

Interactive comment on “Extratropical cyclone statistics during the last millennium and the 21st century” by Christoph C. Raible et al.

Anonymous Referee #3

Received and published: 11 July 2018

This manuscript investigates a long integration of the 1deg version of the CESM, from 850-2100 (with RCP8.5 forcings). The authors use 12-hourly data to track extratropical cyclones (ETCs) over the North Atlantic. Results are dominated by interannual-to-decadal scale fluctuations of cumulative ETC metrics (count, intensity, precipitation) but no obvious external forcing signal is noted. After 2100, strong increases in ETC precipitation and decreases in ETC count are noted, with authors applying a regression analysis to demonstrate that these changes are mostly thermodynamic in nature, in line with previously published work. Some regional variations are also considered, particularly over the Mediterranean and Scandinavia.

In general, I feel the manuscript is clear and crisp, albeit not with overly novel conclusions. As a scientist who deals mostly with future storminess associated with climate

C1

change, I think this type of analysis is relevant to our understanding of climate models and the dynamics of the features themselves within the climate system over long time periods. Where I do have one concern is the results of the tracking algorithm, particularly with regards to CESM, that may be somewhat influencing the results. Before final publication, I feel these should be addressed by either retracking the storms or running a sensitivity analysis. Assuming the authors have a pipeline that performs the subsequent analysis in Figs. 4-8, this should be fairly trivial to undertake.

As one who has used CESM data in the past, if the authors are using the in-line 1000hPa geopotential (Z1000) as a variable (versus calculating Z1000 using the hybrid coefficients and topography) they are likely having issues with the fact that CESM will not automatically interpolate "below ground." Therefore, while the true Z1000 is likely negative over high-terrain areas (e.g., Greenland) the Z1000 reported from CESM is anomalously positive since the code will not go below the lowest model level (at least, according to my recollection). This is a quirk of the CESM in-line interpolation and is likely causing the issues (high cyclone count near high-terrain areas) seen in Fig. 2b since the "background" Z1000 field is biased very high. This can probably be rapidly verified by just comparing the time-mean Z1000 in both ERA-Interim (ERA-I) and CESM. The optimal correction for this would be to use some sort of offline solver with the 3-D Z field and Python/NCL/IDL/etc.

In this vein, it is not clear why the authors are not tracking on sea level pressure (PSL), which is essentially a prognostic quantity in most climate models (technically PS is prognostic, but the correction to PSL primarily uses other prognostic variables like T and the surface topography field). PSL is a much more widely-used quantity when evaluating climate models and would likely alleviate the issues

The fact that CESM simulates far more cyclones than ERA-I is therefore questionable. While there are certainly some differences in effective resolution, etc. of the datasets, a factor of almost 2x (Line 202) in the total number of storms between CESM and ERA-I seems quite high at first blush. The authors hypothesize this is due to weak

C2

storms, but that is not clear to me from Fig. 3. For example, the SLP distribution shows more weak storms for CESM, but also more strong ones. Having a smaller radius distribution is also not necessarily indicative of weaker storms, as aspects of the model configuration such as numerical diffusion and how grids are interpolated may contribute to differences here. This is somewhat hinted at in Figs 3c-d. As an additional example, one could make an argument that 1deg ETCs would be "smaller" than 4deg ETCs, but 1deg ETCs *should* be more intense based on being better resolved.

I would like the authors to consider "retracking" the storms if PSL is available. They could easily modify their algorithm to search for prognostic deficits in PSL as in other trackers within the IMILAST project (of which the lead authors of this manuscript already contributed to). If that is not available, I would like the authors to try and evaluate whether or not the issues of additional ETCs tracked in CESM are related to the Z1000 issue noted above. One option would be to run CESM for a short period (perhaps a few decades) and compare the results of using the inline Z1000 with PSL or a more accurately diagnosed Z1000.

Minor comments:

Line 122: "So called" is too colloquial, would just say "this is the 1deg version of the model used in CMIP-class experiments" or thereabouts.

Line 123-124: Would include a sentence or two about the subgrid physics package used in this version of CESM (in the atmospheric model) since that would have the largest impact on the results here, particularly thermodynamic ones.

Line 245: Are there changes in mean storm-track, basin-wide surface pressure, etc. that may be relevant here?

Line 262: 4deg models certainly underresolve synoptic scale features, which ETCs are.

Line 299: Is this a basin-wide metric? I question a bit about correlating the spatial

C3

pattern with basin-wide metrics as then I'd expect large scale North Atlantic patterns that control ETCs to dominate this analysis (e.g., the NAO).

Line 312: The spatial field remains quite noisy, I would be a bit careful about being too conclusive since, even with a multi-century simulation, I'm not sure we can be completely confident very small ($O(10\text{deg})$) spatial patterns are tremendously significant in a model whose effective resolution is probably $\sim 6\text{deg}$ (see Skamarock 2004 for discussion of effective resolution in numerical models).

Line 321: This reads as a bit "hand-wavy;" I'd formalize and clean up the text a bit.

Line 332: These "barotropic pressure structures" could be underresolved warm core storms (e.g., tropical cyclones) moving to mid-to-high latitudes. 1deg models are capable of starting to simulate these features, albeit far weaker than what is observed (e.g., Wehner et al., 2014, Walsh et al., 2015).

References:

Skamarock, W.C., 2004: Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra. *Mon. Wea. Rev.*, 132, 3019–3032, <https://doi.org/10.1175/MWR2830.1>

Walsh, K.J., S.J. Camargo, G.A. Vecchi, A.S. Daloz, J. Elsner, K. Emanuel, M. Horn, Y. Lim, M. Roberts, C. Patricola, E. Scoccimarro, A.H. Sobel, S. Strazzo, G. Villarini, M. Wehner, M. Zhao, J.P. Kossin, T. LaRow, K. Oouchi, S. Schubert, H. Wang, J. Bacmeister, P. Chang, F. Chauvin, C. Jablonowski, A. Kumar, H. Murakami, T. Ose, K.A. Reed, R. Saravanan, Y. Yamada, C.M. Zarzycki, P.L. Vidale, J.A. Jonas, and N. Henderson, 2015: Hurricanes and Climate: The U.S. CLIVAR Working Group on Hurricanes. *Bull. Amer. Meteor. Soc.*, 96, 997–1017, <https://doi.org/10.1175/BAMS-D-13-00242.1>

Wehner, M. F., K. A. Reed, F. Li, Prabhat, J. Bacmeister, C. Chen, C. Paciorek, P. J. Gleckler, K. R. Sperber, W. D. Collins, A. Gettelman, and C. Jablonowski (2014), The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1, *J. Adv. Model. Earth Syst.*, 6, 980–997, doi: 10.1002/2013MS000276.

C4

