An ANN Model for the Influence of Siding Materials on Single-Family Home Values

Mohammed E. Haque, Ph.D., P.E., Amro Taibah, Ph.D.

Texas A&M University, Texas, USA/ King Abdulaziz University, Saudi Arabia

Abstract

The real estate valuation is a complex process considering the range of variables that are known to play a role in determining such a value. The esthetics and choice of building's façade materials can greatly influence a homebuyer's decision. This paper evaluated such buyer's preferences of various siding materials by estimating the market-clearing prices determined by the buyers of single-family homes (SFHs). The study employs a comparative hedonic estimation and artificial neural network (ANN) model. The study relied on a large sample of SFHs that were sold during the period from 1997 to 2000 in College Station, Texas. The analysis was restricted to homes that are highly homogeneous in their structural attributes in order to eliminate the impact of such attributes on the home values. The main aim of this study was to develop an ANN model to determine the influence of four different siding materials on home values. Several different ANN models with different layers/slabs connections, weights, and activation functions were trained. The presented ANN Jordan-Elman Nets with logistic activation function was the best one among all other trials. It converged very rapidly to produce a very high predictive efficacy. The trained model was evaluated using the data that was not used during the training, which also indicated very good agreement between the actual and ANN predicted price. Additionally the study was aimed to compare the hedonic value estimation and ANN value prediction methods. The results indicate that stucco siding had the most significant impact on the property value. The estimated implicit values of different siding materials were a measure of the importance of such materials to the homebuyer and was resembled in the form of a paid premium. Moreover, in comparing the hedonic results and ANN results, this study found that both analytical methods support one another and have assigned similar weights to the various construction types that have been studied. In addition ANN showed to have a higher predictive accuracy level than did the hedonic estimation. The findings extend the body of literature concerned with real estate value analysis and have significant implications in the realm of fund allocation decision making for a real estate developer.

Introduction

Artificial intelligence (AI) applications have gained a broad interest in civil/construction/ architectural engineering problems. Its applications are very extensive and interdisciplinary. The graduate students civil/construction/architectural students should especially be encouraged to learn various applications of computing techniques including artificial neural network (ANN), genetic algorithm (GA), etc. This paper highlights various applications of AI. As an example of

a graduate project, this paper demonstrated an ANN model to determine the influence of siding materials on residential home values. Estimating the value of a property concerns builders, developers, homebuyers, appraisers, economists, and policy makers among many others. Nonetheless, real estate valuation is a complex process considering the range of variables that are known to play a role in determining such a value. For example, in the case of single-family home (SFH) valuation, Can[1] indicates that the price of a home is determined as a function of various structural, neighborhood, and location variables used to determine property value when using a hedonic estimation procedure.

In addition, it is well established in the literature that the larger portion of a property's value is mainly estimated by its structural attributes that include, but are not limited to, heated area, number of stories, number of bedrooms, type of construction, etc.[2] Hence, knowing to what degree a particular component impacts the overall value can provide valuable information pertaining to the significance of that particular component.

To a real estate developer, such information would provide valuable insight pertaining to allocation of funds decision-making. Generating higher proceeds from financing and/or from eventual sale of a property is the main goal of the real estate developer. He or She accomplishes this by creating a real estate market value far in excess of the real estate development cost. Creating real estate value is the most effective and powerful tool in real estate development used to maximize returns on real estate investments. It is through the reduction of needed equity, increasing returns from financing, and increasing returns from sales proceeds that higher returns are generated[3].

Therefore, understanding the specific impact of construction quality on total value is one important piece of this puzzle. This paper intends to explore the impact of construction quality, resembled by various types of siding materials, on residential property value by employing a comparative hedonic-artificial neural network approach.

Artificial Neural Networks – concepts and definitions

Neural computing is a relatively new field of artificial intelligence (AI), which tries to mimic the structure and operation of biological neural systems, such as the human brain, by creating an Artificial Neural Network (ANN) on a computer. An ANN is a modeling technique that is useful to address problems where solutions are not clearly formulated or to validate the results obtained through other modeling techniques[4]. The network has the ability to learn by example. Patterns in a series of input and output values of example cases are recognized. This acquired "knowledge" can then be used by the network to predict unknown output values for a given set of input values[5]. An ANN is composed of simple interconnected elements called processing elements (PEs) or artificial neurons that act as microprocessors. Fig. 1 illustrates a simple processing element of an ANN with three arbitrary numbers of inputs and outputs[6]. Each PE has an input and an output side. The connections on the input side correspond to the dendrites of the biological original and provide the input from other PEs while the connections on the output side correspond to the axon and transmit the output. Synapses are mimicked by providing connection weights between the various PEs and transfer functions or thresholds within the PEs. Currently, back-propagation is the most popular, effective, and easy to learn model for complex networks[6,

[7]. For the last few years, the first author has been using various ANN back-propagation Multi-layer Perceptron (MLP) modeling techniques in materials science[6, 7], structural/construction engineering[8-10], and construction management[11]. To develop a back-propagation neural network, a developer inputs known information, assigns weight to the connections within the network architecture, and runs in the networks repeatedly until the output is satisfactorily accurate. The weighted matrix of interconnections allows the neural networks to learn and remember[5]. During training, once the optimum weights are reached, the weights and biased values encode the network's state of knowledge. Thereafter, using the network on new cases is merely a matter of simple mathematical manipulation of these values.

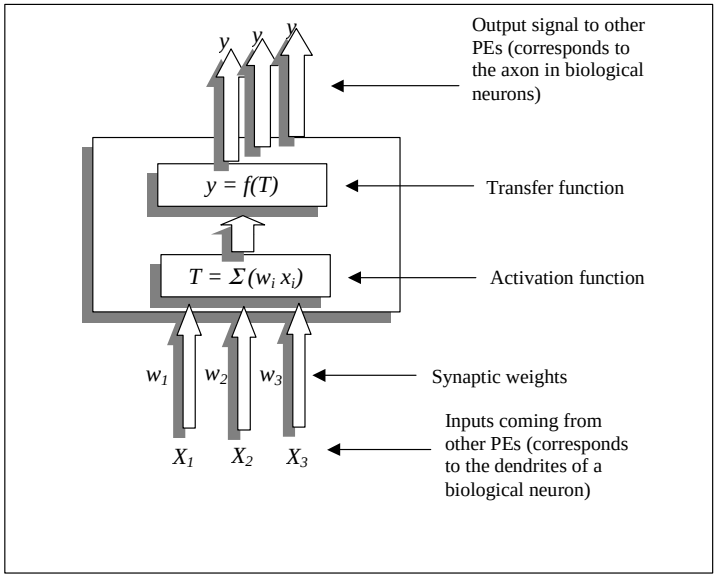

Fig. 1 Processing element of an ANN

Methodology

This study is designed to explore the potentials of combining conventional econometric methods for property value analysis, namely the hedonic price function, with ANN capabilities in assessing change, due to construction type factors, in residential property values. The methodology employed intends to compare results from both approaches to validate one another and to better understand the nature of the relationship between construction type and market value.

Study population: The study population consists of a sample of 145 highly homogeneous confirmed sales of SFHs in the city of College Station, Texas that sold during the period from 1997-2000. SFHs of the same structural characteristics were used in this analysis to eliminate the effects of such variables. This will improve the model performance and accuracy in capturing the effect of construction-type on home values. Also the study used price per square foot (PPSQFT) for this analysis to minimize variation within the sample. An initial regression was conducted including a variable accounting for the year-sold for each observation. All PPSQFT values were then transformed to 2000 value.

Data Collection: The data, obtained from the Multiple Listing Services (MLS), includes price data, structural attributes data, and address information. A local real estate broker provided data strictly for research purposes. The MLS data was validated and integrated with data from the Brazos County Appraisal District (BCAD). After cleaning, validating, and integration, only data that have complete records has been deemed suitable for this analysis.

Study variables: The variables included in the analysis are defined in Table 1, and included sales price per square foot and building age. Fig. 2 shows some of the construction types of residential houses. In addition, the data included the type of construction variables as designated by the MLS data. These included construction-veneer, construction-frame, construction-mix, and

construction-stucco. These categories represented the type of construction system used for each home and particularly the exterior siding material.

Table 1: Variable definitions.

VARIABLES	DEFINITIONS
	DEPENDENT VARIABLE
PPSQFT00	Market price/Heated Area adjusted to 2000 value
	STRUCTURAL ATTRIBUTES
AGE_YR	Dummy--Age of the building at the time it is sold in years
	CONNSTRUCTION ATTRIBUTES
CONST_VENEER	At least 85% of exterior is brick veneer
CONST_FRAME	Asbestos/T111 plywood siding
CONST_MIX	Combination of two materials
CONST_STUCO	Stucco-paint/sheet sidings.

Stucco Siding

Veneer and wood

Brick Veneer

Veneer and stucco

Fig. 2 Some of the construction types of residential houses.

Results of Hedonic Procedure

An analysis of the data was performed using the hedonic estimation procedure in order to ascertain the relationship between property market value and the type of siding material used. The hedonic price function describes the relationship between the observed prices of commodities and the characteristics associated with the commodities, based on the hypothesis that "goods are valued for their utility-bearing attributes or characteristics," [12]. Hence the coefficient estimates represent the premium paid by the homeowner for the aesthetic utility that any siding material might bare. Based on the results of the analysis, the hedonic equation can be written as follows:

$$PPSQFT00 = 52.419 - 0.515*AGE\_YR + 16.595*CONST\_VENEER + 15.734*CONST\_FRAME + 18.323*CONST\_MIX + 37.642*CONST\_STUCO + e$$

Where e = error term.

The F-value of the model was found to be statistically significant at the 0.0001 levels. The F statistic basically tests how well the model, as a whole, accounts for the dependent variable's behavior. The estimation efficacy of the model was found to be moderately high with a coefficient of determination (R^2) of 0.622. This is an indication that the model is able to explain approximately 62% of the total variation in the property's market value. The results of this study indicate that "construction type" is an important attribute that has a significant effect on SFH market value. Moreover the findings indicate that certain types of construction have a more considerable impact than others. For example, based on the study sample, it was found that using various types of stucco siding in exterior finishing could have a noticeably higher impact on future market price than other sidings. Compared to brick veneer, which is the most common exterior finishing material in the College Station market, Stucco finishing was found to reflect an increase difference of approximately $21 per square foot in market price. In addition it was found that plywood sidings had the least premium estimated at approximately $15 per square foot. Homes with mixed siding materials in the sample were mainly brick veneer mixed with either wood or stucco sidings. This category also showed an estimated hedonic price of approximately $18 per square foot.

ANN Model and Results

The parameter inputs considered in the proposed ANN model were the ones that were found to be statistically significant using the hedonic procedure (AGE_YR, CONST_VENEER, CONST_FRAME, CONST_MIX, CONST_STUCO). The output parameter was price per square foot adjusted to 2000 value (PPSQFT00). During training the model, 130 observations were used out of a total of 145 observations. The remaining 15 observations, separated arbitrarily from the total number of observations, were used during testing the trained model. Several different ANN models with different layers/slabs connections, weights, and activation functions (including logistic, linear, Symmetric Logistic, and Gaussian) were trained. In addition, pattern selections including "Rotation" and "Random" were applied with weight updates using Vanilla, Momentum and TurboProp. The presented ANN- Jordan-Elman recurrent network with linear and logistic mixed activation function ("Rotation" for pattern selection, and "Momentum" for weight updates) was the best one among all other trials. In this research, Jordan-Elman recurrent network converged very rapidly to produce a very high predictive efficacy. Fig. 3 shows the

actual and ANN predicted price ($ per Sq. Ft.) during the Network Training Phase, which indicates very low errors between the actual price and the ANN predicted price. The trained model was evaluated using the data that was not used during the training. Fig. 4 depicts the actual and ANN Predicted Price during the network evaluation phase, which indicates very small difference between the actual and ANN predicted price. During network training, R^2 was obtained as 0.7788, and 0.8627 during network evaluation, which were very close to 1.0, indicating a very good fit between the actual and the network prediction. R^2 is a statistical indicator usually applied to multiple regression analysis, and was calculated using the following formulae[13]:

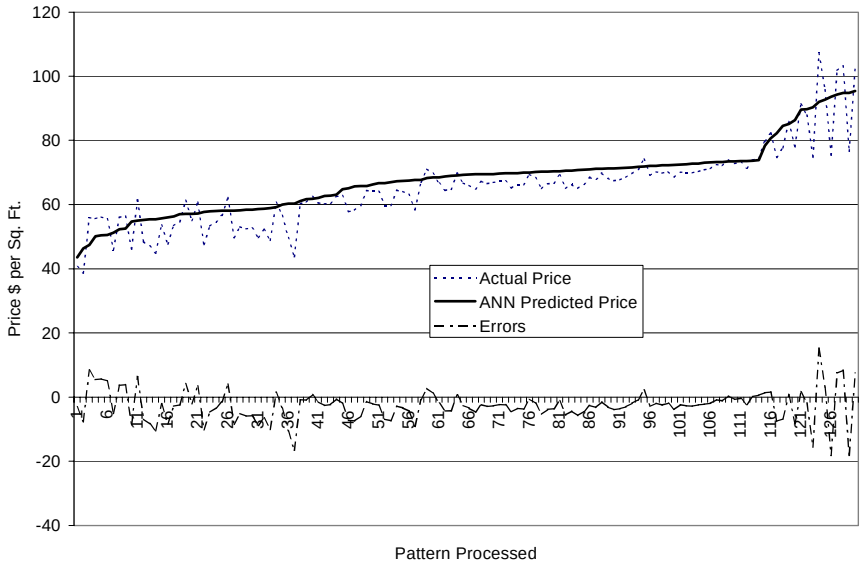

Fig. 3 Actual and ANN Predicted Price during the Network Training Phase

$$R^2 = 1 - (SSE/SS_{yy})$$

Where
$SSE = \Sigma (y - \hat{y})^2$,
$SS_{yy} = \Sigma (y - \bar{y})^2$,
y is the actual value,
$\hat{y}$ is the predicted value of y,
and $\bar{y}$ is the mean of the y values.

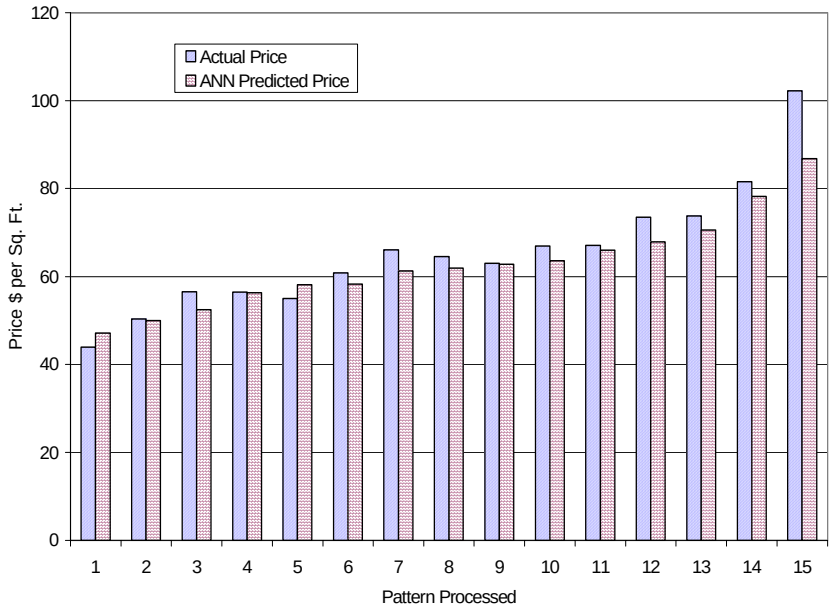

Fig. 4 Actual and ANN Predicted Price during the Network Evaluation

ANN Price Prediction

Hypothetical data was developed in order to compare the performance of the hedonic estimation model and the ANN model. The purpose is to further the validation and evaluate the performance similarities or differences between the two methods. Ten imaginary data points were used with building ages of 5 and 10 years old and siding material characteristics that were chosen arbitrarily. The two

models produced price values that were within acceptable variation of each other. Fig. 5 shows a good comparative agreement between the hedonic estimation and ANN prediction models.

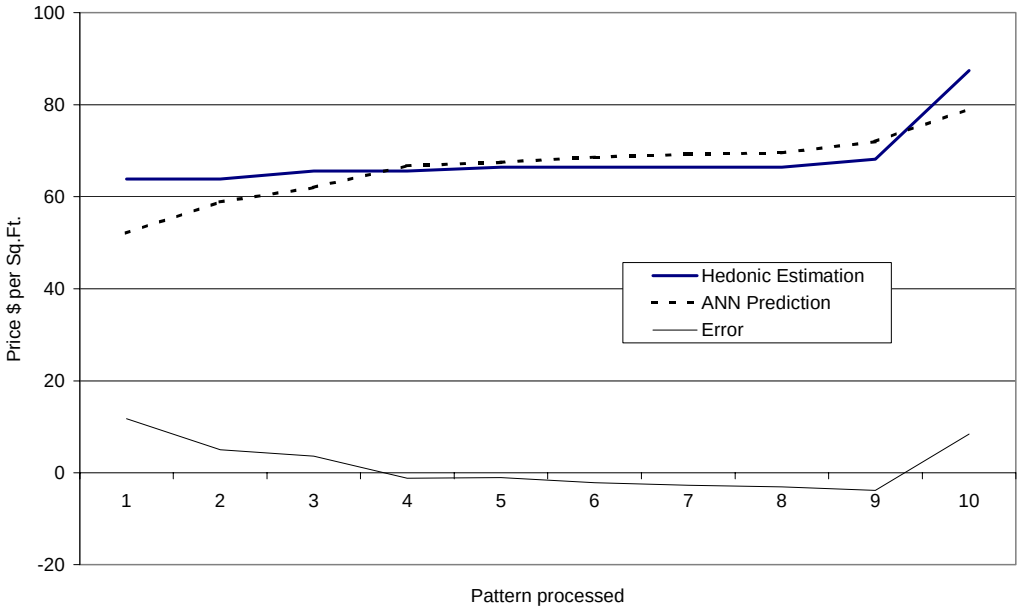

Fig. 5 ANN/hedonic prediction comparison

Educational Significance

It is increasingly important to go beyond traditional departmental course curriculum boundaries for some areas of science and engineering education. Artificial Intelligence (AI) is one such field; its applications are very extensive and interdisciplinary. Neural networks are powerful computing devices. They can process information more readily than traditional computer systems. This is due to their highly parallel architecture inspired by the structure of the brain. Applications and research into the use of neural networks have evolved from their ability to understand complex relationships and hidden patterns within large data sets. The graduate students should especially be encouraged to learn various applications of contemporary computing techniques including artificial neural network (ANN), genetic algorithm (GA), etc. Civil/construction/ architectural engineering has exercised a rapidly growing interest in the application of neurally inspired computing techniques. This paper is an example of ANN's interdisciplinary application in the fields of architectural engineering and real estate value analysis.

Conclusion

In comparing the hedonic results and ANN results this study found that both analytical methods support one another and have assigned similar weights to the various construction types that have been studied. In addition ANN showed to have a higher predictive accuracy level than did the hedonic estimation. The study showed that various construction types have various impacts on market values of SFHs. The estimated implicit values of different siding materials are a measure of the importance of such a material to the homebuyer and is resembled in the form of a paid

premium. The findings extend the body of literature concerned with real estate value analysis and have significant implications in the realm of fund allocation decision making for a real estate developer. This paper clearly demonstrated AI's extensive and interdisciplinary application, especially in the field of Architectural Engineering. This paper will motivate and encouraged graduate students to learn various applications of computing techniques including artificial neural network (ANN), genetic algorithm (GA), etc, and to use them in practical research.

Bibliography

1. Can, A., Residential Quality Assessment: An Alternative Approach Using GIS, Annals of Regional Science, 1992, 26, 97-110.
2. Sirpal, R., Empirical Modeling of the Relative Impacts of Various Sizes of Shopping Centers on the Values of Surrounding Residential Properties, Journal of Real Estate Research, 1994, Fall, 9:4, 487-505.
3. Sharkawy, A., A Physical-Financial Model for Design Economy Trade-Offs, Real Estate Research Issues: Essays in Honor of James A. Graaskamp, Delisle, J., Ed., American Real Estate Society, Madison, Wisconsin, 1994.
4. Chester, M., Neural Networks: A Tutorial, Englewood Cliffs, NJ: Prentice Hall, 1993.
5. Obermeier, K. and Barron, J., Time to Get Fried Up, BYTE, 1989, 14(8), 227-233.
6. Haque, M.E., and Sudhakar, K.V., Prediction of Corrosion-Fatigue behavior of DP Steel through Artificial Neural Network, International Journal of Fatigue, Vol. 23 (1), 2001, pp. 1-4.
7. Haque, M.E., and Sudhakar, K.V. ANN-based Prediction Model for Fatigue Crack Growth in DP Steel. International Journal of Fatigue & Fracture of Engineering Materials and Structures, Vol. 24/1, 2001, pp. 63-68.
8. Haque, M.E. and Mund, A., "Loads on Shores and Slabs during Multistory Structure Construction: An Artificial Neural Network Approach," Proceedings of the American Society for Engineering Education, 2002 ASEE Annual Conference, Section 1421 Construction Engineering, June 16-19, 2002, Montreal, Quebec, Canada.
9. Haque, M.E., "An Artificial Neural Network Model for Preliminary Design of Reinforced Concrete Beam-Column," Proceedings of the American Society for Engineering Education, 2002 ASEE Annual Conference, Section 2306 Architectural Engineering, June 16-19, 2002, Montreal, Quebec, Canada.
10. Haque, M.E. and Karandikar, V., "A Study on Comfort and Safety in a Residential Housing Complex: A Neuro-Genetic Knowledge Model," Proceedings of the Second International Conference on Information Systems in Engineering and Construction (ISEC 2002), June 13-14, 2002, Cocoa Beach, Florida.
11. Choudhury, I. and Haque, M.E. "A Study of Cross-cultural Training in International Construction Using General Linear Model Procedure and Artificial Neural Network Approach," proceedings of the 3rd International Conference on Construction Project Management (3ICCPM) Conference, 2001, pp. 444-453, Singapore.
12. Rosen, S., Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, Journal of Political Economy, January/February, 1974, 82:1, 34-55.
13. Ward Systems Group, NeuroShell 2 User's Manual. Ward Systems Group, Inc., Maryland, USA, 1996.

MOHAMMED E. HAQUE

Dr. Mohammed E. Haque is an associate professor of the Department of Construction Science at Texas A&M University at College Station, Texas. He has over fifteen years of professional experience in analysis, design, and investigation of building, bridges and tunnel structural projects of various city and state governments and private sectors. Dr. Haque is a registered Professional Engineer in the states of New York, Pennsylvania and Michigan, and members of ASEE, ASCE, and ACI. Dr. Haque received a BSCE from Bangladesh University of Engineering and Technology, a MSCE and a Ph.D. in Civil/Structural Engineering from New Jersey Institute of Technology, Newark, New Jersey. His research interests include fracture mechanics of engineering materials, composite materials and advanced construction materials, computer applications in structural analysis and design, artificial neural network applications, knowledge based expert system developments, application based software developments, and buildings/ infrastructure/ bridges/tunnels inspection and database management systems.

AMRO A. TAIBAH

Dr. Amro A. Taibah is an assistant professor in the Department of Landscape Architecture, College of Environmental Design, King Abdulaziz University, in Jeddah, Saudi Arabia. His research interests are in real estate

value analysis, Landscape design values and aesthetics, GIS-based decision support systems, geoinformatic analysis for urban planning and policymaking, and mitigating construction management issues in the Saudi market. Dr. Taibah's interests also include professional consultation in urban design and construction management. Dr. Taibah holds a B.S. in landscape architecture from King Abdulaziz University, a Masters degree in landscape architecture (MLA) from North Carolina State University, Raleigh, N.C., a Master of Science in Construction Management (MSCM) and a Ph.D. in Urban and Regional sciences from Texas A&M University, College Station, TX.