A support vector machine approach to model urban growth in the greater Tirana region, Albania

Lakes, Tobia[1]; John, Isabelle[1]; Müller, Daniel[1, 2]; Krüger, Carsten[1]; Rabe, Andreas[1]

[1] Geomatics Lab, Geography Department, Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
[2] Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany

Introduction

Spatially explicit land use modelling can give new insights on past patterns and processes and may provide information on likely future developments to support decision-making. This is of particular importance in regions where rapid and frequently unregulated urban development as well as short-term changes in demography and migration patterns have led to a dramatic increase of urban area and population such as in the greater Tirana region, Albania. Several modelling techniques have been developed within recent years such as cellular automata, agent-based, system dynamics, neural networks, and logistic regressions (Verburg et al. 2006; Lakes et al. 2009). Among these approaches machine learning techniques seem promising particularly in studies with little knowledge about the land use change process under investigation and data problems regarding quality, availability and comparability. Support vector machines (SVM) are a supervised learning approach for classification or regression of a multidimensional feature space and have been successfully applied for classification of remote sensing data (van der Linden et al. 2008; Burges 1998). The aim of this paper is to test a support vector machine approach for modelling urban land dynamics in the greater Tirana region, Albania.

Material and Methods

We apply the support vector machine imageSVM 2.0.1 implemented in ENVI® to model urban growth in the greater Tirana region, Albania. Land Use data is derived from a binary classification (urban/non-urban) of Landsat data for 1988, 2001 and 2008. A layer on urban growth is then calculated for the periods 1988-2000, 2001-2008 and 1988-2008. Drivers of change are represented by environmental data on elevation, slope, water bodies, road infrastructure, as well as by socio-economic data on population figures and further processed to calculate neighbourhood and density measures. We calibrate the SVM with the urban growth data for the years 1988-2000, 2000-2008 and 1988-2008 using random samples of 1000 per class for each data set. We apply the calibrated models to the respective whole data sets and calculate several statistical accuracy measures. The model runs are validated using a separate data set of again 1000 samples in each class for the respective time periods. To reveal the significance of the integrated drivers we apply a sequential feature forward selection process.

Results and Discussion

Modeling results of urban growth reflect the spatial patterns of urban growth in the investigated region for all 3 time periods. We achieve Kappa values of about 0.6 for all modeling runs, however, due to data shortcomings the modeling results vary slightly between the different time periods. Results suggest that urban land use growth in the greater Tirana region can be modeled in a spatially explicit way using a SVM approach. Modifications in rate and location of change in the investigated time periods was pictured in the modeling results. Modeling results accuracies vary slightly according to the number of driving factors included as features in the SVM to explain the urban growth. The sequential feature forward selection process reveals that driver information derived from population

figures that were implemented with the intention to explain a majority of urban growth have no significant effect on the model results. Instead, it is the information about distance to existing urban agglomerations and the distance to roads which explains most of the urban growth. In this study SVM has offered benefits for overcoming shortcomings in the data sets and exploiting the available information (van der Linden et al. 2009), however, at the same time new knowledge and inferences on the process of urban growth are limited by the data-driven approach.

Conclusions and outlook

We applied a support vector machine approach to model rapid and largely unregulated urban growth in Tirana region, Albania. Results suggest that the urban growth process is driven predominately by the distance to existing urban agglomerations and the distance to roads. The derived information on regional land use change may offer new insights into the predominately unregulated process of urban growth and may provide insights into likely future development (Verburg et al. 2006). Further research on SVM modeling will focus on integrating machine learning approaches with sophisticated modeling approaches to exploit the benefits of each respective modeling technique for modeling likely future urban land use change in a spatially explicit way.

BIBLIOGRAPHY

Burges, C.J.C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, 2, pp. 121-167.

Lakes, T., Mueller, D., Krueger, C., 2009. Cropland change in southern Romania: a comparison of logistic regressions and artificial neural networks. Landscape Ecology, pp. 1195-1206

van der Linden, S., Janz, A., Waske, B., Eiden, M. and Hostert, P., 2007. Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines. Journal of Applied Remote Sensing,1, pp. 1-16.

Verburg, P.H., Kok, K., Pontius Jr., R.G. and Veldkamp, A., 2006. Modeling land-use and land-cover change. In: E.F. Lambin (Editor), Land-use and land-cover change: local processes and global impacts Springer, Berlin, pp. 117-135.