strategies to mitigate them. Overall, this research has the potential to make a significant contribution to the understanding and management of water resources in carbonate basins and help ensure the long-term sustainability of the ecosystems of the basin."

Reviewer's Comment 3

There are also quite a few qualitative statements that are not for the authors to decide about the quality of the modeling results. For example, on L234, the text states "these values are more than acceptable." It is not possible for the authors to make this assessment because they do not know what applications the readers may deem are "acceptable" - this is a qualitative statement based only on the authors' subjective assessment. The results should simply be reported and allow the reader to decide if these results are acceptable for their application or need. Another example is in L254, where the sentence reads, "It is apparent that the model is very good at reproducing the lowest discharges..." This should be changed to read something like, "The model is able to reproduce flows at the lowest discharges..." and then report or reference the accuracy at which the flows can be reproduced.

Authors' Reply:

This point will be considered in the revised version of the manuscript as much as possible. We will polish the manuscript from these ambiguous expressions to make them more scientifically sound.

Reviewer's Comment 4:

The results, interpretations, and discussion all relate specifically to the study area and there appears to be no further attempt to generalize or broaden the findings to the wider audience of HESS. There are also few stations used in the analysis, further limiting the interpretation of the results more widely. It would be helpful to frame these sections with a broader audience in mind beyond the study area.

Authors' Reply:

We do agree in this context with the reviewer. In this regard, we will highlight the following points which make the manuscript more suitable for a wider range of journal readers:

1) In the <u>modified Introduction</u>, we will describe different approaches to take account of external groundwater flux contributions to a basin together with the advantages and drawbacks of each approach. Furthermore, we will describe the challenges of the lumped modeling approaches under the circumstance of data scarcity and then our strategies to address these challenges will be discussed. This will allow the paper to reach a broader audience.

2) We will consider the role of the storage for these types of catchments in sustaining streamflow during years of significant precipitation deficit. Nowadays, the European drought issue is at the core attention

of many researchers, and investigating this issue for the long-memory hydrological catchments (like the Nera River basin) would be interesting. Therefore, Section 5.2 will be totally modified.

3) The study area (Nera River basin) is a good representative of the carbonate basin behavior and makes our research experiment scalable to other carbonate areas across the Mediterranean region. Besides the importance of knowing the discharge of the Upper Nera River as the resource of multipurpose water supplies, this area as a representative of carbonate basins with the problem of groundwater data shortage could be of interest to those dealing with the carbonate basins with similar challenges.

The Introduction will be rewritten as follows (subject to further revisions:

"**Introduction**:

Carbonate/karst landscapes represent approximately 7-12 percent of the Earth's continental area and they provide a significant challenge for hydrologists (Hartmann et al. (2014)). Due to the capability of these landscapes in retaining water for a longer period (i.e., long-term hydrological memory catchments), the storage has an important role in the control of drought propagation and delayed hydrological recovery (Alvarez-Garreton et al. (2021)).

Generally, a carbonate/karst landscape forms when the percolated precipitation dissolves the bedrock in the subterranean carbonate/karst environments and creates extensive fissures, open fractures, conduits, and caves. This can result in a complex network of groundwater flowpaths occurring within the same or adjacent aquifers (Kiraly et al., 1995). Therefore, the hydrological simulations are generally more complex in these areas where external groundwater flow is significant.

For modelling carbonate/karst basins, different approaches have been used. One powerful solution is using distributed, process-based models (e.g., Rooji, 2020; Hartmann et al., 2014), which are based upon groundwater partial differential equation solvers. Yet, the main challenge of this kind of distributed model is that they require a large amount of hydrogeological data and extensive field analysis to set appropriate physical parameter values and correct boundary conditions. On top of that, a large computational power is needed to run these models (Li et al., 2022).

Alternative to these models is black-box models also based on machine learning in which all the details about the structure of the aquifer and the hydrodynamics parameters are not needed (e.g., Tapoglou et al. (2014) and Castilla-Rho et al. (2015)). Although the implementation of these models is easy, specifically for the areas where information is lacking, the main drawback is that their model parameters do not have a physical meaning and just are indirectly related to the characteristics of the carbonated system (Zhou et al., 2019). Furthermore, these models do not explicitly solve the water budget and thus it is not possible to have information about the dynamic of water budget components.

The other types of model are those called lumped models, which are based on a set of ordinary differential equations (ODEs) that conceptualize the whole carbonated system as a system of reservoirs without modeling explicitly its spatial variability (e.g., Hartmann et al., 2014; Rimmer and Hartmann, 2012; Butscher and Huggenberger, 2008; Tritz et al., 2011; Jukic and Denić-Jukić, 2009; Butscher and Huggenberger, 2008; Duboisl et al., 2020). However, even in this case, the definition of correct model parameters relies on calibration and inverse modeling using monitored discharge data (Hartmann et al., 2014). In

particular, there is a considerable amount of lumped modeling studies (e.g., Rimmer and Hartmann, 2012; Dubois et al., 2020) in which the fast and slow drainage from the carbonate system have been modeled by using different types of reservoirs. The parameters of these reservoirs have been obtained through calibration or, indirectly via tracers information (i.e., an artificial tracer is introduced into a sinkhole, and traces are then sought far away from it in the surrounding areas at different times (Hartmann et al., 2014; Zhang et al., 2021; Nanni et al., 2020)). Despite being useful, this technique is time-demanding and cannot be always implementable because of inaccessibility issues.
Techniques that use the correlation between precipitation and discharge can be valuable alternatives to understanding the behavior of carbonate systems, especially where there is a lack of field information about the water circulation in the basin. For instance, Fiorillo and Doglioni (2010) carried out a cross-correlation analysis to cope with the time that water needs to flow through the fissured aquifers. A useful method also has been borrowed from applied economics (Kristoufek, 2014, 2015) by Giani et al. (2021) to successfully estimate the basin response time of the hydrographs concerning the precipitation. However, so far, the later data analysis technique has not been applied to the complex carbonate systems to obtain the hydrological response to the precipitation.

In this study we explore the following three research questions (RQs):

1. Is it possible to model the complex carbonate catchment response to precipitation by relying only upon streamflow and precipitation time series? What kind of modelling solution is suitable to do that? And, is the parsimonious modeling solution suitable for that?

2. What is the impact of the external contributing area on streamflow in catchments characterized by fractured carbonate rocks behavior? And, to what extent does this contributing area impact the total streamflow from small headwater catchments to the main outlet?

3. What is the role of the storage for these types of catchments in sustaining streamflow during years of significant precipitation deficit?

We answer these questions on the Nera River basin, one of the main tributaries of the Tiber River (the second largest river in Italy) which contributes to almost 50% of its total discharge. Nera River is characterized by a significant portion of fissured and fractured carbonate rocks feeding the river discharge by releasing a large amount of groundwater into the river bed by the streambed springs. Thus, this catchment is a good representative of the carbonate catchments to answer the three RQs.

Groundwater data shortage is a problem that is not unique to the Upper Nera River area, and the findings of this study could help to inform water management and policy decisions in other carbonate basins as well. By providing a comprehensive analysis of the water cycle in this area, this study could also help identify potential sources of water stress and suggest strategies to mitigate them. Overall, this research has the potential to make a significant contribution to the understanding and management of water resources in carbonate basins and could help ensure the long-term sustainability of these vital ecosystems."

Reviewer mentioned that "here are also few stations used in the analysis, further limiting the interpretation of the results more widely":

A: In fact, this is the main challenge that we want to solve (i.e., implementing the model over regions with data scarcity issues).

We will add some comments on that in the discussion section of the revised manuscript to remark this.

Reviewer's Comment 5

The conclusions make some interesting points, which actually do emphasize some of the potential novel aspects of the work but they are not emphasized in the manuscript elsewhere. For example, Conclusion #1 and the sentence on L356-357 discuss the insight that the classical approach for delineating basins is not appropriate and a preliminary check on the water balance is needed for karst system, especially if runoff coefficients are high. I am not sure of the novelty of this finding but this is a point that is noted in the title but then not mentioned again until the conclusions. The paper should be reframed with these contributions in mind. I will note again that I am not sure this will improve the novelty of the work but the conclusions are much more clearly stated as to the contribution of the work and it was unfortunate to wait until the end of the paper to understand the potential contributions of this work.

Authors' Reply:

Thanks for the comment leading us to reframe the manuscript and to make everything cleaner.

Other Issues:

Minor comments will be all answered in the revised manuscript. Additional modifications that will be added to the next manuscript are as follows:

1. Based on the reviewer's comments, the Abstract will be changed to accomplish their suggestions.
2. We will improve Fig. 6, 7, and 9 in the original manuscript by highlighting the proposed evaluation method (based on empirical conditional probability) demonstrating that the general classical scores are not enough to evaluate the models.
3. Based on the reviewer's comments, the Conclusions will be modified also in the manuscript.

References

Alvarez-Garreton, C., Boisier, J. P., Garreaud, R., Seibert, J., and Vis, M.: Progressive Water Deficits during Multiyear Droughts in Basins with Long Hydrological Memory in Chile., Hydrology and Earth System Sciences, 25, 429–446, 2021.

Butscher, C. and Huggenberger, P.: Intrinsic vulnerability assessment in karst areas: A numerical modeling approach., Water Resources Research, 44, 1–15, https://doi.org/http://dx.doi.org/10.1029/2007WR006277, 2008.

Castilla-Rho, J. C., Mariethoz, G., Rojas, R., Andersen, M. S., and Kelly, B. F.: An agent-based platform for simulating complex human–aquifer interactions in managed groundwater systems., Environmental Modelling and Software, 73, 305–323, https://doi.org/http://dx.doi.org/10.1016/j.envsoft.2015.08.018, 2015.

Dubois, E., Doummar, J., Pistre, S., and Larocque, M.: Calibration of a lumped karst system model and application to the Qachqouch karst spring (Lebanon) under climate change conditions., Journal of Hydrology, 24, 4275—-4290, 2020.

Duboisl, E., Doummar, J., Pistre, S., and Larocque1, M.: Calibration of a lumped karst system model and application to the Qachqouch karst spring (Lebanon) under climate change conditions., Hydrol. Earth Syst. Sci., 24, 4275–4290, 2020.

Fiorillo, F. and Doglioni, A.: The relation between karst spring discharge and rainfall by cross-correlation analysis (Campania, southern Italy)., Hydrogeology Journal volume, 18, 1881–1895, 2010.

Giani, G., Rico-Ramirez, M. A., and Woods, R. A.: A Practical, Objective, and Robust Technique to Directly Estimate Catchment Response Time., Water Resources Research, 57, https://doi.org/https://doi.org/10.1029/2020wr028201, 2021.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J., and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches., Reviews of Geophysics, 52, 218–242, 2014.

Jukic, D. and Denić-Jukić, V.: Groundwater balance estimation in karst by using a conceptual rainfall-runoff model., Journal of Hydrology, 373, 302–315, 2009.

Kiraly, L., Perrochet, P., and Rossier, Y.: Effect of the epikarst on the hydrograph of karst springs: a numerical approach., Bull. Centre d'Hydrogéol, 14, 199–220, 1995.

Kristoufek, L.: Detrending moving-average cross-correlation coefficient: Measuring cross-correlations between non-stationary series., Physica A: Statistical mechanics and its applications, 406, 169–175, 2014.

Kristoufek, L.: What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis., PLoS ONE, 28, 1–19, https://doi.org/10.1371/journal.pone. 0123923, 2015.

Li, J., Yuan, D., Zhang, F., Liu, J., and Ma, M.: A physically based distributed karst hydrological model (QMG model-V1.0) for flood simulations., Geosci. Model Dev., 15, 6581—-6600, 2022.

Nanni, T., Vivalda, P. M., Palpacelli, S., Marcellini, M., and Tazioli, A.: Groundwater circulation and earthquake-related changes in hydrogeological karst environments: a case study of the Sibillini Mountains (central Italy) involving artificial tracers., Hydrogeology Journal, 28, 2409–2428, https://doi.org/https://doi.org/10.1007/s10040-020-02207-w, 2020.

Rimmer, A. and Hartmann, A.: Simplified Conceptual Structures and Analytical Solutions for Groundwater Discharge Using Reservoir Equations., 2, 217–238, https://doi.org/http://dx.doi.org/10.5772/34803, 2012.

Rooji, R. D.: Towards Improved Numerical Modeling of Karst Aquifers: Coupling Turbulent Conduit Flow and Laminar Matrix Flow under Variably Saturated Conditions., Université de Neuchâtel, 583, 1–13, https://doi.org/https://doc.rero.ch/record/8817, 2020.

Tapoglou, E., Karatzas, G. P., Trichakis, I. C., and Varouchakis, E. A.: A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation., Journal of hydrology, 519, 3193–3203, https://doi.org/http://dx.doi.org/10.1016/j.jhydrol.2014.10.040, 2014.

Tritz, S., Guinot, V., and Jourde, H.: Modelling the Behaviour of a Karst System Catchment Using Non-Linear Hysteretic Conceptual Model., Journal of Hydrology, 397, 250–262, https://doi.org/http://dx.doi.org/10.1016/j.jhydrol.2010.12.001, 2011.

Zhang, z., Chen, X., Cheng, Q., and Soulsby, C.: Using Storage Selection (SAS) functions to understand flow paths and age distributions in contrasting karst groundwater systems., Journal of Hydrology, 602, 218–242, 2021.

Zhou, Q., Sing, V. P., Zhou, J., Chen, X., and Xiong, L.: Rainfall-runoff simulation in karst dominated areas based on a coupled conceptual hydrological model., Journal of Hydrology, 573, 524–533, 2019.