Open Access

Advances in Environmental and Engineering Research

Research Article

Contributing Factors to Forest Loss in Conterminous U.S. for the 1990s and 2000s

Giorgos Mountrakis *, Sheng Yang

State University of New York College of Environmental Science and Forestry, 1 Forestry Dr., Syracuse, NY, 130210, USA; E-Mails: gmountrakis@esf.edu; syang16@syr.edu

* **Correspondence:** Giorgos Mountrakis; E-Mail: gmountrakis@esf.edu

Academic Editor: Zed Rengel

Adv Environ Eng Res
2021, volume 2, issue 4
doi:10.21926/aeer.2104026

Received: March 24, 2021
Accepted: October 13, 2021
Published: October 26, 2021

Abstract

While numerous studies have considered forest loss factors at local scales, there is a gap of comparative quantitative regional modeling at the U.S. national level. Here, we investigated statistical relationships between gross forest cover loss (GFCL) and numerous socioeconomic, biophysical and ownership variables between two decades, the 1990s and the 2000s. A spatial error model was employed to compensate for spatial autocorrelation effects. Models from the 2000s had stronger explanatory power than the 1990s models, especially in the Northeast and the South (R^2 of 0.89 and 0.87 respectively). The amount of forested areas in low slopes was a highly influential factor for high GFCL, followed by urban area cover and mill density. On the other hand, agricultural cover was negatively correlated with GFCL acting as a stabilizing factor in the South and Midwest regions. Our study offers an important insight in regional drivers of GFCL, drivers that should be further examined in the local context to gather better understanding of their contributions.

Keywords

Deforestation; underlying factors; proximate factors; spatial error model; Continental U.S.

1. Introduction

Numerous studies have examined deforestation drivers from local to global scales. These drivers were usually grouped in two categories, namely proximate causes and underlying causes [1]. Proximate causes were defined as human activities or immediate actions at the local level, such as agricultural expansion, that originated from planned land use and directly influence forest cover. Underlying causes represent fundamental social processes, including human population changes and agricultural policies that motivate the proximate causes and either operate at the local level or have an indirect impact from the national or global level.

Internationally, deforestation has been linked to demographic (i.e., population density) and accessibility (i.e., road proximity) variables [2], agricultural expansion [3], poverty [4, 5], illegal logging [6] (Honey-Rosés 2009 in Mexico) and other biophysical variables including altitude, distance to crops, and distance to rivers [7]. In another study, among 41 tropical countries between 2000 and 2005, urban population growth and exports of agricultural products were recognized as significant drivers of forest loss [8]. Education levels and forest land ownership were also found to have impact on deforestation in 128 countries with larger than 50,000 km2 forest [9]. In two other studies population, education levels, and forest ownership (state owned or privately owned), were found to influence deforestation in Africa and Eastern Europe [10, 11]. National-level data from 100 tropical countries were collected and two proximate drivers, commercial agriculture and subsistence agriculture, were identified as significant drivers of deforestation, whereas timber extraction and logging were recognized as proximate drivers of forest degradation [12].

Looking at the U.S. specifically, recent studies have shown pervasive deforestation across the conterminous U.S. [13-16]. In addition, population and income in the U.S. are projected to increase in the first half of twenty first century, resulting in increased residential area and deforestation [17]. Mechanical disturbance, (e.g. logging), non-mechanical disturbance, (e.g., wildfire), and conversion to developed and agriculture lands were the main sources of these forest land losses nationally in U.S. and in several major level II ecoregions in the 1973-2000 period [13, 16]. In another U.S. Geological Survey study, forest land cover was found to be the source of 40% or more of newly-developed urban lands in 35 of the 84 level III ecoregions in the continental U.S. between 1973 and 2011 [15].

Regional studies exist within the U.S. examining forest cover changes and driving forces in ecoregions [16, 18-20]. Studies on deforestation and fragmentation at local scales found that income, urbanization, slope and population were the correlated factors in parts of metropolitan areas in Georgia and California as well as in mountain areas in Arkansas [21-23]. However, only correlation coefficients or descriptive information were reported without a quantitative model showing the influence and/or significance of variables. To date, there have been no studies that examine the driving factors of deforestation at the national level in the U.S. with regional models that can guide understanding and management of deforestation at finer scales, for example the county level. A national level perspective is necessary to put regional scale drivers within a national context and to identify drivers possibly originating outside the study region as the result of national socio-economic events [18]. Fine study units (e.g. counties) are critical as that is one of the scales where decisions related to deforestation frequently operate in additional to national forest policy and private landowner decisions, and on county levels most socio-economic datasets can be derived [24].

To close this critical knowledge gap, we examined deforestation in the conterminous 48 states in the U.S. and driving factors of both direct causes (e.g., logging), and of indirect or underlying causes (e.g., population dynamics). We built regional models with drivers of deforestation to examine which factors are universally influential across the U.S. and which show a more regional effect. Our overarching goal is to use the modeling results to guide further regional research and management decisions.

2. Materials and Methods

2.1 Study Area

The study was conducted in the conterminous U.S. at the county level and investigated two decades, the 1990s and the 2000s. From the original 3,109 counties a subset was selected. Firstly, 23 counties were removed as their boundaries have changed between the two decades. Secondly, a minimum threshold of 10% forest cover was established for each county, which led to the final total of 2,132 counties. This threshold was selected to allow our models to concentrate on forested counties and reduce the variable sensitivity as counties with small total forest could be highly sensitive to minor changes. Figure 1 presents the selected counties along with forest cover percentage. In the following sections we discuss the response variable (forest cover), the explanatory variables (a group of socioeconomic, biophysical and ownership variables) and the statistical models to investigate potential relationships.

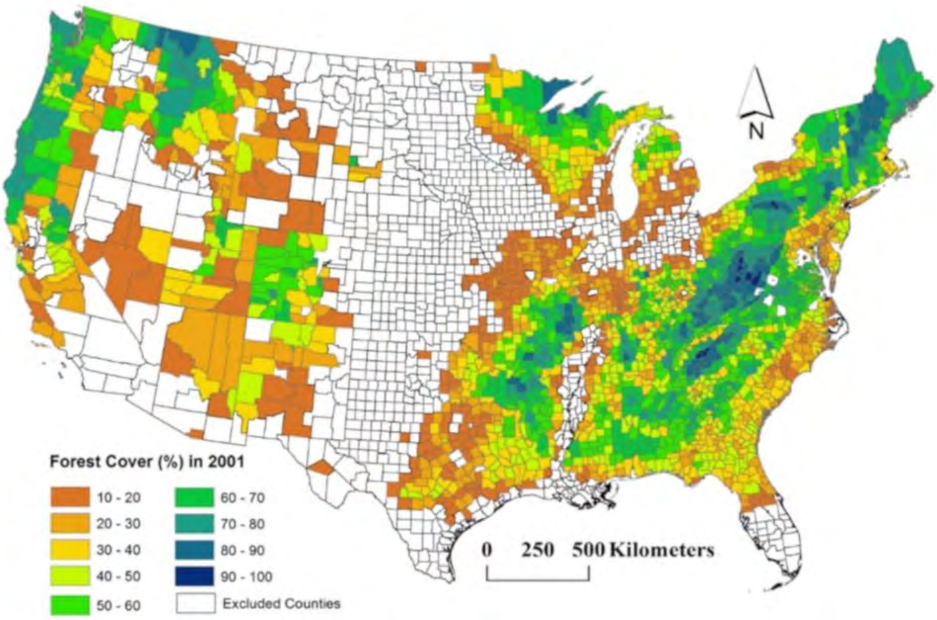

Figure 1 Percentage forest cover in 2001. The data was from NCLD change product and were aggregated in counties [25].

2.2 Response Variable

Forest cover data were obtained from the National Land Cover Database (NLCD) change product of 1992-2001 and the 2011 NLCD product. Both NLCD products are satellite-derived and provide wall-to-wall forest cover datasets over the entire continental U.S. The attributed accuracy for both products is approximately 80% for the forest cover changes in the change product of 1992-2001 [26] and similarly near 80% in the 2011 NLCD for the forest loss [27]. While these accuracies are not ideal for our study, they are not a cause of concern as our analysis was conducted at the county level. This aggregation of 30m cell data, instead of a per pixel comparison, provides higher confidence in the data accuracy. Furthermore, our comparison was done over time for the same counties so local classification errors would be present in both temporal datasets of the same county thus not introducing a considerable bias.

The NLCD mapping products contain multiple classes pertaining to forest cover. For the binary forest map creation the classes of deciduous forest, evergreen forest, and mixed forest were merged into a single forest class. Gross forest cover loss (GFCL), defined as the total amount of forest losses in a county without accounting for forest gains, was aggregated at the county level using zonal statistics. An important clarification for our study is that GFCL contains all forest class transitions, for example it includes changes between the forest and the shrub/grassland class changes. Therefore some of the observed changes may not be permanent or long lasting.

For modeling purposes the square root transformation of the GFCL percentage was applied for both 1990s and 2000s, to ensure normal distribution of the response variables, an essential condition for our subsequent regression analysis. The square root transformation was selected over a range of transformations to normalize the response variable distribution that was skewed to the right (larger values).

2.3 Explanatory Variables

The pool of predictor variables was comprised of four general groups: socioeconomic, ownership, biophysical, and spatial neighborhood variables (Supplementary Table 1). The socioeconomic variables were generated from Census 1990, 2000, and 2010, including population and other related demographic attributes, all aggregated at the county level. These variables represent underlying factors that may drive forest cover loss processes in indirect ways. Ownership variables and biophysical variables were generated by integrating forest land cover data with national datasets on land ownership and biophysical attributes. Ownership variables represent underlying driving factors of forest cover loss, while biophysical variables represent proximate factors which have direct impact on forest cover loss. The spatial neighborhood variables are derived from the first three groups to account for spatial autocorrelation effects among observations in counties.

2.3.1 Population

Population growth has been shown to positively correlate with forest cover loss in tropical countries [8] and it can potentially impact forest cover loss in the U.S. [17]. Decadal forest change 1992-2001 at the state level in the U.S. was positively related to population growth and population density [28]. Population change at the local level was also related to forest cover loss near the Atlanta metropolitan areas in Georgia [21]. In our study, U.S. Census data in 1990, 2000, and 2010 at the county level were processed and utilized for extraction of population data using the National Historical Geographic Information System (NHGIS) database of the Minnesota Population Center (www.nhgis.org). Percentage population changes were also computed for the two decades.

2.3.2 Income and Education

Education and income are key demographic factors that were shown to correlate with forest cover changes in prior studies. For example, the human development index representing education, income and health was shown to be related to forest cover loss in tropical forests [29]. Income increase was the driving factor of forest cover loss in several parts of the U.S. [17, 23]. Education levels have been shown to have impacts on forest cover loss in many parts of the world [10, 11].

Variables related to education and income were extracted from Census data in 1990, 2000, and 2010 from NHGIS. For educational attainment, the percentage of people with high school diploma or higher was calculated by aggregating categories including 'High School Graduate', 'Some college, No degree', 'Associate Degree', 'Bachelor's Degree', 'Master's Degree', 'Professional School Degree', and 'Doctorate Degree'. Furthermore, poverty was often found to be related to forest cover loss in various studies [5]. Poverty status was calculated as the percentage of households below the poverty level in a county. Per capita income was also attained from the census database without further modification. Education attainment, poverty levels, and per capita income were based on responses from a sampled population in the original Census data collection. Although the sampled values are statistical estimations for the actual values, the true value falls between the upper and lower bounds of estimated value with statistical significance at the 90% confidence level [30]. In addition to the static values, percentage change of income, poverty, and education within each decade was computed.

2.3.3 Ownership and Protected Lands

Numerous studies have examined and shown influence of land ownership and protection status on the forest cover loss process [31]. The direction and magnitude of these relations may vary in different countries. Forest land ownership is potentially related to forest cover loss in the U.S. [32]. Protected lands have been recognized to help prevent forest cover loss, thus exhibiting negative relation with forest losses [28]. The protected forests were included based on the protected land dataset Protected Area Database of the United States, PAD-US (CBI Edition) Version 2 from which GAP Status 1-3 were included to extract protected forests. In addition, major ownership classes, (e.g., federal, state, and private lands), were also extracted from this database to differentiate forests from various ownership types. The vectors of different ownership types and protection status were overlaid with forest maps and the amount of forest lands falling within the boundaries of each category was aggregated in every county by applying zonal statistics. Finally, the percentage of forest cover in each category out of total forest cover in the county was computed as input variables of statistical analysis.

2.3.4 Road Network

Road networks and distance to road have been recognized as contributing factors of forest cover loss [21]. Proximity to road is expected to positively relate to forest cover loss due to easier harvesting and human development. Road networks data were extracted from U.S. Census Bureau's TIGER Primary and Secondary Road data available in Geospatial Data Gateway data portal of the U.S. Department of Agriculture Natural Resources Conservation Service (USDA NRCS) to generate distance to road variables. Buffers of 1km and 5km from road networks were generated to represent areas under different road influence. Forest maps were overlaid with road network buffers and the amount of forest within the buffer zones was aggregated in counties by computing zonal statistics. The percentage of forest in each category of 1km and 5km buffers out of total forest cover in each county was then calculated as input variable.

2.3.5 Topography

Elevation and slope have been shown to be critical factors related to forest cover loss in many studies in the U.S. and other parts of the world [33]. Existing forests in low slope areas in the beginning of the decades were more prone to forest cover loss due to higher possibility of residential development and less constraint for harvesting. Slope variables were derived from the National Elevation Dataset (NED) which is the primary elevation data product of the USGS (U.S. Geological Survey). The NED is a wall-to-wall dataset with the best available raster elevation data of the conterminous United States. For consistent cover across the continental U.S. and to match our NLCD data the 30m resolution product was used. It was necessary to integrate slope variables extracted from raster datasets with socioeconomic data from the decennial census that was in vector format. A slope data layer was generated from the elevation raster data layer. Slope values were grouped in two datasets: one binary dataset dividing the U.S. lands above and below a 5% threshold and another binary dataset dividing U.S. lands above or below a 10% threshold. Forest maps were overlaid with these categorical slope data layers and the amount of forest in each category was aggregated in counties by applying zonal statistics. Finally, the percentage of forest cover in each category out of total forest cover in the county was generated as final inputs for statistical modeling.

2.3.6 Urban Clusters

Urbanization and urban population growth as well as proximity to residential areas have been recognized as major factors of forest cover loss in tropical areas and in the U.S. [8]. Forest cover loss in the U.S. is closely related to urban development and subsequent land use changes [16, 17] with over 40% of urban development coming from forest cover changes in many U.S. regions [15]. Urban areas data were processed based on the definition and boundary shape files of urban clusters from U.S. Census Bureau's TIGER files. Urban clusters of larger than 500ha were selected and identified as another category to represent influences of major metropolitan areas. Buffers of 5km and 10km from urban clusters were generated to represent different pressure magnitudes of urban areas. Forest maps were overlaid with urban clusters, and the percent of forest in each category out of total forest cover in the county was aggregated in counties using zonal statistics.

2.3.7 Agriculture

The pressure of agriculture land use on forests have been recognized in many studies in U.S. (Drummond & Loveland 2010b; Sleeter et al. 2013; Napton et al. 2010; Auch et al. 2012). Agriculture land cover datasets were extracted from the NLCD change product, and the amounts of agriculture land were generated for each county by applying zonal statistics. Then, the percentage of agriculture land in the county was generated as a predictor variable.

2.3.8 Mill Density

Timber cutting and harvesting has been recognized as drivers of forest cover loss in many parts of U.S. [16, 20]. In order to incorporate the influence from mechanical disturbance, or harvesting, mill locations in the continental U.S. were extracted from shapefiles that represent information collected by mill data managers in Forest Inventory and Analysis (FIA) Units as well as collaborators in the Texas Forest Service, the Forest Products Laboratory (FPL), and the Focused Science Delivery Program of the Pacific Northwest Research Station. The best available data of national coverage were from the 2005 dataset, which is a point vector dataset representing mills across the U.S. The numbers of mills were aggregated for each county. To have a quantitative estimate of wood extraction effects in each county and from neighboring counties, the kernel densities of pulp mill location points were computed. Kernel density expresses in a distance-weighted fashion the total number of mills that fall within the search area of a kernel. A raster map of 500m cell size was generated showing kernel density. For each 500m cell, the density of mills was computed in 200 km radius neighborhood distances using a Gaussian distance-weighted function and the values were aggregated at the county level by zonal statistics. The entire process was conducted in ESRI's ArcGIS software.

2.3.9 Climate

Regional climatic variability may be influencing GFCL dynamics. To examine this influence, mean precipitation and temperature were derived from averaging precipitation and temperature over the decades of 1990s and 2000s based on monthly precipitation and temperature data from PRISM (http://www.prism.oregonstate.edu/). The Standardized Precipitation Evapotranspiration Index (SPEI) was estimated from SPEIbase v2.4 (http://spei.csic.es/database.html) which contains monthly data of SPEI over the period 1901-2014. The monthly SPEI data were extracted and averaged over the decades of 1990s and 2000s using the National Oceanic Atmospheric Administration (NOAA) Weather and Climate Toolkit. Finally, mean precipitation, mean temperature, and mean SPEI variables were aggregated in counties by zonal statistics.

2.3.10 Neighboring Effects

One important assumption of regression analyses is the independence among variable observations, which is often compromised by correlations between observations. The observations in our study units, i.e., counties, can be affected by Tobler's first law of Geography in that counties nearby are more related than counties further away. To incorporate and address this spatial autocorrelation among neighboring counties, for all the aforementioned variables, another process of neighborhood averaging was taken to reflect the influences of neighboring counties on a given focal county except for mill density which was processed as kernel density over neighboring areas. The new value for a focal county is a weighted average in 100km, 300km, and 500km neighborhood with decreasing weights following a Gaussian function curve within the neighborhood. The analysis used ESRI's ArcGIS software.

2.4 Statistical Modeling

The statistical modeling aimed at identifying the contributing factors for GFCL. Our approach followed two general steps, first variable selection followed by the spatial error model. To select the optimal set of correlated variables, we started with all variables and applied a stepwise regression using ordinary least squares. During this process predictor variables were added and removed one at a time based on minimum entry and removal criteria. Initially, the Bayesian Information Criterion (BIC) was used for performance evaluation and the model with lowest BIC was selected. We further eliminated variables with low statistical significance, that is P-value larger than 0.05. The process of stepwise regression generated a model of the best goodness of fit with the tradeoff between model complexity and accuracy taken into account by having lowest BIC. However, the predictor variables in the resulting model may exhibit high multicollinearity thus making individual variable influence difficult to isolate. To address multicollinearity selected variables were further filtered. Among related predictor variables with response variable (correlation coefficient > 0.1), predictors with the highest correlation with other predictors and lowest estimated coefficients were removed.

Spatial autocorrelation of regression residuals can lead to violation of regression assumptions. We applied a Lagrange Multiplier test to examine if significant spatial autocorrelation existed. A spatial error model (SEM) was implemented, shown in the following equation:

$$Y = X\beta + \lambda W\varepsilon + \xi \tag{1}$$

Where λ is a spatial error parameter and $W\varepsilon$ is a spatially lagged error term. Y is a vector representing the response variable of and X represents the predictor variables, β represents a vector of coefficients of intercept and predictor variables. Finally ε represents the model error with constant mean and variance.

The weight matrix W was generated using the 40km neighborhood distance. This distance was selected as it produced the lowest Moran's I value in 1990s and 2000s residuals. In this case ξ can meet the assumption of independent well-behaved error term if there is spatial autocorrelation in the original error term ε. The spatial error parameter λ is often considered as a nuisance parameter with little interest per se, and λ is necessary to correct for the spatial autocorrelation [34]. The mean of Y is not affected by the spatial error dependence. After correcting spatial autocorrelation by SEM, a Moran's I test was applied to examine if spatial autocorrelation was removed from model residuals.

Finally, in order to capture regional variability, separate models were developed for the East, Midwest, South, and West U.S. regions (see Supplementary Figure 1 for region boundaries). These major regions are defined by federal laws and are widely used for data collection and analysis. This regional segmentation offers a good balance between capturing local dynamics/drivers and having sufficient counties to develop our models and offer reliable results.

3. Results and Discussion

3.1 Gross Forest Cover Loss Spatial and Frequency Distribution

Figure 2 shows the distribution of gross forest cover loss (GFCL) in the two decades. The average GFCL in the selected 2132 counties in 1990s was 5.7%, whereas in 2000s it was 9.9%. Looking further into this GFCL magnitude change between the two decades, only 6 counties experienced GFCL larger than 30% in the 1990s, whereas the same number was 200 counties in the 2000s. This indicates that GFCL, as a percent for each county, has experienced a higher portion of extreme losses in the 2000s.

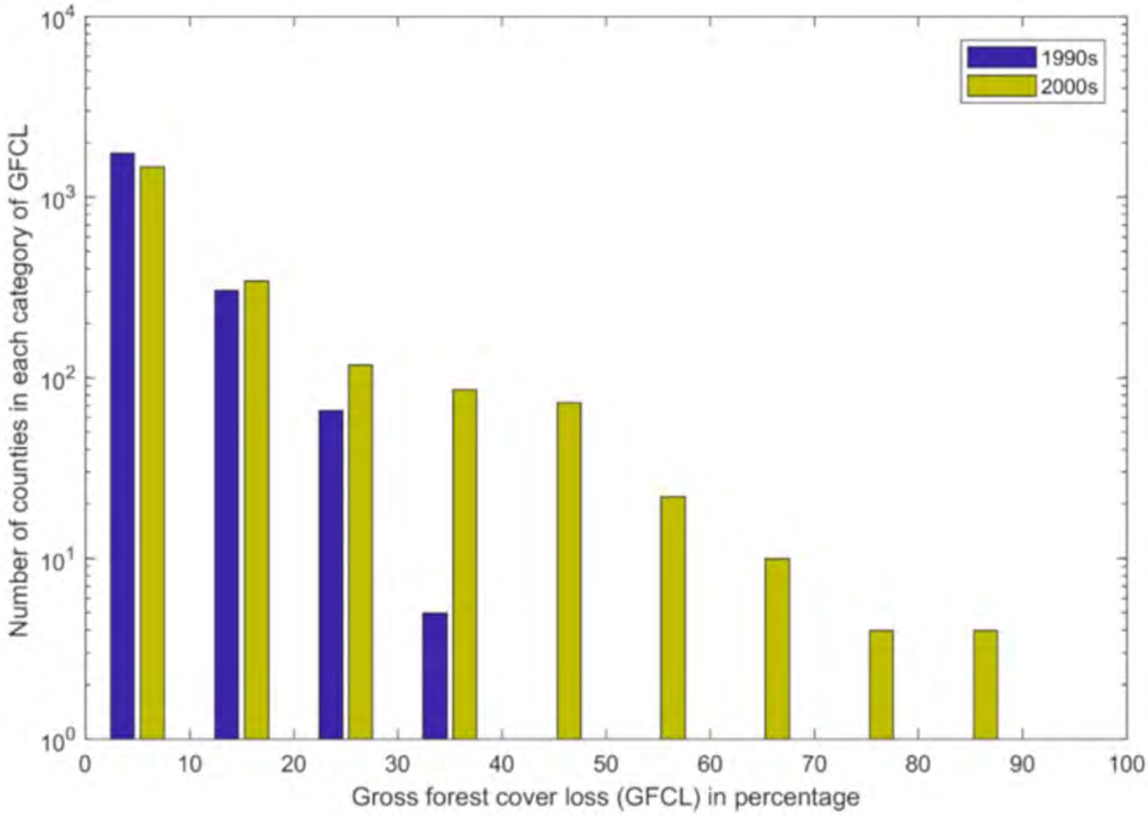

Figure 2 County frequency distribution (log of frequency) across gross forest cover loss groups (%). Percentage losses are relative to forest cover at the start of each corresponding decade.

A map of the gross forest cover loss (GFCL) response variable, the percentage of GFCL, is shown in the two maps from the 1990s and the 2000s in Figure 3. It is clear that the geographic distribution of higher relative forest losses (i.e. loss percentages) showed major differences between 1990s and 2000s. The higher percentage of forest losses during the 1990s were in the southeast coast from North Carolina to Texas. In contrast, GFCL in the 2000s shows more intensified GFCL percentages in the eastern coast, now extending further north to New York City. A hot spot is also identified in Minnesota. It should be emphasized that these maps depict relative, not absolute, forest loss. Therefore, comparisons can be made within each county across time, however regional interpretations should consider the fact that forest cover varies significantly across counties (see Figure 1) therefore absolute losses may be distributed differently.

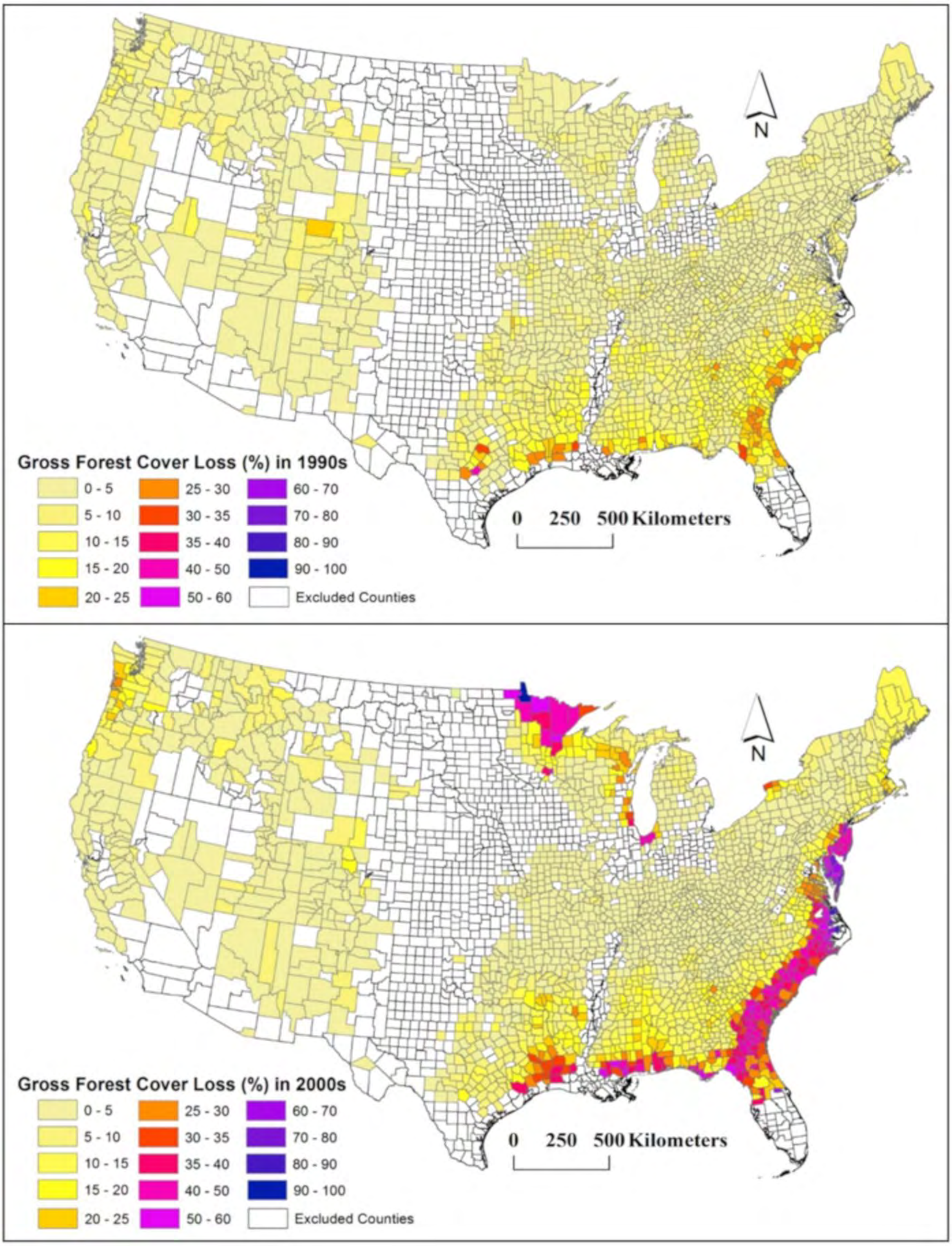

Figure 3 Gross forest cover loss during 1990s and 2000s.

3.2 Drivers Associated with Gross Forest Cover Loss

By applying an Ordinary Least Squares (OLS) stepwise linear regression, predictor variables were found that contribute significantly to forest loss in 1990s and 2000s. In order to assess multicollinearity issues typical metrics were calculated (see Table S2 and Table S3 in supplementary material). Correlation coefficients of all pairs of predictor variables were less than 0.26, indicating low correlation among them. The Variance Inflation Factor (VIF) of all variables was considerably less than 5, a typical threshold for high correlation identification. Likewise, the conditional index for all variables were less than 30 (maximum value was 1.84), which suggests absence of multicollinearity [8].

In the original OLS implementation (non-spatial regression) spatial autocorrelation was present. The Moran's I of the residuals for the national 1990s model was 0.116 (p-value = 0.001) and for the 2000s was 0.165 (p-value = 0.001). Furthermore, the LaGrange multiplier test showed spatial autocorrelation in both the 1990s and the 2000s model residuals with p-value < 0.05. Therefore, the Spatial Error Regression (SER) was selected as it can effectively address spatial autocorrelation in the residuals of the multiple linear regression to avoid violating the assumption of independent residuals in regression modeling.

Using SER as the underlying modeling methodology separate models were developed for each decade and region. The results are presented in table 1. The first observation is the considerably higher variability captured by the 2000s vs. the 1990s models. This could be partially attributed to the increasing dominant influences from harvesting [35] in the eastern U.S. over other disturbance agent classes such as fire, decline, wind, flooding which were more influential in the 1990s and were not fully captured by the 1990s model. This could also be a contributor to the lower performance of our models in the West. The increasing influence of logging in the eastern U.S. in 2000s was also found in regional study of GFCL [36]. The dominant GFCL in eastern U.S. in 2000s was also shown in another global study of GFCL [37]. Below we examine the contributions of each predictor variable.

Table 1 Regional spatial error model results. Empty cells and * indicate coefficients/ R^2 with p-value >0.05. All coefficients were standardized. Both precipitation and temperature variables were significant in South 1990s and Midwest 2000s results but only the higher coefficient variable from the two was selected due to high correlation.

	Northeast		**South**		**Midwest**		**West**	
Decade	**1990s**	**2000s**	**1990s**	**2000s**	**1990s**	**2000s**	**1990s**	**2000s**
Variable	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**
Constant	1.30	3.32	2.1	2.41	1.76	1.52	2.05	2.34
Slope<5% forest (%)	0.33	0.97	0.39	0.93	0.17	0.69	0.36	
Urban (>500ha) 10 km Buffer Forest (%)	0.23	0.18	0.18	0.08	0.06		0.16	
Agriculture (%)			-0.22	-0.34	-0.12	-0.38		
Mill Density (count/km^2)	0.11	0.17	0.2	0.22	0.11			
Mean Precipitation (in.)	-0.28			0.43		-0.72		0.12
Mean Temperature (F)		0.72	0.48		0.52			
Spatial Error λ	0.42	0.75	0.52	0.76	0.58	0.61	0.47	0.62
R^2	0.52	0.89	0.66	0.87	0.44	0.71	0.26*	0.5*

3.2.1 Forest in Low Sloped Area

The most influential was the percent of forest at a 5% or less slope (SLP5). The lower slope forested lands represent fewer constraints for harvesting and development. From 1973 to 2000 in the U.S. nationally, conversion of forests because of mechanical disturbance (i.e., logging) accounted for the largest gross forest loss [13]. Another major type of transition from forests was development during 1973-2000 [13]. For example, in the southeast Piedmont ecoregion, a major forested ecoregion in the east where forest cover is dominant, forest cover loss due to mechanically disturbance of forest was the major type of conversion between 1973 and 2000 [36]. This relation is comparable to the results in a study in Atlanta, Georgia [21] where forestland area change from 1973 to 1999 positively correlated with mean slope gradient at the county level. At census tract level, forestland area change from 1973 to 1999 correlated negatively with mean slope gradient [21]. It should be noted that our modeling results focused on forest cover loss compared with previous studies on net forest cover changes with mixed effects from forest gains.

In addition, the SLP5 coefficients increased considerably from 1990s to 2000s. This potentially reflects increasing pressure on GFCL from harvesting and wood extraction in flat areas, as steep slopes were not as cost-effective to harvest compared to flatter lands [38]. Logging was found to be the primary cause of GFCL in U.S., especially in southeast U.S., during between 2000 and 2005 [37].

3.2.2 Urbanization

The relatively high impact of urbanization was reflected with high coefficients in the Northeast and South models. Nationally in U.S., conversion of forests to developed lands is a major type of land cover change from 1973 to 2000 [13]. In the southeast Piedmont ecoregion with forests as dominant land cover, forest to urban development were the major types of conversions between 1973 and 2000 [36]. On the other hand, forests were also found to be a preferred land cover for urban expansion across U.S. 1973-2011 [15] and in the southeast [16] between 1992 and 2001.

3.2.3 Agriculture

Agriculture showed significant relationship with GFCL in the South and Midwest models acting as an important stabilizing factor. This was initially a surprising finding considering a wide range of studies showing a major transition from forests to agriculture between 1973 and 2000 across the U.S. [13, 18]. Furthermore, internationally agriculture is typically seen as a major deforestation contributor, primarily driven by potential financial benefits. Within the U.S. there has been substantial agricultural intensification, thus reducing overall land conversion pressures. In addition, farming communities are appreciative of land benefits and land sustainability, thus resisting to considerable forest loss.

3.2.4 Mill Density

Mill density, used as a proxy for harvesting intensity, also correlated positively to GFCL, especially for the Northeast and South in 2000s. Commercial forestry is among the most common drivers of land cover changes in southeast U.S. between 1973 and 2000 [18]. Specifically, conversion of forests caused by mechanical disturbance (i.e., logging) accounted for the largest gross forest loss from

1973 to 2000 in the U.S. nationally [13]. The increasing coefficients of mill density possibly reflect the increasing commerce of forest products, forest harvesting, wood extraction in the 2000s. For example, in a highly-forested level III ecoregion Piedmont, forest to mechanically disturbed increased in both the period of 2000 to 2006 and the period of 2006 to 2011, compared with 1990 [36].

3.2.5 Climate

While both precipitation and temperature variables were tested during model development, only one of them was further processed due to high correlation. Mean temperature was significantly correlated with GFCL in Northeast for 2000s and South and Midwest for the 1990s. This correlation could be attributed to temperature effects on tree mortality. Mean precipitation varied between positive and negative contribution to GFCL. The negative sign could be attributed to better than expected precipitation addressing drought issues [39] or extreme precipitation causing flooding and therefore negatively affecting forests [40].

Finally, the spatial parameter lambda was significant with low p-value, suggesting considerable spatial autocorrelation in the error structure of the regression model. The lambda parameter affects the modeling results by correcting the biasing influence of the spatial autocorrelation in the error term. It is treated as nuisance parameter meaning that statistically it is not of interest per se but must be accounted for in the analysis of those parameters that are of interest [34].

4. Conclusions

In this study, we investigated an extensive collection of variables from three categories, namely socioeconomic, ownership, and biophysical variables that potentially relate to GFCL. We found that the amount of forested areas in low slopes was positively correlated to GFCL. This suggests that site access and deforestation cost may be playing an important role in forest cutting selection. Urban area cover was also a contributing factor to forest loss dynamics, a well-documented result of urban sprawl. From the forest harvesting perspective, mill density was also correlated with high forest loss for the Northeast and South regions. This depicts the harvesting pressures which while not as high as other drivers they should be considered. Climatic factors also played a role but it is difficult to provide concrete results as they may partially capturing localization and other drivers not examined in this study (e.g. policy and management). Strong agricultural presence was negatively correlated with GFCL in the South and Midwest regions, statistically quantifying for the first time the importance of agriculture in controlling forest losses at the regional level. This could be a direct effect of the Conservation Reserve Program in 2000s that promoted forest stability in counties with a higher percent of agriculture lands.

For future work, it would be interesting to incorporate additional variables capturing management and decision making. Also expanding the study to include the 2010-2020 decade would offer additional insights. With the proliferation of satellite data, forest change detection is not the limiting factor for expanding this analysis globally, rather the availability of well documented driver data becomes a primary concern.

Additional Materials

The following additional materials are uploaded at the page of this paper.

1. Table S1: List of considered predictor variables. Exp. Corr. represents expected correlation. Location is from the prior studies of forest cover loss and correlation.

2. Figure S1: Boundaries of major regions in the U.S. from Census Topologically Integrated Geographic Encoding and Referencing (TIGER) data.

3. Table S2: Multicollinearity analysis among predictor variables during 1990s. Variance inflation factor (VIF), conditional index, and correlation coefficients were computed for each predictor variable and all pairs of variables.

4. Table S3: Multicollinearity analysis among predictor variables during 2000s. Variance inflation factor (VIF), conditional index, and correlation coefficients were computed for each predictor variable and all pairs of variables.

Author Contributions

Mountrakis designed the experiment and wrote the manuscript. Yang conducted the experiment and produced the figures.

Funding

This work was supported by the National Urban and Community Forestry Advisory Council, U.S. Forest Service and a SUNY ESF Graduate Assistantship.

Competing Interests

The authors have declared that no competing interests exist.

References

1. Geist HJ, Lambin EF. 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.
2. Mena CF, Bilsborrow RE, McClain ME. Socioeconomic drivers of deforestation in the Northern Ecuadorian Amazon. Environ Manage. 2006; 37: 802-815.
3. Waiswa D, Stern MJ, Prisley SP. Drivers of deforestation in the Lake Victoria Crescent, Uganda. J Sustain For. 2015; 34: 259-275.
4. Hübler M. How to curb poverty-related deforestation? Appl Econ Lett. 2017; 24: 374-380.
5. Macqueen DJ. Forest connect: Reducing poverty and deforestation through support to community forest enterprises. Int For Rev. 2008; 10: 670-675.
6. Honey‐Rosés J. Disentangling the proximate factors of deforestation: The case of the Monarch butterfly Biosphere Reserve in Mexico. Land Degrad Dev. 2009; 20: 22-32.
7. Chadid MA, Dávalos LM, Molina J, Armenteras D. A Bayesian spatial model highlights distinct dynamics in deforestation from coca and pastures in an Andean biodiversity hotspot. Forests. 2015; 6: 3828-3846.

8. DeFries RS, Rudel T, Uriarte M, Hansen M. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. Nat Geosci. 2010; 3: 178-181.
9. Koyuncu C, Yilmaz R. Impact of private forest ownership on deforestation and poverty. Qual Quant. 2013; 47: 1657-1664.
10. Cvitanović M, Blackburn GA, Jepsen MR. Characteristics and drivers of forest cover change in the post-socialist era in Croatia: Evidence from a mixed-methods approach. Reg Environ Change. 2016; 16: 1751-1763.
11. Oyekale AS, Adeleke AH. Assessment of land degradation and farm-level deforestation in the Niger Delta Region of Nigeria. Life Sci J. 2012; 9: 118-122.
12. Hosonuma N, Herold M, De Sy V, De Fries RS, Brockhaus M, Verchot L, et al. An assessment of deforestation and forest degradation drivers in developing countries. Environ Res Lett. 2012; 7: 044009.
13. Sleeter BM, Sohl TL, Loveland TR, Auch RF, Acevedo W, Drummond MA, et al. Land-cover change in the conterminous United States from 1973 to 2000. Glob Environ Change. 2013; 23: 733-748.
14. Yang S, Mountrakis G. Forest dynamics in the US indicate disproportionate attrition in western forests, rural areas and public lands. Plos ONE. 2017; 12: e0171383.
15. Auch RF, Drummond MA, Xian G, Sayler KL, Acevedo W, Taylor JL. Regional differences in upland forest to developed (urban) land cover conversions in the conterminous US, 1973-2011. Forests. 2016; 7: 132.
16. Drummond MA, Loveland TR. Land-use pressure and a transition to forest-cover loss in the eastern United States. BioScience. 2010; 60: 286-298.
17. Alig RJ. Land use changes involving forestry in the United States, 1952 to 1997, with projections to 2050. Portland: US Department of Agriculture, Forest Service, Pacific Northwest Research Station; 2003. pp.1-77.
18. Napton DE, Auch RF, Headley R, Taylor JL. Land changes and their driving forces in the Southeastern United States. Reg Environ Change. 2010; 10: 37-53.
19. Sohl TL, Sohl LB. Land-use change in the Atlantic coastal pine barrens ecoregion. Geogr Rev. 2012; 102: 180-201.
20. Auch RF, Napton DE, Kambly S, Moreland Jr TR, Sayler KL. The driving forces of land change in the Northern Piedmont of the United States. Geogr Rev. 2012; 102: 53-75.
21. Lo CP, Yang X. Drivers of land-use/land-cover changes and dynamic modeling for the Atlanta, Georgia metropolitan area. Photogramm Eng Remote Sensing. 2002; 68: 1073-1082.
22. De Chant T, Gallego AH, Saornil JV, Kelly M. Urban influence on changes in linear forest edge structure. Landsc Urban Plan. 2010; 96: 12-18.
23. Wall DL. Deforestation and cattle production in the Ozark mountains of Arkansas: The case of Searcy County. J Rural Stud. 1996; 12: 27-40.
24. Martin D. Geographic information systems: Socioeconomic applications. Hove: Psychology Press; 1996.
25. Mountrakis G. Linking population and forest dynamics over the conterminous US for the 1990s and 2000s. Adv Environ Eng Res. 2021; 2. Doi:10.21926/aeer.2101003.
26. Kim DH, Sexton JO, Noojipady P, Huang C, Anand A, Channan S, et al. Global, Landsat-based forest-cover change from 1990 to 2000. Remote Sens Environ. 2014; 155: 178-193.

27. Wickham J, Stehman SV, Gass L, Dewitz JA, Sorenson DG, Granneman BJ, et al. Thematic accuracy assessment of the 2011 national land cover database (NLCD). Remote Sens Environ. 2017; 191: 328-341.
28. Zheng D, Heath LS, Ducey MJ. Carbon benefits from protected areas in the conterminous United States. Carbon Balance Manag. 2013; 8: 4.
29. Jha S, Bawa KS. Population growth, human development, and deforestation in biodiversity hotspots. Conserv Biol. 2006; 20: 906-912.
30. Grekousis G, Mountrakis G. Sustainable development under population pressure: Lessons from developed land consumption in the conterminous US. PloS ONE. 2015; 10: e0119675.
31. Cho SH, Kim H, Roberts RK, Kim T, Lee D. Effects of changes in forestland ownership on deforestation and urbanization and the resulting effects on greenhouse gas emissions. J For Econ. 2014; 20: 93-109.
32. Alig RJ. A United States view on changes in land use and land values affecting sustainable forest management. J Sustain For. 2007; 24: 209-227.
33. Davis R, Yang Z, Yost A, Belongie C, Cohen W. The normal fire environment-modeling environmental suitability for large forest wildfires using past, present, and future climate normals. For Ecol Manag. 2017; 390: 173-186.
34. Anselin L. Spatial Econometrics. Amsterdam: Kluwer Academic Publishers; 2007. pp.310-330.
35. Cohen WB, Yang Z, Stehman SV, Schroeder TA, Bell DM, Masek JG, et al. Forest disturbance across the conterminous United States from 1985-2012: The emerging dominance of forest decline. For Ecol Manag. 2016; 360: 242-252.
36. Auch RF, Napton DE, Sayler KL, Drummond MA, Kambly S, Sorenson DG. The Southern Piedmont's continued land-use evolution, 1973-2011. Southeast Geogr. 2015; 55: 338-361.
37. Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, et al. High-resolution global maps of 21st-century forest cover change. Science. 2013; 342: 850-853.
38. Wear DN, Flamm R. Public and private forest disturbance regimes in the Southern Appalachians. Nat Resour Model. 1993; 7: 379-397.