Computers, Environment and Urban Systems 50 (2015) 53–65

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Enhancing the calibration of an urban growth model using a memetic algorithm

William Veerbeek [a,b,*], Assela Pathirana [a], Richard Ashley [a,c], Chris Zevenbergen [a,b]

[a] Flood Resilience Group, Dep. Water Science and Engineering, UNESCO-IHE Institute for Water Education, Westvest 7, 2611 AX Delft, Netherlands
[b] Hydraulic Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, Netherlands
[c] Pennine Water Group, Department of Civil and Structural Engineering, University of Sheffield, Sir Frederick Mappin Building, Mappin Street, Sheffield S1 3JD, United Kingdom

ARTICLE INFO

Article history:
Received 28 February 2013
Received in revised form 6 November 2014
Accepted 6 November 2014

This paper is dedicated to Prof. Erik Pasche† who contributed to, supported and continues to inspire the ongoing research work in CORFU.

Keywords:
Urban growth
Calibration
Optimization
Memetic algorithm
Beijing
Cellular automata

ABSTRACT

At present, many approaches and models have been developed to perform spatially explicit simulations that mimic observed land use and land cover changes (LULC) for a given area. Calibration of such models is often performed using comparatively standard 'off-the-shelf' machine-learning algorithms that are not necessarily suited to perform effectively within the model's implementation. This method becomes problematic when the computational costs of applying an evaluation function to determine the goodness-of-fit are high; calibration using 'standard' algorithms often requires many iterations to achieve satisfactory outcomes. Furthermore, in some cases, future LULC projections manifest significant changes in trends, particularly when increasing the number of LULC classes in the simulation and the number of associated transition rules. This study presents an adapted machine-learning algorithm to optimize a parameter set applied in a Dinamica-EGO-based LULC change model. A sequentially applied memetic algorithm is applied to optimize a vast parameter set by extending a genetic algorithm with a local search function. To achieve consistent long-term projections, a 2-stage approach is applied in which the expansion of the urban extent and diversification of urban LULC classes are calculated sequentially. The outcomes repeatedly show a much faster convergence toward a high goodness-of-fit; significantly fewer iterations and a smaller population size can be used to attain a similar performance level than when using a standard GA-enhanced calibration. Furthermore, the observed spatial trends are maintained for long-term projections using 5-year intervals. In the current application, the model is applied to the rapidly growing metropolitan area of Beijing, China.

1. Introduction

Originally used in regional economics (e.g., Allen, 1954; Alonso, 1964; Forrester, 1969), land use change models, particularly urban growth models, attempt to mimic historical and future LULC transitions in a spatially explicit manner. Depending on the selected representation of the spatial components, cells or patches represent discrete LULC classes (i.e., states) that can change over time. These changes are affected by a set of drivers that are conceptualized as transition rules. Contemporary urban growth models are often based on cellular automata (CA) models (e.g., Batty & Xie, 1994; Clarke, Hoppen, & Gayos, 1997; Li & Yeh, 2000; White & Engelen, 1993) that describe cell-based LULC transitions as a function of local interactions, which represent neighboring conditions that drive the formation of spatial urban patterns. Often, these models are combined with 'top-down'-driven transition rules that incorporate fixed physical properties (e.g., slope or elevation) and/or statistically determined growth drivers (e.g., population growth or economic development). Although most models were essentially developed as generic models capable of representing the growth dynamics of any metropolitan area, they can be adjusted to mimic LULC transformations in specific cities or regions. The model calibration and validation stages can be performed manually, but they are frequently automated using historical LULC data as a training set (e.g., Li & Yeh, 2002). During calibration, the relation between the predicted and observed LULC, by using a set of metrics determining the goodness-of-fit (e.g., Næsset, 1995), are combined with an update function that changes the transition rules. When an optimal correlation is found (i.e., no significant improvement in the goodness-of-fit can be obtained), the growth transition rules are applied to prospective years to obtain the projections. This process is depicted in Fig. 1.

* Corresponding author at: Flood Resilience Group, Dep. Water Science and Engineering, UNESCO-IHE Institute for Water Education, Westvest 7, 2611 AX, Delft, Netherlands. Tel.: +31 15 2151844.
E-mail address: w.veerbeek@unesco-ihe.org (W. Veerbeek).

http://dx.doi.org/10.1016/j.compenvurbsys.2014.11.003
0198-9715/

Machine-learning algorithms or other regression methods are frequently used to calibrate LULC change models. Long, Mao, and Dang (2009), as well as many other authors (e.g., Hu & Lo, 2007; Liu & Phinn, 2003), used logistic regression to optimize transition rules. Li and Yeh (2004) applied an artificial neural network to optimize parameters, while Yang, Li, and Shi (2008) used a support vector machine. Recent applications include particle swarm optimization methods (Feng, Liu, Tong, Liu, & Deng, 2011; Rabbani, Aghababaee, & Rajabi, 2012) or ensemble learning strategies in which several machine learning algorithms are executed in parallel (Gong, Tang, & Thill, 2012). All of these stochastic methods are commonly setup as iterative processes that require multiple model runs to obtain convergence to satisfactory solutions. Such approaches work reasonably well when the computational costs of running a calibration and validation sequence are relatively low. For example, when a LULC transition for single cells is calculated, the result is evaluated and an adjustment is made to one of the controlling parameters; thus, the computational requirements are minimal. Yet, some models rely on both local and global comparisons of the LULC change maps for evaluations. Particularly large areas composed of millions of cells are computationally very costly, resulting in a long-duration calibration. While for instance GPU-accelerated calibration (Blecic, Cecchini, & Trunfio, 2014) could cope with the increasing demand for computational power, the underlying methods do not fundamentally change. In addition, while calibration using machine-learning algorithms can produce LULC changes that mimic observed transitions, overfitting the parameters might lead to variable future projections. Thus, the observed spatial development trends in historical data are discontinued when running many subsequent iterations of the calibrated LULC change model.

By building upon a Dinamica-EGO-based LULC change model (Filho, Corradi, Cerqueira, & Araújo, 2003; Filho, Rodrigues, & Costa, 2009), a 2-stage modeling approach is introduced to separate the calculation of the urban-area growth from the diversification of the growth extent into urban LULC classes. This method ensures the production of consistent LULC patterns over long periods. The model is equipped with a customized automatic calibration method based on a genetic algorithm (GA). The GA is extended with a local search function, which significantly reduces the required number of candidate solutions and iterations to produce robust and accurate results. This approach provides an alternative for the often used ‘off-the-shelf’ machine-learning algorithms used in LULC change models. To test the outcomes, the model is initially applied to the Beijing metropolitan area, which is an ideal case study due to the combination of market-driven rapid urban expansion and top-down planning policies (Han, Lai, Dang, Tan, & Wu, 2009). The model is required to adjust to alternative urban development trends that might not simply evolve near the current urban clusters. Furthermore, the relatively large case study area, combined with the applied 30-m spatial resolution, could require substantial computations.

In the first section of the paper, the case study and the Dinamica EGO model are introduced, including a detailed description of the methods used to define transition rules and the evaluation criteria used to estimate the goodness-of-fit of the produced LULC changes. The second part provides the background and context for the development of the 2-stage approach, as well as the GA extended calibration. Subsequently, the outcomes are presented. A comprehensive analysis should provide sufficient evidence for the robustness of the observations and interpretations. Finally, a brief discussion of the underlying assumptions and ongoing issues is presented.

2. The case study

2.1. Beijing

The case study used to test the model is greater Beijing, China. To an extent, Beijing’s urban development is typical for an Asian megacity; since the late 1980s, the city has undergone massive expansion and redevelopment which has doubled the size of the urban extent in the last 15 years. Over 1995–2005, this expansion comprised 19% infill, 75% extension and 6% leapfrogging development (Fig. 2). Although greater Beijing is surrounded by mountains to the north and west, the city’s potential for development is relatively unconstrained. However, Beijing’s future growth is not unlimited; the absence of freshwater bodies (Jiang, 2009) and the increasing traffic congestion (Zhao, 2010) are likely to limit Beijing’s expansion in the long term. Despite the large degree of freedom for development, Beijing remains relatively compact. The majority of the urban extent is contiguous and expanded from the 15th century “Forbidden City”, which forms the geographic center of the city. Nevertheless, Beijing is facing significant expansion due to urban sprawl (Zhao, 2010), which has emerged over the last decade.

In 2005, the Beijing metropolitan area housed approximately 15 million inhabitants (Beijing Statistics Bureau, 2005), which had been expected to increase to 18 million by 2020. However, the current population has already surpassed 19.6 million (National Statistics Bureau, 2011), and the Beijing Academy of Social Sciences recently revised their estimate to 26 million by 2020 (Caixin Online, 2012). These discrepancies show that there is no real consensus regarding the population growth in Beijing and that future containment might be difficult to achieve through policy and urban growth constraints. In contrast to many other rapidly developing megacities, the urban development of Beijing is being facilitated through a succession of regional development

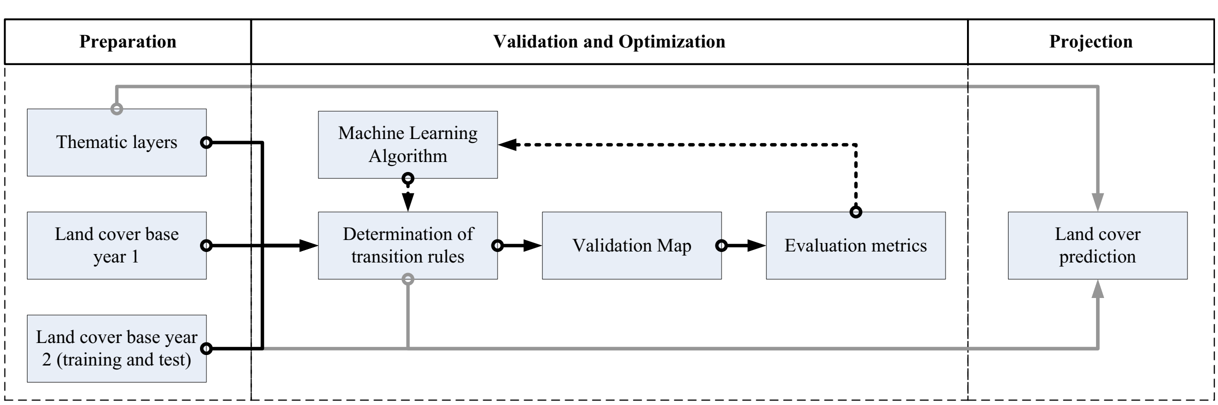

Fig. 1. Typical setup for a LULC change model, including the feedback mechanism for the calibration.