Dear reviewer,

We are very pleased to finish a revised version of the manuscript essd-2023-1 entitled "**Spatiotemporally consistent global dataset of the GIMMS Normalized Difference Vegetation Index (PKU GIMMS NDVI) from 1982 to 2022**". In preparing this revision we have considered all your comments and incorporated most of the suggestions. **Temporal coverage of PKU GIMMS NDVI has been extended from 2020 to 2022.** We greatly appreciate your time and effort spent in reviewing this manuscript, which have improved the revised version of the manuscript.

Substantial improvements have been made based on your comments, including:

(1) We have provided more details on how we performed the time-weighted aggregation method to convert the temporal resolution of the MODIS NDVI product (MOD13C1) from 16 days to half-month.

(2) We have also elaborated the method used to splice the PKU GIMMS NDVI and MODIS NDVI.

Below we provide point-to-point responses, each following the specific comment from the reviewer. All the changes have been marked by red in the revised manuscript.

Sincerely yours,

Zaichun Zhu, Ph. D. (on behalf of the author team)

School of Urban Planning and Design

Peking University

Tel: 86 185 0042 6608

Email: zhu.zaichun@pku.edu.cn

[Comment 1] *This study proposes PKU GIMMS NDVI, a new global long-term NDVI time series data that covers 1982 to 2020 based on AVHRR and MODIS sensors onboard satellite platforms. The PKU GIMMS NDVI extends the GIMMS NDVI3g data and has better data quality. It has better agreement with Landsat NDVI compared to GIMMS NDVI3g, alleviating the orbital drift problem in the AVHRR sensors. The method proposed in this study could be used to generate consistent global NDVI data in the future, which would help study global terrestrial biosphere dynamics.*

[Response 1] We thank the reviewer for helping improve our manuscript. We believe that the framework proposed in this study, which employs massive high-quality Landsat NDVI samples and a data consolidation method, could benefit future work that aims to generate long-term remote sensing-based earth data. We hope that our PKU GIMMS NDVI could provide a more accurate vegetation monitoring in the context of global environmental changes.

[Comment 2] *Detailed comments:*

Line 14: "global-wide" may be simplified as "global"

[Response 2] We thank the reviewer for pointing it out. The term "global-wide" has been simplified as "global" in the new version of the manuscript.

The following changes are made in the revised manuscript:

(a) Abstract Section:

"In this study, we presented a machine learning model that employed massive high-quality global Landsat NDVI samples and a data consolidation method" (Page 1, Line 14-15)

[Comment 3] *Line 26: When introducing NDVI, please cite the original NDVI paper: Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA Spec. Publ, 351(1), p.309.*

[Response 3] Thank you for providing the original NDIV paper from Rouse et al. (1974). It is now cited in the manuscript.

The following changes are made in the revised manuscript:

(a) First paragraph of the Introduction Section:

"The Normalized Difference Vegetation Index (NDVI) characterizes the biophysical, biochemical, and physiological conditions of vegetation (Rouse et al., 1974; Rondeaux et al., 1996; Gao et al., 2000; Yin et al., 2022)." (Page 1, Line 28-29)

[Comment 4] *Line 97: Please provide literature references to the data sources.*

[Response 4] Thanks for the suggestion. The literature references were originally available under sub-sections of specific data. Now, they have also been added in the data overview.

The following changes are made in the revised manuscript:

(a) First paragraph of the Data Section:

"Four global satellite products were used in this study: Landsat Surface Reflectance data (Collection 1 Tier 1) (Masek et al., 2006; Vermote et al., 2016), MODIS Land-Cover Type product (V6.1) (Friedl et al., 2002), GIMMS NDVI3g product (V1.0) (Pinzon and Tucker, 2014), and MODIS NDVI product (V6.1) (Didan, 2021)." (Page 4, Line 101-103)

[Comment 5] *Line 126: How is the time-weighted aggregation performed? Please explain in detail.*

[Response 5] In the revised manuscript, we have provided more details on the time-weighted aggregation method, which converted the temporal resolution of the MODIS NDVI product (MOD13C1) from 16 days to half-month. The method was adopted from Zhu et al. (2013). Its central idea is to assign weights to all MOD13C1 scenes that could temporally intersect with a particular half-month interval, where the weight depends on the possibility of intersection. The half-month NDVI product was finally calculated as the weighted sum of the scenes. More details including schematic illustrations can be found in Zhu et al. (2013).

The following changes are made in the revised manuscript:

(a) Introduction to MOD13C1 in Section 2.3:

"To match the temporal and spatial resolutions, we first performed a time-weighted

aggregation method on MOD13C1 to produce an NDVI product at a temporal resolution of half-month. The method was adopted from Zhu et al. (2013). Its central idea is to assign weights to all MOD13C1 scenes that could temporally intersect with a particular half-month interval, where the weight depends on the possibility of intersection. The half-month NDVI product was finally calculated as the weighted sum of the scenes. We then performed nearest neighbor sampling to upscale the spatial resolution to 1/12°." (Page 5, Line 134-139)

[Comment 6] *Line 127: Maybe the authors want to say "upscale" instead of "downscale."*

[Response 6] The reviewer is right. The spatial resampling from 0.05° to 1/12° (0.083°) should be an upscaling process. Thanks for this and we have modified the wording in the revised manuscript.

The following changes are made in the revised manuscript:

(a) Introduction to MOD13C1 in Section 2.3:

"We then performed nearest neighbor sampling to upscale the spatial resolution to 1/12°." (Page 5, Line 138-139)

[Comment 7] *Line 139: How were the Landsat NDVI samples aggregated to 1/12°? Please explain in detail.*

[Response 7] The reviewer referred to this sentence "The original spatial resolution of MCD12Q1 was 500 m, but it was spatially aggregated with Landsat NDVI samples to 1/12° to match the GIMMS NDVI3g product (Section 3.2)". Here we apologize for the confusion as two MODIS Land-Cover Type products (MCD12Q1 in 500 m and MCD12C1 in 0.05°), rather than only MCD12Q1, were employed in this study. The MCD12Q1 was employed to select sample locations in Landsat NDVI cross-calibration (Section 3.1.1), and the MCD12C1 was employed to establish biome-specific BPNN models (Section 3.2.2).

As for the issue raised by the reviewer, the Landsat NDVI samples were aggregated to 1/12° for training BPNN models with GIMMS NDVI3g (1/12°).

Specifically, 40,000 random sample locations (1/12°) were first generated across the globe. Then at a time step of half-month, we identified sample locations with high-quality GIMMS NDVI3g data (QC=0), searched all available Landsat data, and uniformly placed 9 matrices of 20 × 20 Landsat pixels (30 m resolution) within each sample location. We have added many details in Section 2.4 and Section 3.2.1 to clarify this.

The following changes are made in the revised manuscript:

(a) Introduction to MCD12Q1 and MCD12C1 in Section 2.4:

"The MODIS Land-Cover Type products provide global maps of land cover for each year between 2001–2019 (Friedl et al., 2002)."

"This study employed two MODIS Land-Cover Type products with different spatial resolutions, i.e., 500 m (MCD12Q1) and 0.05° (MCD12C1). The MCD12Q1 was used to select sample locations for Landsat NDVI cross-calibration (Section 3.1.1). The MCD12C1 was used to establish biome-specific BPNN models with GIMMS NDVI3g after being spatially aggregated to 1/12° using the nearest neighbor resampling method (Section 3.2.2)." (Page 5, Line 150-153)

(b) Landsat NDVI sample selection in Section 2.4:

"A total of 40,000 sample locations were randomly selected from the GIMMS NDVI3g product with a spatial resolution of 1/12°. Then at a time step of half-month, we identified sample locations with high-quality GIMMS NDVI3g data (QC=0) and uniformly placed 9 matrices of 20 × 20 Landsat pixels within each location (1/12°). Landsat pixel values were extracted from all available scenes." (Page 8, Line 211-214)

[Comment 8] *Line 144: "temporal" should be "temporally."*

[Response 8] Thanks for pointing out this grammar error. It has been fixed in the revised manuscript.

The following changes are made in the revised manuscript:

(a) First paragraph of the Methodology Section:

"1) Landsat sensor cross-calibration to create temporally consistent Landsat data as a benchmark;" (Page 6, Line 157-158)

[Comment 9] *Line 226: could the authors elaborate more on how they spliced the PKU GIMMS NDVI and MODIS NDVI?*

[Response 9] In the revised manuscript, more details have been provided on the consolidation of PKU GIMMS NDVI and MODIS NDVI, including how we used an 11 × 11 moving window to establish the pixel-wise Random Forests (RF) model during the overlapping periods of 2003–2015, how we tested the significance of the RF model, and how we applied the RF model to the non-overlapping period.

The following changes are made in the revised manuscript:

(a) Consolidation method in Section 3.3:

"In this study, we used a pixel-wise method inspired by Mao et al. (2012) to splice the PKU GIMMS NDVI product (1982–2015) and MODIS NDVI product (2003–2022). The pixel-wised method has been demonstrated more accurate than the global models (Yang et al., 2021). Specifically, the MODIS NDVI was first resampled to have the same spatial resolution (1/12°) and temporal resolution (half a month) as the PKU GIMMS NDVI (see Section 2.3). Then, during the overlapping periods (2003–2015), an 11 × 11 moving window (approximately 1° equivalent) was placed around each pixel. All the neighbors that had the same vegetation biome type with the center pixel were identified and their NDVI values were extracted from both products. This resulted in at most 1573 GIMMS-MODIS NDVI sample pairs (11 × 11 pixels per year in 13 years) for each pixel location. The sample pairs were further screened based on the data quality of PKU GIMMS NDVI (quality information adopted from GIMMS NDVI3g; see Section 2.2) and MODIS NDVI (see Section 2.3). Based on the sample pairs, the Random Forests (RF) regression model was constructed (Breiman, 2001), with explanatory variables of the PKU GIMMS NDVI and the longitude and latitude of samples and target variable of the MODIS NDVI. This study found that the significance of the RF model largely depended on the data quality of PKU GIMMS NDVI and MODIS NDVI. As such, we used 90% of the sample pairs for RF establishment and 10% for validation. R^2 was calculated. The pixel-wise RF model was applied to the non-overlapping period only when $R^2 > 0.2$ with $p < 0.001$; otherwise, the PKU

GIMMS NDVI was adjusted by aligning its mean value to that of the MODIS NDVI." (Page 9, Line 244-258)

[Comment 10] *Line 254: How can seasonal fluctuations in the time series of NDVI bias be removed via the multi-year averaging method? Please explain.*

[Response 10] The seasonal fluctuations were removed by subtracting the multi-year average at a particular time of the year from the original data, i.e.,

$$bias\_deseason_{i,j} = bias\_origin_{i,j} - mean(bias\_origin_j) \quad (1)$$

Where $bias\_origin_{i,j}$ is the original NDVI bias at the time j of the year i (e.g., the first half-month of January in 2005); $mean(bias_origin_j)$is the multi-year average at the time j (e.g., average on the first half-month of January for all years); and $bias\_deseason_{i,j}$ is the NDVI bias after removing the seasonal fluctuation. We have clarified this in the revised manuscript.

The following changes are made in the revised manuscript:

(a) PKU GIMMS NDVI evaluation in Section 3.3:

"Seasonal fluctuations in the time series of NDVI bias were first removed by subtracting the multi-year average at a particular time of the year, i.e.,

$$bias\_deseason_{y,t} = bias\_origin_{y,t} - mean(bias\_origin_t) \quad (5)$$

Where $bias\_origin_{y,t}$ is the original NDVI bias at the time t of the year y (e.g., the first half-month of January in 2005); $mean(bias\_origin_t)$ is the multi-year average at the time t (e.g., the first half-month of January for all years); and $bias\_deseason_{y,t}$ is the NDVI bias after removing the seasonal fluctuation." (Page 10, Line 279-284)

[Comment 11] *Line 257: Maybe it should be "... was evaluated at 1,000 random points ..."*

[Response 11] Thank you. The sentence has been re-written in the revised manuscript.

The following changes are made in the revised manuscript:

(a) PKU GIMMS NDVI evaluation in Section 3.3:

"The consolidation of PKU GIMMS NDVI with MODIS NDVI was evaluated at 1,000

random points for each vegetation biome type." (Page 11, Line 286-287)

[Comment 12] *Figure 2: maybe the authors could also show the regression line and equation in each panel?*

[Response 12] The Figure 2 has been updated to show the regression lines and equations for all panels.

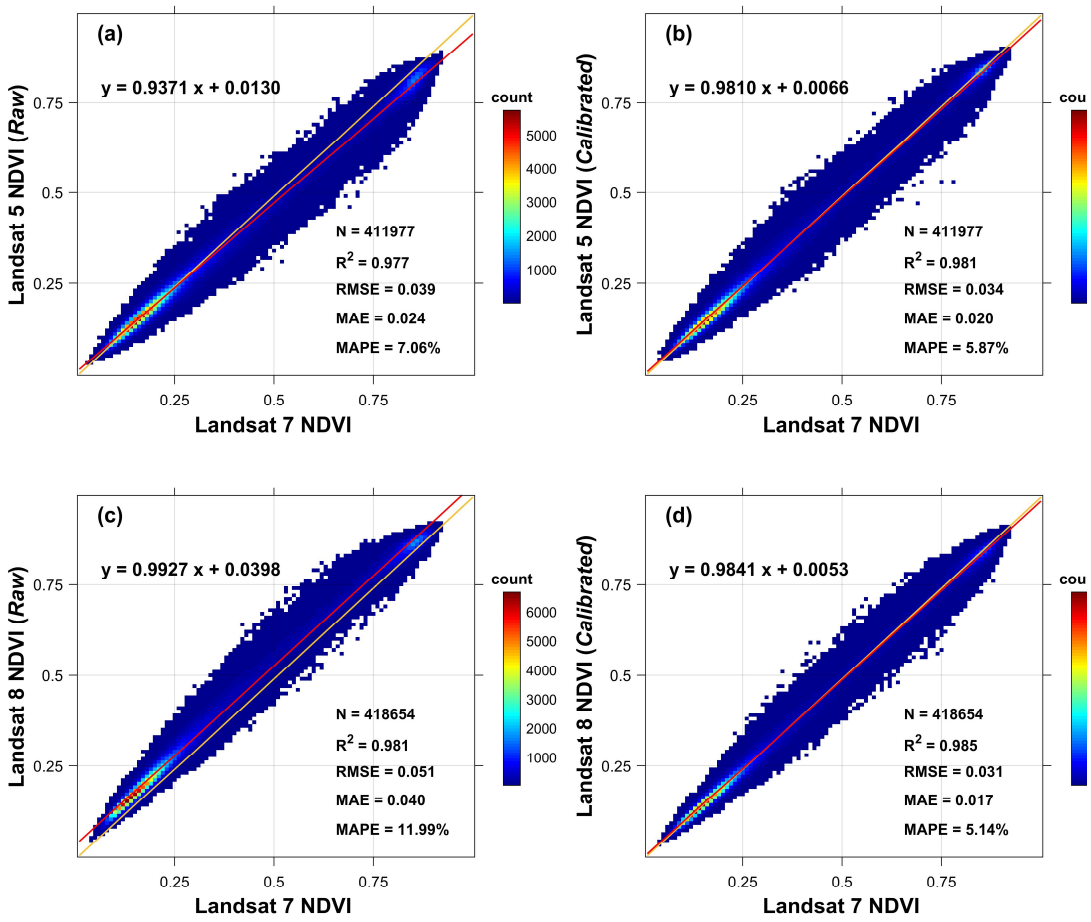

Figure 2. Efficiency of NDVI cross-calibration between Landsat sensors. (a) Landsat 7 NDVI vs. uncalibrated Landsat 5 NDVI. (b) Landsat 7 NDVI vs. calibrated Landsat 5 NDVI. (c) Landsat 7 NDVI vs. uncalibrated Landsat 8 NDVI. (d) Landsat 7 NDVI vs. calibrated Landsat 8 NDVI. Red line is the regression line and orange diagonal line represents a 1:1 relationship. The size of NDVI interval in the maps is 0.01. NDVI intervals with sample number < 10 were omitted.

[Comment 13] *Line 301: The section title could be "Validation of PKU GIMMS NDVI and GIMMS NDVI3g"*

[Response 13] Thanks for this comment. We used the term "Direct validation" instead of "Validation" for the section title because we would like to distinguish our validation analysis based on reference samples from those based on inter-products comparison analysis. After a serious consideration, we decide to keep the term "Direct validation". We sincerely appreciate the suggestion and understanding from the reviewer.

[Comment 14] *Figure 6: please explain in the figure caption how the R2 was computed in detail.*

[Response 14] The R^2 was calculated between Landsat NDVI samples and the GIMMS NDVI3g/PKU GIMMS NDVI products (Section 3.4). We recognized that the current figure caption could be misleading ("Comparison of R^2 between the GIMMS NDVI3g and PKU GIMMS NDVI products in pre-MODIS (1982–2000) and MODIS (2001–2015) period."). In the revised manuscript, the caption has been clarified and the calculation of R^2 has been explained.

The following changes are made in the revised manuscript:

(a) Figure 6 caption:

"Accuracies of the GIMMS NDVI3g and PKU GIMMS NDVI products measured by R^2 for pre-MODIS (1982–2000) and MODIS (2001–2015) period. The R^2 was calculated between the NDVI products and Landsat NDVI samples…"

[Comment 15] *Line 337: The section title could be "Comparison with MODIS NDVI"?*

[Response 15] Thanks for this suggestion. The title of Section 4.3.1 has been changed to "Comparison with MODIS NDVI" in the revised manuscript.

References:

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R., and Myneni, R. B.: Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011, Remote Sens, 5, 927-948, https://doi.org/10.3390/rs5020927, 2013.