Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-90-RC2, 2019

Geoscientific Model Development Discussions

Interactive comment on "Assessment of wavelet-based spatial verification by means of a stochastic precipitation model (wv_verif v0.1.0)" *by* Sebastian Buschow et al.

Joseph Bellier (Referee)

joseph.bellier@noaa.gov

Received and published: 24 May 2019

GENERAL COMMENTS:

This manuscript is the latest contribution to a field that has recently received attention in the verification literature: the use of wavelets for quantifying discrepancies in "texture" between high-resolution precipitation fields, as an alternative to verification tools that are affected by the double penalty (the classical point-wise measures), or to verification methods such as the object-based ones where discrimination capabilities can be found too dependent to intrinsic parameters. The authors expand the work presented in the previous related papers, by testing wavelet-based scores based on

different approaches for data reduction, but also by testing the discrimination capabilities of their scores on a controlled but realistic environment, leveraging an existing stochastic rainfall model.

Moreover, the authors investigate some aspects of the wavelet-based approach more in depth than in the previous papers, such as the choice of the mother wavelet, and the bias correction due the redundancy of the wavelet transform. I personally think that these latter investigations are equally interesting (compared to the development of the new scores), although the manuscript considers them as side results. I therefore encourage the authors to take the opportunity for expanding on these investigations, especially because I have found them missing in the previous related papers, and I am probably not the only one. I wouldn't require new experiments, but rather more in-depth discussions. See my specific comments below.

In overall, I have found the paper very well written, and clearly structured. The experiments are rigorous, the hypotheses are verified, and comparisons to existing alternatives in the literature are conducted. Moreover, the figures are of good quality.

However, I have found on several occasions that the introductions to the basic concepts of the approach were too brief. The consequence is that only a minority can read and understand the paper without having to read other related papers. I understand that the authors don't want to repeat things that are well explained elsewhere, but I would recommend a few more explanations on these concepts so that the paper is more self-sufficient.

All in all, I think this is a work of great quality, and I strongly recommend publication, pending minor corrections.

SPECIFIC COMMENTS:

P6 section 3: I have found the introduction to the wavelet theory pretty limited. I understand that you don't want to provide the mathematical details, but I think it is a difficult start for a reader whose knowledges about wavelets is short. I would therefore recommend that you to start this section with a few sentences presenting (in words) what are wavelets, why are they popular for analyzing signals (or fields in the 2D case), and to which references the reader could refer for a more detailed (and mathematical) introduction.

P7 l5-6: I have two comments here. First of all, you may clarify that the squared weights quantify the energy spectra. Indeed, at this point of the paper you have used several times the term “spectra” but have not defined it, and you haven’t used yet the term “energy”. Second, I think it would be nice here to briefly mention the origin of this bias (the redundancy of the wavelet transforms?), and also which form does it takes (the energy increasing over and over with increasing scales?). In my opinion, how does this bias really affect the energy spectrum is something that has been poorly explained in the previous LS2W papers that you cite, and it would be nice to let the reader know what should he expect in case he doesn’t apply the correction.

P7 l14-32: To my knowledge, your manuscript is the first (among the others having used the LS2W spectra for verifying precipitation fields) that investigates the choice of the mother wavelet. However, this paragraph is hard to grasp for a non-familiar reader, especially the differences between the wavelets. As a suggestion for improvement, I recommend that you add a figure of the plot of the different wavelets, so that it is easier to see the differences in terms of smoothness and support. You could also take the opportunity to refer to this figure at the very beginning of section 3, when introducing the mother wavelet function for the very first time. However, it is possible that the reader doesn’t understand how one can apply a 1D function (the wavelet) to a 2D field, so it might be necessary to explain the process in few words (apply on the rows, then on the columns, etc.).

P8 l13-15: In my opinion, the fact that the amount of negative energy averages out if we choose a wavelet smoother than D1 should not be introduced as “Preliminary experiments have shown that . . .”, but deserves a more detail paragraph and eventu-

ally supporting figures. Indeed, in my opinion, allowing negative energy is one of the biggest issue of the RDWT, so if you show that this problem vanishes by using other wavelets than the Haar wavelet, this is an important result, which should be discussed in more details.

P13 l15-19: It might it is necessary to give a little more explanation about the S component of the SAL (and perhaps a figure), so that your paper is self-sufficient. Moreover, you should say a few words about the ensemble version of SAL as well.

MINOR COMMENTS AND TECHNICAL CORRECTIONS:

p1 l19: “a given rain field is forecast perfectly, but slightly displaced”: If there is a displacement error, then the field is not perfectly forecast. You may replace “field” by “object” or “feature”.

p1 l23: After "four main strategies", the reader expects a descriptive list of each of these strategies. This is actually what you do in the paragraph that follows, but we have to wait until p2 l6 ("the last") to be sure that you are indeed referring to these four strategies. I recommend that you make the description more explicit.

p2 l17: remove the coma

p2 l22: You may briefly mention here the notion of "local stationarity".

p2 l25: As you write "corrected RDWT", you may later in the sentence say: "to obtain an unbiased estimate of the local wavelet spectra" (otherwise we don’t know why you need to correct).

p2 l33: It is not clear why does considering both the ensemble and the deterministic case "avoid the need for further data reduction".

P7 l7: Isn’t it “phi_{j,l,u}” instead of “phi_{j,j,u}”?

P7 l11: You say here that the smoothing is the final step of the spectra estimation procedure. However, in the package LS2W the smoothing takes place before the bias correction by the matrix A-1. Please clarify.

P7 l25-26: I don't understand what do you mean by a "sparse representation". Please clarify.

P7 l28-30: You have defined at l23 the labels "ExP" and "LeA", but these are not used until P17. Maybe you could use them here (instead of "this version of D4").

P8 l9: It may be nice to remind why the invariance under shift is necessary.

P9 l6-7: Maybe it would be better to replace "(i,j)" (and elsewhere where you refer to the coordinates) by other indices such as "(x,y)", to avoiding confusion with J referring to scales.

P9 l11: Please indicate the rationale behind the logarithm transformation.

P9 l25: This introduction to Fig 4a is confusing, because you say "as a function of the scale parameter", but when you look at the Figure you read "scale" for the x-axis, although the scale parameter you are referring to is in the y-axis. Please clarify.

P10 Fig 4: For plots (a) and (b), you may change the style of the black dash line, as we actually don't see the dash.

P12 Equation (6): I'm wondering if readers unfamiliar with the energy score might be confused with your definition. Indeed, the name "Energy Score" has here nothing to do with the "energy" of the spectra you are referring to, and this might be confusing with the fact that you define y and F as the observed and forecast (energy) spectrum, although in the general definition of the energy score, y and F are simply the observation and the forecast, no matter which quantity is being forecast. A more general definition of the score may reduce the risk of confusion.

P12 l5: You never mention clearly in this paragraph that the forecast quantity at hand is a multivariate vector. Even if you bold the observation y and the realizations X and X', you should make crystal clear that it is a multivariate quantity, and give the dimension.

P12 Table 1: Actually, some of your scores (Hemd and Hcd) work for both deterministic and probabilistic forecasts, so maybe you could modify you table by either adding a column "deterministic" that you fill with "yes" or "no", or by modifying the title of your current column and fill it by "deterministic", "probabilistic" or "deterministic and probabilistic".

P13, Variogram score: It is not clear whether you apply the variogram score to quantities that represent the wavelet spectrum or the precipitation field. From p13 l5, I understand that X, y and EF refer to the vector of the spectrum, but later you refer to spatial locations, so that I figure out that your forecast quantities are fields, is that correct? Please clarify. More generally, please clarify which scores are built from the wavelet approach, and which ones from the precipitation fields directly.

P14 l12: I would add "for ensemble forecasts" (after "the established alternatives"), to clarify why you don't consider the RMSE here.

P14 and 15, Fig 5 and 6: the energy score is here referred to as SpEn, although in the text you refer to Spe. Similarly, in Fig 7 you refer to Semd, although in the text you refer to SPemd. In addition to these corrections, I think it would be nice to use the subscripts in the Figures, so that it is fully consistent with the text.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-90, 2019.