Reply to Reviewer #1:

Thank you for the insightful comments and detailed instructions on how to improve the manuscript, and the manuscript has been revised based on the reviewer's comments. Below, the texts with italic font are the reviewer's comments, and the texts with normal font and blue color are authors' responses.

Authors attempted to identify the East Asian temperature extremes contributed by dynamical and thermodynamical processes. If is of scientific significance for a better understanding regional climate anomaly and extremes under warming. The results are clearly presented with logics. I suggest authors to clarify how the critical method applied, and provide more details.

1. *Section 2.3.1, The details how the dynamic adjustment approach was applied are not clear. It is confusing how a target SLP pattern is defined in advance, based on this target field you rank the monthly fields according to their similarity to the target SLP filed, and selected 150 most similar fields. But later, you said the 100 randomly selected to construct the target SLP pattern. Quite confusing, please clarify to avoid misleading.*

Response: Thank you for your suggestion.

(1) "For a given month and year (e.g. December 1990) in each ensemble member, we rank the 2000 (1800) December SLP fields in the PiCTL simulation by their similarity with the target SLP pattern according to Euclidean distance." In this sentence, the SLP field in December 1990 is the target SLP field.

(2) We first select 150 closest fields, and then randomly subsample 100 fields from these 150 fields to compute their optimal linear combination that best fits the target SLP field.

We have modified this part as "For a given "target" month and year (e.g. December 1990) in each ensemble member, we rank the 2000 (1800) December SLP fields in the PiCTL simulation by their similarity with the target SLP pattern according

to Euclidean distance. From the 150 SLP fields with the smallest Euclidean distances, we randomly subsample 100 SLP fields to construct the best estimation of the target SLP pattern by linear combination. The same set of linear coefficients is applied to the accompanying SAT fields to obtain the associated linear combination of SAT. We repeat the subsampling procedure 100 times and average the 100 linear combinations to derive the dynamically induced SAT field in the target month. Deser et al. (2016) illustrate the importance of this iterative random selection process and the reason for the repeated subsampling procedure is to take into account the uncertainty related to internal thermodynamic variability and to ensure the robustness of the results …"

Please see L147-158 in the revised manuscript.

2. *The target atmospheric circulation fields are global or hemisphere or regional? May be of different meanings and limitations.*

Response: The target atmospheric circulation fields are regional. In the manuscript, we use the domain 15°~90°N, 30°~180°E for the SLP analogues. Similar to Deser et al. (2016), we examined the impact of the regional selection and the results are not sensitive to the precise region used (Figure A1).

We have added this in the revised manuscript as "We use the domain 15°~90°N, 30°~180°E for the SLP analogues. The sensitivity to the precise region used is small (Figures not shown; e.g. within ±5° of latitude and ±10° of longitude)." Please see L158-160.

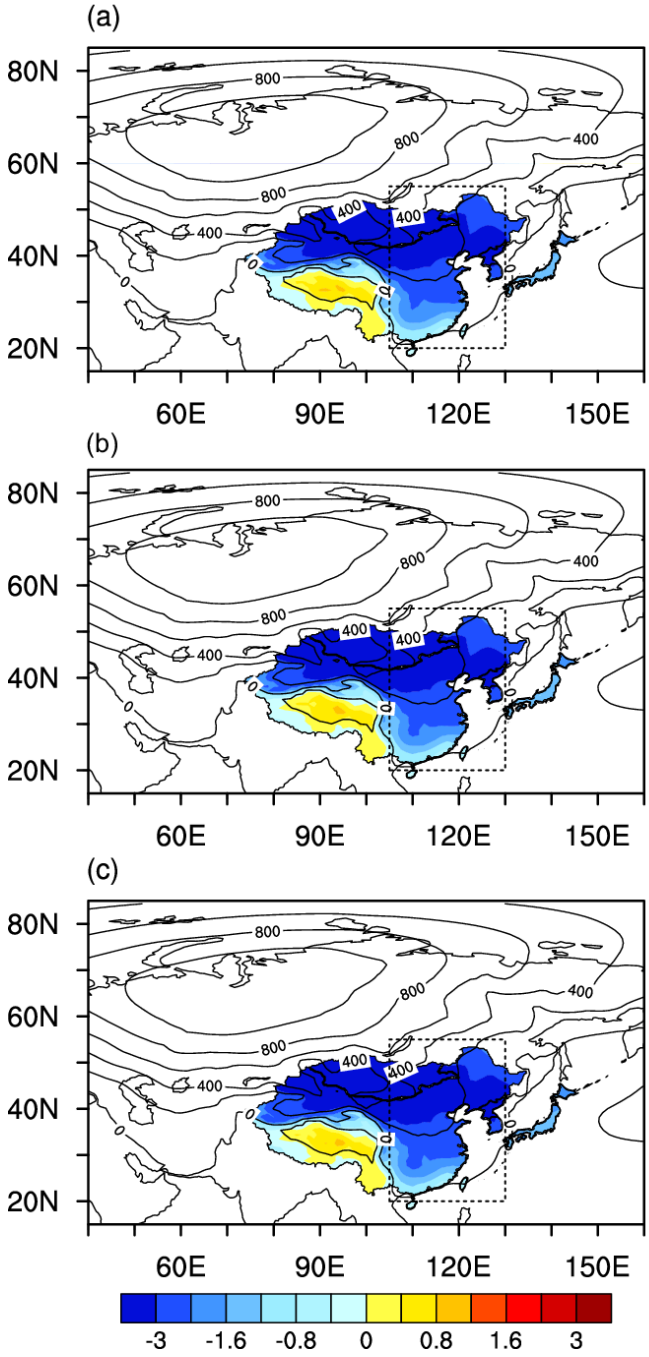

Figure A1 The composites of cold–month SAT anomaly in color shading and SLP anomaly in contours (relative to 1986-2005 boreal winter climatology) in CESM-LE during the period of 1986-2005 boreal winter obtained from different domains for SLP analogues: (a) 10°~90°N, 20°~190°E, (b) 15°~90°N, 30°~180°E and (c) 20°~85°N, 40°~170°E

3. *Randomly select 100 from 150 fields to construct the dynamic field. Why random 100? Why not simply 150 filed combination? Any evidence or estimation for the optimum number of samples?*

Response: Simply using the closest 150 fields for fitting directly is also feasible. The reason for randomly subsampling 100 fields and taking the average value after multiple iterations is to fully take into account the uncertainty related to internal thermodynamic variability, and to increase the robustness of the results (Deser et al., 2016). As for the choice of random subsample size, we examined the differences in estimating the dynamic component of the actual field with different subsample sizes. Figure A2 shows the comparison when the subsample size is chosen as 100 and 10, revealing small differences between the two. However, with a larger subsample size, the details are better captured.

We have also checked the effect of iterations. The results indicate that as the number of iterations increases (approximately larger than 20), the differences in SAT anomalies among different iterations decrease (Figure A3).

We have added the reason for the random subsampling procedure as "Deser et al. (2016) illustrate the importance of this iterative random selection process and the reason for the repeated subsampling procedure is to take into account the uncertainty related to internal thermodynamic variability and to ensure the robustness of the results." Please see L155-158 in the revised manuscript.

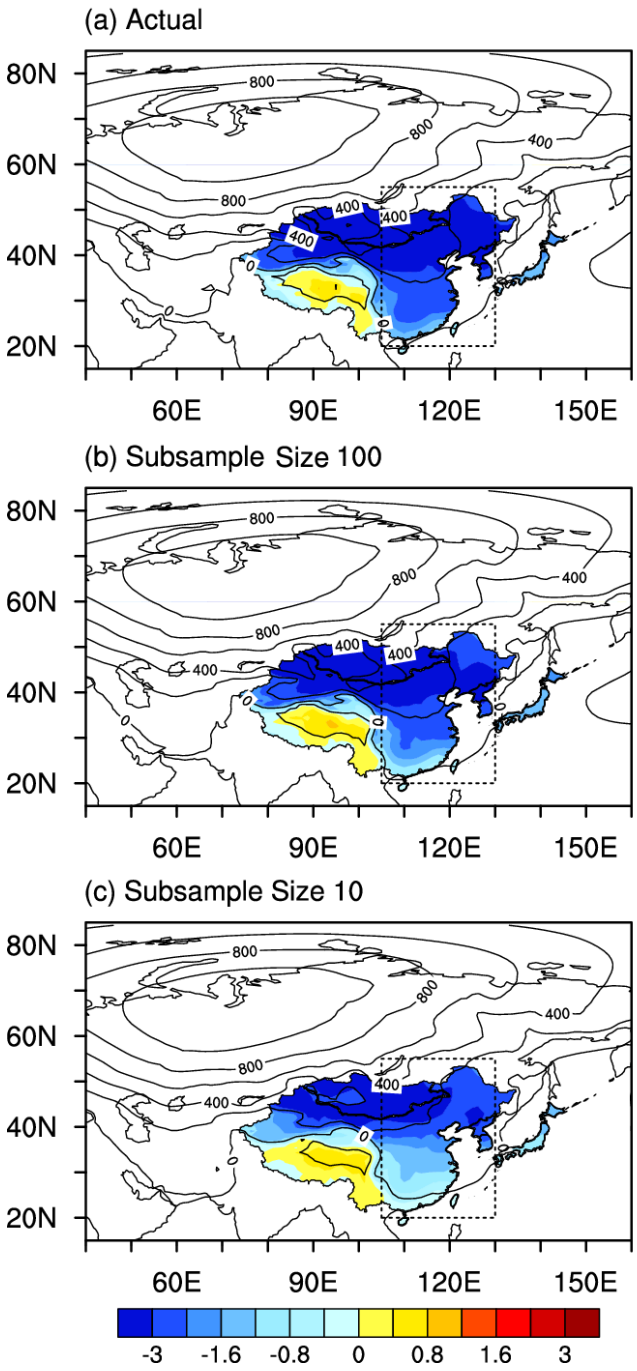

Figure A2 The composites of cold–month SAT anomaly in color shading and SLP anomaly in contours (relative to 1986-2005 boreal winter climatology) in CESM-LE during the period of 1986-2005 boreal winter: (a) total anomaly, (b) estimated dynamic component with subsample size of 100, (c) estimated dynamic component with subsample size of 10.

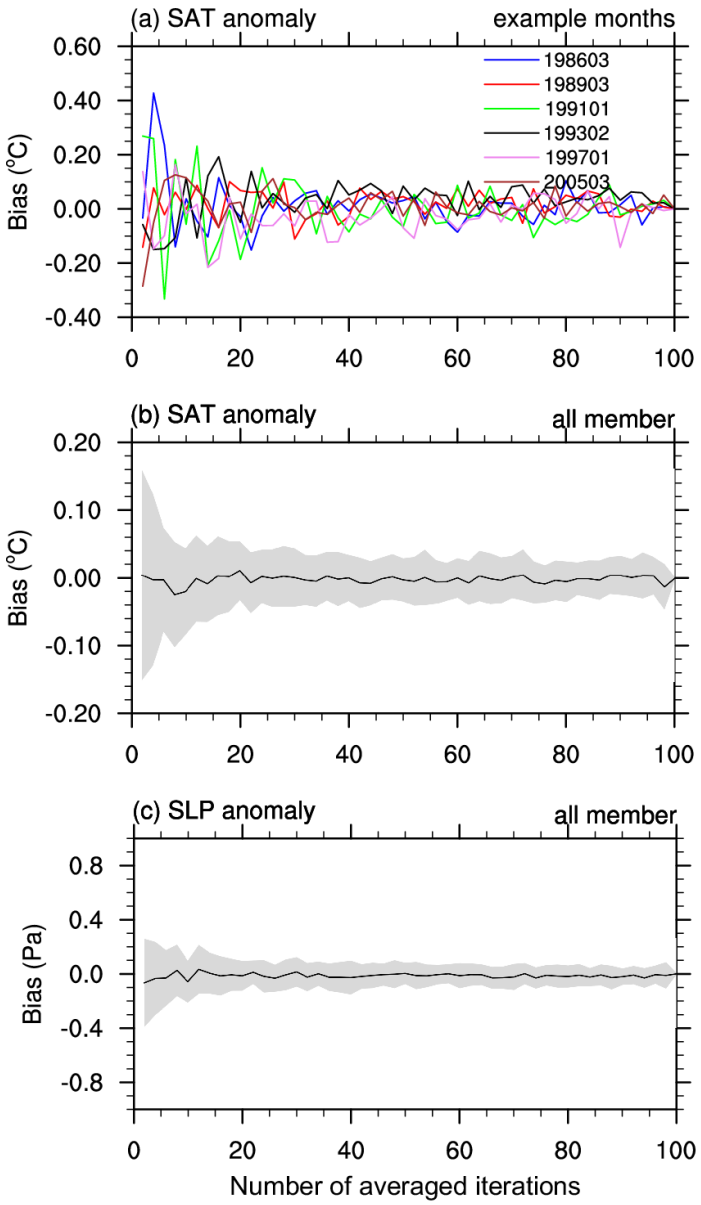

Figure A3 Differences between the results obtained at different numbers of iterations and the results obtained at 100 iterations for (a) example months selected from run 2 of CESM-LE. Subplots (b) and (c) show the result of cold extremes in 1986-2005. The shading shows the range of two standard deviations among the model members of CESM-LE.

4. *Readers may wonder whether the 150 fields totally capture the dynamical contribution or not. If not, the thermodynamical contribution would be overestimated. A discussion or simply analysis would help clarify.*

Response: Thank you for your suggestion. We used CESM-LE to check whether 150 fields can capture the dynamic contribution to cold extremes. 20, 50, 100, 150, 200, 300 fields are used for comparison. As shown in Figure A4, there is no significant difference when the number of selected fields exceeds 100. We have added this as "To test whether 150 selected SLP fields are sufficient to estimate the target SLP, a sensitivity analysis is conducted on the sample size of the selected closest fields. The findings suggest that there is no significant difference when the number of selected fields exceeds 100." Please see L160-163 in the revised manuscript.

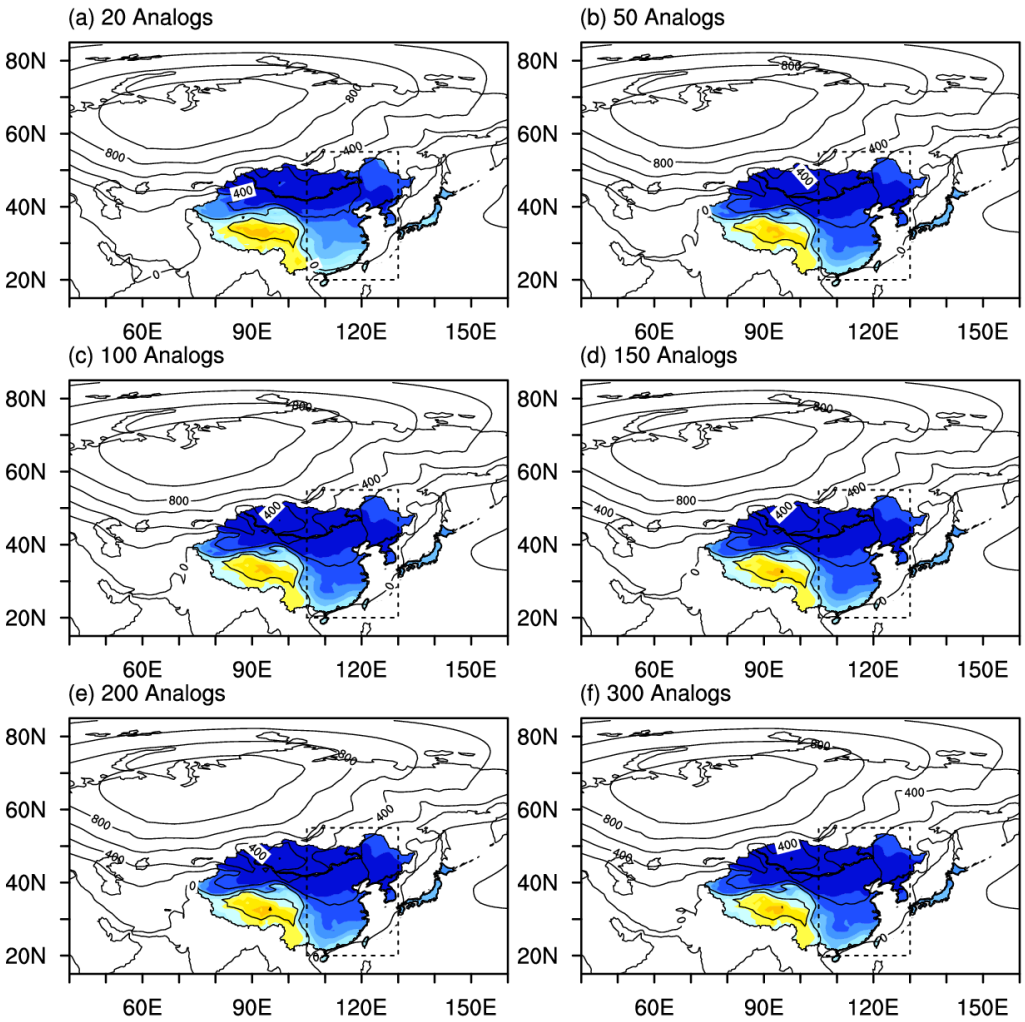

Figure A4 The composites of cold –month SAT anomaly in color shading and SLP anomaly in contours (relative to 1986-2005 boreal winter climatology) in CESM-LE

during the period of 1986-2005 boreal winter obtained from different numbers of selected SLP analogs: (a) 20 SLP analogs, (b) 50 SLP analogs, (c) 100 SLP analogs, (d) 150 SLP analogs, (e) 200 SLP analogs, (f) 300 SLP analogs.

5. *Section 2.3.2, How the dynamic adjustment approach is applied, reads a little complicated, it is better to present in a direct and simply way.*

Response: Thank you for your suggestion. In this section, two points should be emphasized. Firstly, because there is no PiCTL simulation in the observation, it is necessary to first construct an SAT series that is not influenced by external factors (similar to PiCTL simulation). Secondly, since there is only one member in the observation, a separate dynamic adjustment needs to be made for the internal dynamic component, instead of directly subtracting the ensemble mean, as is done in the models. All other steps are the same as those applied in the models.

We enhanced the connectivity of the statements to provide better context and improve the overall clarity of the procedures as follows.

"There is no PiCTL simulation in the observation. Therefore, before computing the dynamic component of SAT, the quadratic trend of the SAT during 1920-2012 is first subtracted to obtain SAT series without anthropogenic forcing. Similar to the application to the model ensembles, for each month and year in the observation, 40 SLP fields subsampled from 60 closest SLP fields are first selected (excluding the target month). Then, dynamic adjustment procedure described in section 3.2.1 is applied to derive the dynamically-induced SAT fields in the observation.

Different from model simulations, there is only one member in the observation, we cannot separate the forced and internal parts by calculating the ensemble mean or subtracting the ensemble mean. To obtain the internal dynamic contribution to the

observed SAT anomaly, a separate dynamic adjustment based on the internal component of the observed SLP anomalies is performed. It is worth noting that, the internal component of the observed SLP anomalies is obtained by subtracting the model ensemble-mean SLP anomaly from the observed SLP anomaly at each time step.

After we get the internal dynamic component of SAT anomaly, the forced dynamic component is calculated by subtracting the internal dynamic component from the total dynamic component. Thermodynamic components are obtained as residuals (total minus dynamic) for both forced and internal components."

Please see L170-187 in the revised manuscript.

References

Deser, C., Terray, L., & Phillips, A. S. (2016). Forced and internal components of winter air temperature trends over North America during the past 50 years: Mechanisms and implications. Journal of Climate, 29(6), 2237-2258. https://doi.org/10.1175/JCLI-D-15-0304.1