To Referee #1

Dear reviewer,

We are very pleased to finish a revised version of the manuscript essd-2023-68 entitled "**Spatiotemporally consistent global dataset of the GIMMS Leaf Area Index (GIMMS LAI4g) from 1982 to 2020**". In preparing this revision we have considered all your comments and incorporated most of the suggestions. We greatly appreciate your time and effort spent reviewing this manuscript, which has improved the revised version of the manuscript.

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

(1) More details have been provided for the two datasets of PKU GIMMS NDVI and Landsat LAI samples. Specifically, statistics on their performance are now available.

(2) The GIMMS LAI4g is now validated by ground LAI measurements.

(3) The saturation issue in different vegetation biomes of GIMMS LAI4g has been discussed.

(4) The vegetation trends before and after 2000 have been analyzed for GIMMS LAI4g and other global long-term LAI products.

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

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] *Cao et al. generated a new global LAI dataset GIMMS LAI4g (hereafter, LAI4g), this dataset overcome the effect of satellite orbital drift and sensor degradation from NOAA series satellites. Its results may explain the LAI trend bias between pre-MODIS era (1982-1999) and MODIS era (2000-present) and it also explains whether the vegetation area is greening or browning over the world. This study fills the gap of how the AVHRR sensors affect the LAI trend (the significant improvement is showed at figure 6), which have not been solved in the previous studies.*

[Response 1] We thank the referee for highlighting the significance of this research. We are happy to provide a new version of the GIMMS LAI product (GIMMS LAI4g) that eliminates the effect of NOAA satellite orbital drift and AVHRR sensor degradation. On top of that, GIMMS LAI4g also demonstrates high absolute accuracies which meet the accuracy target proposed by the Global Climate Observation System. We hope this product could distinguish the true vegetation trend from uncertainties from NOAA/AVHRR and LAI references and help better understand the vegetation dynamic under the current global environmental changes.

[Comment 2] *However, three major points needed to be improved before publication.*

#1 The authors used two datasets, i.e. PKU GIMMS NDVI (Li et al. under review), Landsat LAI samples (Zha et al. in preparation) as the principal data source for generating LAI4g product. However, these two datasets have not finished their peer-review process, so these two datasets are not fully convincing. At least, the authors should show the statistics on how robust these two datasets are. I also suggested the authors adding the ground measured LAI (could be found in Xu et al. 2018 RSE and Ma et al. 2022 RSE) as validation to verify LAI4g product.

Xu, B., Li, J., Park, T., Liu, Q., Zeng, Y., Yin, G., ... & Myneni, R. B. (2018). An integrated method for validating long-term leaf area index products using global networks of site-based measurements. Remote Sensing of Environment, 209, 134-151.

Ma, H., & Liang, S. (2022). Development of the GLASS 250-m leaf area index product (version 6) from MODIS data using the bidirectional LSTM deep learning model.

Remote Sensing of Environment, 273, 112985.

[Response 2] We thank the reviewer for these suggestions. In the revised manuscript, we have provided more details on the generation process of two datasets (PKU GIMMS NDVI and Landsat LAI samples) and the statistics of their performance in absolute accuracy and vegetation trend. The annual variation of Landsat LAI sample size has been added to the updated Figure 2. We have also provided Figure S1 in the supplementary file that uses Evergreen Broadleaf Forests (EBF) as an example to demonstrate how the effect of NOAA satellite orbital drift and AVHRR sensor degradation was eliminated in the PKU GIMMS NDVI.

Also, a total of 113 field LAI measurements from BELMANIP 2.1, DIRECT 2.1, and ORNL were employed to evaluate the GIMMS LAI4g product as well as other global long-term LAI products. The results showed that GIMMS LAI4g (R^2 = 0.69, RMSE = 0.86 m^2/m^2, MAE = 0.61 m^2/m^2, MAPE = 33%) had comparable accuracies with GIMMS LAI3g (R^2 = 0.71, RMSE = 0.78 m^2/m^2, MAE = 0.55 m^2/m^2, MAPE = 30%) and GLASS LAI (R^2 = 0.68, RMSE = 0.82 m^2/m^2, MAE = 0.60 m^2/m^2, MAPE = 33%), and it had the lowest underestimation in terms of the fitting line with a slope of 0.87 and an intercept of 0.05 (Figure 4).

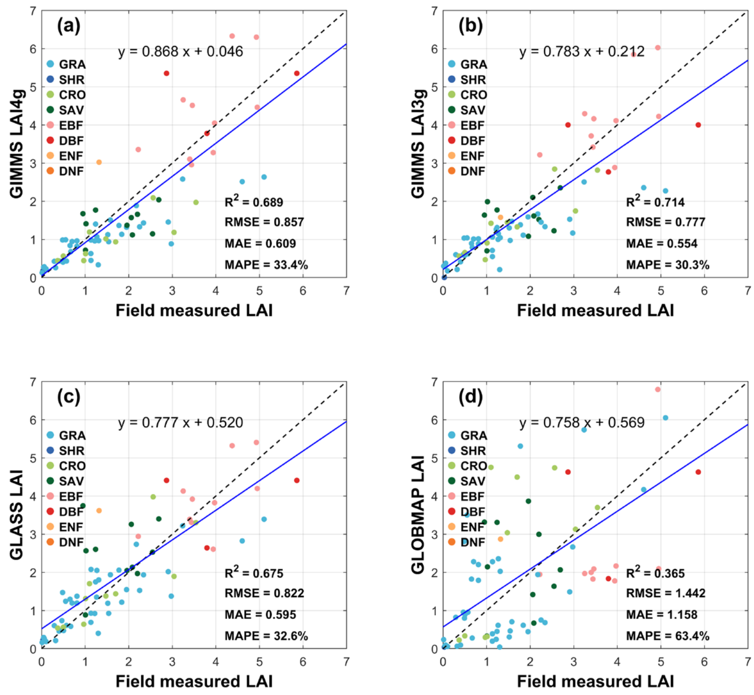

Figure 4. Validation of the (a) GIMMS LAI4g, (b) GIMMS LAI3g, (c) GLASS LAI,

and (d) GLOBMAP LAI products using 113 field LAI measurements from 49 sites in the projects of BELMANIP 2.1, DIRECT 2.1, and ORNL. Sites of different vegetation biome types are marked by colors. The error metrics are R^2, RMSE, MAE, and MAPE. The blue fitting lines and dashed 1:1 lines are drawn.

The following changes are made in the revised manuscript:

(a) Introduction to PKU GIMMS NDVI in Section 2.1:

"In the generation of PKU GIMMS NDVI, Landsat NDVI from Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and Operational Land Imager (OLI) were first cross-calibrated. Then, massive high-quality Landsat NDVI samples were extracted by considering a series of factors including clouds, cloud shadows, water, snow, aerosol, and radiation performance. The Landsat NDVI samples were employed to calibrate the GIMMS NDVI3g product with other explanatory variables using biome-specific machine learning models. The calibrated NDVI product was finally consolidated with the MODIS NDVI product to extend the temporal coverage to the year 2020. The major improvement of PKU GIMMS NDVI over its counterparts is that it well removed the NOAA orbital drift and AVHRR sensor degradation effects, especially in tropical regions (Figure S1). Its overall R^2, Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) are 0.975, 0.033, and 9%, respectively. It is highly consistent with MODIS NDVI in terms of pixel value (R^2 = 0.962, MAE = 0.032, and MAPE = 6.5%) and global vegetation trend. PKU GIMMS NDVI inherited the quality control (QC) information from the GIMMS NDVI3g. A QC value of 0, 1, and 2 indicates NDVI of good quality, NDVI retrieved from spline interpolation, and NDVI retrieved from average seasonal profile, respectively. The PKU GIMMS NDVI record during AVHRR missions from 1982 to 2015 (before consolidation with MODIS NDVI) was used in this study. It is available at https://zenodo.org/record/7441559#.Y7J7y3ZByCo." (Page 4, Line 97-110)

(b) Introduction to Landsat LAI samples in Section 2.2:

"The Landsat LAI sample dataset provides approximately 4.9 million high-quality samples with a spatial resolution of 1/12 ° and a temporal resolution of half a month (Zha, in preparation). It covers the global vegetated area with all vegetation biome types defined in the MODIS land cover product (the third classification scheme; see section 2.4) and a long-time span from 1984 to 2020. In the generation of Landsat LAI samples, 70,000 sample locations for Deciduous Needleleaf Forests [DNF] and 100,000 sample locations for each of the other vegetation biome types were randomly selected based on the MODIS land cover product. At the sample locations, Reprocessed MODIS LAI (in 500 m resolution; see Section 2.3) and Landsat surface reflectance from TM, ETM+, and OLI scenes (20 × 20 pixels in 30 m resolution) were extracted, creating massive sample pairs. The sample pairs were then rigorously screened by criteria that were not limited to those mentioned in Section 2.1 (i.e., clouds, cloud shadows, etc.) but also included Landsat sample purity, NDVI-LAI relationship, and the saturation state of the MODIS LAI. Biome- and Landsat sensor-specific Random Forest models with other explanatory variables were built based on the sample pairs. The models were applied to historical Landsat data at 40,000 random sample locations (1/12°) to create the final Landsat LAI sample dataset. Validation of the dataset through observations from the BEnchmark Land Multisite ANalysis and Intercomparison of Products (BELMANIP) network (Baret et al., 2006) and the Oak Ridge National Laboratory (ORNL) (Scurlock et al., 2001) showed high absolute accuracies (R^2 = 0.76, MAE = 0.45 m^2/m^2, Root Mean Square Error (RMSE) = 0.66 m^2/m^2). The inter-comparison with the Reprocessed MODIS LAI shows a high temporal consistency. This study selected 3.6 million Landsat LAI samples between 1984 and 2015." (Page 4-5, Line 112-127)

(c) Introduction to field LAI measurements in Section 2.8:

"The field LAI measurements were from three projects namely, BELMANIP 2.1 (available at https://calvalportal.ceos.org/web/olive/site-description) (Baret et al., 2006), DIRECT 2.1 (available at https://calvalportal.ceos.org/lpv-direct-v2.1) (Garrigues et al., 2008), and ORNL (available at https://daac.ornl.gov/VEGETATION/guides/LAI_guide.html) (Scurlock et al., 2001). The BELMANIP 2.1 and DIRECT 2.1 provide 3 km × 3 km averaged LAI values derived from sites in networks of FLUXNET, AERONET, VALERI, BigFoot, etc. The upscaling from site-based LAI to 3 km × 3 km LAI used high spatial resolution imageries such as Landsat and SPOT. Most global long-term LAI products have utilized the BELMANIP and DIRECT LAI as ground truth for product evaluation (Myneni et al, 2002; Liu et al., 2012; Xiao et al., 2016; Zhu et al., 2013), yet the LAI measurements in both projects were available only after the late 1990s. Note that

GLASS LAI (version 4) also employed BELMANIP sites for LAI model training (Xiao et al., 2016). This study further incorporated ORNL sites which provide field LAI measurements during 1932−2020 despite possible scaling effects due to spatial heterogeneity. We prudently examined all the sites and measurements in BELMANIP 2.1, DIRECT 2.1, and ORNL, and removed measurements that were acquired from heterogeneous sites or coincident among the three projects. Eventually, 113 field LAI measurements from 49 sites were obtained. Information on selected field LAI measurements can be found in Table S1." (Page 6-7, Line 173-187)

(d) Results of validation using field LAI measurements in Section 4.3:
"Based on field LAI measurements, GIMMS LAI4g generated from the BPNN models presented comparable accuracies (R^2 = 0.69, RMSE = 0.86 m^2/m^2, MAE = 0.61 m^2/m^2, MAPE = 33%) with GIMMS LAI3g (R^2 = 0.71, RMSE = 0.78 m^2/m^2, MAE = 0.55 m^2/m^2, MAPE = 30%), and GLASS LAI (R^2 = 0.68, RMSE = 0.82 m^2/m^2, MAE = 0.60 m^2/m^2, MAPE = 33%) (Figure 4). GIMMS LAI3g had the best performance in error measures (i.e., R^2, RMSE, MAE, and MAPE), but GIMMS LAI4g had the lowest underestimation for the fitting line with a slope of 0.87 and an intercept of 0.05 (Figure 4). GLOBMAP LAI presented the largest discrepancies from the LAI measurements." (Page 14, Line 335-340)

[Comment 3] *#2 The saturation of LAI is common in LAI4g over biomes. For example, figure 3 a1 showed that the Landsat LAI ranged from 0 to 4 m2 m-2, however, the LAI4g is saturated at 3 m2 m-2. Similar conditions can be found at CRO, SHR, SVA, ENF, DNF. If there is a 1:1 line in each subplot, the underestimation of LAI4g in high range of Landsat LAI would be significant. I guess the reason the train data for LAI4g is NDVI, the NDVI may saturated but LAI not, so the model setting may lead to the LAI4g saturation.*

[Response 3] We agree that the saturation of NDVI at high values (dense canopy) could be propagated to downstream products. To minimize the saturation effect in the GIMMS LAI4g, we established biome-specific models that incorporate multiple explanatory variables (longitude, latitude, month, and satellite number and years since launch) besides the PKU GIMMS NDVI, to explain the LAI variations from space, time, biome, and satellite. These explanatory variables were well contributed to the accuracies of LAI models (Figure 3). As a result, the saturation effect for a particular biome was subtle at a majority of sample locations (red dots in Figure 5) and was observable for samples whose LAI values deviated from the average (yellow and blue dots in Figure 5). For sparse vegetation types including GRA, SHR, CRO, and SAV, the GIMMS LAI4 would be underestimated at high values; and for dense vegetation types including EBF, DBF, ENF, and DNF, the GIMMS LAI4 would be underestimated at low and medium values. The 1:1 lines and fitting lines were added in the updated Figure 5. For the global vegetation biome, the LAI fitting line is close to the 1:1 line with a slope of 0.95 and an intercept of 0.07 (Figure 5a-9). It should be noted that the saturation effect can also be observed in EBF, DBF, ENF, and DNF of GLASS LAI, which was not derived from NDVI data but from surface reflectance.

Despite the saturation effect in GIMMS LAI4g and GIMMS LAI3g (based on GIMMS NDVI3g), we believed that the use of GIMMS NDVI products is beneficial as they have properly dealt with the limitations of the AVHRR instrument or operation, e.g., vicarious calibration, atmospheric and cloud correction, bias correction, satellite orbital drift, and sensor degradation (Pinzon and Tucker, 2014; Li et al., in review). The GIMMS LAI3g is one of the core data references in the IPCC Sixth Assessment Report for the assessment of global vegetation changes (Eyring et al., 2021), and has been cited for more than 600 times (from *Web of Science*).

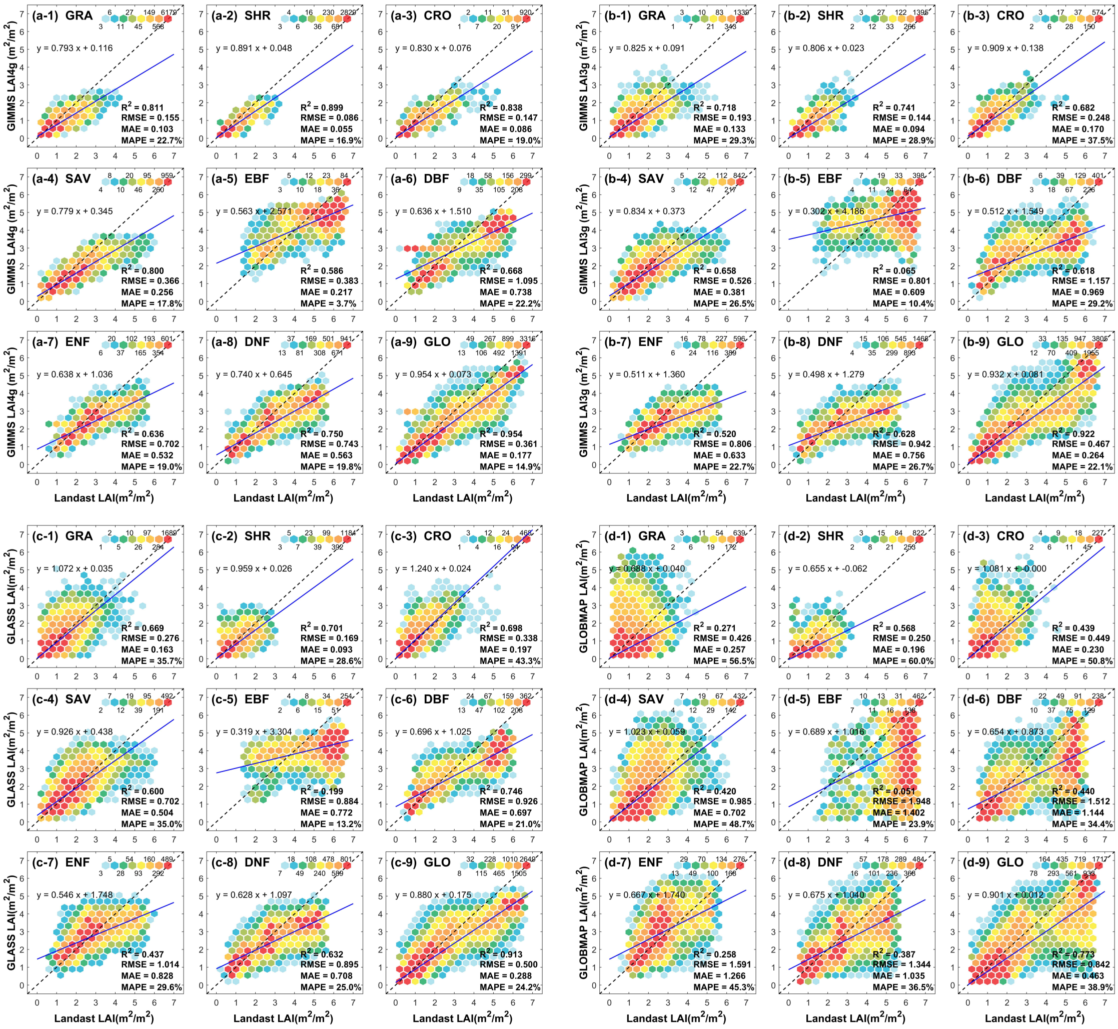

Figure 5. Validation of the (a) GIMMS LAI4g, (b) GIMMS LAI3g, (c) GLASS LAI, and (d) GLOBMAP LAI products in different vegetation biomes using Landsat LAI samples from 1984 to 2015. The error metrics are R^2, RMSE, MAE, and MAPE. GLO represents the global vegetation biome. The color of the dots represents LAI value frequencies in a 0.5 (m^2/m^2) interval.

The following changes are made in the revised manuscript:

(a) Results on the saturation effect in Section 4.3:

"Compared to Landsat LAI samples, the four global LAI products were underestimated in almost all the vegetation types except for CRO of GLASS LAI (Figure 5). We found certain levels of saturation in GIMMS LAI4g and GIMMS LAI3g, such as for high values of GRA, SHR, CRO, and SAV and medium values of EBF, DBF, ENF, and DNF (Figure 5). This could be attributed to the use of NDVI data in LAI models. However, we also observed the saturation effect in EBF, DBF, ENF, and DNF of GLASS LAI, which was not derived from NDVI data. For GIMMS

LAI4g, the saturation was relatively subtle at a majority of sample locations (red dots) and was obvious for locations whose LAI values deviated from the average (yellow and blue dots). The LAI fitting line of the global vegetation biome in GIMMS LAI4g (Figure 5a-9) was close to the 1:1 line." (Page 17, Line 370-376)

(b) Discussions on the saturation effect in Section 5.3:

"The use of PKU GIMMS NDVI could also result in the saturation effect in GIMMS LAI4g as NDVI data tend to saturate at high values (Figure 5). This study established biome-specific models that incorporate multiple explanatory variables besides PKU GIMMS NDVI to account for the LAI variations in space, time, biome, and satellite. This effort could help alleviate the saturation effect though the effect still exists." (Page 27, Line 555-558)

[Comment 4] *The analysis of LAI trend can be expanded. Zhu et al. 2021 showed that the CO2 fertilization effect (CFE) is diversity from satellite LAI products, and they showed different patterns before and after 2000. Since the LAI4g overcome the effect of satellite orbital drift and sensor degradation from NOAA series satellites, I think it could solve the problem raised in Zhu et al. 2021. So, the authors can add the analysis on how the LAI trend change before and after 2000 for LAI4g, which benefits the ecology research community on how the vegetation area is greening or browning over the world.*

Zhu, Z., Zeng, H., Myneni, R. B., Chen, C., Zhao, Q., Zha, J., ... & MacLachlan, I. (2021). Comment on "Recent global decline of CO2 fertilization effects on vegetation photosynthesis". Science, 373(6562), eabg5673.

[Response 4] As suggested by the reviewer, Figure 13 has been augmented in the revised manuscript to illustrate the vegetation trends before and after 2000. Before 2000, GIMMS LAI4g demonstrated generally greening trends except for EBF (-1.65×10^{-3} $m^2m^{-2}yr^{-1}$), although EBF sustained greening since 1991. After 2000, GIMMS LAI4g exhibited continuous greening for all vegetation biome types.

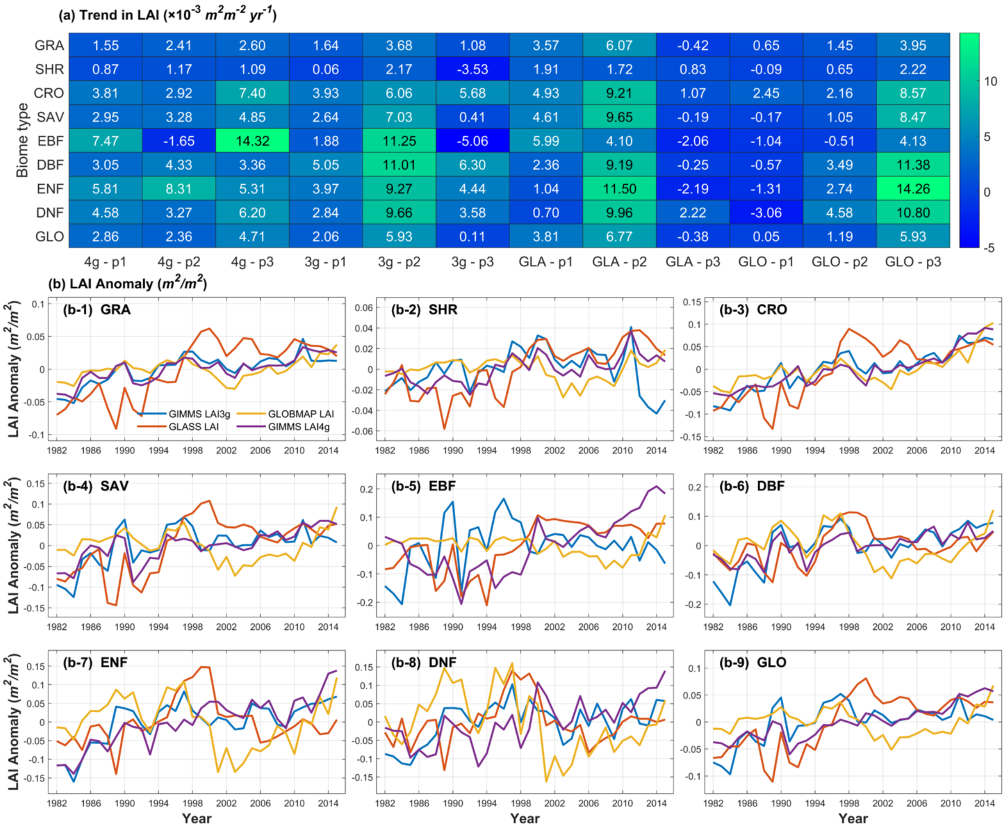

Figure 13. Variations of annual LAI anomaly of different vegetation biomes in the global LAI products during 1982−2015. The LAI products include GIMMS LAI4g, GIMMS LAI3g, GLASS LAI, and GLOBMAP LAI. (a) shows the slope values of the annual LAI during 1982−2020 (p1), 1982−2000 (p2), and 2001−2020 (p3). In the x-axis, 4g, 3g, GLA, and GLO stands for GIMMS LAI4g, GIMMS LAI3g, GLASS LAI, and GLOBMAP LAI, respectively. (b) shows the annual LAI time series.

The following changes are made in the revised manuscript:

(a) Results on vegetation trends before and after 2000 in Section 4:

"Figure 13a shows the annual average LAI trends during 1982−2020 (p1), 1982−2000 (p2), and 2001−2020 (p3) for different vegetation biomes of the four LAI products. For the whole period domain (1982−2020), the GIMMS LAI3g, GIMMSLAI4g, and GLASS LAI products presented a similar greening trend for the global vegetation biome, with a slope value of 2.06×10^{-3} $m^2m^{-2}yr^{-1}$, 2.86×10^{-3} $m^2m^{-2}yr^{-1}$, and 3.81×10^{-3} $m^2m^{-2}yr^{-1}$, respectively." (Page 23-24, Line 474-477)

"Before 2000, the four LAI products generally demonstrated greening trends except for EBF in GIMMS LAI4g (-1.65×10^{-3} $m^2m^{-2}yr^{-1}$) and GLOBMAP LAI (-0.51×10^{-3} $m^2m^{-2}yr^{-1}$). The LAI products showed few agreements in vegetation trends after 2000.

GIMMS LAI4g and GLOBMAP LAI exhibited continuous greening in all biomes, GIMMS LAI3g exhibited browning in SHR and EBF, and GLASS LAI was dominated by a browning trend in GRA, SAV, EBF, DBF, and ENF." (Page 24, Line 480-484)

"For EBF in GIMMS LAI4g, although its trend during 1982−2000 was overall negative (browning) (Figure 13a), it has sustained a greening trend since 1991 (Figure 13b-5)." (Page 24, Line 488-489)

[Comment 5] *Since the line number is not covered every line, so I just give a range of line number here.*

[Response 5] We feel sorry about this inconvenience. However, we used the template provided by ESSD so there is not too much we can do right now.

[Comment 6] *~L45 'However, the accuracies of the current LAI products have been limited by uncertainties primarily in the remote sensing data and the LAI reference data (Fang et al., 2019).' Jiang et al. 2017 GCB is also an important reference here.*

[Response 6] The work from Jiang et al. (2017) has been cited in the revised manuscript. The revised sentence is now "However, the accuracies of the current LAI products have been limited by uncertainties primarily in the remote sensing data and the LAI reference data (Fang et al., 2019; Jiang et al., 2017)." (Page 2, Line 44-45)

[Comment 7] *~ L70 to 75 It is not common to cite two papers in under review process in the introduction section. So, the authors may consider citing other published papers.*

[Response 7] We totally understand that the datasets in review (PKU GIMMS NDVI and Landsat LAI samples) would raise questions about how the datasets were generated and their accuracies. However, no other papers on the two datasets are currently published. We would like to claim that it is our duty to immediately cite the final version of the datasets once they finish the peer review. As compensation and as the reviewers suggested, we have added more details in the revised manuscript to better describe PKU GIMMS NDVI and Landsat LAI samples (Section 2.1 and Section 2.2). We would also like to let the reviewer know that, the manuscript related to PKU GIMMS NDVI is currently in revision as well. All the revisions have been incorporated into the updated version of GIMMS LAI4g, despite the revisions being subtle and no major conclusions altered.

[Comment 8] *Section 2.1 and 2.2, these two datasets are not finished peer reviewed especially the Landsat LAI sample dataset. To ensure the GIMMS LAI4g is reliable, the authors should give enough evidence to show these two data sources are robust.*

[Response 8] Thanks for this comment. More details on the two datasets (PKU GIMMMS NDVI and Landsat LAI samples), including their performance in absolute accuracy and vegetation trend have been provided in Sections 2.1 and 2.2 of the revised manuscript. Please refer to **Response 2**.

[Comment 9] *Section 2.3 I found the authors have not used the ground measurement LAI as directly validate for LAI4g product. I recommended that the authors may use the ground measured LAI dataset in Xu et al. 2018 RSE and Ma et al. 2021 RSE to validate the LAI4g product.*

[Response 9] This comment is also related to **Comment 2**. We have used field LAI measurements from BELMANIP 2.1, DIRECT 2.1, and ORNL to evaluate the GIMMS LAI4g product as well as other global long-term LAI products.

[Comment 10] *Section 2.5 There are two version of GLASS-LAI v50, one is totally based on AVHRR data (http://www.glass.umd.edu/LAI/AVHRR/), this dataset didn't use the MODIS data as data source. The other one is fully based on MODIS with 500m spatial resolution. From my experience, there are no such dataset used both AVHRR and MODIS data as input to generate GLASS LAI product, so the authors should ensure this point.*

[Response 10] After a careful examination, we need to clarify that the GLASS LAI product used in this study belongs to version 4 (Xiao et al., 2016) rather than version

5. We apologize for this mistake, which has been corrected in the revised manuscript. At the time that our project started, GLASS LAI version 5 or newer was not available (Liang et a., 2021; Ma and Liang, 2022). The availability information of the latest GLASS LAI product has also been added in the revised manuscript. Please understand that with new versions of the global long-term LAI products being constantly available, it would be difficult to always use the latest versions in the study. More importantly, we believe that the update of other LAI products would not lower the significance of our GIMMS LAI4g.

The following changes are made in the revised manuscript:

(a) Introduction to GLASS LAI in Section 2.5:

"The GLASS LAI (version 4) with a temporal resolution of …" (Page 6, Line 151)
"The GLASS LAI (V4) product was acquired from ftp://ftp.glcf.umd.edu/. It should be noted that version 5 and version 6 of the GLASS LAI product have been available when our study has been prepared (Liang et a., 2021; Ma and Liang, 2022)." (Page 6, Line 156-158)

(b) Clarification in Section 3.1:

"For example, a typical ANN, …GLASS LAI (version 4) (Xiao et al., 2014), and GIMMS LAI3g (Zhu et al., 2013), respectively." (Page 8, Line 201-202)

[Comment 11] *Section 2.6 There are two version of GLOBMAP product. The previous version showed that the LAI is decreasing after 2000 (see Jiang et al. 2017 GCB figure1). The current version is GLOBMAP V3, which showed an increasing trend after 2000. So which one is the reliable one, it would be an open question. The authors may consider not using this dataset to do the inter-comparison.*

[Response 11] Thanks for this comment. The GLOBMAP LAI product used MODIS surface reflectance as input from 2001−now. For GLOBMAP LAI version 2, the MODIS product version was C5; and for GLOBMAP LAI version 3, the MODIS product version was C6 (10.1029/2012JG002084). The update from MODIS C5 to C6 addressed a series of known issues including the sensor degradation, which was believed to correct the analytical vegetation trend from negative to positive (Lyapustin et al., 2014). With this being clarified, however, we agree with the reviewer that current global long-term LAI products of different versions need comprehensive evaluation. For our study, with the ultimate goal to signify the improvements of GIMMS LAI4g, we may still choose to use the latest versions of popular LAI products including GLOBMAP LAI. The latest versions were determined at the time our study has been prepared.

[Comment 12] *Section 3. A flow chart of how to generate LAI4g is useful.*

[Response 12] A workflow has been added accordingly in the revised manuscript, as below.

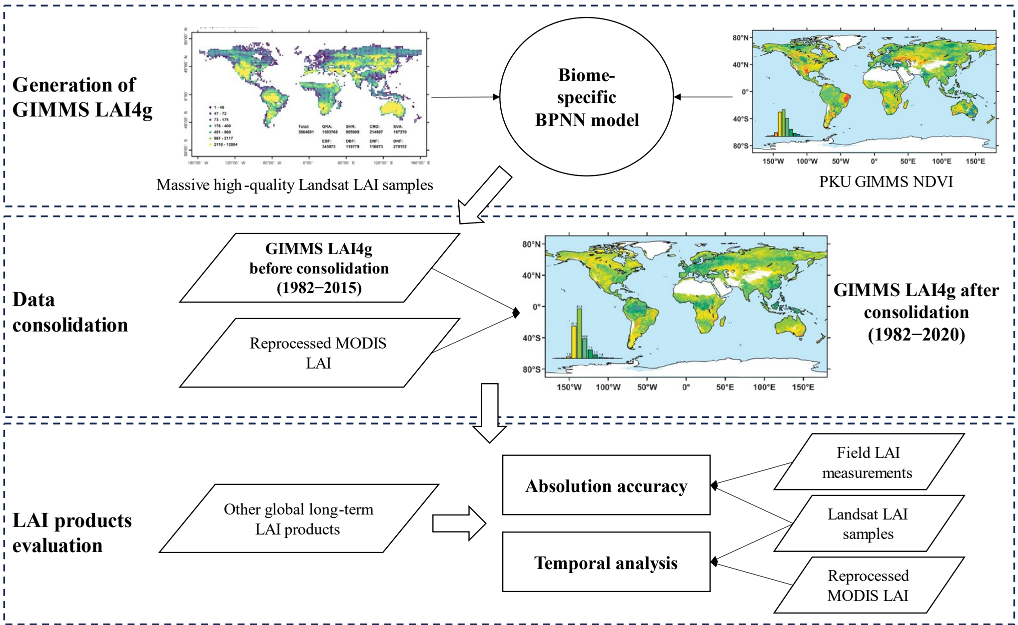

Figure 1. Schematic diagram of the generation and evaluation of the GIMMS LAI4g product.

[Comment 13] *Figure 1a I guess the SVA should be SAVanna? It would be more common to use the abbreviation SAV to represent savanna.*

[Response 13] Thanks for pointing this out. We have replaced all SVA with SAV through all texts and figures.

[Comment 14] *Figure 1c will the LAI distribution of different data source affect LAI4g? The Landsat LAI showed a peak around 0.5 m2 m-2, but other product didn't* *show a similar pattern.*

[Response 14] We used Figure 1 (updated Figure 2; Figure 2 will be used hereafter) to demonstrate that the Landsat LAI samples have a good distribution in space (Figure 2a), time (Figure 2b), and value (Figure 2c). Specifically, Figure 2c shows that the value distribution of the Landsat LAI samples could overall match the global long-term LAI products. Regarding the concern from the reviewer, if the value distribution is significantly mismatched between LAI samples and LAI products, the samples would not be representative and suitable for training the LAI models for GIMMS LAI4g. In this sense, different LAI distributions will affect LAI4g. We recognize that Figure 2c was not properly drawn as the frequency curves overlapped each other. A new figure has been made in the revised manuscript. The new figure shows that both GIMMS LAI3g and Landsat LAI samples peak at 0.5 m^2/m^2 and the peaking values for GLASS LAI and GLOBMAP LAI were smaller.

To further demonstrate the temporal representation of Landsat LAI samples, we have also updated Figure 2b to show the inter-annual variation of the Landsat LAI sample size. During 1984−2015, the Landsat LAI sample size per year increased from 28,323 in 1984, peaked at 200,315 in 2001 when both Landsat 5 and Landsat 7 were available, and then leveled off until 2012 (sample size: 22,106). As Landsat 5's decommissioning, very few images were acquired from November 2011 to May 2012 (https://www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-archives-landsat-4-5-thematic-mapper-tm-level-1-data). Since the launch of Landsat 8 in 2013, the Landsat LAI sample became steadily available again.

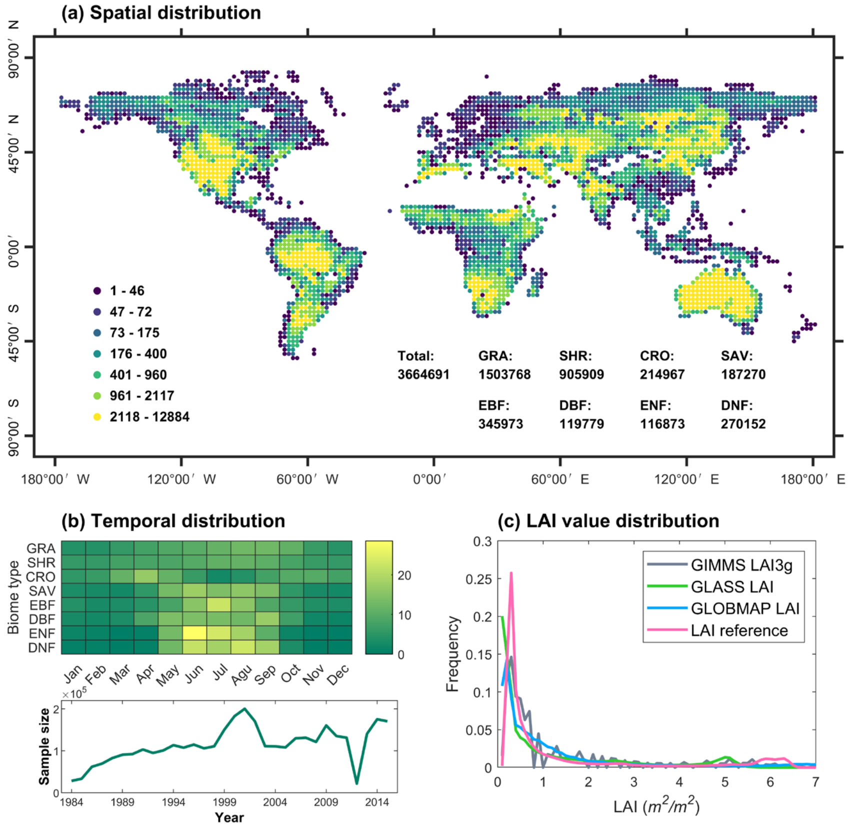

Figure 2. Spatial and temporal distribution of the LAI reference data. (a) The global distribution of LAI samples in 2° grids. The LAI sample size for each vegetation is listed. (b) The temporal distribution of LAI samples for the eight vegetation biome types and the annual variation of LAI sample size. (c) The distribution of LAI values in percentage (bin width = 0.1) for Landsat LAI samples, GIMMS LAI3g, GLASS LAI, and GLOBMAP LAI. It should be noted that 40,000 Reprocessed MODIS LAI samples were introduced at locations and months when Landsat LAI samples were scarce.

The following changes are made in the revised manuscript:

(a) Inter-annual variation of the Landsat LAI sample size in Section 4.1:

“During 1984−2015, the Landsat LAI sample size per year increased from 28,323 in 1984, peaked at 200,315 in 2001 when both Landsat 5 and Landsat 7 were available, and then leveled off until 2012 (sample size: 22,106) (Figure 2b). From November 2011 to May 2012, very few images were acquired with Landsat 5’s decommissioning, (https://www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-archives-landsat-4-5-thematic-mapper-tm-level-1-data). Since the launch of Landsat 8 in 2013, the Landsat LAI sample became steadily available again.” (Page 11, Line 292-297)

[Comment 15] *Figure 3 the saturation effect to LAI4g is significant (see my comment at major point #2).*

[Response 15] A detailed reply to this comment is referred to **Response 3**. The saturation did exist, and it was more significant for samples whose LAI values deviated from the average (yellow and blue dots in Figure 5). For a majority of samples (red dots in Figure 5), the saturation effect was subtle. The fitting line for GLO in updated Figure 5 is close to the 1:1 line.

[Comment 16] *Figure 4 There is data missing (in white color) at southwest China, Euro and southern part of USA.*

[Response 16] This study used 2° × 2° grids to evaluate the relative accuracies among GIMMS LAI4g, GIMMS LAI3g, and GLASS LAI. We assumed the evaluation to be reliable only if the LAI sample count within each grid was larger than 100. We have clarified this in the text and Figure 4 caption of the revised manuscript.

The following changes are made in the revised manuscript:

(a) Results interpretation in Section 4.3:

"Grids with a small Landsat LAI sample size (< 100) were excluded as they may not provide a reliable evaluation." (Page 17, Line 378-379)

(b) Figure 6 caption:

"**Figure 6.** Dominance map of the GIMMS LAI3g, GIMMS LAI4g, and GLASS LAI based on their MAE. The map was drawn in 2° × 2° grids whose colors were composed of reciprocal averages of MAE from the GIMMS LAI4g (green), GIMMS LAI3g (red), and GLASS LAI (blue). Non-vegetated grids and grids with small Landsat LAI sample size (< 100) were filled white. A greener grid, for example, indicates that the GIMMS LAI4g has a lower MAE (or a higher absolute LAI accuracy). " (Page 17-18, Line 387-390)

[Comment 17] *Figure 6 this is a significant improvement for LAI product based on AVHRR data! Very nice result.*

[Response 17] We appreciate it. The removal of effects of NOAA orbital drift and AVHRR sensor degradation is one of the highlights of GIMMS LAI4g.

[Comment 18] *Figure 9 authors should add the statistics for each histogram.*

[Response 18] Thanks for the suggestion. In the revised manuscript, we have labeled all value frequencies at each LAI interval in the histograms.

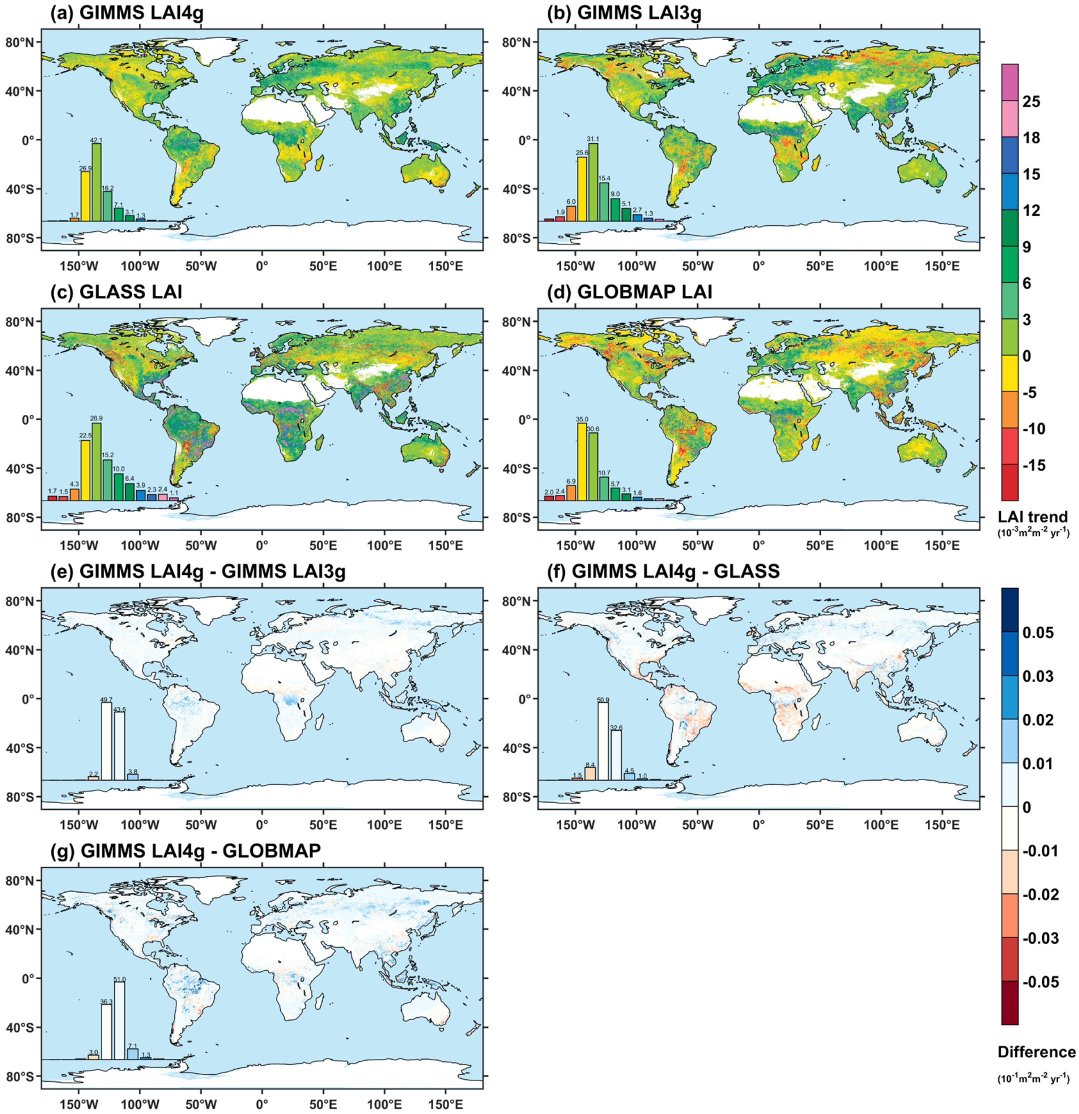

Figure 12. Global maps of LAI trends and their differences between the global LAI products during 1982–2015. The LAI products include GIMMS LAI4g after consolidation (a), GIMMS LAI3g (b), GLASS LAI (c), and GLOBMAP LAI (d). The trend was calculated as the slope of a linearly fitted LAI time series. (e) to (g) show the slope differences between the GIMMS LAI4g and the other three LAI products.

[Comment 19] *Figure 10 authors may add the statistics for the global LAI trend before (pre-MODIS era) and after 2000 (MODIS era). This can provide the insight of whether the vegetated area is greening or browning.*

[Response 19] This comment is related to **Comment 4**. Figure 10a (updated Figure 13a) has been augmented to illustrate the vegetation trends before and after 2000.

[Comment 20] *L 435 'Two other LAI products, namely the GLASS LAI (1982−2018) and GLOBMAP LAI (1982−2020), also incorporated MODIS data (reflectance). ' This is not totally right, the GLASS-LAI product (1982−2018) is only based on AVHRR data (Xiao et al. 2017).*

Xiao, Z., Liang, S., & Jiang, B. (2017). Evaluation of four long time-series global leaf area index products. Agricultural and Forest Meteorology, 246, 218-230.

[Response 20] We apologize again for this mistake. We used version 4 of the GLASS LAI product. In this version, GLASS LAI was generated from AVHRR reflectance during 1981−1999 and from MODIS surface reflectance after 2000 (Xiao et al., 2016). Corresponding corrections have been made, as indicated in a previous reply.

References:

Eyring, V., Gillett, N. P., Achutarao, K., Barimalala, R., Barreiro Parrillo, M., Bellouin, N., Cassou, C., Durack, P., Kosaka, Y., McGregor, S., Min, S.-K., Morgenstern, O., and Sun, Y.: Human Influence on the Climate System, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J.B.R., Maycock, T.K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 423−552, https://doi.org/10.1017/9781009157896.005, 2021.

Ma, H. and Liang, S.: Development of the GLASS 250-m leaf area index product (version 6) from MODIS data using the bidirectional LSTM deep learning model, Remote Sens. Environ., 273, 112985, https://doi.org/10.1016/j.rse.2022.112985, 2022.

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., Zhao, X., and Zhou, J.: The Global Land Surface Satellite (GLASS) Product Suite, Bull. Am. Meteorol., 102, E323–E337. https://doi.org/10.1175/BAMS-D-18-0341.1, 2021.

Li, M., Cao, S., and Zhu, Z.: Spatiotemporally consistent global dataset of the GIMMS Normalized Difference Vegetation Index (PKU GIMMS NDVI) from 1982 to 2020, Earth Syst. Sci. Data, in review.

Lyapustin, A., Wang, Y., Xiong, X., Meister, G., Platnick, S., Levy, R., Franz, B., Korkin, S., Hilker, T., Tucker, J., Hall, F., Sellers, P., Wu, A., and Angal, A.: Scientific impact of MODIS C5 calibration degradation and C6+ improvements, Atmos. Meas. Tech., 7, 4353–4365, https://doi.org/10.5194/amt-7-4353-2014, 2014.

Pinzon, J. E. and Tucker, C. J.: A non-stationary 1981-2012 AVHRR NDVI3g yime series, Remote Sens., 6, 6929-6960, https://doi.org/10.3390/rs6086929, 2014.

Xiao, Z. Q., Liang, S. L., Wang, J. D., Xiang, Y., Zhao, X., and Song, J. L.: Long-time-series global land surface satellite leaf area index product derived from MODIS and AVHRR surface reflectance, IEEE Trans. Geosci. Remote Sens., 54, 5301-5318, https://doi.org/10.1109/tgrs.2016.2560522, 2016.

To Referee #2

Dear reviewer,

We are very pleased to finish a revised version of the manuscript essd-2023-68 entitled "**Spatiotemporally consistent global dataset of the GIMMS Leaf Area Index (GIMMS LAI4g) from 1982 to 2020**". In preparing this revision we have considered all your comments and incorporated most of the suggestions. We greatly appreciate your time and effort spent reviewing this manuscript, which has improved the revised version of the manuscript.

Substantial improvements have been made based on your comments. Significant ones include:

(1) More details have been provided for the two datasets of PKU GIMMS NDVI and Landsat LAI samples. The removal of the NOAA orbital drift and AVHRR sensor degradation effects in PKU GIMMS NDVI is illustrated. The generation and accuracies of the Landsat LAI samples are elaborated.

(2) The saturation issue in GIMMS LAI4g has been discussed in terms of NDVI and BPNN models.

(3) Ground LAI measurements have been used as an independent source to validate the GIMMS LAI4g product.

(4) Global distribution maps of GIMMS LAI4g in January and July have been provided.

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

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] *The authors have introduced an updated version of the GIMMS LAI4g datasets, covering the period between 1982 and 2020. BPNN model was trained using PKU GIMMS NDVI, Landsat LAI samples, and three other explanatory variables to predict the LAI values. Also, a pixel-wise linear fusion based method was further used to consolidate the AVHRR-based LAI with the MODIS LAI product. Despite the dataset's design to effectively mitigate the effects of satellite orbital drift and sensor degradation, I believe there is room for further improvement through addressing or discussing the following questions:*

[Response 1] We appreciate the reviewer for all the comments. The GIMMS LAI4g product has been generated to mitigate the effect of NOAA satellite orbital drift and AVHRR sensor degradation, as well as to further improve the spatiotemporal consistency and accuracy. We hope our following responses could well address your concerns and the product of GIMMS LAI4g has been further improved.

[Comment 2] *Major:*

Reliability of the samples: two core datasets that used to train the BPNN model, the PKU GIMMS NDVI and Landsat LAI samples are from two unpublished/in-preparation research/articles. This may result in controversy if the authors are unable to provide additional information on these two datasets. a) please provide more details on how the PKU GIMMS NDVI well removed the NOAA orbital drift and AVHRR sensor degradation effects; b) The Landsat LAI samples were trained from MODIS LAI, using what method and sampling strategy? What is the accuracy of the Landsat LAI samples against in-site data?

[Response 2] Thanks for this comment. We have realized that the relatively brief introduction to the PKU GIMMS NDVI product and Landsat LAI samples would raise many concerns about the reliability of these datasets. In the new version of the manuscript, we have added more details on the generating process of the datasets and their performance in absolute accuracy and vegetation trend.

Specifically, a) we have stated how we used cross-calibrated Landsat NDVI with other explanatory variables to further calibrate GIMMS NDVI3g. The effect of NOAA satellite orbital drift and AVHRR sensor degradation could thus be eliminated in the PKU GIMMS NDVI. Using Evergreen Broadleaf Forests (EBF) as an example, a new figure is now available in the supplementary file (Figure S1) to demonstrate the efficacy of our strategy. b) We have also clarified in the revised manuscript how to generate random sample locations across the globe and how to establish and screen sample pairs between Landsat data and MODIS LAI. The sample pairs were used to build Landsat LAI models which were then applied to the Landsat archive. Landsat LAI samples were eventually created at 40,000 sample locations and were verified against in-situ LAI measurements (BELMANIP and ORNL).

The following changes are made in the revised manuscript:

(a) Introduction to PKU GIMMS NDVI in Section 2.1:

"In the generation of PKU GIMMS NDVI, Landsat NDVI from Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and Operational Land Imager (OLI) were first cross-calibrated. Then, massive high-quality Landsat NDVI samples were extracted by considering a series of factors including clouds, cloud shadows, water, snow, aerosol, and radiation performance. The Landsat NDVI samples were employed to calibrate the GIMMS NDVI3g product with other explanatory variables using biome-specific machine learning models. The calibrated NDVI product was finally consolidated with the MODIS NDVI product to extend the temporal coverage to the year 2020. The major improvement of PKU GIMMS NDVI over its counterparts is that it well removed the NOAA orbital drift and AVHRR sensor degradation effects, especially in tropical regions (Figure S1). Its overall R^2, Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) are 0.975, 0.033, and 9%, respectively. It is highly consistent with MODIS NDVI in terms of pixel value (R^2 = 0.962, MAE = 0.032, and MAPE = 6.5%) and global vegetation trend. PKU GIMMS NDVI inherited the quality control (QC) information from the GIMMS NDVI3g. A QC value of 0, 1, and 2 indicates NDVI of good quality, NDVI retrieved from spline interpolation, and NDVI retrieved from average seasonal profile, respectively. The PKU GIMMS NDVI record during AVHRR missions from 1982 to 2015 (before consolidation with MODIS NDVI) was used in this study. It is available at https://zenodo.org/record/7441559#.Y7J7y3ZByCo." (Page 4, Line 97-110)

(b) Introduction to Landsat LAI samples in Section 2.2:

"The Landsat LAI sample dataset provides approximately 4.9 million high-quality samples with a spatial resolution of 1/12 ° and a temporal resolution of half a month (Zha, in preparation). It covers the global vegetated area with all vegetation biome types defined in the MODIS land cover product (the third classification scheme; see section 2.4) and a long-time span from 1984 to 2020. In the generation of Landsat LAI samples, 70,000 sample locations for Deciduous Needleleaf Forests [DNF] and 100,000 sample locations for each of the other vegetation biome types were randomly selected based on the MODIS land cover product. At the sample locations, Reprocessed MODIS LAI (in 500 m resolution; see Section 2.3) and Landsat surface reflectance from TM, ETM+, and OLI scenes (20 × 20 pixels in 30 m resolution) were extracted, creating massive sample pairs. The sample pairs were then rigorously screened by criteria that were not limited to those mentioned in Section 2.1 (i.e., clouds, cloud shadows, etc.) but also included Landsat sample purity, NDVI-LAI relationship, and the saturation state of the MODIS LAI. Biome- and Landsat sensor-specific Random Forest models with other explanatory variables were built based on the sample pairs. The models were applied to historical Landsat data at 40,000 random sample locations (1/12°) to create the final Landsat LAI sample dataset. Validation of the dataset through observations from the BEnchmark Land Multisite ANalysis and Intercomparison of Products (BELMANIP) network (Baret et al., 2006) and the Oak Ridge National Laboratory (ORNL) (Scurlock et al., 2001) showed high absolute accuracies (R^2 = 0.76, MAE = 0.45 m^2/m^2, Root Mean Square Error (RMSE) = 0.66 m^2/m^2). The inter-comparison with the Reprocessed MODIS LAI shows a high temporal consistency. This study selected 3.6 million Landsat LAI samples between 1984 and 2015." (Page 4-5, Line 112-127)

[Comment 3] *Model development: As is well known that, NDVI saturates at the dense* *vegetation areas (LAI around 3), can you explain how the BPNN model utilizes the already saturated NDVI to predict LAI in higher ranges (4-7)?*

[Response 3] It is true that NDVI would saturate at dense canopies. To deal with this effect, we introduced other explanatory variables (including longitude, latitude, month, and satellite number and years since launch) and established biome-specific BPNN models to explain the LAI variations in space, time, satellite, and biome. As such, the predicted LAI was not only determined by NDVI values but other explanatory variables. Figure 3 shows the significant contributions from spatial and temporal variables to the BPNN model accuracy. Figure 5 shows that GIMMS LAI4g could be relatively resistant to saturation. The saturation effect for a particular biome was subtle at a majority of sample locations (red dots in Figure 5) and was observable for samples whose LAI values deviated from the average (yellow and blue dots in Figure 5). For sparse vegetation types including GRA, SHR, CRO, and SAV, the GIMMS LAI4 would be underestimated at their high values; and for dense vegetation types including EBF, DBF, ENF, and DNF, the GIMMS LAI4 would be underestimated at their low and medium values. It should be noted that the LAI fitting line for the global vegetation biome is close to the 1:1 line (Figure 5a-9) and the saturation effect can also be observed in EBF, DBF, ENF, and DNF of GLASS LAI, which was not derived from NDVI data but from surface reflectance.

Despite the potential saturation effect, GIMMS-based NDVI products have been among the most reliable and popular ones because they overcome the limitations of the AVHRR instrument or operation, e.g., vicarious calibration, atmospheric and cloud correction, bias correction, satellite orbital drift, and sensor degradation (Pinzon and Tucker, 2014; Li et al., in review). For example, the GIMMS LAI3g (based on GIMMS NDVI3g) is one of the core data references in the IPCC Sixth Assessment Report for the assessment of global vegetation changes (Eyring et al., 2021), and has been cited for more than 600 times (from *Web of Science*).

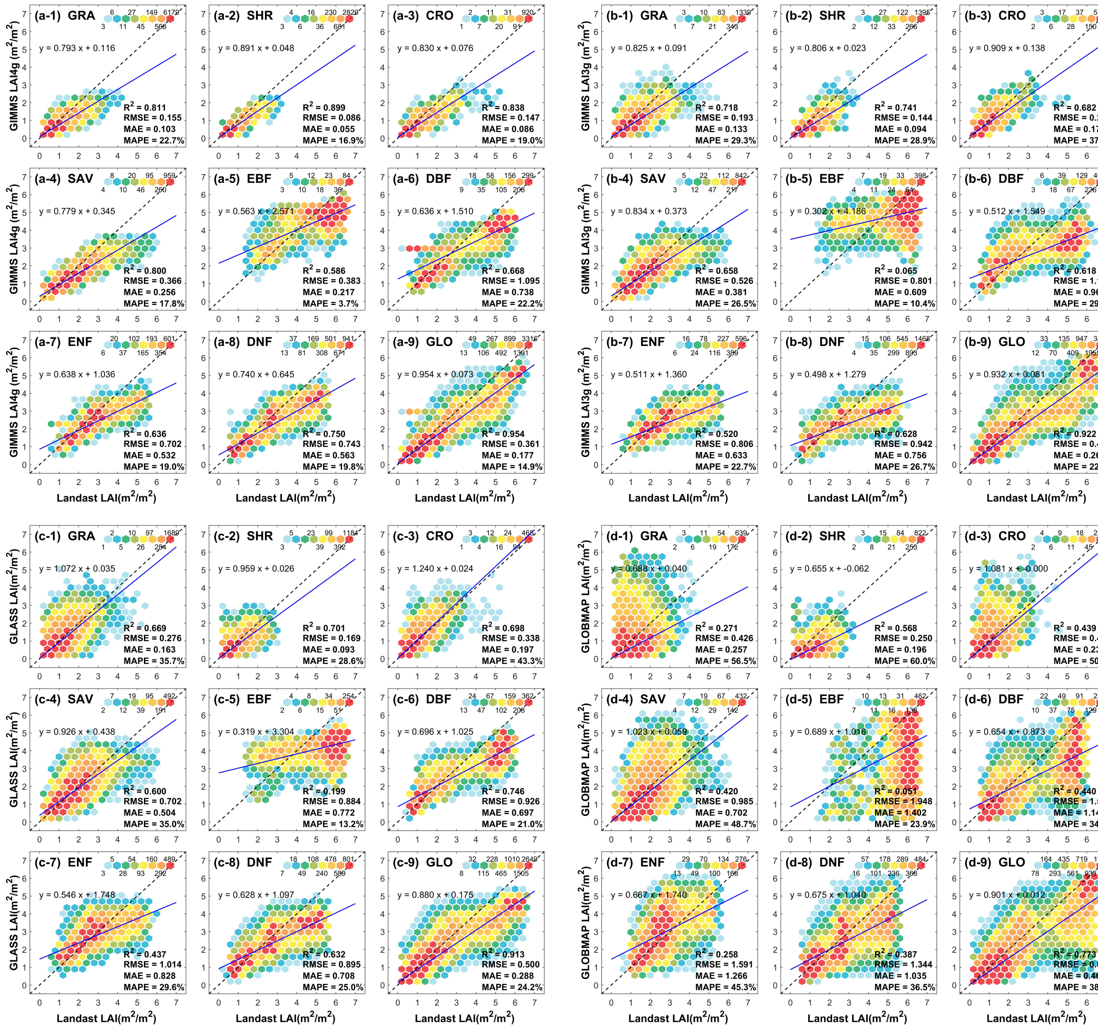

Figure 5. Validation of the (a) GIMMS LAI4g, (b) GIMMS LAI3g, (c) GLASS LAI, and (d) GLOBMAP LAI products in different vegetation biomes using Landsat LAI samples from 1984 to 2015. The error metrics are R^2, RMSE, MAE, and MAPE. GLO represents the global vegetation biome. The color of the dots represents LAI value frequencies in a 0.5 (m^2/m^2) interval.

The following changes are made in the revised manuscript:

(a) Results on the saturation effect in Section 4.3:

"Compared to Landsat LAI samples, the four global LAI products were underestimated in almost all the vegetation types except for CRO of GLASS LAI (Figure 5). We found certain levels of saturation in GIMMS LAI4g and GIMMS LAI3g, such as for high values of GRA, SHR, CRO, and SAV and medium values of EBF, DBF, ENF, and DNF (Figure 5). This could be attributed to the use of NDVI data in LAI models. However, we also observed the saturation effect in EBF, DBF, ENF, and DNF of GLASS LAI, which was not derived from NDVI data. For GIMMS

LAI4g, the saturation was relatively subtle at a majority of sample locations (red dots) and was obvious for locations whose LAI values deviated from the average (yellow and blue dots). The LAI fitting line of the global vegetation biome in GIMMS LAI4g (Figure 5a-9) was close to the 1:1 line." (Page 17, Line 370-376)

(b) Discussions on the saturation effect in Section 5.3:

"The use of PKU GIMMS NDVI could also result in the saturation effect in GIMMS LAI4g as NDVI data tend to saturate at high values (Figure 5). This study established biome-specific models that incorporate multiple explanatory variables besides PKU GIMMS NDVI to account for the LAI variations in space, time, biome, and satellite. This effort could help alleviate the saturation effect though the effect still exists." (Page 27, Line 555-558)

[Comment 4] *Currently, the validation of the GIMMS LAI4g datasets appears to rely heavily on the Landsat LAI sample data, which has not been published and may not necessarily represent the true LAI values on the ground. As a result, there seems to be a lack of independent validation for the proposed dataset. I would like to suggest that the author use the available field data to validate their dataset.*

[Response 4] We have to agree with the reviewer that, the current validation of GIMMS LAI4g based on Landsat LAI samples only could hardly provide solid and comprehensive evidence for its accuracies. In the revised manuscript, independent in-situ LAI data have been introduced. A total of 113 field LAI measurements from BELMANIP 2.1, DIRECT 2.1, and ORNL were employed to evaluate the GIMMS LAI4g product as well as other global long-term LAI products. The results showed that GIMMS LAI4g (R^2 = 0.69, RMSE = 0.86 m^2/m^2, MAE = 0.61 m^2/m^2, MAPE = 33%) had comparable accuracies with GIMMS LAI3g (R^2 = 0.71, RMSE = 0.78 m^2/m^2, MAE = 0.55 m^2/m^2, MAPE = 30%) and GLASS LAI (R^2 = 0.68, RMSE = 0.82 m^2/m^2, MAE = 0.60 m^2/m^2, MAPE = 33%), and it had the lowest underestimation for the fitting line with a slope of 0.87 and an intercept of 0.05 (Figure 4).

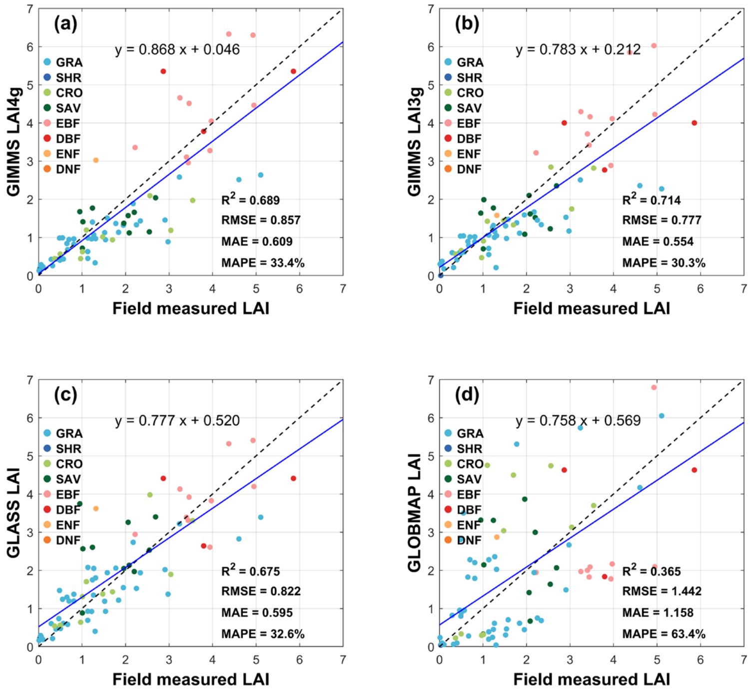

Figure 4. Validation of the (a) GIMMS LAI4g, (b) GIMMS LAI3g, (c) GLASS LAI, and (d) GLOBMAP LAI products using 113 field LAI measurements from 49 sites in the projects of BELMANIP 2.1, DIRECT 2.1, and ORNL. Sites of different vegetation biome types are marked by colors. The error metrics are R^2, RMSE, MAE, and MAPE. The blue fitting lines and dashed 1:1 lines are drawn.

The following changes are made in the revised manuscript:

(a) Introduction to field LAI measurements in Section 2.8:

"The field LAI measurements were from three projects namely, BELMANIP 2.1 (available at https://calvalportal.ceos.org/web/olive/site-description) (Baret et al., 2006), DIRECT 2.1 (available at https://calvalportal.ceos.org/lpv-direct-v2.1) (Garrigues et al., 2008), and ORNL (available at https://daac.ornl.gov/VEGETATION/guides/LAI_guide.html) (Scurlock et al., 2001). The BELMANIP 2.1 and DIRECT 2.1 provide 3 km × 3 km averaged LAI values derived from sites in networks of FLUXNET, AERONET, VALERI, BigFoot, etc. The upscaling from site-based LAI to 3 km × 3 km LAI used high spatial resolution imageries such as Landsat and SPOT. Most global long-term LAI products have utilized the BELMANIP and DIRECT LAI as ground truth for product evaluation (Myneni et al, 2002; Liu et al., 2012; Xiao et al., 2016; Zhu et al., 2013), yet the LAI measurements in both projects were available only after the late 1990s. Note that GLASS LAI (version 4) also employed BELMANIP sites for LAI model training

(Xiao et al., 2016). This study further incorporated ORNL sites which provide field LAI measurements during 1932−2020 despite possible scaling effects due to spatial heterogeneity. We prudently examined all the sites and measurements in BELMANIP 2.1, DIRECT 2.1, and ORNL, and removed measurements that were acquired from heterogeneous sites or coincident among the three projects. Eventually, 113 field LAI measurements from 49 sites were obtained. Information on selected field LAI measurements can be found in Table S1." (Page 6-7, Line 173-187)

(b) Results of validation using field LAI measurements in Section 4.3:
"Based on field LAI measurements, GIMMS LAI4g generated from the BPNN models presented comparable accuracies (R^2 = 0.69, RMSE = 0.86 m^2/m^2, MAE = 0.61 m^2/m^2, MAPE = 33%) with GIMMS LAI3g (R^2 = 0.71, RMSE = 0.78 m^2/m^2, MAE = 0.55 m^2/m^2, MAPE = 30%), and GLASS LAI (R^2 = 0.68, RMSE = 0.82 m^2/m^2, MAE = 0.60 m^2/m^2, MAPE = 33%) (Figure 4). GIMMS LAI3g had the best performance in error measures (i.e., R^2, RMSE, MAE, and MAPE), but GIMMS LAI4g had the lowest underestimation for the fitting line with a slope of 0.87 and an intercept of 0.05 (Figure 4). GLOBMAP LAI presented the largest discrepancies from the LAI measurements." (Page 14, Line 335-340)

[Comment 5] *Providing global distribution maps of GIMMS LAI4g at representative time points and comparing them with other products may be useful and informative for potential users.*

[Response 5] As suggested by the reviewer, we provided LAI distribution maps of multi-year average in January and July for GIMMS LAI4g, GIMMS LAI3g, GLASS LAI, and GLOBMAP LAI (Figure 7).

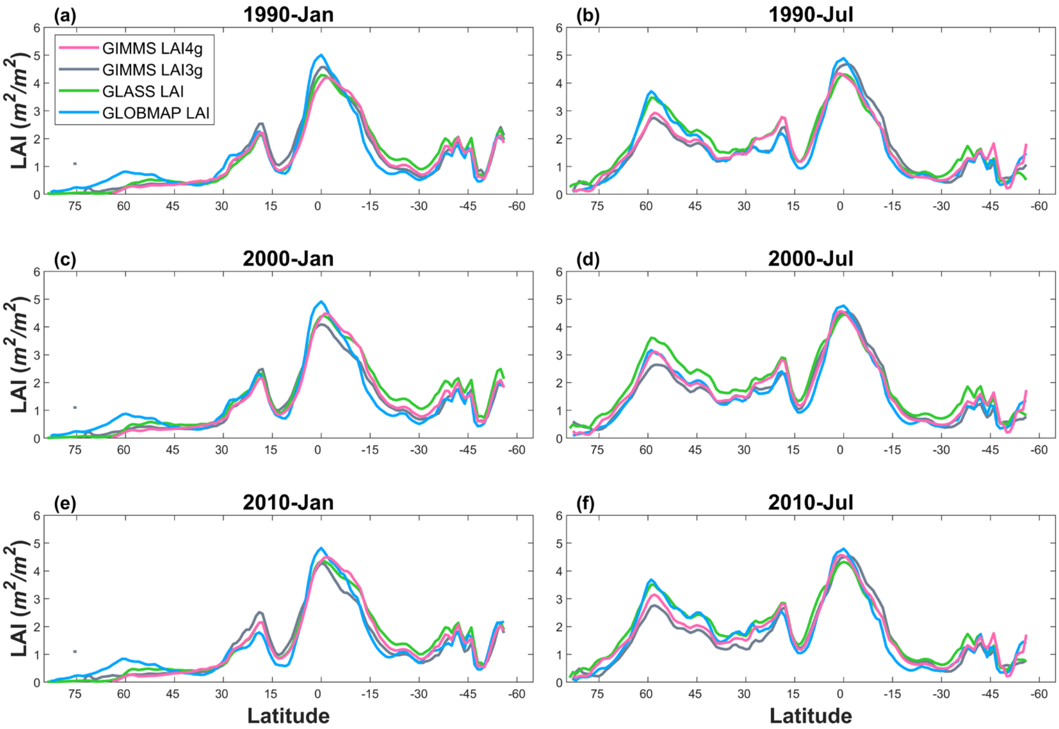

Figure 7. Inter-comparison of spatially averaged LAI along latitude between the GIMMS LAI4g, GIMMS LAI3g, GLASS LAI, and GLOBMAP LAI in January and July of the years 1990, 2000, and 2010. The spatial average was calculated at an interval of 1°.

The following changes are made in the revised manuscript:

(a) GIMMS LAI4g at representative time points in Section 4.3:

"Figure 7 showcases the spatially averaged LAI along latitude in January and July of the years 1990, 2000, and 2010 for GIMMS LAI4g, GIMMS LAI3g, GLASS LAI, and GLOBMAP LAI, respectively. The four LAI products were overall consistent. The GIMMS LAI4g and GIMMS LAI3g had lower values at northern middle latitudes (35°N – 65°N) in July (Figure 7b; Figure 7d; Figure 7f). Also in July, the GLOBMAP LAI and GLASS LAI in the Northern Hemisphere maintained good consistency for the years 1990 and 2010 (Figure 7b; Figure 7f), but the GLOBMAP LAI was systematically lower than GLASS LAI for the year 2000 (Figure 7d). " (Page 18, Line 392-397)

[Comment 6] *I suggest that the authors consider including RMSE as an evaluation parameter in their study. RMSE quantifies the variance of errors and is therefore* *more representative than MAE.*

[Response 6] Thank you for this suggestion. We have included RMSE as an error metric in the revised manuscript. Corresponding results including the texts, figures, and tables have been updated. Please check the revised manuscript as this modification was made in substantial locations.

[Comment 7] *To provide better clarity on potential under or overestimations for high/low values, please add a 1:1 line to Fig. 3.*

[Response 7] This is a good point. The 1:1 lines, LAI fitting lines, and color bars have been added to all sub-figures in Figure 3 (updated Figure 5; see **Response 3**).

[Comment 8] *L430 "We chose not to include the vegetation biome type as an explanatory variable (GIMMS LAI3g did) because the values of the vegetation biome type are deterministic rather than continuous", the NOAA satellite number is also deterministic but why it was adopted as the explanatory variable?*

[Response 8] The reason is that NOAA missions divided the whole time series (1982−2010s) into periods (Figure 8). Building individual models for different NOAA missions would fail to explain the temporal consistency in LAI. We used NOAA satellite number as an explanatory variable so that the BPNN models provided consistent LAI values from 1982−2010s. This has been clarified in the revised manuscript.

The following changes are made in the revised manuscript:

(a) Discussion on BPNN models in Section 5.1:

"The deterministic NOAA satellite number was used as an explanatory variable so that the temporal consistency of the BPNN models can be ensured." (Page 25, Line 521-522)

References:

Li, M., Cao, S., and Zhu, Z.: Spatiotemporally consistent global dataset of the GIMMS Normalized Difference Vegetation Index (PKU GIMMS NDVI) from 1982 to 2020, Earth Syst. Sci. Data, in review.

Pinzon, J. E. and Tucker, C. J.: A non-stationary 1981-2012 AVHRR NDVI3g time series, Remote Sens., 6, 6929-6960, https://doi.org/10.3390/rs6086929, 2014.

To Referee #3

Dear reviewer,

We are very pleased to finish a revised version of the manuscript essd-2023-68 entitled "**Spatiotemporally consistent global dataset of the GIMMS Leaf Area Index (GIMMS LAI4g) from 1982 to 2020**". In preparing this revision we have considered all your comments and incorporated most of the suggestions. We greatly appreciate your time and effort spent reviewing this manuscript, which has improved the revised version of the manuscript.

Substantial improvements have been made based on your comments. Significant ones include:

(1) We have clarified the reason why we consolidate the GIMMS LAI4g with the Reprocess MODIS LAI (MODIS BNU LAI). We finally provide two versions of GIMMS LAI4g, one solely based on AVHRR data and one consolidated with MODIS LAI.

(2) More details have been provided on the Landsat LAI sample dataset including its generation and accuracies.

(3) The removal of NOAA satellite orbital drift and AVHRR sensor degradation is now illustrated for all vegetation biomes.

(4) The bias of vegetation trend in GIMMS LAI4g after consolidation with MODIS BNU LAI has been discussed.

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

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 manuscript describes a new Global Inventory Modeling and Mapping Studies (GIMMS) LAI product (GIMMS LAI4g) for the period 1982-2020. The new product represents a major advance over previous products in that it uses the latest PKU GIMMS NDVI product and a representative subset of Landsat LAI samples to remove sensor degradation and orbital drift that resulted in bias in previous long-term products (e.g., LAI3g). The improved methodology appears to have effectively removed bias and the new product is demonstrated to out-performed other existing widely used LAI products over most of the terrestrial land surface.*

[Response 1] We thank the referee for highlighting the significance of this research.

[Comment 2] *Overall, I feel the manuscript is well written and the methodology is well-described. I commend the authors for this very important undertaking and the major advance of incorporating Landsat observations to help correct bias due to sensor degradation and orbital drift in the long-term AVHRR record. This is a landmark dataset that will be widely used in advancing our understanding of long-term trends in global vegetation dynamics. I have a few major comments below that need to be addressed before I can recommend acceptance of this manuscript for publication.*

[Response 2] We are inspired by this comment. We are happy that the manuscript basically meets the reviewer's standard, and that the reviewer confirms the advantage of eliminating the effect of satellite orbital drift and sensor degradation using the Landsat archive. Efforts have been made to further improve the manuscript and the dataset according to the comments. We hope the revised manuscript could address all the concerns.

[Comment 3] *Major Comments:*

1. The introduction is very well-written and does an excellent job of framing the importance of the research and new product.

[Response 3] Thanks again for this positive comment. Many modifications and improvements have been made in other sections of the manuscript. We hope they are also satisfying.

[Comment 4] *Methods 3.1. The PKU GIMMS NDVI spans the 1982-2020 time period. Why don't the authors simple use the BPNN model to predict LAI values for the full record? Why do the authors instead incorporate MODIS LAI after 2003? This seems to add unnecessary sources of uncertainty into the process. It should hold that the PKU GIMMS NDVI is the best available long-term NDVI product and the LAI reference dataset (1984-2020) would be adequate to build a long-term LAI product using the BPNN model alone.*

[Response 4] We apologize for the confusion. In the current versions of PKU GIMMS NDVI and GIMMS LAI4g, MODIS products were adopted after 2003 or 2004. AVHRR was the only data source before 2000 that provides spatiotemporally continuous observations, but its availability and data quality have been limited since the late 2010s (https://www.usgs.gov/centers/eros/science/usgs-eros-archive-advanced-very-high-resolution-radiometer-avhrr). For instance, the NDVI data in AVHRR-based GIMMS NDVI3g was available until 2015. From this point, it would be inevitable to consolidate AVHRR-based products with more recent products to acquire up-to-date long-term time series. During the period that two products overlapped, a choice has to be made on which one shall be used. As a more advanced sensor, MODIS has been elaborately verified and is believed to provide more reliable data. As such, in both PKU GIMMS NDVI and GIMMS LAI4g, MODIS data were used during the overlapping period.

With this being said, we agree with the reviewer that it would make lots of sense to provide an AVHRR-based LAI product that is independent of MODIS data. Therefore, we have generated the AVHRR-based LAI product (1982−2015) without consolidation with Reprocessed MODIS LAI. This new dataset is also available at https://doi.org/10.5281/zenodo.8035760.

The following changes are made in the revised manuscript:

(a) Added information in the Data availability section:

"In the same repository, we have also provided the version of GIMMS LAI4g that is solely based on AVHRR data, which means that its generation was free from the consolidation with Reprocessed MODISL LAI and it used the version of PKU GIMMS NDVI before consolidation with MODIS NDVI." (Page 28, Line 595-597)

[Comment 5] *Results 4.1. The LAI reference dataset spans 1984-2020 and relates Landsat reflectance to MODIS LAI using machine learning. This paper describing this dataset is currently in review. I think there should be more detail on this product included in the text. For instance, it needs to be clearly stated which MODIS LAI product was used in the generation of the LAI reference dataset so it's clear that there is consistency with the MODIS LAI used in the data consolidation step.*

[Response 5] Thanks for the comment. In the revised manuscript, we have provided more details on the Landsat LAI sample dataset, including its generation process and performance in absolute accuracy and vegetation trend. We hope the updated text could be more informative.

The following changes are made in the revised manuscript:

(a) Introduction to Landsat LAI samples in Section 2.2:

"The Landsat LAI sample dataset provides approximately 4.9 million high-quality samples with a spatial resolution of 1/12 ° and a temporal resolution of half a month (Zha, in preparation). It covers the global vegetated area with all vegetation biome types defined in the MODIS land cover product (the third classification scheme; see section 2.4) and a long-time span from 1984 to 2020. In the generation of Landsat LAI samples, 70,000 sample locations for Deciduous Needleleaf Forests [DNF] and 100,000 sample locations for each of the other vegetation biome types were randomly selected based on the MODIS land cover product. At the sample locations, Reprocessed MODIS LAI (in 500 m resolution; see Section 2.3) and Landsat surface reflectance from TM, ETM+, and OLI scenes (20 × 20 pixels in 30 m resolution) were extracted, creating massive sample pairs. The sample pairs were then rigorously screened by criteria that were not limited to those mentioned in Section 2.1 (i.e., clouds, cloud shadows, etc.) but also included Landsat sample purity, NDVI-LAI relationship, and the saturation state of the MODIS LAI. Biome- and Landsat sensor-specific Random Forest models with other explanatory variables were built based on the sample pairs. The models were applied to historical Landsat data at 40,000 random sample locations (1/12°) to create the final Landsat LAI sample dataset. Validation of the dataset through observations from the BEnchmark Land Multisite ANalysis and Intercomparison of Products (BELMANIP) network (Baret et al., 2006) and the Oak Ridge National Laboratory (ORNL) (Scurlock et al., 2001) showed high absolute accuracies (R^2 = 0.76, MAE = 0.45 m^2/m^2, Root Mean Square Error (RMSE) = 0.66 m^2/m^2). The inter-comparison with the Reprocessed MODIS LAI shows a high temporal consistency. This study selected 3.6 million Landsat LAI samples between 1984 and 2015." (Page 4-5, Line 112-127)

[Comment 6] *Results 4.4. The temporal consistency analysis is critical in demonstrating how effectively sensor degradation and orbital drift bias was removed from the dataset. Currently, only results for EBF are shown (Figure 6). I suggest the authors expand this analysis across biomes and globally. Improved temporal consistency is one of the major advances of this dataset and should be robustly characterized.*

[Response 6] We used EBF as an example for two reasons. First, it suffers from a more obvious effect of NOAA satellite orbital drift and AVHRR sensor degradation than other vegetation biomes (Pinzon and Tucker, 2014; Tian et al., 2015; Zhu et al., 2013). Second, EBF has long been a research focus with an abundance of biomass and intensive vegetation activity. In the supplementary materials of the revised manuscript, we have provided results of the temporal consistency analysis for all other vegetation biomes, where different levels of satellite orbital drift and sensor degradation can be effectively eliminated (Figure S4-S10).

[Comment 7] *After 2003, the GIMMS LAI4g product directly matches the MODIS*

BNU LAI product from Yuan et al., (2011). This could create bias in the trend in LAI4g over the AVHRR time period (1982-2003) and over the MODIS time period (2004-2020). Also, given their reliance on different MODIS products, this could introduce differences in the trends of the PKU NDVI product and the LAI4g product. The MODIS BNU LAI dataset used an algorithm that interpolated and smoothed the original MODIS LAI data. For instance, MODIS BNU LAI is impacted by the follow steps used in the algorithm: 1) MODIS LAI retrievals that experienced saturation were flagged as low-quality and discarded; 2) the upper envelope smoothing process may generate positive bias in the trend. It would be helpful for the authors to show the annual anomalies and trend of the various products used in the derivation of LAI4g including PKU NDVI4g (1982-2020), LAI4g no consolidation (1982-2015), LAI4g consolidation (1982-2020), BNU LAI (2003-2020). This could be show for the full global region in the main paper and PFT level trends could be moved to the SI. Please then explain any differences in the PKU NDVI, BNU LAI, and LAI4g trend over the MODIS era (2003-2020).

Yuan, H., Dai, Y., Xiao, Z., Ji, D., & Shangguan, W. (2011). Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. Remote Sensing of Environment, 115(5), 1171–1187. https://doi.org/10.1016/j.rse.2011.01.001

[Response 7] As mentioned in **Response 4**, in order to acquire spatiotemporally consistent LAI products from 1982 to now, it is inevitable to consolidate the irreplaceable AVHRR-based LAI with the LAI product generated from more recent and advanced sensors. This study selected the MODIS BNU LAI (Reprocessed MODIS LAI in the manuscript) in consolidation because it was argued to have higher accuracy due to the spatiotemporal filtering and the use of quality flag (Samanta et al., 2011; Yuan et al., 2011). Note that the MODIS BNU LAI has also served as the LAI target in generating GIMMS LAI3g and our Landsat LAI samples. As a result, the vegetation trend in GIMMS LAI4g is jointly determined by AVHRR data and MODIS data. The bias could hardly be quantified but as the reviewer suggested, we have compared the annual anomalies and trends of GIMMS LAI4g before (1982−2015) and after (1982−2020) consolidation, MODIS BNU LAI (1982−2020), and PKU

GIMMS NDVI4g (1982−2015) (updated Figure 11). We found a good consistency between them. As suggested by the reviewer, the biome-specific version of Figure 11 is available in Figure S12.

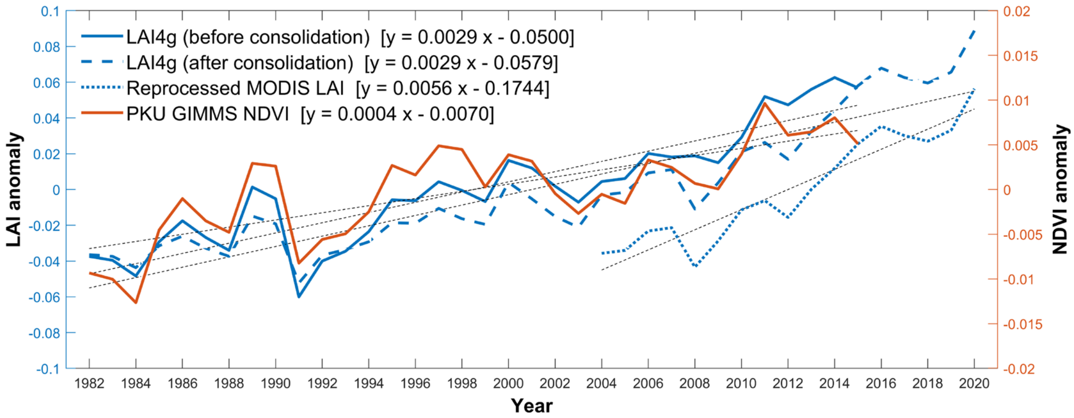

Figure 11. Annual anomalies and trends of GIMMS LAI4g before consolidation (1982−2015), GIMMS LAI4g after consolidation (1982−2020), Reprocessed MODIS LAI (2004−2020), and PKU GIMMS NDVI (1982−2015).

The following changes are made in the revised manuscript:

(a) Comparison of LAI variations between products in Section 4.4:

"Figure 11 demonstrates a good consistency in the overlapping periods between annual variations of the final GIMMS LAI4g product (GIMMS LAI4g after consolidation) and the input and intermediate products. The shapes of the anomalies were similar. The LAI trends for GIMMS LAI4g remained constant before and after consolidation (2.9×10^{-3} $m^2m^{-2}yr^{-1}$), despite Reprocessed MODIS LAI presenting a high greening trend during 2004−2020. The consistency has also been found in a biome-specific manner (Figure S12). This result indicates that both BPNN modeling (PKU GIMMS NDVI vs GIMMS LAI4g before consolidation) and data consolidation (GIMMS LAI4g before consolidation vs GIMMS LAI4g after consolidation) preserved the LAI anomaly and trend." (Page 22, Line 451-457)

[Comment 8] *The GIMMS LAI4g product derived from NDVI4g (before harmonization with MODIS BNU LAI) is a valuable in its own right. This dataset serves as a high-quality estimation of LAI, purely from AVHRR, inter-calibrated with*

Landsat, and independent of other MODIS-based LAI data. I recommend the authors publish this dataset along with the other final version.

[Response 8] The reply to this comment could be referred to **Response 4**. We agree with the reviewer's opinion and have provided the version of GIMMS LAI4g that is purely based on AVHRR data (1982−2015).

[Comment 9] *Minor comments:*

Figure 5. The LAI4g line should be on top of the others and easier to see.

[Response 9] Thank you for this suggestion. In the Figure 5 (updated Figure 7), the curve related to GIMMS LAI4g has been placed on top of others.

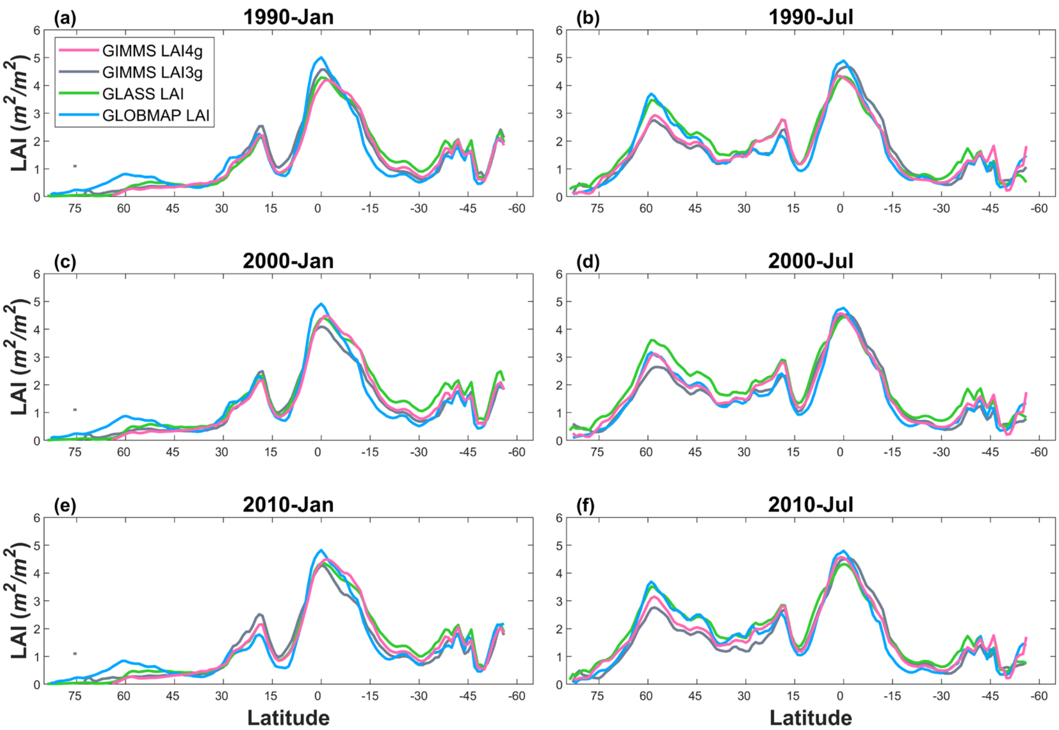

Figure 7. Inter-comparison of spatially averaged LAI along latitude between the GIMMS LAI4g, GIMMS LAI3g, GLASS LAI, and GLOBMAP LAI in January and July of the years 1990, 2000, and 2010. The spatial average was calculated at an interval of 1°.

[Comment 10] *Figure 7. The figure could be reduced to the global timeseries and a few key regions discussed in the paper. The rest of the timeseries could be moved to* *the SI.*

[Response 10] This is a good suggestion. The new figure (updated Figure 9) has included the GIMMS LAI4g time series at the global scale and in selected regions including Europe (Ciais et al., 2005), Amazon (Wang et al., 2013), Congo (Zhou et al., 2014), China, and India (Chen et al., 2019a). Temporal variations of GIMMS LAI4g for vegetation biome were moved to the Supplementary materials. We found a systematic deviation between the GIMMS LAI4g and MODIS LAI products during 2004−2015 in all regions before data consolidation. The deviation has been efficiently eliminated by our pixel-wise consolidation method, especially for the Amazon rainforests. After consolidation, the temporal variations of the GIMMS LAI4g were self-consistent in all periods. Similar results were found regarding the vegetation biome type (Figure S11).

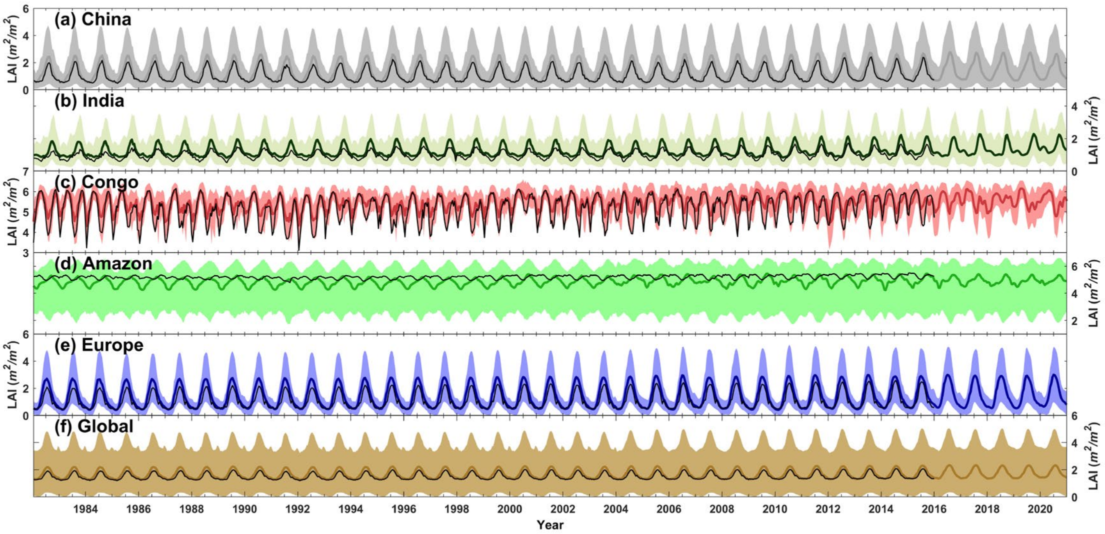

Figure 9. Temporal variations of the GIMMS LAI4g during 1982−2020 in selected hotspot regions of China (a), India (b), Congo (c), Amazon (d), and Europe (e) and at the global scale (f). GLO represents the global vegetation biome. The bold colored line represents the LAI average of GIMMSLAI4g after data consolidation, with shadow covering the value range between 10% and 90% quantiles. The thin black line represents the LAI average of GIMMSLAI4g before consolidation. It should be noted that the GIMMS LAI4g after consolidation shared the same footprint with the Reprocessed MODIS LAI after the year 2004.

The following changes are made in the revised manuscript:

(a) GIMMS LAI4g temporal variation in key regions in Section 4.4:

"Figure 9 shows the GIMMS LAI4g time series before (thin black line) and after (bold colored line) data consolidation from 1982 to 2020 at the global scale and in selected hotspot regions of Europe, Amazon, Congo, India, and China. The GIMMS LAI4g shares the same footprint with the Reprocessed MODIS LAI after the year 2004. Before consolidation, there was a systematic deviation between the GIMMS LAI4g and MODIS LAI products during 2004−2015 in all regions, especially for the Amazon rainforests (Figure 9d). The pixel-wise fusion method has successfully matched the GIMMS LAI4g time series with MODIS LAI, eliminating abnormal shifts in vegetation phenology. This is especially true for the Amazon rainforests, where the phenological curve has been substantially corrected and enhanced by the MODIS LAI (Figure 9d). As a result, the temporal variations of the GIMMS LAI4g after consolidation were self-consistent in all periods. This temporal consistency has also been evaluated regarding the vegetation biome type and similar results were found (Figure S4)." (Page 20, Line 417-425)

References:

Pinzon, J. E. and Tucker, C. J.: A non-stationary 1981-2012 AVHRR NDVI3g yime series, Remote Sens., 6, 6929-6960, https://doi.org/10.3390/rs6086929, 2014.

Samanta, A., Costa, M.H., Nunes, E.L., Vieira, S.A., Xu, L., and Myneni, R.B..: Comment on "Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 Through 2009." Science, 333, 1093–1093. https://doi.org/10.1126/science.1199048, 2011.

Tian, F., Fensholt, R., Verbesselt, J., Grogan, K., Horion, S., and Wang, Y. J.: Evaluating temporal consistency of long-term global NDVI datasets for trend analysis, Remote Sens Environ, 163, 326-340, https://doi.org/10.1016/j.rse.2015.03.031, 2015.

Yuan, H., Dai, Y., Xiao, Z., Ji, D., and Shangguan, W.: Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. Remote Sens. Environ., 115, 1171–1187. https://doi.org/10.1016/j.rse.2011.01.001, 2011.

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.