Science of the Total Environment 646 (2019) 320–335

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

A Random Forest-Cellular Automata modelling approach to explore future land use/cover change in Attica (Greece), under different socio--economic realities and scales

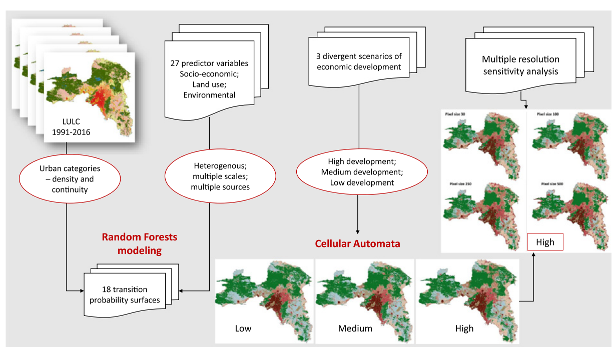
Dimitrios Gounaridis [a,*], Ioannis Chorianopoulos [a], Elias Symeonakis [b], Sotirios Koukoulas [a]

[a] Department of Geography, University of the Aegean, Mytilene, Greece
[b] School of Science and the Environment, Manchester Metropolitan University, United Kingdom

HIGHLIGHTS

- We explore the land use/cover dynamics in Attica region (Greece) under 3 scenarios.
- Scenarios reflect different economic performance realities and planning options.
- Heterogeneous factors expressed in multiple scales, units and resolutions are included.
- Transition potential with RF and CA modelling are used to project the LUC changes.
- Simulation results are subject to a multi-resolution sensitivity analysis approach.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 6 June 2018
Received in revised form 21 July 2018
Accepted 21 July 2018
Available online 23 July 2018

Editor: Ouyang Wei

Keywords:
Land use/cover change
Random forests
Cellular automata
Simulation
Development scenarios
Land use planning

ABSTRACT

This paper explores potential future land use/cover (LUC) dynamics in the Attica region, Greece, under three distinct economic performance scenarios. During the last decades, Attica underwent a significant and predominantly unregulated process of urban growth, due to a substantial increase in housing demand coupled with limited land use planning controls. However, the recent financial crisis affected urban growth trends considerably. This paper uses the observed LUC trends between 1991 and 2016 to sketch three divergent future scenarios of economic development. The observed LUC trends are then analysed using 27 dynamic, biophysical, socio-economic, terrain and proximity-based factors, to generate transition potential maps, implementing a Random Forests (RF) regression modelling approach. Scenarios are projected to 2040 by implementing a spatially explicit Cellular Automata (CA) model. The resulting maps are subjected to a multiple resolution sensitivity analysis to assess the effect of spatial resolution of the input data to the model outputs. Findings show that, under the current setting of an underdeveloped land use planning apparatus, a long-term scenario of high economic growth will increase built-up surfaces in the region by almost 24%, accompanied by a notable decrease in natural areas and cropland. Interestingly, in the case that the currently negative economic growth rates persist, artificial surfaces in the region are still expected to increase by approximately 7.5% by 2040.

1. Introduction

Land use and land cover (LULC) changes are considered to be the most prominent influence of humans on the environment.

* Corresponding author at: SAGISRS Lab, Department of Geography, University of the Aegean, Lofos Panepistimiou, Mytilene 81100, Lesvos, Greece.
E-mail address: gounaridis.d@geo.aegean.gr (D. Gounaridis).

https://doi.org/10.1016/j.scitotenv.2018.07.302

Technological and medical advancements have brought about unprecedented increases in the human population and, consequently, in the need for access to resources. This need has in turn caused substantial and growing transformations to the Earth's surface (Vitousek et al., 1997) with often undesirable impacts and magnitudes that vary from local to global scales. The dual role of humans to actively contribute to LULC changes and, at the same time, be on the receiving end of experiencing the consequences of these changes, emphasizes the need for a better understanding of the human-LULC change nexus.

A wide variety of LULC change models have been developed to meet the scientific community needs for understanding how and why LULC evolves (Schrojenstein Lantman et al., 2011). Generally, LULC models are widely used to analyze the complex structure of linkages and feedbacks between drivers of change, determine their relevance to particular changes and project how much land is used where and for what purpose, under different predefined attributes and conditions. This type of information is then adopted in a meaningful way in order to support policy decision making related to land-use (Mallampalli et al., 2016). However, by definition, LULC models can not exactly replicate the complex interactions and nonlinear relations which are apparent in LULC systems. At a fundamental level, they are, rather, a process that provides a platform that, allows computer experiments to be undertaken (Brown et al., 2013). When the system in question is simple, the processes and interactions that characterize it can be easily determined and the results are somehow expected, while projections and other kinds of extrapolations are a straightforward task. When dealing, however, with inherently complex systems, as is the case with LULC changes, the models are able to represent and exemplify only a small fraction of the whole mechanism in order to highlight important processes.

The recent methodological and technological advancements have paved the way for more articulated LULC models which are able to answer more complex questions. Such questions could be in regard to what the possible outcomes would be if alternative pathways were followed, or which outcome is the most desirable from a list of alternatives, as well as a diverse range of other 'what-if' scenarios. Increasingly, scenario-based analysis is now being adopted by a range of disciplines pertaining to LULC change, as fruitful experiments for exploring the possible future trajectories of historical and current trends (Murray-Rust et al., 2013). Considering that the number of potential futures is actually infinite (Greeuw et al., 2000), scenarios are not used to predict the future in a precise manner, but to explore possible future directions and to consider a range of alternative pathways. To do so, the scenario-based analysis fully recognizes the infinity of potential futures and attempts to focus only to an understandable and manageable set of alternatives. This is achieved by delineating plausible, presumably coherent and internally consistent storylines of different socio-economic development trajectories (Rounsevell and Metzger, 2010).

When modelling LULC, the scale, the spatial resolution and the extent of the study area are important attributes of all spatially explicit models (Agarwal et al., 2002). The term scale refers to the spatial, temporal, quantitative, or analytic dimension used to measure and study the processes that are modelled (Gibson et al., 2000). Scale also involves the terms extent and resolution: extent refers to the magnitude of a dimension used in measuring (e.g. study area boundaries on a map), whereas resolution refers to the precision used in this measurement (e.g. pixel size) (Gibson et al., 2000). Moreover, resolution refers not only to spatial resolution, but also to thematic, which is the level of precision in LULC categories. In addition, the term temporal resolution is used to refer to the time span and frequency of the analysis. Modelling LULC changes, therefore, requires a range of scales to be defined since it is a phenomenon that involves multiple processes that act over different scales. At each scale, different processes have a dominant influence on the outcome (Meentemeyer, 1989; Van Delden et al., 2011). Approaches that do not consider the various scales involved in LULC changes, are prone to aggregation or oversimplification errors and thus fail to reproduce the dominant cross-scale interactions. This is due to the fact that features and processes that operate at local scales are not always observable when dealing with larger areas and coarser spatial resolution data (Verburg et al., 2004). On the other hand, studies that focus solely on the local level often fail to incorporate information about the general context which can only be derived from coarser spatial resolution data (Larondelle and Lauf, 2016). Given that all models are driven by their input data, studies focusing on specific LULC processes, considering only a single scale and using data that are particularly suitable only to a certain area, are not representative, transferable or reproducible to different scales. Therefore, such approaches are characterized by higher levels of uncertainty and depend on a number of critical assumptions (Kok and Veldkamp, 2001; Van Delden et al., 2011; Veldkamp et al., 2001; Verburg et al., 2006). Moreover, it is a common assumption that the modelling results are highly affected by the quality and the technical details, such as the pixel size of inputs and the bias they entail (Kocabas and Dragicevic, 2006; Van Delden et al., 2011).

Models designed to analyze LULC dynamics can be divided into categories according to their perspective, their domain, the methodological framework they apply, their spatial or non-spatial nature etc. (literature reviews by Agarwal et al., 2002, Briassoulis, 2000; Schrojenstein Lantman et al., 2011). A simple, non-exhaustive, classification with regard to their methodological origins would include i) Empirical–statistical models using multivariate regression and geostatistical analysis (e.g. He and Lo, 2007; Poelmans and Van Rompaey, 2010). ii) Stochastic and optimization models (e.g. Brown et al., 2002), which consider one objective or simply convent multiple objectives into one and the optimization takes place with the use of weighting methods (Ma and Zhao, 2017). iii) Dynamic process-based simulation (Veldkamp and Fresco, 1996; Verburg et al., 2002) which often involve multiple models subdivided in modules that capture non-spatial (e.g demand) and spatially explicit (e.g allocation) processes. iv) Agent-based modelling (e.g. Manson, 2005; Robinson et al., 2012), which simulate the actions and interactions of autonomous agents involved in LULC change.

However, LULC models that solely rely on statistical approaches often suffer from limitations such as sensitivity to outliers and noise, collinearity issues and factors compatibility (Dormann et al., 2013; Eastman et al., 2005). On the other hand, a variety of models pertaining to artificial intelligence, such as agent-based models, have been successfully applied for addressing the complex, non-linear behavior of human-nature interactions and decision making. This type of models, however, are suitable to capture processes at the individual, household or neighborhood levels and when it comes to agent behavior they can be very complex and are often parametrized with qualitative social survey data and other types of participatory approaches (Zagaria et al., 2017).

For LULC scenario-based simulation modelling, a growing body of the literature employs Cellular automata (CA) which consist of a dynamic simulation framework where space is represented as a grid of cells and time is considered as discrete unit. The basic principle of CA is that the state of a given pixel is determined by taking into account its previous state, the spatial interactions with the surroundings in a given neighborhood and a set of defined transition rules. These elements dictate the possible change of a cell and can be expert-based or calculated from statistical analysis of historical LULC changes (White and Engelen, 2000). CA models, although very simple, have the strong ability to represent rich LULC patterns and handle nonlinear, stochastic and spatially explicit LULC processes (Sante et al., 2010).

The biggest advantage of CA is that they are fully consistent with Geographic Information Systems (GIS) and remote sensing. Additionally, CA can be coupled with other types of models and thus they are flexible to allow the elaboration and extension of the methodological procedures according to the needs of a case study (Aburas et al., 2016). For instance, CA have been previously combined with a plethora of modelling frameworks such as Markov chains (Arsanjani et al., 2013),