「정보시스템연구」 제10권 제2호
한국정보시스템학회
2001년 12월, pp. 151-164.

Airline In-flight Meal Demand Forecasting with Neural Networks and Time Series Models

Young-Chan Lee Sogang University chanlee@family.sogang.ac.kr
Chang Gab Seo Tongmyong University gabida@tit.ac.kr

<목 차>

Ⅰ. Introduction

The activities of all enterprises should be planned and performed on the basis of demand forecasting. The production of goods or services performed without forecasting could bring on many risks. Therefore, demand forecasting is probably the most fundamental factor in production planning and control. After forecasting is processed, the enterprise can aggregate production planning, capacity planning, scheduling, inventory planning, control, and etc.

Although it is very difficult to forecast under uncertain political, economic, and social changes, enterprises should first collect and analyze the demand data related to the market.

But forecasting is not always perfect and the real data might be different from that of the forecasted data. Like this, demand forecasting has the unfortunate problem in which we are not sure how accurate the data is. Accuracy works as a decision factor in dispute over the necessity of forecasting. On one side, the most important factor which influences the results of forecasting is the selection of forecasting techniques. Even though many forecasting techniques are suggested, they may not all be suitable.

Specifically, most airline in-flight meal services are provided without forecasting the exact demand in need, therefore, may result in the delay of plane schedules and the waste of meals on account of inaccurate forecasting.

Meanwhile, some studies presented a hybrid forecasting technique which combined both neural network and exponential smoothing as an outstanding forecasting technique of time series data for airline in-flight meal demands(Suh & Kwak, 1998, Kwak et al., 1998, Kwak et al., 2000). Specifically, they conducted comparison studies among exponential smoothing, Box-Jenkins, and neural networks. Although Box-Jenkins is known to be the most theoretically perfect technique, they showed that exponential smoothing is superior to Box-Jenkins in case of short-term forecasting, and they finally suggested a mixed neural network model using exponential smoothing expectancy as an input in neural networks. In these studies, they used mean absolute error(MAE) and mean square error(MSE), which was the outstanding method in estimating the accuracy of forecasting technique, while they focused on reducing the error(Mathews & Diamantopoulos, 1994).

In this paper, we also try to show the improved forecasting performance in time series data when used with neural networks in case of airline in-flight meal demand forecasting.

The remainder of this paper is organized as follows: In section 2, we review on neural networks and show recent trends in forecasting. In section 3, we present the methodology and the results. The last section is the conclusion.

Ⅱ. Literature Review on Neural Networks

Recently, neural networks have been widely applied to various areas. It has been shown that the accuracy of the model using neural networks is quite excellent. In this paper, we also tried to use neural networks as a forecasting technique.

Neural networks do have certain drawbacks in mathematical background, but it is still applied to many fields because of the variety in practical applications (Salchenberger et al., 1992; Eddy et al., 1993; Norman & Wang., 1993; Chanda et al., 1994; Philipoom et al., 1994; Jain & Nag, 1995; Markham & Ragsdale, 1995). It especially shows its excellence in time series forecasting which needs robustness against noise or variation(Gorr et al., 1994; Hill et al., 1996; Kim & Noh, 1996; Kwon & Golden, 1996; Jo, 1998; Suh & Kwak, 1998; Kwak et al., 1998; Kwak et al., 2000).

Up to now, many studies on neural networks have been published, but the robustness of neural networks has not reached one precise conclusion. Kwon & Golden(1996) showed neural networks were superior to traditional statistical models. Hill et al.(1996), Jain & Nag(1995), Markham & Ragsdale(1995), Philipoom et al.(1994), and Salchenberger et al.(1992) came up with the same result.

Kim & Noh(1996) conducted comparison study between Box-Jenkins and neural networks. They showed that neural networks are superior to Box-Jenkins and that they had statistical significance. But it was also shown that there was no significant difference among other various neural network models.

On the other hand, Gorr et al.(1994) showed that neural networks are not significant with a statistical test.

But most of the results of these studies showed that the neural networks' accuracy depended on the number of nodes, sample size, and architecture of neural networks. Neural networks have some advantages in business problems. The first is its ease of application. Neural networks have the ability of internal pattern recognition by simple data learning process. The second is robustness. Even though data set is incomplete or distorted, neural networks can make good results using learning algorithm without additional data handling. The third is that it is useful to analyze nonlinear relations among input variables since neural networks with numerous neurons are of parallel structures and nonlinear functions of higher degree.

In the real world, we encounter higher degree nonlinear systems correlated with input variables. Therefore, neural networks are profitable in modeling with complex and correlated variables.

Meanwhile, neural networks have some defects. First, if the data is not sufficient enough or there isn't proper learning function among the data, then a satisfied solution cannot be found. Second, it is hard to explain why this result can be taken for the result of neural networks. The reason is that the result is generated from repetitive calculation of many weights and input patterns, and the weights are determined by complicated learning activity functions. Third, neural networks can use up much time and cost during the data selection, learning, processing, and analysis. Fourth, neural networks have problems in processing time. Connections between each node means multiplication, so the total processing time relates highly with the numbers of connection. Therefore, a small increase in unit numbers can cause a large increase in the total processing time. But fortunately, these problems have been recently solved by several application programs.

On one side, many studies show that neural networks are superior to other techniques when data is sufficient enough. This means that neural networks are inferior when there is lack of data. Therefore, in case of insufficient data for one reason or another, it is not recommended to use neural networks.

Classification problems using neural networks such as discriminant analysis, can be solved effectively by using highly correlated data with factor analysis. But in demand forecasting of time series data, Few researchers use this kind of highly correlated data. Furthermore, most time series analyses related to neural networks simply focus on comparing two techniques. Kim & Noh(1996) simply compare time series analysis with neural network models. Suh & Kwak(1998) showed the usefulness and application possibility of hybrid neural networks which combined both the neural network and traditional time series model.

Recently, the hybrid neural networks were considered as an effective model for the improvement of forecasting performance. Jo's study(1998) is a good example showing improvement of forecasting performances in neural networks. He designed a hybrid neural network model by adding virtual series to time series data. In this study, we focused on improving the forecasting performance by considering additional time series data as input variables.

It has been shown that neural networks using highly correlated data can improve the performance for classification problems. Following this logic, we applied highly correlated data to demand forecasting with time series data.

We compared and analyzed the forecast results of two models, one with highly correlated data and the other without it.

Ⅲ. Analysis of Demand Forecasting Results

In this paper, the data of airline in-flight meals from April 1st to October 15th in 1997 were used for forecasting. To improve the performances of demand forecasting, we employed a multi-layer perceptron model with one hidden layer using backpropagation algorithm.

Three different types of neural network models are used. The first model is called the 'simple neural network model'. Assuming that the past amount of demand for airline meals affect future one, simple neural network model uses four time lagged data. They are previous day's data (lag 1), data of two days before (lag 2), data of three days before (lag 3), and data of one week before (lag 7). For the simple neural network model, we examine four different networks with different number of input nodes. The network with one input node uses lag 1 data. The network with two input nodes uses lag 1 data and lag 2 data. The network with 3 input nodes uses lag 1, lag 2, and lag 3 data. The network with 4 input nodes uses all of the four time lagged data.

The second model is called the 'hybrid neural network model'. In addition to four time lagged data which are used in the first model, the hybrid model uses the forecast data from exponential smoothing which is known to be superior to other forecasting technique such as Box-Jenkins in previous studies(Suh & Kwak, 1998, Kwak et. al., 1998, Kwak et. al., 2000). Each of the four networks in the hybrid model is designed by adding one more input node for the exponentially smooted data to the network examined in the first model. Thus, the numbers of input nodes are 2, 3, 4, and 5.

The third model is called the 'hyper-hybrid neural network model'. The hyper-hybrid model adds 4 more input nodes to the networks examined in the hybrid model. Additional 4 input data are lag 1 data and exponentially smoothed data from each of two other airlines which travel the same route. Therefore, the numbers of input nodes are 6, 7, 8, and 9.

Total number of observed data is 198. First 7 data points are used for time lagging and 45 data points are used for testing. Then, 116 data points are randomly selected and used for training and the other 30 are used for cross validation. <Table 1> presents the description of the data as above mentioned.

<Table 1> Data description

Total data(198)	Usee data(191)	Training data	116
		Validation data	30
		Test data	45
	Excepted data(7)	Time lag	7

We processed 2000 iterations for model learning. In this process, we adopted stopped training (or early stopping, optimal stopping) for cross validation. After this training, forecasting results were produced by using weights that were calculated by activity function when training was stopped. We employed hyperbolic tangent (tanh) as an activity function for tanh is the most desirable transfer function(Brown & Harris, 1994). This nonlinear transfer function is defined by:

$$\tanh(x) = \frac{1-e^{-2x}}{1+e^{-2x}}$$

Input layer has nodes of each independent variable and is connected to the hidden layer. It does not matter how many numbers of nodes are in the hidden layer, but generally the number of nodes in the hidden layer is less than or equal to twice the number of the input nodes. These hidden nodes are connected to the output node. Our models have only one output node because we have only on dependent variable. We considered that the # of nodes in the hidden layer is twice the # of input nodes according to general view. We used default setting of the neural network training programfor the number of hidden nodes. The training program was Neurosolutions ver. 3.0.

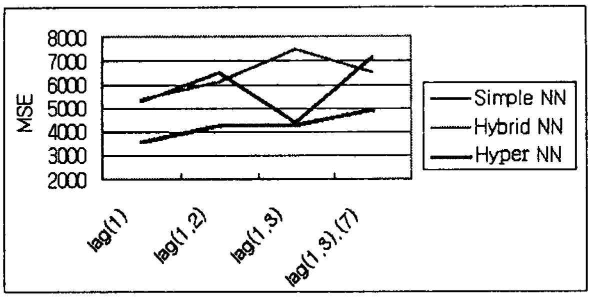
As you see in <Table 2> for airline UA807, the hyper-hybrid neural network model (L model) has the least error during training epochs. It means that the weights calculated by the L model are the best. The L model contains input nodes of all time lag data, forecast data from exponential smoothing, and 4 other data from the two airlines.

It is necessary to divide data set as training data set, cross validation data set and test data set to minimize underfitting or overfitting problems. The training data set is for fitting data, while the cross validation data set is for preventing weights from underfitting and overfitting. The test data set is for performance measurement. We computed the cross validation error periodically during training, and stopped training when the cross validation error started to increase.

<Table 2> Comparison of neural network models at 807E

807E			Data Sets					
			Training Data		Validation Data		Test Data	
Models	Input Noes		Iterate	NMSE	Iterate	NMSE	MAE	MSE
Simple Neural Networks	lag(1)	A	2000	0.0947	1545	0.1190	59.0573	5259.7275
	lag(1-2)	B	2000	0.0688	549	0.1426	68.0232	6471.5474
	lag(1-3)	C	2000	0.0621	7	0.1415	**51.0659**	**4371.1245**
	lag(1-3), (7)	D	2000	0.0434	1510	0.1430	67.442	7110.2612
Hybrid Neural Networks	lag(1), **Exp.**	E	2000	0.0873	1879	0.1494	**54.1973**	**5361.0044**
	lag(1-2), **Exp.**	F	2000	0.0609	556	0.1435	64.2420	6125.2446
	lag(1-3), **Exp.**	G	2000	0.0505	1471	0.0990	65.0600	7464.6646
	lag(1-3), (7), **Exp.**	H	2000	0.0280	763	0.1123	63.5964	6517.3833
Hyper-Hybrid Neural Networks	lag(1), Exp., **Cor.**	I	2000	0.0181	5	0.1275	**45.9372**	**3552.6597**
	lag(1-2), Exp., **Cor.**	J	2000	0.0096	3	0.1294	51.1397	4264.7446
	lag(1-3), Exp., **Cor.**	K	2000	0.0069	5	0.1344	50.8001	4304.6714
	lag(1-3), (7), Exp., **Cor.**	L	2000	**0.0010**	6	0.1486	52.9493	4903.6475

Test results using the weights of cross validation showed that I model had the least mean absolute error and mean square error for airline UA807. So, hyper-hybrid neural network model produced better forecasting performances than other models. Comparing three models in respect of the least error, forecasting accuracy is obtained in the order of hyper-hybrid neural network model, simple neural network model, and hybrid neural network model. If we compare the performance measures among the networks with same input nodes, i.e., with same time lagged data, forecasting accuracy is obtained in the order of hyper-hybrid model, hybrid model, and simple model. Figure 1 shows the mean sqare errors of the models for different input data.



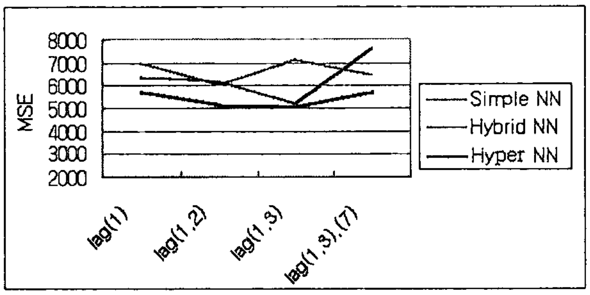
<Figure 1> Comparison of neural network models at each time lag in 807E

In the training process of airline UA808, the hyper-hybrid neural network model (K model) has the least forecasting error as for UA807. K model contains input nodes of all time lag data except for lag 7, forecasting data by exponential smoothing, and 4 other data from the two airlines.

<Table 3> Comparison of neural network models at 808E

808E			Data Sets					
			Training Data		Validation data		Test Data	
Models	Input Nodes		Iterate	NMSE	Iterate	NMSE	MAE	MSE
Simple Neural Networks	lag(1)	A	2000	0.0493	3	0.0978	63.6558	6417.9395
	lag(1-2)	B	2000	0.0455	2	0.0982	62.7859	6195.9771
	lag(1-3)	C	2000	0.0368	2	0.0950	**57.3638**	**5232.5596**
	lag(1-3), (7)	D	2000	0.0263	1	0.0924	69.9925	7616.1089
Hybrid Neural Networks	lag(1), **Exp.**	E	2000	0.0479	1121	0.0977	66.8971	6960.6313
	lag(1-2), **Exp.**	F	2000	0.0444	10	0.0990	**61.9347**	**6033.6870**
	lag(1-3), **Exp.**	G	2000	0.0325	436	0.0981	67.9369	7093.3667
	lag(1-3), (7), **Exp.**	H	2000	0.0276	6	0.0915	64.7215	6448.6104
Hyper-Hybrid Neural Networks	lag(1), Exp., **Cor.**	I	2000	0.0068	491	0.0926	60.6225	5755.7935
	lag(1-2), Exp., **Cor.**	J	2000	0.0040	12	0.1048	**55.0114**	5114.5107
	lag(1-3), Exp., Cor.	K	2000	**0.0026**	19	0.0971	55.5418	**5038.7451**
	lag(1-3),(7), Exp., **Cor.**	L	2000	0.0028	133	0.1019	61.4292	5766.7134

Forecasting results using the weights of cross validation shows that the J model has the least mean absolute error and that the K model has the least mean square error for airline UA808. This means that hyper-hybrid neural network model produces better forecasting results than other models. Comparing three models in respect of the least error for UA808, error, forecasting accuracy is obtained in the order of hyper-hybrid neural network model, simple neural network model, and hybrid neural network model. If we compare the performance measures among the networks with same input nodes, i.e., with same time lagged data, forecasting accuracy is obtained in the order of hyper-hybrid model, hybrid model, and simple model. Figure 2 shows the mean sqare errors of the models for different input data.

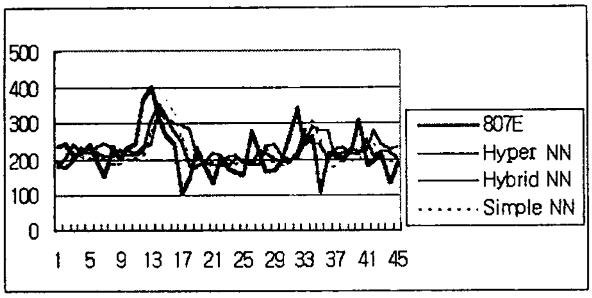


<Figure 2> Comparison of neural network models at each time lag in 808E

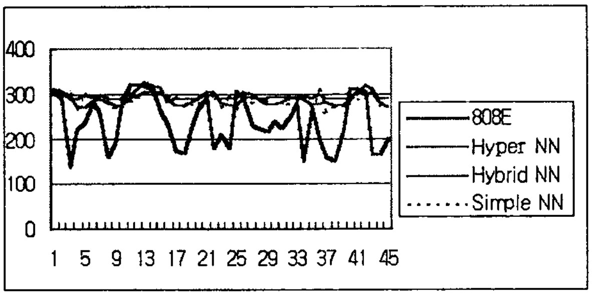
The summary of forecasting is presented in <Table 4>. As you see in the summary table, hyper-hybrid neural network model is better than other neural network models in many cases. So we are able to forecast better by using various correlated data to original time series data. The comparison of forecasting results is presented in <Figure 3> for UA807, <Figure 4> for UA808, respectively.

<Table 4> Results of model comparison at each airline

Airlines	Comparison rule	Model comparison	Example model
807E	Minimum error	Hyper<Simple<Hybrid	I<C<E
	The same time lag	Hyper<Hybrid<Simple	J<F<B, L<H<D
808E	Minimum error	Hyper<Simple<Hybrid	K<C<F
	The same time lag	Hyper<Hybrid<Simple	J<F<B, L<H<D



<Figure 3> Comparison of forecasting results at 807E



<Figure 4> Comparison of forecasting results at 808E

Ⅳ. Conclusions

Although neural networks have some advantages in time series forecasting problem by using highly correlated data, most studies on demand forecasting show only the difference of forecast error between neural networks and traditional time series analysis.

This study shows the possibility of application in neural networks using other highly correlated time series data. To forecast demand, we considered 'simple neural network model', 'hybrid neural network model' adding exponential smoothing expectancy to simple neural network model, and 'hyper-hybrid neural network model' adding forecasting data of other highly correlated time series data to hybrid neural network model.

We compared and analyzed the accuracy of forecasting for each model. Through this study, several important contributions are provided as follows.

Firstly, Suh & Kwak(1998), Kwak et al.(1998, 2000) showed that the hybrid neural network model is superior to the simple neural network model. In this study, however, we showed that the simple neural network model is superior to the hybrid neural network model in some case. When we considered the same time lag, forecast performance among three models depended on each combination of input variables. Mostly hybrid neural network models are superior to simple neural network models, and hyper-hybrid neural network models are superior to hybrid neural network models.

Secondly, hyper-hybrid neural network models are superior to other neural network models, so it is desirable to use forecasting technique adding various data correlated to original time series data. Last of all, in forecasting the exact amount of in-flight meals, each airline company could reduce the waste of meals, and therefore, leads to the reduction of cost. Moreover, it enhanced the cost competitiveness of each airline, keeps the schedules on time, and leads to offering overall services of good quality.

Although the above contribution, this paper has the following limits.

Firstly, we focused only on the demand of airline meals, so we could not consider the possibility of usefulness or application in other fields.

Secondly, recurrent network or time-lag recurrent algorithm is suitable to time series data analysis. But, we could not consider other neural network architectures.

Time lagged recurrent networks(TLRNs) are multi-layer perceptrons(MLPs) extended with short term memory structures. Most real-world data contains information in its time structure, i.e. how the data changes with time. Yet, most neural networks are purely static classifiers. TLRNs are the state of the art in nonlinear time series prediction, system identification and temporal pattern classification. In the future study, we will consider time lagged recurrent networks(TLRNs) as our research model.

REFERENCES

Brown, M. and C. Harris, Neuro-Fuzzy Adaptive Modeling and Control, Prentice-Hall, 1994.

Chanda, G.D., G.S. Dispensa, and S. Ghose, "Comparing the predictive performance of a neural network model with some traditional market response models," International Journal of Forecasting, Vol. 10, No. 2, 1994, pp. 235-244.

Drury, D.H., "Issues in forecasting management," Management International Review, Vol. 9, No. 1, 1990, pp. 317-329.

Duarte, T. and R. Taffler, "Neural Networks and Empirical Research in Accounting," Accounting and Business Research, Vol. 12, No. 2, 1996, pp. 347-355.

Eddy, P., M.Y. Hu, and M.S. Hung, "Two-Group Classification Using Neural Networks," Decision Sciences, Vol. 24, No. 4, 1993, pp. 825-845.

Gorr, W.L., N. Daniel, and J. Szczypula, "Comparative study of artificial neural network and statistical models for predicting student grade point averages," International Journal of Forecasting, Vol. 10, No. 1, 1994, pp. 17-34.

Hill, T., M. O'Connor, and W. Remus, "Neural Network Models for Time Series Forecasts," Management Science, Vol. 42, No. 7, 1996, pp. 1082-1092.

Jain, B.A. and B.N. Nag, "Artificial Neural Network Models for Pricing Initial Public Offerings," Decision Sciences, Vol. 26, No. 3, 1995, pp. 283-301.

Jo, T.C., "Neural Approach to Time Series Prediction by Generating Virtual Series," Proceedings of the 1998 Fall Conference of the Korean Institute of Industrial Engineers, 1998, pp. 1210-1214.

Kim, S.H. and H. Noh, "A comparative Study of ARIMA and Neural Network Models: Case Study in Korean Corporate Bond Yields," Proceedings of KORMS '96, 1996, pp. 19-22.

Kwak, S.H., Y.C. Lee, and K.T. Park, "A Study on Improving Forecasting Performance in Time Series Data with Neural Networks," Proceedings of KES '98, 1998, pp. 196-204.

Kwak, S.H., Y.C. Lee, and K.H. Yang, "A Study on Improving Forecasting Performance in Time Series Data with Neural Networks," Proceedings of INFORMS-KORMS Seoul 2000 International Conference, 2000.6.18-21.

Kwon, O. and B. Golden, "A Neural Network Model for Predicting Atlantic Hurricane Activity," Proceedings of KORMS '96, 1996, pp. 39-42.

Markham, I.S. and C.T. Ragsdale, "Combining Neural Networks and Statistical Predictions to Solve the Classification Problem in Discriminant Analysis," Decision Sciences, Vol. 26, No.2, 1995, pp. 229-241.

Mathew, B.P. and A. Diamantopoulos, "Towards a Taxonomy of Forecast Error Measures," Journal of Forecasting, 1994, Vol. 13, No. 2, pp. 405-422.

Norman, P. and S. Wang, "Application of the Back Propagation Neural Network Algorithm with Monotonicity Constraints for Two-Group Classification Problems," Decision Sciences, Vol. 24, No. 1, 1993, pp. 60-75.

Philipoom, P.R., R.L. Paul and L. Wiegmann, "Using Neural Networks to Determine Internally-Set Due-Date Assignments for Shop Scheduling," Decision Sciences, Vol. 25, No. 5/6, 1994, pp. 825-851.

Salchenberger, L.M., C.E. Mine, and N.A. Lash, "Neural Networks: A New Tool for Predicting Thrift Failures," Decision Sciences, Vol. 23, No. 4, 1992, pp. 899-916.

Suh, C.J. & S.H. Kwak, "Airline meal demand forecasting with neural network and time series," Journal of the Korean Production and Operations Management Society, Vol. 9, No. 2, 1998, pp. 67-95.

<초록>

인공신경망을 이용한 항공기 기내식 수요예측의 예측력 개선 방안에 관한 연구

이 영 찬 서강대학교 경영학연구원 chanlee@family.sogang.ac.kr
서 창 갑 동명정보대학교 경영정보학과 gabida@tit.ac.kr

현재의 항공사 기내식 수요예측 시스템으로는 항공기 운항의 지연이나 초과 주문으로 인한 손실 문제를 해결하기 힘든 것으로 알려져 있다. 이러한 문제를 해결하기 위해 본 연구에서는 항공기 기내식 시계열 자료만을 입력변수로 사용한 단순인공신경망모형(simple neural network model), 단순인공신경망모형에 전통적인 시계열 기법(본 연구에서는 지수 평활법)의 예측 결과를 입력변수로 추가한 혼합인공신경망모형(hybrid neural network model), 그리고 혼합인공신경망 모형에 상관관계가 높은 다른 시계열 자료(본 논문에서는 유사 노선의 다른 항공기 기내식 시계열 자료)를 인공신경망의 입력변수로 추가시킨 하이퍼혼합인공신경망모형(hyper hybrid neural network model)을 새로운 항공기 기내식 수요예측 기법으로 제안하고, 이들 모형의 예측력을 비교·분석하였다. 분석 결과 하이퍼혼합인공신경망 모형의 예측력이 가장 우수한 것으로 나타나, 인공신경망을 기반으로 한 수요예측에 있어 상관관계가 높은 다른 시계열 자료를 입력변수로 추가함으로써 인공신경망모형의 예측력을 개선시킬 수 있음을 알 수 있었다.