

Interactive comment on “Multivariate stochastic bias corrections with optimal transport” by Yoann Robin et al.

M Muskulus (Referee)

michael.muskulus@ntnu.no

Received and published: 3 September 2018

Synopsis

The manuscript studies bias-correction methods for climate models (that are here considered as dynamical systems). A new method based on optimal transport theory is suggested: given results from two models X and Y and considering these as probability distributions, a joint probability law is determined from optimal transport theory that couples the two distributions for X and Y in a certain optimal sense (i.e., least work in transforming X into Y). The ensuing joint distribution can then be used to obtain stochastic corrections for data samples, by sampling from the conditional distribution of Y , given X (or vice versa) - which is an interesting idea. The authors then continue

C1

to suggest corrections for the case of observing two models at two different points in time. Here they propose to estimate the optimal transport from Y_0 to X_0 and from X_0 to X_1 . Together, these two joint distributions are used to obtain values of Y_1 from Y_0 , by adding random realizations of the differences between X_0 and X_1 to Y_0 , scaled by the covariance matrix between X_0 and Y_0 . Simulations with a nonstationary, perturbed variant of the Lorenz model show the potential for reconstruction of the (known) distribution of Y_1 . Results comparing two different climate models are somewhat less satisfactory and show that time evolution in both models seems to differ in some important aspects that are not captured too well by the method then. However, the correction method at single time instances still works well.

Assessment:

The manuscript is well written and the topic is suitable for HESS. The explanation of the proposed methods is well executed, being correct and relevant, while being concise and leaving out unnecessary mathematical details. All in all a manuscript that I enjoyed reading. It is actually quite thought-provoking, since the correction method proposed for the "nonstationary" case could also be done differently - maybe it would be worth to investigate these other options also? I can recommend the manuscript for publication, but I would like to invite the authors to comment on a few issues (see detailed comments below), mostly in order to further increase the relevance.

Detailed comments:

1. Equation 1: The authors use the quadratic Wasserstein distance. This has a number of nice theoretical properties, but in statistics the L1 Wasserstein distance (i.e. without the square, then known as the "Kantorovich-Rubinstein" distance) would be a more robust choice — although potentially loosing uniqueness of the solution then — and might be considered. Please add a few words about the choice of the distance here.

C2

- The way the "nonstationary" case is addressed is very interesting, but also somewhat controversial. Both the CDF-t method and the authors' work is based on assuming that time evolution is somehow the same for the two models/systems considered, i.e., that

$$T_{Y^1, Y^0} = T_{X^0, X^1}.$$

This might be justified from a dynamical point of view, but from a statistical (or data science) point of view it seems more reasonable to actually consider that the bias between the two systems remains the same, i.e., that

$$T_{Y^1, X^1} = T_{Y^0, X^0}$$

The transformation would then be the opposite, $G^1 = F^1 \cdot (F^0)^{-1} \cdot G^0$ instead of $G^1 = G^0 \cdot (F^0)^{-1} \cdot F^1$ as for the above. Has nobody considered this so far in the literature? Why not? The paper would benefit a lot from a (short) discussion of this second possibility! (and maybe even a few results with it)

- page 7, line 18: Related to the previous item, "because we want to keep the evolution of the model" is a somewhat unscientific statement. Why do you want to assume that the time evolution is the same - what are the reasons that make this a suitable assumption here, especially for complex climate models?
- page 8, line 14: "whereas it increases between ..." - It does, but only here for this example, not in general. Please mention this, to avoid confusion.
- page 8, step 3: The proposed adaption seems somewhat unnatural to me. Looking at Figure 3, I would think that Figure 3a is a more appropriate reconstruction than Figure 3b, since it captures the important fact that the uncertainty about the values increases when transferring the assumed dynamical evolution. This seems actually desirable! But of course it all depends on the goal here. Please comment!

C3

- Finally, it could be nice to discuss in the conclusions the relationship with copulas - which are functions that capture the dependence structure between two random variables X, Y, as does the transport plan here - so there is an underlying "optimal transport copula". Mentioning this connection could maybe make the work presented here interesting for a larger readership.

Minor comments

- page 3, line 13: Just a comment: the nomenclature is a bit strange (this seems to have historic roots in this field, so is not the authors' fault), the name "transfer function" does not seem a good choice as it means something quite different in dynamics. It would be more appropriate to simply call this a "map".
- page 3, line 15: maybe add "deterministic" before "transfer function"?
- page 12, line 19: "close" instead of "closed"
- page 8, line 22-23: A matrix is not "definite", so you probably mean "positive-definite" here in both cases?
- page 14, line 25: "significant" instead of "significative"
- page 14: The discussion of the results shown in Figure 5 is quite dense, it could benefit from a few more words?