Journal of the Society of Korea Industrial and Systems Engineering
Vol. 25, No. 2, pp.77-83, April 2002.

보정된 가솔린 수요예측치: 인공신경망적 접근

염창선

부경대학교 경영학부

Adjusted Gasoline Demand Forecasts: Artificial Neural Networks Approach

Chang Sun Yum

Division of Business Administration, Pukyong National University

본 연구에서는 가솔린 시계열 예측전문가들이 수요를 예측하고, 더 나아가 직감적으로 행하고 있는 보정과정을 자동화하기 위해 신경망을 사용한다. 가솔린 수요 예측분야에서 보정을 위해 사용되는 전형적인 판단요소는 정부 에너지 절약 정책, 에너지 산업의 파업, 공휴일 등이 있다. 주요 추세가 순환신경망에 의해 예측되고 이들 판단요소의 효과가 다층신경망에 의해 탐지되어 보정된다. 가솔린 수요에 대한 실험결과는 보정과정을 갖는 예측구조가 하나의 신경망을 사용하는 예측구조 보다 더 나은 예측력을 보였다. 그리고 본 연구에서 제시한 접근방법이 순환신경망이나 ARIMA 모델을 사용하는 것보다 더 나은 결과를 가졌다.

KeyWords: Gasoline Demand Forecasting, Artificial Neural Networks, Adjustment

1. Introduction

During the past few decades, various forecasting models have been developed, and the forecasting accuracy has been improved substantially. Among the many models studied and tested, time series models are the most popular. But, the time series models are deficient in the sense that they merely extrapolate past patterns in the data without reflecting the expected irregular and infrequent future events. Even the neural network can not overcome this limitation with its traditional architecture.

A clue to overcome this hurdle can come from the observation that in practice, forecasting experts judgmentally adjust the statistical forecasts. A set of research supports the rationale of such human expert's judgmental adjustments. Georgoff and Murdick(1986) claimed that the forecaster should incorporate subjective judgments in dynamic situations when the statistical models can not reflect significant internal and external changes. Edmundson et al.(1988) commented that the well-structured judgmental process can consistently outperform the statistical model-based extrapolation. Lee et al.(1990) suggested the adjustment process using expert systems. From the other angle, Wolfe and Flores (1990) have also shown that the ARIMA model based forecasts can be enhanced by adopting the Satty's Analytical Hierarchy Approach for the judgmental adjustment. Flores et al.(1991) have compared the performance of two methods: AHP and centroid based methods. Lee and Yum(1998) have empirically shown the superiority of post-adjustment forecasts using neural networks.

To improve the Korean gasoline forecasts, this study adopts an architecture that is denoted as the model Main(RNN) + Judgmental(NN). In this architecture, the main trend is forecasted by a RNN(Recurrent Neural Network) model. Judgmental effects are estimated by the NN(Neural Network) model as the multi-layered perceptrons, and the effect is additively adjusted. To validate the proposed architecture, we compared the performance with the RNN model and ARIMA model.

For the empirical tests, 190 gasoline data points for January 1984 - October 1999 were collected. 168 data points for January 1984 - December 1997 were used for training. The remaining 22 data points for January 1998 - October 1999 were used for prediction. We have forecasted for two time points, one-month ahead for the operational schedule and three-month ahead for the gasoline purchase planning. To evaluate the performance of the forecasting, Mean Square Error(MSE), Mean Absolute Error(MAE) and Mean Absolute Percentage Error(MAPE) are adopted as measures of accuracy. The RNN model and the NN model have adopted the feedforward architecture with the backpropagation algorithm.

The remaining sections of this paper are organized as follows. In Section 2, the judgmental adjustment process is formally defined. In Section 3, the main trend of Korean gasoline demand is analyzed using recurrent neural networks. In Section 4, the procedures of the NN based adjustment are explained step by step with the Korean gasoline case. In Section 5, comparative performance evaluations are described. Section 6 concludes this paper.

2. Judgmental adjustment process

Judgmental factor is defined as a factor that cannot be fully incorporated into the time series models, and thus cannot be effectively identified by the extrapolation of past patterns in the data set(Lee et al. 1990). That is, the judgmental factors can be regarded as the factors whose effects cannot be subsumed by the past data series. The concept of judgmental factor is comparable with the notation of causality of Granger(1969). According to his notation, although a variable may affect a dependent variable, if the effect is subsumed by the past data series, the causality cannot be explicitly grasped. In this sense, the judgmental factors can be regarded as the factors whose effects cannot be subsumed by the past data series.

Common characteristics of judgmental factors are:

1. The factors occur irregularly and infrequently. Thus the number of data point is usually too small for statistical modeling.
2. Nevertheless, the impact is too significant to neglect.
3. The effect is transient.
4. The occurrence of coming events can be recognized in advance although it is not easy to judge their impact precisely.

To formalize the adjustment process, the following notational form is useful:

The traditional time series process can be defined as Eq. (1):

$$\hat{Y}(t) = f(Y(t-1),\ Y(t-2),\ \cdots) \qquad (1)$$

where $Y(t-k)$ = actual data for the period $t-k$ (k is a time lag), and $\hat{Y}(t)$ = forecast for the period t. In contrast, the judgmentally adjusted times series process can be formally defined as Eq. (2) and (3):

$$\hat{Y}(t) = f(Y(t-1),\ Y(t-2),\ \cdots) \text{ and } \qquad (2)$$

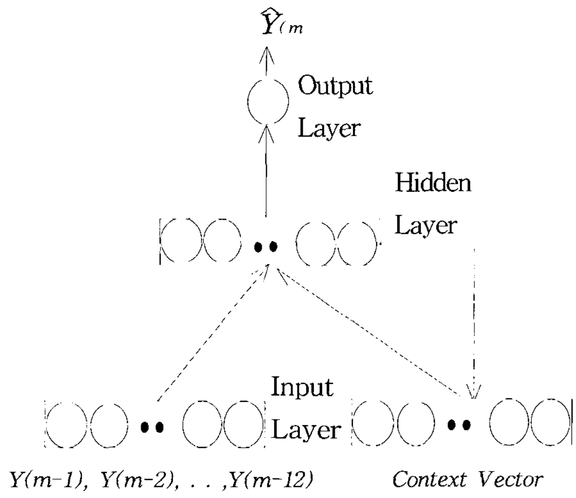
$$\hat{Y}_a(t) = \hat{Y}(t) + \triangle \hat{Y}(t) \qquad (3)$$

where $\triangle \hat{Y}(t)$ = estimated magni- tude of judgmental adjustment for the period t, and $\hat{Y}_a(t)$= adjusted fore- cast for the period t.

If we can construct a huge forecasting model where all the relevant variables are included, there will be no need to consider judgmental factors. The following reasons prevent us from constructing such a huge forecasting model(Oh, 1990).

1. Number of data points: we cannot have enough data points of some relevant variables for statistical analysis. In practice, most organizations have data banks which do not permit even univariate time series modeling.
2. Time span of data points: some variables have

different time span of data points from those of others
3. Data availability: we cannot obtain the data which are required to generate a corresponding model forecast. Also, a lot of data is missing since a file system is a widely used means to manipulate data.
4. Difficulty in formulating mathematical form of social phenomena: some social phenomena cannot be analyzed by mathematical models.
5. Impossibility or inefficiency of identification and estimation of huge model: Huge model cannot always be identified and estimated.



<Figure 1> Architecture of a RNN by Elman

3. Main trend forecasting using the recurrent neural networks

To analyze the time series gasoline main trend, we adopted a RNN suggested by Elman(1988). Fig. 1 shows the RNN with a context vector. The activations in the hidden nodes at time t-1 are copied into the context vector which is the input to the network for time t. This is equivalent to having the hidden nodes completely and recurrently connected, and backpropagating one step in time along the recurrent connections. Therefore the reaction of the network to the new input is a function of both the new input and the preceding context. What is stored in the context vector at any given time is a compressed trace of all preceding inputs, and this compressed trace influences the manner in which the network reacts to each succeeding input. The above characteristic of the context vector agrees with the basic assumption of time series analysis and requires smaller input variables that result in a small network and thus reduce the overfitting problem.

We designed several RNN architectures to determine the best one. As input variables, we selected three input variables (demands at *m-1*, *m-2* to reflect recency, *m-12* to reflect the seasonal cycle) along with the optional ones from *m-3*, *m-4*, *m-5*, *m-6*, and *m-7*. For the hidden nodes, we again selected even numbers of hidden nodes between 1 and $2n + 1$ where n is the number of input variables, according to the Hecht-Nielsen theorem(Fausett, 1994).

We trained all of the recurrent neural network for 500, 1000, 1500, and 2000 epochs with 112 out of 168 data points. Table 1 illustrates the sample data set for training. The trained recurrent neural network models were tested with the remaining 56 data points. Among the best models, the RNN with 6 input variables for *m-1*, *m-2*, *m-3*, *m-4*, *m-5*, *m-12*, 5 hidden nodes, 1 output node, and 1500 epochs was

<Table 1> Sample data set for training (Unit: thousand barrels)

Pattern	Inputs	Output
	Data set at (m-12,m-5,m-4,m-3, m-2, m-1)	*Data at m*
Pattern 85	*(1981, 2743, 2579, 2575, 2638, 2619)*	*2609*
Pattern 86	*(1800, 2579, 2575, 2638, 2619, 2609)*	*2479*
Pattern 87	*(2119, 2575, 2638, 2619, 2609, 2479)*	*2706*
Pattern 88	*(2323, 2638, 2619, 2609, 2479, 2706)*	*2799*
Pattern 89	*(2372, 2619, 2609, 2479, 2706 2799)*	*2900*
Pattern 90	*(2408, 2609, 2479, 2706 2799, 2900)*	*3088*
Pattern 91	*(2557, 2479, 2706 2799, 2900 3088)*	*2951*
Pattern 92	*(2743, 2706 2799, 2900 3088, 2951)*	*3205*

selected as the main trend model because it had the minimum MSE. The model is retrained with the 168 data points from the whole learning period. This trained model was used as Main(RNN) model to predict the main trend forecasts. The MSE, MAE and MAPE for one-month ahead forecasting were 11780, 93.7 and 2.80 respectively, while 14095, 111.3 and 3.16 were the values for three-months ahead forecasting.

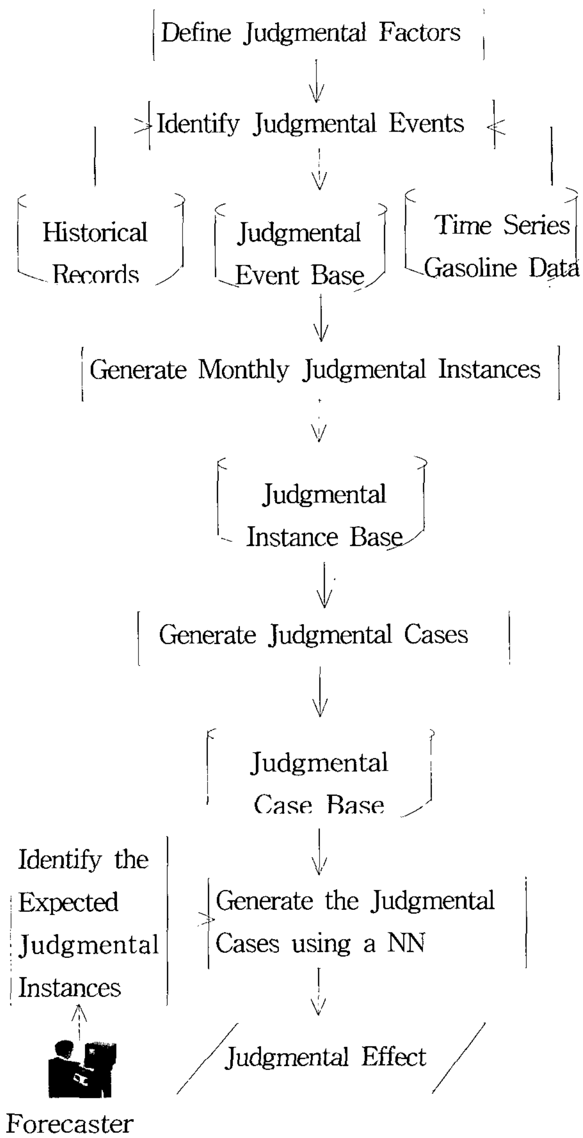
4. Estimation of judgmental effects using the neural network

In this section, the procedure for estimating judgmental effects is proposed. The overall procedure is depicted in Fig. 2. Key steps in the procedure are the following:

1. Definition of judgmental factors and identification of historical judgmental events.
2. Generation of judgmental monthly instances and cases.
3. Generalization of judgmental cases and Estimation of the judgmental effect.

4.1 Definition of judgmental factors and identification of historical judgmental events

The judgmental factors are specified according to the forecasting domain. The judgmental factors for the Korean gasoline domain are the following: governmental energy saving regulations, special holidays, strikes of energy intensive industries. Attributes of these judgmental factors are explained in Table 2. The governmental energy saving regulations is denoted as the $R_m(k, i, d)$, where m = monthly time tag of the occurring month, k = regulation type, i = intensity of regulation, d = impacted days. The special holidays is denoted as the $H_m(k, d)$, where k = holiday type, d = impacted days. The strikes of energy intensive industries is denoted as the $S_m(k, i, d)$, where k = industry type, i = intensity of strike, d = impacted days.



<Figure 2> Procedure for estimating judgmental effects

The realization of judgmental factors is judgmental

<Table 2> Attributes of judgmental factors

Judgmental factors		Attributes of judgmental factors
Name	Notation	
Regulation	$R_m(k, i, d)$	k = regulation type, i = intensity of regulation, d = impacted days
Holiday	$H_m(k, d)$	k = holiday type, d = impacted days
Strike	$S_m(k, i, d)$	k = industry type, i = intensity of strike, d = impacted days

events. Each judgmental factor has multiple judgmental events. An judgmental event of the governmental energy saving regulations, $R_{57}(1, 0.2, 23)$ at September 1988 (57th month from January 1984) is the restriction of driving either odd or even number plated cars within Seoul during the Olympic game. This kind of regulation is classified as type *1*. The intensity *(0.2)* is measured by multiplying the percentage of restricted cars (0.5) by the fraction of car holding regions in the country (0.4), and the duration is *23* days from September 15 to October 7, 1988. In September 1988, a special holiday judgmental event $H_{57}(2, 1)$ is occurred whose type = *2* represents Thanksgiving Day. These judgmental historical judgmental events can usually be identified by examining their historical records and the outliers historical of data.

4.2 Generation of monthly judgmental instances and cases

A judgmental event may affect the demand more than one month, while the effect should be evaluated for each month. So the judgmental event needs to be split into relevant monthly judgmental instances. Since the judgmental event $R_{57}(1, 0.2, 23)$ is overlaid over the 57th month and the 58th month, it should be partitioned into two monthly judgmental instances: $r_{57}(1, 0.2, 16)$ and $r_{58}(1, 0.2, 7)$. Since the impact lasted only one day, $H_{57}(2, 1) = h_{57}(2, 1)$. Since $r_{57}(1, 0.2, 16)$ had happened simultaneously with $h_{57}(2, 1)$ on September 1988, the impact -106 is measured as the composite effect of two instances. The impact 106 computed by delineating the difference between the forecast by the main trend model and its actual value on September 1988. Additionally, the impact can be computed by an external formula. For instance, if a forecaster could obtain the information necessary to compute the impact of car regulation (the total number of cars affected and average daily gasoline consumption per car), she or he could compute the expected amount of decreased consumption quite precisely by the following formula:

Effect = number of affected cars * average daily gasoline consumption per car * impacted duration.

The judgmental case is the pair of occurrences of judgmental monthly events and their impact. Table 3 shows the sample of judgmental cases. The judgmental case on September 1988 (judgmental case 20) is described as $\{r_{57}(1, 0.2, 16), h\ r_{57}(2,1), s_{57}(0, 0, 0), -106\}$ in the first row of Table 3 that is showed sample judgmental cases. We have identified 61 judgmental cases for the training period from January 1984 to December 1997.

<Table 3> Sample judgmental cases

Case number	Judgmental case
Case 20	$\{r_{57}(1, 0.2, 16), h_{57}(2, 1), s_{57}(0, 0, 0), -106\}$
Case 21	$\{r_{58}(1, 0.2, 7), h_{57}(0, 0), s_{57}(0, 0, 0), -80\}$
Case 22	$\{h_{62}(1, 2), h_{57}(0, 0), s_{57}(0, 0, 0), 64\}$

4.3 Generalization of the judgmental cases and estimation of the judgmental effect

The judgmental cases can be generalized by adopting a NN model. Since the data points of the historical judgmental adjustments are usually not sufficient for statistical analysis, the NN is more suitable(Connor et al. 1994; Jhee and Lee, 1993, Lu et al. 1993). The input of a NN model corresponds to

<Table 4> Performances of Main(RNN) + Judgmental(NN), Main(RNN), and ARIMA

Forecasting time	Criteria	Main(RNN) + Judgmental(NN)	Main(RNN)	ARIMA	*Effect of Judgmental(NN)*
One-month ahead	MSE	11780	14359	16103	-2579 (-17.9%)
	MAE	93.7	109.1	121.2	-15.4 (-14.1%)
	MAPE	2.80	3.18	3.21	-0.38 (-11.9%)
Three-months ahead	MSE	14095	16469	27281	-3274 (-14.4%)
	MAE	111.3	128.6	145.8	-17.3 (-13.4%)
	MAPE	3.16	3.69	4.06	-0.53 (-14.3%)

<Table 5> p-Values for t-test

Model	Forecasting time	Criteria	Main(RNN)	ARIMA
Main(RNN) + Judgmental(NN)	One-month ahead	MSE	0.14	0.02**
		MAE	0.21	0.01**
		MAPE	0.17	0.01**
	Three-months ahead	MSE	0.13	0.03**
		MAE	0.18	0.01**
		MAPE	0.10*	0.02**

[**: 5% significant level, *: 10% significant level.]

the judgmental instances along with a monthly time tag, and the output corresponds to the monthly impact. For generalizing the judgmental cases, their best NN model Judgmental(NN) should be constructed according to the NN construction procedure with backpropagation algorithm.

A forecaster should identify the expected monthly judgmental instances to estimate their impacts on the prediction periods. Upcoming judgmental instances can usually be identified through internal/external information such as news, government announcements and management policies before the event occurs. The effect of upcoming monthly judgmental instances can be estimated simply by inputting monthly judgmental instances and their monthly time tag into the generalized neural network Judgmental(NN) for their judgmental instances. Such judgmental effects are added to the time series forecasts computed by the Main(RNN) model.

5. Comparative Performance Evaluation

We evaluated the performances of Main(RNN) + Judgmental(NN) model, Main(RNN) model and ARIMA model on the gasoline demand. The performances of three models are summarized in Table 4. We intended to test the following hypothesis.

Hypothesis 1: Judgmental adjustment can enhance forecasting accuracy.

This hypothesis can be validated by the pairwise t-test of "Main(RNN) + Judgmental(NN) model and Main(RNN) models".

As we can see from Table 5, p-values for MSE, MAE and MAPE are 0.14, 0.21 and 0.17 respectively for one-month ahead forecasting, 0.13, 0.18 and 0.10 respectively for three-month ahead forecasting. According to this result, we cannot accept the Hypothesis 1 with 10% of significance, but we could find strong evidence that the judgmental adjustment is beneficial by all error measures. The 6th column of table 4 shows the effects of Judgmental(NN). We can see that the forecasts after the adjustment have less errors in all aspects of error measurement. Note the reduced MSE by the additive judgment are 17.9% and 14.4% in one- and three-months ahead prediction models respectively. This implies that judgmental adjustment can contribute to performance enhancement of the Main(RNN).

Hypothesis 2: The Main(RNN) + Judgmental(NN) model outperforms the ARIMA model. This hypothesis can be validated by the pairwise t-test of "Main(RNN) + Judgmental (NN) model and the ARIMA model. The identified model ARIMA with logarithmic value of the gasoline data was ARIMA(6,1,1)*$(0,1,1)_{12}$.

According to the result in Table 5, the hypothesis 2 can be accepted with a significant of 0.05 for all criteria for two forecasting time points. Note the p-values for MSE, MAE and MAPE are 0.02, 0.01 and 0.01 respectively for one-month ahead forecasting, 0.03, 0.01 and 0.02 respectively for three-months ahead forecasting. Therefore, it turned out that the Main(RNN) + Judgmental(NN) model outperforms the ARIMA model.

6. Conclusion

To improve the Korean gasoline forecasts, we proposed an architecture with judgmental adjustment process. In the architecture, the main trend is forecasted by a recurrent neural network model, while judgmental effects are estimated by the neural network model, and the judgmental effect is additively adjusted. We have seen that the architecture outperforms the ARIMA model. We have also seen that the effect of the judgmental adjustments to reflect the irregular factors is most critical in building reliable gasoline forecasting systems.

This approach might be effective for many other industries as long as there exist some explicit irregular patterns, which is the case in most real world problems.

References

[1] Connor J., Martin, D., Atlas L.; "Recurrent Neural Networks and Robust Time Series Prediction", IEEE Transactions on Neural Networks, 5(2): 240-254, 1994.

[2] Edmundson, R.H., Lawrence, M., O'Connor, M.J.; "The Use of Non-Time Series Information in Sales Forecasting: A Case Study", Journal of Forecasting, 7: 201-211, 1988.

[3] Elman, J. L.; *Finding Structure in Time*, technical report 8801, Center for Research on Language, University of California, San Diego, 1988.

[4] Fausett L.; *Fundamentals of Neural Networks*, Prentice-Hall., 1994.

[5] Flores, B.E., Olson, D.L., Wolfe, C.; "Judgmental Adjustment of Forecasts: A Comparison of Methods", International Journal of Forecasting, 7: 421-433, 1991.

[6] Georgoff, D.M., Murdick, R.G.; "Manager's Guide to Forecasting", Harvard Business Review, 1986.

[7] Granger C.W.J.; "Investigating causal relations by econometric and cross - spectral models" Econometrica, 37: 1969.

[8] Jhee, W.C., Lee, Jae K.; "Performance of Neural Networks in Managerial Forecasting, Intelligent Systems in Accounting", Finance and Management, 2: 55-71, 1993.

[9] Lee, Jae K., Oh, S.B., Shin, J.C.; "UNIK-FCST: Knowledge-assisted Adjustment of Statistical Forecasts", Expert Systems with Applications: An International Journal, 1(1): 39-49, 1990.

[10] Lee, Jae K., Yum, C.S.; "Judgmental Adjustment Time Series Forecasting Using Neural Networks", Decision Support Systems, 22(1): 135-154, 1998.

[11] Lu, C.N., Wu, N.T., Vemuri, S.; "Neural Network Based Short Term Load Forecasting", IEEE Transactions on Power Systems, 8(1): 336-342, 1993.

[12] Oh, S.B.; "An Intelligent Decision Support System for Forecasting Time Series Data: Expert System Approach", Ph.D dissertation, KAIST, 1990.

[13] Wolfe, C., Flores, B.; "Judgmental Adjustment of Earnings Forecasts", Journal of Forecasting, 9: 389-405, 1990.