

Article

Consistent Classification of Landsat Time Series with an Improved Automatic Adaptive Signature Generalization Algorithm

Matthew P. Dannenberg ^{1,*}, Christopher R. Hakkenberg ² and Conghe Song ^{1,2,*}

¹ Department of Geography, University of North Carolina, Chapel Hill, NC 27599, USA

² Curriculum for the Environment and Ecology, University of North Carolina, Chapel Hill, NC 27599, USA; hakkenbe@live.unc.edu

* Correspondence: mdannenb@live.unc.edu (M.P.D.); csong@email.unc.edu (C.S.); Tel.: +1-919-843-4764

Academic Editors: Soe W. Myint, Xiaofeng Li and Prasad S. Thenkabail

Received: 27 June 2016; Accepted: 15 August 2016; Published: 24 August 2016

Abstract: Classifying land cover is perhaps the most common application of remote sensing, yet classification at frequent temporal intervals remains a challenging task due to radiometric differences among scenes, time and budget constraints, and semantic differences among class definitions from different dates. The automatic adaptive signature generalization (AASG) algorithm overcomes many of these limitations by locating stable sites between two images and using them to adapt class spectral signatures from a high-quality reference classification to a new image, which mitigates the impacts of radiometric and phenological differences between images and ensures that class definitions remain consistent between the two classifications. We refined AASG to adapt stable site identification parameters to each individual land cover class, while also incorporating improved input data and a random forest classifier. In the Research Triangle region of North Carolina, our new version of AASG demonstrated an improved ability to update existing land cover classifications compared to the initial version of AASG, particularly for low intensity developed, mixed forest, and woody wetland classes. Topographic indices were particularly important for distinguishing woody wetlands from other forest types, while multi-seasonal imagery contributed to improved classification of water, developed, forest, and hay/pasture classes. These results demonstrate both the flexibility of the AASG algorithm and the potential for using it to produce high-quality land cover classifications that can utilize the entire temporal range of the Landsat archive in an automated fashion while maintaining consistent class definitions through time.

Keywords: Landsat; land cover and land use change (LCLUC); classification; random forest; signature extension; topography; National Land Cover Database (NLCD); time series

1. Introduction

The areal extents and distributions of different land cover types play a key role in many of the ecological and biophysical processes operating within the Earth system, including the absorption of solar radiation by the Earth surface [1], partitioning of absorbed radiation to latent and sensible heat fluxes [2], hydrologic processes such as infiltration and overland flow [3], biodiversity [4], and the movement of organisms within and among landscapes [5]. In addition, land cover maps serve as critical inputs for constraining many models of Earth system processes [6–8]. Remote sensing provides the means to map and monitor land cover and land use change (LCLUC) over large regions, at scales ranging from the regional [9] to the global [10–13].

Remote sensing of LCLUC possesses several challenges, particularly when classifying images from multiple dates. Developing accurate reference data for training and accuracy assessment of land

cover maps is an expensive and time-consuming task with tradeoffs that often force researchers to sacrifice ideal data collection procedures in order to remain within budget and time constraints [14]. The challenges inherent in the classification of land cover from single-date remotely sensed imagery are compounded when attempting to monitor LCLUC using multi-temporal imagery. For example, when classifiers are trained independently for each image, inconsistent class definitions among training datasets increase the potential for semantic errors in change detection. Differences in cover class designation between two maps from different dates could therefore result either from a genuine change in land cover or from spurious change due to inconsistent class definitions between the two maps. Automated approaches based on signature extension use class spectral libraries to map land cover in multiple images based on a single set of class-specific spectral signatures. These approaches attempt to enforce spectral consistency in class definitions between images, but are very sensitive to variation in illumination, sun-sensor geometry, and atmospheric and phenological conditions between images [15,16]. These signature extension approaches therefore depend on time-consuming atmospheric correction procedures as well as anniversary date imagery which may not be available in some regions due to frequent or persistent cloud cover. Unlocking the potential of remote sensing for mapping and monitoring LCLUC hinges upon the development of algorithms that overcome these limitations in a consistent and efficient manner.

The automatic adaptive signature generalization (AASG) algorithm [17] offers the potential to develop automated, multi-temporal land cover classifications while maintaining consistent class definitions and adjusting to radiometric variation among images. AASG uses a simple change detection technique to identify pixels that remain stable between two image dates, using these as the basis for extracting class-specific spectral signatures from subsequent, spatially coincident imagery and a high-quality reference classification. These generalized signatures adapt to the radiometric idiosyncrasies of the target imagery and serve as automated training data from which a new classification can be obtained, thus maintaining consistency in class definitions between images despite potentially large radiometric differences. Gray and Song [17] found that, unlike traditional signature extension, AASG can produce accurate land cover maps even when atmospheric effects went uncorrected and when large phenological differences existed between images (e.g., summer-winter image pairs in a temperate region).

In this study, we refined the AASG algorithm initially developed by Gray and Song [17] in order to increase accuracy and consistency of multi-temporal image classification, specifically through the use of stable site identification parameters that are tuned for each individual land cover class rather than globally defined thresholding parameters. We also incorporated improved input data, including multi-season imagery and additional topographic information, and a random forest (RF) classifier in place of maximum likelihood classification. We tested the updated version of AASG, hereafter referred to as AASG2, in the Raleigh-Durham-Chapel Hill region of North Carolina, USA using Landsat 5 TM imagery (path 16, row 35), a digital elevation model (DEM) from the Shuttle Radar Topography Mission (SRTM), and the National Land Cover Database (NLCD) for reference land cover classifications and comparison.

2. Materials and Methods

2.1. Automatic Adaptive Signature Generalization

The automatic adaptive signature generalization (AASG) algorithm automates the workflow for conducting multi-temporal land cover classifications with minimal reference data requirements [17]. AASG ensures consistency in class definitions between images by automatically adapting class spectral signatures to each image, thereby reducing the potential for semantic errors in LCLUC detection despite potentially large radiometric differences between images. Because frequent cloud cover may limit the availability of anniversary-date imagery, AASG's signature generalization approach allows temporally irregular imagery, including scenes in very different phenological states, to guide land cover change

detection. AASG operates under the assumption that within a sufficiently large scene, conversion from one land cover class to another is relatively rare, and thus the majority of the landscape will remain stable between the dates of a reference image (I_1) and a spatially coincident target image (I_2).

Stable sites are selected in two steps. First, a larger group of potential stable sites are selected from those pixels located near the mean of the image difference histogram ($I_2 - I_1$; Figure 1). Specifically, Gray and Song [17] identified stable pixels as those falling within a predefined number of standard deviations from the mean of the image difference histograms, such that $\mu_k \pm c \times \sigma_k$, where μ_k and σ_k are the mean and standard deviation of the difference values for class k , respectively, and c is a globally defined thresholding parameter determining the width of the interval. Second, due to the potential for confusion arising from image misregistration and edge effects, especially along class boundaries in multi-temporal imagery, a class-specific erode filter based on a user-defined kernel is employed to ensure that only core pixels are retained as stable sites.

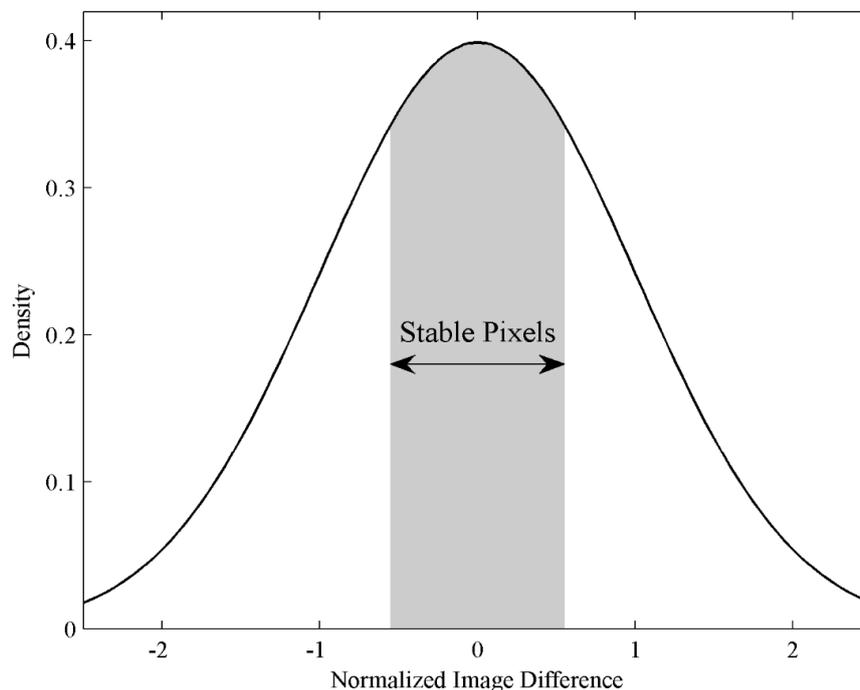


Figure 1. Identification of stable pixels based on image difference histograms ($I_2 - I_1$) for each land cover class. Stable pixels are those within a given distance from the mean (μ_k) of class k .

Once the sets of pixels corresponding to stable sites between the training and target images are determined, they can be coupled with their respective class labels from a high-quality reference classification (C_1) and used to develop class-specific spectral signatures in the target image (I_2). While the spectral signatures of stable cover classes may have changed greatly between the two image dates due to variation in atmospheric and ground moisture conditions, vegetation phenology, illumination/view geometry, etc., the identity of those classes will have remained unchanged. Thus, the class-specific spectral signature sampled from stable sites in the target image (I_2) can be used as adaptive training data (in conjunction with the consistent class definitions determined from the reference classification) to parameterize a unique classifier for the entire target image (including non-stable sites) (Figure 2). The generalized signature derived for each target image is adaptive in that it reflects the specific radiometric profile of the target image, and automated because sampling in all subsequent imagery is entirely algorithmic and obviates the need for user-delineated training data and image-specific radiometric correction procedures. The accuracy and efficiency of AASG for performing multi-temporal image classification makes it especially attractive for large-scale operational approaches.

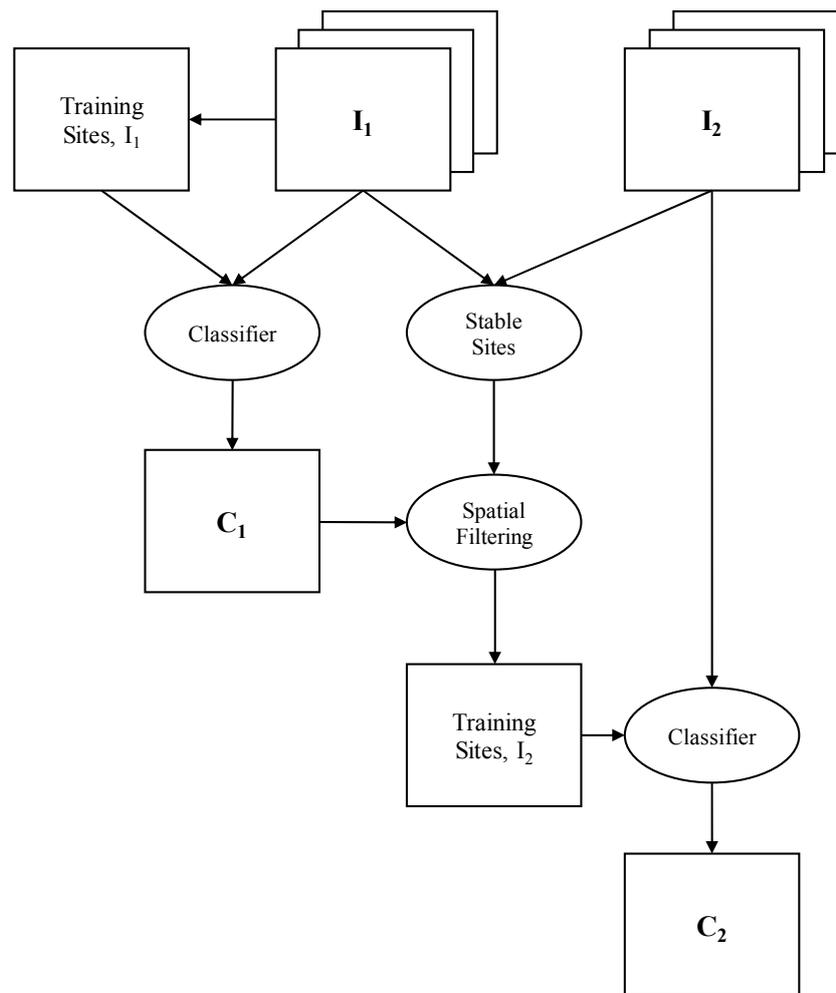


Figure 2. Workflow for the AASG method. Adapted from Gray and Song [17].

2.2. Class-Specific Thresholds for Stable Site Identification

In the initial version of AASG, hereafter referred to as AASG1, Gray and Song [17] tested global c thresholding parameters of 0.25, 0.5, and 1.0. However, for scenes with large variation in both the prevalence and within-class heterogeneity of land cover classes, these globally defined parameters may be inadequate. Rather, the thresholds for identifying stable pixels can be tuned for each land cover class to balance the needs for an adequate training sample size while minimizing the probability of contaminating the training set with mislabeled pixels. We therefore tested a simple approach that scales the proportion of pixels identified as stable for a given land cover class as a function of the total number of pixels in the scene that are assigned to that class in the reference classification. For land cover class k with N_k pixels in the reference classification, we define a set proportion, p_k , of N_k to identify as stable sites:

$$p_k = \begin{cases} 0.5, & \text{if } N_k \leq 100 \\ b - a \times \log_{10} N_k, & \text{if } 100 < N_k \leq 10 \cdot n_{max} \\ \frac{n_{max}}{N_k}, & \text{if } N_k > 10 \cdot n_{max} \end{cases} \quad (1)$$

where n_{max} is a user-defined maximum number of pixels that can be identified as stable, and a and b are parameters that scale p_k between 0.1 and 0.5. The proportion of pixels considered stable therefore declines as the prevalence of a given class increases, reflecting the need to achieve an adequately large sample size for less prevalent land cover classes while reducing the probability of incorrectly

including unstable pixels in the training set as class prevalence increases. Thus, for land cover classes that constitute a very small portion of the landscape ($N_k \leq 100$), half of the pixels are used as candidate training pixels in AASG2 (Figure 3). However, we note that for such classes there may be high variance in the training set due to small N_k and a higher likelihood of including unstable and mislabeled pixels, and it may be appropriate to consider removing such a class from further analyses or to combine it with a similar class. For land cover classes that comprise a large proportion of the entire scene ($N_k > 10 \cdot n_{max}$), a maximum of n_{max} pixels are identified as stable and used as training candidates. For classes with intermediate representation in the landscape, we scale the proportion of pixels used as candidates from 0.5 to 0.1 as N_k increases. Once p_k has been determined, the $p_k \times N_k$ pixels nearest to the class mean of the image difference histogram are considered to be stable. As in AASG1, the stable site maps are eroded using a 3 pixel by 3 pixel filter window to minimize errors due to potential image misregistration and mixed pixels along class boundaries.

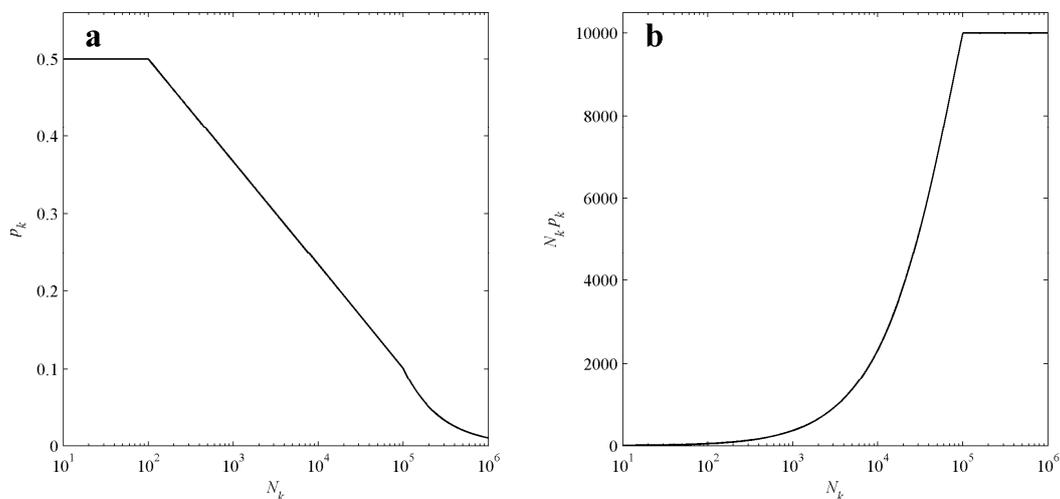


Figure 3. Determination of class-specific thresholding factors. (a) The proportion of pixels used as stable sites (p_k) as a function of the total number of pixels per class (N_k); (b) The total number of pixels used as stable sites ($N_k p_k$) as a function of the total number of pixels per class (N_k). Both graphs (a) and (b) are shown with $n_{max} = 10,000$ pixels.

2.3. Topographic Metrics

Topography is an important driver of soil moisture [18], vegetation distribution [19,20], and human land use [21]. Since topography plays such an important role in the movement of water over and through landscapes, topographic indices can improve classification of land cover classes, particularly those characterized by surface water [22] and wetlands [23,24]. In AASG1, the topographic wetness index (TWI) was used as ancillary input data. In AASG2, we used a 30 m DEM from SRTM to derive an expanded suite of topographic indices for use as predictor variables. In addition to elevation and slope, we also used the upslope accumulated area (UAA), a modified UAA index (mUAA), and the TWI as predictor variables in AASG2 classifications. UAA and TWI capture important aspects of water flow and accumulation and are frequently used in watershed hydrology and spatial scaling of ecosystem fluxes [25–28]. We calculated UAA, mUAA, and TWI using the System for Automated Geoscientific Analyses (SAGA), a freely available geographic information system [29].

2.4. Multi-Season Imagery

In AASG2, we utilize imagery from multiple seasons per year as training data to increase the robustness of classification predictions of target imagery. The use of multi-season, rather than single-date imagery, mitigates against the potential for confusion and spurious classification especially

for land cover classes that demonstrate large intra-annual variability and ephemeral reflectance characteristics [9,30]. The use of multi-season imagery to inform class spectral signatures is especially effective for classifications of phenologically variable cover types like agriculture. Without the ability of multi-season imagery to capture a potentially short growing season for certain crops, a single-date image from a pre-leaf-out or post-harvest date would likely fail to adequately capture crop cover [31]. Similarly, the inclusion of both leaf-off and leaf-on imagery from temperate regions allows for better prediction of sub-canopy cover types, such as ephemeral water bodies or impervious surfaces, that would otherwise be obscured by the upper canopy in leaf-on imagery [32–34]. The inclusion of multi-season imagery for training and prediction is in keeping with other aspects of the AASG algorithm which seek to minimize ephemeral reflectance characteristics and instead capture a generalized spectral signature that is consistent for the entire time period of the training dataset.

2.5. Random Forest Classification

In AASG2, we replaced the maximum likelihood classifier used in AASG1 with a random forest (RF) classifier. The RF, which consists of an ensemble of classification trees, is a machine-learning technique that can be used in a variety of contexts for classification of multidimensional datasets [35]. Each “tree” in the “forest” is grown from a bootstrapped sample of the observations with the remaining observations left “out-of-bag” (OOB). The final classification for a new or withheld observation is then determined based on the majority vote of the trees. The OOB observations are also used to estimate the importance of each predictor variable by randomly permuting each predictor, thus destroying its information content, and measuring the increase in error of OOB observations as compared to the original classifier. RF classifiers are not as susceptible as individual classification trees to over-fitting or “overtraining” [35–38], and they are especially good at modeling nonlinearity and interactions among predictor variables [39]. RF classifiers have demonstrated higher predictive accuracy than more traditional classification methods in both ecology [39] and remote sensing [40], and have demonstrated comparable predictive accuracy to other machine-learning methods (e.g., support vector machines) [36,41]. In addition, RF models are relatively robust to error or noise in training datasets [35,37,42], making them particularly appealing for mitigating potential errors in stable site identification or assignment of class labels to stable sites.

2.6. Study Area and Remotely Sensed Data

We evaluated the performance of AASG2 in a 60 km × 60 km study area of the Research Triangle region in North Carolina, USA. The Triangle has experienced rapid population growth and land cover change over the past several decades, making it an ideal study area for validation of AASG2. Specifically, over the past thirty years, the Triangle Region experienced accelerating LCLUC trends driven by urbanization and interacting trends in timber and agricultural markets [43]. Changes were unevenly distributed throughout the study area, with the greatest cover change occurring due to urbanization near cities like Raleigh, Durham, Cary, and Apex, primarily at the expense of surrounding agricultural fields and secondary forests [9]. Agricultural and forest land cover—including planted pine, oak-pine mixed, and upland hardwood forests—experienced large transformations prior to and during the study period owing to urban development as well as conversion among the two cover types following market forces [9,43].

We obtained atmospherically corrected Landsat 5 TM surface reflectance over our study area from the LEDAPS archive [44,45]. While atmospheric correction and conversion to surface reflectance are not necessary for the development of high-quality land cover classifications with AASG [17], it is necessary for the accurate calculation of spectral vegetation indices. We selected early-, mid-, and late-season images from three single-year windows centered on the target classification dates of 2006 and 2011 (Table 1). While we attempted to collect anniversary date images for each season, image dates were offset by a month or more due to the presence of clouds. Cloud-free summer images were unavailable in 2006 and 2011, so we instead used images from 2007 and 2010, respectively. We transformed each

image in the multi-season stack into “brightness” (KT1), “greenness” (KT2), and “wetness” (KT3) components using the Kauth-Thomas tasseled cap transformation [46,47].

Table 1. Landsat 5 TM images used in this study.

Year	Season		
	Early	Mid	Late
2001		18 September 2001	
2006	9 April 2006	18 August 2007	5 December 2006
2011	7 April 2011	10 August 2010	1 November 2011

For reference and comparison with the AASG classifications, we obtained the National Land Cover Database (NLCD) classifications of our study region for the years 2001 [48], 2006 [49], and 2011 [50]. We excluded developed open space (class 21) from further analyses since it is spectrally similar to other grass-dominated classes [51,52] and classification requires additional contextual information such as municipal boundaries [9]. This class was masked from the reference classification, and any pixels classified as developed open space in NLCD were reassigned to their most likely alternative class based on the spectral signatures of the remaining NLCD classes.

2.7. Analyses

AASG requires a high-quality reference classification in order to extract class spectral signatures for a new image. Here, we used the 2001 NLCD as the reference classification (C_1) with a 2001 mid-season Landsat reference image (I_1). We then used both AASG1 and AASG2 to classify land cover for 2006 and 2011. First, for AASG1, we followed the original version proposed by Gray and Song [17]. We used a global c thresholding parameter of 0.25 to identify stable sites in image difference histograms based on the reflectance in the red band (Landsat 5 TM band 3). In AASG1, we used a maximum likelihood classifier with ten predictors, including mid-season surface reflectance from the six mid-season solar reflective bands of Landsat TM plus four additional spectral transformations and ancillary datasets used in Gray and Song [17]: the normalized difference vegetation index, the simple ratio vegetation index, the structural index [53], and TWI from SRTM. Second, for AASG2, we used the updated procedures detailed in sections 2.2–2.5. Stable sites were determined from image difference histograms calculated from mid-season KT1 using class-specific thresholding parameters. We tested several values of n_{max} for determining stable sites. AASG2 was based on an RF classifier with 14 predictors: three KT indices from each image for the three seasons plus five topographic indices (elevation, slope, UAA, mUAA, and TWI).

Gray and Song [17] showed that AASG is capable of producing accurate land cover maps, with overall accuracies that meet or exceed those of traditional signature extension approaches. To demonstrate that AASG2 further improves the ability to automatically update existing land cover classifications, we compared predicted land cover from AASG1 and AASG2 to NLCD land cover for 2006 and 2011, which were not used in the training process. While the NLCD also contains error and cannot be regarded as a “true” source of reference data, these comparisons provide a good benchmark for testing whether AASG2 improves the ability to extend high-quality classifications to other time periods compared to AASG1. Following standard accuracy assessment procedures, we present these results in confusion matrices, which are normalized by the proportion of the landscape that is mapped as each class [54]. From these confusion matrices, we calculated overall agreement (OA), user’s agreement (UA), and producer’s agreement (PA) between the AASG and NLCD classifications for each year.

3. Results

AASG1 and AASG2 were generally successful in reproducing the land cover patterns of the NLCD in both 2006 (Figure 4) and 2011 (Figure 5). The major water bodies, which include Falls Lake in

the northeast and Jordan Lake in the southwest, were clearly evident in all classifications. Both AASG algorithms reproduced extensive areas of woody wetlands to the north of Jordan Lake and west of Falls Lake in both years, though with varying degrees of success, as discussed below. The AASG models also reproduced extensive agricultural fields, primarily hay/pasture, in the northwestern portion of the study region. The major metropolitan areas of Raleigh, Durham, and Chapel Hill were successfully mapped by both AASG models.

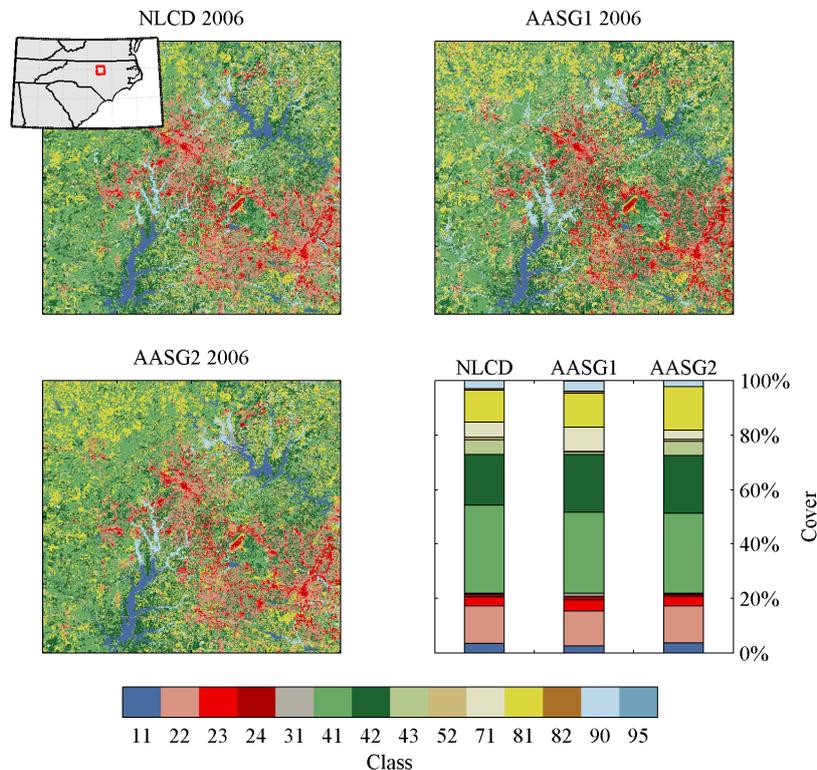


Figure 4. 2006 land cover from NLCD, AASG1, and AASG2.

With few exceptions, the total areal extent of each class as mapped by AASG2 was more similar to the NLCD class extents than AASG1 (Figures 4 and 5, and Tables 2 and 3). This improvement is particularly evident for low intensity development, mixed forest, grassland/herbaceous, and woody wetlands. The extents of low intensity development and mixed forest were substantially underestimated by AASG1, while AASG2 and the NLCD mapped very similar proportions of these cover classes in both 2006 and 2011. The extent of grassland/herbaceous was overestimated by AASG1 and underestimated by AASG2 in both 2006 and 2011 compared to the NLCD. AASG1 overestimated the extent of woody wetlands in 2006 but underestimated its extent in 2011, while AASG2 produced woody wetland extents that were more stable from year to year and more similar to the NLCD. The most notable differences in proportional land cover between the NLCD and AASG2 occurred in the two agricultural classes. Relative to the NLCD, AASG2 overestimated the extent of hay/pasture and underestimated the extent of cultivated crops.

Overall agreement between AASG2 and the NLCD was higher than that between AASG1 and the NLCD in both 2006 (70.2% vs. 60.2%; Table 2, Tables S1 and S2) and 2011 (67.9% vs. 60.1%; Table 3, Tables S3 and S4). When aggregated from the modified Anderson Level 2 classification system to Anderson Level 1 (which combines the sub-divided developed, forest, agriculture, and wetland classes into a single class each), the agreement with the NLCD increased to 74.8% in 2006 and 76.1% in 2011 for AASG1 and 81.8% in 2006 and 80.9% in 2011 for AASG2. Overall agreement between the NLCD and AASG2 exhibited a monotonic increase with increasing n_{max} (Figure 6), suggesting that the RF

model benefits from having a large training sample set that captures a large range of within-class variability. The ability to identify a large training dataset is a major advantage of AASG. Traditional manual delineation of training sets typically attempts to find the purest pixels representing a land cover type, therefore losing the inherent within-class variability. The high within-class variability included in large training samples would be a major challenge for conventional parametric classifiers, such as the maximum likelihood classifier. However, since the RF is a nonparametric machine-learning algorithm, the large within-class variability can be used in the classification, leading to improved classification accuracy.

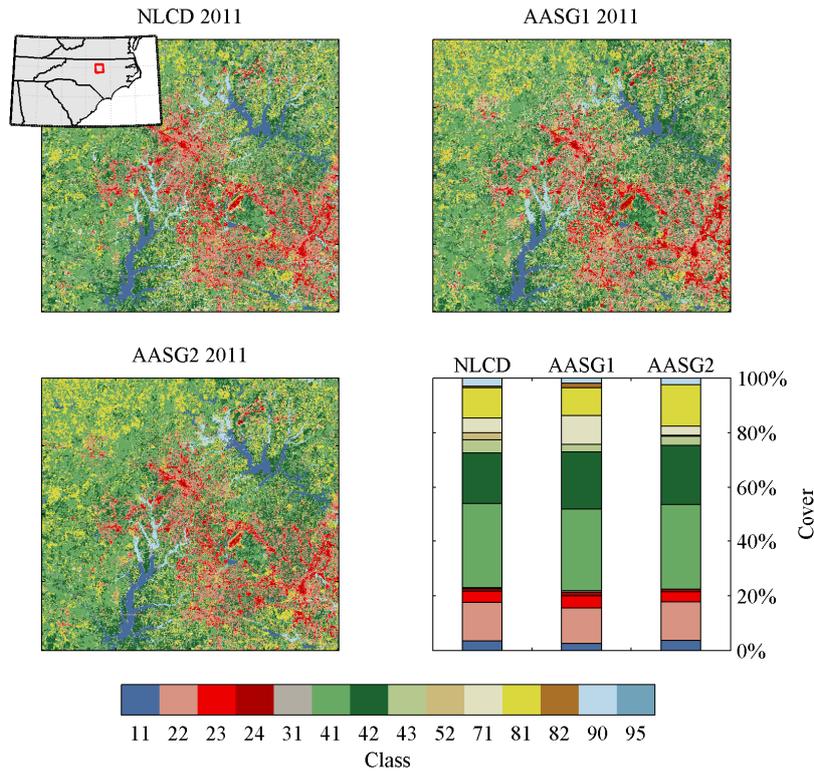


Figure 5. 2011 land cover from NLCD, AASG1, and AASG2.

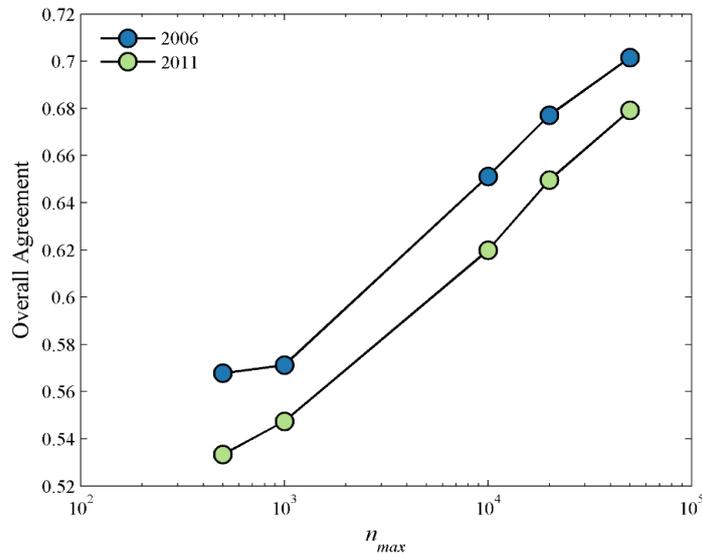


Figure 6. Overall agreement between NLCD and AASG2 as a function of n_{max} .

Table 2. Comparison of proportional land cover, overall agreement (OA), producer's agreement (PA), and user's agreement (UA) in 2006. Full confusion matrices are presented in Tables S1–S2.

Class	NLCD	AASG1 (OA = 60.2%)			AASG2 (OA = 70.2%)		
	Cover	Cover	PA	UA	Cover	PA	UA
11 Open water	3.4%	2.5%	70.4%	97.5%	3.6%	89.7%	86.9%
22 Low intensity developed	13.7%	12.8%	63.8%	68.4%	13.6%	73.0%	73.4%
23 Medium intensity developed	3.4%	4.0%	55.9%	46.9%	3.4%	60.6%	61.0%
24 High intensity developed	0.9%	1.3%	57.7%	40.5%	0.9%	60.3%	61.5%
31 Barren land	0.3%	1.2%	12.7%	3.2%	0.4%	9.4%	7.2%
41 Deciduous Forest	32.5%	29.9%	68.0%	74.1%	29.6%	75.8%	83.3%
42 Evergreen forest	18.6%	21.3%	69.4%	60.8%	21.1%	79.5%	70.2%
43 Mixed forest	5.3%	0.9%	2.5%	14.8%	5.2%	36.2%	36.7%
52 Shrub/scrub	1.0%	0.1%	0.4%	6.2%	0.7%	16.7%	23.6%
71 Grassland/herbaceous	5.6%	8.9%	29.3%	18.4%	3.4%	23.5%	38.7%
81 Pasture/hay	11.6%	12.5%	73.4%	68.0%	15.9%	87.4%	63.6%
82 Cultivated crops	0.7%	0.8%	9.4%	8.5%	0.0%	2.6%	53.0%
90 Woody wetlands	2.9%	3.7%	41.4%	32.1%	2.2%	50.7%	65.7%
95 Emergent herbaceous wetlands	0.1%	0.2%	11.1%	5.2%	0.0%	3.9%	20.5%

Table 3. Comparison of proportional land cover, overall agreement (OA), producer's agreement (PA), and user's agreement (UA) in 2011. Full confusion matrices are presented in Tables S3–S4.

Class	NLCD	AASG1 (OA = 60.1%)			AASG2 (OA = 67.9%)		
	Cover	Cover	PA	UA	cover	PA	UA
11 Open water	3.5%	2.6%	70.6%	95.3%	3.6%	92.6%	88.3%
22 Low intensity developed	14.1%	13.0%	64.9%	70.3%	14.1%	73.5%	73.1%
23 Medium intensity developed	4.0%	4.4%	54.4%	49.8%	3.5%	55.7%	62.8%
24 High intensity developed	1.1%	1.4%	57.6%	43.4%	1.0%	55.8%	60.6%
31 Barren land	0.3%	0.5%	6.0%	2.9%	0.1%	3.8%	12.2%
41 Deciduous Forest	31.1%	30.1%	70.9%	73.4%	31.3%	76.4%	76.0%
42 Evergreen forest	18.5%	21.0%	72.3%	63.7%	21.7%	78.3%	66.8%
43 Mixed forest	4.9%	2.7%	6.6%	12.2%	3.3%	22.6%	34.0%
52 Shrub/scrub	2.6%	0.0%	0.0%	5.2%	0.3%	3.8%	31.7%
71 Grassland/herbaceous	5.4%	10.6%	30.0%	15.3%	3.4%	19.6%	31.1%
81 Pasture/hay	11.1%	10.3%	66.0%	71.4%	15.2%	86.0%	62.6%
82 Cultivated crops	0.6%	1.6%	12.0%	4.3%	0.0%	2.7%	63.9%
90 Woody wetlands	2.9%	1.9%	31.3%	48.3%	2.4%	50.7%	60.2%
95 Emergent herbaceous wetlands	0.1%	0.1%	4.4%	9.7%	0.0%	1.7%	41.9%

Both AASG1 and AASG2 exhibited relatively high agreement with the NLCD for the surface water class ($UA \geq 86.9\%$ and $PA \geq 70.4\%$), which tends to be spectrally distinct from other classes and exhibits low within-class spectral variability, though AASG1 had a much higher UA than PA for this class. AASG1 had relatively low agreement with NLCD on low intensity developed ($UA \leq 70.3\%$ and $PA \leq 64.9\%$). Pixels classified as low intensity developed by the NLCD were most often confused with medium intensity developed, evergreen forest, and grassland/herbaceous by AASG1. For both versions of AASG, there was considerable confusion among forest classes (particularly with mixed forest), though AASG2 had higher UA and PA than AASG1 for every forest class in both 2006 and 2011. When aggregated to a single forest class, both the AASG1 and the AASG2 classifications showed high agreement with NLCD classifications ($UA \geq 87.4\%$ and $PA \geq 81.7\%$). AASG2 offered considerable improvement over AASG1 in the ability to map woody wetlands, which were frequently confused with deciduous forest by AASG1. Both AASG1 and AASG2 frequently disagreed with NLCD on mapping the least prevalent classes: barren land, shrub/scrub, grassland/herbaceous, cultivated crops, and emergent herbaceous wetland. For a given land cover class, the magnitudes of UA and PA from AASG2 classifications were typically quite similar. Asymmetries between UA and PA most often occurred for classes with low prevalence in the landscape (e.g., cultivated crops and emergent herbaceous wetland),

with the exceptions of evergreen forest and pasture/hay, each of which had higher total areal extents in the AASG2 classifications than in NLCD and which had higher PA than UA.

The OOB training data from the RF model indicated that the mid-season “greenness” index (KT2) was the most important predictor variable for classification of land cover in our study region while, on average, topographic indices did not improve agreement between AASG2 and the NLCD (Figure 7). However, predictor variable importance varied substantially among classes. For classifying surface water, the accuracy of AASG2 declined by more than 20% when the KT1 “brightness” index from any of the three seasons was excluded. For the developed and forest classes, mid-season KT2 contributed most to AASG2 accuracy, though early- and late-season KT2 were also important predictors for the high intensity developed class. Multi-season imagery—including early- and mid-season KT1, mid-season KT3, and late-season KT2—was particularly important for mapping hay/pasture. The annual cycles of planting and harvesting likely imposed a phenological signal in this class that distinguished it from spectrally similar classes like grassland/herbaceous and low intensity developed, and the multi-season imagery was therefore able to separate these classes in AASG2.

While topography played a relatively minor role in mapping most classes, it was among the most important variables for mapping woody wetlands. This class tends to be spectrally similar to the other forest classes but occurs in distinctly different, lower elevation landscape positions where water is likely to accumulate and allow periodic saturation. Since high mUAA and TWI indicate regions where water is likely to accumulate, these two variables, along with elevation, improved the agreement between the NLCD and AASG2 (UA \geq 60.2%; PA \geq 50.7%). While this class was still frequently confused with other forest classes (particularly deciduous forest), the topographic indices included in AASG2 offered substantial gains in agreement over AASG1, which exhibited very little skill in mapping woody wetlands (UA \leq 48.3%; PA \leq 41.4%).

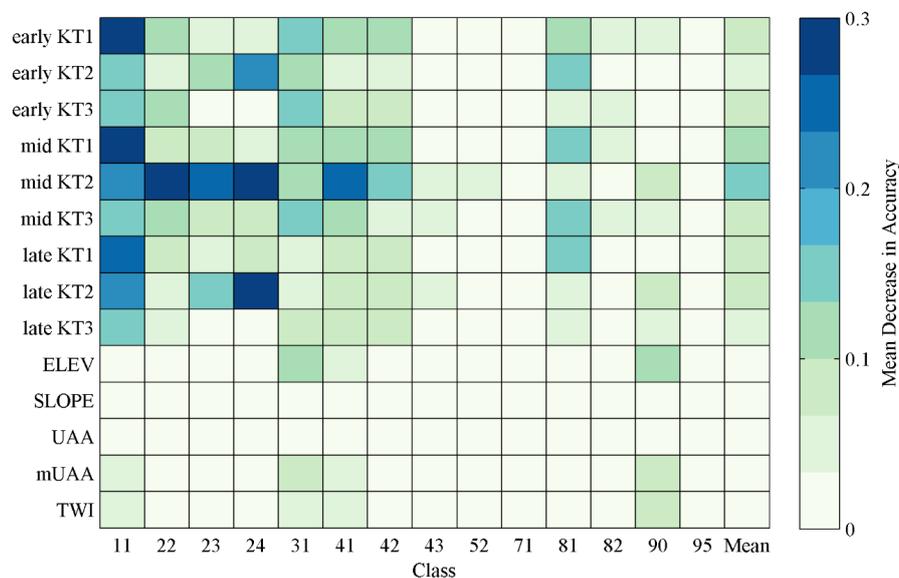


Figure 7. Mean decrease in agreement between NLCD and AASG2 when a given variable is excluded.

4. Discussion

Fully realizing the potential of the global remotely sensed archive for land cover classification depends on the development of efficient, accurate algorithms that can overcome challenges associated with radiometric differences among scenes, phenological differences in non-anniversary date imagery, and semantic differences among class definitions that occur when classifiers are trained independently. AASG overcomes many of the challenges inherent in classification of multi-temporal remotely sensed imagery by mitigating the needs for atmospheric correction and anniversary date imagery through the

adjustment of spectral signatures to each individual image, by maintaining consistent class definitions throughout the image time series, and by providing an automated workflow.

The viability of AASG for multi-temporal image classification has been demonstrated in previous research [17]. Compared to a photo-interpreted reference dataset, the initial version of AASG (AASG1) achieved an overall accuracy of at least 66%, which matched or exceeded the accuracy achieved with traditional signature extension approaches [17]. In this study, we demonstrated that our updated AASG2 workflow produced maps with higher overall agreement with the withheld NLCD classifications in both 2006 and 2011 compared to AASG1. Additional experiments confirm that both the updated input data (i.e., multi-season imagery and additional topographic metrics) and the updated methodology (i.e., class-specific stable site identification parameters and the RF classifier) contributed to the overall improvement of AASG2 (Tables S5–S6). Most of the improved performance in 2006 was attributable to updated input data (Table S5), while the improvement in 2011 was mostly attributable to the updated stable site identification and classification methodologies (Table S6).

The improvements made to the AASG workflow further enhance the potential for developing consistent, temporally dense time series of land cover classifications from high-quality reference maps, such as the NLCD. Applications that require frequent land cover maps could therefore use AASG to extend the NLCD, which has a large spatial extent but low temporal frequency, over the entire Landsat archive in an automated manner that maintains consistent class definitions among different image dates. This research also highlights the flexibility of the AASG method for application with a variety of potential classifiers, including traditional parametric methods (e.g., maximum likelihood classification) and nonparametric machine-learning methods like random forests, neural networks, or support vector machines. AASG also handles as many or as few input variables as are available to the user, so long as those inputs are spatially compatible with one another (i.e., same resolution, extent, and projection).

While we explored a simple approach for generating class-specific stable site thresholding parameters, future refinements of AASG could implement approaches that optimize the selection of stable sites to maximize agreement with withheld reference data or to minimize overlap of spectral signatures among classes. Additionally, to date, AASG has been applied as a “hard” classifier (i.e., where each pixel is mapped as a single class). Previous research has demonstrated that sub-pixel land cover fractions can be derived from fuzzy membership functions [55,56] or the posterior probabilities of class membership from hard classifiers [57–59]. The AASG algorithm developed in this research could therefore be adapted for prediction of sub-pixel land cover fractions using the posterior “vote” distributions from the individual classification trees in ensemble classifiers like the RF.

5. Conclusions

In this study, we improved the automatic adaptive signature generalization (AASG) algorithm through incorporation of a simple procedure for adapting stable site identification parameters to each land cover class (in place of globally defined thresholding parameters). We also incorporated improved input data (multi-season imagery and an expanded suite of topographic data) and a nonparametric random forest machine-learning classifier. We tested the updated algorithm in the Research Triangle region of North Carolina, where our refinements to AASG resulted in substantial increases in agreement with the National Land Cover Database (NLCD) maps from 2006 and 2011 compared to the original formulation. The areal extent of each class in the withheld NLCD classifications was better replicated by AASG2 than by AASG1, and these areal extents tended to remain more consistent and stable from year to year in the AASG2 classifications. Compared to AASG1, overall agreement between the withheld NLCD classifications and AASG2 was 10.0% higher in 2006 and 7.8% higher in 2011. The increase in agreement was concentrated within several classes (including woody wetlands, low intensity developed, and mixed forests) that AASG1 struggled to classify but that AASG2 skillfully reproduced. The inclusion of additional topographic indices improved the ability to distinguish woody wetlands from other forest classes, and the use of multi-season imagery improved the ability of AASG to classify water, developed classes, forest classes, and hay/pasture. These results demonstrate both the flexibility of the AASG approach for multi-temporal classification as well as the potential for

extending high-quality land cover classifications forward and backward in time with a simple and automated method.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/8/8/691/s1, Table S1: Comparison of 2006 NLCD to 2006 AASG1, Table S2: Comparison of 2006 NLCD to AASG2, Table S3: Comparison of 2011 NLCD to 2011 AASG1, Table S4: Comparison of 2011 NLCD to 2011 AASG2, Table S5: Overall agreement for different combinations of AASG data and methodology in 2006, Table S6: Overall agreement for different combinations of AASG data and methodology in 2011.

Acknowledgments: This study was partially supported by the U.S. National Science Foundation (Grant No. DEB-1313756). The authors thank three anonymous reviewers, whose comments and suggestions improved the manuscript. User guide and source code for AASG can be downloaded at <http://csong.web.unc.edu/downloads/>.

Author Contributions: Matthew Dannenberg, Christopher Hakkenberg, and Conghe Song conceived and designed the experiments; Matthew Dannenberg performed the experiments and analyzed the data; Matthew Dannenberg, Christopher Hakkenberg, and Conghe Song discussed and interpreted results; Matthew Dannenberg and Christopher Hakkenberg wrote the paper.

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

References

1. Bonan, G.B. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* **2008**, *320*, 1444–1449. [[CrossRef](#)] [[PubMed](#)]
2. Bounoua, L.; Defries, R.; Collatz, G.J.; Sellers, P.; Khan, H. Effects of land cover conversion on surface climate. *Clim. Change* **2002**, *52*, 29–64. [[CrossRef](#)]
3. Bronstert, A.; Niehoff, D.; Bürger, G. Effects of climate and land-use change on storm runoff generation: Present knowledge and modelling capabilities. *Hydrol. Process.* **2002**, *16*, 509–529. [[CrossRef](#)]
4. Newbold, T.; Hudson, L.N.; Hill, S.L.L.; Contu, S.; Lysenko, I.; Senior, R.A.; Borger, L.; Bennett, D.J.; Choimes, A.; Collen, B.; et al. Global effects of land use on local terrestrial biodiversity. *Nature* **2015**, *520*, 45–50. [[CrossRef](#)] [[PubMed](#)]
5. Nagendra, H.; Munroe, D.K.; Southworth, J. From pattern to process: Landscape fragmentation and the analysis of land use/land cover change. *Agric. Ecosyst. Environ.* **2004**, *101*, 111–115. [[CrossRef](#)]
6. Bonan, G.B.; Oleson, K.W.; Vertenstein, M.; Levis, S.; Zeng, X.; Dai, Y.; Dickinson, R.E.; Yang, Z.-L. The land surface climatology of the community land model coupled to the NCAR community climate model. *J. Clim.* **2002**, *15*, 3123–3149. [[CrossRef](#)]
7. Bonan, G.B.; Levis, S.; Sitch, S.; Vertenstein, M.; Oleson, K.W. A dynamic global vegetation model for use with climate models: Concepts and description of simulated vegetation dynamics. *Glob. Chang. Biol.* **2003**, *9*, 1543–1566. [[CrossRef](#)]
8. Running, S.W.; Nemani, R.R.; Heinsch, F.A.; Zhao, M.; Reeves, M.; Hashimoto, H. A continuous satellite-derived measure of global terrestrial primary production. *Bioscience* **2004**, *54*, 547–560. [[CrossRef](#)]
9. Sexton, J.O.; Urban, D.L.; Donohue, M.J.; Song, C. Long-term land cover dynamics by multi-temporal classification across the Landsat-5 record. *Remote Sens. Environ.* **2013**, *128*, 246–258. [[CrossRef](#)]
10. 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. [[CrossRef](#)] [[PubMed](#)]
11. Loveland, T.R.; Reed, B.C.; Brown, J.F.; Ohlen, D.O.; Zhu, Z.; Yang, L.; Merchant, J.W. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens.* **2000**, *21*, 1303–1330. [[CrossRef](#)]
12. Hansen, M.C.; DeFries, R.S.; Townshend, J.R.G.; Sohlberg, R. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* **2000**, *21*, 1331–1364. [[CrossRef](#)]
13. 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. [[CrossRef](#)]
14. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2009.
15. Song, C.; Woodcock, C.E. Monitoring forest succession with multitemporal Landsat images: Factors of uncertainty. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2557–2567. [[CrossRef](#)]

16. Pax-Lenney, M.; Woodcock, C.E.; Macomber, S.A.; Gopal, S.; Song, C. Forest mapping with a generalized classifier and Landsat TM data. *Remote Sens. Environ.* **2001**, *77*, 241–250. [[CrossRef](#)]
17. Gray, J.; Song, C. Consistent classification of image time series with automatic adaptive signature generalization. *Remote Sens. Environ.* **2013**, *134*, 333–341. [[CrossRef](#)]
18. Western, A.W.; Grayson, R.B.; Blöschl, G.; Willgoose, G.R.; McMahon, T.A. Observed spatial organization of soil moisture and its relation to terrain indices. *Water Resour. Res.* **1999**, *35*, 797–810. [[CrossRef](#)]
19. Kreft, H.; Jetz, W. Global patterns and determinants of vascular plant diversity. *Proc. Natl. Acad. Sci. USA.* **2007**, *104*, 5925–5930. [[CrossRef](#)] [[PubMed](#)]
20. Currie, D.J.; Paquin, V. Large-scale biogeographical patterns of species richness of trees. *Nature* **1987**, *329*, 326–327. [[CrossRef](#)]
21. Wear, D.N.; Bolstad, P. Land-use changes in southern Appalachian landscapes: Spatial analysis and forecast evaluation. *Ecosystems* **1998**, *1*, 575–594. [[CrossRef](#)]
22. Feng, M.; Sexton, J.O.; Channan, S.; Townshend, J.R. A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic-spectral classification algorithm. *Int. J. Digit. Earth* **2016**, *9*, 113–133. [[CrossRef](#)]
23. Ozesmi, S.L.; Bauer, M.E. Satellite remote sensing of wetlands. *Wetl. Ecol. Manag.* **2002**, *10*, 381–402. [[CrossRef](#)]
24. Corcoran, J.M.; Knight, J.F.; Gallant, A.L. Influence of multi-source and multi-temporal remotely sensed and ancillary data on the accuracy of random forest classification of wetlands in northern Minnesota. *Remote Sens.* **2013**, *5*, 3212–3238. [[CrossRef](#)]
25. Beven, K.J.; Kirkby, M.J. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. J.* **1979**, *24*, 43–69.
26. Band, L.E.; Patterson, P.; Nemani, R.; Running, S.W. Forest ecosystem processes at the watershed scale: Incorporating hillslope hydrology. *Agric. For. Meteorol.* **1993**, *63*, 93–126. [[CrossRef](#)]
27. Emanuel, R.E.; Epstein, H.E.; McGlynn, B.L.; Welsch, D.L.; Muth, D.J.; D’Odorico, P. Spatial and temporal controls on watershed ecohydrology in the northern Rocky Mountains. *Water Resour. Res.* **2010**, *46*. [[CrossRef](#)]
28. Riveros-Iregui, D.A.; McGlynn, B.L. Landscape structure control on soil CO₂ efflux variability in complex terrain: Scaling from point observations to watershed scale fluxes. *J. Geophys. Res.* **2009**, *114*. [[CrossRef](#)]
29. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* **2015**, *8*, 1991–2007. [[CrossRef](#)]
30. Wolter, P.T.; Townsend, P.A. Multi-sensor data fusion for estimating forest species composition and abundance in northern Minnesota. *Remote Sens. Environ.* **2011**, *115*, 671–691. [[CrossRef](#)]
31. Lucas, R.; Rowlands, A.; Brown, A.; Keyworth, S.; Bunting, P. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 165–185. [[CrossRef](#)]
32. Latifovic, R.; Pouliot, D. Multitemporal land cover mapping for Canada: Methodology and products. *Can. J. Remote Sens.* **2005**, *31*, 347–363. [[CrossRef](#)]
33. Townsend, P.; Walsh, S. Remote sensing of forested wetlands: Application of multitemporal and multispectral satellite imagery to determine plant community composition and structure in southeastern USA. *Plant Ecol.* **2001**, *157*, 129–149. [[CrossRef](#)]
34. Yuan, F.; Sawaya, K.E.; Loeffelholz, B.C.; Bauer, M.E. Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sens. Environ.* **2005**, *98*, 317–328. [[CrossRef](#)]
35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
36. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [[CrossRef](#)]
37. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [[CrossRef](#)]
38. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* **2006**, *9*, 181–199. [[CrossRef](#)]
39. Cutler, D.R.; Edwards, T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, J.G.; Lawler, J.T. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [[CrossRef](#)] [[PubMed](#)]

40. 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. [[CrossRef](#)]
41. Sesnie, S.E.; Finegan, B.; Gessler, P.E.; Thessler, S.; Ramos Bendana, Z.; Smith, A.M.S. The multispectral separability of Costa Rican rainforest types with support vector machines and random forest decision trees. *Int. J. Remote Sens.* **2010**, *31*, 2885–2909. [[CrossRef](#)]
42. 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. [[CrossRef](#)]
43. Wear, D.N.; Greis, J.G. *The Southern Forest Futures Project: Technical Report; General Technical Report SRS-GTR-178*; USDA-Forest Service, Southern Research Station: Asheville, NC, USA, 2013.
44. Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72. [[CrossRef](#)]
45. Masek, J.G.; Vermote, E.F.; Saleous, N.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T.K. LEDAPS Calibration, Reflectance, Atmospheric Correction Preprocessing Code, Version 2. Available online: <http://dx.doi.org/10.3334/ORNLDAAC/1146> (accessed on 1 March 2013).
46. Kauth, R.J.; Thomas, G.S. The tasseled cap—A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, West Lafayette, Indiana, 29 June–1 July 1976; pp. 41–51.
47. Crist, E.P.; Cicone, R.C. Application of the tasseled cap concept to simulated thematic mapper data. *Photogramm. Eng. Remote Sens.* **1984**, *50*, 343–352.
48. Homer, C.; Dewitz, J.; Fry, J.; Coan, M.; Hossain, N.; Larson, C.; Herold, N.; McKerrow, A.; VanDriel, J.N.; Wickham, J. Completion of the 2001 national land cover database for the conterminous United States. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 337–341.
49. Fry, J.A.; Xian, G.; Jin, S.; Dewitz, J.A.; Homer, C.G.; Yang, L.; Barnes, C.A.; Herold, N.D.; Wickham, J.D. Completion of the 2006 National land cover database for the conterminous United States. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 858–864.
50. Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K. Completion of the 2011 national land cover database for the conterminous United States—Representing a decade of land cover change information. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 345–354.
51. Wickham, J.D.; Stehman, S.V.; Fry, J.A.; Smith, J.H.; Homer, C.G. Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. *Remote Sens. Environ.* **2010**, *114*, 1286–1296. [[CrossRef](#)]
52. Wickham, J.D.; Stehman, S.V.; Gass, L.; Dewitz, J.; Fry, J.A.; Wade, T.G. Accuracy assessment of NLCD 2006 land cover and impervious surface. *Remote Sens. Environ.* **2013**, *130*, 294–304. [[CrossRef](#)]
53. Fiorella, M.; Ripple, W.J. Determining successional stage of temperate coniferous forests with Landsat satellite data. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 239–246.
54. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
55. Foody, G.M.; Cox, D.P. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *Int. J. Remote Sens.* **1994**, *15*, 619–631. [[CrossRef](#)]
56. Foody, G.M. Estimation of sub-pixel land cover composition in the presence of untrained classes. *Comput. Geosci.* **2000**, *26*, 469–478. [[CrossRef](#)]
57. Atkinson, P.M.; Cutler, M.E.J.; Lewis, H. Mapping sub-pixel proportional land cover with AVHRR imagery. *Int. J. Remote Sens.* **1997**, *18*, 917–935. [[CrossRef](#)]
58. Bastin, L. Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities as tools for unmixing coarse pixels. *Int. J. Remote Sens.* **1997**, *18*, 3629–3648. [[CrossRef](#)]
59. Foody, G.M. Relating the land-cover composition of mixed pixels to artificial neural network classification output. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 491–499.

