This manuscript deals with a new approach (TSMBC) of how to incorporate the time as an additional variable into a multivariable bias correction. The approach can be conducted with existing multivariate BC methods such as MBCn, R2D2, or MRrec. Here, the dOTC approach is followed and the results are compared to a “naive” method, the “Random Bias Correction” (RBC).

The method is first tested on a synthetic dataset, following a VAR process, before applying it to “real” climate data, based on a pseudo reality approach, i.e. treating the RCM results as observations.

The approach could potentially be interesting and innovative. It seems that this is the first time that the time is treated as separate variable in the bias correction. However, I have some doubts that the results are reliable for the application with the real data case (see detailed comments below). Moreover, I think that the evaluation of the TSMBC using synthetic data based on the VAR process is of limited value. It did not convince me technically and scientifically, nor did it help me to better understand the proposed procedure.

On the other hand, more information is required to understand the potential value of the TSMBC. Authors did not convincingly present the methodological background. Critical questions remain unanswered, e.g. what is a VAR process? How is the sampling from the VAR process done? How does the dOTC works?

The Wasserstein metric is also not well introduced in the method section.

Major issues:

- It remains spurious how and why the increase of the numbers of dimensions (could be time lags or other “variables”) affects the stability of the approach. It is just mentioned that the dimensionality should not exceed 10.
- I have some concerns about applying a BC using climate simulations (based on GCMs and not on reanalysis data) if the temporal sequence of variables is addressed,
however, in this case I think it would be acceptable, since the reference is not
observation data but downscaled results of the same forcing GCM.

- My main concern stems from Figure 1 (right, top line). It seems that the mean
precipitation and temperature fields do not correspond to the coast line, as I would
strongly assume. Due to the coarse resolution, you would expect some distortions in
the overlay, but this looks really erroneous. It seems that the projection of GCM and
RCM is wrong, it could be reversed left to right.

- Unfortunately, this would have tremendous impacts on the results and interpretations
in the following (e.g. the spatial dependencies given in Figure 6). For instance, please
explain the statement in lines 300-302. Why is the evolution of GCM variables so
different from that of the RCM? Indeed, the RCM includes more spatially-detailed
“processes”, but is still driven by the GCM. Since the domain of the RCM is rather small,
the impact of the forcing is expected to dominate the RCM simulations.

Moreover, I cannot understand the differences the different performances of the
calibration and the projection period (Figure 4 & 5). I would expect very similar
performances. What is leading to the big discrepancies between the different periods?

- The evaluation results of the TSMBC using synthetic data based on the VAR process are
not convincing (whole section 3) and - at least for me - not fully understandable. For
the revisions, I would suggest to leave out this synthetic exercise. Rather, I would
focus on better explain the applied methods, i.e. the bias corrections approach applied
here (dOTC), the Wetterstein-based metric, and how the naïve RBC (reference
approach) works. I am also wondering if this naïve approach is really suitable for fair
comparison.

- The introduction should be improved, e.g. the statement given in line 28 (... (ii) from
inherent biases in the model simulations.”) is not very helpful. Potential reasons for the
biases shall be mentioned. More and more recent references are required, e.g. for
strong statements given in lines 39 & 40.