Sustainable Cities and Society 90 (2023) 104390

Contents lists available at ScienceDirect

Sustainable Cities and Society

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

Spatio-temporal modeling of parcel-level land-use changes using machine learning methods

Emre Tepe [a,*], Abolfazl Safikhani [b]

[a] Department of Urban and Regional Planning, University of Florida, 444 Architectural Building, P.O. Box 115706, Gainesville, FL, 32611, USA
[b] Department of Statistics, George Mason University, Fairfax, VA 22030, USA

ARTICLE INFO

Keywords:
Land-use change
Spatio-temporal modeling
Machine learning
Big data
GPU parallel processing

ABSTRACT

Spatio-temporal modeling of parcel-level land development dynamics is essential to maintain sustainable urban growth. Modeling parcel-level urban development controlling contemporaneous and historical conditions involve computational challenges since data sizes grow quickly beyond the capabilities of conventional statistical-based spatial models. Machine Learning (ML) methods provide computationally feasible methods for large-scale data sets. This paper introduces new ML applications using advanced algorithms and GPU parallel processing to model large-scale urban land developments. Special attention is given to accelerating the construction of spatial weight matrices and training ML models. Specifically, artificial neural networks and random forests are applied to the state of Florida's land-use data, which contains nearly 9 million parcels, to predict parcels with changes in their land use based on historical and neighborhood data. The adaptive Hashing algorithm coupled with GPU parallel processing accelerates the average processing time for identifying the fixed number of nearest neighbors used for accounting spatial autocorrelation, by almost 16,000 times. Also, ML model training times are shortened by 49–547 times using GPU. Further, our best ML model achieves approximately 92% accuracy while outperforming some competing methods, including logistic regression. Such a high prediction accuracy helps policymakers adjust budget allocations to meet local land-use change projections.

1. Introduction

Rural and urban lands are constantly altered due to human activities and natural events. Some changes are irreversible. For example, converting agricultural land into residential land causes the removal of fertile soil on the ground, and the land cannot be easily reused for agricultural purposes. Similarly, new land developments towards environmentally sensitive areas such as forests, wetlands, and coastal areas also adversely impact the ecological systems. Government agencies responsible for growth management require long-term planning and sufficient resources to reduce the potential impacts of urban expansions and maintain sustainable environments. Predicting future land developments is challenging due to various uncertainties and complex dynamics. Land-use change models are used to mimic human activities using proxy information. However, current modeling approaches have three main limitations to achieving high accuracy: (1) accounting for dynamic relationships (spatial and temporal), (2) controlling non-linear relationships among features, and (3) constructing regional high-resolution models.

In recent years, factors influencing land-use changes have been investigated by various researchers from Planning, Geography, Economics, and Environmental Science (Bhat et al., 2015; Carrion-Flores & Irwin, 2004; Chomitz & Gray, 1996; Deng & Srinivasan, 2016; Nahuelhual et al., 2012; Verburg et al., 2004). Some studies indicate the significance of incorporating contemporaneous and historical land developments to achieve accurate models (Bhat et al., 2015; Huang et al., 2009; Irwin et al., 2003; Tepe & Guldmann, 2017, 2020). Also, two studies show the importance of modeling an entire county at the parcel level to achieve even higher model accuracy (Tepe & Guldmann, 2017, 2020). Land developments are affected by investments nearby and dynamics in the region. However, a statewide model accounting for regional dynamics requires computationally feasible methodologies and high-resolution spatial data sets. The increasing availability of spatially explicit data offers new opportunities for large-scale modeling (Cressie & Wikle, 2011; Hefley et al., 2017; Wikle & Hooten, 2010). Computational advantages of supervised Ensemble Learning methods, such as Random Forest (RF) and Artificial Neural Networks (ANN), can be considered for analyzing such big spatial data sets (Bahadori et al., 2014;

* Correspondence to: Department of Urban and Regional Planning, University of Florida, 1480 Inner Road, Gainesville, FL 32611, USA.
E-mail addresses: emretepe@ufl.edu (E. Tepe), asafikha@gmu.edu (A. Safikhani).

https://doi.org/10.1016/j.scs.2023.104390
Received 7 June 2022; Received in revised form 9 December 2022; Accepted 1 January 2023
Available online 6 January 2023
2210-6707/

Ceci et al., 2017; Delasalles et al., 2019; Ren & Wu, 2014; Supinie et al., 2009). Also, these methods provide highly accurate predictions due to incorporating non-linear relations. However, the existing approaches to incorporate spatial dependencies are computationally infeasible when the data sets exceed 100 K spatial observations due to the required excessive computation power and computer memory.

The main goal of this study is to introduce a computationally feasible ML-based spatio-temporal land-use change (LUC) modeling framework using advanced programming algorithms and GPU parallel processing. The introduced modeling framework is suitable for developing statewide LUC models. Statewide models can simulate the future impacts of many planning policies, such as the decentralization of city centers in the post-pandemic period or future land developments in coastal areas at the risk of rising sea levels. In addition to its application in planning, the developed computationally feasible modeling approaches can be easily applied to a broad spectrum of research areas, including hedonic pricing, traffic forecasting, urban energy consumption, ecological systems, disease spread, trade relations, and business networks. For instance, historical traffic volume data have a similar data generation process.

In this research, the introduced modeling framework is applied to the Florida Parcel Database, published by UF GeoPlan Center (Par, 2019), consisting of almost 9 million unique parcels. The database contains information about parcel geometry, actual construction year of buildings, land use, sale record, and the number of buildings. Historical land development conditions can be generated using the exact construction year information, and essential parcel and neighborhood characteristics can be retrieved from the data set. Historical records can be traced back to 1650; the most recent year is 2019. All model training and parameter estimations are completed using UF HiPerGator 3.0 supercomputer.

In this study, significant accelerations are achieved in constructing spatial weight matrix nearly 16,000 times and training ML-based models by 49–547 times utilizing the developed adaptive Hashing algorithm and GPU parallel processing. Such computational advancements allow growth management agencies to perform multiple simulations where such tasks would take years to complete without solving computational challenges. Our research also highlights the importance of spatio-temporal components in land-use change modeling to achieve accurate predictions.

The remainder of the paper is organized as follows. Section 2 critically reviews the relevant literature, while Section 3 summarizes machine learning methods applied to our parcel data set and computational improvements to deal with extensive spatial data. A brief introduction to the Florida statewide parcel data is provided in Section 4, while model results and performance measures are presented in Section 5. Section 6 discusses the research findings. Finally, some concluding remarks and future research directions are discussed in Section 7.

2. Literature review

Different methodological approaches for modeling land-use changes at various geographical scales have been proposed over the last two decades by a diverse group of researchers with backgrounds in economics, urban planning, natural sciences, and geography. The main focus of these studies is on the driving factors of land-use changes in urban environments (Bhat et al., 2015; Carrion-Flores & Irwin, 2004; Chomitz & Gray, 1996; Deng & Srinivasan, 2016; Huang et al., 2009; Nahuelhual et al., 2012; Tepe & Guldmann, 2017, 2020; Verburg et al., 2004). This literature review aims to critically review LUC models incorporating spatial and temporal dynamics, Cellular Automata (CA) applications, and new approaches using ML methods.

Decision-making processes involved in land-use changes depend on various demographic, socio-economic, and spatial characteristics. Modeling these complexities provides researchers, practitioners, and policy-makers insights into land development dynamics. In land development dynamics, we expect that an investment decision on a given property depends on investment decisions on neighboring properties. Neglecting such dynamics in models provide poor results. Previous studies ignoring the spatial dependencies and implementing random cross-validation in ML applications to spatial data sets have provided underestimated residuals due to the autocorrelation in the data (Bahn & McGill, 2013; Gasch et al., 2015; Gudmundsson & Seneviratne, 2015; Juel et al., 2015; Meyer et al., 2018, 2019; Micheletti et al., 2014; Roberts et al., 2017). Nahuelhual et al. (2012) highlight that spatio-temporal dynamics play an essential role in timber plantation expansion in south-central Chile over two separate periods (1975–1990, 1990–2007) using an autologistic regression model while Huang et al. (2009) account for spatial and temporal dynamics when modeling land-use changes in New Castle County, Delaware, over three separate periods (1984–1992, 1992–1997, 1997–2002). Similarly, Ferdous and Bhat (2013) explicitly control spatial interactions and temporal lags by implementing a spatial panel ordered-response probit model. Finally, Tepe and Guldmann (2017, 2020) formalize spatio-temporal binary and multinomial spatio-temporal autologistic regression models to model land-use changes at the parcel level in Delaware county, Ohio.

Accounting for spatial dependencies in LUC models is necessary to control spatial heterogeneity in land development dynamics, while temporal dynamics are required to achieve accurate models. Incorporating spatial components explicitly in LUC modeling results in computational challenges. Recent statistical models included a form of the explicit spatial part that has less than 3000 sample size due to computational difficulties (Bhat et al., 2015; Deng & Srinivasan, 2016; Ferdous & Bhat, 2013; Huang et al., 2009; Nahuelhual et al., 2012), while Tepe and Guldmann (2017, 2020) introduce computational feasibility of modeling approach for LUC modeling. However, such a modeling approach still requires improvements when the number of observations increases substantially.

CA approach has been widely used in land use and land cover change models (Guan et al., 2011; Herold et al., 2003; Jahanishakib et al., 2018; Lau & Kam, 2005; Lu et al., 2020; Maria de Almeida et al., 2003; Pan et al., 2010; Pinheiro et al., 2020; Stevens & Dragićević, 2007; Ulloa-Espíndola & Martín-Fernández, 2021). In CA models, the state of a cell is a function of the cell's immediate proximity based on predefined neighborhood rules. Therefore, this approach successfully mimics local spatial relationships. Herold et al. (2003) apply the CA approach to model spatial–temporal dynamics in Santa Barbara, California, highlighting the importance of spatial metrics to achieve reliable urban growth predictions. Similarly, Pinheiro et al. (2020) and Ulloa-Espíndola and Martín-Fernández (2021) develop CA-based LUC models in their studies. Some CA applications integrate the Markov Chain approach to introduce temporal relationships in LUC models (Guan et al., 2011; Jahanishakib et al., 2018; Yang et al., 2012). Many CA-based LUC models use grid cells to reduce computational complexity. In this approach, each grid cell must have a homogeneous state. However, spatial structures are discontinuous and irregular in shape. Vector and irregular grid-based CA models are introduced to overcome the spatial scale dependency problem (Lu et al., 2020; Moreno et al., 2008; Stevens & Dragićević, 2007). Calibrations of CA models are critical to achieving robust results. Pan et al. (2010) discuss the optimal cell size to achieve accurate results. Maria de Almeida et al. (2003) stress the predefined transition rules to successfully calibrate CA-based land use models. In the CA approach, transition rules are given, whereas the primary goal in LUC modeling is to estimate these transition rules. Therefore, statistical models are mainly preferred as transition rules. Multiple hybrid models combine statistical and CA models to achieve robust results (Berry et al., 1996; Flamm & Turner, 1994; Hazen & Berry, 1997; Veldkamp & Fresco, 1996).

ML methods such as Random Forest (RF), Deep Learning (DL), and Support Vector Machine (SVM) are integrated into CA-based LUC models, where these ML methods are used for transition rules. Gounaridis et al. (2019) introduce a hybrid CA-based LUC model using the RF approach for classifying detailed land-use categories accounting for environmental, physical, accessibility, and socio-economic indicators in Attica, Greece. Karimi et al. (2019) used the SVM approach to model Guilford County's urban expansions between 2001 and 2006. Their LUC model incorporates site-specific, proximity, and neighborhood characteristics to estimate a given cell's status (vacant or built-up). Their findings highlight the importance of spatial clustering patterns of the same land use categories. Xing et al. (2020) integrate the DL method into their CA-based spatio-temporal LUC model, where spatio-temporal dynamics such as land cover proportion, site-specific, and proximity measures are collected using Landsat images and road networks. Finally, Li et al. (2022) introduce a CA optimization method using the firefly algorithm to predict future urban growth accurately.

There are also alternative approaches to CA-based LUC models. Zhou et al. (2020) use Markov Chains in their CA model, where land-use changes in Shanghai are investigated by incorporating site-specific, proximity, and socio-economic characteristics and planning guidelines. Fu et al. (2022) also introduce a dynamic modeling approach using the CA-based Markov model to analyze spatio-temporal patterns of land-use changes in Mianzhu City, China, while Yu et al. (2021) discuss the lack of sufficient historical information to calibrate CA-based LUC models. This study implements a Markov Chain Monte Carlo approximation using Bayesian computation to calibrate CA models. In addition, Okwuashi and Ndehedehe (2021) integrate Markov chains with the SVM in their CA-based modeling of urban changes. Alternatively, Lv et al. (2021) introduce a gravity-based approach in their RF-integrated CA-based model to account for spatial interactions between cities. In this study, urban and non-urban land use classifications are examined using a set of predictor variables consisting of economic, social, educational characteristics and infrastructure, and environmental conditions. Such gravitational model approaches improve traditional CA-based models by accounting for travel costs. Finally, Shafizadeh-Moghadam et al. (2021) implement a Forward Feature Selection algorithm for RF transition models in their CA-based LUC models. Their study investigates a given grid's status (urban growth and non-urban persistence) based on a set of characteristics: slope, altitude, and distances from roads, crops, greenery, urban, and barren. Accounting for proximity factors in the absence of socio-economic factors is highlighted.

A few recent studies implemented ML applications in their LUC modeling. At the same time, these methods are used more often in image processing and pattern recognition in the context of land cover changes (Wang et al., 2022). Ron-Ferguson et al. (2021) model non-linear relationships in land development dynamics using the RF approach accounting for a wide range of explanatory factors, including socio-economic, built environment characteristics, and landscape metrics, while Talukdar et al. (2021) implement Bagging and RF methods to model spatio-temporal dynamics of land cover changes among the water body, agricultural land, vegetation, sand bar, bare land, and built-up area categories in Bangladesh, accounting for a set of landscape metrics. The bagging model provides higher accurate estimations because of higher levels of tree depths than the RF model. The model successfully captures land cover fragmentation in the study area. Soares-Filho et al. (2013) introduce a heuristic modeling approach based on a Genetic Algorithm (GA) to improve the accuracy of LUC models. Further, Zhai et al. (2020) implement a Convolution Neural Network (CNN) approach in their Vector-based CA models at the parcel level, where site-specific and proximity characteristics are incorporated in their modeling. This novel approach mimics local neighborhood dynamics using the convolution kernel and local connectors. Finally, Kim et al. (2022) introduce ML applications to model spatio-temporal dynamics of land-use changes at the block group level. Their study reveals the significance of temporal lags of contemporaneous neighborhoods for predicting future urban growth.

ML methods provide robust model results because of accounting for non-linear relations (Bahadori et al., 2014; Delasalles et al., 2019). Overall model accuracy of LUC models using ML approaches ranges between 78% and 94% (Gounaridis et al., 2019; Karimi et al., 2019; Lv et al., 2021; Ron-Ferguson et al., 2021; Xing et al., 2020). Basse et al. (2014) discuss the benefit of integrating the ANN approach with CA-based modeling to increase model accuracy. ML methods also allow working with large data sets when spatial components are incorporated. Among the recent works mentioned, Zhai et al. (2020) worked with data consisting of approximately 125,000 spatial objects, while Tepe and Guldmann (2017, 2020) introduced a computationally feasible approach for spatio-temporal modeling of land-use changes using statistical methods where the data set consists of almost 73,000 parcels, where all computations are conducted on a supercomputer. Finally, CA models can successfully mimic local dynamic spatial structures. However, this approach requires a robust transition rule to define the next state of a given cell. ML and DL models incorporating spatial dependencies can also account for more complex dynamic spatial relationships while estimating accurate results due to their non-linear structures and unbiased estimators.

3. Methodology

This section briefly introduces our approach to integrating spatio-temporal dependencies into ML- and DL-based methods and our computational improvements to deal with large spatial data. In this study, we train models to classify two outcomes (change in land use status or continuation of the current status). A binary response model like the Logistic Regression (LR) framework can be considered as a potential approach. Due to LR's parametric structure, we can introduce the spatial and temporal parameters in the systematic component. Eq. (1) represents an LR framework accounting for spatio-temporal dynamics in land-use changes:

$$\ln\left[\frac{p(y_{i,t}|x_{i,t},y_{j,t})}{1-p(y_{i,t}|x_{i,t},y_{j,t})}\right]=\beta_0+\sum_{k=1}^{K}\beta_k x_{i,t,k}+\sum_{j\in\mathcal{N}_i}\rho y_{j,t}+\sum_{l=1}^{L}\sum_{j\in\mathcal{N}_i}\theta_l y_{j,t-l}, \quad (1)$$

where $y_{i,t}$ is a status change (1 for change and 0 for no change) in parcel i at time t; Logit $(\ln[p(.)/(1-p(.))])$ is the link function; $p(.)$ is conditional probability of change in status of parcel i at time t; $x_{i,t,k}$ is the kth exogenous variable for parcel i at time t; β_k is the kth regression coefficient; $\mathcal{N}_i$ is set of the indices of the neighbors of parcel i; j is the index of ith parcel's neighbors; ρ is spatial autoregressive coefficient; l is the index of spatio-temporal lags $(1,\dots,L)$; and θ_l is the lth spatio-temporal autoregressive coefficient.

3.1. Machine learning models

In this section, RF and ANN methods are briefly introduced. These methods are widely used in classification tasks and provide accurate predictions due to their non-linear structures. In this research, based on the given information, binary responses (change/no change in land use status) are modeled using RF and ANN modeling approaches.

RF is a widely used ensemble learning method for classification where the algorithm consists of many decision trees and bootstrap samples in each decision tree (Breiman, 2001). The main procedure of an RF model to classify land-use change status is illustrated in Fig. 1. The first step in the RF classification procedure is to create bootstrap samples from the training data set to grow trees from these bootstrap samples. A randomly selected set of features is used in each tree node to build the next node. Each RF tree identifies an output class (change or no change in land use status) based on the tree structure. These identified outputs from trees are also considered as votes. In the final phase, the RF output is computed based on the majority voting, i.e., the final output will be the dominantly identified output class.

ANN is also another recently implemented method for LUC modeling. Fig. 2 presents a simple form of ANN (i.e., feed-forward NN)

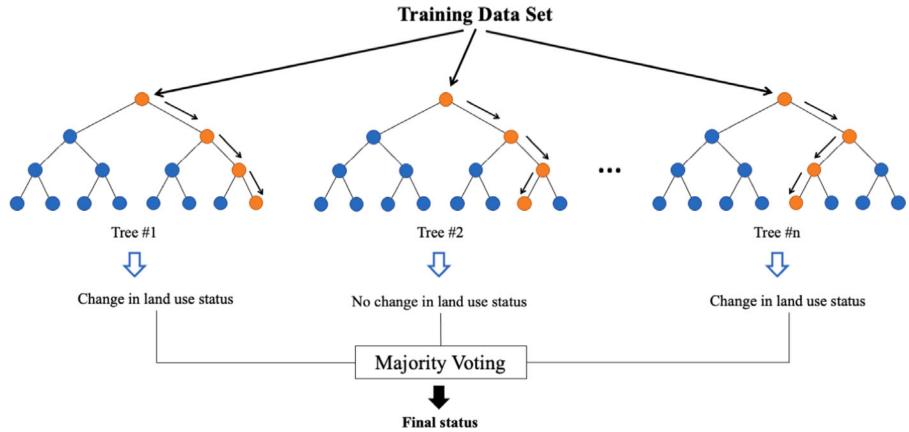


Fig. 1. Random forests.

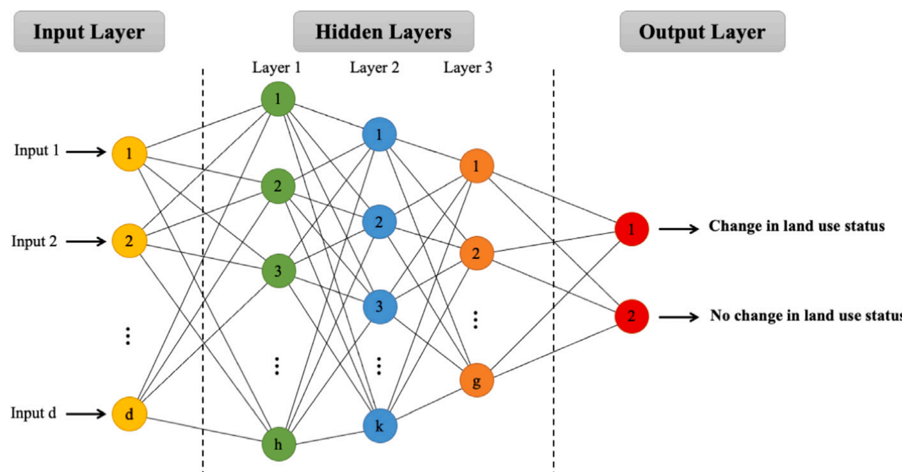


Fig. 2. Artificial neural networks.

applied to LUC modeling, with three hidden layers and two potential land-use categories. ANN is a mathematical formulation of the human-brain process (Hastie et al., 2009). ANN comprises multiple layers, including the input layer, one or multiple hidden layers, and the output layer. Each layer has numerous nodes, which are interconnected with nodes in other layers. A given node may be activated and obtains a weight using a pre-defined activation function dedicated to each layer. *Sigmoid* (see Eq. (2)) and *Softmax* (see Eq. (3)) activation functions are used in this study. The forward propagation method uses the output from a layer as input in the next layer. After obtaining outputs, the ANN model is further trained by backpropagation to reduce bias and improve prediction accuracy (Fan et al., 2019).

$$\Phi(z) = \frac{1}{1+e^{-z}}, \tag{2}$$

$$\Sigma(z)_i = \frac{e_i^z}{\sum_{j=1}^{K} e^{z_j}}, \tag{3}$$

where z is the input value/vector, e is the standard exponential function, and K is the number of classes.

3.2. Computationally feasible approach

Spatial components are incorporated into RF and ANN models using the Spatial Weight Matrix ($\boldsymbol{W}$) approach, widely used in spatial statistical models (Anselin, 1988). This method allows us to conceptualize spatial relationships using mathematical notations. In the matrix $\boldsymbol{W}$, weights are assigned for all pairs of spatial objects in the data set based on a pre-defined neighboring rule. These rules can be summarized under three main groups: fixed distance-, fixed nearest neighbors-, and contiguity-based rules. Combinations of these rules can also be used.

Computations of pairwise distances using the Euclidean distance formula are necessary for constructing a $\boldsymbol{W}$ matrix. However, serial processing leads to the quadratic computational complexity $O(n^2)$, where n is the number of spatial units. To mitigate this issue, We develop a computationally feasible algorithm to reduce the complexity of computations required for constructing a $\boldsymbol{W}$ matrix, which is the fundamental component to account for spatial dependencies explicitly in ANN and RF models. Our algorithm utilizes Hashing Algorithm, an effective technique to search for similarities in large-scale data sets, and GPU parallel computing techniques (Keckler et al., 2011) to accelerate the construction of these matrices.

Identifying the Kth nearest neighbors is computationally more challenging than the fixed distance-based rule. Therefore, this section introduces our computational solution for this rule. Hashing Algorithm allows one to search for potential neighbors using a subset of the complete data set for a given spatial point instead of searching the entire data. Since point densities vary over space, our algorithm starts with an optimal hash size and expands the search region based on the point density. Fig. 3 illustrates the adaptive search procedure using the Hashing Algorithm. For a given point i, we start searching for neighbors in the hash bin containing the point itself and this bin's neighboring hash bins. The initial search region is expanded by adding adjacent hash bins of the hash bins surrounding the central bin until a sufficient

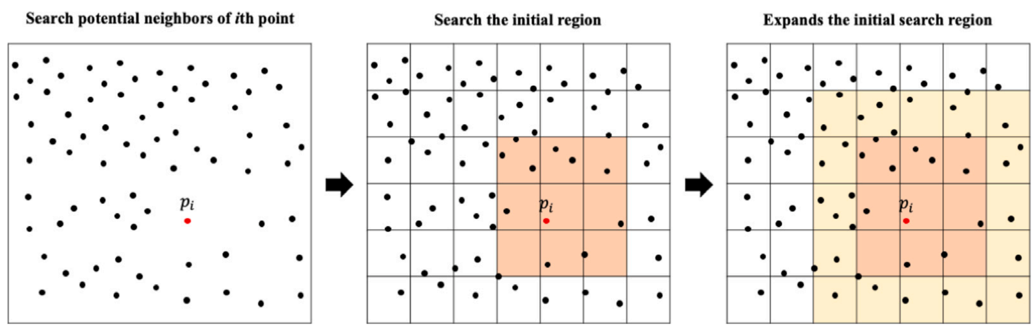


Fig. 3. Adaptive searching using the Hashing Algorithm.

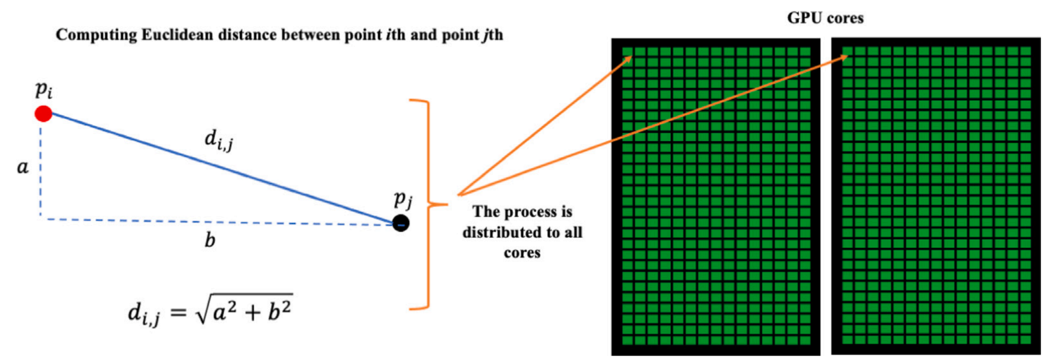


Fig. 4. GPU parallel processing.

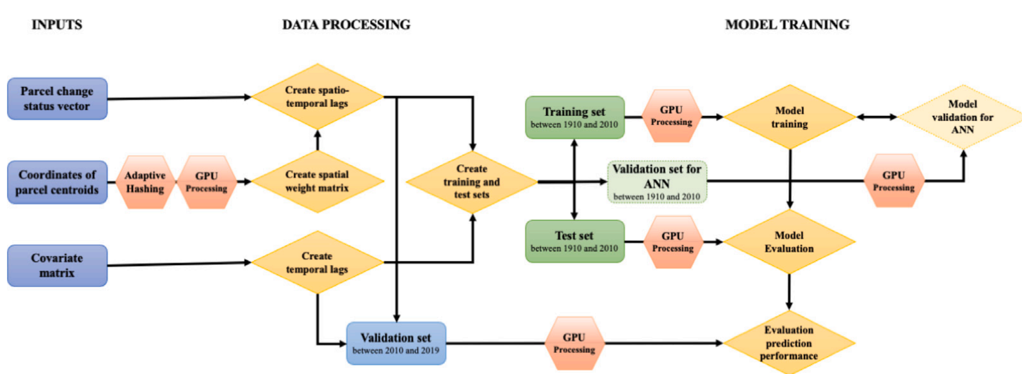


Fig. 5. Flowchart of model training and evaluation using Adaptive Hashing and GPU processing.

number of potential neighbors are identified. Eq. (4) represents the method used to compute the initial hash size:

$$r = \sqrt{\frac{((\max X - \min X) \times (\max Y - \min Y)) \times k}{\pi \times n}}, \tag{4}$$

where r is the initial radius for hash size; X is a vector of X (horizontal) coordinates of vertices of spatial objects, Y is a vector of Y (vertical) coordinates of vertices of spatial objects; k is the degree of nearest neighbors; and n is the number of spatial objects.

After identifying potential neighbors using the adaptive Hashing Algorithm, Euclidean distances between the ith point and other points in the search region are computed using GPU parallel processing (see Fig. 4). After the distances are calculated, Kth nearest neighbors are identified after the subset is sorted in ascending order based on distance values. The pseudo-code of the proposed identification of K-nearest neighbors is summarized in Algorithm 1 located in the supplementary document. Fig. 5 summarizes steps in model training and evaluation and shows tasks utilize Adaptive Hashing and GPU processing.

4. Data set

In this research, Florida is selected as the study area. Florida is one of the fastest-growing states in the United States. With a population of 765,000 in 1910 (U.S. Census Bureau, 1910), Florida's total population has increased by more than 28 times over 110 years, to 21.5 million in 2019 (U.S. Census Bureau, 2019). Historical conditions of land development were generated using the actual construction year information. Historical records can be traced back to 1650; the most recent year is 2019. We focused on the most recent 110 years due to the significant land developments in Florida during this period.

In the 2019 Florida Parcel Database, published by UF GeoPlan Center (Par, 2019), there are 8,995,663 individual parcel records where 6,577,320 among these parcels changed their land-use status as a result of human actions between 1910 and 2019. In this data, all urban and rural land-use activities are included. Fig. 6 depicts annual changes in parcel status in Florida between 1910 and 2019. The most considerable change in parcels' status was observed in 2006. The average number of changes in parcel status was 60,892 over the period, which is less than 1% of the total parcels. In this study, annual parcel data is aggregated

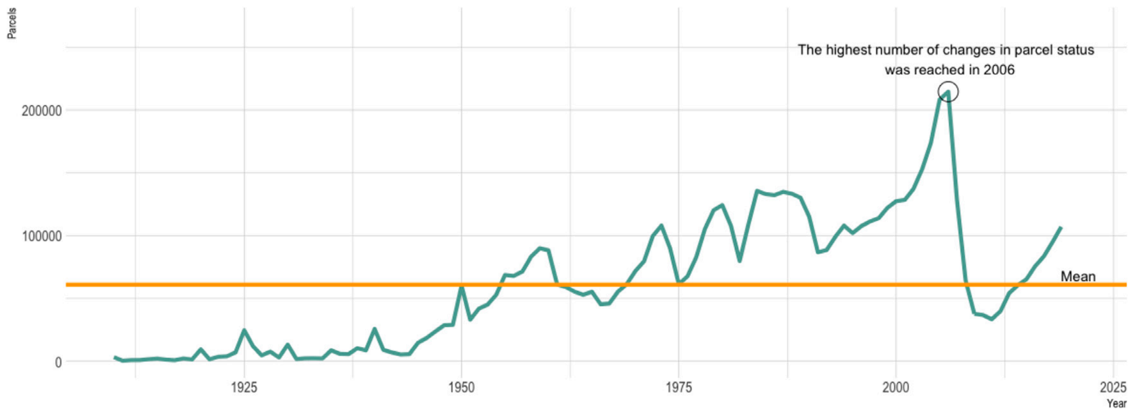


Fig. 6. Numbers of parcels changed status in Florida between 1910 and 2019.

to 10-year data to reduce the imbalance of binary categories where the number of zeros is substantially larger than the number of ones. Further descriptive statistics and details about spatial neighborhood structures are presented in the supplementary document.

5. Main results

This section presents model results of LR-, RF- and ANN-based LUC models. Models are constructed using the 10-year aggregated parcel data. After the aggregation, there are 11-time points (1910–1919, 1920–1929, …, 2010–2019). The most recent time point (2010–2019) is excluded from model training for evaluating the predictive ability of the introduced model. Previous studies show the importance of historical neighborhood characteristics in modeling land development dynamics (Kim et al., 2022; Tepe & Guldmann, 2017, 2020). In this study, lagged covariates account for delayed impacts of neighborhood conditions on land development potentials. Also, a spatio-temporal variable is incorporated to account for lagged spatial dependencies in land development dynamics. We tested up to three temporal lags to investigate the impact of temporal dependencies on land development dynamics. However, we did not include spatial autocorrelation coefficient (i.e., $\rho = 0$) in the LR, RF, and ANN models since the goal is to perform out-of-sample prediction.

Due to the considerable variation in parcel density across the state, we preferred the K-nearest neighbors rule to conceptualize spatial relationships, where each spatial object has the same number of neighboring parcels. We tested 100th, 500th, and 1000th nearest neighbor rules, including fixed effect and inverse distance weighted versions, to find the best rule to represent the spatial relationships. All spatial weight matrices are row-standardized to avoid the identification problem. Table 1 summarizes the tested model settings. Table S3 summarizes important metrics and estimated coefficients for tested 18 LR-based land-use change models. Model 5 is considered the best model based on the highest AUC (0.721), moderately high F1 (0.002), and high overall accuracy (0.925) scores.

In the RF-based LUC modeling framework, 20% of the whole parcel data is randomly selected and used for testing purposes. The remaining 80% of the data is included in model training. RF has hyper-parameters that must be pre-defined. In the model training, the maximum tree depth is set as 2000 to allow sufficient tree depth. RF estimators between 10 and 100 with 10 increments are tested to optimize the model. Table S4 represents important metrics for tested 18 RF-based models where each model is trained using 10 different estimators. Model 2 with 100 estimators is the best performing model based on accuracy measure, where the overall model accuracy reaches 94.8% utilizing the test data. Model 14 with 90 estimators is the best-fitting model for the data based on the F1 score (0.576), while Model 4 with 100 estimators is the best-fitting model for the data based on AUC metrics (0.931).

Table 1
Model settings.

Model	Kth degree	Standardization method for W	Temporal lags
Model 1	100 nearest neighbors	Fixed effect	1 lag
Model 2	100 nearest neighbors	Inverse distance weighting	1 lag
Model 3	500 nearest neighbors	Fixed effect	1 lag
Model 4	500 nearest neighbors	Inverse distance weighting	1 lag
Model 5	1000 nearest neighbors	Fixed effect	1 lag
Model 6	1000 nearest neighbors	Inverse distance weighting	1 lag
Model 7	100 nearest neighbors	Fixed effect	2 lags
Model 8	100 nearest neighbors	Inverse distance weighting	2 lags
Model 9	500 nearest neighbors	Fixed effect	2 lags
Model 10	500 nearest neighbors	Inverse distance weighting	2 lags
Model 11	1000 nearest neighbors	Fixed effect	2 lags
Model 12	1000 nearest neighbors	Inverse distance weighting	2 lags
Model 13	100 nearest neighbors	Fixed effect	3 lags
Model 14	100 nearest neighbors	Inverse distance weighting	3 lags
Model 15	500 nearest neighbors	Fixed effect	3 lags
Model 16	500 nearest neighbors	Inverse distance weighting	3 lags
Model 17	1000 nearest neighbors	Fixed effect	3 lags
Model 18	1000 nearest neighbors	Inverse distance weighting	3 lags

In LUC modeling based on the ANN framework, 20% of the entire parcel data is randomly selected and used for testing purposes. The remaining 80% of the data is used to train ANN models, while 20% of the training data is dedicated to validation purposes used in back-propagation processes during the training. Multiple hidden dense layers account for non-linear relationships (Nair & Hinton, 2010; Nwankpa et al., 2018). The introduced ANN model has three dense hidden layers. The activation functions and parameters are a *softmax* function with 300 parameters, another *softmax* function with 100 parameters, and a *sigmoid* function with 1 parameter. We tested a higher number of parameters for the *softmax* function. However, the trained models resulted in the over-fitting issue. Therefore, we reduced the number of parameters used in the function. Further, 100 Epochs were used during training for ANN models where the batch size was 1024. The optimizer "*Adam*" was used to estimate neuron weights, while the binary cross-entropy was used as the loss function. *Adam* is a gradient-based optimization algorithm for stochastic objective functions (Kingma & Ba, 2014). Essential metrics for training, validation, and test data are presented for tested 18 ANN-based LUC models in Table S5. We use an alternative metric for the F1 score in ANN models because the Keras library does not provide an F1 score option in model training. Alternatively, we calculated the Precision–Recall Curve (PRC). However, F1 scores are computed for evaluating ANN model performance using the test data. Model 1 is the best-fitting model based on test accuracy (92.6%). In contrast, Model 7 is considered the best-performing model based on the PRC (0.122), and Model 15 is the best model based on AUC metrics (0.597). Further comparisons of identified best models are discussed in Section 5.1. Overall, the RF-based LUC models provide higher accuracy for the minority category than LR and ANN models.

5.1. Model comparisons

The best models identified based on metrics are further investigated in this section. ANN-based models are further trained to calibrate hyperparameters (see the Supplementary document Section S3.1). Table S6 in the supplementary document presents the identified 9 best model results. Model metrics like Accuracy, F1 score, and AUC are included to compare these models. The confusion matrices for 9 best models are provided in the supplementary document. Because of the excessive number of zeros in the data set, the overall accuracy measure may provide false confidence in the model. Alternatively, accurate prediction of no change in land use status is as important as predicting a change in land use status. RF-based models 2 & 4 (see Figures S7d & S7e) correctly predict the significant majority of no land-use changes (nearly 99% accuracy for this category) and the most status changes in land use (slightly over 40% for the category) compared to other model results. Therefore, in the next section, these models are used to evaluate the predictive ability of the introduced modeling framework.

We also computed feature importance for each model parameter using an appropriate method to investigate the impacts of included parameters. Since LR is considered a linear model and provides coefficients for parameters, we computed model coefficients based on the standardized training data. For RF-based models, feature importance is computed as the mean and standard deviation of accumulation of the impurity decrease within each tree (further details can be found in Breiman (2001)). Eq. (5) shows the Gini impurity calculation for a classification task. Finally, the importance of jth feature in ANN models is computed using Eq. (6). The value of the difference (Δ_j) indicates the importance of feature j (please see Wei et al. (2015) for more details).

$$\text{Gini Impurity} = \sum_{i=1}^{C} f_i(1 - f_i), \tag{5}$$

$$\Delta_j = \left| \Phi^* - \frac{1}{M}\sum_{i=1}^{M} \Phi_{i,j} \right|, \tag{6}$$

where f_i is the frequency of class i and C is the number of unique classes. Δ_j is the importance of feature j; Φ^* is the base model accuracy; $\Phi_{i,j}$ is the model accuracy for jth feature at mth trial; M is the number of trials.

Feature importance is summarized in Fig. 7. Share of single-family residential lands within a 2-mile radius is the most significant feature in 7 out of 9 models. It is identified as the essential second variable in the remaining two models. The share of vacant lands is the second most important feature. Small multi-family residential lands and one-story store lands are essential features in these best models. Irwin et al. (2003) and Tepe and Guldmann (2020) highlight the importance of residential shares in land development. Shares of agricultural and institutional lands are also identified as significant features (Carrion-Flores & Irwin, 2004; Deng & Srinivasan, 2016; Irwin et al., 2003; Liao & Wei, 2014). Finally, these models' first and second spatio-temporal lags are other powerful features. Many previous studies also support the importance of spatio-temporal lags in LUC modeling (Ferdous & Bhat, 2013; Nahuelhual et al., 2012; Tepe & Guldmann, 2017, 2020).

5.2. Predictive ability of the model

In this research, we performed two types of predictions. In the previous section, we conducted a spatial out-of-sample prediction using the models trained 80% of the data covering the periods between 1910 and 2009. Such prediction allows us to estimate the parcel status for an unknown location, similar to the *Kriging* method. Since we also have temporal lags in our modeling framework, we can also perform temporal out-of-sample predictions. Fig. 8 illustrates the conceptual framework for predicting future changes in parcel status using a trained model, where only one spatio-temporal lag is included to simplify the illustration. Change in parcel status at time $t+1$ depends on the lagged covariates and spatio-temporal variable at time t. The spatio-temporal variable consists of the spatial weight matrix and changes in parcel status at time t. Once changes in parcel status at time $t+1$ are estimated, the spatio-temporal variable is updated using the predicted values and estimations of changes in parcel status at time $t+2$ is dependent on the lagged covariates (at time t) and the updated spatio-temporal variable (at time $t+1$).

The recent time point (2010–2019) data were excluded from the initial modeling training to evaluate our modeling frameworks' temporal out-of-sample prediction ability. When the best-fitting models (Models 2, 4, and 14) are used to predict land-use changes between 2010 and 2019, the overall model accuracies are around 94%, where the F1 score is very low (roughly 0.02). These models provided weak performances for predicting changes in parcel status. Therefore, we applied an alternative training strategy to improve temporal out-of-sample prediction. Random splitting of training and test sets is not suitable for temporal dynamics. Therefore, we applied a rolling cross-validation approach to train our RF-based models (Nicholson et al., 2017). In this method, models are trained using previous years' data, then evaluated using future time points. We split the data covering between 1910 and 2009 into three consecutive data sets (1910–1949, 1950–1979, and 1980–2009). In addition, we introduce a new binary variable that indicates the parcel change in status in previous years to account for the land development rates in Florida over time. Furthermore, a set of site-specific, proximity, and socio-economic variables is added to the covariate matrix to improve temporal out-of-sample predictions because previous studies discussed in the literature review section highlighted the importance of these variables (see Section S3.3 in the supplementary document for more details). The new training method substantially improves the model accuracy measures. Such a variable helps to increase the overall model prediction accuracy. Table S7 in the supplementary document presents all tested models' metrics. In the table, Figure of Metric (FoM) scores are also computed for all models (Martino et al., 2019).

The best-performing model is selected based on accuracy scores. The RF-based LUC model 13 is based on the data covering the period between 1980 and 2009 with three spatio-temporal lags where the fixed effect is used for the spatial matrix, and the number of estimators is 5. Model 7 is considered as the second best-performing model. Model 13 was used to predict new developments between 2010 and 2019. The best-performing model accurately predicts 92.0% of the changes in land use statuses, where the F1 score is 0.373, the AUC is 0.902, and the FoM is 0.053. Fig. 9 illustrates the actual and predicted land-use status changes in Florida at the parcel level between 2010 and 2019. Based on the actual data, 544,845 parcels changed their land-use status between 2010 and 2019, while the model correctly identifies 39.5% of them. In Florida, 8,450,818 parcels did not alter their land use status during the same period. The model accurately predicts 95.7% of these no-status change cases.

The proposed RF-based LUC modeling framework is further evaluated due to imbalanced data. The parcel-level data is divided into balanced subgroups using random sampling. 10,000 equally-distributed observations are drawn from the full data set for each time point, so a sample data contains 30,000 parcels. We created ten balanced samples to evaluate the consistency of accuracy scores. Table S8 shows the average accuracy scores of all 18 models using the sets of ten equally-sized samples, where standard deviations of these accuracy scores are extremely small, indicating the stability of the proposed model with different samples. Based on the four metrics presented in the Table, Model 13 also provides the best results. Especially the model accurately predicts both outcomes (change or no change in status).

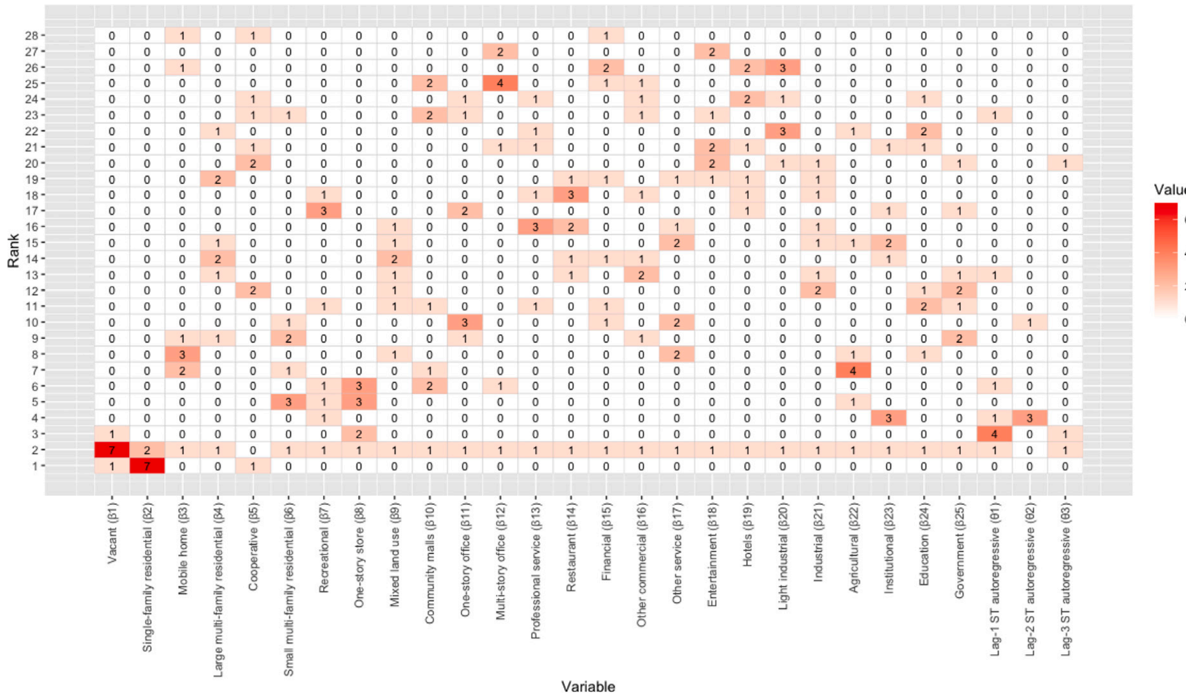


Fig. 7. Rankings of features among the best models based on feature importance.

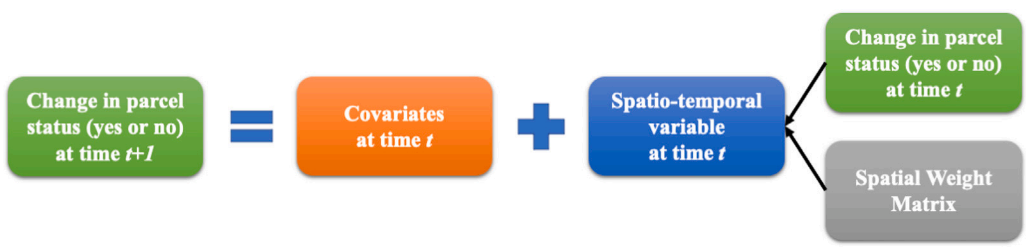
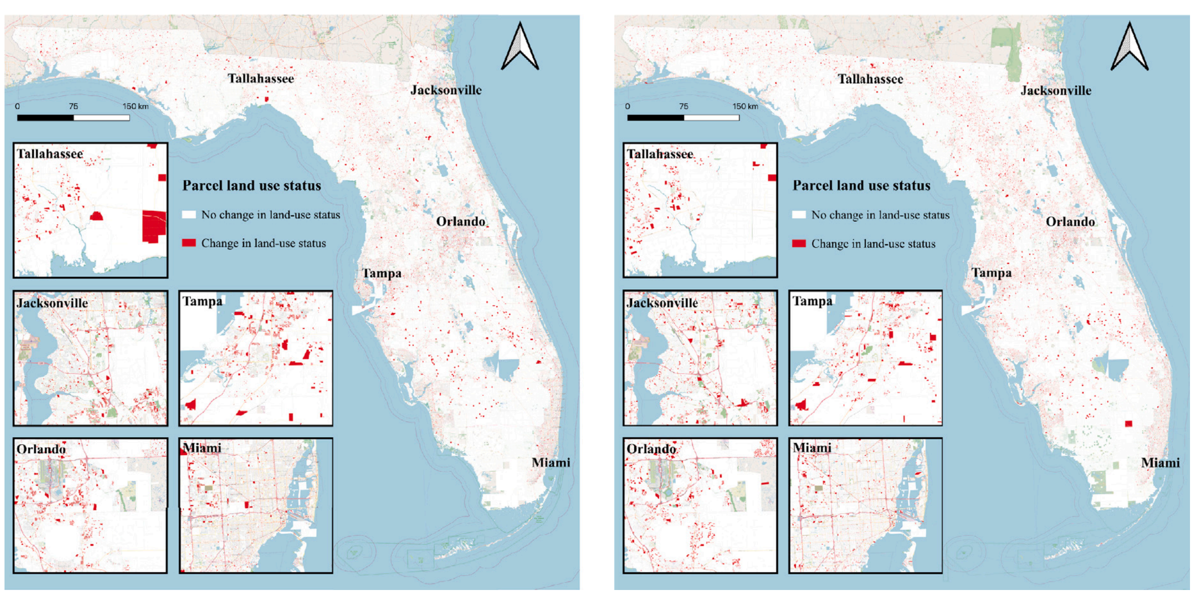


Fig. 8. Predicting future changes in parcel status using a trained model.



(a) Actual parcel land use status (b) Predicted parcel land use status

Fig. 9. Parcel land use status between 2010 and 2019.

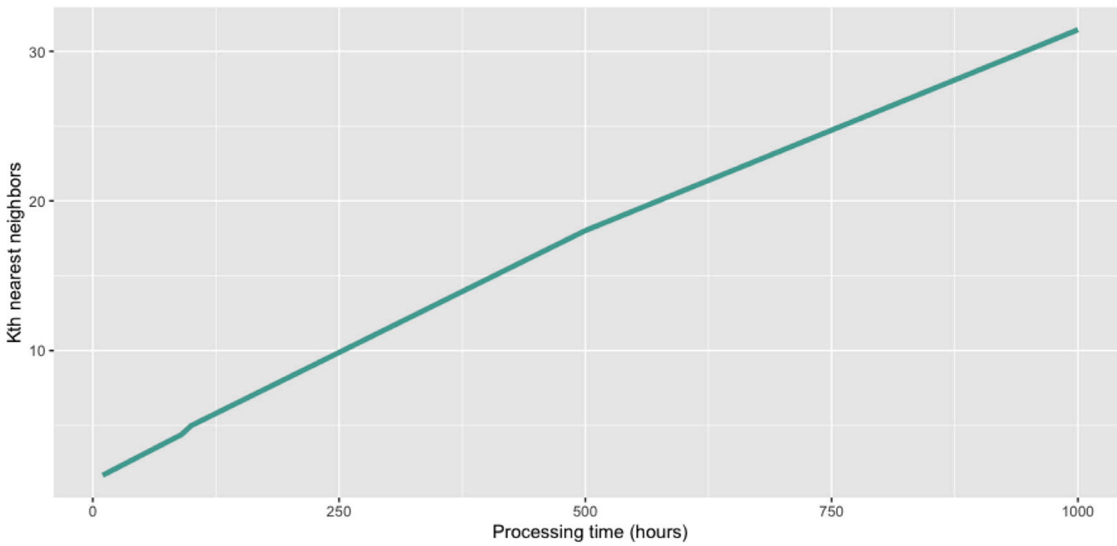


Fig. 10. Processing time of spatial weight matrices.

5.3. Computational advancements

The computational challenges of constructing a spatial weight matrix are presented in Section 3.2. Fig. 10 presents the processing times of constructing spatial weight matrices based on a range of the Kth nearest neighbor rule using our adaptive Hashing algorithm. All tests are performed using HiPerGator NVIDIA DGX™ A100 nodes (Research Computing, 2021). If the exact computation is completed using serial processing, the estimated computation time is 167,075 h for any number of Kth nearest neighbors. Our algorithm accelerates the processing times substantially for small Kth values and slows down as the Kth value increases. The average acceleration is 15,704 times in our tests where Kth values range between 10 and 1000, with an increment of 10.

Matrix multiplications of lagged changes in parcel status ($\boldsymbol{Y}_{t-1}$), and the spatial weight matrix ($\boldsymbol{W}$) are performed using *SciPy* sparse matrix functions to reduce excessive memory usage. We also use GPU parallel processing to speed up model training, and all model computations are performed using HiPerGator NVIDIA DGX™ A100 nodes. We used *cuML* GPU machine learning algorithms introduced under the *RAPID* project (Raschka et al., 2020) for LR and RF models and GPU-enabled TensorFlow (Abadi et al., 2015) module for ANN models. Fig. 11 shows the gained speed-up in model training for LR, RF, and ANN models when GPU parallel processing is enabled. We performed benchmark tests on a single supercomputer node using the same data set. For LR, the maximum iterations are set as 1000. For RF, the maximum tree depth is defined as 2000, where the number of estimators, bins, and streams are 30, 15, and 8, respectively. For the ANN model, we tested a sophisticated model with a total number of parameters is 1,000,000, while the epoch is 10 to evaluate computational gain. The most considerable acceleration is observed for LR training. GPU-based processing accelerates LR model training approximately 527 times faster than CPU equivalents. We also significantly improve the processing of RF model training by utilizing GPU parallel processing. GPU-enabled process speeds up the training almost 122 times faster than its CPU equivalent. Finally, ANN model training also benefits from GPU parallel processing. Our benchmark test shows a 49 times speed-up in the processing. The benefit of GPU parallel processing becomes more significant as ANN models get more sophisticated. In other words, acceleration rates increase as the number of model parameters rises.

The introduced adaptive Hashing algorithm substantially reduces computational challenges and necessary memory space when the number of neighbors or the size of fixed distance in neighborhood rules is relatively small to the extent of data space. GPU parallel processing significantly changes processing time when a sizeable Kth degree or radius is used for constructing spatial weight matrices. The performance of the proposed approach is also tested using multiple nodes and GPUs using our statewide data. Our performance tests show a slight improvement in computation time when multiple nodes and GPUs are used. Therefore, our introduced approach can also run large spatial data sets on personal computers, preferably with CUDA-enabled GPU cards.

6. Discussion

In recent years, spatial and temporal modeling methods have been applied to various LUC models. Incorporating spatial and temporal dynamics provides robust models with limited information because land investment decisions depend on the other investors' decisions in the neighborhood (network) and past conditions. However, an explicit involvement of spatial relationships in statistical and ML models is computationally challenging due to the requirement of searching all possible pairs in the data while constructing the spatial weight matrix. In statistical methods, parameter estimations are also computationally complex due to the computation of the inverse of spatial weight matrices which are generally squared matrices. However, ML and DL methods are computationally feasible alternatives, which can account for non-linear dynamics required for mimicking actual dynamics in models. However, accounting for spatial dynamics in ML/DL methods has been recently introduced, and these methods generally require large training data sets to achieve desired model accuracy levels. Therefore, a feasible modeling approach is required. Our proposed modeling framework successfully deals with the computational challenges in constructing spatial weight matrices and estimating necessary model parameters. The best-fitted spatial and temporal out-of-sample predictions reach around 94% and 92% overall accuracies, respectively. As discussed in the literature section, controlling spatial and temporal dependencies also provide more reliable accuracy measures. Such a high accuracy using only neighborhood characteristics provides excellent opportunities for further development of LUC models to predict future land developments. Accurate expectations about future land developments reduce many uncertainties for local governments and enable them to make long-term investments confidently and use their financial resources efficiently.

Our model results show that spatial and temporal dynamics at the parcel level are essential factors in modeling land-use changes using limited proxy information. Parcel-level data sets provide rich information about heterogeneous land development dynamics. A statewide

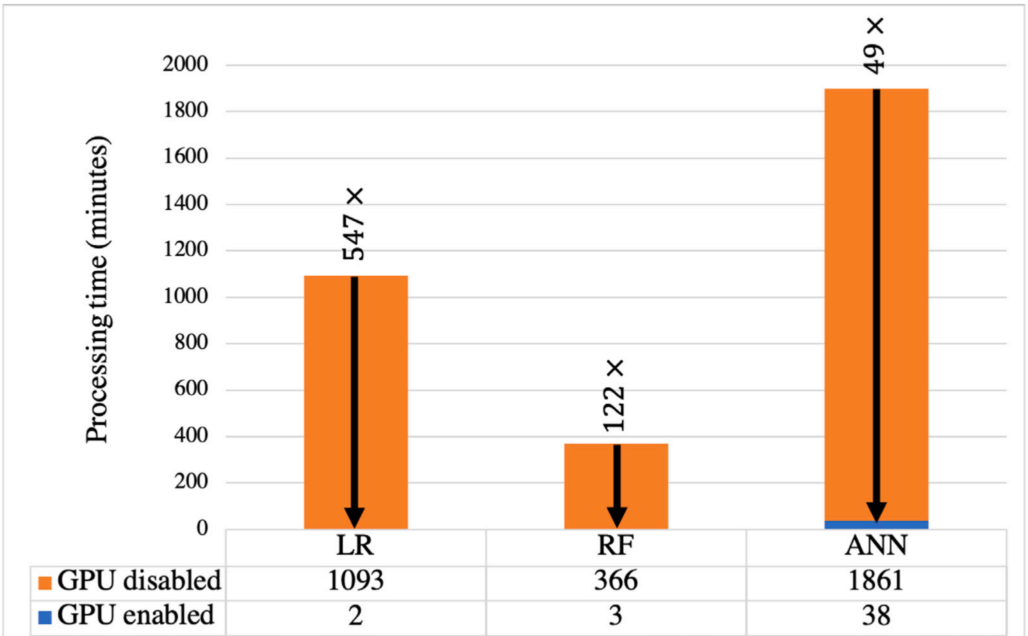


Fig. 11. Comparison of model training processing times using GPU parallel processing and GPU disabled.

model successfully captures many local and regional dynamics and improves models' spatial out-of-sample prediction accuracy powers. We also concluded that the spatio-temporal models should be trained differently to improve temporal out-of-sample prediction accuracy. The rolling cross-validation method significantly improved the model accuracy for the minority class. However, there are still possibilities to improve the model's accuracy. As we investigate historical dynamics, our modeling framework did not capture some external impacts, such as the financial crises and significant changes in demographics and socio-economic structures. Demographic changes such as population increase, median age, and household size should be included in LUC models. In addition, certain socio-economic variables such as median income, home values, education attainments, and employment ratio would provide insightful information to LUC models. Therefore, model prediction power will benefit from the involvement of more demographic and socio-economic variables where these variables summarize overall macro and micro economic trends.

7. Conclusion

This paper introduced a computationally feasible spatio-temporal LUC modeling framework using adaptive Hashing algorithms and GPU parallel processing. The proposed modeling framework address the three main limitations in LUC modeling; (1) accounting for spatial and temporal relationships, (2) controlling non-linear dynamics among features, and (3) constructing high-resolution models at the regional scale. Based on our best knowledge, our modeling framework is the first parcel-level statewide LUC model accounting for spatio-temporal dynamics, where the data set contains information about nearly 9 million parcels between 1910 and 2019. We used UF HiPerGator 3.0 supercomputer to train LR, RF, and ANN models. We accelerated the construction of the spatial weight matrix by almost 16,000 times using our adaptive Hashing algorithm and GPU parallel processing. Also, we gain substantial improvements in computation time for model training using GPU parallel processing. Our computational advancements can also benefit statistical and simulation-based LUC models such as CA. In addition, the introduced computational improvements enable researchers to complete many large-scale modeling tasks using preferably GPU-enabled personal computers.

In the empirical model, 25 neighborhood characteristic variables are incorporated. The model results provided accurate results only using neighborhood characteristics and spatio-temporal components. Including spatial and temporal dependencies significantly increases the overall accuracy because these dynamics provide insights to model unobserved information in land development dynamics. The best-fitting model achieved almost 92% overall temporal out-of-sample accuracy. Therefore, our modeling framework has many potentials for predicting future land-use changes or evaluating future land developments based on various scenarios. Also, our introduced computationally feasible modeling approaches can be easily applied to a broad spectrum of research areas, such as hedonic pricing, traffic forecasting, urban energy consumption, ecological systems, disease spread, trade relations, and business networks.

Among the advantages mentioned, the proposed method has two limitations which are interesting future research questions: (1) modeling multiple land use categories where the data will be highly imbalanced; (2) incorporating contemporaneous spatial components for out-of-sample predictions similar to spatial econometric models like Spatial Autoregressive Model (Anselin, 1988).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Emre Tepe reports financial support was provided by National Science Foundation.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the University of Florida Research [grant number AGR DTD 12-02-20, 2020]; the National Science Foundation, United States [grant number 2124507, 2021]. The authors would like to thank anonymous reviewers for helpful comments.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.scs.2023.104390.

References

Abadi, Martín, Agarwal, Ashish, Barham, Paul, Brevdo, Eugene, Chen, Zhifeng, Citro, Craig, Corrado, Greg S., Davis, Andy, Dean, Jeffrey, Devin, Matthieu, Ghemawat, Sanjay, Goodfellow, Ian, Harp, Andrew, Irving, Geoffrey, Isard, Michael, Jia, Yangqing, Jozefowicz, Rafal, Kaiser, Lukasz, Kudlur, Manjunath, ... Zheng, Xiaoqiang (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. URL https://www.tensorflow.org/. Software available from tensorflow.org.

Anselin, Luc (1988). *Spatial econometrics: methods and models*. Springer Netherlands.

Bahadori, Mohammad Taha, Yu, Qi (Rose), & Liu, Yan (2014). Fast multivariate spatio-temporal analysis via low rank tensor learning. In *Advances in neural information processing systems* (pp. 3491–3499).

Bahn, Volker, & McGill, Brian J. (2013). Testing the predictive performance of distribution models. *Oikos*, *122*(3), 321–331. http://dx.doi.org/10.1111/j.1600-0706.2012.00299.x, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2012.00299.x.

Basse, Reine Maria, Omrani, Hichem, Charif, Omar, Gerber, Philippe, & Bódis, Katalin (2014). Land use changes modelling using advanced methods: Cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale. *Applied Geography*, [ISSN: 0143-6228] *53*, 160–171. http://dx.doi.org/10.1016/j.apgeog.2014.06.016, URL https://www.sciencedirect.com/science/article/pii/S0143622814001325.

Berry, Michael W., Hazen, Brett G., MacIntyre, Rhonda L., & Flamm, Richard O. (1996). LUCAS: A system for modeling land-use change. *IEEE Computational Science & Engineering*, *3*, 24–35.

Bhat, Chandra R., Dubey, Subodh K., Alam, Mohammad Jobair Bin, & Khushefati, Waleed H. (2015). A new spatial multiple discrete-continuous modeling approach to land use change analysis. *Journal of Regional Science*, *55*(5), 801–841.

Breiman, Leo (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Carrion-Flores, Carmen, & Irwin, Elena G. (2004). Determinants of residential land-use conversion and sprawl at the rural-urban fringe. *American Journal of Agricultural Economics*, *86*(4), 889–904.

Ceci, Michelangelo, Corizzo, Roberto, Fumarola, Fabio, Malerba, Donato, & Rashkovska, Aleksandra (2017). Predictive modeling of PV energy production: How to set up the learning task for a better prediction? *IEEE Transactions on Industrial Informatics*, *13*, 956–966.

Chomitz, Kenneth M., & Gray, David A. (1996). Roads, land use, and deforestation: A spatial model applied to Belize. *The World Bank Economic Review*, *10*(3), 487–512.

Cressie, N., & Wikle, C. K. (2011). *Coursesmart series*, *Statistics for spatio-temporal data*. Wiley, ISBN: 9780471692744, URL https://books.google.it/books?id=-kOC6D0DiNYC.

Delasalles, Edouard, Ziat, Ali, Denoyer, Ludovic, & Gallinari, Patrick (2019). Spatio-temporal neural networks for space-time data modeling and relation discovery. *Knowledge and Information Systems*, *61*, 1241–1267.

Deng, Yu, & Srinivasan, Sumeeta (2016). Urban land use change and regional access: A case study in Beijing, China. *Habitat International*, *51*, 103–113.

Fan, Jianqing, Ma, Cong, & Zhong, Yiqiao (2019). A selective overview of deep learning. arXiv preprint arXiv:1904.05526.

Ferdous, Nazneen, & Bhat, Chandra R. (2013). A spatial panel ordered-response model with application to the analysis of urban land-use development intensity patterns. *Journal of Geographical Systems*, *15*, 1–29.

Flamm, R. O., & Turner, M. G. (1994). Alternative model formulations for a stochastic simulation of landscape change. *Landscape Ecology*, *9*(1), 37–46.

Fu, Fei, Deng, Shuman, Wu, Dan, Liu, Wenwen, & Bai, Zhonghua (2022). Research on the spatiotemporal evolution of land use landscape pattern in a county area based on CA-Markov model. *Sustainable Cities and Society*, [ISSN: 2210-6707] *80*, Article 103760. http://dx.doi.org/10.1016/j.scs.2022.103760, URL https://www.sciencedirect.com/science/article/pii/S2210670722000919.

Gasch, Caley K., Hengl, Tomislav, Gräler, Benedikt, Meyer, Hanna, Magney, Troy S., & Brown, David J. (2015). Spatio-temporal interpolation of soil water, temperature, and electrical conductivity in 3D + T: The cook agronomy farm data set. *Spatial Statistics*, [ISSN: 2211-6753] *14*, 70–90. http://dx.doi.org/10.1016/j.spasta.2015.04.001, URL https://www.sciencedirect.com/science/article/pii/S2211675315000251. Spatial and Spatio-Temporal Models for Interpolating Climatic and Meteorological Data.

Gounaridis, Dimitrios, Chorianopoulos, Ioannis, Symeonakis, Elias, & Koukoulas, Sotirios (2019). A random forest-cellular automata modelling approach to explore future land use/cover change in attica (Greece), under different socio-economic realities and scales. *Science of the Total Environment*, [ISSN: 0048-9697] *646*, 320–335. http://dx.doi.org/10.1016/j.scitotenv.2018.07.302, URL https://www.sciencedirect.com/science/article/pii/S0048969718328006.

Guan, DongJie, Li, HaiFeng, Inohae, Takuro, Su, Weici, Nagaie, Tadashi, & Hokao, Kazunori (2011). Modeling urban land use change by the integration of cellular automaton and Markov model. *Ecological Modelling*, [ISSN: 0304-3800] *222*(20), 3761–3772. http://dx.doi.org/10.1016/j.ecolmodel.2011.09.009, URL https://www.sciencedirect.com/science/article/pii/S0304380011004698.

Gudmundsson, L., & Seneviratne, S. I. (2015). Towards observation-based gridded runoff estimates for Europe. *Hydrology and Earth System Sciences*, *19*(6), 2859–2879. http://dx.doi.org/10.5194/hess-19-2859-2015, URL https://hess.copernicus.org/articles/19/2859/2015/.

Hastie, Trevor, Tibshirani, Robert, & Friedman, Jerome (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Hazen, B. C., & Berry, M. W. (1997). The simulation of land-cover change using a distributed computing environment. *Simulation Practice and Theory*, *5*, 489–514.

Hefley, Trevor J., Hooten, Mevin B., Hanks, Ephraim M., Russell, Robin E., & Walsh, Daniel P. (2017). Dynamic spatio-temporal models for spatial data. *Spatial Statistics*, [ISSN: 2211-6753] *20*, 206–220. http://dx.doi.org/10.1016/j.spasta.2017.02.005, URL https://www.sciencedirect.com/science/article/pii/S2211675316300689.

Herold, Martin, Goldstein, Noah C., & Clarke, Keith C. (2003). The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment*, [ISSN: 0034-4257] *86*(3), 286–302. http://dx.doi.org/10.1016/S0034-4257(03)00075-0, URL https://www.sciencedirect.com/science/article/pii/S0034425703000750. Urban Remote Sensing.

Huang, Bo, Zhang, Li, & Wu, Bo (2009). Spatiotemporal analysis of rural-urban land conversion. *International Journal of Geographical Information Science*, *23*(3), 379–398.

Irwin, Elena G., Bell, Kathleen P., & Geoghegan, Jacqueline (2003). Modeling and managing urban growth at the rural-urban fringe: A parcel-level model of residential land use change. *Agricultural and Resource Economics Review*, *32*(1), 83–102.

Jahanishakib, Fatemeh, Mirkarimi, Seyed Hamed, Salmanmahiny, Abdolrassoul, & Poodat, Fatemeh (2018). Land use change modeling through scenario-based cellular automata Markov: improving spatial forecasting. *Environmental Monitoring and Assessment*, *190*(6), 332.

Juel, Anders, Groom, Geoffrey Brian, Svenning, Jens-Christian, & Ejrnæs, Rasmus (2015). Spatial application of random forest models for fine-scale coastal vegetation classification using object based analysis of aerial orthophoto and DEM data. *International Journal of Applied Earth Observation and Geoinformation*, [ISSN: 1569-8432] *42*, 106–114. http://dx.doi.org/10.1016/j.jag.2015.05.008, URL https://www.sciencedirect.com/science/article/pii/S0303243415001178.

Karimi, Firoozeh, Sultana, Selima, Shirzadi Babakan, Ali, & Suthaharan, Shan (2019). An enhanced support vector machine model for urban expansion prediction. *Computers, Environment and Urban Systems*, [ISSN: 0198-9715] *75*, 61–75. http://dx.doi.org/10.1016/j.compenvurbsys.2019.01.001, URL https://www.sciencedirect.com/science/article/pii/S0198971518304332.

Keckler, Stephen W., Dally, William J., Khailany, Brucek, Garland, Michael, & Glasco, David (2011). GPUs and the future of parallel computing. *IEEE Micro*, *31*(5), 7–17. http://dx.doi.org/10.1109/MM.2011.89.

Kim, Yuna, Safikhani, Abolfazl, & Tepe, Emre (2022). Machine learning application to spatio-temporal modeling of urban growth. *Computers, Environment and Urban Systems*, [ISSN: 0198-9715] *94*, Article 101801. http://dx.doi.org/10.1016/j.compenvurbsys.2022.101801, URL https://www.sciencedirect.com/science/article/pii/S019897152200045X.

Kingma, Diederik P., & Ba, Jimmy (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Lau, Kwok Hung, & Kam, Booi Hon (2005). A cellular automata model for urban land-use simulation. *Environment and Planning B: Planning and Design*, *32*(2), 247–263. http://dx.doi.org/10.1068/b31110.

Li, Qingmei, Feng, Yongjiu, Tong, Xiaohua, Zhou, Yilun, Wu, Peiqi, Xie, Huan, Jin, Yanmin, Chen, Peng, Liu, Shijie, Xv, Xiong, Liu, Sicong, & Wang, Chao (2022). Firefly algorithm-based cellular automata for reproducing urban growth and predicting future scenarios. *Sustainable Cities and Society*, [ISSN: 2210-6707] *76*, Article 103444. http://dx.doi.org/10.1016/j.scs.2021.103444, URL https://www.sciencedirect.com/science/article/pii/S2210670721007174.

Liao, Felix H. F., & Wei, Y. H. Dennis (2014). Modeling determinants of urban growth in Dongguan, China: a spatial logistic approach. *Stochastic Environmental Research and Risk Assessment*, *28*(4), 801–816. http://dx.doi.org/10.1007/s00477-012-0620-y.

Lu, Yi, Laffan, Shawn, Pettit, Chris, & Cao, Min (2020). Land use change simulation and analysis using a vector cellular automata (CA) model: A case study of Ipswich City, Queensland, Australia. *Environment and Planning B: Urban Analytics and City Science*, *47*(9), 1605–1621. http://dx.doi.org/10.1177/2399808319830971, URL https://doi.org/10.1177/2399808319830971.

Lv, Jianjun, Wang, Yifan, Liang, Xun, Yao, Yao, Ma, Teng, & Guan, Qingfeng (2021). Simulating urban expansion by incorporating an integrated gravitational field model into a demand-driven random forest-cellular automata model. *Cities*, [ISSN: 0264-2751] *109*, Article 103044. http://dx.doi.org/10.1016/j.cities.2020.103044, URL https://www.sciencedirect.com/science/article/pii/S0264275120313925.

Maria de Almeida, Cláudia, Batty, Michael, Vieira Monteiro, Antonio Miguel, Câmara, Gilberto, Soares-Filho, Britaldo Silveira, Cerqueira, Gustavo Coutinho, & Pennachin, Cássio Lopes (2003). Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation. *Computers, Environment and Urban Systems*, [ISSN: 0198-9715] *27*(5), 481–509. http://dx.doi.org/10.1016/S0198-9715(02)00042-X, URL https://www.sciencedirect.com/science/article/pii/S019897150200042X.

Martino, Alessio, De Santis, Enrico, Baldini, Luca, & Rizzi, Antonello (2019). Calibration techniques for binary classification problems: A comparative analysis. In *IJCCI 2019*, *Proceedings of the 11th international joint conference on computational intelligence* (pp. 487–495). Setubal, PRT: SCITEPRESS - Science and Technology Publications, Lda, ISBN: 9789897583841, http://dx.doi.org/10.5220/0008165504870495.

Meyer, Hanna, Reudenbach, Christoph, Hengl, Tomislav, Katurji, Marwan, & Nauss, Thomas (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, [ISSN: 1364-8152] *101*, 1–9. http://dx.doi.org/10.1016/j.envsoft.2017.12.001, URL https://www.sciencedirect.com/science/article/pii/S1364815217310976.

Meyer, Hanna, Reudenbach, Christoph, Wöllauer, Stephan, & Nauss, Thomas (2019). Importance of spatial predictor variable selection in machine learning applications – moving from data reproduction to spatial prediction. *Ecological Modelling*, [ISSN: 0304-3800] *411*, Article 108815. http://dx.doi.org/10.1016/j.ecolmodel.2019.108815, URL https://www.sciencedirect.com/science/article/pii/S0304380019303230.

Micheletti, Natan, Foresti, Loris, Robert, Sylvain, Leuenberger, Michael, Pedrazzini, Andrea, Jaboyedoff, Michel, & Kanevski, Mikhail (2014). Machine learning feature selection methods for landslide susceptibility mapping. *Mathematical Geosciences*, *46*(1), 33–57. http://dx.doi.org/10.1007/s11004-013-9511-0.

Moreno, Niandry, Wang, Fang, & Marceau, Danielle J. (2008). An object-based land-use cellular automata model to overcome cell size and neighborhood sensitivity. In *International society for photogrammetry and remote sensing - GEOBIA 2008 - pixels, objects, intelligence geographic object based image analysis for the 21st century, Vol. XXXVIII*.

Nahuelhual, Laura, Carmona, Alejandra, Lara, Antonio, Echeverria, Cristian, & Gonzalez, Mauro E. (2012). Land-cover change to forest plantations: Proximate causes and implications for the landscape in south-central Chile. *Landscape and Urban Planning*, *107*, 12–20.

Nair, Vinod, & Hinton, Geoffrey E. (2010). Rectified linear units improve restricted boltzmann machines. In *ICML*.

Nicholson, William B., Matteson, David S., & Bien, Jacob (2017). VARX-L: Structured regularization for large vector autoregressions with exogenous variables. *International Journal of Forecasting*, [ISSN: 0169-2070] *33*(3), 627–651. http://dx.doi.org/10.1016/j.ijforecast.2017.01.003, URL https://www.sciencedirect.com/science/article/pii/S0169207017300080.

Nwankpa, Chigozie, Ijomah, Winifred, Gachagan, Anthony, & Marshall, Stephen (2018). Activation functions: Comparison of trends in practice and research for deep learning. arXiv preprint arXiv:1811.03378.

Okwuashi, Onuwa, & Ndehedehe, Christopher E. (2021). Integrating machine learning with Markov chain and cellular automata models for modelling urban land use change. *Remote Sensing Applications: Society and Environment*, [ISSN: 2352-9385] *21*, Article 100461. http://dx.doi.org/10.1016/j.rsase.2020.100461, URL https://www.sciencedirect.com/science/article/pii/S2352938520306364.

Pan, Ying, Roth, Andreas, Yu, Zhenrong, & Doluschitz, Reiner (2010). The impact of variation in scale on the behavior of a cellular automata used for land use change modeling. *Computers, Environment and Urban Systems*, [ISSN: 0198-9715] *34*(5), 400–408. http://dx.doi.org/10.1016/j.compenvurbsys.2010.03.003, URL https://www.sciencedirect.com/science/article/pii/S0198971510000244.

2019 Florida parcel data statewide. (2019). URL https://www.fgdl.org/metadataexplorer/explorer.jsp.

Pinheiro, E. A. L., Camini, N. A., Soares, M. R. S., & Sumida, S. S. (2020). Cellular automata model - landscape dynamics simulation tool in the process of change in land use and cover in the city of gacha do norte – mt. In *2020 IEEE Latin American GRSS & ISPRS remote sensing conference* (pp. 225–229). http://dx.doi.org/10.1109/LAGIRS48042.2020.9165590.

Raschka, Sebastian, Patterson, Joshua, & Nolet, Corey (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. arXiv preprint arXiv:2002.04803.

Ren, Yuanfang, & Wu, Yan (2014). Convolutional deep belief networks for feature extraction of eeg signal. In *2014 international joint conference on neural networks (IJCNN)* (pp. 2850–2853).

Research Computing (2021). University of Florida HiPerGator. URL http://researchcomputing.ufl.edu.

Roberts, David R., Bahn, Volker, Ciuti, Simone, Boyce, Mark S., Elith, Jane, Guillera-Arroita, Gurutzeta, Hauenstein, Severin, Lahoz-Monfort, José J., Schröder, Boris, Thuiller, Wilfried, Warton, David I., Wintle, Brendan A., Hartig, Florian, & Dormann, Carsten F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. http://dx.doi.org/10.1111/ecog.02881, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.02881.

Ron-Ferguson, Nathan, Chin, Jae Teuk, & Kwon, Youngsang (2021). Leveraging machine learning to understand urban change with net construction. *Landscape and Urban Planning*, [ISSN: 0169-2046] *216*, Article 104239. http://dx.doi.org/10.1016/j.landurbplan.2021.104239, URL https://www.sciencedirect.com/science/article/pii/S0169204621002024.

Shafizadeh-Moghadam, Hossein, Minaei, Masoud, Pontius, Robert Gilmore, Jr., Asghari, Ali, & Dadashpoor, Hashem (2021). Integrating a forward feature selection algorithm, random forest, and cellular automata to extrapolate urban growth in the tehran-karaj region of Iran. *Computers, Environment and Urban Systems*, [ISSN: 0198-9715] *87*, Article 101595. http://dx.doi.org/10.1016/j.compenvurbsys.2021.101595, URL https://www.sciencedirect.com/science/article/pii/S0198971521000028.

Soares-Filho, Britaldo, Rodrigues, Hermann, & Follador, Marco (2013). A hybrid analytical-heuristic method for calibrating land-use change models. *Environmental Modelling & Software*, [ISSN: 1364-8152] *43*, 80–87. http://dx.doi.org/10.1016/j.envsoft.2013.01.010, URL https://www.sciencedirect.com/science/article/pii/S1364815213000236.

Stevens, Daniel, & Dragićević, Suzana (2007). A GIS-based irregular cellular automata model of land-use change. *Environment and Planning B: Planning and Design*, *34*(4), 708–724. http://dx.doi.org/10.1068/b32098.

Supinie, Timothy A., McGovern, Amy, Williams, John, & Abernathy, Jennifer (2009). Spatiotemporal relational random forests. In *2009 IEEE international conference on data mining workshops* (pp. 630–635). http://dx.doi.org/10.1109/ICDMW.2009.89.

Talukdar, Swapan, Eibek, Kutub Uddin, Akhter, Shumona, Ziaul, Sk, Towfiqul Islam, Abu Reza Md., & Mallick, Javed (2021). Modeling fragmentation probability of land-use and land-cover using the bagging, random forest and random subspace in the Teesta river basin, Bangladesh. *Ecological Indicators*, [ISSN: 1470-160X] *126*, Article 107612. http://dx.doi.org/10.1016/j.ecolind.2021.107612, URL https://www.sciencedirect.com/science/article/pii/S1470160X21002776.

Tepe, Emre, & Guldmann, Jean-Michel (2017). Spatial and temporal modeling of parcel-level land dynamics. *Computers, Environment and Urban Systems*, [ISSN: 0198-9715] *64*, 204–214. http://dx.doi.org/10.1016/j.compenvurbsys.2017.02.005, URL https://www.sciencedirect.com/science/article/pii/S0198971516301880.

Tepe, Emre, & Guldmann, Jean-Michel (2020). Spatio-temporal multinomial autologistic modeling of land-use change: A parcel-level approach. *Environment and Planning B: Urban Analytics and City Science*, *47*(3), 473–488. http://dx.doi.org/10.1177/2399808318786511.

Ulloa-Espíndola, René, & Martín-Fernández, Susana (2021). Simulation and analysis of land use changes applying cellular automata in the south of Quito and the Machachi valley, province of Pichincha, Ecuador. *Sustainability*, [ISSN: 2071-1050] *13*(17), http://dx.doi.org/10.3390/su13179525, URL https://www.mdpi.com/2071-1050/13/17/9525.

U. S. Census Bureau (1910). *1910 census: Volume 1. Population, general report and analysis*: *Technical report*, URL https://www.census.gov/library/publications/1913/dec/vol-1-population.html.

U. S. Census Bureau (2019). State population totals: 2010–2019. *Technical report*, URL https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-total.html.

Veldkamp, A., & Fresco, L. O. (1996). CLUE: a conceptual model to study the conversion of land use and its effects. *Ecological Modelling*, *85*, 253–270.

Verburg, Peter H, van Eck, Jan R Ritsema, de Nijs, Ton C M, & Martin J Dijst, Paul Schot (2004). Determinants of land-use change patterns in The Netherlands. *Environment and Planning B: Planning and Design*, *31*, 125–150.

Wang, Junye, Bretz, Michael, Dewan, M. Ali Akber, & Delavar, Mojtaba Aghajani (2022). Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. *Science of the Total Environment*, [ISSN: 0048-9697] *822*, Article 153559. http://dx.doi.org/10.1016/j.scitotenv.2022.153559, URL https://www.sciencedirect.com/science/article/pii/S0048969722006519.

Wei, Pengfei, Lu, Zhenzhou, & Song, Jingwen (2015). Variable importance analysis: A comprehensive review. *Reliability Engineering & System Safety*, [ISSN: 0951-8320] *142*, 399–432. http://dx.doi.org/10.1016/j.ress.2015.05.018, URL https://www.sciencedirect.com/science/article/pii/S0951832015001672.

Wikle, C. K., & Hooten, M. B. (2010). A general science-based framework for dynamical spatio-temporal models. *Test*, *19*(3), 417–451. http://dx.doi.org/10.1007/s11749-010-0209-z, URL http://pubs.er.usgs.gov/publication/70037651.

Xing, Weiran, Qian, Yuehui, Guan, Xuefeng, Yang, Tingting, & Wu, Huayi (2020). A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation. *Computers & Geosciences*, [ISSN: 0098-3004] *137*, Article 104430. http://dx.doi.org/10.1016/j.cageo.2020.104430, URL https://www.sciencedirect.com/science/article/pii/S0098300419307708.

Yang, Xin, Zheng, Xin-Qi, & Lv, Li-Na (2012). A spatiotemporal model of land use change based on ant colony optimization, Markov chain and cellular automata. *Ecological Modelling*, [ISSN: 0304-3800] *233*, 11–19. http://dx.doi.org/10.1016/j.ecolmodel.2012.03.011, URL https://www.sciencedirect.com/science/article/pii/S0304380012001184.

Yu, Jingyan, Hagen-Zanker, Alex, Santitissadeekorn, Naratip, & Hughes, Susan (2021). Calibration of cellular automata urban growth models from urban genesis onwards - a novel application of Markov chain Monte Carlo approximate Bayesian computation. *Computers, Environment and Urban Systems*, [ISSN: 0198-9715] *90*, Article 101689. http://dx.doi.org/10.1016/j.compenvurbsys.2021.101689, URL https://www.sciencedirect.com/science/article/pii/S019897152100096X.

Zhai, Yaqian, Yao, Yao, Guan, Qingfeng, Liang, Xun, Li, Xia, Pan, Yongting, Yue, Hanqiu, Yuan, Zehao, & Zhou, Jianfeng (2020). Simulating urban land use change by integrating a convolutional neural network with vector-based cellular automata. *International Journal of Geographical Information Science*, *34*(7), 1475–1499. http://dx.doi.org/10.1080/13658816.2020.1711915.

Zhou, Liang, Dang, Xuewei, Sun, Qinke, & Wang, Shaohua (2020). Multi-scenario simulation of urban land change in shanghai by random forest and CA-Markov model. *Sustainable Cities and Society*, [ISSN: 2210-6707] *55*, Article 102045. http://dx.doi.org/10.1016/j.scs.2020.102045, URL https://www.sciencedirect.com/science/article/pii/S2210670720300329.