Review - *An updated assessment of past and future warming over France based on a regional observational constraint* by Ribes et al (https://doi.org/10.5194/esd-2022-7)

This paper reports on the application of a Bayesian technique that provides estimates of the distributions of past and future regional warming conditional on observed changes in regional and global mean temperatures. Conditioning on the observed changes produces historical warming estimates for France that are lower than recorded in the observations, but higher than obtained by calculating the CMIP6 multi-model mean. It also produces projections of future change that tend to warm more than the CMIP6 multi-model mean and have lower uncertainty than the regional projections from the CMIP6 models. The paper also considers how the constrained warming estimates can be used to provide estimates of updated 30-year climatologies for moving 30-year windows that account for the evolving climate. Precipitation is very briefly considered in the supplementary information.

While the paper is very nicely written, it feels as if it has both too much material, and in some senses, too little material.

First, there is very little discussion of the assumption that the "models are indistinguishable from the truth", but this is at very core of the technique. This is important to discuss, because it implies both model democracy and model equity with observations. This seems counter to approaches that would identify constraints based on model performance relative to observations and concerns that the very high climate sensitivity models found in CMIP6 may not be realistic (e.g., Hausfather et al, 2022, https://www.nature.com/articles/d41586-022-01192-2).

Second, the authors rely on previous papers that they have published (Ribes et al, 2021) or submitted (Qasmi and Ribes, 2021) to present the methodology. While I don't have strong objections to this, it does mean that this paper is not as self-contained as it might be. This creates some challenges for readers who are interested in the methodological details since some of the details of the application of the method are different in this paper than in Ribes et al (2021) and Qasmi and Ribes (2021).

Third, the paper does not provide very much information about the observed regional temperature time series and how it is computed. A hard to obtain paper by Mestre et al (2013) is cited describing the homogenization software that is used in the production of the dataset, but apparently homogenizations is only applied to the early part of the station records that go into the French national thermal index – data from 1947 onwards are not homogenized. When I went searching for the French « Indice Thermique Nationale », I found it surprisingly difficult to find a description. I did, ultimately, find a map showing the locations of the 30 long-running stations that comprise this temperature indicator, and wasn't surprised to see that many of those stations appear to be located in or near urban centers. A potential concern about the data, therefore, is whether these stations are affected by the gradual expansion or intensification of urban heat islands, in which case one might worry that the recorded warming may not be representative of the warming experienced by Metropolitan (mainland) France as a

whole. The paper hints that declining aerosol burdens may have something to do with the rapid warming in the region, but it may also be the case that the data are not as homogeneous as thought.

Finally, I think it should be recognized that the technique apparently does not account for uncertainty in covariance and mean vector estimates that enter into calculation of the posterior distribution for regional mean temperatures. To do so, wouldn't it be necessary to use a more complex hierarchical Bayes model that uses priors on these parameters to describe their potential uncertainty?

Some specific comments:

72-73: This sounds like it might be unnecessarily complex. Since the interest is just in the Metropolitan France spatial mean temperature, does this interpolation scheme simply amount to weighting each climate model grid box by the fraction of the box that lies within Metropolitan France? Is there perhaps a step here that is not described, such as using only the 30 10-km grid boxes that contain the 30 stations that contribute to the ITh to calculate the model version of the ITh?

82: It does seem a bit ironic that, on the one hand, the authors repeat concerns about pattern scaling (lines 42-47), but then rely on it quite heavily throughout the rest of the paper (note that I personally would do so as well).

99: Is the fact that the 30 stations are all at low elevation a limitation?

120-126: I was initially misled by the structure of equation (1) and the explanation of the matrix H. It some time for me to realize that H's function was simply to select a piece of the vector x (i.e., it consists of zeros, except for one or more embedded identity sub-matrices that live within H). Describing the methodology more completely within the paper so that one doesn't have to go searching in related papers would greatly help the reader. This could be done in the supplement, for example, which would probably be better used for that purpose than for dealing quickly with precipitation in any case) .

154: I'm not sure that "measurement uncertainty" would be quite the right description of the information that is represented by the spread of the HadCRUT5 ensemble. Instrumental measure error might be part of that spread, but uncertainty associated with incomplete instrumental coverage and missing data, the adjustment of different types of SST data, the calculation of local anomalies, etc, all contribute, do they not? What we learn, I think, is something about the uncertainty of the estimated global mean temperature anomaly for each year.

158-160: In this case, I think it is appropriate to talk about measurement error, and to assume that it is small relative to internal variability (provided that the stations have very little if any missing data).

242-255: This seems to be about attribution, and yet the next subsection has the title "Attribution". Section 3.2 could perhaps be better entitled "Contributions to warming from different forcing agents" or some such.

295: The discussion of the climate normals adds an interesting twist to the paper, but before making a lot of the apparent lack of uniformity of warming and discussing evidence for a change of seasonality, shouldn't one add uncertainty bands to Figure 4?

332: It would be good to add a reference to a paper that describes and analyzes this "well-known" enhancement.

464-469: Should one also ask a question about the observations? Using a classical D&A technique, Sun et al. (2016, doi: 10.1038/NCLIMATE2956), showed that about 1/3rd of the recorded warming in China between 1951 and 2010 could be attributed to urban warming, and that the remaining warming of about 1°C was much more consistent with the global land mean warming over that period.