 doi:10.20944/preprints202205.0367.v1

Article

A New Spatial Downscaling Algorithm to Generate A Long-Term Global 250-m Resolution NDVI Product

Zhimin Ma[1], Chunyu Dong[1,2*], Kairong Lin[1], Yu Yan[1], Jianfeng Luo[1], Dingshen Jiang[1] and Xiaohong Chen[1*]

[1] School of Civil Engineering, Sun Yat-sen University, and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519082, China
[2] Professorship in Hydrology and Climatology, Institute of Geography, Heidelberg University, Heidelberg 69120, Germany
*Corresponding author: dongchy7@mail.sysu.edu.cn (C.D.); eescxh@mail.sysu.edu.cn (X.C.)

Abstract: Satellite-based Normalized Difference Vegetation Index (NDVI) time-series data are useful for monitoring the changes of vegetation ecosystems in the context of global climate change. However, there are currently no ideal NDVI datasets that reconcile long-term series with high spatial resolution. Here, we have developed a simple and novel data downscaling algorithm based on the coefficient of variation (CV) statistics, which combines the detailed spatial features of MODIS data with the long-term temporal information of AVHRR data. The proposed data fusion method helps generate a global monthly NDVI database that has a 250 m-resolution and covers the long period of 1982–2018. We evaluated the accuracy of the fused data against MODIS NDVI using the metrics of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Pearson's correlation coefficients (R). Validation suggests a high performance of the downscaling algorithm and a high accuracy of the new NDVI database. We further applied the downscaled data to monitor NDVI changes of various vegetation types and in areas having high vegetation heterogeneity, and we obtained stable results similar to MODIS data. The whole data downscaling and validation processes were completed on the Google Earth Engine platform, and here we provide a code for users to easily get the data for any part of the world. The downscaled global-scale NDVI time series has high potential in monitoring the temporal and spatial dynamics of the terrestrial ecosystems under changing environments.

Keywords: downscaling; NDVI; MODIS; AVHRR; GEE

1. Introduction

Global climate change has had a significant impact on the biological, physical, and chemical processes of terrestrial ecosystems around the world over the last half-century[1,2]. Terrestrial vegetation, as a key component of the global ecosystem, is particularly vulnerable to climate change[3]. Satellite remote sensing has become an important tool for measuring and monitoring the dynamics of large-scale terrestrial ecosystems due to its broad coverage, high temporal and spatial precision, and consistency[4,5]. Satellite spectral vegetation indices (VI) are numerical markers of vegetation growth and biomass that might help explain how plants evolve[6,7]. The photosynthetic potential (or greenness) of a vegetation canopy is represented by the satellite-derived Normalized Difference Vegetation Index (NDVI) and higher NDVI scores usually indicate that the vegetation is more vigorous. In the context of severe global climate change[8,9]. It is critical to employ vegetation indices to research the fluctuations and interactions between vegetation and the entire globe using long-term satellite-observed data.

People began to launch a series of satellite sensors, such as the Landsat series, Advanced Very High-Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS), and others, in the 1970s and 2000s to record NDVI with spatial resolution ranging from 30 m to 8 km[10,11]. However, almost every product has certain limitations. Almost every product, however, has its limitations. The Terra and Aqua satellites

 doi:10.20944/preprints202205.0367.v1

orbit the earth every day, providing higher spatial resolution images, enhanced atmospheric corrections, and a more precise geo-registration NDVI dataset for MODIS products, but they only begin in 2000, making it impossible to study three or four decades of vegetation variations[12,13]. The AVHRR sensors onboard the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellite series can provide a constant long-term dataset with the start date from 1981. However, the spatial resolution of the AVHRR sensors is coarse and cannot capture the fine-scale features required for monitoring land cover and ecosystem changes across heterogeneous areas[14].

Therefore, people have started to focus on how to develop methods to combine the high spatial-resolution and long time-span NDVI data with composite multi-source satellite data, and these methods can be divided into five categories: approaches based on weight functions, unmixing, Bayesian, learning, and hybrid ways[15]. In the early period when the number of high-resolution satellites was relatively small, the researchers generally used the linear mixture model to downscale the NDVI product[16]. After the advent of high-precision satellites in the 2000s, people began to combine data from many sensors with varying resolutions to create products with high spatial and temporal resolution[17]. Some researchers proposed a spatially and temporally adaptive reflection fusion model (STARFM) to blend Landsat and MODIS data to predict daily surface reflectance at Landsat spatial resolution and MODIS temporal resolution, and the Enhanced STARFM based on the STARFM (ESTARFM)[18,19]. The Spatial-Temporal Adaptive Algorithm for Mapping Reflectance Change (STAARCH) employed an algorithm capable of identifying the area of change and the precise moment, as well as improving fusion accuracy by selecting data at the best time[20]. Recently, machine learning and deep learning approaches have also been used for NDVI data downscaling. For example, some researchers used an artificial neural network (ANN) to map the NDVI indices from AVHRR to match and extend MODIS NDVI data at 1° and other researchers propose a deep learning spatiotemporal data fusion approach based on Very Deep Super-Resolution (VDSR) to fuse the NDVI retrievals from Sentinel-2 and Landsat 8 images[10,21]. However, a superior global NDVI database that possesses both high spatial resolution and a long period has not been developed thus far, which hinders the detailed investigation of vegetation response to global climate change.

To achieve the goal of the long-term and higher-resolution vegetation variations observations, we performed VIs fusion and pay more attention to the MODIS NDVI and AVHRR NDVI. MODIS NDVI product is considered to be the perfection of the AVHRR NDVI database, which improves the spatial resolution and chlorophyll sensitivity, eliminates the interference of atmospheric water vapor, reduces the radiometric calibration geometric distortion, and adjusts the synthesis method [22,23]. Many scholars have investigated the distinction between these two items. They discovered that the trends of AVHRR NDVI data were in good agreement with the trends of MODIS NDVI data overall, and that both time-series data can effectively describe the growth dynamics of alpine grassland. Although MODIS performed better than AVHRR, the discrepancies between them are slight[24-26]. Therefore, MODIS NDVI is consistent in time and space with AVHRR NDVI, allowing for the mixing of MODIS high-resolution data with AVHRR low-resolution data to provide continuous, long-term, and high-resolution NDVI datasets.

The GEE cloud platform is supported by Google's Cloud Infrastructure, a platform developed by Google for online visual computing, analysis, and processing of a large number of global-scale earth science data (especially satellite image data)[27,28]. The platform provides a highly straightforward approach for users all over the world to either use GEE's shared enormous quantity of data or uploads their own private data to the cloud platform, where it would be preprocessed[29,30]. Preprocessed data will be stored in the public data area and will be available for use at any time. Researchers can quickly broaden the spatial and temporal scales of their study without the restriction of downloading data. We chose GEE as our study platform since we aim to generate a dataset that spans over 40 years and downloading such large amounts of data would be difficult.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

Here we use the Google Earth Engine (GEE) platform to execute the downscaling process, and thus all data can be easily downloaded from the GEE platform. It is a perfect way of using the GEE platform to process long-term and large-scale data, and readers may easily use the code and generate their own data. The major goal of this research is to develop a GEE-based flexible, operable, and efficient NDVI downscaling approach. Two objectives are expected to be achieved:

1) Create a global-scale 250 m-resolution downscaled NDVI product and test its accuracy; 2) Discuss the product's reliability and uncertainty at both local and global scales.

2. Materials and Methods

2.1. Materials

2.1.1. NDVI products

Firstly, we used the MODIS level 3 NDVI product (MOD13Q1 Version 6) from NASA's Terra polar-orbiting sun-synchronous satellite (10:30 AM local time) with a revisit of 16 days and a spatial resolution of 250 meters from 2001 to 2018. MODIS NDVI products are calculated based on atmospherically corrected bidirectional surface reflectance and hidden by water, clouds, heavy aerosols, and cloud shadows [22].

The NOAA Climate Data Record (CDR) of AVHRR NDVI contains gridded daily NDVI at a resolution of 0.05° (about 5 km) derived from the NOAA AVHRR Surface Reflectance product from 1981 to the present [31]. In Version 5, incorrect data in the time, latitude, and longitude variables have been corrected. Due to the sensor degrading beginning in late 2018, striped images and the missing images have occurred in the southern hemisphere. Thus, we used the AVAHRR NDVI from 1982 to 2018 as the whole downscaled dataset period.

2.1.2. Landcover products

The auxiliary input data were the landcover maps from Copernicus, the MODIS Fire_cci Burned Area pixel products, and the GAP land cover classification map. The Copernicus Landcover maps are provided for the period 2015-2019 over the entire Globe with a resolution of 100 meters and reached an accuracy of 80% at Level1 overall years. The fire product of version 5.1 is a monthly global ~250 m spatial resolution dataset containing information on the burned area as well as ancillary data. The GAP landcover classification map in 2011 provided detailed vegetation types including the Conterminous U.S., Alaska, Hawaii, and Puerto Rico. We have filtered the fire, farmland, and building areas from the FireCCI51 map and the Copernicus landcover map before the downscaling steps, which can help extract the areas of significant land cover change to avoid having large uncertainty results.

All of these datasets are available on the Google Earth Engine platform, and can be derived from: https://developers.google.com/earth-engine/datasets.

2.2. Downscaling method

2.2.1. Data pre-processing

Before the data downscaling, we collected all the burned areas from 2001 to 2018, as well as the farmland and building areas from 2015 to 2019. We assume these areas have abrupt NDVI changes and are not suitable for the proposed data downscaling algorithm. Hence, we combined these pixels as a masking layer and screened out all these specific areas from the NDVI dataset. Then, for each reserved pixel, we composited the MODIS and AVHRR NDVI products as monthly time series using the Maximum Value Composite (MVC) technique, i.e., extracting the pixel-wise maximum NDVI value among all the available NDVI observations within each month. Then we resampled the AVHRR data to 250 m resolution using the bicubic spatial interpolation method. The MVC approach is easy and often works well because most NDVI observation errors are negative due to the obstructing effects of atmospheric layers and clouds[32,33].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

2.2.2. Downscaling algorithm

Since remote sensing data contain both temporal and spatial information, we downscaled the NDVI data at these two scales[34]. Previous studies suggest that, in an annual phenological cycle, combining the spatial characteristics of a single fine resolution image and the temporal characteristics (phenology) of a time series of coarse resolution images can yield NDVI images with high spatial and temporal resolution[35,36]. In the light of the above progress, here we seek for the spatial change information in the fine-resolution MODIS NDVI and the temporal change information in the long-term AVHRR NDVI in the downscaling process.

(1) Changes at the spatial scale

First, we needed to know how these two NDVI products differed in terms of feature spatial variations. To illustrate the diverse changing information in space, we used the monthly coefficient of variation (CV) in each pixel. The AVHRR products were divided into two portions, one matched to MODIS data from 2001 to 2018, and the other from 1982 to 2000. MODIS CV has a higher value than AVHRR CV in general because MODIS CV has more detailed spatial changing information. For example, while downscaling the coarse-resolution AVHRR data, the NDVI change degree should be increased for pixels containing a variety of vegetation with clear seasonal NDVI dynamics. However, pixels with a high proportion of bare soil or buildings only exhibit minimal NDVI variations over time, resulting in compressed NDVI variability. Hence, we corrected the dimensional difference between MODIS and AVHRR products at each pixel using the $R_{CVx,y,m}$ ratio.

We expressed the change degree of NDVI variability for the various features between the fine-resolution MODIS data and the coarse-resolution AVHRR data using the proportional parameter $R_{CVx,y,m}$ from 2001 to 2018 (Eq. 1). In addition, we used the proportional parameter $R_{CVx,y,n}$ between the pre-2001 AVHRR CV (1982–2000) and the post-2000 AVHRR CV (2001–2018) to represent the long-term AVHRR NDVI spatial changes with dramatic changes in landcover between the two periods (Eq. 2).

$$R_{CVx,y,m} = MODIS\_CV / AVHRR\_CV_{post} \quad (1)$$

$$R_{CVx,y,n} = AVHRR\_CV_{pre} / AVHRR\_CV_{post} \quad (2)$$

where $MODIS\_CV$ is the pixelwise monthly coefficient of variation (CV) of the MODIS NDVI (2001–2018), and $AVHRR\_CV_{pre}$ and $AVHRR\_CV_{post}$ are the pixelwise monthly CV of the AVHRR NDVI for 1982–2000 and 2001–2018, respectively.

(2) Changes at the temporal scale

The per-pixel monthly NDVI median from 2001 to 2018 was calculated as the baseline value in MODIS and AVHRR NDVI products. Then, for each month of the time series, we compared each monthly AVHRR NDVI value to the baseline median of that month to determine the change degree. To define the temporal change, we employ the parameter $K_{x,y,t}$, which can capture the temporal information in the coarse resolution AVHRR NDVI data at the time scale (Eq. 3).

$$K_{x,y,t} = (NDVI_{L,x,y,t} - NDVI_{L,x,y,bl}) / NDVI_{L,x,y,bl} \quad (3)$$

where $NDVI_{L,x,y,t}$ is the pixelwise monthly AVHRR NDVI in the entire time series, and $NDVI_{L,x,y,bl}$ is the baseline median NDVI in different months during the period of 2001–2018.

(3) Bringing together data on both spatial and temporal changes

The final phase of downscaling combined the AVHRR product's long-term temporal information with the MODIS product's fine-scale spatial information (Eq. 4, 5). The corresponding monthly MODIS NDVI median was the fine-scale background NDVI at each pixel. We multiplied the high-resolution baseline NDVI with the modifying parameters

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

of $K_{x,y,t}$, and $R_{CVx,y,\mathrm{m}}$, and then added them to the baseline NDVI to reflect relative changes in both temporal and spatial scales from 2001 to 2018. For periods before 2000 where MODSI data was not available, we multiplied by an additional factor $R_{CVx,y,\mathrm{n}}$ to account for the temporal variations of AVHRR data. Finally, we can get the downscaled data using the formulas:

$$NDVI_{H,x,y,t} = NDVI_{H,x,y,bl} \times (1 + K_{x,y,t} \times R_{CVx,y,m}) + \varepsilon_{x,y,t} \quad (4)$$

$$NDVI_{H,x,y,t} = NDVI_{H,x,y,bl} \times (1 + K_{x,y,t} \times R_{CVx,y,m} \times R_{CVx,y,n}) + \varepsilon_{x,y,t} \quad (5)$$

where $NDVI_{H,x,y,t}$ is the downscaled high-resolution NDVI, $NDVI_{H,x,y,bl}$ is the monthly median MODIS NDVI (baseline NDVI), and $\varepsilon_{x,y,t}$ is the random error generated in the downscaling process.

2.2.3. Error validation

(1) GEE Implementation in a regional area

To begin, we wanted to test the method in a regional setting, thus we chose the Idaho state, in the northwest United States, as a case study location to demonstrate the NDVI downscaling technique due to its diversified geography and range of vegetation types. (Fig. 1**Error! Reference source not found.**). We utilize the topographic diversity (*D*) as a metric to describe the range of temperature and moisture conditions that species can encounter in their local habitats (Fig. 1b) [37]. The *D* value of most places is approaching 1.0, indicating that Idaho has a wide range of topo-climate environments and a great diversity of plants. Idaho's landscape can be divided into three regions: 1) the northern narrow strip and the mountainous parts, which are rich in timber; 2) the Snake River Plain, which runs through the state and is the state's main agricultural sector; and 3) the southern mountainous region.

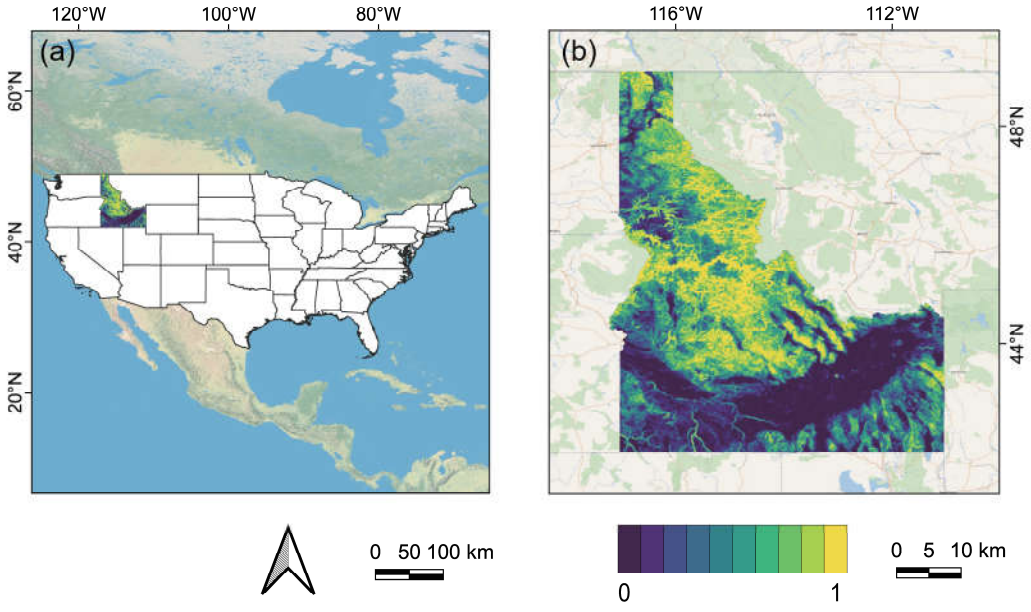


Figure 1. (a) Base map of the contiguous United States with state borders in black and a red polygon to indicate the location of the Idaho state. (b) A topographic diversity map of the Idaho state from the Global SRTM Topographic Diversity dataset.

(2) Evaluation Indices

We used three standard indices to validate the error between MODIS NDVI data and the downscaled NDVI data, and we used a timeframe of around 60 months from 2014 to 2018 to avoid data overfitting[38,39]. The evaluation indices are root mean square error

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

(RMSE), mean absolute error (MAE), and Pearson's correlation coefficient (R) and were calculated at the pixel level. RMSE is a commonly used measure of the difference between values and represents the sample standard deviation of the difference between the predicted and observed values. As a result, it offers a comprehensive evaluation of recalculation, including data retrieval accuracy and precision. We provided the RMSE map and calculated the mean for all pixels. MAE indicates the mean of the absolute error between the predicted and observed values and the value is determined directly for residuals. We calculated the MAE in each pixel and made the map. The Pearson's correlation coefficient R is a frequently used metric for determining the degree of linear association between two variables. As a result, the accuracy at the spatial scale was estimated using the Pearson's correlation coefficient between downscaled NDVI and MODIS NDVI at the pixel level.

(4) Validation of accuracy for various vegetation types

To evaluate the overall applicability of the downscaling technique at the global scale, we chose 8 typical vegetation patches of about 100 km² in size from the world vegetation map, and randomly selected one thousand pixels' NDVI values in both the MODIS and the downscaled NDVI images for the validation period I to create linear fit lines. The vegetation types were derived from the Copernicus Landcover maps.

We assume high-resolution images can often provide more accurate spatial information, and we examined the capabilities of the three NDVI data sets to investigate NDVI variations in various vegetation types. We retrieved the median NDVI values for three types of vegetation in each month from all pixels in their respective ranges between July 2014 and June 2015, which indicated the average levels of greenness for each vegetation type. Then we analyzed the NDVI time series to see if these three sources of data could accurately differentiate the three vegetation types.

The purpose of this study was to generate monthly downscaled NDVI from 1982 to 2018. The main strategy was to create a statistical temporal relationship from the low spatial-resolution AVHRR NDVI and then input high-spatial-resolution data into the statistical relationships to get the downscaled NDVI. These three main formulas were created to develop the link between MODIS NDVI and AVHRR NDVI in step 3. To summarize, the downscaling approach has four major steps as shown in **Error! Reference source not found.**2, which were described as follows:

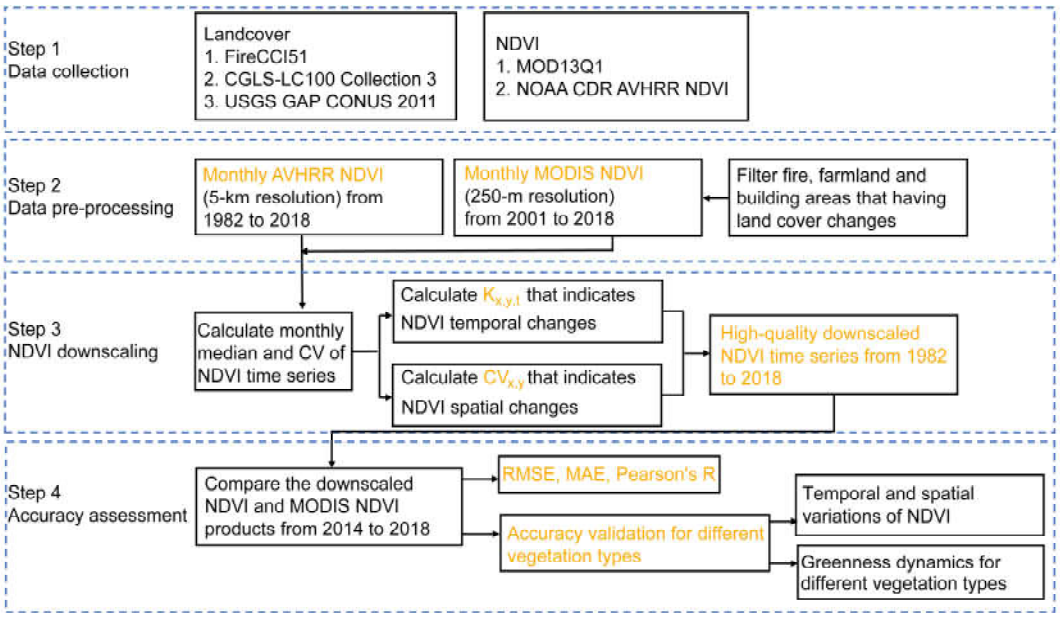


Figure 2. Processes of the downscaling approach.

 doi:10.20944/preprints202205.0367.v1

1. During the data pre-processing step, we first collected the fire, agricultural, and building regions as the mask, and then calculated the monthly MODIS NDVI, and AVHRR NDVI data, which were then resampled to 250 m resolution to be consistent with MODIS data.
2. We completed the entire downscaling process on the GEE platform, and we used the proposed empirical formulas to capture the statistical relationship between AVHRR NDVI and MODIS NDVI and merged the temporal and spatial information in them.
3. A rigorous accuracy assessment was conducted to validate the performance of the downscaling algorithm for different vegetation types and at both the regional and global scales.

3. Results

3.1. Validation at the regional scale

There were some examples of visual comparisons of the downscaled and the standard MODIS NDVI, and we presented areas encompassing the entire state of Idaho as well as randomly selected smaller locations in Fig.3. Since the growing season offers more information about vegetation changes, we chose three downscaled NDVI images for the growing season (i.e., May, July, and September) in 2016 to compare to the MODIS NDVI product. First, we carried out a comparison examination of the entire region, indicating the general difference between the two products was modest and the NDVI value distribution ranges were consistent with some subtle deviations in the spatial distribution of NDVI in some spots. Then, when we examined the dynamic changes in the vegetation growing season from May to September, we found that the greenness differences were noticeable from spring to autumn.

Then, we chose three small portions (P1−P3 in Fig. 3) around 30 km^2 in the vegetation-rich area to assess the accuracy of the downscaled outputs. Please note that the areas in patch 1 and 2 (P1, P2) were mainly dominated by evergreen needle leaves, and the area in patch 3 (P3) was mainly grass. Some slight discrepancies between the MODIS NDVI and the downscaled NDVI could be seen in these three small vegetation patches. However, the seasonal greenness dynamics between the two products were quite similar over time, with all of the inaccuracies being minor at the three vegetation patches (Fig. 3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

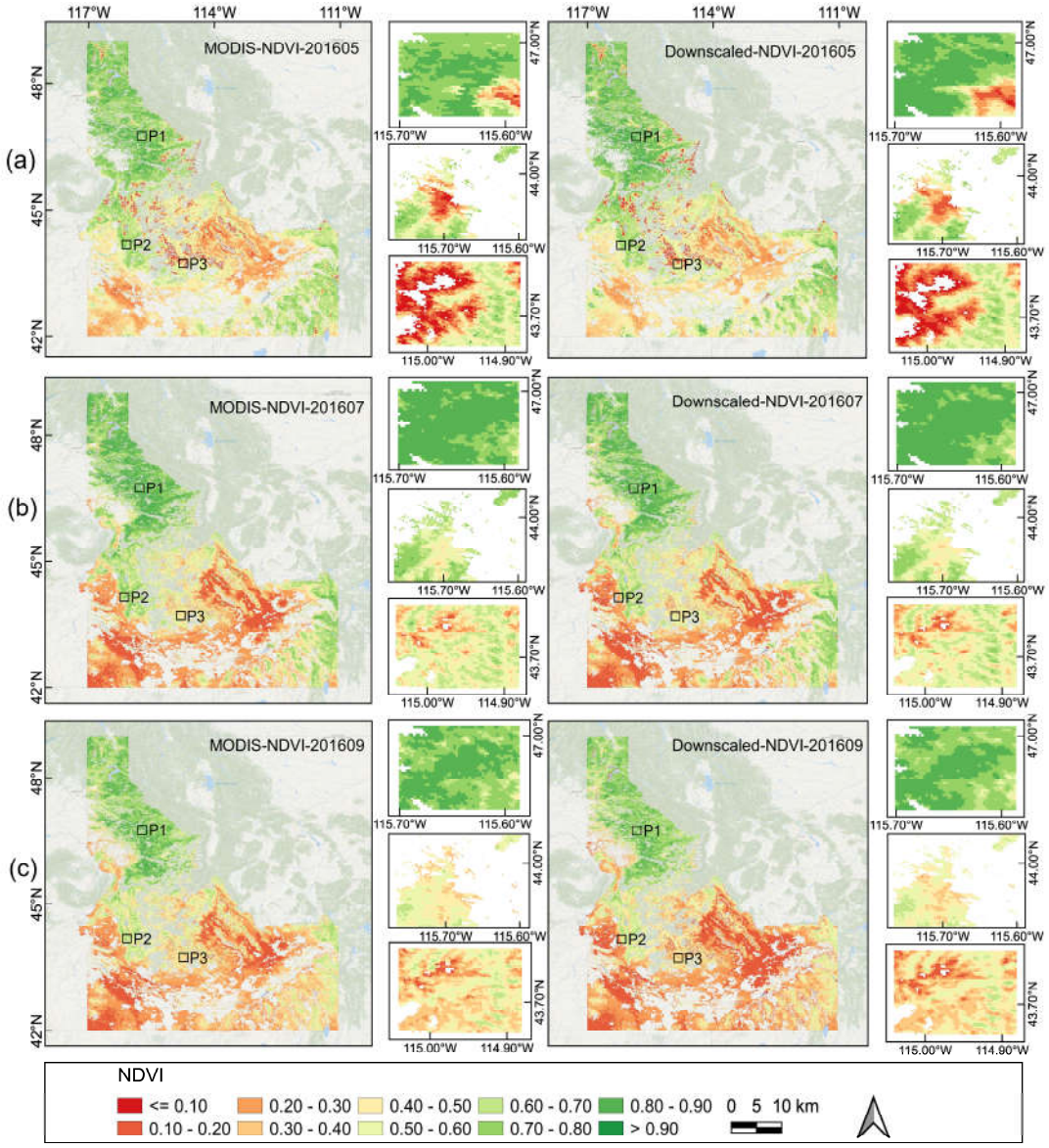


Figure 3. Comparisons of NDVI between the downscaled and MODIS datasets for the Idaho state and at three selected vegetation patch areas in May (a), July (b), and September (c) in 2016.

The means of the three error indices in Idaho were then determined, and the results are displayed below in Table 1. We obtained the mean value from all pixels in each location; in particular, for the calculation of Pearson's R^2 in Idaho, we randomly picked 1000 points about 1000 times to obtain the mean due to the large number of pixels in the entire state. The overall MAE, RMSE, and R^2 are 0.039, 0.055, and 0.86 respectively, all showing acceptable accuracy of the downscaled NDVI products. The biggest MAE and RMSE are less than 0.1, while the lowest R^2 is greater than 0.5, indicating that the downscaling algorithm worked effectively in Idaho.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

Table 1. Error statistics for the comparison between downscaled NDVI and MODIS NDVI in Idaho. The NDVI images are shown in Fig. 3.

Time/Location	MAE	RMSE	R^2
201605 Idaho	0.047	0.069	0.923
201607 Idaho	0.033	0.052	0.949
201609 Idaho	0.052	0.076	0.858
201605-P1	0.062	0.097	0.759
201607-P1	0.015	0.020	0.866
201609-P1	0.016	0.023	0.882
201605-P2	0.061	0.077	0.856
201607-P2	0.035	0.045	0.897
201609-P2	0.034	0.046	0.676
201605-P3	0.053	0.071	0.903
201607-P3	0.030	0.041	0.895
201609-P3	0.034	0.048	0.860
Mean	0.039	0.055	0.860

At each pixel of Idaho, we tested all the three error indices of the downscaled NDVI dataset against the simultaneous MODIS NDVI, which reflected the errors produced by the downscaling procedure [10,40]. The validation maps of the RMSE, MAE, Pearson-r and RGB composite maps of these three indicators are shown in Figure 4. The RMSEs and MAEs are low, mostly ranging between 0 to 0.2 (Fig. 4a, b), and the correlation r values are high, ranging mostly from 0.7 to 1 (Fig. 4c), implying that the downscaled NDVI errors are minor. The blue pixels in the RGB composite error map indicate that the downscaling technique is highly accurate (i.e., low RMSE, MAE, and high Pearson-r), whereas the yellow pixels show areas with proportionally more errors (i.e., higher RMSE, MAE, and lower Pearson-r), and the black pixels indicate that the values of the three indices are all small, indicating that the original data is lost. (Fig. 5d)[41]. The majority of Idaho are in blue and a few black pixels refer to filtered crops. Besides, some yellow pixels are mainly in the north, where the landscape is more undulating with a lot of forests, and therefore these places have relatively more uncertainty in the NDVI downscaling.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

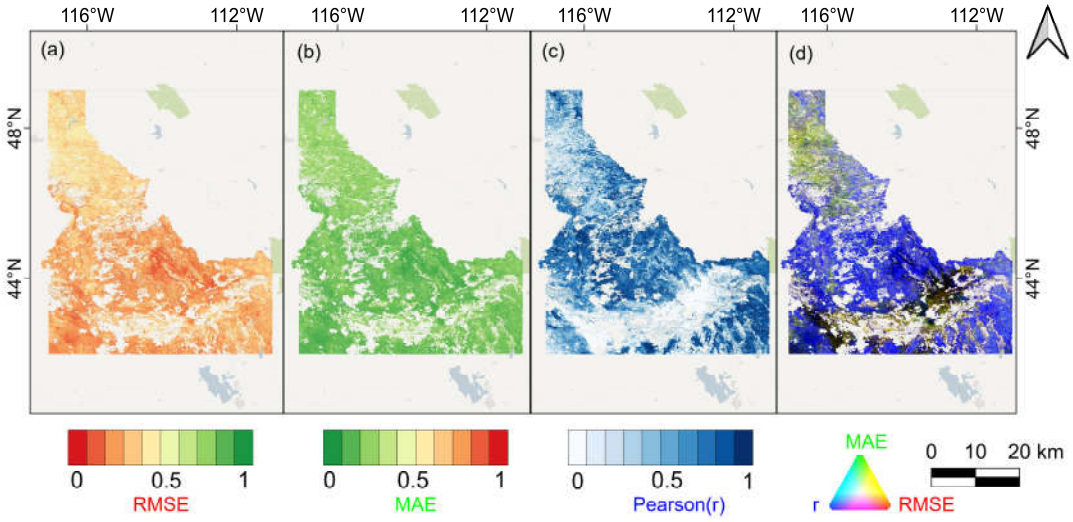


Figure 4. Spatial variation of the goodness of fit indices. (a) RMSE, (b) MAE, (c) Pearson (r), and (d) RGB composite image of the three metrics (Red: RMSE, Green: MAE, and Blue: Pearson-r). Blue colors in the panel (d) refer to areas having high accuracy with low RMSE and MAE, and high Pearson (r), yellow colors suggest decreased accuracy with higher RMSE and MAE and lower Pearson (r), and black colors indicate areas having few vegetation covers and showing low RMSE, MAE, and Pearson (r) [41].

3.2. Validation at the global scale

In addition to the regional-scale study, we also performed the NDVI data downscaling on a global scale using GEE. Then we further expanded the validation to the entire globe and evaluated the uncertainties of the data fusion algorithm for every pixel over the world (Fig. 5). Validation suggests that the majority of RMSE and MAE values fall between 0 and 0.1, according to the first observation (Fig. 5a & b). Second, the majority of Pearson-r values are reasonably high and greater than 0.6, except for the tropics around the equator, arid deserts, wastelands in Asia and Africa, and scant vegetation wilderness in Australia (Fig. 5c).

The accuracy analysis was the same as for the Idaho error index maps; however, the RGB map in Fig. 5d, which combined the three error indices, provided more evident error information. As explained above, the different colors in the RGB composite error map reflect different levels of precision, so we focused on the black and yellow pixels.

First, we noticed that the black areas were typically found in sparsely vegetated areas, such as the central and eastern Asian desert regions and the Australian inland desert belts. Second, we saw that yellow pixels were concentrated in the tropics, indicating bigger inaccuracies, which may be explained by the following two reasons. The first is related to a shift in land cover over the studied period. Several studies have shown that many forests in Southeast Asia and South America are shrinking and being reclaimed as farmland since 2000, and frequent fires also cause forest loss, which implies that spatial changes are highly abrupt, and the sorts of ground objects change significantly [42,43]. Another reason is that the vegetation covers are dense in these areas, so NDVI is easily saturated and cannot distinguish more seasonal changes in vegetation greenness. Indeed, NDVI time series have been revealed to have strong data noise in the tropics, which could be caused by sub-pixel cloud contamination and a failure in the cloud filtering process [44,45]. Furthermore, some places of high altitudes and latitudes have larger RMSE and MAE and lower Pearson-r (yellow in Fig. 5d), likely owing to a high frequency of cloud and snow, which impedes satellite observations [41].

Overall, all the three goodness of fit indicators (RMSE, MAE, and Pearson-r) suggest that the downscaling algorithm performs well (blue in Fig. 5d) in most parts of the world. As a result, using the proposed NDVI product for ecological applications at various scales

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

should be safe. When running the code in areas with prolonged cloudiness and snow, as well as areas with dense rainforest or sparse vegetation, researchers should pay more attention to the uncertainties and errors of the approach, and strict validation is needed before applying the downscaled NDVI products.

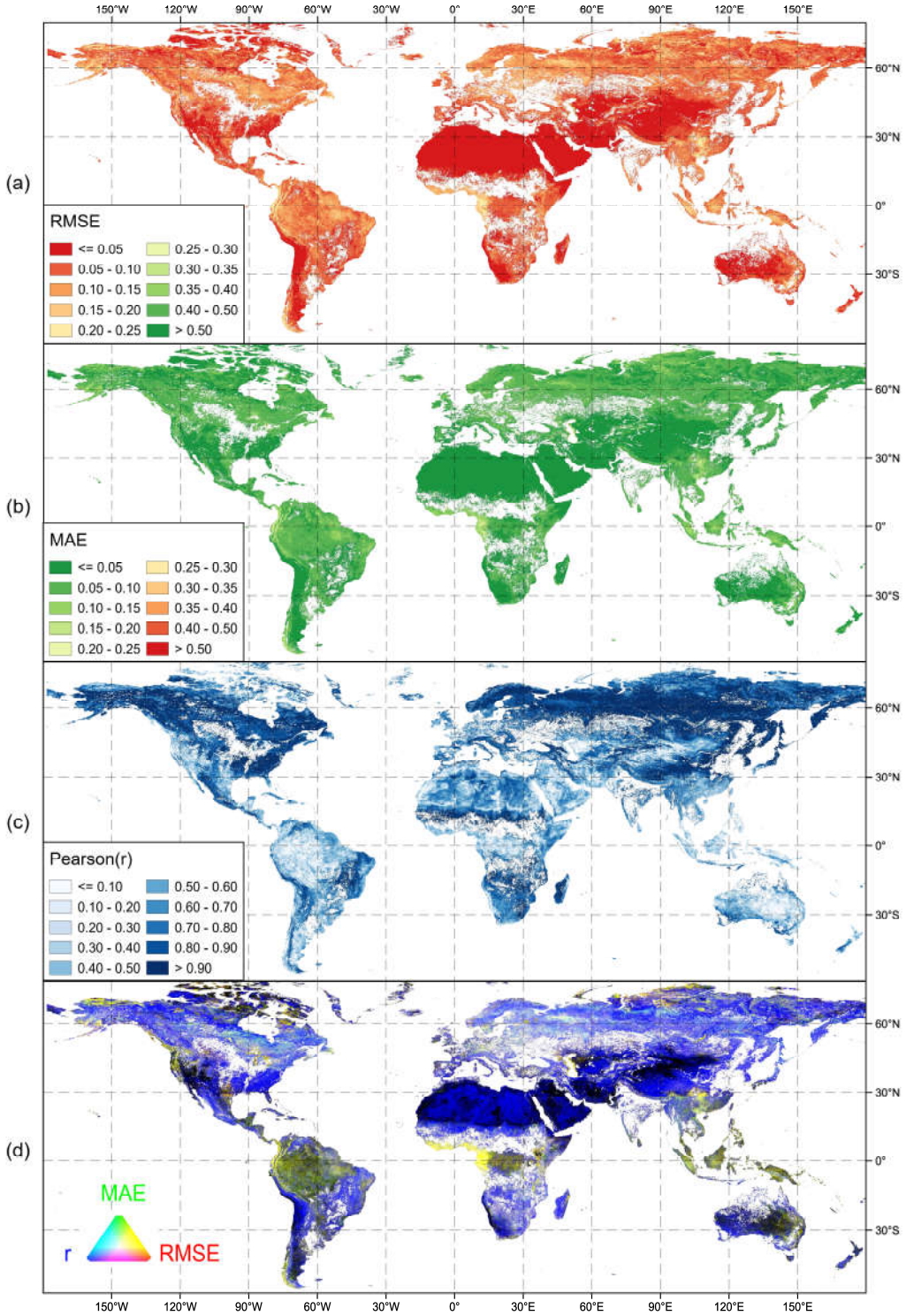


Figure 5. Same as Fig. 4, but for the whole world.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

3.3. Validation for different vegetation types

3.3.1. Comparison of downscaled and MODIS NDVI datasets for global major vegetation types

We used the Copernicus Global Land Cover dataset to obtain a global vegetation type map and then selected different representative vegetation types across continents to compare MODIS NDVI and the downscaled NDVI. The validation was carried out at 8 verification areas including many vegetation types with an area of approximately 100 km^2 around the world (Fig. 6). The main vegetation types were coniferous forest, broad-leaved forest, grassland, and tropical shrub.

We investigated the accuracy of the downscaled NDVI in applying to different vegetation types in two ways: (1) To demonstrate the capability of the downscaled data in synchronously capturing the seasonal greenness dynamics, we calculated and compared the monthly NDVI median time series derived from the two products from 2014 to 2018 (Fig. 7); (2) To assess the overall performance of the downscaled data in accurately demonstrating the NDVI spatial changes, we randomly selected 1000 pixels' NDVI values in both MODIS and the downscaled images for every month from 2014 to 2018 to make linear fit lines (Fig. 8) [46].

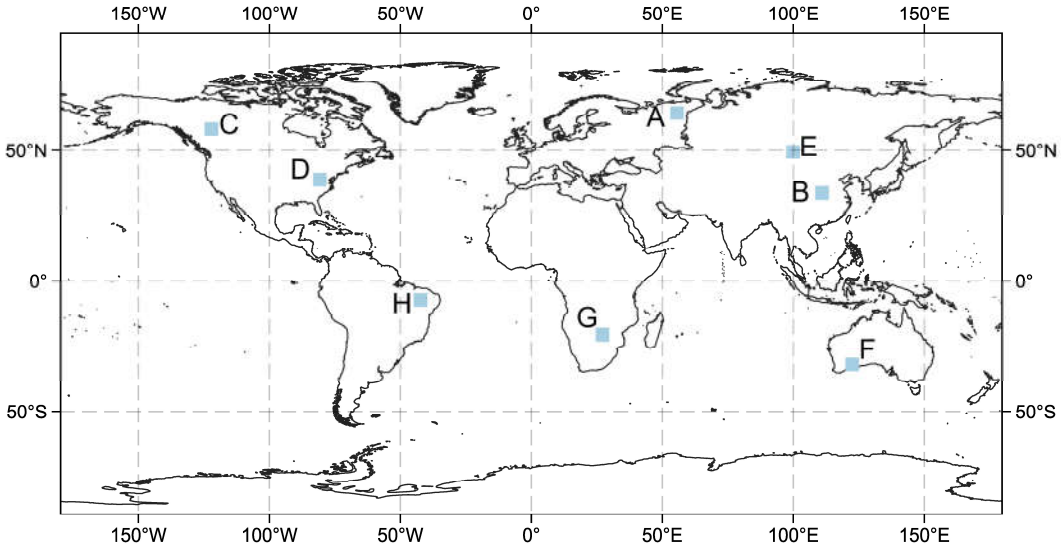


Figure 6. Locations of the eight selected verification areas that have eight representative vegetation types. (A) Siberia Evergreen Needleleaf Forests, (B) Asian Evergreen Broadleaf Forests, (C) North American Evergreen Needleleaf Forests, (D) North American Deciduous Broadleaf Forests, (E) Mongolian Steppe, (F) Australian Savanna, (G) African Shrub, (H) South American Tropical Shrub.

Since terrestrial vegetation has such a large impact on the carbon cycle, energy exchange, and water balance of terrestrial ecosystems, studying the global and regional vegetation dynamics of NDVI on an annual and seasonal basis is critical[47,48]. Here, to make a temporal comparison, we chose the validation period from 2014 to 2018 and retrieved the monthly median as the basic level. Figure 7 depicted the eight types of vegetation changes through time, with the green lines representing the MODIS NDVI and the red lines representing the downscaled NDVI. We could observe that the two lines were consistent overall. However, there were some noticeable biases between the two in 2015 and 2018 for some vegetation patches (e.g., Australian Savanna and African Shrub, Fig. 7f & 7g), indicating that there may be some disturbance factors affecting vegetation greenness, such as drought, fire, and so on[49].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

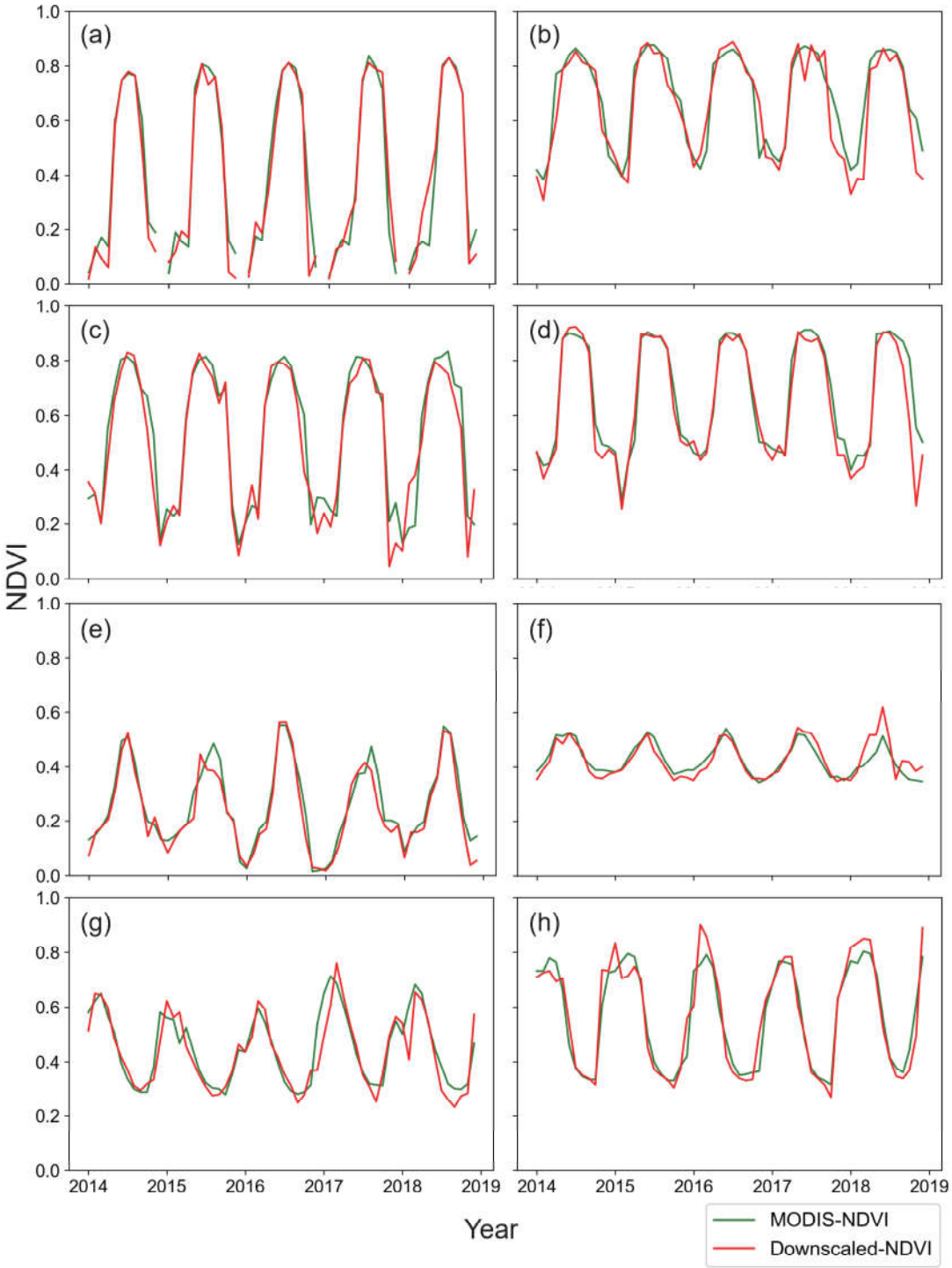


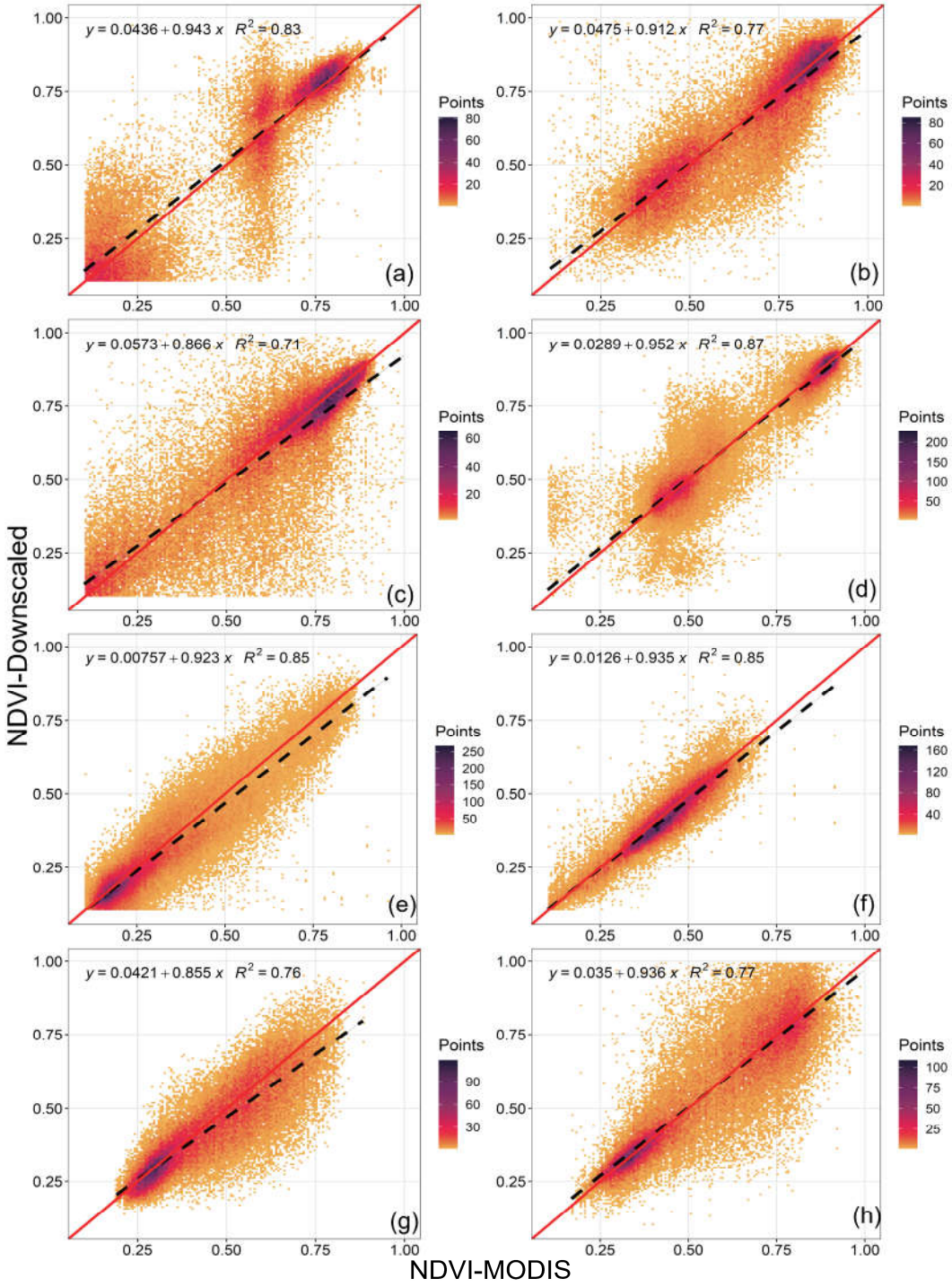
Figure 7. The long-term dynamic changes comparison between the downscaled and MODIS NDVI for different vegetation types. Comparison of NDVI for (a) Siberia Evergreen Needleleaf Forests, (b) Asian Evergreen Broadleaf Forests, (c) North American Evergreen Needleleaf Forests, (d) North American Deciduous Broadleaf Forests, (e) Mongolian Steppe, (f) Australian Savanna, (g) African Shrub, (h) South American Tropical Shrub.

Figure 8 depicts the linear fitting lines of the downscaled versus MODIS NDVI for the representative regions and vegetation types around the world. The fuchsia-point concentrated areas highlighted where the majority of the NDVI values in each type of vegetation were in the plots, as well as how lush each vegetation community was. The most intensive range of grassland NDVI values, for example, was approximately 0 to 0.4 (Fig. 8e & 8f), while the range of the forest was around the upper limit of NDVI (0.8~1.0) (Fig.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

8a-d). To make this comparison, we randomly sampled 1000 pixels while considering both the geographical and temporal changes; hence, the results should be reliable. The downscaled NDVI data matched well with the MODIS NDVI with reasonably high R^2 values (0.7~0.9). At the same time, the verifications suggest that the downscaled product can well capture the fine-scale NDVI dynamics and universal vegetation features.

As described above, the comparison and verification were carried out in time and space based on the NDVI seasonal time series and random sampling points across typical vegetation types over the globe. We might fairly infer that the downscaled NDVI product, which achieved both high resolution (250 m) and a long time-frame (1982–2018), could be utilized as a good greenness indicator for monitoring much of the world's vegetation communities.



 doi:10.20944/preprints202205.0367.v1

Figure 8. The correlation verification between the downscaled and MODIS NDVI for different vegetation types. The vegetation types (a-h) are same as those in Figure 7.

3.3.2. Comparison of the three NDVI datasets for areas with mixed vegetation types

We further compared the performance of the three NDVI datasets (MODIS, downscaled, and AVHRR) for areas with mixed vegetation types. This helps determine the discriminating capability of the NDVI products. The GAP/LANDFIRE National Terrestrial Ecosystems database contains more detailed vegetation and land cover classifications, allowing us to delineate the extent of three major vegetation types in Idaho: the Northern Rocky Mountain Dry-Mesic Montane Mixed Conifer Forest (MCF), the Northern Rocky Mountain Lower Montane, Foothill and Valley Grassland (GL), and the Northern Rocky Mountain Montane-Foothill Deciduous Shrubland (DS) [50].

We found that the higher spatial resolution of the standard MODIS NDVI and the downscaled NDVI products both showed better performance than the AVHRR product in discriminating forest from other vegetation types, as suggested by the more visible curve differences between different vegetation types for the former two datasets (Fig. 9a & 9b). As GL and DS have similar features, the differences were not generally apparent; however, MODIS and the downscaled NDVI still exhibit bigger differences than AVHRR NDVI. In general, the downscaled and MODIS NDVI products with higher spatial resolution can both accurately recognize greenness dynamics of different vegetation types, and their performance is similar. The AVHRR product displays a poor performance in applying to mixed-vegetation areas, which might be caused by its lower spatial resolution and failure in characterizing vegetation changes for heterogeneous regions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

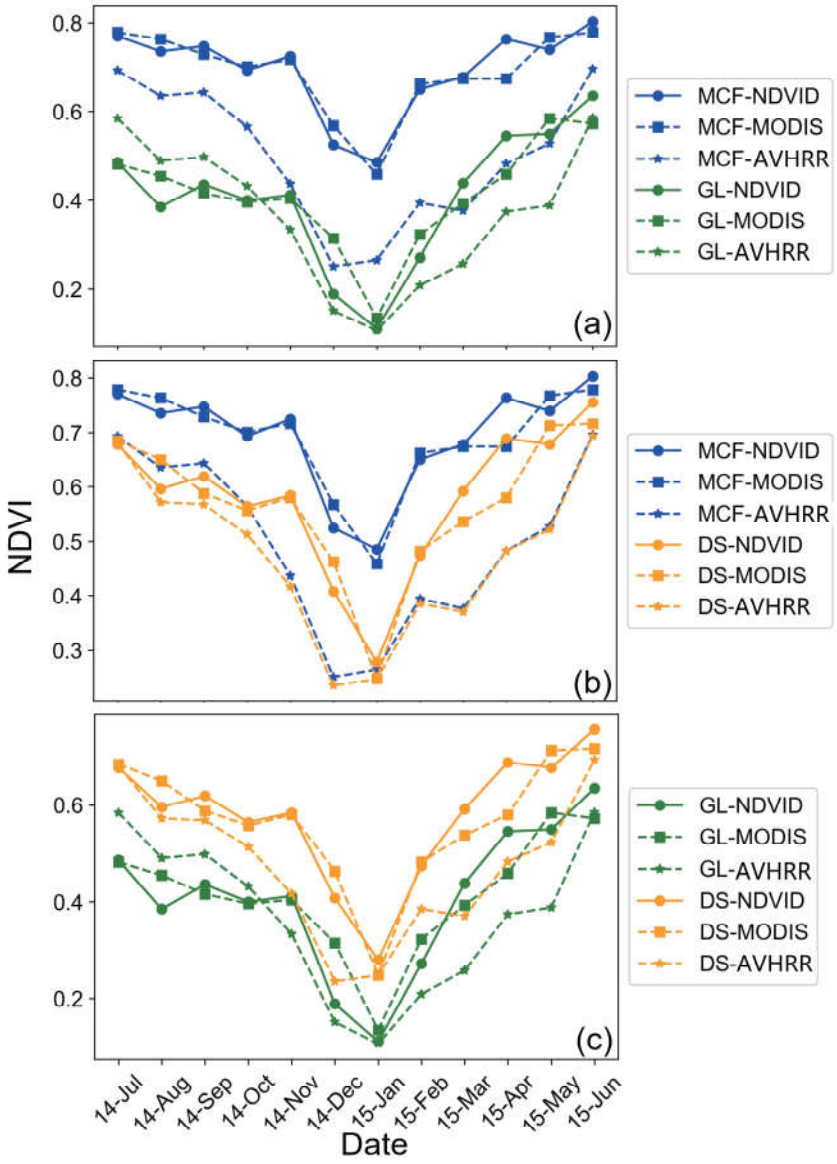


Figure 9. Performance comparison of the three NDVI datasets for distinguishing different vegetation types in the Idaho state. (a) mixed conifer forest (MCF) versus grassland (GL), (b) mixed conifer forest (MCF) versus deciduous shrubland (DS), and (c) grassland (GL) versus deciduous shrubland (DS). Each point represents the NDVI estimates from the MODIS, downscaled (NDVID), or AVHRR NDVI datasets for every month from one site.

4. Discussion

4.1. Value of the downscaled NDVI product

Satellite-based vegetation indices (VIs) have made a great contribution to the global to regional scale monitoring of terrestrial ecosystem dynamics. Many studies have revealed that over 25% to 50% of the global vegetated area displays a significant greening trend as a response to the rapid global environmental change, e.g., CO_2 fertilization, nitrogen deposition, climate change, and land cover change (LCC) etc.[51, 52]. However, some recent studies suggest that the interannual variability of vegetation greenness has significantly increased over time[53], and widespread greening-to-browning reversals are hidden in the overall vegetation greening[54,55,56]. Nearly all of these important studies relied on the long-term but coarse resolution AVHRR time series, or the fine resolution but short-term MODIS product. The drawback in the vegetation database hinders accurate estimation of the plant activity under a changing climate. The inferior data might play a role in causing contradictory conclusions regarding the vegetation greenness trends in the literature.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 May 2022 doi:10.20944/preprints202205.0367.v1

Here in this paper, we proposed a novel data fusion approach which shows good performance in downscaling the coarse-resolution AVHRR NDVI (~5 km) to the MODIS resolution (250 m). The downscaled long-term (1982–2018) and high resolution (250 m) global NDVI database has a high potential in accurately depicting the dynamics of trends, variability, and seasonality in greenness of different vegetation communities at both global and regional scales, and it is particularly useful in the mixed-vegetation areas.

4.2. Potential causes of discrepancies among different products

To test the accuracy of the downscaled NDVI product, we conducted rigorous validation of the product against the standard MODIS NDVI database at both the regional and global scales, and for different vegetation types. All the validation suggests high accuracy of the downscaled NDVI product and the data fusion algorithm. The majority of RMSE and MAE values fall between 0 and 0.1 (Fig. 5a & b), and the Pearson-r values reach > 0.7 for all the typical vegetation types over the globe (Fig. 8). However, larger discrepancies between the two products are evident over the sparse-vegetation areas, and the equatorial belt (Fig. 5d).

The large error in the sparse-vegetation areas likely related to an inherent defect of the algorithm. Since the downscaling approach relies on analyzing the difference in the coefficient of variation (CV) between AVHRR and MODIS NDVI database, and thus a slight change in the mean NDVI value for a sparse-vegetation area may lead to a big change in CV. As a result, the algorithm will enlarge the difference between the two products and cause large bias in the downscaled product.

The high deviations of the downscaled product in the equatorial belt should be related to the shortcomings of NDVI itself, i.e., NDVI is easy to reach its saturated status in dense vegetated areas such as the tropical rainforest, which has been widely documented in the literature[44,45]. EVI is a good alternative vegetation index other than NDVI, and it appears to be superior in discriminating subtle differences in areas of high vegetation density, which can be attributed to the correction for atmospheric and background effects in the EVI algorithm[57]. However, EVI time series is only available since entering the MODIS era, and its value is weakened by the short time span of the data, compared to NDVI database that has accumulated observations for several decades.

4.3. Limitations of the downscaling algorithm

A major limitation of the downscaling algorithm is that it is not suitable for the areas that has experienced dramatic land use/cover changes (LUCC), and thus we have removed the areas showing abrupt NDVI changes before the analyses. One important assumption of the proposed data fusion approach is that the vegetation type has not been replaced or removed during the studied period. It means the variability of a vegetation community in greenness did not change much or kept a relatively stable change.

However, a long-term and gradual change in CV has been considered by the downscaling algorithm, as we have split the entire period to two portions (1982–2000 and 2001–2018), and an adjusting parameter ($R_{CVx,y,n}$) has been used to quantify the long-term changes in CV of vegetation greenness (Eq. 2). Hence, the downscaled NDVI product should be useful in capturing both the short-term dynamics and long-term trends in vegetation greenness, which reflects vegetation flourishing, deterioration, and restoration due to climate change, drought, and competition etc.

LUCC has become an important phenomenon of the global environmental change, which have attracted much attention of the worldwide researchers. LUCC can be induced by both human activities and natural disturbance. On the one hand, human activities affect vegetation cover in two directions: destruction and restoration. For example, China has been conducting many large-scale ecological restoration projects (ERPs) for the past two decades[58]. At the same time, about 82 billion m^2 of farmland had been established in the Southeast Asia's highlands, and a huge amount of former forest lands had been turned into agriculture globally between 2001 and 2019[43,59]. On the other hand, the global climate

change has caused latitude and altitude changes in vegetation in boreal, temperate, and tropical ecosystems through prolonged drought, extreme rainfall, and wildfires[60,61]. Thus, it is of high significance to develop more robust data fusion techniques for satellite-based vegetation indices, which is feasible for the areas experiencing dramatic LUCCs.

Another limitation of the study is that here we downscaled the AVHRR NDVI database as monthly time series. The low temporal resolution is not appropriate for phenological studies. The original AVHRR and MODIS datasets are daily time series. However, frequent cloud obstructions induce a lot of data gaps or noises in the daily time series, and thus the daily database is not reliable in monitoring the long-term dynamics and trends of vegetation greenness. We composited the NDVI images as monthly time series using the Maximum Value Composite (MVC) algorithm can largely reduce the uncertainties due to the data quality. The monthly high-resolution NDVI database is particularly useful in analyzing the fine-scale vegetation's response to global climate change and drought.

In addition, we did not extend the time span of the downscaled NDVI product beyond 2018, since the AVHRR sensor has degraded since the late 2018, and the NOAA AVHRR NDVI database display widespread data gaps in the southern hemisphere after then[62]. However, the major purpose of this study is to extend the MODIS resolution NDVI database to the pre-2000 period. Besides, a 38-year and high resolution NDVI time series should have many advantages than the 22-year MODIS NDVI product in analyzing the long-term trends of the global vegetation dynamics.

5. Conclusion

GEE is used as the operation platform in this work to undertake a long-term and global-scale data fusion of remotely sensed vegetation index data. The primary idea behind our technique was to extract the fine-scale NDVI spatial information from the high-resolution MODIS images, and then integrate it with the long-term NDVI temporal information from the AVHRR database. Finally, we created a downscaled 250-m resolution global NDVI dataset from 1982 to 2018, and the data quality was compared to the standard MODIS NDVI products at both the regional and global scales. The RMSE and MAE are less than 0.1 in most locations, and the Pearson's r is typically greater than 0.6, according to the validation results. However, in locations having abrupt landcover changes, the downscaled products may contain bigger biases. The good performance of the downscaling algorithm likely suggests that the coefficient of variation (CV) could be a valuable intermediary for combining the fine and coarse-resolution remote sensing data. The downscaled global NDVI product offers a great potential for investigating the global plant response to climate change and monitoring the fine-scale (250 m) and long-term (four decades) vegetation dynamics in facing to different environmental stress.

Author Contributions: C.D. developed the downscaling algorithm, C.D. and Z.M. designed the study, Z.M. generated the data, C.D. and Z.M. analyzed the data and wrote the paper. All co-authors contributed to writing and commented on the manuscript.

Acknowledgments: The authors acknowledge the financial supports from the National Key Research and Development Program of China (Grants No. 2021YFC3001000) and the National Natural Science Foundation of China (Grants No. 41801254).

Conflicts of Interest: The authors declare no conflict of interest.

References

1 Rustad, L. E. The response of terrestrial ecosystems to global climate change: Towards an integrated approach. *Science of The Total Environment* **404**, 222-235, doi:https://doi.org/10.1016/j.scitotenv.2008.04.050 (2008).

2 Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. & Nemani, R. R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **386**, 698-702, doi:10.1038/386698a0 (1997).

3 Li, D., Wu, S., Liu, L., Zhang, Y. & Li, S. Vulnerability of the global terrestrial ecosystems to climate change. *Global Change Biology* **24**, 4095-4106, doi:https://doi.org/10.1111/gcb.14327 (2018).

4 Peters, A. J. *et al.* Drought monitoring with NDVI-based Standardized Vegetation Index. *Photogrammetric Engineering and Remote Sensing* **68**, 71-75 (2002).

5 Gu, Y. & Wylie, B. K. Downscaling 250-m MODIS Growing Season NDVI Based on Multiple-Date Landsat Images and Data Mining Approaches. *Remote Sensing* **7**, 3489-3506 (2015).

6 Candiago, S., Remondino, F., De Giglio, M., Dubbini, M. & Gattelli, M. Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. *Remote Sensing* **7**, 4026-4047 (2015).

7 Jönsson, A. M., Eklundh, L., Hellström, M., Bärring, L. & Jönsson, P. Annual changes in MODIS vegetation indices of Swedish coniferous forests in relation to snow dynamics and tree phenology. *Remote Sensing of Environment* **114**, 2719-2730, doi:https://doi.org/10.1016/j.rse.2010.06.005 (2010).

8 Yoder, B. J. & Waring, R. H. The normalized difference vegetation index of small Douglas-fir canopies with varying chlorophyll concentrations. *Remote Sensing of Environment* **49**, 81-91, doi:https://doi.org/10.1016/0034-4257(94)90061-2 (1994).

9 Robinson, N. P. *et al.* A Dynamic Landsat Derived Normalized Difference Vegetation Index (NDVI) Product for the Conterminous United States. *Remote Sensing* **9**, 863 (2017).

10 Brown, M. E., Lary, D. J., Vrieling, A., Stathakis, D. & Mussa, H. Neural networks as a tool for constructing continuous NDVI time series from AVHRR and MODIS. *International Journal of Remote Sensing* **29**, 7141-7158, doi:10.1080/01431160802238435 (2008).

11 Tarnavsky, E., Garrigues, S. & Brown, M. E. Multiscale geostatistical analysis of AVHRR, SPOT-VGT, and MODIS global NDVI products. *Remote Sensing of Environment* **112**, 535-549, doi:https://doi.org/10.1016/j.rse.2007.05.008 (2008).

12 Rao, Y., Zhu, X., Chen, J. & Wang, J. An Improved Method for Producing High Spatial-Resolution NDVI Time Series Datasets with Multi-Temporal MODIS NDVI Data and Landsat TM/ETM+ Images. *Remote Sensing* **7**, 7865-7891 (2015).

13 Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G. & Worthy, L. D. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment* **105**, 142-154, doi:https://doi.org/10.1016/j.rse.2006.06.018 (2006).

14 Stellmes, M., Udelhoven, T., Röder, A., Sonnenschein, R. & Hill, J. Dryland observation at local and regional scale — Comparison of Landsat TM/ETM+ and NOAA AVHRR time series. *Remote Sensing of Environment* **114**, 2111-2125, doi:https://doi.org/10.1016/j.rse.2010.04.016 (2010).

15 Zhu, X., Cai, F., Tian, J. & Williams, T. K.-A. Spatiotemporal Fusion of Multisource Remote Sensing Data: Literature Survey, Taxonomy, Principles, Applications, and Future Directions. *Remote Sensing* **10**, 527 (2018).

16 Kerdiles, H. & Grondona, M. O. NOAA-AVHRR NDVI decomposition and subpixel classification using linear mixing in the Argentinean Pampa. *International Journal of Remote Sensing* **16**, 1303-1325, doi:10.1080/01431169508954478 (1995).

17 Meng, J., Du, X. & Wu, B. Generation of high spatial and temporal resolution NDVI and its application in crop biomass estimation. *International Journal of Digital Earth* **6**, 203-218, doi:10.1080/17538947.2011.623189 (2013).

18 Feng, G., Masek, J., Schwaller, M. & Hall, F. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing* **44**, 2207-2218, doi:10.1109/TGRS.2006.872081 (2006).

19 Zhu, X., Chen, J., Gao, F., Chen, X. & Masek, J. G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment* **114**, 2610-2623, doi:https://doi.org/10.1016/j.rse.2010.05.032 (2010).

20 Hilker, T. *et al.* A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment* **113**, 1613-1627, doi:https://doi.org/10.1016/j.rse.2009.03.007 (2009).

21 Htitiou, A., Boudhar, A. & Benabdelouahab, T. Deep Learning-Based Spatiotemporal Fusion Approach for Producing High-Resolution NDVI Time-Series Datasets. *Canadian Journal of Remote Sensing* **47**, 182-197, doi:10.1080/07038992.2020.1865141 (2021).

22 Didan, K., Munoz, A. B., Solano, R. & Huete, A. MODIS vegetation index user's guide (MOD13 series). *University of Arizona: Vegetation Index and Phenology Lab* (2015).

23 van Leeuwen, W. J. D., Orr, B. J., Marsh, S. E. & Herrmann, S. M. Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications. *Remote Sensing of Environment* **100**, 67-81, doi:https://doi.org/10.1016/j.rse.2005.10.002 (2006).

24 Fan, X. & Liu, Y. A global study of NDVI difference among moderate-resolution satellite sensors. *ISPRS Journal of Photogrammetry and Remote Sensing* **121**, 177-191, doi:https://doi.org/10.1016/j.isprsjprs.2016.09.008 (2016).

25 Fensholt, R. & Proud, S. R. Evaluation of Earth Observation based global long term vegetation trends — Comparing GIMMS and MODIS global NDVI time series. *Remote Sensing of Environment* **119**, 131-147, doi:https://doi.org/10.1016/j.rse.2011.12.015 (2012).

26 Fontana, F., Rixen, C., Jonas, T., Aberegg, G. & Wunderle, S. Alpine Grassland Phenology as Seen in AVHRR, VEGETATION, and MODIS NDVI Time Series - a Comparison with In Situ Measurements. *Sensors* **8**, doi:10.3390/s8042833 (2008).

27 Gorelick, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* **202**, 18-27, doi:https://doi.org/10.1016/j.rse.2017.06.031 (2017).

28 Tian, F., Wu, B., Zeng, H., Zhang, X. & Xu, J. Efficient Identification of Corn Cultivation Area with Multitemporal Synthetic Aperture Radar and Optical Images in the Google Earth Engine Cloud Platform. *Remote Sensing* **11**, 629 (2019).

29 Elnashar, A. *et al.* Downscaling TRMM Monthly Precipitation Using Google Earth Engine and Google Cloud Computing. *Remote Sensing* **12**, 3860 (2020).

30 Tavakkoli Piralilou, S. *et al.* A Google Earth Engine Approach for Wildfire Susceptibility Prediction Fusion with Remote Sensing Data of Different Spatial Resolutions. *Remote Sensing* **14**, 672 (2022).

31 Vermote, E. *et al.* NOAA Climate Data Record (CDR) of normalized Difference Vegetation Index (NDVI), Version 4. *NOAA Natl. Clim. Data Cent* (2014).

32 Chen, P. Y., Srinivasan, R., Fedosejevs, G. & Kiniry, J. R. Evaluating different NDVI composite techniques using NOAA-14 AVHRR data. *International Journal of Remote Sensing* **24**, 3403-3412, doi:10.1080/0143116021000021279 (2003).

33 Pettorelli, N. *et al.* Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution* **20**, 503-510, doi:https://doi.org/10.1016/j.tree.2005.05.011 (2005).

34 Zhai, Y., Qu, Z. & Hao, L. Land Cover Classification Using Integrated Spectral, Temporal, and Spatial Features Derived from Remotely Sensed Images. *Remote Sensing* **10**, 383 (2018).

35 Bindhu, V. M. & Narasimhan, B. Development of a spatio-temporal disaggregation method (DisNDVI) for generating a time series of fine resolution NDVI images. *ISPRS Journal of Photogrammetry and Remote Sensing* **101**, 57-68, doi:https://doi.org/10.1016/j.isprsjprs.2014.12.005 (2015).

36 Tian, J. *et al.* Improving the accuracy of spring phenology detection by optimally smoothing satellite vegetation index time series based on local cloud frequency. *ISPRS Journal of Photogrammetry and Remote Sensing* **180**, 29-44, doi:https://doi.org/10.1016/j.isprsjprs.2021.08.003 (2021).

37 Theobald, D. M., Harrison-Atlas, D., Monahan, W. B. & Albano, C. M. Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. *PloS one* **10**, e0143619 (2015).

38 Gu, Y. *et al.* An Optimal Sample Data Usage Strategy to Minimize Overfitting and Underfitting Effects in Regression Tree Models Based on Remotely-Sensed Data. *Remote Sensing* **8**, 943 (2016).

39 Bartlett, P. L., Long, P. M., Lugosi, G. & Tsigler, A. Benign overfitting in linear regression. *Proceedings of the National Academy of Sciences* **117**, 30063-30070, doi:doi:10.1073/pnas.1907378117 (2020).

40 Nomura, R. & Oki, K. Downscaling of MODIS NDVI by Using a Convolutional Neural Network-Based Model with Higher Resolution SAR Data. *Remote Sensing* **13**, doi:10.3390/rs13040732 (2021).

41 Shiff, S., Helman, D. & Lensky, I. M. Worldwide continuous gap-filled MODIS land surface temperature dataset. *Scientific Data* **8**, 74, doi:10.1038/s41597-021-00861-7 (2021).

42 Xu, X., Jia, G., Zhang, X., Riley, W. J. & Xue, Y. Climate regime shift and forest loss amplify fire in Amazonian forests. *Global Change Biology* **26**, 5874-5885, doi:https://doi.org/10.1111/gcb.15279 (2020).

43 Feng, Y. *et al.* Doubling of annual forest carbon loss over the tropics during the early twenty-first century. *Nature Sustainability*, doi:10.1038/s41893-022-00854-3 (2022).

44 Zhu, X. & Liu, D. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. *ISPRS Journal of Photogrammetry and Remote Sensing* **102**, 222-231, doi:https://doi.org/10.1016/j.isprsjprs.2014.08.014 (2015).

45 Hmimina, G. *et al.* Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. *Remote Sensing of Environment* **132**, 145-158, doi:https://doi.org/10.1016/j.rse.2013.01.010 (2013).

46 Zhang, Y. *et al.* NDVI-based vegetation changes and their responses to climate change from 1982 to 2011: A case study in the Koshi River Basin in the middle Himalayas. *Global and Planetary Change* **108**, 139-148, doi:https://doi.org/10.1016/j.gloplacha.2013.06.012 (2013).

47 Wu, C. *et al.* Present-day and future contribution of climate and fires to vegetation composition in the boreal forest of China. *Ecosphere* **8**, e01917, doi:https://doi.org/10.1002/ecs2.1917 (2017).

48 Ning, T., Liu, W., Lin, W. & Song, X. NDVI Variation and Its Responses to Climate Change on the Northern Loess Plateau of China from 1998 to 2012. *Advances in Meteorology* **2015**, 725427, doi:10.1155/2015/725427 (2015).

49 Solórzano, J. V. & Gao, Y. Forest Disturbance Detection with Seasonal and Trend Model Components and Machine Learning Algorithms. *Remote Sensing* **14**, 803 (2022).

50 Homer, C. *et al.* Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing* **81**, 345-354 (2015).

51 De Jong, R., de Bruin, S., de Wit, A., Schaepman, M. E., & Dent, D. L. Analysis of monotonic greening and browning trends from global NDVI time-series. *Remote Sensing of Environment*, 115(2), 692-702 (2011).

52 Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., ... & Zeng, N. Greening of the Earth and its drivers. *Nature Climate Change*, 6(8), 791-795 (2016).

53 Chen, C., He, B., Yuan, W., Guo, L., & Zhang, Y. Increasing interannual variability of global vegetation greenness. *Environmental Research Letters*, 14(12), 124005 (2019).

54 Pan, N., Feng, X., Fu, B., Wang, S., Ji, F., & Pan, S. Increasing global vegetation browning hidden in overall vegetation greening: Insights from time-varying trends. *Remote Sensing of Environment*, 214, 59-72 (2018).

55 Ding, Z., Peng, J., Qiu, S., & Zhao, Y. Nearly half of global vegetated area experienced inconsistent vegetation growth in terms of greenness, cover, and productivity. *Earth's Future*, 8(10), e2020EF001618 (2020).

56 Cortés, J., Mahecha, M. D., Reichstein, M., Myneni, R. B., Chen, C., & Brenning, A. Where are global vegetation greening and browning trends significant?. *Geophysical Research Letters*, 48(6), e2020GL091496 (2021).

57 Kamel, D., Munoz, A. B., Ramon, S., & Huete, A. MODIS Vegetation Index User's Guide. Vegetation Index and Phenology Lab of The University of Arizona, The University of Arizona: Tucson, AZ, USA (2015).

58 Wang, C., Gao, Q., Wang, X. & Yu, M. Spatially differentiated trends in urbanization, agricultural land abandonment and reclamation, and woodland recovery in Northern China. *Scientific Reports* **6**, 37658, doi:10.1038/srep37658 (2016).

59 Zeng, Z. *et al.* Highland cropland expansion and forest loss in Southeast Asia in the twenty-first century. *Nature Geoscience* **11**, 556-562, doi:10.1038/s41561-018-0166-9 (2018).

 doi:10.20944/preprints202205.0367.v1

60 Rosenzweig, C. *et al.* Attributing physical and biological impacts to anthropogenic climate change. *Nature* **453**, 353-357, doi:10.1038/nature06937 (2008).
61 Gonzalez, P., Neilson, R. P., Lenihan, J. M. & Drapek, R. J. Global patterns in the vulnerability of ecosystems to vegetation shifts due to climate change. *Global Ecology and Biogeography* **19**, 755-768, doi:https://doi.org/10.1111/j.1466-8238.2010.00558.x (2010).
62 Otón, G., Lizundia-Loiola, J., Pettinari, M. L., & Chuvieco, E. Development of a consistent global long-term burned area product (1982–2018) based on AVHRR-LTDR data. International Journal of Applied Earth Observation and Geoinformation, 103, 102473 (2021).