Neural Networks In Business Time Series Forecasting: Benefits And Problems

Lin Zhao, Ph.D., Purdue University Calumet, USA

ABSTRACT

Many studies examine the use of Neural Networks (NNs) as a tool for business time series forecasting, but the findings have been mixed and inconsistent. This paper explores the conditions under which NNs can improve business time series forecasting based on studies that compare NNs with traditional statistical models. The findings are that NNs generally outperform alternatives when data are nonlinear or discontinuous, building effective NNs for time series forecasting, including designing and selecting the structure, simulation functions, stopping rules, training algorithms and evaluation criteria remains challenging. A case study is discussed to illustrate these findings, and implications for future research and practice are also provided.

Keywords: Time Series, Forecasting, Neural Networks, Validation

INTRODUCTION

Neural networks (NNs) are mathematical models inspired by biological neural networks. The first application of NNs to forecasting dates back to 1964, when, in his thesis, Hu (1964) applied Widrow's adaptive linear network to weather forecasting. The results were limited due to the lack of effective training algorithms. Until the invention of the back propagation algorithm (Webos, 1974, 1988; Rumelhart et al., 1986), the application of NNs for forecasting stagnated. Over the last two decades we have seen a rapid rise in the application of NNs to a range of business problems including managerial forecasting applications (Gorr, 1994; Sharda, 1994; Hill et al., 1996; Swanson & White, 1996; Zhang, 2001). To assess the benefits of NNs, most empirical research has compared NNs with traditional statistical models using accuracy as the criterion. The results from these comparisons are mixed and confounding (Adya & Collopy, 1998).

Tang, Almeida & Fishwick (1991) and Tang & Fishwick (1993) reported a comparative study of experiments with time series of different complexity in which NNs outperformed the Box-Jenkins model for time series with short memory, and were comparable for series with long memory. Kang (1991) compared NNs and Autobox on 50 of the M-Competition series, and found that NNs often performed better when predicting forecasting horizons beyond some periods ahead, although some researchers don't agree with this. Hill, O'Connor & Remus. (1996) found that when compared with six statistical methods using the 111 time series of the M-competition, NNs did significantly better across monthly and quarterly time series. In contrast, Faraway & Chatfield (1998) reported on a case study in which they used the Bayesian information criterion to compare NNs with the Box-Jenkins and Holt-Winters methods. Their research suggested that NNs which often fitted well gave poor out-of-sample forecasts, and were difficult to interpret. Recently, in the M3-Competition, NNs didn't perform as well as simple methods (Makridakis & Hibon, 2000). With these mixed outcomes, it seems important to systematically identify conditions under which NNs can improve forecasting accuracy.

BACKGROUND

In the Business Source Premier database, 184 articles on the application of NNs to business forecasting can be identified for the period between 1990 and 2001. 41 papers closely associated NNs with business time series forecasting become our focus for this study. Some researchers argue that NNs do a good job in forecasting nonlinear series which are common in business (Hill et al., 1996; Zhang et al., 1998), but some others provide negative findings in their empirical studies (Faraway & Chatfield, 1998). These mixed results support Chamberlin's view

(1965) that "it is perhaps absurd to assume that any specific method of instructional procedure is more effective than all others, under any and all circumstances". Thus, this paper will not argue whether NNs are better for business forecasting or not, but review previous studies to figure out under what conditions, NNs are more likely to outperform traditional methods for business time-series forecasting following condition-seeking research strategies (Greenwald et al., 1986).

First, NNs seems to outperform traditional forecasting methods in the short term for time series with noise. Sharda and Patil (1990) compared a neural network model and AutoBox, a Box-Jenkins forecasting expert system. By using a sample of 75 series from the M-competition, they found that simple NNs trained by the back propogation algorithm, performed as well as the Autobox procedure. Later, Sharda and Patil (1992) and Tang et al. (1991, 1993) found that NNs outperformed Autobox for time series with short memory and/or with noise, and produced similar forecasting results for series with long memory. Marquez et al. (1991) used stimulated data to compare NNs with regression models and found that NNs perform best under conditions of high noise and low sample size. But with less noise or larger sample sizes, they become less accurate. Ansuj et al. (1996) concluded that the forecasts obtained using back propagation NNs were more accurate than those of the ARIMA model with interventions. In 1999, Hu et al. investigated out-of-sample performance of NNs in predicting the weekly British pound/US dollar exchange rate. They concluded that NNs were superior to random walk models when the forecast horizon is short, and that the accuracy of NNs was not very sensitive to the sampling variation of four evaluation criteria (RMSE, MAE, MAPE and MdAPE). Zhang (2001), using both simulated and real data, concluded that the capability of NNs for linear time-series forecasting is quite comparable to Box-Jenkins' ARMA (p, q) processes, when the time series contain noise. Hwang (2001) concluded that back propagation NNs trained with a particular noise level tend to perform better than ARMA (p, q) structures.

Second, NNs seems to outperform traditional forecasting methods for monthly and quarterly time series. As shown in Appendix 1, Chakraboty et al. (1992) found using world observations of monthly flour prices in three cities, NNs outperformed ARMA in multivariate time-series forecasting. Caire et al. (1992) also found that the forecasting quality of NNs is equivalent, but not superior to the result with using the ARIMA approach for daily data. Refenes et al. (1993) compared NNs to multiple linear regression using monthly stock rankings. They found that both the in-sample and out-of-sample performance of the network gave a better fit than linear regression. Nam & Schaefer (1995) found that NNs were superior to the linear regression and Winters' exponential models for monthly airline passenger traffic. Kohzadi et al. (1996) compared NNs and ARIMA models to forecast US monthly live cattle and wheat prices. Their results showed that NNs were considerably more accurate than traditional ARIMA models. Hann & Steurer (1996) found that NNs outperformed linear regression in weekly data, and could produce similar predictions in monthly data with regard to the criteria of Theil's U, the annualized return and the Sharpe Ratio. Later, Hill, O'Connor and Remus (1996) designed a comprehensive comparison published in Management Science. They used a sample of 111 time series from the M-competition in their research. They compared NNs with six traditional forecasting models, including deseasonalized exponential smoothing, Box-Jenkins, deseasonalized Holt's, graphical (human judgment), a combination of six methods and a naïve model. Their study showed that NNs did significantly better than traditional statistical and human judgment methods based on the absolute percentage error (APE), when forecasting quarterly and monthly data. However, with annual data, NNs did not outperform traditional models but were comparable. Subsequently, Goh (1998) compared the accuracy of demand models and found that NNs outperformed the univariate Box-Jenkins approach and multiple loglinear regression on quarterly data, using relative errors. In Leung et al. (2000)'s study, NNs had better forecasting accuracy than multivariate transfer function models and random walk models for currency exchange rate forecasting of monthly data. Moreover, this difference was statistically significant.

Third, NNs seem to outperform traditional forecasting methods for discontinuous and non-linear time series. Hill, O'Connor and Remus (1996) found that NNs were particularly effective on discontinuous time series. Baker & Richards (1999) compared linear regression methods and NNs for forecasting educational spending. Their results showed that general regression neural networks (GRNN) yielded the only results suggesting non-linear sensitivity. Kohzadi et al. (1996) found that NNs can capture more turning points than ARIMA.

Fourth, evaluation criteria matter. Researchers used a variety of accuracy criteria in their empirical studies. MAPE, RMSE and MdAPE seem to be the most popular evaluation criteria. Kuo & Reitsch (1995) compared NNs and conventional methods of forecasting including multiple regression, exponential smoothing and the Box-Jenkins procedure AR (1). As compared with multiple regression, NNs produced a lower standard error of estimate but a higher mean absolute deviation (MAD). As Compared with AR (1) and exponential smoothing, NNs produced lower level of errors when measured by both standard error of estimate and MAD.

Fifth, the structure of NNs matters. Although NNs can be regarded as a powerful forecasting method based on the above description, it is difficult to get the successful results without careful design of their structures as several researchers emphasized. Faraway & Chatfield (1998) indicated that without a careful choice of the architecture, the activation functions and appropriate starting values for the weights, fitting routines may not converge, may converge to a local minimum or may lead to ridiculous forecasts. Thus, some NNs which achieve a good fit with training data may give poor out-of-sample forecasts.

Hoptroff et al. (1991) found that NNs trained by the concurrent descent technique outperformed linear regression when the available data was sparse or noisy. Hansen & Nelson (1997) applied a genetic algorithm (GA) to generate NN architectures that projected a significant strengthening in the nonagricultural employment growth rate that ARIMA and exponential smoothing methods missed. In Baker & Richards' study (1999), only the specified GRNN suggested non-linear sensitivity compared with AR (1). Hansel et al. (1999) applied a GA to determine an architecture that allows NNs to outperform statistical models on the Box-Jenkins data sets. Leung et al. (2000), as mentioned earlier, also presented GRNN which had a higher degree of forecasting accuracy than multi-layered feedforward networks (MLFN). Choudhury et al. (2002) specified a fuzzy associative memory NN model that provided better performance than ARIMA in yearly manpower forecasting.

In addition, NNs without hidden layers may provide performance comparable to some traditional forecasting methods as several studies found. Sharda and Patil (1990) found that NNs without a hidden layer are functionally similar to Box-Jenkins ARIMA models. Carvalho et al. (1998) compared the use of back propagation NNs to forecast travel demand from disaggregate discrete choice data with logit models. They found that NNs with no hidden layers exhibit almost identical performance to logit models in all three cases. Hwarng & Ang (2001) argued that a simple two-layered network, with proper input window size, using a semi-linear transfer function and trained by the back propagation algorithm, is able to consistently outperform the MLBP network, and that the two layered network is comparable to or better than the Box-Jenkins modeling approach for a majority of the time series corresponding to ARMA (p, q) structure.

CASE STUDY

Of all of the attempts to apply neural networks to the problem of extrapolating time series, one of the most successful studies is probably that of Hill, O'Connor and Remus (1996). They used simple neural networks to forecast multiple horizons for the 111 series subset of the M-Competition (Makridakis et al., 1982) data. They compared forecasts from their network with those from well-established alternative methods (deseasonalized single exponential smoothing, Box-Jenkisn, deseasonalized Holt's exponential smoothing, a judgmental method, a combination of six methods, and naïve forecasts). They used absolute percentage errors (APEs) to make the comparisons on 1451 forecasts. They complied, in other words, with all of the procedures for making forecasting comparisons that Collopy, Adya, and Armstrong (1994) argued are essential for drawing meaningful conclusions. In addition, they analyzed the conditions under which the neural network performed well. In short, their research design was exemplary.

One of the ways in which science progresses most rapidly is through the replication and extension of important studies, if we take the definition of replication as the process of "re-searching" an observation, investigation, or experimentation to compare findings (Berthon, Pitt, Ewing, & Carr, 2002). It is obviously valuable to replicate Hill, O'Connor and Remus study as the first step to better understand their research, even further to get their findings generalized and extended. So we have attempted to directly replicate that study. The pattern of our results matches that of the original study and supports many of its conclusions, but we were not able to obtain nearly the magnitude of the improvements reported there. The stopping rules proposed in the original study have high variance associated with them, and do not seem reasonable for such applications. In addition, the original study's networks, and in particular its stopping rules, are not described in a way that permits ready replication. However, Hill et al.'s conclusion that neural networks represent a good approach to extrapolating discontinuous series and certain nonlinear series was supported by our replication.

DISCUSSIONS AND IMPLICATIONS

NNs are not a final solution for time series forecasting under any conditions, but they seem to improve forecasting especially in short time series with noise. The advantages of NNs discussed suggested that the future looks bright for the use of NNs for business time series forecasting. However, it is necessary to pick the evaluation criteria and structures of NNs carefully.

The most important implication of our research is that it is considerably harder to obtain substantial improvements in extrapolative forecasting with neural networks than might be assumed reading this earlier study. This results in part from the "black box" character of neural networks. Since we cannot understand precisely which of the features in Hill et al.'s networks account for their apparent superior performance, it is difficult to know where to put one's energy when using them in other environments.

Given the pattern of results obtained in the case study, though, it does appear that neural networks may be an appropriate extrapolation technique for nonlinear and discontinuous series. This is important, because those have been difficult to forecast successfully with other methods. For series that can be identified as nonlinear and discontinuous, it may be possible to gain significant improvements over other methods even by using very simple neural network architectures and stopping rules.

As for practice, we can only recommend that those building forecasting systems based around neural networks proceed cautiously. Clearly, over-fitting is a hazard to be avoided. Though the heuristics provided in this study represent one approach to this problem, it is not clear that they will deliver the desired results. Other approaches should be explored, and until some robust principles establish themselves in the literature, practitioners should remain somewhat skeptical about extrapolative forecasts produced by neural networks and the claimed benefits.

ACKNOWLEDGEMENT

The author would like to thank Dr. Fred Collopy and Dr. Miles Kennedy at Case Western Reserve University for their continuous help and support to develop this paper.

AUTHOR INFORMATION

Lin Zhao, Ph.D., is an Assistant Professor of Information Systems in the School of Management at Purdue University Calumet. Her research focuses on Human-Computer Interaction (HCI) and business decision making. Her teaching interests include principles of information systems, system analysis & design, system development, database management, business intelligence, and E-business strategies. She received her Ph.D. in Management Information Systems from Case Western Reserve University in 2008. Before joining Purdue University Calumet, Dr. Zhao taught one year in the Department of Computer and Information Science at Arkansas Tech University.

REFERENCES

1. Adya, M. & Collopy, F. (1998) How Effective are Neural Networks at Forecasting and Prediction? A Review and Evaluation. *Journal of Forecasting*, 17, 481-495.
2. Ansuj, A. P., Camargo, M. E., Radharamanan, R. & Petry, D. G. (1996). Sales forecasting using time series and neural networks. *Computer and Industrial Engineering*, 31, ½, 421-424.
3. Baker, B. D. & Richards, C. E. (1999). A comparison of conventional linear regression methods and neural networks for forecasting educational spending. *Economics of Education Review*, 18, 405-415.
4. Berthon, P., Pitt L., Ewing, M. & Carr, C. L. (2002). Potential Research Space in MIS: A Framework for Envisioning and Evaluating Research Replication, Extension, and Generation. *Information Systems Research*, 13, 4, 416-427.

5. Caire, P., Hatabian, G. & Muller, C. (1992). Progress in forecasting by neural networks. Proceedings of the International Joint Conference on Neural Networks, 2, 540–545.
6. Carvalho, M., Dougherty, M., Fowkes A. & Wardman, M. (1998). Forecasting travel demand: a comparison of logit and artificial neural network methods. *Journal of the Operational Research Society*, 49, 717-722.
7. Chakraborty, K., Mehrotra, K., Mohan, C. K. & Ranka S. (1992). Forecasting the behavior of multivariate time series using neural networks. *Neural Networks*, 5, 961-970.
8. Chamberlin, T. C. (1965). The method of multiple working hypotheses. *Science*, 148, 754-759.
9. Choudlury, J. P., Sarkar, B. & Mukherjee, S. K. (2002). Forecasting of engineering manpower through fuzzy associative memory neural network with ARIMA: a comparative study. *Neurocomputing*, 47, 241-257.
10. Collopy, F., Adya, M. & Armstrong, J. S. (1994). Principles for Examining Predictive Validity: The Case of Information-Systems Spending Forecasts. *Information Systems Research,* 5, 2, 170-179.
11. Faraway, J & Chatfield, C. (1998). Time series forecasting with neural networks: a comparative study using the airline data. *Applied Statistics*, 47, 2, 231-250.
12. Goh, B. (1998). Forecasting residential construction demand in Singapore: a comparative study of the accuracy of time series, regression and artificial neural network techniques. *Engineering, Construction and Architectural Management,* 5, 3, 261-275.
13. Gorr, W. (1994). Research Perspective on Neural Networks. *International Journal of Forecasting*, 10, 1, 1-4.
14. Greenwald, A. G., Pratkanis, A. R., Leippe, M. R. & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? *Psychological Review*, 93, 2, 216-229.
15. Hann, T. H. & Steurer, E. (1996). Much ado about nothing? Exchange rate forecasting: neural networks vs. linear models using monthly and weekly data. *Neurocomputing*, 10, 323-339.
16. Hill, T., O'Connor, M. & Remus, W. (1996). Neural Network Models for Time Series Forecasts. *Management Science*, 42, 7, 1082-1092.
17. Hoptroff, R. G., Bramson, M. J. & Hall, T. J. (1991). Forecasting economic turning points with neural nets. *International Joint Conference on Neural Networks*, 1, 347-352.
18. Hu, M. J. C. (1964). Application of the adaline system to weather forecasting. *Master Thesis, Technical Report 6775-1*. Stanford, CA: Stanford Electronic Laboratories.
19. Hwarng, H. B. & Ang, H. T. (2001). A simple neural network for ARMA(p, q) time series. *Omega*, 29, 319-333.
20. Kang, S. (1991). An investigation of the Use of Feedforward Neural Networks for Forecasting. Ph. D. Thesis, Kent State University.
21. Kohzadi, N., Boyd, M. S., Kaastra, I., Kermanshahi, B. S. & Scuse, D. (1995). Neural networks for forecasting: an introduction. *Canadian Journal of Agricultural Economics*, 43, 463-474.
22. Kuo, C. & Reitsch, A. (1995). Neural networks vs. conventional methods of forecasting. *Journal of Business Forecasting*, winter, 17-22.
23. Leung, M. T., Chen, A. & Daouk H. (2000). Forecasting exchange rates using general regression neural networks. *Computers & Operations Research*, 27, 1093-1110.
24. Makridakis, S., Anderson, A., Carbone R., Fildes, R., Hibon M., Lewandowski, R., Newton, J., Parzen, E. & Winkler, R. (1982). The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. *Journal of Forecasting*, 1, 111-153.
25. Makridakis, S. & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, 16, 451-476.
26. Marquez, L., Hill, T., Worthley, R. & Remus, W. (1991). Neural network models as an alternative to regression. Proceedings of the IEEE 24th Annual Hawaii International Conference on Systems Sciences, 4, 129-135.
27. Refenes, A. N., Azema-Rarac, M. & Zapranis, A. D. (1993). Stock ranking: neural networks vs. multiple linear regression. IEEE International Conference on Neural Networks, 3, 1419-1426.
28. Rumelhart, D. E. & J. McClelland. (1986). *Parallel Distributed Processing*. Cambridge, MA: MIT Press.
29. Sharda, R. & Patil, R. B. (1990). Neural networks as forecasting experts: an empirical test. *International Joint Conference on Neural Networks*, 1, 491-494.

30. Sharda, R. & Patil, R. B. (1992). Connectionist approach to time series prediction: an empirical test. *Journal of Intelligent Manufacturing*, 3, 317-323.
31. Sharda, R. (1994). Neural networks for the MS/OR analyst: An application bibliography. *Interface* 24, 2, 116-130.
32. Swanson, N. R., White, H. (1997). Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. *International Journal of Forecasting*, 13, 439-461.
33. Tang, Z., C. de Almeida, & Fishwick P. A. (1991). Time Series Forecasting using Neural Networks vs. Box-Jenkins Methodology. *Simulation*, 57, 5, 303-310.
34. Tang, Z. & Fishwick P. A. (1993). Feed-Forward Neural Nets as Models for Time Series Forecasting. *ORSA Journal of Computing*, 5, 4, 374- 386.
35. Werbos, P. J. (1974). Beyond regression: new tools for prediction and analysis in the behavioral science. Ph.D. Thesis, Harvard University
36. Werbos, P. J. (1988). Generalization of back propagation with application to a recurrent gas market model. *Neural Networks*, 1, 339-356.
37. Zhang, G. (2001). An investigation of neural networks for linear time-series forecasting. *Computers & Operations Research*, 28, 1183-1202.