Remote Sensing of Environment 199 (2017) 201–217

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A time series of annual land use and land cover maps of China from 1982 to 2013 generated using AVHRR GIMMS NDVI3g data

Yaqian He *, Eungul Lee, Timothy A. Warner

Department of Geology and Geography, West Virginia University, Morgantown, WV 26506, United Sates

A R T I C L E I N F O

Article history:
Received 27 June 2016
Received in revised form 29 June 2017
Accepted 15 July 2017

Keywords:
Land use and land cover
GIMMS NDVI3g
MODIS MCD12Q1
Phenology
China

A B S T R A C T

A time series of annual land use and land cover (LULC) maps that cover an extended period of time is a key dataset for climatological studies investigating land-atmosphere interaction. Change in LULC can influence regional climate by altering the surface roughness, soil moisture, heat flux partition, and terrestrial carbon storage. Although annual global LULC maps are generated from Moderate-resolution Imaging Spectroradiometer (MODIS) data, the earliest MODIS LULC map is for 2001, which limits the potential time period for climatological analyses. This study produced a continuous series of annual LULC maps of China from 1982 to 2013 using random forest classification of 19 phenological metrics derived from Advanced Very High Resolution Radiometer (AVHRR) Global Inventory Modeling and Mapping Studies (GIMMS) third generation NDVI (NDVI3g) data. The classifier was trained using reference data derived from the MODIS land cover type product (MCD12Q1). Based on a comparison with Google Earth images, the overall accuracy of a simplified eight-class version of our 2012 LULC map is 73.8%, which is not significantly different from the accuracy of the MODIS map of the same year. Our maps indicate that for the three decades studied, the area of croplands and forests in China increased, and the area of grasslands decreased. These annual maps of land cover will be an important dataset for future climate studies, and the methodologies used in this study can be applied to other geographical regions where availability of continuous time series of LULC maps is limited.

1. Introduction

Land surface condition is an important factor in determining climate through both biophysical and biogeochemical processes (Foley et al., 2003). Recent changes in land surface conditions (e.g., albedo, soil moisture, and surface roughness) and atmospheric composition (e.g., CO_2 and methane) due to land use and land cover change (LULCC) have had significant effects on regional and global climates (Bonan et al., 1992; Foley et al., 2005; Lee et al., 2015; Mahmood et al., 2014; McPherson, 2007; Pielke, 2005). For example, Bonan et al. (1992) found that replacing bare ground and tundra with boreal forest results in warming of both winter and summer air temperatures. In a more recent study, He and Lee (2016) found that vegetation growth in the Sahel may have induced the recent trend of increasing rainfall in that region.

There have been many attempts to explore the effects of LULCC on climate based on observational (Kaufmann and Stern, 1997; Lee et al., 2009; Lee et al., 2015) and modeling (Douglas et al., 2006; Eltahir, 1996; Lawrence and Chase, 2010; Lawrence et al., 2012; Lee et al., 2011) studies across the globe (Pielke et al., 2011). Both observational and modeling studies require as an input land use and land cover (LULC) maps that characterize the pattern of changing LULC over time. However, due to the limited length of time for which such LULC maps are available, previous climate studies have often made simplifying assumptions, for example, using a potential vegetation map (Fu, 2003) or a single map of a specific year (Zhu, 2012) to represent the LULC of the entire study period. For instance, Zhu (2012) employed a single map of irrigation for the year 2000 to examine the impacts of irrigation on climate during the period of 1978 to 2008, despite the tremendous change in irrigation infrastructure during those three decades. Therefore, obtaining a continuous sequence of annual LULC maps over an extended time period, for at least multiple decades, is critical for quantifying the effects of LULCC on climate.

Townshend et al. (1991) and Running et al. (1994) have noted that only remotely sensed data can potentially provide accurate and repeatable global land use and land cover for monitoring change. Since then, a number of studies have generated LULC maps in China using remotely sensed data. For instance, Wang et al. (2012) classified urban areas based on Landsat TM/ETM+ data during the years 1986–1994, 1999–2002, and 2008–2010, and examined urban expansion in China from the 1990s to 2010s. Liu et al. (2014) used Landsat and Huanjing-1 satellite data from the late 1980s, 1995, 2000, 2005, and 2010 to generate a sequence of LULC maps of China every five years, which they used to investigate the spatiotemporal change patterns of LULC. However, these

* Corresponding author.
E-mail address: yahe@mix.wvu.edu (Y. He).

http://dx.doi.org/10.1016/j.rse.2017.07.010
0034-4257/

datasets have relatively low temporal frequency of acquisition. Finding cloud-free images to cover all of China may require imagery from several different years or a combination of data from different satellites. These problems may result in inconsistent periods of time in the analysis and biased results because of inconsistent data sources. Another remotely sensed dataset that is particularly useful for such work is cloud-minimized multi-temporal composites (Holben, 1986) derived from high temporal frequency images acquired at a coarse/moderate spatial scale (pixels 250 m or larger). This coarse/moderate spatial scale is better suited to the coarse scale typically used in climate modeling (e.g. 0.05° (approximately 5 km) or larger (Lawrence and Chase, 2007)) compared to the finer scale sensors with global coverage, such as Landsat (Kim et al., 2014; Sexton et al., 2013). An example of a dataset generated from a moderate-scale sensor that has been used for climate work is the Moderate-resolution Imaging Spectroradiometer (MODIS) land cover type product (MCD12Q1) (Lawrence and Chase, 2007). Unfortunately, however, global LULC data from MODIS is only available since 2001, thus limiting the time span that can be studied, an important consideration in climate studies.

Fortunately, there is a dataset with coarse spatial resolution and high temporal frequency that does provide global data for an extended period: NOAA Advanced Very High Resolution Radiometer (AVHRR) imagery (James and Kalluri, 1994). Often combined with MODIS data, AVHRR data have been widely used for monitoring land surface conditions at coarse scales (Andres et al., 1994; Lunetta et al., 2006). The early AVHRR instruments had just 4 spectral bands, although this was later increased to 5 in some subsequent sensors (Tucker et al., 2005). One of the most important derived datasets from AVHRR data is based on the Normalized Difference Vegetation Index (NDVI), a normalized ratio of the visible red and near infrared spectral bands (Rouse et al., 1974). Water is generally associated with negative NDVI values, bare soil with values near zero (Sabins and Lulla, 2007), and vegetation with high positive values that are broadly indicative of the amount of photosynthesizing vegetation present (Dappen, 2003).

There is a long history of using AVHRR imagery for continental (Townshend et al., 1987; Tucker et al., 1985) and even global scale mapping (DeFries et al., 1995; DeFries and Townshend, 1994; Hansen et al., 2000; Loveland et al., 2000). For example, Gitas et al. (2004) used AVHRR imagery to map burned areas in the Spanish Mediterranean coast region after a large forest fire. More recently, Zhang et al. (2016) generated a global climatic vegetation map from AVHRR imagery. These studies typically used NDVI data alone, or in combination with reflectance values in each spectral band as well as temperature variables from the AVHRR thermal bands. However, in a study of the importance of different variables in discriminating classes for producing a global land cover map using AVHRR, DeFries et al. (1995) found summary metrics describing NDVI phenology were by far the most important for most vegetation classes. Nevertheless, for LULC classes for which vegetation phenology is of limited diagnostic use, such as the urban, snow, water, and barren classes, ancillary data such as digital cartographic information was found to be necessary (e.g. Loveland et al., 2000). Significantly, these early efforts focused on producing single maps, and to our knowledge, AVHRR data has not yet been used to produce a time series of LULC maps over broad regions. In recent years less attention has been paid to AVHRR data classification due to the availability of improved, higher signal to noise, and higher spatial resolution remotely sensed data, such as from MODIS (Schneider et al., 2009; Muhammad et al., 2015). However, AVHRR data comprise the longest global image time-series, and thus provide the potential to generate a long-term time-series of LULC maps, a key input for climatological analysis.

The aim of this study is to produce annual land use and land cover maps of Mainland China for the three decades covering the period from 1982 to 2013. The maps are produced based on a random forest classification using phenological metrics derived from the AVHRR Global Inventory Modeling and Mapping Studies (GIMMS) NDVI third generation (NDVI3g) dataset and trained using land cover information from the MODIS MCD12Q1 dataset. The key attribute of our classified land cover maps is that they comprise a continuous time series covering three decades, which contrasts with the limited temporal information previously available for use in climate studies.

2. Data and methods

2.1. Data

The primary dataset in this study is AVHRR GIMMS NDVI3g, first version, with data covering the period from 1982 to 2013, which were acquired from https://nex.nasa.gov/nex/projects/1349/. GIMMS NDVI3g data have been normalized to account for issues such as sensor calibration loss, orbital drift, and atmospheric effects such as volcanic eruptions (Pinzon and Tucker, 2014). The spatial resolution of the data is 1/12°. Each layer in the dataset is a bimonthly (15 days) composite produced using the maximum NDVI value for each pixel (Holben, 1986). We reprojected NDVI3g data onto a geographic grid, with WGS 1984 spheroid.

The second major dataset used is MODIS MCD12Q1 collection 5 data (Channan et al., 2014), which we utilized as a reference source for identification of training areas and class labels for the AVHRR classification. The MODIS MCD12Q1 data with a WGS 1984 spheroid were obtained from University of Maryland http://glcf.umd.edu/data/lc/. The data comprise annual maps of land use classes keyed to the International Geosphere-biosphere Programme (IGBP) classification system, covering the period 2001 to 2012. MODIS MCD12Q1 collection 5 data are generated from MODIS bands 1–7 and enhanced vegetation index data using an ensemble supervised classification algorithm (Friedl et al., 2010). The original MODIS MCD12Q1 data were resampled by the University of Maryland to 1/12° pixels, the resolution of the AVHRR GIMMS NDVI3g dataset, using a majority aggregation method (Channan et al., 2014). In this approach, each new pixel was labeled as the class that most frequently occurred in the original resolution data, for the area encompassing that new pixel.

The Chinese Land-Use/cover (CLU) dataset for 1995, 2000, 2005, and 2010 were obtained from the Data Center for Resources and Environmental Sciences (RESDC), Chinese Academy of Sciences (http://www.resdc.cn). CLU data were compared with our classified LULC maps in order to assess the 32-year time series of LULC dataset more robustly. The CLU data are produced mainly from 30 m Landsat TM data, as well as 30 m Huangjing-1 satellite imagery and 20 m China-Brazil Earth Resources Satellite-1 imagery using a human-computer interactive interpretation method (Liu et al., 2003a; Liu et al., 2010; Liu et al., 2014). The CLU data have 6 classes: cropland, woodland, water body (which includes water, snow, and ice), built-up land, and unused land. The accuracy of the six classes of land use is about 94.3% (Liu et al., 2014). RESDC provides the CLU data with spatial resolution of 1 km. To be consistent with our classified maps, we resampled the data to 1/12° spatial resolution, using a majority aggregation approach.

2.2. Data pre-processing

Although the temporal compositing process used in producing the NDVI3g dataset greatly reduces cloud and other atmospheric effects, residual noise remains (de Jong et al., 2011; Reed et al., 1994). Cleaning and smoothing NDVI data is therefore necessary (Fig. 1) (A second version of GIMMS NDVI3g dataset, including data up to 2015, has recently been made available. However, this new dataset is not directly compatible with the original dataset, and for that reason, we did not incorporate the new data. Specifically, the binary VI3g data format of the first version was changed to the Network Common Data Form (NetCDF) for the second version, and the 1 to 7 range of the quality flag for the first version was adjusted to 0 to 2 for the second version.)

1) *Cleaning AVHRR GIMMS NDVI3g data* - The quality information flags for NDVI3g data range from 1 to 7. Flag values of 1 and 2 represent good data, 3 indicates the application of a spline interpolation (i.e.

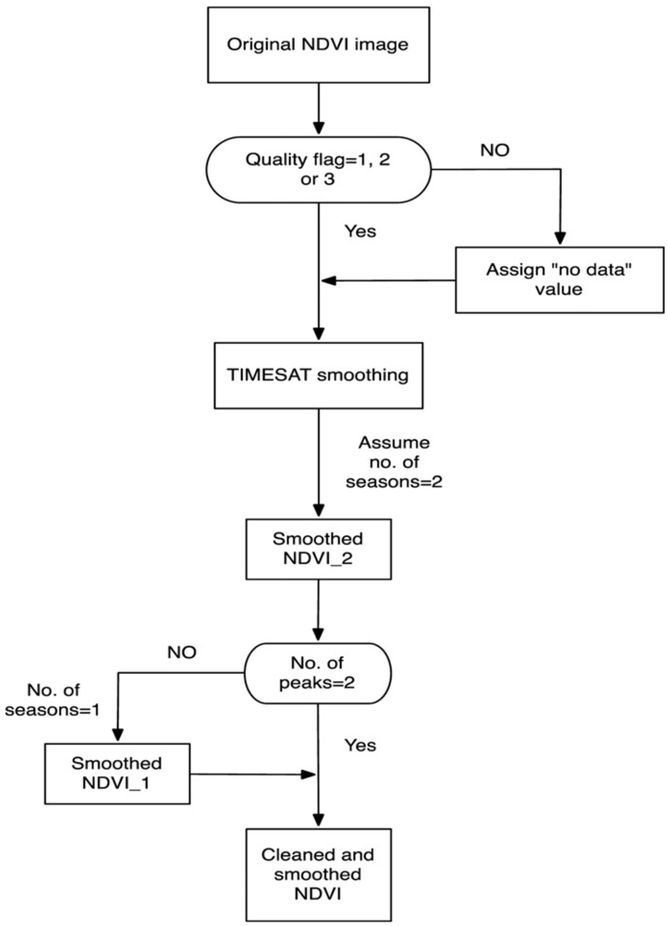

Fig. 1. Flowchart for pre-processing the GIMMS NDVI3g data.

a data gap that has been filled), 4 and 6 indicate possible snow, 5 indicates a gap filled through averaging the seasonal profile, and 7 indicates missing data. We retained data with flag values of 1, 2, and 3, following Chen et al. (2016), and excluded data with flag values of 4 to 7 by assigning those locations a "no data" value.

2) *Smoothing the cleaned AVHRR GIMMS NDVI data* - Even after removing pixels that are flagged as having low quality, smoothing is required to reduce the noise. This was done using the program TIMESAT (Jönsson and Eklundh, 2002, 2004). TIMESAT is a software package for analyzing time-series of satellite sensor data and available from http://web.nateko.lu.se/timesat/timesat.asp. Numerous studies have employed TIMESAT for phenological analysis (Boyd et al., 2011; Heumann et al., 2007; Palacios-Orueta et al., 2012; Palmer et al., 2015). TIMESAT offers multiple outlier removal methods: median filtering, approaches using weights from the Seasonal-Trend decomposition procedure based on Loess (STL)-decomposition (Cleveland et al., 1990) or weights from STL-decomposition multiplied with the original weights assigned based on the ancillary data, and three smoothing functions (a Savitzky-Golay filter, an asymmetric Gaussian filter, and a double logistic smoothing function) (Jönsson and Eklundh, 2004). Following He et al. (2015), we selected median filtering to remove outliers that deviate more than two standard deviations from the median in a moving window (Eklundh and Jönsson, 2015) and the double logistic method to smooth the time-series.

TIMESAT smoothing requires the user to specify the number of growing season per year. In China, two growing seasons are common in some locations (e.g., cropland areas in Southern China), and only one growing season predominates elsewhere (e.g., deciduous forest and cropland in Northern China). In order to identify the appropriate number of seasons, we first smoothed the entire NDVI time-series with two growing seasons (Smoothed NDVI_2) (Fig. 1). Each resulting smoothed time-series was checked to determine if there were at least four points of increasing NDVI before the NDVI peak (i.e., where the first derivative is zero) and four points of decreasing NDVI points after the peak, in each year. We chose these criteria as the potential growing season for cropland in China is approximately four months. If the time-series for a particular year did not meet these criteria, it was replaced with a smoothing based on the assumption of a single growing season (Smoothed NDVI_1) (Fig. 1).

For the time series of a single pixel, any year with missing data that comprises a continuous period longer than 0.2 years, is not smoothed by TIMESAT. Similarly, the entire 32 year time-series for any pixel is also not smoothed if 25% of the data is missing (Gao et al., 2008). These pixels were labeled "no data" in our classified LULC maps, and covered 3.7% of the area of China.

Fig. 2 displays the spatial patterns of cleaned, and cleaned and smoothed NDVI for Julian day 75 of 2011 as an example. Pixels with flag values 4 to 7, which were excluded in the cleaning processes, are

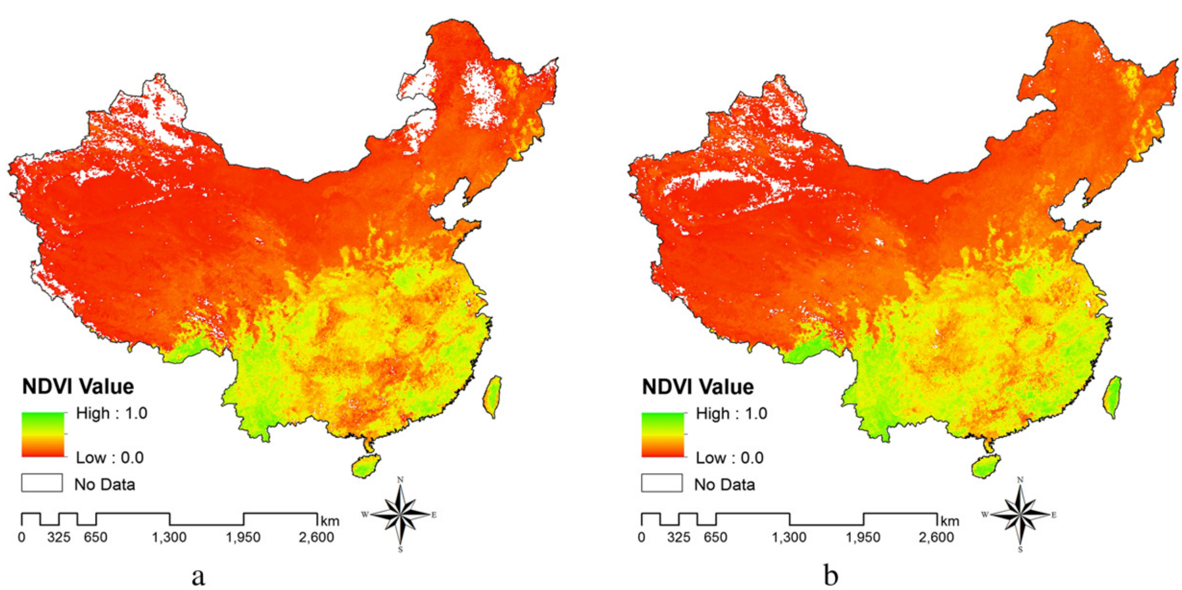

Fig. 2. Spatial patterns for Julian day 75 of 2011 of (a) cleaned and (b) cleaned and smoothed NDVI.

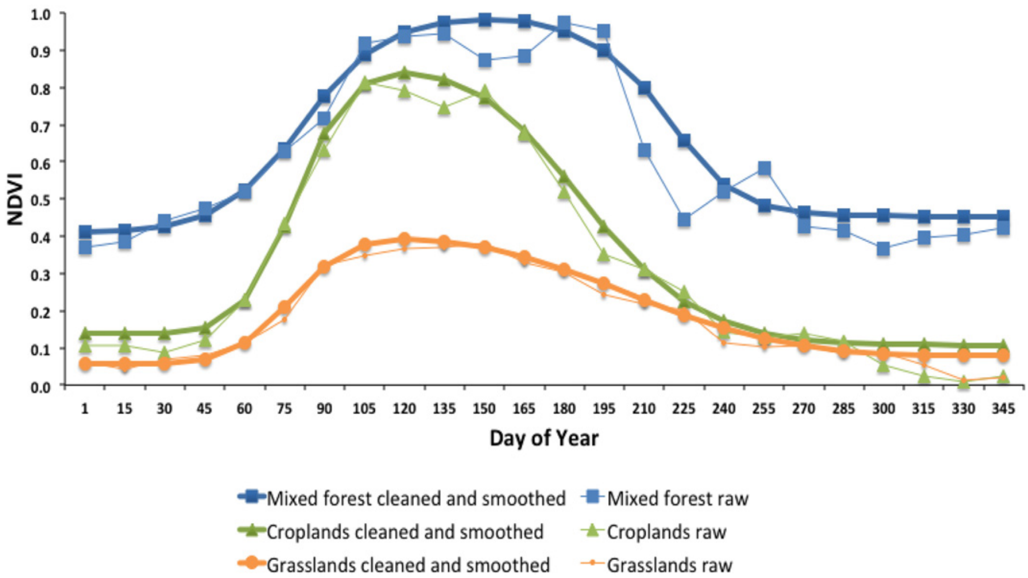

Fig. 3. Raw, and cleaned and smoothed NDVI time-series of mixed forest, croplands, and grasslands.

mostly found in Northeast China and Western China, as shown in white in Fig. 2 (a). The smoothing processes discussed above improves the completeness of the NDVI, reducing, but not entirely eliminating, the number of no data pixels, as shown in Fig. 2 (b). This is due to the limitation of TIMESAT, as mentioned above.

Fig. 3 displays the raw NDVI and cleaned and smoothed NDVI time-series for a single year of randomly selected individual pixels from Northeast China representing mixed forest, croplands, and grasslands. After cleaning and smoothing, the NDVI profiles provide generalized overall patterns of the NDVI time-series.

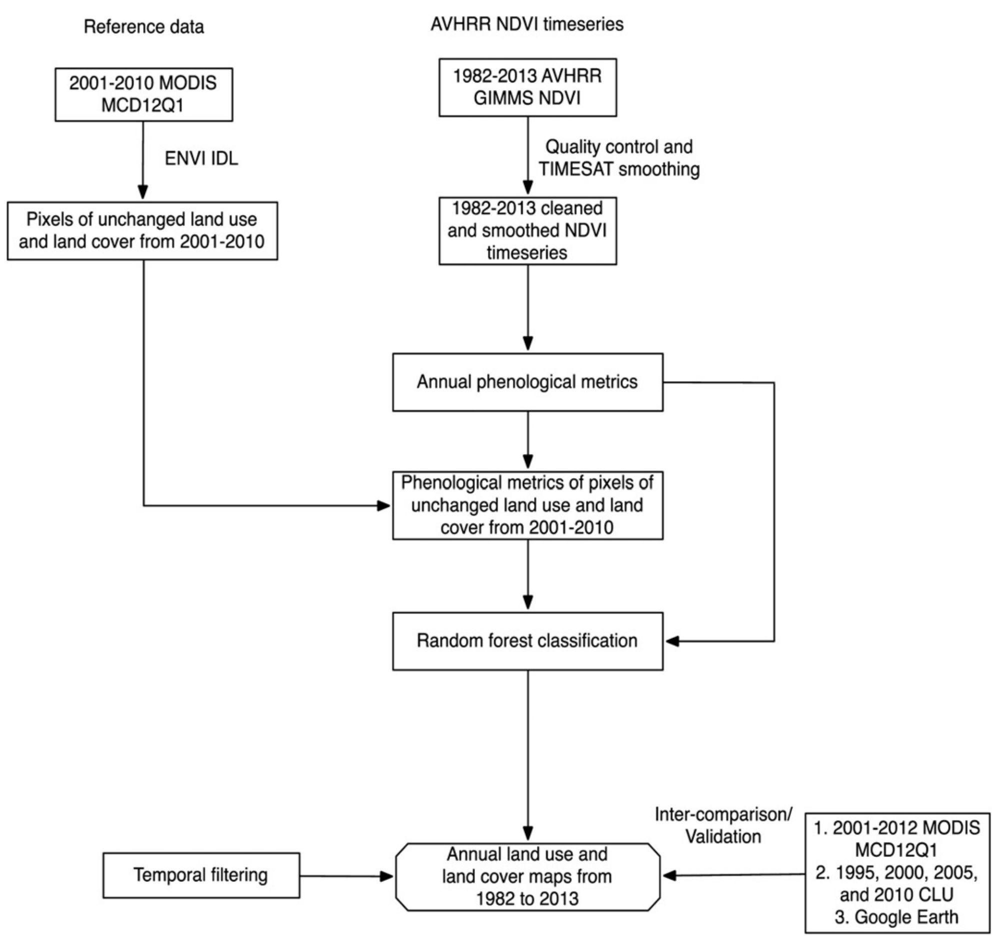

Fig. 4. Flowchart of the land use and land cover classification approach.

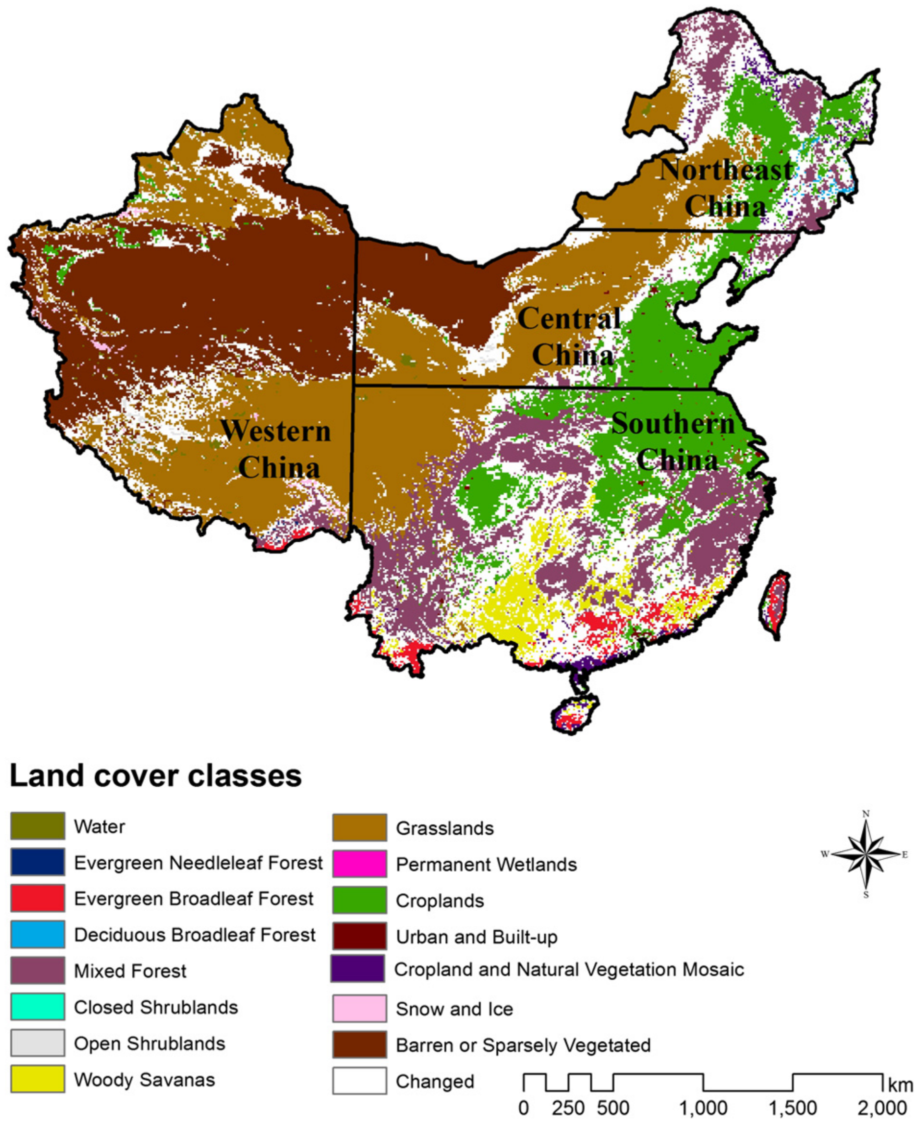

Fig. 5. Unchanged pixels of LULC for Mainland China, 2001 to 2010 derived from MODIS MCD12Q1, and used for training the random forest classifier.

Table 1
Number of pixels for each unchanged LULC type in each region in China. These pixels were randomly split, with 25% used for training, and 75% for validation.

MODIS MCD12Q1 class	Number of pixels				Class for AVHRR classification
	Western	Northeast	Central	Southern	
Water	313	60	135	260	Water
Evergreen needleleaf forest	20	0	0	0	Evergreen needleleaf forest
Evergreen broadleaf forest	141	0	0	1432	Evergreen broadleaf forest
Deciduous needleleaf forest	0	0	0	0	
Deciduous broadleaf forest	0	117	75	0	Deciduous broadleaf forest
Mixed forest	496	2702	1832	10,032	Mixed forest
Closed shrublands	0	0	2	0	
Open shrublands	210	0	43	0	Open shrublands
Woody savannas	0	1	0	2949	Woody savannas
Savannas	0	0	0	0	
Grasslands	16,062	5387	12,530	3628	Grasslands
Permanent wetlands	0	0	0	2	
Croplands	450	3947	7751	6177	Croplands
Urban and built-up	4	25	138	233	Urban and built-up
Cropland and natural vegetation mosaic	7	556	102	521	Cropland and natural vegetation mosaic
Snow and ice	432	0	1	5	Snow and ice
Barren or sparsely vegetated	23,881	0	5159	0	Barren or sparsely vegetated

Table 2
The 19 phenological metrics used as input for the random forest classification.

Phenological metrics
1: Maximum NDVI value
2: Minimum NDVI value
3: Julian day of maximum NDVI value
4: Julian day of minimum NDVI value
5: Integral of NDVI between Day 105 and Day 315
6: Integral under the NDVI curve
7: Maximum derivative of NDVI curve
8: Minimum derivative of NDVI curve
9: Julian day of maximum derivative of NDVI curve
10: Julian day of minimum derivative of NDVI curve
11: Julian day of start season
12: Julian day of end season
13: NDVI value of start season
14: NDVI value of end season
15: Integral between maximum derivative and minimum derivative
16: Integral between start season and maximum value
17: Integral between end season and maximum value
18: Maximum NDVI value - minimum NDVI value
19: Maximum NDVI value/integral under the NDVI curve

2.3. Land use and land cover classification methods

Pixels identified as having two growing seasons (as described in Section 2.2) were directly labeled as croplands, since natural vegetation should have only one growing season. The remaining pixels, which comprise those with one growing season, were classified following the procedures in Fig. 4.

2.3.1. Classification

Our aim in selecting a method for obtaining training data for the classification was to develop a method that provided many examples of each land cover class, from multiple years, in order to capture both geographic and temporal variability. We therefore used as a reference data source pixels of unchanged land cover during the period from 2001 to 2010 in the MODIS MCD12Q1 dataset (Fig. 4). The reference data were randomly split, with 25% used as training data (25%) and the remaining 75% used as validation data (75%) for an initial evaluation of the classification.

Prior to carrying out the classification, China was divided into four separate regions generally based on vegetation zones, as suggested by Hou (1981) (Fig. 5), and each region was classified separately. This segmentation was applied because the characteristics of phenological patterns of the mapping classes vary geographically. For example, croplands in Central and Southern China may have two growing seasons and spring green-up occurs typically in March, while croplands in Northeast China only have one growing season and green-up occurs only in May (Wu et al., 2010). Similarly, natural vegetation spring green-up occurs later, moving from south to north (Zhang et al., 2006). The boundaries of the sub-regions were chosen along arbitrary N-S and E-W lines, and not along previously mapped ecological boundaries. Our choice for doing so reflected our philosophical concern not to impose a simple view of sharp ecotonal boundaries that have not moved over the more than three decades of the study. An added advantage of using simple, relatively arbitrary sub-region boundaries is that it made for simpler processing, which could potentially easily be applied to the production of a global map, where the issue of identifying simple sharp ecotonal boundaries between regions would be even more problematic. A potential drawback with our approach as it perhaps increases the chance for classification inconsistencies across the sub-region boundaries.

Table 4
Class for comparing with CLU data.

Class for comparing with CLU data	Class for AVHRR classification
Water	Water
	Snow and ice
Forest	Evergreen needleleaf forest
	Evergreen broadleaf forest
	Deciduous broadleaf forest
	Mixed forest
	Open shrublands
	Woody savannas
Grasslands	Grasslands
Croplands	Croplands
	Cropland and natural vegetation mosaic
Urban and built-up	Urban and built-up
Barren or sparsely vegetated	Barren or sparsely vegetated

The four zones chosen were: Western China, Northeast China, Central China, and Southern China. In Western China, barren or sparsely vegetated land and grasslands dominate, with some minor croplands. In Northeast China, the main vegetation types are grasslands, croplands, and mixed forest. In Central China, the dominant classes are barren or sparsely vegetated, grasslands, and croplands. Southern China has the most diverse vegetation types, including grasslands, croplands, forest, and savanna (Fig. 5).

The numbers of pixels for each unchanged LULC type in each region are shown in Table 1. For some LULC types, the numbers of pixels are zero or close to zero. For example, there are no unchanged pixels in the deciduous needleleaf forest class and only two pixels of permanent wetlands and closed shrublands. These LULC types of limited extent were excluded from further analysis, reducing the original 17 classes down to 13 LULC classes that were mapped (Table 1, class for AVHRR classification).

The cleaned and smoothed 1982–2013 AVHRR GIMMS NDVI dataset (Section 2.2) was used to generate a total of 19 phenological metrics

Table 3
Allowed and disallowed class transitions.

Class		Year $n+1$									
		Water	Forest	Open shrublands	Woody savannas	Grasslands	Croplands	Urban and built-up	Cropland and natural vegetation mosaic	Snow and ice	Barren or sparsely vegetated
Year n and $n+2$	Water	Yes	No	No	No	No	No	No	No	No	No
	Forest	No	Yes[a]	No	No	No	No	No	No	No	No
	Open Shrublands	No	No	Yes	Yes	Yes	Yes	No	Yes	No	No
	Woody Savannas	No	No	Yes	Yes	Yes	Yes	No	Yes	No	No
	Grasslands	No	No	No	No	Yes	Yes	No	Yes	No	Yes
	Croplands	No	No	No	No	Yes	Yes	No	Yes	No	Yes
	Urban and Built-up	No	No	No	No	No	No	Yes	No	No	No
	Cropland and Natural Vegetation Mosaic	No	No	No	No	Yes	Yes	No	Yes	No	No
	Snow and Ice	No	No	No	No	No	No	No	No	Yes	No
	Barren or Sparsely Vegetated	No	No	No	No	Yes	No	No	No	No	Yes

[a] Transitions of the same forest type: “Yes”, transitions between different types of forest: “No”.

Table 5
Class for Google Earth validation.

Class for Google Earth validation	Class for AVHRR classification
Water	Water
Forest	Evergreen needleleaf forest
	Evergreen broadleaf forest
	Deciduous broadleaf forest
	Mixed forest
	Open shrublands
	Woody savannas
Grasslands	Grasslands
Croplands	Croplands
Urban and built-up	Urban and built-up
Cropland and natural vegetation mosaic	Cropland and natural vegetation mosaic
Snow and ice	Snow and ice
Barren or sparsely vegetated	Barren or sparsely vegetated

(Table 2), including start of growing season, end of growing season, and maximum and minimum NDVI values, for each pixel. Maximum and minimum NDVI value, Julian day of maximum and minimum NDVI value, NDVI value of start and end of season, and Julian day of start and end of the season are commonly used phenological variables for land cover characterization (Knight et al., 2006; Vuolo et al., 2011; Xue et al., 2014; Yan et al., 2015) because they capture the gross pattern of the annual NDVI cycle. For instance, in Fig. 3, the maximum NDVI value is greatest for mixed forest, followed by croplands, and then grasslands, which have the lowest value. The croplands in Fig. 3 are distinguished by an earlier green-up than mixed forest and grasslands. Additional key summary metrics were generated from integrated NDVI values. According to Liu et al. (2015) and Vuolo et al. (2011), integrated phenological metrics are important in classifying crops, thus we included several integrated values, such as the integral under the NDVI curve, and the integral between start season and maximum value. Other metrics, for example those based on the maximum and minimum derivative of the NDVI curve, were chosen in order to capture properties related to the rate and timing of phenological change, such as green-up (Alcantara et al., 2012; DeFries and Townshend, 1994; Klein et al., 2012; Nellis et al., 2009).

Classification was carried out using the R randomForest package (Liaw et al., 2009). Ensemble learning algorithms such as random forests have received increasing attention, because they are simple to implement, with few user-specified parameters, tend not to be sensitive to noise or overtraining, and therefore do not need pruning, and are generally found to be more robust than single classifiers (Gislason et al., 2006; Pal, 2005; Rodriguez-Galiano et al., 2012). The random forest classifier consists of a combination of a large number of classification trees, which "vote" to produce a single outcome for each pixel (Breiman, 2001). Each individual tree is generated from a random subset of the training data, as well as a random subset of the variables. In this way, the individual trees have reduced accuracy, but also reduced correlation, resulting in a more reliable overall classification. The random forest classifier can handle thousands of variables without variable deletion (Rodriguez-Galiano et al., 2012), and can even be applied when the number of variables is much larger than the number of samples (Dahinden, 2011). A further benefit is that the classifier provides an estimate of the importance of each variable by summarizing the accuracy of trees that don't use that variable. The random forest classifier (Breiman, 2001) has been widely used in many fields, including remote sensing (Baudron et al., 2013; Cutler et al., 2007; Maxwell and Warner, 2015; Maxwell et al., 2016; Speiser et al., 2015). Random forest classification requires two user-defined parameters: the number of decision trees produced (*ntree*) and the number of variables available for splitting at each node (*mtry*). In general, the value of *ntree* simply has to be large enough to give a stable result; we chose a value of 500 based on prior experience (Maxwell et al., 2016). For *mtry*, we chose the default value, in the randomForest package (Liaw et al., 2009), the square root of the number of predictor variables (i.e., 4), following Liu et al. (2016), though Shi and Yang (2016) advocate for a larger number of variables, combined with a smaller number of trees. The importance of the 19 phenological variables was measured using the mean decrease in accuracy (MDA) derived from the random forest classifier. The larger mean decrease in accuracy means the more the accuracy of the random forest decreased due to the exclusion of a variable, thus the greater the assumed importance of that variable (Breiman, 2001). Separate random forest classifications were generated to map the LULC for each of the four regions.

2.3.2. Temporal filtering

In order to try to improve the overall quality of the map time series, short-term, unreasonable land cover transitions were identified and suppressed (Baker et al., 2013; Clark et al., 2010). For example, it would be unlikely that a forested pixel would be converted to urban cover, and then subsequently changed back to forest cover in the following year. Therefore, we used a temporal filter with a 3-year moving window to remove the disallowed land use and land cover transitions (Clark et al., 2010). Specifically, we tested to see if the classes from year n and $n + 2$ were the same. If the classes were the same and class $n + 1$ was a disallowed transition as specified by Table 3, then class $n + 1$ was replaced with the class from year n.

2.3.3. Inter-comparisons and accuracy evaluation

To assess the reliability of the classified maps, we chose multiple approaches because the 32-year time series of land use and land cover

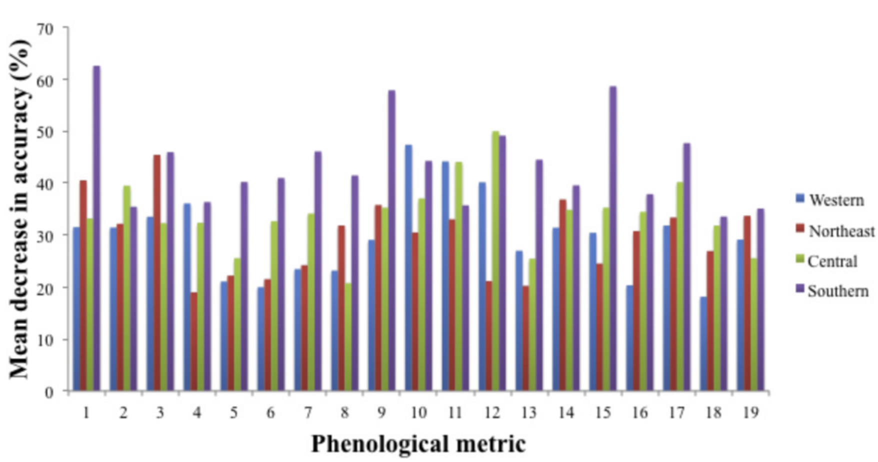

Fig. 6. Relative importance of 19 phenological metrics as indicated by mean decrease in accuracy (larger values indicate higher importance). (See Table 2 for associated metric for each metric number).

maps represent such a complex dataset. These included inter-comparison with the validation data (75%), the entire 2001–2010 MODIS MCD12Q1 LULC maps (years that were used in training the classifier), as well as the entire MODIS MCD12Q1 LULC maps for 2011 and 2012 (years that were excluded from the reference dataset), and the 1995, 2000, 2005, and 2010 CLU data. In addition, we undertook a more traditional error evaluation (Olofsson et al., 2014) using high resolution 2012 Google Earth images as a reference source.

We chose to compare our maps against the MODIS data in order to benchmark our approach against the input MODIS data. This comparison can potentially provide insight regarding how successful the AVHRR NDVI data are in reproducing the overall patterns as identified with MODIS, a sensor with superior spectral and radiometric resolution (Tucker et al., 2005). We compared our maps with CLU dataset, because it was a typical LULC dataset of China generated from high resolution imageries, such as Landsat and Huanjing − 1 data. The CLU dataset has a classification system that differs from ours, as mentioned in Section 2.1. Thus, we combined our classes of evergreen needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, mixed forest, open shrubland, and woody savannas, into a single forest class; water and snow and ice into the water class; and croplands and the cropland and natural vegetation mosaic into the croplands class (Table 4, see column

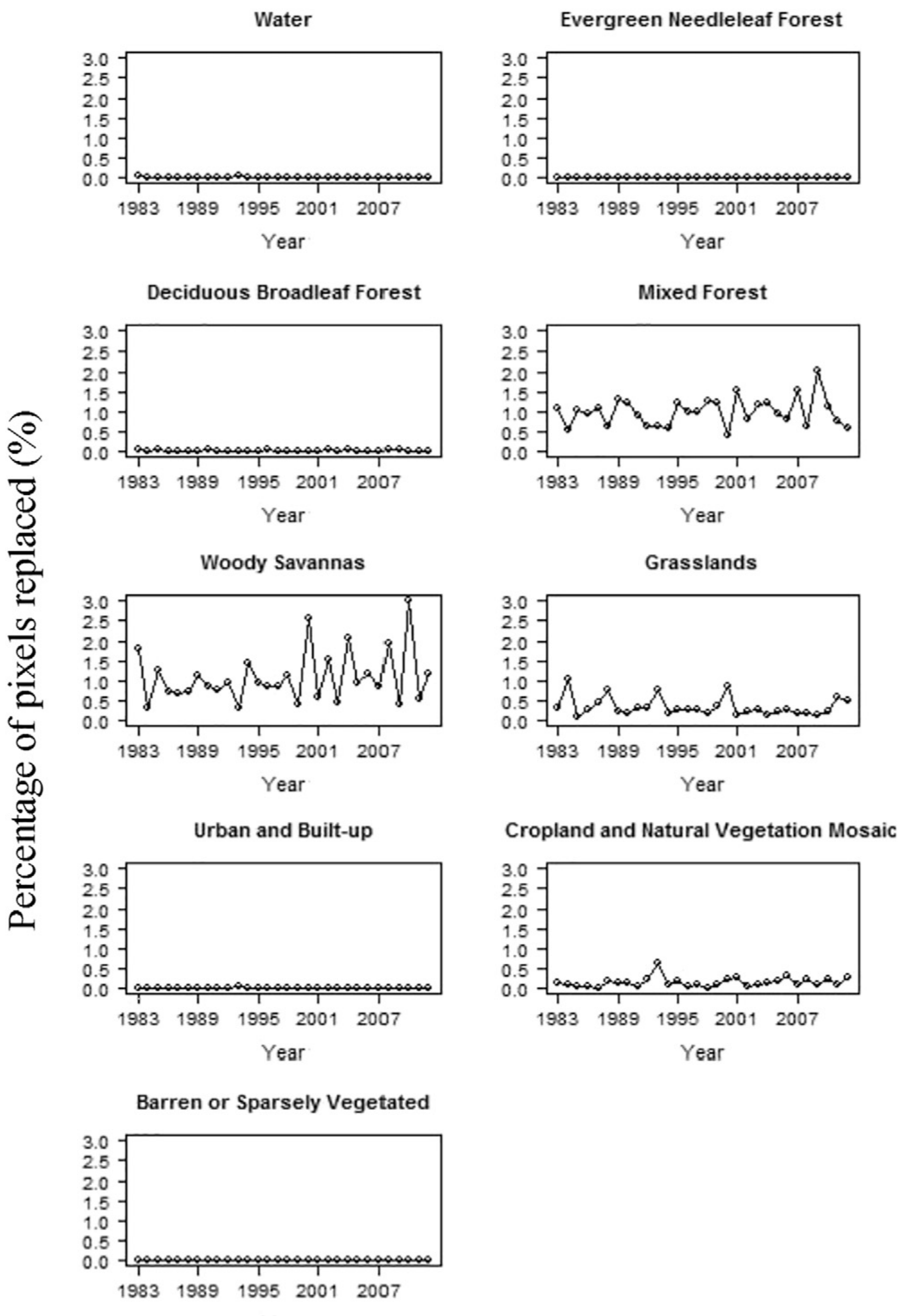

Fig. 7. Percentage of pixels for each class for which the filtering operation changed the labeled class, per year from 1983 to 2012 (Note: the filtering operation does not affect the first and last year data, i.e., 1982 and 2013).

"class for comparing with CLU data"). It is important to note, however, that the inter-comparisons with MODIS and CLU data are not an accuracy evaluation, since the MODIS and CLU data themselves have errors. For example, the global overall accuracy of MODIS MCD12Q1 has been estimated as approximately 75% (Zhao et al., 2013).

Comparisons with other land cover classifications provide useful insight into the similarities of the results with different sensors. However, only an assessment using independent reference data can provide an estimate of the map accuracy. We therefore undertook an accuracy evaluation using a manual interpretation of Google Earth imagery, focusing on land cover for the year 2012. This year was chosen because there were relatively abundant images in Google Earth for that year, and also this was a year not used in training the random forest classifier used to produce the AVHRR LULC maps.

A random sampling strategy with stratification was chosen to select sample points (Olofsson et al., 2014). The strata were the mapped classes. Based on multinomial sampling theory, we estimated a minimum of 150 random samples would be required in order to generate an estimate with 10% precision and 15% confidence (Jensen, 2016). An initial random sample of 300 sample points with a minimum of 15 points for each stratum was selected across the study area. Sample points for the final analysis after points without appropriate high resolution imagery for 2012 in Google Earth were removed, leaving a final total of 256 points, many more than the 150 points we set as a minimum. Although we had concerns that only using locations for which 2012 imagery was available might introduce bias, there was no obvious pattern to the availability of such imagery.

A visual estimate was made of the dominant land cover within a $1/12° \times 1/12°$ square, representing the AVHRR pixel dimensions, which was drawn around each sample point in Google Earth. Because it was not always possible to visually differentiate between all classes, we combined evergreen needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, mixed forest, open shrubland, and woody savannas to form a single forest class (Table 5, class for Google Earth validation). Thus, although the original map has 13 classes, the accuracy evaluation is based on only eight of those classes, and therefore the accuracy we estimated is for a simplified map that does not differentiate forest classes. The accuracy of the map with the original 13 classes will of course be lower than that of the eight classes map we evaluated.

Because the sampling design is stratified random using the map classes as strata, the cell entries of the error matrix are estimated using (Olofsson et al., 2014):

$$P_{ij} = W_i \frac{n_{ij}}{n_{i+}} \tag{1}$$

where P_{ij} denotes the proportion of area for the population that is class i according to the classification information, and class j according to the reference information. W_i is the proportion of area mapped as class i. n_{ij} is the number of samples in class i according to the classification, and class j according to the reference information. n_{i+} denotes the row totals.

Recent research has called into questioning the value of the kappa statistic (Pontius and Millones, 2011), consequently we instead calculated allocation disagreement and quantity disagreement, measuring which divide overall error into components related to errors in class location and proportion, respectively (Pontius and Millones, 2011). Based on the error matrix generated from Eq. (1), the overall allocation disagreement (A) and overall quantity disagreement (Q) were calculated as follows (Pontius and Millones, 2011; Warrens, 2015):

$$a_i = 2\ \min(p_{i+}, p_{+i}) - 2p_{ii} \tag{2}$$

$$A = \frac{1}{2}\sum_{i=1}^{C} a_i \tag{3}$$

$$q_i = |\ p_{i+} - p_{+i}\ | \tag{4}$$

$$Q = \frac{1}{2}\sum_{i=1}^{C} q_i \tag{5}$$

where C is the number of classes. p_{i+} and p_{+i} denote the row and column totals, respectively. a_i is the allocation disagreement for class i. q_i is the quantity disagreement for class i.

The same points were used to estimate the accuracy of our 2012 AVHRR map and the MODIS MCD12Q1 2012 map. McNemar's test (de Leeuw et al., 2006) was used to assess whether our classified LULC was significantly different from that of the MODIS classification.

3. Results and discussion

3.1. Relative importance of 19 phenological metrics

The importance of the 19 phenological metrics as predictors measured using the mean decrease in accuracy of the random forest classifier is showed in Fig. 6. It is notable that the lowest mean decrease in accuracy is approximately 18%, indicating all the metrics appear to be useful for all regions, and that there is not a great deal of redundancy in the 19 metrics. This finding is most evident for Southern China, where excluding any single metric seemed to have a particularly large effect (no <30% mean decrease in accuracy). The other major observation from Fig. 6 is that there is little consistency in the importance of individual metrics for the different regions of China. However, in general, the most important metrics are Julian dates of phenological events, such

Table 6
Random forest user's accuracy (UA), producer's accuracy (PA), and overall accuracy for each region based on the validation data (75%).

	Western China		Northeast China		Central China		Southern China	
	UA	PA	UA	PA	UA	PA	UA	PA
Water	44.4%	2.1%	66.7%	8.3%	75.0%	5.3%	25.0%	3.2%
Evergreen needleleaf forest	25.0%	7.1%						
Evergreen broadleaf forest	76.0%	51.4%					64.2%	46.4%
Deciduous broadleaf forest			42.6%	29.9%	85.7%	20.7%		
Mixed forest	76.5%	79.0%	89.0%	97.9%	85.0%	89.9%	79.1%	88.4%
Open shrublands	40.0%	3.8%			100.0%	3.6%		
Woody savannas							66.9%	58.1%
Grasslands	89.3%	92.9%	94.9%	97.2%	94.0%	96.6%	85.4%	86.5%
Croplands	71.5%	50.0%	91.7%	89.2%	92.9%	92.5%	85.3%	84.6%
Urban and built-up			50.0%	5.9%	25.0%	1.0%	52.4%	15.4%
Cropland and natural vegetation mosaic			62.8%	37.4%	63.0%	22.4%	55.4%	37.0%
Snow and ice	66.7%	3.5%						
Barren or sparsely vegetated	93.6%	94.9%			96.7%	95.5%		
Number of validation pixels	28,502		9518		19,653		16,518	
Overall accuracy	91.4%		91.5%		93.5%		79.0%	

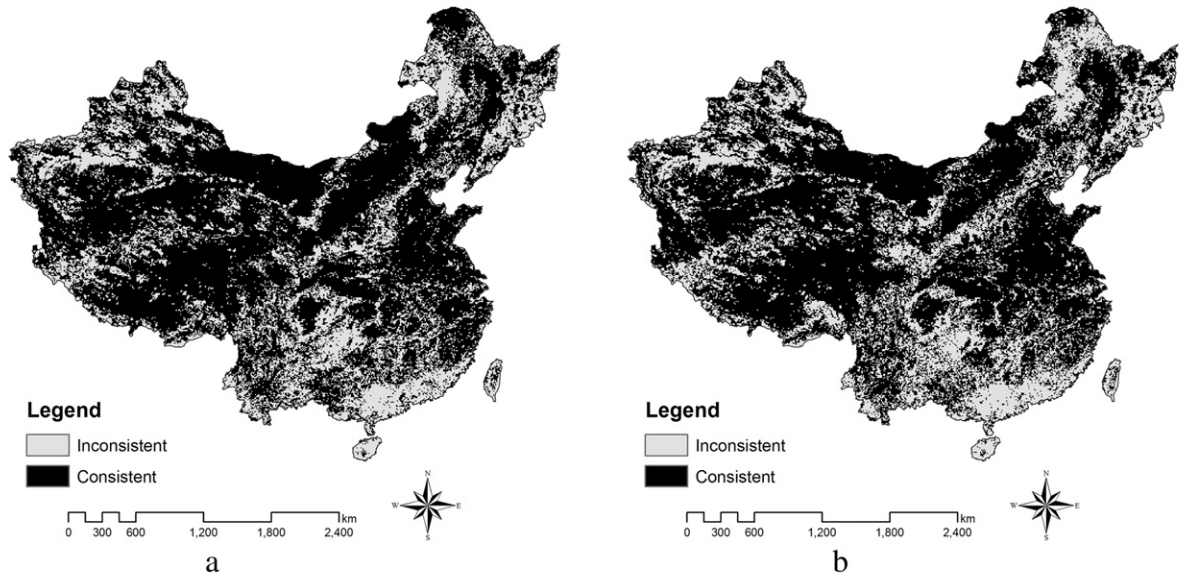

Fig. 8. Comparisons of the classified LULC maps with MODIS MCD12Q1 in (a) 2011 and (b) 2012.

as the start of season (metric number 11), probably due to differences in the growth calendars of different vegetation cover types. For example, croplands start green-up earlier than mixed forest and grasslands (Fig. 3). In Northeast China and Southern China, actual NDVI values are also important (e.g. the maximum NDVI value, metric number 1). As Northeast China and Southern China have mixed forest and other vegetation classes (Fig. 5), maximum NDVI may differentiate less productive non-forest biomes from more highly productive forests. In Central China, the integral of NDVI over time (e.g. between the maximum value and the end of season, metric number 17) is also important. Barren or sparsely vegetated areas, grasslands, and croplands are dominant in Central China (Fig. 5). Integrated NDVI values, such as integral between maximum value and the end of season, may separate croplands from other vegetation cover types, such as grasslands, since croplands are usually characterized by a high rate of green-up and senescence.

3.2. Temporal filtering of time-series of LULC maps

Comparing the classified maps before and after temporal filtering, the percentage of pixels replaced varies among different classes and years (Fig. 7). Water, evergreen needleleaf forest, deciduous broadleaf forest, open shrublands, urban and built-up, snow and ice, and barren or sparsely vegetated are replaced less than other classes, while mixed forest, woody savannas, grasslands, and croplands are replaced more frequently. This is likely due to the relatively large area of the latter classes.

3.3. Inter-comparison and accuracy evaluation

In this section, we first evaluate the reliability of the AVHRR classifications. This is done by a comparison with the validation data (75%), and comparisons with the entire 2001–2012 MODIS classifications and 1995, 2000, 2005, and 2010 CLU data. These inter-comparisons with the validation data (75%) and MODIS data are based on 13 classes (i.e., the column "class for AVHRR classification" in Table 1), the inter-comparison with the CLU data is based on six classes (i.e., the column "class for comparing with CLU data" in Table 4). After these comparisons, we then report the results of the more traditional accuracy evaluation, which is based on the eight classes in the column "class for Google Earth validation" in Table 5. After the accuracy evaluation we summarize the geographic and temporal trends in the 32-year time series in the following section. The focus of temporal trend analysis is on the areas of land cover classes in individual date and not change maps.

3.3.1. Inter-comparison of classified LULC with validation data (75%)

The user's accuracy and producer's accuracy vary among different classes and different regions (Table 6). It is apparent that the user's and producer's accuracies tend to be lower for most classes in Southern China compared to the other regions, possibly a result of cloud contamination. Some classes, for example, the barren or sparsely vegetated, grasslands, and mixed forest, are consistently mapped with relatively high accuracy (defined as here as >75% user's and producer's accuracies). In contrast, water and evergreen needleleaf forest are mapped generally (though not always) with lower reliability (user's and producer's accuracies <50%). It is notable that the classes with higher accuracies, such as mixed forest, tend to cover larger areas, and thus have larger number of training samples (Table 1). The overall accuracy of validation data (75%) is also shown in Table 6. All of regions have high accuracy (>91%), except for southern China, which has 79.0% overall accuracy.

3.3.2. Inter-comparison of classified LULC with MODIS MCD12Q1

The comparison of our classification with the entire MODIS MCD12Q1 maps for the years not used in training the classifier (2011 and 2012) indicates a consistency of 71.0% for 2011, and 69.3% for 2012. Inconsistency is high in Northeast China and Southern China (Fig. 8), whereas the two datasets are generally much more consistent in Western China. It is notable that the areas of inconsistent land cover are common where cloud cover is more frequent, including

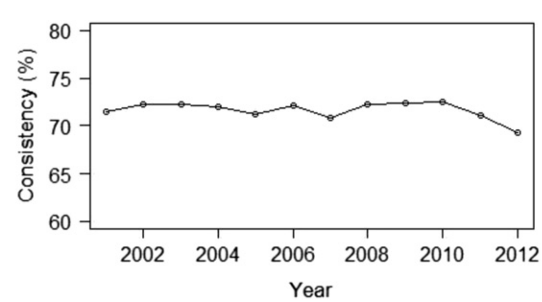

Fig. 9. Consistency between classified LULC maps and MODIS MCD12Q1 from 2001 to 2012.

parts of the humid Southeastern China, and some of the relatively mountainous regions of Northeastern and Western China, as well as the places where land use and land cover change appears to be more common. However, for the Tibetan Plateau, inconsistent regions are more common on the edges of the region than the interior. This may be due to the fact that the interiors of region tend to be dominated by more homogenous LULC (e.g., grasslands), while the edges of the region are transitional areas for different LULC types. For example, in Southeast edge of Tibetan Plateau, mixed forest and grasslands coexist. Furthermore, the inconsistency between these two land use and land cover datasets may be also related to the different data sources (i.e., NDVI in this study and reflectance of bands 1–7 in MCD12Q1) and methodologies (i.e., random forest in this study and ensemble supervised classification algorithm in MCD12Q1) for producing these two datasets.

Notably, the consistency values for 2011 and 2012 are only slightly lower than the average of the consistency values observed for 2001–2010, years that were used for training the classifier (Fig. 9). This result is encouraging, because it provides some evidence that the classifier is able to extrapolate to years other than those used in training. If the training data, or the classifier, were not adequate to capture the overall patterns, we would expect a much greater drop in years not used in training, when the annual patterns of rainfall or temperature, for example, might be slightly different than the years used for training.

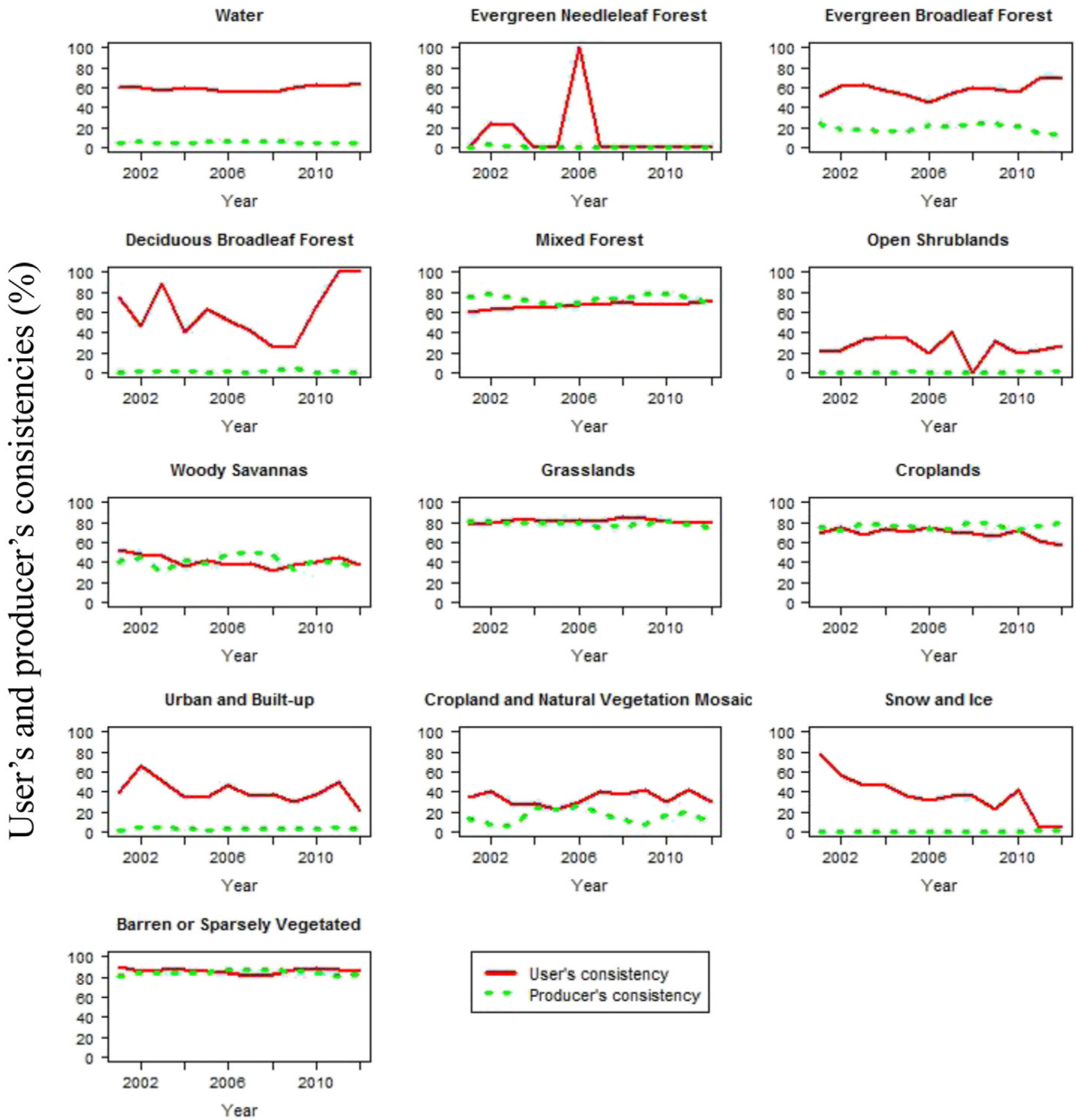

Fig. 10. User's and producer's consistencies for each class between classified LULC maps and MODIS MCD12Q1 from 2001 to 2012.

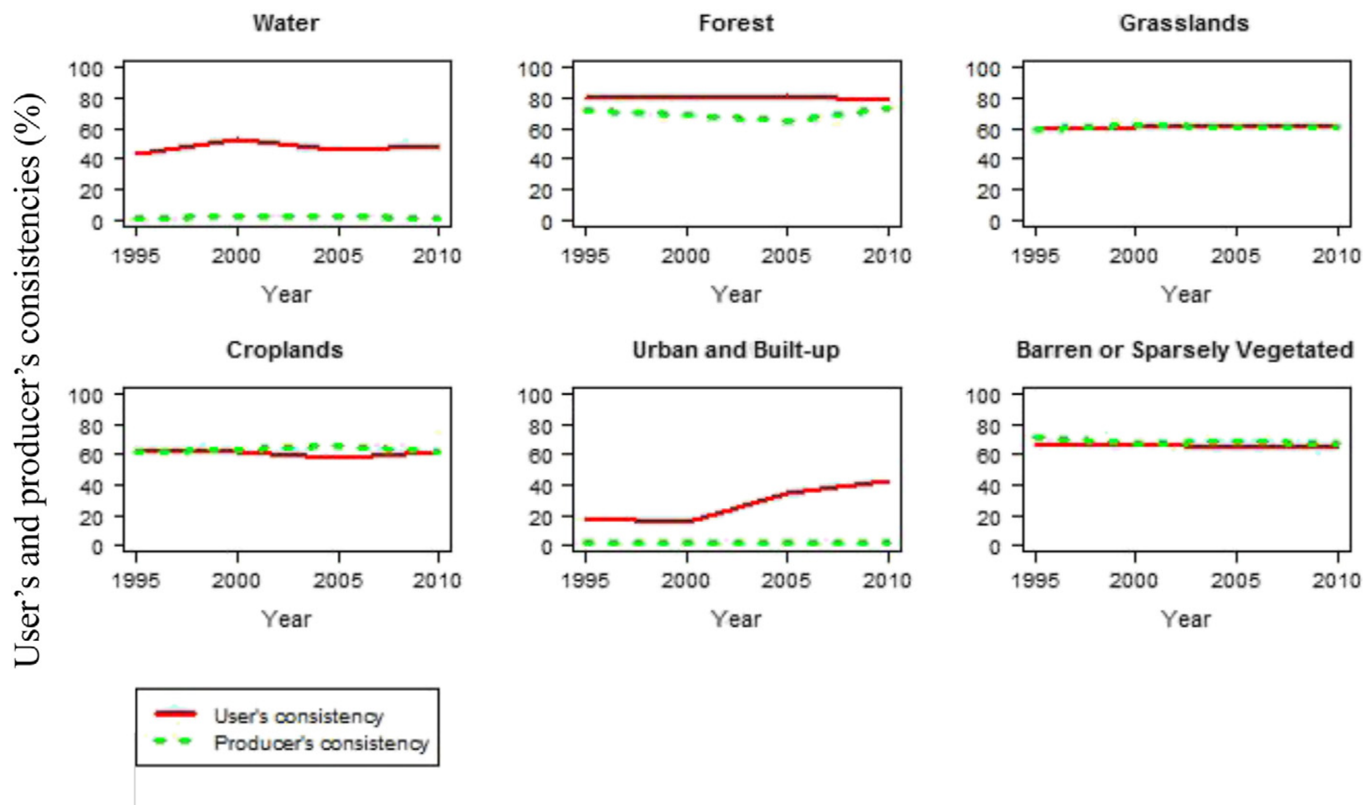

Fig. 11. User's and producer's consistencies for each class between classified LULC and CLU maps for 1995, 2000, 2005, and 2010.

To further explore the consistency for different classes between these two LULC datasets, we display a time-series of user's consistency (calculated as the consistent pixels for class *i*/all pixels for class *i* in our classified map) and producer's consistency (calculated as the consistent pixels for class *i*/all pixels for class *i* in MODIS data) for each class from 2001 to 2012 (Fig. 10). The classes with the highest user's and producer's consistencies for each year tend to be those of mixed forest, woody savannas, grasslands, croplands, and barren or sparsely vegetated. This may be due to the better performance of the random forest classification for the classes with a larger number of training samples (Table 1 and Table 6). The classes with the lowest consistencies are water, evergreen needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, open shrublands, urban and built-up, and snow and ice.

3.3.3. Inter-comparison of classified LULC with CLU dataset

The overall consistency values between our classified LULC and the CLU dataset, for years 1995, 2000, 2005, and 2010, are 64.3%, 64.3%, 63.0%, and 64.4%, respectively. Compared to the consistency values of our classified LULC maps and MODIS MCD12Q1, the consistency values for the CLU maps are generally lower. The very different spatial resolutions (i.e., 30 m for original CLU maps and 1/12° for our classified LULC maps) and different classification systems may have contributed to the lower consistency values.

Fig. 11 summarizes the user's and producer's consistencies for our classified LULC and CLU maps for each class for 1995, 2000, 2005, and 2010, calculated as in Fig. 10. Forest, grasslands, croplands, and barren or sparsely vegetated classes have relatively higher consistencies, while water and urban and built-up have lower consistencies.

3.3.4. Accuracy evaluation of the 2012 classification using Google earth imagery

The error matrix for the accuracy assessment of the 2012 AVHRR map, using the visual interpretation of eight classes from 256 samples of Google Earth images from 2012 as reference data, is shown in Table

Table 7
Error matrix of Google Earth data and classified LULC.

		Reference LULC (from Google Earth interpretation)									
		Water	Forest	Grasslands	Croplands	Urban and built-up	Cropland and natural vegetation mosaic	Snow and Ice	Barren or sparsely vegetated	Total	User's accuracy
Classified LULC	Water	0.00049	0.00000	0.00000	0.00000	0.00004	0.00004	0.00000	0.00000	0.00057	86%
	Forest	0.00000	0.15101	0.00000	0.04066	0.00000	0.02323	0.00000	0.00000	0.21491	70%
	Grasslands	0.00000	0.01921	0.19211	0.07684	0.00000	0.00480	0.00000	0.01441	0.30737	63%
	Croplands	0.00000	0.03800	0.01900	0.16623	0.00475	0.01425	0.00000	0.00000	0.24223	69%
	Urban and built-up	0.00000	0.00011	0.00004	0.00008	0.00015	0.00011	0.00000	0.00000	0.00049	31%
	Cropland and natural vegetation mosaic	0.00000	0.00000	0.00000	0.00082	0.00000	0.00905	0.00000	0.00000	0.00988	92%
	Snow and Ice	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00053	0.00177	0.00230	23%
	Barren or sparsely vegetated	0.00000	0.00000	0.00427	0.00000	0.00000	0.00000	0.00000	0.21799	0.22226	98%
	Total	0.00049	0.20833	0.21542	0.28463	0.00494	0.05149	0.00053	0.23417	1.00000	
	Producer's accuracy	100%	72%	89%	58%	3%	18%	100%	93%		73.8%

Table 8
Error matrix of Google Earth data and MODIS LULC.

		Reference LULC (from Google Earth interpretation)								
		Water	Forest	Grasslands	Croplands	Urban and built-up	Cropland and natural vegetation mosaic	Barren or sparsely vegetated	Total	User's accuracy
MODIS LULC	Water	0.00619	0.00000	0.00000	0.00000	0.00077	0.00000	0.00000	0.00696	89%
	Forest	0.00000	0.15929	0.01493	0.02489	0.00000	0.04480	0.00498	0.24889	64%
	Grasslands	0.00441	0.03531	0.18098	0.05738	0.00000	0.02207	0.01766	0.31782	57%
	Croplands	0.01035	0.00000	0.00345	0.13106	0.01035	0.01380	0.00000	0.16900	78%
	Urban and built-up	0.00000	0.00000	0.00000	0.00119	0.00238	0.00000	0.00000	0.00358	67%
	Cropland and natural vegetation mosaic	0.00000	0.00332	0.00000	0.00996	0.00000	0.01661	0.00000	0.02989	56%
	Barren or sparsely vegetated	0.00000	0.00000	0.00367	0.00367	0.00000	0.00000	0.21652	0.22386	97%
	Total	0.02095	0.19793	0.20303	0.22816	0.01350	0.09728	0.23915	1.00000	
	Producer's accuracy	30%	80%	89%	57%	18%	17%	91%		71.3%

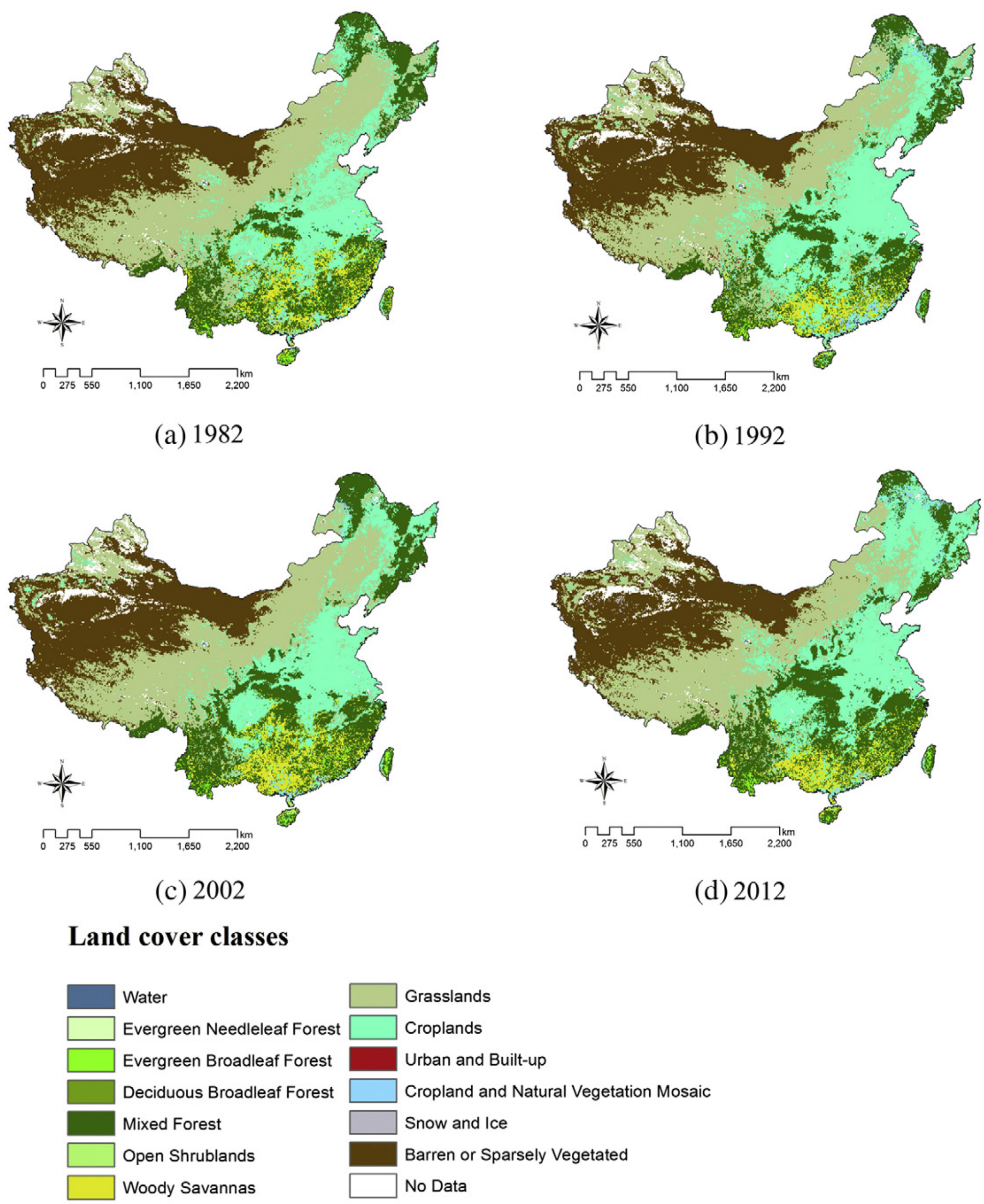

Fig. 12. Annual LULC maps of China, produced by random forest classification. (a) 1982, (b) 1992, (c) 2002, and (d) 2012.

7. The overall accuracy is 73.8%. The producer's and user's accuracies for urban and built-up, as well as producer's accuracy for cropland and natural vegetation mosaic and user's accuracy for snow and ice, are low, but these estimates are based on a relatively small number of samples, and thus the uncertainty in these estimates is comparatively large. Another possible reason for the low accuracy of urban and built-up class may be the coarse spatial resolution of GIMMS NDVI data. As mentioned in Liu et al. (2003c) and Loveland et al. (2000), the urban class may easily be confused with other classes due to the complex mixtures of surface materials within each pixel. The low accuracy for cropland and natural vegetation mosaic and snow and ice may be due to the similarity of phenological characteristics with other classes (i.e., croplands and barren or sparsely vegetated, respectively). For instance, Guan et al. (2014) noted that in Africa, most croplands were fragmented and mixed with natural savannas, resulting in generally similar phenology patterns. The barren or sparsely vegetated class, which has a distinctive low NDVI year-round, has user's and producer's accuracies >90%. Overall, the allocation disagreement (16.2%) is higher than the quantity disagreement (10.0%), which indicates the disagreement between our classified LULC map and Google Earth reference data is mainly due to errors in the location of the mapped classes, rather than their proportions in the map. Improving the quality of the training data might improve the spatial consistency of the map and thus improve the overall accuracy of our classified map.

Using the same samples from Google Earth as reference data, the MODIS MCD12Q1 2012 error matrix is shown in Table 8. Because there is no snow and ice class for the sample points in 2012 MODIS data, it is impossible to calculate the proportion of area P_{ij} for snow and ice in error matrix. Therefor, we deleted the snow and ice class, remaining seven classes with 253 sample points for validating 2012 MODIS MCD12Q1. As with our 2012 AVHRR map, the MODIS producer's accuracies for the urban and built up land, as well as cropland and natural vegetation mosaic classes, are very low. The allocation disagreement and quantity disagreement are 12.1% and 16.6%, respectively, indicating the disagreement between MODIS map and Google Earth reference data is mainly due to errors in the proportions in the map. The overall accuracy of the 2012 MODIS LULC for the seven classes is estimated as 71.3%, 2.5% lower than the accuracy of our classified LULC. However, the McNemar's statistic (de Leeuw et al., 2006) based on the comparison of the two accuracy assessments is approximately 0.4, indicating that this difference is not statistically significant at the 95% confidence level.

3.4. Spatial patterns and temporal trends of annual LULC in China

Example classifications for 1982, 1992, 2002, and 2012 are shown in Fig. 12. It is worth noting that the classification was carried out in four separate regions, with arbitrary boundaries (Fig. 5). Nevertheless, a close examination of the final classifications (Fig. 12) indicates no obvious evidence of artifacts or errors across these boundary lines. The maps show that in Western China, LULC appears broadly similar during the period 1982 to 2012. Croplands in Xinjiang province increase a little from 1982 to 2002, and then decrease from 2002 to 2012. The initial increase in croplands in Xinjiang may be attributed to the successful promotion of modern agronomic technology (Yin, 2008), whilst the recent decrease, also observed by Liu et al. (2008), may reflect the conversion of croplands to built-up land, associated with the policy of increased Western China development. In Northeast China croplands increase notably during the three decades, while grasslands and mixed forest decrease. In Central China, croplands first increase from 1982 to 1992, then decease along the upper reach of Yellow River basin from 1992 to 2002, followed by an increasing trend during the last decade. In contrast, grasslands in Central China decrease from 1982 to 1992, followed by an increasing trend from 1992 to 2002 and a decreasing trend from 2002 to 2012. The recent increase in croplands, and reduction in other classes, is supported by the observations of Xu et al. (2015) in Central China, who noted, in a study that focused on the period since 2000, a conversion of wetlands, barren areas, and woody shrubland to croplands. In Southern China, croplands area decreases during the entire period from 1982 to 2012. Mixed forest and evergreen broadleaf forest increase, especially along the Yangtze river and tropical regions, such as the south part of Yunnan province. Woody savannas first increase from 1982 to 2002, then decrease during the last decade. The overall changing patterns in croplands documented in these maps, an increase in Northern China and decrease in Southern China, are broadly consistent with Liu et al. (2014). However, although a thorough analysis of the underlying reasons for different patterns of LULC over different regions in China is needed, this is beyond the scope of this paper.

Based on abovementioned analyses, croplands, forest, and grasslands show the clear spatial change patterns. In Fig. 13, we show the

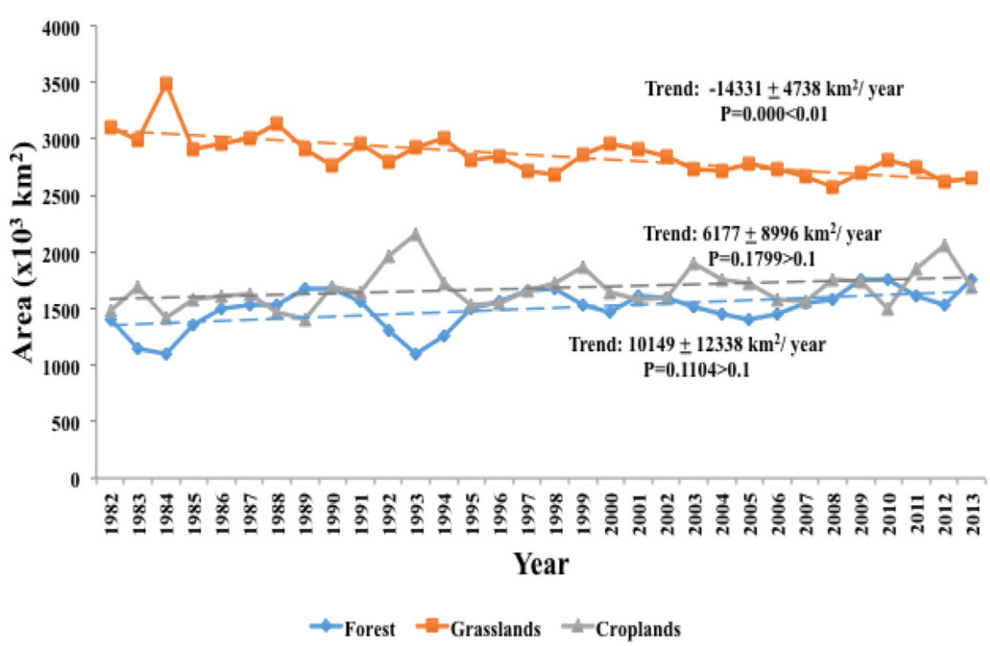

Fig. 13. Temporal changes in area of grasslands, croplands, and forest classes from 1982 to 2013. Dotted lines represent overall trend.

overall trends of these three classes across the entire 32 years of the study for further analysis. In here, the forest class means evergreen needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, and mixed forest in the column "class for AVHRR classification" in Table 1. We used the R nlme package (Pinheiro et al., 2014) to conduct the linear regression trend analysis accounting for temporal autocorrelation. The significance of the trend was tested by Student's *t*-test. The overall trend for grasslands is a significant decrease at the 1% level. This is consistent with Liu et al. (2014), who used Landsat TM/ETM + data acquired in intervals of five years to explore LULCC in China since the late 1980s. The overall trend for forest is increasing, which may be due to afforestation efforts, such as the "Grain for Green" and "Three-North Shelterbelt" projects (Liu et al., 2014). The overall trend of croplands is an increase during the three decades. However, the standard errors of the trends for forest and croplands are larger than that of grasslands, and thus the uncertainty of these trends may be bigger. The trends for both forest and croplands are not significant at the 10% level. While the overall trends of these three classes are generally consistent with previous studies, there are some short-term variations in the three times series that seem inconsistently large, such as the large temporary increase in grasslands in 1984 as well as the decline in forest and increase in croplands in 1993. The latter short-term anomaly was also identified by He and Shi (2015), but the reasons for these variations do need to be explored further. Absent an obvious physical explanation, such as unusual weather patterns, or other national policy changes or pressures (Liu et al., 2003b), our assumption is that these variations may be artifacts, such as the result of temporal inconsistent variations of GIMMS NDVI time-series due to sensor drift. Previous studies have revealed the temporal inconsistency of GIMMS NDVI dataset (Detsch et al., 2016; Fensholt et al., 2009; Fensholt and Proud, 2012; Tian et al., 2015). For example, Tian et al. (2015) found that the abrupt increase of NDVI around 1994 coincided with the sensor shift from NOAA-11 to NOAA-9. The coarse spatial and temporal resolution of GIMMS dataset may also contribute to the artifacts. Alcaraz-segura et al. (2010) stated that GMMMS NDVI dataset failed to capture long-term ecosystem changes in some places, such as central Canada, while they were evident by using higher spatial resolution NDVI datasets, such as Canadian Centre for Remote Sensing (CCRS) NDVI dataset.

4. Conclusions

Both observational and modeling studies have shown that LULCC can significantly affect the climate system (Takata et al., 2009; Webster, 1987). This may happen through biogeophysical (changes in water and energy balance) and biogeochemical (changes in CO_2 and methane) processes that modify surface wetness, partition surface energy between sensible and latent heat fluxes, alter roughness of the land surface, and change terrestrial carbon storage (Foley et al., 2003; McPherson, 2007). China has experienced extensive LULCC, including cropland expansion, desertification, deforestation, afforestation, and urbanization (Ge et al., 2004; Houghton and Hackler, 2003; Lin and Ho, 2003; Liu et al., 2005a; Liu et al., 2005b). However, due to the limited period for which time-series of annual land use and land cover maps are available, LULCC information in China has been normally used in climate modeling in only a simplified manner.

To address this need for an extended time-series of LULC data, we constructed a three-decade continuous time series of annual land use and land cover maps of China from 1982 to 2013 using AVHRR GIMMS NDVI3g data. The reference data for training the classifier was a 25% sample of the pixels of constant LULC class in the MODIS MCD12Q1 annual land cover maps from 2001 to 2010. Classes for which the number of reference pixels were zero, or close to zero, were excluded, reducing the number of classes from 17 to 13 (Table 1). 19 phenological features were derived from the AVHRR data, and used as attributes in the random forest classification.

Based on the validation data (75%), the overall accuracy of the AVHRR classification for 2001 to 2010 was > 91% for each region, except Southern China, for which it was 79%. This result is strong evidence that the AVHRR phenological features are broadly able to differentiate the different MODIS LULC classes. The higher user's accuracy and producer's accuracy for mixed forest, grasslands, croplands, and barren or sparsely vegetated in all of the four regions indicate the performance of random forest classifier is better in separating classes with larger reference data.

In the comparison with the MODIS classification across all of China for 2011 and 2012, years not used in training the classifier, consistency was 71.0% and 69.3%, respectively. It is not surprising that these numbers are lower than the accuracies for the unchanged pixels in the 2001 to 2010 reference data, since the latter data are presumably mostly pixels with consistent land cover with relatively distinct remote sensing spectral characteristics. Furthermore, we would expect slightly lower accuracy of land cover classification in years not used for training, since the climatic variations and thus phenological patterns from 2001 to 2010 may not have captured the entire range of possible conditions. However, comparisons of our AVHRR classification and the MODIS MCD12Q1 data from 2001 to 2010 (years that were used for training) show only 0–3% improvement, indicating that this effect is small. The relatively low consistency for Southern China may be due to extensive clouds in the relatively humid Southern China. For instance, An et al. (2015) found that poor relationships between MODIS and SPOT NDVI datasets in Southern China may be attributed to greater cloud cover in that area. In summary, we regard the broad consistency with MODIS data as one line of evidence of the ability of the classifier to be extended over time. The comparisons between our classified maps and CLU data show the lower consistency values, ranging from 63.0% to 64.4%. It is in our expectation, because converting the very different spatial resolutions (i.e., 30 m for original CLU and 1/12° for our classified LULC) and classification systems may induce additional inconsistencies.

An overall accuracy assessment of the AVHRR classification was carried for 2012, a year not used for training. The reference data were derived from a visual interpretation of stratified random samples of Google Earth imagery. This accuracy assessment combined all the forest classes, resulting in a simplified map with just eight classes. The overall accuracy of this eight-class map was 73.8%. In comparison, the MODIS MCD12Q1 product, which has been estimated to have a global accuracy of approximately 75% (Zhao et al., 2013), was found to have an accuracy of 71.3% for China in 2012. The McNemar's test indicated no significant difference between the MODIS LULC and AVHRR LULC accuracies. We regard this finding as particularly notable. MODIS is a sensor with greater spectral and radiometric resolution than AVHRR (Tucker et al., 2005), and thus achieving an accuracy with AVHRR that is similar to the MODIS product is encouraging. On the hand, it is important to acknowledge that the MODIS MD12Q1 is generated through a global classification, whereas our classification is based on four regional classifications, a much simpler mapping task.

Several areas of future work seem promising. First, using as reference data areas in the MODIS MD12Q1 maps that do not change over an extended period of time, an approach previously also successfully employed by others, including Klein et al. (2012) and Wohlfart et al. (2016), provides a simple approach that can easily be applied to other regions. In particular, we plan to investigate ways to scale the method up to a global approach, since having annual land cover maps of the entire world for > 30 years could be particularly valuable to the climate modeling community. A second line of research would be to consider alternative methods for generating the reference data. Although the MODIS data provide a very effective method for selecting large numbers of training samples over multiple years, the reliability of the MODIS data is not that high (Zhao et al., 2013). Moreover, since pixels of some unchanged classes are not present in some of the regions in the training data (e.g., deciduous broadleaf forest in Western and Southern China) (Table 1), the final classification maps in these regions do not have these classes. In addition, the reference data does not distribute evenly across different classes. Some classes have larger numbers of reference data, such as grasslands, while other classes have much less reference data, such as water (Table 1). These influence the classification accuracies of the classes (Table 6). It is possible that if we could generate more reliable training data, we might improve the accuracy of the AVHRR classification. The final area of possible future research would be to establish statistically robust approaches to investigate the time series of LULC maps for its potential to produce change maps. This would require an evaluation of the accuracy of the change information itself, rather than only an evaluation of the accuracy of the individual dates, as we have done in this study.

In summary, this study generated annual land use and land cover maps in China from 1982 to 2013 using AVHRR GIMMS NDVI3g data. The overall accuracy from random forest classifier was high for all of the regions, except for Southern China. Based on a comparison of visual interpretation of images from Google Earth, the overall accuracy of the simplified eight-class LULC map was 73.8%, which was not statistically different from that of the simplified seven-class MODIS MCD12Q1 LULC map (71.3%). Based on temporal evolution of areas for forest, grasslands, and croplands during the last three decades, the overall trend was consistent with previous studies (He and Shi, 2015; Liu et al., 2014). These thirty-two years of annual maps of land cover will be an important dataset for quantifying the associations of recent LULCC with changes in the regional climate systems in East Asia, and the preprocessing, classification, and validation methods used in this study could be applied to other geographical regions where the availability of continuous LULC maps is limited.

Acknowledgement

We thank NASA for providing the GIMMS NDVI3g and MODIS MCD12Q1 data. We also appreciate Dr. Aaron Maxwell of West Virginia University for providing help with the random forest R package. We are grateful to three anonymous reviewers and Dr. Emilio Chuvieco for their constructive and helpful comments, which greatly strengthened this manuscript.

References

Alcantara, C., Kuemmerle, T., Prishchepov, A.V., Radeloff, V.C., 2012. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sens. Environ. 124, 334–347.
Alcaraz-segura, D., Chuvieco, E., Epstein, H.E., Kasischke, E.S., Trishchenko, A., 2010. Debating the greening vs. browning of the North American boreal forest: differences between satellite datasets. Glob. Chang. Biol. 16, 760–770.
An, Y., Gao, W., Gao, Z., Liu, C., Shi, R., 2015. Trend analysis for evaluating the consistency of Terra MODIS and SPOT VGT NDVI time series products in China. Front. Earth Sci. 9, 125–136.
Andres, L., Salas, W.A., Skole, D., 1994. Fourier analysis of multi-temporal AVHRR data applied to a land cover classification. Int. J. Remote Sens. 15, 1115–1121.
Baker, B.A., Warner, T.A., Conley, J.F., McNeil, B.E., 2013. Does spatial resolution matter? A multi-scale comparison of object-based and pixel-based methods for detecting change associated with gas well drilling operations. Int. J. Remote Sens. 34, 1633–1651.
Baudron, P., Alonso-Sarría, F., García-Aróstegui, J.L., Cánovas-García, F., Martínez-Vicente, D., Moreno-Brotóns, J., 2013. Identifying the origin of groundwater samples in a multi-layer aquifer system with random Forest classification. J. Hydrol. 499, 303–315.
Bonan, G.B., Pollard, D., Thompson, S.L., 1992. Effects of boreal forest vegetation on global climate. Nature 359, 716–718.
Boyd, D.S., Almond, S., Dash, J., Curran, P.J., Hill, R.A., 2011. Phenology of vegetation in Southern England from Envisat MERIS terrestrial chlorophyll index (MTCI) data. Int. J. Remote Sens. 32, 8421–8447.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Channan, S., Collins, K., Emanuel, W., 2014. Global Mosaics of the Standard MODIS Land Cover Type Data. University of Maryland and the Pacific Northwest National Laboratory, College Park, Maryland, USA.
Chen, T., Wang, G., Yuan, W., Li, A., Liu, Y.Y., 2016. Asymmetric NDVI trends of the two cropping seasons in the Huai River basin. Remote Sens. Lett. 7, 61–70.
Clark, M.L., Aide, T.M., Grau, H.R., Riner, G., 2010. A scalable approach to mapping annual land cover at 250 m using MODIS time series data: a case study in the Dry Chaco ecoregion of South America. Remote Sens. Environ. 114, 2816–2832.
Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I., 1990. STL: a seasonal-trend decomposition procedure based on loess. J. Off. Stat. 6, 3–73.
Cutler, D.R., Edwards Jr., T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.
Dahinden, C., 2011. An improved random forest approach with application to the performance prediction challenge datasets. Hands-on Pattern Recognition, Challenges in Machine Learning. 1, pp. 223–230.
Dappen, P., 2003. Using Satellite Imagery to Estimate Irrigated Land: A Case Study in Scotts Bluff and Kearney Counties. Center of Advanced Land Management Information Techonologies. University of Nebraska-Lincoln, NE,USA.
DeFries, R., Townshend, J., 1994. NDVI-derived land cover classifications at a global scale. Int. J. Remote Sens. 15, 3567–3586.
DeFries, R., Hansen, M., Townshend, J., 1995. Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. Remote Sens. Environ. 54, 209–222.
Detsch, F., Otte, I., Appelhans, T., Hemp, A., Nauss, T., 2016. Seasonal and long-term vegetation dynamics from 1-km GIMMS-based NDVI time series at Mt. Kilimanjaro, Tanzania. Remote Sens. Environ. 178, 70–83.
Douglas, E.M., Niyogi, D., Frolking, S., Yeluripati, J., Pielke, R.A., Niyogi, N., Vörösmarty, C., Mohanty, U., 2006. Changes in moisture and energy fluxes due to agricultural land use and irrigation in the Indian Monsoon Belt. Geophys. Res. Lett. 33, L14403.
Eklundh, L., Jönsson, P., 2015. TIMESAT 3.2 with Parallel Processing Software Manual. http://web.nateko.lu.se/timesat/docs/TIMESAT32_software_manual.pdf Accessed 16. 05. 02.
Eltahir, E.A., 1996. Role of vegetation in sustaining large-scale atmospheric circulations in the tropics. J. Geophys. Res. Atmos. 101, 4255–4268.
Fensholt, R., Proud, S.R., 2012. Evaluation of earth observation based global long term vegetation trends—comparing GIMMS and MODIS global NDVI time series. Remote Sens. Environ. 119, 131–147.
Fensholt, R., Rasmussen, K., Nielsen, T.T., Mbow, C., 2009. Evaluation of earth observation based long term vegetation trends—Intercomparing NDVI time series trend analysis consistency of Sahel from AVHRR GIMMS, Terra MODIS and SPOT VGT data. Remote Sens. Environ. 113, 1886–1898.
Foley, J.A., Costa, M.H., Delire, C., Ramankutty, N., Snyder, P., 2003. Green surprise? How terrestrial ecosystems could affect earth's climate. Front. Ecol. Environ. 1, 38–44.
Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., 2005. Global consequences of land use. Science 309, 570–574.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182.
Fu, C., 2003. Potential impacts of human-induced land cover change on East Asia monsoon. Glob. Planet. Chang. 37, 219–229.
Gao, F., Morisette, J.T., Wolfe, R.E., Ederer, G., Pedelty, J., Masuoka, E., Myneni, R., Tan, B., Nightingale, J., 2008. An algorithm to produce temporally and spatially continuous MODIS-LAI time series. IEEE Geosci. Remote Sens. Lett. 5, 60–64.
Ge, Q., Dai, J., He, F., Zheng, J., Man, Z., Zhao, Y., 2004. Spatiotemporal dynamics of reclamation and cultivation and its driving factors in parts of China during the last three centuries. Prog. Nat. Sci. 14, 605–613.
Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27, 294–300.
Gitas, I.Z., Mitri, G.H., Ventura, G., 2004. Object-based image classification for burned area mapping of Creus Cape, Spain, using NOAA-AVHRR imagery. Remote Sens. Environ. 92, 409–413.
Guan, K., Wood, E.F., Medvigy, D., Kimball, J., Pan, M., Caylor, K.K., Sheffield, J., Xu, X., Jones, M.O., 2014. Terrestrial hydrological controls on land surface phenology of African savannas and woodlands. J. Geophys. Res. Biogeosci. 119, 1652–1669.
Hansen, M., DeFries, R., Townshend, J.R., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int. J. Remote Sens. 21, 1331–1364.
He, Y., Lee, E., 2016. Empirical relationships of sea surface temperature and vegetation activity with summer rainfall variability over the Sahel. Earth Interact. 20, 1–18.
He, H., Shi, X., 2015. Spatio-temporal characteristics of land cover changes in China during 1990-2010. Int. J. Geogr. Inf. Sci. 17, 1323–1332.
He, Y., Bo, Y., de Jong, R., Li, A., Zhu, Y., Cheng, J., 2015. Comparison of vegetation phenological metrics extracted from GIMMS NDVIg and MERIS MTCI data sets over China. Int. J. Remote Sens. 36, 300–317.
Heumann, B.W., Seaquist, J., Eklundh, L., Jönsson, P., 2007. AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982–2005. Remote Sens. Environ. 108, 385–392.
Holben, B.N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data. Int. J. Remote Sens. 7, 1417–1434.
Hou, X., 1981. A further discussion on the principle and scheme for vegetation regionalization of China. Acta Phytoecologica et Geobotanica Sinica. 5, pp. 290–301.
Houghton, R., Hackler, J., 2003. Sources and sinks of carbon from land-use change in China. Glob. Biogeochem. Cycles 17, 1034.
James, M., Kalluri, S.N., 1994. The pathfinder AVHRR land data set: an improved coarse resolution data set for terrestrial monitoring. Int. J. Remote Sens. 15, 3347–3363.
Jensen, J.R., 2016. Introductory Digital Image Processing: A Remote Sensing Perspective. 4th Edition. Pearson, New Jersey.
de Jong, R., de Bruin, S., de Wit, A., Schaepman, M.E., Dent, D.L., 2011. Analysis of monotonic greening and browning trends from global NDVI time-series. Remote Sens. Environ. 115, 692–702.
Jönsson, P., Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE Trans. Geosci. Remote Sens. 40, 1824–1832.
Jönsson, P., Eklundh, L., 2004. TIMESAT—a program for analyzing time-series of satellite sensor data. Comput. Geosci. 30, 833–845.
Kaufmann, R.K., Stern, D.I., 1997. Evidence for human influence on climate from hemispheric temperature relations. Nature 388, 39–44.
Kim, D.-H., Sexton, J.O., Noojipady, P., Huang, C., Anand, A., Channan, S., Feng, M., Townshend, J.R., 2014. Global, Landsat-based forest-cover change from 1990 to 2000. Remote Sens. Environ. 155, 178–193.

Klein, I., Gessner, U., Kuenzer, C., 2012. Regional land cover mapping and change detection in Central Asia using MODIS time-series. Appl. Geogr. 35, 219–234.

Knight, J.F., Lunetta, R.S., Ediriwickrema, J., Khorram, S., 2006. Regional scale land cover characterization using MODIS-NDVI 250 m multi-temporal imagery: a phenology-based approach. GISci. Remote. Sens. 43, 1–23.

Lawrence, P.J., Chase, T.N., 2007. Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0). J. Geophys. Res. Biogeosci. 112 G01023.

Lawrence, P.J., Chase, T.N., 2010. Investigating the climate impacts of global land cover change in the community climate system model. Int. J. Climatol. 30, 2066–2087.

Lawrence, P.J., Feddema, J.J., Bonan, G.B., Meehl, G.A., O'Neill, B.C., Oleson, K.W., Levis, S., Lawrence, D.M., Kluzek, E., Lindsay, K., 2012. Simulating the biogeochemical and biogeophysical impacts of transient land cover change and wood harvest in the Community Climate System Model (CCSM4) from 1850 to 2100. J. Clim. 25, 3071–3095.

Lee, E., Chase, T.N., Rajagopalan, B., Barry, R.G., Biggs, T.W., Lawrence, P.J., 2009. Effects of irrigation and vegetation activity on early Indian summer monsoon variability. Int. J. Climatol. 29, 573–581.

Lee, E., Sacks, W.J., Chase, T.N., Foley, J.A., 2011. Simulated impacts of irrigation on the atmospheric circulation over Asia. J. Geophys. Res. Atmos. 116 (1984–2012). D08114.

Lee, E., He, Y., Zhou, M., Liang, J., 2015. Potential feedback of recent vegetation changes on summer rainfall in the Sahel. Phys. Geogr. 36, 449–470.

de Leeuw, J., Jia, H., Yang, L., Liu, X., Schmidt, K., Skidmore, A., 2006. Comparing accuracy assessments to infer superiority of image classification methods. Int. J. Remote Sens. 27, 223–232.

Liaw, A., Wiener, M., Breiman, L., Cutler, A., 2009. Package "Randomforest". https://cran.r-project.org/web/packages/randomForest/randomForest.pdf Accessed 16. 04. 12.

Lin, G., Ho, S.P., 2003. China's land resources and land-use change: insights from the 1996 land survey. Land Use Policy 20, 87–107.

Liu, J., Liu, M., Zhuang, D., Zhang, Z., Deng, X., 2003a. Study on spatial pattern of land-use change in China during 1995–2000. Sci. China Ser. D Earth Sci. 46, 373–384.

Liu, J., Zhang, Z., Zhuang, D., Wang, Y., Zhou, W., Zhang, S., Li, R., Jiang, N., Wu, S., 2003b. A study on the spatial-temopral dynamic changes of land-use and driving forces analyses of China in the 1990s. Geogr. Res. 22, 1–12.

Liu, J., Zhuang, D., Luo, D., Xiao, X., 2003c. Land-cover classification of China: integrated analysis of AVHRR imagery and geophysical data. Int. J. Remote Sens. 24, 2485–2500.

Liu, J., Liu, M., Tian, H., Zhuang, D., Zhang, Z., Zhang, W., Tang, X., Deng, X., 2005a. Spatial and temporal patterns of China's cropland during 1990–2000: an analysis based on Landsat TM data. Remote Sens. Environ. 98, 442–456.

Liu, J., Tian, H., Liu, M., Zhuang, D., Melillo, J.M., Zhang, Z., 2005b. China's changing landscape during the 1990s: large-scale land transformations estimated with satellite data. Geophys. Res. Lett. 32, L02405.

Liu, X., Lv, X., Luo, Q., 2008. Analysis of cultivated land change in Xinjiang from 1996 to 2005. Res. Soil Water Conserv. 15, 128–134.

Liu, J., Zhang, Z., Xu, X., Kuang, W., Zhou, W., Zhang, S., Li, R., Yan, C., Yu, D., Wu, S., Jiang, N., 2010. Spatial patterns and driving forces of land use change in China during the early 21st century. J. Geogr. Sci. 20, 483–494.

Liu, J., Kuang, W., Zhang, Z., Xu, X., Qin, Y., Ning, J., Zhou, W., Zhang, S., Li, R., Yan, C., 2014. Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s. J. Geogr. Sci. 24, 195–210.

Liu, X., Bo, Y., Zhang, J., He, Y., 2015. Classification of C3 and C4 vegetation types using MODIS and ETM+ blended high Spatio-temporal resolution data. Remote Sens. 7, 15244–15268.

Liu, J., Heiskanen, J., Aynekulu, E., Maeda, E.E., Pellikka, P.K., 2016. Land cover characterization in West Sudanian Savannas using seasonal features from annual Landsat time series. Remote Sens. 8, 365.

Loveland, T., Reed, B., Brown, J., Ohlen, D., Zhu, Z., Yang, L., Merchant, J., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int. J. Remote Sens. 21, 1303–1330.

Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., Worthy, L.D., 2006. Land-cover change detection using multi-temporal MODIS NDVI data. Remote Sens. Environ. 105, 142–154.

Mahmood, R., Pielke, R.A., Hubbard, K.G., Niyogi, D., Dirmeyer, P.A., McAlpine, C., Carleton, A.M., Hale, R., Gameda, S., Beltrán-Przekurat, A., 2014. Land cover changes and their biogeophysical effects on climate. Int. J. Climatol. 34, 929–953.

Maxwell, A.E., Warner, T.A., 2015. Differentiating mine-reclaimed grasslands from spectrally similar land cover using terrain variables and object-based machine learning classification. Int. J. Remote Sens. 36, 4384–4410.

Maxwell, A.E., Warner, T.A., Strager, M.P., 2016. Predicting Palustrine wetland probability using random forest machine learning and digital elevation data-derived terrain variables. Photogramm. Eng. Remote. Sens. 82, 437–447.

McPherson, R.A., 2007. A review of vegetation—atmosphere interactions and their influences on mesoscale phenomena. Prog. Phys. Geogr. 31, 261–285.

Muhammad, S., Niu, Z., Wang, L., Aablikim, A., Hao, P., Wang, C., 2015. Crop classification based on time series MODIS EVI and ground observation for three adjoining years in Xinjiang. Spectrosc. Spectr. Anal. 35, 1345–1350.

Nellis, M.D., Price, K.P., Rundquist, D., 2009. Remote sensing of cropland agriculture. In: Warner, T.A., Nellis, B.D., Foody, G.M. (Eds.), The SAGE Handbook of Remote Sensing. SAGE publications Ltd., California, pp. 368–380.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57.

Pal, M., 2005. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 26, 217–222.

Palacios-Orueta, A., Huesca, M., Whiting, M.L., Litago, J., Khanna, S., Garcia, M., Ustin, S.L., 2012. Derivation of phenological metrics by function fitting to time-series of spectral shape indexes AS1 and AS2: mapping cotton phenological stages using MODIS time series. Remote Sens. Environ. 126, 148–159.

Palmer, S., Odermatt, D., Hunter, P., Brockmann, C., Presing, M., Balzter, H., Tóth, V., 2015. Satellite remote sensing of phytoplankton phenology in Lake Balaton using 10 years of MERIS observations. Remote Sens. Environ. 158, 441–452.

Pielke, R.A., 2005. Land use and climate change. Science 310, 1625–1626.

Pielke, R.A., Pitman, A., Niyogi, D., Mahmood, R., McAlpine, C., Hossain, F., Goldewijk, K.K., Nair, U., Betts, R., Fall, S., 2011. Land use/land cover changes and climate: modeling analysis and observational evidence. Wiley Interdiscip. Rev. Clim. Chang. 2, 828–850.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., 2014. Package "nlme". https://cran.r-project.org/web/packages/nlme/nlme.pdf Accessed 16. 12. 02.

Pinzon, J.E., Tucker, C.J., 2014. A non-stationary 1981–2012 AVHRR NDVI3g time series. Remote Sens. 6, 6929–6960.

Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32, 4407–4429.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. J. Veg. Sci. 5, 703–714.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogramm. Remote Sens. 67, 93–104.

Rouse Jr., J.W., Haas, R., Schell, J., Deering, D., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS. Proceedings of the Thrid ERTS Symposium, NASA, Washington, DC, USA, pp. 309–317.

Running, S.W., Loveland, T.R., Pierce, L.L., 1994. A vegetation classification logic based on remote sensing for use in global biogeochemical models. Ambio 23, 77–81.

Sabins, J., Lulla, K., 2007. Remote Sensing: Principles and Interpretation. 3rd edition. Waveland, Illinois.

Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environ. Res. Lett. 4, 044003.

Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., DiMiceli, C., 2013. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. Int. J. Digital Earth 6, 427–448.

Shi, D., Yang, X., 2016. An assessment of algorithmic parameters affecting image classification accuracy by random forests. Photogramm. Eng. Remote. Sens. 82, 407–417.

Speiser, J.L., Durkalski, V.L., Lee, W.M., 2015. Random forest classification of etiologies for an orphan disease. Stat. Med. 34, 887–899.

Takata, K., Saito, K., Yasunari, T., 2009. Changes in the Asian monsoon climate during 1700–1850 induced by preindustrial cultivation. Proc. Natl. Acad. Sci. 106, 9586–9589.

Tian, F., Fensholt, R., Verbesselt, J., Grogan, K., Horion, S., Wang, Y., 2015. Evaluating temporal consistency of long-term global NDVI datasets for trend analysis. Remote Sens. Environ. 163, 326–340.

Townshend, J.R., Justice, C., Kalb, V., 1987. Characterization and classification of south American land cover types using satellite data. Int. J. Remote Sens. 8, 1189–1207.

Townshend, J., Justice, C., Li, W., Gurney, C., McManus, J., 1991. Global land cover classification by remote sensing: present capabilities and future possibilities. Remote Sens. Environ. 35, 243–255.

Tucker, C.J., Townshend, J.R., Goff, T.E., 1985. African land-cover classification using satellite data. Science 227, 369–375.

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

Vuolo, F., Richter, K., Atzberger, C., 2011. Evaluation of time-series and phenological indicators for land cover classification based on MODIS data. Processdings of Remote Sensing for Agriculcuture, Ecosystem, and Hydrology, SPIE, Prague, Czech. 8174.

Wang, L., Li, C., Ying, Q., Cheng, X., Wang, X., Li, X., Hu, L., Liang, L., Yu, L., Huang, H., 2012. China's urban expansion from 1990 to 2010 determined with satellite remote sensing. Chin. Sci. Bull. 57, 2802–2812.

Warrens, M.J., 2015. Properties of the quantity disagreement and the allocation disagreement. Int. J. Remote Sens. 36, 1439–1446.

Webster, P.J., 1987. The Elementary Monsoon. Wiley, New York.

Wohlfart, C., Liu, G., Huang, C., Kuenzer, C., 2016. A River Basin over the course of time: multi-temporal analyses of land surface dynamics in the Yellow River Basin (China) based on medium resolution remote sensing data. Remote Sens. 8, 186.

Wu, W., Peng, Y., Tnag, H., Zhou, Q., Chen, Z., Shibasaki, R., 2010. Characterizing spatial patterns of phenology in cropland of China based on remotely sensed data. Agric. Sci. China 9, 101–112.

Xu, L., Li, B., Yuan, Y., Gao, X., Liu, H., Dong, G., 2015. Changes in China's cultivated land and the evaluation of land requisition-compensation balance policy from 2000 to 2010. Resource Sci. 37, 1543–1551.

Xue, Z., Du, P., Feng, L., 2014. Phenology-driven land cover classification and trend analysis based on long-term remote sensing image series. IEEE J. Sel. Top. Applied Earth Observ. Remote Sens. 7, 1142–1156.

Yan, E., Wang, G., Lin, H., Xia, C., Sun, H., 2015. Phenology-based classification of vegetation cover types in Northeast China using MODIS NDVI and EVI time series. Int. J. Remote Sens. 36, 489–512.

Yin, X., 2008. Analysis on the change of land use by remote sensing technology in Manas county. J. Shihezi Univ. (Nat. Sci.) 26, 402–406.

Zhang, X., Friedl, M.A., Schaaf, C.B., 2006. Global vegetation phenology from moderate resolution imaging spectroradiometer (MODIS): evaluation of global patterns and comparison with in situ measurements. J. Geophys. Res. Biogeosci. 111, G04017.

Zhang, X., Wu, S., Yan, X., Chen, Z., 2016. A global classification of vegetation based on NDVI, rainfall and temperature. Int. J. Climatol. http://dx.doi.org/10.1002/joc.4847.

Zhao, X., Xu, P., Zhou, T., Li, Q., Wu, D., 2013. Distribution and variation of forests in China from 2001 to 2011: a study based on remotely sensed data. Forests 4, 632–649.

Zhu, X., 2012. The Impact of Agricultural Irrigation on Land Surface Characteristics and near Surface Climate in China. University of Maryland, USA.