Dear Dr. Peleg,

We thank the reviewers and the commentators for acknowledging the value of our work, their feedback, and their constructive comments. We appreciate the wide range of inputs, which allowed us to enrich our introduction and discussion section. In addition to many useful references, the reviewers and commentators point out where and how the stochastic streamflow generator could be reformulated or extended. While each of the points risen is valid, most of them would make the model more complex, less flexible, and less generalizable. With the stochastic simulator presented in the manuscript, we aim at proposing a simple and flexible tool, which can be adapted to different contexts. This is facilitated by the provision of the simulation procedure as an R-package. In order to guarantee flexibility and generalizability, we combine phase randomization, which is a nonparametric approach for the generation of a time dependence structure, with the flexible four-parameter kappa distribution. Making the model more complex or more parametric would imply a loss in flexibility. We therefore would like to retain the main features of the model proposed. However, we agree that a more profound discussion of its limitations and potential extensions is necessary and valuable. We also agree that the issue with the replacement of negative values by zero values needs to be addressed and that the use of an empirical marginal distribution can in some cases be sufficient.

Below, we address the points risen by the two reviewers Ashish Sharma and Demetris Koutsoyiannis and state how we would like to address them in a revised version of the manuscript. Our replies to the reviewers' comments are written in blue and italic to distinct them from the reviewers' comments.

On the behalf of all co-authors,

Yours sincerely,

Manuela Brunner

Reviewer 1 (Ashish Sharma)

My congratulations to the authors on this excellent paper. Very glad to see a clever adopted to frequency domain alternatives in formulating a stochastic streamflow generator. My comments below are aimed to enhance the presentation and I am in support of publication once these have been addressed. Comments are:

Reply: *Thank you for acknowledging the value of our work and for the constructive comments, which help to enrich the introduction and discussion section.*

line 2/9 - The authors are missing the works by Keylock (10.1029/2012WR011923). This work performed resampling to an existing time series using phase randomization in the frequency domain. If I remember correctly, it had some nice inclusion of ICA to tackle the multivariate issue, and wavelets to get around nonstationarity in the data that cannot be handled using a fourier transformation alone. I think they need to read those papers (I am familiar with the above one but there may be more since) and acknowledge them here, and also try and show how their work distinguishes itself from the above paper.

Reply: *The work by Keylock (2007) indeed shows many parallels to the approach presented in this paper. His approach is not directly based on the Fourier transformation but rather based on the wavelet decomposition of a signal. Instead of the phases of the Fourier transform, the wavelet coefficients are (partly) randomized. The randomized series are then backtransformed to the time domain by using a rank-ordering procedure as* *presented in the approach used in our manuscript. Keylock (2012) later extended the procedure to the joint simulation at multiple sites. The work by Keylock will be acknowledged in the introduction and discussion section.*

line 3/21: I think the work by Mehrotra (10.1029/2005JD006637) should be acknowledged here as it represents essentially something analogous to a ARMAX type of a model even though it is cast as a stochastic downscaling approach. A mention should be made on the ability to preserve low frequency variability, which I believe the proposed approach will be able to address as well.
Reply: *The work by Mehrotra and Sharma (2006) will be acknowledged as an approach allowing for the extension of Markov chains to multiple sites by using spatially correlated random numbers.*

Line 3/35: Even though it relates to the problem of correcting systematic biases, given the use of phase transformation (not randomisation), the approaches of Nguyen should perhaps be acknowledged for completeness. The rationale behind these approaches and the one here has a lot in common. (10.1007/s00382-018-4191-6, 10.1016/j.jhydrol.2016.04.018).
Reply: *Thank you for pointing out these references. We will acknowledge the work of Nguyen et al. (2019) in the discussion section where we talk about options of how to improve the representation of the cross-correlation in simulated series.*

line 5/21: The authors may want to look through the details of (10.1007/s00382-018-4191-6, 10.1016/j.jhydrol.2016.04.018) as they performed another level of preprocessing - they fit a Thomas Feiring type model to the monthly data and after that structure was removed, the Fourier transformation was performed. This was done after trying with the steps referred to above, as it was found to exhibit clear advantages.
Reply: *We experimented with different types of deseasonalization techniques and found that the normalization at daily scale served the purpose of removing seasonality in the data well. Compared to using a Thomas-Fiering model, the approach used here is non-parametric and does not assume any temporal seasonality structure. Deseasonalizing by a Thomas-Fiering model and re-adding this seasonality at the end, might be valuable if the reproduction of the lag-1 autocorrelation was an issue, which was not the case here. However, it requires the fitting of a parametric model which is data dependent. Our routine works independent of the time resolution of the data and is easily adjustable to different contexts. We show that the ACF of the observed data is nicely preserved by the approach employed in our study.*

line 6/21: Setting negatives to zero is not a clean option. Please refer to the Keylock paper above again on how they restricted their approach to resampling to avoid having to set negatives to zero.
Reply: *We agree that setting negative values to zero is indeed not very elegant. We will change the algorithm in order to avoid this. Instead of replacing negative values by zero, we will replace these values by a value sampled from a uniform distribution in the interval [0, min(Q_obs_day)], where min(Q_obs_day) represents the minimum of the observed values corresponding to the day under consideration.*

line 11/10: Underestimation of cross-correlations is I think addressed well in (10.1007/s00382-018-4191-6). The trick that is used is to not randomly generate phases for all variables, but for a "key" variable (say biggest streamflow mean location). And then maintain the phase difference between alternate sites. The phase difference in space helps capture the cross-dependence attributes.
Reply: *The approach proposed by* Nguyen et al. (2019) *for a good representation of the cross-correlation between two or multiple time series in the context of bias correction could also be adopted in the stochastic simulation framework presented in our manuscript.*

The discussion section will be extended by the phase-difference correction functions introduced by Nguyen et al. (2019)*.*

Lastly, I feel not addressing the issue of non-stationarity in a stochastic generation paper under our present climate should be discouraged. The issue of nonstationarity can be addressed in the sense of a discussion by thinking of adding an exogenous predictor variable set in the formulation, which can impart the changes needed. Some discussion to that effect would be good to include in the paper before it is published.
Reply: *We agree that addressing non-stationarity, if present, is important. The manuscript therefore contains a note stating that the stochastic generator could be applied using discharge time series simulated with a hydrological model driven by meteorological data simulated with a GCM (and RCM) (p 13. L20-22 in the original manuscript). We will slightly extend the discussion by discussing more options of how to adjust the phase randomization approach to non-stationary conditions.*

Reviewer 2 (Demetris Koutsoyiannis)

1. The Technical Note by Brunner et al. (2019) implements a useful idea for easy stochastic simulation of daily streamflow, based on spectral representation and phase randomization. The method has several limitations (see below) but it is practical and useful, and it certainly deserves publication. I believe several issues can be improved before final publication and therefore I am providing some suggestions. I also appreciate the commentaries by Francesco Serinaldi, Ioannis Tsoukalas and Panayiotis Dimitriadis, who provided a lot of information to the authors. I think this information is useful to optimize their Note and also to put it in the context of modern and older literature, some of which is missing in the literature review. I believe that not everything suggested in the commentaries needs to be addressed, as this would change the orientation of the Note. However, with several changes in the formulations and a few expansions, rather than additional analyses, the Note could be improved. My own suggestions, which I list in the following points, fall in two categories: (a) recognition of the limitations of the method and (b) improvements in formulations, phraseology and terminology.
Reply: *Thank you for acknowledging the usefulness and practicality of our approach. We would like to retain the main characteristics of the stochastic simulation procedure proposed here, which make it a flexible and generalizable tool. However, we will extend the discussion section in order to discuss its limitations more in depth.*

2. A first limitation, which in the current version is not stated clearly, is the severe dependence of the method on the sample size of observations. The synthetic series has the same length as the observed series. The authors properly recognize the importance of respecting long-range dependence (LRD) in simulation. However, to study its effect in hydrosystems we need synthetic series much longer than the observed. The use of ensembles of small-length time series may not be equivalent with using a long time series as each member of the ensemble is independent from the others.
Reply: *Thank you for pointing out the need to address this limitation. We agree that producing an ensemble of time series of the same length as the observed one might not be equivalent to the generation of one very long time series if long-range dependence features are present which exceed the length of the observed series. However, such features cannot be generated anyway since the model is fitted based on a limited number of years of observations. We added this issue to the discussion.*

3. A second limitation is the absence of a model for time dependence. While the authors correctly adopt a model for the marginal distribution (e.g. they state "Using the empirical distribution instead of the Kappa distribution would prevent us from obtaining values that go beyond the range of observed data…") their method misses to do so for the dependence structure. The empirical autocorrelogram and periodogram are affected by significant bias and huge noise (see references provided by Panayiotis Dimitriadis) and if we do not use a model, then we reproduce a particular random realization, in terms of autocorrelogram and periodogram, in all our simulated series. I believe authors' statement "The periodogram, the empirical counterpart of the power spectrum, shows high values at those frequencies which correspond to strong periodic components" is only partly true and perhaps misleading. The periodogram could be regarded a realization of a stochastic process per se (on the frequency domain) and its peaks do not necessarily reflect a real peak in the "true" power spectrum. The same thing happens with the autocorrelogram. For example, the ups and downs in the empirical autocorrelograms in Fig. 5 may well be sampling artefacts, which we do not need to reproduce—but the method does reproduce them. If the authors have difficulty to accept my comment, I would suggest doing an experiment with a particular (smooth) autocorrelation function and see the ups and downs in the produced autocorrelogram and periodogram of a single realization.

Reply: *As correctly stated above, the approach brought forward here employs a nonparametric approach for the stochastic generation of the time dependence structure. In theory, one could try to find a suitable parametric model to represent this time dependence. We are convinced, however, that doing so would be far from straightforward since it would be very difficult to represent the complexity in the time dependence structure at different time scales ranging from short to long range. Furthermore, the choice of a model would be dependent on the catchment area. Besides its flexibility in reproducing dependence structures at different ranges, the approach presented here has the advantage of being applicable to any dataset of interest without having to fit a parametric time dependence model. The nonparametric time dependence model is flexible, applicable in any catchment, and easy to apply.*

4. A third limitation is the lack of parsimony of the entire methodology. From the statement "We fit a separate distribution for each day to take into account seasonal differences in the distribution of daily streamflow values" one can imagine that the overall method encompasses lots of parameters. Apparently, it is nowadays easy to do calculations with lots of parameters but, in my view, stochastics goes beyond calculations and algorithms. Parsimony in stochastic modelling is always important (see Koutsoyiannis 2016).

Reply: *We agree that the stochastic model presented in our study involves a fair amount of parameters. However, this is necessary if we would like to achieve a proper representation of the distributions of the daily discharge values in addition to being able to generate values outside the range of the observed values. If the user is satisfied with values within the observed values, he/she might forgo the use of a theoretical distribution and use the empirical distribution of the observed values for backtransformation instead. This option will be implemented in the new version of the R-package PRsim and the issue will be addressed in the revised version of the manuscript.*

5. A final limitation for the particular time scale of modelling, i.e. daily, is the lack of explicit modelling of time irreversibility (an issue also mentioned in the comment by Francesco Serinaldi). This would not be an issue if the time scale was monthly or longer, but I suspect that it is for the daily scale (see Koutsoyiannis 2019 and also Müller et al. 2017). I clarify here that I do not suggest changing the method to overcome the limitations (e.g. to become more parsimonious or to take irreversibility into account). Rather, I just recommend stating them in a clear and explicit manner.

Reply: *Thank you for pointing out the issue of time irreversibility. We agree that this aspect* *is not explicitly considered in the modeling strategy. Neither is it considered in most existing modeling approaches. The lack of consideration of the time irreversibility issue is mentioned in the discussion section.*

6. Now coming to the second category of my suggestions, I would recommend avoiding the name kappa distribution for the chosen distribution. It is true that in hydrological literature this name is in common use, but if we wish to facilitate communication with other disciplines, we should be aware that the name kappa distribution has another meaning in statistical thermodynamics—namely it is used to describe Cauchy-type (or Student-like) distributions in motion of particles (e.g. Olbert, 1968; Livadiotis and McComas 2013). The specific distribution used in the Note (which I do not think is a generalization of GEV as suggested by Ioannis Tsoukalas), is commonly (in most disciplines) referred to as the Dagum distribution—see https://en.wikipedia.org/wiki/Dagum_distribution. In addition, in terms of sign conventions in eqn. (4), I would suggest changing the signs of k and h and replacing the two minus signs in front of them with plus signs. This will make the expression more convenient and intuitive, and also complying to the standard notation used in other disciplines (e.g. as seen in the above web site).

Reply: *Thank you for highlighting that there was some confusion about the use of the term kappa distribution. The development of the name of the kappa distribution introduced by Hosking in 1994 has indeed an interesting history and there is a huge potential for confusion.* Mielke (1973) *introduced a three-parameter kappa distribution.* Hosking (1994) *generalized this distribution to a four-parameter distribution and called it "the four-parameter kappa distribution", which is a generalization of the generalized logistic, GEV, and generalized Pareto distributions. The Dagum distribution is also related to the three-parameter distribution by Mielke (1973) in the sense that it has the same properties but uses a different parameterization* (Kleiber, 2008)*. We here used the four-parameter kappa distribution by Hosking, which offers more flexibility compared to the three-parameter distribution. We retain the notation introduced by Hosking (1994) to stress the link to this original publication. We will add a remark to the text highlighting the link between the article by Hosking and the article by Mielke.*

7. The phrase "Stochastically generated time series mimic the characteristics of observed data and represent sets of plausible but as yet unobserved streamflow sequences" (my emphasis) may distort the meaning of what stochastic simulation is. It is not a matter of something that is "yet unobserved" but expected to be observed in the future. It is a matter of producing artificial "realizations" from the stochastic model. A model, stochastic or otherwise, is not identical to the real world.

Reply: *The corresponding sentence will be rephrased.*

8. The term "deseasonalization" needs to be used with care and clarification; otherwise it may mislead people to think that, by techniques like that used in the Note, we can get rid of seasonality. This, however, is quite difficult—if ever possible. With transformations of the time series, either linear (as in standardizing by mean and variance of each period) or nonlinear (as in fitting a separate distribution for each period, like what is done in this Note), we can only remove the seasonal effect on the marginal distribution, not that of the joint distribution of a cyclostationary stochastic process. (For example, differences in autocorrelation coefficients in different seasons are not removed by techniques such as the above mentioned). Therefore I suggest replacing "deseasonalization" with "deseasonalization of the marginal distribution."

Reply: *Thank you for suggesting this clarification. Deseasonalization will be replaced by deseasonalization of the marginal distribution.*

9. The notion of “nonparametric” techniques referred to in the literature review is, in my opinion, problematic when we deal with stochastic processes with time dependence. As opposite to iid statistics, in which the first “i" (independent) is taken for granted, in stochastics there cannot be “nonparametric” methods; something of parametric type is always present, albeit sometimes hidden. Furthermore, the “bootstrap approaches” also mentioned in the Note are unsuitable for stochastic processes as they distort the stochastic structure—particularly in the presence of LRD. Therefore, I suggest making these clarifications and limiting the references to such types of models (as well as to ARMA-type models whose value is only historical, I believe). Instead, I suggest extending the review to other models, more appropriate for hydrological applications, such as those suggested by other commenters.
Reply: *The term nonparametric has been used in the literature for certain types of models used for the generation of stochastic time series (Salas and Lee, 2010). We highlight that these approaches are not suitable for the reproduction of long-range dependence (see p:3, l:16-17). The disadvantages of the ARMA models are also clearly stated (see p:3, l:1-4). We will extend the literature review by more advanced models, which allow for more flexible time dependence structures (e.g.* Tsoukalas et al. (2018)*).*

10. Could the authors double check their equations? Is an imaginary unit missing somewhere in equation (2)? Could they correct the notation in eqn (3)? (Is ‘rand’ meant to be a subscript?).
Reply: *We checked the equations and there was indeed an imaginary unit missing (we misspelled j instead of i). The equations will be corrected in the revised version of the manuscript.*

11. Finally, I uphold the other commenters in congratulating the authors and I particularly second Panayiotis Dimitriadis in congratulating them for using modest phraseology. I would add in the reasons for congratulation the fact that they do not follow the cliches and fashionable paths: for example they limit their mentions to climate impacts and nonstationarity, a notion that has become a must in hydrological papers—often by authors who do not know what it actually is (see Koutsoyiannis and Montanari 2015; Serinaldi and Kilsby, 2015).
Reply: *Thank you.*

Selected comments from the comments by Francesco Serinaldi, Ioannis Tsoukalas, and Panayiotis Dimitriadis

The comments provided by the three commentators are highly appreciated and will be used to enrich the introduction, methods, and discussion sections. More specifically, we will address several approaches which could be employed to allow for an improved representation of the cross-correlation in simulated time series; we will address the problem of temporal asymmetry; we will add a comment on the SPARTA model by Tsoukalas et al. (2018) to the introduction; we will specify the estimation method used for the estimation of the parameters of the kappa distribution; we tested the Burr type XII (validation results see Figure 1) and generalized Gamma distributions, which were, however, not flexible enough to model the marginal distributions of the daily discharge values and led to unrealistically extreme high flows in the simulations and will therefore not be considered as alternatives to the kappa distribution.

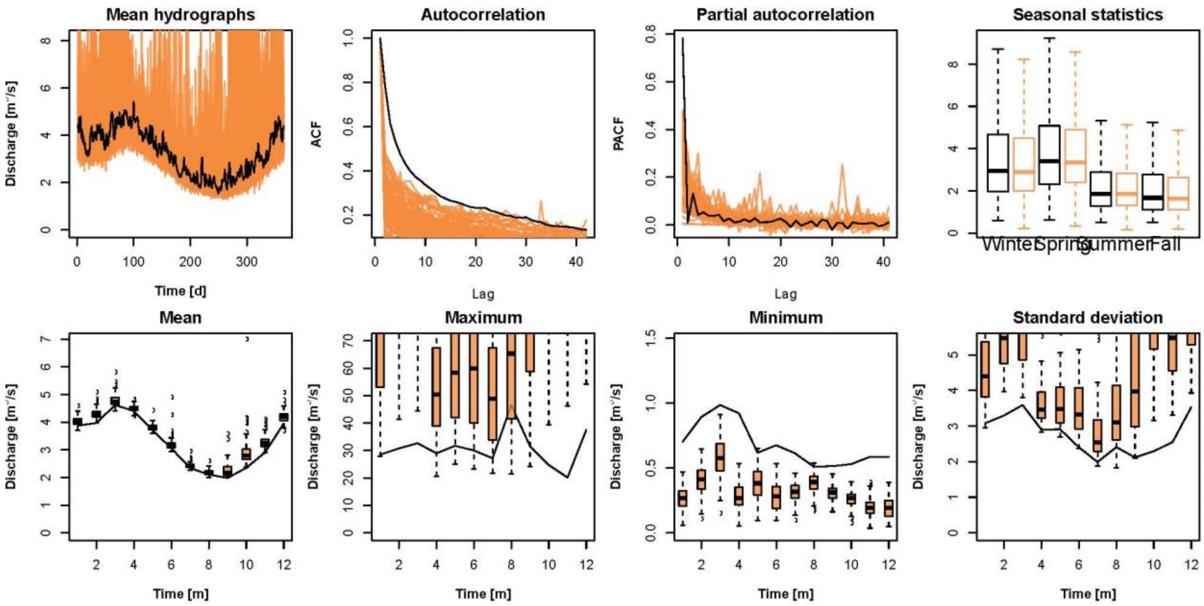

Figure 1: Validation plots for the Birse catchment on discharge time series generated using the Burr type XII distribution.

References used in the answers to the reviewers

Hosking, J.R.M., 1994. The four-parameter kappa distribution. IBM J. Res. Dev. 38, 251–258.

Keylock, C.J., 2012. A resampling method for generating synthetic hydrological time series with preservation of cross-correlative structure and higher-order properties. Water Resour. Res. 48, 1–18. https://doi.org/10.1029/2012WR011923

Keylock, C.J., 2007. A wavelet-based method for surrogate data generation. Phys. D Nonlinear Phenom. 225, 219–228. https://doi.org/10.1016/j.physd.2006.10.012

Kleiber, C., 2008. A auide to the Dagum distributions, in: Chotikapanich, D. (Ed.), Modeling Income Distributions and Lorenz Curves. Springer, Berlin-New York, pp. 97–117. https://doi.org/10.1007/978-0-387-72796-7_6

Mehrotra, R., Sharma, A., 2006. A nonparametric stochastic downscaling framework for daily rainfall at multiple locations. J. Geophys. Res. Atmos. 111, 1–16. https://doi.org/10.1029/2005JD006637

Mielke, P.W., 1973. Another family of distributions for describing and analyzing precipitation data. J. Appl. Meteorol. 12, 275–280.

Nguyen, H., Mehrotra, R., Sharma, A., 2019. Correcting systematic biases across multiple atmospheric variables in the frequency domain. Clim. Dyn. 52, 1283–1298. https://doi.org/10.1007/s00382-018-4191-6

Salas, J.D., Lee, T., 2010. Nonparametric simulation of single-site seasonal streamflows. J. Hydrol. Eng. 15, 284–296. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000189

Tsoukalas, I., Efstratiadis, A., Makropoulos, C., 2018. Stochastic periodic autoregressive to anything (SPARTA) modeling and simulation of cyclostationary processes with arbitrary marginal distributions. Water Resour. Res. 54, 161–185. https://doi.org/10.1111/j.1752-1688.1969.tb04897.x