Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-181-RC1, 2022

Comment on gmd-2022-181

Anonymous Referee #1

Referee comment on "The Euro-Mediterranean Center on Climate Change (CMCC) decadal prediction system" by Dario Nicolì et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-181-RC1, 2022

In the paper “The CMCC Decadal Prediction System”, Nicolì et al. document the CMCC DPS and evaluate its predictive skill with initialized hindcast simulations that have been contributed to CMIP6. The system distinguishes itself by its use of the CMIP6-generation Earth system model CMCC-CM2-SR5 as well as its full-field initialization strategy that covers all components and generates realistic spread in the subsurface ocean. Comparing the CMCC DPS results to uninitialized historical CMCC-CM2-SR5 simulations, the authors demonstrate global skill benefits from initialization that are particularly pronounced in the North Atlantic despite a substantial forecast drift in that region. The authors highlight skillful multiyear prediction of North Atlantic atmospheric circulation variability and attribute modest skill over land to the limited ensemble size. Given the uniqueness and performance of the system, the authors make a strong case that the CMCC DPS is a valuable addition to the suite of existing decadal prediction systems. The data analysis and figures of the paper are clear and innovative, including probabilistic skill measures and skill dependence on lead time and average period. The authors succeeded in keeping the paper short by selecting a few key metrics. Overall, the paper is well and concisely written, the argumentation is sound, and the paper fits well the scope of the journal. I therefore recommend its publication in GMD after having addressed a few minor comments listed below.

1. L89 “Finally,”

Repeated in the next sentence.

2. L94 “15-member ensembles”

A bit unclear how the spread for all 15 initial conditions is obtained from two land surface conditions and five ocean conditions. Are the 10 states of Init utilized in addition? Or, did you apply any small initial perturbations? Please elaborate.

3. L97 "ERA-Interim (1979–2020)"

To my knowledge, ERA-Interim is only available until August 2019. Considering that you describe CMCC DPS as an operational system, could you please detail which product you use for the initialization in 2020 and later?

4. L101 "atmospheric forcing datasets: CRUNCEP version 7 (Viovy, 2016) and GSWP3 (Kim, 2017)"

According to a quick search, CRUNCEP version 7 is available until 2016 and GSWP3 until 2010. Do these products indeed cover the entire 1960¬–2020 period and are they updated in the future? If not the case, how does CMCC DPS initialize the land component beyond their coverage?

5. L101–102 "provide four instantaneous 2-meter air temperature and humidity, 10-meter winds and surface pressure every six hours"

Do you mean "four" realizations of instantaneous fields (i.e., a mini-ensemble)? Or, do you mean four times per day (i.e., every six hours)? Please clarify.

6. L104–105 "An ensemble of 5 ocean initial states is used to initialize the ocean and sea ice components"

Could you add more detail on how the ocean initial state is used to initialize the sea ice component? Does CMCC DPS initialize sea ice thickness in addition to sea ice concentration?

7. L116 "The predictive skill for both initialized reforecasts and uninitialized projections is assessed against observational products."

Please specify the temporal coverage of the assessment. Is it always 1961–2021 (for LY1 and LY1–5) and 1966–2021 (for LY6-10)? If the coverage varies for the different

validation products, then it should be stated for each product separately (either in the main text or in the figure captions).

8. L130 "observed variability over the historical 1960–2014 period targeted by the decadal reforecasts".

Correct? Or should it be something like 1961–2021?

If 2014 is indeed the last verification year, then all hindcasts with start dates in 2014 and later are not included in the analysis.

9. L137 "MSE_HO (MSE_PO) is the mean square error evaluated for the initialized (uninitialized) ensemble mean against observations"

Are the respective climatologies (i.e., the bias) subtracted prior to the MSE computation as in Goddard et al. 2013 who defined anomalies "relative to their respective climatologies (which is equivalent to the removal of mean bias)"?

To my understanding, the MSSS used by Goddard et al. is designed to reflect phase and amplitude errors but not climatological bias.

10. L159–160 "The statistical significance is assessed with a one-tailed Student's T test (Wilks, 2011), accounting for auto-correlation in the time series (Bretherton et al., 1999)."

Please specify explicitly if local or field significance is tested (I assume "local" but got a bit confused by the citation to Bretherton et al. on the effective number of "spatial" degrees of freedom).

Also, could you please re-state the formula used to account for auto-correlation or cite the equation in Bretherton et al. 1999? Is it the same as used in Vecchi et al. 2017 (https://doi.org/10.1175/JCLI-D-17-0335.1)?

11. L161 "reference period 1981–2010, in agreement with WMO recommendation"

According to latest WMO recommendation the period for the computation of climate normals is "the most-recent 30-year period finishing in a year ending with 0" i.e. currently 1991–2020 (https://library.wmo.int/doc_num.php?explnum_id=4166). I'd therefore suggest removing "in agreement with WMO recommendation".

12. L178 "The ensemble spread envelope"

Please state how you defined this envelope (either in the main text or the figure captions). For example, is it defined as min/max or a certain percentile range of the data?

13. L180–182 "The initialization contributes to the reduction of the Init ensemble spread, which is about half the envelope of the NoInit for lead-year 1. This is not so unexpected since, when a simulation is initialized the observed internal variability is imposed thus reducing the uncertainty related to systematic errors (Doblas-Reyes et al., 2013)"

What do you mean by "systematic errors" here? To me, the sentence seems a bit unclear/confusing and not quite in line with the reasoning of Doblas-Reyes at al. who wrote "The initialization can, in addition to providing information about the phase of internal variability, correct systematic errors in the model response to external forcings".

First, a correction of the model response to external forcings (i.e., the forced climate trend) does not necessarily lead to a reduction in ensemble spread. The forced trend should be the same or very similar for all members given the model parameters are not perturbed.

Second, imposing the observed "internal" variability should not significantly change the forced climate trend. To my understanding, Doblas-Reyes et al. argue that initialization benefit is partly due to imposing the observed forced trend in addition to synchronizing the internal variability.

I would have expected that the reduction in ensemble spread is simply a consequence of the synchronization of internal climate variability. But maybe I have misunderstood, and this is what the authors intended to say. Please clarify.

14. L192 "The added value of initialization"

Regarding the attribution of ACC differences and MSSS to added value of initialization: Did you verify that these metrics are not overly sensitive to the use of different ensemble

sizes of NoInit (10) and Init (15)?

15. L199–200 "the skill undergoes a clear deterioration over the tropical and northern part of the Pacific Ocean (Fig. 2c)"

A clear deterioration relative to what? Relative to the first-year skill or relative to the skill of NoInit? Please clarify.

I would say that Fig. 2d does not show a clear deterioration relative to the NoInit skill. That the multiyear skill is deteriorated relative to the first-year skill in the tropical and northern part of the Pacific Ocean is maybe not too surprising given the limited predictability of ENSO and its strong influence on the PNA variability. Or would you expect more multiyear skill in that region?

16. L268 "large ACC values are also found over regions linked to the AMV through remote teleconnections"

For which variables?

17. L278 "Init well captures the NAO variability despite the limited ensemble size (ACC=0.80 applying an 8-year running mean to the model index)"

This is interesting and could indicate that the ratio of predictable components (RPC) is lower than found in other prediction systems (e.g., Scaife and Smith 2018). Could you say anything about how well the amplitude is predicted? A small amplitude (i.e., RPC >> 1) could also be a likely explanation for modest skill over land despite obtaining high ACCs for NAO.

Scaife, A.A., Smith, D. A signal-to-noise paradox in climate science. npj Clim Atmos Sci 1, 28 (2018). https://doi.org/10.1038/s41612-018-0038-4

18. L282 "TPI with significant skill peaking at lead-years 4–10"

How would explain the skill emergence for higher lead times? Would you, for example, attribute it to initialization shock, sampling uncertainty (i.e., different end points of evaluation period for different lead times) or a combination of both?

19. L291 "a set of 15-member hindcasts, covering the period 1960–2020"

Maybe "initialized every year during 1960–2020" would be more precise?

If I have understood correctly, the entire forecast period covers 1961–2030 (plus December 1960). Is that right?

20. L302–303 "The poor response in MSSS is also confirmed by a cold bias"

"confirmed" somewhat implies a causal relation between the poor MSSS response and cold bias. Could you elaborate on that? If the relationship is not known, then writing "coincides with" could be more appropriate.

Following Goddard et al. 2013, the MSSS is computed from anomalies "relative to their respective climatologies (which is equivalent to the removal of mean bias)". Hence, the MSSS panelizes amplitude and phase errors but not directly biases. Still, it is plausible that biases in high latitude ocean temperatures and sea ice cover also affect the representation of climate variability there.

21. L348–349 "inability of the DPS to reproduce the observed variability"

Do you mean inability to reproduce the historical temporal evolution of the observed variability or the inability to reproduce the dynamics and other characteristics of the observed variability? "inability" suggests you anticipate a higher real-world climate predictability over the North Pacific than the hindcasts of CMCC DPS would suggest. If so, why would you think so?

22. Sections 3.4 and 4

Can you provide a rational for why you show and discuss the NoInit results in sections 3.1–3.3 but not in 3.4 and 4. Was it mainly to keep the paper short and/or the benefit of initialization turned out to be less clear for the ROC and climate indices?