Article

Estimation of Tree Cover in an Agricultural Parkland of Senegal Using Rule-Based Regression Tree Modeling

Stefanie M. Herrmann *, Andrew J. Wickhorst and Stuart E. Marsh

Arizona Remote Sensing Center, School of Natural Resources and the Environment,
The University of Arizona, 1955 E. Sixth St., Tucson, AZ 85719, USA;
E-Mails: drewman@email.arizona.edu (A.J.W.); smarsh@email.arizona.edu (S.E.M.)

* Author to whom correspondence should be addressed; E-mail: stefanie@ag.arizona.edu;
Tel.: +1-520-626-0830.

Abstract: Field trees are an integral part of the farmed parkland landscape in West Africa and provide multiple benefits to the local environment and livelihoods. While field trees have received increasing interest in the context of strengthening resilience to climate variability and change, the actual extent of farmed parkland and spatial patterns of tree cover are largely unknown. We used the rule-based predictive modeling tool Cubist® to estimate field tree cover in the west-central agricultural region of Senegal. A collection of rules and associated multiple linear regression models was constructed from (1) a reference dataset of percent tree cover derived from very high spatial resolution data (2 m Orbview) as the dependent variable, and (2) ten years of 10-day 250 m Moderate Resolution Imaging Spectrometer (MODIS) Normalized Difference Vegetation Index (NDVI) composites and derived phenological metrics as independent variables. Correlation coefficients between modeled and reference percent tree cover of 0.88 and 0.77 were achieved for training and validation data respectively, with absolute mean errors of 1.07 and 1.03 percent tree cover. The resulting map shows a west-east gradient from high tree cover in the peri-urban areas of horticulture and arboriculture to low tree cover in the more sparsely populated eastern part of the study area. A comparison of current (2000s) tree cover along this gradient with historic cover as seen on Corona images reveals dynamics of change but also areas of remarkable stability of field tree cover since 1968. The proposed modeling approach can help to identify locations of high and low tree cover in dryland environments and guide ground studies and management interventions aimed at promoting the integration of field trees in agricultural systems.

---

1. Introduction

1.1. Importance of Tree Cover in West African Dryland Farming Systems

Agricultural parklands are a traditional land use system in semi-arid West Africa that is characterized by deliberate retention of scattered trees on cultivated land [1,2]. These field trees constitute an integral part of the ecological and livelihood systems. They contribute to the maintenance of soil fertility and water conservation and provide food, fodder, fuel, medicinal products, and building materials to farmers [3]. Demographic, economic, climatic, and environmental developments of the past decades have put increasing pressure on this land use system and losses of trees have been reported, particularly since the devastating droughts of the late 1960s, e.g., Gonzales [4]. At the same time, there has been a growing awareness of the importance of field trees as integral components of dryland agricultural systems, particularly in the contexts of preventing or reversing land degradation and strengthening resilience to climate variability and change [5,6]. As a result, the knowledge base on agroforestry systems has been greatly advanced over the years, spanning research from tree-soil-crop biophysical interactions to social aspects of tree production and indigenous knowledge. However, most studies have been carried out at local scales and little quantitative data exist on the extent and spatial patterns of field tree cover at regional scales [7,8]. Scientists and development specialists have expressed great interest in assessments of the spatial extent and conditions of this traditional land use system, which is considered an asset in the portfolio of adaptation strategies to the unpredictable risks of climate change. A systematic assessment of current field tree cover is required to identify target areas for promoting regeneration of field trees, allocating funds for interventions, and monitoring long term program success.

1.2. Role of Sub-Pixel Remote Sensing for Assessing Tree Cover

With its synoptic perspective and consistent observing capabilities, remote sensing is a viable tool for systematic mapping of tree cover. Indeed, individual trees and clusters of trees can be readily identified on very high spatial resolution (0.5 m–4 m) images provided by commercial satellite sensors. However, the cost and data volume associated with very high spatial resolution images make their use problematic for covering larger geographic areas. On the other hand, medium, and coarse resolution data, while easily and freely available, do not provide enough spatial detail to resolve trees, leaving us with the challenge of inferring tree cover using sub-pixel analysis methods, which extract information on one or more components of mixed pixels [9].

Several of these methods were used in previous research to estimate tree and other green vegetation cover fractions at regional or global scales. They can be broadly divided into two categories: (1) linear mixture models, which rely on pure endmembers for model calibration, and (2) regression tree algorithms, which make use of the entire continuum of tree cover as training data to produce a set of linear models.

De Fries *et al.* (2000) used linear mixture modeling to generate a 1-km global percent tree cover dataset from Advanced Very High Resolution Radiometer (AVHRR) data [10]. Hansen *et al.* (2003) produced a 500 m global percent tree canopy cover map from Moderate Resolution Imaging Spectrometer (MODIS) data using a regression tree algorithm and training data derived from Landsat [11]. Rokhmatuloh *et al.* (2005) mapped tree cover for Africa at 1-km resolution, also using a regression tree algorithm [12]. A number of finer resolution studies were conducted at regional scales with multispectral Landsat data based on linear mixture [13–15] as well as regression tree [16] modeling approaches. The vast majority of the regional studies emphasize forested ecosystems, and global scale tree cover products also seem to perform better in regions of denser tree cover than in sparsely vegetated drylands. For our area of interest, the semiarid Sahel zone of West Africa with its rather sparse tree cover, neither the 500 m MODIS vegetation continuous fields (VCF) product [10] nor the most current 250 m VCF product (Collection 5) appear to accurately portray the spatial pattern of tree cover. Its limitations for describing patterns of tree cover in savannas are also acknowledged by Favier *et al.* (2012) [17].

1.3. Objectives

Our primary objective is to test a locally calibrated, rule-based modeling approach and produce a regionally valid map of current (2000s) tree cover for the west central agricultural region of Senegal, which is accurate enough to be useful for land use/management planning purposes and to function as a baseline for monitoring changes in tree cover. A second objective is to show the usefulness and potential application of this product for assessing long-term changes in tree cover at the example of a longitudinal transect.

2. Data and Methods

2.1. Study Area

Our study area is the west central agricultural region of Senegal [18], which extends over *ca.* 26,000 km^2 (Figure 1). It has been a center of intense cash crop production of peanuts since French colonial times and is also referred to as the "Old Peanut Basin". Originally a mosaic of wooded savanna and open woodlands, this ecoregion has been transformed into a cultural landscape by centuries of human activity. Today, a quasi-continuous agricultural parkland dominated by the leguminous tree *Faidherbia albida* (also *Acacia albida*, or "kad" in Wolof) has replaced almost all remnants of natural vegetation (Figure 2). *Faidherbia albida* is characterized by a reverse phenology, shedding its leaves at the peak of the rainy season, while being in full leaf throughout the dry season. Other indigenous woody species that are present in the region include *Acacia macrostachya, Acacia Senegal, Adansonia digitata, Sterculia setigera* (all deciduous), *Anogeissus leiocarpus, Combretum micranthum, Combretum nigricans, Guiera senegalensis, Tamarindus indica* (all semi-evergreen), *Balanites aegyptiaca, Boscia senegalensis, Combretum glutinosum, Piliostigma reticulatum* (all evergreen). Two non-indigenous evergreen trees that are commonly found in village locations as shade trees are the mango tree (*Mangifera indica*) and the neem tree (*Azadirachta indica*).

Despite fluctuations in the extent of cultivated area resulting from economic and policy factors, rain-fed agriculture remains the most important land use and a mainstay of people's livelihoods. In the past two

decades, cultivation of the food staple millet has increased at the expense of the cash crop peanuts [2]. This shift in cropping patterns can be partly explained by a cutback of fertilizer subsidies, which has also renewed awareness of the importance of field trees for improving and conserving soil fertility.

Figure 1. Location map of the west central agricultural region in Senegal.
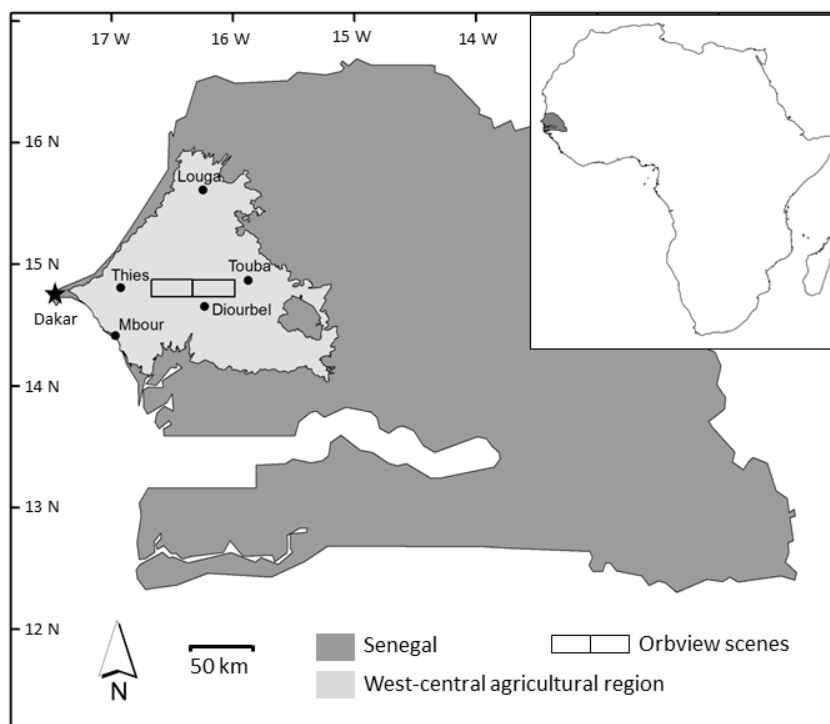
Figure 2. Agricultural parkland landscape with a field tree canopy cover of 5–8% in the administrative region of Thies, January 2011. The *Faidherbia albida* trees are in full leaf, whereas the ground is bare of photosynthetically active herbaceous or crop cover in this season, which facilitates extraction of tree cover from remote sensing data (Photo: Gray Tappan, USGS).


The climate in the study area is sahelo-sudanian with mean annual rainfall ranging from 300 mm in the north to 700 mm in the south, concentrated in the summer months (July to October) and highly

variable from year to year. Hence, the region is considered marginal for rainfed agriculture, and production has been hit by several severe droughts in the past [19]. Under future climate change scenarios, higher temperatures and potentially lower and more variable rainfall are expected to place additional stress on the region's land and water resources [20], adding to loss of soil fertility and propensity to erosion from decades of peanut cultivation [21]. In the interest of improving resilience and productivity of agricultural production systems, maintaining or increasing field tree cover has been promoted as one pillar of "climate smart agriculture" in this type of ecosystem [22].

2.2. Data

We used a multi-resolution approach, which combines very high spatial resolution data with high temporal resolution data, in order to compensate for the tradeoffs inherent in each individual remote sensing dataset [23]. The datasets and processing steps are illustrated in Figure 3 and will be described in the following sections.

Figure 3. Illustrated diagram of the data processing stream leading to the estimated percent tree cover map of the study region. Cubist® was used to derive a set of decision rules from training cases, each of which is linked to a multiple linear regression model.

Processing stream for % tree cover mapping
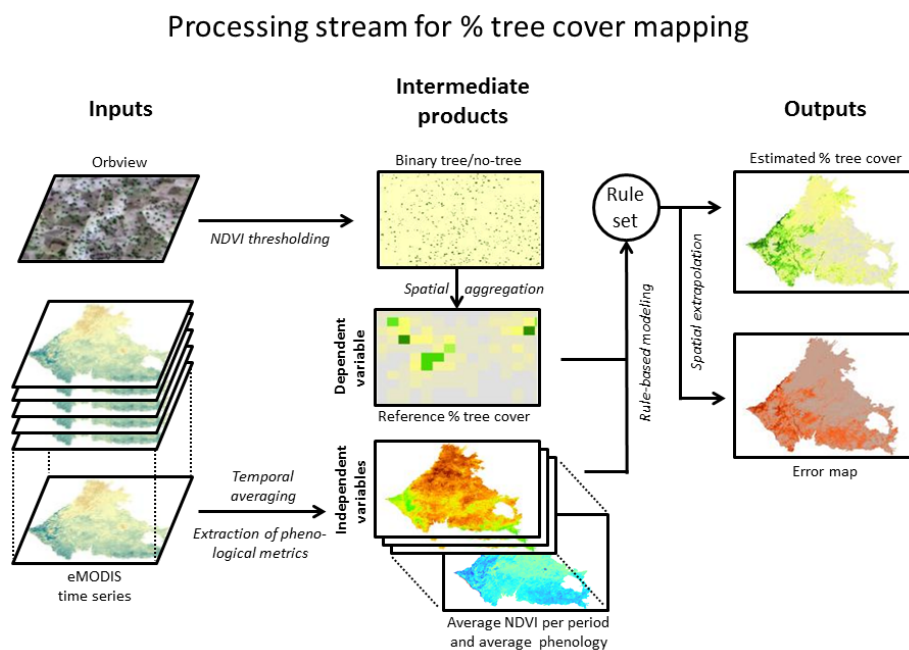
2.2.1. High Spatial Resolution Data: Orbview

Two adjacent 35 × 15 km Orbview scenes located in the center of the study area (Figure 1) were used to identify trees and create a reference dataset of percent tree cover (*i.e.*, canopy cover area per pixel area) for training and validation purposes. Both scenes date from 22 March 2011, a time of year when herbaceous cover is dormant and cultivated fields are harvested, whereas most trees are in full leaf. At a spatial resolution of 2 m, this multispectral (visible and near infrared wavelengths) image product from the commercial GeoEye satellite is well suited for distinguishing individual and clusters of trees.

In addition to the two Orbview scenes that were used in the tree cover modeling, high resolution imagery accessed in Google Earth provided supplemental supporting information for visual validation of the modeling results beyond the area covered by the Orbview scenes. This imagery was acquired between 2000 and 2011 by the GeoEye, Digital Global, and SPOT sensors.

2.2.2. High Temporal Resolution Data: MODIS

The Moderate Resolution Imaging Spectrometer (MODIS) is characterized by a high temporal frequency of acquisitions—the MODIS sensors on the Terra and Aqua satellite platforms provide daily morning and afternoon overpasses respectively—as well as significantly improved spatial resolutions, geolocational accuracy, and atmospheric corrections compared to the legacy Advanced Very High Resolution Radiometer (AVHRR) sensor [24]. Since its launch in 2000, MODIS data products have been used by the land remote sensing community for a number of research applications including vegetation monitoring, crop yield estimation, and burn scar identification [25]. One of the most widely used data products from this sensor is the Normalized Difference Vegetation Index (NDVI), a normalized ratio of reflectances in the red and near infrared portions of the electromagnetic spectrum ((NIR − red)/(NIR + red)), which is sensitive to chlorophyll content and can be used as an indicator of the amount of actively photosynthesizing, green vegetation [26,27].

In this study, we used the eMODIS (expedited MODIS) NDVI, a preprocessed dataset for Africa provided by the US Geological Survey Earth Resources Observation Systems (EROS) Data Center in response to needs expressed by the vegetation monitoring community for a more user-friendly regional scale satellite data set. In historical processing mode, the eMODIS NDVI is a temporally smoothed [28] time series composited in 10-day intervals every five days at 250 m spatial resolution for the period of 2001 to 2010 [29], resulting in 72 compositing periods per year. Its shorter compositing intervals compared to the standard MODIS NDVI (16 days) make this product particularly useful for capturing rapid changes in vegetation phenology and subtle differences in the phenologies of woody and herbaceous cover.

2.2.3. Historical Space Photography: Corona

Declassified satellite photography from the Corona intelligence project [30] dating from 31 January 1968, was used for visually estimating tree cover in the late 1960s, a time period which is often taken as a reference point of the landscape condition before the great Sahelian droughts. The panchromatic space photos, taken on elongated film strips covering a ground area 14 × 188 km, have a spatial resolution of up to 1.8 m, but show geometrical distortions due to the movement of the satellite as the lens scans [31]. No calibration or ancillary data are available. While the Corona data do not lend themselves to automated detection or extraction of tree cover because of their limited spectral resolution, they are a valuable source of information for visual interpretation [10] and, as such, have been used in local-scale land use and land cover change studies to extend the baseline back to the 1960s [32–34]. Fields in the study area are mostly harvested by late January, which affords high contrast between darker trees and the brighter soil background. The distinction between trees and shrubs, however, can be challenging in the panchromatic realm and has to rely on observed differences in texture and brightness.

2.3. Data Preparation

2.3.1. Orbview-Derived Reference Map

In order to produce a reliable reference map of percent tree cover—that is, the percentage of tree canopy cover in a MODIS-sized pixel—we first had to accurately distinguish tree from non-tree pixels in the Orbview scenes. To this end, we computed the NDVI from the red and near infrared bands of the Orbview data (Figure 4a). As evergreen and semi-evergreen trees and shrubs were the only green vegetation present in the scenes in the late dry season, they clearly stood out against their surroundings by a higher NDVI. However, a minor number of deciduous trees present in the scene were leafless at the time of acquisition of the Orbview image and, hence, could not be detected by the NDVI. A simple decision tree classifier was used to find the most appropriate threshold to differentiate tree crowns from the smaller shrubs, the shadows and the soil background. An NDVI threshold of 0.5 was found to work best across the images and was applied to produce binary maps of tree *versus* non-tree cover for both scenes (Figure 4b), resulting in a classification accuracy of 95%, based on accuracy assessment of 100 random points in the tree class and another 100 random points in the non-tree class. Only very few locations—burnt and irrigated areas misclassified as trees—were subsequently manually corrected in order to ensure an optimal reference dataset. The binary maps were then reprojected to match the eMODIS geographic projection and aggregated to the resolution of 250 m using bilinear resampling in order to yield maps of percent tree cover per MODIS pixel (Figure 4c). The two adjacent percent tree cover maps were mosaicked into one reference map for subsequent tree cover modeling.

Figure 4. 1 km^2 subset of (**a**) false-color composite of Orbview image, (**b**) extracted tree canopy cover superimposed on Orbview image, and (**c**) percent tree cover reference data at 250 m resolution, aggregated from the extracted tree canopy cover (superimposed). Lowest percent tree cover is shown in light yellow, highest in dark green tones, see full color legend in figure in Section 3.2.

      (**a**)   (**b**)   (**c**)
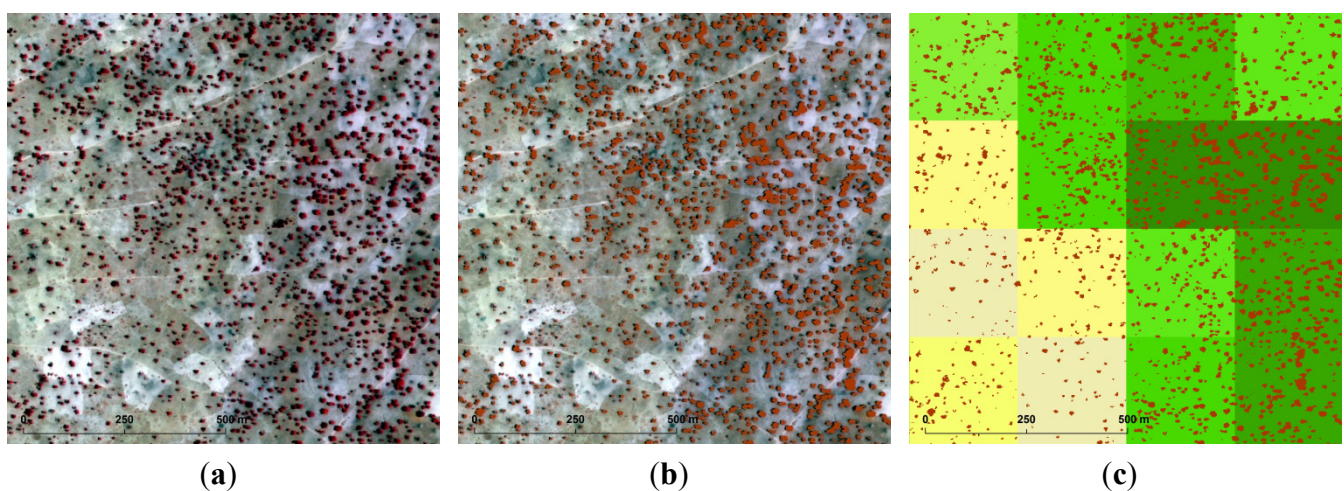
2.3.2. MODIS-Derived Input Variables

The estimation of percent tree cover from eMODIS NDVI data depends on slight differences in overall greenness and seasonality between trees and non-tree land/soil cover types. The considerable

interannual variability in timing and amounts of rainfall, however, leads to an equally variable vegetation response from year to year, which can exceed the often subtle differences between land/soil cover types. In order to correct for the effects of rainfall variability, we averaged the NDVI of each compositing period over multiple years, with ten-year averages yielding the best results in the subsequent modeling effort. In this point, our approach differs from prior studies [10,11,12], which all rely on only one year of input data. We believe that the shortcoming of limiting any temporal change studies to 10-year periods is outweighed by the advantage of producing more accurate results for our area of interest.

In addition to characterizing vegetation by the average NDVI per compositing period, we used the Timesat 3.0 software [35] to extract a number of phenological metrics from the eMODIS NDVI time series. The potential of using slight differences in phenological, or seasonality, characteristics of trees and grasses for modeling the distribution of those two growth forms in savanna environments is highlighted by Wagenseil and Samimi [36]. The metrics describe key parameters of the phenological curve and have been widely used in scientific applications, including mapping environmental changes in Africa [37–40]. Timesat fits local polynomial functions to time series of NDVI data, from which the metrics are extracted based on a definition for the start and end of the growing season, in our case a proportion (0.2) of the seasonal amplitude. The complete set of phenological metrics obtained from Timesat includes four seasonality metrics (dates of the start, end, and middle of the growing season, as well as length of the growing season), five magnitude metrics (base level NDVI, maximum NDVI, NDVI amplitude, small seasonal NDVI integral, large seasonal NDVI integral) and two change metrics (rate of increase of NDVI at the beginning of the growing season, rate of decrease of NDVI at the end of the growing season). We extracted all 11 metrics for each year of the time series and then computed the 10-year averages of each of them.

In order to get a preliminary estimate of the most promising input data, we performed simple linear correlation analyses between the target percent tree cover from the reference map and each of the potential independent variables (average NDVI of 72 compositing periods and 11 phenological metrics). Although some variables showed higher correlations than others, with the peak of the dry season (January–February) compositing periods and the phenological metrics most strongly correlated with percent tree cover, we included the entire set of all compositing periods and phenological metrics in the final rule-based model.

2.3.3. Georectification of Corona Image

We manually oriented and georectified one Corona image in four segments, using 15–20 ground control points (GPCs) per segment (approximately 14 × 45 km) that could be identified in both the Corona image and Google Earth, such as road intersections. A second order polynomial warping procedure was then applied to each of the Corona image segments in order to rectify them to the geographic coordinate system using these GPCs [41]. As the topography of our study area is relatively flat, this method produced satisfactory results with locational errors ranging from 5 m to 30 m and a near seamless fit between the four image segments.

2.4. Rule-Based Modeling and Spatial Application

A rule-based predictive model was developed, which estimated tree cover in the west central agricultural region of Senegal using a classification and regression tree (CART) approach [42]. The CART approach is a data mining method to inductively identify patterns between a dependent variable (the tree cover reference dataset) and a collection of independent predictor variables (the MODIS-derived coarse resolution input variables). A regression tree contains a set of conditional rules that partition the data space into smaller regions, each of which is linked to a multiple linear regression model predicting the dependent from the independent variables [43].

In a first step, the mosaicked reference map was divided into training and validation data by a stratified random sampling procedure, which resulted in 8,229 training pixels and 7,420 validation pixels spread evenly across the extent of the map. We then used the Cubist algorithm (Rulequest Research) for generating a rule-based regression tree from the training data and the corresponding independent predictor variables. To optimize the accuracy of the model prediction, committee models with 5–10 members—each representing one regression tree—were constructed, where each subsequent member adjusts for inaccuracies in the predictions of the previous member in a boosting-like scheme. The final modeling result is a simple average of the predictions of each committee member and yields higher overall accuracies than a model without the committee option [43].

The committee model resulting in the most accurate results for the training and validation data was then applied to the pixels of the entire study area, producing a map of estimated percent tree cover. Due to the nature of the input data, the resulting tree cover map represents the average tree cover for the period 2001 to 2010, which will be referred to as the current time period.

2.5. Model Validation

The model was evaluated by its performance on both the training data, which were used in generating the model, and the validation data, which the model had not seen before [44]. Three measures express different aspects of the model's performance: (1) the linear correlation coefficient (Pearson product moment correlation coefficient R) between reference/target and model-estimated tree cover, a dimensionless measure of the strength of linear association between two variables; (2) the mean absolute error (MAE) between reference/target and model-estimated tree cover, which is unambiguous and therefore more appropriate as a dimensioned measure of average error than the root mean square error [45]; and (3) the relative error, here defined as the ratio of the mean absolute error to the error that would result from predicting the mean value of percent tree cover at each pixel. The latter expresses how useful the model is to meaningfully describe the modeled variable.

A spatially explicit error layer (2nd figure in Section 3.2) was produced in conjunction with the map of model-estimated tree cover to show the spatial distribution of potential error beyond the area of the reference map, based on the estimated error of the rules used to predict each pixel value [42]. In addition to these quantitative measures of model accuracy, geographic patterns of model-estimated tree cover were visually compared to tree cover on recent (circa 2010) high resolution Google Earth imagery across the entire study area.

2.6. Change Detection Application

We used a 200-km longitudinal transect, which traverses the study area from west to east at 14.75°N (1st figure in Section 3.3), to compare tree cover at two different time periods and assess changes in spatial patterns. Model-estimated tree cover for the current time period was averaged over continuous 1 × 1 km sample areas along the transect. For the same sample areas, tree cover in the late 1960s was visually estimated from the Corona imagery using a comparison chart for visual estimation of percent cover similar to charts used in ecological field research [46]. The sample size of 1 km^2, rather than the 250 m resolution of the tree cover map, was chosen to better match the variability of tree cover across the transect and minimize the impact of geo-referencing inaccuracies.

3. Results and Discussion

3.1. Model Accuracy and Variables Used

A 5-member committee model, which produced between 16 and 35 rules per committee member, was selected as the best model based on its performance on the training and validation data. For the training data, a correlation coefficient r of 0.88 between reference and model-estimated tree cover was achieved, with a mean absolute error of 1.07% tree cover and a relative error of 0.45. For the validation data, the correlation coefficient r was 0.77 and the mean absolute and relative errors amounted to 1.03% tree cover and 0.53 respectively. The correlation coefficients are highly significant, and the low absolute and relative errors indicate that the model performs well for both the training and unseen validation data. Figure 5 plots the observed or reference tree cover against model estimated tree cover and shows a good correspondence between the two datasets, with a slight tendency for underestimation of the model in the higher tree cover areas. The vast majority of pixels, however, are clustered in the lower tree cover areas (96% of the pixels in the training and validation data have a tree cover of ten percent and lower and 84% have a tree cover of five percent and lower). Table 1 details the distribution of error by tree cover categories in 5% increments.

Figure 5. Graphs of observed and modeled percent tree cover for the training and validation datasets (MAE = Mean absolute error, RE = Relative error).
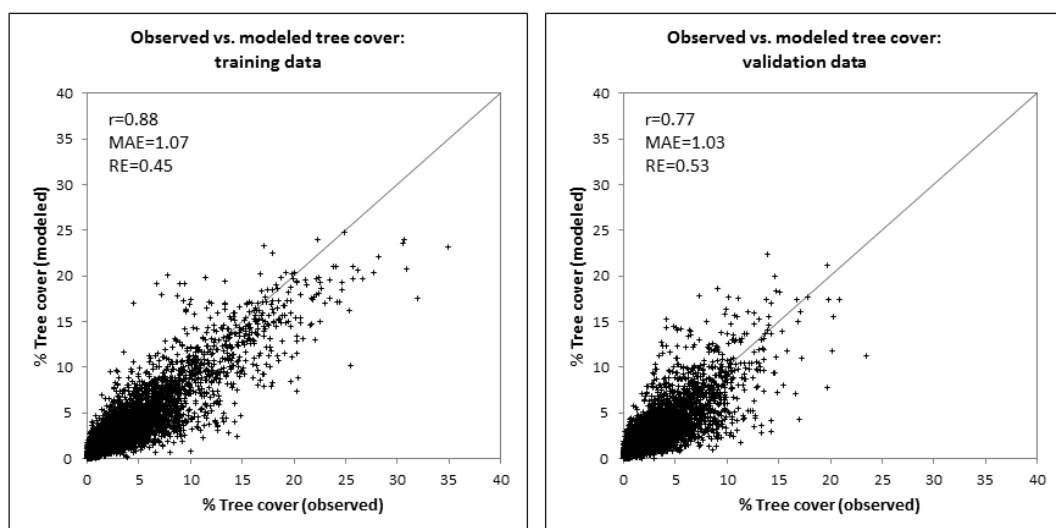
Table 1. Distribution of mean absolute errors (MAE) of estimated tree cover by categories (5% increments of observed/reference tree cover) for training and validation data (N=number of pixels per category).

Tree Cover Category	MAE in Training Data	MAE in Validation Data
5% tree cover and less	0.74% (N = 6722)	0.81% (N = 6469)
5%–10% tree cover	2.11% (N = 1049)	2.32% (N = 814)
10%–15% tree cover	3.15% (N = 276)	3.70% (N = 119)
15%–20% tree cover	3.21% (N = 129)	4.56% (N = 14)
20% tree cover and more	6.68% (N = 54)	7.20% (N = 4)

The independent variables contributed to the conditional rules and multiple linear regressions of the model to different extents. All but two variables were used by the model, with the growing season and early dry season compositing periods being used in more of the regressions than the late dry season compositing periods. As expected, the phenological metrics contributed quite importantly to the model. The start, end, and length of season metrics as well as the maximum NDVI, amplitude and small integral were used in over 90% of the model equations. However, building the model on the phenological metrics alone would have reduced model accuracy.

3.2. Spatial Patterns of Tree Cover and Their Interpretation

High model accuracy for the training and validation data does not guarantee equally good performance when the model is extrapolated to a larger area. However, with a mean of 4.7% tree cover and a majority of pixels falling into the 3%–4% tree cover range, the modeling results for the west central agricultural region appear realistic for this environment (Figure 6). Highest tree densities are found in the western part of the study area, in the peri-urban areas between Dakar and Thies and in the Niayes zone along the Atlantic coast. The proximity to urban markets and the vulnerability of annual crops to drought has given rise to the development of horticulture and arboriculture in these areas [47]. Fruit tree orchards contribute to the high tree cover found there. However, the model overestimated tree cover in some of the depressions of Niayes zone where intensive vegetable gardening predominates, due to phenological similarities with evergreen trees. Other areas of higher than average tree cover include the high density *Faidherbia albida* agricultural parklands in the Thies region, among them the homelands of the ethnic Serer, who have maintained this agroforestry system for centuries for its benefits to millet cultivation and cattle raising [48]. Localities of particularly high field tree cover were found east of Khombole (14.773°N, 16.663°W; 10%–15% tree cover), northwest of Tivouane (14.980°N, 16.861°W; 7%–10% tree cover), around Pekes (15.112°N, 16.416°W; 7%–15% tree cover), northeast of Mbour (14.486°N, 16.938°W; 7%–12% tree cover)—whereas east of Mbour tree cover seems overestimated. Depressions and riparian areas, for example the Sine valley, also stand out by denser tree cover than their surroundings. Across the entire map, an irregular dot pattern of high concentrations of trees indicates the presence of settlements. They are mainly shade trees, fruit trees and live fences as well as trees of cultural and spiritual significance, such as groves of baobab trees, by which long abandoned settlements are still identifiable.

Figure 6. Map of model estimated tree cover in the west central agricultural region of Senegal (**a**) Enlarged area shows reference tree cover, (**b**) modeled tree cover (**c**) and MODIS Vegetation Continuous Fields (VCF) product (Collection 5) of 2008 for the same area (**d**). The 2008 VCF product was chosen for display, because the VCF quality layers indicate that it is the best of the available 11 years of VCF data for this area. Reference and modeled tree cover range from 0 to 22%, whereas the VCF tree cover ranges from 0 to 3%.

(**a**)

(**b**)

(**c**)

(**d**)
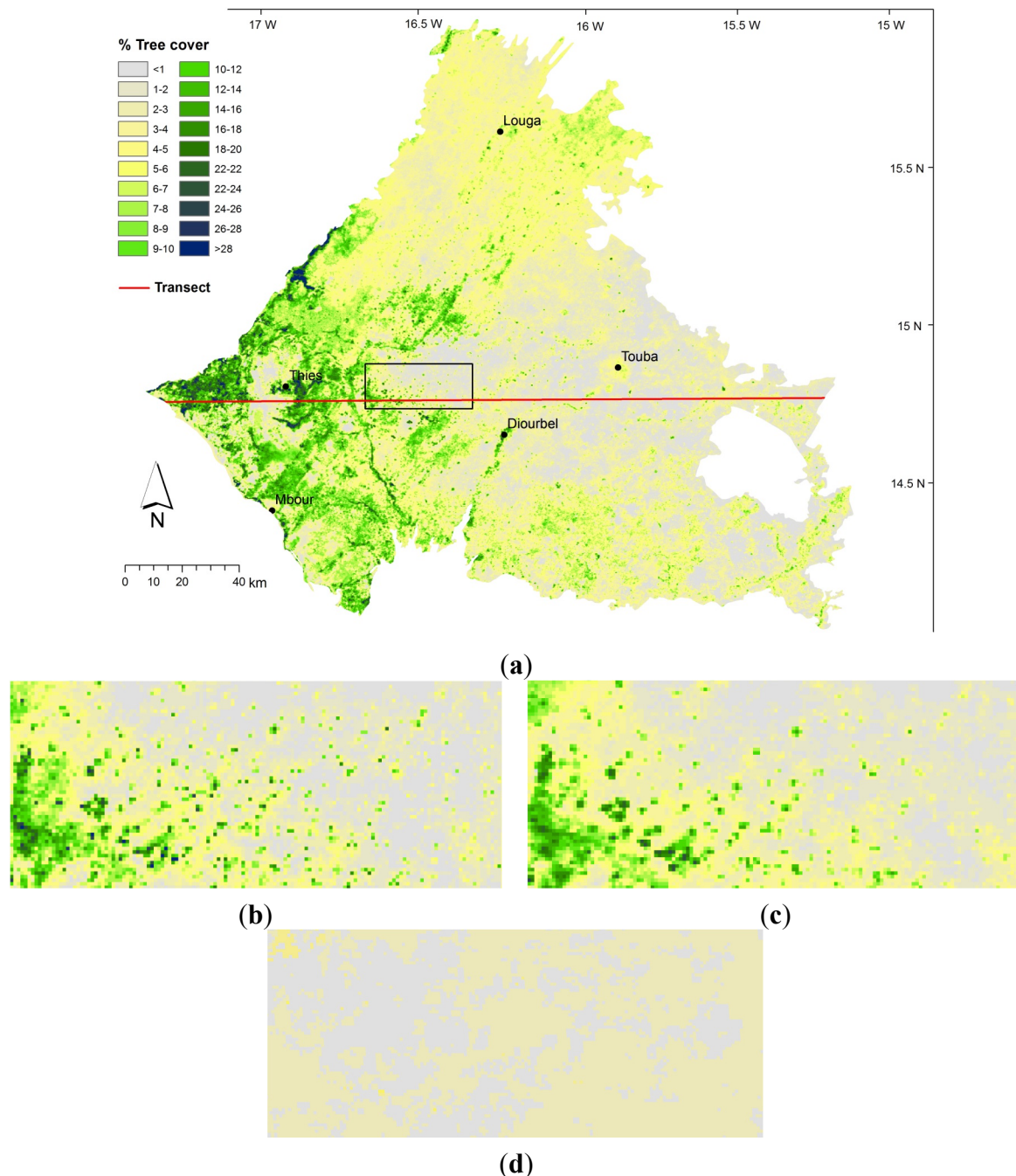
Overall, the map seems to accurately depict the spatial pattern of tree cover across the study area. It performs better for our area of interest than the MODIS Vegetation Continuous Fields (collection 5) product [49], which is aimed at showing patterns of tree cover at global to continental rather than regional to local scales. Some instances of overestimation of the tree cover in our map can be attributed to the challenge of distinguishing between trees and shrubs. Although we made an effort to

define a threshold for the reference map that includes trees and excludes smaller, less green shrubs, the model was not always successful in distinguishing between the two woody vegetation types. Another source of error is the presence of irrigation in horticulture, which results in a crop phenology resembling that of trees due to the prolonged growing season. Whereas errors might be highest in absolute terms in the high tree cover areas (Figure 7 and Table 1), in relative terms they are more significant where tree cover is lowest. This confirms findings by Hansen [17] that low tree cover is difficult to accurately estimate, as the soil background dominates over the tree signal.

Figure 7. Error map showing the spatial distribution of potential error beyond the area of the reference map. For each pixel, the number represents the difference between the value that is predicted and the value that could exist, based on the estimated error of the rule created from the training data and used to predict the pixel value.
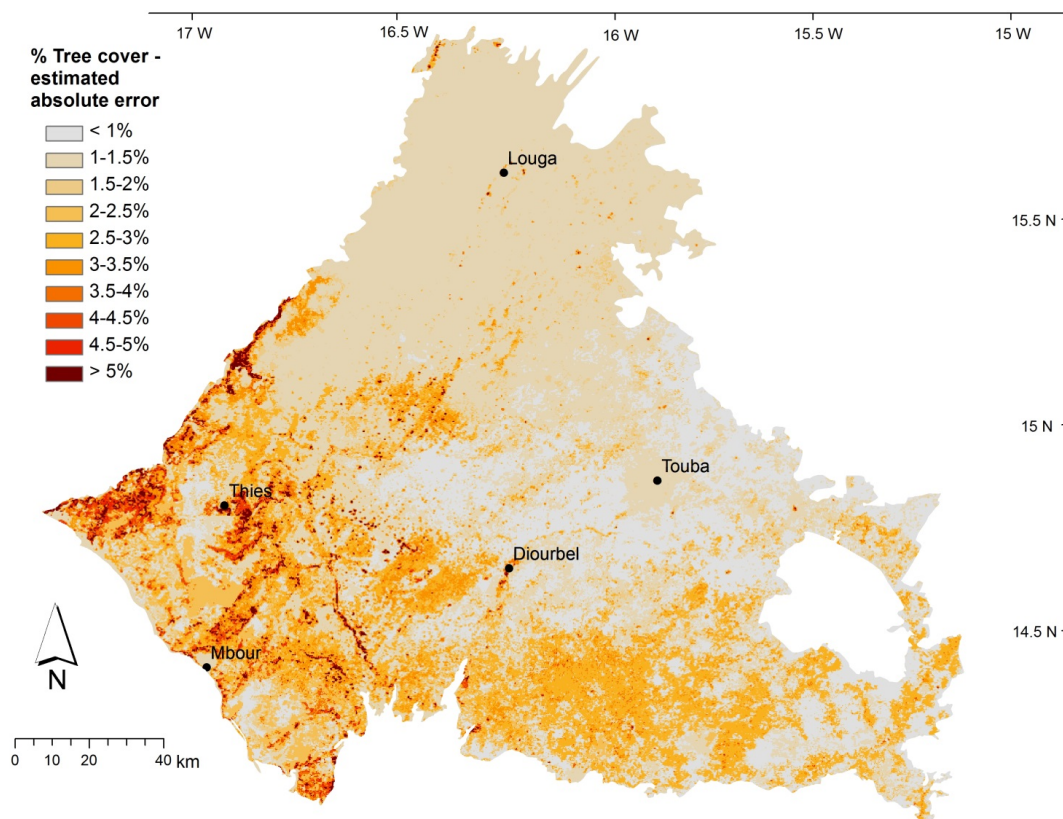
## 3.3. Stability and Change Since 1968

The map of tree cover provided a starting point for identifying areas of interest and examining change in tree cover over time. While we cannot apply the rule-based modeling approach to estimate tree cover at earlier time periods, for which MODIS data are not available, the Corona mission afforded us with a unique source of data from the late 1960s that could be exploited visually, on a sample basis. We compared current tree cover as estimated by the Cubist model, and verified on Google Earth, to tree cover visually interpreted from Corona imagery along a 200 km longitudinal transect (Figures 6, 8, and 9). The seven-sample moving average in Figure 8 reduces the effects of potential geolocational errors between the two data sources and flattens spikes caused by individual outliers, facilitating appreciation of the overall pattern of change, both in time and in space.

Figure 8. West-east transect of tree cover at 14.75°N. The current tree cover was modeled (see paragraph 2.4), and the 1968 tree cover was estimated visually from Corona images. The thicker lines represent 7-sample moving averages (averages of the data from point N − 3 to point N + 3 at each point N) of the original curves.
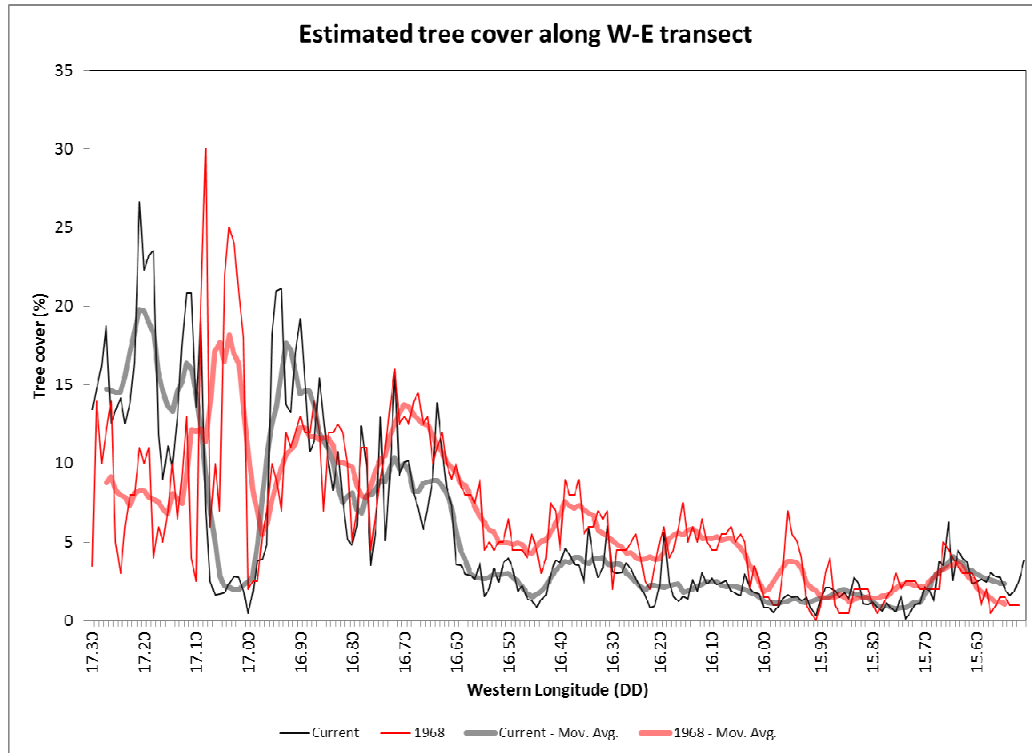
Both curves in Figure 8 show a gradient of decreasing tree cover from west to east, though the gradient appears steeper in current times than in the late 1960s. In the westernmost part of the study area, current tree cover is considerably higher than in 1968 due to the expansion of arboriculture. From 17.10°W to 17.00°W, current tree cover declines sharply. This corresponds to the area of the Protected Forest (forêt classée) of Thies, where tree cover has been decimated to almost zero as a result of the high demand for fuel wood by the growing urban populations [50]. What remains are smaller shrubs and patches of bare ground, as well as a phosphate surface mine at 17.00°W. In 1968, by contrast, parts of the forest still had relatively dense tree cover, making the 1968 curve peak in this area. Fruit tree orchards in the southeastern outskirts of the city of Thies cause another peak in current tree cover around 16.90°W, which exceeds field tree cover found there in 1968. From 16.80°W eastward, *Faidherbia albida*, which make up the traditional parkland tree cover, dominate. Field tree densities appear to have declined overall since 1968, but there are a few locations where they have been stable or even increasing, particularly in the vicinity of villages. The estimated average tree cover along the entire transect was 6.53% in 1968 and has dropped slightly to a current average of 5.75%—a decrease which could be within the margin of error of the estimation. The changes in the pattern of tree cover from a gentle to a steeper west-east gradient are more pronounced. Figure 9 illustrates increasing, stable and decreasing tree cover along the transect by contrasting high-resolution images of 1 × 1 km sites in 1968 and *circa* 2010.

Figure 9. Examples of increasing, stable and decreasing tree cover, contrasting current very high resolution satellite imagery (top row) and Corona images from 1968 (bottom row) for selected 1 km² samples: (**a**) 14.76°N, 17.05°W, south of Pout in the region of Thies: arboriculture in formerly sparsely covered land, (**b**) 14.76°N, 16.66°W, east of Khombole in the region of Diourbel: relatively stable high tree density agricultural parkland (some of the smaller trees/shrubs have disappeared), (**c**) 14.78°N 16.1°W, west of Touba in the region of Diourbel: firewood needs of the city of Touba have led to a decline in tree cover.
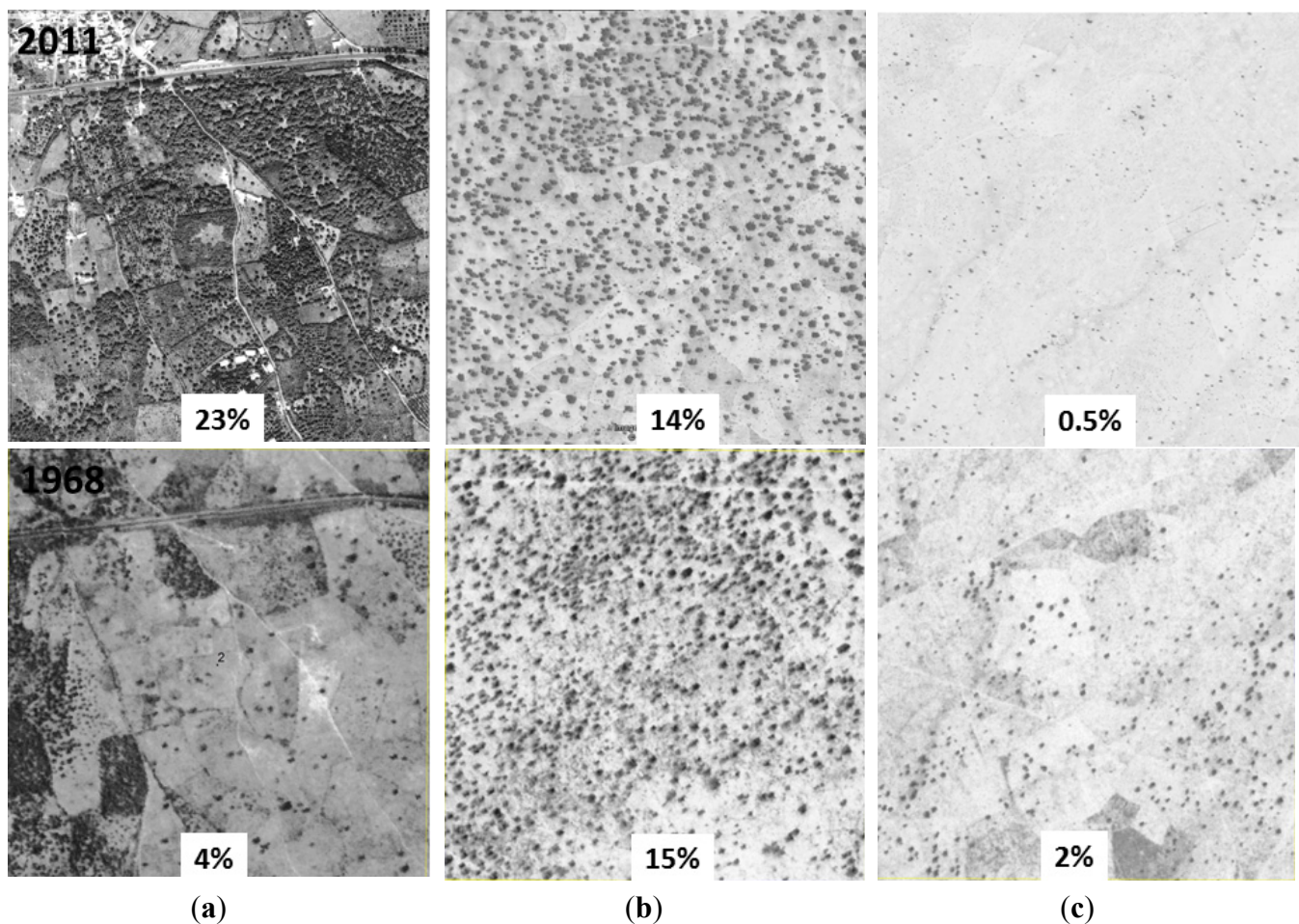
3.4. Applicability to Other Geographic Areas

Our results demonstrate that rule-based regression tree modeling has potential to estimate tree cover in dryland environments with acceptable accuracy, if high spatial resolution data can be obtained for a subset of the area of interest. The quality of results achievable with this modeling approach depends on the complexity of the landscape, the types of trees and their phenology, and the timing of acquisition of the very high resolution imagery. For the west-central agricultural region of Senegal, these three factors came together favorably to facilitate the production of a high quality tree cover map: (1) the landscape is rather uniform and has little relief; (2) the dominant tree, *Faidherbia albida*, is in full leaf in the dry season when other green vegetation is largely absent; and (3) Orbview images were available during the peak of the dry season. Preliminary analyses applying the same modeling approach in a similar environment in Northern Ghana show promising results, albeit of slightly lower accuracy than for the Senegal case owing to difficulties of deriving an equally accurate reference map in

this more complex environment with an undulating relief. However, the spatial pattern of tree cover derived from the locally-calibrated rule-based model is more realistic than that of the global VCF dataset for the same region, corroborating the importance of locally calibrated tree cover maps. It remains to be examined whether including ancillary data, such as elevation, slope, and soil type, would help improve the accuracy of the reference map and of the model-estimated tree cover in more complex landscapes, where diversity of relief, soils, and vegetation make extracting tree cover more challenging.

4. Conclusions

We produced a map of tree cover for the west-central agricultural region of Senegal by training a rule-based model with an Orbview-derived reference dataset to estimate percent tree cover from coarser resolution eMODIS data. The estimated accuracy of this locally calibrated map—with absolute mean errors of just above 1% tree cover—exceeds that of any regional or global percent tree cover products known to us. That is, it reliably identifies pockets of high tree cover within the agricultural parkland as well as areas where field trees are virtually absent. Potential errors of estimation were highest in the high tree cover areas in absolute terms and in the low tree cover areas in relative terms. A comparison of current tree cover with historic cover along a transect shows a slight overall decrease in tree cover from an average of 6.53% to 5.75% over the past four decades; however, this decrease is not uniform in space and there are areas where tree cover has remained stable or even increased, in some locations dramatically. The most striking increases are mostly explained by the expansion of fruit tree plantations in the vicinity of urban centers. Areas where *Faidherbia albida* field tree cover was maintained or increased might indicate success stories of good land management practices by rural communities. Understanding the motivations and benefits of maintaining or promoting field trees in these communities could help upscale such success stories to other communities, which face similar environmental and economic conditions but have maintained fewer field trees. Due to its higher accuracy than readily available global tree cover products, this tree cover map can be used to guide ground studies to investigate the driving forces behind the spatial and temporal variability of field tree cover in this agricultural parkland and help identify areas that might particularly benefit from outreach activities and resource management interventions. It also lends itself as a baseline for longer term monitoring of tree cover, e.g., at decadal intervals, including long term evaluation of reforestation projects. Changes at shorter time intervals, however, cannot be resolved. Considering that trees are a slow variable, as opposed to the herbaceous cover, the constraints resulting from this are limited.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Boffa, J.-M. *Agroforestry Parklands in Sub Saharan Africa*; FAO: Rome, Italy, 1999.

2. Boffa, J.-M. West African agroforestry parklands: keys to conservation and sustainable management. *Unasylva* **2000**, *51*, 11–17.

3. Pullan, A. Farmed parkland in West Africa. *Savanna.* **1974**, *3*, 119–152.

4. Gonzalez, P. Desertification and a shift of forest species in the West Africa Sahel. *Clim. Res.* **2001**, *17*, 217–228.

5. Haglund, E.; Ndjeunga, J.; Snook, L.; Pasternak, D. Dry land tree management for improved household livelihoods: Farmer managed natural regeneration in Niger. *J. Environ. Manage.* **2011**, *92*, 1696–1705.

6. Garrity, D.P.; Akinnifesi, F.K.; Ajayi, O.C.; Weldesemayat, S.G.; Mowo, J.G.; Kalinganire, A.; Larwanou, M.; Bayala, J. Evergreen Agriculture: a robust approach to sustainable food security in Africa. *Food Secur.* **2010**, *2*, 197–214.

7. Reij, C.; Tappan, G.; Smale, M. *Agroenvironmental. Transformation in the Sahel: Another Kind of "Green Revolution"*; IFPRI Discussion Paper 914; International Food Policy Research Institute: Washington, DC, USA, 2009.

8. Zomer, R.J.; Trabucco, A.; Coe, R.; Place, F. *Trees on Farm: Analysis of Global Extent and Geographical Patterns of Agroforestry*; ICRAF Working Paper No. 89; World Agroforestry Centre: Nairobi, Kenya, 2009.

9. Chuvieco, E.; Huete, A. *Fundamentals of Satellite Remote Sensing*; FLCRC Press: Boca Raton, FL, USA, 2010.

10. DeFries, R.S.; Hansen, M.C.; Townshend, J.R.G.; Janetos, A.C.; Loveland, T.R. A new global 1-km dataset of percentage tree cover derived from remote sensing. *Glob. Change Biol.* **2000**, *6*, 247–254.

11. Hansen, M.C.; DeFries, R.S.; Townshend, J.R.G.; Carroll, M.; Dimiceli, C.; Sohlberg, R.A. Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm. *Earth Interact.* **2003**, *7*, 1–15.

12. Rokhmatuloh, D.N.; Al Bilbisi, H.; Tateishi, R. Percent tree cover estimation using regression tree method: A case study of Africa with very high resolution QuickBird images as training data. *IEEE Trans. Geosci. Remote* **2005**, *3*, 2157–2160.

13. Souza, C.; Barreto, P. An alternative approach for detecting and monitoring selectively logged forests in the Amazon. *Int. J. Remote Sens.* **2000**, *21*, 173–179.

14. Lu, D.S.; Moran, E.; Batistella, M. Linear mixture model applied to Amazonian vegetation classification. *Remote Sens. Environ.* **2003**, *84*, 456–469.

15. Wang, C.; Qi, J.; Cochrane, M. Assessment of tropical forest degradation with canopy fractional cover from Landsat ETM+ and IKONOS imagery. *Earth Interact.* **2005**, *9*, 1–18.

16. Huang, C.; Yang, L.; Wylie, B.; Homer, C. A strategy for Estimating Tree Canopy Density Using Landsat 7 ETM+ and High Resolution Images over Large Areas. In Proceedings of Third International Conference on Geospatial Information in Agriculture and Forestry, Denver, CO, USA, 5–7 November 2001.

17. Favier, C.; Aleman, J.; Bremond, L.; Dubois, M.A.; Freycon, V.; Yangakola, J.-M. Abrupt shifts in African savanna tree cover along climatic gradient. *Glob. Ecol. Biogeogr.* **2012**, *21*, 787–797.

18. Tappan, G.G.; Sall, M.; Wood, E.C.; Cushing, M. Ecoregions and land cover trends in Senegal. *J. Arid Environ.* **2004**, *59*, 427–462.

19. Mbow, C.; Mertz, O.; Diouf, A.; Rasmussen, K.; Reenberg, A. The history of environmental change and adaptation in eastern Saloum, Senegal. Driving forces and perceptions. *Glob. Planet. Change* **2008**, *64*, 210–221.

20. IPCC. *Climate Change 2007—Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the IPCC*; Cambridge University Press: Cambridge, UK, 2007.

21. Tschakert, P.; Tappan, G. The social context of carbon sequestration: considerations from a multi-scale environmental history of the Old Peanut Basin of Senegal. *J. Arid Environ.* **2004**, *59*, 535–564.

22. FAO. *"Climate-Smart" Agriculture. Policies, Practices and Financing for Food Security, Adaptation and Mitigation*; FAO: Rome, Italy, 2010.

23. Homer, C.G.; Aldridge, C.L.; Meyer, D.K.; Schell, S.J. Multi-scale remote sensing sagebrush characterization with regression trees over Wyoming, USA: Laying a foundation for monitoring. *Int. J. Appl. Earth Obs.* **2012**, *14*, 233–244.

24. Townshend, J.R.G.; Justice, C.O. Towards operational monitoring of terrestrial systems by moderate-resolution remote sensing. *Remote Sens. Environ.* **2002**, *83*, 351–359.

25. Garcia-Moran, T.J.; Mas, J.-F.; Hinkley, E.A. Land cover mapping applications with MODIS; a literature review. *Int. J. Digit. Earth* **2012**, *9*, 63–87.

26. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150.

27. Tucker, C.J.; Nicholson, S.E. Variations in the size of the Sahara Desert from 1980 to 1997. *Ambio* **1999**, *28*, 587–591.

28. Swets, D.L.; Reed, B.C.; Rowland, J.D.; Marko, S.E. A weighted least-squares approach to temporal NDVI smoothing. In Proceedings of ASPRS Annual Conference: From Image to Information, Portland, Oregon, 1999.

29. USGS. *eMODIS Africa Product Guide Version 1.0*; USGS EROS Data Center: Sioux Falls, SD, USA, 2011.

30. McDonald, R.A. Corona: success for space reconnaissance, a look into the Cold War, and a revolution in intelligence. *Photgramm. Eng. Remote Sens.* **1995**, *61*, 689–720.

31. Bitelli, G.; Girelli, V.A. Metrical use of declassified satellite imagery for an area of archaeological interest in Turkey. *J. Cult. Herit.* **2009**, 10S, 35–40.

32. Hamandawana, H.; Eckardt, F.; Chanda, R. Linking archival and remotely sensed data for long-term environmental monitoring. *Int. J. Appl. Earth Obs.* **2005**, *7*, 284–298.

33. Wardell, D.A.; Reenberg, A.; Tottrup, C. Historical footprints in contemporary landuse systems: forest cover changes in savannah woodlands in the Sudano-Sahelian zone. *Glob. Environ. Chang.* **2003**, *13*, 235–254.

34. Tappan, G.G.; Hadj, A.; Wood, E.; Lietzow, R.W. Use of Argon, Corona, and Landsat imagery to assess 30 years of land resource changes in West-Central Senegal. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 727–735.

35. Eklundh, L.; Joensson, P. *Timesat 3.0 Software Manual*; Lund University: Lund, Sweden, 2010.

36. Wagenseil, H.; Samimi, C. Woody Vegetation Cover in Namibian Savannahs: A Modelling Approach Based on Remote Sensing. *Erdkunde* **2007**, *61*, 325–334

37. Eklundh, L.; Olsson, L. Vegetation index trends for the African Sahel 1982–1999. *Geophys. Res. Lett.* **2003**, *30*, doi:10.1029/2002GL016772.

38. Olsson, L.; Eklundh, L.; Ardoe, J. A recent greening of the Sahel—Trends, patterns and potential causes. *J. Arid Environ.* **2005**, *63*, 556–566.

39. Seaquist, J.W.; Olsson, L.; Ardoe, J.; Eklundh, L. Broad-scale increase in NPP quantified for the African Sahel, 1982–1999. *Int. J. Remote Sens.* **2006**, *27*, 5115–5122.

40. Heumann, B.W.; Seaquist, J.W.; Eklundh, L.; Joensson, P. AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982–2005. *Remote Sens. Environ.* **2007**, *108*, 385–392.

41. Schowengerdt, R.A. *Remote Sensing*: *Models and Methods for Image Processing*, 3rd ed.; Academic Press: London, UK, 2007.

42. Breiman, L.; Olshen, J.; Stone, C. *Classification and Regression Trees*; Chapman and Hill: New York, NY, USA, 1984.

43. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013.

44. Atkinson, P.M; Foody G.M. Uncertainty in Remote Sensing and GIS: Fundamentals. In *Uncertainty in Remote Sensing and GIS*; Foody, G.M., Atkinson, P.M., Eds.; Wiley: Chichester, West Sussex, UK, 2002; Chapter 1, pp. 1–18.

45. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82.

46. Terry, R.D.; Chilingar, G.V. Comparison charts for visual estimation of percent composition. Allen Hancock Foundation. Los Angeles, CA. Reprinted from *J. Sediment. Petrol*. **1955**, *23*, 226–234.

47. Touré Fall, S.; Fall, A.S.; Cisse, I. Urban livestock systems in the Niayes zones in Senegal. *Urban. Agric. Mag.* **2000**, *2*, 17–19.

48. Hirai, M. A vegetation-maintaining system as a livelihood strategy among the Seerer, West Central Senegal. *Afr. Study Monogr.* **2005**, *Suppl. 30*, 183–193.

49. Townshend, J.R.G.; Carroll, M.; Dimiceli, C.; Sohlberg, R.; Hansen, M.; DeFries, R. *Vegetation Continuous Fields MOD44B, 2001–2010 Percent Tree Cover, Collection 5*; University of Maryland, College Park, MD, USA, 2011. Available online: http://glcf.umd.edu/data/vcf/ (accessed on 21 January 2013).

50. Ba, Y. *Problématique de la lutte contre la dégradation des ressources naturelles dans la communauté rurale de Fandène. (département de Thies.)*; Université Cheikh Anta Diop: Dakar, Senegal, 2010.