Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks

Authors Nghiep Nguyen and Al Cripps

Abstract

This article compares the predictive performance of artificial neural networks (ANN) and multiple regression analysis (MRA) for single family housing sales. Multiple comparisons are made between the two data models in which the data sample size, the functional specification and the temporal prediction are varied. ANN performs better than MRA when a moderate to large data sample size is used. For the application, this moderate to large data sample size varied from 13% to 39% of the total data sample (506 to 1,506 observations out of 3,906 total observations). The results give a plausible explanation why previous papers have obtained varied results when comparing MRA and ANN predictive performance for housing values.

Introduction

This study compares the predictive performance of multiple regression analysis (MRA) and backpropagation feed forward artificial neural network (ANN) for single family residential property value. The two data models, MRA and ANN, are compared using different functional model specifications, sample (training) data and evaluation criteria. The same specification improvements benefits both the ANNs and the MRAs, and a plausible explanation is provided as to why other studies that have compared the MRA and ANN data modeling tools to predict the value of residential property have had varied results.

To fairly compare the data models, one should address the possible methodological problems associated with each data model that might distort its performance. The studies that compare the MRA and ANN data models are identified. In addition, several studies dealing with the MRA data model specification are also identified and applied to the data set used in this study. Inherent problems (in implementing the MRA and ANN data models) that may affect the performance of each model and hence affect the results of previous studies are discussed. These findings are then applied to a new data set for comparison of the two data models. The constraints and results of the comparison are then given along with conclusions.

Although the standard feedforward neural network with backpropagation learning is used for the comparison in this article, experiments were conducted with multiple learning variations such as enhanced backpropagation, backpropagation with weight decay, quickpropagation, resilient backpropagation, backpercolation and counterpropagation. Various ANN architectures such as ARTMAP, GAUSSIAN, regression neural, etc. were also examined. After hundreds of experiments and multitudes of architectures, the standard backpropagation was found to perform better than the other neural network architectures examined for this application. In addition, since several other studies compared standard backpropagation neural network with MRA with varied results, a change of the ANN model would not have allowed a direct comparison between the results and those of previous studies. Thus, the question and focus of this study is why have some studies concluded MRA is better while others have concluded that standard backpropagation ANN is better for predicting sold property value.

Implementation Issues

Some studies have demonstrated the superiority of ANN over MRA in predicting housing values (Tsukuda and Baba, 1990; Do and Grudnitski, 1992; Tay and Ho, 1991/1992; and Huang, Dorsey and Boose, 1994;). Other studies (Allen and Zumwalt, 1994; and Worzala, Lenk and Silva, 1995), however, have noted that ANNs are not necessarily superior. Because of an ANN's ability to learn and to recognize complicated patterns without being programmed with preconceived rules, it can easily be applied with little knowledge (statistically) of the data set. Unlike regression analysis, an ANN does not need a predetermined functional form based on the determinants. This feature of an ANN is important, since several studies (Grether and Mieskowski, 1974; Jones, Ferri and McGee, 1981; and Do and Grudnitski, 1993) found age has a nonlinear relationship with housing value (for the data set used in their study). Other studies have found that in addition to age, the square footage of living area has a nonlinear relationship with housing value (Goodman and Thibodeau, 1995). Because of the findings of previous studies and the theoretical strength of an ANN, one would anticipate ANN's performance to be better than that of MRA.

When using MRA, the methodological problems of functional form misspecification, nonlinearity, multicollinearity and heteroskedasticity should be addressed. When faced with possible nonlinear functional form, often one can convert a nonlinear relationship to a linear one before applying regression analysis (Kmenta, 1971:454). As noted, some studies have found that age and square footage of living area have a nonlinear relationship with housing value. Multicollinearity does not affect the predictive ability of MRA or that of ANN (Neter, Wasserman and Kutner, 1990:300) because the inferences are made within the jointly defined region of the observations. Multicollinearity, however, does make it infeasible to disentangle the effects of the supposedly independent variables, but determining the effects of each variable on selling price is not relevant to this article. Heteroskedasticity is normally present when cross-sectional data is used as in this study. In addition to the model methodological problems, leaving out a relevant explanatory variable is another source of error when using MRA and ANN. This is often due to the unavailability of data.

When using a feed forward ANN with back propagation learning, methodological problems such as number of hidden layers, number of neurons in each hidden layer, selection of training set, size of training set, selection of validation set, size of validation set and overtraining must be addressed. Generally, the level of training and number of hidden neurons affect the memorization and generalized predictability of the model. The more extensive the training and the more hidden neurons used, then the better the model is able to produce the correct results for the training set (memorization). On the other hand, the ANN is less likely to accurately predict new data (generalization), *i.e.*, the ability of the ANN to generalize is weakened when over training occurs. To avoid over training, a heuristic method as described in Hecht-Nielsen (1990:117) may be used. Although limited, there is some theoretical basis to assist one in determining the number of hidden layers and neurons to use. For further discussion, see Bishop (1995:121–26). In most situations, there is no way to determine the best number of hidden units without training several networks and estimating the generalization error of each. If the ANN has too few hidden units, then the training error and generalization error will be high due to underfitting and high statistical bias. If the ANN has too many hidden units, then the training error will be low, but the generalization error will be high due to overfitting and high variance (for more details, see Geman, Bienenstock and Doursat, 1992). If the training set is not representative of the data set (statistically), then there is no basis from which the ANN can learn. Normally, a representative training set is formed by using a random sample of the data set. If the training set is too small, then the ANN will tend to memorize the training patterns too specifically and extreme data points (noise) will have an inordinate influence on the model. This can be corrected, however, by using a k-fold cross validation training method (Stone, 1977; and Goutte, 1997).

Previous Applicable Studies

Do and Grudnitski (1992) used both MRA and ANN to predict residential housing value based on the eight inputs: age in years, number of bedrooms, number of bathrooms, square footage of living area, number of garage stalls, number of fireplaces, number of stories and lot size. Their MRA model is $SP_i = f(S_{ij})$ where SP_i is the property i selling price and S_{ij} is the set of explanatory housing price variables for property i. Their ANN model consists of an input layer of eight neurons (corresponding to the eight explanatory variables of the MRA), a hidden layer of three neurons and an output layer of one neuron representing the estimated value of the property. Their results show that the ANN is nearly twice as accurate as the MRA model in estimating residential property values (an ANN mean absolute error of 6.9% as compared to the MRA mean absolute error of 11.26%).

Relative to the ANN, the findings of Do and Grudnitski are reliable only if specialized training methods were used. Their sample size is very small (58 total and from just one neighborhood) while the ANN (fully connected) uses 27 weights. Without specialized training, the generalized predictability of the model is severely limited and questionable (Baum and Haussler, 1989). A special training method such as k-fold cross validation training is often used for small sample sizes (Stone, 1977; and Goutte, 1997).

Tay and Ho (1991/1992) used a data set of residential apartment properties in Singapore to test the predictive performance of ANN and MRA. They found the ANN model to have a mean absolute error of 3.9% while that of the MRA model to be 7.5%. However, their study by did not use specialized training, which is normally necessitated by a small sample size.

The above-mentioned studies generally support the superiority of ANN over MRA in predictive ability. There are also studies supporting the superiority of the MRA over ANN while other studies show inconclusive results. Worzala, Lenk and Silva (1995) used the same methodology as that of Borst (1992) and Do and Grudnitski (1992) on a different data set. The study conducted by Worzala et al. did not show the ANN superior to the MRA.

Relative to data model specification for housing value, several studies have identified various MRA functional specification improvements for predicting housing value, whereas studies using ANNs typically have not addressed the issue of inputs (functional specification). In what follows, various studies offering improvements to the MRA data model for predicting housing value are identified. In return, these are used (in experiments) to show that they also improve the ANN data model. Typical MRA functional forms used in other studies have been linear, semi-log and log-log (Goodman, 1978; Palmquist, 1979; Linneman, 1980; Halvorsen and Pollakowski, 1981; and Goodman and Thibodeau, 1995). Palmquist actually compared linear, semi-log, and log-linear, inverse semi-log and Box-Cox transformations to correct house price depreciation and selected the semi-log for usage in his study. In regard to nonlinearity, other studies have shown a nonlinear relationship between housing value and age (Grether and Mieszkowski, 1974; Jones, Ferri and McGee, 1981; and Do and Grudnitski, 1993) and the square footage of living area (Goodman and Thibodeau, 1995). Generally, these investigations used quadratic, cubic and quartic values for age and/or square footage.

Data

For this study, a total of 3,906 observations of sold single family residential properties from the Multiple Listing Service for the Rutherford County, Tennessee market were collected for the eighteen-month period from January 1, 1993 through June 30, 1994. The property attributes available and used for this study are: the square feet of living area (*sqft*), the number of bedrooms (*bed#*), the number of baths (*bath#*), the number of years since the property was built (*age*), the quarter the property sold (*quarter#*) and whether or not the property has a garage or carport (*garage_cp*).

From the 3,906 observations, random sample sizes of 306, 506, 706, 906, 1,106, 1,306, 1,506, 1,706, 1,906, 2,106, 2,306, 2,506, 2,706, 2,906, 3,106, 3,306, 3,506 and 3,706 are selected. These sets are referred to as training set one through eighteen (T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, T12, T13, T14, T15, T16, T17, and T18), respectively. Each training set is an extension of the previous, *i.e.*, $T1 \subset T2 \subset T3 \subset \cdots \subset T18$. The compliment of each training set, relative to the data set of 3,906 observations, gives a correspond validation set, *i.e.*, V1 through V18, respectively. Exhibit 1 provides some basic statistical information for the data set. Each of the training sets T1 through T18 contains a uniform distribution of the quarters of the eighteen-month period (quarter one through quarter six). These sample sets are used to discount any biases associated with each data model related to sample size. For both the MRA and ANN models, the training samples used for each data model are identical, as are the validation samples. The validation set V1 is used with training set T1; validation set V2 is used with training set T2; validation set V3 is used with training set T3, etc.

Description of MRA and ANN Models

To contrast the predictive performance of the MRA and ANN models, 108 comparisons were made in which the training size, the functional specification and the temporal prediction were varied. The MRA and ANN models (used in this study) address the issues noted earlier. These include functional form misspecification, nonlinearity and heteroskedasticity for MRA models. For ANN models, these include number of hidden layers, number of neurons in each hidden layer, selection of training set, size of training set, selection of validation set, size of validation and over training.

As noted, previous studies have provided several improvements to the MRA data model when predicting selling price. ANN improvements have been used to predict selling price. In what follows, six MRA functional specifications (based on previous studies) are used to design corresponding ANN models. This process is used to show that improvements to the MRA model also improve the ANN model (see Exhibit 2). The resulting data models are compared using the eighteen training sets and corresponding eighteen validation sets. The MRA functional forms used in this study are linear, semi-log and log-log.

Whenever cross-sectional data is used as in this study, one does not expect the same variance of residual errors across the sample, and hence heteroskedasticity is possible. To test for heteroskedasticity, the Goldfield and Quandt test (Goldfield and Quandt, 1965) is performed. Heteroskedasticity in age was found for each training set. After correcting for heteroskedasticity, the improved quantified relationship for each training set is shown in Exhibit 3 (T1, T9 and T18 for MRA) and Exhibit 4 (for ANN).

Exhibit 1 | Basic Statistical Information

Selling Price	
Minimal	120,000
Maximum	385,000
Mean	87,644
Median	79,500
Std. Dev.	3,814
Square Footage	
Minimal	250
Maximum	9,254
Mean	1,652
Median	1,450
Std. Dev.	604
Bedrooms	
Minimal	1
Maximum	12
Mean	3.1
Median	3
Std. Dev.	0.60
Bathrooms	
Minimal	1
Maximum	6
Mean	2.2
Median	2
Std. Dev.	0.62
Age	
Minimal	1
Maximum	125
Mean	9.9
Median	6
Std. Dev.	12.4

Note: The randomly selected training and validation subsets were representative of this larger 3,906 parent population.

Six model specifications are given for both the MRA and the ANN. The model specifications offer a series of improvements and a comparison of the data models under various conditions. The MRA data models are applied to the corresponding validation data sets V1 through V18. In a similar manner, the ANNs specified in Exhibit 4 are trained individually on the data sets T1 through T18 and applied to the corresponding validation data set V1 through V18. The comparisons are made between each data model's performance on the validation data sets (see Exhibit 2).

Exhibit 2 | The Comparative Forecasting Performance of MRA and ANN

Data Model	Train Set (quantity)	Validation Set (quantity)	MAPE (%)	FE 5% (%)	FE 15% (%)	Over 15% (%)	Superior Model* MAPE FE
MRA M1	T1 (306)	V1 (3600)	14.1	25.2	48.0	26.8	MRA
ANN M1	T1	V1	14.9	26.8	41.2	32.1	MIXED
MRA M1	T2 (506)	V2 (3400)	16.8	28.7	36.8	34.6	ANN
ANN M1	T2	V2	13.9	31.8	41.6	26.6	ANN
MRA M1	T3(706)	V3(3200)	16.7	29.5	37.1	33.4	ANN
ANN M1	T3	V3	14.5	32.1	40.6	27.3	ANN
MRA M1	T4(906)	V4(3000)	17.7	30.0	37.3	32.7	ANN
ANN M1	T4	V4	11.7	36.6	39.8	23.7	ANN
MRA M1	T5(1106)	V5(2800)	17.1	30.6	38.2	31.3	ANN
ANN M1	T5	V5	12.3	35.0	40.5	24.5	ANN
MRA M1	T6(1306)	V6(2600)	16.8	32.2	37.4	30.4	ANN
ANN M1	T6	V6	11.2	36.4	42.1	21.5	ANN
MRA M1	T7(1506)	V7(2400)	17.2	31.1	37.8	31.1	ANN
ANN M1	T7	V7	11.3	34.5	44.1	21.4	ANN
MRA M1	T8(1706)	V8(2200)	18.0	32.4	36.3	32.3	ANN
ANN M1	T8	V8	10.0	38.2	43.7	18.0	ANN
MRA M1	T9(1906)	V9(2000)	18.2	31.1	35.9	32.9	ANN
ANN M1	T9	V9	10.0	39.0	42.6	18.4	ANN
MRA M1	T10(2106)	V10(1800)	17.8	30.4	36.9	32.6	ANN
ANN M1	T10	V10	9.3	41.0	42.6	16.3	ANN
MRA M1	T11(2306)	V11(1600)	17.3	30.8	37.2	32.0	ANN
ANN M1	T11	V11	9.9	40.2	42.2	17.6	ANN
MRA M1	T12(2506)	V12(1400)	18.3	30.6	37.1	32.3	ANN
ANN M1	T12	V12	9.1	43.5	40.8	15.7	ANN
MRA M1	T13(2706)	V13(1200)	17.2	30.7	37.0	32.3	ANN
ANN M1	T13	V13	8.6	44.2	41.6	14.2	ANN
MRA M1	T14(2906)	V14(1000)	16.9	31.1	37.2	31.6	ANN
ANN M1	T14	V14	8.6	44.1	42.2	13.7	ANN
MRA M1	T15(3106)	V15(800)	16.9	31.0	37.2	31.8	ANN
ANN M1	T15	V15	8.4	46.7	39.6	13.6	ANN
MRA M1	T16(3306)	V16(600)	16.9	30.7	37.3	32.0	ANN
ANN M1	T16	V16	8.7	46.9	39.1	14.0	ANN
MRA M1	T17(3506)	V17(400)	16.8	30.9	36.9	32.2	ANN
ANN M1	T17	V17	8.5	48.9	36.3	14.8	ANN

Exhibit 2 | (continued)

The Comparative Forecasting Performance of MRA and ANN

Data Model	Train Set (quantity)	Validation Set (quantity)	MAPE (%)	FE 5% (%)	FE 15% (%)	Over 15% (%)	Superior Model* MAPE FE
MRA M1	T18(3706)	V18(200)	16.6	30.8	37.6	31.6	ANN
ANN M1	T18	V18	7.5	55.4	33.6	11.0	ANN
MRA M2	T1 (306)	V1 (3600)	13.2	25.2	50.0	24.8	MRA
ANN M2	T1	V1	15.5	27.6	38.9	33.5	MIXED
MRA M2	T2 (506)	V2 (3400)	12.0	30.8	44.3	24.9	MRA
ANN M2	T2	V2	14.2	32.2	41.0	26.8	MIXED
MRA M2	T3(706)	V3(3200)	12.9	29.2	45.3	25.5	MRA
ANN M2	T3	V3	14.6	31.9	40.6	27.5	MIXED
MRA M2	T4(906)	V4(3000)	14.3	29.9	44.5	25.6	ANN
ANN M2	T4	V4	11.3	36.9	39.9	23.2	ANN
MRA M2	T5(1106)	V5(2800)	13.9	30.4	43.2	26.4	ANN
ANN M2	T5	V5	11.4	36.7	42.2	21.1	ANN
MRA M2	T6(1306)	V6(2600)	13.5	32.3	41.2	26.5	ANN
ANN M2	T6	V6	11.0	37.0	41.6	21.4	ANN
MRA M2	T7(1506)	V7(2400)	13.9	32.3	40.4	27.3	ANN
ANN M2	T7	V7	11.5	34.4	42.7	22.9	ANN
MRA M2	T8(1706)	V8(2200)	14.9	30.9	40.3	28.7	ANN
ANN M2	T8	V8	10.0	38.3	43.9	17.8	ANN
MRA M2	T9(1906)	V9(2000)	15.0	30.2	40.9	28.9	ANN
ANN M2	T9	V9	9.9	39.3	42.2	18.5	ANN
MRA M2	T10(2106)	V10(1800)	14.9	31.2	40.1	28.7	ANN
ANN M2	T10	V10	9.2	41.3	42.7	16.0	ANN
MRA M2	T11(2306)	V11(1600)	14.7	31.3	39.9	28.8	ANN
ANN M2	T11	V11	9.7	42.0	40.5	17.6	ANN
MRA M2	T12(2506)	V12(1400)	14.9	30.8	40.0	29.2	ANN
ANN M2	T12	V12	8.8	42.8	41.9	15.3	ANN
MRA M2	T13(2706)	V13(1200)	14.9	31.0	39.7	29.3	ANN
ANN M2	T13	V13	8.4	44.3	42.6	13.1	ANN
MRA M2	T14(2906)	V14(1000)	14.6	31.3	39.8	28.9	ANN
ANN M2	T14	V14	8.6	44.6	42.4	13.0	ANN
MRA M2	T15(3106)	V15(800)	14.5	31.3	40.0	28.7	ANN
ANN M2	T15	V15	8.2	47.1	40.3	12.6	ANN
MRA M2	T16(3306)	V16(600)	14.5	31.0	39.9	29.1	ANN

Exhibit 2 | (continued)

The Comparative Forecasting Performance of MRA and ANN

Data Model	Train Set (quantity)	Validation Set (quantity)	MAPE (%)	FE 5% (%)	FE 15% (%)	Over 15% (%)	Superior Model* MAPE FE
ANN M2	T16	V16	8.7	47.2	38.1	14.7	ANN
MRA M2	T17(3506)	V17(400)	14.6	30.8	39.7	29.5	ANN
ANN M2	T17	V17	8.3	49.2	38.5	12.3	ANN
MRA M2	T18(3706)	V18(200)	14.4	31.3	39.8	29.0	ANN
ANN M2	T18	V18	7.2	55.9	33.1	11.0	ANN
MRA M3	T1 (306)	V1 (3600)	13.2	26.8	48.0	25.2	MRA
ANN M3	T1	V1	15.2	27.3	40.6	32.0	MIXED
MRA M3	T2 (506)	V2 (3400)	12.1	30.6	44.5	24.9	MRA
ANN M3	T2	V2	13.9	32.2	42.4	25.3	MIXED
MRA M3	T3(706)	V3(3200)	12.7	29.0	45.5	25.5	MRA
ANN M3	T3	V3	14.4	31.3	41.6	27.1	MIXED
MRA M3	T4(906)	V4(3000)	13.9	29.5	46.1	24.4	ANN
ANN M3	T4	V4	11.2	36.0	42.0	22.0	ANN
MRA M3	T5(1106)	V5(2800)	13.5	30.3	44.6	25.1	ANN
ANN M3	T5	V5	11.6	36.0	41.6	22.3	ANN
MRA M3	T6(1306)	V6(2600)	13.0	32.5	42.0	25.6	ANN
ANN M3	T6	V6	11.0	36.5	42.2	21.3	ANN
MRA M3	T7(1506)	V7(2400)	13.4	33.1	40.4	26.5	ANN
ANN M3	T7	V7	11.4	34.5	43.3	22.1	ANN
MRA M3	T8(1706)	V8(2200)	14.5	31.3	40.3	28.4	ANN
ANN M3	T8	V8	10.1	38.3	43.0	18.6	ANN
MRA M3	T9(1906)	V9(2000)	14.4	31.6	40.0	28.4	ANN
ANN M3	T9	V9	9.8	38.5	44.1	17.5	ANN
MRA M3	T10(2106)	V10(1800)	14.2	32.1	39.5	28.4	ANN
ANN M3	T10	V10	9.3	41.8	41.9	16.3	ANN
MRA M3	T11(2306)	V11(1600)	14.0	31.7	40.5	27.8	ANN
ANN M3	T11	V11	9.4	41.7	42.4	15.9	ANN
MRA M3	T12(2506)	V12(1400)	14.1	30.2	41.8	28.0	ANN
ANN M3	T12	V12	8.8	43.6	41.8	14.6	ANN
MRA M3	T13(2706)	V13(1200)	14.1	30.8	40.6	28.6	ANN
ANN M3	T13	V13	8.4	44.6	42.6	12.8	ANN
MRA M3	T14(2906)	V14(1000)	13.8	31.3	40.6	28.1	ANN
ANN M3	T14	V14	8.6	44.6	42.1	13.3	ANN

Exhibit 2 | (continued)

The Comparative Forecasting Performance of MRA and ANN

Data Model	Train Set (quantity)	Validation Set (quantity)	MAPE (%)	FE 5% (%)	FE 15% (%)	Over 15% (%)	Superior Model* MAPE FE
MRA M3	T15(3106)	V15(800)	13.6	31.8	40.5	27.7	ANN
ANN M3	T15	V15	8.3	46.3	41.0	12.6	ANN
MRA M3	T16(3306)	V16(600)	13.6	31.5	40.6	27.9	ANN
ANN M3	T16	V16	8.5	46.6	39.7	13.7	ANN
MRA M3	T17(3506)	V17(400)	13.6	31.3	40.6	28.2	ANN
ANN M3	T17	V17	8.5	47.6	38.4	14.0	ANN
MRA M3	T18(3706)	V18(200)	13.4	31.5	40.7	27.7	ANN
ANN M3	T18	V18	7.2	55.4	33.6	11.0	ANN
MRA M4	T1 (306)	V1 (3600)	11.7	35.3	42.2	22.5	MRA
ANN M4	T1	V1	15.5	28.7	38.8	32.5	MRA
MRA M4	T2 (506)	V2 (3400)	10.9	37.4	40.5	22.1	MRA
ANN M4	T2	V2	14.3	32.6	40.9	26.5	MRA
MRA M4	T3(706)	V3(3200)	11.7	32.9	43.9	23.2	MRA
ANN M4	T3	V3	14.3	30.7	41.6	27.8	MRA
MRA M4	T4(906)	V4(3000)	12.9	34.3	45.0	20.6	ANN
ANN M4	T4	V4	11.0	35.6	43.6	20.8	MIXED
MRA M4	T5(1106)	V5(2800)	12.4	34.3	45.2	20.5	ANN
ANN M4	T5	V5	11.2	35.9	42.8	21.3	MIXED
MRA M4	T6(1306)	V6(2600)	12.0	34.5	45.6	19.9	ANN
ANN M4	T6	V6	10.9	37.3	42.0	20.6	MIXED
MRA M4	T7(1506)	V7(2400)	11.9	35.5	44.4	20.1	ANN
ANN M4	T7	V7	11.3	35.2	43.0	21.8	MRA
MRA M4	T8(1706)	V8(2200)	12.3	33.2	45.8	21.0	ANN
ANN M4	T8	V8	10.2	38.4	42.9	18.7	ANN
MRA M4	T9(1906)	V9(2000)	12.2	33.1	46.1	20.8	ANN
ANN M4	T9	V9	9.8	39.9	42.4	17.7	ANN
MRA M4	T10(2106)	V10(1800)	12.1	33.4	45.9	20.7	ANN
ANN M4	T10	V10	9.1	42.2	42.6	15.3	ANN
MRA M4	T11(2306)	V11(1600)	12.1	32.8	46.2	21.0	ANN
ANN M4	T11	V11	9.9	41.6	41.0	17.4	ANN
MRA M4	T12(2506)	V12(1400)	12.1	33.0	46.2	20.8	ANN
ANN M4	T12	V12	8.8	43.2	41.2	15.6	ANN
MRA M4	T13(2706)	V13(1200)	12.2	33.0	45.8	21.2	ANN

Exhibit 2 | (continued)

The Comparative Forecasting Performance of MRA and ANN

Data Model	Train Set (quantity)	Validation Set (quantity)	MAPE (%)	FE 5% (%)	FE 15% (%)	Over 15% (%)	Superior Model* MAPE FE
ANN M4	T13	V13	8.5	45.1	40.4	14.4	ANN
MRA M4	T14(2906)	V14(1000)	12.0	33.4	45.4	21.2	ANN
ANN M4	T14	V14	8.6	45.0	41.9	13.1	ANN
MRA M4	T15(3106)	V15(800)	12.0	33.1	45.6	21.3	ANN
MRA M4	T17(3506)	V17(400)	11.8	33.2	45.5	21.3	ANN
ANN M4	T17	V17	8.6	49.2	36.0	14.8	ANN
MRA M4	T18(3706)	V18(200)	11.8	34.1	44.8	21.1	ANN
ANN M4	T18	V18	7.3	56.9	32.6	10.5	ANN
MRA M5	T1 (306)	V1 (3600)	12.4	26.5	47.4	21.1	MRA
ANN M5	T1	V1	13.4	31.9	41.7	26.4	MIXED
MRA M5	T2 (506)	V2 (3400)	11.9	29.2	45.5	25.3	MRA
ANN M5	T2	V2	13.5	33.2	41.3	25.5	MIXED
MRA M5	T3(706)	V3(3200)	12.4	29.6	44.9	25.5	MRA
ANN M5	T3	V3	13.1	33.1	41.9	25.0	ANN
MRA M5	T4(906)	V4(3000)	13.6	30.0	45.9	24.1	ANN
ANN M5	T4	V4	10.9	36.5	42.0	21.5	ANN
MRA M5	T5(1106)	V5(2800)	13.1	29.5	46.0	24.5	ANN
ANN M5	T5	V5	11.0	37.4	41.4	21.2	ANN
MRA M5	T6(1306)	V6(2600)	12.7	30.0	45.9	24.1	ANN
ANN M5	T6	V6	10.7	37.8	42.1	20.1	ANN
MRA M5	T7(1506)	V7(2400)	12.8	28.6	46.5	24.9	ANN
ANN M5	T7	V7	11.1	36.0	42.4	21.3	ANN
MRA M5	T8(1706)	V8(2200)	13.1	28.3	46.4	25.3	ANN
ANN M5	T8	V8	9.7	38.9	42.9	18.2	ANN
MRA M5	T9(1906)	V9(2000)	12.9	29.1	46.0	24.9	ANN
ANN M5	T9	V9	9.7	40.7	41.8	17.5	ANN
MRA M5	T10(2106)	V10(1800)	12.9	28.7	46.6	24.7	ANN
ANN M5	T10	V10	9.0	42.6	41.4	16.0	ANN
MRA M5	T11(2306)	V11(1600)	12.8	28.8	46.8	24.5	ANN
ANN M5	T11	V11	9.5	40.7	42.3	17.0	ANN
MRA M5	T12(2506)	V12(1400)	12.8	29.2	46.8	24.0	ANN
ANN M5	T12	V12	9.0	43.6	40.8	15.6	ANN
MRA M5	T13(2706)	V13(1200)	12.9	29.0	46.9	24.1	ANN

Exhibit 2 | (continued)

The Comparative Forecasting Performance of MRA and ANN

Data Model	Train Set (quantity)	Validation Set (quantity)	MAPE (%)	FE 5% (%)	FE 15% (%)	Over 15% (%)	Superior Model* MAPE FE
ANN M5	T13	V13	8.3	45.2	41.1	13.7	ANN
MRA M5	T14(2906)	V14(1000)	12.7	28.8	47.2	24.0	ANN
ANN M5	T14	V14	8.5	43.5	42.4	14.1	ANN
MRA M5	T15(3106)	V15(800)	12.7	29.0	46.6	24.4	ANN
ANN M5	T15	V15	8.4	46.3	40.3	13.4	ANN
MRA M5	T16(3306)	V16(600)	12.6	29.0	46.8	24.2	ANN
ANN M5	T16	V16	8.5	45.8	41.7	12.5	ANN
MRA M5	T17(3506)	V17(400)	12.6	29.3	46.7	24.1	ANN
ANN M5	T17	V17	8.4	48.4	38.5	13.0	ANN
MRA M5	T18(3706)	V18(200)	12.5	29.5	46.7	23.9	ANN
ANN M5	T18	V18	7.1	56.9	34.6	8.5	ANN
MRA M6	T1 (306)	V1 (3600)	12.3	34.3	39.5	26.1	MRA
ANN M6	T1	V1	13.4	35.3	40.2	24.5	ANN
MRA M6	T2 (506)	V2 (3400)	11.3	36.8	39.5	23.7	MRA
ANN M6	T2	V2	15.6	32.1	37.3	30.6	MRA
MRA M6	T3(706)	V3(3200)	12.0	32.6	42.2	25.2	MRA
ANN M6	T3	V3	14.3	34.6	37.3	28.0	MIXED
MRA M6	T4(906)	V4(3000)	13.2	34.7	42.9	22.4	ANN
ANN M6	T4	V4	11.9	36.3	38.6	25.0	MIXED
MRA M6	T5(1106)	V5(2800)	12.8	33.4	44.4	22.2	ANN
ANN M6	T5	V5	11.5	37.1	39.8	23.2	MIXED
MRA M6	T6(1306)	V6(2600)	12.3	33.1	45.2	21.7	ANN
ANN M6	T6	V6	10.9	38.4	41.4	20.2	ANN
MRA M6	T7(1506)	V7(2400)	12.2	33.8	44.3	21.9	ANN
ANN M6	T7	V7	11.4	34.4	43.1	22.5	ANN
MRA M6	T8(1706)	V8(2200)	12.6	32.1	45.4	22.5	ANN
ANN M6	T8	V8	10.0	40.5	41.7	17.8	ANN
MRA M6	T9(1906)	V9(2000)	12.5	31.8	45.7	22.5	ANN
ANN M6	T9	V9	9.9	40.6	42.3	17.1	ANN
MRA M6	T10(2106)	V10(1800)	12.4	31.8	46.1	22.1	ANN
ANN M6	T10	V10	9.5	42.0	41.0	17.0	ANN
MRA M6	T11(2306)	V11(1600)	12.4	32.8	44.6	22.5	ANN
ANN M6	T11	V11	10.2	40.7	40.3	19.0	ANN

Exhibit 2 | (continued)

The Comparative Forecasting Performance of MRA and ANN

Data Model	Train Set (quantity)	Validation Set (quantity)	MAPE (%)	FE 5% (%)	FE 15% (%)	Over 15% (%)	Superior Model* MAPE FE
MRA M6	T12(2506)	V12(1400)	12.4	32.4	45.2	22.3	ANN
ANN M6	T12	V12	9.1	43.6	39.4	17.0	ANN
MRA M6	T13(2706)	V13(1200)	12.5	32.2	45.0	22.8	ANN
ANN M6	T13	V13	9.0	44.6	38.8	16.5	ANN
MRA M6	T14(2906)	V14(1000)	12.3	32.3	44.7	23.0	ANN
ANN M6	T14	V14	8.7	46.0	38.8	15.2	ANN
MRA M6	T15(3106)	V15(800)	12.3	33.2	44.3	22.5	ANN
ANN M6	T15	V15	8.8	45.8	38.3	15.9	ANN
MRA M6	T16(3306)	V16(600)	12.2	33.0	44.6	22.5	ANN
ANN M6	T16	V16	9.3	43.0	39.5	17.5	ANN
MRA M6	T17(3506)	V17(400)	12.1	33.7	44.2	22.1	ANN
ANN M6	T17	V17	8.0	46.9	38.0	15.0	ANN
MRA M6	T18(3706)	V18(200)	12.1	34.3	43.6	22.1	ANN
ANN M6	T18	V18	7.6	53.4	36.6	10.0	ANN

Note: *For each training set, the first entry in this column represents the superior model when evaluated using MAPE criterion while the second entry represents the superior model using the FE criterion.

MRA Model Specification

Based on previous studies, the MRA models (1 through 6) are created and used for comparison. Model 1 specifications are based on studies that show that *sqft*, *age*, *bed#* and *bath#* influence selling price (see, for example, Newsome and Zietz, 1992; Nguyen and Rogers, 1992; and Do and Grudnitski, 1993). Other factors, such as a garage, fireplace, number of stories, sewer connection to city and lot size, are relevant factors. However, due to the lack of data, only information as to whether the property has a garage and/or carport is used (as a dummy variable). The measurement units for *sqft* and *age* are *sqft*/100 and *age*/10, respectively. These linear changes provide more manageable data values when using quadratic, cubic and quartic values for age and/or square footage. The five-quarter dummy variables in models 1 through 5 are included to reflect the seasonality and to incorporate the price shift from year 93 to year 94.

Exhibit 3 | Regression Models for Training Sets T1, T2 and T18

	Training Set T1		Training Set T9		Training Set T18	
Variable	Coeff.	*t*-Stat.	Coeff.	*t*-Stat.	Coeff.	*t*-Stat.
Panel A: MRA Model 1 (linear)						
Intercept	777.62	−1.75	6717.79	−7.68	6672.23	−10.80
Living Area/100	5080.55	24.06	5739.49	61.90	5789.68	92.07
Age/10	−6654.90	0.11	−13408.42	2.73	−12752.49	4.05
Bed#	−4869.18	−1.97	−5095.06	−6.04	−4245.54	−7.41
Bath#	9375.99	4.78	5004.69	6.00	2947.16	5.29
Garage_cp	9673.66	4.09	8471.80	7.56	8565.13	11.48
Quarter1	9829.80	−3.89	−9516.18	−8.61	−9007.06	−12.00
Quarter2	−7845.90	−3.09	−8446.08	−8.15	−7065.44	−9.92
Quarter3	1827.51	0.78	−6121.61	−6.16	−4724.49	−6.86
Quarter4	−7392.69	−3.19	−5667.74	−6.00	−4500.95	−6.87
Quarter5	1972.77	0.81	−3237.87	−3.07	−2777.41	−3.85
Number of Observations	306		1906		3706	
Adj. R^2	0.97		0.96		0.97	
Panel B: MRA Model 2 (non-linear)						
Intercept	30107.84	−0.80	30190.17	−6.39	27379.11	−9.33
Living Area/100	1782.20	2.71	3251.78	14.06	3676.46	21.58
$(Living\ Area/100)^2$	73.10	5.28	55.52	11.69	49.04	13.33
Age/10	−5063.70	3.49	−18559.74	9.88	−18052.76	12.51
$(Age/10)^2$	−137.78	−0.06	2910.05	2.60	2670.60	3.68
Bed#	−4939.36	−2.08	−5307.68	−6.52	−4593.16	−8.21
Bath#	9280.76	4.92	4683.33	5.83	2665.05	4.91
Garage_cp	12252.39	5.28	10985.44	9.93	10298.21	13.86
Quarter1	−10493.23	−4.33	−9081.70	−8.52	−8618.20	−11.77
Quarter2	−5937.05	−2.41	−7532.22	−7.52	−6451.62	−9.27
Quarter3	1504.32	0.67	−5286.13	−5.50	−4147.19	−6.17
Quarter4	−7032.90	−3.17	−4818.86	−5.28	−4109.85	−6.43
Quarter5	2407.61	1.03	−2384.06	−2.34	−2176.57	−3.08
Number of Observations	306		1906		3706	
Adj. R^2	0.97		0.97		0.97	
Panel C: MRA Model 3 (non-linear)						
Intercept	28242.10	0.54	32039.47	−4.78	28975.10	−6.71
Living Area/100	1740.73	2.64	3285.78	14.19	3698.76	21.71
$(Living\ Area/100)^2$	73.70	5.30	54.84	11.53	48.62	13.22
Age/10	8575.37	3.18	−30896.83	10.10	−28459.25	12.82
$(Age/10)^2$	−15506.77	−0.91	15288.50	2.44	12791.91	3.20
$(Age/10)^3$	4257.64	0.87	−2858.48	−1.79	−2228.02	−2.23
$(Age/10)^4$	−290.86	−0.80	146.30	1.49	111.64	1.78
Bed#	−5032.37	−2.12	−5287.16	−6.50	−4581.24	−8.20
Bath#	9446.39	4.98	4637.52	5.77	2648.82	4.88
Garage_cp	12264.18	5.28	10889.55	9.84	10226.11	13.77
Quarter1	−10284.49	−4.21	−9103.40	−8.55	−8638.01	−11.80

Exhibit 3 | (continued)

Regression Models for Training Sets T1, T2 and T18

Variable	Training Set T1 Coeff.	Training Set T1 *t*-Stat.	Training Set T9 Coeff.	Training Set T9 *t*-Stat.	Training Set T18 Coeff.	Training Set T18 *t*-Stat.
Panel C: MRA Model 3 (non-linear)						
Quarter2	−5918.15	−2.40	−7508.79	−7.50	−6435.74	−9.25
Quarter3	1577.67	0.70	−5298.57	−5.52	−4138.94	−6.16
Quarter4	−7018.37	−3.16	−4819.19	−5.28	−4108.78	−6.43
Quarter5	2378.50	1.02	−2356.43	−2.31	−2161.36	−3.07
Number of Observations	306		1906		3706	
Adj. R^2	0.97		0.97		0.97	
Panel D: MRA Model 5 (log-log)						
Intercept	−0.12	86.97	−0.11	214.89	−0.11	305.64
Log(*Living Area*)	0.86	17.90	0.82	41.47	0.86	62.68
Log(*Age*)	3.87	−7.38	3.84	−17.42	3.82	−25.16
Log(*Bed#*)	−0.14	−1.93	−0.002	−0.09	−0.03	−1.61
Log(*Bath#*)	0.07	1.40	0.07	3.12	0.05	3.54
Garage_cp	0.06	5.59	0.07	15.37	0.06	19.97
Quarter1	−0.01	−3.13	−0.01	−5.19	−0.01	−7.97
Quarter2	−0.01	−2.38	−0.01	−4.69	−0.01	−6.82
Quarter3	−0.004	−0.99	−0.01	−3.96	−0.01	−4.55
Quarter4	−0.01	−2.26	−0.005	−2.85	−0.005	−4.10
Quarter5	−0.003	−0.71	−0.003	−1.82	−0.004	−2.91
Number of Observations	306		1906		3706	
Adj. R^2	0.77		0.77		0.78	
Panel E: MRA Model 4 (semi-log)						
Intercept	4.94	0.63	4.78	−4.39	4.67	−6.31
Living Area / 100	0.04	13.86	0.03	35.53	0.03	48.34
(*Living Area / 100*)^2	−0.0004	−6.88	−0.0003	−15.65	−0.0003	−19.78
Age / 10	0.04	49.64	−0.12	153.45	−0.11	240.94
(*Age / 10*)^2	−0.07	−1.12	0.04	1.69	0.04	2.39
(*Age / 10*)^3	0.02	0.99	−0.01	−1.08	−0.01	−1.54
(*Age / 10*)^4	−0.001	−0.87	0.0004	0.87	0.0003	1.21
Bed#	−0.18	−5.15	−0.11	−10.81	−0.07	−11.63
Bath#	0.21	4.99	0.12	10.04	0.07	10.05
Bath# / Bed#	−0.58	−4.54	−0.34	−9.31	−0.21	−9.68
Garage_cp	0.07	7.93	0.07	14.18	0.06	19.62
Quarter1	−0.05	−5.42	−0.04	−8.79	−0.04	−11.90
Quarter2	−0.03	−3.12	−0.03	−7.82	−0.03	−9.64
Quarter3	−0.01	−0.86	−0.03	−6.31	−0.02	−6.50
Quarter4	−0.03	−3.50	−0.02	−4.68	−0.01	−5.56
Quarter5	0.01	0.87	−0.01	−2.63	−0.01	−3.43
Number of Observations	306		1906		3706	
Adj. R^2	0.99		0.99		0.99	

Exhibit 3 | (continued)

Regression Models for Training Sets T1, T2 and T18

Variable	Training Set T1 Coeff.	Training Set T1 *t*-Stat.	Training Set T9 Coeff.	Training Set T9 *t*-Stat.	Training Set T18 Coeff.	Training Set T18 *t*-Stat.
Panel F: MRA Model 6 (semi-log)						
Intercept	4.96	1.17	4.77	−4.22	4.67	−6.11
Living Area / 100	0.04	13.16	0.03	34.58	0.03	47.07
$(Living\ Area / 100)^2$	−0.0004	−6.56	−0.0003	−14.78	−0.0003	−18.80
Age / 10	0.07	46.57	−0.12	149.39	−0.11	235.41
$(Age / 10)^2$	−0.11	−1.62	0.04	1.58	0.04	2.27
$(Age / 10)^3$	0.03	1.45	−0.01	−1.00	−0.01	−1.45
$(Age / 10)^4$	−0.002	−1.30	0.0003	0.80	0.0003	1.14
Bed#	−0.19	−5.02	−0.11	−10.85	−0.07	−11.99
Bath#	0.21	4.81	0.12	9.88	0.07	10.06
Bath# / Bed#	−0.62	−4.55	−0.35	−9.35	−0.22	−9.92
Garage_cp	0.06	6.77	0.07	14.25	0.06	19.69
Number of Observations	306		1906		3706	
Adj. R^2	0.99		0.99		0.99	

Note: The majority of coefficients are statistically significant at the 99% confidence level.

The variable *bath#/bed#* is introduced in models 4 and 6 because of the significance shown by Newsome and Zietz (1992). The Newsome and Zietz study used data from the same market (Rutherford County) as this study. The use of all three variables—bed#, bath# and bath#/bed#—may introduce multicollinearity, but as previously discussed, this is irrelevant in the predictive performance of the model.

Exhibit 4 | ANN Nodels

Training Set T1 through T18 — Number of Neurons

Ann Model	Input Layer	Hidden Layer	Output Layer
1	11	7	1
2	13	12	1
3	15	12	1
4	16	12	1
5	11	6	1
6	10	8	1

Other studies have shown a nonlinear relationship between housing value and age (Grether and Mieszkowski, 1974; Jones, Ferri and McGee, 1981; and Do and Grudnitski 1993) and the square footage of living area (Goodman and Thibodeau, 1995). Generally, these investigations use quadratic, cubic and quartic values for age and/or square footage. Hence, the nonlinear values of *sqft* and *age* are used in models 2, 3, 4 and 6.

Model 1 MRA Specification:

$$
\begin{aligned}
price \equiv b_0 + b_1 * \frac{sqft}{100} + b_2 * \frac{age}{10} + b_3 * bed\# + b_4 * bath\# \\
+ b_5 * garage\_cp + \sum_{i=6}^{10} b_i * quarter_{i-5}.
\end{aligned}
\tag{1}
$$

Model 2 MRA Specification:

$$
\begin{aligned}
price \equiv b_0 + b_1 * \frac{sqft}{100} + b_2 * \left(\frac{sqft}{100}\right)^2 + b_3 * \frac{age}{10} \\
+ b_4 * \left(\frac{age}{10}\right)^2 + b_5 * bed\# + b_6 * bath\# \\
+ b_7 * garage\_cp + \sum_{i=8}^{12} b_i * quarter_{i-7}.
\end{aligned}
\tag{2}
$$

Model 3 MRA Specification:

$$
\begin{aligned}
price \equiv b_0 + b_1 * \frac{sqft}{100} + b_2 * \left(\frac{sqft}{100}\right)^2 + b_3 * \frac{age}{10} \\
+ b_4 * \left(\frac{age}{10}\right)^2 + b_5 * \left(\frac{age}{10}\right)^3 + b_6 * \left(\frac{age}{10}\right)^4 \\
+ b_7 * bed\# + b_8 * bath\# + b_9 * garage\_cp \\
+ \sum_{t=10}^{14} b_i * quarter_{i-9}.
\end{aligned}
\tag{3}
$$

Model 4 MRA Specification:

$$\begin{aligned}\log(price) \equiv\ & b_0 + b_1 * \frac{sqft}{100} + b_2 * \left(\frac{sqft}{100}\right)^2 + b_3 * \frac{age}{10} \\ & + b_4 * \left(\frac{age}{10}\right)^2 + b_5 * \left(\frac{age}{10}\right)^3 + b_6 * \left(\frac{age}{10}\right)^4 \\ & + b_7 * bed\# + b_8 * bath\# + b_9 * \frac{bath\#}{bed\#} \\ & + b_{10} * garage\_cp + \sum_{i=11}^{15} b_i * quarter_{i-10}. \end{aligned} \tag{4}$$

Model 5 MRA Specification:

$$\begin{aligned}\log(price) \equiv\ & b_0 + b_1 * \log(sqft) + b_2 * \log(age) \\ & + b_3 * \log(bed\#) + b_4 * \log(bath\#) \\ & + b_5 * garage\_cp + \sum_{i=6}^{11} b_i * quarter_{i-5}. \end{aligned} \tag{5}$$

Model 6 MRA Specification:

$$\begin{aligned}\log(price) \equiv\ & b_0 + b_1 * \frac{sqft}{100} + b_2 * \left(\frac{sqft}{100}\right)^2 + b_3 * \frac{age}{10} \\ & + b_4 * \left(\frac{age}{10}\right)^2 + b_5 * \left(\frac{age}{10}\right)^3 + b_6 * \left(\frac{age}{10}\right)^4 \\ & + b_7 * bed\# + b_8 * bath\# + b_9 * \frac{bath\#}{bed\#} \\ & + b_{10} * garage\_cp. \end{aligned} \tag{6}$$

Each of the MRA models are regressed on the training sets T1 through T18. The resulting quantified relationships (coefficients) for training sets T1, T9 and T18 are given in Exhibit 3.

ANN Model Specification

Feed forward ANNs with inputs corresponding to each MRA model specification,[1] one hidden layer and one output are used to predict selling price (see Exhibit 4). The linear transformation in Equation (7) is applied to each ANN input and the linear transformation in Equation (8) is applied to the selling price. These linear transformations are necessary since the ANN[2] software requires all inputs to be in the interval $(-1,1)$ and outputs to be in the interval $(0,1)$. The following linear transformation is used to translate the inputs:

$$\textit{translated\_value} \equiv \frac{\textit{data field value} - \textit{average data field value}}{\textit{maximum data field value} - \textit{minimum data field value}}. \tag{7}$$

For training and validation, the target output (selling price) is translated to the interval (0,1) by using the minimum selling price, the maximum selling price and the following linear transformation:

$$\textit{translated\_price} \equiv \frac{\textit{selling price} - \textit{minimum selling price}}{\textit{maximum selling price} - \textit{minimum selling price}}. \tag{8}$$

For ANN, there are "rules of thumb" for net size, number of patterns and error size (Baum and Haussler, 1989) and for the number of hidden layers and number of neurons per hidden layer. Based on Baum's results, one can approximate the number of patterns using the formula:

$$\textit{Number of patterns} \approx \frac{\textit{number of weights}}{\textit{accuracy of classification expected}}. \tag{9}$$

Theoretical results show that one hidden layer is sufficient for a backpropagation net to approximate any continuous mapping from the input data to the output data to an arbitrary degree of accuracy (Hecht-Nielsen, 1987). In general, two hidden layers may make training easier in some situations (Fausett, 1994), but the number of hidden layers and number of neurons per hidden layer for each application must be determined via experimentation. One hidden layer with the number of neurons dependent on the model specification was found to perform the best (see Exhibit 4).

To train each ANN, the standard back propagation learning method is used. To avoid overtraining, a heuristic method based on an increase in the sum square error is used (Hecht-Nielsen, 1990:117). For ANN training, the same random sample training sets, T1 through T18 used in each MRA model, was used to train each ANN model. The resulting ANN data model specifications are given in Exhibit 4.

Empirical Evaluation Criteria

Two error measuring criteria are used to evaluate the MRA and ANN models for the validation data sets. The first is the Mean Absolute Percentage Error (MAPE), and the second is the absolute percentage error.

The MAPE is defined as:

$$MAPE \equiv \left(\sum_{i=1}^{n} \left| \frac{(P_i - A_i)}{A_i} * 100 \right| \right) \div n, \tag{10}$$

where P_i and A_i are the predicted selling price and the actual selling price of property i in the set of n properties. The data model with a smaller MAPE is deemed superior. This error measurement attempts to produce a single number that represents the total error for all properties. This error measurement fails, however, to provide information as to how the error deviates between the properties. For example, if a model provides extremely accurate results for 90% of the properties tested while providing horribly inaccurate results for 10% of the properties tested, the MAPE value for this model may be comparable to another model with unacceptable results (*i.e.*, a large standard deviation in error, but with a comparable MAPE). To measure how the error deviates, a second error measurement criterion was also used.

The absolute percentage error (hereafter referred to as forecasting error, FE) classifies properties into three categories: (1) those with an FE of less than 5%; (2) those with an FE between 5% and 15%; and (3) those with an FE greater than 15%. These forecasting error ranges are chosen based on the understanding that 5% is acceptable to most investors, 5% to 15% is in a fuzzy area and is a somewhat unreliable indicator, while more than 15% is unacceptable.

Forecasting error for property i is defined as:

$$FE \equiv \left| \frac{(P_i - A_i)}{A_i} * 100 \right|. \tag{11}$$

As an illustration, an FE of 19% in the less than 5% range (for either data model MRA or ANN), means that 19% of the properties tested have the predicted selling price within 5% of the actual selling price. It follows that the superior model is the model with the higher percentage of most accurately predicted properties.

Empirical Results

Using the evaluation criteria given in Equations (10) and (11), Exhibit 2 summarizes the comparative forecasting performance of MRA and ANN. The best performing MRA (when comparing MRA to MRA) occurs for model M4 when using both MAPE and FE for evaluation (*i.e.*, generally MAPE is the lowest and FE 5% is the highest for this model). The log-log model, M5, is the poorest performing model for MRA when both MAPE and FE 5% are used for evaluation. The ANN data model tends to overcome functional model misspecification if sufficient data sample size is available. This generally ranged from 13% to 39% of the total sample size.

The results in Exhibit 2 generally demonstrate that as the functional model specification improves, the performance of MRA data model improves where as the performance of the ANN model improves as the training size increases. Although this correlation is to be expected, it also shows that the best data model (regardless of whether MRA or ANN is used) fluctuates as the model specification and training size vary. The fluctuation in the ANN model's performance can be attributed to the large number of possible parameter settings and the absence of a methodical approach to choosing the best settings. For example, experiments must be conducted to determine the best data representation, model specification, number of hidden layers, number of neurons on each hidden layer, learning rate and number of training cycles. All of these interrelate to give the best ANN model. Although exhaustive testing of the various parameters is impossible, one must conduct extensive experimentation. Experimentation always leads one to question whether sufficient tests and combinations of parameters were conducted to obtain the best model. Failure of other studies to conduct thorough experimentation to determine these parameter settings would cause their ANN to perform poorly.

Based on Exhibit 2, the ANN performs better (using both criteria) than the MRA when a moderate to large data sample size is used. In this study, this moderate to large data sample size varied from 13% to 39% of the total data sample and is dependent on the functional model specification (Baum and Haussler, 1989). The MRA performs better (using the MAPE evaluation criterion) than the ANN when a small data sample size is used. For each functional model specification, the MRA's performance is somewhat constant as the sample size varies, whereas the ANN's performance significantly improves as the data sample size increases.

Conclusion

In comparing the MRA and ANN models, attempts have been made to ensure a fair comparison of their predictive performance. Multiple comparisons have been made in which the training size is varied, the functional specification is varied, the temporal prediction is varied and the issues noted in the Implementation Issues section are addressed. There were 108 comparisons made. Two criteria have been used to evaluate the models—MAPE and FE. Based on Exhibit 2, the MAPE and FE criteria may differ in the selection of the best data model (*e.g.*, see M4 T5 in Exhibit 2). Thus, one should be cautious as to which performance criterion is used for evaluating forecasting accuracy as well as the data sample size used. When a moderate to large data sample size is used, the ANN performs better (using both criteria) than the MRA. For this application, the data sample size is 506 to 1,506 observations (from a total of 3,906 observations) before the ANN outperformed the MRA (using both criteria). In general, as the ANN model functional specification becomes more complex, the training sample size must be increased in order for the ANN to perform better than the corresponding MRA model. The MRA performs better (using the MAPE evaluation criterion) than the ANN when a small data sample size is used. For each functional model specification, the MRA's performance is somewhat constant as the sample size varies, whereas the ANN's performance significantly improves as the data sample size increases.

The fluctuation in the ANN model's performance can be attributed to the large number of possible parameter settings and the absence of a methodical approach to choosing the best settings. For example, experiments must be conducted to determine the best data representation, model specification, number of hidden layers, number of neurons on each hidden layer, learning rate, and number of training cycles. All of these interrelate to give the best ANN model. Failure to conduct such experiments may result in a poorly specified ANN model.

If other input variables such as fireplace, number of stories, siding materials, lot size, sewer connection, and financing type are included, then the outcome might be different. This is a limitation of this study (*i.e.*, if more/different input variables are used, then the outcome of the comparison may change). Since the data used in this study is only from sold properties and other studies have shown that the sale set is not representative of the unsold set, the reader should be cautious in applying these conclusions to an unsold set of properties.

The results give a plausible explanation why previous studies have obtained varied results when comparing MRA and ANN predictive performance for housing values. The predictive performance depends on the evaluation criteria (MAPE and FE) used in combination with the training size and model specification. Fluctuation in the ANN model's performance may be due to the larger number of parameters settings chosen via experimentation and dependent on training sample size.

In conclusion, if one provides sufficient data training size and appropriate ANN parameters, then ANN performs better than MRA. Otherwise, the results vary. For practical purposes, the ANN is recommended when there is sufficient sample data set and/or when there is no theoretical basis for the data model functional form. Otherwise, the MRA is recommended.

Endnotes

[1] The ANN models one through five are trained using input variables quarter one through quarter six whereas the MRA models are trained using dummy variables quarter one through quarter five.

[2] The software used for creating, training, and testing the ANN is Stuttgart Neural Network Simulator (SNNS) Version 4.2. This software can be obtained from Univestität Stuttgart, ftp://ftp.informatik.uni-stuttgart.de/pub/SNNS/.

References

Allen, W. C. and J. K. Zumwalt, Neural Networks: A Word of Caution, Unpublished Working Paper, Colorado State University, 1994.

Baum, E. B. and D. Haussler, What Size Net Gives Valid Generalization?, *Neural Computation*, 1989, 1:1, 151–60.

Bishop, C. M., *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

Borst, R. A., Artificial Neural Networks: The Next Modeling/Calibration Technology for the Assessment Community, *Artificial Neural Networks*, 1992, 10, 69–94.

Do, Q. and G. Grudnitski, A Neural Network Approach to Residential Property Appraisal, *The Real Estate Appraiser*, 1992, 58, 38–45.

——., A Neural Network Analysis of the Effect of Age on Housing Values, *Journal of Real Estate Research*, 1993, 253–64.

Fausett, L., *Fundamentals of Neural Networks*, Prentice Hall, 1994.

Geman, S., E. Bienenstock and R. Doursat, Neural Networks and the Bias/Variance Dilemma, *Neural Computation*, 1992, 4, 1–58.

Goldfield, R. M. and R. E. Quandt, Some Tests for Homoskedasticity, *Journal of the American Statistical* Association, 1965, 60, 539–47.

Goodman, A. C., Hedonic Prices, Price Indices, and Housing Markets, *Journal of Urban Economics*, 1978, 5:4, 471–84.

Goodman, A. C. and T. G. Thibodeau, Age-Related Heteroskedasticity in Hedonic House Price Equations, *Journal of Housing Research*, 1995, 6, 25–42.

Goutte, C., Note on Free lunches and Cross-validation, *Neural Computation*, 1997, 9, 1211–15.

Grether, D. and P. Mieszkowski, Determinants of Real Values, *Journal of Urban Economics*, 1974, 1:2, 127–45.

Halvorsen, R. and H. Pollakowski, Choice of Functional Form for Hedonic Price Equations, *Journal of Urban Economics*, 1981, 10:1, 37–49.

Hecht-Nielsen, R., Kolmogorov's Mapping Neural Network Existence Theorem, Paper presented at IEEE First International Conference on Neural Networks, San Diego, CA, 1987.

——., *Neurocomputing*, Reading, MA: Addison-Wesley, 1990.

Huang, C-S., R. E. Dorsey and M. A. Boose, Life Insurer Financial Distress Prediction: A Neural Network Model, *Journal of Insurance Regulation*, 1994, 13:2, 131–67.

Jones, W., M. Ferri and L. McGee, A Competitive Testing Approach to Models of Depreciation in Housing, *Journal of Economics and Business*, 1981, 33:3, 202–11.

Kmenta, J., *Elements of Econometrics*, New York, NY: Macmillan Publishing, 1971.

Linneman, P., Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market, *Journal of Urban Economics*, 1980, 8:1, 47–68.

Mark, J. and M. A. Goldberg, Multiple Regression Analysis and Mass Assessment: A Review of the Issues, *Appraisal Journal*, 1988, 56, 89–109.

Neter, Wasserman and Kutner, *Applied Linear Statistical Models*, Third edition., McGraw-Hill, 1990.

Newsome, B. A. and J. Zietz, Adjusting Comparable Sales Using Multiple Regression Analysis—The Need For Segmentation, *The Appraisal Journal*, 1992, 129–35.

Nguyen, N. H. and W. B. Rogers, The Economic Impacts of the New Tax Laws on the Rutherford County Real Estate Market: An Empirical Study, *Journal of Real Estate Appraisal and Economics*, 1992, 6:1, 43–7.

Palmquist, R. B., Hedonic Price and Depreciation Indexes for Residential Housing: A Comment, *Journal of Urban Economics*, 1979, 6:2, 267–71.

Stone, C. J., Consistent Nonparametric Regression, *Annals of Statistics*, 1977, 5, 595–645.

Tay, D. P. H. and D. K. K. Ho, Artificial Intelligence and the Mass Appraisal of Residential Apartments, *Journal of Property Valuation & Investment*, 1991/1992, 10, 525–540.

Tsukuda, J. and S-I. Baba, Predicting Japanese Corporate Bankruptcy in Terms of Financial Data Using Neural Networks, *Computers & Industrial Engineering*, 1994, 27:1-4, 445–48.

Worzala, E. M., M. M. Lenk and A. Silva, An Exploration of Neural Networks and Its Application to Real Estate Valuation, *Journal of Real Estate Research*, 1995, 10:2, 185–202.

This research was funded in part by a grant from Middle Tennessee State University. We wish to express our appreciation.

Nghiep Nguyen, Middle Tennessee State University, Murfreesboro, TN 37132-0048 or nguyen@mtsu.edu.

Al Cripps, Middle Tennessee State University, Murfreesboro, TN 37132-0048 or csal@mtsu.edu.