AN ARTIFICIAL NEURAL NETWORKS APPROACH TO FORECAST SHORT-TERM RAILWAY PASSENGER DEMAND

Tsung-Hsien TSAI
Ph. D. candidate
Department of Transportation and
Communication Management Science
National Cheng Kung University
1, University Rd., Tainan,
70101 Taiwan, R.O.C.
Fax: +886-6-275-3882
E-mail: thtsai@mail.tcm.ncku.edu.tw

Chi-Kang LEE
Professor
Department of Transportation and
Communication Management Science
National Cheng Kung University
1, University Rd., Tainan,
70107 Taiwan, R.O.C.
Fax: +886-6-275-3882
E-mail: cklee@mail.ncku.edu.tw

Chien-Hung WEI
Professor
Department of Transportation and
Communication Management Science
National Cheng Kung University
1, University Rd., Tainan,
70101 Taiwan, R.O.C.
Fax: +886-6-275-3882
E-mail: louisewei@mail.ncku.edu.tw

ABSTRACT

This paper experiences a three-phrase back-propagation neural network approach to forecast short-term railway passenger demand. The first phase involves the selection of variables, the size of training data set, and the modification of stochastic outliers, under a specific origin-destination (O/D) pair of a given train service. In the second phase, in order to verify the robustness of developed approach, we construct two aggregated models, in which each model applies different temporal aggregations of demand. In the third phase, we construct three integrated models by considering multiple train services simultaneously to enhance the future application. The approach shows encouraging results and most forecasting performances are under 20% of mean absolute percentage error (MAPE). In addition, the approach is able to forecast railway passenger demand effectively under various scenarios of train services. The outcomes of the models can offer detailed demand prediction for railway operation planning, such as train scheduling and seat allocations.

Key Words: Short-term Forecasting, Railway Passenger demand, Back-propagation Neural Network.

1. INTRODUCTION

Short-term forecasting is important for dealing with the daily operational problem in many practical fields. (Davis, G. A. et al., 1990; Smith and Demetsky, 1994; Dougherty and Cobbett, 1997; Ledoux, 1997) construct respective short-term traffic forecasting models to help traffic control centers adopt appropriate strategies for avoiding traffic congestion. (Sun, X. et al., 1997) constructs a passenger-forecasting model to implement seat allocation and yield management for maximizing revenues. (Peng, T. M. et al., 1992; Charytoniuk and Chen, 2000) construct load-forecasting models to predict the hourly load demand for establishing the power supply policy. (Atiya, A. et al., 1997) conducts a monthly flow-forecasting model to predict the potential damage of flood for establishing the agriculture water supply policy. All of these models suggest the essentiality and importance of short-term demand forecasting when considering short-term operational plans. For railway companies, if an effective demand forecasting model is available, it is helpful to conduct appropriate strategies for coping with daily operational problems.

Railway passenger demand is dynamic over time and space, and its pattern is hard to capture. Each origin-destination (O/D) pair has its own daily, weekly, and seasonally demand trait. It is obviously not a static phenomenon. However, most daily operational activities are carried out based on the static demand information. Taking seat allocation as an example, it is pre-set and inflexible, no matter what the day of week is, and no matter whether the peak season it is in or not, the number of seat allocation for a specific O/D pair of a given train service is always the same. Train services and seat allocations yield the major product for Taiwan Railway Administration (TRA), but they are apparently not maneuvered very well. The load factor of the most popular train service of TRA in 2002 is nearly 10 % less than the last year, and it may be worse after the commercial operation of Taiwan high speed rail system in two years time. In order to understand passengers demand pattern and help operators conduct a passenger-oriented planning, a short-term railway passenger demand forecasting model is essential. However, a railway company may have hundreds of trains in service every day, and many O/D pairs for each train trip. It is definitely not easy to predict the demand of many O/D pairs at the same time, or integrating them into a model.

The autoregressive integrated moving average (ARIMA) model is typically applied to cope with short-term forecasting problems (Moorthy and Ratcliffe, 1988; Tang, Z. et al., 1991). Kalman filtering theory (Okutani, 1984) and nonparametric regression (Davis and Nihan, 1991), are applied to solve the problem as well. In recent decades, artificial neural networks (ANN) is widely applied and attains promising results. Among all ANN structures, Back-propagation neural network (BPN) is the most popular, and has been used to solve many transportation problems (Dougherty, 1995). BPN stresses the non-linear mapping between dependent variables and independent variables (Zurada, 1995), and has advantages of parallel computing, multiple outputs generating, and adaptability to complex situations, even without any assumption of data distribution. In this paper, we focus on BPN technique to formulate short-term railway passenger demand forecasting models. Three phrases have been experienced to achieve the goal.

The first phrase is the construction of the approach and involves three steps: selection of variables, adoption of training samples, and modification of outliers. In the short-term forecasting, cross section data is usually unavailable, and time series data is usually applied to formulate forecasting models. The selection of input variables is essential because the more variables applied in a forecasting model, the less efficiency for model learning and predicting, and the less potential for future applications. In this paper, we select influential variables through the autoregressive function (ACF) to reach efficiency and maintain performances. Another reason of inefficiently training by network is the adoption of redundant training samples. It is helpful to filter the useless training samples without sacrificing performances. After these two experiments, we modify some stochastic and unexplained outliers in the series to enhance the performances. In the second phase, a peak-morning aggregated model and a daily aggregated model are implemented to verify the robustness of the proposed BPN approach. The performances show the proposed approach also functions well. In the third phase, an integrated network structure is constructed to predict multiple train services at the same time. The overall results confidently show the ability of the proposed approach to deal with the complex demand problem in the railway transportation.

2. PROCEDURE OF BACK-PROPAGATION NEURAL NETWORK

The basic element of BPN is a node or processing element. Processing elements defined by the researching problem constitute layers. Processing elements in different layers connected to one and another constitute the whole network structure. The networks used in the study involve three layers of nodes, and they are the input layer, the hidden layer, and the output layer. The input layer has one or more nodes for each independent variable. The output layer has one or more nodes that correspond to the predicted values. The hidden layer consists of several nodes that receive inputs from nodes in the input layer and feed their outputs to nodes in the output layer. The number of nodes in the hidden layer in the study is simply the arithmetic average of the number of nodes in the input layer and that in the output layer. Each link between nodes has an associated weight, which represents the strength of the connection. Three steps identify the operation of a node, including receiving input signals and connection weights, summing up the information, and transforming it by the activation function to produce an output signal. The activation function used in the study is the widely used sigmoid function. In the training process, a large number of examples dated before the predicting period is required. Using the current weights, BPN computes a set of outputs with the example input data. These network outputs are then compared against their corresponding values in the example set by computing the sum of square error. After that, the weights are updated by a partial derivative function, which propagates the errors back to the input layer. If the operation in the learning process is successful, the global error reduces gradually, and the process eventually reaches a convergent result. The whole procedure can be shown as Figure 1.

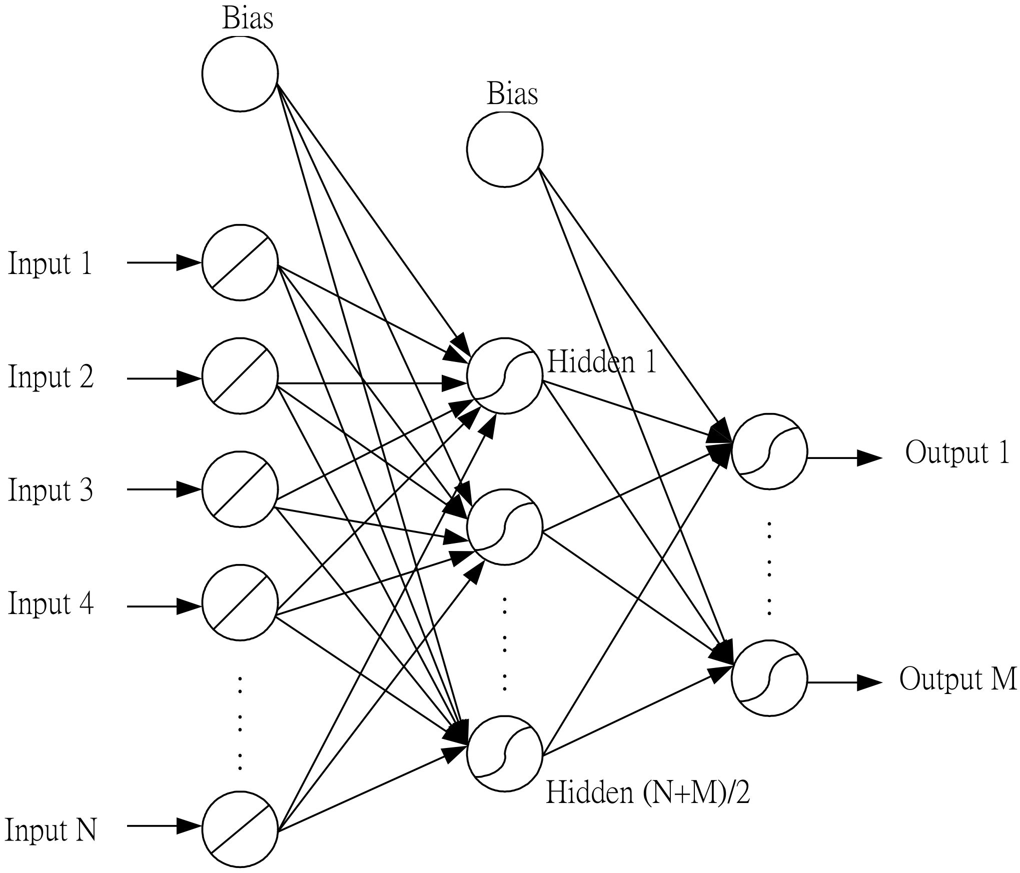

Figure 1: The procedure of back-propagation neural network

Mean Absolute Percentage Error (MAPE), usually applied in evaluating forecasting error, is used in the testing phase as the measure for building practical ANN models. The formula is shown in equation (1), where $av(i,k)$ and $pv(i,k)$ are the actual value and predicting value of example i for output neuron k respectively. The measure is computed with the testing data set, that contains examples in the predicting period. According to [Lewis, 1982], if MAPE is within 20%, then the model is categorized as good forecasting.

$$MAPE = \frac{1}{M*S} * \sum_{k=1}^{M} \sum_{i=1}^{S} \left| \frac{av(i,k) - pv(i,k)}{av(i,k)} \right| * 100\% \qquad (1)$$

3. DATA PROPERTY

All analyses reported in the paper are conducted with ticket sales data of TRA on the western line of inter-city railway system. The length of western line is 406 kilometers, and it is located on the mostly developed area with 95% population of Taiwan. There are 20 stations of A class, and many other stations of B and C classes. Major train services on the western line consist of limited express, express, ordinary train and commuter train. They are denoted as train service A, B, C, and D. Train A in general stops at some stations of A class, train B stops at some stations of A and B classes, and train C stops at most stations on the line. The total number of trains per day per direction on the western line is about 60, in which train A is 27%, train B 20%, train C 10%, and train D 43%. From the initial station to its final station, the total operating time of train A, is about 4 to 5 hours; that of train B is about 5.5 to 6.5 hours; that of train C is about 6 to 7 hours.

There are at least three types of difficulty to deal with the data and make the short-term passenger demand forecasting accurate. First, the time series pattern of passenger volumes for a specific O/D pair and various train service is nonlinear. For example, a time series data, shown in Figure 2, is the passenger volumes from Kaohsiung to Taipei in 1999, for train A departing at 8:00, train B departing at 10:00, and train C departing at 9:00. In addition, the chaotic variations happened especially at holidays and seasonal vacations. Second, the time series patterns of different O/D pairs of a given train service are rarely similar. For example, Figure 3 reveals three O/D pairs with long, medium and short distance. Third, passenger behavior is changeable even for the same train service and the same O/D pair. For example, Figure 4 shows the passenger volumes from Kaohsiung to Taipei in 1999 and 2000, for train A departing at 8:00. The casual relationships among these patterns are difficult to understand. Hence, it brings significant difficulties to build a good integrated model representing some patterns or several passenger volume variables simultaneously.

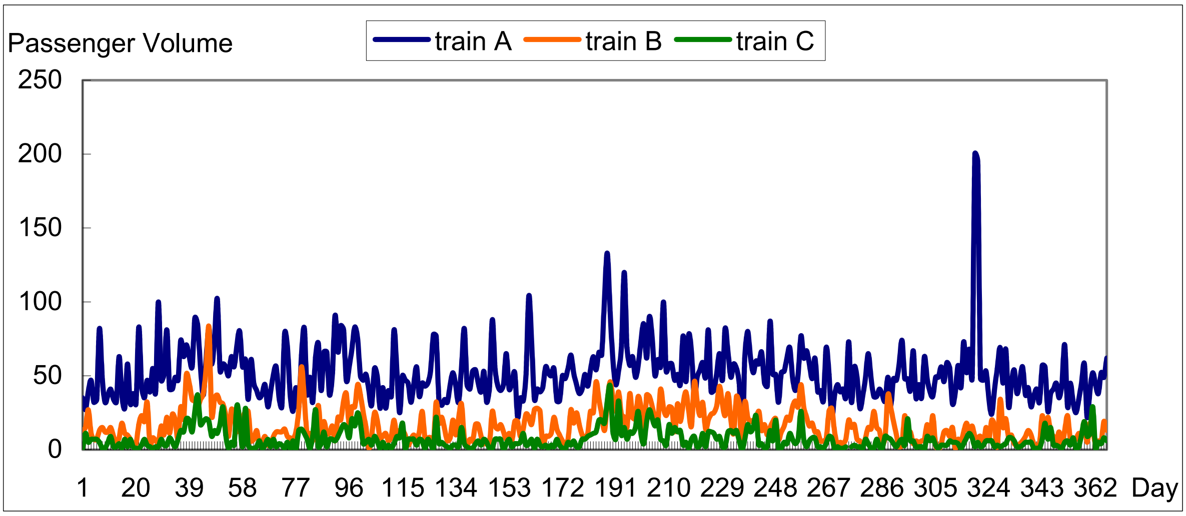

Figure 2. Passenger volumes from Kaohsiung to Taipei of different train services

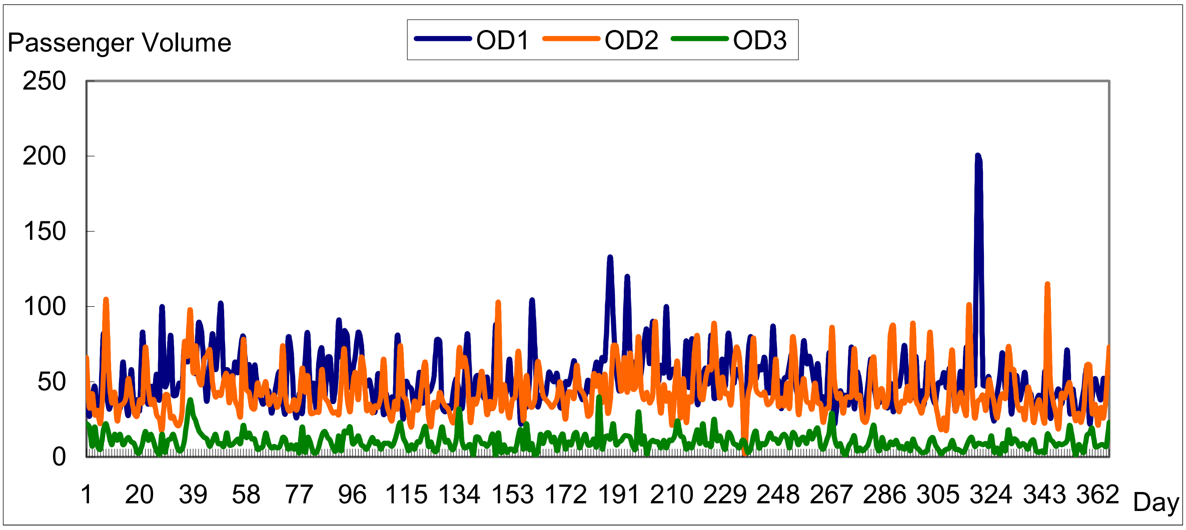

Figure 3. Passenger volumes with various O/D pairs of train A

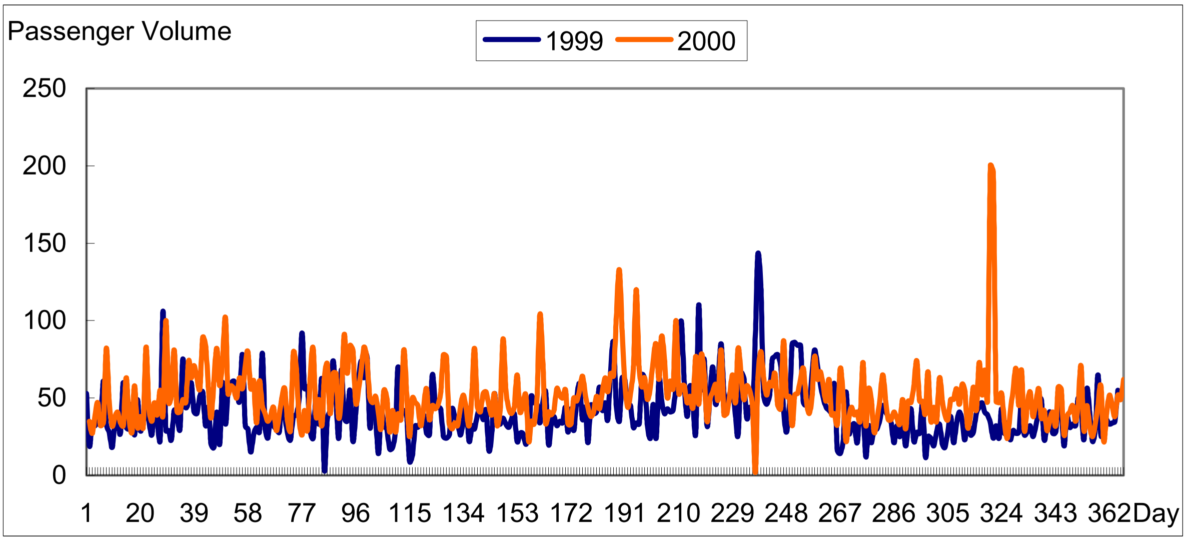

Figure 4. Passenger volumes with the same service and the same O/D pair in two years

4. BACK-PROPAGATION NEURAL NETWORK APPROACH

This section presents an approach, which can be applied to forecast the short-term passenger demand of TRA. A trial and error process is essential for the BPN model to examine the forecasting performance. The basic model structure has been first established as a basis for developing aggregated models and integrated models.

4.1 Basic Model Structure

Short-term passenger demand is a function of time and space. In order to focus on the influence of time, we start from a specific O/D pair for a given train class A. The basic model structure is established based on input variable selection and the size of training data set. The training data set is the passenger volumes of train class A departing at 8:00 A. M. from Kaohsiung to Taipei during 12 months in 1999, and the testing data set is the extension of the same data series during January 2000.

4.1.1 Selection of Input Variables

Using a parsimonious variable set can increase the learning efficiency without sacrificing the performance. This paper applies the autocorrelation function (ACF), as given in equation (2), to select the influential variables. Four scenarios of one-year, half-year, one-quarter and one-month samples dated before the testing data set, are examined for testing the stability of the parsimonious variable set. Its major purpose is to find a proper range for presenting the trait of passenger demand. According to ACF values and priori knowledge, influential variables are chosen as the parsimonious variable set. Table 1 shows the parsimonious variable set and the corresponding performance under these four scenarios. We find that the more number of samples is applied to calculate ACF values, the more reasonable result and good forecasting performances are accomplished.

$$\rho = \frac{Cov(X_{ij}^{cn}(t), X_{ij}^{cn}(t+k))}{\sqrt{Var(X_{ij}^{cn}(t))}\sqrt{Var(X_{ij}^{cn}(t+k))}} \tag{2}$$

where $X_{ij}^{cn}(t)$ are the passenger volumes from origin i to destination j by train class c and run n at time t, and k is the time span.

Table 1: Parsimonious variable sets of different scenarios and corresponding performances

	One-year	Half-year	One-quarter	One-month
ACF High	$x_{ij}^{cn}(t-1)$	$x_{ij}^{cn}(t-1)$	$x_{ij}^{cn}(t-9)$	$x_{ij}^{cn}(t-26)$
	$x_{ij}^{cn}(t-14)$	$x_{ij}^{cn}(t-14)$	$x_{ij}^{cn}(t-1)$	$x_{ij}^{cn}(t-29)$
	$x_{ij}^{cn}(t-21)$	$x_{ij}^{cn}(t-9)$	$x_{ij}^{cn}(t-14)$	$x_{ij}^{cn}(t-23)$
Low	$x_{ij}^{cn}(t-35)$	$x_{ij}^{cn}(t-7)$	$x_{ij}^{cn}(t-23)$	$x_{ij}^{cn}(t-22)$
	$x_{ij}^{cn}(t-7)$	$x_{ij}^{cn}(t-15)$	$x_{ij}^{cn}(t-7)$	$x_{ij}^{cn}(t-18)$
MAPE (%)	17.2	17.1	17.6	20.6

4.1.2 The Size of Training Data Set

Using the right and moderate number of training samples for network learning is another way to improve the efficiency. To identify the appropriate size of training data set, we reduce the number of one-month samples from the farthest month one after another. Table 2 shows the performance of applying previous month data dated before the predicting period is as good as that of applying longer training samples.

Table 2: Performances of different number of training samples (Unit: one month)

	12s	11s	10s	9s	8s	7s	6s	5s	4s	3s	2s	1
MAPE(%)	17.2	17.8	17.9	19.2	18.6	19.3	19.6	18.5	16.4	17.2	14.4	15

4.1.3 Model Extension

In order to reassure the basic model structure, we apply the approach to forecast all 12 months in 2000 consecutively. The whole process is becoming a dynamic rolling state, known as the moving window data learning method [Peng, T. M. et al., 1990; Shimodaira, 1996]. In addition, we introduce two dummy variables. One represents seasonal vacations and the other festivals, in order to enhance the effect of peak dates. Table 3 shows the parsimonious variable set of each month, and the corresponding forecasting performance in 2000. Although the parsimonious variable sets are plausible, several performances of predicting months are not good. In the following steps, two improvements are applied to enhance the forecasting performances.

Table 3: Parsimonious variable sets and performances for each month in 2000

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$x_{ij}^{cn}(t-1)$	V	V	V	V	V	V	V	V	V	V	V	V
$x_{ij}^{cn}(t-7)$	V	V	V	V	V	V	V	V	V	V	V	V
$x_{ij}^{cn}(t-14)$	V	V	V	V	V	V	V	V	V	V	V	V
$x_{ij}^{cn}(t-15)$	V	V										
$x_{ij}^{cn}(t-21)$	V	V	V	V	V	V	V					
$x_{ij}^{cn}(t-28)$	V	V	V	V	V	V	V	V	V	V	V	V
$x_{ij}^{cn}(t-35)$	V	V	V	V	V	V	V		V	V	V	V
Peak Date	V	V	V	V	V	V	V	V	V	V	V	V
MAPE(%)	15.8	27.4	58.0	24.5	19.5	23.3	28.9	205	28.7	18.7	23.3	42

4.2 Improvements of Basic Model Structure

Checking the distribution of actual value and network output, we find that outliers are the major source of distorting the forecasting performances. However, we cannot find any appropriate interpretation for these outliers. In order to detect and modify outliers, we set up varied upper and lower bound based on different periods of seasonal vacations. According to the frequency distribution of passenger volumes in 1999 and 2000, we roughly observe that the distribution of passenger volumes is similar to the normal distribution. Consequently, we make the mean value plus double standard deviation as the upper bound, and that minus double standard deviation as the lower bound within each period. The average of remaining passenger volumes (Chen, H. et al., 1999) based on the day of week within each period is used to replace the outliers. Figure 5 exhibits the detection of outliers within each period.

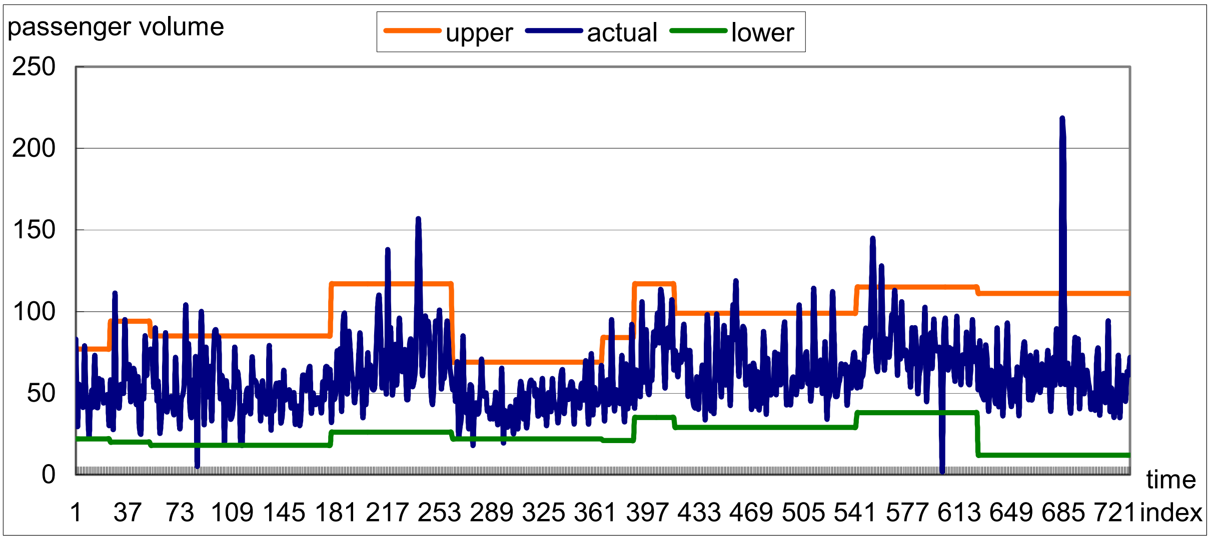

Figure 5: The detection of extreme values

Although we applied dummy variables to represent the peak dates in the basic model, some months with worse forecasting performances almost appear before or after the seasonal vacations. The most possible reason is that the information within the training data set is not enough to represent the variation of seasonal vacations and festivals. Consequently, we use a month, the same as the predicting month of one year ago, into the training data set.

Table 4 shows the result after implementing these two improvements. We can find that all predicting months achieve better performances. In addition, in order to reassure the influence of the size of training data set, we implement one-year training data set to identify the forecasting performances of all months again. However, there is no further significant improvement. Figure 6 shows the proposed approach based on the above experiments.

Table 4: Performances after implementing two improvements

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
MAPE(%)	16.9	22.2	36.1	19.0	15.7	15.6	14.6	17.9	18	19	18.4	23

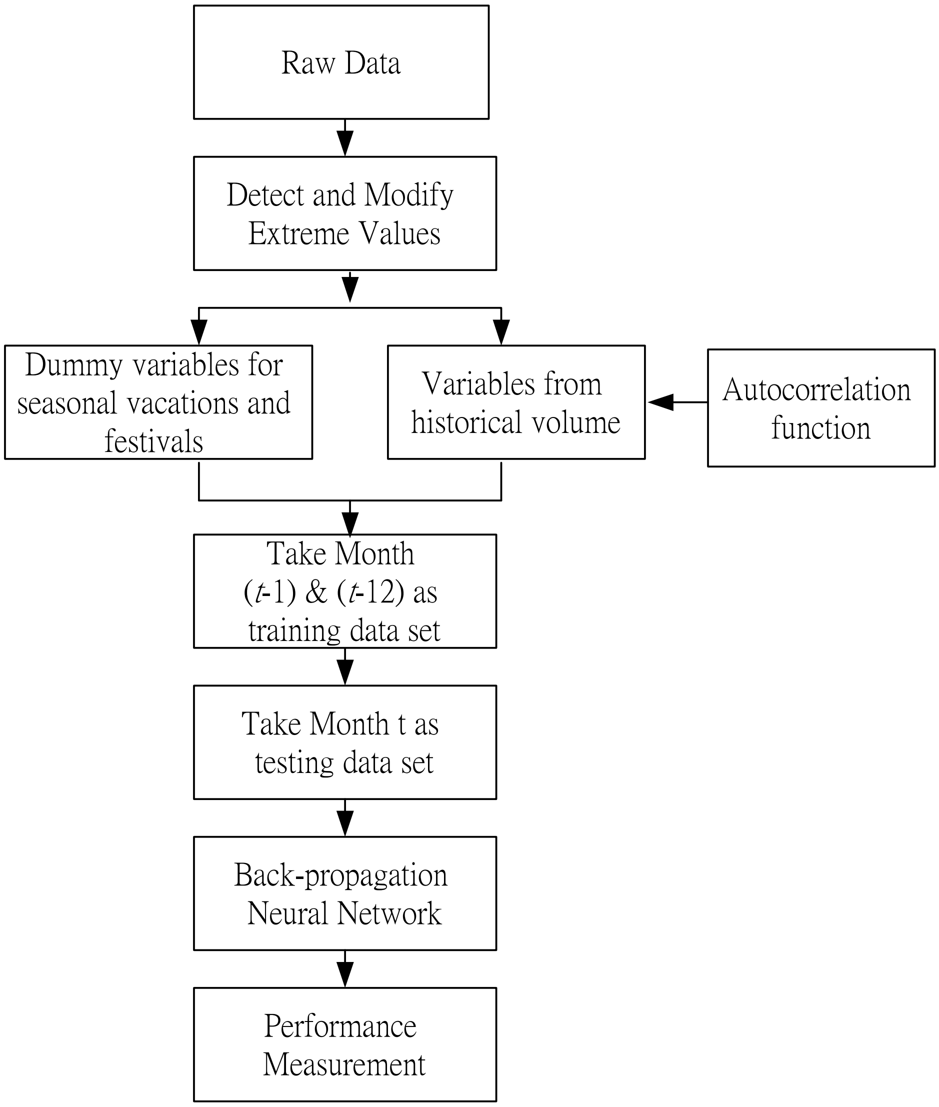

Figure 6: The approach for the short-term railway passenger forecasting

5. AGGREGATED AND INTEGRATED MODELS

Based on the proposed approach, we implement two aggregated models, one is the peak-morning aggregated model and the other is the daily aggregated model to verify the capability of proposed approach. We also construct three integrated models considering multiple train services to enhance the future applications.

5.1 Aggregated Model

In reality, passengers of the same O/D pair compete for their ideal schedule service, involving the train class and the departure time. If ideal schedule service is not available, the most likely situation that reflects in the ticket sales data is the switch of passenger volumes to other train services. Through aggregation, the chaos of data can be reduced.

Table 5 and Table 6 exhibit the respective parsimonious variable sets and corresponding performances of aggregated models. In most cases, the longer temporal aggregated data is applied, the better performance is achieved.

Table 5: Parsimonious variable sets and performances of peak-morning aggregated models

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$x_{ij}^{cn}(t-1)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-7)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-14)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-21)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-28)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-35)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
Peak Date	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
MAPE(%)	13.9	17.8	26.3	17.8	12.8	13.6	13.2	12.0	16.8	15.9	10.8	15.3

Table 6: Parsimonious variable sets and performances of daily aggregated models

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$x_{ij}^{cn}(t-1)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-3)$			∨	∨	∨	∨	∨	∨				
$x_{ij}^{cn}(t-7)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-10)$				∨	∨	∨						
$x_{ij}^{cn}(t-14)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-21)$				∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-28)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
$x_{ij}^{cn}(t-32)$			∨									
$x_{ij}^{cn}(t-35)$	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
Peak Date	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨	∨
MAPE(%)	9.7	23.3	16.3	15.4	12.6	14	8.9	10.6	14	12.9	11	11.8

5.2 Integrated Model

The realistic train services of TRA constitute multiple train classes and many O/D pairs. Three integrated models are constructed to integrate multiple train services. The first model is the schedule service of multiple O/D pairs to integrate the difference of O/D pairs. The second model is multiple schedules of the same train class and the same O/D pair to integrate the difference of departure time. The third model is multiple schedules of the various train classes and the same O/D pair to integrate the difference of train classes. Figure 7 shows the concept of integrated models and Table 7 exhibits the performance of June 2000, as an example. We find that the integrated models still perform well. This phenomenon shows the potential of the proposed BPN approach for further service integration.

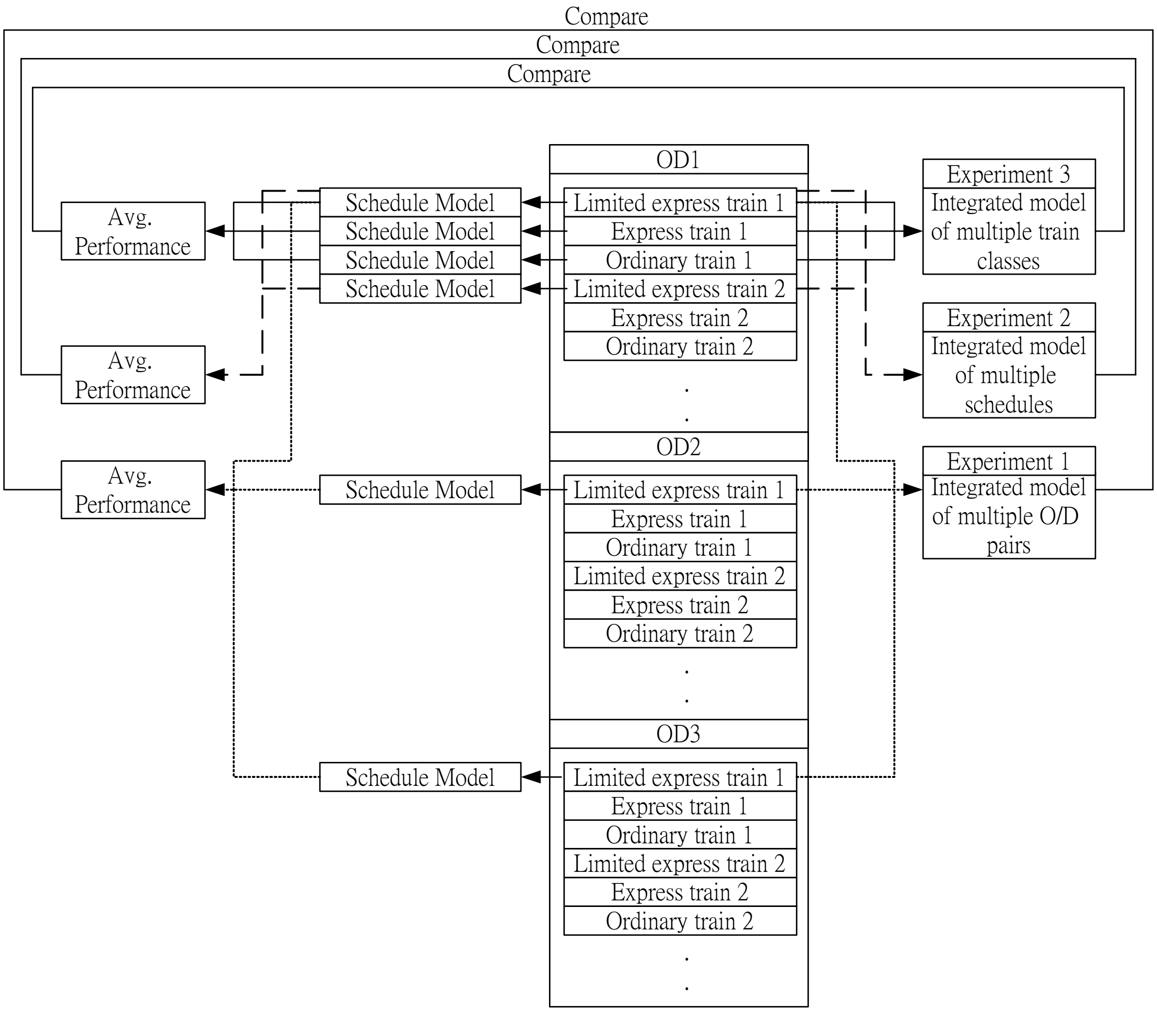

Figure 7: The process of integrated models

Table 7: Performances of integrated models

Experiment 1	O/D1model	O/D2 model	O/D3 model	Average	Integrated model
MAPE (%)	15.6	22.3	12.1	16.7	17.5

Experiment 2	LE 1 model	LE 2 model	LE 3 model	Average	Integrated model
MAPE (%)	19.5	15.6	19.1	18.1	18.0

LE: Limited Express

Experiment 3	LE model	E model	O model	Average	Integrated model
MAPE (%)	12	20.3	20.2	17.5	19.4

LE: Limited Express E: Express O: Ordinary Train

6. CONCLUSION

Short-term forecasting has been discussing widely in the traffic flow-forecasting and load-forecasting. However, it is seldom addressed in the railway transportation. An effective short-term passenger-forecasting model is essential for daily operation and planning of railway transportation. The description in the data property indicates passenger demand is varied over time and space. This is the major obstacle to achieve effective forecasting performance. Throughout the experiments, promising performances show the potential of BPN to forecast the short-term railway passenger demand. The proposed approach also offers a guideline to extract useful information from large data set. Most performances of basic models are under 20% of MAPE and aggregated models show better forecasting performances than basic models. Integrated models show the possibility of considering multiple train services simultaneously, thus reinforcing the practicability in solving real operation problems. However, this approach still has two restrictions. First, due to the data filtering before the training, the model cannot explain the effect of special events. The approach is suitable for the general situation. Second, due to the use of variable $x_{ij}^{cn}(t-1)$, we just can do the prediction one day before the departure unless we rid this variable. However, this may make the prediction less effective.

ACKNOWLEDGEMENTS

The authors are grateful to National Science Council for the financial support of NSC91-2211-E006-116.

REFERENCES

Van Arem, B., Kirby, H. R., Van Der Vlist, M. J. M. and Whittaker, J. C. (1997) Recent advances and applications in the field of short-term traffic forecasting, **International Journal of Forecasting, Vol. 13, No. 1,** 1-12.

Atiya, A, El-Shoura, S., Shaheen, S. and El-Sherif, M. (1997) Application of neural networks to the problem of forecasting the flow of the river NILE, **Neural Network for Signal Processing VII, Proceeding of the 1997 IEEE Workshop,** 598-606.

Charytoniuk, W. and Chen, M. S. (2000) Neural network design for short-term load forecasting, **International Conference on Electric Utility Deregulation and**

Restructuring and Power Technologies 2000, 554-561.

Chen, H., Mussone, L., Montgomery, F. and Susan, G. M. (1999) The effect of missing data on neural network performance in forecasting flow. In Z. Yang, C. P. Wang and B. Mao (eds.), **Traffic and Transportation Studies.** American Society of Civil Engineers, Reston.

Davis, G. A., Nihan, N. L., Hamed, M. M., and Jacobson, L. N. (1990) Adaptive forecasting of freeway traffic congestion, **Transportation Research Record 1287**, 29-33.
Davis, G. A. and Nihan, N. L. (1991) Nonparametric regression and short-term freeway traffic forecasting, **Journal of Transportation Engineering, Vol. 117, No. 2,** 178-188.

Dougherty, M. (1995) A review of neural networks applied to transport, **Transportation Research Part C, Vol. 3, No. 4,** 247-260.

Dougherty, M. S. and Cobbett, M. R. (1997) Short-term inter-urban traffic forecasts using neural networks, **International Journal of Forecasting, Vol. 13, No. 1,** 21-31.

Drezga, I. and Rahman, S. (1998) Input variable selection for Ann-based short-term load forecasting, **IEEE Transactions on Power Systems, Vol. 13, No. 4,** 1238-1244.

Lachtermacher, G. and Fuller, J. D. (1995) Backpropagation in time-series forecasting, **Journal of Forecasting, Vol. 14,** 381-393.

Ledoux, C. (1997) An urban traffic flow model integrating neural networks, **Transportation Research Part C, Vol. 5, No. 5,** 287-300.

Lewis, C.D. (1982) **Industrial and Business Forecasting Method**, Butterworth Scientific, London.

Kaastra, I. and Boyd, M. (1996) Designing a neural network for forecasting financial and economic time series, **Neurocomputing, Vol. 10, No. 3,** 215-236.

Moorthy, C. K. and Ratcliffe, B. G. (1988) Short term traffic forecasting using time series methods, **Transportation Planning and Technology, Vol. 12**, 45-56.

Peng, T.M., Hubele, N. F. and Karady, G. G. (1992) Advancement in the application of neural networks for short-term load forecasting, **Transaction on power systems, Vol. 7, No. 1,** 250-257.

Okutani, I. (1984) Dynamic prediction of traffic volume through Kalman filtering theory, **Transportation Research Part B, Vol. 18, No. 1,** 1-11.

Shimodaira, H. (1996) A method for selecting similar learning data in the prediction of time series using neural networks, **Expert Systems With Applications, Vol. 10, No. 3/4,** 429-434.

Smith, B. L and Demetsky, M. J. (1994) Short-term traffic flow prediction: Neural network approach, **Transportation Research Record 1453**, 98-104.

Sun, X., Brauner, E. and Hormby, S. (1997) A large scale neural network for airline forecasting in revenue management. In G. Yu (eds.), **Operations Research in the Airline Industry.** Kluwer Academic Publishers, Boston.

Tang, Z., Almeida, C. D. and Fishwick, P. A. (1991) Time series forecasting using neural networks vs. Box-Jenkins methodology, **Simulation, Vol. 57, No. 5,** 303-310.

Wei, C. H. (2001) Analysis of artificial neural networks models for freeway ramp metering control, **Artificial Intelligence in Engineering, Vol. 15, No. 3,** 241-252.

Wei, W. S. (1994) **Time Series Analysis.** Addison-Wesley, California.

Yun, S.Y., Namkoong, S., Rho, J.H., Shin, S.W. and Choi, J.U. (1998) A performance evaluation of neural network models in traffic volume forecasting, **Mathematic Computing Modelling, Vol. 27, No.9-11,** 293-310.

Zhu, Z. and Yang, Z.(1999), Dynamic prediction of traffic flow using Backpropagation neural network, In Z. Yang, C. P. Wang and B. Mao (eds.), **Traffic and Transportation Studies.** American Society of Civil Engineers, Reston.

Zurada, J.M. (1995) **Introduction to artificial neural systems.** West, St. Paul.