Response to comments

Paper #: essd-2023-220

Title: A coarse pixel scale ground "truth" dataset based on the global in situ site measurements to support validation and bias correction of satellite surface albedo products

Journal: Earth System Science Data

Thank you for providing us with so many valuable suggestions and they do help improve the paper. According to the reviewers' comments and suggestions, we revised the paper carefully and tried to give satisfactory answers to the reviewers' questions. The corresponding modifications are highlighted in red font in the revised paper.

The summaries of the revision for this paper are as follows:

First, we have reorganized the data and added all available sites. Moreover, parts of results and discussion, main findings and conclusion, as well as the abstract were rewritten based on the complete dataset.

Second, the necessity for upscaling models was further elucidated by integrating the work of other researchers in **Introduction** and **Conclusion**. Furthermore, we discussed the applicability of upscaling models at various sites and provided an objective statement about the role and significance of the pixel scale ground "truth" dataset. Its relationship with existing satellite albedo products and ground measurements was also explained.

Third, we have added the quantification of uncertainty of upscaling models for each site in **Section 4.1**. Moreover, we have described how we addressed the issue of varying footprint sizes at distinct sites, as well as the rationale for implementing ETM+ imagery.

Fourth, we have explained the spatial and temporal resolution of the different data used in the methodology and conclusions, and added a detailed description of the illumination geometry, including black-sky albedo, and white-sky albedo, for the albedo products used. Additionally, we have clarified the sample size for the boxplots and re-examined the implications regarding sample size in the **Results** and **Discussion** section.

Fifth, we have explained the reason for the methodology Section being similar to those of Wu et al.(2020), and emphasized the importance of the content of our work.

Sixth, we have corrected typing errors; complemented supporting evidence and literature; improved charts and figures; and corrected spelling and grammatical errors in this paper.

For the specific comments for each reviewer, we have made a detailed reply as follows.

Reviewer #2

The authors constructed a global albedo database in coarse pixel scale based on the high-resolution Landsat7 ETM+ images and 368 *in situ* sites from sparsely distributed observation networks globally. The results showed that the new database overcomes the shortcoming of *in situ* albedo measurements and can be

used as ground truth, which captures spatiotemporal variations of surface albedo. However, there are many mistakes in the current manuscript which are due to the carelessness of the authors. Moreover, some parts of the content have indications of plagiarism.

Therefore, before the current manuscript can be published, the authors should reply to the following comments diligently.

As described by the authors, one criterion of the methodology in this manuscript is the spatial resolution of high-resolution albedo observation should be equivalent to the footprint of in-situ observation (lines 205-207). However, the authors also highlighted that the footprints of in-situ sites are not fixed. It depends on the height of the albedometers (Lines 113-115). Have the authors compared the size of the footprints of a total of 368 in-situ sites with that of the Landsat7 ETM+ (30 m)? How about the results? Please discuss this issue with figures or tables.

Re: The footprint of *in situ* sites is a function of measurement heights of the albedometers from the underlying surface and the field of view of the sensors. The former typically depends on the height of tower and height of the canopy top (different at different time), which are generally different from one site to another. The latter is not fully consistent due to the ideal and non-ideal cosine response of the sensors (Balzarolo et al., 2011; Cescatti et al., 2012; Song et al., 2019; Marion, 2021). Therefore, the footprints of *in situ* sites are not fixed. However, it is difficult to make a comparison between the footprints of *in situ* observation and the spatial resolution of high-resolution albedo observation. Because the footprints of *in situ* sites are various. Even for the same site, the footprint of *in situ* site is not consistent at different time due to the change of underlying surface (e,g., vegetation growth). But the effect of the spatial scale difference between *in situ* measurements and high-resolution data is believed to be negligible since the selection of high resolution data follows strict rules:

First, its spatial resolution should be minimal to maintain surface homogeneity within the fine pixel scale and ensure stable radiation acquisition.

Second, according to the albedo data observed at the FLUXNET site, approximately 80% of the energy in the observed signal originates from within 10-20 meters of the flux tower (Cescatti et al. 2012; Wang et al., 2014). Hence, the spatial resolution of the data should be near the footprint of *in situ* sites.

Third, since the upscaling coefficients were determined by long-time series high-resolution albedo maps and then were applied to long time series *in situ* measurements, the high-resolution albedo maps should cover at least one full cycle period, typically a year, to account for seasonal changes in surface heterogeneity caused by phenology and to guarantee the stability of the upscaling coefficients.

For these reasons, the Landsat ETM+ albedo data were adopted in this study. In the revised manuscript, we have added these explanations in **Section 3.1**.

In the manuscript, the coarse spatial resolution of the albedo product is 500 m (MCD43A3 V061) and the high resolution of the albedo is 30 m (Landsat7 ETM+). Therefore, the authors retrieved the upscaling coefficients to upscale the surface albedo from a high resolution of 30 m to a coarse resolution of 500 m. However, since 500 cannot be divided by 30, there should be some high-resolution observations partially covered at the edge of the coarse grid. How to deal with this issue? Please explain.

Re: In fact, we have used the 17×17 ETM+ pixels (an approximately 510 m ×510 m area) centered at MODIS pixel to calculate the pixel scale ground "truth". Namely, the spatial resolution of the ground "truth" is 510 m. The difference between the spale scale of MCD43A3 V061 and pixel scale ground "truth" is negligibly small, because the spatial response is very small at the margin areas of the pixel (Peng et al., 2015). To clarify this point, we have added the sentence as "*Secondly, it facilitated coarse pixel-level aggregation within a 17×17 window (an approximately 510 m ×510 m area, considered as a coarse scale pixel), serving to be the reference value of the coarse pixel albedo.*" in **Section 2.2**.

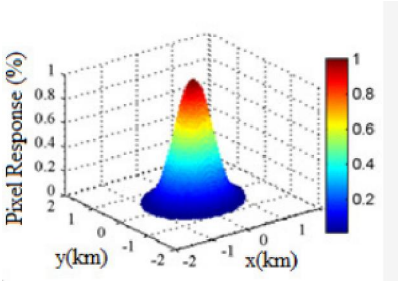


Figure. The point spread function of MODIS albedo products (Peng et al., 2015).

References:

Peng, J., Liu, Q., Wang, L., Liu, Q., Fan, W., Lu, M., and Wen, J.: Characterizing the Pixel Footprint of Satellite Albedo Products Derived from MODIS Reflectance in the Heihe River Basin, China, Remote Sensing, 7(6), 6886-6907, https://doi.org/10.3390/rs70606886, 2015.

Figure 3: The label of x-axis is wrong. According to line 292, Fig. 3 is the scatter plot of $\theta_{upscaling}$ and $\theta_{reference}$, none of them should be the "Pixel scale ground truth". Please check.

Re: Great thanks for pointing out this mistake. The mistake has been corrected as:

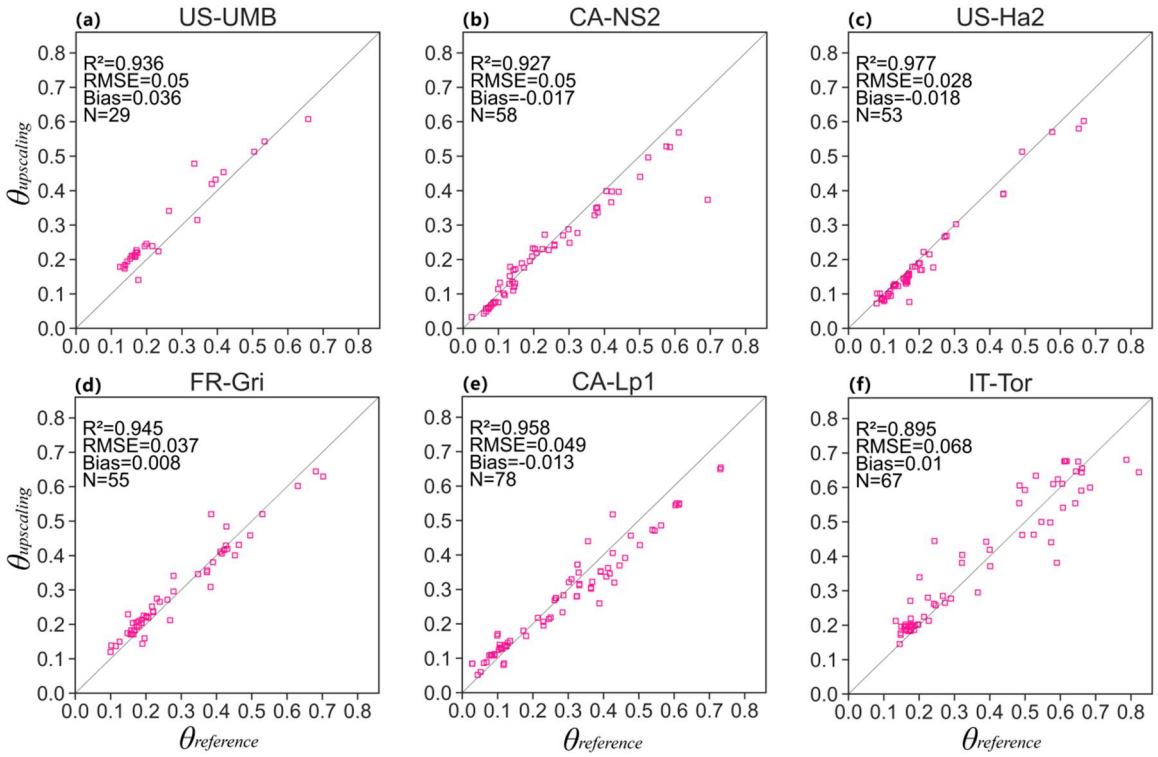


Figure 3: The scatter plots between the upscaling results ($\theta_{upscaling}$) with the upscaling models and the coarse pixel scale reference ($\theta_{reference}$).

Meanwhile, the six subpanels represented six land cover types according to the caption of Fig. 3. However, the authors didn't mention their locations (Lon/Lat) as well as the land cover types. Please add.

Re: We have added information about the *in situ* sites that correspond to the six subpanels in **Section 4.1.**

Table 1: Description of the *in situ* sites used in the model performance analysis.

Networks	US-UMB	CA-NS2	US-Ha2	FR-Gri	CA-Lp1	IT-Tor
Location(lon, lat)	(-84.7138, 45.5598)	(-98.5247, 55.9058)	(-72.1779, 42.5393)	(13.51259, 50.9500)	(-122.8414, 55.1119)	(7.5781, 45.8444)
Spatial heterogeneity	0.0133079	0.0640852	0.0065224	0.5564959	0.18694994	1.01929451
Elevation(m)	236.72682	271.09771	367.29669	377.65914	749.265564	2162.78979
Land cover type	DBF	EBF	MF	CRO	WSA	GRA

Figure 4: Please add line x=0 in the subpanel of Bias. Meanwhile, I don't agree with the expression "the biases concentrated around 0" in the conclusion (Line 482). Please revise the relevant content.

Re: As suggested by the reviewer, the line x=0 in the subpanel of Bias in Figure 4 has been added.

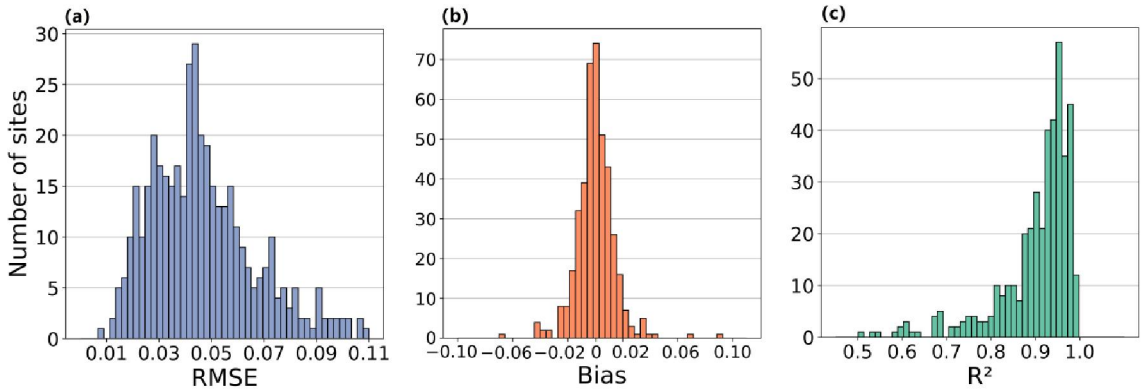


Figure 4. Distribution of RMSE (a), Bias (b), and R^2 (c) of the upscaling coefficients. The histograms presented here combine the results of the 416 *in situ* sites.

The expression "*the biases concentrated around 0*" in the Conclusion has been revised. The related sentence has been rephrased as "*The suitability of the upscaling model for applying to the in situ measurements was initially evaluated globally. The upscaling coefficients displayed an acceptable overall accuracy, with 90 % of bias following a normal distribution within the range of ± 0.02.* ".

Figures 5-9: I cannot find the description of the mean of the boxplot. What is the meaning of the line in the center box? The mean of median value? Please describe it clearly.

Re: As suggested by the reviewer, we have added the description of the mean of the boxplot. The black lines denote the median values. Taking Figure 6 as an example, the revised figure is shown as follows.

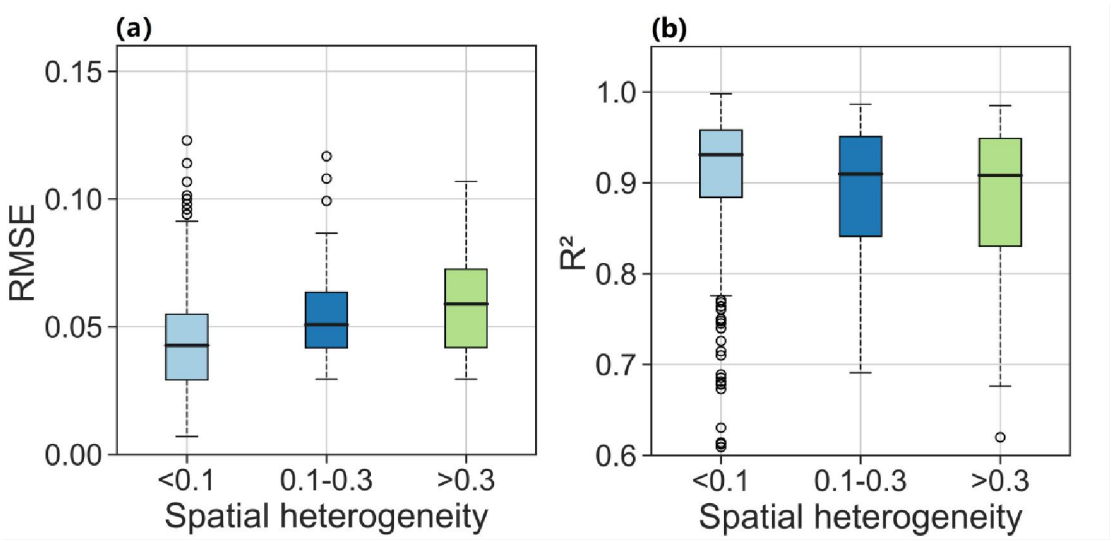


Figure 6: Boxplots showing the dependence of RMSE (a) and R^2(b) of the upscaled albedo on spatial heterogeneity. Three different degrees of spatial heterogeneity are marked by different colors. Black lines indicate median values. Outliers are values that are farther than 1.5 interquartile ranges. The accuracy response of the upscaling model to different spatial heterogeneity. The number of *in situ* sites with spatial heterogeneity of [0,0.1], [0.1-0.3], and [0.3-1.5] are 337, 49, and 30, respectively.

Meanwhile, what's the sample number of each boxplot? Please add the description and tables.

Re: In the revised manuscript, we have added the number of *in situ* sites for each level of spatial heterogeneity (Figure 6), each level of elevation (Figure 7), and each land cover type(Figure 9).

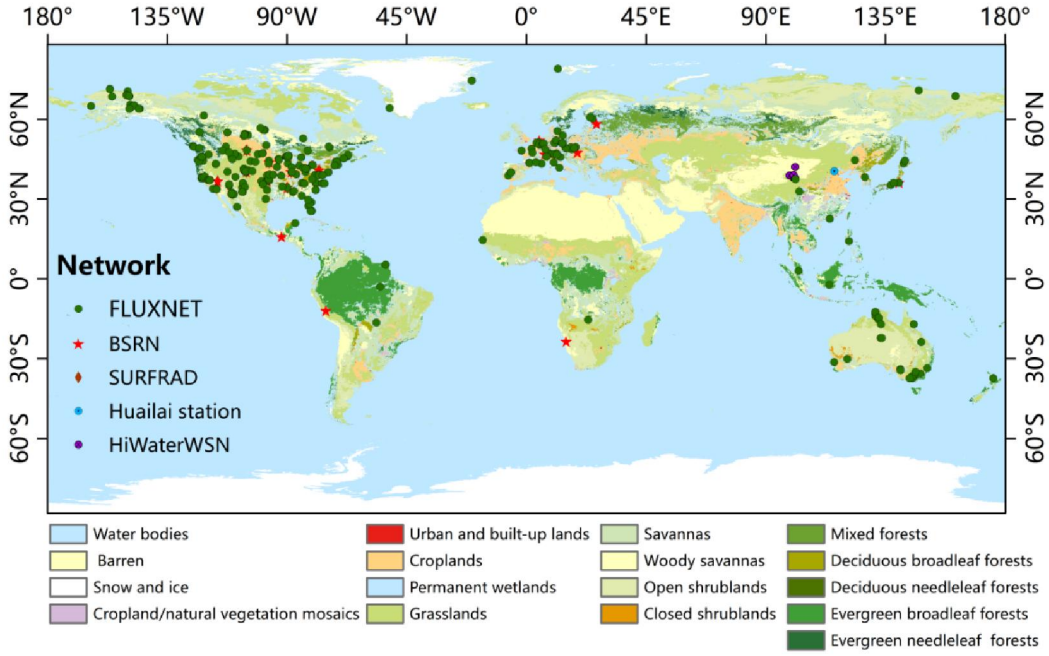
Lines 225-226: how to choose the $\theta_{ETM+\_in\_situ}$? Do you mean the nearest Landsat7 ETM+ pixel to the *in situ* site? Please explain.

Re: $\theta_{ETM+\_in\ situ}$ denotes the ETM+ pixel albedo time series containing the *in situ* site. Namely, it refers to the ETM+ pixel in which *in situ* site is located.

According to Fig. 1, there is a large portion of regions without *in situ* sites, especially for the regions covered with snow (e.g., Siberia) or with high elevation (e.g., Tibet). Therefore, how can the authors announce that their database can be used globally (in the abstract and conclusion)? Please explain.

Re: In the revised manuscript, we have added the *in situ* albedo measurements over Australia in the revised manuscript. Moreover, the *in situ* measurements over Siberia and other regions with effective measurements were also included in the dataset. The number of *in situ* sites increased to 416 for the dataset. It is true that the number of *in situ* sites is more than 416 within the globe. However, some sites were

145 excluded either due to the lack of incoming radiation information or the small data size after quality control. The distribution of these *in situ* sites is shown as follows. Given that these *in situ* sites are widely distributed on the globe and cover a wide range of environmental conditions (atmospheric model, aerosol model, spatial homogeneity and heterogeneity, temporal variation characteristics), they were believed to be representative of the globe.



150

Figure 1. The distribution of the 416 *in situ* sites over different land cover types.

Results and Discussion: The bias and RMSE of the upscaling results seems equivalent to the typical uncertainty of the surface albedo coarse resolution satellite products. Why are the authors satisfied with
155 their results? Please explain.

Re: It is true that the upscaling model itself has errors because it suffers from its own source of uncertainty. Therefore, over homogeneous surfaces where *in situ* site measurements are spatially representative, using this upscaling model will bring no benefits or even counteract due to the errors of the upscaling model. Nevertheless, over heterogeneous surface where *in situ* sites are lack of spatial
160 representativeness, the benefits outweigh disadvantages. The accuracy assessment results of pixel scale ground "truth" dataset demonstrate that the accuracy of reference data can be enhanced by 17.09 % over the regions with strong spatial heterogeneity. However, the degree of improvement with this dataset displays a decreasing trend as the reduction of spatial heterogeneity. In order to clarify this point, we have added the paragraph "*……For instance, the in situ measurements can be directly used as the pixel scale reference over*
165 *homogeneous surfaces or in the case that the satellite acquisition and in situ measurement footprints are similar, and the upscaling model is not necessary as it has its own source of uncertainty. But the upscaling model is useful for heterogeneous areas when in situ measurement footprints are less than satellite pixel size, because it increases the representativeness of the sampling for direct validation. The accuracy assessment results of pixel scale ground "truth" dataset demonstrate that the accuracy of reference data can be*

enhanced by 17.09 % over the regions with strong spatial heterogeneity……" in **Conclusion**.

As regards to the accuracy of the current coarse resolution surface albedo satellite products, their accuracy (between 0.03 and 0.05) is usually assessed over relatively homogeneous land surfaces. And the validation works over heterogeneous are still rare currently. The spatial scale mismatch over heterogeneous surfaces remains to be challenging to fully understand the overall accuracy of satellite products in different areas. Hence, our dataset can be considered as an important addition to the reference data on the coarse pixel scale over heterogeneous land surfaces.

Methodology: The content and the structure of the methodology in the current manuscript are quite similar to those of Wu et al., (2020). I also find the reference "Peng et al. (2015)" in line 240 is not included in the References part of the current manuscript. So, I believe the author who wrote the current manuscript plagiarized the whole content of methodology from Wu et al., (2020) and just modified some keywords. I leave the decision to the editor to decide whether to reject the current manuscript.

Re: We really appreciate the rigorous scientific attitude of the reviewer. In fact, the upscaling methodology of Wu et al., (2020) was developed by our research group, and the authors of Wu et al. (2012) are also the main contributors to this paper. However, the paper of Wu et al. (2020) merely proposed the upscaling method and did not comprehensively assess the effectiveness of this upscaling method. Moreover, this upscaling method has never been applied to the single *in situ* site measurements of the sparsely globally distributed observation networks (e.g., SURFRAD, BSRN, and Fluxnet) except for Huailai and Heihe River Basin, China. As a result, its transferability to *in situ* sites all over the world is still unknown. As the continuation and deepening of our previous work (Wu et al., 2020), this study puts emphasis on the comprehensive evaluation and extensive use of this upscaling method. Furthermore, a coarse pixel scale ground "truth" dataset was provided for validation and bias correction of satellite surface albedo products.

To counter and prevent misunderstanding, we have added the sentence as "*To overcome the representative errors of in situ measurements and promote utilization ratio of in situ sites from these sparse networks in validation, Wu et al. (2020) have proposed an upscaling method specified for the single site in situ measurements. However, the effectiveness of this method has not been comprehensively assessed and its transferability to in situ sites all over the world is still unknown. As the continuation and deepening of our previous work (Wu et al., 2020), this study puts emphasis on the comprehensive evaluation and extensive use of this upscaling method based on 416 in situ sites throughout the world. Furthermore, a coarse pixel scale ground "truth" dataset was provided for validation and bias correction of satellite surface albedo products. The potential usage of this dataset was also discussed.*" in **Introduction** of the revised manuscript.

The reference of Peng et al. (2015) has been added to the reference list.

The current manuscript should be polished before resubmission.

Re: Great thanks for the comment. The manuscript has been polished by a native speaker.

Minor comments:

Please check the number of equations throughout the manuscript. I found two "equation (4)" and "equations (10-12)". Moreover, I found the size of the equation numbers is different. Please explain the reason.

Re: We have corrected these errors in the revised manuscript.

Line 220: the right side of this equation is wrong. A comma is missing in the subscript. Please refer to the paper Wu et al., (2020), and fix it.

Re: This mistake has been corrected.

Line 226: the size of the words "indicates the" is smaller than the others, please explain the reason.

Re: The font size has been made consistent.

Line 237: What does the $\theta_{\text{in situ}}$ stand for? Please describe it in the main content clearly.

Re: $\theta_{\text{in situ}}$ denotes *in situ* site measurement. To describe it more clearly, this sentence has been revised as "*When the upscaling coefficients were determined, they were applied to in situ site measurements ($\theta_{in\ situ}$) to simulate the in situ reporting of surface albedo ($\theta_{in\ situ\_ETM+}$)……*" in the revised manuscript.

Line 240: I cannot find the reference "Peng et al. (2015)" in your "References".

Re: The reference of Peng et al. (2015) has been added to the reference list.

Line 273: the metric "coefficient of determination (R2)" was introduced in line 269, but the equation only gave "R". Please explain the reason.

Re: The coefficient of determination (R^2) was employed in this paper. The equation (15) has been revised as:

$$R^2 = \frac{[\sum_{d=1}^{L}(\theta_{\text{upscaling}}(d)-\overline{\theta_{\text{upscaling}}})(\theta_{\text{reference}}(d)-\overline{\theta_{\text{reference}}})]^2}{\sum_{d=1}^{L}(\theta_{\text{upscaling}}(d)-\overline{\theta_{\text{upscaling}}})^2 \sum_{d=1}^{l}(\theta_{\text{reference}}(d)-\overline{\theta_{\text{reference}}})^2} \quad (15)$$

Line 292: Please make sure it is "Fig.2" or "Fig. 3"? The same problem also can be found in line 299 (Fig. 3 or Fig. 4).

Re: The formulation (e.g., Fig. 2, Fig. 3, Fig. 4) has been made consistent throughout the paper.

Line 335: the lowest RMSE around "0.3"? Are you sure?

Re: We are sorry for this mistake. It should be 0.03. We have thoroughly checked the revised manuscript to avoid typos.