Environ Monit Assess (2015) 187: 57
DOI 10.1007/s10661-015-4298-8

Modeling land use and land cover changes in a vulnerable coastal region using artificial neural networks and cellular automata

Yi Qiang · **Nina S. N. Lam**

Received: 27 June 2014 / Accepted: 11 January 2015 / Published online: 3 February 2015

Abstract As one of the most vulnerable coasts in the continental USA, the Lower Mississippi River Basin (LMRB) region has endured numerous hazards over the past decades. The sustainability of this region has drawn great attention from the international, national, and local communities, wanting to understand how the region as a system develops under intense interplay between the natural and human factors. A major problem in this deltaic region is significant land loss over the years due to a combination of natural and human factors. The main scientific and management questions are what factors contribute to the land use land cover (LULC) changes in this region, can we model the changes, and how would the LULC look like in the future given the current factors? This study analyzed the LULC changes of the region between 1996 and 2006 by utilizing an artificial neural network (ANN) to derive the LULC change rules from 15 human and natural variables. The rules were then used to simulate future scenarios in a cellular automation model. A stochastic element was added in the model to represent factors that were not included in the current model. The analysis was conducted for two sub-regions in the study area for comparison. The results show that the derived ANN models could simulate the LULC changes with a high degree of accuracy (above 92 % on average). A total loss of 263 km^2 in wetlands from 2006 to 2016 was projected, whereas the trend of forest loss will cease. These scenarios provide useful information to decision makers for better planning and management of the region.

Keywords Coastal sustainability · Land use land cover change · Coupled natural-human dynamics · Artificial neural network · Cellular automata

Y. Qiang (✉) · N. S. N. Lam
Department of Environmental Sciences, Louisiana State University, 1273 Energy, Coast, and Environment Building, Baton Rouge, Louisiana 70803, USA
e-mail: yi_qiang@hotmail.com

Introduction

Land use land cover (LULC) change, as manifested by snapshots of remote sensing images, is considered an important indicator of the processes on the earth surface. Considerable amount of work has been conducted in modeling LULC changes to tackle problems such as deforestation (Lambin et al. 2003; Perez-Vega et al. 2012), urban expansion (Batty et al. 1999; Murray-Rust et al. 2013; Dewan and Yamaguchi 2009), and ecosystem service changes (Berger 2001; Wang et al. 2006). Lately, LULC change modeling has been applied to analyze the coupling of natural and human systems (CNH) using a variety of GIS and spatial modeling techniques with promising results (Brown et al. 2002; Liu et al. 2012). However, there remain many technical and application issues in using LULC change modeling to capture the interactions between natural and human systems, including issues on the selection of variables, modeling techniques, and usefulness of the results (Crooks et al. 2008). More studies are needed to help improve the modeling methods so that they can be used to address significant societal problems through a better understanding of the underlying natural and human processes (Leemans and Eickhout 2004; Liu et al. 2010).

This paper examines the LULC dynamics in the Lower Mississippi River Basin (LMRB), a coastal region considered to be among the most vulnerable in the USA (Fig. 1). In recent years, increasing attention has been paid to coastal areas due to their high exposure and vulnerability to natural hazards. A critical challenge facing these vulnerable coastal regions is how to achieve sustainability in the face of climate change and sea-level rise. To better address the challenge, an understanding of how human and natural factors interact in the system is necessary (Liu et al. 2007; Turner 2010). Hence, a well-designed LULC model to accurately represent and simulate the coupling process is the first step.

The Lower Mississippi River Basin is a region of high economic, cultural, and ecological significance. However, the disappearance of land is a critical problem along the southern coastal areas. Although there have been numerous studies on the region (Reyes et al. 2000; Day et al. 2007), most of them focused on either a specific process or a local scale. A system-level model, such as the LULC model described in this paper, is necessary to capture the dynamic linkages among the major components from the natural and human subsystems and to quantify the linkages through empirical validation. Moreover, during the last decade (2000–2010), there was a steady population growth in the northern part of the region (called the North hereafter), in contrast with a significant decline in population in the southern part (the South). This contrasting pattern of population growth and decline between the North and the South has accelerated after the strike of Hurricane Katrina in 2005 (Lam et al. 2009a, b). With these observations, we post two questions in this study: (1) How did the LULC change processes differ between the

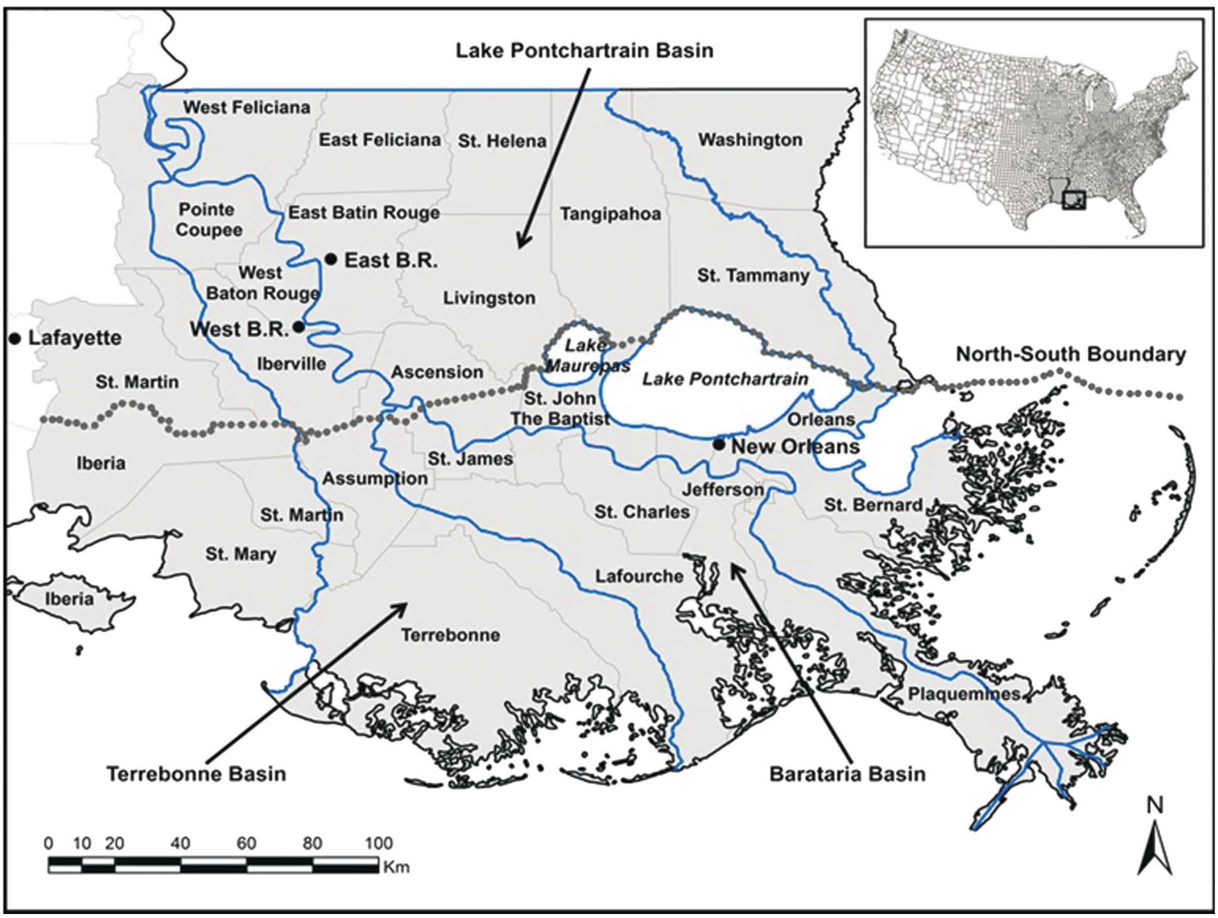

Fig. 1 The Lower Mississippi River Basin (LMRB) and the hypothetical north-south boundary

North and the South? (2) How will the future LULC patterns in the two sub-regions look if no intervention or mitigation measures are implemented? Answers to these questions should help policy makers in developing effective strategies to make the region more sustainable.

Since LULC changes are typically driven by many internal and external forces interacting dynamically with their surrounding environment, modeling the LULC changes using static equations alone would not be adequate. Cellular automata (CA) models have been proven as an effective approach to modeling LULC changes (Parker et al. 2003; Schweitzer et al. 2011). A CA represents complex spatial processes by defining simple local interactions (Lauf et al. 2012; Vicari et al. 2007). The rules of the local interaction can be derived by either a top-down or a bottom-up manner. The top-down approach deduces the rules from existing theories, and then calibrates them by trial and error, whereas the bottom-up approach derives the transitional rules of LULC changes from empirical data. Prior assumptions and knowledge about the LULC mechanism are the precondition of the top-down approach, which unfortunately is seldom the case. Hence, most LULC modeling uses the bottom-up approach to derive the rules. Data mining methods based on artificial intelligence techniques, such as rule set learning (Li and Yeh 2004) and artificial neural network (Li and Yeh 2002; Okwuashi et al. 2012), have been commonly used in deriving the rules. Also, many optimization techniques, such as genetic algorithm (Stewart et al. 2004) and simulated annealing (Feng and Liu 2012), have been applied to automatically search for the best parameter combination (García et al. 2013).

This study applies an artificial neural network (ANN) approach to derive the transitional rules for the North and the South individually using time-series remote sensing images. The ANN models include both natural and human variables that are deemed to be important in modifying the landscape. Unlike most previous CA models, this study adds a stochastic element in the model to represent factors that are not included in the model. The transitional rules derived from the ANN models are then used in a CA model to simulate future LULC patterns. This paper aims to contribute to the scientific literature by documenting how a combined and modified ANN and CA approach could be used to incorporate both natural and human factors to better understand and simulate complex LULC dynamics in a very vulnerable coastal region. The methodological issues involved in the modeling will be highlighted. The methods and approaches used in this study should be highly applicable to analyzing other coastal regions facing similar threats and vulnerabilities. From a management perspective, the results from this study will answer the two questions posted above, which are how did the land change processes differ between the North and the South and how will the future land use land cover patterns look? The resulting model could be used to help answer questions posed by the local communities, such as whether land loss along the coast will accelerate due to increasing fragmentation of land and whether certain urban areas will continue to expand, so that better planning strategies to adapt such changes can be made.

Study area

General information

The study area, broadly recognized as the Lower Mississippi River Basin (LMRB), is located in southeastern Louisiana, USA, and extends from the parishes (i.e., counties) north of Lake Pontchartrain to the coast (Fig. 1). The study area was selected as a result of combining both the natural boundaries (watersheds) and political boundaries (parishes). This region encompasses 26 Louisiana parishes. We used Lake Pontchartrain as an approximate boundary where parishes north of the Lake are considered "the North," whereas the parishes south of Lake Pontchartrain are called "the South." The South is considered one of the most vulnerable coastal systems in the USA, subjecting to multiple threats, including land subsidence, flooding, oil spills, and hurricanes (Lam et al. 2009a, b, 2012). According to the US 2010 census statistics, during the last decade (2000–2010), a gradual population growth (15.4 %) in the North has been observed, in contrast with a significant decline (−11.2 %) in the South. The study area is highly engineered with various coastal restoration plans to restore wetlands (Coastal Protection and Restoration Authority of Louisiana CPRA 2012; Stokstad 2005). A US Geological Survey estimate shows that Louisiana has lost 4900 km^2 of coastal wetlands since 1932. Climate change introduces new risks to coastal Louisiana residents in the form of sea-level rise and more intense and frequent storms and flooding events. While there is no single cause of the wetland loss, human actions are a primary factor (CPRA 2012). Debates continue on whether the high cost of coastal restoration (>9 billion dollars) is cost-effective given the continuing threats of sea-level rise and regional subsidence. All these unsolved questions and uncertainties make future planning and management of the region extremely challenging.

LULC changes

The LULC data of the study area were obtained from NOAA Coastal Service Center, under the Coastal Change Analysis Program (C-CAP). These data were classified from Landsat 5 Thematic Mapper (TM) images through extensive field sampling, validation, and standard quality-control review procedures, which have guaranteed high classification accuracy and consistency (Dobson et al. 1995). The spatial resolution of these data is 30 m by 30 m. The entire study area includes 53,384,656 pixels (The North 22,162,275 pixels, the South 31,222,381 pixels). The C-CAP data use the USGS classification scheme (Anderson et al. 1976), which includes seven LULC classes in the study area, including urban, agriculture, rangeland, forest, water, wetland, and barren. Currently, NOAA has published LULC data covering the study area at three time stamps, i.e., 1996, 2001, and 2006. We selected the LULC data of 1996 and 2006 to develop the model because of two reasons. First, a longer time span (10- vs. 5-year) is considered more reliable in capturing the LULC changes. Second, the 1996 and 2006 LULC data were obtained from remote-sensing images taken in the same season, which should minimize the classification inconsistencies caused by seasonal changes.

Figure 6a and Table 1 show the loss and gain of the seven LULC types between 1996 and 2006 for the two sub-regions. The North was found to have more LULC changes than the South. In the North, the conversion from forest to rangeland is significant (i.e., 1059 km^2), despite the fact that a considerable amount of rangeland has changed back to forest (i.e., 412 km^2). The unbalanced exchange between forest and rangeland resulted in significant rangeland gain and forest loss in the North. In the South, the majority of changes took place in water, wetland, and barren land, with a significant loss of wetland to water and barren land. In both the North and the South, there was a net gain of urban land.

Methodology

Artificial neural network

Due to its accuracy, efficiency, and ability of modeling non-linear relations, artificial neural network (ANN) are now widely used for classification, pattern recognition, function approximation, and optimization in various applications (Jain et al. 1996). Li and Yeh (2002, 2004) applied ANNs to calibrate a LULC model to simulate the landscape scenarios in a city of China, which has generated satisfactory results. A number of researchers have also reported success in applying ANNs to calibrate LULC models in specific regions (Mahajan and Venkatachalam 2009; Okwuashi et al. 2012; Ju and Lam 2007; Zhou 2006). An ANN consists of an input layer, one or more hidden layers, and an output layer. Each layer consists of a number of neurons. Each neuron in the input layer accepts one of the input variables and generates an output value to the next layer. In such a way, the input variables are like signals that pass through the layers in the ANN and eventually generate output values.

Figure 2 illustrates the architecture of a three-layer ANN. Equation 1 describes how a neuron in the receiver layer receives values from the neurons in the sender layer, where I_i^n is the input value from the ith neuron in the sender layer and I_j^{n+1} is the output generated by the jth neuron in the receiver layer. $W_{i,j}$ denotes the weights of the input values and b is a bias value added to the summation of all inputs. There are a few options for the transition function f. The selection of f is dependent on the characteristics of the data to be trained. In the input layer, all variables need to be rescaled into the range between 0 and 1 for standardization.

$$I_j^{n+1} = f\left(\sum\nolimits_i w_{ij} I_i^n + b_j\right) \tag{1}$$

In this study, the training of the ANN was carried out in the Matlab® neural network toolbox. The default feedforward model was used. The steps of training are shown in Figure 3. Through training, the algorithm iteratively searches for the optimal weight combination for the ANN. According to the user setting, the program divides the sample set into a training set, a validation set, and a test set. In each iteration, the ANN is updated to fit the training set, and the result is verified in the validation set. The algorithm stops training when the error of the ANN in the validation set stops reducing or the maximum number of iterations has been reached. The test set

Table 1 LULC convers 1996-2006 (unit km^2)

		Change from (1996)							Total (2006)
		Urban	Agriculture	Rangeland	Forest	Water	Wetland	Barren	
The North									
Change to (2006)	Urban	1343 (99 %)	24.8 (1 %)	31.9 (2 %)	38.2 (1 %)	0.5 (0 %)	24.4 (0 %)	2.2 (2 %)	1464.9
	Agriculture	1 (0 %)	3897.4 (95 %)	16.9 (1 %)	41.1 (1 %)	0.7 (0 %)	42.7 (1 %)	1.6 (1 %)	4001.6
	Rangeland	0.2 (0 %)	135.8 (0.3 %)	1306.3 (72 %)	1085.8 (31 %)	2.3 (0 %)	212.4 (3 %)	11.1 (13 %)	2752.9
	Forest	0.1 (0 %)	3.2 (0 %)	412.5 (23 %)	2242.3 (65 %)	0.2 (0 %)	25.2 (0 %)	0.35 (0 %)	2683.7
	Water	0 (0 %)	6.4 (0 %)	3.2 (0 %)	4.3 (0 %)	1375.4 (99 %)	25.1 (0 %)	2.8 (3 %)	1417.2
	Wetland	0.45 (0 %)	24.5 (1 %)	31.3 (2 %)	37.1 (1 %)	6.1 (0 %)	7427.8 (96 %)	1.1 (1 %)	7528.2
	Barren	0.1 (0 %)	3.5 (0 %)	5.9 (0 %)	8.1 (0 %)	6.0 (0 %)	8.3 (0 %)	66.3 (78 %)	98.2
	Total (1996)	1344 (100 %)	4094.5 (100 %)	1808.1 (100 %)	3456.9 (100 %)	1391.4 (100 %)	7765.9 (100 %)	85.3 (100 %)	19,946.8
The South									
Change to (2006)	Urban	1119.1 (100 %)	9.1 (0 %)	3.7 (4 %)	0.2 (1 %)	0.1 (0 %)	8.6 (0 %)	0.4 (0 %)	1141.1
	Agriculture	0.1 (0 %)	2054.3 (99 %)	0.3 (0 %)	0 (0 %)	0 (0 %)	15.9 (0 %)	0.3 (0 %)	2070.8
	Rangeland	0.1 (0 %)	7.5 (0 %)	97.8 (94 %)	1.3 (7 %)	0 (0 %)	0.3 (0 %)	1.8 (1 %)	108.7
	Forest	0 (0 %)	0 (0 %)	0.8 (0 %)	17.3 (90 %)	0 (0 %)	0.1 (0 %)	0 (0 %)	18.3
	Water	0 (0 %)	0.6 (0 %)	0.1 (0 %)	0	12,440.2 (99 %)	171.0 (1 %)	22.9 (10 %)	12,634.9
	Wetland	1 (0 %)	2.4 (0 %)	1.3 (1 %)	0.5 (3 %)	27.4 (0 %)	11,753.9 (98 %)	5.7 (2 %)	11,791.3
	Barren	0.8 (0 %)	0 (0 %)	0.1 (0 %)	0	111.4 (0 %)	18.1 (0 %)	204.8 (87 %)	335.3
	Total (1996)	1120.1 (100 %)	2073.7 (100 %)	104.2 (100 %)	19.3 (100 %)	12,579.1 (100 %)	11,967.9 (100 %)	235.9 (100 %)	28,100.3

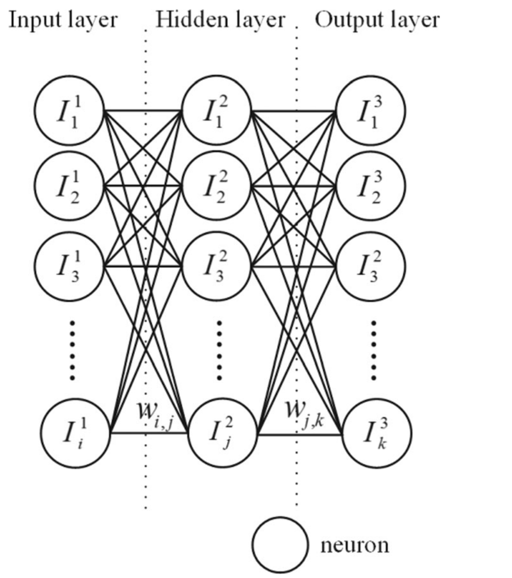

Fig. 2 The structure of a three-layer ANN (modified from Basheer and Hajmeer (2000))

provides an independent evaluation of the ANN performance after the training has stopped. The derived ANN model maps the complex relations between input variables and targets. The number of hidden layers, number of neurons in the hidden layers, and number of iterations may affect the training performance. The derived ANN can be used to simulate the targets with a different set of inputs, with the assumption that the relations between the inputs and targets are the same.

Cellular automata

Cellular automata (CA) are a type of discrete model commonly used in modeling spatial-temporal processes

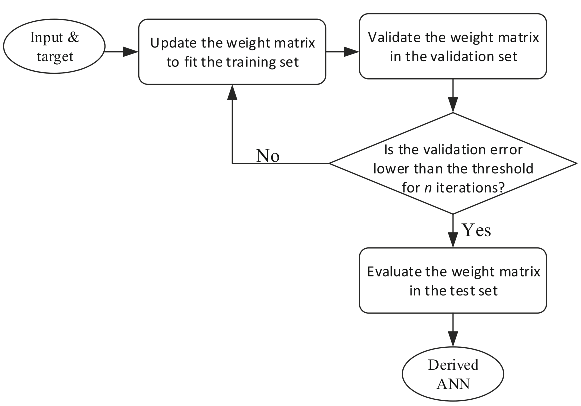

Fig. 3 The steps of training an ANN in the neural network toolbox in Matlab®

(Clarke et al. 2007; Clarke and Gaydos 1998). A cellular automaton includes five components: space, state, time step, neighborhood, and transitional rule (Wang 2007). The space in a CA is depicted as a lattice, composed of equally sized and shaped cells. The cells in the lattice interact with each other with predefined transitional rules, which take into account the spatial relation between the cells, such as neighborhood and distance. The cells interact with each other and update their states once at every time step. A CA can have more than one layer so that the interaction can be based on multiple variables of each cell. Through repeated interactions between the cells, a CA can model spatial dynamics at different scales. Using this approach, complex spatial processes can be simulated by defining local rules. As CA is inherently a spatial model, it is directly compatible with raster data and can be easily implemented in a GIS. Nowadays, it is widely applied in modeling environmental and ecological processes (Wu and Webster 1998; Batty et al. 1999; Ward et al. 2000). In this study, the LULC data and related variables were transformed into raster datasets, which can be modeled in a CA. With the transitional rules derived by the ANN, the CA is able to simulate the LULC dynamics at different time steps in a raster space.

Implementation

Variable set

A total of 15 variables, selected based on the literature to represent major natural and human factors, were used as input for the ANN modeling. The variables can be classified into three categories, including land properties, proximity to entities of interest (EOI), and neighboring LULC (see Table 2). The land properties include elevation, soil type, and original LULC. In coastal areas, elevation greatly affects the hydrological processes and hazard exposure (Day et al. 2007). Soil type affects the vegetation cover, land subsidence, and urban construction (Burkett et al. 2002). The original LULC is also an important factor, since most land tends to stay on its previous state. Many studies have shown that LULC changes, particularly urban expansion, are greatly dependent on the proximity to transportation infrastructure (Badoe and Miller 2000; Feng and Liu 2012). We separated primary roads and secondary roads, as literature shows that they have different influences to surrounding

Table 2 The input variables and data processing

	Variables	Original format and source	Data preparation
Land properties	Elevation	Raster dataset from National elevation dataset (NED) from US Geological Survey (USGS), updated in 2004	*Clip* the study area
	Soil type	Polygon shapefile from US General Soil Map (STATSGO2) data from USGS, updated in 1998	1. *Clip* the study area 2. Use *polygon to raster* to convert the shapefile to raster
	Original LULC[a]	The Coastal Change Analysis Program (C-CAP) datasets from National Oceanic and Atmospheric Administration (NOAA), updated in 1996 and 2006	*Clip* the study area
Proximity to EOI	Distance to primary roads Distance to secondary roads	Polyline shapefile TIGER/Line® files from the US Census Bureau, updated in 2000 and 2006	1. *Clip* the study area 2. Using *Euclidean distance* to calculate the proximity to roads
	Distance to urban area[a]	The Coastal Change Analysis Program (C-CAP) datasets from National Oceanic and Atmospheric Administration (NOAA), updated in 1996 and 2006	1. *Clip* the study area 2. Select urban cells by *raster calculator* 3. Use *focal statistics* to sum urban cells in a circle with 50-cell radius 4. Using *raster calculator* to select the cells having more than 40 urban cells within the radius
	Distance to open water[a]		1. *Clip* the study area 2. Select water cells by *raster calculator* 3. Use *focal statistics* to sum urban cells in a circle with 7-cell radius 4. Using *raster calculator* to select the cells that have more than 24 water cells within the radius
	Distance to pipelines	Pipelines in Louisiana, from USGS, Biological Resources Division's, National Wetlands Research Center (NWRC), updated in 1999	1. *Clip* the study area 2. Use *Euclidean distance* to calculate proximity to secondary roads
LULC in neighborhood	Number of urban cells[a] Number of agriculture cells[a] Number of rangeland cells[a] Number of forest cells[a] Number of open water cells[a] Number of wetland cells[a] Number of barren cells[a]	The Coastal Change Analysis Program (C-CAP) datasets from National Oceanic and Atmospheric Administration (NOAA), updated in 1996 and 2006	1. *Clip* the study area 2. Select cells in a particular LULC by *raster calculator* 3. Use *focal statistics* to sum the LULC cells in a circle with 7-cell radius

Text in italics are ArcMap® tools used for the data preparation. All variables were linearly rescaled to the range from 0 to 1

[a] Dynamic variables that update in each iteration in the simulation

LULC changes (Brown et al. 2002; Lo and Yang 2002). As newly developed urban areas tend to be adjacent to existing urban areas (Batty 2007), the proximity to urban areas is considered as well. Through a sequence of spatial operations (see Table 2), the scattered small urban areas were filtered out, keeping only those continuous urban areas. Also, proximities to open water area and energy infrastructure (i.e., pipelines) have certain impacts on LULC change (González and Tornqvist 2006). Positive spatial autocorrelation had been observed in most LULC datasets (Overmars et al. 2003; Wear and Bolstad 1998), which means that the LULC of a land cell is closely related to its neighboring cells. Thus, the numbers of neighbor cells in the seven LULC types were also included as input for our experiment. The original input datasets are in diverse formats and from different sources, which were then converted to raster datasets with the same cell size and in the same spatial extent by ArcGIS® toolbox. All the selected input variables were linearly rescaled to the range from 0 to 1.

ANN training

The 15 raster datasets were converted to a text file containing a number of arrays for the ANN training in Matlab®. In this file, each array corresponds to a cell in the raster space and includes a set of variables of that cell. Given the input variables, an ANN was trained to estimate the probabilities (ranging from 0 to 1) that a land cell converts to a certain LULC type. Two ANNs were trained separately for the South and the North. Loading all data for training takes an extremely long time for computation. Previous studies showed that there is a logarithmic relation between the accuracy of the derived ANN and the training time (Li and Yeh 2002), which means that loading more data only increases the accuracy linearly, but causes an exponential increase of the training time. Considering the trade-off between computational time and desired performance, 10 % of the dataset were randomly selected for the ANN training. Of the 10 % sampled data, 70 % were used as the training set, and 15 % each were used for the test set and the validation set. The training was carried out in the ANN training tool in Matlab® which is based on a back-propagation training algorithm (Foody 1996). This study utilized a three-layer feedforward ANN model, which has been proven effective in solving most learning tasks (Gong 1996; Zhou and Civco 1996). The hidden layer consisted of 20 neurons. Since the computational efficiency of the simulation model was not the goal of this study, we did not systematically record the processing time and the number of iterations. However, we set the maximum number of iterations to 1000 and found that the model usually converged before the maximum number of iterations was reached, mostly at about 800 iterations. The processing time for each simulation was approximately 8–12 h (overnight) using a computer with a 2.4-GHz processor.

CA simulation

The derived ANN was then applied as transitional rules in a cellular automaton to simulate the LULC changes for the two sub-regions. Given the same input variables, the trained ANN model can calculate the probability that a land cell converts to one of the seven LULC types. Previous studies usually define that a land cell changes to a LULC type with the highest conversion probability. If the LULC with the highest conversion probability is the same as the original type, the land cell stays unchanged. However, this approach will make the changes with moderate conversion probabilities very unlikely to happen. Consequently, such simulation will lead to a result that the dominant changes would become even more dominant and the less frequently occurred changes would become even less likely to occur. Thus, to overcome this shortcoming and to add a stochastic element in this experiment, the conversion of a land cell was determined by a weighted random selection of the top three LULC types that have the highest conversion probabilities. Equation 2 explains how the weights are defined.

$$w'_i = \frac{p_i}{\sum_{i=1,2,3} P_i} \tag{2}$$

where P_i stands for the ith highest conversion probability. Equation 2 basically rescales the probabilities of the top three LULC types into 0-1 so that the random generator can be used to create chances that a pixel may change to a LULC type with the second or third transitional probability. This approach preserves the stochastic characteristic of the LULC change processes and could make the categorical distribution of different types of changes match better with the reality. Also, this setting is reasonable because a land cell does not necessarily change to a LULC type with the highest conversion probability, if it is not much higher than the second highest, and so on. The random selection could include more than the top three conversion probabilities, which would unduly increase computational time and stochastic disturbance.

To start the simulation, the ANN loads the initial variables as input and computes the LULC at the next time stamp. The computed LULC cells are then used as input variables for the next iteration. The iteration stops when a certain threshold is reached. We used the criterion that the iteration stops when the total number of pixels changed is the same as that of the last iteration. The simulation was implemented by a Python script written to loosely couple ArcGIS® and Matlab®. In each iteration, the Python program uses the ArcGIS® libraries to process the input raster datasets and convert them into a text file that can be loaded by Matlab. Next, the Python program calls the Matlab® program through the Matlab's Component Object Model (COM) interface to load the text file and simulate the output LULC using the derived ANN. The COM interface allows external application to control the objects and functions in Matlab. In the next iteration, the Python program loads the LULC text file exported from Matlab to update the input datasets and convert them into a text file. In this manner, ArcGIS does the data processing jobs and Matlab runs the simulation, exchanging data in text files. The general workflow of ANN training and CA simulation is illustrated in Fig. 4, whereas the coupling between ArcGIS and Matlab in the simulation is shown in Fig. 5.

Experiment results

The experiment consisted of the validation and simulation phases. First, the 2006 LULC was simulated using the 1996 datasets as the initial state, which was then compared with the actual LULC in 2006 to evaluate the accuracy of the ANN. Second, the future scenario was simulated using the actual 2006 datasets as the initial state.

Validation

The performance of the ANN model was assessed by comparing pixel-by-pixel the simulated and actual 2006 datasets. Table 3 shows that the simulation result yielded more than 90 % accuracy for most LULC types, with an overall accuracy of 91.8 and 97.1 % and Kappa statistics of 89.4 and 95.3 % in the North and the South, respectively.

Figure 6 compares the gain and loss for each LULC type between the simulated and the actual datasets; it shows that the categorical distribution of the simulated changes generally matched that of the actual changes. The LULC types that increased in the real dataset also increased in the simulation, and vice versa. The only exception was agriculture land in the South, which decreased in the real data but was simulated as an increase. The validation result confirms that (1) the selected variables were reasonable to describe the LULC changes and (2) the ANN model built from the 10 % sampling of the dataset could simulate the LULC changes in the rest of the study area with an acceptable accuracy.

Future prediction

After the validation, the next step was to simulate the LULC changes in 2016, assuming the transitional rules between 1996 and 2006 will continue during the next time interval. In this simulation, the LULC and input variables acquired in (or near) 2006 were used as input to simulate the LULC in 2016. The same stopping rule was used, which was the total number of pixels changed being the same as between 1996 and 2006.

Figure 7 shows that the simulated categorical distribution of the changes between 2006 and 2016 generally followed that of the previous time interval.

Figure 8 maps the simulated future changes. Since the changed area only occupied a small part of the study area, the spatial distribution of specific LULC changes was difficult to see. In order to make the changed pixels more visible at the study area scale, the net changes (i.e., gain minus loss) of a LULC type were aggregated within a neighborhood of each cell. The neighborhood used was a circle with a radius of 100 pixels (i.e., 3000 m). Through color coding, the spatial distribution of the LULC changes became more visible in Fig. 9.

From Figs. 7 and 9, we can observe several trends that are likely to occur in the next 10 years. (1) In general and as expected, more LULC changes were projected in the North than in the South. (2) Urban growth in both the North and the South will continue to increase (increased by 112 %, 127 km^2 in the North and 166 %,

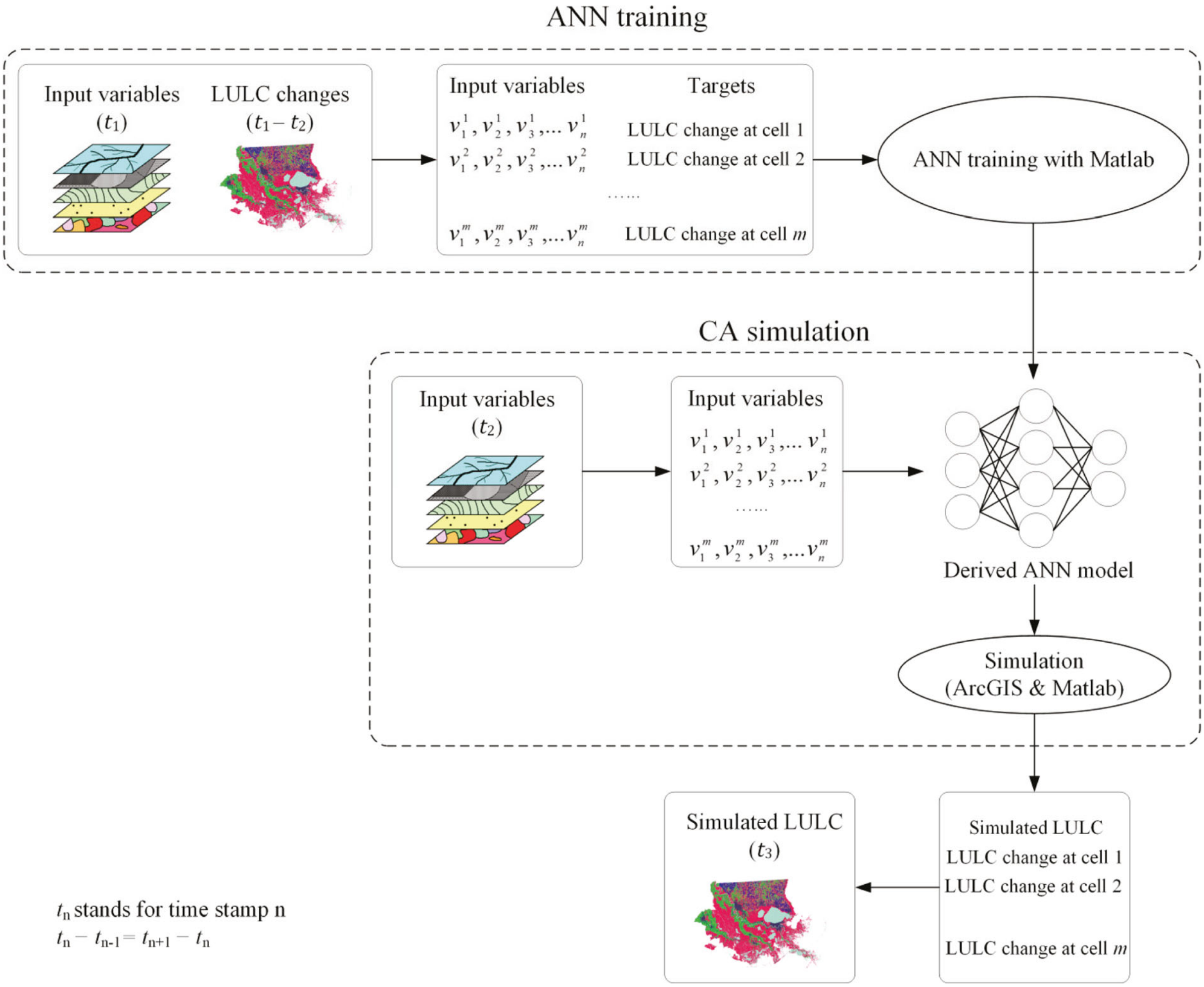

Fig. 4 The workflow of the training and simulation processes

56 km^2 in the South). Urban growth in the North will occur mostly in areas around major cities such as Lafayette and Baton Rouge. In the South, moderate urban growth was still projected near New Orleans, as well as along rivers and canals. (3) The spatial distribution of agricultural land change remains similar between the two time periods, with agricultural land in the North projected to decline 82 % more than the previous 10 years, amounting to a loss of 84 km^2. (4) Water gain (land loss) in the South will increase 38 %, amounting to an area of 76.8 km^2 (0.4 % of the total land of the South), with major pockets of water predicted to occur near the east side of Atchafalaya River and the mouth of Mississippi River. (5) Wetland loss will continue in a similar spatial pattern, but the rate of loss will accelerate by 14 % (amounting to a loss of 135 km^2 which is 3.2 % of total wetland) in the North, whereas wetland loss will slow down by 37 % (amounting to a loss of 128 km^2 which is 1.5 % of total wetland) in the South. (6) The trend of forest loss in the North will almost cease, changing from a loss of 386.6 km^2 to only 16 km^2 in this time period.

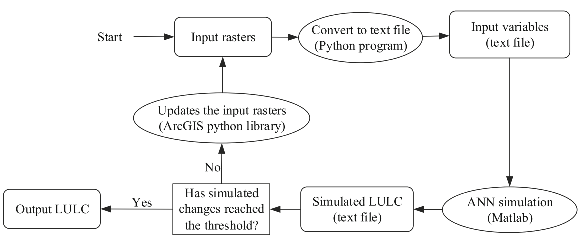

Fig. 5 The interaction between ArcGIS® and Matlab® in the simulation

Table 3 The confusion matrices between the actual and simulated LULC in 2006 (unit km^2)

	Actual LULC								
		Urban	Agriculture	Rangeland	Forest	Water	Wetland	Barren	User accuracy (%)
The North									
Simulated LULC	Urban	1414.6	55.2	156.6	85.8	7.3	99.8	9.2	77.4
	Agriculture	28.5	3821.1	101.2	2.4	7.5	48.4	4.7	95.2
	Rangeland	8.8	84.6	1947.9	60.6	4.9	55.3	6.4	89.8
	Forest	3.2	11.5	331.4	2483.3	3.9	68.5	0.8	85.6
	Water	0.9	4.1	3.7	0.5	1343.3	15.0	8.6	97.6
	Wetland	7.0	21.8	201.0	50.6	47.0	7237.4	4.6	95.6
	Barren	1.6	3.1	11.2	0.6	2.7	3.5	63.9	73.9
	Producer accuracy (%)	96.6	95.5	70.8	92.5	94.8	96.1	65.1	
The South									
Simulated LULC	Urban	1110.7	21.1	3.7	0.5	2.2	14.9	5	95.9
	Agriculture	16.2	2019.9	7.6	0.3	1.7	17.6	0.7	97.9
	Rangeland	2.1	2	83.6	0.6	0.4	8.3	0.3	86.0
	Forest	0.3	0.1	1.1	15.1	0	0.2	0	89.8
	Water	1	8.9	1	0.1	12,161.7	101.8	44	98.7
	Wetland	10.1	16.9	10.7	1.1	461	11,637.5	30.1	95.6
	Barren	0.7	1.9	1.1	0.6	6.7	11.1	255.2	92.1
	Producer accuracy (%)	97.3	97.5	76.8	82.7	96.3	98.7	76.1	

The Kappa statistics for the North and the South were 89.4 and 95.3 %, respectively

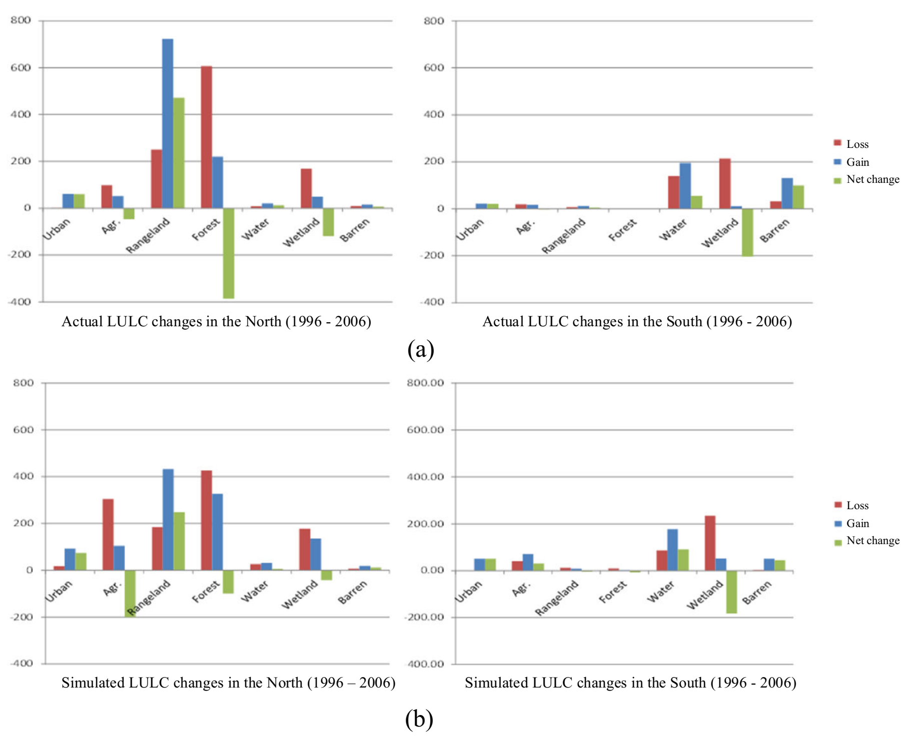

Fig. 6 Categorical distribution of the actual (**a**) and simulated (**b**) LULC changes between 1996 and 2006 (unit km^2)

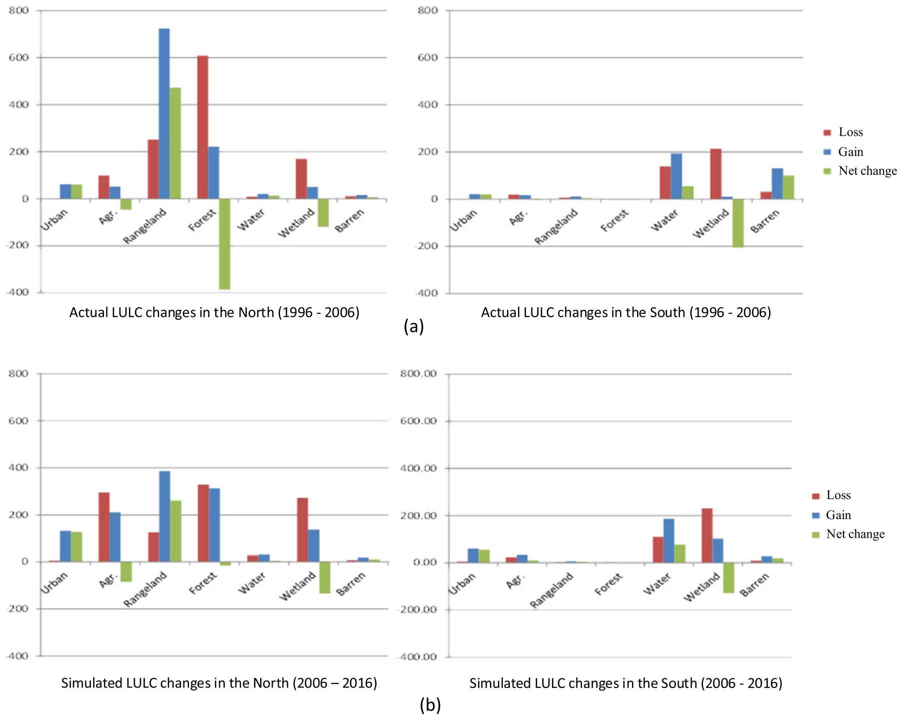

Fig. 7 Categorical distributions of the actual changes between 1996 and 2006 (**a**) and the simulated changes between 2006 and 2016 (**b**) (unit km^2)

Discussion

Modeling LULC changes is a very challenging task, as the changes may be caused by different processes and related to a wide range of variables. This study focused on one of the most vulnerable coastal areas where the human and natural systems are closely interacting. Given the complexity of the study region, the ANN would seem to be the most suitable modeling approach, as it can model non-linear relationships and researchers can build the model from ground zero. Since the study area has different landscapes and social-economical characteristics, the LULC processes in different areas would be different, which reasonably leads to the assumption that the LULC transitional rules would vary spatially. Thus, a global ANN derived from the whole study area would not apply accurately to different parts of the study area. This was confirmed through several trials where the globally derived ANN generated unrealistic simulations. This study solves this problem by partitioning the study area into the North and South, each sub-region is expected to have its own sets of natural and human processes. The results show that the two ANNs derived separately from the two sub-regions provided much better performance. Furthermore, unlike most other ANN simulations, this study incorporates a

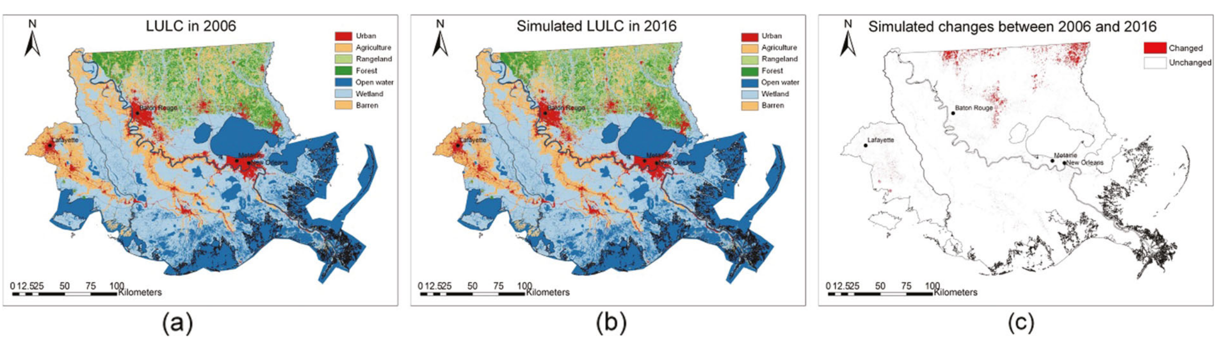

Fig. 8 Comparison between **a** the actual LULC in 2006, **b** simulated LULC in 2016, and **c** the changed areas between 2006 and 2016

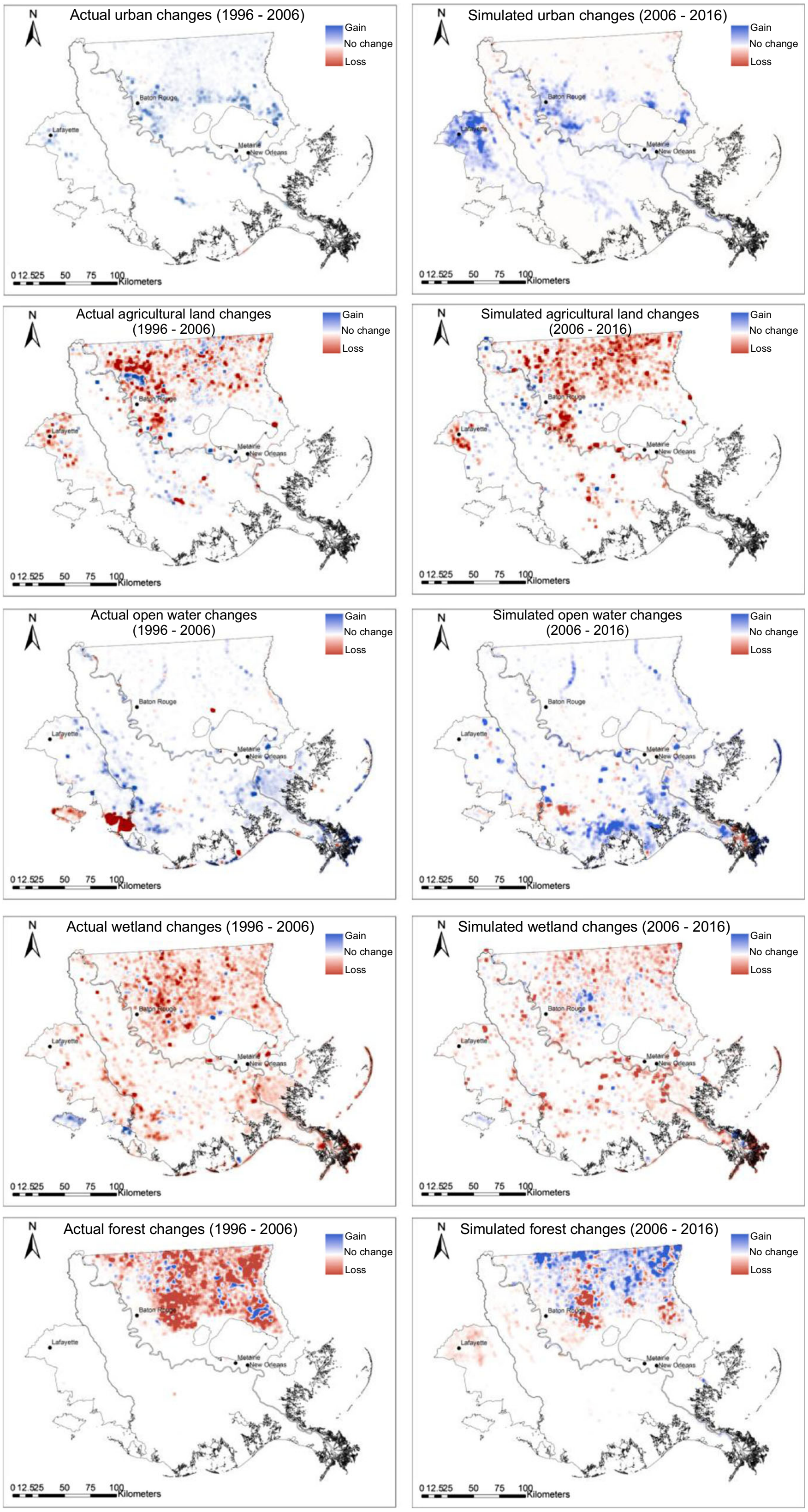

Fig. 9 Actual LULC changes 1996–2006 (*left column*) and simulated LULC changes 2006–2016 (*right column*). The two maps in the same row use the same color scheme

stochastic element in the simulation by using a weighted random selection of the top three LULC types based on the conversion probability. The stochastic element is designed to symbolize factors that were not included in the model; hence, the modeling approach used here could be considered a conceptual advancement.

There are several possible future improvements. First, although most of the variables (10 out of 15) were updated in the iterations, five variables including elevation, soil type, distance to primary roads, distance to secondary roads, and distance to pipelines stayed unchanged throughout the entire simulation. Updating these variables into dynamic variables would certainly help make the simulation more realistic. However, the updated data for these variables are not easily available, and they become a source of uncertainties in the simulation of future scenarios. Due to this shortcoming, simulation to longer time periods would be less and less reliable. It is noted that the five variables stayed unchanged in the simulation were variables referring to the infrastructure and physical environment, and we can reasonably assume that they did not change substantially in a 10-year period. Moreover, we could use the model to simulate “business-as-usual” scenarios, and then vary some of the parameters such as elevation changes as a result of sea-level rise to evaluate their impacts on the future landscape.

Second, our study assumed that the overall rate of LULC change was the same in the two time intervals, which was used as a criterion to stop the simulation. However, this assumption does not necessarily hold, as the rate of LULC changes may vary through time. Thus, the scenario simulated by the ANN has uncertainty in its exact temporal location. To solve this problem, additional information on how much LULC changes are expected in a 10-year period is needed to establish the temporal scale of the simulation. Third, the ANN is a black box data mining approach. The ANN does not tell which variable is most influential to the output and how one variable affects the other variables and the output. In addition, the current ANN model seemed to have a tendency to smooth out the magnitude of changes; thus, large LULC changes such as from forest to rangeland and urban areas in 1996–2006 in the North were not accurately captured (see Figs. 6a or 7a). As described earlier, the North has experienced significant population growth over the past decade, which would result in loss of forest land to other human activities. Further investigations about the underlying processes that trigger the changes should be conducted in the future. Other modeling approaches, such as decision trees and logistic regression will need to be explored and compared to derive the underlying landscape change rules.

The ANN training and simulation were implemented in both Matlab® and ArcGIS® platforms. Both platforms provide abundant tools that have greatly facilitated the model development process. However, a great deal of programming work was spent on bridging between the two platforms. For example, the LULC and input variable rasters had to be converted into a text file in order for them to be processed by Matlab®. In the simulation, the raster input variables were again needed to be formatted as value strings for the ANN calculation. The data in the two platforms could not be exchanged directly due to the memory issues. To the best of our knowledge, there is no GIS packages equipped with robust data mining tools. With the increasing demands in solving geospatial problems with data mining techniques, integrating data mining toolboxes into GIS packages will greatly benefit many users. In addition, techniques of parallel computing would be needed for simulation of a longer time span.

Conclusions

This study has demonstrated the use of ANN and CA in modeling LULC changes in a vulnerable coastal region, the Lower Mississippi River Basin in Louisiana. The study shows that ANN is an effective approach to derive the rules of LULC changes in an area with complex social and natural characteristics. Unlike traditional approaches, this study built separate models for the two sub-regions (the North and the South) with distinctive characteristics and incorporated a stochastic element for simulation. The results show that the ANN models, which were derived from a 10 % sample using 15 variables representing both natural and human characteristics, could simulate the LULC changes of the entire area from 1996–2006 with satisfactory degrees of accuracy, 91.8 % for the North and 97.1 % for the South, respectively. Based on the same models, a simulation to 2016 reveals several future trends of LULC changes in this region, in addition to identifying where the changes will most likely to occur. The simulation scenario shows that in the next 10 years, (1) urban land growth in both the South and the North will double (increased by 112 %

in the North and 166 % in the South); (2) agriculture land in the North will decline 84 % more than the past 10 years, amounting to a loss of 84 km^2; (3) wetland loss will accelerate by 14 % (amounting to a loss of 135 km^2 which is 3.2 % of total wetland) in the North, but will slow down by 37 % (still amounting to a loss of 128 km^2 which is 1.5 % of total wetland) in the South; (4) the trend of forest loss in the North will almost cease, changing from a loss of 386.6 km^2 to only 16 km^2; and (5) water expansion (land loss) will increase by 38 % in the South, which is equivalent to a land loss of 76.8 km^2 (0.4 % of the total land of the South). The study results provide a baseline condition of how and where LULC changes in relation to the 15 variables in a vulnerable coastal region. The findings could provide useful guidance for future refinement of the model and detailed investigation of the LULC processes. The scenario will be helpful to the planning and management in the region as it strives to become more sustainable.

Acknowledgments This material is based upon work supported by a research grant from the US National Science Foundation under the Dynamics of Coupled Natural Human Systems (CNH) Program (grant number: 1212112). Any opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agency.

References

Anderson, J.R., Hardy, E.E., Roach, J.T. & Witmer, R.E. (1976). *A land use and land cover classification system for use with remote sensor data*. Washington: United States Government Printing Office.

Badoe, D. A., & Miller, E. J. (2000). Transportation–land-use interaction: empirical findings in North America, and their implications for modeling. *Transportation Research Part D: Transport and Environment, 5*(4), 235–263.

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods, 43*(1), 3–31.

Batty, M. (2007). *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. Cambridge: The MIT press.

Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems, 23*(3), 205–233.

Berger, T. (2001). Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. *Agricultural Economics, 25*(2–3), 245–260.

Brown, D.G., Goovaerts, P., Burnicki, A., & Li, M.Y. (2002). Stochastic simulation of land-cover change using geostatistics and generalized additive models. *Photogrammetric Engineering and Remote Sensing, 68*(10), 1051–1061.

Burkett, V. R., Zilkoski, D. B., & Hart, D. A. (2002). Sea-level rise and subsidence: implications for flooding in New Orleans, Louisiana. In US Geological Survey Subsidence Interest Group Conference: Proceedings of the Technical Meeting, Galveston, Texas, 27-29 November 2001.

Clarke, K. C., & Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science, 12*(7), 699–714.

Clarke, K.C., Dietzel, C., & Goldstein, N.C. (2007). A decade of SLEUTHing: lessons learned from applications of a cellular automaton land use change model. *Classics in IJGIS: twenty years of the international journal of geographical information science and systems*, 413-427.

Coastal Protection and Restoration Authority of Louisiana (CPRA) (2012). Louisiana's coastal master plan for a sustainable coast. State of Louisiana, Baton Rouge, p. 189.

Crooks, A., Castle, C., & Batty, M. (2008). Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems, 32*(6), 417–430.

Day, J. W., Boesch, D. F., Clairain, E. J., Kemp, G. P., Laska, S. B., Mitsch, W. J., et al. (2007). Restoration of the Mississippi delta: lessons from hurricanes Katrina and Rita. *Science, 315*(5819), 1679–1684.

Dewan, A. M., & Yamaguchi, Y. (2009). Using remote sensing and GIS to detect and monitor land use and land cover change in Dhaka Metropolitan of Bangladesh during 1960-2005. *Environmental Monitoring and Assessment, 150*(1–4), 237–249.

Dobson, J.E., Ferguson, R.L., Field, D.W., Wood, L.L., Haddad, K.D., Iredale, I.H., et al. (1995). *NOAA Coastal Change Analysis Program (C-CAP): guidance for regional implementation.* US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service.

Feng, Y., & Liu, Y. (2012). A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing. *International Journal of Geographical Information Science, 27*(3), 449–466.

Foody, G. M. (1996). Relating the land-cover composition of mixed pixels to artificial neural network classification output. *Photogrammetric Engineering and Remote Sensing, 62*(5), 491–498.

García, A. M., Santé, I., Boullón, M. & Crecente, R. (2013). Calibration of an urban cellular automaton model by using statistical techniques and a genetic algorithm. Application to a small urban settlement of NW Spain. *International Journal of Geographical Information Science*, 1-19.

Gong, P. (1996). Integrated analysis of spatial data from multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping. *American Society for Photogrammetry and Remote Sensing, 62*(5).

González, J. L., & Tornqvist, T. E. (2006). Coastal Louisiana in crisis: subsidence or sea level rise? *Eos. Transactions American Geophysical Union, 87*(45), 493–498.

Jain, A. K., Jianchang, M., & Mohiuddin, K. M. (1996). Artificial neural networks: a tutorial. *Computer, 29*(3), 31–44.

Ju, W., & Lam, N. (2007). Urban land use classification: applying texture analysis and artificial intelligence. *Imaging Notes, 22*(3), 26–30.

Lam, N.S.N., Arenas, H., Li, Z., Liu, K.B. (2009a). An estimate of population impacted by climate change along the U.S. Coast. *Journal of Coastal Research*, 1522-1526.

Lam, N.S.N., Arenas, H., Pace, K., LeSage, J., & Campanella, R. (2012). Predictors of business return in New Orleans after Hurricane Katrina. *PloS One, 7*(10), e47935.

Lam, N.S.N., Pace, K., Campanella, R., LeSage, J., & Arenas, H. (2009b). Business return in New Orleans: decision making amid post-Katrina uncertainty. *PloS One, 4*(8), e6765.

Lambin, E. F., Geist, H. J., & Lepers, E. (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources, 28*(1), 205–241.

Lauf, S., Haase, D., Hostert, P., Lakes, T., & Kleinschmit, B. (2012). Uncovering land-use dynamics driven by human decision-making—a combined model approach using cellular automata and system dynamics. *Environmental Modelling & Software, 27–28*, 71–82.

Leemans, R., & Eickhout, B. (2004). Another reason for concern: regional and global impacts on ecosystems for different levels of climate change. *Global Environmental Change, 14*(3), 219–228.

Li, X., & Yeh, A. G.-O. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science, 16*(4), 323–343.

Li, X., & Yeh, A. G.-O. (2004). Data mining of cellular automata's transition rules. *International Journal of Geographical Information Science, 18*(8), 723–744.

Liu, J., Dietz, T., Carpenter, S.R., Folke, C., Alberti, M., Redman, C.L., et al. (2007). Coupled human and natural systems. *AMBIO: A Journal of the Human Environment, 36*(8), 639–649.

Liu, X., Li, X., Shi, X., Huang, K., & Liu, Y. (2012). A multi-type ant colony optimization (MACO) method for optimal land use allocation in large areas. *International Journal of Geographical Information Science, 26*(7), 1325–1343.

Liu, X., Li, X., Shi, X., Zhang, X., & Chen, Y. (2010). Simulating land-use dynamics under planning policies by integrating artificial immune systems with cellular automata. *International Journal of Geographical Information Science, 24*(5), 783–802.

Lo, C. P., & Yang, X. (2002). Drivers of land-use/land-cover changes and dynamic modeling for the Atlanta, Georgia metropolitan area. *Photogrammetric Engineering and Remote Sensing, 68*(10), 1073–1082.

Mahajan, Y., & Venkatachalam, P. (2009). Neural network based cellular automata model for dynamic spatial modeling in GIS. In O. Gervasi et al. (Eds.), *Computational science and its applications—ICCSA 2009* (pp. 341–352). Heidelberg: Springer Berlin.

Murray-Rust, D., Rieser, V., Robinson, D.T., Milicic, V., & Rounsevell, M. (2013). Agent-based modelling of land use dynamics and residential quality of life for future scenarios. *Environmental Modelling & Software, 46*, 75–89.

Okwuashi, O., Isong, M., Eyo, E., Eyoh, A., Nwanekezie, O., & Olayinka, D.N., et al. (2012). GIS Cellular automata using artificial neural network for land use change simulation of Lagos, Nigeria. *Journal of Geography and Geology, 4*(2).

Overmars, K. P., De Koning, G. H. J., & Veldkamp, A. (2003). Spatial autocorrelation in multi-scale land use models. *Ecological Modelling, 164*(2–3), 257–270.

Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M.J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of the Association of American Geographers, 93*(2), 314–337.

Perez-Vega, A., Mas, J. F., & Ligmann-Zielinska, A. (2012). Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environmental Modelling & Software, 29*(1), 11–23.

Reyes, E., White, M. L., Martin, J. F., Kemp, G. P., Day, J. W., & Aravamuthan, V. (2000). Landscape modeling of coastal habitat change in the Mississippi delta. *Ecology, 81*(8), 2331–2349.

Schweitzer, C., Priess, J. A., & Das, S. (2011). A generic framework for land-use modelling. *Environmental Modelling & Software, 26*(8), 1052–1055.

Stewart, T. J., Janssen, R., & Van Herwijnen, M. (2004). A genetic algorithm approach to multiobjective land use planning. *Computers & Operations Research, 31*(14), 2293–2313.

Stokstad, E. (2005). Louisiana's wetlands struggle for survival: new focus. *Science, 310*(5752), 1264–1266.

Turner, B. L. (2010). Vulnerability and resilience: coalescing or paralleling approaches for sustainability science? *Global Environmental Change, 20*(4), 570–576.

Vicari, A., Alexis, H., Del Negro, C., Coltelli, M., Marsella, M., & Proietti, C. (2007). Modeling of the 2001 lava flow at Etna volcano by a cellular automata approach. *Environmental Modelling & Software, 22*(10), 1465–1471.

Wang, F. (2007). *Land-cover and land-use study using genetic algorithms, petri nets, and cellular automata.* (Ph.D. dissertation). Louisiana State University.

Wang, Z.M., Zhang, B., Zhang, S.Q., Li, X.Y., Liu, D.W., & Song, K.S., et al. (2006). Changes of land use and of ecosystem service values in Sanjiang Plain, northeast China. *Environmental Monitoring and Assessment, 112*(1–3), 69–91.

Ward, D. P., Murray, A. T., & Phinn, S. R. (2000). A stochastically constrained cellular model of urban growth. *Computers, Environment and Urban Systems, 24*(6), 539–558.

Wear, D. N., & Bolstad, P. (1998). Land-use changes in southern Appalachian landscapes: spatial analysis and forecast evaluation. *Ecosystems, 1*(6), 575–594.

Wu, F., & Webster, C. J. (1998). Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and Design, 25*(1), 103–126.

Zhou, G. (2006). *Detecting the social-economic conditions of urban neighborhoods through a combined methodology of wavelet transform and artificial neural networks.* (Ph.D. dissertation). Louisiana State University.

Zhou, J., & Civco, D. L. (1996). *Using genetic learning neural networks for spatial decision making in GIS*. Bethesda: ETAT S-UNIS: American Society for Photogrammetry and Remote Sensing.