THE SOCIOECONOMIC DRIVERS OF LAND COVER DYNAMICS IN THE LI RIVER BASIN, CHINA

Yuan Zhang, Qiming Qin, Ruofeng Xu, Cong Lin, Jun Li
Institution of Remote Sensing and GIS, Peking University, 5 Yiheyuan Road, Beijing 100871, China
Email: zhangyuan76@gmail.com, qmqinpku@163.com, xuruofengpku@163.com, lincong520@163.com, newton2069@126.com

KEY WORDS: Cross Correlation; Land Cover Classification; Fixed Effects; Relationship Analysis; Regression

ABSTRACT: Socioeconomic factors play an important part in the land cover change especially after the Industrial Revolution. This paper focuses on analyzing what and how socioeconomic factors drive the land cover change in the Li River Basin, China with a unique tourism-driven economy. Three types of analyses were conducted to fulfill this purpose: generating the correlation coefficients between antecedent socioeconomic variables and land cover proportions; using fixed effects stepwise regression to model their relationship for four different land cover types; using cross correlation to explore the lagged relationship. These analyses demonstrate that variations of socioeconomic variables are closely related to the quantities of land cover change, and they are extremely important in the urbanization process.

1. INTRODUCTION

The extent and pace of human activity have been significantly expanding and accelerating. The land cover change, as one of the most prominent consequences of intensified human activity, has become more rapid and extensive since the 20th century (Koomen et al., 2007). This makes the land cover change a very hot foci. The land cover change is very complex in that it's a combination of gradual modifications and sudden conversions. It's not only influenced by socioeconomic factors, but also by physical factors. It has already become a global issue instead of just a local issue. The advent of remote sensing technology greatly assists with the land cover change research, by providing land cover classification data and other physical variables. This paper focuses on the perspective of how socioeconomic variables influence the quantities of land cover change. Socioeconomic factors play a strong role in the land cover change, especially since the late 20th century, when global economic activity increased nearly sevenfold between 1950 and 2000, while global population doubled in roughly the past 40 years (Lambin and Geist, 2008). Analyses were conducted in this paper to reveal how proportions of different land cover types respond to variations of antecedent socioeconomic variables during the period of 1991 through 2013.

2. METHODOLOGY

2.1 Study Area

The Li River Basin is located in the northeastern Guangxi Zhuang Autonomous Region, China. Because the socioeconomic variables are based on counties, the study area adopted in this paper is the four counties in the Li River Basin, including Guilin County, Lingchuan County, Xingan County and Yangshuo County (Figure 1).
The study area of the Li River Basin is geographically complex, and its economy is primarily tourism-driven, which makes this area unique and worthwhile for uncovering what and how socioeconomic variables drive the land cover dynamics and its hidden mechanisms.

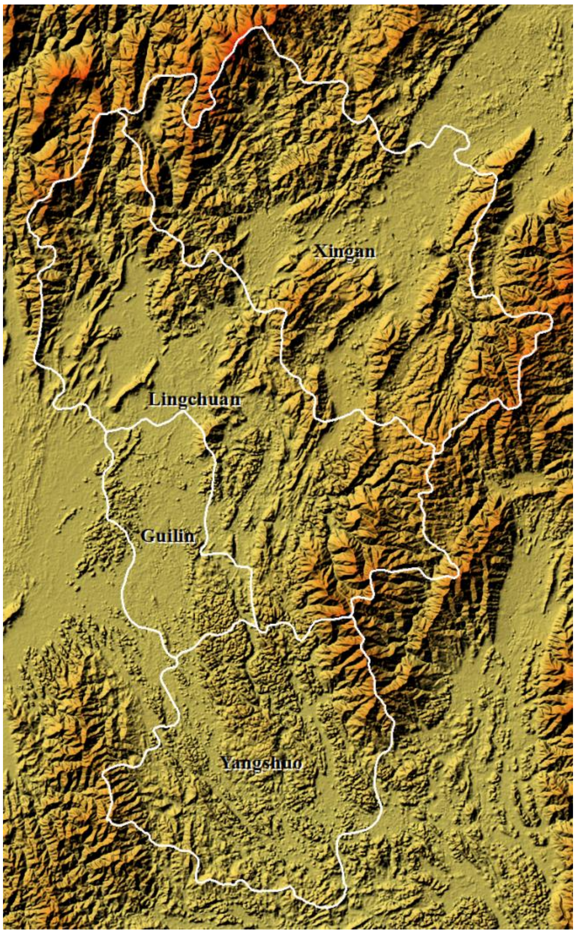

Figure 1 The study area of the Li River Basin in Guangxi Province, China

2.2 Data Sources and Preprocessing

The data used in this paper include the 30-meter Landsat land cover classification data, and the county-based socioeconomic data.

The Landsat images of the year 1991, 1996, 1999, 2002, 2009, and 2013 were classified into five classes: forest, water, urban, unclassified and clouds/shadows. The class clouds/shadows contains no useful information, and therefore regarded as no-data area. The rest classes form a four-class classification system, and this system was thereby consistently adopted for the classification of the six Landsat images in the paper. The object-oriented classification algorithm (implemented by eCognition) was used to obtain the land cover classification images.

The socioeconomic data were obtained from “The Socioeconomic Yearbook of Guilin District”, which includes the annual GDP, the GDP of the first, second and third sector, and population from 1988 to 2012.

2.3 Analysis

Firstly, the relationship was quantitatively assessed, using Pearson’s correlation coefficient and the stepwise regression model, between per-pixel lagged socioeconomic variables (SE variables) and the proportion of each land cover type. This helps to reveal how the quantity of land cover change is driven by socioeconomic development. The relationship between the per-pixel SE variables and the proportion of each land cover type was analyzed, instead of that between the more intuitive county-wide SE variables and area of each land cover type, to avoid the bias caused by variation in county areas.

The trends of land cover change were then analyzed. The cross correlation analysis was used to derive the instantaneous and lagged relationships between the above derived trends and different SE variables. The cross correlation analysis, which was applied on two time series to derive the lagged correlations between them (Venables and Ripley, 2002), was used in this paper to analyze the influence of SE variables on land cover proportions at different lags.

3. RESULTS AND DISCUSSIONS

3.1 Correlations Coefficients between SE Variables

The correlations between different SE variables must be analyzed before exploring how the set of SE variables, as a whole, influence land cover change to avoid the multicollinearity during modeling. The SE variables at the lag of one to three years were used in the analysis, and the pairwise correlations between SE variables at different lags are illustrated in Figure 2.

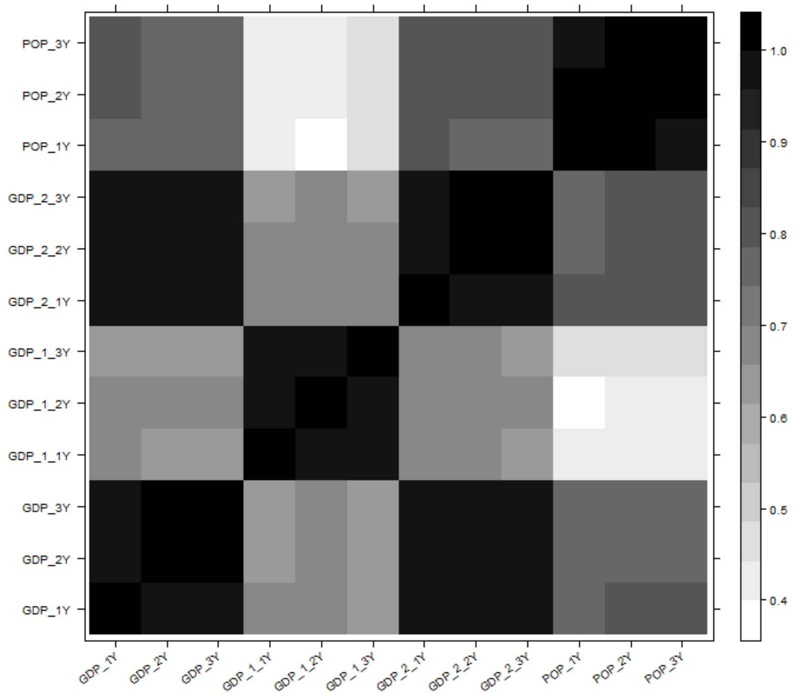

Figure 2 Pairwise correlations between socioeconomic variables (*abbreviations: VAR_xY indicates the corresponding variable VAR at the lag of x year)

Figure 2 indicates that: the same SE variable at different lags are extremely highly correlated (the black squares along the lower-left upper-right diagonal); the GDP variables and the GDP of the second sector variables are also extremely highly correlated; the POP variables are highly correlated with the GDP and the GDP of the second sector variables; the rest correlations are moderate ranging from 0.4 to 0.6.

3.2 Correlations between Socioeconomic Variables and Land Cover Proportions

Firstly considering the high correlation between the same SE variable at different lags, the pairwise correlation between the per-pixel SE variables at only the lag of the first year and the proportions of four land cover types were analyzed. The results are illustrated in Figure 3.

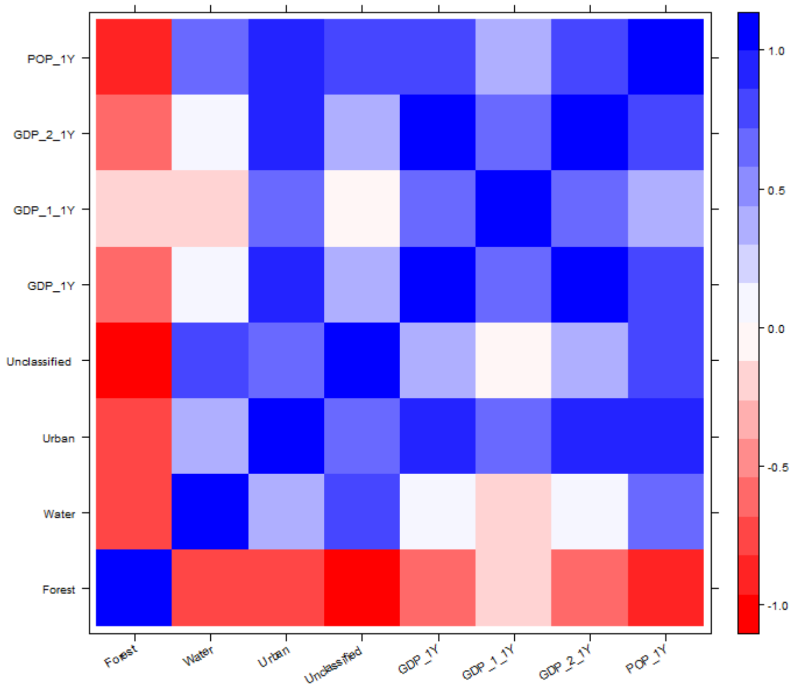

Figure 3 The pairwise correlation between the per-pixel socioeconomic variables at the lag of one year and proportions of four land cover types

The relationships as indicated in Figure 3 were analyzed respectively for four land cover types:

Forest proportion is very highly negatively correlated with the unclassified proportion, indicating the conversion of forest to unclassified land cover types (such as croplands) is a major form of deforestation. Forest proportion is moderately to highly negatively correlated with GDP, GDP of the second sector and population, with the only exception being GDP of the first sector. This suggests that the economic development, especially in the second sector, and the population increase can lead to the shrink of forest. Among the four variables, the correlation between forest proportion and population is the highest, reaching a correlation coefficient of greater than 0.9.

The water proportion, as compared with proportions of other land cover types, is much less correlated with socioeconomic variables. The only moderate correlation for water proportion is the one with population (about 0.59). This moderate positive correlation can be possibly attribute to human induced water and environment conservation or the tradition that human tends to live close to water.

The correlation between urban proportion and all SE variables are the highest, implying the strong influence of economic development and population increase on the urbanization. Its correlation with GDP, GDP of the second sector, and population are all greater than 0.9. Its correlation with GDP of the first sector is lower, which is about 0.66.

Unclassified proportion is moderately correlated with GDP and GDP of the second sector, and is highly correlated with population. The correlation between unclassified and GDP of the first sector is very low.

In sum, the correlation coefficients for forest proportion are all negative, while those for the proportions of other land cover types are all positive. This indicates that the economic growth (signified by the increase of GDP and population) leads to the forest shrink and the expansion of other land cover types.

3.3 Regression Models with Fixed Effects between SE Variables and Land Cover Proportions

A Variation inflation factor (VIF) and partial correlation based removal procedure was applied on the SE variables at different lags to reduce multicollinearity while retaining the variables that are more correlated with land cover proportions. The remaining variables after the removal procedure were all included in the regression model. The fixed effects, i.e. county effects that were able to explain the time-invariant variations among counties, were formated as four dummy variables, and were chosen to enter the regression model in a stepwise manner.

Four regression models were thereby developed for the four land cover types as described above. The SE variables and dummy variables that were included in the four models and their respective R squared values are shown in Table 1. The SE variables in Table 1 are ranked by importance quantified by beta weights (standardized coefficients).

Table 1 Variables and R squared of the four regression models

Model	SE and dummy variables	R squared
Forest	POP, GDP, County_3 (Yangshuo), and GDP_1	94.9%
Water	County_4 (Guilin), POP, County_2 (Xingan), GDP_1 and GDP	89.6%
Urban	POP, GDP_1 and GDP	99.1%
Unclassified	POP, GDP, GDP_1 and County_3	90.6%

County 3 (Yangshuo County) was included in the regression models built for forest and urban, indicating a slightly different response pattern of the proportions of both land cover types to the variations in GDP, GDP_1 and POP. This was possible due to the distinct GDP composition and its trend of Yangshuo County. The GDP of Yangshuo County was predominantly the GDP of the first sector in the early 1990s, and gradually developed into n third-sector dominant economy recently. The GDP of other counties are either first sector or second sector dominate.

3.4 Cross Correlations between SE variables and Land Cover Proportions

The cross correlation analysis was conducted between each of the per-pixel SE variables and proportion of each land cover type. Figure 4, as an example, shows the cross correlation between GDP and the proportions of four land cover types.

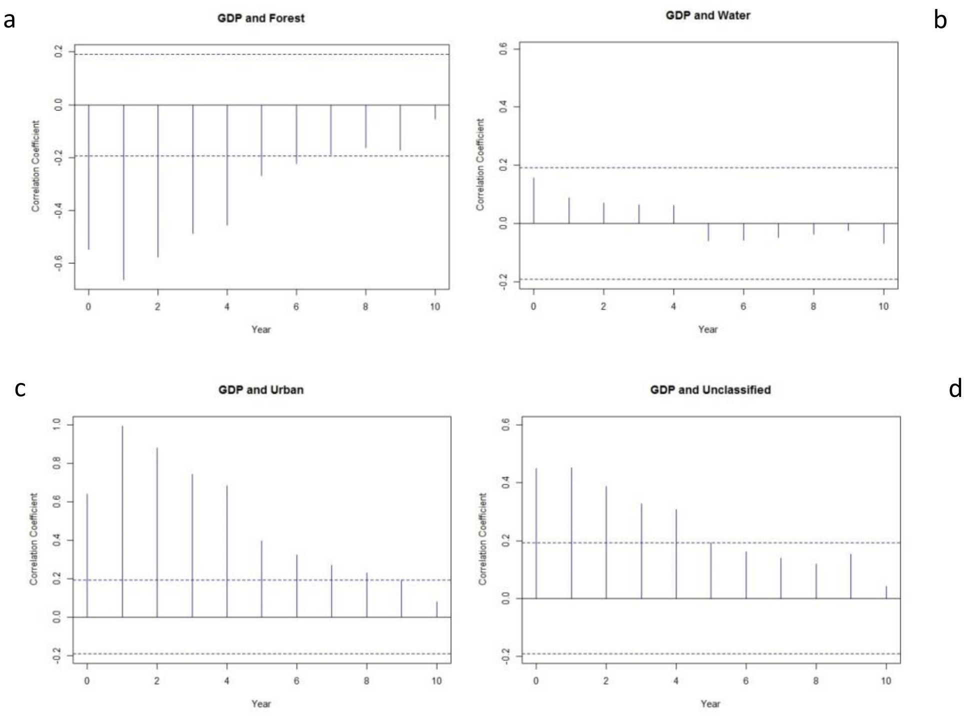

Figure 4 Cross correlation between the total GDP and proportions of forest (a), water (b), urban (c), and unclassified (d) (correlation coefficients beyond the dashed line are significant)

As indicated in Figure 4, GDP has no significant correlation with water; correlation dramatically decreases when lag is greater than four years for the other three land cover types (forest, urban and unclassified); GDP is highly correlated with urban proportions at the lag of one to four years (0.685 – 0.995), and the correlation decreases when lag increases.

For GDP of the first sector, it only significantly correlated with urban proportions, yet with a much lower coefficient as compared with the total GDP. Cross correlations between GDP_2 and land cover proportions are very much similar with those between GDP and land cover proportions, due to the extremely high correlation between GDP and GDP_2 (0.994). Correlation between urban proportion and lagged GDP_2, as compared with that between urban proportion and lagged GDP, is slightly lower at the lag of one year and higher at lags of two to four years, indicating a slower decreasing rate when lag increases. Correlations between urban proportion and lagged GDP_2 are generally higher. GDP_3 also shows a similar cross correlation as compared with the total GDP.

The land cover proportions respond much faster and stronger to population, than to GDP variables. Specifically, the proportions of forest, water and unclassified respond maximally to population without any lags, indicating the instantaneous influence of population. All land cover proportions are moderately to highly correlated with population. Water proportion are influenced more by population, and much less by GDP variables.

4. CONCLUSIONS

The relationships between the socioeconomic variables (including the total GDP, GDP of the three sectors and population) and proportions of four land cover types (forest, water, urban and unclassified) were analyzed by calculating the correlation coefficient, building the fixed-effects stepwise regression model and generate the cross correlation plots. It can be concluded that economic growth has a negative effect on forest expansion while has a positive effect on the expansions of other three land cover types; urban proportion are under the strongest influence of SE variables; water proportion are the most weakly correlated with SE variables, and it's much more correlated with population than with GDP variables; all land cover type proportions respond more strongly and rapidly to population than to GDP variables.

ACKNOWLEDGEMENT: This research is funded by National Science and Technology Pillar Program of China (Grant No. 2012BAC16B04).

REFERENCES

Koomen, E.,Stillwell, J., Bakema, A. and Scholten, H. J., 2007. Modelling Land-use Change Progress and Applications,Springer.

Lambin, E. F. and Geist, H. J., 2006. Land-use and Land-cover Change, Springer.

Venables, W. N. and Ripley, B. D., 2002. Modern Applied Statistics with S (Fourth Edition), Springer-Verlag.