

Responses to Review Comments

Dear Editor,

Thank you and all the reviewers for their careful review and valuable comments. We have carefully addressed all of the comments and revised our manuscript accordingly. Below please find our item-by-item responses to all comments raised by two reviewers and the topic editor. Please note that the review comments are in black, **our responses** are highlighted in blue, **the extracts from the manuscript** are in red, and **the new texts that have been added or changed** are in bold. Thank you very much for your time.

Best regards,

Ming Luo and Yongquan Zhao, on behalf of all authors

Comments by Reviewer #1

Major comments.

The dataset is very import to study the climate change during this period over this region. The high spatial resolution and the monthly temporal resolution are basically reasonable, based on the previous datasets. The use of homogenous meteorological observation data is the right way to evaluate. This research is novel and belongs to the application of big data for large volume of Satellite data used, and machine learning algorithm. The authors are from 8 universities/colleges. The results supply important reference for other research fields like the field of geoscience, biology, sociology, and medicine and engineering.

Response: Many thanks for your appreciation of our work and the constructive comments. We have carefully revised our manuscript according to your comments and gave the following responses for your further review.

1) Line 415, harm, it is suggested to change the word. All 11 indices are calculated based on SAT, so the accuracy of this dataset is up to the accuracy of the dataset (Zhang 2022b) and the algorithm. SAT has important influence to the other 11 indices. How to understand these?

Response: Thank you for your comment. We have changed “harm” to “influence”. The 11 human thermal indices were calculated from SAT at a daily scale (see Table 2 in the main text) for each station. Then, the daily thermal indices were monthly averaged and used for training together with the other predictors (e.g., LST, land cover, topography, and temporal variation) to estimate the spatial distribution of the 11 monthly indices at the 1 km grids. Ideally, the spatial prediction should be conducted based on monthly or even daily SAT. However, the SAT observations with full spatial coverage are not available, so we used monthly LST as an alternative to predict the spatial distribution of the monthly human thermal indices. In this case, the accuracy of monthly LST may influence the accuracy of our HiTIC-monthly dataset. We thereby performed multiple

accuracy assessments at various spatial (e.g., individual stations, the mainland of China, and its urban agglomerations) and temporal scales (e.g., monthly, yearly, and overall). The assessment results indicate that our dataset shows a desirable performance and exhibits good agreement with the observations in both spatial and temporal dimensions (see Section 4.1 of the main text), demonstrating the broad applicability of our dataset.

2) (Zhang 2022b) SAT is LST, but not Tair (1.5 meters above the surface). Why use LST but not Tair in Table 1 to compute the other 11 indices?

Response: As shown in our response to your comment #1, we first computed daily thermal indices observed at the weather stations based on daily SAT and other meteorological observations (see Table 2), and then aggregated them to monthly values. These spatially-discrete (station-based) monthly indices were then used for predicting spatially-continuous human thermal indices on 1 km grids, which adopted seamless monthly LST (seamless SAT is not available) and other related variables as predictors. Previous studies have shown that LST is highly correlated with SAT and can be used as a proxy to predict SAT and SAT-related human thermal indices (Shamir & Georgakakos, 2014; W. Zhu, Lú, & Jia, 2013; X. Zhu, Zhang, Xu, Sun, & Hu, 2019). Our results also demonstrate that using LST can generate a desirable accuracy in predicting monthly human thermal indices.

3) Line 157-158 “Section 6 compares our products with two existing datasets, and the main findings of this paper are summarized in Section 7”. Where is section 6 and section 7? They are data availability and conclusions.

Response: Thank you for pointing out this issue. We have revised the description as (Lines 157–159): **“Comparisons on our products with two existing datasets are in Section 5, data availability is provided in Section 6. The main findings of this paper are summarized in Section 7.”**

4) Why not keep Figures S1-S9, Tables S1-S5 as formal ones? Please consider again which to keep in the manuscript.

Response: Thanks a lot for your advice. We have moved Tables S1–S2 to the main text in the current revision, and we decided to keep Figures S1–S9 and Tables S3–S5 in the supplementary material because there are already 14 figures and 4 tables in the main text. Including more figures and tables may not meet the requirement of the journal of ESSD and influence the readability of our manuscript.

5) How to get monthly data from MODIS daily mid-daytime 13:30 and mid-nighttime 01:30 LST? The monthly mean is different considering of the diurnal variation and satellite observations from ascending orbits and descending orbits.

Response: Thank you for the comment. Our monthly LST values were calculated by averaging daily LST in the corresponding month in the calendar years of 2003–2020. The daily mean LST values were obtained by averaging four MODIS observations in a day, which correspond to mid-daytime and mid-nighttime observations from ascending and descending orbits. In the current revised version, we have added the following explanations (Lines 178–183): **“We used monthly LST as one of the inputs to predict the spatial distribution of 12 thermal indices. Monthly LST values were calculated by averaging daily LST, which was obtained by averaging four observations in a day, including mid-daytime and mid-nighttime observations from ascending and descending orbits of MOD11A1 (Terra) and MYD11A1 (Aqua). More details about the LST data are described in Zhang et al. (2022b).”**

6) Line 190-191, how to compute and use the covariates?

Response: Thank you for the question. The covariates directly come from the data providers or producers, and these information is listed in Table 1. All these covariates

were pre-processed to have the same spatial extent, projection, and spatial resolution for predicting human thermal indices with full spatial coverage at the 1 km resolution.

7) HiTIC? What is it? Is it HTI (human thermal index)? No “I” after “human” and before “thermal”.

Response: “HiTIC” stands for **High** spatial resolution **Thermal Index Collection**. We have changed the title to “**HiTIC-Monthly: A Monthly High Spatial Resolution (1 km) Human Thermal Index Collection over China during 2003–2020**” to avoid misunderstanding.

Minor comments.

1) Add abbr. of LGBM when it is first used. Light Gradient Boosting Machine. Check others, please.

Response: Thank you for your suggestion. The full name of LGBM has been added (Lines 38–39): “**In this collection, 12 commonly-used thermal indices were generated by the Light Gradient Boosting Machine (LGBM) learning algorithm from multi-source gridded data.**”

2) Author contribution. It seems only H.Z is responsible for analysis and data processing, others are all involved in writing. For computations, are there anyone else in the author list?

Response: The first author H.Z. and two corresponding authors of M.L. and Y.Z. are responsible for data analysis and computation. We also wish to clarify that this work involves a large number of tasks (including but not limited to the selection of predictors, the accuracy assessment, the discussion and interpretation of the results, and the comparison with other datasets), and all other authors in the list contributed to discussion, investigation, and writing.

3) Line 439. The unit of the dataset is degree or 0.01 degree? Is 0.01 the scale factor?

Response: Yes, the unit of the dataset is 0.01 degree Celsius (°C). The values are stored in integer type (Int16) for saving storage and need to be multiplied by 100 when in use. The following updates have been made in the main text (Lines 472–474): **“The unit of the dataset is 0.01 degree Celsius (°C), and the values are stored in an integer type (Int16) for saving storage space, and need to be divided by 100 to get the values in degree Celsius when in use.”**

4) Please revise the followings.

- Line 247-248, as the figure displays

Response: Revised.

- Line 384, Figures S8j&m. “Figure S8”, not figures, and add a blank before j.

Response: Revised.

It is suggested to delete equations (1-2) and (3-6), add the reference.

Response: Thanks for the suggestion. With respect, we believe that these equations are helpful for readers to understand the details more easily, and thus can be kept in the main text. As suggested, we have added the relevant references in the current revision:

Lines 205–206: **“ E_a is derived from T and RH rather than directly observed at meteorological stations (Eqs. 1~2; Bolton (1980)).”**

Lines 238–239: **“Four statistic metrics, namely, determination coefficient (R^2), Mean Absolute Error (MAE), RMSE, and Bias (Rice, 2006).”**

- If necessary, please keep the same description of the time period for there are many kinds in the manuscript, like 2003~2020, from 2003 to 2020, during the year

2003~2020, (2003 to 2020), during 2003~2020, from January 2003 to December 2020, 2003-2020.

Response: Thanks for your suggestion. We have changed all descriptions of the time period to “**during 2003–2020**”.

- Line 755. Delete the last sentence. Line 759, change : to . after 2003~2020.

Response: Revised.

- Add longitude and latitude (or the numerical scope signs) in Fig 1, 4-7, 9-12, 14.

Response: Added.

Figure 8. Prediction accuracies of 12 human thermal indices... Add 12. In individual years, change the description. Time series? Add the description of (a) (b) (c).

Response: Added.

- Figure 10. Spatial distributions of 12 human thermal indices... Add 12.

Response: Added.

- Line 780. Figure 11. National average, are you sure to write like this? Check it in other places in the manuscript. Just use average is fine. Delete “straight”. Line 789, Figure 13, national, delete this. Add 12, 12 human thermal indices.

Response: Thank you. We have revised these sentences accordingly. For example, we have revised the caption of Figure 11 as (Lines 843–844): “**Temporal changes of the 12 annually-averaged human thermal indices over the mainland of China during 2003–2020.**”

- Line 785, Figure 12, add 12. “the trends of annual mean”, please consider how to express better. Inter-annual variation?

Response: Thank you for the comment. The sentence has been rewritten as (Lines 849–850): “**Spatial distributions of the linear trends (unit: °C per decade) in the 12**

annually-averaged human thermal indices over the mainland of China during 2003–2020.”

- Line 795, Figure 14. HITIC, is it “i”? in mainland China, is it over mainland of China? Please check the description of “mainland China” in the manuscript. In July 2018, move to the end of four major UAs. Delete i.e., change to :.

Response: Thank you for the comment. We have revised these sentences as follows (Lines 858-861):

“Comparison of the spatial patterns among HDI_0p25_1970_2018 (HDI), HiTiSEA, and **HiTIC-Monthly** for AT_{in} **over the mainland of China and its four largest UAs** in July 2018: Beijing-Tianjin-Hebei (BTH), Yangtze River Delta (YRD), middle Yangtze River Valley (mYRV) and Pearl River Delta (PRD). Colored circles indicate the observed AT_{in} values at individual meteorological stations.”

Comments by Reviewer #2

The authors have produced a high-resolution (1 km×1 km) thermal index collection at a monthly scale (HiTIC-Monthly) in China during 2003 to 2020, with 12 widely used human thermal indices. The authors have created a high-resolution products for quantifying thermal index in China, which is valuable to the scientific community. I have some comments to be addressed by the authors.

Response: Thank you very much for your helpful comments and suggestions. We have carefully revised our manuscript based on your comments. Below please find our item-by-item responses.

1. The biggest concern is the temporal resolution. Why do the authors choose monthly resolution, rather than daily? Daily products would be extremely useful to characterize extreme events, which are of societal importance.

Response: Thank you very much for your comment. We agree with you that daily products could be useful to characterize extreme events. However, a daily thermal indices dataset with full spatial coverage requires daily covariates that have full spatial coverage as well, and these data are not available. We produced a long-term monthly thermal indices dataset with high accuracies. The monthly dataset could be useful for the studies of climate change and urban climate and environment across China and has the potential for investigations of heat-related illnesses and deaths. A daily high-resolution human thermal index collection (HiTIC-Daily) will be produced and released in our future studies.

2. Lines 168-169: How about the impacts of precipitation on thermal indices? Have the authors considered precipitation as a covariate?

Response: Thank you for bringing our attention to precipitation. We did not include precipitation as a covariate because the precipitation data are not normally distributed.

More importantly, they exhibit many zero values in many regions of China (especially in the dry season), which would increase the uncertainty of the spatial prediction.

3. Figures 7 and 8: The results indicate spatial variability of bias in the thermal indices. What factors drive the spatial variability of the bias? Meanwhile, there is also temporal variability in the bias (Figure 8), and what is the drivers of this variability? Are the spatial and temporal variabilities of the bias related to background climates?

Response: Thank you for the comment. Yes, the spatial and temporal variabilities of the bias are related to background climates. As shown in Figure 7, the biases exhibit zonal variations across the space, i.e., positive bias values tend to distribute in northern China and negative values are mainly located in the south. The spatial variability is likely caused by the generally lower temperatures in the north and higher temperatures in the south. The extremely small values in the north and the extremely large values in the south may be overestimated and underestimated to some extent, respectively. The overestimation and underestimation issues caused by limited training samples of extreme values are quite common in machine learning (Cho, Yoo, Im, & Cha, 2020; Li, Li, Li, & Liu, 2020; Uddin, Nash, Mahammad Diganta, Rahman, & Olbert, 2022; Wu et al., 2022), due to limited samples of extreme low and high values compared to the rest of the samples. The temporal variability of the bias is relative to climate warming. Under climatic warming, the lower temperatures appear in early periods (e.g., 2003–2005) while relatively higher temperatures occur in more recent periods (e.g., 2016–2019). The overestimation of lower temperature in early periods and the underestimation of higher temperature in recent periods result in the temporal variability of the bias (Figure 8). It should be noted that, although the bias has spatial and temporal variabilities, these variations are quite small (i.e., ranging from $-0.3\text{ }^{\circ}\text{C}$ to $+0.3\text{ }^{\circ}\text{C}$). Overall, the estimations in our study are reliable (see the evaluation results in Section 4.1).

In this revision, we have added the following discussions:

Lines 282–290: “Positive *Bias* values tend to distribute in northern China while negative values are mainly located in the south. This spatial variability is likely caused by the generally lower temperatures in the north and higher temperatures in the south. In particular, the extremely small values in the north and the extremely large values in the south may be overestimated and underestimated to some extent, respectively, due to limited samples of extremely small and large values (compared with the rest of the samples) when training the machine learning model. The overestimation and underestimation issues caused by limited training samples of extreme values are quite common in machine learning (Cho et al., 2020; Li et al., 2020; Uddin et al., 2022; Wu et al., 2022).”

Lines 297–304: “*Biases* vary between -0.04 °C and 0.04 °C across all years. This temporal variability of the *Bias* is related to the yearly climate variations, and is characterized by a marginal overestimation of lower temperatures that mainly appeared in early periods (e.g., 2003–2005) and the underestimation of higher temperatures mostly in recent periods (e.g., 2016–2019). Under climatic warming over the past decades, the lower temperatures tended to appear in early periods while relatively higher temperatures more likely occurred in more recent periods. Extremely small values of temperature in earlier periods and the large values in the later periods may be slightly overestimated (i.e., with positive *Bias* values) and underestimated (i.e., with negative *Bias* values), respectively, thereby characterizing the temporal variations of the *Bias*.”

4. One way of evaluating the quality of these products is to evaluate the EOFs of these products. For example, what are the first three EOFs in each product? How do the temporal coefficients change over time across these products? Such spatial-temporal evaluation would be desirable.

Response: Thank you very much for this suggestion. We have followed your suggestion and evaluated the first three EOF modes of the 12 thermal indices. The spatial patterns

and corresponding temporal coefficients of these EOFs are depicted in Figures R1–R4 (see also Figure S10–S13 in the Supporting Information). As Figure R1 shows, the leading EOF (EOF1) of all 12 thermal indices exhibit highly consistent spatial distribution with higher values in the north and lower values in the south. Their temporal variations are also similar to each other (Figure R2). The second and third EOF modes (EOF2 and EOF3) are also similar among different thermal indices (except EOF3 of NET, Figures R2–R4). These results demonstrate the desirable quality of our products. We have thus included the following discussions in the revised manuscript (Lines 350-356):

“The dominant modes of these indices are further examined by applying the empirical orthogonal function (EOF) analysis (Figures S10–S13). As Figure S10 shows, the leading EOF (EOF1) of all 12 thermal indices exhibit highly consistent spatial distribution with higher values in the northern region and lower values in the south. Their temporal variations are also similar to each other (Figure S11). The second and third EOF modes (EOF2 and EOF3) are also similar among different thermal indices (except EOF3 of NET, Figures S11–S13). These results demonstrate the desirable quality of our products.”

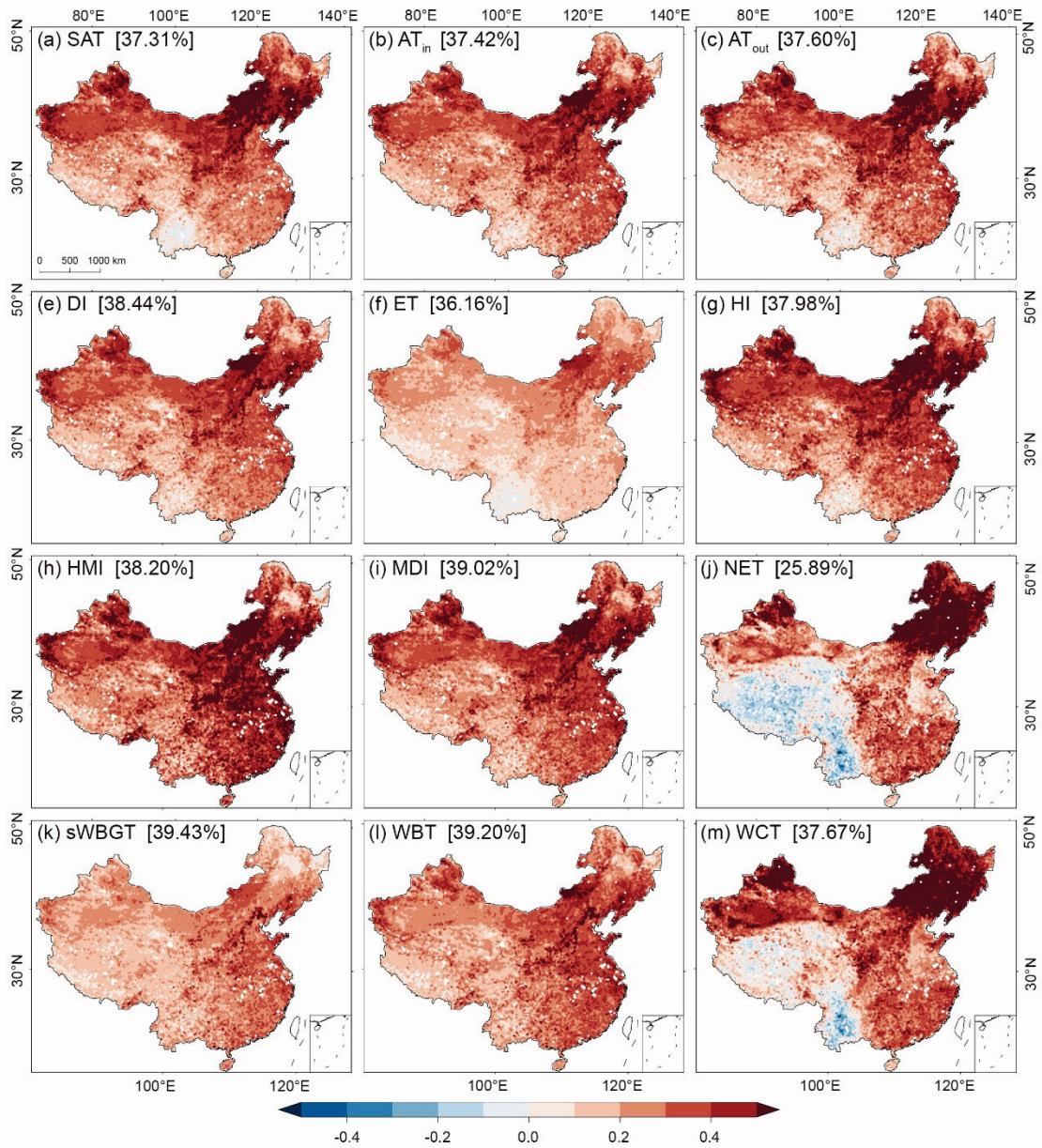


Figure R1. Spatial distributions of the leading empirical orthogonal function (EOF1) of the 12 human thermal indices over the mainland of China. The numbers in square brackets represent the percentage of the variance explained by the corresponding EOF mode.

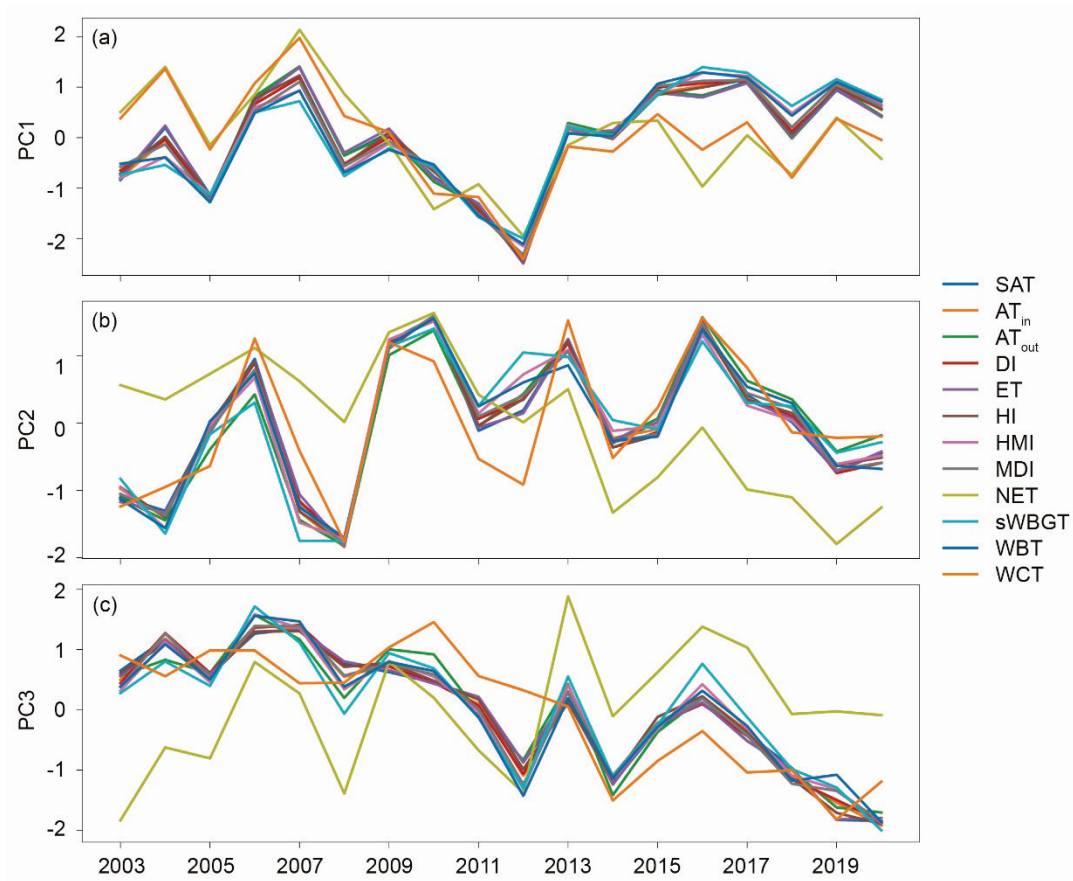


Figure R2. Time series of the principal components (PCs) corresponding to the first three EOF modes of the 12 human thermal indices over the mainland of China during 2003–2020.

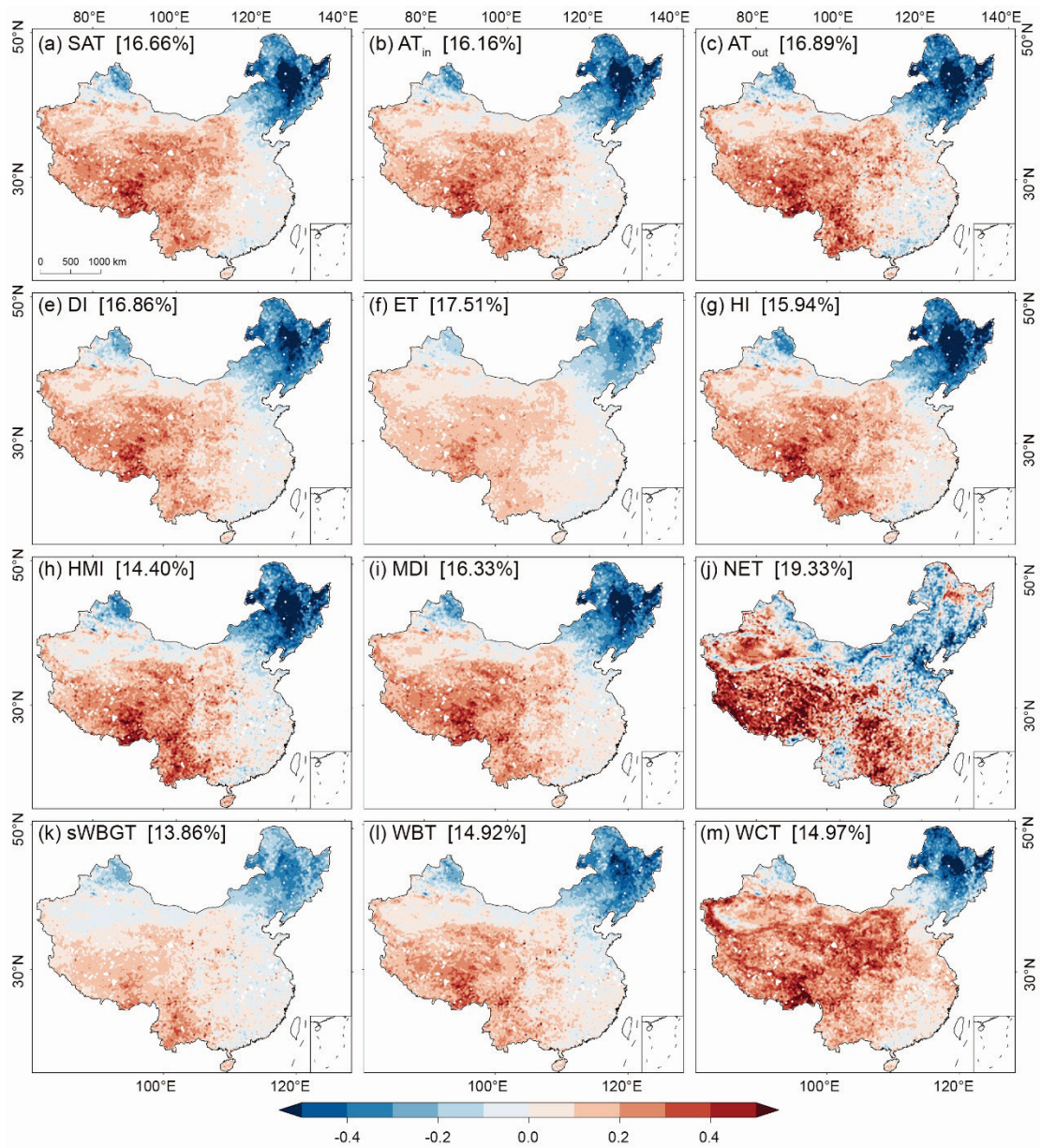


Figure R3. As Figure R1 but for the second EOF (EOF2).

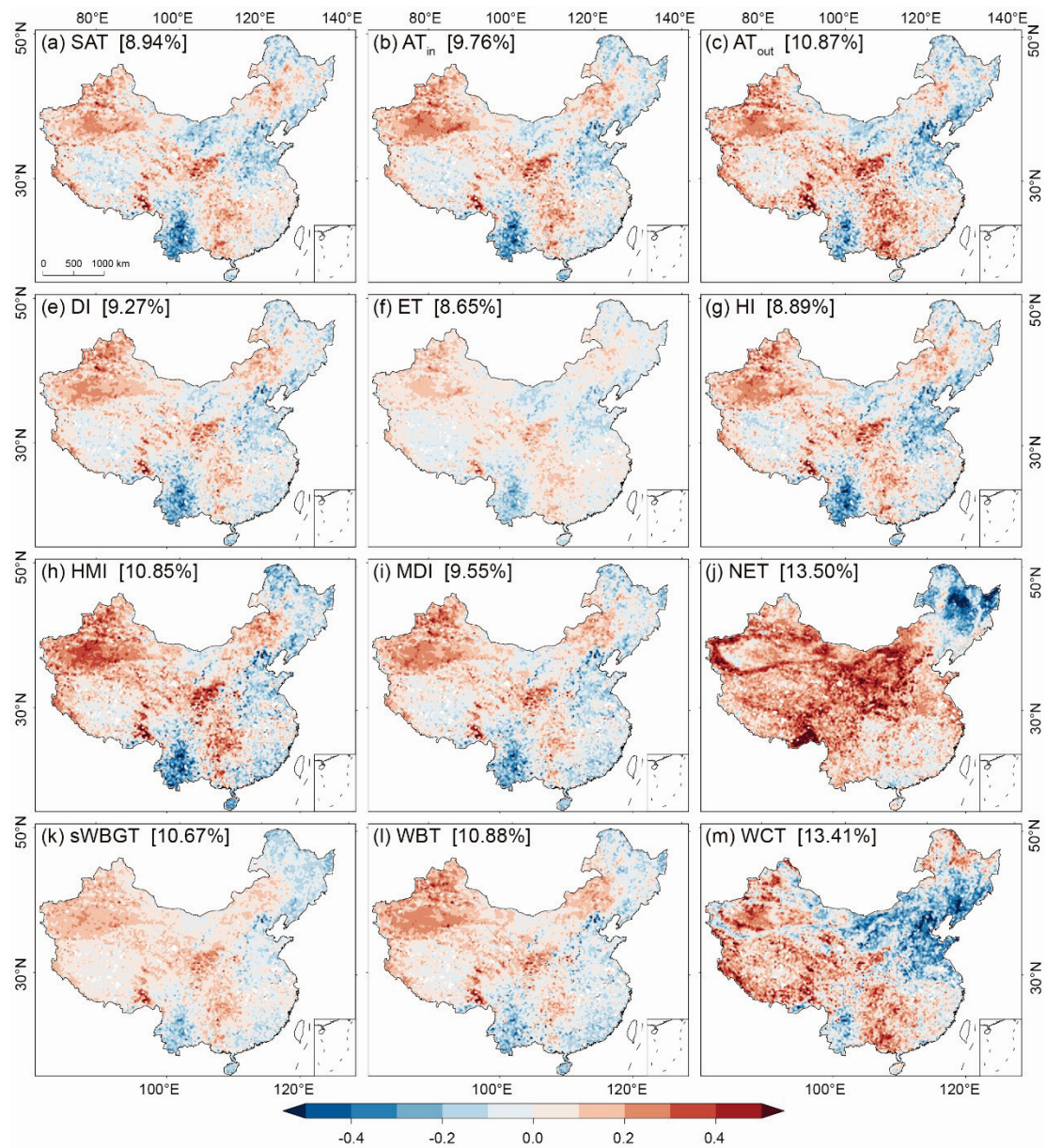


Figure R4. As Figure R1 but for the third EOF (EOF3).

Second-round Comments by Reviewer #1

1. Line 157-158, suggest: Comparisons on our products with two existing datasets are in Section 5, data availability is provided in Section 6, ...

Response: Revised per your suggestion.

2. The 3rd response to the reviewer2 (spatial variability, temporal variability, background climates), Please consider how to answer the question directly.

Response: Thank you very much for your comment. Below please see our updated and direct responses (please note that our response to reviewer #2 has also been updated accordingly in the current version):

Yes, the spatial and temporal variabilities of the bias are related to background climates. As shown in Figure 7, the biases exhibit zonal variations across the space, i.e., positive bias values tend to distribute in northern China and negative values are mainly located in the south. The spatial variability is likely caused by the generally lower temperatures in the north and higher temperatures in the south. The extremely small values in the north and the extremely large values in the south may be overestimated and underestimated to some extent, respectively. The overestimation and underestimation issues caused by limited training samples of extreme values are quite common in machine learning (Cho et al., 2020; Li et al., 2020; Uddin et al., 2022; Wu et al., 2022), due to limited samples of extreme low and high values compared to the rest of the samples. The temporal variability of the bias is relative to climate warming. Under climatic warming, the lower temperatures appear in early periods (e.g., 2003–2005) while relatively higher temperatures occur in more recent periods (e.g., 2016–2019). The overestimation of lower temperature in early periods and the underestimation of higher temperature in recent periods result in the temporal variability of the bias (Figure 8). It should be noted that, although the bias has spatial and temporal variabilities, these variations are quite small (i.e., ranging from -0.3 °C to

+0.3 °C). Overall, the estimations in our study are reliable (see the evaluation results in Section 4.1).

In this revision, we have added the following discussions:

Lines 282–290: “Positive *Bias* values tend to distribute in northern China while negative values are mainly located in the south. This spatial variability is likely caused by the generally lower temperatures in the north and higher temperatures in the south. In particular, the extremely small values in the north and the extremely large values in the south may be overestimated and underestimated to some extent, respectively, due to limited samples of extremely small and large values (compared with the rest of the samples) when training the machine learning model. The overestimation and underestimation issues caused by limited training samples of extreme values are quite common in machine learning (Cho et al., 2020; Li et al., 2020; Uddin et al., 2022; Wu et al., 2022).”

Lines 297–304: “*Biases* vary between -0.04 °C and 0.04 °C across all years. This temporal variability of *Bias* is related to the yearly climate variations, and is characterized by a marginal overestimation of lower temperatures that mainly appeared in early periods (e.g., 2003–2005) and the underestimation of higher temperatures mostly in recent periods (e.g., 2016–2019). Under climatic warming over the past decades, the lower temperatures tended to appear in early periods while relatively higher temperatures more likely occurred in more recent periods. Extremely small values of temperature in earlier periods and the large values in the later periods may be slightly overestimated (i.e., with positive *Bias* values) and underestimated (i.e., with negative *Bias* values), respectively, thereby characterizing the temporal variations of the *Bias*.”

Comments by Topic Editor

There is still a minor comment from one of the reviewers. In order to save time for the publication of the MS, I forward his opinion to you here. Please consider it.

The 3rd response to the reviewer 2 is still not directly answered. The question from the reviewer 2 is good, but the authors almost use the same answer on Oct 1st. What factors drive the spatial variability of the bias? What are the drivers of the temporal variability? Are the spatial and temporal variabilities of the bias related to background climates? If the answers are in the listed references, how about answer the questions directly and add some discussion or explanations in the manuscript?

Response: Thank you very much for your comment. Below please see our updated and direct responses. Please be reminded that, in the current version we have also updated our responses to the corresponding comment raised by reviewers #2 and #3.

Yes, the spatial and temporal variabilities of the bias are related to background climates. As shown in Figure 7, the biases exhibit zonal variations across the space, i.e., positive bias values tend to distribute in northern China and negative values are mainly located in the south. The spatial variability is likely caused by the generally lower temperatures in the north and higher temperatures in the south. The extremely small values in the north and the extremely large values in the south may be overestimated and underestimated to some extent, respectively. The overestimation and underestimation issues caused by limited training samples of extreme values are quite common in machine learning (Cho et al., 2020; Li et al., 2020; Uddin et al., 2022; Wu et al., 2022), due to limited samples of extreme low and high values compared to the rest of the samples. The temporal variability of the bias is relative to climate warming. Under climatic warming, the lower temperatures appear in early periods (e.g., 2003–2005) while relatively higher temperatures occur in more recent periods (e.g., 2016–2019). The overestimation of lower temperature in early periods and the underestimation of higher temperature in recent periods result in the temporal variability of the bias (Figure 8). It should be noted that, although the bias has spatial and temporal variabilities, these

variations are quite small (i.e., ranging from $-0.3\text{ }^{\circ}\text{C}$ to $+0.3\text{ }^{\circ}\text{C}$). Overall, the estimations in our study are reliable (see the evaluation results in Section 4.1).

In this revision, we have added the following discussions:

Lines 282–290: “Positive *Bias* values tend to distribute in northern China while negative values are mainly located in the south. This spatial variability is likely caused by the generally lower temperatures in the north and higher temperatures in the south. In particular, the extremely small values in the north and the extremely large values in the south may be overestimated and underestimated to some extent, respectively, due to limited samples of extremely small and large values (compared with the rest of the samples) when training the machine learning model. The overestimation and underestimation issues caused by limited training samples of extreme values are quite common in machine learning (Cho et al., 2020; Li et al., 2020; Uddin et al., 2022; Wu et al., 2022).”

Lines 297–304: “Biases vary between $-0.04\text{ }^{\circ}\text{C}$ and $0.04\text{ }^{\circ}\text{C}$ across all years. This temporal variability of *Bias* is related to the yearly climate variations, and is characterized by a marginal overestimation of lower temperatures that mainly appeared in early periods (e.g., 2003–2005) and the underestimation of higher temperatures mostly in recent periods (e.g., 2016–2019). Under climatic warming over the past decades, the lower temperatures tended to appear in early periods while relatively higher temperatures more likely occurred in more recent periods. Extremely small values of temperature in earlier periods and the large values in the later periods may be slightly overestimated (i.e., with positive *Bias* values) and underestimated (i.e., with negative *Bias* values), respectively, thereby characterizing the temporal variations of the *Bias*.”

References

- Cho, D., Yoo, C., Im, J., & Cha, D. H. (2020). Comparative Assessment of Various Machine Learning-Based Bias Correction Methods for Numerical Weather Prediction Model Forecasts of Extreme Air Temperatures in Urban Areas. *Earth and Space Science*, 7(4). doi:<https://doi.org/10.1029/2019ea000740>
- Li, Y., Li, M., Li, C., & Liu, Z. (2020). Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. *Sci Rep*, 10(1), 9952. doi:<https://doi.org/10.1038/s41598-020-67024-3>
- Shamir, E., & Georgakakos, K. P. (2014). MODIS Land Surface Temperature as an index of surface air temperature for operational snowpack estimation. *Remote Sensing of Environment*, 152, 83-98. doi:<https://doi.org/10.1016/j.rse.2014.06.001>
- Uddin, M. G., Nash, S., Mahammad Diganta, M. T., Rahman, A., & Olbert, A. I. (2022). Robust machine learning algorithms for predicting coastal water quality index. *J Environ Manage*, 321, 115923. doi:<https://doi.org/10.1016/j.jenvman.2022.115923>
- Wu, J., Fang, H., Qin, W., Wang, L., Song, Y., Su, X., & Zhang, Y. (2022). Constructing High-Resolution (10 km) Daily Diffuse Solar Radiation Dataset across China during 1982–2020 through Ensemble Model. *Remote Sensing*, 14(15). doi:<https://doi.org/10.3390/rs14153695>
- Zhu, W., Lü, A., & Jia, S. (2013). Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment*, 130, 62-73. doi:<https://doi.org/10.1016/j.rse.2012.10.034>
- Zhu, X., Zhang, Q., Xu, C. Y., Sun, P., & Hu, P. (2019). Reconstruction of high spatial resolution surface air temperature data across China: A new geo-intelligent multisource data-based machine learning technique. *Sci Total Environ*, 665, 300-313. doi:<https://doi.org/10.1016/j.scitotenv.2019.02.077>