Landscape Ecol (2009) 24:1195–1206
DOI 10.1007/s10980-009-9404-2

RESEARCH PAPER

Cropland change in southern Romania: a comparison of logistic regressions and artificial neural networks

Tobia Lakes · Daniel Müller · Carsten Krüger

Received: 30 September 2008 / Accepted: 18 August 2009 / Published online: 17 September 2009

Abstract Changes in cropland have been the dominating land use changes in Central and Eastern Europe, with cropland abandonment frequently exceeding cropland expansion. However, surprisingly little is known about the rates, spatial patterns, and determinants of cropland change in Eastern Europe. We study cropland changes between 1995 and 2005 in Argeş County in Southern Romania with two distinct modeling techniques. We apply and compare spatially explicit logistic regressions with artificial neural networks (ANN) using an integrated socioeconomic and environmental dataset. The logistic regressions allow identifying the determinants of cropland changes, but cannot deal with non-linear and complex functional relationships nor with collinearity between variables. ANNs relax some of these rigorous assumptions inherent in conventional statistical modeling, but likewise have drawbacks such as the unknown contribution of the parameters to the outcome of interest. We compare the outcomes of both modeling techniques quantitatively using several goodness-of-fit statistics. The resulting spatial predictions serve to delineate hotspots of change that indicate areas that are under more eminent threat of future abandonment. The two modeling techniques address two controversial issues of concern for land-change scientists: (1) to identify the spatial determinants that conditioned the observed changes and (2) to deal with complex functional relationships between influencing variables and land use processes. The spatially explicit insights into patterns of cropland change and in particular into hotspots of change derived from multiple methods provide useful information for decision-makers.

Keywords Land use change · Human-environment system · Spatial analysis · Logistic regression · Neural network · Eastern Europe · Romania

T. Lakes (✉) · D. Müller · C. Krüger
Geomatics Lab, Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
e-mail: Tobia.Lakes@geo.hu-berlin.de

D. Müller
e-mail: mueller@iamo.de

C. Krüger
e-mail: Carsten.Krueger@geo.hu-berlin.de

D. Müller
Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany

Introduction

Massive changes in land use have taken place following the collapse of the socialist system in Central and Eastern Europe, as well as in the Commonwealth of Independent States (Bicik et al. 2001; Lerman et al. 2004; Peterson and Aunap 1998). These land use changes have been dominated by large-scale cropland

and pasture abandonment (Kuemmerle et al. 2008; Kuemmerle et al. 2006; Müller and Munroe 2008; Palang et al. 2006). At the same time, some areas have experienced agricultural expansion (Verburg et al. 2006). Eastern Europe experienced an intensive broad-scale shock with the transformation from centrally-planned to market-oriented economies beginning in the early 1990s. The resulting transition process instigated widespread alterations in agricultural land use in Eastern Europe that were caused by (but are not limited to) the collapse of state support for agricultural production, changing ownership structures, the emergence of additional income opportunities, benefits from EU agricultural policies and a new geographic mobility, which resulted in massive emigration from rural areas (Elbakidze and Angelstam 2007; Macours and Swinnen 2008; Müller and Munroe 2008).

While the bearing of these broad-scale factors is palpable, little is known about the rates, spatial patterns and underlying causes of agricultural land use dynamics in Eastern Europe. Few consistent and spatially explicit assessments have considered small-scale determinants and patterns of cropland changes in the dynamic landscapes of Eastern Europe. Romania, the second largest new member state of the EU, is particularly interesting as it holds the largest share of farmland (65%) and labor force employed in agriculture (32%) among the new member states (Müller et al. 2009). In this study we focus on cropland, which refers to all areas cultivated with annual or perennial crops, but excludes managed and unmanaged grasslands. Cropland use is not only important due to its socioeconomic relevance, but also carries considerable ecological importance. Romania is characterized by high endemic biodiversity connected to agricultural areas and to the Carpathian mountain range, Europe's largest temperate forest ecosystem (Cremene et al. 2005; Ioras 2003; Oszlanyi et al. 2004).

Spatially explicit modeling approaches have been successfully applied to investigate land use change processes. Modeling methods include statistical regressions (Chomitz and Gray 1996), artificial neural networks (Pijanowski et al. 2005), or fuzzy modeling (Wieland and Mirschel 2008), among others. In this paper, we apply and compare two techniques to model cropland change in southern Romania. First, we use spatially explicit logistic regressions to identify the spatial determinants of cropland change. Logistic regressions are a statistical technique to determine the significance, direction and strength of an independent variable (covariate) on a dichotomous dependent variable (Hosmer and Lemeshow 2000). Logistic regressions have been successfully applied to analyze the causes of deforestation (e.g. Chomitz and Gray 1996; Mertens and Lambin 2000) and cropland abandonment (Müller and Munroe 2008). Second, we calibrate artificial neural networks (ANN), which have attracted growing interest in land change science in recent years (Almeida et al. 2008; Müller and Mburu 2009; Pijanowski et al. 2002). Artificial neural networks are a pattern recognition tool that can describe the complex and non-linear character of environmental processes (May et al. 2008; Salazar-Ruiz et al. 2008; Wieland and Mirschel 2008). As such, they attempt to achieve the highest fit for data sets without explicitly identifying the functional relationships between input and output variables. Artificial neural networks are of particular importance in studies with data shortcomings (Fischer and Abrahart 2000; Openshaw 1998).

The aim of our study is to use the modeling techniques to investigate the spatial patterns and the location of hotspots of cropland change in Argeş, Romania. We accomplish this by comparing the insights gained from logistic regression analysis and artificial neural networks modeling. Cropland changes are categorized into two mutually exclusive land cover conversions: cropland abandonment and cropland expansion. To analyze changes in cropland, we employ an integrated spatial dataset consisting of land cover maps, information derived from a socio-economic census, geo-biophysical variables and accessibility measures. An improved understanding of past changes and their determinants can support policy and management decisions. The resulting insights may be used as input into spatially targeted land use planning activities in rural Romania.

Materials and methods

Study area

Our study area is Argeş County, situated on the southern foothills of the Carpathian mountain range (Fig. 1). The county covers 6,826 km^2 and is subdivided into 99 communes with an average of 70 km^2

per commune. Argeş encompasses a wide range of environmental conditions, with a mean annual rainfall between 550 mm in the southern plains and 1,100 mm in the northern mountains, and a mean annual temperature between one and nine degrees Celsius (data from the Romanian National Institute for Meteorology, Hydrology, and Water Management; see Kuemmerle et al. (2009) for a more detailed description of the study area).

Elevations in the county range from 100 to more than 2,500 m above sea level. The mountainous north is characterized by rugged terrain and includes the foothills of the Carpathian mountain range, as well as Romania's highest mountain, the Moldoveanu Peak, at 2,544 m. The hilly midlands contain the county's capital and major market center, Piteşti, which in 2003 accounted for 27% of the county's population, or 174,000 of a total of 650,000 people (NIS 2004). Romania's first highway, built in 1960, links Piteşti with the capital Bucharest in the southeast of the country. The southern part of Argeş consists of a plain between 100 and 150 m above sea level.

Rural land cover consists of 39% forest, 23% grass and shrubland, and 28% cropland (Kuemmerle et al. 2009). Cropland production is dominated by wheat and other grains as well as maize. Land cover is closely linked to the North–South gradient. Cropland is the dominant land cover in the southern plains and the northern parts are predominately covered by forests and mountain meadows. In the hilly areas of central Argeş there is a close mix of heterogeneous agricultural land uses and urban settlements. Cropland prevails in the valley bottoms of the hilly areas, but decreases as one moves away from the road network (Kuemmerle et al. 2009). During socialist times, large parts of the agricultural land outside the cities had been collectivized. With the collapse of the political system new ownership structures developed throughout the county.

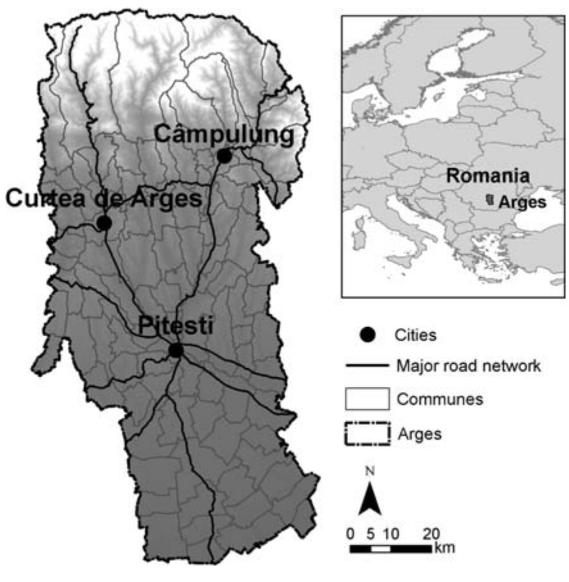

Fig. 1 Study area: Argeş County, Romania

Data

Land cover data was derived from Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) images for 1995 and 2005. A hybrid classification approach was used to identify land cover changes. The cropland class has a high producer accuracy of 92% (90%) with a conditional Kappa of 0.93 (0.92) in 1995 (2005) (see Kuemmerle et al. (2009) for details of the classification methods). We calculated cropland abandonment as those locations covered with cropland in 1995 but not in 2005. Cropland expansion equals one for all cells that were not cropland in 1995 but were converted into cropland by 2005.

We derived elevation, slope and the range of the slope in a three-by-three pixel window from the Shuttle Radar Topographic Mission (SRTM, Slater et al. 2006). The topographic data also serves as a proxy for climatic variation and for soil data, as both variables are strongly correlated due to the north–south gradient in the study area. Unlike Müller et al. (2009), we used official census data provided at the commune level by the National Institute of Statistics of Romania in 2003 and 1996. This data covers all of Argeş, communes, including municipalities and towns. We derived socioeconomic determinants of cropland change from this commune census and merged the commune with the raster data based on the communes' administrative boundaries. Population, the number of people employed in education, the number of people employed in agriculture, crop yield, and the number of livestock in each commune were derived from this census. We calculated livestock units using the livestock unit coefficients for countries belonging to the Organisation for Economic Co-operation and Development (OECD) (Chilonda and Otte 2006). A self-conducted census of all rural communes in Argeş augmented the dataset with village-level information on population (Müller et al. 2009). Spatially explicit population density maps were created by interpolating the villages' population data using inverse distance weighting.

Several measures captured the accessibility of each location: We estimated the transportation costs from every location in the study area to each village centroid, as well as to Piteşti and to the built-up area derived from the CORINE land cover. We also calculated the transportation costs to the road network. Cropland density was included as an additional driver, calculated as the number of pixels covered by cropland at the start of the change period (1995) in a three-by-three pixel neighborhood. We further included the density of grassland, forest, and settlements in a three-by-three neighborhood. Finally, we included the area of each commune in square kilometers to control for variations in the size of the communes. All data was stored in raster format, referenced to UTM/WGS 84 and resampled to a spatial resolution of 100 m.

Methods

Logistic regression

Logistic regressions are the dominant statistical tool for describing linear relationships between a dichotomous outcome variable and a set of explanatory variables. In the terminology of statistical pattern recognition, a logistic regression is a single-layer network where the coefficients are the parameters of the linearly combined input variables that are transformed using a non-linear, logistic activation function (Bishop 1995). We estimated reduced-form, spatially explicit logistic regressions to extract the direction and strength of the influences of underlying factors that led to a change in the extent of cropland from 1995 to 2005.

We selected the set of explanatory variables based on *a priori* knowledge of the study area that we gained during several months of an intensive field campaign during summer and autumn 2005. We also conducted expert interviews during this field season as well as a thorough literature review. Variables that were hypothesized to influence cropland change were entered into the regressions in order to test their statistical significance and the direction of their influence. To control for endogeneity issues, we only selected predetermined variables that express the state at the start of the change period. In this way, we assumed that the state of a variable has a bearing on subsequent land change (Perz and Skole 2003). While this may not totally eliminate endogeneity (Gellrich et al. 2007), it seems reasonable to assume that subsequent cropland changes are a response to the conditions at the beginning of the change period. The descriptive statistics of the resulting final sample are listed in Table 1 for both change processes.

We devised one regression model for cropland abandonment and one for cropland expansion. Both regression models followed identical sampling procedures. First, we regularly sampled locations 500 m away from the closest selected neighbor (Besag 1974) to reduce spatial autocorrelation in the dependent variable. This resulted in 27,284 observations, from which we selected all cells that were covered by cropland in 1995 for the abandonment model and all cells not covered by cropland for the cropland expansion model. In this way, we incorporated the temporal dependencies of the two processes on the presence or absence of cropland at the start of the change period (e.g. cells not covered by cropland could not become abandoned cells). We then conducted unequal sampling and randomly selected 1,000 observations from the respective change class (i.e., from pixels that were abandoned or expanded) and from the pixels that did not undergo change (i.e., stable cropland or stable non-cropland). The logistic regressions were estimated based on these 2,000 observations. We also removed variables that had pair-wise correlation coefficients higher than 0.7 to avoid high multicollinearity among the explanatory variables. These included slope, transportation costs to Piteşti and the commune centers, the area of a commune, as well as the Euclidean distance to the dwellings. To calculate the predicted probabilities, we adjusted the constant terms to account for the unequal sampling rates (Maddala 1983). Finally, we adjusted for potential correlations of observations within communes using the Huber and White estimator (Williams 2000). This technique yields robust standard errors and partly accounts for potential spatial autocorrelations in the model residuals (Gellrich et al. 2007).

Table 1 Descriptive statistics of independent variables

	Abandonment		Stable cropland		Expansion		Stable non-cropland	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Elevation (100 m)	3.5	1.5	1.9	0.9	4.3	2.3	8.0	5.0
Slope range (avg degrees in 3 × 3 window)	5.7	4.1	1.8	2.7	5.8	5.2	9.6	5.8
Cost distance to built-up area[a]	13.7	17.1	19.0	22.4	19.3	31.9	44.6	47.7
Cost distance to road network[a]	43.0	28.8	25.7	11.1	46.9	38.8	88.7	52.7
Forest in 3 × 3 window, 1996 (pixels)	1.0	2.0	0.2	0.9	1.8	2.9	5.0	3.8
Cropland in 3 × 3 window, 1996 (pixels)	4.6	2.8	7.5	2.2	2.0	2.4	0.5	1.3
Grassland in 3 × 3 window, 1996 (pixels)	2.3	2.4	0.6	1.2	3.8	3.0	1.7	2.6
Settlements in 3 × 3 window, 1996 (pixels)	1.0	1.8	0.6	1.6	1.3	2.2	1.0	2.4
Population density, 1996 (interpolated)	1,392	5,837	2,150	9,906	1,742	5,175	1,591	10,982
Net migration, 1996 (100 persons)	−0.7	1.8	−1.0	2.9	−1.1	2.8	−0.8	2.1
Maize yield, 1996 (100 kg/ha)	35.6	8.6	36.9	8.3	33.4	9.0	29.5	9.7
Livestock units (100), 1996	12.9	5.9	15.6	8.5	14.6	6.9	13.3	4.7
Employees in education, 1996 (persons)	102.5	132.3	101.3	30.4	122.5	281.1	178.3	431.1
Employees in agriculture, 1996 (persons)	60.8	107.5	118.6	124.9	58.9	102.2	27.7	67.4

[a] Cost distance is calculated using slope, land use type and road category

The goodness-of-fit of the logistic regressions was assessed by the percent of observations predicted correctly (PC), Cohen's Kappa and the area under the curve (AUC) of the receiver operating characteristic. Cohen's Kappa assesses the accuracy of location and PC describes the ratio of correctly predicted cells from the total number of cells. The AUC measures the performance of a model compared to a random model while the cut-off threshold is varied from zero to one (Pijanowski et al. 2006; Schneider and Pontius 2001).

Prediction maps were generated by applying the calculated coefficients from the sampled data to the full dataset to calculate the predicted probabilities for the entire county. Predicted probabilities were calculated using the common threshold of 50%, where all probabilities higher than 50% were predicted as presence and all probabilities below this threshold were labeled as predicted absence.

Artificial neural networks

In this study we use the Land Transformation Model (LTM, Pijanowski et al. 2002), an artificial neural network (ANN) model that relies on the Stuttgart Neural Network Simulator (SNNS, Zell et al. 2000). Artificial neural networks can determine linear as well as non-linear functional relationships based on pattern recognition. We applied a multilayer perceptron, which is a supervised classification method that trains the network on the known output of a dataset (Dreiseitl and Ohno-Machado 2002). We selected a 3-layer structure of neurons with several input layers (in statistical terminology explanatory variables), one output layer (the outcome or response variable) and hidden layers in between (Bishop 1995). As suggested by Pijanowski et al. (2005), the same number of hidden and input layers were used. We calibrated one model for each change process (abandonment and expansion). These feed-forward networks repeatedly present the data to the network and calculate the mean squared error for each cycle. This error describes the difference between the calculated and the expected output (Bishop 1995). This error is then minimized by iteratively adjusting the weights.

We applied the resilient back propagation (RPROP) training algorithm, which adapts the network weights in each learning step based on the signs of the partial derivatives, but not based on their magnitude (Riedmiller and Braun 1993). In this way, the learning is less likely to be trapped in local minima, the number of necessary learning steps can be significantly reduced and the network is more robust and therefore more resilient against the choice of input variables (Riedmiller and Braun 1993). Hence, this algorithm may offer particular advantages for complex modeling approaches with a larger number of input variables as in our study. Figure 2 illustrates the structure of the ANN for the example of cropland abandonment.

Ample flexibility in the choice of the input variables exists in the neural network estimations. For the sake of comparability, we selected the same variables that were used in the logistic regressions (see Sect. 2.3.1 and Table 1). In addition, we included all those variables that were removed from the regressions due to multicollinearity.

For both the abandonment and the expansion simulations, we randomly selected 10% of all observations to simultaneously reduce training time and avoid over-fitting the network which can result in biased accuracy assessments and an inability of the model to generalize (Bishop 1995; Pijanowski et al. 2006). The neural network was trained on every second cell in the randomly selected observations and was run for 250,000 cycles. After training, the network was tested with the data which had not been used for training. We recorded the overall Kappa, AUC, and the PC in each training cycle. The cycle with the best fit based on the three metrics was applied to the entire dataset to calculate a map of suitability values for each pixel (neural nets do not create probabilities per se, see e.g. Bishop 1995). These suitability values can be used to derive the propensity for change, with high values representing areas with a high likelihood of change.

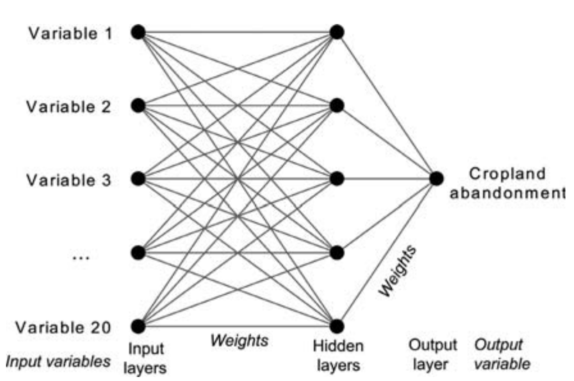

Fig. 2 Structure of the artificial neural network for cropland abandonment

Comparison of the modeling results

Based on the logistic regression as well as on the ANN results for the suitability of cropland change, we produced a map of likely future hotspots of both abandonment and expansion. To delineate the hotspots, we summed up the output values for abandonment and expansion in a one square kilometer moving window and classified the respective results into ten quantiles. The highest quantile corresponds to the highest spatial concentration and intensity of cropland change. The identified hotspots were qualitatively validated with expert knowledge and field data.

Results

Cropland change

Some 17% (356 km^2) of cropland was abandoned between 1995 and 2005 (Müller et al. 2009). During the same period, 12% (229 km^2) of the locations which were not covered by cropland in 1995 were cover by cropland in 2005. Figure 3 illustrates the spatial pattern of cropland changes between 1995 and 2005.

Cropland abandonment was particularly widespread in the central and hilly areas of Argeş, often on sloping land in the valleys north and east of Piteşti. Cropland usage remained largely stable in the southern area. Cropland expansion was scattered

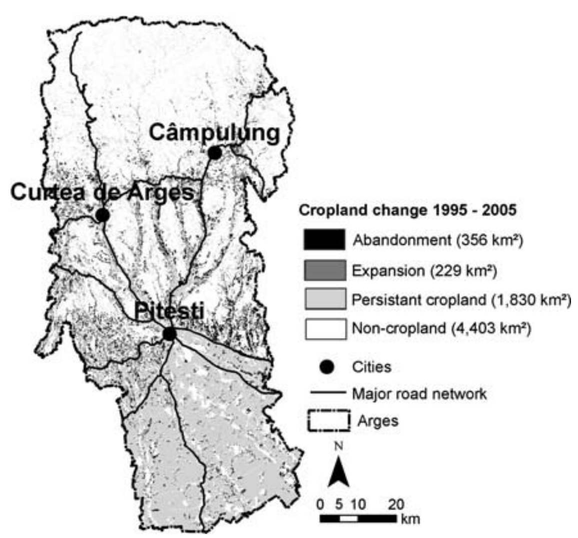

Fig. 3 Observed cropland change between 1995 and 2005

across the county, with minor agglomerations in the north of the study area, close to the major roads and on a few patches in the southern plains.

Results from logistic regressions: determinants and predicted probabilities of cropland change

The goodness-of-fit of the two logistic regression models differed radically (Table 2). While the AUC was 0.72 for the abandonment model, the expansion equations exhibited an AUC of 0.5, equal to a random model. Results of the abandonment equations for PC and Kappa were reasonably good, while the fit for the expansion equations was again close to random. Therefore, we have no confidence in the determinants that were derived for cropland expansion and do not interpret the respective coefficients.

Results from the spatially explicit logistic regressions for abandonment indicate that topographic characteristics, market accessibility and the spatial structure of land use are the most important spatial determinants of cropland abandonment (Table 3). For example, abandonment was more likely at higher elevations and on steeper slopes. For every 100 m of altitude, the risk of abandonment increased by 97% and an additional slope degree in a three-by-three window rendered abandonment 14% more likely. Cropland abandonment was more likely closer to built-up areas and at locations that were more difficult to reach from the road network. A large amount of cropland in neighboring locations had a strong negative bearing on abandonment and one additional pixel covered by cropland reduced abandonment by 24%. The presence of settlements in the neighborhood (300 × 300 m) decreased abandonment by 19% at the 10% significance level. Finally, more livestock units per square kilometer decreased the likelihood of abandonment.

The far left image in Fig. 4 depicts prediction maps for cropland abandonment (no prediction maps were created for the regression results of cropland expansion, for the reasons given above).

Table 2 Goodness-of-fit of logistic regressions (cropland abandonment/expansion between 1995 and 2005)

Process	PC	AUC	Kappa
Abandonment	0.70	0.70	0.40
Expansion	0.53	0.51	0.01

Results from ANN: propensities of cropland change

The ANN model fit was moderate for cropland abandonment, with an AUC of 0.87 and a Kappa of 0.54. As expected, these values were lower for cropland expansion, with an AUC of 0.84 and a

Table 3 Logistic regressions results (odds ratios) (cropland abandonment/expansion between 1995 and 2005)

	Abandonment	Expansion
Elevation (100 m)	1.970***	0.899***
Slope range (avg degrees in 3 × 3 window)	1.142***	1.000
Cost distance to build-up area	0.993*	0.995**
Cost distance to road network	1.022***	0.996**
Forest in 3 × 3 window, 1996 (pixels)	0.910	1.264***
Cropland in 3 × 3 window, 1996 (pixels)	0.760**	1.745***
Grassland in 3 × 3 window, 1996 (pixels)	1.009	1.575***
Settlements in 3 × 3 window, 1996 (pixels)	0.814*	1.300***
Population density, 1996 (interpolated)	1.000	1.000
Net migration, 1996 (100 persons)	1.005	0.938**
Maize yield, 1996 (100 kg/ha)	0.996	1.004
Livestock units (100), 1996	0.977***	1.016
Employees in education, 1996 (persons)	1.001	1.000
Employees in agriculture, 1996 (persons)	1.000	0.999
Constant	0.632	0.078***
Number of observations	2000	1904

* $P < 0.10$, ** $P < 0.05$, *** $P < 0.01$

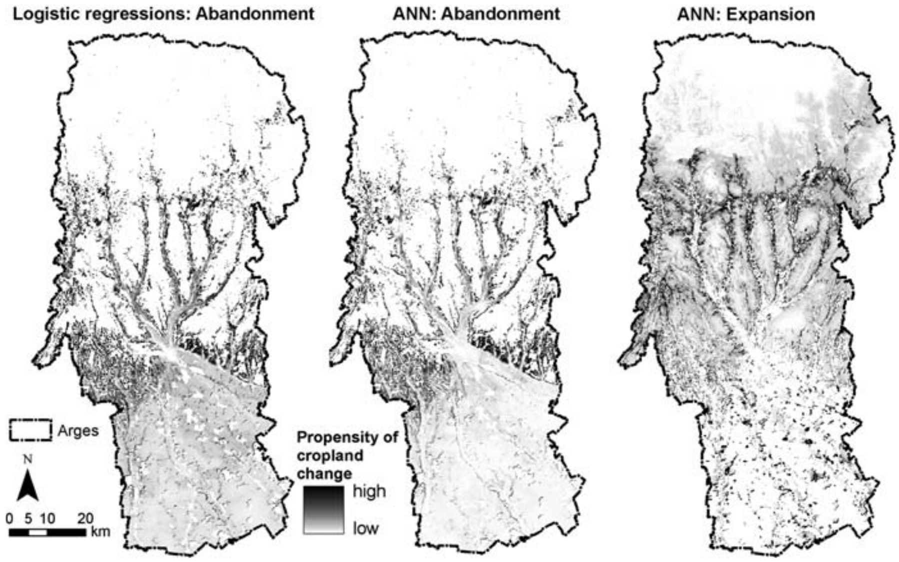

Fig. 4 Propensity for cropland change

Table 4 Goodness-of-fit of neural networks (cropland abandonment/expansion between 1995 and 2005)

Process	PC	AUC	Kappa
Abandonment	0.61	0.87	0.54
Expansion	0.34	0.84	0.30

Kappa of 0.3 (see Table 4). The highest accuracies according to PC and Kappa were achieved for cropland abandonment at cycle 1,500 and for cropland expansion at cycle 900. The patterns of these training cycles were used to derive the maps of the propensity to transition out of cropland (abandonment) or to be converted into cropland (expansion) in Argeş County (Fig. 4).

Spatial predictions and hotspots of change from logistic regressions and ANN

The resulting spatial predictions of the logistic regression and the ANN yielded similar spatial patterns for abandonment with the highest likelihood for abandonment in the hilly areas in the center of Argeş and along the valley bottoms.

Abandonment pathways follow the major roads that traverse Argeş from southeast to northwest and from southwest to northeast. The propensities were much lower in the plains, but still affected a substantial number of pixels. The northern areas are not affected by abandonment since the overall amount of cropland is small. The highest propensities for cropland expansion derived from the ANN were evident in the valleys of the hilly area.

Hotspots of cropland change show a distinct spatial pattern for abandonment and expansion. Abandonment hotspots are particularly concentrated in the hilly areas. Hotspots of abandonment and expansion repeatedly occur next to each other, especially in the hilly parts. These areas are probably under a more imminent threat of change.

Discussion

Integrated analysis of cropland change

Cropland changes in Argeş, Romania, were investigated using an integrated data set of cropland maps, socioeconomic census data, accessibility measures and geophysical variables. Satellite-derived land use maps allowed us to study the spatial patterns of cropland change at a regional scale. Our results demonstrate the importance of an integrated approach combining socioeconomic and environmental data to study the complex nature of land use processes. The regressions for cropland abandonment showed that abandonment takes place in locations that have a low natural potential (e.g. high elevation and steep slopes). In addition to the environmental determinants, abandoned locations are characterized by adverse market access and the proximity of other cropland and settlements. This agrees with findings in Western Europe (Gellrich et al. 2007; Tasser and Tappeiner 2002). The new political system in Romania that followed the collapse of socialism probably led to the increased importance of market access. Demographic variables such as migration and population numbers were not significant explanatory variables for abandonment, partly because remote sensing measures the extent of cropland and not the land use intensity on a given plot of land, which may have declined due to large reductions in labor inputs. Neighboring land use exerted a large influence on patterns of cropland occurrence. Small and isolated areas of cropland were more likely to be abandoned than those in close proximity to other cropland areas. This indicates a tendency towards a homogenous cropland structure.

The maps of propensity for abandonment show the relationship between environmental and socioeconomic factors and cropland change and mirror the observed patterns of cropland abandonment. Abandonment dominates in the hilly areas where complex terrain and poor market access influence cropland use. Pathways of abandonment follow the major roads that traverse Argeş from southeast to northwest and from southwest to northeast. In contrast, the southern plains show much lower propensity for abandonment due to high environmental suitability and homogeneous patterns of cropland use. The northern areas are not affected by abandonment as they are predominately covered by mountain meadows and forests with very few cropland areas to start with.

Since 1995 cropland expansion has occurred across Arges in a relatively uniform manner with few characteristic patterns. The spatially explicit logistic regressions could only explain a relatively small share of the cropland dynamics. Illustrated by the low explanatory value of the expansion models' statistical findings. The results from the neural network nevertheless yield some useful findings about patterns of cropland expansion. The maps of propensity for expansion show the greatest likelihood of expansion on land adjacent to existing cropland. This was particularly evident in the south where high propensities for expansion exist for the few remaining fragments of non-cropped land. Expansion is also very likely in the valley bottoms and on hillsides, where tracts of unused land are suitable for crop production.

Both cropland change processes frequently occurred in adjacent regions in the hilly areas. This corresponds with the findings of other studies carried out in geographically similar settings where intensification and abandonment occur in areas with similar environmental characteristics (Nagendra et al. 2003). In order to visualize the patterns of abandonment and expansion we identified contiguous areas of cropland change (Fig. 5). These 'hotspots' of cropland change are concentrated in areas with heterogeneous land use types in the hilly region of Argeş. In the case of abandonment, hotspots are situated in heterogeneous terrain, with low environmental suitability located far away from markets. Hotspots of expansion are extending across the northern part of the hilly area which is most likely to be re-used as cropland due to environmental suitability, availability of unused land as well as due to the high proportion of cropland at nearby locations.

The identification of these hotspots allowed us to avoid some of the locational mistakes that are common to pixel-based analysis. Moreover, it enabled us to make a better visual judgment of areas that are under a more imminent threat of change. This is useful when implementing spatially targeted policy

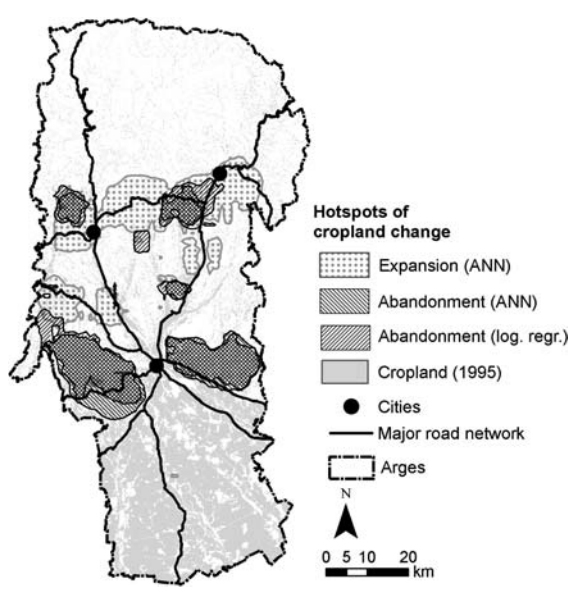

Fig. 5 Hotspots of cropland change

and management measures (Müller and Mburu 2009). The hotspot results may, for example, serve as an input into regional land use planning activities and specify areas with a need for more detailed investigation. The hotspot maps of cropland changes in Argeş further stress the importance of spatial planning that addresses different paths of development trajectories within one region.

Comparison of the two modeling approaches

From a methodological point of view, spatial logistic regression models and ANN share several commonalities, yet have different objectives, benefits and challenges. We used the logistic regressions to analyze the contribution of bio- and geophysical factors, accessibility measures, as well as socioeconomic variables to cropland change. The logistic regressions allowed us to effectively describe the relationships between the variables, rank the importance of influencing factors and test hypotheses about the underlying processes of land change (Munroe and Müller 2007). This is a powerful way to link hypothesized processes of land use change with land cover changes observed on the ground. Logistic regressions thus provide a good means of retrospective analysis within the limits of spatially explicit regression models that concern problems of optimal scale of available data, different model specifications implying different spatial correlation structures and the linear combination of input variables (Anselin 2002).

Conversely, ANNs seek to achieve the highest goodness-of-fit based on machine learning. They do not address cause and effect relationships nor do they improve process understanding. But the ANNs allow us to partly fill knowledge and data gaps, which proved particularly useful in the case of cropland expansion. The poor results of the logistic regression suggest that the functional relationship between cropland expansion and the selected set of independent variables can not be approximated with a linear combination of independent variables with the logit of expansion (Hosmer and Lemeshow 2000). Artificial neural networks are not constrained by the occurrence of non-linear relationships between inputs and output layers. Because their activation functions permit the modeling of more complex relationships (Bishop 1995). This can be particularly beneficial for the analysis of environmental land change processes (Pijanowski et al. 2005; Wieland and Mirschel 2008). Moreover, ANNs are not limited by incorrect sampling, multicollinearity between variables, spatial or temporal autocorrelation, or the insignificance of single variables (Bishop 1995) since their aim is to achieve the highest fit without identifying functional relationships of individual variables. This makes ANN particularly suitable for a range of applications in land change science that deal with large human-environment data sets from multiple sources of varying accuracy and precision.

However, the drawbacks of ANNs are related to their data-driven approach. ANNs do not require a priori knowledge of underlying processes. Instead, they aim to recognize patterns that result from such processes and hence do not offer new insights into functional relationships. Furthermore, resulting patterns may assign false importance to relationships if major drivers have not been included in the model. This leads to a major limitation of ANNs; the results cannot be fully retraced and any new insights into land use change processes are limited to the interpretation of the patterns derived.

The mathematical similarity of the two approaches (single versus multi-layer perceptron) is responsible for some common drawbacks of both models. One limitation of both modeling approaches is the inability to model feedbacks between variables, a challenge that has yet to be satisfactorily achieved in land use modeling (Verburg 2006). Another limitation of both approaches is the calculation of two separate models for abandonment and expansion, without the consideration of interactions between these processes or of additional, possibly competing, land use changes. Modeling many land change processes simultaneously is more data-intensive and often impractical for the analysis of large areas. In addition, such modeling also comes at the expense of comprehensibility and reproducibility.

We believe that the comparison of the two modeling approaches presented here yielded both explanatory and predictive findings. Both are important for understanding land change processes in regions where empirical evidence for the underlying causes and possible future developments is scarce, such as in the case of the postsocialist countries of Central and Eastern Europe.

Conclusions

In Argeş, Romania, land changes processes are dominated by changes in the extent of cropland. We used an integrated data set to analyze the changes of cropland and their potential developments. Knowledge about changes in the extent of cropland is important for understanding its impact on the multifunctional nature of agricultural production (OECD 2001) and on numerous ecosystem services and biodiversity.

We compared two modeling techniques that exploit the benefits inherent in explanatory statistics and pattern recognition. We believe that both approaches generated complementary insights and can support decision-making. Such analysis may further assist regional spatial planning efforts, environmental monitoring and biodiversity conservation in areas that experience significant changes in land use. Further research should include scenarios of likely future pathways of cropland usage to investigate the shift in political and socioeconomic variables instigated by Romania's membership in the European Union. This is particularly important because agricultural extensification and cropland abandonment are expected to be the largest land-use changes in the enlarged European Union, particularly in marginal areas (Verburg et al. 2006).

Acknowledgments We are very grateful for the valuable comments of two reviewers and Eleanor Milne from the Office for Integration and Modeling that significantly improved the paper. We thank Tobias Kuemmerle and Patrick Griffiths for the remote sensing analysis of land cover changes. We also thank the Global Land Project and the guest editors for organizing this special issue. Funding of the data collection from the Deutsche Forschungsgemeinschaft (DFG) under the Emmy-Noether Programme is gratefully acknowledged.

References

Almeida CM, Gleriani JM, Castejon EF, Soareasfilho BS (2008) Using neural networks and cellular automata for modelling intra-urban land-use dynamics. Int J Geogr Inf Sci 22:943–963

Anselin L (2002) Under the hood: issues in the specification and interpretation of spatial regression models. Agric Econ 27:247–267

Besag JE (1974) Spatial interaction and the statistical analysis of lattice systems. J R Stat Soc 36:192–236

Bicik I, Jelecek L, Stepanek V (2001) Land-use changes and their social driving forces in Czechia in the 19th and 20th centuries. Land use policy 18:65–73

Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, Oxford

Chilonda P, Otte J (2006) Indicators to monitor trends in livestock production at national, regional and international levels. Livest Res Rural Dev 18. Accessed 22 Nov 2006 at http://www.cipav.org.co/lrrd/lrrd18/8/chil18117.htm

Chomitz KM, Gray D (1996) Roads, lands use, and deforestation: a spatial model applied to Belize. World Bank Econ Rev 10:487–512

Cremene C, Groza G, Rakosy L, Schileyko AA, Baur A, Erhardt A, Baur B (2005) Alterations of Steppe-like grasslands in Eastern Europe: a threat to regional biodiversity hotspots. Conserv Biol 19:1606–1618

Dreiseitl S, Ohno-Machado L (2002) Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 35:352–359

Elbakidze M, Angelstam P (2007) Implementing sustainable forest management in Ukraine's Carpathian Mountains: the role of traditional village systems. For Ecol Manage 249:28–38

Fischer MM, Abrahart RJ (2000) Neurocomputing—tools for geographers. In: Openshaw S, Abrahart RJ (eds) Geocomputation. Taylor & Francis, London, pp 187–218

Gellrich M, Baur P, Koch B, Zimmermann NE (2007) Agricultural land abandonment and natural forest re-growth in the Swiss mountains: a spatially explicit economic analysis. Agric Ecosyst Environ 118:93–108

Hosmer DW, Lemeshow S (2000) Applied logistic regression. Wiley, New York

Ioras F (2003) Trends in Romanian biodiversity conservation policy. Biodivers Conserv 12:9–23

Kuemmerle T, Radeloff VC, Perzanowski K, Hostert P (2006) Cross-border comparison of land cover and landscape pattern in Eastern Europe using a hybrid classification technique. Remote Sens Environ 103:449–464

Kuemmerle T, Hostert P, Radeloff V, van der Linden S, Perzanowski K, Kruhlov I (2008) Cross-border comparison of post-socialist farmland abandonment in the Carpathians. Ecosystems 11:614–628

Kuemmerle T, Müller D, Griffiths P, Rusu M (2009) Land use change in Southern Romania after the collapse of socialism. Reg Environ Change 9:1–12

Lerman Z, Csaki C, Feder G (2004) Evolving farm structures and land use patterns in former socialist countries. Q J Int Agr 43:309–335

Macours K, Swinnen JFM (2008) Rural–urban poverty differences in transition countries. World Dev 2008

Maddala GS (1983) Limited-dependent variables in econometrics. Cambridge University Press, New York

May RJ, Maier HR, Dandy GD, Fernando TMKG (2008) Non-linear variable selection for artificial neural networks using partical mutual information. Environ Model Softw 23(10–11):1312–1326

Mertens B, Lambin EF (2000) Land-cover change trajectories in southern Cameroon. Ann Assoc Am Geogr 90:467–494

Müller D, Mburu J (2009) Forecasting hotspots of forest clearing in Kakamega Forest, Western Kenya. For Ecol Manage 257:968–977

Müller D, Munroe DK (2008) Changing rural landscapes in Albania: cropland abandonment and forest clearing in the postsocialist transition. Ann Assoc Am Geogr 98:855–876

Müller D, Kuemmerle T, Rusu M, Griffith P (2009) Lost in transition: determinants of cropland abandonment in postsocialist Romania. J Land Use Sci 4:109–128

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

Nagendra H, Southworth J, Tucker C (2003) Accessibility as a determinant of landscape transformation in western Honduras: linking pattern and process. Landscape Ecol 18:141–158

NIS (2004) Database of localities. National Institute of Statistics, Bucharest

OECD (2001) Multifunctionality: towards an analytical framework. Organisation for Economic Co-operation and Development (OECD), Paris

Openshaw S (1998) Neural network, genetic, and fuzzy logic models of spatial interaction. Environ Plan A 30: 1857–1872

Oszlanyi J, Grodzinska K, Badea O, Shparyk Y (2004) Nature conservation in Central and Eastern Europe with a special emphasis on the Carpathian Mountains. Environ Pollut 130:127–134

Palang H, Printsmann A, Gyuró ÉK, Urbanc M, Skowronek E, Woloszyn W (2006) The forgotten rural landscapes of Central and Eastern Europe. Landscape Ecol l 21: 347–357

Perz SG, Skole DL (2003) Social determinants of secondary forests in the Brazilian Amazon. Soc Sci Res 32:25–60

Peterson U, Aunap R (1998) Changes in agricultural land use in Estonia in the 1990s detected with multitemporal Landsat MSS imagery. Landsc Urban Plan 41:193–201

Pijanowski BC, Brown DG, Shellito BA, Manik GA (2002) Using neural networks and GIS to forecast land use changes: a land transformation model. Comput Environ Urban Syst 26:553–575

Pijanowski BC, Pithadia S, Shellito BA, Alexandridis K (2005) Calibrating a neural network-based urban change model for two metropolitan areas in the upper midwest of the united states. Int J Geogr Inf Sci 19:197–215

Pijanowski BC, Alexandridis KT, Müller D (2006) Modelling urbanization patterns in two diverse regions of the world. J Land Use Sci 1(2):83–108

Riedmiller M, Braun H (1993) A direct adaptive method for faster backpropagation learning: the RPROP algorithm. In IEEE International Conference on Neural Networks, San Francisco

Salazar-Ruiz E, Ordieres JB, Vergara EP, Capuz-Rizo SF (2008) Development and comparative analysis of tropospheric ozone prediction models using linear and artificial intelligence-based models in Mexicali, Baja California (Mexico) and Calexico, California (US). Environ Model Softw 23:1056–1069

Schneider LC, Pontius RGJ (2001) Modeling land-use change in the Ipswich watershed, Massachusetts, USA. Agric Ecosyst Environ 85:83–94

Slater JA, Garvey G, Johnston C, Haase J, Heady B, Kroenung G, Little J (2006) The SRTM data "finishing" process and products. Photogramm Eng Remote Sensing 72:237–247

Tasser E, Tappeiner U (2002) Impact of land use changes on mountain vegetation. Appl Veg Sci 5:173–184

Verburg PH (2006) Simulating feedbacks in land use and land cover change models. Landscape Ecol 21:1171–1183

Verburg PH, Schulp CJE, Witte N, Veldkamp A (2006) Downscaling of land use change scenarios to assess the dynamics of European landscapes. Agric Ecosyst Environ 114:39–56

Wieland R, Mirschel W (2008) Adaptive fuzzy modeling versus artificial neural networks. Environ Model Softw 2:215–224

Williams RL (2000) A note on robust variance estimation for cluster-correlated data. Biometrics 56:645–646

Zell A, Mamier G, Vogt M, Mache N, Hübner R, Döring S, Herrmann K, Soyez T, Schmalzl M, Sommer T, Hatzigeorgiou A, Posselt D, Schreiner T, Kett B, Clemente G, Wieland J, Gatter J (2000) SNNS—Stuttgart Neural Network Simulator. Institute for Parallel and Distributed High Performance Systems, University of Stuttgart, Wilhelm Schickard Institute for Computer Science, University of Tübingen, Stuttgart