Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-463-SC1, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Hydrological signatures describing the translation of climate seasonality into streamflow seasonality" *by* Sebastian J. Gnann et al.

Ryan Teuling

ryan.teuling@wur.nl

Received and published: 10 November 2019

This review was prepared as part of graduate program course work at Wageningen University, and has been produced under supervision of dr Ryan Teuling by a student that prefers to stay anonymous. The review has been posted because of its good quality, and likely usefulness to the authors and editor. This review was not solicited by the journal.

Peer review on "Hydrological signatures describing the translation of climate seasonality into streamflow seasonality" by Gnann et al.

The manuscript "Hydrological signatures describing the translation of climate seasonality into streamflow seasonality" by Gnann et al. proposes two new hydrological signatures: the amplitude ratio and phase shift between the climatic forcing and the streamflow. The aim of this research is to use these signatures to quantify the catchment response to climatic forcing and use them for model evaluation. To determine the amplitude ratio and the phase shift, a sine function is fitted through both the climatic forcing and the streamflow. The climatic forcing is defined as the precipitation minus the potential evapotranspiration. The signatures are interpreted with the response (signatures) of linear reservoirs in series or parallel to climatic forcing. To test if the signature values are hydrologically interpretable, signatures for catchments in the UK and the US are defined and related to catchment characteristics to see if there is a pattern. Two models are discussed based on the signature range that they can produce. The authors conclude that the signatures can be used for model evaluation and to help model builders decide on the model configuration. The use of hydrological signatures to define a model configuration is a novelty, it would be interesting to look for other hydrological signatures and further investigate the abilities of this method. The phase shift is an interesting signature because it could quantify the time delay between climatic forcing and streamflow. However, my main concern is on the way the signatures are used here to evaluate models. The method is not appropriate, the model evaluation is not complete and no comparison is made with other evaluation methods. Furthermore, I also have some critical remarks on the proposed new signatures. They have a low accuracy and are not widely applicable. My last concern is about the conclusions, which are all based on visual interpretation instead of statistical analysis. Because of these reasons, I do not see the added value of this manuscript to the existing body of literature and therefore I recommend to reject the manuscript.

To start with, I will explain my main concern on the model evaluation using the proposed method. In the paper a new way of model evaluation is proposed, namely looking at range of values of signatures (phase shift and amplitude ratio) that different models can produce. To test how large the range of produced signatures by the models is, a Monte-

Carlo sampling experiment is done. The authors state that this new method could be more meaningful and fit-for-purpose than already existing model evaluation methods: "Signatures rooted in hydrological theory offer a potentially more meaningful and fit-for-purpose alternative to the typically used statistical metrics such as the Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe". I do not agree with this argument, I will discuss the flaws of this method in the next paragraph. First, I would like to raise attention to the fact that only two models are tested and no comparison is made with already existing model evaluation methods. I think a much more extensive approach is needed if they want to propose this as an alternative for the already existing model evaluation methods. More models need to be tested and the outcome of this evaluation method needs to be compared with outcomes of other model evaluation methods to see if they are in line and whether this method really gives more meaningful outcomes.

First of all, in the manuscript only two figures show the results of the model evaluation with this new method. These figures alone, are not enough to evaluate the two models. Quantitative statements on the model functioning are needed, i.e. how well does the model predict the streamflow? All conclusions are based on visual interpretation, but graphs can sometimes be misleading, statistical analysis would be much more appropriate to compare different models on their functioning. In this manuscript only two models are tested, but if a lot of models need to be tested, numbers would make it easier to tell which model is best instead of comparing a lot of graphs. Secondly, the choice of catchments used for model evaluation influences the outcome. For this experiment, 40 catchments in the UK are used. However, the UK catchments show better relationships between the signatures and catchment relationships (see figure 6 & 7), so the choice of using UK catchments instead of US catchments influences the outcome of testing this method. Thirdly, the number of parameters differs for the two models. Whether the difference between the signature space of the models is due to model functioning instead of the used range for different parameters, is questionable. My suspicion increases when reading line 481-483 "The actual reason...in Figure 2." and line 500-501 "Particularly the flow...than 60 days.", it seems that the signature out-

put is determined by the parameter range instead of the model functioning, so how will this method evaluate models in an objective way then? The conclusion that the signatures are a diagnostic tool because GR4J is not capable of modelling the whole signature space (Line 427-428), is thus not valid in my opinion!

Lastly, only a small part of model predictability aspects is evaluated. Pechlivanidis et al. (2011) summarized different model evaluation methods, where they discuss different objective approaches. Objective functions are here defined as numerical measures of the difference between the model simulated output and the observed (measured) catchment output. The Nash-Sutcliffe Efficiency (NSE) and Kling and Gupta Efficiency (KGE) are examples of this approach. The proposed method here is an objective function as well, since produced signatures by models are compared with observed signatures of catchments. The KGE has been introduced to overcome some limitations of NSE, this method analyses the correlation, the bias, and a measure of relative variability in the simulated and observed values (Gupta et al., 2009). This method evaluates thus more aspects of model functioning than the new proposed method here, which only gives an indication of the ability of a model to attenuate the climate forcing into a streamflow signal with a right time delay (if the signatures are correct!), but not if the model can produce the right streamflow variability and mean, peak and low flows.

The authors could improve the method by evaluating the models based on more hydrological signatures and quantify the model functioning. Furthermore, they could do test more models and compare the outcomes with other model evaluation methods. They could also improve the transparency of this method by adding a table with the changed parameters and the range. Finally, they should argue why although different parameters of the two models are changed, the model outcomes can still be compared.

My second major concern is about the signatures, they have a low accuracy and are not widely applicable. First, I will address the accuracy of the signatures. To determine the phase shift and amplitude ratio, a sine function is fitted on the climate forcing (P-ETp) and the streamflow. The method of fitting a sine function through the forcing and streamflow time series does not seem adequate to me. Most of the catchment regimes do not show a clear sinusoidal yearly cycle. This is well visible in the 16 different regime types, after Weingartner and Aschwanden (1992). For example, catchments that show two discharge peaks in one year cannot be described well by a sine function, this will lead to an error in the phase shift. In the paper two examples are given were a sinusoidal function is fitted on the climate forcing and streamflow. The timing of the sine function (phase) on the forcing and streamflow seems to be quite good in these cases. However, the sine function does not follow peak discharge and low discharges. This is clearly visible in the middle figure of the East Avon at Upavon catchment, the discharge peaks in 2001 and 2003 are not represented in the sine function (discharge is double the fitted discharge!), also the sine function does not follow the discharge in 2005/2006 when there is a low discharge. This shows that the sine fitting leads to errors in the amplitude ratio. Since the signatures are used for model evaluation, these errors could also lead to errors in the outcome of the model evaluation.

Furthermore, the use of the potential evapotranspiration (ETp) leads to errors (and thus lower accuracy) in the signatures for semi-arid and arid catchments, since the potential evapotranspiration deviates from the actual evapotranspiration in these areas. This problem is mentioned by the authors in line 380-385. This problem can be solved by including a model to estimate the actual evapotranspiration (also mentioned by the authors). This would also help interpreting the signatures with the catchment characteristics. For example, in line 436-438 the authors state that the signatures of the US catchments show a relation with the moisture index. This conclusion is made based on visual interpretation of figure 7a. But I think this conclusion is not valid because the signatures of the dry catchments on the left side have a large uncertainty because of the use of ETp instead of ETact.

The other disadvantage of the signatures is that they are not widely applicable. The problem of using the signatures for arid and semi-arid areas is already mentioned, but this could be solved by using the actual evapotranspiration. However, these signatures

are also not valid for catchments with precipitation falling as snow. Since catchments with precipitation as snow show a typically seasonal cycle, the need of leaving these out of consideration is a major lack of the proposed signatures. Furthermore, the signatures are also not valid for climates with a less distinct seasonal pattern, so this will further limit the applicability of the signatures. Because the signatures can only be used for a certain type of catchments, it is the question whether they contain new information on the streamflow seasonality of these catchments. There are already hydrological signatures that describe the response of streamflow to climatic forcing, for example the flow duration curve. A steep slope in the flow duration curve indicates a fast response of the streamflow to climate forcing whereas a flatter curve indicates a relatively damped response and higher storage (Yadav, 2007). Only the timing component might add new information, but since the method of determining the phase shift is not accurate, I do not see the added value.

The last thing I would like to point out is that all conclusions based on visual interpretation instead of statistical analysis. For the sine fitting method, I would like to see the goodness of it or the sum of squared errors (SS), to know how well the fit of the sine function to the climatic forcing and streamflow is. For the relationships between the signatures and catchment characteristics, it would be better to calculate the correlation coefficient instead of only the visual interpretation, since this might be misleading. The same goes for the model evaluation method, it would be nice to have a quantitative statement on how well the model works. This would also make it easier to compare more models, as mentioned before.

Minor issues and typo's:

Minor issue 1: Line 68-70 “All of these ... streamflow seasonality.” I am not convinced. For example the slope of the flow duration curve can say something about the translation of climatic forcing into streamflow seasonality. A steep slope in the flow duration curve indicates a fast response of the streamflow to precipitation inputs whereas a flatter curve indicates a relatively damped response and higher storage (Yadav, 2007).

Minor issue 2: Line 94-96 “The amplitude might ... seasonal component alone.” Why stating this if it is not done for this research, is it a follow up research suggestion? Then it should be placed in the discussion.

Minor issue 3: “catchment form” can better be replaced by catchment characteristics (For example in line 100). Catchment form suggest you are looking at the effect of a small river with a lot of branches or a stretched river.

Minor issue 4: The aim could be stated much more clearly, “test whether the seasonal signatures are useful for modelling practice (line 101)” not specific enough.

Minor issue 5: Line 110/111 “We use Ep ... would be needed.” Not a valid argument, how much would the uncertainty increase if you add another model?

Minor issue 6: In line 124 a small remark is made on the method of the sine fitting. This could be elaborated a bit more. Why use the sine fitting method? Which methods did you compare and why did you choose for the linear regression method (it is now in the supplement, but I think it is better to include it in the text)?

Minor issue 7: A reference is needed to support line 200 “The upper limit...shape parameter equal to 2.” Minor issue 8: About figure 3, could you explain the form of the curve when ?2 becomes larger and ?1 and fraction going to second reservoir are constant.

Minor issue 9: Line 235, explain the choice for Latin Hypercube sampling.

Minor issue 10: Table 1, add more information on range variables. For example for moisture index: -1= most arid and 1= most humid.

Minor issue 11: Figure 4: add color indication to description, climatic forcing (blue) and streamflow (orange).

Minor issue 12: Figure 5: Based on what criteria are the benchmark catchments chosen (grey dots)? Same goes for the two red dots, random or do they represent a certain

type of catchments?

Minor issue 13: Line 284, missing reference to table 1. Catchment attributes

Minor issue 14: Line 300, missing reference to table 1. Catchment attributes

Minor issue 15: Line 304-305 “Yet generally, ...in figure 6).” Statement is not explained in discussion, why are the US phase shift larger than for the UK catchments?

Minor issue 16: Figure 9b, Higher probability for high BFI for GR4J than IHACRES, but GR4J lower phase shift (max 60 days)!! Why? I would expect a larger phase shift when a larger part of the flow is slow flow.

Minor issue 17: Line 393-395 “Since the BFI... seasonal signatures.” I do not agree, the BFI cannot be used as a cause for observed patterns, but it can be related to the observed pattern. A higher base flow means more slow flow so this could be related to a larger phase shift.

Typo’s:

Line 17: sensitive Line 64: minimum Line 73: seasonality Line 278: reproduce Line 496: outputs

References:

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91

Pechlivanidis, I. G., Jackson, B. M., McIntyre, N. R., & Wheater, H. S. (2011). Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. Global NEST journal, 13(3), 193-214.

Weingartner, R., & Aschwanden, H. (1992). Discharge regime–the basis for the estimation of average flows. Hydrological Atlas of Switzerland, Plate, 5, 26.

Yadav, M., Wagener, T., and Gupta, H. (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Advanced Water Resources, 30, 1756–1774

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-463, 2019.