MODELING URBAN DYNAMICS WITH ARTIFICIAL NEURAL NETWORKS AND GIS

Chris Weisner
RAM Mobile Data USA
cwiesner@smtplink.ram.com

David J. Cowen
Department of Geography & Liberal Arts Computing Lab
University of South Carolina, USA
cowend@sc.edu

ABSTRACT

The modeling of dynamic urban systems has been of interest to spatial analysts for the better part of the past four decades, and the development of geographic information systems (GIS) has sparked continued interest in spatial process modeling. Recent research in a number of far reaching disciplines has shown artificial neural networks (ANN) to be powerful tools for modeling many dynamic systems (Vemuri and Rogers, 1994). This research investigated the possibilities for ANN's as spatial analytic tools. To this end, an artificial neural network was linked with a GIS for the purpose of modeling urban growth in sub-regions of a metropolitan area. The validity of the ANN model was tested against a linear regression model. The results of this research support the hypothesis that ANN are in fact useful spatial analytic tools and can be used to accurately model dynamic urban systems.

INTRODUCTION

Modeling and prediction of urban growth have been of interest to researchers for the better part of the past four decades (Chapin and Weiss, 1968, Batty and Longley, 1994). Much of the rationale behind this research was to determine the cause and effect of the urban form on transportation patterns and to use this knowledge for the planning of future transportation networks. Researchers were interested in the potential for computer models to enable the testing of changes in policy and urban resources on transportation networks, and thus the models proposed were deductive in nature.

The model put forth in this research was an inductive approach to the urban modeling problem, and incorporated an artificial neural network (ANN) in conjunction with a geographic information system (GIS) to model a spatio-temporal database of single family residential building permits. The model was based on the assumption that the time of occurrence and magnitude of urban growth in a sub-region of a metropolitan area is a function of the development already occurring in the sub-region and within it's neighboring areas.

METHODOLOGY

The spatial data structure created for this research was an arbitrarily defined tessellation of 2.6 square mile regular hexagons covering the two county Columbia SC study area. The benefits of using a regular hexagon tessellation was that neighborhood relations, shape, size and orientation are held constant throughout the surface. Building permits are indicators of the morphology of the urban landscape Halls, Cowen and Jensen (1994). This study used the single family residential housing units subset of a building permits database for an eleven year period. For this study, the training set included the building permit data from the years 1981 through 1989. The test set contained the data from 1990 and 1991. For this study a hexagon had to have had at least 10 permits issued at least one of the years during the period. This ensured that there was enough training set data for the neural network to find a pattern of development (fig.1).

Based on previous research it was determined that artificial neural networks model the time series of nonlinear dynamic systems by mapping the state of the system at time t, x(t), to some future state, x(t + Δt). Chakraborty et al (1992) demonstrated improved results by incorporating the time series of comparable objects (cities) as inputs, and this approach was adapted to this study by including the states of the "neighborhood" (the six surrounding hexagons) with the state of each hexagon as inputs to the ANN model (fig 2). The spatial relationships between the hexagon and its neighbors are built into the structure of the ANN through the arrangement of the input nodes and the weight connections between the input layer and the hidden layer . This arrangement is held constant throughout the study area. This has the effect of defining a regular semi-lattice organization over the entire surface (fig. 3). In terms of the specifics of the dynamic urban system, this is the urban organization argued for by Alexander (1965) in his two part essay "A City is Not a Tree". Few models have adopted this structure, opting instead for a simpler hierarchical tree-like structure.

Since the spatial relationships between neighboring hexagons are hard coded into the ANN structure, one set of network weights for the entire area would not necessarily provide the best model. In fact, the relationship between each hexagon and it's neighbors changes with respect to the central business district (CBD) throughout the study area (Fig. 4). Once the permits were partitioned in space and

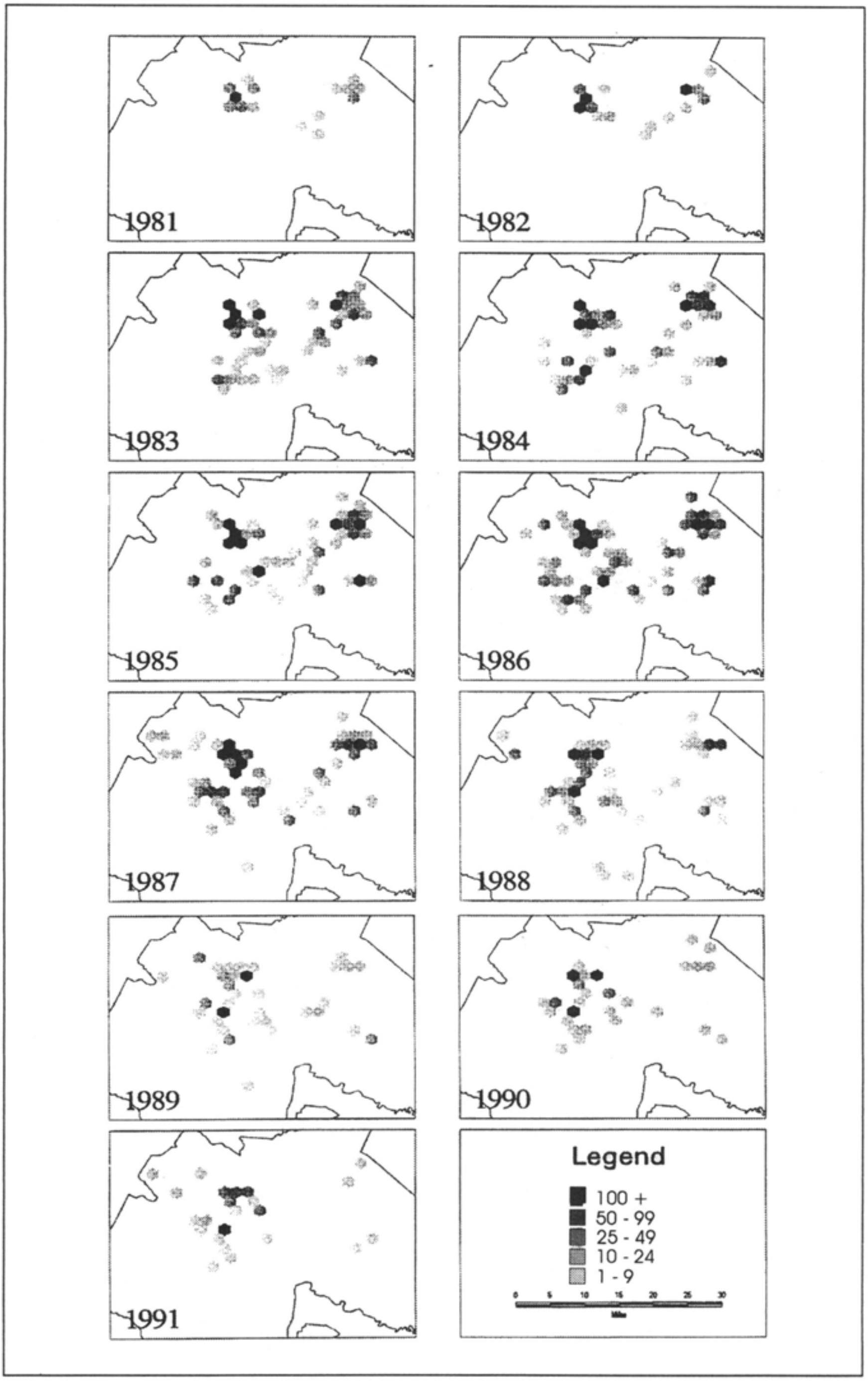

Figure 1 Distribution of building permits in hexagonal data structure.

time, their theoretical time series would start with a period of no growth corresponding to the period when the area was in non-urban land use, a short period of active growth as the urban fringe passes through the area and a final period of no development occurring when the available space in the area has become saturated with development.

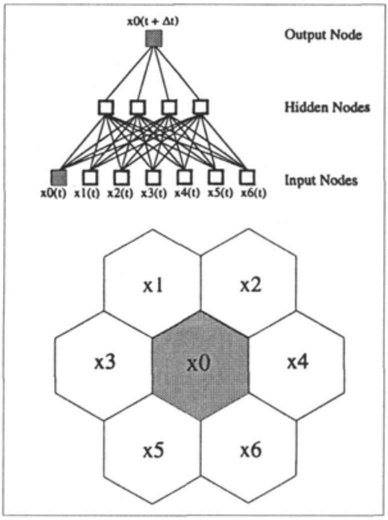

Fig. 2 Hexagonal ANN

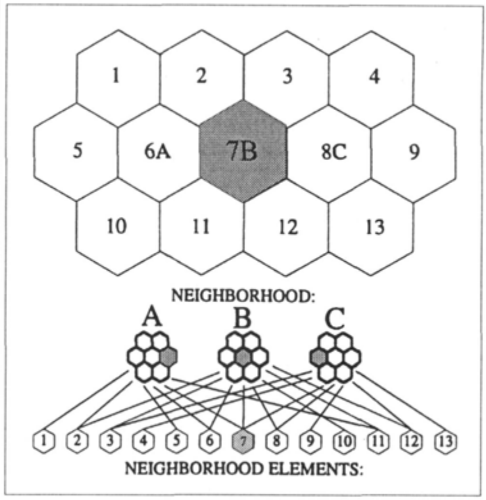

Figure 3 ANN Neighborhood

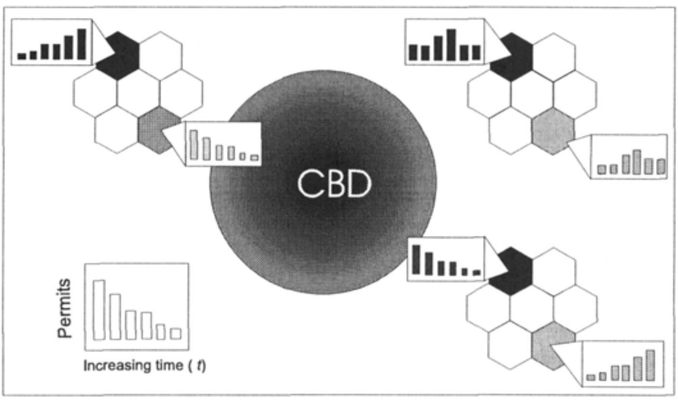

Fig. 4 Hypothetical time series from CBD

The last issue to be resolved was the development of the specifics of the artificial neural network model(s). This process involved the selection of an appropriate number of hidden nodes, the scaling of the data from real world values to ANN values, and the training threshold of the sum of the square error to be used in training the ANN. A number of researchers have indicated the difficulty of determining appropriate network architecture's (i.e. the number of hidden nodes) for modeling a wide range of data sets (Heermann and Khazenie, 1992; ; Fletcher and Goss, 1993; Lodewyck and Deng, 1993). Lowe and Webb (1991) have suggested that the number of hidden nodes represent a Euclidean dimension into which the dimension of the *attractor* of the system (which may be of fractal dimension) is embedded. For this research, initial feedback indicated that between four and two hidden nodes were adequate to produce acceptable results. Of the ninety-four ANN models chosen, forty-eight (48) utilized four hidden nodes, twenty-five (25) utilized three hidden nodes and twenty-one (21) utilized two hidden nodes. The scaling used in this study incorporated the following rules:

1. The scaling values were between 0.2 and 0.8.
2. The minimum and maximum values for each training set were determined from the center hexagon in the seven hexagon neighborhood. Values in the surrounding six hexagons which were less than the *minimum* were assigned the minimum value and values which were greater than the *maximum* were assigned the maximum value.
3. In many instances the range of activity was still quite large with many small values and a few instances of large values. In these cases, experimentation indicated that taking the log of the data values produced desirable results.

Thus, for each model that was developed, two approaches were used - ordinary linear and log-linear scaling. Twenty-four ANN models were developed for each of the ninety-four hexagons in the study area. The twenty-four models correspond to variations in the number of hidden nodes (2,3,4), the scaling method used (linear, log-linear), and the learning criteria used to end the training phase of the ANN model development. The log-linear scaling method was used by sixty (60) of the ninety-four ANN models and the straight linear scaling was used by thirty-four (34). A final consideration in specifying the model was the learning threshold for the sum of the squared error that must be reached before the training of the network weights ends. It is generally agreed that small sum-of-the-square errors attained during the training phase result in networks which have "learned" the idiosyncrasies of the training data, and may result in poor generalization to the test data set and other data the network has not previously "seen". To adjust for this, this study tested four different training thresholds, 0.1, 0.25, 0.4 and 0.55, at which time adjustment of the network weights stopped. The "best" model for each hexagon was chosen as the one with the lowest sum of the square error on the test data set (1990,1991). This model was then used in all subsequent analysis.

MODEL EVALUATION

The ANN predictions for the years 1982 through 1990 were generated by iterating the model forward one year using actual data from the year before as inputs. For example, actual 1981 data was used to produce predictions for 1982 and so forth. The 1991 ANN prediction is a two iteration case in which the model predictions for 1990 were used as inputs to produce the prediction. The linear trend model is plotted as the regression line representing the trend of the data between 1981 and 1989 extended through the test set years of 1990 and 1991. For each hexagon, the "best" model was chosen as the one which had the lowest sum of the square error on the test data set (fig. 5). In most cases the ANN model was able to lock into a pattern of development in the training data and produce predictions which were superior to the linear trend model. The ANN models with small learning criteria (< 0.1) approximate the trend of training set data quite well, while those models with larger learning criteria do not represent the training data as well.

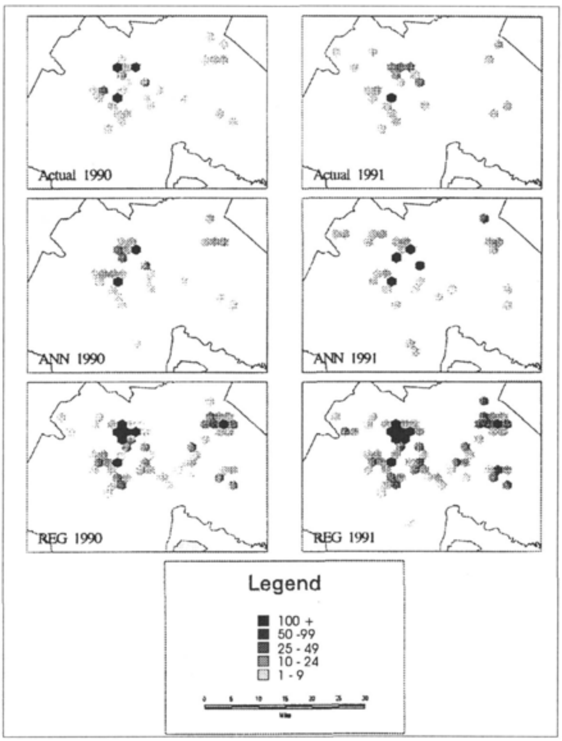

Fig. 5 Predicted building permits versus actual

A few cases illustrate how some of the models predicted the building permit data (fig. 6). In case 1 an ANN model was able to accurately approximate the nonlinear trend of the building permit data, including the test data set. Case 2 is an interesting case in which the trend of the building permit data does not match the hypothetical time series. The ANN model was able to pick up on the appropriate pattern and predict the increase in permits occurring in 1991. The prediction is not drastically different from the linear trend model in this year, but the ability of the ANN model to adjust for this upswing is evident. Case 3 is justification for using higher learning criteria during the training phase of the model development. The increased learning criteria allowed the ANN model to ignore what may be noise in the training set data and still have the ability to model the general trend of the data and produce desirable results on the test set data. This property gives ANN's a distinct advantage over linear trends when modeling dynamic systems. Case 4 illustrates a problem involved with modeling dynamic systems. In this case the model has picked up on an inappropriate trend in the data and has projected the growth upwards in 1991 when in actuality it is tending towards zero.

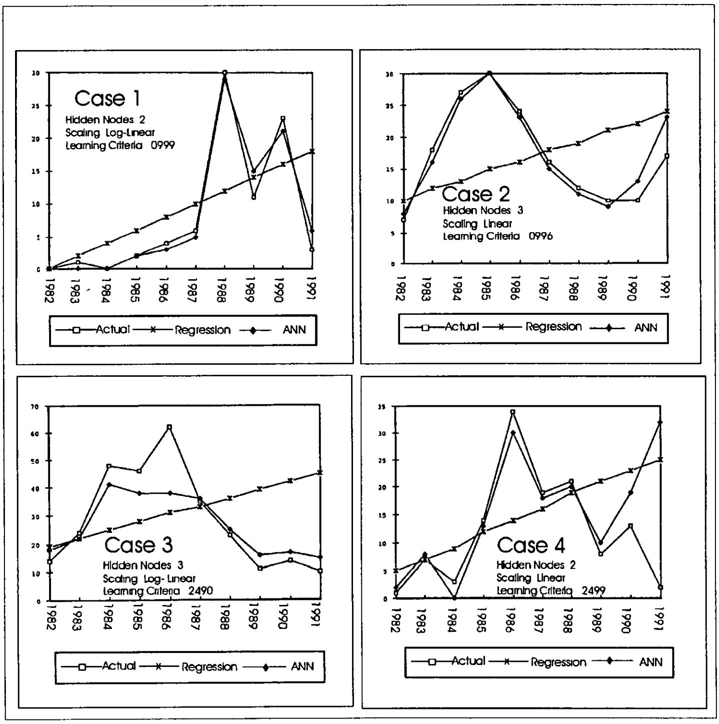

Fig. 6 Four examples of ANN and linear predictions

The final evaluation was based on regressing each model's (ANN and regression) predictions against the actual building activity occurring within each hexagon in both 1990 and 1991 (fig. 7). This resulted in four bivariate regression equations . Over the 94 hexagons in the study area, the ANN models were shown to be superior to the linear trend models in that the ANN models produced predictions which were closer to the actual data values than did the linear trend model. For the 1990 ANN one iteration model the regression parameters for the one iteration ANN model predictions for 1990 the r^2 was 0.83. In contrast the regression model had an r^2 of 0.61. The intercept for the regression estimate was 9.49 which was significantly different than zero. These parameters indicate that the linear trend predictions for 1990 consistently over-predicting the number of building permits throughout the study area.

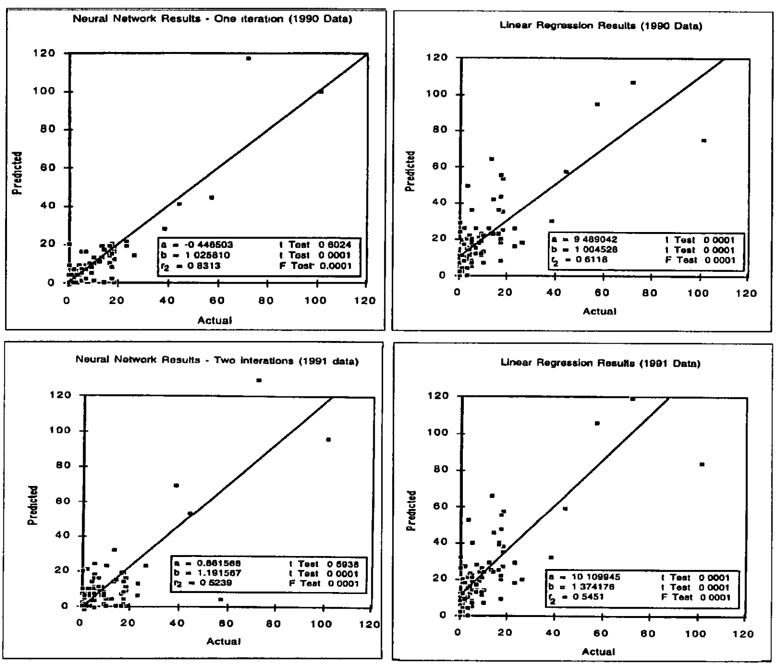

Fig. 7 Regression results of ANN versus linear model.

The regression model for the two iteration ANN model predictions for 1991 and the actual number of permits had an r^2 of only 0.52 and a slope of 1.19. The expected growth of the error term between the trajectory of the ANN model predictions and the actual data values is apparent in these results. The fact that the two iteration case had an intercept of zero and a slope near 1 which indicate that while the models do not perform consistently for all hexagons, they do produce results which are around the desired values. The comparison of the linear trend predictions for 1991 and the actual number of permits generated an r^2 of 0.55 with a slope of 1.37. As was the case with the results of the linear trend on the 1990 test data, the linear predictions for 1991 are consistently higher than the actual number of permits. The difference between the two approaches is most clearly demonstrated by the fact that the intercept of the ANN model is approximately zero (0.86) and the intercept of the linear trend model significantly greater than zero (10.11).

CONCLUSIONS

A spatial temporal database was constructed by aggregating a database of building permit data to a tessellation of regular hexagons. Artificial neural network techniques were used to develop models that replicated the time series for each hexagon. These models were evaluated by comparing the predictions of the models for two years of data the model did not see during the training phase of model development against the predictions from a linear trend model. The results of the study provide strong evidence for power of an ANN to model non-linear trends. For the one iteration case, the ANN model was able to produce predictions over the entire study area which closely resembled the actual values, while the linear trend model produced results which consistently overestimated the actual number of permits. The models were able to adjust to the variations in the building permit data without being aware of fluctuations in the local economy, available land for development, accessibility, etc. The success of the approach used here is encouraging for modeling systems, such as urban dynamics, for which the relationships between the underlying mechanisms are not well understood and for which precise data is not available.

BIBLIOGRAPHY

Alexander, C. (1965) "A City is Not a Tree", *Architectural Forum* **122** (1), 58-61 and (2), 58-62.

Batty, M, and Xie, Y. (1994) "Urban Analysis in a GIS Environment" from *Spatial Analysis and GIS*, edited by Fotheringham, S. and Rogerson, P., Taylor & Francis, Bristol, PA, pp. 189-219.

Chakraborty, K., Mehrotra, K., Mohan, C.K. and Ranka, S. (1992) "Forecasting the Behavior of Multivariate Time Series Using Neural Networks", *Neural Networks*, vol. 5, no. 6, pp. 961-970.

Chapin F.C. and Weiss S.F. (1968) "A Probabilistic Model for Residential Growth", *Transportation Research* vol. 2, pp. 375-390.

Fletcher, D. and Goss, E. (1993) Forecasting with neural networks: An application using bankruptcy data, Information and Management, 24, Elsevier .

Halls, J.N., Cowen, D.J. and Jensen, J.R. (1994) "Predictive Spatio-Temporal Modeling in GIS", *Proceedings of the Sixth International Symposium on Spatial Data Handling*, vol. 1, pp. 431-448.

Heermann, P.D. and Khazenie, N. (1992) Classification of Multispectral Remote Sensing Data Using a Back-Propagation Neural Network, IEEE Transactions on Geoscience and Remote Sensing, Vol. 30, No. 1, January 1992, IEEE.

Lodewyck, R.W. and Deng, P, (1993) Experimentation with a back-propagation neural network: An application to planning end user system development, Information and Management, 24, Elsevier Science Publishers.

Lowe, D. and Webb, A.R. (1991) "Time series prediction by adaptive networks: a dynamical systems perspective", *IEEE Proceedings-F*, vol. 128, no. 1, February 1991, pp. 17-24.

Vemuri, V.R.. and Rogers, R.D.. (1994) *Artificial Neural Networks: Forecasting Time Series*, IEEE Computer Society Press, Los Alamitos, CA.