Open Access
Earth System
Science
Data
Discussions

Spatiotemporally consistent global dataset of the GIMMS Normalized Difference Vegetation Index (PKU GIMMS NDVI) from 1982 to 2020

Muyi Li[1,†] , Sen Cao[1,†], Zaichun Zhu[1,2,*]

[1]School of Urban Planning and Design, Shenzhen Graduate School, Peking University, Shenzhen 518055, China
5   [2]College of Urban and Environmental Sciences, Peking University, Beijing 100871, China

[†]These authors contributed equally to this work

Correspondence: Zaichun Zhu (zhu.zaichun@pku.edu.cn)

Abstract: Global products of remote sensing Normalized Difference Vegetation Index (NDVI) are critical to assessing the vegetation dynamic and its impacts and feedbacks on climate change from local to global scales. The previous versions of the

10   Global Inventory Modelling and Mapping Studies (GIMMS) NDVI product derived from the Advanced Very High Resolution Radiometer (AVHRR) provide global biweekly NDVI data starting from the 1980s, being a reliable long-term NDVI time series that has been widely applied in Earth and environmental sciences. However, the GIMMS NDVI products have several limitations (e.g., orbital drift and sensor degradation) and cannot provide continuous data for the future. In this study, we presented a machine learning model that employed massive high-quality and global-wide Landsat NDVI samples and a data

15   consolidation method to generate a new version of the GIMMS NDVI product, i.e., PKU GIMMS NDVI (1982−2020), based on AVHRR and Moderate-Resolution Imaging Spectroradiometer (MODIS) data. A total of 3.6 million Landsat NDVI samples that well spread across the globe were extracted for vegetation biomes in all seasons. The PKU GIMMS NDVI exhibits higher accuracy than its predecessor (GIMMS NDVI3g) in terms of R^2 (0.975 over 0.942), mean absolute error (MAE: 0.033 over 0.074), and mean absolute percentage error (MAPE: 9% over 20%). Notably, PKU GIMMS NDVI effectively eliminates the

20   evident orbital drift and sensor degradation effects in tropical areas. The consolidated PKU GIMMS NDVI has a high temporal consistency with MODIS NDVI in describing vegetation trends ($R^2 = 0.962$, MAE = 0.032, and MAPE = 6.5%). The PKU GIMMS NDVI product can potentially provide a more solid data basis for global change studies. The theoretical framework that employs Landsat data samples can facilitate the generation of remote sensing products for other land surface parameters.

Keywords: PKU GIMMS NDVI; Landsat; MODIS; Back Propagation Neural Network

1 Introduction

25

The Normalized Difference Vegetation Index (NDVI) characterizes the biophysical, biochemical, and physiological conditions of vegetation (Rondeaux et al., 1996; Gao et al., 2000; Yin et al., 2022). As a normalized ratio of the Near-infrared (NIR) and red bands, it minimizes many forms of multiplicative noise, including soil background, atmosphere, and sun–target–sensor geometry (Rondeaux et al., 1996; Gao et al., 2000; Yin et al., 2022). Due to the long archive, simplicity, and robustness,

30 NDVI is one of the most popular vegetation indices (VIs) that have been widely used in the quantification of vegetation dynamics (Badgley et al., 2017; Gamon et al., 2016; Joiner et al., 2018; Li et al., 2019), ecosystems carbon and water cycles (Zhu et al., 2021; Wang et al., 2021; Schubert et al., 2012; Cui et al., 2021), and environmental stress and disturbances (Aghakouchak et al., 2015; Qin et al., 2021; Peng et al., 2020). NDVI can be acquired from satellite sensors since the 1970s, but it wasn't until the late 1990s that NDVI data of different temporal and spatial resolutions became steadily available from

35 better designed and calibrated sensors such as the Moderate-Resolution Imaging Spectroradiometer (MODIS) (Didan, 2021), 'Satellite Pour l'Observation de la Terre' (SPOT) VEGETATION (SPOT-VGT) (Maisongrande et al., 2004), and Visible Infrared Imaging Radiometer (VIIRS) (Cao et al., 2014). For a long time before the late 1990s, the Advanced Very High Resolution Radiometer (AVHRR) sensor onboard NOAA satellites has been the only NDVI data source that provides frequent and continuous global observations. Several sets of global long-term time-series NDVI products have been released based on

40 AVHRR, such as the Global Inventory Modelling and Mapping Studies (GIMMS) NDVI3g (Pinzon and Tucker, 2014), Long Term Data Record version 4 (LTDR4) NDVI (Pedelty et al., 2007), and Vegetation Index and Phenology version 3 (VIP3) NDVI (Pedelty et al., 2007). These products have provided great insights into how ecological processes of vegetation influence and respond to ongoing climate change (Wang et al., 2021; Zhu et al., 2021; Zhang et al., 2017; Piao et al., 2020; Myers-Smith et al., 2020; Chen et al., 2019). However, uncertainties in the NDVI products have also led to inconsistency not only between

45 different products but also for the same product in different time periods, placing many studies in a dilemma (Wang et al., 2022; Zeng et al., 2022; Fensholt and Proud, 2012).

There are several sources of uncertainties in AVHRR-based NDVI products. The first comes from the discrepancies in band settings (e.g., center wavelength and spectral response function) within AVHRR sensors (i.e., AVHRR-2 and AVHRR-3) as well as with other sensors (such as MODIS and VIIRS) (Yang et al., 2021; Trishchenko et al., 2002; Pinzon and Tucker,

50 2014; Fan and Liu, 2016). Second, NDVI inconsistencies could also occur between the same AVHRR sensors onboard different NOAA satellites. In this case, the sensors would have different image acquisition times and sun–target–sensor geometries, yielding a "jump" (a sudden change in values) phenomenon in NDVI time series (Tian et al., 2015; Frankenberg et al., 2021; Jiang et al., 2017; Los, 1998). For example, the AVHRR sensor onboard NOAA-11 has a considerably larger NDVI than preceding and subsequent AVHRR sensors (Beurs and Henebry, 2004). Third, uncertainties could be introduced

55 by the NOAA satellite orbital drift and AVHRR sensor degradation due to the harsh environments in space (Wang et al., 2022). Artificial signals from the orbital drift in humid areas were evident for the AVHRR-based NDVI products (e.g., VIP3 NDVI, LTDR4 NDVI, and GIMMS NDVI3g) and downstream products such as the GIMMS Leaf Area Index (LAI)3g (Zhu et al., 2013).

For long-term vegetation trend analysis, an accurate global NDVI product requires us to well address the abovementioned

60 uncertainties, particularly the ones related to temporal inconsistency. Some efforts have thus been made in past years (Tucker et al., 2005; Jiang et al., 2008; Doelling et al., 2007; Cao et al., 2004). One strategy performed NDVI calibration using the data acquired when NOAA orbital drift or AVHRR sensor degradation had not occurred. For example, Jiang et al. (2008) used NDVI in the inaugural year of NOAA satellites as a baseline to calibrate NDVI of other years. The other strategy calibrated

AVHRR NDVI with other sensors that have overlapping periods of observation with AVHRR. Pinzon et al. (2014) used
65 SeaWiFS NDVI data as a benchmark to evaluate the consistency of GIMMS NDVIg data with a Bayesian approach. Other
studies have employed SPOT-VGT NDVI data (Tucker et al., 2005), Meteosat-8 NDVI data (Doelling et al., 2007), MODIS
NDVI data (Cao et al., 2004), or VIIRS data (Yang et al., 2021) to calibrate the other NDVI products derived from AVHRR
sensors. The basic assumption behind the two strategies is that the calibration models and parameters derived from one or
more overlapping periods must be static through time. This is not necessarily true because the performance of satellite sensors
70 could be a function of multiple factors that are not limited to their internal settings and the seasonality (Kogan, 1995). Without
a sufficient understanding of product accuracy in all periods, uncertainties in AVHRR NDVI calibration can hardly be
determined.

The Landsat data have the potential to evaluate and calibrate global NDVI products in all periods. As one of the earliest
satellite missions, Landsat satellite series have provided the longest space-based record of Earth's land since the 1970s (Roy
75 et al., 2016; Wulder et al., 2019; Wulder et al., 2016). Landsat sensors have a high spatial resolution, low frequencies of sensor
change, and in particular, high temporal consistency in reflectance (Zhang et al., 2021; Weng et al., 2014; Dong et al., 2020).
Verification results from Pseudo-Invariant Calibration Points (PICS) (such as desert, water, ice, and snow) showed that the
temporal variations of Top-Of-Atmosphere (TOA) reflectance were less than 2% for most Landsat sensors (except for Landsat
5 TM at SWIR 2 which is 3%) during their orbit time (Helder et al., 2013). Although the relatively small field-of-view and
80 long revisit period has limited Landsat for global applications (Maisongrande et al., 2004), its excellent temporal consistency
has aided some important studies of vegetation trend via sample analysis, such as in the Arctic region from 1984 to 2016
(Berner et al., 2020). In recent years, an increasing number of studies have used Landsat data for global dataset production via
tools such as the Google Earth Engine (GEE) platform (Zhang et al., 2022; Cao et al., 2021).

In this context, this study uses the long-term Landsat sample data to develop a new version of the GIMMS NDVI product
85 (PKU GIMMS NDVI) (1982–2020) from the GIMMS NDVI3g (current version) (1982–2015) and MODIS NDVI products
(2001–2020). We first cross-calibrate NDVI data from different Landsat missions and extract a mass of high-quality Landsat
NDVI samples worldwide for all time periods (1984–2015). Based on the samples, we generate the PKU GIMMS NDVI using
biome-specific Back Propagation Neural Network (BPNN) models with GIMMS NDVI3g data and selected explanatory
variables (the longitude and latitude, associated month, and the NOAA number and years since launch). Then, the temporal
90 coverage of PKU GIMMS NDVI is extend to the year 2020 by consolidating with the MODIS NDVI product using a pixel-
by-pixel linear regression method. Results of Landsat NDVI cross-calibration are reported. We directly validate PKU GIMMS
NDVI's accuracy via individual Landsat NDVI samples and compare it with GIMMS NDVI3g. We also examine the accuracy
distribution in space for both products and demonstrate the performance of PKU GIMMS NDVI in alleviating uncertainties
from the orbital drift and sensor degradation. The consolidation of PKU GIMMS NDVI with MODIS NDVI, and the
95 performance of consolidated PKU GIMMS NDVI in characterizing vegetation trends are also evaluated.

2 Data

Four global satellite products were used in this study: Landsat Surface Reflectance data (Collection 1 Tier 1), MODIS Land-Cover Type product (V6.1), GIMMS NDVI3g product (V1.0), and MODIS NDVI product (V6.1). The Landsat Surface Reflectance data were used to generate NDVI samples. The MODIS Land-Cover Type product was used to label NDVI samples with vegetation biome types. The GIMMS NDVI3g was the main data source from which our PKU GIMMS NDVI product would be created. The MODIS NDVI product was used to extend the temporal coverage of the generated PKU GIMMS NDVI product.

2.1 Landsat Surface Reflectance Data (Collection 1 Tier 1)

We obtained Landsat Surface Reflectance data between 1984 and 2015 with a spatial resolution of 30 m from the Google Earth Engine (GEE) platform. These data comprised the Collection 1 (Tier 1) datasets of Landsat 5 (TM), 7 (ETM+), and 8 (OLI), produced by the United States Geological Survey (USGS). Landsat 5 was launched in March 1984 and retired in January 2013. Landsat 7 and 8 were launched in April 1999 and February 2013, respectively, and are still in operation. The USGS uses the Landsat Ecosystem Disturbance Adaptive Processing System (Masek et al., 2006) to perform atmospheric and terrain corrections for Landsat 5 and Landsat 7 and uses the Landsat 8 Surface Reflection Code (Vermote et al., 2016) to perform corrections for Landsat 8. Previous studies have revealed that Landsat reflectance data have good temporal consistency that can be used to produce a set of long-term stable benchmarks (Helder et al., 2013). However, this study found a systematic deviation between Landsat 5/Landsat 8 and Landsat 7. The correction method for the systematic deviation is described in Section 3.1.

2.2 GIMMS NDVI3g (V1.0) Product

This study selected the latest version (third generation) of the GIMMS NDVI dataset (GIMMS NDVI3g, V1.0) generated from AVHRR sensors onboard a series of NOAA satellites (NOAA 7, 9, 11, 14, 16, 17, and 18) (Pinzon and Tucker, 2014). The GIMMS NDVI3g dataset has a spatial resolution of 1/12°. Half-month maximum NDVI composite was used to eliminate the atmospheric effects on the NDVI magnitude. This compositing scheme resulted in two maximum NDVI values per month. The entire available GIMMS NDVI3g record, which extends from July 1981 to December 2015, was used in this study. Pixels with negative NDVI values referred to snow and other contaminated data (e.g., pixels with large inland water bodies) and pixels of bad quality, determined by the Quality control (QC) layer, were removed from all analyses.

2.3 MODIS Vegetation Index Product (MOD13C1, V6.1)

The MODIS Vegetation Index product (MOD13C1) (Didan, 2021) is accessible at NASA's Earth Observing System Data and Information System (EOSDIS) (https://search.earthdata.nasa.gov/). The product has a temporal resolution of 16 days and a spatial resolution of 0.05°. To match the temporal and spatial resolutions with the GIMMS NDVI3g product, we first
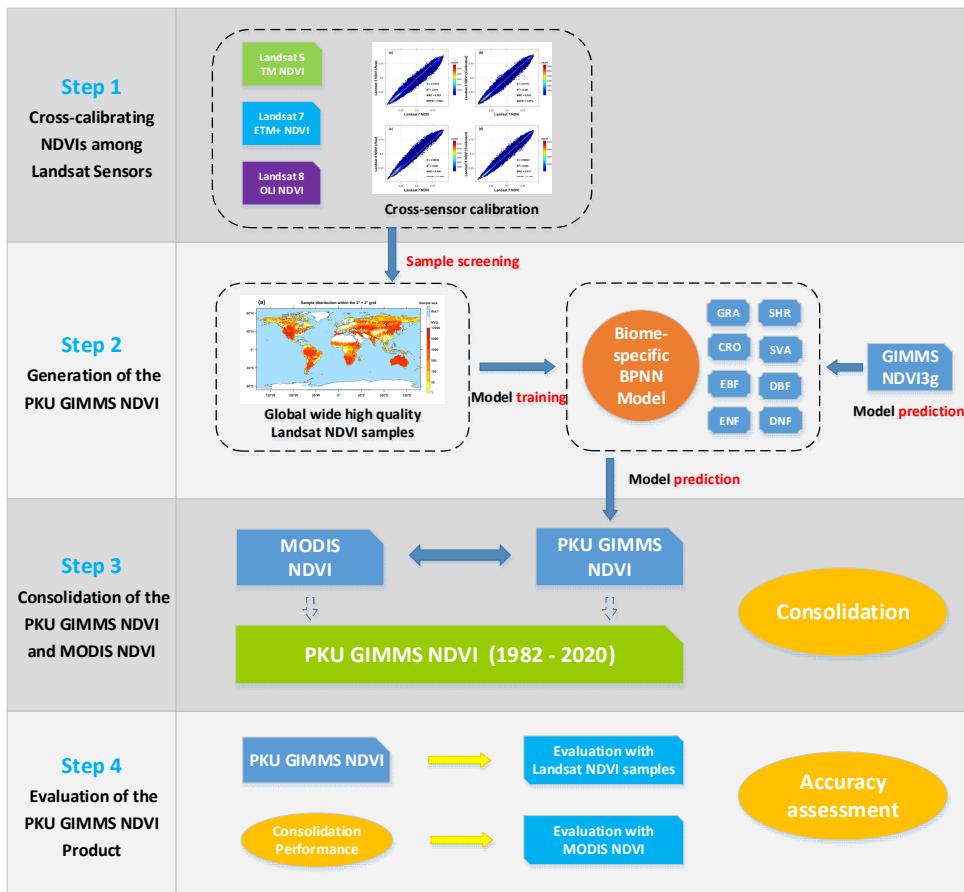
performed time-weighted aggregation on MODIS NDVI to produce an NDVI product at a temporal resolution of half-month. Then we performed nearest neighbour sampling to downscale the spatial resolution to 1/12°.

2.4 MODIS Land-Cover Type Product (MCD12Q1, V6.1)

The MODIS Land-Cover Type product (MCD12Q1) provides global maps of land cover with a spatial resolution of 500 m for each year between 2001–2019 (Friedl et al., 2002). It has five legacy classification schemes, including International Geosphere-Biosphere Program (IGBP) classification system, University of Maryland (UMD) classification system, Leaf Area Index (LAI) classification system, BIOME-Biogeochemical Cycles (BGC) classification system, Plant Functional Types (PFT) classification system, and FAO-Land Cover Classification System (LCCS) classification system. The LAI classification scheme was used in this study. The LAI classification scheme has 11 classes, including eight natural vegetation types (evergreen needleleaf forests [ENF], evergreen broadleaf forests [EBF], deciduous needleleaf forests [DNF], deciduous broadleaf forests [DBF], shrublands [SHR], savannas [SAV], grasslands [GRA], croplands [CRO]) and three non-vegetated lands (water bodies [WAT], non-vegetated lands [NVG], and urban and built-up lands [URB]). In data analysis, we also merged the natural vegetation types into one global vegetation biome [GLO]. 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). The vegetation biome type with the highest frequency from 2001–2019 was considered as the vegetation biome type from 1982–2020. This could be a margin of error but it is the best option.

3 Methodology

A schematic overview of the methodology involves four key steps as illustrated in Figure 1: 1) Landsat sensor cross-calibration to create temporal consistent Landsat data as a benchmark; 2) Generation of PKU GIMMS NDVI from GIMMS NDVI3g using per-biome Landsat NDVI samples, Back Propagation Neural Network (BPNN) models and other explanatory variables; 3) Consolidation of PKU GIMMS NDVI with MODIS NDVI to extend the temporal coverage of PKU GIMMS NDVI to the year 2020; and 4) Evaluation of PKU GIMMS NDVI in terms of its performance in temporal and spatial accuracies and in eliminating the orbital drift and sensor degradation.

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

3.1 Cross-calibrating NDVIs among Landsat Sensors

150   Systematic deviation exists in NDVIs between Landsat 5, Landsat 7, and Landsat 8 (Berner et al., 2020). Specifically, NDVI derived from Landsat 5 is smaller than that from Landsat 7 and NDVI from Landsat 7 is smaller than that of Landsat 8 (Berner et al., 2020) (Figure 2a and Figure2c). As the Landsat NDVI served as a benchmark in this study, the systematic deviations were first removed. We adopted the method by Berner et al. (2020) that used BPNN to calibrate Landsat 5 and Landsat 8 to the Landsat 7 level. The reason for considering Landsat 7 as the benchmark is that Landsat 7 has overlapping

155   periods with both Landsat 5 and Landsat 8.

3.1.1 Sample locations

For the Landsat sensor cross-calibration, one hundred thousand (100,000) sample locations were randomly selected for each vegetation biome type from the MCD12Q1. For each sample (500 m resolution), a matrix of 20 × 20 Landsat pixels (30 m resolution) was extracted at the sample centre from Landsat 5, Landsat 7, and Landsat 8 images acquired between 1984 and

Open Access Earth System Science Data Discussions

160 2015. The Landsat pixels at each sample location were further refined to guarantee that only high-quality clear-sky measurements were included in our study.

First, all Landsat data during August 1991 and December 1992 when Mount Pinatubo erupted were excluded. Second, the abundance of aerosols and thin clouds was used to determine the quality of the sample location (and associated Landsat pixels). If many of the pixels had a high atmospheric opacity (provided by Landsat products), the whole sample location was
165 removed. For Landsat 5 and Landsat 7, the threshold of average atmospheric opacity was set to 0.3. For Landsat 8, the percentage of clear pixels (which have an atmospheric opacity index of 1) must be higher than 80% (320 pixels). Third, the quality of the Landsat scene, cloud contamination, and radiation magnitude were used to determine the quality of individual pixels. A pixel was marked as low-quality if (1) the associated Landsat scene had excessive cloud coverage (> 80 %), (2) the pixel was contaminated by clouds, cloud shadows, water, or snow judged by the CF Mask algorithm (Foga et al., 2017; Zhu
170 et al., 2015), or (3) the pixel had implausibly high (>1) or extremely low (0.001) surface reflectance due to radiation saturation and atmospheric adjustment. In this study, the whole sample location was removed if the percentage of high-quality pixels was lower than 90% (360 pixels).

NDVI was calculated and averaged from high-quality pixels at the remaining sample locations. The sample locations were divided into eighty percent for model training and twenty percent for model evaluation.

175 **3.1.2 Cross-calibration using BPNN models**

BPNN is one of the most popular and established Artificial Neural Network (ANN) algorithms used in ecological studies (Hong et al., 2021; Meng et al., 2020; Yang et al., 2018). An ANN is a machine-learning algorithm inspired by the structure and function of biological neural networks. A typical ANN comprises input (explanatory variables), output (target variable), and hidden layers, each containing artificial neurons whose numbers range from several to hundreds. In the model
180 training of BPNN, signals flow from the input layer to the output layer, after likely passing through several hidden layers. Errors in the output layer propagate backward to the previous layers until they satisfy the user-defined threshold, and the network attempts to minimize the discrepancies between observations and predictions.

This study used NDVI sample locations (500 m resolution) in the overlapping periods between Landsat 7 and Landsat 5/Landsat 8 to train BPNN models. The models were then extrapolated to calibrate Landsat 5 and Landsat 8 in non-overlapping
185 periods. The extrapolation to non-overlapping periods was reliable on the basis that the optical sensors onboard Landsat satellites are temporally consistent with themselves and the reflectance data have been well geometrically and radiometrically calibrated (Irons et al., 2012; Wulder et al., 2019; Wulder et al., 2016). Specifically, NDVI values from Landsat 7 and Landsat 5/Landsat 8 were paired at each sample location (Section 3.1.1) if their acquisition times were less than 10 days. In total, 12,718,863 sample pairs were obtained for all vegetation biome types. When training the BPNN model, NDVI of Landsat
190 5/Landsat 8 was used as the explanatory variable, and NDVI of Landsat 7 was the target variable. We also included the image acquisition time (day of the year) and the sample location's spatial coordinates (longitude and latitude) as covariates to explain potential seasonal and regional variations in the samples.

3.2 Generation of the PKU GIMMS NDVI

3.2.1 Landsat NDVI samples

195   The cross-calibrated Landsat data were used to calibrate the GIMMS NDVI3g product. A total of 40,000 sample locations were randomly selected from the GIMMS NDVI3g product with a spatial resolution of 1/12°. Then, nine Landsat pixels from all available single Landsat 5, Landsat 7, and Landsat 8 scenes were uniformly placed within each sample pixel (1/12°). The Landsat pixel quality was examined in the same way as Section 3.1.1. The sample location was considered valid if the number of high-quality pixels exceeded 5 (out of nine). Its NDVI value was calculated as the average from high-quality

200   Landsat pixels. The samples were also divided into eighty/twenty percent for model training and evaluation.

3.2.2 BPNN models with GIMMS NDVI3g and other explanatory variables

With the Landsat NDVI samples (1/12° resolution), the BPNN model was also used to calibrate the GIMMS NDVI3g product. In the BPNN model, GIMMS NDVI3g data from 1984 to 2015 were used as an explanatory variable, and the Landsat NDVI was the target variable. We also included other explanatory variables associated with spatial, temporal, and satellite

205   information. The spatial information (longitude and latitude) accounts for the spatial autocorrelation in image samples; temporal information (month) accounts for vegetation dynamics; and satellite information (NOAA satellite number and years since its launch) accounts for issues from NOAA orbit drift and AVHRR sensor degradation. One BPNN model was built for each of the eight vegetation biome types. GIMMS NDVI3g was first explored as a single explanatory variable in the BPNN model, and other explanatory variables were added in an enumerative order. In detail, five feature combinations were set up

210   to evaluate their impacts on the BPNN model: (S1) NDVI alone; (S2) NDVI and spatial information (longitude and latitude); (S3) NDVI, spatial information, and time information (month); (S4) NDVI, spatial information, time information, and NOAA satellite number; and (S5) NDVI, spatial information, time information, NOAA satellite number and years since its launch. The optimal parameters for each enumeration were derived through ten-fold cross-verification. The final BPNN model for NDVI calibration was determined with an appropriate set of explanatory variables and the optimal parameters.

215 ### 3.3 Consolidation of the PKU GIMMS NDVI and MODIS NDVI

Over the past two decades, GIMMS NDVI3g products have been extensively utilized for spatiotemporal dynamic monitoring of vegetation, carbon and water cycles of ecosystems, and other related studies. They have provided powerful data support for several significant conclusions in Earth and environmental sciences. However, the latest data in GIMMS NDVI3g is until 2016 and no further upgrades will be provided. This study extended the temporal coverage of GIMMS NDVI3g so that

220   the investigation of recent responses and feedback of vegetation to climate change can be possible. The MODIS NDVI product has excellent precision, temporal consistency, and a long-time span. It is considered the best medium–high resolution global NDVI produced over the past two decades. It could be utilized as an effective extension of the PKU GIMMS NDVI.

However, the band settings of MODIS are different from that of AVHRR. A simple combination of these two products would lead to systematic inconsistencies in NDVI values. Some methods have been proposed to deal with this issue, such as
225 maximum-minimum stretching (Yang et al., 2021), histogram matching (Jiang et al., 2008), and machine learning ((Berner et al., 2020)). In this study, we selected a pixel-scale linear fusion method (Mao et al., 2012) to splice the PKU GIMMS NDVI product (1982–2015) and MODIS NDVI product (2001–2020). This pixel-wised method has demonstrated more accurate than the global models (Yang et al., 2021). 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. Then, NDVI data in the overlapping periods (2004–2015) were
230 extracted from both products. The pixel-wised linear regression model was constructed based on the extracted NDVI. This study found that the linear regression model might not be significant in EBF due to frequent clouds. As such, we applied the regression model to the non-overlapping periods 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. The final PKU GIMMS NDVI product comprised the NDVI product derived from GIMMS NDVI3g between 1982 and 2003 and the MODIS NDVI product between 2004 and 2020.

235 **3.4 Evaluation of the PKU GIMMS NDVI Product**

This study used a direct verification method to evaluate our product of PKU GIMMS NDVI (Justice et al., 2000). For different vegetation biome types, the PKU GIMMS NDVI product was compared to Landsat NDVI values at the remaining twenty percent of the sample locations (1/12°). As a comparison, the GIMMS NDVI3g was also evaluated in the same manner. Five metrics were calculated for accuracy assessment, i.e., sample number (N), R^2, Mean Absolute Error (MAE), Mean
240 Absolute Percentage Error (MAPE), and bias. R^2 measures the percentage of variations that models can explain; MAE and MAPE measure absolute and relative error values at the sample level; and bias measures errors at the global level. The spatial distribution of R^2 was analysed for the PKU GIMMS NDVI and GIMMS NDVI3g products in 2°×2° grids. To highlight the differences between AVHRR-2 and AVHRR-3, NDVI products were evaluated in two separate periods (AVHRR-2: 1982–2000 and AVHRR-3: 2001–2015).

245 We also evaluated the performance of our PKU GIMMS NDVI product in alleviating the effects of orbital drift and sensor degradation; and compared it to the GIMMS NDVI3g product. Tian et al. (2015) observed that the GIMMS NDVI3g product showed a noticeable artefact in humid areas, which may have been caused by the NOAA satellite orbit drift and AVHRR sensor degradation. Zhu et al. (2013) also documented the significant orbital drift in the tropics. However, their conclusions either lacked a quantitative analysis or were solely based on statistical observations at a regional scale because
250 long-term, continuous, and time-consistent benchmark data before 2000 were lacking. This study used NDVI bias in the tropical vegetation type of EBF to measure the magnitude of the orbital drift and sensor degradation effect. The bias was calculated as the mean value of NDVI deviation relative to Landsat NDVI in percentage (Helder et al., 2013) (Eq. 1).

$$bias \% = \frac{1}{N} \sum_{i=1}^{N} \frac{(NDVI_{GIMMS} - NDVI_{Landsat})}{NDVI_{Landsat}} \times 100\% \qquad (1)$$
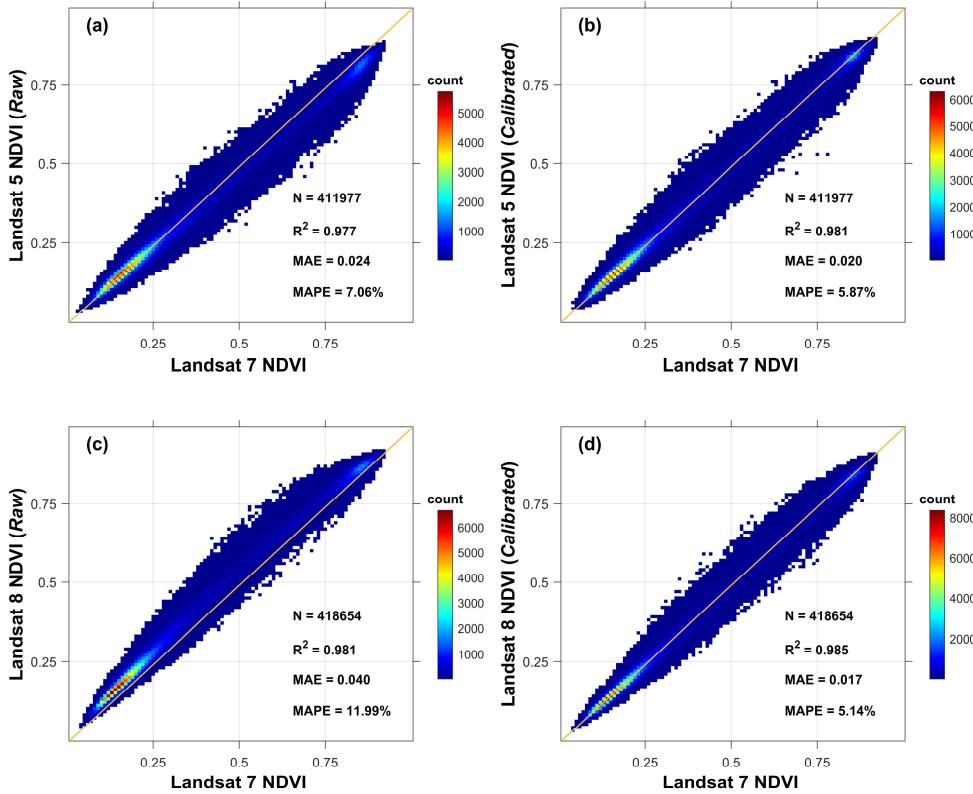
where N denotes the number of Landsat NDVI samples. If there is orbital drift or sensor degradation, the bias will drastically fluctuate; otherwise, it remains constant. Seasonal fluctuations in the time series of NDVI bias were first removed via the multi-year averaging method. Then, inter-annual trends of the bias were extracted via the Ensemble Empirical Mode Decomposition (EEMD) approach (Huang et al., 1998).

The consolidation of PKU GIMMS NDVI with MODIS NDVI was evaluated by 1,000 random points for each vegetation biome type. Using MODIS NDVI as the reference, R^2, MAE, and MAPE for PKU GIMMS NDVI before and after consolidation were calculated and compared during the overlap period (2003–2015). To evaluate the performance of PKU GIMMS NDVI in characterizing vegetation trends (greening or browning), we performed linear regression analysis on the time series of annual average NDVI between 2004 and 2015 at each pixel. The linear regression slope could represent a green trend (positive slope value) or a browning trend (negative slope value). Trends from multiple NDVI products, i.e., GIMMS NDVI3g, MODIS NDVI, and PKU GIMMS NDVI (before and after consolidation), were compared.

4 Results

4.1 Cross-calibration between Landsat 7 and Landsat 5/Landsat 8

More than 12 million Landsat sample pairs (600 m resolution) were acquired for Landsat sensor cross-calibration. Based on the samples, sixteen BPNN models were established to calibrate Landsat 5 NDVI and Landsat 8 NDVI for eight vegetation biome types. Figure 2b and Figure 2d show the NDVI calibration results of Landsat 5 and Landsat 8 against Landsat 7. Both relationships were strong with high R^2 ($R^2 = 0.981$ for Landsat 5 and 0.985 for Landsat 8), low MAE (MAE = 0.020 for Landsat 5 and 0.017 for Landsat 8), and low MAPE (MAPE = 5.87% for Landsat 5 and 5.14% for Landsat 8). Compared to uncalibrated data (Figure 2a and Figure 2c), negative deviation in Landsat 5 NDVI and positive deviation in Landsat 8 NDVI have been efficiently eliminated.

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. 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.
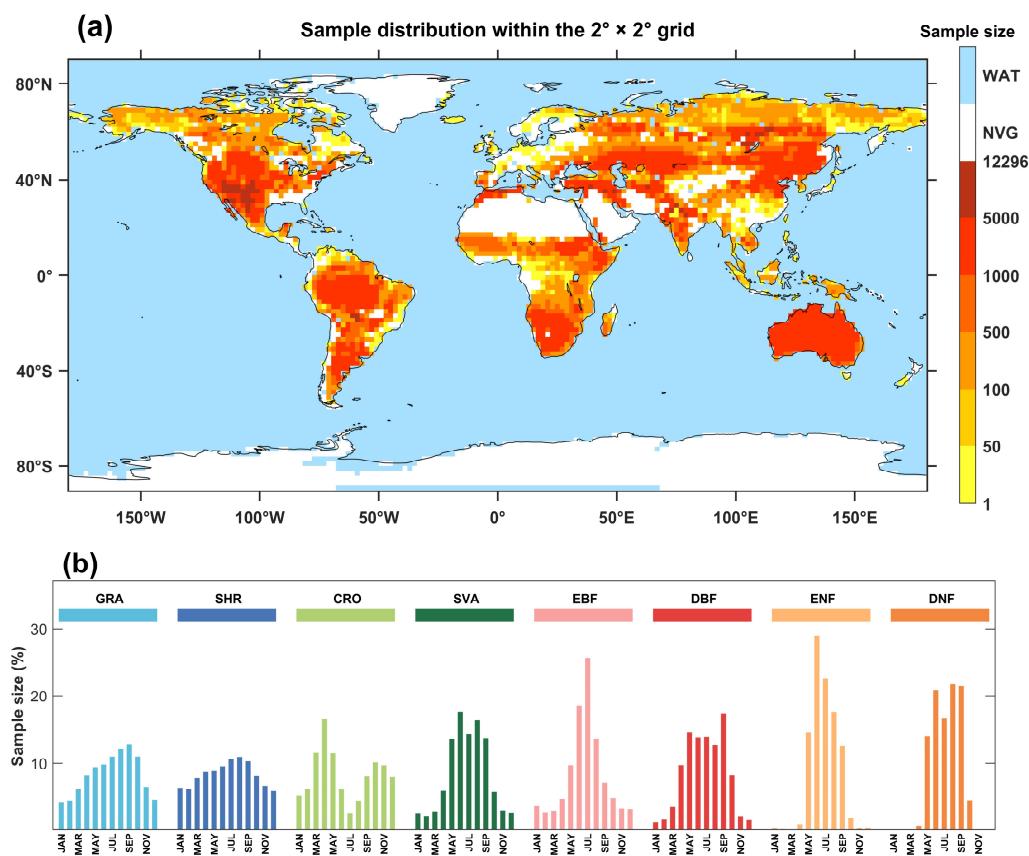
4.2 The PKU GIMMS NDVI product

4.2.1 Spatiotemporal Representativeness of the Landsat NDVI Samples

275   Approximately 3.6 million Landsat NDVI samples (1/12°) from 1984 to 2015 were obtained for GIMMS NDVI3g calibration. The count and spatiotemporal distribution of the samples primarily depended on the availability of Landsat images, which were affected by clouds, cloud shadows, aerosols, climatic conditions, and other factors. The sample count per vegetation biome type was approximately proportional to its total area of coverage (Figure 3a and Figure 3b).

In the spatial domain, our samples covered most vegetated regions worldwide (Figure 3a). Meanwhile in some regions,
280   the number of high-quality samples was relatively small. These regions include (1) northern high latitudes, where suffer the polar night phenomenon, high solar zenith angle, and high observation zenith angle, (2) tropical rainforest areas with abundant precipitation and clouds which lower the quality of remote sensing data, and (3) the Sichuan Basin in Southwest China and

11

areas with a temperate marine climate (e.g., Western European continent, British Isles, and west coast of North and South America). In the time domain (Figure 3b), the samples of vegetation biome types showed single (for most biomes except CRO) or double (CRO) peaks depending on the time of their growing seasons. This guaranteed sufficient samples in the growing season for accurate NDVI prediction with BPNN. For the biomes of ENF and DNF that are primarily distributed in the high northern latitudes, the number of samples in winter (October to April) was < 500. We resolved this problem by reducing the explanatory variables in the BPNN model. We did not select samples from non-vegetated regions because we were not concerned with non-vegetated areas.

Figure 3. Spatial and temporal distribution of refined Landsat NDVI samples (3.6 million). (a) Distribution of Landsat NDVI samples within the 2° × 2° grid. (b) Percentage of samples among the eight vegetation biome types in each month.
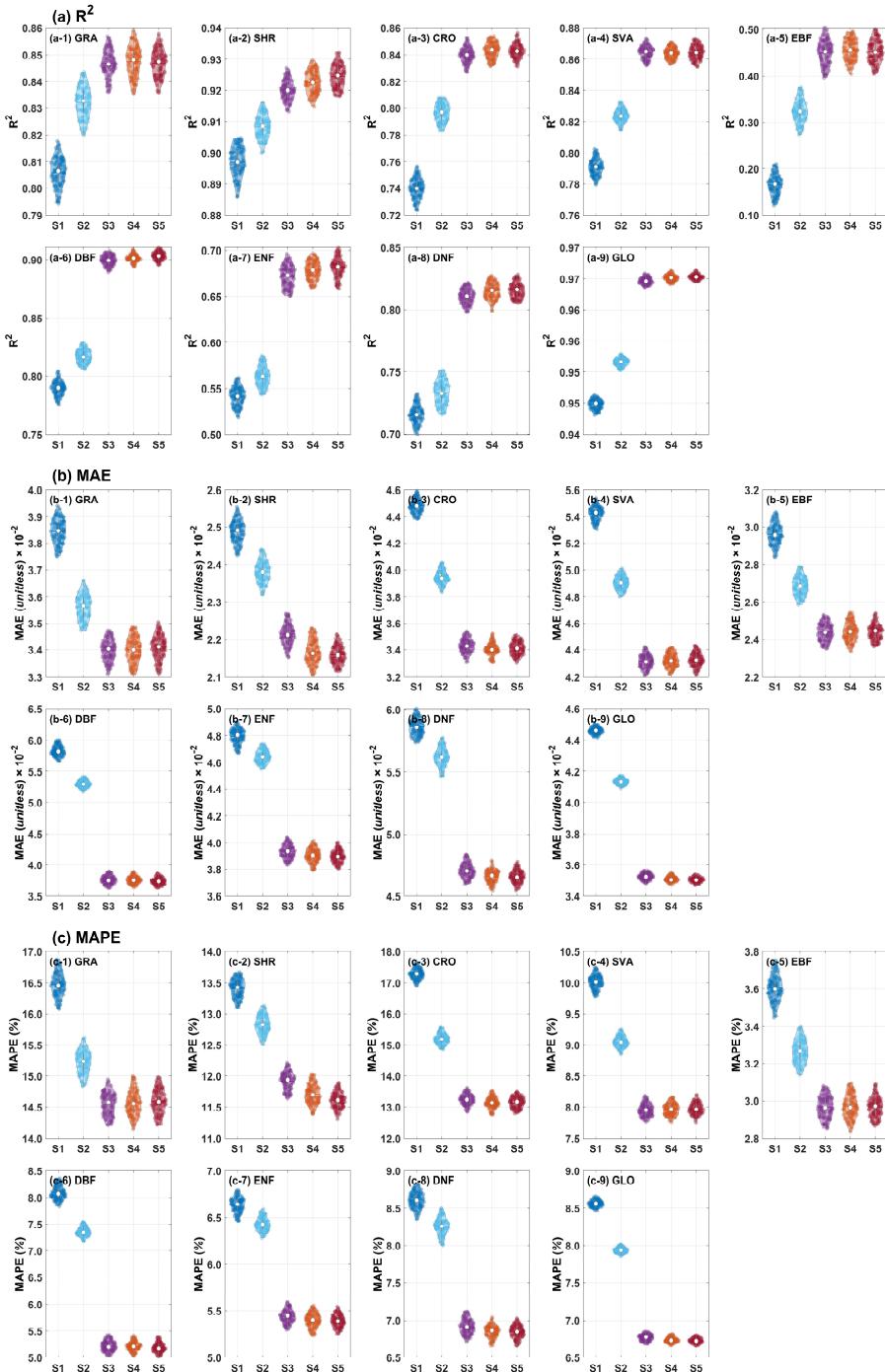
4.2.2 Selection of Explanatory Variables for the BPNN Model

The accuracy of BPNN models under different combinations of explanatory variables (S1 to S5) is shown in Figure 4. The addition of spatial location significantly improved the accuracy of predicted NDVI for the vegetation biome types that are distributed worldwide. The improvement has not been observed for vegetation biome types that are relatively concentrated (e.g., ENF and DNF). The addition of temporal information improved the accuracy of vegetation types with prominent seasonal

295   variations such as DBF and DNF. Finally, adding the NOAA satellite number and orbit time could also improve the accuracy of BPNN models, especially for SHR.

For the combination containing all explanatory variables (S5), the R^2 of most vegetation biome types except for EBF and ENF was > 0.8. For vegetation biomes overall, the R^2 reached 0.96, and the relative error was only 11.35%. Therefore, all available explanatory variables, i.e., the NDVI, longitude, latitude, month, NOAA satellite number, and years since the NOAA

300   satellite's launch, contributed to the BPNN model in this study.

Figure 4. Performance of different combinations of explanatory variables (S1 to S5) for BPNN models. (a), (b) and (c) show the R^2, MAE, and MAPE of the BPNN models, respectively. GLO represents the global vegetation biome. The combinations of explanatory variables are: (S1) NDVI alone; (S2) NDVI and spatial information (longitude and latitude); (S3) NDVI, spatial information, and time
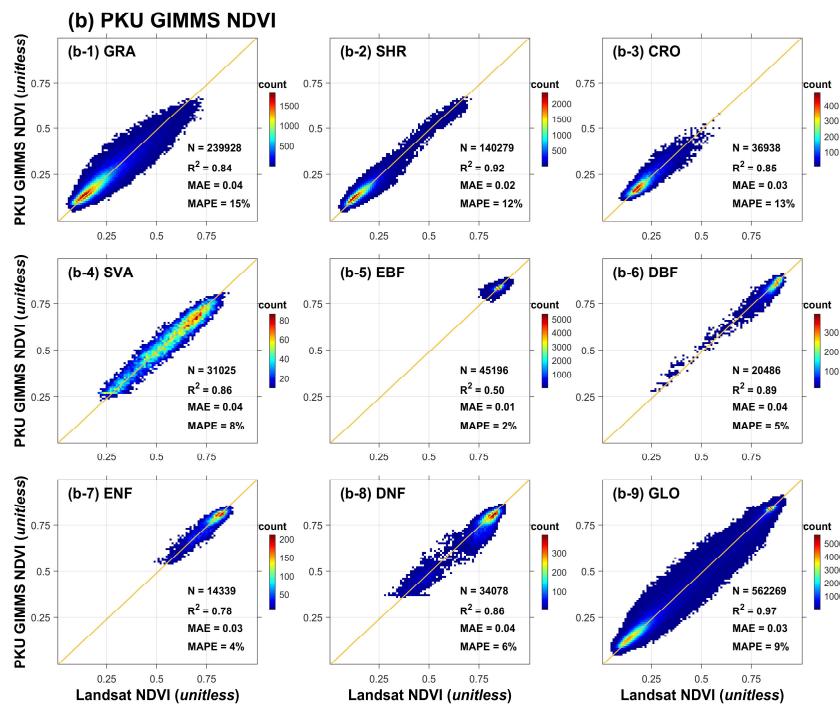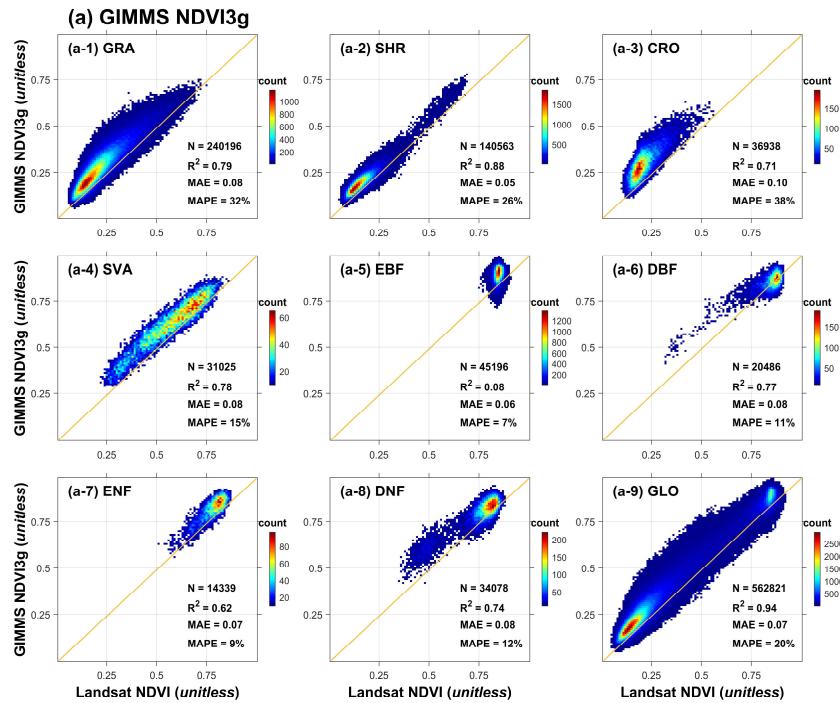
information (month); (S4) NDVI, spatial information, time information, and NOAA satellite number; and (S5) NDVI, spatial information, time information, NOAA satellite number and years since its launch.

4.2.3 Direct Validation of PKU GIMMS NDVI and GIMMS NDVI3g

Our PKU GIMMS NDVI product and the GIMMS NDVI3g product were directly verified with the remaining twenty percent of the Landsat NDVI samples from 1984 to 2015 (Figure 5). Overall, the accuracy of the PKU GIMMS NDVI (R^2 = 0.97, MAE = 0.03, MAPE = 9%) was higher than that of the GIMMS NDVI3g (R^2 = 0.94, MAE = 0.07, MAPE = 20%) in all
305 metrics. Among different vegetation biome types, the NDVI quality of SHR (R^2 = 0.88 for GIMMS NDVI3g and 0.92 for PKU GIMMS NDVI) was higher than that of other biome types. The accuracy of the EBF was relatively low for both products (R^2 = 0.08 and 0.50). The reason was that EBF is primarily distributed in tropical areas where the quality of remote sensing data is poor due to frequent clouds and rains.

For the GIMMS NDVI3g product, the accuracy differences between vegetation biome types were evident (Figure 5).
310 NDVI of SHR (MAE = 0.05), SAV (MAE = 0.08), ENF (MAE = 0.07), and DNF (MAE = 0.08) has been systematically overestimated (Figure 5a). GRA and CRO were also overestimated, mainly when NDVI values were high (Figure 5a). NDVI of the EBF has a rather low accuracy in the GIMMS NDVI3g products, with an R^2 of only 0.08. For the PKU GIMMS NDVI product, its performance in different vegetation biome types was more stable (R^2: 0.50 to 0.92; MAE: 0.05 to 0.10; MAPE: 2% to 15%) and the scatter points remained near the 1:1 line (Figure 5). In particular, the R^2 of the EBF was improved to 0.50
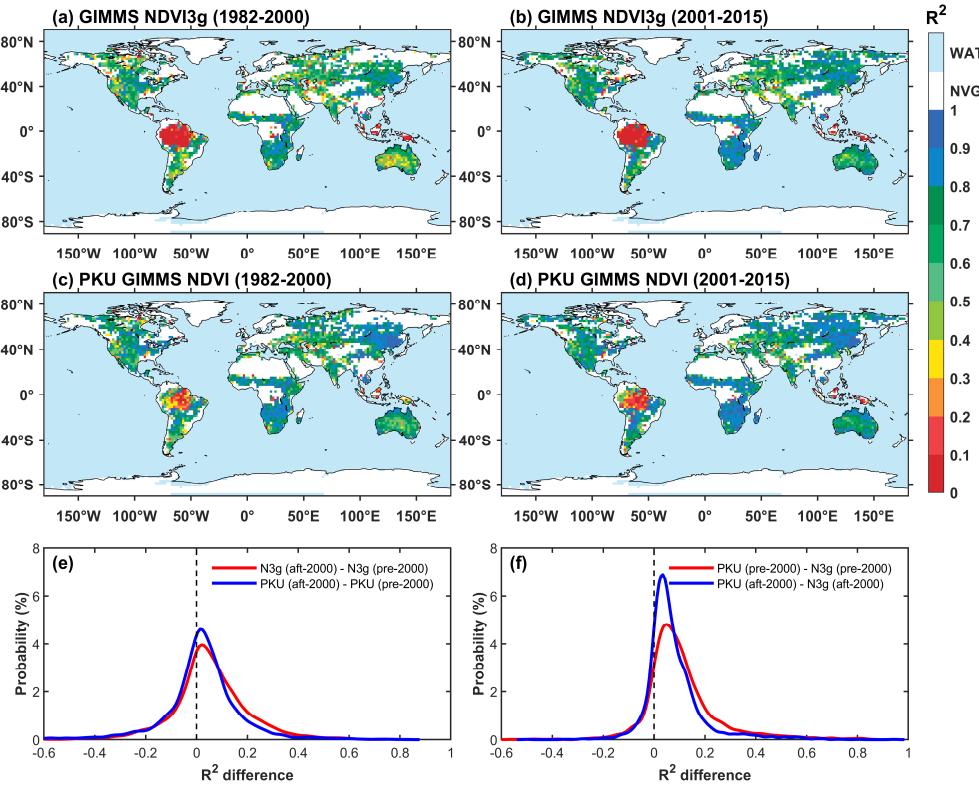315 in the PKU GIMMS NDVI.

Figure 5. Direct validation of the (a) GIMMS NDVI3g and (b) PKU GIMMS NDVI products. Individual Landsat NDVI samples from 1984 to 2015 were used in the validation at a 1/12° resolution. Orange lines represent a 1:1 line. GLO represents the global vegetation biome.

4.2.4 Accuracy of PKU GIMMS NDVI and GIMMS NDVI3g in space

The accuracies of the PKU GIMMS NDVI and GIMMS NDVI3g products exhibited strong spatial heterogeneity (Figure 6). The low-accuracy areas were primarily concentrated in the tropics and high northern latitudes, and the high-accuracy regions were concentrated in the mid-latitudes of the Eurasian continent, the Great Plains of the United States, and savanna-dominated areas of Africa and Australia. In the tropical rainforest area where both products had relatively low accuracies, the PKU GIMMS NDVI better performed especially in Southeast Asia and the northwestern Amazon region. However, the improvement of PKU GIMMS NDVI over GIMMS NDVI3g was not significant along the western coast of the European continent and Southeast China, probably due to the small number of training samples.
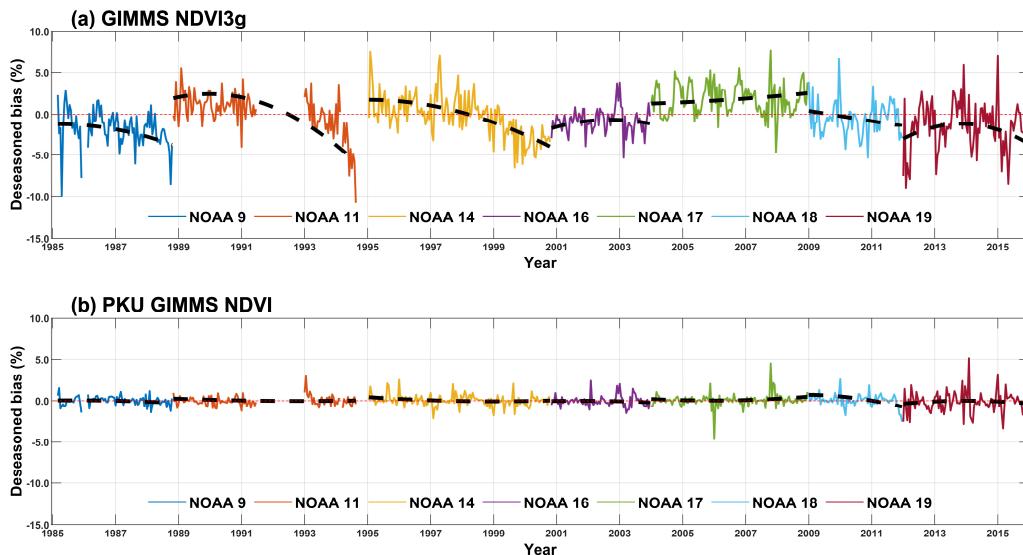
Probability density diagrams were drawn to show NDVI differences between two periods (before and after 2000) (Figure 6e) and between two products (Figure 6f). The accuracy of both NDVI products after 2000 was generally higher than before 2000. The difference was more evident for the GIMMS NDVI3g product (Figure 6e). The PKU GIMMS NDVI improved the accuracies over the GIMMS NDVI3g, especially for the period before 2000 (Figure 6f).

Figure 6. Comparison of R^2 between the GIMMS NDVI3g and PKU GIMMS NDVI products in pre-MODIS (1982–2000) and MODIS (2001–2015) period. (a) to (d) shows the spatial distributions of R^2 in $2° \times 2°$ grids. Non-vegetated grids and grids with less than 20 validation samples are marked in white. (e) and (f) shows the probability distribution of R^2 differences between the two periods (before 2000 and after 2000) and between the two products (GIMMS NDVI3g and PKU GIMMS NDVI), respectively.

4.2.5 Alleviation of the Orbital Drift and Sensor Degradation Effect

As shown in Figure 7, the GIMMS NDVI3g product exhibited evident false signals in the EBF region, which agreed
with the previous findings (Tian et al., 2015). The NDVI bias from different NOAA satellites significantly varied, which may
cause the "jump" phenomenon between NOAA missions. Before 2000, the effect of orbital drift and sensor degradation were
evident at the last phases of satellite launch. This is especially true for the NOAA11 satellite (Figure 7a). The effect became
relatively small for NOAA satellites launched after 2000. In the PKU GIMMS NDVI product, the impact from orbital drift
and sensor degradation has been effectively rectified (Figure 7b). NDVI bias did not change significantly over time, indicating
that the PKU GIMMS NDVI product had good temporal consistency.

Figure 7. Temporal variations of NDVI bias% in EBF for (a) the GIMMS NDVI3g product and (b) the PKU GIMMS NDVI product. The black dash line represents the interannual trend extracted by the EEMD method. Values from different NOAA satellite missions were distinguished with colors.
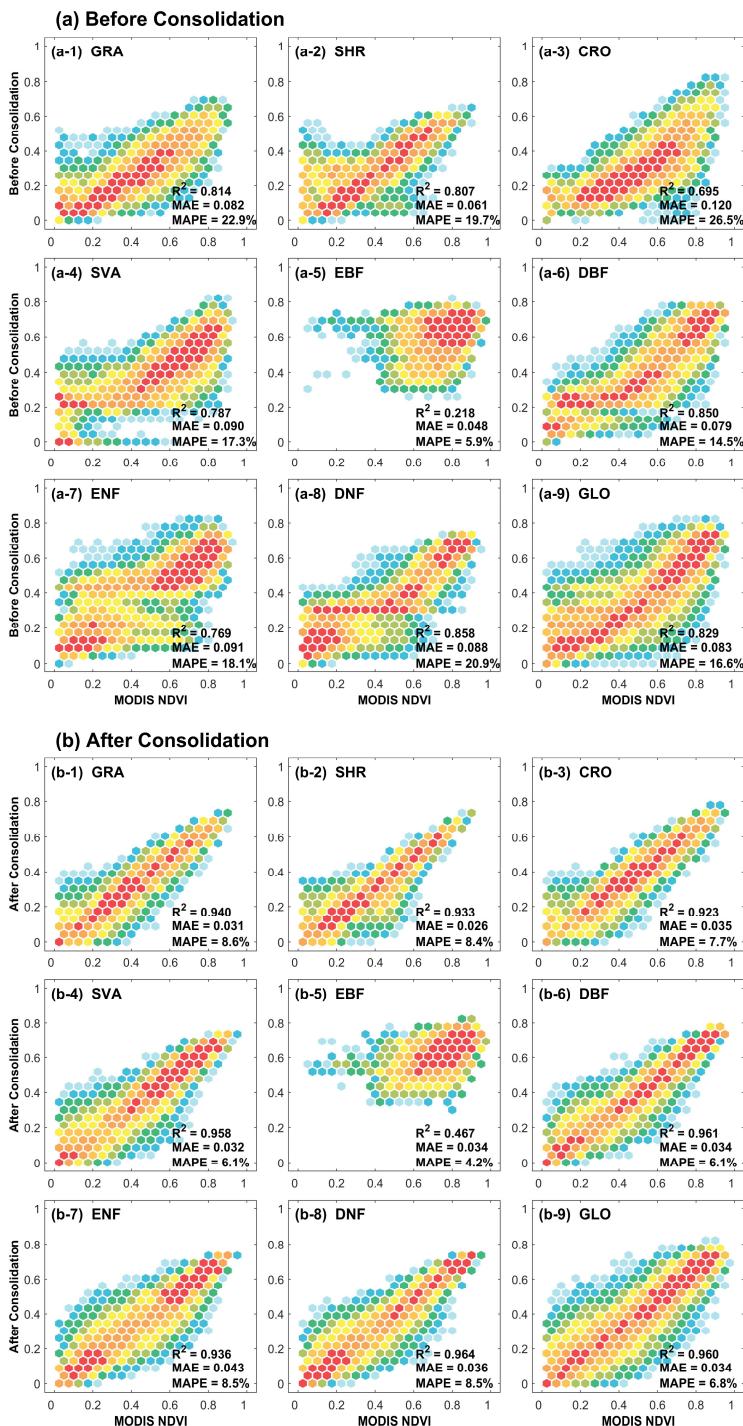
4.3 Consolidated PKU GIMMS NDVI

4.3.1 Direct Validation with MODIS NDVI

The consolidation process improved the consistency level between PKU GIMMS NDVI and MODIS NDVI from acceptable ($R^2 = 0.829$, MAE = 0.083, and MAPE = 16.6%) (Figure 8a-9) to high ($R^2 = 0.962$, MAE = 0.032, and MAPE = 6.5%) (Figure 8b-9). Specifically, the PKU GIMMS NDVI before data consolidation was systematically lower than MODIS NDVI; but the relationship approached 1:1 after consolidation. The improvement in consistency was different among vegetation biome types. CRO, GRA, and ENF had the greatest improvement, as their MAPE decreased from 26.5%, 22.9%, and 20.9% to 7.7%, 8.6%, and 8.5%, respectively (Figure 8a and Figure 8b). The probability distribution densities of R^2, MAPE, and bias were also analysed based on NDVI values befor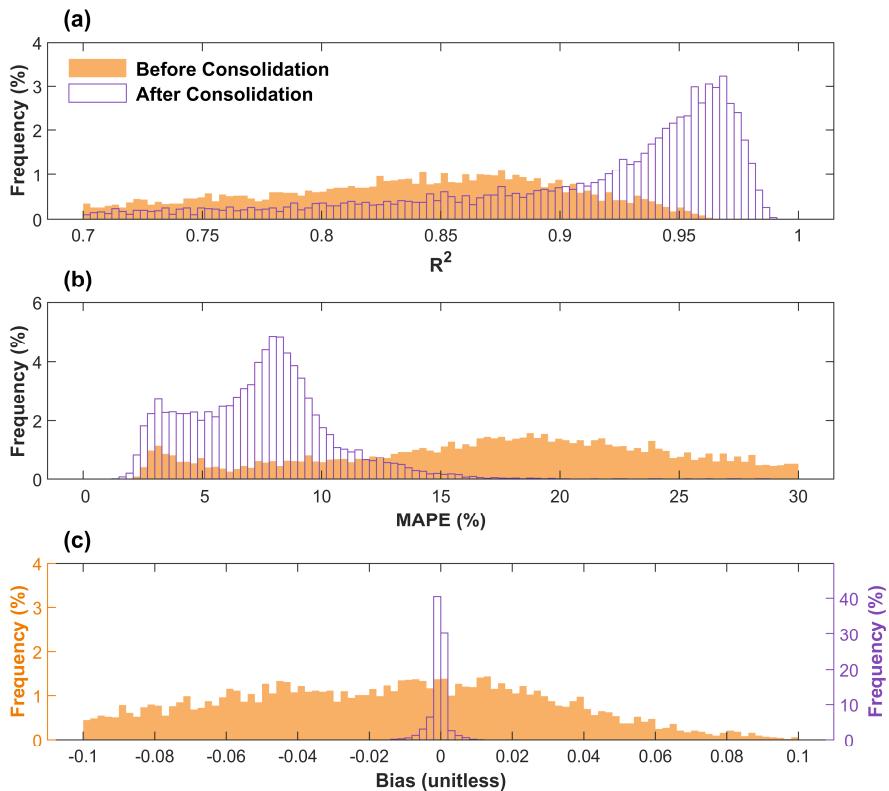e and after consolidation at all samples (8,000) (Figure 9). The results show that the R^2 was slightly improved (Figure 9a), and the MAPE was significantly decreased (Figure 9b) after consolidation.
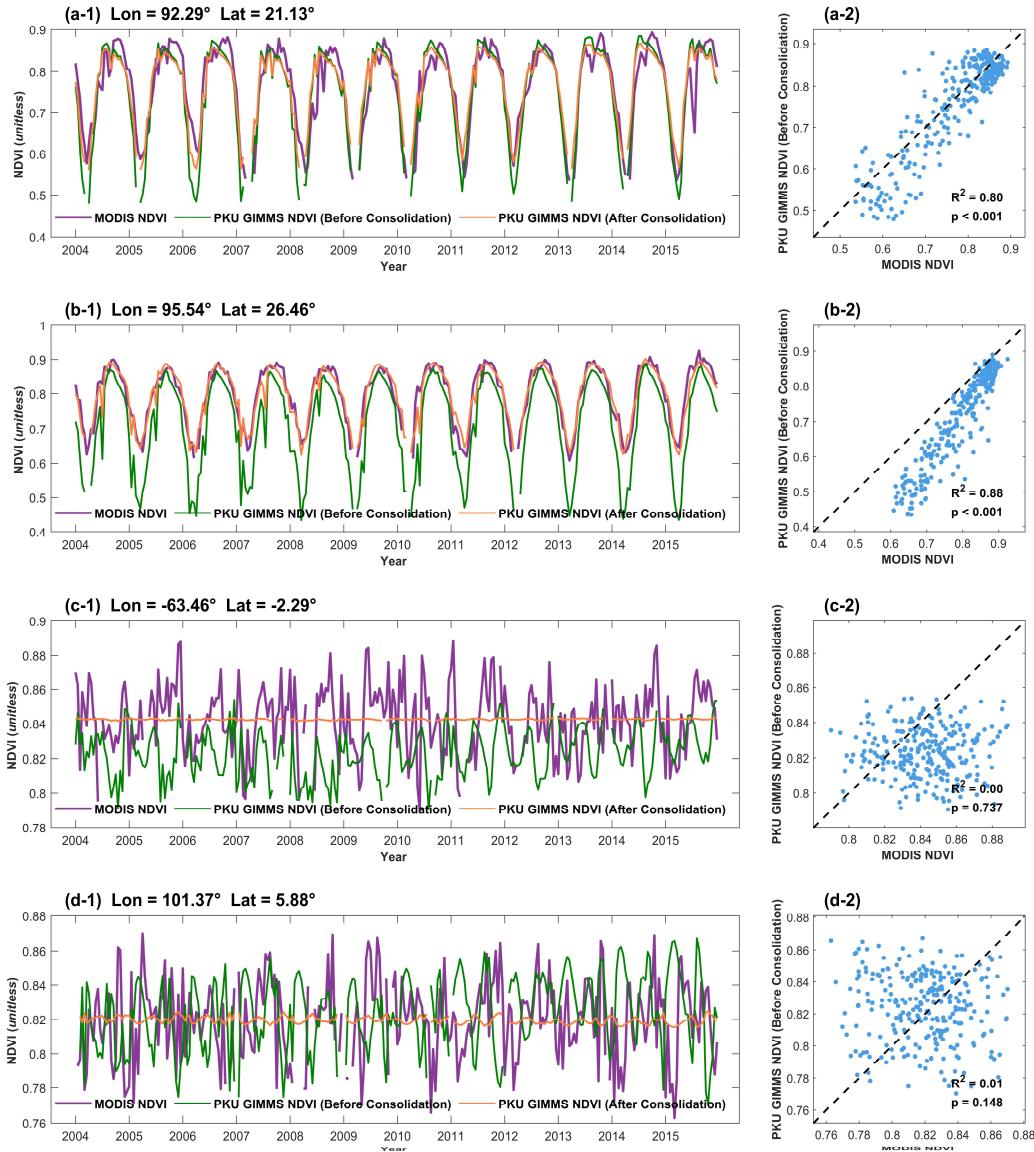
For EBF, the improvement in consistency after consolidation was relatively small (5.9% to 4.2%). The pixel-wised linear regression model for data consolidation was significant for selected EBF locations (Figure 10a and Figure 10b); but it was not for most locations. Due to frequent clouds and rains, both PKU GIMMS NDVI and MODIS NDVI could present maximum noises (Figure 10c and Figure 10d). In this case, it was determined that individual MODIS NDVI values were no longer reliable as benchmark in the regression model and a simple mean-value aligning method was adopted.

Figure 8. Direct validation of the PKU GIMMS NDVI product (a) before and (b) after consolidation. The validation was performed using 1,000 MODIS NDVI samples at a 1/12° resolution for each vegetation biome type from 2004 to 2015.

Figure 9. Distribution of the accuracy metrics before and after consolidation. The accuracy metrics include (a) R^2, (b) MAPE, and (c) bias. The distribution was calculated using 8,000 random sample points from 2004 to 2015 with MODIS NDVI as the true value.
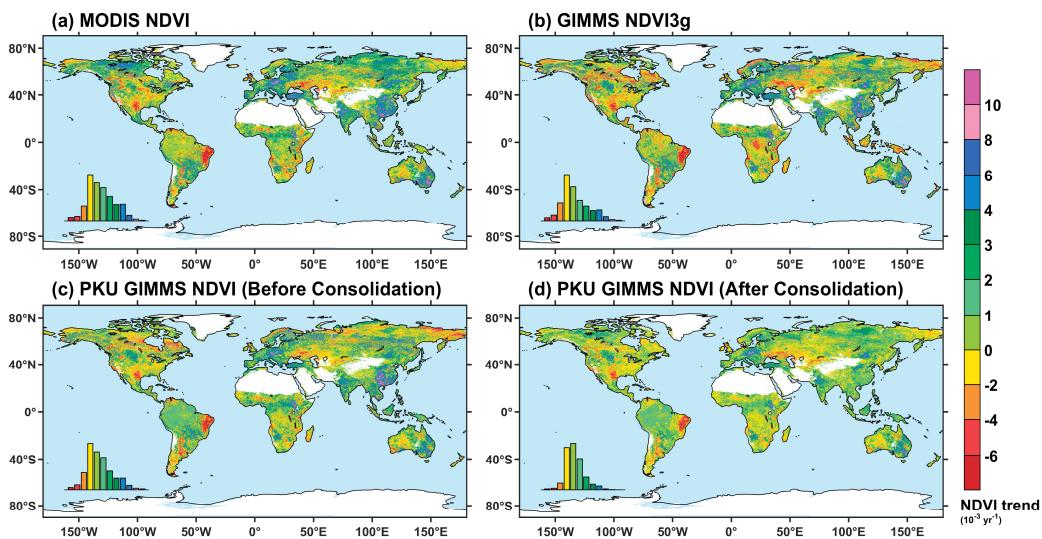
Earth System
Science
Data
Open Access
Discussions

Figure 10. Selected EBF locations illustrating the consistency between the PKU GIMMS NDVI product (before and after consolidation) and the MODIS NDVI product. (a) and (b) showcase the locations with significant pixel-wised linear regression models. (c) and (d) showcase the locations with insignificant models. NDVI values outside two standard deviations were treated as outliers and removed.

4.3.2 Vegetation trend analysis

355   Figure 11 shows the distribution of slopes in the linear regression for the GIMMS NDVI3g, MODIS NDVI, and PKU GIMMS NDVI (before and after consolidation) products at 1/12° grids. Overall, these products showed similar spatial patterns
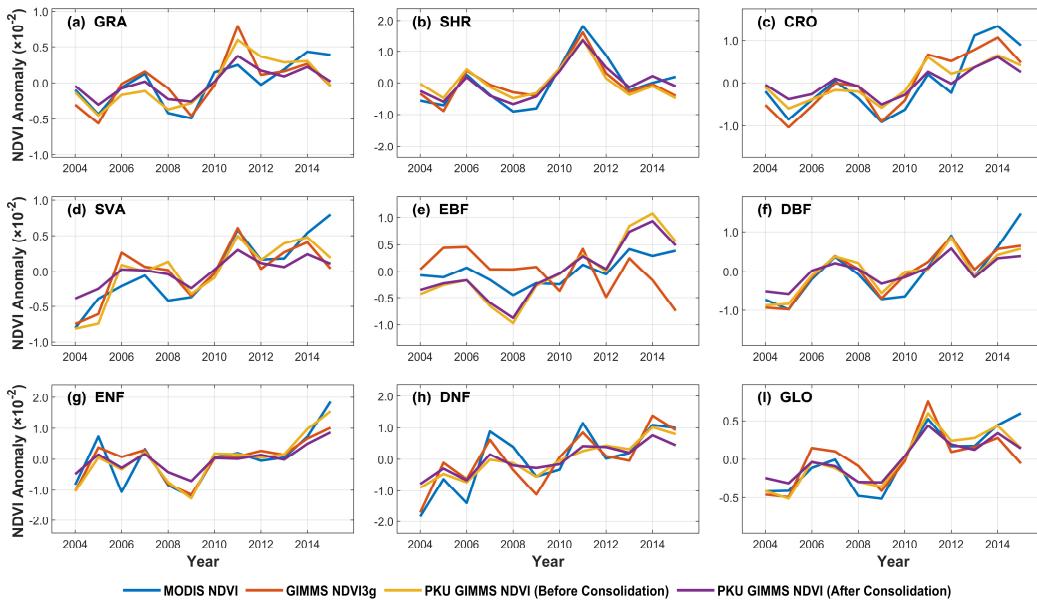
in some global hot spots. For example, all of them capture the greening trend in China and India. However, the GIMMS NDVI3g products showed the opposite trend against MODIS NDVI and PKU GIMMS NDVI products in the tropical evergreen broadleaf forest regions of Africa and Southeast Asia (Figure 11).

Figure 11. Distribution of the NDVI trend from 2004 to 2015. The NDVI trend is calculated based on the time series of annual average NDVI from (a) MODIS NDVI, (b) GIMMS NDVI3g, (c) PKU GIMMS NDVI (before consolidation), and (d) PKU GIMMS NDVI (after consolidation). Non-vegetation area is marked in white.

360

The time series of annual NDVI anomalies that were calculated as an area-weighted average for each vegetation biome type are shown in Figure 12. Overall, PKU GIMMS NDVI (before consolidation) (0.757×10^{-3} yr^{-1}) shows a similar trend to MODIS NDVI (0.883×10^{-3} yr^{-1}) (Figure 12i). The trend of GIMMS NDVI3g was slightly lower (0.502×10^{-3} yr^{-1}). Among vegetation biome types except for EBF, GIMMS NDVI3g, PKU GIMMS NDVI (before consolidation) and MODIS NDVI showed consistent greening or browning trends and interannual fluctuation characteristics (Figure 12). In EBF area, GIMMS NDVI3g showed a browning trend due to the impact of orbital drift and sensor degradation (Figure 12e and Figure 11a), which was consistent with the research by Wang et al. (2022). In PKU GIMMS NDVI products, the effect of orbital drift and sensor degradation has been alleviated. It showed a greening trend in EBF, consistent with MODIS NDVI (Figure 12e).

Figure 12. Biome-specific NDVI anomalies from 2004 to 2015 in MODIS, GIMMS NDVI3g, PKU GIMMS NDVI (before consolidation), and PKU GIMMS NDVI (after consolidation). The NDVI anomalies were calculated as area-weighted annual averages. GLO represents the global vegetation biome.

5. Discussion and Conclusion

5.1 Improvements over other long-term Global NDVI Products

The generation of long-term global NDVI products has been challenging due to the uncertainties associated with the impacts of satellite orbital drift and sensor degradation/calibration, artefacts related to data processing and analysis, and effects from atmosphere, BRDF, scale, topography, etc. (Zeng et al., 2022). While a lot of work remains to do, the current NDVI products have made substantial progress in dealing with selected uncertainties and improving the data quality. Depending on the type of uncertainty, many methods and data have been introduced with different degrees of succession. For example, the GIMMS NDVI3g product has primarily focused on identifying the systematic deviations between AVHRR-2 and AVHRR-3 caused by differences in the sensor characteristics, and on resolving the inconsistency between AVHRR NDVI products (Pinzon and Tucker, 2014).

The significance of the PKU GIMMS NDVI product, was the use of massive representative high-quality Landsat NDVI samples across the globe from 1984 to 2015. More than 12 million samples of different vegetation biome types were extracted to create temporally consistent Landsat data (Figure 2); and approximately 3.6 million Landsat NDVI samples were obtained to generate PKU GIMMS NDVI. The number, temporal coverage, and spatial coverage of the samples (Figure 3) have been unparalleled compared to preceding long-term NDVI products that also employed samples from other sensors, most of which

became available only after late 1990s (e.g., SeaWiFS: 1997−2010; SPOT-VGT: 1999-2014; MODIS: since 2001; VIIRS:

385 since 2012). The massive Landsat NDVI samples paved the way for accurate Landsat sensor cross-calibration (Figure 2) and

PKU GIMMS NDVI generation (Figure 5 and Figure 6). They helped efficiently remove the uncertainties from NOAA orbital

drift and AVHRR sensor degradation (Figure 7).

5.2 Uncertainty Source

Despite our efforts to improve the number and quality of the Landsat NDVI samples, uncertainties remain. Besides

390 those related to the Landsat NDVI itself (such as the geometric and radiometric errors), one major source of uncertainty is the

sample size. For certain vegetation biomes, the number of samples in some regions (such as EBF and northern high latitudes)

may be insufficient due to the constraints in solar zenith angle and environmental conditions. Our results indicated that the

accuracy of BPNN models was lower (despite acceptable) in the regions with fewer samples (Figure 3 and Figure 6). Future

research in these regions should include samples from other satellite data as supplements.

395   Another source of uncertainty is the BNPP model. We developed individual BPNN models for vegetation biome types.

As such, errors in the static vegetation biome map and in the MODIS land cover product, and the heterogeneity of earth surface

might all diminish the model performance. As an input in the BNPP model, GIMMS NDVI3g could also transmit its

uncertainties to the PKU GIMMS NDVI. Furthermore, subsequent research can include other explanatory variables in the

BPNN such as the environmental variables (e.g., solar radiation, temperature, and precipitation) better explain NDVI

400 variations.

5.3 A summary of the PKU GIMMS NDVI product

In this study, a new version of GIMMS NDVI product, i.e., the PKU GIMMS NDVI, was developed using the BPNN

model featured by employing a large number of global high-quality Landsat NDVI samples and a data consolidation method

that employed MODIS NDVI. The PKU GIMMS NDVI covers a time span of 1982 to 2020, with a spatial resolution of 1/12°

405 and a temporal resolution of half-month. The high reliability and high accuracy of the PKU GIMMS NDVI product is

demonstrated by:

- The Landsat NDVI samples used to generate the PKU GIMMS NDVI were abundant (3.6 million), well distributed, representative, and high-quality. The inter-comparison between NDVI of different Landsat sensors after calibration showed high R^2 (> 0.98), low MAE (< 0.02) and low MAPE (< 6%).

410
- Assessing against the Landsat NDVI, the PKU GIMMS NDVI had high overall accuracies (R^2 = 0.97, MAE = 0.03, MAPE = 9%) and high accuracies for individual vegetation biomes (R^2: 0.50 to 0.92; MAE: 0.05 to 0.10; MAPE: 2% to 15%). It better performed than the GIMMS NDVI3g (overall: R^2 = 0.94, MAE = 0.07, MAPE = 20%), especially for the tropical evergreen broadleaf forests (EBF) (R^2 = 0.08 in the GIMMS NDVI3g; R^2 = 0.50 in the PKU GIMMS NDVI).

415

Earth System
Science
Data
Open Access
Discussions

- The PKU GIMMS NDVI efficiently removed the effects of NOAA orbital drift and AVHRR sensor degradation.
- During the overlapping period, the consolidated PKU GIMMS NDVI has a good agreement with MODIS NDVI in values (R^2 = 0.962, MAE = 0.032, MAPE = 6.5%) and in vegetation trend (MODIS NDVI: 0.883×10^{-3} yr^{-1}; PKU GIMMS NDVI: 0.757×10^{-3} yr^{-1}).

420

The long-term, continuous, and reliable PKU GIMMS NDVI for the past 40 years can provide a more accurate vegetation monitoring in the context of climate change. The framework proposed in this study which used high-quality Landsat samples with BPNN and other explanatory variables (the longitude and latitude, associated month, and the NOAA number and years since launch) can also benefit the development of other remote sensing products of land surface parameters (e.g., the

425 development of LAI products).

Data availability

The spatiotemporally consistent global dataset of the GIMMS Normalized Difference Vegetation Index (PKU GIMMS NDVI) generated in this study is openly available at https://doi.org/10.5281/zenodo.7441559 (Li et al., 2022). It covers the whole global vegetation area at half month temporal resolution and 1/12° spatial resolution from 1982 to 2020. It is available

430 in Geographic Lat/Lon projection and TIFF format.

Author contributions. ZZ conceptualized and supervised the project. ZZ, ML, and SC designed the workflow of methodology to product the dataset. ML led most of the work in data acquisition and processing. ML, ZZ, and SC performed data analysis. ML, SC, and ZZ prepared the manuscript. SC and ZZ reviewed and edited the draft. All authors contributed to the interpretation

435 of the results.

Competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this study.

440 **Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Earth System
Open Access Science Discussions
Data

Reference

AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., and Hain, C. R.: Remote
455   sensing of drought: Progress, challenges and opportunities, Rev geophys, 53, 452-480, https://doi.org/10.1002/2014rg000456, 2015.

Badgley, G., Field, C. B., and Berry, J. A.: Canopy near-infrared reflectance and terrestrial photosynthesis, Sci Adv, 3, e1602244, https://doi.org/10.1126/sciadv.1602244, 2017.

Berner, L. T., Massey, R., Jantz, P., Forbes, B. C., Macias-Fauria, M., Myers-Smith, I., Kumpula, T., Gauthier, G., Andreu-
460   Hayles, L., Gaglioti, B. V., Burns, P., Zetterberg, P., D'Arrigo, R., and Goetz, S. J.: Summer warming explains widespread but not uniform greening in the Arctic tundra biome, Nat Commun, 11, 1-12, https://doi.org/10.1038/s41467-020-18479-5, 2020.

Beurs, K. M. d. and Henebry, G. M.: Trend analysis of the Pathfinder AVHRR Land (PAL) NDVI data for the deserts of central Asia, IEEE Geosci Remote S, 1, 282-286, https://doi.org/10.1109/LGRS.2004.834805, 2004.

465 Cao, B., Yu, L., Naipal, V., Ciais, P., Li, W., Zhao, Y., Wei, W., Chen, D., Liu, Z., and Gong, P.: A 30 m terrace mapping in China using Landsat 8 imagery and digital elevation model based on the Google Earth Engine, Earth syst sci data, 13, 2437-2456, https://doi.org/10.5194/essd-13-2437-2021, 2021.

Cao, C., Weinreb, M., and Xu, H.: Predicting Simultaneous Nadir Overpasses among Polar-Orbiting Meteorological Satellites for the Intersatellite Calibration of Radiometers, J Atmos Ocean Tech, 21, 537-542, https://doi.org/10.1175/1520-
470   0426(2004)021<0537:PSNOAP>2.0.CO;2, 2004.

Cao, C., De Luccia, F. J., Xiong, X., Wolfe, R., and Weng, F.: Early On-Orbit Performance of the Visible Infrared Imaging Radiometer Suite Onboard the Suomi National Polar-Orbiting Partnership (S-NPP) Satellite, IEEE T Geosci Remote, 52, 1142-1156, https://doi.org/10.1109/tgrs.2013.2247768, 2014.

Chen, C., Park, T., Wang, X. H., Piao, S. L., Xu, B. D., Chaturvedi, R. K., Fuchs, R., Brovkin, V., Ciais, P., Fensholt, R.,
475   Tommervik, H., Bala, G., Zhu, Z. C., Nemani, R. R., and Myneni, R. B.: China and India lead in greening of the world through land-use management, Nat sustain, 2, 122-129, https://doi.org/10.1038/s41893-019-0220-7, 2019.

Cui, Y. K., Jia, L., and Fan, W. J.: Estimation of actual evapotranspiration and its components in an irrigated area by integrating the Shuttleworth-Wallace and surface temperature-vegetation index schemes using the particle swarm optimization algorithm, Agr forest meteorol, 307, 108488, https://doi.org/10.1016/j.agrformet.2021.108488, 2021.

480 Didan, K.: MODIS/Terra Vegetation Indices Monthly L3 Global 0.05Deg CMG V061 (V061), NASA EOSDIS Land Processes DAAC [dataset], https://doi.org/10.5067/MODIS/MOD13C2.061, 2021.

Doelling, D. R., Garber, D. P., Avey, L. A., Nguyen, L., and Minnis, P.: The calibration of AVHRR visible dual gain using Meteosat-8 for NOAA-16 to 18, Conference on Atmospheric and Enviromental Remote Sensing Data Processing and Utilization III: Readiness for GEOSS, San Diego, CA, Aug 17-30, WOS:000251483900008, 61-71, 10.1117/12.736080,
485   2007.

Dong, J., Fu, Y., Wang, J., Tian, H., Fu, S., Niu, Z., Han, W., Zheng, Y., Huang, J., and Yuan, W.: Early-season mapping of winter wheat in China based on Landsat and Sentinel images, Earth syst sci data, 12, 3081-3095, https://doi.org/10.5194/essd-12-3081-2020, 2020.

Fan, X. and Liu, Y.: A global study of NDVI difference among moderate-resolution satellite sensors, ISPRS J Photogramm,
490   121, 177-191, https://doi.org/10.1016/j.isprsjprs.2016.09.008, 2016.

Fensholt, R. and Proud, S. R.: Evaluation of Earth Observation based global long term vegetation trends - Comparing GIMMS and MODIS global NDVI time series, Remote Sens Environ, 119, 131-147, https://doi.org/10.1016/j.rse.2011.12.015, 2012.

Foga, S., Scaramuzza, P. L., Guo, S., Zhu, Z., Dilley, R. D., Beckmann, T., Schmidt, G. L., Dwyer, J. L., Hughes, M. J., and
495   Laue, B.: Cloud detection algorithm comparison and validation for operational Landsat data products, Remote Sens Environ, 194, 379-390, https://doi.org/10.1016/j.rse.2017.03.026, 2017.

Frankenberg, C., Yin, Y., Byrne, B., He, L. Y., and Gentine, P.: Comment on "Recent global decline of CO_2 fertilization effects on vegetation photosynthesis" COMMENT, Science, 373, eabg2947, https://doi.org/10.1126/science.abg2947, 2021.

500   Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.: Global land cover mapping from MODIS: algorithms and early results, Remote Sens Environ, 83, 287-302, https://doi.org/10.1016/s0034-4257(02)00078-0, 2002.

Gamon, J. A., Huemmrich, K. F., Wong, C. Y. S., Ensminger, I., Garrity, S., Hollinger, D. Y., Noormets, A., and Penuelas, J.: A remotely sensed pigment index reveals photosynthetic phenology in evergreen conifers, P Natl Acad Sci USA, 113,
505   13087-13092, https://doi.org/10.1073/pnas.1606162113, 2016.

Gao, X., Huete, A. R., Ni, W. G., and Miura, T.: Optical-biophysical relationships of vegetation spectra without background contamination, Remote Sens Environ, 74, 609-620, https://doi.org/10.1016/s0034-4257(00)00150-4, 2000.

Helder, D., Thome, K. J., Mishra, N., Chander, G., Xiong, X. X., Angal, A., and Choi, T.: Absolute Radiometric Calibration of Landsat Using a Pseudo Invariant Calibration Site, IEEE T Geosci Remote, 51, 1360-1369,
510   https://doi.org/10.1109/tgrs.2013.2243738, 2013.

Hong, X.-C., Wang, G.-Y., Liu, J., Song, L., and Wu, E. T. Y.: Modeling the impact of soundscape drivers on perceived birdsongs in urban forests, J clean prod, 292, 125315, https://doi.org/10.1016/j.jclepro.2020.125315, 2021.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. L. C., Shih, H. H., Zheng, Q. N., Yen, N. C., Tung, C. C., and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, P Roy Soc
515   a-math phy, 454, 903-995, https://doi.org/10.1098/rspa.1998.0193, 1998.

Irons, J. R., Dwyer, J. L., and Barsi, J. A.: The next Landsat satellite: The Landsat Data Continuity Mission, Remote Sens Environ, 122, 11-21, https://doi.org/10.1016/j.rse.2011.08.026, 2012.

Jiang, C. Y., Ryu, Y., Fang, H. L., Myneni, R., Claverie, M., and Zhu, Z. C.: Inconsistencies of interannual variability and trends in long-term satellite leaf area index products, Global Change Biol, 23, 4133-4146,
520   https://doi.org/10.1111/gcb.13787, 2017.

Jiang, L., Tarpley, J. D., Mitchell, K. E., Zhou, S., Kogan, F. N., and Guo, W.: Adjusting for long-term anomalous trends in NOAA's global vegetation index data sets, IEEE T Geosci Remote, 46, 409-422, https://doi.org/10.1109/tgrs.2007.902844, 2008.

Joiner, J., Yoshida, Y., Zhang, Y., Duveiller, G., Jung, M., Lyapustin, A., Wang, Y. J., and Tucker, C. J.: Estimation of
525   Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data, Remote Sens, 10, 1346, https://doi.org/10.3390/rs10091346, 2018.

Justice, C., Belward, A., Morisette, J., Lewis, P., Privette, J., and Baret, F.: Developments in the 'validation' of satellite sensor products for the study of the land surface, Int J Remote Sens, 21, 3383-3390, https://doi.org/10.1080/014311600750020000, 2000.

530   Kogan, F. N.: Application of vegetation index and brightness temperature for drought detection, in: Natural Hazards: Monitoring and Assessment Using Remote Sensing Technique, edited by: Singh, R. P., and Furrer, R., Advances in Space Research-Series, 11, 91-100, 10.1016/0273-1177(95)00079-t, 1995.

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 (V1.0), zenodo [dataset], https://doi.org/10.5281/zenodo.7441559, 2022.

535   Li, X., Zhou, Y., Meng, L., Asrar, G. R., Lu, C., and Wu, Q.: A dataset of 30 m annual vegetation phenology indicators (1985-2015) in urban areas of the conterminous United States, Earth syst sci data, 11, 881-894, https://doi.org/10.5194/essd-11-881-2019, 2019.

Los, S. O.: Estimation of the ratio of sensor degradation between NOAA AVHRR channels 1 and 2 from monthly NDVI composites, IEEE T Geosci Remote, 36, 206-213, https://doi.org/10.1109/36.655330, 1998.

540   Maisongrande, P., Duchemin, B., and Dedieu, G.: VEGETATION/SPOT: an operational mission for the Earth monitoring; presentation of new standard products, Int J Remote Sens, 25, 9-14, https://doi.org/10.1080/0143116031000115265, 2004.

Mao, D., Wang, Z., Luo, L., and Ren, C.: Integrating AVHRR and MODIS data to monitor NDVI changes and their relationships with climatic parameters in Northeast China, Int j appl earth obs, 18, 528-536, https://doi.org/10.1016/j.jag.2011.10.007, 2012.

545   Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., Gao, F., Kutler, J., and Lim, T. K.: A Landsat surface reflectance dataset for North America, 1990-2000, IEEE Geosci Remote S, 3, 68-72, https://doi.org/10.1109/lgrs.2005.857030, 2006.

Meng, X., Bao, Y., Liu, J., Liu, H., Zhang, X., Zhang, Y., Wang, P., Tang, H., and Kong, F.: Regional soil organic carbon prediction model based on a discrete wavelet analysis of hyperspectral satellite data, Int j appl earth obs, 89, 102111,
550   https://doi.org/10.1016/j.jag.2020.102111, 2020.

Myers-Smith, I. H., Kerby, J. T., Phoenix, G. K., Bjerke, J. W., Epstein, H. E., Assmann, J. J., John, C., Andreu-Hayles, L., Angers-Blondin, S., Beck, P. S. A., Berner, L. T., Bhatt, U. S., Bjorkman, A. D., Blok, D., Bryn, A., Christiansen, C. T., Cornelissen, J. H. C., Cunliffe, A. M., Elmendorf, S. C., Forbes, B. C., Goetz, S. J., Hollister, R. D., de Jong, R., Loranty, M. M., Macias-Fauria, M., Maseyk, K., Normand, S., Olofsson, J., Parker, T. C., Parmentier, F. J. W., Post, E.,
555   Schaepman-Strub, G., Stordal, F., Sullivan, P. F., Thomas, H. J. D., Tommervik, H., Treharne, R., Tweedie, C. E., Walker, D. A., Wilmking, M., and Wipf, S.: Complexity revealed in the greening of the Arctic, Nat clim change, 10, 106-117, https://doi.org/10.1038/s41558-019-0688-1, 2020.

Pedelty, J., Devadiga, S., Masuoka, E., Brown, M., Pinzon, J., Tucker, C., Roy, D., Ju, J. C., Vermote, E., Prince, S., Nagol, J., Justice, C., Schaaf, C., Liu, J. C., Privette, J., Pinheiro, A., and Ieee: Generating a Long-term Land Data Record from
560   the AVHRR and MODIS instruments, IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Barcelona, SPAIN, Jul 23-27, WOS:000256657301039, 1021-1024, 10.1109/igarss.2007.4422974, 2007.

Peng, J., Dadson, S., Hirpa, F., Dyer, E., Lees, T., Miralles, D. G., Vicente-Serrano, S. M., and Funk, C.: A pan-African high-resolution drought index dataset, Earth syst sci data, 12, 753-769, https://doi.org/10.5194/essd-12-753-2020, 2020.

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., Bjerke, J. W., Chen, A., Ciais, P., Tommervik, H., Nemani, R. R., and
565   Myneni, R. B.: Characteristics, drivers and feedbacks of global greening, Nat rev earth env, 1, 14-27, https://doi.org/10.1038/s43017-019-0001-x, 2020.

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.

Qin, Y., Xiao, X., Wigneron, J.-P., Ciais, P., Brandt, M., Fan, L., Li, X., Crowell, S., Wu, X., Doughty, R., Zhang, Y., Liu, F.,
570   Sitch, S., and Moore, B.: Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon, Nat clim change, 11, 442-448, https://doi.org/10.1038/s41558-021-01026-5, 2021.

Rondeaux, G., Steven, M., and Baret, F.: Optimization of soil-adjusted vegetation indices, Remote Sens Environ, 55, 95-107, https://doi.org/10.1016/0034-4257(95)00186-7, 1996.

Roy, D. P., Kovalskyy, V., Zhang, H. K., Vermote, E. F., Yan, L., Kumar, S. S., and Egorov, A.: Characterization of Landsat-
575   7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity, Remote Sens Environ, 185, 57-70, https://doi.org/10.1016/j.rse.2015.12.024, 2016.

Schubert, P., Lagergren, F., Aurela, M., Christensen, T., Grelle, A., Heliasz, M., Klemedtsson, L., Lindroth, A., Pilegaard, K., Vesala, T., and Eklundh, L.: Modeling GPP in the Nordic forest landscape with MODIS time series data-Comparison with the MODIS GPP product, Remote Sens Environ, 126, 136-147, https://doi.org/10.1016/j.rse.2012.08.005, 2012.

580 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.

Trishchenko, A. P., Cihlar, J., and Li, Z.: Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors, Remote Sens Environ, 81, 1-18, https://doi.org/10.1016/S0034-4257(01)00328-5,
585   2002.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., Vermote, E. F., and El Saleous, N.: An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, Int J Remote Sens, 26, 4485-4498, https://doi.org/10.1080/01431160500168686, 2005.

Vermote, E., Justice, C., Claverie, M., and Franch, B.: Preliminary analysis of the performance of the Landsat 8/OLI land
590   surface reflectance product, Remote Sens Environ, 185, 46-56, https://doi.org/10.1016/j.rse.2016.04.008, 2016.

Wang, S. H., Zhang, Y. G., Ju, W. M., Chen, J. M., Cescatti, A., Sardans, J., Janssens, I. A., Wu, M. S., Berry, J. A., Campbell, J. E., Fernandez-Martinez, M., Alkama, R., Sitch, S., Smith, W. K., Yuan, W. P., He, W., Lombardozzi, D., Kautz, M., Zhu, D., Lienert, S., Kato, E., Poulter, B., Sanders, T. G. M., Kruger, I., Wang, R., Zeng, N., Tian, H. Q., Vuichard, N., Jain, A. K., Wiltshire, A., Goll, D. S., and Penuelas, J.: Response to Comments on "Recent global decline of CO_2
595   fertilization effects on vegetation photosynthesis" COMMENT, Science, 373, eabg7484, https://doi.org/10.1126/science.abg7484, 2021.

Wang, Z., Wang, H., Wang, T., Wang, L., Liu, X., Zheng, K., and Huang, X.: Large discrepancies of global greening: Indication of multi-source remote sensing data, Glob ecol conserv, 34, e02016, https://doi.org/10.1016/j.gecco.2022.e02016, 2022.

600   Weng, Q., Fu, P., and Gao, F.: Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data, Remote Sens Environ, 145, 55-67, https://doi.org/10.1016/j.rse.2014.02.003, 2014.

Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., Fosnight, E. A., Shaw, J., Masek, J. G., and Roy, D. P.: The global Landsat archive: Status, consolidation, and direction, Remote Sens Environ, 185, 271-283, https://doi.org/10.1016/j.rse.2015.11.032, 2016.

605   Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G., Anderson, M. C., Belward, A. S., Cohen, W. B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosillo, T., Hipple, J. D., Hostert, P., Hughes, M. J., Huntington, J., Johnson, D. M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T. A., Schaaf, C., Schott, J. R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J. C., Wynne, R. H., and Zhu, Z.: Current status of Landsat program, science, and applications, Remote Sens Environ, 225, 127-147, 610   https://doi.org/10.1016/j.rse.2019.02.015, 2019.

Yang, S., Feng, Q., Liang, T., Liu, B., Zhang, W., and Xie, H.: Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region, Remote Sens Environ, 204, 448-455, https://doi.org/10.1016/j.rse.2017.10.011, 2018.

Yang, W., Kogan, F., Guo, W., and Chen, Y.: A novel re-compositing approach to create continuous and consistent cross-615   sensor/cross-production global NDVI datasets, Int J Remote Sens, 42, 6025-6049, https://doi.org/10.1080/01431161.2021.1934597, 2021.

Yin, G., Verger, A., Descals, A., Filella, I., and Peñuelas, J.: A Broadband Green-Red Vegetation Index for Monitoring Gross Primary Production Phenology, Journal of Remote Sensing, 2022, 9764982, https://doi.org/10.34133/2022/9764982, 2022.

620   Zeng, Y. L., Hao, D. L., Huete, A., Dechant, B., Berry, J., Chen, J. M., Joiner, J., Frankenberg, C., Bond-Lamberty, B., Ryu, Y., Xiao, J. F., Asrar, G. R., and Chen, M.: Optical vegetation indices for monitoring terrestrial ecosystems globally, Nat rev earth env, 3, 477-493, https://doi.org/10.1038/s43017-022-00298-5, 2022.

Zhang, X., Liu, L., Chen, X., Gao, Y., and Jiang, M.: Automatically Monitoring Impervious Surfaces Using Spectral Generalization and Time Series Landsat Imagery from 1985 to 2020 in the Yangtze River Delta, Journal of Remote 625   Sensing, 2021, 9873816, https://doi.org/10.34133/2021/9873816, 2021.

Zhang, X., Xu, M., Wang, S., Huang, Y., and Xie, Z.: Mapping photovoltaic power plants in China using Landsat, random forest, and Google Earth Engine, Earth syst sci data, 14, 3743-3755, https://doi.org/10.5194/essd-14-3743-2022, 2022.

Zhang, Y., Song, C., Band, L. E., Sun, G., and Li, J.: Reanalysis of global terrestrial vegetation trends from MODIS products: Browning or greening?, Remote Sens Environ, 191, 145-155, https://doi.org/10.1016/j.rse.2016.12.018, 2017.

630   Zhu, Z., Wang, S. X., and Woodcock, C. E.: Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images, Remote Sens Environ, 159, 269-277, https://doi.org/10.1016/j.rse.2014.12.014, 2015.

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 635   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.

Zhu, Z. C., Zeng, H., Myneni, R. B., Chen, C., Zhao, Q., Zha, J. J., Zhan, S. M., and MacLachlan, I.: Comment on "Recent global decline of CO_2 fertilization effects on vegetation photosynthesis" COMMENT, Science, 373, eabg5673, https://doi.org/10.1126/science.abg5673, 2021.