land

MDPI

Article

Prediction of Land Use Change in Long Island Sound Watersheds Using Nighttime Light Data

Ruiting Zhai [1], Chuanrong Zhang [1,2,*], Weidong Li [1,2], Mark A. Boyer [1,2] and Dean Hanink [1]

1 Department of Geography, University of Connecticut, 215 Glenbrook Rd., Storrs, CT 06269, USA; ruiting.zhai@uconn.edu (R.Z.); weidong.li@uconn.edu (W.L.); mark.boyer@uconn.edu (M.A.B.); dean.hanink@uconn.edu (D.H.)

2 Center for Environmental Sciences and Engineering, University of Connecticut, 3107 Horsebarn Hill Rd., U-4210, Storrs, CT 06269, USA

* Correspondence: chuanrong.zhang@uconn.edu; Tel.: +1-860-486-2610

Academic Editor: Andrew Millington

Received: 3 November 2016; Accepted: 2 December 2016; Published: 7 December 2016

Abstract: The Long Island Sound Watersheds (LISW) are experiencing significant land use/cover change (LUCC), which affects the environment and ecosystems in the watersheds through water pollution, carbon emissions, and loss of wildlife. LUCC modeling is an important approach to understanding what has happened in the landscape and what may change in the future. Moreover, prospective modeling can provide sustainable and efficient decision support for land planning and environmental management. This paper modeled the LUCCs between 1996, 2001 and 2006 in the LISW in the New England region, which experienced an increase in developed area and a decrease of forest. The low-density development pattern played an important role in the loss of forest and the expansion of urban areas. The key driving forces were distance to developed areas, distance to roads, and social-economic drivers, such as nighttime light intensity and population density. In addition, this paper compared and evaluated two integrated LUCC models—the logistic regression–Markov chain model and the multi-layer perception–Markov chain (MLP–MC) model. Both models achieved high accuracy in prediction, but the MLP–MC model performed slightly better. Finally, a land use map for 2026 was predicted by using the MLP–MC model, and it indicates the continued loss of forest and increase of developed area.

Keywords: land use/cover change; Long Island Sound Watersheds; nighttime lights; logistic regression; multi-layer perception; Markov chain

1. Introduction

The Long Island Sound (LIS) is one of the nationally most important estuaries and one of the world's most productive and utilized water bodies. The water quality of the Sound is highly affected by the conditions of its watersheds. A scientific understanding of its watersheds is critical to making effective water policy and management. What has been built on the watersheds and what people have done on the land can have significant influences on LIS and its tributaries. Understanding how land use/cover has changed in its watersheds is critical for effectively managing the coastal water quality.

For the watersheds, land transition is the most significant factor influencing water quality and runoff [1]. The Long Island Sound Watersheds (LISW) cover more than 16,000 square miles and include portions of six states (New York, Connecticut, Rhode Island, Massachusetts, New Hampshire, and Vermont) and the Province of Quebec in Canada [2]. The area is inhabited by 32 million people and includes areas with development levels from the most urbanized to extremely rural (even dedicated wilderness) in North America. It represents a socio-ecological system, the dynamics of which have been affected strongly by changes in land use/cover [3]. The land use activities of the millions of people who live within the LISW have a tremendous impact on the natural habitats of many species and the water bodies of LIS. Recognizing the importance of LISW, a policy [4] of improving management of the watersheds has been adopted in this region. The policy includes reducing impervious surfaces and restoring and protecting vegetation along streams and lakes, which are strongly connected to land use/cover in the watersheds.

Information on historical and potential future watershed land cover is vitally important in watershed management. Many studies have shown that land use/cover change (LUCC) in LISW has highly affected the watersheds by impacting the metabolism and productivity of LIS [1] and increased the scarcity and contamination of water resources [5–7]. In addition to its direct influences on water bodies, LUCC also affects climate change [8], habitat loss [9,10], the spread of invasive species [11], and biota [12] via numerous and complex pathways in the watersheds.

In fact, the land use/cover of this region has gone through tremendous changes over the past four centuries. There were forest clearance and agricultural expansion in the seventeenth century, which reached a peak from 1820 to 1880 [13], and up to 90% forests were cleared for farming by the mid-1800s [14]. Reforestation on abandoned fields began in 1850 and increased progressively through the early twentieth century, and much of the land reverted back to mixed hardwood forest [13]. The reforestation has significant implications for the environment and society, for example, terrestrial carbon storage [15].

However, due to increased forest cutting and the trends of urbanization, additional reforestation and net-positive forest change have been diminished in recent years. Studies show that decreasing forest cover in many locales is related to the expansion of residential and other development [16,17]. Another important reason is the low-density development in rural areas. Brown et al. [18] analyzed the rural land-use trends in the United States within the period from 1950 to 2000 and found that the new pattern of developed land was the increasing attractiveness of nonmetropolitan areas and the decreasing density of settlement. Based on their calculation, by 2000, the area of higher density urbanized development was only 6.7% of the low-density development area. A similar pattern was found in this research area; many rural and forested lands in the region have changed to houses and industrial development in the last fifty years [19–21].

Although the general historical picture of LUCCs is well known and the vulnerability of remaining natural areas to the LUCCs has been assessed, the study of recent trends of LUCC in this region, especially after 2000, is lacking, and what will happen to the landscape of this region under the low-density development pattern has not been extensively studied. The land-use pattern and forest loss have impacts on biodiversity and ecosystem health, and forest change, especially, has significant impacts on carbon sequestration. Therefore, the prediction of future land change is an important part of potential management and policy options. Efforts toward the management of the extent, location, rate, and intensity of human-caused deforestation have become more important in land-management strategies [22]. In this paper, we predicted the future LUCCs in this region under the trend from 1996 to 2006 in order to provide useful information on the potential rates and causes of land change, especially the transitions between development and forest.

Based on existing knowledge and data, prediction of future LUCCs in the region remains difficult and there are uncertainties because LUCCs are not simple processes [23,24]. Where changes will potentially occur may be predicted by using LUCC models [25], which provide predictions through analyzing the factors that may contribute to the changes [26]. In fact, modeling land use/cover changes is a rapidly growing scientific field. Significant progress has been made in developing LUCC models, and there are many different ones in the literature. Recent reviews of the various LUCC models are provided by [27,28]. In this study, we used two integrated models—the logistic regression–Markov chain (LR–MC) model (a combination of logistic regression and Markov chains) and the multi-layer perception-Markov chain (MLP–MC) model (a combination of multi-layer perception and Markov chains), both of which are available in IDRISI Selva [29] to predict the future LUCC in this region. IDRISI Selva was developed by Clark Labs at Clark University, and it is a combination of a geographic information system and remote sensing software. The prediction ability of these two models has been demonstrated in many studies [30–34]. In order to refine our prediction we first compared these two models, and then used the better one (MLP–MC model) for prediction.

The driving forces of LUCC can help us understand the causes of change, and they are also a very important part of prediction. Some relatively comprehensive reviews of common factors involved in modeling LUCC can be found in [35,36]. Three categories of drivers were used in this study. They are biophysical drivers, socio-economic drivers, and proximity characteristics—for example, elevation data, income per capita, and distance to roads, respectively. Those drivers have been used and verified by many studies [37,38]. Moreover, we added nighttime light (NTL) data, which is a good indicator of the economic and urban development, into the driving forces of LUCC. The NTL data can be obtained from the Defense Meteorological Satellite Program/Operational Linescan System (DMSP-OLS) or the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument [39]. The NTL data of 2006 from the DMSP-OLS nighttime light time series dataset was used for this study. The files from this dataset are the combinations of all the available cloud-free data from DMSP-OLS for calendar years of 1992 to 2013. However, the data need to be processed among different years before being used due to the absence of on-board calibration in OLS. To reduce discrepancies, an intercalibration process is necessary [40]. In this paper, because only one year NTL data was used, the intercalibraton step was ignored. The NTL time series dataset has been used in many studies, for example, mapping urbanization dynamics [41], estimating GDP (Gross Domestic Product) growth [42,43], estimating in-use steel stocks in civil engineering and buildings [44], and monitoring economic development from space [45]. Although NTL data has been used in many urban growth and economic activity studies, to the best of our knowledge, its usage in predicting future LUCC is rare.

Although information from LUCCs in LISW is important for sustainable development of LIS and a number of studies provided information on historical LUCCs in LISW, it remains uncertain how the land use/cover will change in this region in the future. The main goal of this study is to gain a sufficient understanding of the future LUCC in LISW. To achieve this goal, we (1) identified the major drivers of LUCC in this area and used NTL as a prediction driver, (2) compared the abilities of the LR–MC model and the MLP–MC model for predicting LUCC in LISW, and (3) predicted the land use/cover in 2026. We intend to gain insights into the following questions: (1) What are the most relevant drivers of land use change in LISW? (2) How are the patterns of land use in LISW changing today? (3) What change do we expect in the next 10 years?

2. Materials and Methods

2.1. Study Area

The study area encompasses over 1500 square miles, 93% of the whole Long Island Sound Watersheds (LISW), which includes diverse landscapes in Connecticut, Massachusetts, New Hampshire, Rhode Island, and Vermont (Figure 1). Within the study area, hundreds of local watersheds drain into streams and rivers, which eventually flow into the Sound. The basin of the Connecticut River is the major component of LISW. It begins in Canada and empties into the LIS. To understand the nutrient dynamics, water quality, and habitats of the Sound, it is necessary to understand the LUCC in LISW. The study area is heavily forested and more than 70% of its surface is covered by forest. The major land-cover types in the watersheds include forest, residential/commercial land, wetland and open land. This region was nearly completely covered by forest before the 17th century and most of the land was cleared for farmlands during the 18th century and the early 19th century [46]. After the widespread farm abandonment, this landscape reached its apex of reforestation. Recently, this area has been under the rapid suburbanization and it is facing a second phase of deforestation caused by urban expansion and land use intensification.

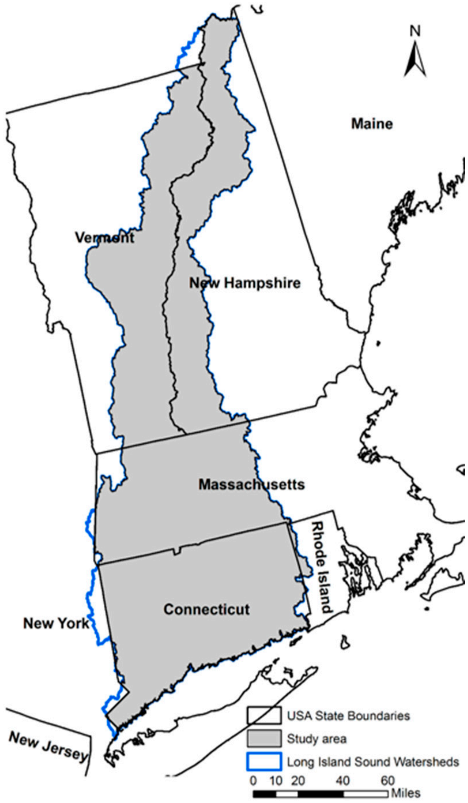


Figure 1. Location of the study area.

2.2. Data

A thematically consistent land use/cover dataset for the years 1996, 2001, and 2006 was created using post-classification processing of the original land-use/cover maps, which were downloaded from the Coastal Change Analysis Program and the National Land Cover Dataset. The created maps have a 60m × 60m pixel resolution. Seven land cover/use types were considered: low-density development, medium-density development, high-density development, forest land, scrub/shrub land, crop/grass land, and other land. Forest land includes deciduous, evergreen, and mixed forests; and other land includes wetland and waterbody. Low-density development is the areas where impervious surfaces account for 20% to 49% percent of total cover. Medium-density development is the areas where impervious surfaces account for 50% to 79% of the total cover. Low-density/medium-density developments are the areas with a mixture of constructed materials and vegetation. High-density development is the areas where impervious surfaces account for 80% to 100% of the total cover and people reside or work in high numbers.

Driving forces are generally divided into three groups [47]: socio-economic drivers, biophysical drivers, and proximity causes (land management variables). The land use change drivers with their data sources utilized in this study are listed in Table 1. These driving forces have been used in many studies [48]. Elevation is important in this landscape because it is prone to flooding. Slope and aspect are vital to land developers who want to minimize landscaping costs. Combining a range of socio-economic drivers is also important for better prediction of LUCC. Recently, with an increase of population, this region was under the processes of urbanization and suburbanization. Since LUCCs such as deforestation are most commonly linked to population growth and income, socio-economic drivers including per capita income, population density, and housing density were used in this study to model the LUCC. The NTL data was also used in this study because it is a good indicator of economic activity [49]. Newly developed areas are often close to the commercial areas, big cities, and roads. Therefore, proximity characteristics (e.g., distance to road and distance to city) were also used as important drivers in this study. Although land-use policies play an important role in driving land-use changes, they were not considered in this study due to the difficulty of quantifying them. Other important drivers may include climate variables; however, climate data were not used in this study because of their low resolutions and poor performance at the scale of analysis used here. Other factors such as housing price are important, but those data are not available for 1996, 2001, and 2006.

Figure 2 contains the maps of the input explanatory drivers. It can be seen that the soil types do not have much spatial variation in the research area. High NTL values, high per capita income, high population density, and high housing density occur along the coastal areas. NTL intensity, population density, and housing density also have comparatively high values along the Connecticut River. Please note that there are no data for population density and housing density near the United States–Canadian border.

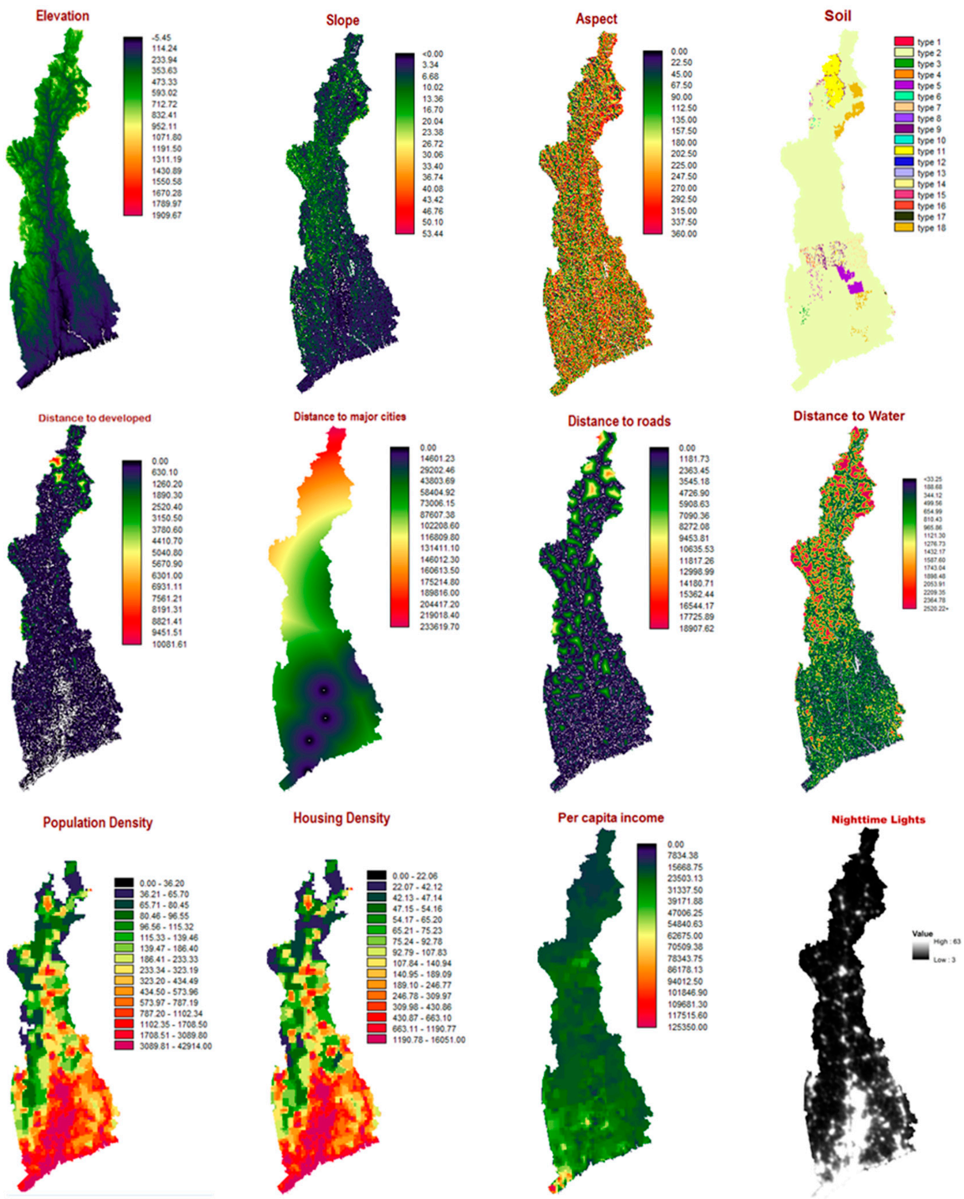


Figure 2. Maps of the input explanatory drivers.

Table 1. Explanatory drivers and their data sources.

	Processed Data	Data Sources
Biophysical drivers	Elevation	USGS National Elevation Dataset (NED)
	Slope	Calculated from USGS Elevation data
	Aspect	Calculated from USGS Elevation data
	Soil type	National Historical Geographic Information System (NHGIS)
Socio-economic drivers	Nighttime light intensity	National Centers for Environmental Information (NCEI)
	Per capita income (gridded)	National Historical Geographic Information System (NHGIS)
	Population density (gridded)	National Historical Geographic Information System (NHGIS)
	Housing density (gridded)	National Historical Geographic Information System (NHGIS)
Proximity causes	Roads (primary and secondary)	US Census Bureau TIGER files
	Distance to road	US Census Bureau TIGER files
	Distance to water	U.S. Geological Survey (USGS)
	Distance to major city	U.S. Geological Survey (USGS)
	Distance to developed area	U.S. Geological Survey (USGS)

2.3. LUCC Prediction

2.3.1. LR–MC Model

The LR–MC model is an integrated model, which combines a logistic regression model and a temporal one-dimensional Markov chain model to predict LUCC. Due to its spatial explicitness and explanatory power, logistic regression has been a very effective model for land use change analyses [50]. However, it lacks the capability of describing the temporal dynamics of LUCC and quantifying change, thus it is integrated with the Markov chain model to overcome these limitations. In the integrated LR–MC model, the logistic regression is first used to investigate the main driving forces determining land use change and generate the probability surface of future land change. Then based on the resulting probability surface of future land change, the Markov chain model is used to estimate the quantity of land use change.

Logistic regression analysis is one of the most frequently used methods for predicting LUCC. It is a method to discover the nonlinear relationship between the dependent variable and independent variables. In this study the dependent variable is land use/cover classes and the independent variables are the driver forces of LUCC, which include biophysical drivers, socio-economic drivers, and proximity cause drivers. Here the logistic regression yields mathematical formulas to quantify the relationships between different land classes and their drivers. Based on a set of scores on the independent variables, the probability of a land-use/cover class change that occurs on any piece of land can be estimated by logistic regression analysis. For instance, the probability of change of a specific land class i, based on a set of variables, can be calculated with the following formula:

$$P\,(y = i|X) = \frac{\exp(B_0 + \sum BX)}{1 + \exp(B_0 + \sum BX)} \tag{1}$$

where P is the probability of the occurrence of land class i at a grid cell; X represents the set of independent variables with $X = (x_1, x_2,..., x_k)$; B_0 is an intercept of the model; B represents the estimated parameters with $B = (b_1, b_2,..., b_k)$.

As aforementioned, the logistic regression model suffers from some limitations in change quantification and temporal analysis [38]. In order to quantify land use change and produce temporal outputs from the logistic regression model, the Markov chain model has to be integrated with logistic regression. The key input parameter of the Markov chain model is the transition probability matrix, which describes the probabilities associated with various land use/cover state changes. The future land use/land cover L_{t+1} can be predicted using the transition probability matrix P and historical land use/land cover L_t by $L_{t+1} = P \times L_t$.

2.3.2. MLP–MC Model

The MLP–MC model is an integrated model which combines the multi-layer perception (MLP) and the temporal one-dimensional Markov chain model to predict the LUCC. In the MLP–MC hybrid approach, the MLP is used to establish functional relationships between the LUCC driving forces. The products of MLP are probability surfaces of each transition, such as the transition from low-density development to medium-density development or the transition from forest to low-density development. The probability surfaces are grid pixels with the same resolution as the land cover maps of 2006 (60 m × 60 m). Each pixel has a value range from 0 to 1 to show the possibility that the transition occurs in the pixel. Then the Markov chain model is used to project the likely total quantity of change and a competitive land allocation to extrapolate land cover into the future.

MLP is one of the most widely used feed-forward artificial neural networks. It has become widely used due to its ability to learn and sort patterns by trial and error. It can model non-linear complex land-use/cover patterns by taking nonlinear complex relationships among the driving forces and LUCC into account. The MLP process contains three layers in a unidirectional process: input, hidden, and output [51]. Each layer consists of nodes in a directed graph and fully connects to the next layer. Except for the input nodes, each node is a neuron (or a processing element) with a nonlinear activation function. The network of the MLP is trained by a supervised learning technique called the backpropagation algorithm, which involves spreading the errors from the output layers to the input layers iteratively in order to correct the values of the weights. The MLP calculates weights for input values, input layer nodes, hidden layer nodes, and output layer nodes using a feed forward manner, which propagates input through the hidden layers and the output layers. The signals transmit from node to node and are modified by weights associated with each connection. The receiving node sums the weighted inputs from all of the nodes connected to it from the previous layer.

Compared to logistic regression, multi-layer perception has two important benefits for LUCC analyses: One is that the input variables do not need to be independent of each other; the other is that it can model several or all the land use/cover transitions at the same time [52]. The MLP–MC hybrid approach can model the spatial and temporal change of land-use/cover by taking the advantages of both the MLP model and the Markov chain model.

3. Results and Analyses

3.1. LUCC Driving Forces

In order to select the driving-force variables, we computed Cramer's V coefficients, which can indicate the degree to which each explanatory variable is associated with the distribution of land cover classes. A driving-force variable is selected provided that it contributes significantly to the explanation of the spatial distribution of the land cover classes of interest. Cramer's V coefficients were computed using the Land Change Modeler software, which was provided by IDRISI Selva [30]. Cramer's V [53] is the most commonly used statistic among the chi-square-based measures of strength of the association between one nominal variable and either another nominal variable or an ordinal variable. The values of Cramer's V range from 0 to 1. A high Cramer's V indicates that the potential explanatory value of the variable is good, but it does not guarantee a strong performance since it cannot account for the mathematical requirements of the modeling approach used and the complexity of the relationship. However, if the Cramer's V is low, it is a good indication that the explanatory variable should be discarded. Variables with a Cramer's V of about 0.15 or higher are useful while those with values of 0.4 or higher are good [54].

Table 2 lists the computed Cramer's V coefficients for the 1996–2001 period. From the overall Cramer's V values in the table, it can be seen that the strongest explanatory variable is the distance to developed area. It has a high association with medium-density development and forest. So it is reasonable for a new developed area to occur usually near the original developed area, thus causing deforestation. The weakest explanatory variables are soil type and per capita income, which have overall values lower than 0.10. Soil type shows limited association with the land cover distribution, because it has little variation over most parts of the study area. The limited association of per capita income with the land cover classes may be caused by the relatively even income in the research area. This does not mean that land cover classes have lower association with economic drivers, because NTL, housing density, and population density show high association with the land cover distribution.

Table 2. Cramer's *V* coefficients over the 1996–2001 period, indicating the quantitative association levels of the explanatory variables (drivers) and the studied land use/cover distribution.

	Low-Density Development	Medium-Densiy Development	High-Density Development	Forest	Crop/Grass	Scrub/Shrub	Other	Overall
Elevation	0.2072	0.2171	0.1778	0.3990	0.1964	0.0276	0.1552	0.1900
Slope	0.1112	0.1305	0.0952	0.3582	0.1739	0.0387	0.3567	0.1923
Aspect	0.0168	0.0260	0.0134	0.1161	0.0556	0.0115	0.3410	0.1414
Soil type	0.0447	0.0634	0.0249	0.1237	0.0847	0.0239	0.0655	0.0579
NTL	0.2844	0.2682	0.2459	0.4254	0.1917	0.0316	0.0779	0.2207
Per capita income	0.0758	0.0924	0.0583	0.1519	0.0925	0.0351	0.0536	0.0756
Housing density	0.2788	0.2484	0.2716	0.3824	0.1443	0.0215	0.0565	0.2108
Population density	0.2794	0.2464	0.2620	0.3765	0.1374	0.0231	0.0476	0.2071
Distance to city	0.1175	0.1406	0.0957	0.2290	0.1265	0.0160	0.0322	0.1080
Distance to developed area	0.1875	0.3969	0.1955	0.4806	0.2240	0.0238	0.0451	0.2903
Distance to road	0.2056	0.3851	0.0908	0.4331	0.2066	0.0263	0.0557	0.2148
Distance to water	0.0658	0.1015	0.0364	0.2681	0.0861	0.0090	0.55476	0.2373

3.2. Model Comparison and Validation

Land cover/use maps obtained through the LR–MC model and MLP–MC model for 2006 are presented in Figure 3, along with the actual 2006 land cover/use map. The two predicted maps forecast the land-cover/use distribution in 2006, based on change parameters between 1996 and 2001. There is a high similarity between projected maps and the actual map. These two models both have good performance. However, there is lower accuracy at the upper edge, which is near the border of the United States and Canada, and the major incorrect predictions happened between forest and scrub/shrub, probably due to the lack of data in this area and the low association between drivers and the scrub/shrub class. In addition, the MLP–MC model has a better performance at the upper edge.

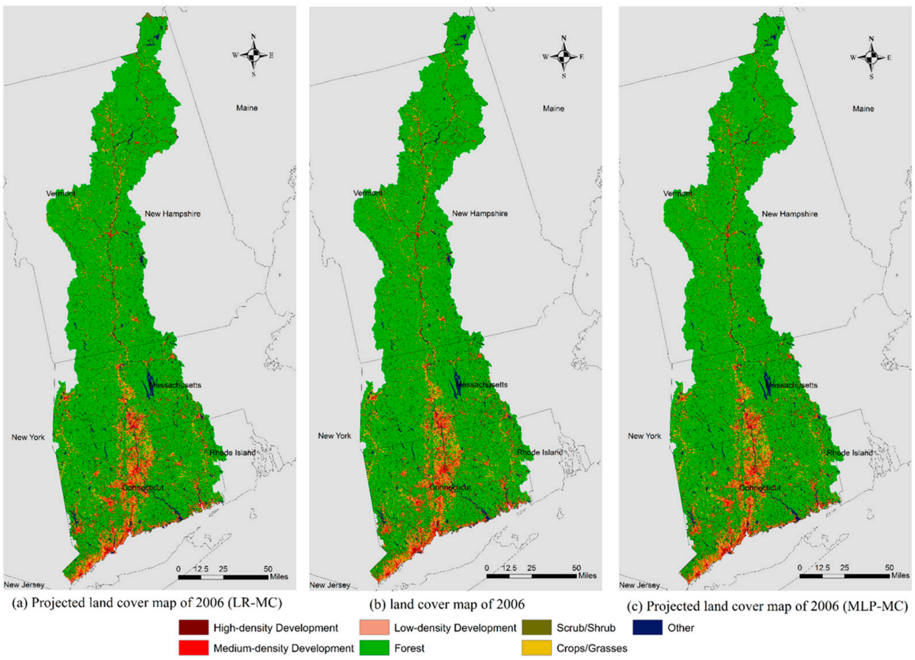


Figure 3. (**a**) Projected 2006 land cover/use map by LR–MC model, (**b**) actual 2006 land cover/use map, and (**c**) projected 2006 land cover/use map by MLP–MC model.

The land-cover/use map generated by the LR–MC model has a total overall accuracy of 98.88% and a Kappa index of 0.993. The land cover/use map generated by the MLP–MC model has a total overall accuracy of 99.04% and a Kappa index of 0.994. Both methods achieved a very high accuracy in LUCC prediction, while the MLP–MC model performed slightly better than the LR–MC model did. The high percentage of unchanged land area is the main reason for the high accuracy.

3.3. LUCC Analysis

Table 3 shows the quantities of each land-cover/use class in different years and their changes in terms of hectares and percentage. The major land-cover/use class in LISW is forest land, accounting for over 73% of the whole area. There was an increase in developed area, scrub/shrub land and crop/grass land, and a decrease in forest land over the two five-year periods of 1996–2001 and 2001–2006. Approximately 7569 ha of forest was lost within the period 1996–2001, while almost twice that area, 12,831 ha of forest, was lost within the next five-year period, so it appears that there was an increasing trend of losing forest. At the same time, low-density development area had an increase of 247 ha during 1996–2001, and its increase was 22 ha during 2001–2006. The medium-density development and the high-density development had a higher area increase during 2001–2006 than during the period 1996–2001. The comparatively high increase of low-density development in the former period might have been caused by population growth and lower land prices in rural areas. Some of the low-density development area was changed to medium-density or high-density development area in the latter period, which might have been caused by the development of amenities and more houses.

Table 3. Quantities of land cover/use changes over time in terms of hectare and percentage of each class.

Class	1996		2001		2006		1996–2001	2001–2006
	ha	%	ha	%	ha	%		
Low-density development	85,229	2.13%	85,476	2.14%	85,497	2.14%	247	22
Medium-density development	308,424	7.72%	308,426	7.72%	309,159	7.74%	3	733
High-density development	22,734	0.57%	22,953	0.57%	24,118	0.60%	219	1165
Forest	2,974,558	74.43%	2,966,988	74.24%	2,954,157	73.92%	−7569	−12,831
Scrub/shrub land	71,335	1.79%	76,536	1.92%	81,241	2.03%	5201	4705
Crop/grass land	403,105	10.09%	404,577	10.12%	410,134	10.26%	1471	5557
Other	130,906	3.28%	131,335	3.29%	131,985	3.30%	429	649

Figure 4 shows the contributions of land-cover/use classes to net changes of developed areas at different densities. Within these two periods, forest had the highest contributions to the increases of all density levels of development areas, except for the high-density development during 1996–2001 when medium-density development had the highest contribution. A remarkable difference between these two periods is that the total contribution of forest to development areas was 359 ha (sum of 194 ha, 114 ha, and 51 ha) in the former period, and it increased to 1602 ha (sum of 235 ha, 886 ha, and 481 ha) in the latter period. The loss of forest caused by development was increasing. Another difference is that there was no transition from low-density development to medium/high-density development in the former period, but the transitions were obvious in the latter period. Consequently, although the net change of low-density development was only a little (22 ha) (Table 3) during the latter period, the change from forest to low-density development was high in this period, even higher than that in the former period. This means that the spatial distribution of low-density development might not change, and the development of amenities and new houses changed much low-density development area to medium/high-density development area in the latter period.

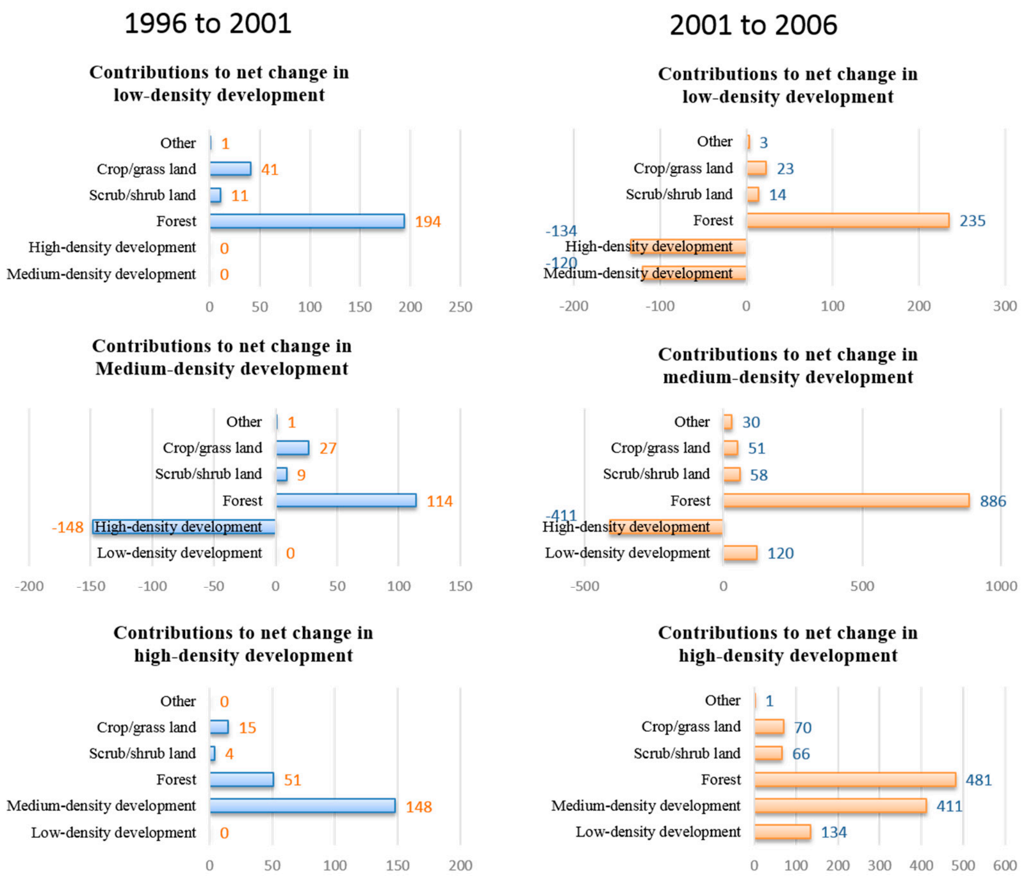
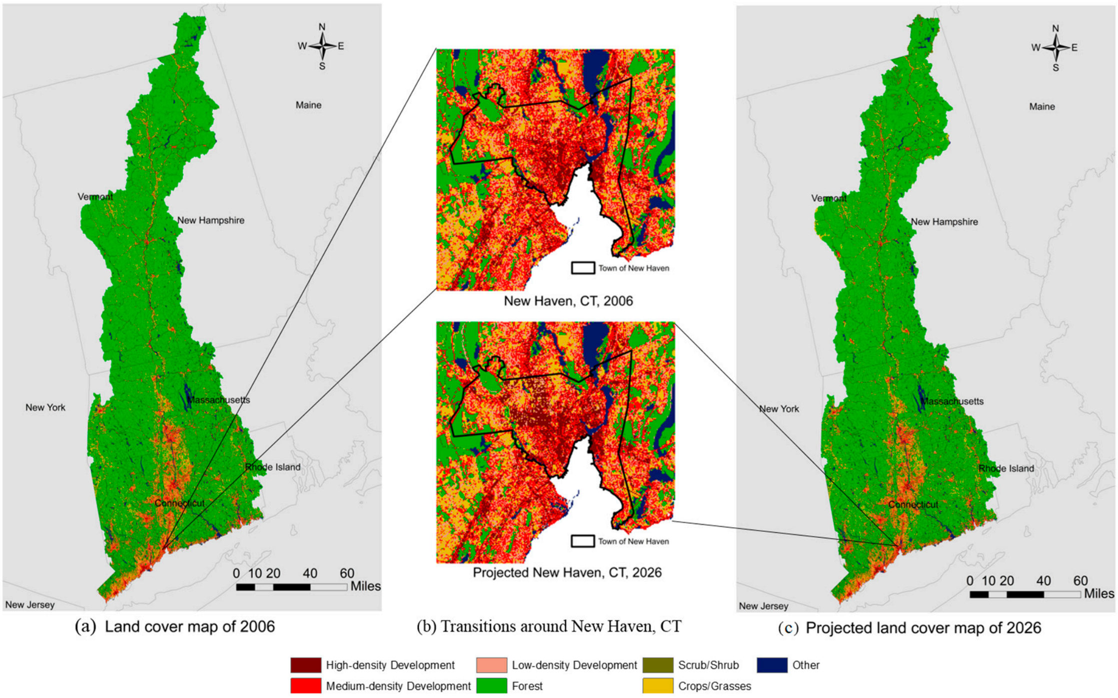


Figure 4. Contributions of land cover/use classes to net changes of developed areas at different density levels (Left column (light blue bars) shows the contributions to net change from 1996 to 2001, and right column (light orange bars) shows the contributions to net change from 2001 to 2006).

Land-cover/use projections for 2026 were carried out by applying the MLP–MC model to analyze possible future LUCCs (Figure 5). The projection was based on the change trends from 1996 to 2001 and from 2001 to 2006, and the combined transition probability matrices from these two periods. Great similarity between the predicted map and the actual 2006 land cover/use map created from remotely sensed imagery can be observed. Urban expansion happens along the coastal area and Connecticut River, where there are high population densities and high NTL values. Predicted results (Table 4) indicate that 50,422 ha of forest will be lost from 2006 to 2026. The increase of high-density development is 5444 ha, which is almost 4 times the increase from 1996 to 2006. Medium-density development contributes the most to this transition, and it is predicted to have a 377 ha loss. The increase of low-density development is also remarkable, with almost five times the increase from 1996 to 2006. These change trends are visible when comparing the changes around major cities between the actual 2006 land cover/use map and the predicted 2026 land cover/use map. For example, the transitions around New Haven are very large; much of medium-density development changes to high-density development, and some forest cover at the urban edge converts to low-density development (e.g., at the top-left corner of Figure 5b).



(a) Land cover map of 2006 (b) Transitions around New Haven, CT (c) Projected land cover map of 2026

Figure 5. (**a**) 2006 land cover/use map, (**b**) transitions around New Haven, CT, and (**c**) projected 2026 land cover/use map by MLP–MC model.

Table 4. Quantities of the predicted 2026 land cover/use classes in terms of hectare and percentage of each class.

Class	Projected 2026 Data		Change from 2006 to 2026 (ha)
	ha	%	
Low-density development	86,929	2.18%	1432
Medium-density development	308,783	7.73%	−377
High-density development	29,562	0.74%	5444
Forest	2,903,735	72.66%	−50,422
Scrub/shrub	99,232	2.48%	17,991
Crop/grass	432,899	10.83%	22,765
Other	135,152	3.38%	3167

4. Discussion

Nighttime light (NTL) data has a high association with land cover/use distribution, which indicates that it can be used in LUCC prediction. High NTL values occur along the coastal area and the Connecticut River. This should be similar to other economic drivers. There are several shortcomings in NTL data, including coarse spatial resolution, limited dynamic range and lack of in-flight calibration [39]. However, compared with other socio-economic drivers, NTL data has some obvious advantages. First, it is easier to access (it is open to everyone and can be downloaded for free). Second, it has global coverage and is not restricted to specific areas. Third, NTL data from the new source (VIIRS) is available for almost every day, except for some situations influenced by clouds or moonlight. Therefore, it can be used as a supplement to economic drivers, or used as a proxy of economic drivers when they are not available in some regions or in a specified year.

The expansion of development areas is a major cause of declining forest cover in many locales [18,55]. The same situation exists in LISW, where low-density development within a commuting distance to metropolitan areas caused the fragmentation of forest. The development of amenities and new houses in low-density development areas pushed the urban edge even farther. Moreover, vegetable production and diary production were pushed further from the expanding urban edge, consequently causing more loss of forest cover. Although surplus dairy production might have led to the abandonment of pasture and the natural conversion to forest [56], the conversion was quite slow. Due to the slower growth rate of the forest in the study area, most mechanically disturbed lands could not directly convert back to forest land. They usually transitioned to shrub land or grass land in some time intervals. Therefore, compared with the loss of forest cover, the reforestation area was quite small. The decrease of forest in this region eventually may cause a loss in carbon storage and sequestration potential. The land cover change also has a direct impact on the Long Island Sound ecosystem. In the watersheds, the rain can carry pollutants from impervious surfaces and flow into the Sound. The increase of development area may lead to more polluted runoff. Polluted runoff can cause low levels of oxygen and high counts of pathogens that lead to the close of beaches and the loss of biodiversity [2]. Landscapes create a broad range of valuable ecosystem services, which should not be ignored while making land-use decisions.

Both integrated models (i.e., LR–MC and MLP–MC) achieved high accuracy in LUCC prediction, but this does not mean that they are perfect for such modeling. In the study area, the number of changed grid cells is much smaller than that of unchanged grid cells, leading to a high accuracy in prediction. Logistic regression has a good performance in modeling the relationships between the drivers and LUCCs. However, quantifying all the potential interactions among the different drivers of LUCC in a logistic regression model is difficult, because of (1) the lack of understanding of all of those factors, (2) the lack of sufficient information, and (3) the restrictions of the functional form of the logistic regression model. Such drawbacks may be overcome by combining it with a Markov chain model. However, the quantification power of a Markov chain model gradually declines as the projected date moves forward [57]. For example, the projected land cover/use map for 2026 cannot achieve as high an accuracy as the projected land cover/use map for 2006 does.

The advantage of MLP is that it is a system capable of modeling complex relationships among variables. Nevertheless, MLP has a "black-box" process—it defines the relationships between drivers and land cover/use change in the hidden layer(s), which makes the integration of expert knowledge difficult. Due to the use of the neural networks, it is difficult to modify the relationship between explanatory variables and change potential with MLP when developing alternative scenarios. This is the limitation of MLP. Models that can incorporate dynamic changes (e.g., different Markov matrices), heterogeneous and nonlinear relationships between LUCC and the underlying drivers, and expert knowledge would allow us to obtain an ensemble of land-change scenarios. Assessments of the performances of different models in predicting potential LUCC are important because inappropriate models may lead to erroneous or inaccurate land conservation and zoning policies.

There are two main limitations of this study that may be remedied in future analyses. The first is that all of the drivers are static; that is, their dynamic changes are not considered. LUCC are complex processes, which are shaped by dynamic and nonlinear interactions of various change drivers. While the relationship between LUCC and some explanatory variables, such as elevation, soil and slope, may be relatively stable over time, the relationship between LUCC and other variables such population density, income, and distance to road may show temporal dynamics. The second is that it did not incorporate the effect of public policies and other potentially important qualitative variables such as cultural values, individual behaviors, and socio-demographic survey data, which are not available. Agent-based land use modelling may be a better approach to effectively incorporate human behavior-related driving forces in LUCC prediction.

5. Conclusions

Mapping land-use/cover change (LUCC) in the Long Island Sound Watersheds (LISW) is important for effective management of the Sound, because land use/cover in the Sound's watersheds has a close relationship with its water quality as the LUCC in the Sound's watersheds may degrade the quality of water flowing through them. However, few studies to date have been undertaken to analyze and predict LUCC in this area. In addition, assessing different approaches for modeling LUCC in this area is also important for understanding the processes that determine the changes. Two integrated models, the LR–MC model and the MLP–MC model, were compared for modeling the LUCC in the LISW in terms of their predictive power and prediction accuracy. While it is impossible to validate the predicted maps for the future land use/cover, we verified the two integrated models using the land-use/cover map for 2006. The validation results show that both methods have good performance and are capable of incorporating environmental and socioeconomic factors in LUCC prediction, while the prediction result of the MLP–MC model has a slightly higher accuracy. The most difficult to predict is the transition between scrub/shrub land and forest land, due to the low correlations of scrub/shrub land with input drivers and that its change may be more dependent on climate factors.

An analysis of past, present and future LUCCs in LISW shows that the increase of developed land area has happened and will continue, similar to the loss of forest area in LISW. Some of the forest loss in the study area was due to residential and commercial development such as construction of houses, other buildings, and golf courses. Forest loss at a high rate was found in regions with high population density. Areas with fast population growth were linked to drastic forest loss. The changes of low-density development pushed the urban edge further and increased the fragmentation of forest. Hence, the key drivers of land transitions in this study region are social-economic drivers and proximity causes. Distance to developed areas has the highest association, followed by distance to road, NTL intensity, population density, and housing density. Importantly, 80% of forest lands in this area are privately owned [58]. The ownership of many forests changed in the last 20 years, and many owners may plan to sell their acreage in the next several years. The changing ownership indicates that, without better management, loss of forest will definitely happen, and consequently threaten air and water quality and wildlife habitat. Even if there are transitions from scrub/shrub and crop/grass to forest, its progress is slow. The reforestation area was quite small compared with the deforestation area.

Since the research area has been undergoing rapid suburban development and the input explanatory variables lack dynamic information, the loss of forest and the increase of developed area may be underestimated. In general, resource management and other related governmental agencies should prepare for the possible LUCCs in the LISW in order to mitigate the impact on water quality and wildlife habitat. The stability of current and future forest services, like carbon stocks and biodiversity, could also benefit from improved analysis of the trends of LUCC in this region.

Acknowledgments: This research is partially supported by USA NSF grant No. 1414108. Authors have the sole responsibility to all of the viewpoints presented in the paper. The publication cost is supported by USA NSF grant No. 1414108.

Author Contributions: Ruiting Zhai, Chuanrong Zhang, Weidong Li contributed to the research design. Ruiting Zhai, Chuanrong Zhang and Weidong Li contributed to the writing of the paper. Ruiting Zhai undertook analysis of datasets. Ruiting Zhai, Chuanrong Zhang, Weidong Li, Mark A. Boyer, and Dean Hanink contributed to the revisions and critical reviews from the draft to the final stages of the paper.

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

References

1. Hopkinson, C.S.; Vallino, J.J. The relationships among man's activities in watersheds and estuaries: A model of runoff effects on patterns of estuarine community metabolism. *Estuaries* **1995**, *18*, 598–621. [CrossRef]
2. Long Island Sound Study. *Protection & Progress, Long Island Sound Study Biennial Report 2011–2012*; R.C. Brayshaw and Company: Warner, NH, USA, 2013.
3. Long Island Sound Study. *Long Island Sound Study CCMP*; Environmental Protection Agency: Washington, DC, USA, 1994.
4. 2003 Long Island Sound Agreement. Available online: http://longislandsoundstudy.net/about/our-mission/sound-agreements/2003-long-island-sound-agreement/ (accessed on 28 September 2016).
5. Bhaduri, B.; Harbor, J.; Engel, B.; Grove, M. Assessing watershed-scale, long-term hydrologic impacts of land-use change using a GIS-NPS model. *Environ. Manag.* **2000**, *26*, 643–658. [CrossRef] [PubMed]

6. Tong, S.T.; Chen, W. Modeling the relationship between land use and surface water quality. *J. Environ. Manag.* **2002**, *66*, 377–393. [CrossRef]
7. Meador, M.R.; Goldstein, R.M. Assessing water quality at large geographic scales: Relations among land use, water physicochemistry, riparian condition, and fish community structure. *Environ. Manag.* **2003**, *31*, 0504–0517. [CrossRef] [PubMed]
8. Meyer, J.L.; Sale, M.J.; Mulholland, P.J.; Poff, N.L. Impacts of climate change on aquatic ecosystem functioning and health[1]. *J. Am. Water Resour. Assoc.* **1999**. [CrossRef]
9. Strayer, D.L.; Beighley, R.E.; Thompson, L.C.; Brooks, S.; Nilsson, C.; Pinay, G.; Naiman, R.J. Effects of land cover on stream ecosystems: Roles of empirical models and scaling issues. *Ecosystems* **2003**, *6*, 407–423. [CrossRef]
10. Akber, M.A.; Shrestha, R.P. Land use change and its effect on biodiversity in Chiang Rai province of Thailand. *J. Land Use Sci.* **2015**, *10*, 108–128. [CrossRef]
11. Scott, M.C.; Helfman, G.S. Native invasions, homogenization, and the mismeasure of integrity of fish assemblages. *Fisheries* **2001**, *26*, 6–15. [CrossRef]
12. Quinn, J.M.; Cooper, A.B.; Davies-Colley, R.J.; Rutherford, J.C.; Williamson, R.B. Land use effects on habitat, water quality, periphyton, and benthic invertebrates in Waikato, New Zealand, hill-country streams. *N. Z. J. Mar. Freshw. Res.* **1997**, *31*, 579–597. [CrossRef]
13. Foster, D.R. Land-use history (1730–1990) and vegetation dynamics in central New England, USA. *J. Ecol.* **1992**. [CrossRef]
14. Foster, D.R.; Aber, J. *Forests in Time. Ecosystem Structure and Function as a Consequence of 1000 Years of Change. Synthesis volume of the Harvard Forest LTER Program*; Yale University Press: New Haven, CT, USA, 2004.
15. Houghton, R.; Hackler, J. Changes in terrestrial carbon storage in the United States. 1: The roles of agriculture and forestry. *Glob. Ecol. Biogeogr.* **2000**, *9*, 125–144. [CrossRef]
16. Heimlich, R.E.; Anderson, W.D. *Development at the Urban Fringe and Beyond: Impacts on Agriculture and Rural. Economic Research Service*; US Department of Agriculture: Washington, DC, USA, 2001.
17. Hall, B.; Motzkin, G.; Foster, D.R.; Syfert, M.; Burk, J. Three hundred years of forest and land-use change in Massachusetts, USA. *J. Biogeogr.* **2002**, *29*, 1319–1335. [CrossRef]
18. Brown, D.G.; Johnson, K.M.; Loveland, T.R.; Theobald, D.M. Rural land-use trends in the conterminous United States, 1950–2000. *Ecol. Appl.* **2005**, *15*, 1851–1863. [CrossRef]
19. Mosher, E.S.; Silander, J.A., Jr.; Latimer, A.M. The role of land-use history in major invasions by woody plant species in the northeastern North American landscape. *Biol. Invasions* **2009**, *11*, 2317–2328. [CrossRef]
20. Foster, D.; Aber, J.; Cogbill, C.; Hart, C.; Colburn, E.; D'Amato, A.; Donahue, B.; Driscoll, C.; Ellison, A.; Fahey, T.; et al. *Wildlands and Woodlands: A vision for the New England Landscape*; Harvard University Press: Cambridge, MA, USA, 2010.
21. Jeon, S.B.; Olofsson, P.; Woodcock, C.E. Land use change in New England: A reversal of the forest transition. *J. Land Use Sci.* **2014**, *9*, 105–130. [CrossRef]
22. Drummond, M.A.; Loveland, T.R. Land-use pressure and a transition to forest-cover loss in the Eastern United States. *BioScience* **2010**, *60*, 286–298. [CrossRef]
23. Brown, D.G.; Duh, J.D. Spatial simulation for translating from land use to land cover. *Int. J. Geogr. Infor. Sci.* **2004**, *18*, 35–60. [CrossRef]
24. Brown, D.G.; Aspinall, R.; Bennett, D.A. Landscape models and explanation in landscape ecology—A space for generative landscape science? *Prof. Geogr.* **2006**, *58*, 369–382. [CrossRef]
25. Veldkamp, A.; Lambin, E.F. Predicting land-use change. *Agric. Ecosyst. Environ.* **2001**, *85*, 1–6. [CrossRef]
26. Overmars, K.; de Koning, G.; Veldkamp, A. Spatial autocorrelation in multi-scale land use models. *Ecol. Model.* **2003**, *164*, 257–270. [CrossRef]
27. Brown, D.G.; Verburg, P.H.; Pontius, R.G.; Lange, M.D. Opportunities to improve impact, integration, and evaluation of land change models. *Curr. Opin. Environ. Sustain.* **2013**, *5*, 452–457. [CrossRef]
28. Rosa, I.; Ahmed, S.E.; Ewers, R.M. The transparency, reliability and utility of tropical rainforest land-use and land-cover change models. *Glob. Chang. Biol.* **2014**, *20*, 1707–1722. [CrossRef] [PubMed]
29. Eastman, J.R. *IDRISI Selva*; Clark University: Worcester, MA, USA, 2012.
30. Pueyo, Y.; Beguería, S. Modelling the rate of secondary succession after farmland abandonment in a Mediterranean Mountain Area. *Landsc. Urban Plan.* **2007**, *83*, 245–254. [CrossRef]

31. Oñate-Valdivieso, F.; Sendra, J.B. Application of GIS and remote sensing techniques in generation of land use scenarios for hydrological modeling. *J. Hydrol.* **2010**, *395*, 256–263. [CrossRef]
32. Pérez-Vega, A.; Mas, J.F.; Ligmann-Zielinska, A. Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environ. Model. Softw.* **2012**, *29*, 11–23. [CrossRef]
33. Arsanjani, J.J.; Helbich, M.; Kainz, W.; Boloorani, A.D. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 265–275. [CrossRef]
34. Wang, W.; Zhang, C.; Allen, J.M.; Li, W.; Boyer, M.A.; Segerson, K.; Silander, J.A. Analysis and prediction of land use changes related to invasive species and major driving forces in the state of Connecticut. *Land* **2016**. [CrossRef]
35. Poelmans, L.; Van Rompaey, A. Complexity and performance of urban expansion models. *Comput. Environ. Urban Syst.* **2010**, *34*, 17–27. [CrossRef]
36. Dubovyk, O.; Sliuzas, R.; Flacke, J. Spatio-temporal modelling of informal settlement development in Sancaktepe district, Istanbul, Turkey. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 235–246. [CrossRef]
37. Musaoglu, N.; Tanik, A.; Kocabas, V. Identification of land-cover changes through image processing and associated impacts on water reservoir conditions. *Environ. Manag.* **2005**, *35*, 220–230. [CrossRef] [PubMed]
38. Hu, Z.; Lo, C. Modeling urban growth in Atlanta using logistic regression. *Comput. Environ. Urban Syst.* **2007**, *31*, 667–688. [CrossRef]
39. Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proc. Asia-Pacific Adv. Netw.* **2013**, *35*, 62–69. [CrossRef]
40. Elvidge, C.D.; Ziskin, D.; Baugh, K.E.; Tuttle, B.T.; Ghosh, T.; Pack, D.W.; Erwin, E.H.; Zhizhin, M. A fifteen year record of global natural gas flaring derived from satellite data. *Energies* **2009**, *2*, 595–622. [CrossRef]
41. Zhang, Q.; Seto, K.C. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sens. Environ.* **2011**, *15*, 2320–2329. [CrossRef]
42. Ghosh, T.; Powell, R.L.; Elvidge, C.D.; Baugh, K.E.; Sutton, P.C.; Anderson, S. Shedding light on the global distribution of economic activity. *Open Geogr. J.* **2010**. [CrossRef]
43. Henderson, J.V.; Storeygard, A.; Weil, D.N. Measuring economic growth from outer space. *Am. Econ. Rev.* **2012**, *102*, 994–1028. [CrossRef] [PubMed]
44. Hsu, F.C.; Elvidge, C.D.; Matsuno, Y. Exploring and estimating in-use steel stocks in civil engineering and buildings from night-time lights. *Int. J. Remote Sens.* **2013**, *34*, 490–504.
45. Keola, S.; Andersson, M.; Hall, O. Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. *World Dev.* **2015**, *66*, 322–334.
46. Cronon, W. *Changes in the Land: Indians, Colonists, and the Ecology of New England*; Hill & Wang: New York, NY, USA, 2011.
47. Turner, B.L., II; Skole, D.; Sanderson, S.; Fischer, G.; Fresco, L.; Leemans, R. *Land-Use and Land-Cover Change, Science/Research Plan*; IIASA: Laxenburg, Austria, 1995.
48. Rodríguez Eraso, N.; Armenteras-Pascual, D.; Alumbreros, J.R. Land use and land cover change in the Colombian Andes: Dynamics and future scenarios. *J. Land Use Sci.* **2013**, *8*, 154–174. [CrossRef]
49. Mellander, C.; Lobo, J.; Stolarick, K.; Matheson, Z. Night-time light data: A good proxy measure for economic activity? *PloS ONE* **2015**. [CrossRef] [PubMed]
50. Munroe, D.K.; Southworth, J.; Tucker, C.M. Modeling spatially and temporally complex land-cover change: the case of western Honduras *. *Prof. Geogr.* **2004**, *56*, 544–559.
51. Mas, J.F.; Puig, H.; Palacio, J.L.; Sosa-López, A. Modelling deforestation using GIS and artificial neural networks. *Environ. Model. Softw.* **2004**, *19*, 461–471. [CrossRef]
52. Eastman, J.R. *IDRISI Taiga Guide to GIS and Image Processing*; Clark Labs, Clark University: Worcester, MA, USA, 2009.
53. Liebetrau, A.M. *Measures of Association*; Sage: Newbury Park, CA, USA, 1983.
54. Eastman, J.R. *IDRISI Andes Guide to GIS and Image Processing*; Clark University: Worcester, MA, USA, 2006; pp. 87–131.
55. Heimlich, R.E.; Anderson, W.D. *Development at the Urban Fringe and Beyond: Impacts on Agriculture and Rural*; US Department of Agriculture: Washington, DC, USA, 2001.

56. Alig, R.; Mills, J.; Butler, B. Private timberlands: Growing demands, shrinking land base. *J. For.* **2002**, *100*, 32–37.
57. Mas, J.F.; Kolb, M.; Paegelow, M.; Olmedo, M.T.C.; Houet, T. Inductive pattern-based land use/cover change models: A comparison of four software packages. *Environ. Model. Softw.* **2014**, *51*, 94–111. [CrossRef]
58. The Nature Conservancy. A Policy Agenda for Conserving New England's Forests. 2013. Available online: http://www.nature.org/ourinitiatives/regions/northamerica/unitedstates/massachusetts/newsroom/policy-agenda-ne-forests-13.pdf (accessed on 1 October 2016).