Response to Referee #1

Responses are written in blue.

The focus of this study is on seasonality of forcings (i.e., watershed inputs) and streamflow (i.e., outputs) and how the former is translated into the latter through watersheds functioning. To understand the role of watersheds in dampening of forcings seasonality, authors develop two signatures (namely, the amplitude ratio and the phase shift) and show how combinations of linear models result in certain values for these two signatures. Subsequently, they calculate values for the same signatures using data from several watersheds in the UK and US and overlay the results on top of linear model findings. In this way, they could devise a perceptual model for a given watershed, e.g., two parallel linear reservoirs show to be suitable to model streamflow in some catchment. Finally, authors assess two hydrologic models to figure out whether or not they could properly reproduce expected variations of these two signatures. This task helps evaluate structural adequacy of a given model. The paper is really well-written, and has high quality presentations. Because this research also provides theoretical foundations for the analyses in this paper, I consider it a great contribution. I believe that the proposed methodology has many applications in the field of watershed modeling and water resources management. Still, I have a few comments that are provided below, which might help improve the quality of this interesting manuscript. I would recommend minor revision.

We thank reviewer #1 for the helpful and encouraging feedback.

Comments: Maybe my most major comment is about similarity in concepts between this study and previous studies. Authors themselves also point out that several previous research have essentially relayed the same type of information, but maybe using different techniques (such as unit hydrograph, transit time distributions, etc.). I still do not completely understand what the benefits of the proposed method are, and this requires a dedicated section in the paper. Basically, any other quantitative tools that highlight the differences between the time series characteristics of inputs and outputs could be used here too. For example, we could simply use lag time between forcings and streamflow time series, or maybe variance of these time series, to investigate watershed functioning. For instance, if the ratio between normalized variance of inputs and outputs is really small, watershed might be groundwater dominated. Such a situation would be actually the case with low amplitude ratio under the proposed method. My question is, 'what makes this method unique or better in comparison to other methods?

Thank you for pointing that out. We have indeed pointed out similarities to other techniques, we however think that they do not necessarily relate to the same type of information. Transit times focus on the velocity of water particles and therefore yield different insights. Many other methods (unit hydrograph, lag time, variance of time series) focus on shorter time scales. We believe that the focus on seasonal dynamics can yield related yet additional information compared to methods focusing on event scales. Furthermore, we chose the approach because there are analytical solutions for how sine waves are propagated by linear systems. This allows for example to interpret the results in terms of configurations of linear reservoirs and to estimate their associated time constants. The suggested ratio of normalised variances will probably be related to the seasonal signatures, yet how exactly can such a number be interpreted beyond a qualitative statement like "this watershed might be groundwater dominated"? We will clarify the motivation for our approach in a revised manuscript.

Line 358-359: regarding limitations of this study, authors here mention that “In other climates with a less distinct seasonal pattern, or with two seasons per year our approach will not work”. I would argue that there are other limiations that need to be mentioned here too. For example, the proposed method requires quite long records of data.

From the SI it can be seen that 10 years are enough to obtain a robust result for most places. But of course, we require at least a couple of years (i.e. seasonal cycles) to meaningfully fit a sine curve. We will add a sentence about data limitations.

Authors claim that ‘inference from observed values of the signatures’ is a potential outcome of this method, but as I said, data is needed for this purpose, right?

The reviewer is correct that data is required for this purpose. We will clarify the sentence to make it clearer regarding what can be inferred from the signatures.

Moreover, most likely the method won’t work for sub-annual time scales (because there are lots of hydrological non-linearities at smaller time scales.

We agree with the reviewer here. We decided to focus on the annual time scale because it has a clear physical meaning (see lines 106-110) and because the seasonal flow regime is of importance to many applications. We will emphasise that in a revised manuscript.

Note that SI 1.4 briefly investigates non-linear reservoirs.

Maybe, elaborate on different limitation aspects of this research in a separate section.

We will add a discussion of the limitations you mentioned to Section 5.1 and change the title of that section. We think that another separate section on limitations might not necessarily be helpful. For example, we discuss the limitations of the modelling exercise in Section 5.4 (line 507-519), where we think it fits best.

Other minor comments: Line 125: explain how multiple linear regression method will be used. I haven’t seen any material so far that explains how linear regression could be useful.

We used multiple linear regression to fit sine waves to data. This is explained in SI 2.1.2. We will add a clearer reference to that in the text.

Line 546: ‘reduce the need for calibration’. . .I don’t think so. Maybe, signatures calculated in this research could be used as additional calibration metrics to improve the probability of getting the right answer for the right reasons. . .but not replacing the calibration process.

Once a certain arrangement of linear reservoirs is chosen, the signatures are associated with time constants of these reservoirs. For example, if we chose a model consisting of two reservoirs in series, the theory can be used to obtain the two time constants of the reservoirs. This might not replace the calibration process completely, but it could be used to limit parameter ranges or to fix certain parameters. Since we haven’t tested that yet, we can’t say whether that will be useful in practice. Yet in any case, as you have said, the signatures might be used as an additional calibration metric (which is also indicated by our modelling experiment). We will revise the paragraph to clarify this.

I have to say that, to me, the most interesting finding in this research is (lines 448-450: the attribute "fraction of highly productive fractured aquifers", which is a hydrogeological classification available for the UK, shows a much clearer pattern than any soil or geology attributes in the US.). This has great applications in model development for ungauged catchments.

Thank you. The question remains of how to get such a classification for other places than the UK.

Minor: Line 16: give a very brief meaning for the word ‘seasonality’. . .later you use terms such as ‘mean seasonal regime’ or ‘seasonal streamflow regime’ or ‘seasonal signatures’, which will make more sense if a clear description of seasonality is provided at the beginning

We will revise the first paragraph to clarify the meaning of the word seasonality.

Line 44-45: Shafii and Tolson (2015) is another reference that needs to be cited here

We will add that reference.

Line 73-74: this sentence is a bit unclear: ‘a signature describing how climate seasonality is translated into streamflow seasonaltiy adds a timing component with a focus on seasonal and thus slower dynamics.’

The obtained phase shift tells us how long – on average – the seasonal forcing peak is delayed before it becomes the seasonal streamflow peak. This time lag (e.g. 1 month) is what we mean by timing component. We will revise that sentence.

Line 237: please explain what you mean by ‘fast flow routing delay (1 to 5 days)’

We will add a more detailed description of the model parameters in the SI.

Thank you

Thank you for your review!