Response letter to Referee Reviewer 1

REVIEWER COMMENTS

Referee Reviewer 1

This study evaluates the sensitivity of the PGW method to regionally uniform vs. gridpoint scale perturbations and the inclusion of different perturbed meteorological variables. The authors study the results in the context of 3 different flood events that occurred in the Northeastern U.S. using WRF-simulations forced by ERA5 and ERA5 perturbed with CESM-LENS.

I thought the basis of this study was interesting, given that there is little consistency in how PGW simulations are designed, but the lack of clarity in the methods, the lack of including key PGW literature, and arguments that did not make sense resulted in a paper that had too many issues to make a strong argument for what the authors were trying to show. For this reason, this paper cannot be accepted unless substantial revision is undergone. I will provide more specific comments below to explain my rationale.

Dear reviewer, thank you very much for your questions and suggestions. We admit this paper had some issues you mentioned, but we think most issues are not fundamental and are caused by misunderstanding and lack of enough clarification. Therefore, we respond to all general issues and some major specific issues you mentioned and hope it can address your concern. We will write a point-to-point response after the interactive review process and revise the paper correspondingly.

General:

1. I am confused about how you evaluate the influence of different meteorological variables in your paper for each flood event. The figure captions merely state "2055 October", "2056 May" and "2056 June". However, each flood event corresponding to those months only lasts a few days of the month. So, when you present your Figures 9-13, are these fields being averaged over the timeframe of the month that contains each flood or is it averaged over the time period of the flood? This must be clearly stated because it muddles the interpretation of your figures.

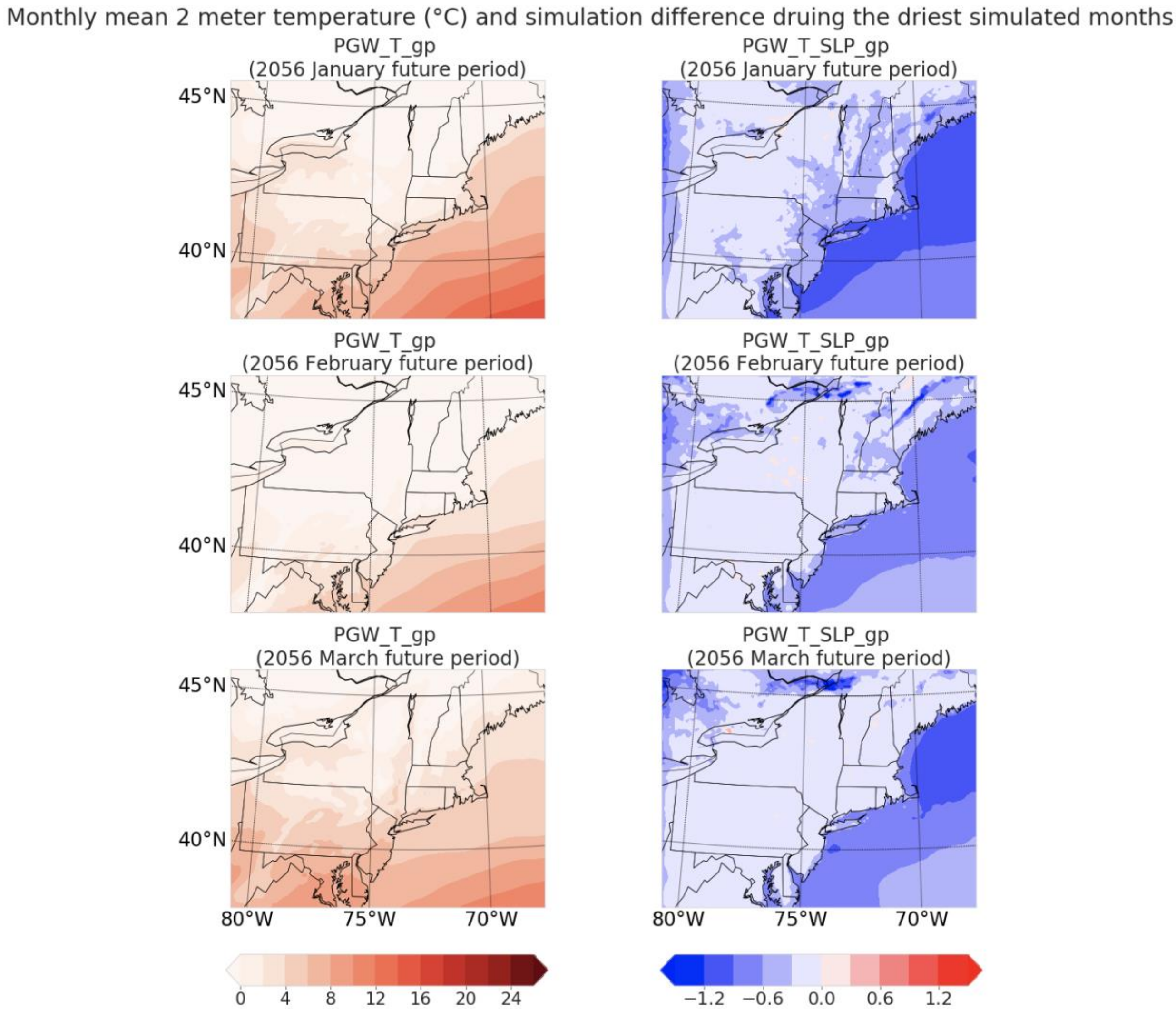
I am concerned that if you take the mean over the month in which the flood occurs in the future simulation, the changes in fields like pressure and temperature are not just a result of the future perturbation, but are also being modified by the mesoscale processes within the storm-producing flood (and any other precipitation events that occurred during that month).

If these figures are just averaged over the time period of the flood (including when it is raining), again, you are introducing mesoscale modifications of the perturbations, including mesolows, outflows, and cool pools. It is impossible to evaluate these figures and what the changes in pressure, temperature, and CAPE mean if they are already being contaminated by the presence of storms.

Sorry for the misunderstanding and confusion. In Ln 99-100, we said "This period is chosen as it includes the three major flood events in Table 1." And as we defined in Table 1, the 2005 Oct Flood, 2006 May Flood, and 2005 June flood periods refer to the Northeast U.S. flooding of October 2005, the New England Flood of May 2006, and the 2006 Mid-Atlantic United States flood, which has the duration from October 7th to 17th 2005, May 12th to May 20th, 2006 and June 23rd to July 5th, 2006. Therefore, in all analysis and corresponding figures (for example Figures 9-13), the "2055 October", "2056 May" and "2056 June" refer to the period of October 7th to 17th 2005, May 12th to May 20th, 2006 and June 23rd to July 5th, 2006. And the period mean refers to the average values of these three periods. We will add clarification in the manuscript.

We do average over the period of the flood, and it's no doubt that the climate perturbations in the WRF simulation will be different from the original climate perturbation (from GCMs) we added to the PGW runs. Actually, in WRF simulations, there always exist atmospheric phenomena and circulation that can alter the original climate perturbations to some degree. In fact, we do want to include the impacts of different storms on the climate perturbations, which make different floods have different sensitivity to PGW methods, and that's why we studied three different storm events and concluded that the 2005 Oct Flood which is caused by the cold front system, is more sensitive to several climate perturbations.

More importantly, we can confirm that the occurrence of the storms will not significantly change the impacts of sea level pressure perturbation on surface air temperature. For example, in Figure R1, we plot the monthly mean 2 meter temperature simulation difference between PGW_T_gp and PGW_T_SLP_gp during the months with lowest monthly mean precipitation (2056 January to March) in our simulation period. And we can see that PGW_T_SLP_gp still has consistently lower 2 meter temperature over the sea. This illustrates that our conclusions regarding the SLP perturbation hold even in the absence of storms



Monthly mean 2 meter temperature (°C) and simulation difference during the driest simulated months

2. This paper lacks a complete understanding of the PGW literature and is missing some key papers, which is necessary if you are evaluating the utility of this method. Papers to include (not exhaustive):

Dougherty, E., and K. L. Rasmussen, 2020: "Changes in flash flood–producing storms in the United States. J. Hydrometeor., 22, 2221–2236, https://doi.org/10.1175/JHM-D-20-0014.1."

Dougherty, E., and K. L. Rasmussen, 2021: "Variations in flash flood-producing storm characteristics associated with changes in vertical velocity in a future climate in the Mississippi River Basin. J. Hydrometeor., 21, 671–687, https://doi.org/10.1175/JHM-D-20-0254.1.

Mahoney, K., D. Swales, M. J. Mueller, M. Alexander, M. Hughes, and K. Malloy, 2018: An examination of inland-penetrating atmospheric river flood event under potential future thermo-dynamic conditions. J. Climate, 31, 6281–6297, https://doi.org/ 10.1175/JCLI-D-18-0118.1.

Lackmann, G. M., 2013: The south-central U.S. flood of May 2010: Present and future. J. Climate, 26, 4688–4709, https://doi.org/10.1175/JCLI-D-12-00392.1.

Liu, C., and Coauthors, 2016: Continental-scale convection-permitting modeling of the current and future climate of North America. Climate Dyn., 49, 71–95, https://doi.org/10.1007/s00382-016-3327-9.
Prein, A. F, C. Liu, K. Ikeda, S. B. Trier, R. M. Rasmussen, G. J. Holland, and M. P. Clark, 2017: Increased rainfall volume from future convective storms in the US. Nat. Climate Change, 7, 880–884, https://doi.org/10.1038/s41558-017-0007-7.
Rasmussen, K. L., A. F. Prein, R. M. Rasmussen, K. Ikeda, and C. Liu, 2017: Changes in the convective population and thermodynamic environments inconvection-permitting regional climate simulations over the United States.

Given constraints on space we limited our review to about 20 PGW-related papers, although we do admit that citing more papers can make this paper more comprehensive and we will cite the papers you mentioned. But the absence of these citations doesn't seem to be a fundamental flaw in the work, and does not alter our conclusions.

3. I think the gravity wave noise section needs to be reevaluated. Previous PGW studies including Lackmann (2013) and Mahoney et al. (2018) have found that gravity wave adjustments in their studies are relatively small and only apparent during the early spinup periods. This makes me skeptical of what you are arguing, given that I do not see any evidence of gravity waves in your figures and the presence of wave-like noise in the presence of storms implies that you are seeing storm-generated gravity waves, which are physical and not just a model artifact.

Two of the papers mentioned above (namely, Lackmann (2013) and Mahoney et al. (2018)) do mention that the gravity wave adjustments are only strong during the early stages of the simulation. But both of them only mention it in one sentence and do not provide any evidence supporting these results (such as figures or wave analysis). For example, Lackmann (2013) only state: "Any imbalance between the wind and mass field is sufficiently small to preclude strong gravity wave adjustment early in the future simulation." Mahoney et al. (2018) state: "Gravity wave adjustment between the wind and mass fields early in the simulations is accordingly short-lived." Because both of them do not provide any further explanation and figures, we cannot comment on their conclusions. We guess that they examined the mean field (like the daily mean), where the gravity wave noise is invisible, instead of examining the simulation at high temporal resolution.

In our simulations, we also find that the gravity waves are not obvious if we look at the period mean (for example as shown in Figures S6 and S7); however, if we look at the sea level pressure at high temporal resolution, as it evolves (the animation in the citation in Ln 203 and Figure R2 here, available at https://zenodo.org/record/6544880#.Y0YyQuzMJ9s), wave-like noise during the flood period and in the PGW simulations are much stronger when compared with the historical run. Also, the wave-like noise is more significant during storm events and that's why we say "the observed noise is enhanced by strong advection during these extreme weather events, as it is much more obvious during these storm events." More importantly, as shown in Figure R3 as attached, we can see the wave-like noise in the magnitude spectrum after Fourier Transform in the PGW runs (the high-frequency signals are shown as light circles in the magnitude spectrum) but do not observe such a signal in the historical run. This further confirms that the wave-like noise does exist in our simulation during the flood periods and that it is not obvious (or present to the same degree) in the control simulation.

Please refer to https://zenodo.org/record/6544880#.Y0YyQuzMJ9s

Figure. R2 The animation of sea level pressure during the 2005 October Flood and returned 2055 October Flood

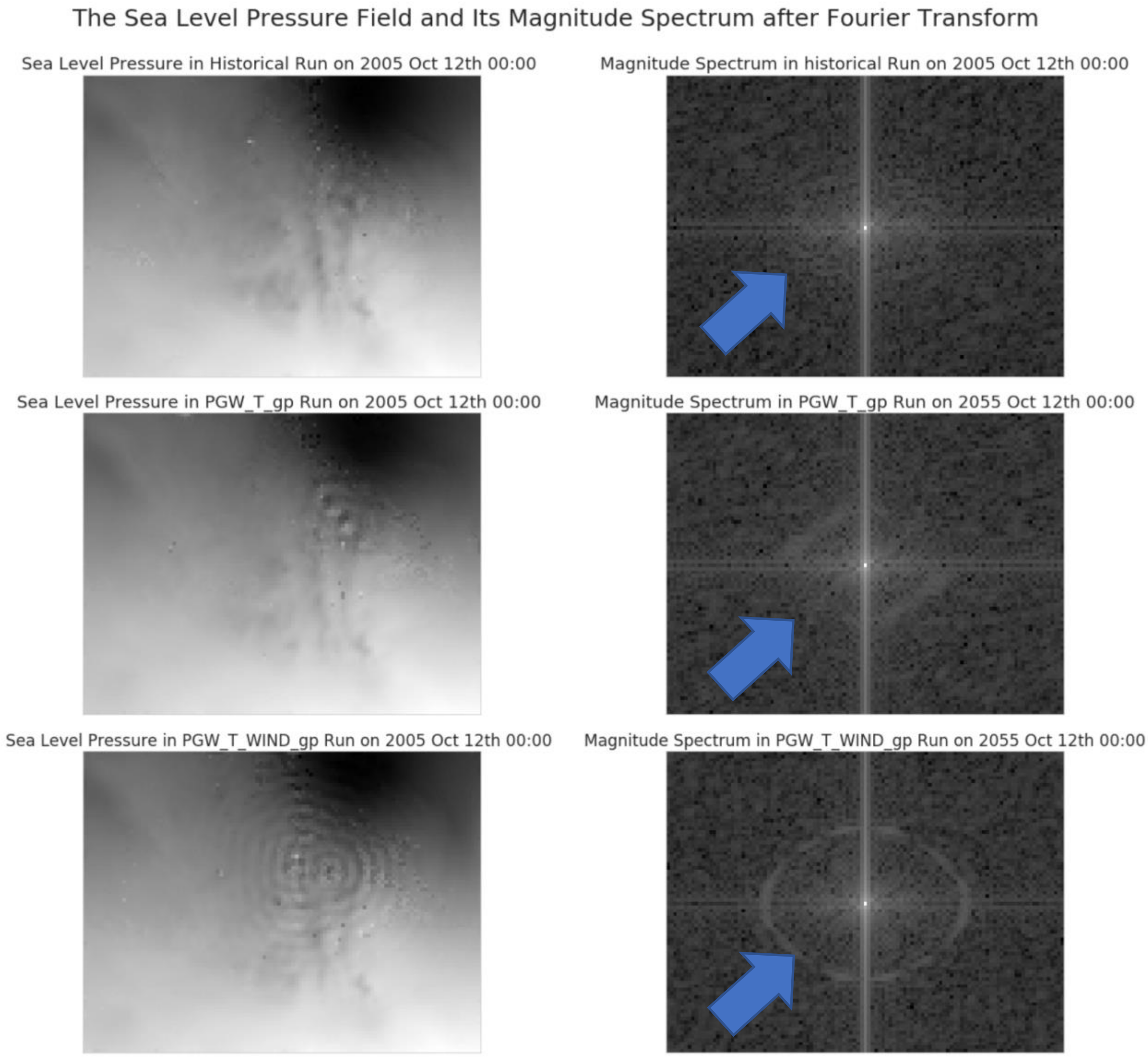


Figure. R3 The sea level pressure field and its magnitude spectrum after Fourier transform

Major specific comments

Lines 89–91: I am curious why you decided not to perturb moisture in this study–can you please explain this in your methods?

We perturb the moisture condition by fixing the relative humidity and applying the warming perturbations instead of directly changing the specific humidity. This is a common approach under the PGW method (e.g., the original PGW papers of Schär et al., 1996 and Frei et al., 1998), since an examination of future and historical relative humidity over this region suggests that changes to relative humidity are not statistically significant. If specific humidity were perturbed then a grid cell where relative humidity was originally 100% could drop below 100% leading to suppression of precipitation. Perturbing specific humidity could also lead to relative humidity being increased above 100%, which is unphysical. For these reasons specific humidity is not perturbed directly.

Line 105: Why is CLM rather than Noah-MP used for a land model?
Lines 105–107: I think the choice of the land model is actually quite important. I understand it is not central to your study, but Barlage at el. (2021) showed the warm, dry bias in the Central U.S. from PGW simulations could be reduced by adding groundwater to Noah-MP in CONUS-wide PGW simulations.

There are a number of reasons why CLM was selected for this study: As described in the paper, both the physical parameterizations and the land model (CLM) we used are from CESM1, which is widely regarded as a high-performance climate model (Kay et al., 2015; Sillmann et al., 2013; Karmalkar et al., 2019). So, CLM is more consistent with the physical parameterizations we used. Also, the CLM model is also the most complicated and expensive of the available options in WRF, and one that shows reasonable performance across a variety of geographies (Jin et al., 2010; Case et al., 2008; Ullrich et al., 2018). Indeed, although Barlage at el. (2021) showed the warm, dry bias in the Central U.S. from PGW simulations could be reduced by adding groundwater to Noah-MP in CONUS-wide PGW simulations, it did not confirm that Noah-MP can outperform the CLM land model. More importantly, this study focuses on the dry periods over the Central U.S. but our study focuses on flood periods over NEUS. We know that model performance varies from region and phenomenon.

Lines 112–114: Can you specify the spatial scales here? Did you employ nudging through the entire vertical domain or just above the boundary layer?

The spectral nudging uses guv, gt, gq and gph equaling to 0.0003. The xwavenum and ywavenum are equal to 3. The nudging is only applied above the boundary layer.

Table 1: How were the start and end dates of the flood determined?

The start and end dates of flood periods are informed by the reference papers in the Meteorological Cause column in Table 1. To get the exact dates, we define the start of the flood as the day when regional daily mean precipitation over the inner domain is larger than 1.8 mm/day, and the end of the flood as the day when two consecutive days after that day is less than 1.8 mm/day.

Line 203–205: When you are saying that dynamical fields are assumed to be unchanged, I think you are speaking very specifically about uniform temperature perturbations in the regional sense. However, I do want to clarify that you have to be careful about making these statements since this is not the case in all PGW simulations. I think it is more accurate to say that the PGW method assumes that the dynamical changes are much smaller order magnitude than thermodynamic changes (Liu et al. 2016; O'Gorman 2015). Furthermore, it is impossible to assume the dynamics don't change due to change in thermodynamics (i.e., temperature), because these interact with each other. For example, Dougherty and Rasmussen (2021) show that storms with stronger updrafts show greater increases in rainfall in 4-km PGW simulations, likely due to the latent heat feedback suggested by Trenberth (1999).

Sorry for the misunderstanding. By saying that dynamical fields are assumed to be unchanged, we mean that the dynamical field in the WRF input is not changed. Certainly the modified thermodynamic fields will change the original dynamical fields as a result of mechanisms like the geostrophic adjustment. We will clarify this in the paper and cite the reference you mentioned.

Lines 209–210: How do you know PGW_T_regional is overestimating precipitation when there is no future "truth" or observations to compare it to? What is your baseline for this statement?

While we cannot know for sure if PGW_T_regional overestimates the precipitation, we mean to say that PGW_T_regional overestimates precipitation compared with the gridpoint perturbation methods. However, there are good reasons to believe that PGW_T_regional is behaving unphysically. Because the PGW_T_gp employs the time-varying temperature perturbations at each gridpoint and pressure level (3D space + 1D time), it can capture the horizontal spatial variance of the temperature perturbation. But PGW_T_regional only uses the time-varying temperature perturbations at the regional mean scale at each pressure level (2D space + 1D time), and so does not include any information about the horizontal spatial variance of the temperature perturbation. There is ample evidence to indicate that the temperature perturbation over the land will be larger than the temperature perturbation over the sea, and so it is likely that the regional mean modification leads to future temperatures over the sea that are higher than climate model projections indicate. The overestimated temperature perturbation in PGW_T_regional can be seen in Figure 4 (in the main paper). Since the relative humidity is fixed in our PGW runs, the change of specific humidity is determined by the temperature perturbation, and is subsequently larger in the regional mean simulation. Since precipitation in this region is largely driven by on-shore flow from the Atlantic, the enhanced specific humidity leads to enhanced vapor transport and increased moisture convergence (and subsequently higher precipitation in Figure 5 and 6 of the paper). Therefore, we claim for this choice of domain, the simulated precipitation in PGW_T_gp is more reliable compared with PGW_T_regional.

Technical comments:

Thank you very much for your comments. We will correct them correspondingly.

Reference

Schär, C., Frei, C., Lüthi, D., and Davies, H. C.: Surrogate climate-change scenarios for regional climate models, Geophysical Research Letters, 23, 669–672, 1996.

Frei, C., Schär, C., Lüthi, D., and Davies, H. C.: Heavy precipitation processes in a warmer climate, Geophysical Research Letters, 25, 1431–1434, 1998.

Karmalkar, A. V., Thibeault, J. M., Bryan, A. M., and Seth, A.: Identifying credible and diverse GCMs for regional climate change studies-case study: Northeastern United States, Climatic Change, 154, 367–386, 2019.

Sillmann, J., Kharin, V., Zhang, X., Zwiers, F., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research: Atmospheres, 118, 1716–1733, 2013.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S., Danabasoglu, G., Edwards, J., et al.: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability, Bulletin of the American Meteorological Society, 96, 1333–1349, 2015.

Ullrich, P., Xu, Z., Rhoades, A., Dettinger, M., Mount, J., Jones, A., & Vahmani, P. (2018). California's drought of the future: A midcentury recreation of the exceptional conditions of 2012–2017. Earth's Future, 6 (11), 1568–1587.

Jin, J., Miller, N. L., & Schlegel, N. (2010). Sensitivity study of four land surface schemes in the WRF model. Advances in Meteorology, 2010.

Case, J. L., Crosson, W. L., Kumar, S. V., Lapenta, W. M., & Peters-Lidard, C. D. (2008). Impacts of high-resolution land surface initialization on regional sensible weather forecasts from the WRF model. Journal of Hydrometeorology, 9 (6), 1249–1266.