International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

A spatial dynamic model of population changes in a vulnerable coastal environment

Kenan Li & Nina S. N. Lam

To cite this article: Kenan Li & Nina S. N. Lam (2018) A spatial dynamic model of population changes in a vulnerable coastal environment, International Journal of Geographical Information Science, 32:4, 685-710, DOI: 10.1080/13658816.2017.1407415

To link to this article: https://doi.org/10.1080/13658816.2017.1407415

Published online: 27 Nov 2017.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE, 2018
VOL. 32, NO. 4, 685–710
https://doi.org/10.1080/13658816.2017.1407415

ARTICLE

A spatial dynamic model of population changes in a vulnerable coastal environment

Kenan Li and Nina S. N. Lam

Department of Environmental Sciences, Louisiana State University, Baton Rouge, LA, USA

ABSTRACT

This study developed a spatial dynamic model to examine the coupled natural–human responses in the form of changes in population and associated developed land area in the Lower Mississippi River Basin region. The goal was to identify key socio-economic factors (utility) and environmental factors (hazard damages, elevation, and subsidence rate) that affected population changes, as well as to examine how population changes affected the local utility and the local environment reciprocally. We first applied areal interpolation techniques with the volume-preserving property to transform all the data at Year 2000 into a unified 3 km by 3 km cellular space. We then built an Elastic Net model to extract 12 variables from a set of 33 for the spatial dynamic model. Afterward, we calibrated the neighborhood effects with a genetic algorithm and use the spatial dynamic model to simulate population and developed land area in 2010. Furthermore, we took a Monte Carlo approach for analyzing the uncertainty of the model outcome. Our accuracy assessment shows that the model on average slightly overpredicts the number of population and the developed land percentage at 2010, as indicated by the low values of mean absolute deviation (MAD) due to quantity. On the other hand, the MADs due to allocation are larger than the MADs due to quantity, with most outliers found in the New Orleans region where population and urban development declined significantly during 2000–2010 after Hurricane Katrina. The proposed model sheds light on the complex relationships between coastal hazards and human responses and provides useful insights to strategic development for coastal sustainability.

ARTICLE HISTORY
Received 30 June 2016
Accepted 16 November 2017

KEYWORDS
Coupled natural–human system; population dynamics; coastal hazards; resilience; Mississippi Delta

1. Introduction

Coastal areas are subject to natural coastal hazards, including land loss, land subsidence, coastal erosion, coastal flooding, tsunami, sea level rise, storm surges, and hurricanes. All these hazards have negative effects on the sustainable development of the communities along the coast. In the United States, the Lower Mississippi River Basin (LMRB) in southern Louisiana is highly vulnerable to coastal hazards (Figure 1). During 2005–2015, the region experienced five hurricanes (Katrina, Rita, Gustav, Ike, and Isaac), which caused substantial loss of human lives and damages to properties (Lam *et al*. 2009a, 2009b, 2012, LeSage *et al*., 2011a, 2011b, 2011c), plus the fact that this region was also experiencing serious land subsidence and land loss

CONTACT Kenan Li kli4@lsu.edu

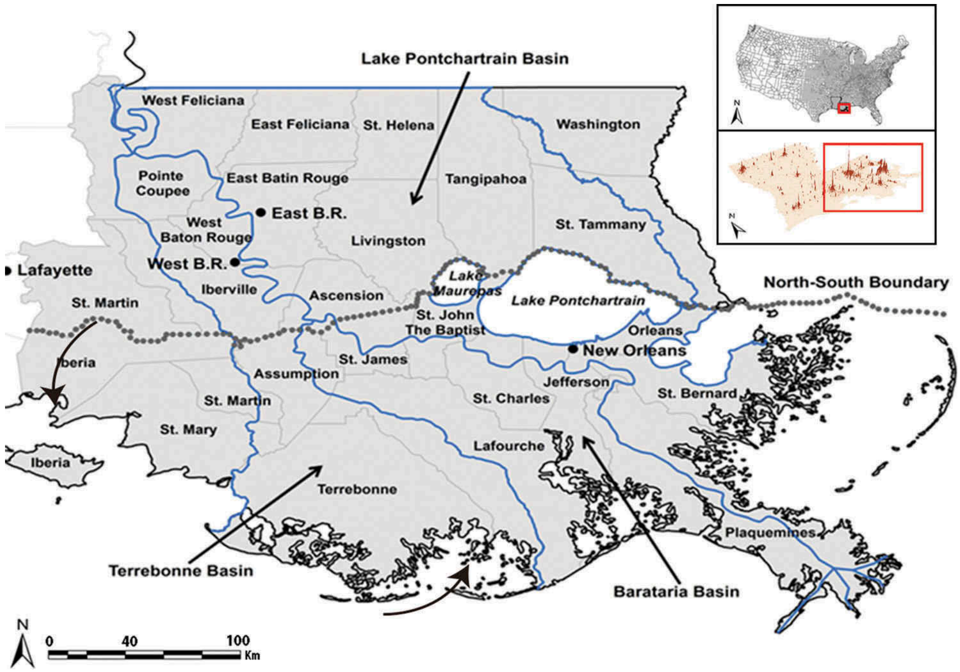

Figure 1. The study area with a hypothetical north–south boundary and its population distribution in 2006 (based on LandScan data).

problem (Zou *et al.* 2016). With the impending threats of climate change and sea level rise, coastal Louisiana is facing a serious challenge to protect the land from disappearing while maintaining economic development, or in other words, to achieve coastal sustainability(Cai *et al.* 2016, Lam *et al.* 2016).

There has been extensive research on the causes and consequences of increased land loss vulnerability in the LMRB region (Blum and Roberts 2009). Moreover, substantial efforts from governmental and nongovernmental agencies have also been made to develop long-term master plans for coastal protection and restoration (Coastal Protection and Restoration Agency (CPRA) of Louisiana, 2012). However, existing literature on coastal protection and restoration has largely focused on understanding the natural environment, and few studies addressed the issues of human–environmental interactions. Furthermore, very few publications examined the coupled natural and human (CNH) systems in this coastal environment (Twilley *et al.* 2016). Studying how the natural and human environments interact in the study region will contribute to a better understanding of the dynamics of the coastal vulnerability and sustainability problem, not only in this region but also in coastal regions worldwide (Liu J. *et al.*, 2007a, 2007b, Collins *et al.* 2011; Kates 2011). In particular, a gradual population growth has been observed over the past decade in the northern part of the LMRB, in contrast with a dramatic population decline in the southern coastal area (Qiang and Lam 2015, 2016). In fact, some of the population increase in the northern LMRB came from the southern part of the region (Plyer 2013). This phenomenon signals a voluntary inland migration that is not part of any governmental coastal restoration plan. Eliciting the decision factors that make people move in a vulnerable coastal region is critical to the overall research on climate change adaptation and coastal sustainability.

In summary, the objective of this study is to develop a spatial dynamic model to understand the CNH responses as reflected in changes in population and developed land area in the LMRB during 2000–2010. With this objective, this study seeks to answer three related questions: (1) what are the key socioeconomic factors (defined as utility hereafter) that trigger population changes? (2) how important are environmental factors (such as exposure and damages from natural coastal hazards, elevation, and subsidence rate) in affecting population changes in the region? and (3) reciprocally, how do population changes affect local utility and local natural environment, such as by increasing developed land area with more roads or canals, and consequently fragmenting the landscape, and therefore leading to more land loss. Answers to these questions can provide valuable insights into the development of better strategies for coastal sustainability. Finally, once a reliable spatial dynamic model is developed and validated, it can be used by planners and decision makers to simulate patterns of population change and evaluate long-term resilience and sustainability in the region under different climate change or policy scenarios.

2. Literature review and conceptual framework development

In assessing the resilience and sustainability of a place-based system, one must consider both the natural and the human systems and evaluate how both systems are coupled (Bolin *et al*. 2000, Liu *et al*. 2007a, 2007b, Collins *et al*. 2011; Kates 2011). Developing a spatial dynamic model to understand and assess the coupling of human and natural environments in a manner that could lead to resilience and sustainability of a region will be very useful. The term CNH system was evolved to describe a branch of interdisciplinary study that examines how human components interact with natural components in an integrated ecological and social system (Liu *et al*. 2007a, 2007b). With consideration of the autonomous human factor, the traditional 'top-down' approach of coastal restoration from the planners and decision makers may not be effective in coping with the resilience and sustainability issues. In planning for coastal sustainability, we must consider the 'bottom-up' facts of human responses and their decisions can provide complementary insights into land resource management and economic development.

2.1 *'Bottom-up' methods in modeling CNH dynamics*

Literature on population changes and CNH dynamics is extensive. Two major modeling tools have been widely employed to understand the complexity: cellular automata (CA) (Batty *et al*. 1994, 1997, Clarke and Gaydos 1998, Malanson *et al*. 2006a, 2006b, Qiang and Lam 2015) and agent-based modeling (ABM) (Brown and Xie 2006).

A CA model uses cells with different states as the smallest modeling units, and defines the rules for updating the states over time with consideration of the neighborhood effects. In a CA model, each cell has a state space, which value is controlled by a set of CA rules to determine the state of each cell at each time step according to its previous state and its neighbor cells' previous states. In some CA models (Boccara and Cheong 1992, 1993), the cells can also have mobility with certain defined site-exchange rules. Figure 2 shows the structure of a CA model.

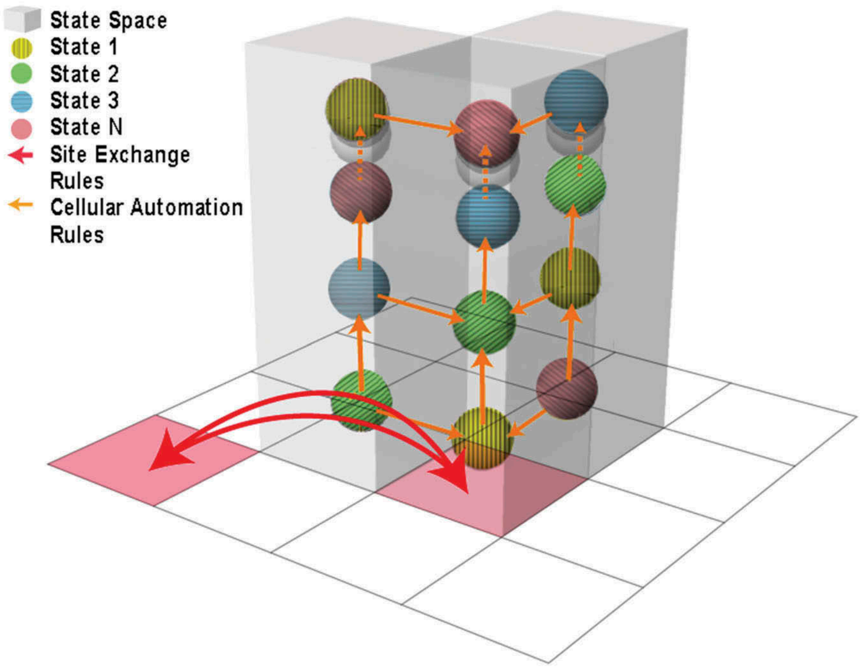

Figure 2. Structure of a cellular automation model.

CA models have been extensively used to model urban development, such as the Slope, Land-use, Exclusion, Urban extent, Transportation, and Hill-shade (SLEUTH) model (Clarke *et al*. 1997), the dynamic urban evolution model (Xie 1997), the multi-criteria evaluation-CA model (Wu and Webster 2000), the multi-agent system-CA model (Ligtenberg *et al*. 2001), the GeoCA-Urban model (Zhou *et al*. 1999), the Voronoi-CA model (Shi and Pang 2000), and the Markov-CA model (Vaz *et al*. 2012.). The major differences among these cell-based methods are the factors used for the development prediction. For example, the SLEUTH model is based on historical trends of land use change data to derive transition rules to simulate future land use under different developmental conditions. However, most of these land use change models are not built in a CNH context, and they seldom consider the coupled feedbacks between urban development, population growth, and socioeconomic factors like the one used in this study.

ABM is an approach using agents to simulate the behavior of a complex system. There is no universal agreement of the definition of the term 'agent' in the context of ABM. Some studies considered any type of independent component, either a software component or a sub-model, as an agent (Bonabeau 2002). Other researchers insist that an agent's behavior must include adaptive components to represent the responses of the agents to the environment (Conte 2002). Jennings (2000) from a computer science perspective emphasized the autonomous behavior of an agent an essential characteristic in an ABM. Figure 3 shows the basic structure of an ABM, which emphasizes some common points shared by these various definitions. These common characteristics of ABMs include the following: (1) agents have the ability to change their physical locations

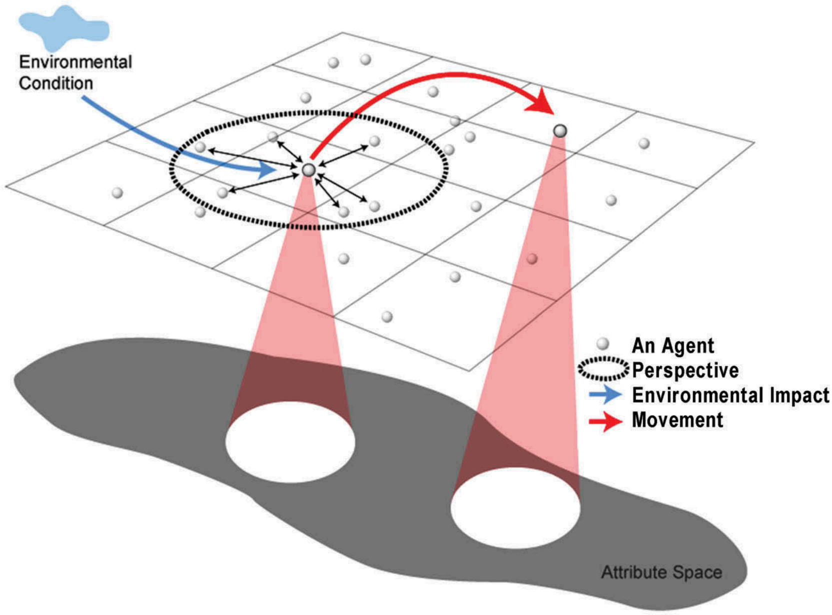

Figure 3. Structure of an agent-based model.

and modify their attributes, (2) agents can actively sense the environment conditions and respond accordingly, and (3) agents can interact with other agents within their perspectives.

There are substantial overlaps between CA and ABM approaches. Both share similar issues in modeling CNH dynamics from the bottom-up. Two major problems are (1) what are the factors to be included in the dynamics and (2) how to explicitly define the parameters of the dynamics using empirical information. For the first problem, different studies in CNH used their different favorite decision factors, such as human life cycle, economic motivation, neighborhood facilities, environmental amenities, and/or housing quality (Lindberg *et al*. 1992, Nijkamp *et al*. 1993, Dökmeci and Berköz 2000; Torrens 2001, Yin and Muller 2007, Niedomysl 2008). Fontaine and Rounsevell (2009) integrated household life-cycle events in their HI-LIFE ABM. However, with few exceptions (Schultz and Elliott 2012), few studies modeled population changes and CNH dynamics in the context of hazard, vulnerability, or resilience.

For the second problem, Li and Liu (2007) determined the parameters of their ABM according to multi-criteria evaluation techniques. However, their method assumed all the modeling units performed the same way as experts. Recent urban residential dynamics studies extended traditional urban economic theory by incorporating human decisions and ABM to model urban development (Krugman 1991, Fujita and Thisse 2002, Irwin and Bockstael 2002, Cavailhès *et al*. 2004, Caruso *et al*. 2007, Munroe 2007, Li and Liu 2008, Parker and Filatova 2008, Chen *et al*. 2012). Notwithstanding that these studies revealed the dynamics of human decisions, they have difficulties in validation using empirical information. Hence, to overcome the empirical validation issues, some empirical statistical models gain popularity by specifying the local spatial dynamics. For example, the discrete choice model uses conditional logit regression equation to predict parameters (McFadden 1978, Bruch and Mare 2012). Inspired by the above methods, we developed an Elastic Net model in this study to extract the rules of the coupled spatial dynamics and applied them to the spatial units, and afterward we calibrated the neighborhood effects by using genetic algorithms (GAs)

2.2 *Conceptual framework development*

The conceptual framework of the spatial dynamic model in this study includes three parts: the modeling units, the system dynamic (SD) functions, and the essential characteristics of the modeling units, as shown in Figure 4.

The design of the modeling units faces two challenges. First, the typical large number of bottom-level entities will take up a large amount of computation time on hundreds of runs that are required for model testing. Second, the data for individual entities on human dynamics are often available only in aggregative forms, such as at the county, census tract, or census block level. In order to build an empirical model for a relatively large geographic scale such as the LMRB in this study, modeling units defined in an aggregation form are necessary for the sake of computational loads

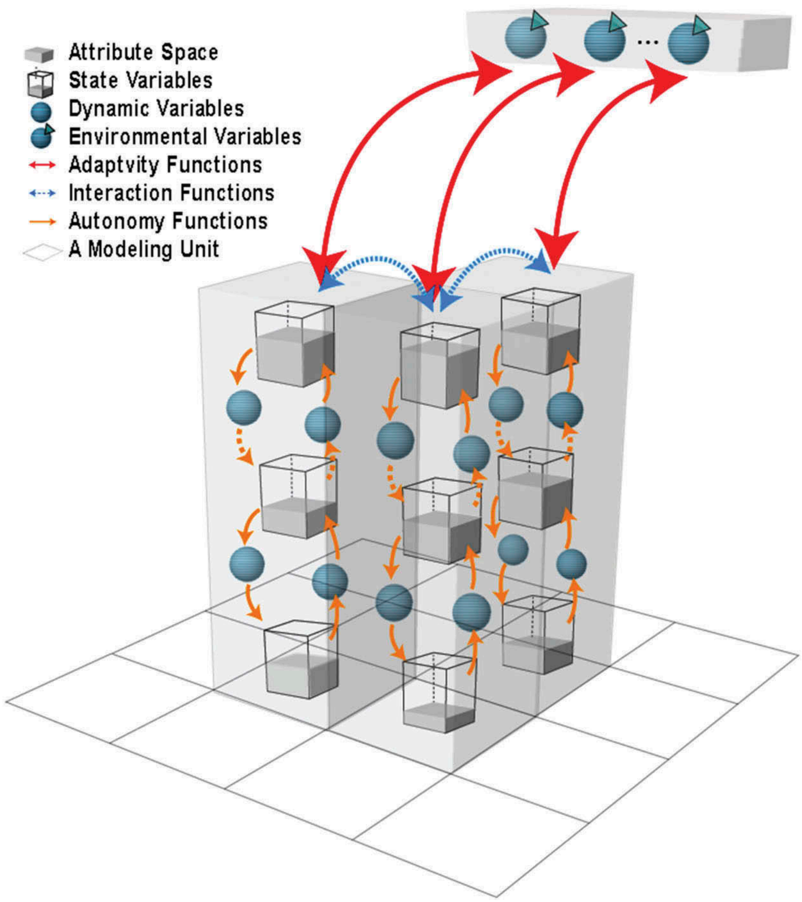

Figure 4. Structure of the spatial dynamic model used in this study.

(Scheffer *et al*. 1995, Hellweger 2008). For those reasons, the modeling units in this study are defined as 3 km by 3 km spatially situated cells. In essence, a modeling unit is a spatial land unit with homogeneous social and economic characteristics, and it is used as the coherent unit of analysis for modeling the CNH dynamics.

The SD procedure was used to understand the behavior of the modeling units. SD is a methodology originated by Forrester (1961, 1969) to frame, model, and understand complex systems. In SD, state variables (dependent variables of the updating functions) are the internal attributes of a modeling unit, of which the changes are to be updated by the SDs. The attributes that contribute to the changes of the state variables are called dynamic variables, which are independent variables in the updating functions. This study utilizes Elastic Net as the method to derive the SD functions between the state variables and the dynamic variables using empirical data.

The proposed model reflects three principal characteristics of ABMs: autonomy, interaction ability, and adaptivity. The autonomous characteristic is that a spatial unit is able to control its internal state. As for the latter two characteristics, the spatial units are able to interact and communicate with their neighbors, and are able to actively perceive their environment and react accordingly. The study focuses on the three principles to develop the SDs, but the model is different from ABMs, because the spatial units are situated cells and lack mobility.

3. Data and methods

3.1 *Data description*

All the data used in this study were collected from the following publicly available databases. Natural hazard damage data were obtained from the 'Storm Data and the Storm Events Database' maintained by National Oceanic & Atmospheric Administration's (NOAA) National Climate Data Center. These hazard damage data are aggregated to one of the three spatial units: counties or zones (a combination of several counties), metropolitan areas, or point locations with longitudes and latitudes. This study considers five major types of coastal hazards: coastal (which includes coastal flooding and storm surge), flood, hurricane, thunderstorm, and tornado. Demographic and socioeconomic data were obtained from the U.S. Census Bureau. Population count data were obtained at the census block level, whereas the other social and economic variables were obtained at the census tract level, as many of them are not available at the census block level. Environmental data, including elevation, water body locations, and road transportation networks, were obtained from the National Map View developed by the U.S. Geological Survey. Land cover and land use data were obtained from the Multi-Resolution Land Characteristics (MRLC) Consortium. Subsidence Rates data were from the NOAA Technical Report 50: Rates of Vertical Displacement at Benchmarks in the Lower Mississippi Valley and the Northern Gulf Coast. Energy structure data such as the location of pipelines and oil/gas wells were from the Louisiana Department of Natural Resources Oracle database. All the variables derived from these datasets for the model building are listed in Table 1.

Table 1. Acronyms and descriptions of the variables used in this study.

Categories	Acronym	Description	Year
Housing	Occupied	Percent; Occupied Housing Units	2000, 2010
	NonVehicle	Percent; Occupied Housing Units with No Vehicles Available	2000, 2010
	NonFuel	Percent; Occupied Housing Units with No House Heating Fuel Used	2000, 2010
	NonPlumb	Percent; Occupied Housing Units Lacking Complete Plumbing Facilities	2000, 2010
	NonKitchen	Percent; Occupied Housing Units Lacking complete Kitchen Facilities	2000, 2010
	NonTele	Percent; Occupied Housing Units with No Telephone Service	2000, 2010
	NonMtg	Percent; Specified Owner-Occupied Units without Mortgage	2000, 2010
	OwnerR	Percent; Owner-Occupied Housing Units	2000, 2010
	MedValue	Number; Median Value of Specified Owner-Occupied Units (Dollars)	2000, 2010
	MedRent	Number; Median Gross Rent of Specified Renter-Occupied Units (Dollars)	2000, 2010
Households	OCST20	Percent; Owner Cost as a Percentage of Household Income Less than 15%	2000, 2010
	OCST35	Percent; Owner Cost as a Percentage of Household Income More than 35%	2000, 2010
	Rent15	Percent; Gross Rent as a Percentage of Household Income Less than 15%	2000, 2010
	Rent35	Percent; Gross Rent as a Percentage of Household Income More than 35%	2000, 2010
	OCSTWMTG	Number; Median Selected Monthly Owner Costs with a Mortgage (Dollars)	2000, 2010
	OCSTNMRG	Number; Median Selected Monthly Owner Costs without a Mortgage (Dollars)	2000, 2010
	HhSize	Number; Average Household Size	2000, 2010
	MeanTime	Number; Mean Travel Time to Work (Minutes)	2000, 2010
	MedIcm	Number; Households Median Income (Dollars)	2000, 2010
Individuals	Female	Percent; Female Population	2000, 2010
	Population	Number; Total Population	2000, 2010
	Under5	Percent; Population under 5 Years Old	2000, 2010
	Over65	Percent; Population over 65 Years Old	2000, 2010
	HighSch	Percent; Population over 25 Years Old with High School Graduation or Higher	2000, 2010
	Married	Percent; Population over 15 Years Old and Now Married (Except Separated)	2000, 2010
	Employed	Percent; Population over 16 Years Old Employed	2000, 2010
	Poverty	Percent; Individuals below Poverty Level	2000, 2010
Human Structures	Road	Number; Road Density	2007
	Pipeline	Number; Pipeline Density	2007
	GasWell	Number; Oil and Gas Injection Wells Density	2007
Environmental	Damages	Number; Property Damages in 2010 Inflation Rate (Dollars)	2000 to 2010
	Subsidence	Number; Subsidence Rate Interpolated by Empirical Bayesian Kriging Using BenchMarks	2004
	Elevation	Number; Mean Elevation Calculated from LIDAR Images (Meters)	2000 to 2010
Land Use and Land Cover	Developed	Percent; Percent of Developed Land Use Area	2001, 2011
	Water	Percent; Percent of Open Water Land Use Area	2001, 2011

3.2 *Areal interpolation*

The data obtained in this study were in different temporal and spatial scales, such as subsidence data in point form, census data at the block, block-group, or census-tract scale, and land use and land cover data in pixels. Moreover, the geographic boundaries of the census data, especially at the block level, changed notably over the study time span. The study applied areal interpolation methods to incorporate all these data from heterogeneous sources and units into a unified set of geographic spatial units to make them spatially and temporally comparable.

The term 'areal interpolation' was first coined in Goodchild and Lam (1980) to denote the problem of transforming data defined in one set of areal units (source zones) to another (target zones), where the two sets of boundaries do not coincide (Lam 1980, 1983, 2009). The target zones for the areal interpolation in this study were the 3 km by 3 km modeling units. The areal interpolation procedure used in this study was an 'intelligent' area-weighting method that has the volume-preserving property (Shu *et al*. 2010, Cromley *et al*. 2012). The developed land cover area was used as an intelligent or ancillary layer, also called a 'control' variable. Areal interpolation with additional control variables is very similar to the principle of dasymetric mapping, a mapping technique designed to reflect within-zone variations (Lam 1983, 2009).

Ancillary data can substantially improve the accuracy of areal interpolation. Instead of using the total area of the source zone to derive the weights in area-weighting interpolation, the study only considered the developed area under the assumption that only developed land has the socioeconomic characteristics of interest to population change. The interpolated value of a given target zone is the areally weighted mean of the values of the source zones intersected with it, and the weights are determined by the total areas of these intersected areas on the ancillary layer. This areal interpolation method was applied only to the variables defined with source-zone boundaries such as the census variables. For the variables that are originally in raster format such as road density, zonal aggregation determines the call values. Figure 5 shows the areal interpolation process with ancillary data in this study.

In Figure 5, the different tints in the source-zone layer represent the original values of the variable of interest. The ancillary layer is a binary layer with the shaded area representing the control variable, which is the developed land cover in this study. Figure 6 shows the values of the source zones (by census block) before the areal interpolation and the values of the target zones (defined in 3 km by 3 km grid cells) after the interpolation. The interpolated pattern resembles closely to the original pattern.

3.3 *Model building*

Elastic Net, a new regularization and variable selection method proposed by Zou and Hastie (2005), was used to select the variables and derive relationships for building the spatial dynamics model. Elastic Net is a hybrid of the Lasso regression (Tibshirani 1996) and the Ridge regression (Hoerl and Kennard 1988), both of which add a penalty term in the fitting function of ordinary linear regression, and the penalty terms are often referred to as L_1 and L_2 penalty, respectively. Lasso regression selects variables by

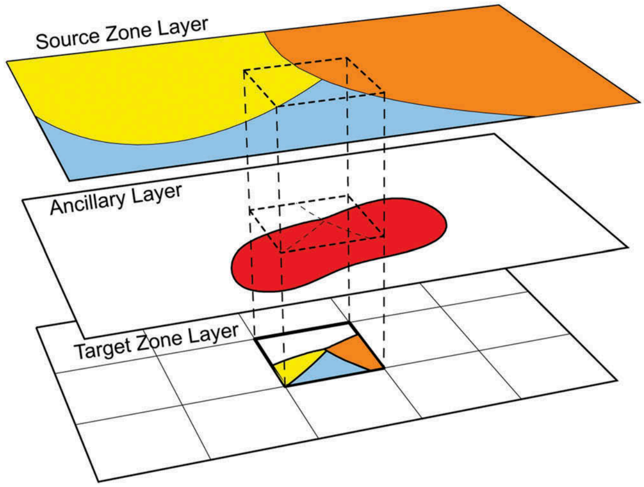

Figure 5. The areal interpolation method with an ancillary layer.

minimizing multicollinearity, whereas Ridge regression minimizes the overfitting problem. As a combination, Elastic Net automatically selects the variables to reduce multicollinearity as the Lasso and keeps a high prediction performance when collinearity exists as the Ridge (Zou and Hastie 2005). The regularization problem of Elastic Net is defined as minimizing the parameters β_0 and β_k in Expression 1.

$$\frac{1}{2N}\sum_{i=1}^{N}\left(y_i-\beta_0-\sum_{k=1}^{m}X_{ki}\beta_k\right)^2+\lambda\sum_{k=1}^{m}((1-\alpha)\beta_k^2+\alpha\mid\beta_k\mid) \tag{1}$$

where N is the number of observations, y_i is the value of the dependent variable at observation i. X_{ki} is the value of the kth independent variable at observation i, λ is a positive regularization parameter, β_0 is an intercept to be estimated, β_k is the coefficient of the kth independent variable to be estimated, m is the number of independent variables, and α (ranging from 0 to 1) is the parameter that determines if the regression model is more like a Lasso or a Ridge regression. The Elastic Net is the same as a Lasso regression when α equals 1. On the contrary, as α shrinks toward zero, the Elastic Net approaches a Ridge regression. The parameter λ determines the number of independent variables to be included. As λ increases, the number of nonzero β_k decreases, which means the number of variables selected decreases, and vice versa. Studies have shown that Elastic Net has an advantage over stepwise regression because it generalizes better to new samples (Whittingham *et al*. 2006). Hence, we used Elastic Net in this study to select variables to develop a simpler utility model.

In this study, population count, developed land area percentage, and utility (defined later) were the state variables in the SDs. Their temporal changes between 2000 and 2010 were calculated and used as dependent variables in the Elastic Net. The other variables at 2000 were independent variables (Table 1). In order to eliminate the data scale impact, all the variables, including their temporal changes, were standardized before running the Elastic Net.

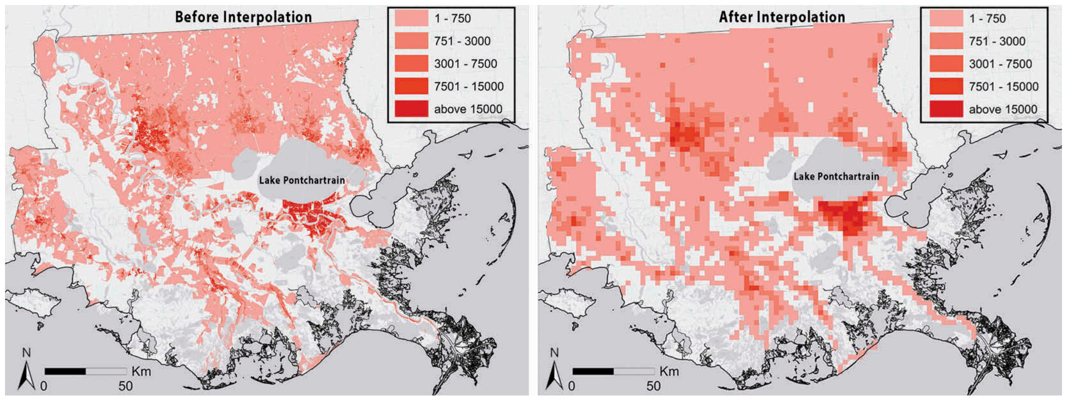

Figure 6. Population count before (by census blocks) and after areal interpolation (by 3 km by 3 km grids).

We used the derived Elastic Net regression equations as the difference functions to update the modeling units, with their coefficients divided by 10 (the number of years in this study) to project the changes of the stock variables for 10 years. Thus, each time step of running the difference equations in the model represents 1 year. Afterward, a neighborhood effect term was added to each difference equation to reflect and test the neighborhood interactivity. We defined the neighborhood as the Moore neighborhood (the eight contiguous neighbors), and used a GA to calibrate the neighborhood effect parameters. The fitness function was the total mean squared errors (MSEs) between the real and the simulated population count and developed land area percentage. By using this fitness function, we assumed that the simulation accuracies of population count and developed area percentage are equally important on training the neighborhood impact parameters. We set the tuning ranges of these parameters from 0 to 1, with a value of 0 indicating no direct neighborhood impacts, whereas a value of 1 implying no autonomy (in other words, no self-control).

The spatial dynamic model is subject to three additional constraints: (1) if the population of a spatial unit is below zero, then it cannot receive negative flows to its population in the next time step; (2) if the developed area percentage of a spatial unit is over 100%, then it cannot receive positive flows to its developed area percentage in the next time step; (3) since the historical data show no decrease in developed area, it was assumed that the developed area percentage cannot have negative flows. If a negative flow is projected by the difference functions, then it is set to zero.

The study area consists of a total of 5890 modeling units (cells). When building the Elastic Net, the value of α was set to 0.5 to consider the equal contributions (50%) of L_1 and L_2 penalties. We used a tenfold cross-validation procedure for iterative calibration and validation. The procedure divides the whole dataset in 10 portions, 9 of them are training datasets for calibration and the 10th one is the test dataset for validation. The cross-validated MSE was used to determine the value of λ and the degree of freedom (number of predictor variables selected). Due to data availability, this validation procedure was not based on time, but across space (Pontius *et al*. 2008a, 2008b). A more robust time-based validation method needs to include data from at least two temporal periods (e.g. 2000–2010 and 2010–2020).

3.4 *Accuracy assessment*

The simulated changes during 2000–2010 were compared with the reference (real) changes in the same period to provide an accuracy assessment of the proposed spatial dynamic model. We employed the accuracy assessment method developed by Pontius *et al*. (2008a, 2008b, 2011) and separated the model error inferred by the total MAD (*MAD_total*) into two components: *MAD_quantity* and *MAD_allocation*. *MAD_quantity* is the absolute difference between the mean real value and the mean simulated value, whereas *MAD_allocation* is the difference between *MAD_quantity* and *MAD_total*. Equations (2)–(4) illustrate how *MAD_total, MAD_quantity*, and *MAD_allocation* are derived:

$$MAD\_total = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \tag{2}$$

$$MAD\_quantity = \left| \frac{\sum_{i=1}^{n} y_i}{n} - \frac{\sum_{i=1}^{n} \hat{y}_i}{n} \right| \tag{3}$$

$$MAD\_allocation = MAD\_total - MAD\_quantity \tag{4}$$

where n is the number of spatial units, y_i is the true population count (or developed land area percentage) of the ith cell, and $\hat{y}_i$ is the predicted value of the ith cell. The MAD values for each comparison (simulated changes vs. reference changes) and for each variable (population and developed land area percentage) were computed and plotted. The accuracy assessment results are summarized below in the 'Results' section.

3.5 *Monte Carlo uncertainty analysis*

We applied a Monte Carlo probabilistic uncertainty analysis using simple random sampling to examine the propagation of uncertainty errors and identify the most important contributors to uncertainty (Doubilet *et al.* 1985). The uncertainty analysis was conducted on the simulated patterns for every 10 years from 2010 to 2050. We first standardized the total values of the three stock variables (population, developed area percentage, and utility) over all the spatial units into z-scores and used them to assess the uncertainty errors. Second, we set the endpoints recording the values of the assessment variables (population, developed area percentage and utility) at the 10th, 20th, 30th, and 40th time step (every 10 years until the end of the simulations in 2050). Finally, the tuning ranges of the parameters were set between 50% below their original values to 50% above their original values. After all these settings, we used the uniform probability distribution for the Monte Carlo simulation to pick the values for the parameters within their tuning ranges independently for 500 trials. We then recorded the values of the assessment variables of each spatial unit for each trial at the endpoints.

4. Results

The Elastic Net method identified 12 socioeconomic variables that were highly correlated with population change. We combined these variables using their coefficients in the Elastic Net function into a single variable and termed it as 'utility'. In essence, utility was a combination of all the socioeconomic externalities that correlated with population change, excluding the environmental externalities. As a result, positive utility will trigger population growth, whereas negative utility will lead to population loss. We used utility as a state variable in the SDs, and calculated the utility of a spatial unit ij ($Utility_{ij}$) using Equation (5):

$$\begin{aligned} Utility_{ij} = & -0.028 \cdot std\left(Pipeline_{ij}\right) - 0.0634 \cdot std\left(Road_{ij}\right) + 0.0745 \cdot std\left(MedValue_{ij}\right) \\ & + 0.0945 \cdot std\left(MedRent_{ij}\right) - 0.0059 \cdot std\left(OCST35_{ij}\right) - 0.0140 \cdot std\left(NonMtg_{ij}\right) \\ & - 0.0234 \cdot std\left(NonFuel_{ij}\right) - 0.0109 \cdot std\left(NonKitchen_{ij}\right) + 0.0243 \\ & \cdot std\left(NonPlumb_{ij}\right) + 0.0160 \cdot std\left(NonTele_{ij}\right) - 0.3163 \cdot std\left(NonVehicle_{ij}\right) \\ & + 0.1141 \cdot std\left(Under5_{ij}\right) \end{aligned} \tag{5}$$

where the function *std*() denotes the standardization function. After building the Elastic Net for all the three stock variables, the neighborhood impact terms (P_1, P_2, and P_3) were added and calibrated (see below), and the final functions are shown as Equations (6)–(8).

$$\begin{aligned} Population_{ij}(t+1) = {} & Population_{ij}(t) + (1-P_1) \\ & \begin{pmatrix} f_3 \cdot Population_{ij}(t) + e_3 \cdot Developed_{ij}(t) + a_3 \cdot Damages_{ij} \\ +b_3 \cdot Elevation_{ij} + g_3 \cdot Water_{ij}(t) + d_3 \cdot Utility_{ij}(t) \end{pmatrix} \\ & + P_1 \left(\frac{1}{8} \sum_{N_k=1}^{8} \begin{pmatrix} f_3 \cdot Population_k(t) + e_3 \cdot Developed_k(t) + a_3 \cdot Damages_k \\ +b_3 \cdot Elevation_k + g_3 \cdot Water_k(t) + d_3 \cdot Utility_k(t) \end{pmatrix} \right) \end{aligned} \tag{6}$$

$$\begin{aligned} Developed_{ij}(t+1) = {} & Developed_{ij}(t) + (1-P_2)(f_2 \cdot Population_{ij}(t) + e_2 \cdot \\ & Developed_{ij}(t) + g_2 \cdot Water_{ij}(t) + d_2 utility_{ij}(t)) + \\ & P_2 \left(\frac{1}{8} \sum\nolimits_{N_k=1}^{8} (f_2 \cdot Population_k(t) + e_2 \cdot Developed_k(t) + g_2 \cdot Water_k(t) + d_2 \cdot utility_k(t)) \right) \end{aligned} \tag{7}$$

$$\begin{aligned} Utility_{ij}(t+1) = {} & Utility_{ij}(t) \\ & + (1-P_3)\big(f_1 \cdot Population_{ij}(t) + a_1 \cdot Damages_{ij} + b_1 \cdot Elevation_{ij} + c_1 \cdot Subsidence_{ij} + d_1 \cdot Utility_{ij}(t)\big) \\ & + P_3 \left(\frac{1}{8} \sum_{N_k=1}^{8} (f_1 \cdot Population_k(t) + a_1 \cdot Damages_k + b_1 \cdot Elevation_k + c_1 \cdot Subsidence_k + d_1 \cdot Utility_k(t)) \right) \end{aligned} \tag{8}$$

where subscript *ij* refers to a spatial unit in row *i* and column *j, t* denotes a certain time point, *k* denotes the *k*th neighbor of the spatial unit *ij*, and the description of all the variable acronyms is listed in Table 1. Table 2 lists the basic descriptive statistics of the original values of the three stock variables, the four environmental variables, and the 12 socioeconomic variables at 2000 (tabulated from the 3 km by 3 km grids). Figure 7 plots the SDs identified from the Elastic Net equations.

The neighborhood parameters P_1, P_2 and P_3(also referred as the second set of parameters) were calibrated by a GA, whereas the other parameters in Equations (6)–(8) (the first set of parameters) were derived from Elastic Net. The reasons for not using GAs to calibrate the first set of parameters are fourfold. First, the first set of parameters were

Table 2. Descriptive statistics of the variables (in 2000) used in the system dynamics model (tabulated from the 3 km by 3 km grids).

Variable	Minimum	Maximum	Mean	Standard Deviation
Utility	−2.89	1.03	0.00	0.22
Population	0.00	38,334.36	385.75	1862.52
Developed	0.00	100	0.058	0.14
Damage	0.00	999,000,000.00	9,887,979.00	39,880,882.00
Elevation	−2.23	112.45	8.93	19.87
Subsidence	−0.02	0.00	−0.01	0.00
Water	0.00	1.00	0.33	0.39
Pipeline	0.00	13.21	0.52	0.78
Road	0.00	0.02	0.00	0.00
MedValue	0.00	426,314.90	40,571.19	46,649.48
MedRent	0.00	1448.38	192.32	215.24
OCST35	0.00	0.36	0.07	0.08
NonMtg	0.00	0.77	0.23	0.25
NonFuel	0.00	0.08	0.00	0.00
NonKitchen	0.00	0.03	0.00	0.01
NonPlumb	0.00	0.04	0.00	0.01
NonTele	0.00	0.19	0.03	0.04
NonVehicle	0.00	0.49	0.04	0.06
Under5	0.00	0.11	0.03	0.04

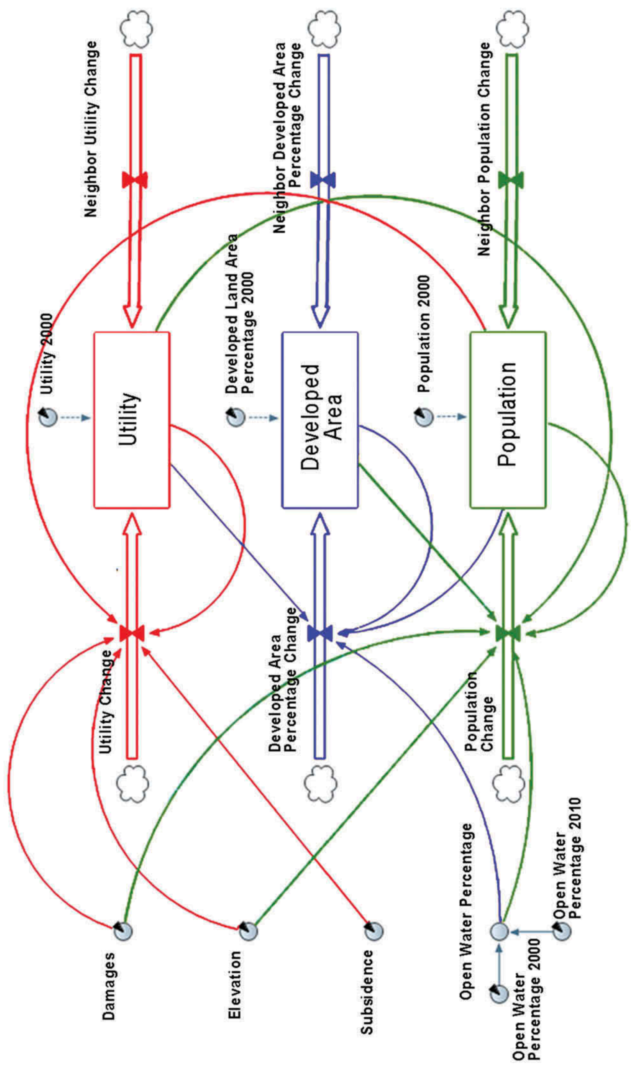

Figure 7. The coupled system dynamics diagram derived in this study.

already empirically estimated by Elastic Net using the real data. Second, the definition of the utility was derived by Elastic Net, and tuning the first set of parameters will cause the linear combination of the socioeconomic variables represented by the utility no longer be the best predictors for population change. Third, although evolutionary algorithms such as GAs can reduce the computational time of running the model considerably in contrast with exhaustive methods, the number of iterations required to find a solution in GAs with many parameters could be enormous (Wright and Alajmi 2005). Thus, testing the convergence with fewer parameters is easier and often preferred. Last, GAs need tuning ranges of the parameters to identify their searching spaces for solutions, and unlike the second set of parameters which has a logical tuning range (from 0 to 1), the first set of parameters does not have tuning ranges. Applying GAs by creating tuning ranges from empirically identified values by another data mining method (such as Elastic Net in this study) would be redundant and trivial.

In applying the GA algorithm, we set the initial population to 100 chromosomes (i.e. P_1, P_2, P_3) and the crossover rate and the mutant rate to 10%. In each iteration, 25% of the population of the current generation were the winners selected from the last generation, and the rest 75% were created by applying crossover and mutant operators on these 25% winners. One hundred iterations were generated, and at each iteration we documented the best chromosome (solution) to study the convergence. The results of the GA are shown in Figure 8.

In Figure 8, the gray line represents the MSE of each iteration, whereas the thick black line represents the best chromosome (feasible solution) of each iteration. To sum up, we found that the best solution is where the contribution percentages of neighborhood impacts were 0% for population change ($P_1 = 0$), 50% for developed land area change ($P_2 = 0.5$), and 0% for utility change ($P_3 = 0$). This means that the best simulation accuracy was achieved when there were no direct neighborhood impacts from the utility change and the population change, but half neighborhood impacts from the developed land area percentage change. Figure 9 compares the real population and developed land area of the study area at 2000, the real population and developed land area at 2010, and the simulated population and developed area at 2010, using both calibration and validation data.

Following the procedure described in Pontius *et al*. (2008a), we created two sets of plots for the two variables of population and developed land using both calibration and validation data (Figure 10). The variable 'utility' was not compared in this analysis because it is a derived variable from a combination of variables and is not directly

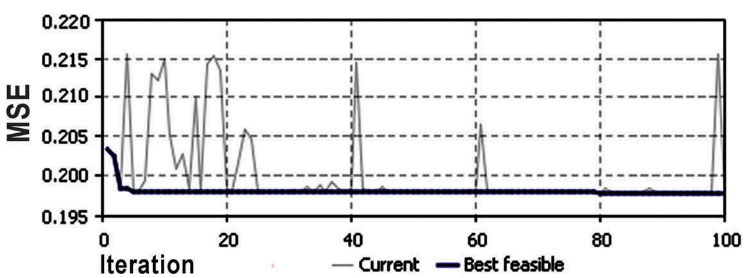

Figure 8. Genetic algorithm results for calibrating the neighborhood impacts.

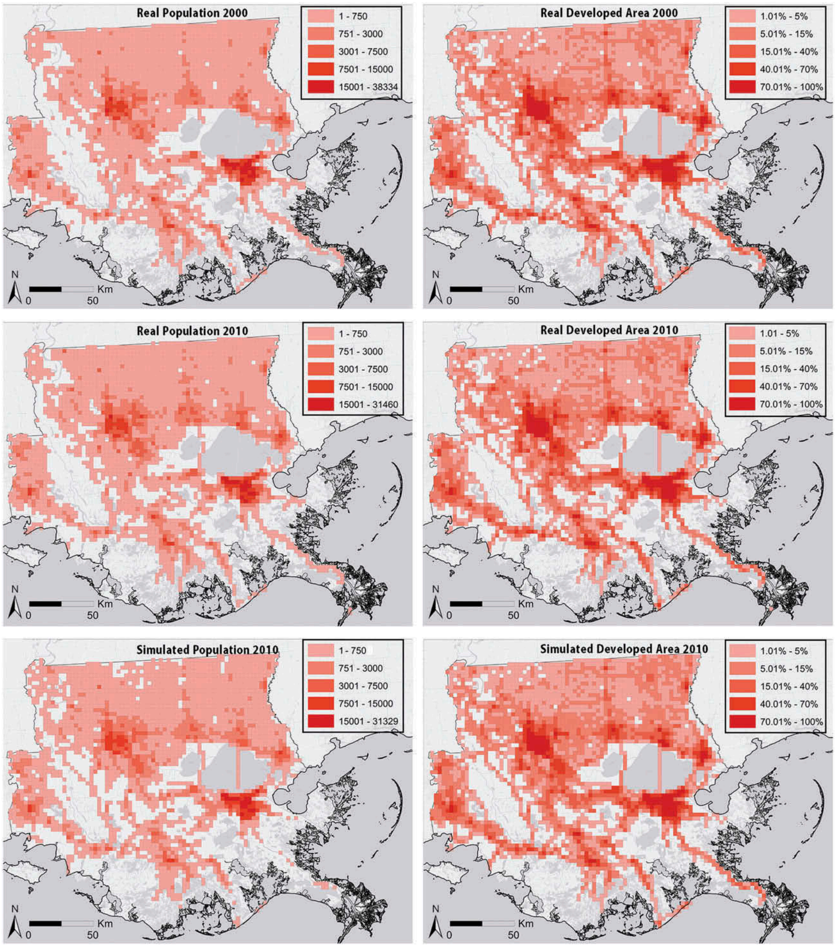

Figure 9. Comparison between the real patterns at 2000 and 2010 and the simulated results at 2010 for population and land area percentage (using both calibration and validation data).

observable. For the variable of population, the first plot (Figure 10(a)) displays the real population at 2010 on the *x*-axis versus the simulated population at 2010 on the *y*-axis, and the second plot (Figure 10(b)) displays the real population at 2010 on the *x*-axis versus the real population at 2000 on the *y*-axis. The first plot shows how closely the model simulates the population at 2010, with the $x = y$ line indicating complete agreement between *x* and *y* values. The second plot shows how much the population changed actually between 2000 and 2010, with the $x = y$ line indicating no changes between the two time points. Therefore, we used the second plot as a persistence model for comparison.

Table 3 lists the *MAD_total*, *MAD_quantity*, *MAD_allocation*, and the mean values of the *x* and *y* axes (Pontius 2008a). According to Table 3 and Figure 10 (subplots (a) and (c)), we found that the model on average slightly overpredicts population and developed land area. For the population variable, the *MAD_total* of the simulated results is smaller than that of the persistence model (Table 3 comparing (a) and (b), 94 < 199); therefore, the simulation model undoubtedly has a higher accuracy than the persistence

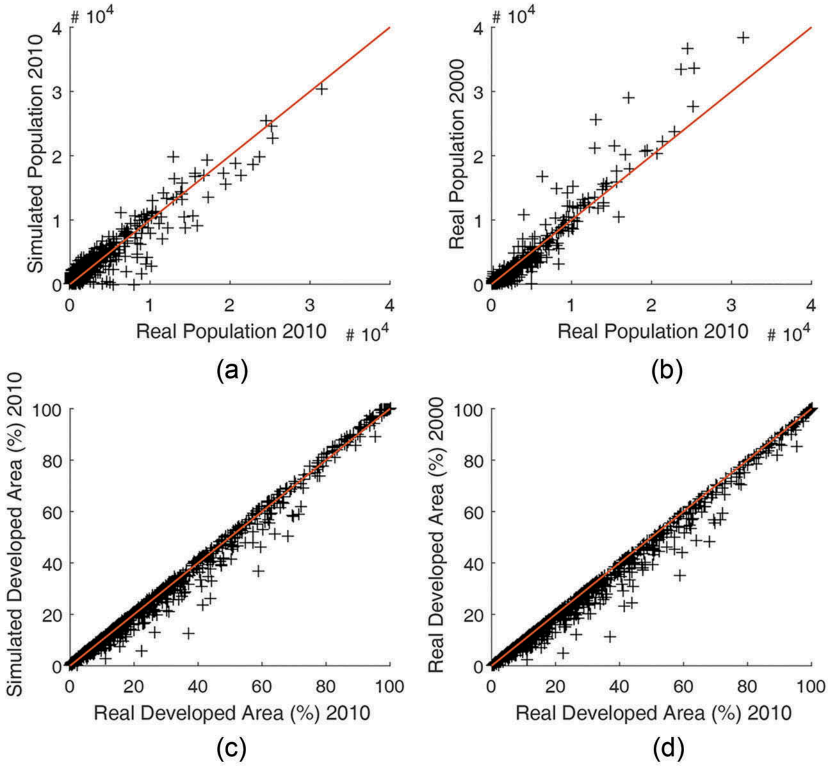

Figure 10. Subplots (a) and (c) compare the simulated and the real results at 2010 for population and developed land area; subplots (b) and (d) compare the real population between 2000 and 2010 (persistence model). All subplots are based on both calibration and validation data. See text for more explanation.

Table 3. Accuracy assessment of the simulation model (5,890 modeling units).

Subplots of Figure 10	Mean *X*	Mean *Y*	*MAD_-total*	*MAD_quantity*	*MAD_allocation*
(a)	398	401	94	3	91
(b)	398	386	199	12	187
(c)	602	604	26	2	24
(d)	602	576	35	26	9

model. When comparing the two components of *MAD_total*, the results show that the *MAD_quantity* is much lower than the *MAD_allocation* (3 vs. 91) for the simulation model, and both are much lower than those of the persistence model. In like manner, the same pattern is true for the variable of developed land, except that the simulated result is less accurate than that of the persistence model in terms of *MAD_allocation* (25 vs. 9). For the developed land area variable, although the simulation model has a higher accuracy than the persistence model in terms of quantity, its accuracy in terms of allocation is lower. Furthermore, a close visual comparison between the simulated and real population maps at 2010 (Figure 9) shows that there are more discrepancies between the real and the simulated patterns for the population variable than for the developed land area variable. We found the most notable discrepancies in the New

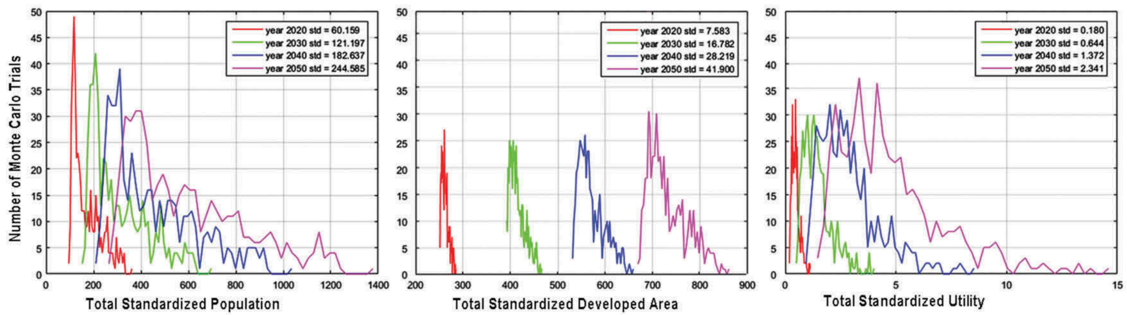

Figure 11. Histograms of standardized total population, total developed area percentage, and total utility.

Table 4. Pearson correlations between the stock variables and selected parameters.

Assessment variable	Parameter[a]	Pearson correlation coefficient			
		Year 2020	Year 2030	Year 2040	Year 2050
Total Population	e_3	0.78	0.77	0.77	0.76
	f_3	0.47	0.50	0.52	0.54
Total Developed	d_2	0.41	0.36	0.32	0.30
Area Percentage	e_2	0.76	0.80	0.83	0.85
Total Utility	f_1	0.68	0.61	0.58	0.57
	e_3	0.52	0.56	0.57	0.57
	f_3	0.32	0.37	0.39	0.40

[a]The parameter symbols are the same as used in Equations (3)–(5).

Orleans region where Hurricane Katrina (in 2005) had caused a dramatic decline in population and urban development during the study period.

We created histograms of all the 500 trials from the uncertainty analysis for the three assessment variables (population, developed land area, utility). All the histograms used 50 equal-size bins, and their center values of the bins are plotted in Figure 11. The probability distributions of the simulated total population and the simulated total developed land area percentage are one-tailed to the right. The reason for this phenomenon is the constraints added in the simulation settings. In contrast, the probability distribution of the simulated total utility is more like two-tailed. The propagated uncertainty errors represented by the probability distributions of the total population and total developed land area percentage (in z-scores) increased almost linearly over the simulation time. For the total utility, the uncertainty increased dramatically from the first endpoint to the second (2020–2030), after that, the increasing speed was also almost linear over the simulation time.

We calculated the Pearson's correlation coefficients between the assessment variables and the parameters to reveal the approximate relative contribution of each parameter to the variance of each assessment variables. The parameters having the greatest effects are considered to be the parameters with the highest correlation coefficients, and they are summarized in Table 4.

5. Discussion

Results from this study indicate that the proposed modeling approach is feasible and can be extended to investigate the long-term population dynamics in a vulnerable coastal region. However, we note several limitations and future improvements of the model. First, in terms of the CNH dynamics, this study mainly focused on changes of the system due to the human component. The dynamics of the three environmental variables, such as elevation, subsidence rate, and land loss rate (open water percentage change), were not updated in the simulation because reliable time-series data and their driving forces on some of the variables (such as subsidence rate) were not available. Nevertheless, the same modeling methodology can be used to uncover the dynamics of these natural variables when their driving factors have been empirically and scientifically identified (e.g. accretion rate, biomass diversity, land building process). In this study, only the developed land use area percentage was modeled. Other types of land cover (e.g. land to water) can also be simulated. Adding more dynamics of the natural part in the model will make the model more comprehensive and help increase our knowledge of the complex CNH dynamics in this coastal region.

Second, we chose Elastic Net to extract the relations among the variables. This 'white-box' data mining approach is preferred whenever possible to reveal the structure and relationships of the natural–human system and extract the rules for building the system dynamics. But as has been noted, we built the Elastic Net using a space-for-time substitution for generating sample data points. Thus, it was assumed that the dependency of a dependent variable on each independent variable is the same among all the spatial units. For this reason, the model could be improved by avoiding the imposition of this underlying space-for-time assumption, but this needs more data for each spatial unit, which means that there should be several temporal data points for each variable in each spatial unit.

Third, results from the GA demonstrate the existence of neighborhood effects to some extent in the final calibrated model. In building such a 'bottom-up' model, parameters setting is an important task. We used the GA as a calibration procedure because it led to quick convergence on tuning the parameters. However, the final calibrated parameters are somewhat confined, leading to no direct neighborhood effects on population and utility (P_1 = 0 and P_3 = 0). Furthermore, as in any evolutional optimization algorithm, there is no guarantee that the final solution is a global optimum. To alleviate this uncertainty, it is possible to test and experiment with different sets of starting values and compare the results. Although there is still no guarantee that a global optimum can be reached, we should revisit this calibration part in future research. In addition, different neighborhood sizes could be tested.

Fourth, we lack information from a third time point due to data availability, hence we could not separate calibration data from validation across time, and as a result we have no measurement of how well the performance of the model predicts through time. Consequently, we applied a substitutional method of validating the data across space, thus the goodness of fit of the developed model only reflects how well the model simulates across space within a short 10-year period. In future studies, with more data collected at different time points, we could validate and increase the prediction accuracy of the model through time.

Fifth, more information on the bottom-up decision-making process, such as those empirically derived from household surveys of people's perception of risk and migration factors, will help us refine the utility variable and increase the accuracy of the model. Sixth, stochastic elements could be included in the model to account for the unknowns or imperfect knowledge that typically exist in complex system modeling. Seventh, all the relationships were calibrated globally without considering spatial nonstationary, future studies could test and quantify local relationships to improve the prediction accuracy. Finally, as in any spatial modeling, scale is a critical issue that can affect the analysis results (Lam and Quattrochi 1992, Lam 2012). A multiscale analysis in the future will reduce the uncertainties of the findings and advance our knowledge of complex SDs in the region.

6. Conclusion

In closing, this study developed a spatial dynamic model to simulate changes in population and developed land area in the LMRB, a coastal region highly vulnerable to sea-level rise and coastal hazards. Using Elastic Net and GAs, 12 socioeconomic and four environmental variables were extracted from 33 variables to build the SD model. Results show that the model on average outstands the persistence model, despite slightly overpredicting both population and developed area. The *MAD_quantity* values confirm that the model has higher accuracy than the persistence model. However, the model did not perform well in terms of *MAD_allocation*, with most outliers occurred in the New Orleans region where Hurricane Katrina had led to a dramatic decline in population and urban development during the study period.

Findings from this study have several significant implications. First, the study provides important empirical baseline information on the study region regarding population dynamics and the impact variables. The 12 variables extracted to represent utility could further be evaluated and compared with the resilience literature in follow-up studies. Second, from the modeling perspective, the uniqueness of modeling coastal vulnerability and population dynamics is that it typically requires a host of natural and human factors, and the data come in diverse forms and have different inherent properties. This study demonstrates how different methods can be combined to develop a CNH dynamic model with a feedback mechanism to examine a compelling societal problem. Third, the model developed in this research focuses on uncovering the social and economic emergent properties in a vulnerable coastal region, which arises from 'bottom-up' interactions in a complex system and not from the existence of a 'top-down' central controller or a planning initiative. The model can be used to monitor the existence of such emergences due to its ability of continuously updating the values of the variables. Fourth, the model offers a foundation to compare the different emergent properties between the northern (more inland) part and the southern (more coastal) part of the study region. By utilizing the simulated results, different impact levels of the extracted variables on the more-coastal and more-inland areas can be analyzed in follow-up studies.

Acknowledgment

This article is supported by two research grants from the U.S. National Science Foundation: one under the Dynamics of Coupled Natural Human Systems (CNH) Program [Award No.121211], and the other under the Coastal Science, Engineering and Education for Sustainability (Coastal SEES) Program [Award No. 1427389]. However, any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the funding agencies.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation [1212112, 1427389].

ORCID

Nina S. N. Lam http://orcid.org/0000-0002-5344-9368

References

Batty, M., Couclelis, H., and Eichen, M., 1997. Urban systems as cellular automata. *Environment and Planning B: Planning and Design*, 24, 159–164. doi:10.1068/b240159

Batty, M., Xie, Y., and Sun, Z.L., 1994. Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems*, 23, 205–233. doi:10.1016/S0198-9715(99)00015-0

Benenson, I. and Torrens, P.M., 2004. *Geosimulation: automata-based modelling of urban phenomena*. West Sussex, UK: Wiley.

Blum, M.D. and Roberts, H.H., 2009. Drowning of the Mississippi Delta due to insufficient sediment supply and global sea-level rise. *Nature Geoscience Letters*, 2 (7), 488–490. doi:10.1038/ngeo553

Boccara, N. and Cheong, K., 1992. Automata network SIR models for the spread of infectious diseases in populations of moving individuals. *Journal of Physics A: Mathematical and General*, 25, 2447–2461. doi:10.1088/0305-4470/25/9/018

Boccara, N. and Cheong, K., 1993. Critical behaviour of a probabilistic automata network SIS model for the spread of an infectious disease in a population of moving individuals. *Journal of Physics A: Mathematical and General*, 26, 3707–3717. doi:10.1088/0305-4470/26/15/020

Bolin, B., *et al.*, 2000. Core questions of science and technology for sustainability.Available from: http://sustainabilityscience.org/content.html?contentid=776/ [Accessed 2 Nov 2014].

Bonabeau, E., 2002. Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99, 7280–7287. doi:10.1073/pnas.082080899

Brown, D.G. and Xie, Y., 2006. Spatial agent-based modelling. *International Journal of Geographical Information Science*, 20 (9), 941–943. doi:10.1080/13658810600830491

Bruch, E. and Mare, R., 2012. Methodological issues in the analysis of residential preferences, residential mobility and neighborhood change. *Sociological Methodology*, 42 (1), 103–154. doi:10.1177/0081175012444105

Cai, H., *et al.*, 2016. Assessing community resilience to coastal hazards in the Lower Mississippi River Basin. *Water*, 8 (2), 46. doi:10.3390/w8020046

Caruso, G., *et al.*, 2007. Spatial configurations in a peri-urban city. A cellular automata based microeconomic model. *Regional Science and Urban Economics*, 37, 542–567. doi:10.1016/j.regsciurbeco.2007.01.005

Cavailhès, J., *et al.*, 2004. Where Alonso meets Sierpinski: an urban economic model of a fractal metropolitan area. *Environment and Planning A*, 36, 1471–1498. doi:10.1068/a36126

Chen, Y., *et al.*, 2012. Defining agents' behaviour based on urban economic theory to simulate complex urban residential dynamics. *International Journal of Geographical Information Science*, 26, 1155–1172. doi:10.1080/13658816.2011.626780

Clarke, K.C. and Gaydos, L., 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, 699–714. doi:10.1080/136588198241617

Clarke, K.C., Hoppen, S., and Gaydos, L.A., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design*, 24, 247–261. doi:10.1068/b240247

Collins, S.L., *et al.*, 2011. An integrated conceptual framework for long-term social-ecological research. *Frontiers in Ecology and the Environment*, 9 (6), 351–357. doi:10.1890/100068

Conte, R., 2002. Agent-based modeling for understanding social intelligence. *Proceedings of the National Academy of Sciences of the United States of America PNAS*, 99 (Supplement 3), 7189–7190. doi:10.1073/pnas.072078999

CPRA, Coastal Protection and Restoration Authority of Louisiana, 2012. *Louisiana's comprehensive master plan for a sustainable coast*. Baton Rouge, LA: CPRA, State of Louisiana.

Cromley, R.G., Hanink, D.M., and Bentley, G.C., 2012. A quantile regression approach to areal interpolation. *Annals of the Association of American Geographers*, 102 (4), 763–777. doi:10.1080/00045608.2011.627054

Dökmeci, V. and Berköz, L., 2000. Residential-location preferences according to demographic characteristics in Istanbul. *Landscape and Urban Planning*, 48 (1–2), 45–55. doi:10.1016/S0169-2046(99)00080-8

Doubilet, P., *et al.*, 1985. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. *Medical Decision Making*, 5, 157–177. doi:10.1177/0272989X8500500205

Fontaine, C.M. and Rounsevell, M.D.A., 2009. An agent-based approach to model future residential pressure on a regional landscape. *Landscape Ecology*, 24 (9), 1237–1254. doi:10.1007/s10980-009-9378-0

Forrester, J.W, 1961. *Industrial dynamics*. Pegasus Communications, MA: the M.I.T. press.

Forrester, J.W., 1969. *Urban dynamics*. Cambridge, MA: the M.I.T. press.

Fujita, M. and Thisse, J.F., 2002. *Economics of agglomeration: cities, industrial location and regional growth*. Cambridge: Cambridge University Press.

Goodchild, M.F. and Lam, N.S.N., 1980. Areal interpolation: a variant of the traditional spatial problem. *Geo-Processing*, 1, 297–312.

Hellweger, F., 2008. Spatially explicit individual-based modeling using a fixed super-individual density. *Computers & Geosciences*, 34, 144–152. doi:10.1016/j.cageo.2006.12.006

Hoerl, A. and Kennard, R., 1988. *Ridge regression. In encyclopedia of statistical sciences*. Vol. 8. New York: Wiley, 129–136.

Irwin, E.G. and Bockstael, N.E., 2002. Interacting agents, spatial externalities and the evolution of residential land use patterns. *Journal of Economic Geography*, 2, 31–54. doi:10.1093/jeg/2.1.31

Jennings, N., 2000. On agent-based software engineering. *Artificial Intelligence*, 117, 277–296. doi:10.1016/S0004-3702(99)00107-1

Kates, R.W., 2011. From the unity of nature to sustainability science: ideas and practice. CID Working Paper (218). Cambridge, MA: Center for International Development, Harvard University. Available from: http://www.hks.harvard.edu/centers/cid/publications/faculty-working papers/cid-working paperno.-218 [Accessed 15 Nov 2011].

Krugman, P., 1991. Increasing returns and economic geography. *Journal of Political Economy*, 99 (3), 483–499. doi:10.1086/261763

Lam, N.S.N., 1980. *Methods and Problems of Areal Interpolation*. PhD Dissertation. University of Western Ontario, London, Ontario, Canada.

Lam, N.S.N., 1983. Spatial interpolation methods: a review. *The American Cartographer*, 10 (2), 129–150. doi:10.1559/152304083783914958

Lam, N.S.N., *et al.*, 2009a. An estimate of population impacted by climate change along the U.S. coast. *Journal of Coastal Research*, Special Issue 56, 1522–1526.

Lam, N.S.N., *et al.*, 2009b. Business return in New Orleans: decision making amid post-Katrina uncertainty. *Public Library of Science (Plos ONE)*, 4 (8), 6765.

Lam, N.S.N., 2012. Geospatial methods for reducing uncertainties in environmental health risk assessment: challenges and opportunities. *Annals of the Association of American Geographers*, 102 (5), 942–950. doi:10.1080/00045608.2012.674900

Lam, N.S.N., *et al.*, 2012. Predictors of business return in New Orleans after Hurricane Katrina. *PLoS One*, 7 (10), e47935: 1–8. doi:10.1371/journal.pone.0047935

Lam, N.S.N., *et al.*, 2016. Measuring community resilience to coastal hazards along the Northern Gulf of Mexico. *Natural Hazards Review*. doi:10.1061/(ASCE)NH.1527-6996.0000193

Lam, N.S.N. and Quattrochi, D.A., 1992. On the issues of scale, resolution, fractal analysis in the Mapping Sciences. *The Professional Geographer*, 44 (1), 88–98. doi:10.1111/j.0033-0124.1992.00088.x

Lam, N.S.N., 2009. Spatial interpolation. *In*: R. Kitchin and N. Thrift, eds. *International encyclopedia of human geography*. Vol. 10. Oxford: Elsevier, 369–376.

LeSage, J.P., *et al.*, 2011a. New Orleans business recovery in the aftermath of Hurricane Katrina. *Journal of Royal Statistics Society A*, 174 (4), 1007–1027. doi:10.1111/j.1467-985X.2011.00712.x

LeSage, J.P., *et al.*, 2011b. Space-time modeling of natural disaster impacts. *Journal of Economic and Social Measurement*, 36, 169–191.

LeSage, J.P., *et al*., 2011c. Do what the neighbors do: reopening businesses after Hurricane Katrina. *Significance*, 8 (4), 160–163. doi:10.1111/j.1740-9713.2011.00520.x

Li, X. and Liu, X.P., 2007. Defining agents' behaviors to simulate complex residential development using multi-criteria evaluation. *Journal of Environmental Management*, 85 (4), 1063–1075. doi:10.1016/j.jenvman.2006.11.006

Li, X. and Liu, X.P., 2008. Embedding sustainable development strategies in agent-based models for use as a planning tool. *International Journal of Geographical Information Science*, 22 (1), 21–45. doi:10.1080/13658810701228686

Ligtenberg, A., Bregt, A.K., and Lammeren, R.V., 2001. Multi-actor-based land use modelling: spatial planning using agents. *Landscape and Urban Planning*, 56, 21–33. doi:10.1016/S0169-2046(01)00162-1

Lindberg, E., *et al*., 1992. Residential location preferences across the life-span. *Journal of Environmental Psychology*, 12, 187–198. doi:10.1016/S0272-4944(05)80070-2

Liu, J.G., *et al*., 2007a. Complexity of coupled human and natural systems. *Science*, 317 (5844), 1513–1516. doi:10.1126/science.1144004

Liu, J.G., *et al*., 2007b. Coupled human and natural systems. *AMBIO: A Journal of the Human Environment*, 36 (8), 639–649. doi:10.1579/0044-7447(2007)36[639:CHANS]2.0.CO;2

Malanson, G.P., Zeng, Y., and Walsh, S.J., 2006a. Complexity at advancing ecotones and frontiers. *Environment and Planning A*, 38, 619–632. doi:10.1068/a37340

Malanson, G.P., Zeng, Y., and Walsh, S.J., 2006b. Landscape frontiers, geography frontiers: lessons to be learned. *The Professional Geographer*, 58, 383–396. doi:10.1111/j.1467-9272.2006.00576.x

McFadden, D. (1978). Choice of residential location. *Cowles Foundation Discussion Paper* 477.

Munroe, D.K., 2007. Exploring the determinants of spatial pattern in residential land markets: amenities and disamenities in Charlotte, NC, USA. *Environment and Planning B: Planning and Design*, 34, 336–354. doi:10.1068/b32065

Niedomysl, T., 2008. Residential preferences for interregional migration in Sweden: demographic, socioeconomic, and geographical determinants. *Environment and Planning A*, 40, 1109–1131. doi:10.1068/a39177

Nijkamp, P., Van Wissen, L., and Rima, A., 1993. A household life cycle model for residential relocation behaviour. *Socioeconomic Planning Sciences*, 27, 35–53. doi:10.1016/0038-0121(93)90027-G

Parker, D.C. and Filatova, T., 2008. A conceptual design for a bilateral agent-based land market with heterogeneous economic agents. *Computers, Environment and Urban Systems*, 32, 454–463. doi:10.1016/j.compenvurbsys.2008.09.012

Plyer, A., 3 May 2013. Facts for Features: hurricane Katrina Recovery. GNOCDC. Greater New Orleans Community Data Center, 30 Jan. 2013. Web. Available from: http://www.gnocdc.org/Factsforfeatures/HurricaneKatrinaRecovery/.

Pontius Jr, R.G., *et al*., 2008b. Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science*, 42 (1), 11–37. doi:10.1007/s00168-007-0138-2

Pontius Jr, R.G. and Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32 (15), 4407–4429. doi:10.1080/01431161.2011.552923

Pontius Jr, R.G., Thontteh, O., and Chen, H., 2008a. Components of information for multiple resolution comparison between maps that share a real variable. *Environmental and Ecological Statistics*, 15 (2), 111–142. doi:10.1007/s10651-007-0043-y

Qiang, Y. and Lam, N.S.N., 2015. Modeling land use and land cover changes in a vulnerable coastal region using artificial neural networks and cellular automata. *Environmental Monitoring and Assessment*, 187. doi:10.1007/s10661-015-4298-8

Qiang, Y. and Lam, N.S.N., 2016. The impact of Hurricane Katrina on urban growth in Louisiana: an analysis using data mining and simulation approaches. *International Journal of Geographical Information Science*, 30, 1832–1852. doi:10.1080/13658816.2016.1144886

Scheffer, M., *et al*., 1995. Super-individuals a simple solution for modelling large populations on an individual basis. *Ecological Modelling*, 80, 161–170. doi:10.1016/0304-3800(94)00055-M

Schultz, J. and Elliott, J.R., 2012. Natural disasters and local demographic change in the United States. *Population Environment*, 34 (3), 293–312. doi:10.1007/s11111-012-0171-7

Shi, W. and Pang, M.Y., 2000. Development of Voronoi-based cellular automata -an integrated dynamic model for GeographicalInformation Systems. *International Journal Geographic Information Sciences*, 14, 455–474. doi:10.1080/13658810050057597

Shu, Y., Lam, N.S., and Reams, M., 2010. A new method for estimating carbon dioxide emissions from transportation at fine spatial scales. *Environmental Research Letters*, 5 (4), 044008. doi:10.1088/1748-9326/5/4/044008

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *Journal R Statistic Social B*, 58, 267–288.

Torrens, P.M., 2001. *Can geocomputation save urban simulation? Throw some agents into the mixture, simmer, and wait*. University College of London, Centre for Advanced Spatial Analysis, London, UK: Working Paper series 32.

Twilley, R.R., *et al*., 2016. Co-evolution of wetland landscapes, flooding and human settlement in the Mississippi River Delta Plain. *Sustainability Science*, 11, 711–731. doi:10.1007/s11625-016-0374-4

Vaz, E., *et al*., 2012. A multi-scenario forecast of urban change: a study on urban growth in the Algarve. *Landscape and Urban Planning*, 104, 201–211. doi:10.1016/j.landurbplan.2011.10.007

Whittingham, M.J., *et al*., 2006. Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, 75 (5), 1182–1189. doi:10.1111/j.1365-2656.2006.01141.x

Wright, J.A., and Alajmi, A.F., 2005. The robustness of genetic algorithms in solving unconstrained building optimization problems. In: 9th International IBPSA Building Simulation Conference, 15-18 August 2005. Montréal, Canada.

Wu, F. and Webster, C.J., 2000. Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. *International Journal of Geographical Information Science*, 14, 625–648. doi:10.1080/136588100424945

Xie, Y.C., 1997. A generalized model for cellular urban dynamics. *Geographical Analysis*, 28, 350–373. doi:10.1111/gean.1996.28.issue-4

Yin, L. and Muller, B., 2007. Residential location and the biophysical environment: exurban development agents in a heterogeneous landscape. *Environment and Planning B: Planning and Design*, 34, 279–295. doi:10.1068/b31182

Zhou, C.H., Sun, Z.H., and Xie, Y.C., 1999. *The research of geographical cellular automata*. Beijing, China: Science Press, 1–163.

Zou, H. and Hastie, T., 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67 (2), 301–320. doi:10.1111/rssb.2005.67.issue-2

Zou, L., *et al*., 2016. Evaluating land subsidence rates and their implications for land loss in the Lower Mississippi River Basin. *Water*, 8 (1). doi:10.3390/w8010010.15p