

Article

Evaluation of Driving Forces of Land Use and Land Cover Change in New England Area by a Mixed Method

Ruiting Zhai ^{1,*}, Chuanrong Zhang ², Weidong Li ², Xiang Zhang ³ and Xueke Li ¹¹ Department of Geography, University of Connecticut, Storrs, CT 06269, USA; xueke.li@uconn.edu² Department of Geography and Center for Environmental Sciences and Engineering, University of Connecticut, Storrs, CT 06269, USA; chuanrong.zhang@uconn.edu (C.Z.); weidong.li@uconn.edu (W.L.)³ Department of Physics, University of Connecticut, Storrs, CT 06269, USA; Xiang.zhang@uconn.edu

* Correspondence: ruiting.zhai@uconn.edu

Received: 26 April 2020; Accepted: 25 May 2020; Published: 27 May 2020

Abstract: Understanding the driving forces of land use/cover change (LUCC) is a requisite to mitigate and manage effects and consequences of LUCC. This study aims to analyze drivers of LUCC in New England, USA. It combines meta-study, GIS, and machine learning to identify the important factors of LUCC in the area. Firstly, we conducted a meta-study of the research on LUCC in the New England area and specifically focused on the driving forces analysis. The meta-analysis revealed that the LUCC studies in the research area were highly related with many other research topics, and population and economic factors were the most mentioned drivers of the LUCC. The drivers of LUCC in this study area for the past several decades were relatively well analyzed. However, the study of the main driving forces of recent LUCC is lacking. Then, the determinants of LUCC for the recent years were quantitatively assessed using the random forests (RF) model along with geospatial data processing. Two planning regions in Connecticut and one planning region in Massachusetts were selected to serve as the case study areas. Investigated variables included environmental and biophysical variables, location measures of infrastructure and existing land use, political variables, and demographic and social variables. These drivers were examined for their relations with LUCC processes. Their importance as driving forces was ranked by the RF method. The results show both consistency and inconsistency between the meta-analysis and the RF method. We found that this mixed method can enhance our understanding of driving forces of LUCC and improve the selection quality of important drivers for modeling LUCC. With more solid information, better land management advices for sustainable development may also be provided.

Keywords: land use/cover change; driving forces; meta-study; random forests

1. Introduction

Land use/cover change (LUCC) is a main component of global environmental change, which is mainly an intended or unintended outcome of human activities. LUCC has profound effects on climate, biodiversity, soil condition, water flows, and ecosystem services at local, regional, and global scales [1–4], which in turn affect land-use decisions. A better understanding of processes, trends, causes, and consequences of LUCC is of crucial importance to land planners, ecologists, and others [5,6]. Moreover, identification of driving forces that cause the dominant LUCC is essential for establishing management strategies and policies to mitigate or prevent negative effects of LUCC or for predicting future changes using models [7]. Identification of driving forces that cause the

dominant land use and land cover transitions will further help managers to establish policies that mitigate or prevent negative effects of LUCC.

The driving forces of LUCC are usually a mix of anthropogenic (e.g., demographic, political, economic, technological, and cultural) and biophysical factors (e.g., climate, soil, and topography) with direct or indirect impacts. The anthropogenic drivers may include well-defined factors, such as population growth [8], and the complex interactions as a result of institutional or cultural influences [9]. Biophysical drivers of LUCC, such as elevation and slope [10], have been primarily studied due to availability of data. Comparatively, climate as one of the important drivers of LUCC is difficult to detect and quantify in a short-term or in a local research, because of its long-term consequences and low resolution of climate data. Although the drivers of LUCC are recognized as important, the complex and diverse interactions among nature, economy, and social systems make them difficult to be identified and quantified.

With development of remote sensing (RS) and geographic information system (GIS) technologies and the reliable sources of data for assessing spatial-temporal LUCC, the research on the driving forces based on the biophysical factors and socio-economic factors has been improved. Previous research often investigated the driving forces of LUCC through a form of descriptive analysis [11,12], or most recently correlation or multiple regression analysis [13,14]. However, these studies either did not integrate both socio-economic and biophysical factors to examine their effects on LUCC or were not able to identify which drivers play important roles in LUCC. Thus, the driver forces of LUCC are still contentious issues and further research is necessary [15].

The understanding of the drivers behind landscape changes has been at the center of recent landscape research, which had been used as a framework for understanding the causes, processes, and outcomes of landscape change [16]. Due to the strong variation of existing research over diverse spatial and temporal scales, the understanding of driving forces of landscape changes is still limited [17]. As the promising approaches that can synthesize and upscale local-level studies to a more general insight, systematic review and meta-analysis techniques have raised attention. A meta-study in land use science is a valuable tool, which can synthesize knowledge and extract broader scale drivers of change from various publications and various individual case studies [18]. In recent years, meta-analyses have provided valuable insights into drivers of land use change; for example, a study identified the proximate and underlying drivers of land change across the whole Europe based on a systematic synthesis of 144 studies [16]. On the other hand, given the complexity of LUCC, measuring correlations between driving forces and LUCC is difficult, and there is no agreement on optimal tools yet. RF, a technique based on bootstrap aggregation, has great potential in this area, due to its capability of identifying complex interactive and non-linear relationships and deriving importance of the predictors [19]. RF has been used widely in land cover classification [20,21]; however, there were few applications in studying driving forces.

The set of “new directions” to tackle the challenges of landscape change studies was defined by [17]. One of the new directions is the integrative approach, which requires quantitative data to be used together with qualitative information and narrative elements for studying the cause and effect relation between driving forces and landscape change to find the general patterns of change. Here in this study, with the guidance of this direction, a mixed method for detecting drivers of LUCC was suggested. This study combined meta-study, GIS, and machine learning to detect drivers of LUCC in the New England area. It started with a meta-study to obtain preliminary insight into LUCC and the drivers of LUCC in the study area. Remote sensing and geospatial data, the reliable data sources for understanding LUCC and determining the drivers of change, were then used for modeling and data analyses. In addition, we applied the random forests (RF) method to extract importance of the drivers for the study area. To the best of our knowledge, the integration of meta-study, GIS, and RF for seeking a deep understanding of the driving forces of landscape change has been barely carried out in early studies. The research area, New England, USA, has a well-documented land use history. Until around 1650, the region was almost completely covered with forest and under the non-anthropogenic disturbances and successional dynamics [22]. Since the Europeans settled in early 1700, more than half of the land was cleared to create croplands and pastures and remaining forest

areas were cut extensively for timber harvest for construction [23]. The forest clearance and agricultural expansion reached a peak, with up to 60%–80% of the land being cleared, in the mid-1800s. Around the same time, farming began to decrease and regrowth of secondary forests on abandoned agricultural land started. As quickly as the land was cleared, much of the land reverted to forest in the early 1990s. New England has been regarded as a primary example of forest transition, which refers to expansion of forest areas. However, New England is facing forest loss and more complex land cover change in recent years [24,25]. The recent land cover change and forest change in New England are not completely unknown. What remains uncertain though are the driving forces of land change during the last two decades.

The objectives of our study are to identify and characterize the most relevant factors responsible for LUCC, especially for recent years, in the New England area, through a new framework that involves a meta-study, GIS, and machine learning. The New England area is an important region for land transition with a diverse landscape. We tried to answer the following questions: (1) How and why is land use/cover changing in the New England area based on the previous studies? (2) What are the parameters that drive LUCC? Which ones are the most relevant driving forces of LUCC? (3) What are the main drivers for recent LUCC in the New England area? How is the consistency between the meta-study result and the RF method result? (4) What are the advantages and disadvantages of the mixed method in our case study?

2. Methods

2.1. Study Area

The study area, the New England region ranges from 41°50' N to 47°29' N latitude and 66°54' W to 73°45' W longitude, is located in the Northeastern United States and includes the states of Connecticut (CT), Rhode Island (RI), Massachusetts (MA), New Hampshire (NH), Vermont (VT), and Maine (ME) (Figure 1). On the one hand, New England is heavily forested, and New Hampshire is the second most forested state of the United States. On the other hand, New England includes some states with high population density, such as Rhode Island and Massachusetts. The meta-study focused on the land use/cover change studies in which the study areas are located in the New England area, cover the whole New England area, or contain parts of the New England area. At the meantime, two planning regions in Connecticut (Northeastern and Southcentral) and one planning region in Massachusetts (Metrowest) were selected for our case study for quantifying the determinants of recent LUCC using the random forests approach (Figure 2). These three planning regions can provide comparison between the relatively rural area (Northeastern) and urban area (Southcentral) within the same state (CT) and comparison between two relatively urban areas (Southcentral and Metrowest) in the two different states (CT and MA). The planning regions are the focus of the RF approach because they can provide a convenient scale for examining important land transition processes, and yet represent a feasible area for data collection and analysis, which ensure evaluation of most available drivers.

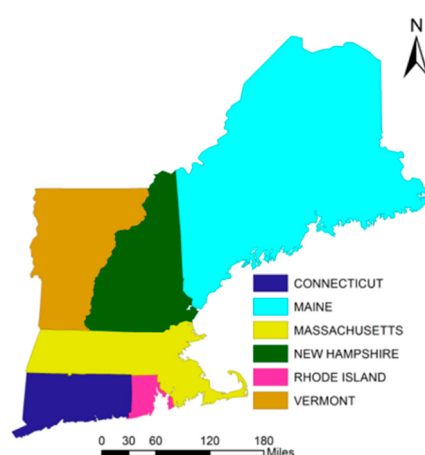
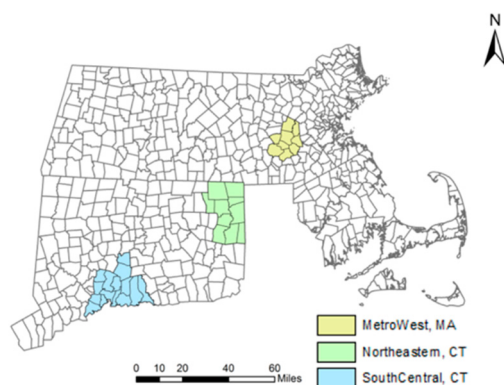


Figure 1. The New England.**Figure 2.** The three planning regions in Massachusetts (MA) and Connecticut (CT).

2.2. Meta-Study

As a study mentioned, “the term meta-studies includes meta-analyses, systematic reviews, and other secondary studies that aim to synthesize case-study findings [18].” The meta-study in this research does not include rigorous statistical treatment of case studies. The robust statistical treatment is usually not possible in land use study due to the differences of case studies in spatial scale, time period and design, complexity of empirical case studies, diversity in data and methods, and different perspectives of researchers from different disciplines [18]. Meta-studies have been used to assess, for instance, long-term urbanization across the globe [26], wetland conversion [27], drivers of typical deforestation [28], causes of desertification [29], and tropical agriculture intensification [30]. However, the meta-studies of LUCC processes and driving forces are still limited.

In our meta-study, we collected case study information on land use/cover change from peer-reviewed articles, reports, book chapters, proceedings of conferences, and PhD theses. We conducted comprehensive search of Elsevier Scopus databases and Google Scholar databases using word combinations of “land use change”, “land cover change”, “land use/cover change”, “New England”, “Northeastern”, “Maine”, “Vermont”, “Massachusetts”, “Rhode Island”, “Connecticut”, and “New Hampshire” in English. These studies were first examined to ensure that the location of study site is in, is part of, or contains New England area. Around 100 papers were collected. We further restricted ourselves to studies that are within a suitable time interval, at least including LUCC after 1980s. Then we excluded the studies just about impact of LUCC (e.g., [31]), composition change of forest (e.g., [22]), and methodology studies (e.g., [32,33]). Thus, totally 54 papers were selected finally, which are all included in Supplementary.

2.3. The Random Forests (RF) Approach

The RF approach is one of the ensemble methods and was classified as one of machine learning approaches at the end of the nineties [34]. It is a popular and efficient algorithm to handle both classification and regression problems. It is based on bootstrap aggregation, which can overcome the overfitting problem of the decision tree model and can rank the explanatory variables using the random forests score of importance [19]. Recently, the RF approach is receiving highlighted interest in land-cover classification using multispectral and hyperspectral satellite sensor imagery [35,36], and lidar and radar data [37–40]. Meanwhile, its power on detecting variable importance has also received a growing attention. For example, [41] applied the RF approach to identify drivers of cropland and urban land changes in Jiangxi province, China, and confirmed its capability for assessing of the complex drivers of LUCC. Despite its popularity and efficiency, the application of RF to the analysis of the spatial driving forces of LUCC data is rare.

The RF method was chosen for evaluating the driving forces of LUCC in the research area for several reasons. First, it has been proved to be able to deal with complex nonlinear problems and

address different types of response variables, such as numerical data (e.g., population density) and categorical data (e.g., soil types). Second, it can also make faithful data descriptions without strong model assumptions. It does not require the normal distribution of the dependent variable and is insensitive to collinearity issues. Another advantage of RF is that it can be trained in parallel. Therefore, in a large LUCC dataset, training of the RF model can be fast. In addition, the RF method can gain good prediction accuracy. Finally, the results of the RF method can be easily interpreted, especially for the feature importance. The feature importance generated by the RF method is normalized to a value between 0 and 1 by dividing the sum of all feature importance values. A higher value indicates a larger impact on the LUCC. The sum of all feature importance values is equal to one.

2.4. Driver Classification and Data

Based on previous studies, proximate causes and underlying driving forces are the factors that affect LUCC processes. Proximate causes are the actual process of LUCC, such as deforestation or urbanization, and underlying driving forces are basic societal or environmental factors that cause these changes, such as income or elevation variation (for more information, see [28,42,43]). Sometimes, the proximate causes and underlying driving forces were mixed together; therefore, generalization of driving factors is not easy. In fact, proximate causes may represent the outcomes of human decisions, which may be affected by the underlying driving factors. For example, a study summarized the worldwide urban expansion studies and found that the observed urban land expansion mainly depended on the interaction with demographic and economic drivers, such as population and GDP growth [26]. Hence, the focus of our study is the underlying driving forces, which has been divided into economic, demographic, political, biophysical, and location drivers (Table 1).

Table 1. Category, descriptions, and examples of the underlying driving forces.

Category	Descriptions	Examples
Demographic drivers	Demographic factors that affect size and composition of population and households.	Emigration/immigration, population
Economic drivers	Economic factors that are the components of economic developments and determinants.	Incomes/households
Political drivers	Diverse policies that affect land use/cover change at various government levels.	Owners/planning polices
Biophysical drivers	Natural factors including topography and climatic conditions that affect land use/cover change.	Elevation/slope Climate
Location drivers	Drivers that focus on location, which may include human activities or immediate actions at the local level.	Distance to road Distance to water

Site-specific studies are needed in order to comprehensively assess the driving forces in New England areas. Three planning regions in CT and MA are selected in this study. To ensure the best data quality, we used different data sources. The thematically-consistent land use/land cover (LULC) datasets of the years 1990 and 2010 for MetroWest, MA were created using post-classification processing of the original LULC maps from the MassGIS [<https://www.mass.gov/orgs/massgis-bureau-of-geographic-information>], and the thematically-consistent LULC datasets for Northeastern, CT and SouthCentral, CT were created using post-classification processing of the original LULC maps from the CLEAR [<https://clear.uconn.edu/index.htm>] (Figure 3). The explanatory variables used in the importance assessment of drivers by the RF approach are listed in Table 2. The explanatory variables used in this study were also collected from the MassGIS (for MetroWest, MA) and CLEAR (for Northeastern, CT and SouthCentral, CT) with the same spatial resolution 30 m × 30 m as the LULC maps.

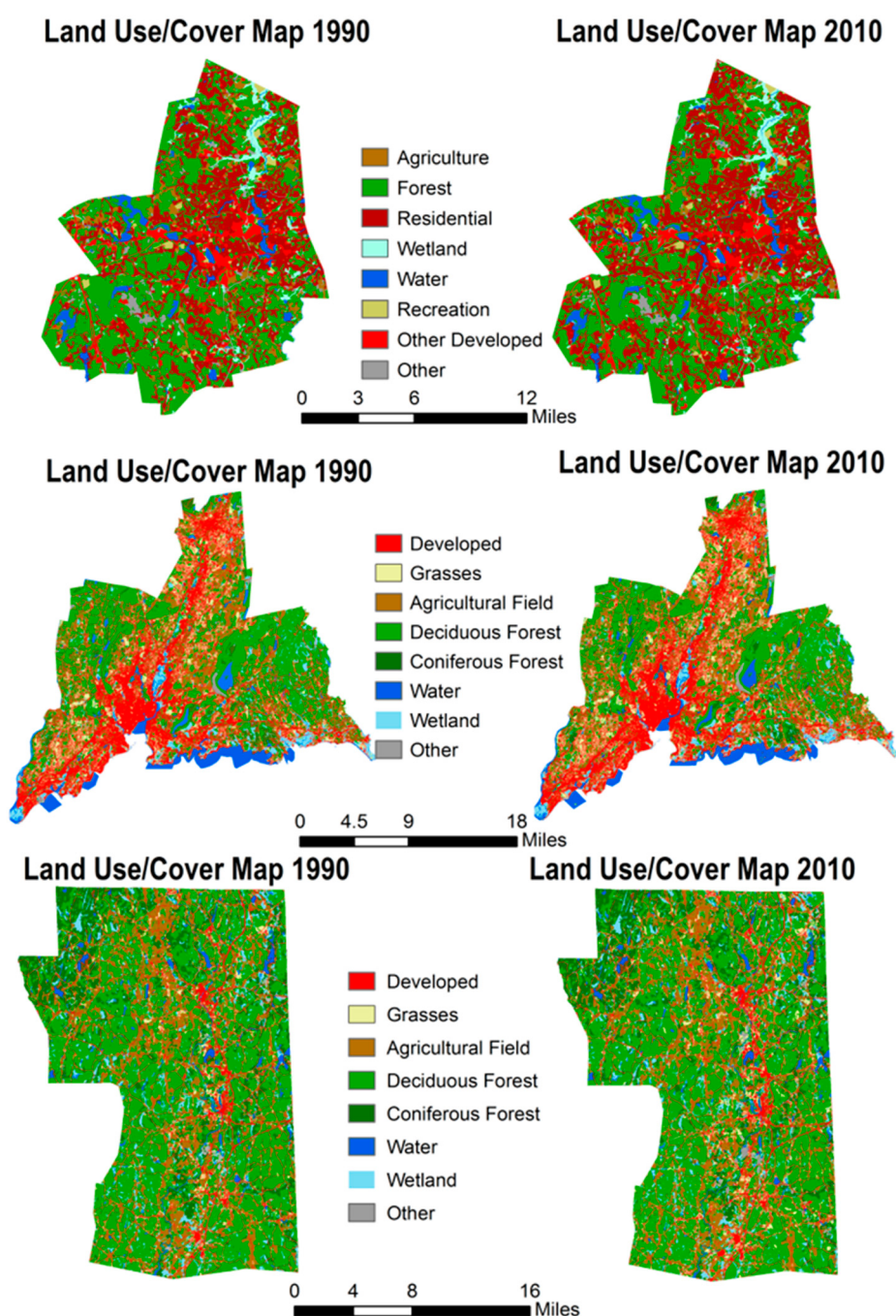


Figure 3. The land use/land cover data of the three planning regions, top: MetroWest, MA; middle: SouthCentral, CT; bottom: Northeastern, CT.

Table 2. Acronyms, names, and descriptions of the driving forces used in this study.

Acronyms	Drivers	Descriptions
DEM	Elevation	Elevation in m
SL	Slope	Slope in degrees
AS	Aspect	Aspect in degrees
ST	Soil type ^a	Categorical soil type data
POPE	Population density earlier ^b	Population density of earlier year
POPL	Population density Later ^c	Population density of later year
POPC	Population density change	Changes in population density from the earlier year to the later year
MDA	Median age ^d	Median age of the population at census block level in 1990

PCI	Per capital income	Averaged income earned per person in 1990
MHI	Median house income ^e	Median household income at census block level in 1990
MHV	Median house value	Median house value at census block level in 1990
PA	Protected area	Reserved area, no development allowed
DW	Distance to water	Euclidean distance from water
DR	Distance to roads	Euclidean distance from roads
DD	Distance to development	Euclidean distance from existed developed area
RD	Roads	Binary roads data
RI	Rivers	Binary rivers data
HUSE	House density earlier ^b	House density of the earlier year
HUSL	House density Later ^c	House density of the later year
HUSC	House density change	Changes in House density from the earlier year to the later year

^a No soil type data of MetroWest. ^b The earlier year of three planning regions is 1990. ^c The later year of MetroWest is 2000 and the later year of Northeastern and SouthCentral is 2010. ^d No median age data of Northeastern and SouthCentral. ^e No median house income data of Northeastern and SouthCentral.

3. Results

3.1. Meta-Study Results

The analysis was limited to LUCC reported in case study areas ranging from small towns to large multi-state regions within or involving the New England (NE) area. Following the study areas of different cases, we found that 14 (i.e., 26%) of all cases occur at the whole NE level (including study areas slightly larger than NE; for example, a study includes NE and New York state [44]); 16 (i.e., 30%) include more than one state in NE but does not cover the entire NE; 7 (13%) are located in MA; 6 (11%) are located in CT; 4 (7%) are located in ME; 3 (6%) are located in RI; 3 (5%) are located in NH; and 1 (2%) is located in VT (Figure 4). More than half of the cases (54%) are at regional scales. 24 (i.e., 44%) of all cases occur at the state level and among the six states, MA and CT have larger numbers of cases. In order to test whether there are differences among case locations, the information on LUCC and its underlying drivers were analyzed based on the case study locations. Totally four location sections were grouped, the NE, the partial NE, the south three states (CT, MA, and RI) of NE, and the rest three states of NE.

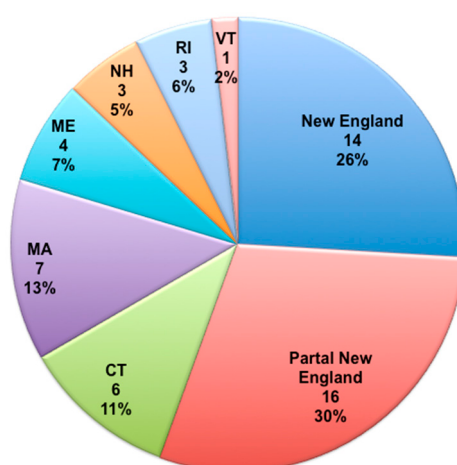


Figure 4. Study areas of the 54 cases.

After reviewing all the 54 papers, we summarized the information on the land cover types involved in each paper, the proportion of the LUCC analysis, and the magnitude of drivers included (Table 3). The cases involving all land cover types tend to be equally widespread (50% to 55%) among regional cases, except for north NE. The cases analyzing all land cover types in north NE hold a higher proportion (88%) than the other regional cases. The cases that only analyzed forest land (30% of all)

and other land cover types (15% of all, such as urban land and protect area) occupy nearly half of all the cases. Meanwhile, regional variations across the cases are significant. The forest-only study is more common in the NE and partial NE cases, while the cases with other land types just occur in partial NE and south NE; this situation is reasonable since the south NE is more developed than the north NE and the other land types are mainly urban land and impervious landscape. Papers with less than 50% content and more than 50% content focusing on LUCC hold similar shares except for south NE, where a higher proportion of cases contain more than 50% content related to LUCC. Papers with less than 50% content for LUCC mainly focus on the research aspects highly related to LUCC, such as forest composition, nonnative species, and ecosystem. In addition, not all the study cases include the driving forces analysis, and nearly one fourth of all cases (24%) do not mention any explanatory factors of LUCC. At the same time, the study cases just including a simple analysis of driving forces of LUCC (e.g., some contain only several sentences about one or two explanatory factors) exist among the four location sections in the range of 25% to 50%. Only 43% of all the study cases involve a relatively comprehensive driver analysis. They do not vary to a considerable degree across the four location sections.

Table 3. Land cover types involved, proportion of land use/cover change (LUCC) analysis, and magnitude of drivers included in all the 54 cases.

1)	All cases (N = 54)		NE (N = 14)		Partial NE (N = 16)		South NE (N = 16)		North NE (N = 8)	
	No.	pct	No.	pct	No.	pct	No.	pct	No.	pct
Land cover type										
All types	30	55%	7	50%	8	50%	8	50%	7	88%
Forest-only	16	30%	7	50%	5	31%	3	19%	1	12%
Other	8	15%	0	0%	3	19%	5	31%	0	0%
Proportion of the LUCC study										
<50% content	26	48%	8	57%	7	44%	6	38%	4	50%
>50% content	28	52%	6	43%	9	56%	10	62%	4	50%
Driver included										
Comprehensive drivers	23	43%	8	57%	6	38%	6	38%	3	38%
Simple drivers	18	33%	5	36%	4	25%	5	31%	4	50%
No driver	13	24%	1	7%	6	37%	5	31%	1	12%

Among all the study cases, we selected the cases that include driver analysis and then summarized these 41 cases based on the driver study methods and the occurrence of specific underlying driving forces in these cases (Table 4). The driver study methods are generally grouped into two types: the descriptive methods and the quantitative methods. The papers with descriptive methods are the papers studying possible driving forces of LUCC based on their observations, reviewing related materials, or personal experiences, and the papers with quantitative methods are the papers quantifying the associations between LUCC (or land cover classes) and driving factors. The cases with descriptive methods (30 or 73%) are more common, about three times more than the cases with quantitative methods (11 or 27%), and the cases with descriptive methods show comparatively low location variations. The top three most commonly mentioned drivers are population, economics, and topography. The population here includes all its similar expressions, such as population density, population growth, and population change. The economics includes some vague words, such as economic activity and social-economic drivers in the descriptive cases, and some distinct words, such as per capita income and median household income in the quantitative cases. The topography mainly includes elevation and slope. Population is considered in 85% of all cases and it is the most pronounced driver in partial NE cases (90%). Economics is considered in slightly more than three fifths of all cases (63%). More cases that demonstrate economic drivers leading to LUCC are researches for NE and partial NE compared with the single state cases.

Topography is less reported than the former two drivers, accounting for 39% of all cases. It appears that cases with topography driven LUCC are the least widespread in North NE (14%).

Table 4. Driver study methods and occurrence of specific underlying driving forces in 41 cases.

2)	All cases including drivers (N = 41)		NE (N = 13)		Partial NE (N = 10)		South NE (N = 11)		North NE (N = 7)	
	No.	pct	No.	pct	No.	pct	No.	pct	No.	pct
Study Methods										
Descriptive methods	30	73%	11	85%	6	60%	8	73%	5	71%
Quantitative methods	11	27%	2	15%	4	40%	3	27%	2	29%
Specific drivers										
Population	35	85%	11	85%	9	90%	9	82%	6	86%
Economics	26	63%	10	77%	8	80%	5	45%	3	43%
Topography	16	39%	5	38%	4	40%	6	55%	1	14%

3.2. Results of the Random Forest Approach

The RF approach was employed to assess the driving forces of LUCC in terms of the importance of individual explanatory variables, that is, the predictive power for capturing the changes (Figure 4). The driver “Distance to developed area (DD)” was consistently ranked as the most important explanatory variable for all the three sample sites. Its contributions to LUCC were 20.44% for MetroWest, MA, 14.00% for SouthCentral, CT, and 23.61% for Northeastern, CT. In contrast, the driver “Protected area (PA)” was ranked as the least important variable for the two sample sites in CT with relative importance lower than 0.3%, and the driver “Aspect (AS)” was ranked as the least important variable (0.29%) for the MetroWest, MA. Obviously, the drivers of two sites in CT have more similar importance compared with those of the site in MA. Besides DD, distance to water (DW) and distance to roads (DR) also play an important role in explaining the land change in the two sites of CT. The location drivers in the two sites of CT explain more land change than social-economic drivers, such as house density change (HUSC) and population density change (POPC). In contrast, the social-economic factors, including HUSC, population density (POPL), and POPC, show strong relationship with land change in Metrowest, MA, with the percentages of variance explained by them being 15.78%, 15.28%, and 7.41%, respectively (Figure 5). Hence, the variation of driver importance among different locations does exist.

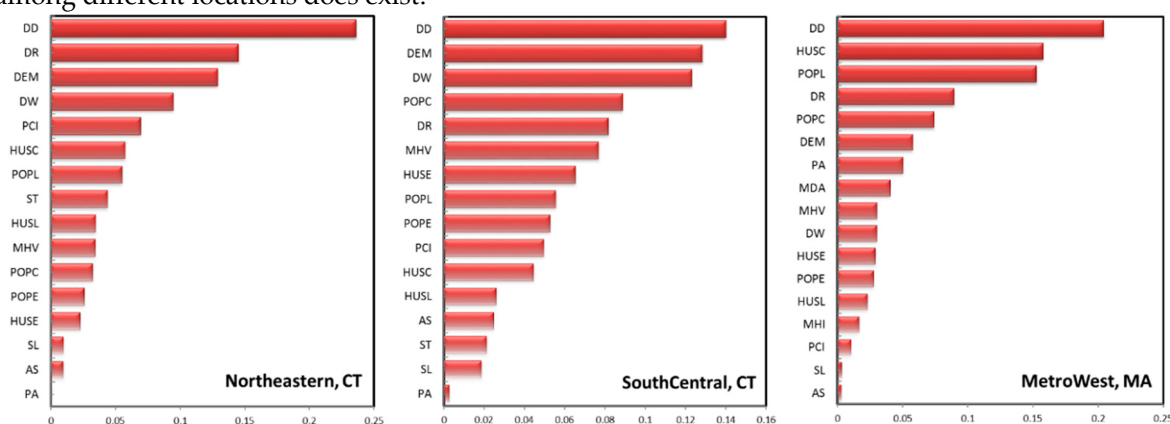


Figure 5. Results of the random forests (RF) approach for assessing the driving forces of LUCC.

The results show that the RF method may serve as an effective tool for evaluating the relationships between land changes and explanatory variables. The RF can predict whether there is a change of the land type at each pixel over this study period using different drivers that could affect

this change, such as population density and distance to the road. The model evaluation method, cross-validation, has been widely accepted to be able to offer a robust measure of predictive power [45]. In order to utilize all the data, 10-fold cross-validation was adopted here to evaluate the performance of the RF method. Using the 10-fold cross-validation method, we split all the data to ten folds evenly, and each time we trained the random forest model based on nine folds and tested the model performance against the rest one-fold. This process was repeated ten times such that every fold served as the rest one-fold once to test the model. The averaged evaluation metric was used as the indicator of the overall model performance. We calculated the prediction of land use change for the three research areas, Northeastern, Southcentral, and Metrowest, and the accuracies from the 10-fold cross-validation are 96.20%, 95.33%, and 95.10%, respectively. Hence, the random forest model, with all the available explanatory variables, can explain over 95% of the land changes for all of the three planning regions.

4. Discussion

4.1. Land Use/Cover Change Studies in New England Area From the Meta-Study

Understanding land use/cover change, which is not a new topic of concern, plays an important role in the study of global environmental change. Past works show that land use/cover change study is closely associated with multiple other research areas, such as urbanization, vegetation dynamics, nonnative species, and climate change in the New England area. LUCC study is not only an important research area for its own, but also serves as the foundation for multidisciplinary studies or the connection between multiple research topics. For example, a study assessed the water quality variability in Sebago Lake, Maine, with the climatic factors and land use change, and indicated an influential role of land use change in determining water quality [1]. Most of the past LUCC studies in the study area are multi-disciplinary studies, and nearly half of them (48% of all studies) have less than 50% content focusing on LUCC analysis based mainly on the LUCC information from pervious LUCC studies. Even the rest studies, which have more than 50% content for LUCC analysis, hardly have comprehensive LUCC analysis. In addition, many studies (30% of all studies) considered only forest land. This was determined by the characteristics of the research area where the major land class is forest. Therefore, a comprehensive LUCC study in the research area is still in demand and, moreover, it could also provide supportive information for related multi-disciplinary studies.

To better understand LUCC, an in-depth analysis of the relationships between the changes and the causes (the driving forces) is needed. The knowledge of the relationships between the changes and the causes may greatly improve modeling and projection of various kinds of LUCC and prevent the possible negative influence. Analysis of the driving forces for the research area is still limited. Nearly one fourth of the previous studies (24%) did not mention the drivers of LUCC and one third of all the studies (33%) simply discussed the drivers with only a few sentences. Even the remaining 43% of all the studies barely had a real comprehensive driving forces analysis. For those studies that mentioned the drivers of LUCC, the descriptive method (73%) was their major research technique. In the research area, an association is evident between LUCC and population growth over the long term. The associations between LUCC and economy and topography can also be found with variations. However, the association between LUCC and population growth is not a simple linear relationship. For example, urban sprawl and deforestation in the past several decades were directly associated with population growth, while they are recently more related to the low-density residential development. Yet these important contributions from the past studies only mapped the broad relationships between LUCC and its driving forces. Moreover, the studies that focused on LUCC and included a relatively comprehensive driver analysis only account for 30% of all the studies.

4.2. Consistency and Difference of the Meta-Study and the RF Method

RF explained over 95% of the LUCC for all three planning regions in our case study. Driver importance of LUCC in the three planning regions within the research area was evaluated and compared. In general, the RF model results show that LUCC is the aggregated effect of anthropogenic

and biophysical drivers, which is identical to the findings from the meta-study. On the other hand, the importance of anthropogenic and biophysical drivers in the three planning regions varies to some extent, which was hardly detected by the meta-study. The RF feature importance reported that socioeconomic factors including population density and house density were predominant determinants of LUCC in the relatively developed regions such as the SouthCentral and the MetroWest, and the biophysical factors were relatively more important in explaining LUCC in the Northeastern, CT. The consistent finding from both meta-study and the RF method is that the location factors, such as distance to developed area and distance to road, were important drivers for all the three planning regions. However, the location drivers were barely mentioned in the previous qualitative studies in this research area. In summary, both consistency and inconsistency between the findings from meta-study and the RF method exist.

4.3. Model Suitability and Reliability

Identification of the driving forces of LUCC is a challenging task, due to the complexity and nonlinearity of LUCC, diversity of drivers at different temporal and spatial scales, and interactions and feedbacks among drivers. A mixed-method approach may provide more comprehensive information. In our study, the meta-study provided an initial idea about the research content. The strength of the meta-study lies in its capability to pool results from individual studies for generalization of the driving forces of LUCC in the research area. However, related LUCC studies in the New England area were rare and at different temporal and spatial scales. Moreover, differences in the research focus and methods could lead to bias. However, the geospatial analysis in the meta-study may provide information about rates of LUCC and spatial distributions of drivers, which may be combined with the RF model results to quantify associations between LULC and driving factors. The quantitative method of the RF model may reduce the subjectivity weakness in the meta-study. In general, quantitative methods are often seen as more reliable than qualitative methods since objective quality criteria are provided by quantitative methods [46,47]. However, many qualitative spatial data as indirect drivers are needed for the quantitative study of LUCC. Therefore, a mixed-method approach that combines qualitative and quantitative analyses should benefit the research through including both direct and indirect driving forces, which may increase the probability of discovering important drivers of LUCC. In addition, it may complement the methodological shortcoming of one method through the advantage of another method [48]. Furthermore, the mixed-method approach allows comparison of different methods about the discordance and consistency of findings. In summary, information provided by multiple sources may be more powerful than that from the single source.

5. Conclusion

The study started with a meta-study to reveal the general information of LUCC in New England and then focused on the driving forces analysis. Through combining geospatial analysis results from the meta-study and the RF model, this mixed method provides a condensed analysis of driving forces of LUCC from both qualitative and quantitative data in the study area. Our study revealed that the population and economic factors are important drivers of LUCC in the research area, with a complex nonlinear relationship. The value of the mixed method is that it provides a more comprehensive identification of the drivers by complementing the findings of one method through the advantage of another method. For instance, in our study the RF model indicated the importance of the location factors, which was barely noticed in the meta-study. Hence, this mixed method may be meaningful to the LUCC study by enhancing understanding of the driving forces and providing the drivers' selection criteria for modeling the changes. Moreover, better land management advice for sustainable development could be made based on more solid information using the mixed method. However, the mixed method is not feasible for all studies. It highly depends on previous studies, which should have a decent amount to support a systematic review. On the other hand, to conduct a quantitative analysis of driving forces, the availability of important data, such as population and economic datasets, is a prerequisite. Moreover, the new approach aims to improve our ability to understand

the driving forces of LUCC, while linking findings from different aspects is always challenging. In future work, with the ongoing improvement of data, the mixed method should be able to better serve the LUCC studies. In addition, research on a more standard way to combine qualitative and quantitative methods is also necessary.

Supplementary Materials: The following are available online at www.mdpi.com/2220-9964/9/6/350/s1.

Author Contributions: Writing—original draft, Ruiting Zhai; Writing – review & editing, Ruiting Zhai, Chuanrong Zhang, Weidong Li and Xueke Li; Supervision, Chuanrong Zhang and Weidong Li; Methodology, Ruiting Zhai and Xiang Zhang. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Newbold, T.; Hudson, L.; Hill, S.L.L.; Contu, S.; Lysenko, I.; Senior, R.; Börger, L.; Bennett, D.J.; Choimes, A.; Collen, B.; et al. Global effects of land use on local terrestrial biodiversity. *Nature* **2015**, *520*, 45–50.
2. Biro, K.; Pradhan, B.; Buchroithner, M.; Makeschin, F. Land use/land cover change analysis and its impact on soil properties in the northern part of gadarif region, sudan. *Land Degrad. Dev.* **2011**, *24*, 90–102.
3. Clerici, N.; Paracchini, M.L.; Maes, J. Land-cover change dynamics and insights into ecosystem services in European stream riparian zones. *Ecohydrol. Hydrobiol.* **2014**, *14*, 107–120.
4. Rogger, M.; Agnoletti, M.; Alaoui, A.; Bathurst, J.C.; Bodner, G.; Borga, M.; Chaplot, V.; Gallart, F.; Glatzel, G.; Hall, J.; et al. Land use change impacts on floods at the catchment scale: Challenges and opportunities for future research. *Water Resour. Res.* **2017**, *53*, 5209–5219.
5. Lambin, E.F.; Turner, B.; Geist, H.J.; Agbola, S.B.; Angelsen, A.; Bruce, J.W.; Coomes, O.T.; Dirzo, R.; Fischer, G.; Folke, C.; et al. The causes of land-use and land-cover change: Moving beyond the myths. *Glob. Environ. Chang.* **2001**, *11*, 261–269.
6. Mallinis, G.; Koutsias, N.; Arianoutsou, M. Monitoring land use/land cover transformations from 1945 to 2007 in two peri-urban mountainous areas of Athens metropolitan area, Greece. *Sci. Total. Environ.* **2014**, *490*, 262–278.
7. Veldkamp, A.; Lambin, E.F. *Predicting Land-Use Change*; Elsevier: Amsterdam, The Netherlands, 2001.
8. Alexander, P.; Rounsevell, M.D.; Dislich, C.; Dodson, J.; Engström, K.; Moran, D. Drivers for global agricultural land use change: The nexus of diet, population, yield and bioenergy. *Glob. Environ. Chang.* **2015**, *35*, 138–147.
9. Campbell, D.J.; Lusch, D.P.; Smucker, T.A.; Wangui, E.E. Multiple Methods in the Study of Driving Forces of Land Use and Land Cover Change: A Case Study of SE Kajiado District, Kenya. *Hum. Ecol.* **2005**, *33*, 763–794.
10. Bajocco, S.; Ceccarelli, T.; Smiraglia, D.; Salvati, L.; Ricotta, C. Modeling the ecological niche of long-term land use changes: The role of biophysical factors. *Ecol. Indic.* **2016**, *60*, 231–236.
11. Patarasuk, R.; Binford, M.W. Longitudinal analysis of the road network development and land-cover change in Lop Buri province, Thailand, 1989–2006. *Appl. Geogr.* **2012**, *32*, 228–239.
12. Teixeira, Z.; Teixeira, H.; Marques, J.C. Systematic processes of land use/land cover change to identify relevant driving forces: Implications on water quality. *Sci. Total. Environ.* **2014**, *470*, 1320–1335.
13. Chen, L.; Ren, C.; Zhang, B.; Wang, Z.; Liu, M. Quantifying Urban Land Sprawl and its Driving Forces in Northeast China from 1990 to 2015. *Sustainability* **2018**, *10*, 188.
14. Kolb, M.; Mas, J.-F.; Galicia, L. Evaluating drivers of land-use change and transition potential models in a complex landscape in Southern Mexico. *Int. J. Geogr. Inf. Sci.* **2013**, *27*, 1804–1827.
15. Beilin, R.; Lindborg, R.; Stenseke, M.; Pereira, H.; Llausàs, A.; Slätmo, E.; Cerqueira, Y.; Navarro, L.M.; Rodrigues, P.; Reichelt, N.; et al. Analysing how drivers of agricultural land abandonment affect biodiversity and cultural landscapes using case studies from Scandinavia, Iberia and Oceania. *Land Use Policy* **2014**, *36*, 60–72.
16. Plieninger, T.; Draux, H.; Fagerholm, N.; Bieling, C.; Bürgi, M.; Kizos, T.; Kuemmerle, T.; Primdahl, J.; Verburg, P.H. The driving forces of landscape change in Europe: A systematic review of the evidence. *Land Use Policy* **2016**, *57*, 204–214.

17. Bürgi, M.; Hersperger, A.M.; Schneeberger, N. Driving forces of landscape change—Current and new directions. *Landsc. Ecol.* **2005**, *19*, 857–868.
18. Van Vliet, J.; Magliocca, N.R.; Büchner, B.; Cook, E.; Benayas, J.M.R.; Ellis, E.C.; Heinemann, A.; Keys, E.; Lee, T.M.; Liu, J.; et al. Meta-studies in land use science: Current coverage and prospects. *Ambio* **2016**, *45*, 15–28.
19. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
20. Ge, G.; Shi, Z.; Zhu, Y.; Yang, X.; Hao, Y. Land use/cover classification in an arid desert-oasis mosaic landscape of China using remote sensed imagery: Performance assessment of four machine learning algorithms. *Glob. Ecol. Conserv.* **2020**, *22*, e00971.
21. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J. Random Forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300.
22. Thompson, J.R.; Carpenter, D.N.; Cogbill, C.V.; Foster, D.R. Four Centuries of Change in Northeastern United States Forests. *PLoS ONE* **2013**, *8*, e72540.
23. Jeon, S.B.; Olofsson, P.; Woodcock, C.E. Land use change in New England: A reversal of the forest transition. *J. Land Use Sci.* **2013**, *9*, 105–130.
24. Zhai, R.; Zhang, C.; Li, W.; Boyer, M.A.; Hanink, D.M. Prediction of Land Use Change in Long Island Sound Watersheds Using Nighttime Light Data. *Land* **2016**, *5*, 44.
25. Zhai, R.; Zhang, C.; Allen, J.M.; Li, W.; Boyer, M.A.; Segerson, K.; Foote, K.E. Predicting land use/cover change in Long Island Sound Watersheds and its effect on invasive species: A case study for glossy buckthorn. *Ann. GIS* **2018**, *24*, 83–97.
26. Seto, K.C.; Fragkias, M.; Güneralp, B.; Reilly, M.K. A Meta-Analysis of Global Urban Land Expansion. *PLoS ONE* **2011**, *6*, e23777.
27. Van Asselen, S.; Verburg, P.H.; Vermaat, J.E.; Janse, J.H. Drivers of Wetland Conversion: A Global Meta-Analysis. *PLoS ONE* **2013**, *8*, e81292.
28. Geist, H.J.; Lambin, E.F. Proximate Causes and Underlying Driving Forces of Tropical Deforestation Tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. *BioScience* **2002**, *52*, 143–150.
29. Geist, H.J.; Lambin, E.F. Dynamic Causal Patterns of Desertification. *AIBS Bull.* **2004**, *54*, 817.
30. Keys, E.; McConnell, W.J. Global change and the intensification of agriculture in the tropics. *Glob. Environ. Chang.* **2005**, *15*, 320–337.
31. Wang, H.; Stephenson, S.R. Quantifying the impacts of climate change and land use/cover change on runoff in the lower Connecticut River Basin. *Hydrol. Process.* **2018**, *32*, 1301–1312.
32. Zhang, X.; Schaaf, C.B.; Gao, F.; Friedl, M.A.; Strakler, A.H.; Hodges, J.C. Mapping land cover and green vegetation abundance using MODIS-like data: A case study of New England. In Proceedings of the IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120), Honolulu, HI, USA, 24–28 July 2000; Volume 5, pp. 2005–2007.
33. Zhai, R.; Li, W.; Zhang, C.; Zhang, W.; Wang, W. The transiogram as a graphic metric for characterizing the spatial patterns of landscapes. *Landsc. Ecol.* **2018**, *34*, 2103–2121.
34. Dietterich, T.G. Ensemble Methods in Machine Learning. In *International Workshop on Multiple Classifier Systems*; Springer: Berlin/Heidelberg, Germany, 2000.
35. Chan, J.C.-W.; Paelinckx, D. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotone mapping using airborne hyperspectral imagery. *Remote Sens. Environ.* **2008**, *112*, 2999–3011.
36. Ghimire, B.; Rogan, J.; Miller, J. Contextual land-cover classification: Incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sens. Lett.* **2010**, *1*, 45–54.
37. Guo, L.; Chehata, N.; Mallet, C.; Boukir, S. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 56–66.
38. Latifi, H.; Nothdurft, A.; Koch, B. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: Application of multiple optical/LiDAR-derived predictors. *Forestry* **2010**, *83*, 395–407.

39. Martinuzzi, S.; Vierling, L.A.; Gould, W.; Falkowski, M.J.; Evans, J.; Hudak, A.T.; Vierling, K.T. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sens. Environ.* **2009**, *113*, 2533–2546.
40. Waske, B.; Braun, M.H. Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 450–457.
41. Wang, Q.; Ren, Q.; Liu, J. Identification and apportionment of the drivers of land use change on a regional scale: Unbiased recursive partitioning-based stochastic model application. *Agric. Ecosyst. Environ.* **2016**, *217*, 99–110.
42. Qasim, M.; Hubacek, K.; Termansen, M. Underlying and proximate driving causes of land use change in district Swat, Pakistan. *Land Use Policy* **2013**, *34*, 146–157.
43. Van Vliet, J.; De Groot, H.L.; Rietveld, P.; Verburg, P.H. Manifestations and underlying drivers of agricultural land use change in Europe. *Landsc. Urban Plan.* **2015**, *133*, 24–36.
44. Lilieholm, R.J.; Meyer, S.; Johnson, M.L.; Cronan, C.S. Land Conservation in the Northeastern United States: An Assessment of Historic Trends and Current Conditions. *Environ. Sci. Policy Sustain. Dev.* **2013**, *55*, 3–14.
45. Helmreich, J.E. *Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2015.
46. Amaratunga, D.; Baldry, D.; Sarshar, M.; Newton, R. Quantitative and qualitative research in the built environment: Application of “mixed” research approach. *Work. Study* **2002**, *51*, 17–31.
47. Malina, M.A.; Nørreklit, H.S.; Selto, F.H. Lessons learned: Advantages and disadvantages of mixed method research. *Qual. Res. Account. Manag.* **2011**, *8*, 59–71.
48. Brewer, J.; Hunter, A. *Multimethod Research: A Synthesis of Styles*; Sage Publications Inc.: New York, NY, USA, 1989.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).