

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/hess-2021-453-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-453

Anonymous Referee #1

Referee comment on "Stochastic daily rainfall generation on tropical islands with complex topography" by Lionel Benoit et al., Hydrol. Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/hess-2021-453-RC1>, 2021

The authors propose a new (I think) multisite daily precipitation generator, designed for use in tropical islands which are characterised by complex topography and associated orographic effects on rainfall, together with relatively stable prevailing weather directions. Their approach is to classify each day into one of several "rainfall types" which each has its own spatial joint distribution of rainfall.

Some of the ideas in the paper are potentially interesting. However, the existing literature on weather generators (particularly for daily precipitation) is vast: to justify yet another new approach therefore, it is necessary to demonstrate that it improves on existing methods in some way (it could be the performance of the method, its ease of implementation, its range of applicability, its computational feasibility etc.). The authors do not provide any such demonstration, perhaps because they don't seem aware of the state of the art in the area. Literature that seems particularly relevant includes that on Hidden Markov Models (e.g. Hughes et al. 1999; Ailliot et al. 2009): this uses the same basic idea of classifying each day on the basis of the joint spatial distribution of precipitation, but does so in what seems to be a more principled way than the present paper. Moreover, approaches using generalised linear models with topographical indices as covariates (e.g. Ambrosino et al. 2014; Chandler 2020) address the issue of topographical variability directly: it's not obvious to me that such approaches would fail in a tropical island setting. I am also surprised that the paper doesn't cite Maraun et al. (2010) which has become almost the canonical reference for anyone working in this area.

In view of the concerns above, as well as some technical issues (see detailed comments below), I don't think the paper merits publication in its current form. To make the case, the authors need to demonstrate that their approach improves on existing methods in some way as described above. Ideally, this would be done by carrying out an informed and fair comparison with a leading alternative method: if this isn't possible then the authors should explain why, and should offer some informed discussion of how their approach might reasonably be expected to compare.

These concerns are quite major. I therefore haven't worried too much about details such as the choice of meteorological covariates: there's no point in thinking about those until I'm convinced that the modelling approach is sound and necessary. Nonetheless, I have a few more detailed comments on the paper, that should be addressed in any revision. They are as follows:

- Line 31: this dismissal of "detailed physical modelling of rain generation processes" seems quite one-sided and poorly informed. It is true that stochastic models are computationally faster than physical ones, but physical models do have their own advantages which are not acknowledged here.
- Line 79: what proportion of the data have been "gap-filled"? How is the gap-filling distributed across time and between stations? How was the filling done? I will add that gap-filling in highly variable situations is, in my view, difficult and potentially dangerous because it will tend to underestimate variability. In my view therefore, for a stochastic rainfall model to be suitable for widespread use, it must be capable of handling incomplete datasets (in particular, what would you do if you needed to generate precipitation at a site for which you have no data? This requirement is common in many realistic applications). Lines 216-217 suggest that the proposed methodology cannot handle incomplete datasets: this is a serious limitation that needs to be acknowledged clearly and openly.
- Lines 101-105: although the spatial variation in precipitation statistics initially seems dramatic here, a more considered inspection reveals that the statistics are more or less constant along NW-SE transects and that the predominant variation is basically along the axis of the trade winds. If one were to scale the seasonal cycles at the specimen locations to a common scale (e.g. proportion of annual rainfall falling in each month), I doubt whether they would be dramatically different. I don't see anything here that really challenges state-of-the-art multisite weather generators, therefore – and hence nothing that really necessitates the development of a new modelling framework.
- Line 133: what's the justification for the formulation in equation (1)? It seems a bit ad hoc. Also, as defined, Z_i isn't latent because it's directly connected to observable quantities.
- Line 142: similarly, what's the justification for the distance-weighting in equation (2)? This again seems a bit arbitrary, although it's certainly true that what is sometimes called the "spatial intermittence" problem is hard to resolve satisfactorily. Nonetheless, in my view it's not clear that the authors' proposal improves on that by, say, Stehlik and Bárdossy (2002) – another example of literature that they seem unaware of, or at least haven't considered as carefully as they should have done. And once again, it's not a "latent" field if it's defined in terms of interpretable quantities.
- Lines 149-150: there's a claim here that the complex spatial distribution "prevents the use of a simple parametric form ... for the spatial copulas". It's not entirely clear to me what this means, but I assume it's something like: you can't find a "standard" spatial covariance model to represent the dependence structure between the $\{Z_i\}$. I'm prepared to believe this (although it could be partly an artefact of the artificial and deterministic distance-dependence structure of equation (2)), but it would be helpful to see some plots to justify it.
- Lines 158-159: the ability to calculate principal components of the "latent" field shows that it isn't latent. See earlier comments on this.
- Sec 2.2.3: I find this "rain typing" approach, which seems to be derived from techniques that are popular in the machine learning community, to be rather clumsy compared with the much more principled approach taken by other authors using HMMs. Even if one were to accept that the approach is worth considering, the classification in lines 172-174 is inappropriate because it fails to account for the uncertainty in each day's weather type (e.g. if you had three states then you could find that the

probabilities are always (0.34, 0.33, 0.33) in which case you would assign every day to state 1 which is clearly nonsense – exactly like the electoral system in some recently failed democracies).

- Line 179: there seems to be something wrong with this equation. If I understand correctly, it is something like a kernel estimate of the transition probabilities based on the Mahalanobis distance in the space of meteorological covariates; but if so, then summing the right-hand side over j should give 1 for any value of i or MC_d (it's supposed to define a conditional probability distribution). It doesn't, at least if γ_{ij} is a "baseline" transition probability (which isn't stated here, but seems to be the case based on a later description – if that's the case, then it's true that $\sum_j \gamma_{ij} = 1$ for all i but the exponential factor in line 179 messes things up).
- Line 184: how does the conditioning inform *seasonality*? This is another slightly dangerous assertion: relationships between precipitation and meteorological covariates can themselves be seasonally varying, for good physical reasons (think precipitation and temperature in temperate latitudes - they are positively associated in winter but negatively associated in summer).
- Lines 189-192: if the models are to be used for downscaling climate model projections, there are other requirements as well e.g. that the covariates are well represented by climate models and capture the climate change signal (see, for example, Maraun and Widmann 2018, Section 11.5).
- Lines 224 and 226: "GMM model" is a tautology. See also my previous comment about the allocation to the "most probable" state.
- Lines 232-233: I don't understand what you're doing here. In lines 218-220 you said that you already estimated the parameters of ψ : why are you doing it again, therefore? Also, any simplistic method of bandwidth selection such as equation (7) is almost guaranteed to fail in a large proportion of applications: the authors' use of such an approach suggests that they don't really understand the potential pitfalls of such an approach – or, indeed, the availability of alternatives.
- Line 244: here, for the first time (I think) we discover that separate empirical copulas are being estimated for each day. What's the basis / justification for this? Are you not just resampling the original data, with a bit of smoothing? [actually this is mentioned on line 351, but I think it should be acknowledged upfront in the methodology].
- Line 281: do you really believe there are 22 distinct rainfall regimes on the island? You've only got 86 stations, so this classification doesn't seem to be reducing the spatial dimension as much as one might hope. As a slightly peripheral (but important) comment: Figure 3(b) will be inaccessible to the $\sim 5\%$ of male readers who suffer from red-green colourblindness.
- Lines 321-326: although it's good to test the model using a variety of measures, I'm not always convinced by the way that this has been done here – and I don't fully understand all aspects of the plot. What are the grey bands in column (b)? What do you mean by "each statistic is estimated as the median across the 50 realizations"? [I ask this because you have dashed lines indicating the quantiles, suggesting that you're showing the distribution rather than the median – in any case, it isn't appropriate to compare a single observation to the median of 50 simulations because they will have different statistical properties even if the model is correct]. There are published papers that deal with this issue correctly: the authors need to familiarise themselves with the literature. The issue is particularly acute in the Q-Q plots of column (d): if you really understood what a Q-Q plot represents, you'd realise that you don't need to duplicate the observations 50 times (I hope you haven't used the median of the simulations here as well ...). Similar comments apply to Figure 5.
- Lines 340-341: I don't know what are the likely applications of precipitation modelling in this location, but if the potential stakeholders include farmers then I think they may be justifiably sceptical of your claim that underestimation of persistence can be tolerated for mathematical convenience. Precipitation modelling is not a mathematical exercise, it relates to lives and livelihoods.
- Lines 362-364: I'm not convinced by this dismissal of the curvature in the Q-Q plot of

Figure 5(b). It is well-known that by compressing the tails of both observed and simulated distributions, Q-Q plots can make it hard to see discrepancies in the tails that may have substantial implications in applications. It would be helpful to consider alternative approaches to visualising this particular comparison (e.g. my guess is that if you were to plot the observed and simulated densities of proportions of dry gauges then you would be a bit more concerned about the simulation performance).

- Line 388: in what sense is the model structure "hierarchical"? It may be worth noting that "hierarchical modelling" has a precise technical meaning in statistics: this isn't what you are doing here. It's probably worth rephrasing for avoidance of ambiguity, therefore.
- Line 392: here, the definition of weather types "based on rain features only" is offered as an apparent advantage of the proposed methodology. I would say that this is a distinct disadvantage because it ignores the physical processes that are operating. In my view, one of the key challenges in stochastic weather generation is to find mathematically tractable ways of capturing the footprints of the fundamental physical processes: this remark from the authors suggests that they either haven't thought about this, or that they consciously disagree with me. At the very least, some more explanation is needed as to why this feature may be considered desirable and defensible.
- Line 424: I disagree that the authors' model represents "a new tool". I think it represents a first step towards reinventing some existing tools (e.g. HMMs), but without acknowledging their existence (or perhaps not understanding what they're capable of).