Modeling Earth Systems and Environment
https://doi.org/10.1007/s40808-020-00781-2

ORIGINAL ARTICLE

Assessment of drivers of forest changes using multi-temporal analysis and boosted regression trees model: a case study of Nyeri County, Central Region of Kenya

Nancy Mwangi[1] · Hunja Waithaka[1] · Charles Mundia[2] · Mwangi Kinyanjui[3] · Felix Mutua[1]

Received: 30 December 2019 / Accepted: 10 April 2020

Abstract

The Central Region of Kenya has undergone significant changes in land cover due to a broad range of drivers. These changes are more pronounced in forestland conversions. Past researches within the study area have identified drivers of land cover change without quantifying the influence of these drivers. Predictor variables include population density, precipitation, elevation, slope, forest fires, soil texture, proximity to roads, rives and towns. Land cover changes were analyzed using multi-temporal land cover maps between year 1990 and 2014. Boosted regression trees model was applied to determine the significant drivers and quantify their relative influence on key forestland transitions. The local and spatial influence of the drivers has further been analyzed by geographical weighted regression using coefficients determined at each sample point. Significant land cover changes continuously occurred over the study period. Forestland reduced from 38.90% in 1990 to 38.14% in 2014. Grassland reduced from 32.59 to 22.57%, cropland increased from 28.05 to 38.83% and wetland changed from 0.07 to 0.04%. Other land which constitutes of bare land and built up increased from 0.38 to 0.42%. The results show population density had the highest contribution to forestland changes throughout the study period, with a minimum contribution of 20.02% to a maximum of 26.04%. Other significant variables over the study period are precipitation, slope, elevation and the proximity variables. The results indicate that the relative influence of the drivers to forestland conversion varies with time, location and type of transition.

Keywords Change analysis · Drivers of change · Relative influence · Spatial influence · Central Kenya

✉ Nancy Mwangi
nwmwangi@jkuat.ac.ke

Hunja Waithaka
hunja@eng.jkuat.ac.ke

Charles Mundia
ndegwa.mundia@gmail.com

Mwangi Kinyanjui
mwangikinyanjui@gmail.com

Felix Mutua
fnmutua@jkuat.ac.ke

[1] Department of Geomatic Engineering and Geospatial Information Systems (GEGIS), Jomo Kenyatta University of Agriculture and Technology (JKUAT), P.O. Box 62000, Nairobi 00200, Kenya

[2] Institute of Geomatics, GIS and Remote Sensing (IGGReS), Dedan Kimathi University of Technology, P.O. Box 657, Nyeri 10100, Kenya

[3] Department of Natural Resources, Karatina University, P.O. Box 1957 Karatina, Kenya

Introduction

Land use land cover conversion and human-driven alterations of the surface of the earth (Foley 2005) have affected spatial and temporal patterns and processes of global ecosystem (Turner et al. 2007). One of the critical land use land cover conversions is in the forestland. Forestland conversion is mostly in form of deforestation and degradation, and it is heavily attributed to human-driven activities and climate change. Deforestation and degradation have considerable influence on carbon and greenhouse gas (GHG) fluxes within an ecosystem (Sohl et al. 2012). GHG emissions are a global concern due to the negative impacts on global climate systems, and any given land use change scenario may emit or sequester carbon (Baccini et al. 2012; Dunn et al. 2013). Recognizing the negative impacts of GHG emissions on global climate system, the international community is currently negotiating initiatives to reduce emissions from deforestation and forest degradation in developing countries.

Published online: 06 May 2020

Forest ecosystems are vital for biogeochemical processes, but also for the livelihood of forest-dependent communities (Igu 2017). Kenya's total forest cover is currently estimated at 7.4% (Government of Kenya 2018), which is below 10% international threshold. It is estimated that forestry contributes about 3.6% of Kenya's GDP and directly or indirectly support other key productive and service sectors. Forests further account for 75% of the country's renewable surface water sources (Government of Kenya 2014). In the last two decades, significant socioeconomic changes have continued to occur in the country, exerting pressure on forestland. The result is decreasing forest cover mainly attributed to unsustainable utilization and conversion of forestland to other land uses (Government of Kenya 2018).

Previous studies have demonstrated the extent of forest loss in Kenya and identified some underlying drivers (Government of Kenya 2010), but there remains a glaring dearth of research that quantifies the influence of these drivers on forestland conversion in Kenya. Combination of climatic, topographic, soil quality, demographic and accessibility factors determine the likelihood of land cover change (Morrison et al. 2018; Were et al. 2014). These factors have largely been analyzed by traditional global regression models in the previous studies (Campbell et al. 2005). A typical global regression model applied to spatial data assumes a stationary process, and the parameters obtained in calibration of such a model are constant over space. However, relationships between predictor and response variables are intrinsically different across space (Saefuddin et al. 2012). Assessment of determinants of forestland conversion is complicated by the nonlinear relationship between the factors and the response variables, and the interactions between predictor variables (Zhang et al. 2016; Kolb et al. 2013). The conventional regression models have scientific merit because they are easy to understand and interpret and provide numerous options to estimate the parameters that relate the input data to the output data (Munroe and Müller 2007). However, problems with unknown and possibly nonlinear relationships between input and output variables are difficult to consider in these regression frameworks (Turner et al. 2007; Verburg et al. 2006).

This study applies boosted regression trees model (BRT) which combines regression trees and boosting to generate non-parametric statistical models that can capture non-linear relationships and interactions between variables (Tonkin et al. 2015). The models offer considerable gains over conventional regression techniques due to their capability of fitting interactions among predictor variables and fitting complex nonlinear relationships (Elith et al. 2008; Leathwick et al. 2006; Zhang et al. 2016). Geographically weighted regression (GWR) applied in this study allows the measured relationships to vary over space and addresses the effect of spatial non-stationarity of data (Saefuddin et al. 2012). The model accounts for heterogeneity through calculation of coefficients at each measurement location point (Kirui et al. 2017). This study demonstrates the use of the models to quantify the contribution of each variable and determine spatial heterogeneity in the influence of the drivers across the study area which is critical in guiding intervention measures.

The datasets used to evaluate the drivers of land cover change in Nyeri County were clipped from the datasets prepared for a previous analysis for the entire Central Region of Kenya. The results of change analysis for the whole of Central Region of Kenya indicated conversion from forestland to cropland was the key transition followed by forest conversion to grassland across all the years. However, analysis of Nyeri County (one county out of the eight in Central Region) indicates a different pattern of change as indicated in the results of analysis of Nyeri County.

Materials and methods

Study area

Nyeri County is one of the eight counties in Central Region of Kenya, covering 2361 sq. km, as shown in Fig. 1. It is located 150.8 km north of Nairobi in the central highlands. The current population is 759,164 people, with approximate population density of 321.54. The study area is characterized by great topographic variability with altitude ranging from 1200 to 5000 m. The slope ranges between 0 and 68%. The annual mean temperature ranges from 12° to 27°. Rainfall ranges from 500 to 2400 mm. It is the only county that constitutes parts of the two major forests in the country, Mount Kenya and the Aberdare ranges. Agriculture forms the main source of livelihood. The county has undergone significant loss of forest cover from 38.9% in 1990 to 20.00% in 2018. The area is characterized with land fragmentation, and surrounding population exploits the forests through logging, charcoal burning, agriculture and encroachment of settlements (Government of Kenya 2010).

Data

Various datasets have been used as shown in Table 1. Land cover change maps were generated from the country's classified land cover maps for years 1990, 1995, 2000, 2005, 2010 and 2014 acquired from Department of Resource Surveys and Remote Sensing (DRSRS). The land cover maps are created from LandsatTM images with a resolution of 30 m. The overall accuracy ranges from 74.88% in 1990 to 85.27% in 2014.

Biophysical and socioeconomic variables considered included slope, elevation, precipitation, soil texture,

Table 1 Datasets

Data	Source	Format
Land cover images	Department of resource surveys and remote sensing (DRSRS)	Raster
Elevation	SRTM 30 m digital elevation model	Raster
Slope	SRTM 30 m digital elevation model	Raster
Precipitation	WorldClim	Raster
	Africa RFE	
Population	Kenya National Bureau of Statistics (KNBS) census data	Shapefile
Roads	Kenya Urban Roads Authority (KURA).	Shapefile
Towns	International Livestock Research Institute (ILRI).	Shapefile
Rivers	International Livestock Research Institute (ILRI).	Shapefile
Soil	Kenya Agricultural and Livestock Research Organization (KALRO)	Shapefile
Forest fires	MODIS forest data from Kenya Forest Services (KFS)	Shapefile

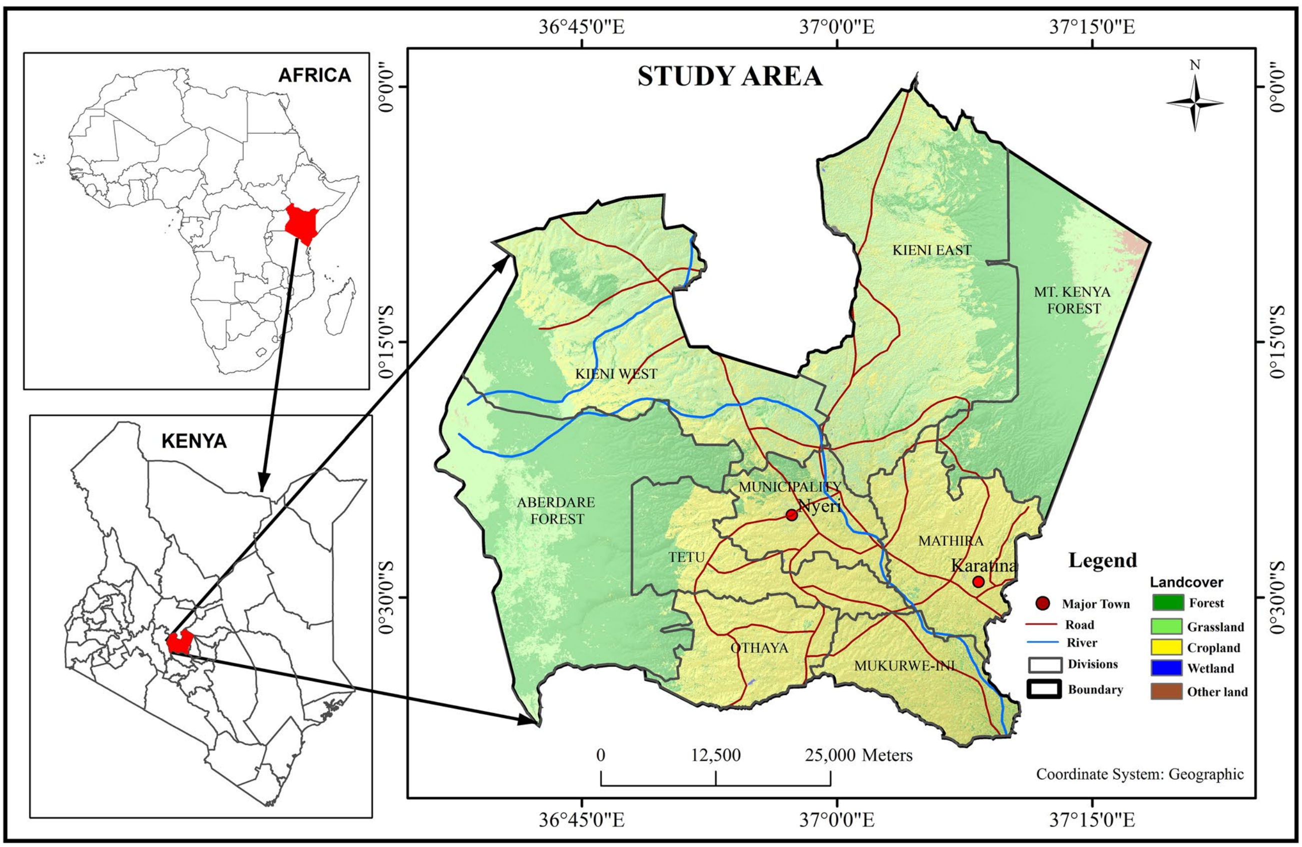

Fig. 1 Study area

distance to major roads, distance to major rivers, distance to towns, population density and forest fires. The SRTM 30 m Digital Elevation Model was used to generate the slope and elevation. Population density was derived from the Kenya National Bureau of Statistics (KNBS) census data for 1989, 1999 and 2009. The unit of measurement is sublocation. Projected population data were provided by KNBS for years when census was not carried out. Road data were acquired from Kenya Urban Roads Authority (KURA). Data for towns and rivers were acquired from International Livestock Research Institute (ILRI). Distance from major roads, distance from towns and distance from rivers surfaces were created using the roads, towns and rivers layers, respectively, using ArcGIS software. Precipitation data were acquired from both WorldClim and Africa RFE sites with a resolution of 1 km. Kenya Forest Services (KFS) provided the shapefiles for MODIS forest fires data. Kenya Agricultural and Livestock Research

Organization (KALRO) provided shapefiles for soil data. All datasets were resampled to 30-m resolution consistent with the resolution of land cover maps.

Methods

The overall overview of the methodology is as shown in the methodology flowchart diagram in Fig. 2. About 15,000 cluster points were randomly sampled and used to extract values for both response and predictor variables. The response variable was extracted from land cover change maps. Points where forestland conversion occurred were assigned a value of one (1) while points where no conversion occurred were assigned a value of zero (0). Continuous surfaces for each of the predictor variables were created in ArcGIS and values extracted for all the sample points.

Boosted regression trees (BRT) modeled the relationship between the predictor variables and the response variable based on the extracted values from random sample points. BRT combines regression trees and boosting to generate non-parametric statistical models that can capture nonlinear relationships and interactions between variables. It accommodates complex linear and nonlinear responses to multiple categorical and continuous predictors while being relatively insensitive to collinearity problems (Zhang et al. 2016). Predictor variables are subjected to a recursive binary split that fits a simple model to each resulting section, until a stopping criterion is reached, such as minimization of prediction error (Hastie et al. 2009). The importance of a predictor variable is quantified based on the relative influence of the variable and the partial dependence plots (Gu et al. 2019). BRT estimates the relative influence/contribution of each predictor variable on the response variable based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman 2003; Thorn et al. 2016).

$$\hat{I}_j^2 = \frac{1}{M}\sum_{m=1}^{M} I_j^2(T_m)$$

$\hat{I}_j^2$ relative influence of input variable, M total number of trees, I_j^2 squared improvement in the model, T_m individual tree.

Relative influence is scaled such that the sum adds to 100, with the higher value indicating stronger influence on the response (Müller et al. 2013). Significance of a variable is either low or high and can be determined in two ways. The first approach is the random chance given by $100\%/n$, where n is the number of input variables. Variables with value below random chance are categorized as of low importance (Thorn et al. 2016). In the second approach, variables whose influence exceeds the median of each model are classified as highly important and those below the median are classified as of low importance (Gu et al. 2019).

Partial dependence plots (PDPs) summarize the overall relationship between input variables and the probability of forestland conversion to other land cover classes, by plotting the modeled relationship between one predictor variable and the response variable when all other predictor variables are held constant at their mean values (Thorn et al. 2016). The model was fitted with Bernoulli distribution to deal with presence or absence of forest conversion. Several combinations of model parameters were tested, learning rate (l_r) of 0.025, 0.05 and 0.1, tree complexity (t_c) of 3, 4, 5, 6 and bag fraction of 0.5 and 0.75. These parameters were calibrated using tenfold cross-validation. The model with the highest cross-validated receiver operating characteristic area under the curve (ROC AUC) score was selected as the most optimal settings (Hastie et al. 2009). The model was parameterized with l_r of 0.005, t_c of 6 and bag fraction of 0.75 which produced the least deviance and the highest AUC using tenfold cross-validation. BRT analysis was performed using *R* packages "gbm" and dismo developed by Ridgeways and Elith (Elith et al. 2008).

Relationships between predictor and response variables are intrinsically different across space. GWR applied in this study allows the measured relationships to vary over space and addresses the effect of spatial non-stationarity of data (Saefuddin et al. 2012). The model accounts for heterogeneity through calculation of coefficients at each measurement location point (Kirui et al. 2017). The significant variables obtained from BRT were used in spatial prediction to determine non-stationarity of each predictor variable using GWR

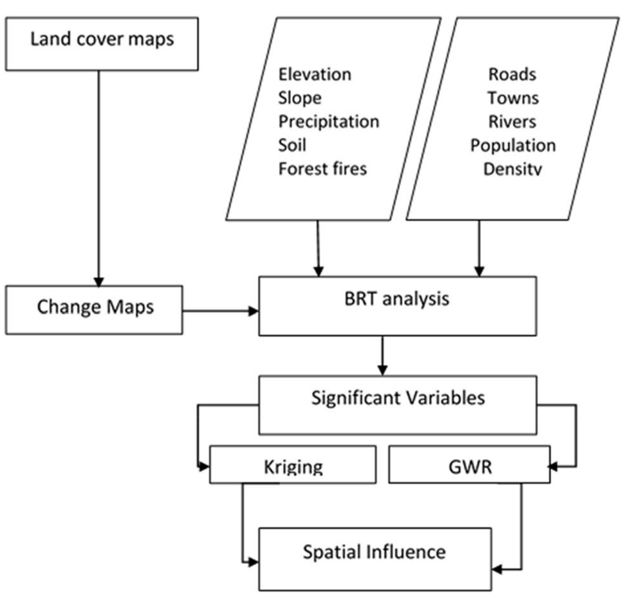

Fig. 2 Methodology

and regression kriging. Regression kriging combines regression and spatial interpolation thus improving its performance above other interpolation methods (Gia Pham et al. 2019).

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_1$$

where β_{0i} and β_{1i} are local model parameters specific to location at (u_i, v_i) coordinate.

GWR further determines a weighting function for estimating local model parameters based on distance function. Observations closer in space are assumed to have greater effect on local parameters.

$$w_j(u_i, v_i) = \exp\left[-{}^1\!/_2\left(d_{ij}/b\right)^2\right]$$

where b is bandwidth.

A bandwidth of 2342.596 and minimum value of Akaike Information Criterion (AIC) of 1900.756 with adjusted R-squared 0.40 were the best achieved used to fit the model. The coefficients at each measurement location point determined from GWR were used in regression kriging to interpolate surfaces depicting the spatial variation of relative influence of each predictor variable (Kirui et al. 2017).

Results

Land cover change analysis

Land cover analysis for the study period showed varying land cover types as shown in Fig. 3. The land cover types include forestland, grassland, cropland, wetland and other land (bare and built up). The percentage change for the individual land cover classes varies with temporal change as shown in Table 2. In the period 1990–1995, there was 5.6% loss of forestland, 16.39% gain in grassland and 12% loss in cropland. In the period between year 1995–2000 and

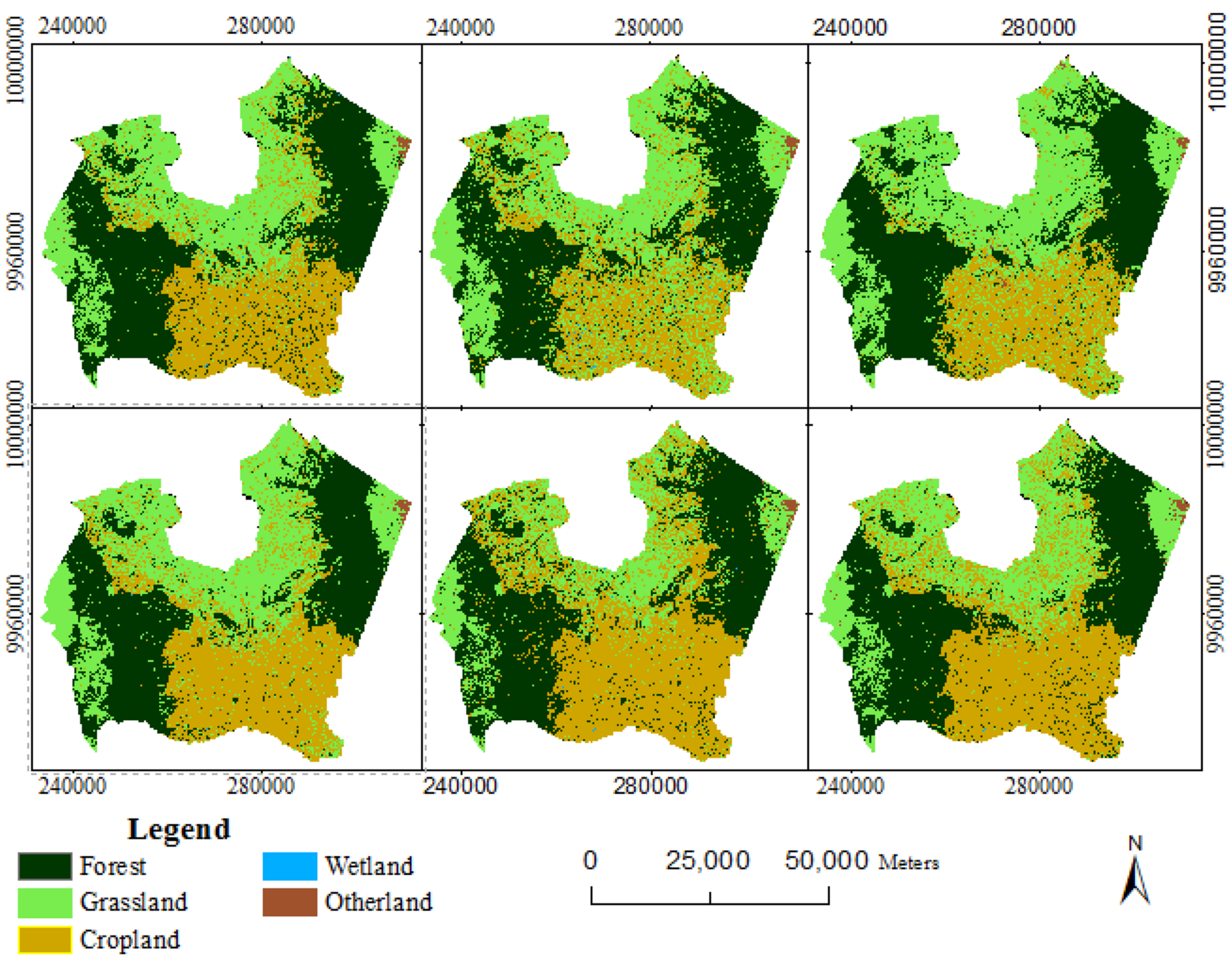

Fig. 3 Land cover types 1990–2014

Table 2 % Land cover change for Nyeri County for the period 1990–2014

Land cover type	1990–1995 % change	1995–2000 % change	2000–2005 % change	2005–2010 % change	2010–2014 % change
% Land cover change between 1990 and 2014					
Forestland	−5.60	0.83	−4.07	7.44	−4.07
Grassland	16.39	3.85	−12.67	−34.56	26.90
Cropland	−12.00	−7.61	29.99	31.85	−12.15
Wetland	100.33	−54.39	−69.26	115.25	−47.18
Other land	0.15	41.62	−33.48	16.80	29.39

2005–2010, the forestland is gained by 0.83% and 7.44%, respectively. Between the year 2000–2005 and 2010–2014, the results indicate equal forestland loss of 4.07%. Forestland experienced gain or loss or persistence in the locations shown in Fig. 4.

The results reveal forestland conversions to cropland and grassland are the key transitions in Nyeri County over the study period. The trends of percentage conversion of forestland to cropland and forestland to grassland between 1990 and 2014 are shown in Fig. 5. Conversion of forestland to cropland and forestland to grassland was highest in the period between 1995 and 2000 with a total percentage conversion of 7.13% and 13.88%, respectively.

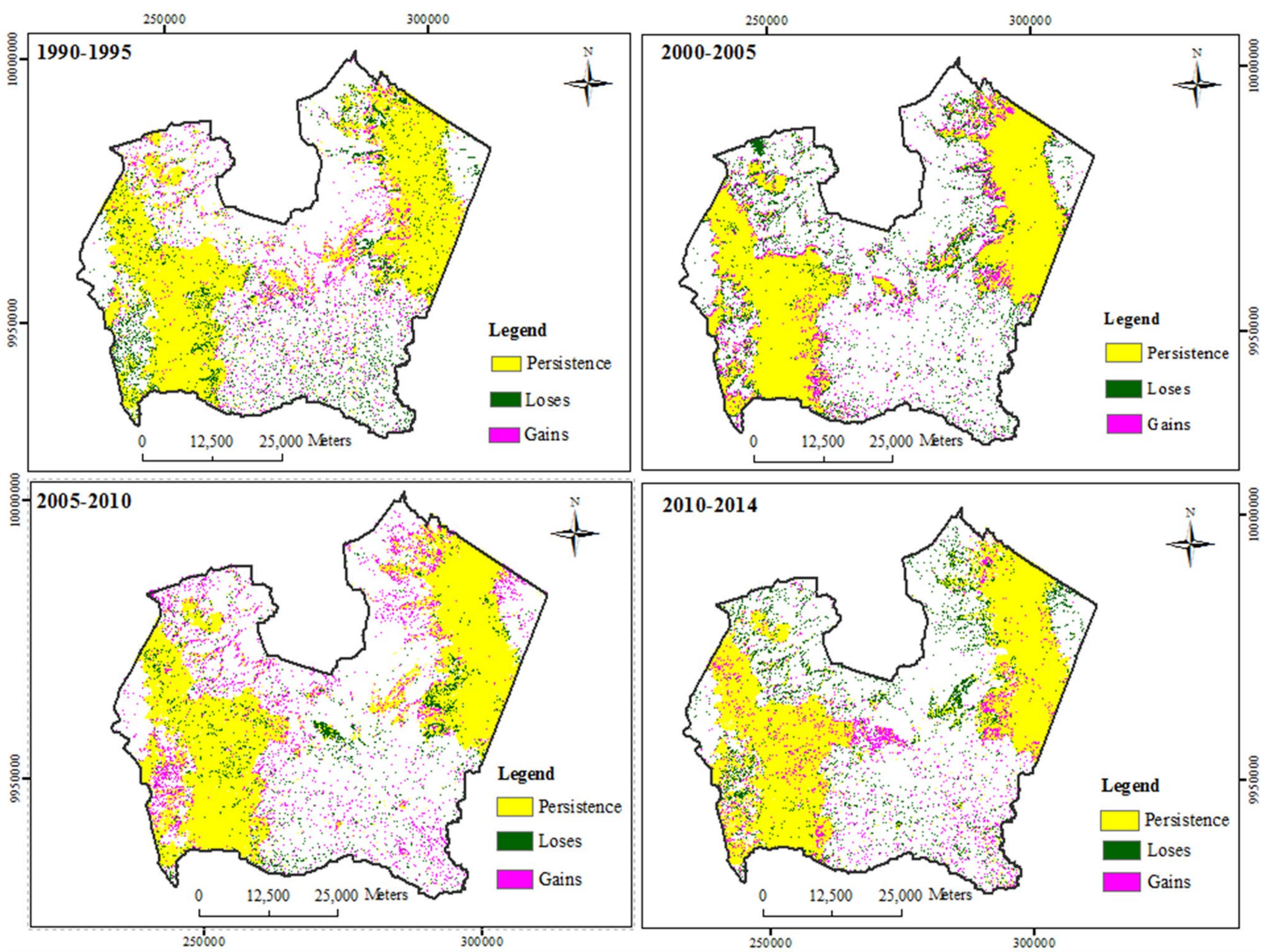

Fig. 4 Forestland changes for the period 1990–2014

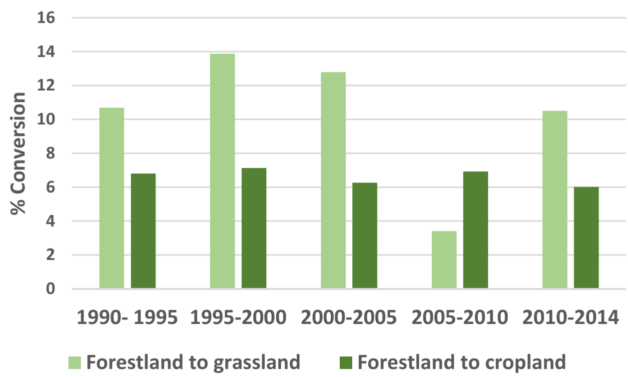

Fig. 5 Forest conversion for Nyeri County between year 1990 and 2014

Analysis of significant drivers of forestland conversion

Nine variables were analyzed in the period between year 1990 and 2014. The percentage contribution of each variable to forestland conversion is shown in Tables 3 and 4. Population density, elevation, slope and precipitation are the four most significant drivers in the study period. Relative contribution of population density to forestland conversion decreased from 26.04 to 17.84% for conversion to cropland, and from 25.43 to 22.43% for conversion to grassland between 1990 and 2014. Soil texture is the least significant driver with relative influence ranging from 0.93 to 3.78% for conversion to cropland and 1.55% to 3.45% for conversion to grassland. The model indicated no variability with forest fires. The results show that distance to towns, distance to roads and distance to rivers contribute significantly to the key conversions with higher influence on forestland conversion to cropland.

Partial dependency plots (PDPs)

Partial dependency plots summarize the overall relationship between a predictor variable and the probability of forest conversion when other variables are kept constant. The *y*-axis represents the probability of forest conversion and the *x*-axis gives the data range of the predictor variable. Population density, elevation, precipitation, slope, distance to roads and distance to towns demonstrate a strong relationship with both conversions for the period 1995–2000 as shown in Figs. 6 and 7. Distance to rivers additionally exhibits a strong relationship with forestland conversion to grassland. 1995–2000 is the period when forestland conversions were highest over the study period as shown in Fig. 5.

The PDPs demonstrate strong relationship of population density on the probability of forestland converting to cropland. Probability of forest conversion to cropland increased for densities ranging 200–1000/km^2 and decreased for densities above 3000/km^2. Probability of conversion increased with increased rainfall with high probability for values ranging 1400 mm and above. Areas with low rainfall demonstrate little or no likelihood of conversion to cropland. Areas with values of elevation ranging between 1500 and 2500 m were more likely to have forestland converted to cropland. The

Table 3 Contribution of drivers to forestland conversion to grassland in Nyeri County

Period	Elevation	Distance to rivers	Distance to roads	Slope	Soil	Distance to towns	Population density	Precipitation
% Relative contribution of variables to conversion from Forestland to Grassland								
1990–1995	25.40	6.10	6.38	8.83	1.55	7.19	25.43	19.14
1995–2000	26.48	9.02	6.00	13.75	2.34	6.02	24.09	12.30
2000–2005	24.75	5.78	9.59	11.09	3.45	7.25	25.11	12.96
2005–2010	20.26	7.73	8.47	14.04	2.98	9.10	23.75	13.67
2010–2014	28.85	6.16	8.60	8.83	2.10	8.96	22.43	14.06

Table 4 Contribution of drivers to forestland conversion to cropland in Nyeri County

Period	Elevation	Distance to rivers	Distance to roads	Slope	Soil texture	Distance to towns	Population density	Precipitation
% Relative contribution of variables to conversion of forestland to cropland								
1990–1995	19.40	7.04	6.75	11.51	0.93	11.21	26.04	17.12
1995–2000	18.09	8.73	10.38	14.35	1.43	10.03	22.77	14.23
2000–2005	16.75	8.79	9.57	14.09	1.31	12.13	25.74	11.62
2005–2010	17.89	7.07	6.85	19.74	3.78	11.66	20.02	12.99
2010–2014	18.60	8.08	9.16	18.42	1.89	10.40	17.84	15.61

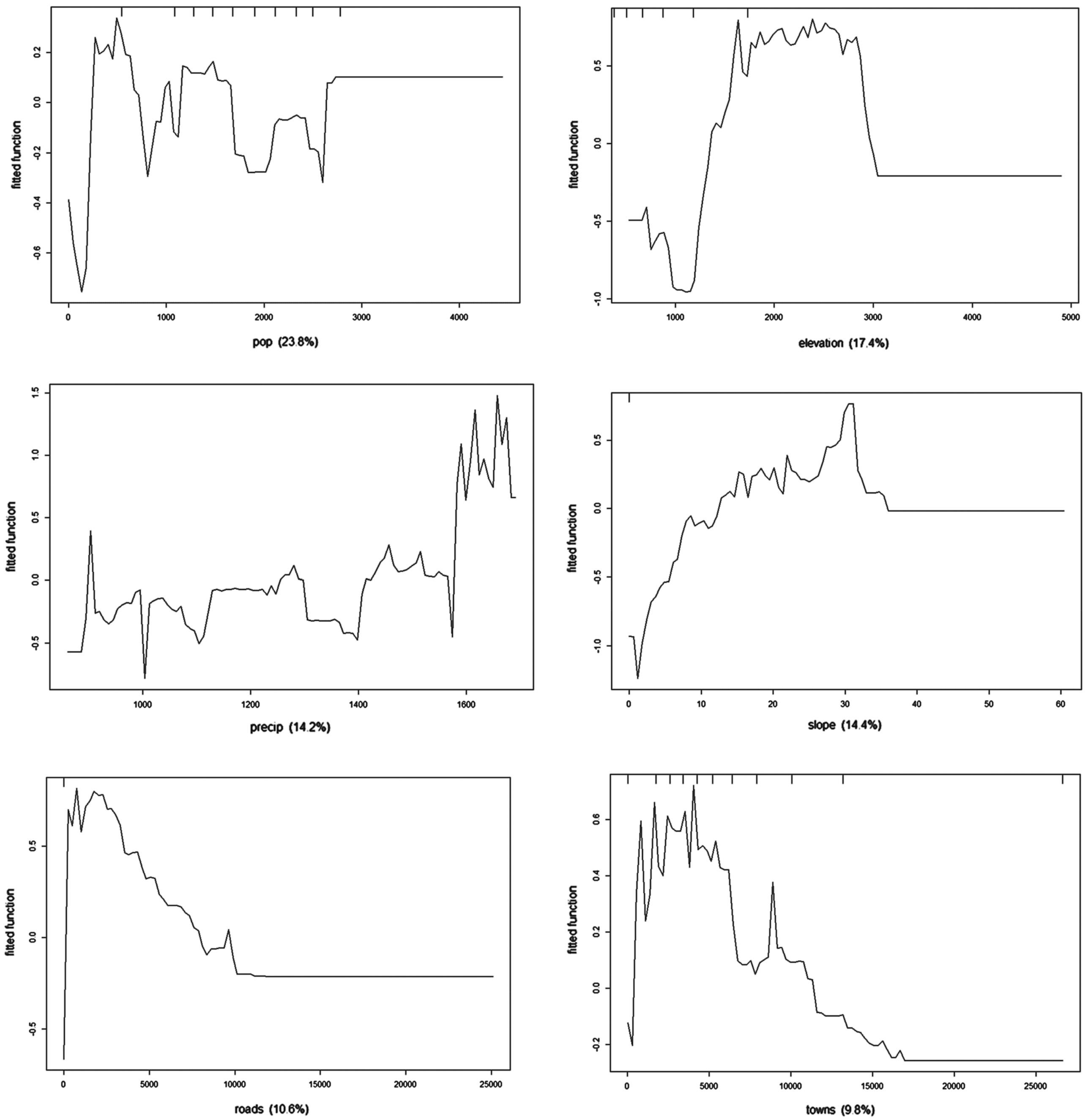

Fig. 6 PDPs of the predicted probability of forestland conversion to cropland between the year 1995 and 2000

probability decreases for areas with 3000 m above mean sea level and above. Probability of forest conversion to cropland decreased with increased distance to towns, distance to roads and distance to rivers.

Probability of forestland conversion to grassland was higher for population density below 1000 km^2. There was less likelihood of conversion where distance to roads and distance to towns was between 5 and 20 km. Conversion to grassland is likely to occur 1–5 km from the rivers. The effect of slope and elevation appears consistent in both conversions. Probability of conversion to grassland was lower in areas with high rainfall, differing with the conversion to cropland where areas with high rainfall are more likely to experience forestland conversion to cropland.

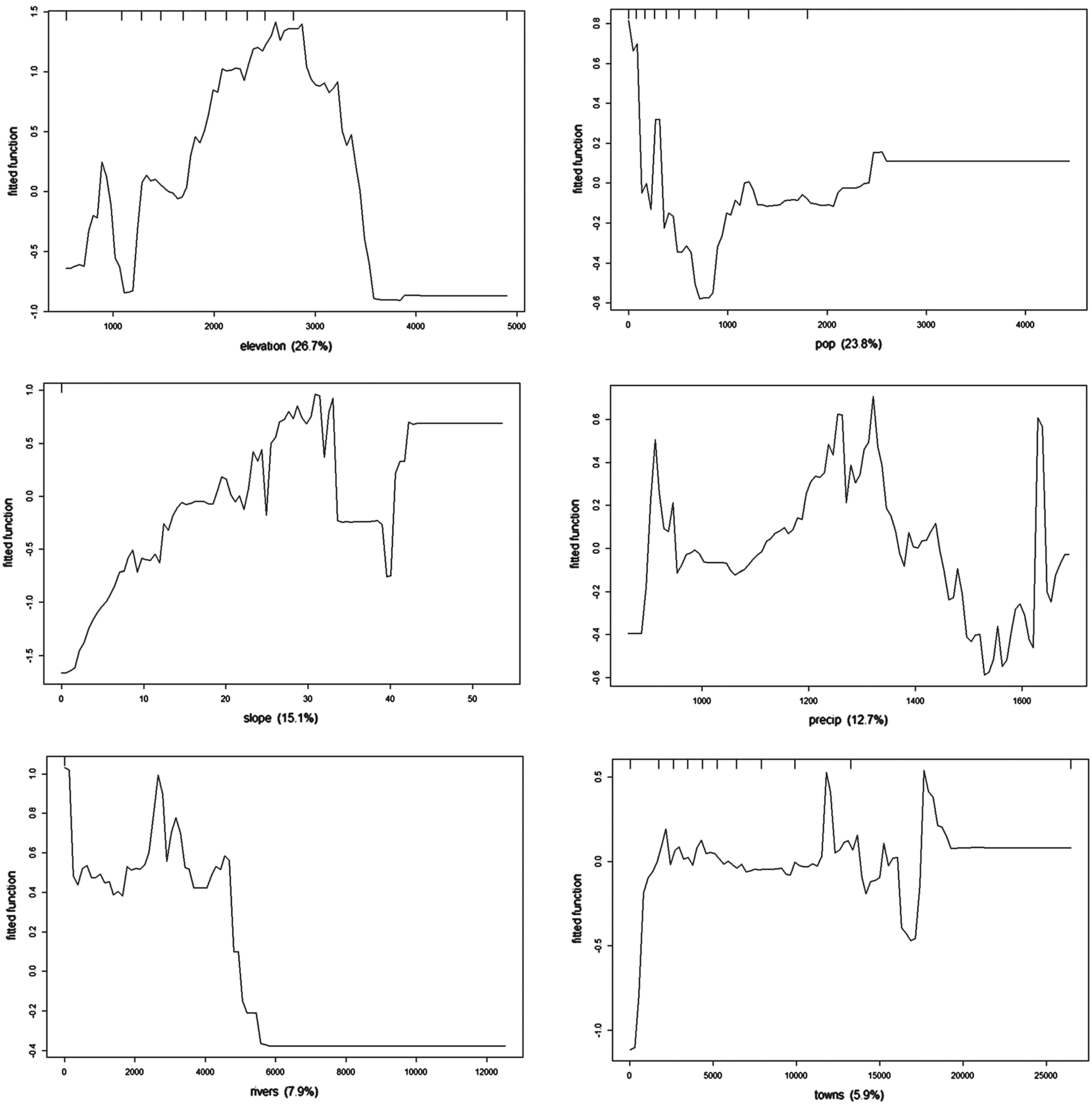

Fig. 7 PDPs of the predicted probability of forestland conversion to grassland in the year 1995–2000

Spatial influence of explanatory variables

The GWR analysis demonstrates variation of influence of each variable on forestland conversion depending on spatial location, i.e., spatial heterogeneity as shown in Figs. 8 and 9.

The maps show non-stationarity relationship of key drivers and the forest conversion to cropland. Regression coefficients results from GWR indicate both negative and positive correlation of the variables with the forest transition at different locations in the study area. The surfaces show the direction of influence which is important in informing policy. In the year 2000, the influence of population density on the forestland conversion to cropland ranged between low and high across the county. Comparison with year 2010 demonstrates a similar pattern where population density influenced forestland conversion to cropland in the entire county with an exception of a small area towards the north eastern border of the county. Areas where the roads demonstrated low influence in 2000 transited to higher influence in the year 2010. The direction of influence follows the

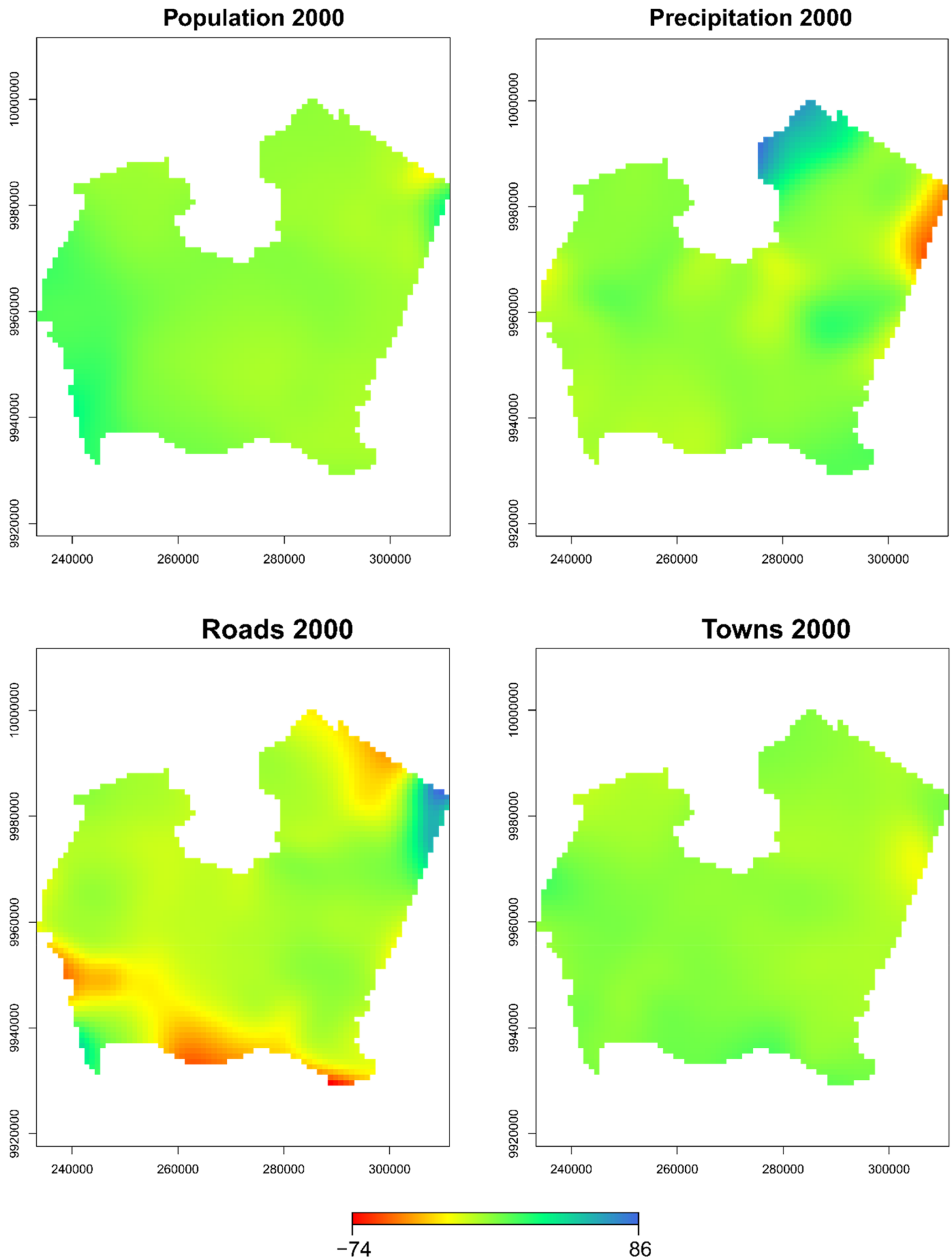

Fig. 8 Spatial influence of forestland conversion to cropland in the year 2000

availability of public access. The results further show the influence of rainfall on forest transition to cropland is more uniform in year 2010 and diverse in year 2000. The direction of higher influence due to rainfall changes from the upper region in the year 2000 to the lower region in the year 2010.

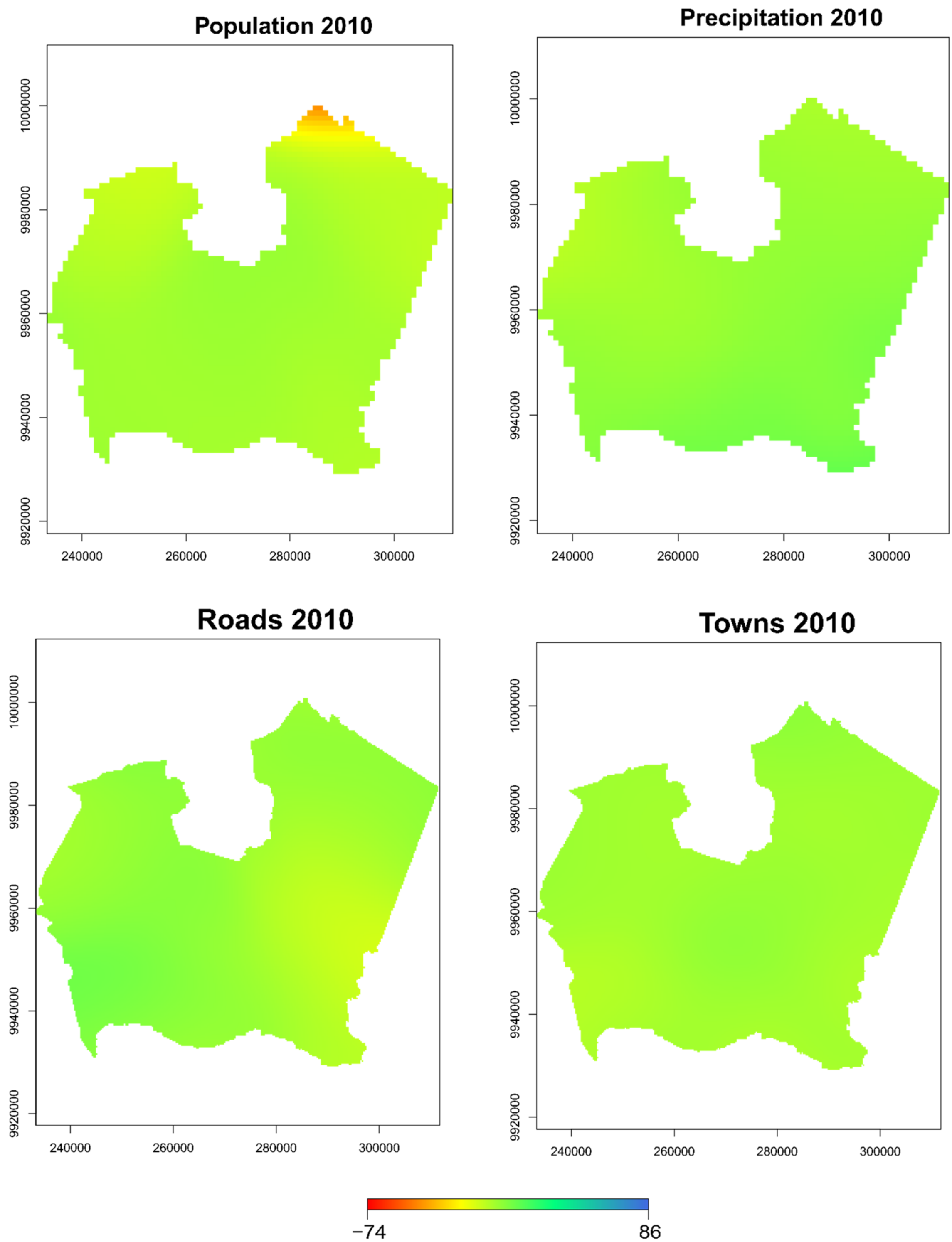

Fig. 9 Spatial influence of forestland conversion to cropland in the year 2010

Discussion of results

Spatial statistical tools provide techniques to quantify the influence of explanatory variables on land use land cover change (Adhikari et al. 2017). The results show consistency with previous studies that the relationship between the drivers and forestland conversion is dependent on the geographical scale, region of study and time (Government of Kenya 2010; Kindu et al. 2015). The contributions of the drivers are different in all the five epochs considered in this study. Forestland conversion to grassland emerged as the key conversion in Nyeri County, a departure from the results of the whole Central Region. There is a significant variation in results from the entire Central Region to a specific county, thus demonstrating effect of the scale of analysis on the influence of drivers on forestland conversion.

Population density has both positive and negative relationship with deforestation exhibiting high value of relative contribution in most epochs. This is partly attributed to population growth in Nyeri County from 607,292 people in year 1989 census to 693,558 people in year 2009 census. The most intense areas of forest loss to cropland occurred at population densities of 200–1000/km^2 and at a radius of about 10 km from towns and roads as demonstrated by the PDPs. This is explained by the fact that most of the conversion occurred on the higher areas of the county near the forests where crop farming is the main economic activity. The results compare well with findings by Thorn et al. (2016). On the contrary, the relationship between conversion and grassland, and population density was stronger in areas where population density was below 500/km^2. These are the semiarid areas of the county characterized by sparse population and are about 20 km and above away from the forestland. The surfaces show influence was high near the Aberd are ranges in 1990 compared to other areas. This is the period a lot of forestland was degazetted for tea farming in this region (Government of Kenya 2010). By 2010, the influence was moderate and uniform across the study area except on the upper region.

Topography exhibits a direct relationship with forestland conversion with significant relative contributions of both elevation and slope in all the epochs. This is contrary to the general idea that topography is a constraint to deforestation. Lack of positive or definite relationship between elevation and slope with deforestation has not been uncommon in previous studies (Peterson et al. 2009). The positive relationship of elevation and slope in this study is logical as the forests in the study area are located in areas with higher elevations and steep slopes. This contradicts the general idea that deforestation occurs in flatter areas due to easy accessibility. However, there is a possibility that something else that overcomes the limitations posed by higher elevations and steeper slopes could be the cause of deforestation in these areas (Adhikari et al. 2017). One of the most likely cause is forest fires which the model results indicated no variability. This is attributed to the duration of 5 years which is long enough for regrowth of affected areas. Analysis with shorter duration is recommended to accurately assess the effect of forest fires.

Proximity variables (distance to roads, distance to towns and distance to rivers) exhibit relationship with deforestation that is expected as shown in the PDPs. Probability of deforestation decreases with increased distance to roads, towns and rivers. Distance to major roads provides easier access to the forestland. There is an increase in the magnitude of influence to both conversions from 1990 to 2014 which is explained by continued improvement in public roads in the region. Direction of influence changed from low to high on the lower region of the county. Recently (Government of Kenya 2018) identified development of major public infrastructures as one of the major factors currently attributing to deforestation and recommended adherence to National Plan Policies. Distance to towns had greater influence on conversion to cropland than conversion to grassland. This is explained by the fact that towns constitute human settlements and the population here largely depends on agriculture for livelihood. The slight decline in influence of proximity to towns between the year 1990 and 2014 is attributed to measures undertaken against illegal logging, charcoal burning and cultivation of land in the forests (Government of Kenya 2010). Influence is moderate and uniform across the study area.

Nyeri County comprises of regions with moderate-to-high rainfall and dry areas with relatively low rainfall. The conversion to cropland occurs in areas of moderate and high rainfall as depicted in the PDPs. These are the regions where large-scale and small-scale agriculture are practiced. These regions are further characterized by fragmentation of land into very small units due to population growth. In the semiarid areas, the main agricultural activity is livestock farming. The PDPs correctly indicate conversion to grassland is more likely in areas of low-to-moderate rainfall where livestock farming is practiced. The relative influence of precipitation decreases with time. The influence had great variation across the space 1990 but uniformly varied in 2010. Soil texture has low influence on both conversions compared with the rest of the variables. The regions with loamy and very clay types of soil texture exhibit more likelihood of conversion to cropland than grassland.

Policy makers establish protected areas to conserve forests. The effectiveness of protected areas in conserving biodiversity has been questioned in previous studies (Andam et al. 2008; Chape et al. 2005). A study by Morrison et al. (2018) indicates fencing protected areas reduce but does not completely eliminate interference of forests by human activities. Aberdare ranges and Mount Kenya are gazetted protected areas in the country, managed and conserved under Forest Management and Conservation Act 2016 (initially Forest Act 2005). However, the findings from this study indicate continued forest loss with population density contributing the highest to forest changes. Several policies and institutional frameworks were enacted in the year 2005 to ensure management and conservation of the forests. This could partly explain the significant decline in forestland conversion to cropland between year 2005 and 2010 and lower rates of conversion from year 2005. On the contrary, deforestation continues to occur as evidenced by land cover change analysis and supported by a recent report on Forest Resources Management (Government of Kenya 2018). This then implies a need for review of existing policies, enforcement and implementation processes.

GWR and BRT models used provided substantial information that explains the relationship between the drivers and deforestation in Nyeri County. The capability of the models to analyze spatial datasets at local scale improves previous results on study of drivers of deforestation in Kenya by demonstrating the spatial variability of the influence of the drivers. Conversely, influence of factors that could not be expressed spatially like policy and political interference could not be quantified in the models.

Conclusion

These results broaden our knowledge of drivers of continued forest loss in Mount Kenya region. Relative influence of these drivers on forestland conversion varied with time, location and nature of conversion. The contribution also differed in magnitude and direction. The key drivers to forest loss were population density, elevation, precipitation, slope, distance to towns, distance to roads and distance to rivers. The results further show that the trend of conversions for the entire region differed with the trend for a specific county. Future policies and research on forest conversions should be time and space specific. Unavailability of historical data for most of the socioeconomic variables resulted in exclusion of some factors in our analysis. The study was based on 5-year epochs which could explain the unexpected results for forest fires due to regrowth. Future studies are recommended with less time interval and with inclusion of more socioeconomic factors.

Acknowledgements We acknowledge the Ministry of Environment and Forestry, Kenya for providing Landsat images and land cover maps. We are grateful to two anonymous reviewers for valuable comments on the initial draft of the manuscript.

Authors contribution NM contributed to formal analysis, methodology and writing of first draft. HW supervised, reviewed and edited the manuscript. CM and JK contributed to conceptualization of project and supervision of research. FN contributed to project administration and acquisition of funding. All authors commented on the manuscript.

Funding This study was part of System for Land-based Emissions Estimation in Kenya (SLEEK) Project, funded by the Australian Government in partnership with the Government of Kenya.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

References

Adhikari S, Fik T, Dwivedi P (2017) Proximate causes of land-use and land-cover change in Bannerghatta National Park: a spatial statistical model. Forests 8:342. https://doi.org/10.3390/f8090342

Andam KS, Ferraro PJ, Pfaff A, Sanchez-Azofeifa GA, Robalino JA (2008) Measuring the effectiveness of protected area networks in reducing deforestation. Proc Natl Acad Sci 105:16089–16094. https://doi.org/10.1073/pnas.0800437105

Baccini A, Goetz SJ, Walker WS, Laporte NT, Sun M, Sulla-Menashe D, Hackler J, Beck PSA, Dubayah R, Friedl MA, Samanta S, Houghton RA (2012) Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nat Climate Change 2:182–185. https://doi.org/10.1038/nclimate1354

Campbell DJ, Lusch DP, Smucker TA, Wangui EE (2005) 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 33:763–794. https://doi.org/10.1007/s10745-005-8210-y

Chape S, Harrison J, Spalding M, Lysenko I (2005) Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. Philos Trans R Soc B 360:443–455. https://doi.org/10.1098/rstb.2004.1592

Dunn JB, Mueller S, Kwon H, Wang MQ (2013) Land-use change and greenhouse gas emissions from corn and cellulosic ethanol. Biotechnol Biofuels 6:51. https://doi.org/10.1186/1754-6834-6-51

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77:802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Foley JA (2005) Global consequences of land use. Science 309:570–574. https://doi.org/10.1126/science.1111772

Friedman JH, Meulman JJ (2003) Multiple additive regression trees with application in epidemiology. Stat Med 22:1365–1381. https://doi.org/10.1002/sim.1501

Gia Pham T, Kappas M, Van Huynh C, Hoang Khanh Nguyen L (2019) Application of ordinary kriging and regression kriging method for soil properties mapping in hilly region of central Vietnam. IJGI 8:147. https://doi.org/10.3390/ijgi8030147

Government of Kenya (2010) Analysis of drivers of deforestation and forest degradation in Kenya

Government of Kenya (2014) Forest policy, 2014. Ministry of Environment, Water and Natural Resources Forest

Government of Kenya (2018) Taskforce report on forest resources management and logging activities in Kenya. Ministry of Environment and Forestry

Gu H, Wang J, Ma L, Shang Z, Zhang Q (2019) Insights into the BRT (boosted regression trees) method in the study of the climate-growth relationship of Masson pine in subtropical China. Forests 10:228. https://doi.org/10.3390/f10030228

Hastie T, Tibshirani R, Friedman J (2009) Unsupervised learning. In: The elements of statistical learning. Springer, New York, pp 485–585. https://doi.org/10.1007/978-0-387-84858-7_14

Igu NI (2017) Swamp forest use and loss in the Niger Delta: contextual and underlying issues. Open J For 07:34–47. https://doi.org/10.4236/ojf.2017.71003

Kindu M, Schneider T, Teketay D, Knoke T (2015) Drivers of land use/land cover changes in Munessa-Shashemene landscape of the south-central highlands of Ethiopia. Environ Monit Assess. https://doi.org/10.1007/s10661-015-4671-7

Kirui PK, Kenduiywo BK, Waithaka EH (2017) Comparison of spatial and conventional regression models in determination of trachoma prevalence and associated risk factors. Geoinform Geostat Overv. https://doi.org/10.4172/2327-4581.1000170

Kolb M, Mas J-F, Galicia L (2013) Evaluating drivers of land-use change and transition potential models in a complex landscape in Southern Mexico. Int J Geogr Inf Sci 27:1804–1827. https://doi.org/10.1080/13658816.2013.770517

Leathwick J, Elith J, Francis M, Hastie T, Taylor P (2006) Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Mar Ecol Prog Ser 321:267–281. https://doi.org/10.3354/meps321267

Morrison J, Higginbottom T, Symeonakis E, Jones M, Omengo F, Walker S, Cain B (2018) Detecting vegetation change in response to confining elephants in forests using MODIS time-series and BFAST. Remote Sens 10:1075. https://doi.org/10.3390/rs10071075

Müller D, Leitão PJ, Sikor T (2013) Comparing the determinants of cropland abandonment in Albania and Romania using boosted regression trees. Agric Syst 117:66–77. https://doi.org/10.1016/j.agsy.2012.12.010

Munroe DK, Müller D (2007) Issues in spatially explicit statistical land-use/cover change (LUCC) models: examples from western Honduras and the Central Highlands of Vietnam. Land Use Policy 24:521–530. https://doi.org/10.1016/j.landusepol.2005.09.007

Peterson LK, Bergen KM, Brown DG, Vashchuk L, Blam Y (2009) Forested land-cover patterns and trends over changing forest management eras in the Siberian Baikal region. For Ecol Manag 257:911–922. https://doi.org/10.1016/j.foreco.2008.10.037

Saefuddin A, Setiabudi NA, Fitrianto A (2012) On comparison between logistic regression and geographically weighted logistic regression: with application to Indonesian poverty data. World Appl Sci J 19:205–210

Sohl TL, Sleeter BM, Sayler KL, Bouchard MA, Reker RR, Bennett SL, Sleeter RR, Kanengieter RL, Zhu Z (2012) Spatially explicit land-use and land-cover scenarios for the Great Plains of the United States. Agr Ecosyst Environ 153:1–15. https://doi.org/10.1016/j.agee.2012.02.019

Thorn A, Thompson J, Plisinski J (2016) Patterns and predictors of recent forest conversion in New England. Land 5:30. https://doi.org/10.3390/land5030030

Tonkin JD, Shah DN, Kuemmerlen M, Li F, Cai Q, Haase P, Jähnig SC (2015) Climatic and catchment-scale predictors of chinese stream insect richness differ between taxonomic groups. PLoS ONE 10:e0123250. https://doi.org/10.1371/journal.pone.0123250

Turner BL, Lambin EF, Reenberg A (2007) The emergence of land change science for global environmental change and sustainability. Proc Natl Acad Sci 104:20666–20671. https://doi.org/10.1073/pnas.0704119104

Verburg PH, Kok K, Pontius RG, Veldkamp A (2006) Modeling land-use and land-cover change. In: Lambin EF, Geist H (eds) Land-use and land-cover change. Springer, Berlin, pp 117–135. https://doi.org/10.1007/3-540-32202-7_5

Were K, Dick ØB, Singh BR (2014) Exploring the geophysical and socio-economic determinants of land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya. GeoJournal 79:775–790. https://doi.org/10.1007/s10708-014-9525-2

Zhang W, Du Z, Zhang D, Yu S, Hao Y (2016) Boosted regression tree model-based assessment of the impacts of meteorological drivers of hand, foot and mouth disease in Guangdong, China. Sci Total Environ 553:366–371. https://doi.org/10.1016/j.scitotenv.2016.02.023