Parcel and Sub-Parcel Land Use Change in Coastal South Carolina: A Case Study in Beaufort County

Kang Shou Lu, Jeffery Allen, John Morgan and Mathew Sedecki

The South Carolina Water Resources Center

Strom Thurmond Institute

Clemson University

And

Department of Geography

Towson University

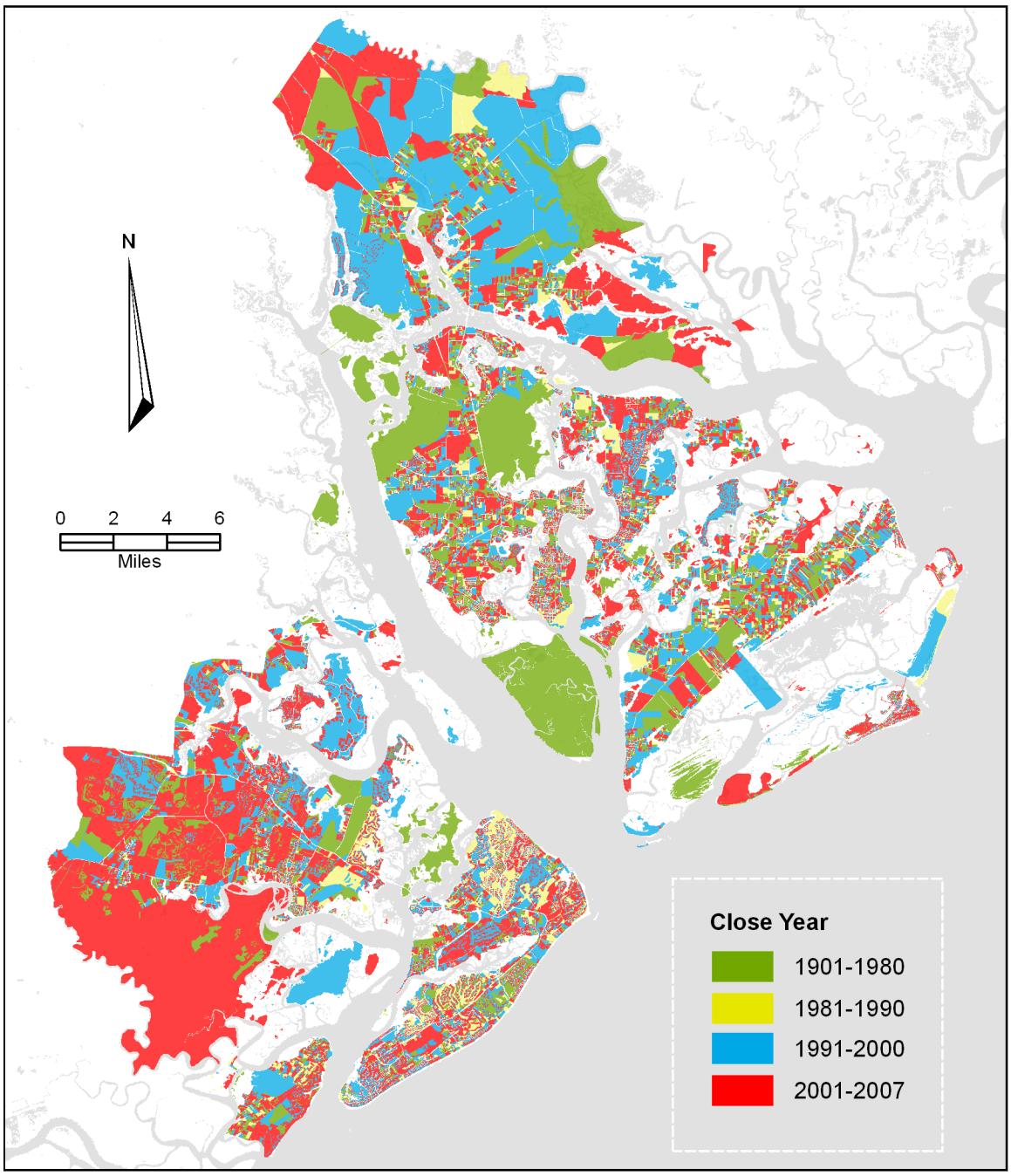
The complexity of land use systems makes it difficult for conventional statistical models to yield desirable predictions. For instance, logistic regression, binary or multinomial, has been the core of many land use models, but it often leads to low success rates or prediction accuracies even when assessed against the calibration dataset, particularly when multiple categories are involved. This often creates more doubt than certainty when the models are used for predicting future land use change.

This paper proposes an artificial neural network (ANN) model that has an enhanced predictive power. Literature suggests that ANNs have the architecture appropriate for modeling complex systems. They have superb approximation power and pattern identification capability desired for spatially explicit modeling and image classification. Recent studies have demonstrated ANNs' ability to improve the prediction accuracy for binary land use change. It is thus hypothesized that ANNs will succeed in a similar fashion when used for predicting the spatio-temporal change of multi-category land use systems.

In this research, we used a three-layer perception network with multiple input units, hidden units, and output units. Weights for units of the hidden layer and output layer were automatically derived from a sample dataset through a robust backpropagation training process. We applied a Bayesian Regularization (BR) algorithm for network training in order to select an optimal model sufficiently trained for greater approximating power and yet general enough for better predictive capability.

We tested the ANN land use model in a coastal area with nearly 10,000 parcels, 33% of which are new developments over nearly two decades between 1990 and 2008. The ANN model was built with 18 variables (input units) that determine the parcel use change among five categories (output units). A large training sample dataset was extracted by using a pseudo stratified random sampling method to minimize any possibility of faulty representation of the full population due to variable shapes and sizes of parcels. The ANN model was validated spatially and temporally against two sample sets, one net-change set, and the full set. Multinomial logistic regression was also performed as a benchmark model for comparative purposes.

Results indicate that the performance of the ANN model was quite impressive and prediction accuracies range between 81.95-98.51% and kappa coefficients between 81.38-97.48% across five different categories. The ANN model also outperformed the benchmark multinomial logistic regression (MLR) model by average 17.5 percentage points across the five use categories and by 9.2 percentage points in overall prediction accuracy. The greatest improvements, by 33.62 and 32.23 percentage points respectively for high-density urban and low density residential, occurred where MLR's predictions are poor. Additional model validation was performed against four different datasets and prediction accuracies (77.6-97.7%) dropped only slightly as expected. Some data issues and modeling techniques were discussed as we recommend ANNs as useful alternatives for modeling and predicting land use changes.



Change in land ownership in Beaufort County, 1901-2007.