Article

Seasonal Land Cover Dynamics in Beijing Derived from Landsat 8 Data Using a Spatio-Temporal Contextual Approach

Jie Wang [1], Congcong Li [2], Luanyun Hu [2], Yuanyuan Zhao [2], Huabing Huang [1] and Peng Gong [1,2,3,*]

[1] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China; E-Mails: wangjie@radi.ac.cn (J.W.); huanghb@radi.ac.cn (H.H.)

[2] Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, Beijing 100084, China; E-Mails: licc129@163.com (C.L.); huluanyun@gmail.com (L.H.); yyzhao12345@gmail.com (Y.Z.)

[3] Joint Center for Global Change Studies, Beijing 100875, China

* Author to whom correspondence should be addressed; E-Mail: penggong@tsinghua.edu.cn; Tel.: +86-10-6279-5385; Fax: +86-10-6279-7284.

Abstract: Seasonal dynamic land cover maps could provide useful information to ecosystem, water-resource and climate modelers. However, they are rarely mapped more frequent than annually. Here, we propose an approach to map dynamic land cover types with frequently available satellite data. Landsat 8 data acquired from nine dates over Beijing within a one-year period were used to map seasonal land cover dynamics. A two-step procedure was performed for training sample collection to get better results. Sample sets were interpreted for each acquisition date of Landsat 8 image. We used the random forest classifier to realize the mapping. Nine sets of experiments were designed to incorporate different input features and use of spatial temporal information into the dynamic land cover classification. Land cover maps obtained with single-date data in the optical spectral region were used as benchmarks. Texture, NDVI and thermal infrared bands were added as new features for improvements. A Markov random field (MRF) model was applied to maintain the spatio-temporal consistency. Classifications with all features from all images were performed, and an MRF model was also applied to the results estimated with all features. The best overall accuracies achieved for each date ranged from 75.31% to 85.61%.

1. Introduction

Land cover and land cover change have great impact on the biogeochemical and biogeophysical processes as well as climate change [1]. Some climate models require information about the time-evolving paths of land cover [2]. Traditionally, land cover was held static on an annual or decadal basis in land process models. However, a climate modeler survey showed that seasonal and monthly dynamics of land cover are preferable [3].

The use of satellite data is one of the most efficient approaches for monitoring land cover and its changes [4]. Many regional and global land cover mapping efforts have been performed using various satellite data (e.g., [5–9]). Information on single categories of land cover, such as the 30 m continuous fields of tree cover [10] and the 30 m forest cover, change in the 21st century [11] were also derived from Landsat data at the global scale. In all these efforts, land cover types were deemed static in a year. None of them is mapped at seasonal or monthly scales.

Long-term land cover dynamics were studied by some researchers. Multiple growing season Landsat images for every year between 1985 and 2007 were used in detecting trends in forest disturbance and recovery [12]. Dense Landsat time series were used to detect forest disturbances [13]. Multi-seasonal Landsat-5 data were used to map long-term land cover dynamics [14]. In these efforts only forest disturbance or inter-annual land cover changes were considered. Zhu *et al.* [15], Zhu and Woodcock [16], and Verbesselt *et al.* [17] tried to get land cover maps and changes by decomposing time series of spectral bands into different components, such as multiple frequency component. Multi-temporal data with short-time intervals and long-time range were required for these methods, and they focused mainly on land cover disturbances caused by fires or insect attacks, and phenological changes. Dynamics of seasonal land cover status were not considered. High-frequency maps were obtained for water bodies over 629 lakes in China using eight-day MODIS composite data from 2000 to 2010 [18].

To map seasonal land cover dynamics, seasonal snow/ice, leaf-on/leaf-off status of vegetation, and other intra-annual changes need to be mapped from multi-seasonal images. It is different from inter-annual land cover change detection and phenology estimation. Temporally dense image series are needed for phenology estimation. Phenology estimation from Landsat-like data is hard when few of the data are cloud-free. Moreover, water, seasonal snow/ice cover and other land cover changes cannot be obtained from phenological information only. Although seasonal snow cover can be tracked by using snow cover models [19], monitored snow-cover dynamics can serve as the calibration and validation data. Because land cover mapping based on a single-date mid-resolution image can be reliable under various circumstances, it is possible to map land cover dynamics with multi-temporal images. Landsat 8 is a continuation of the Landsat series of satellites, and the acquired data are sufficiently consistent with those from previous Landsat missions [20], with improved signal to noise ratio. In this study, we test a set of procedures to map dynamic land covers in Beijing, China using Landsat 8 data. Different combinations of features for land cover classification with the newly acquired data were tested. A spatio-temporal consistency model, Markov Random Fields (MRF), was also tested in this research.

2. Study Area and Data

2.1. Study Area

Our study area is part of the administrative district of Beijing (Figure 1). Beijing is located between latitudes 39°26′ N and 41°03′ N, and longitudes 115°25′ E and 117°30′ E, in the northeastern edge of the North China Plain. The south and east of the city are covered by plains, while the north, northwest and west are surrounded by mountains. Major land covers in this area are impervious layers, croplands, forests and shrublands. It is hot and humid in the summer and cold and dry in the winter.
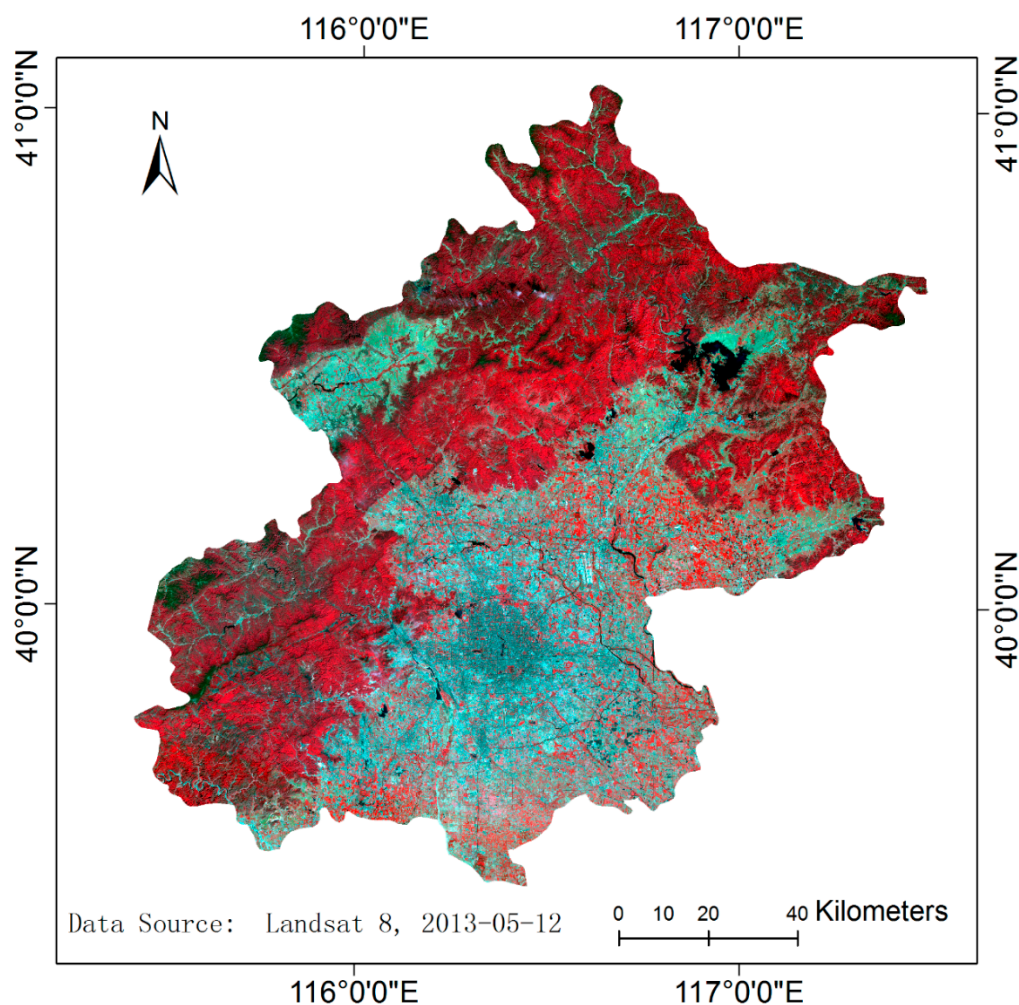

Figure 1. Study area.

2.2. Landsat Data

Landsat 8 images from two scenes in the Worldwide Reference System (WRS)-2 (rows 32 and 33 along path 123) were downloaded from the United States Geological Survey (USGS) website (http://earthexplorer.usgs.gov/). Totally, eighteen images were used in this research. They were acquired on nine different dates. Table 1 shows regional cloud cover proportions for the eighteen images. The row lists day-of-year (DOY) of each image. These images were selected so that there was as little cloud as possible in our study area, and the seasonal land cover dynamics within one year can be mapped.

Although cloud cover proportions for three of the images (path-row 123-33) were over half of the image area, most of the subregions located in our study area were cloud free.

Table 1. Cloud cover proportions (range from 0 to 1) for Landsat 8 images used in this research.

Path-Row	DOY (the Last One Is in 2014, Others Are in 2013)								
	132	164	212	244	276	308	324	340	39
123-32	0.00	0.12	0.17	0.02	0.06	0.01	0.08	0.21	0.37
123-33	0.01	0.13	0.53	0.00	0.12	0.03	0.36	0.60	0.58

Landsat 8 image contains eight Operational Land Imager (OLI) multispectral bands (1–7, 9) at 30 m resolution, one OLI panchromatic band (8) at 15 m resolution, and two Thermal Infrared Sensor (TIRS) bands (10–11) collected at 100 m and resampled to 30 m resolution. We used all bands in our land cover mapping experiments, except the ninth OLI band (1360–1380 nm), which is for cirrus monitoring.

2.3. Data Preprocessing

Data preprocessing includes radiometric processing, image mosaicking, feature stacking and cloud masking (Figure 2). The resultant data set included nine images, each with 16 layers of features (nine radiometrically corrected bands, six texture features and one Normalized Difference Vegetation Index (NDVI)).
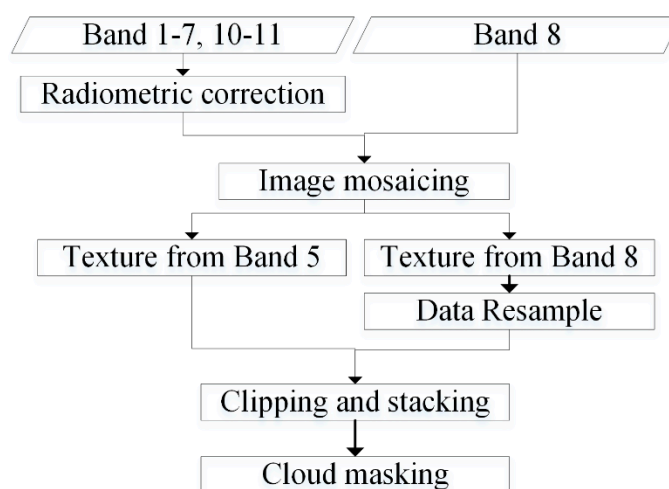

Figure 2. Block diagram of data preprocessing.

The radiometric processing was done automatically using the GM software package. Parameters were automatically estimated or set by information from metadata. The GM package was developed for various satellite data preprocessing. An introduction to it can be found in [9]. Bands 1–7 of all images were first calibrated to top of atmosphere reflectance, and then converted to surface reflectance after atmospheric correction and topographic correction. Bands 10–11 were converted to brightness temperature. Image mosaicking was performed for every two images acquired on the same date. We got nine mosaicked images; each corresponds to an acquisition date listed in Table1.

For every mosaicked image, features for dynamic land cover mapping include nine radiometrically corrected bands (band 1–7, 10,11), one Normalized Difference Vegetation Index (NDVI) channel and six texture features. NDVI is calculated as: NDVI = (NIR − RED) / (NIR + RED) [21], where NIR is band 5 of Landsat 8, and RED is band 4. Three texture features, mean, contrast and entropy of grey-level co-occurrence matrix (GLCM) [22–24], were calculated from bands 5 and 8, respectively. Parameters for co-occurrence matrix were set as follows: grey levels = 16, window size = 7, and offset = (1, 1). Textures extracted from band 8 were resampled to 30 m resolution. Vegetation indices other than NDVI can be included as features. More texture features extracted from multiple bands can also be used. We used the chosen features merely as a compromise between land-cover mapping accuracy and computation load. All layers (nine images, 16 layers of features per image) were registered and clipped according to the extent of Beijing City.

Thick clouds and dark shadows were interpreted scene-by-scene in the stacked image sequence. Interpreted cloud and shadow regions in any image were masked out from all corresponding features.

3. Mapping Dynamic Land Cover Types

3.1. Classification System

At the global scale, two land cover classification systems are widely used, the land cover classification systems of the United Nations Food and Agriculture Organization (FAO) Land Cover Classification System (LCCS) and the classification system of the International Geosphere-Biosphere Programme (IGBP). Gong *et al.* [9] designed a flexible two-level land cover classification system, which could be easily cross-walked to these two classification systems. Additional layers of information, such as fractional cover and vegetation height are needed for cross-walking to other classification systems.

The land cover classification system used here was derived from [9]. Some modifications were made, and only Level 1 classes were used for dynamic land-cover mapping in Beijing. The revised classification system is composed of 12 Level 1 classes, as shown in Table 2. There is no "Tundra" class, and the "Snow/ice" is a transient class in our study area. The "Bare croplands", "Forests, leaf-off", "Grasslands, leaf-off", "Shrubs, leaf-off" classes were defined to indicate out of growing-season conditions of their corresponding vegetation types. Definitions for all the other classes were the same as in [9]. The leaf-on/leaf-off change corresponds to two major stages of the "phenological change" of some vegetation species. Since their status changes with season and such changes alter the surface characteristics, we considered them seasonal dynamics of land cover in this research. In addition, the dynamics of snow cover and water level and other land-cover changes with seasonal variation of climate are all referred to as seasonal land-cover dynamics.

Table 2. Classification system.

1	2	3	4	5	6
Croplands	Forests	Grasslands	Shrublands	Water bodies	Impervious
7	8	9	10	11	12
Barren lands	Snow/ice	Bare croplands	Forests, leaf-off	Grasslands, leaf-off	Shrubs, leaf-off

3.2. Validation and Training Samples

Validation samples and training samples were collected independently. Totally, nine validation sample sets and nine training sample sets were obtained, each corresponding to an acquisition date of Landsat 8 as listed in Table 1. Sample class labels were interpreted based on high-resolution images from Google Earth, field survey photos, field knowledge and the corresponding Landsat 8 image.

Validation samples were collected from 500 uniformly distributed random positions in the study area. A sample point in a Landsat 8 image would be removed were it located in a region covered by clouds or shadows. Points located in regions mixed by multiple land cover types were also removed. Sizes of the nine validation sample sets are all between 400 and 500, with slight differences.

Representative training sample collection is essential in land cover classification efforts. In this research, our training samples were obtained in two steps. The first step was initial interpretation, and the second step was to add more representative samples. Two types of training sample units, big sample units and single-pixel sample units, were collected. The big training sample units were homogeneous in areas greater than 100 m × 100 m (3 × 3 pixels), and pixels within the 3 × 3 window were all extracted and included as training sample units. In the first step, more than 100 big sample positions and more than 100 single-point sample positions were selected, so that there were at least 10 sample points for each class in each of the Landsat 8 image. The second step is presented in Section 3.3.2.

3.3. Methodology

3.3.1. Random Forest and Markov Random Fields

The random forest (RF) [25] classifier was used for land cover classification and probability estimation of Landsat 8 data. RF maintains accuracy when a proportion of the data are missing, and can handle high dimensionality effectively [26,27]. RF had proven effective for land cover classification when using multi-seasonal imagery and multi-seasonal texture [28]. Therefore, we could train RF classifiers with many layers of features, with no need for feature selection. We could estimate land cover types in cloud and shadow regions masked out during data preprocessing when multi-temporal Landsat 8 data were used as input features. The number of trees was set to 200 for all RF classifiers used in this research. Default values were taken for other parameters: the depth of the trees was set as unlimited, the seed value was set as 1, and the number of attributes to be used in random selection was set as the square root of the number of features.

A spatio-temporal consistency model, Markov random fields (MRF), was used to improve land cover classification in this research. It was defined and computed in the same way as in [29,30]. The MRF model is summarized as follows. A pixel in our Landsat 8 data sequence can be located by its spatio-temporal position (s, t). The resultant label of the pixel is computed by:

$$L^*_{(s,t)} = \arg \max_{L_{(s,t)}} \left\{ P\left(L_{(s,t)} \mid L_{\mathbb{N}(s,t)}\right) \times f\left(X_{(s,t)} \mid L_{(s,t)}\right) \right\} \tag{1}$$

where $L_{(s,t)}$ is the class label of the pixel located at position (s, t), $L_{\mathbb{N}(s,t)}$ is the class labels of the pixel's spatio-temporal neighbors, X is the input features, f is the marginal likelihood function estimated by the

classifier as probabilities for all given labels at pixel (s, t), and P is the conditional prior distribution computed from the spatio-temporal relationship:

$$P\left(L_{(s,t)} \mid L_{\mathbb{N}(s,t)}\right) = \frac{1}{Z} \exp\left\{-\left[U_s + U_{T1} + U_{T2}\right]\right\} \tag{2}$$

where Z is a normalizing constant, U_S, U_{T1}, U_{T2} are called the spatial and temporal energy functions, defined as:

$$U_s = \sum_{(s',t') \in \mathbb{N}_{s(s,t)}} \left[-\beta_1 I\left(L_{(s,t)} = L_{(s',t')}\right)\right] \tag{3}$$

$$U_{T1} = \sum_{(s',t') \in \mathbb{N}_{T1(s,t)}} \left[-\beta_2 P\left(L_{(s,t)} \mid L_{(s',t')}\right) + \beta_3 I\left(L_{(s',t')} \nRightarrow L_{(s,t)}\right)\right] \tag{4}$$

$$U_{T2} = \sum_{(s',t') \in \mathbb{N}_{T2(s,t)}} \left[-\beta_4 P\left(L_{(s',t')} \mid L_{(s,t)}\right) + \beta_5 I\left(L_{(s,t)} \nRightarrow L_{(s',t')}\right)\right] \tag{5}$$

where $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are nonnegative coefficients, I(X) is equal to 1 if X is true and 0 otherwise, $L_{(s',t')} \nRightarrow L_{(s,t)}$ means an illogical transition (for example, Impervious $\nRightarrow$ Forests in a short time), and $P(L_2|L_1)$ is the temporal transition probability from L_1 to L_2. The temporal transition probability matrices were estimated using the global transition model [31,32]. The Iterated conditional modes (ICM) algorithm [33] was used to get the result labels for all pixels.

3.3.2. Expanding Training Samples

The method for expanding training samples in this research can be viewed as a simplified version of the active learning algorithms [34]. Much higher overall accuracies were achieved by training sample expansion. The nine training sample sets obtained in the first step were merged into a new training sample set. A random forest classifier of 200 trees was trained on the new sample set. All $16 \times 9 = 144$ layers from the nine images were used as input features. We assumed that pixels hard to classify with the trained classifier were informative, and new training samples should be interpreted from these pixels.

The classifier was applied to all the pixels to estimate probabilities for all classes. An uncertainty measure, entropy, was calculated for each pixel. Entropy is defined as:

$$H(x) = -\sum_{c \in C} p(x_c) \log p(x_c) \tag{6}$$

where c is the class label set, and $p(x_c)$ is the probability of the pixel x belonging to class c. A high entropy value means high uncertainty. A threshold was set to get the most uncertain 30% pixels. Then, 100 positions were randomly located from the most uncertain 30% of pixels. New training samples were interpreted at these positions. More than 20 big sample units and more than 40 single-point sample units were added for each of the nine acquisition dates.

3.3.3. Classification Experiments

Nine sets of experiments (P1–P9) of dynamic land-cover classifications were carried out, as shown in Figure 3, where the training sets, the resultant maps and the improved results are shown from left to right. In each process, nine land cover maps were obtained by training on the nine training sample sets.

Land cover mapping results were obtained with different feature combinations (P1–P6), or with all features from the image sequence (P8). Features used in our classification processes are listed in Table 3, where the NDVI served as an enhancement of vegetation information. Result improvements by using MRF models (P7 and P9) were also performed.
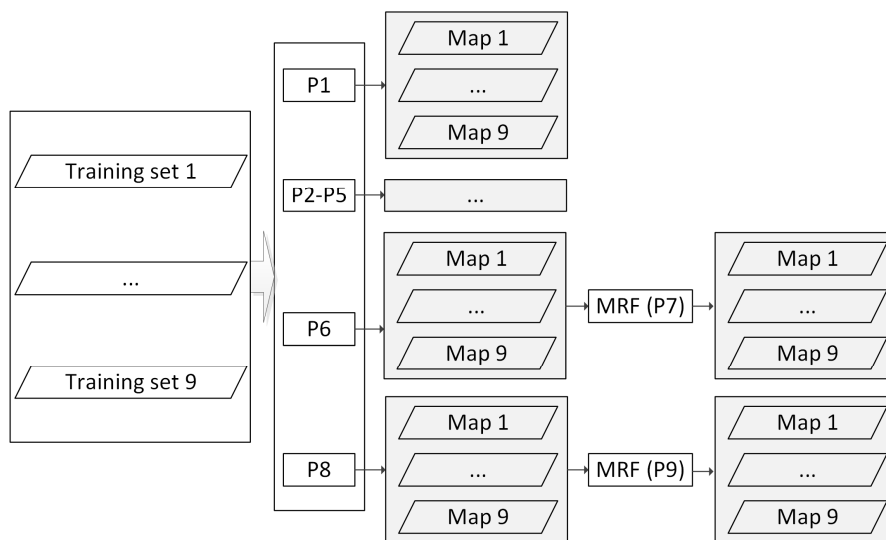

Figure 3. Dynamic land-cover classification processes.

Table 3. Features used in classification processes (P7 is an improvement for P6, and P9 is an improvement for P8).

Process	Features	Process	Features
P1	OLI band 1–7	P6	OLI + NDVI + Textures + TIRS
P2	OLI + Texture	P7	Probabilities and map results from P6
P3	OLI + NDVI	P8	Time series of OLI + NDVI + Textures + TIRS
P4	OLI + TIRS	P9	Probabilities and map results from P8
P5	OLI + NDVI + Texture		

We assessed the impact of features on land cover classification with different feature combinations. Features in an image were grouped as OLI bands (bands 1–7 of Landsat 8 L1T data), texture features, NDVI and TIRS (bands 10–11 of Landsat 8 L1T data). Land cover maps obtained with only the OLI bands were treated as benchmarks. Results obtained by combining other features with the OLI bands were compared with the benchmarks. Totally, six kinds of features were used as input data (P1–P6): (1) the OLI bands 1–7, (2) the OLI with the texture, (3) the OLI with the NDVI, (4) the OLI with the TIRS, (5) the OLI with both the texture and the NDVI, and (6) the OLI with all the other features.

For resultant maps obtained with the OLI and all other features from single-date images, class probabilities were also estimated for each pixel by using the trained random forest classifiers. An MRF model (P7) was applied to improve the accuracy and spatio-temporal consistency of the dynamic land-cover results. The estimated probabilities were used as marginal likelihood estimations, and the temporal transition probability matrices in the MRF model were estimated by counting pixel pairs between every two temporally neighboring images.

In addition to using features from single-date images, classifications using all features from the image sequence were also performed (P8). Each of the nine classifiers was trained on the corresponding training set separately, using all the $16 \times 9 = 144$ features of the nine images. The resultant maps could accurately reflect the land-cover status on the corresponding dates, because the training samples were interpreted based on the land covers at that point in time. Features extracted from the corresponding image would be chosen as the main information source by the RF classifier while features extracted from other images were used as auxiliary information. Features of the masked out pixels were set as missing values. Land cover types were still estimated for these pixels unless all the 144 features were missing values. Therefore, almost all clouds and dark shadows were "padded" in the resultant maps. An MRF model (P9) was applied to the results to further improve the classification, in the same way as in the previous process.

4. Results

Totally, 81 land cover maps were obtained from the nine dynamic classification processes described in Section 3.3.3. The nine maps classified with all the 144 features (P8) are shown in Figure 4. Vegetation was turning green on DOY 132 (Figure 4a), was greener on DOY 164 (Figure 4b), and reached the greenest on DOY 212 and 244 (Figure 4c,d). Some trees, shrubs and grasses were under leaf-off conditions on DOY 276 (Figure 4e). Most vegetation types were under leaf-off conditions in the winter (DOY 308, 324, 340 in 2013 and DOY 39 in 2014, Figure 4(f–i)). It is worth mentioning that the land covers in the last map (Figure 4i) were characterized by seasonal snow/ice.

For the resultant maps of each process, accuracies were assessed using the 9 validation sample sets. All estimated overall accuracies are listed in Table 4. Values in column DOY indicate image acquisition dates, and values in columns P1–P9 (the dynamic land-cover classification processes P1–P9 as described in Section 3.3.3) are estimated overall accuracies. The P1 column is the benchmark whose values are overall accuracies for resultant maps obtained by using the OLI features only. Values in P2–P6 are overall accuracies for resultant maps obtained by adding more features. The single-date image classification results were improved by using the MRF model in process P7. Values in P8 are overall accuracies for resultant maps obtained by using all features from the image sequence, and values in P9 column are overall accuracies for the results in process P9. As a comparison, Table 5 shows overall accuracies for maps obtained in P8 by using only training samples from the first step of the two-step training sample collection. Clearly, considerable improvement in training sample augmentation can be obtainable.

Table 4. Overall accuracies (DOY = Day-of-year, P1–P9 are the nine dynamic land cover classification processes P1–P9 described in Section 3.3.3).

DOY	P1 (%)	P2 (%)	P3 (%)	P4 (%)	P5 (%)	P6 (%)	P7 (%)	P8 (%)	P9 (%)
132	66.27	65.79	62.20	64.83	66.99	65.79	70.81	**77.51**	75.12
164	59.75	63.46	59.51	62.96	64.69	64.69	68.64	74.57	**75.31**
212	60.83	59.12	63.75	63.75	61.31	65.21	76.64	82.73	**85.16**
244	66.25	67.00	66.25	69.23	67.74	71.46	77.67	82.63	**85.61**
276	61.81	67.06	63.48	67.06	68.26	71.36	79.24	79.47	**80.43**
308	59.45	61.69	60.95	62.69	62.94	66.92	71.14	**75.62**	74.38
324	57.89	61.90	57.89	61.40	65.66	67.42	72.18	**77.69**	76.19

Table 4. *Cont.*

DOY	P1 (%)	P2 (%)	P3 (%)	P4 (%)	P5 (%)	P6 (%)	P7 (%)	P8 (%)	P9 (%)
340	58.64	56.02	58.64	58.12	61.26	66.23	68.85	**79.06**	76.44
39	69.51	70.03	71.32	72.35	72.09	74.16	75.71	77.00	**78.29**
Average	62.27	63.56	62.67	64.71	65.66	68.14	73.43	78.48	**78.55**

Bold values highlight the highest accuracy for each image.
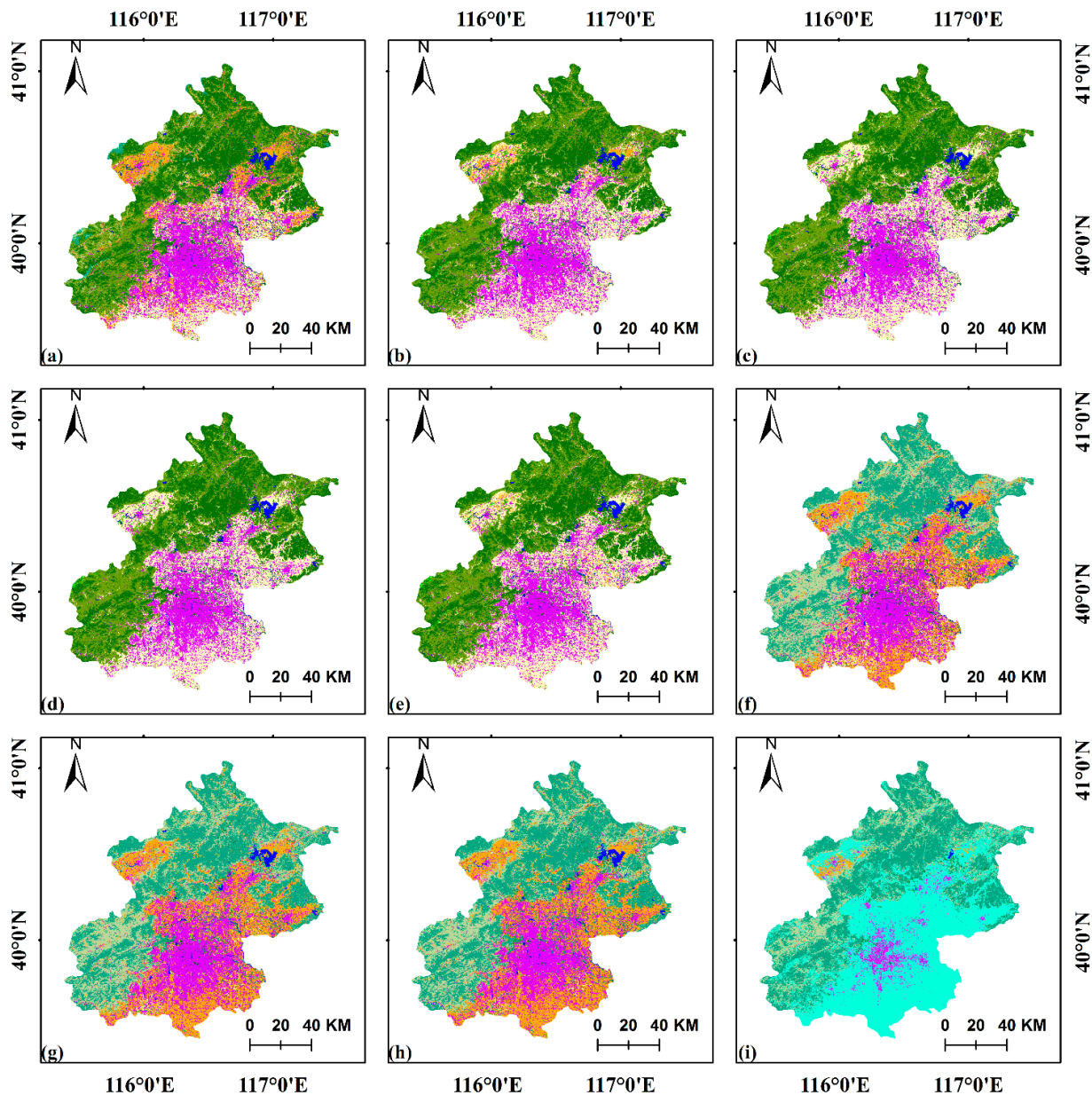
Figure 4. Land cover maps classified with all the 144 features. (**a**–**h**) are the maps for the eight dates (day of year (DOY) 132-356) in 2013. (**i**) is the map for DOY 39 in 2014.

Table 5. Overall accuracies for maps obtained in process P8 by using only training samples collected in the first step.

DOY	132	164	212	244	276	308	324	340	39	Average
OA(%)	68.17	67.87	76.71	77.40	72.34	67.88	70.24	71.45	64.83	70.77

Confusion matrices were computed for all maps. A confusion matrix for a result in growing season is shown in Table 6, and a matrix for a result partly out of growing season is shown in Table 7. The OTH class in Table 7 is an abbreviation for leaf off Grasslands/Barren lands/Grasslands, because there were only a few points for them. In fact, only two barren lands points were correctly classified for these classes. Croplands, forests, shrublands and impervious are the four major land cover types. Confusions between forests and shrublands are the main error sources for the result in the growing season (Table 6). In Table 7, confusions among leaf-on/leaf-off classes were the major error sources.

Table 6. Confusion matrix for a land cover map on day of year (DOY) 244 in process P9 (Overall Accuracy: 85.61%. CR = Croplands, BCR = Bare croplands, FR = Forests, GR = Grasslands, SHR = Shrublands, WB = Water bodies, IMP = Impervious, BL = Barren lands, UA = user accuracy, PA = producer accuracy).

Name	CR	BCR	FR	GR	SHR	WB	IMP	BL	UA (%)
CR	101	2	4	4	1	0	6	0	85.59
BCR	0	0	0	0	0	0	0	0	0.00
FR	1	0	101	0	3	0	0	0	96.19
GR	0	0	0	0	0	0	0	0	0.00
SHR	7	0	23	0	68	0	0	0	69.39
WB	0	0	0	0	0	9	0	0	100.00
IMP	0	1	0	3	0	0	64	3	90.14
BL	0	0	0	0	0	0	0	2	100.00
PA (%)	92.66	0.00	78.91	0.00	94.44	100.00	91.43	40.00	85.61

Table 7. Confusion matrix for a land cover map on day of year (DOY) 308 in process P8 (Overall Accuracy: 75.62%. CR = Croplands; BCR = Bare croplands; FR = Forests; FRN = Forests, leaf-off; SHR = Shrublands; SHRN = Shrubs, leaf-off; WB = Water bodies; IMP = Impervious; OTH = Grasslands/Barren lands/Grasslands, leaf-off; UA = user accuracy; PA = producer accuracy).

Name	CR	BCR	FR	FRN	SHR	SHRN	WB	IMP	OTH	UA (%)
CR	10	14	0	0	0	0	0	0	4	35.71
BCR	9	64	1	0	0	0	0	2	3	81.01
FR	4	1	8	1	0	0	0	0	0	57.14
FRN	0	0	14	72	0	8	0	0	0	76.6
SHR	0	0	0	0	0	0	0	0	0	0
SHRN	1	5	6	15	1	67	0	0	0	70.53
WB	0	0	0	0	0	0	10	0	0	100
IMP	0	4	0	0	0	0	0	71	4	89.87
OTH	0	1	0	0	0	0	0	0	2	66.67
PA (%)	41.67	71.91	27.59	81.82	0	89.33	100	97.26	15.38	75.62

5. Discussion

Texture, NDVI and TIRS are both useful for getting more accurate land cover mapping results. The Values in P5 of Table 4 are greater than the corresponding values in P1, but the values in P2–P4 are not always greater. The values in P6 are the greatest, except for DOY 132, which is slightly lower. NDVI only served as an enhancement of vegetation information, and did not bring new information. Texture did not add as much value as did TIRS; neither did NDVI. However, when texture and NDVI were combined, they were more effective in improving classification accuracies. When all features were used (P6), the overall accuracies were clearly greater than those in the corresponding benchmarks.

The spatio-temporal relationships were not considered in the original classification. Large improvements can be obtained by applying the spatio-temporal consistency constraint, the MRF model, on the classification results for single-date images. The values in P7 of Table 4 are all considerably greater than those in P6. The greatest increase is >10%.

The highest values are in columns P8–P9 of Table 4. They are overall accuracies estimated for classification results obtained by using all features from the image sequence (P8), and for results obtained by applying the MRF model to those from all features (P9). In dynamic land-cover classification, land cover types are determined by their real time conditions at the time of image acquisition. It is different from those where land covers are deemed static in a range of time or in a year. The spatio-temporal neighborhood information is still helpful in mapping dynamic land covers. As an example, the NDVI values for the validation samples of the four major classes, croplands, forests, shrublands and impervious, are boxplotted in Figure 5. Classes are easier to distinguish by multi-date NDVI values than by single-date NDVI values. The values in column P9 are greater than the values in column P8 on DOY 164, 212, 244, 276 and 39, but not on the other dates. Errors are mainly due to confusions among leaf-on/leaf-off classes on those dates when the overall accuracy values in column P9 are not greater. This type of error cannot be corrected by applying the spatio-temporal consistency constraints.
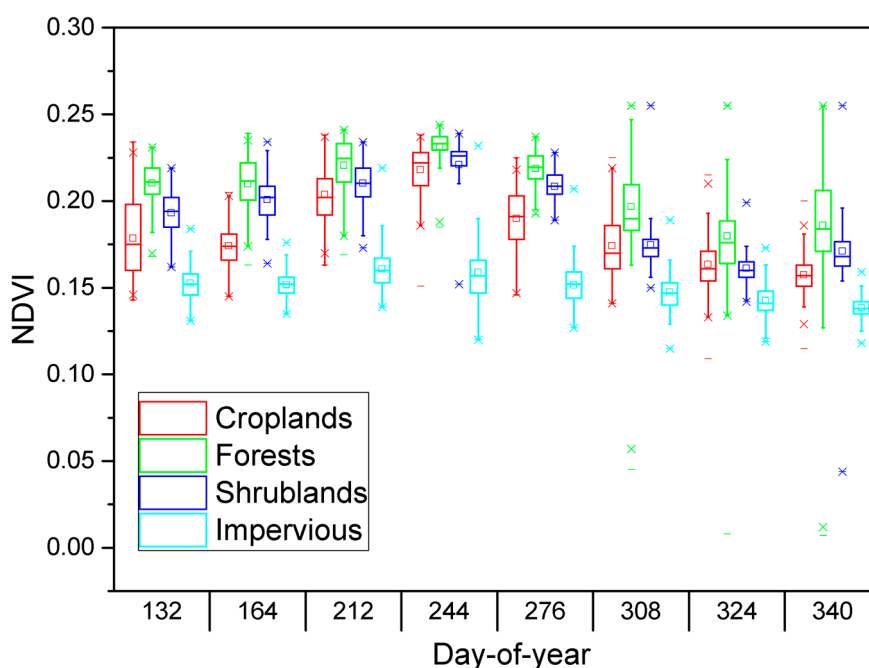

Figure 5. NDVI values for the validation samples of the four major land cover types.

Summer images have the greatest separability of land cover classes. From confusion matrices shown in Tables 6 and 7, it is clear that confusions between forests and shrublands, and confusions among leaf-on/leaf-off classes are the major errors. More training samples for forests and shrublands, especially for samples that are spectrally confusing, will be helpful in reducing confusions between these two categories. Adding vegetation height data in the classification will also reduce the confusions [35]. Classification and accuracy validation for the leaf-on and leaf-off classes were all based on interpreted samples. Quantitative data for vegetation growth status are required for making objective decisions. NDVI or fractional vegetation cover data can be applied to quantitatively indicate vegetation growth status.

The classifiers in process P8 were all trained with the same feature set, all the 144 features of the image sequence. The resultant maps still accurately reflected the land covers on the corresponding dates, because the training samples were interpreted based on the land covers at that point in time. A regional example is shown in Figure 6. False color images and the corresponding results from data acquired on DOY 132 and DOY 276 are listed in Figure 6(b–e). Water bodies are highly dynamic, and there are many mixed pixels in this region. Classification errors often arise in such areas. For example, a few pixels near the water (below the red ellipse in Figure 6c) were misclassified as impervious. However, most pixels were still correctly classified in the two images. It is notable that grasslands, leaf-off in the red ellipse of Figure 6c have been correctly classified as water bodies in Figure 6e. Another factor for the high accuracy is our two-step training sample collection method. Pixels in highly dynamic areas were more likely to be selected in the second step. These sample units represent different land cover types at different points in time. The effect of training sample augmentation can also be shown by comparison between P8 of Tables 4 and 5.
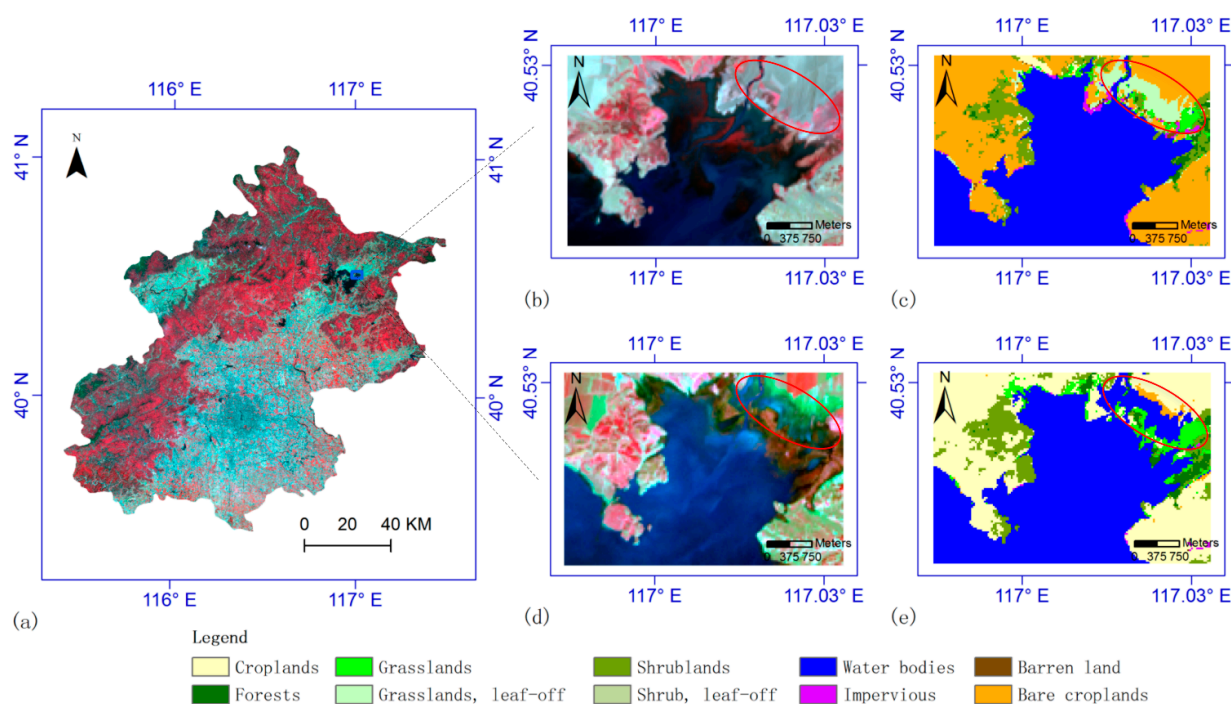

Figure 6. Regional results classified with all the 144 features. The blue region in study area (**a**) indicates the area shown in (**b–e**). The false color images (R: band 5, G: band 4, B: band 3) in (**b**) and (**d**) are from data acquired on day of year (DOY) 132 and DOY 276. And the mapping results are in (**c**) and (**e**). Red ellipses indicate the most dynamic areas.

Clouds and dark shadows were "padded" in processes P8–P9 (see Figure 4), but the padding had no effect on our overall accuracy estimation. Because all validation sample points located in these regions had been removed. Clouds and dark shadows were preprocessed manually. The impact of the preprocessing on the results need to be evaluated in future work. Besides missing value estimation used in this research, the MRF model can also be used for "data padding" by using spatio-temporal neighborhood information [36]. The MRF models can also be used to "pad" NDVI to cloud covered regions and detect temporal changes in NDVI [37].

Spatio-temporal neighborhood information had been used for feature selection and missing value estimation when training random forest with all the 144 features. Better spatio-temporal consistency was obtained by using the MRF model. Land cover dynamics along temporal neighborhood were more reasonable, and pepper-and-salt points were removed in the resultant maps.

Mapping seasonal land-cover dynamics is different from inter-annual land-cover change detection and phenological estimation. Water level dynamics, seasonal snow/ice, phenology change, and other intra-annual changes need to be mapped from multi-seasonal images. Data acquired from the newly launched Landsat 8 within a whole year were used in this research. Only seasonal dynamics in one year were mapped. Long-term data sets are needed, and our methods need to be extended for mapping both intra-annual and inter-annual land-cover variations, as what had been done in [12–17].

6. Conclusions

This research aimed to map seasonal land cover dynamics in Beijing with nine Landsat 8 images by using spatio-temporal information. A two-step procedure was performed for training sample collection to get better results. The average overall accuracy (OA) increased 7% after the training sample augmentation in the second step. Nine sets of classification experiments for mapping seasonal land cover dynamics were evaluated. The random forest classifier was used for classification and probability estimation. Land cover maps obtained using the Operational Land Imager (OLI) multispectral bands from single-date images only were used as benchmarks for comparison (average OA was 62.27%). Texture, Normalized Difference Vegetation Index (NDVI) and Thermal Infrared Sensor (TIRS) were all helpful when included in image classification. Average OA was 68.14% when all features were included. More accurate results were obtained by applying an MRF model (average OA was 73.43%). The highest overall accuracies were achieved by using the last two methods. The average OA for Land cover maps classified with all the 144 features from all images was 78.48%. Further improvements were achieved by using an MRF model (average OA was 78.55%).

Clouds and dark shadows were masked out during data preprocessing. Land cover types were still estimated for pixels located in these areas when mapping with all the features from the image sequence. This did not affect our accuracy estimation. Because validation sample points located in these regions were removed.

Some forests are easy to be confused with shrublands in Beijing. Adding more training samples will reduce the confusion [27]. Confusions among leaf-on and leaf-off classes are major error sources in mapping seasonal land cover dynamics. Classification for leaf-on and leaf-off classes in this research were based on training samples that are subjectively determined by image analysts. Quantitative data for vegetation growth status are required for more accurate leaf-on/leaf-off discriminations.

Acknowledgements

Author Contributions

Peng Gong and Jie Wang conceived and designed this research. Jie Wang did most of the analysis. All authors contributed extensively to data collection and writing of this paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Feddema, J.J.; Oleson, K.W.; Bonan, G.B.; Mearns, L.O.; Buja, L.E.; Meehl, G.A.; Washington, W.M. The importance of land-cover change in simulating future climates. *Science* **2005**, *310*, 1674–1678.
2. Moss, R.H.; Edmonds, J.A.; Hibbard, K.A.; Manning, M.R.; Rose, S.K.; van Vuuren, D.P.; Carter, T.R.; Emori, S.; Kainuma, M.; Kram, T.; *et al.* The next generation of scenarios for climate change research and assessment. *Nature* **2010**, *463*, 747–756.
3. Bontemps, S.; Herold, M.; Kooistra, L.; van Groenestijn, A.; Hartley, A.; Arino, O.; Moreau, I.; Defourny, P. Revisiting land cover observation to address the needs of the climate modeling community. *Biogeosciences* **2012**, *9*, 2145–2157.
4. Yang, J.; Gong, P.; Fu, R.; Zhang, M.; Chen, J.; Liang, S.; Xu, B.; Shi, J.; Dickinson, R. The role of satellite remote sensing in climate change studies. *Nat. Clim. Change* **2013**, *3*, 875–883.
5. Friedl, M.A.; McIver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; *et al.* Global land cover mapping from MODIS: Algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302.
6. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182.
7. Arino, O.; Bicheron, P.; Achard, F.; Latham, J.; Witt, R.; Weber, J.L. GlobCover the Most Detailed Portrait of Earth. Available online: http://www.esa.int/esapub/bulletin/bulletin136/bul136d_arino.pdf (accessed on 16 November 2014).
8. Arino, O., Ramos, J., Kalogirou, V., Defourny, P. and Achard, F. Globcover 2009. In Proceedings of the Living Planet Symposium, Bergen, Norway, 28 June–2 July 2010.
9. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.C.; Zhao, Y.Y.; Liang, L.; Niu, Z.G.; Huang, X.M.; Fu, H.H.; Liu, S.; *et al.* Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654.

10. Sexton, J.O.; Song, X.-P.; Feng, M.; Noojipady, P.; Anand, A.; Huang, C.; Kim, D.-H.; Collins, K.M.; Channan, S.; DiMiceli, C.; *et al.* 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. Digit. Earth* **2013**, *6*, 427–448.

11. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **2013**, *342*, 850–853.

12. Kennedy, R.E.; Yang, Z.G.; Cohen, W.B. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. Landtrendr—Temporal segmentation algorithms. *Remote Sens. Environ.* **2010**, *114*, 2897–2910.

13. Huang, C.Q.; Coward, S.N.; Masek, J.G.; Thomas, N.; Zhu, Z.L.; Vogelmann, J.E. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sens. Environ.* **2010**, *114*, 183–198.

14. Sexton, J.O.; Urban, D.L.; Donohue, M.J.; Song, C.H. Long-term land cover dynamics by multi-temporal classification across the landsat-5 record. *Remote Sens. Environ.* **2013**, *128*, 246–258.

15. Zhu, Z.; Woodcock, C.E.; Olofsson, P. Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sens. Environ.* **2012**, *122*, 75–91.

16. Zhu, Z.; Woodcock, C.E. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* **2014**, *144*, 152–171.

17. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 106–115.

18. Sun, F.D.; Zhao, Y.Y.; Gong, P.; Ma, R.H.; Dai, Y.J. Monitoring dynamic changes of global land cover types: Fluctuations of major lakes in China every 8 days during 2000–2010. *Chin. Sci. Bull.* **2014**, *59*, 171–189.

19. Loth, B.; Graf, H.F.; Oberhuber, J.M. Snow cover model for global climate simulations. *J. Geophys. Res. Atmos.* **1993**, *98*, 10451–10464.

20. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; *et al.* Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172.

21. Rouse, J.W., Haas, R.H., Deering, D.W., Schell, J.A.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Type III Final Report; NASA/GSFC: Greenbelt, MD, USA, 1974.

22. Haralick, R.M.; Shanmuga.K; Dinstein, I. Textural features for image classification. *IEEE Syst. Man Cy. Soc.* **1973**, *Smc3*, 610–621.

23. Gong, P.; Marceau, D.J.; Howarth, P.J. A comparison of spatial feature-extraction algorithms for land-use classification with SPOT HRV data. *Remote Sens. Environ.* **1992**, *40*, 137–151.

24. Caridade, C.M.R.; Marcal, A.R.S.; Mendonca, T. The use of texture for image classification of black & white air photographs. *Int. J. Remote Sens.* **2008**, *29*, 593–607.

25. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.

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

27. Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote Sens.* **2014**, *6*, 964–983.

28. Rodriguez-Galiano, V.F.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.M.; Jeganathan, C. Random forest classification of mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93–107.

29. Liu, D.; Cai, S. A spatial-temporal modeling approach to reconstructing land-cover change trajectories from multi-temporal satellite imagery. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 1329–1347.

30. Wang, J.; Zhao, Y.; Li, C.; Yu, L.; Liu, D.; Gong, P. Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250 m resolution. *ISPRS J. Photogramm. Remote Sens.* 2014, in press.

31. Liu, D.; Kelly, M.; Gong, P. A spatial-temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery. *Remote Sens. Environ.* **2006**, *101*, 167–180.

32. Liu, D.; Song, K.; Townshend, J.R.G.; Gong, P. Using local transition probability models in markov random fields for forest change detection. *Remote Sens. Environ.* **2008**, *112*, 2222–2231.

33. Besag, J. On the statistical-analysis of dirty pictures. *J. Roy. Stat. Soc. B Met.* **1986**, *48*, 259–302.

34. Tuia, D.; Volpi, M.; Copa, L.; Kanevski, M.; Munoz-Mari, J. A survey of active learning algorithms for supervised remote sensing image classification. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 606–617.

35. Huang, H.B.; Li, Z.; Gong, P.; Cheng, X.A.; Clinton, N.; Cao, C.X.; Ni, W.J.; Wang, L. Automated methods for measuring DBH and tree heights with a commercial scanning LiDAR. *Photogramm Eng. Remote Sens.* **2011**, *77*, 219–227.

36. Zhu, X.L.; Liu, D.S. Map-MRF approach to Landsat ETM plus SLC-off image classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1131–1141.

37. Bolin, D.; Lindstrom, J.; Eklundh, L.; Lindgren, F. Fast estimation of spatially dependent temporal vegetation trends using gaussian markov random fields. *Comput. Stat. Data Anal.* **2009**, *53*, 2885–2896.