High-Speed Railway Passenger Flow Volume Prediction Model Based on Factor-Response Simulation Method

Du Xuedong, Shang Rui and Gao Ziyu
Engineering Center, College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao 266510, China

Abstract: A new high-speed railway Passenger Flow Volume (PFV) prediction model was proposed according to the factor-response simulation method, and the percentage of the five factors' influence. The five factors were confirmed from many factors that affect PFV. It can be concluded that the proposed prediction model was precise and practical through an experimental test.

Key words: Factor-response simulation, mean absolute percentage error, consumption function, regional economic conditions

INTRODUCTION

Now-a-days, people's trip generation rate has grown rapidly along with the improvement of people's income and values (Shu et al., 2007). A good prediction model of PFV is helpful for the National Railway Administration to draw up an effective railway operation plan to achieve the greatest economic benefit.

However, PFV is affected by a lot of nonlinear factors in real life which increase the complexity of PFV predictive analysis tremendously (Deurreni and Phnn, 1993; Wang et al., 2002). Considering the importance of PFV prediction, many people try to create or use all kinds of nonlinear factors to upgrade the prediction of PFV. Traditional prediction techniques includes multiple linear regression, stochastic time series, general exponential smoothing, state space and Kalman filter (Davis and Nihan, 1991; Moghram and Rahman, 1989; Okutami and Stephanedes, 1984). Traditional prediction techniques make it hard for the researchers to test. Recently, someone try to find the similarities between some popular networks models and transportation, they come up with prediction models like Neural Tree model (NT), Artificial Neural Networks (ANN), Back-Propagation Neural Network (BPN) (The effect of tannic acid on in vitro gas production and rumen fermentation of sunflower meal (Dougherty, 1995; Zurada, 1995; Lan et al., 2008; Lan, 2009; Zafar et al., 2006) and they have acquired remarkable achievement. Especially for the popular BPN which has been successfully used in many transportation fields. Nevertheless, these researchers all try to match dynamic and changeable PFV with fixed existing models which make it passive and hard to predict PFV precisely.

Five factors and their influence percentage on PFV are discussed in this paper and a new PFV prediction model (Averill and Kelton, 2000) is proposed, at the same time some correction is made on the proposed PFV prediction model. Through the test of the prediction model on Qingdao's PFV data, the Mean Absolute Percentage Error (MAPE) of the model is under 20% totally. Therefore, the PFV prediction model is persuasive and practical (Tsung-Hsien and Chi-Kang, 2003; Lewis, 1982, Yu et al., 2008, Steven and Keh, 2011).

PFV INFLUENCE FACTOR-RESPONSE SIMULATION

Factor-response simulation method is a technique that spread out as many interrelated factors as possible and makes use of the factorial design method to get the influence percentage of every factor. Due to limited space, Table 1 only shows the selected factors and their properties that have a greater impact to PFV.

In Table 1, the factors are classified into three levels A, B and C by the deciding of whether the factor is qualitative or quantitative and controllable. The influence A, B and C reduces by turns. It is not hard to

Corresponding Author: Du Xuedong, Engineering Center, College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao 266510, China
Tel: +86 0532-86057622 Fax: 86-532-86057622

1761
Fig. 1: Relationship between factors and PFV

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>External factor</td>
<td>National railway policies</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>B</td>
<td>Speed raised</td>
</tr>
<tr>
<td>Economic development level</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>B</td>
<td>Stabilization</td>
<td></td>
</tr>
<tr>
<td>Income level</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>B</td>
<td>Increase steadily</td>
<td></td>
</tr>
<tr>
<td>Consumption rate</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>A</td>
<td>Increase steadily</td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>A</td>
<td>Increase rapidly</td>
<td></td>
</tr>
<tr>
<td>Traveling fever</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>B</td>
<td>Flatten</td>
<td></td>
</tr>
<tr>
<td>Internal factor</td>
<td>Ticket price</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>A</td>
<td>Appropriately increase</td>
</tr>
<tr>
<td>Time (speed)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>C</td>
<td>Speed increase</td>
<td></td>
</tr>
<tr>
<td>Ticket-selling service</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>C</td>
<td>Quality improved</td>
<td></td>
</tr>
<tr>
<td>Safety</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>A</td>
<td>Higher safety factor</td>
<td></td>
</tr>
<tr>
<td>Comfort</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>B</td>
<td>More comfortable</td>
<td></td>
</tr>
<tr>
<td>Vehicle maintenance</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>C</td>
<td>In time</td>
<td></td>
</tr>
<tr>
<td>Traffic density</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>C</td>
<td>Increase</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Factor design matrix

<table>
<thead>
<tr>
<th>Factor combination</th>
<th>$E_{1}$</th>
<th>$E_{2}$</th>
<th>$E_{3}$</th>
<th>$E_{4}$</th>
<th>$E_{5}$</th>
<th>$E_{6}$</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{1}$</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{2}$</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{3}$</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{4}$</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{5}$</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{6}$</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{7}$</td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$R_{8}$</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>$R_{9}$</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>$R_{10}$</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>$R_{11}$</td>
</tr>
<tr>
<td>12</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>$R_{12}$</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>$R_{13}$</td>
</tr>
<tr>
<td>14</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>$R_{14}$</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>$R_{15}$</td>
</tr>
<tr>
<td>16</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>$R_{16}$</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>$R_{17}$</td>
</tr>
<tr>
<td>18</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>$R_{18}$</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>$R_{19}$</td>
</tr>
<tr>
<td>20</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>$R_{20}$</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>$R_{21}$</td>
</tr>
<tr>
<td>22</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>$R_{22}$</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>$R_{23}$</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>$R_{24}$</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{25}$</td>
</tr>
<tr>
<td>26</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{26}$</td>
</tr>
<tr>
<td>27</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{27}$</td>
</tr>
<tr>
<td>28</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{28}$</td>
</tr>
<tr>
<td>29</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{29}$</td>
</tr>
<tr>
<td>30</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{30}$</td>
</tr>
<tr>
<td>31</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{31}$</td>
</tr>
<tr>
<td>32</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$R_{32}$</td>
</tr>
</tbody>
</table>

see from the Table 1 that there are six main factors: Economic development level, geographic locations, population density, ticket price, time and safety. The economic development level and different geographic locations are replaced as one factor that is regional economic conditions. Therefore, the factors become regional economic conditions ($F_{ec}$), population density ($F_{pd}$), ticket price ($F_{tp}$), time ($F_{tm}$) and safety ($F_{sa}$). Factor-design matrix based on $2^{k}$ (k is the number of factors, here it equals to 5) will be given below to analyze the influence from factors to PFV while $R_{i}$ indicates the average PFV of every month (Table 2), here + means the relevant factor has a positive affect on the PFV and means the relevant factor has a negative affect on the PFV.

The relationships between factors and PFV can be easily seen in Fig. 1 and the influence percentages of every factor can be calculated as follows:

$$R_{i} = \frac{\sum_{t=1}^{12}(m_{it})}{5}$$

$e_{it}$ indicates the influence of factor $F_{i}$ to average PFV of every month, their expressions are:
Table 3: Parameters table

<table>
<thead>
<tr>
<th>Factor</th>
</tr>
</thead>
</table>
| F      | ±  
| α      |  
| β      |  

Table 4: Coding chart for factors

| C      | Autonomous consumption parameter  
| β      | Marginal propensity to consume parameter (β ∈ [0,1])  
| μ      | Constant  
| t      | Period  
| Y      | Total income of t period  

\[
\begin{align*}
\sigma &= \frac{\sum_{i=1}^{n} \left( R_{i} - R_{i-1} \right)}{16} \\
\beta &= \frac{\sum_{i=1}^{n} \left( R_{i} + R_{i-1} - R_{i-2} - R_{i-3} \right)}{16} \\
\gamma &= \frac{\sum_{i=1}^{n} \left( R_{i} + R_{i-1} - R_{i+1} - R_{i+2} \right)}{16} \\
\epsilon &= \frac{\sum_{i=1}^{n} \left( R_{i} - R_{i-1} + R_{i-2} - R_{i-3} \right)}{16} \\
\omega &= \frac{\sum_{i=1}^{n} \left( R_{i} - R_{i-1} \right)}{16} \\
\end{align*}
\]

Table 3 shows the meaning of +, - , it means the change of PFV when the value of factor f changes from α to β. It will be easy to predict the PFV when we figure out the change law of f (i). Analysis of 5 factors and their expressions will be given in the next section.

**FACTORS ANALYSIS**

PFV prediction is affected by many complex nonlinear factors. Qingdao’s regional economic conditions, population density, ticket price, time and safety five factors are chosen here for comprehensive analysis and obtain Qingdao’s PFV prediction model.

**REGIONAL ECONOMIC CONDITIONS**

Regional economic is mainly measured by consumption level of the specific region and consumption level can be described by the consumption function. Consumption function is a function that determines consuming behavior and the key indicator that reflect regional economic conditions. Consumption function was first proposed by English economist J.M. Keynes (Keynes, 1936; Davis, 1952).

Hall (1978) introduce rational expectation consumption into consumption function in 1978 and proposed adaptive expectations consumption function model:

\[ C_t = \hat{\lambda} a + (1-\lambda) C_{t-1} + \lambda \hat{Y}_t + \mu, \quad (t=1,2,3,...,T) \tag{3} \]

Table 4 shows the meanings of every parameters in the Eq. 3:

Using generating function method to recursively solve the Eq. 3. Assume H (x) = C_0 + C_1 x + C_2 x^2 + ... , multiply (1-λ) x to, H (x), it becomes (1-λ) x H (x) = (1-λ) x C_0 + (1-λ) C_1 x + (1-λ) C_2 x^2 subtract these two expressions above, extract common factor and divide (1-λ) x both sides of the expression, it becomes:

\[ H(x) = \sum_{i=1}^{\infty} \left( a + \lambda \hat{Y}_i + \mu_i \right) \frac{1 - (1-\lambda)^{i}}{\lambda} x^i, \]

since \( H(x) = \sum_{i=0}^{\infty} C_i x^i \), then:

\[ C_i = \hat{\lambda} a + \lambda \hat{Y}_i + \mu_i \frac{1 - (1-\lambda)^{i}}{\lambda} \quad (i=1,2,3,...,T) \tag{4} \]

**POPULATION DENSITY**

Population density will affect travel number of people. This study use Logistic Curve Model to realize the prediction of Qingdao permanent population:

\[ Y_t = \frac{k}{1 + e^{\beta t}} \quad (t=1,2,3,...,T) \tag{5} \]

where, k, α and β are unknown parameters in the expression, t indicates time. Using modified exponential curve parameters method (Fangjin and Wu, 2008), because:

\[ S_i = \sum_{m=1}^{n_i} Y_i, S_0 = \sum_{m=1}^{n_0} Y_0 \quad \text{and} \quad S_i = \sum_{m=1}^{n_i} Y_i \]

Substitute \( S_i, S_0, S_i \) into Eq. 5, therefore:

\[
\begin{align*}
\beta &= \frac{(S_i - S_0)^{0.5}}{S_i - S_0} \\
\alpha &= \frac{m \beta - 1}{S_0(S_i - S_0)(\beta^m - 1)} \\
\kappa &= \frac{m \beta^m - 1}{S_0 \beta^m - S_i} \\
\end{align*}
\]

Choose \( m = 6 \) and select as one month’s permanent population of Qingdao randomly, they are 9091.99, 8731.22, 8397.81. Substitute values into Eq. 6 we get:

\[
\begin{align*}
\beta &= \frac{8397.81 - 8731.22}{8731.22 - 9091.99} = 0.986 \\
\alpha &= \frac{8397.81 - 8731.22}{9091.99 - (360.77) \times (0.924 - 1)} = 4.1321 \times 10^{-4} \\
\kappa &= \frac{9091.99 \times 0.924 - 8731.22}{6 \times (0.924 - 1)} = 1.38 \times 10^{-3} \\
\end{align*}
\]

1763
Get:

\[ Y = \frac{1.38 \times 10^5}{1 + 4.1321 \times 10^4 \times 0.986} \]  

\[ Y = \frac{1.38 \times 10^5}{1 + 4.1321 \times 10^4 \times 0.986} \]  

Qingdao's permanent population of every year and its prediction (Table 5) and population growth is given below.

**TIME, TICKET PRICE AND SAFETY**

Travel time ($T_t$) and waiting time ($W_t$) are the important factors that affect the selection of means of transportation from the view of passengers (Ruhe and Yan, 2005). Passenger's time-consuming in one travel can be expressed as ($\mu$, $\eta$ are parameters):

\[ T_{\text{sum}} = \mu T + \eta W_t \]  

Passengers' total consuming time $T_{\text{sum}}$ in one transportation has a passive influence on passengers' choose of this means of transportation. There are three types of trains EMU, express and local train in Qingdao railway station. $T_t$ is approximately equal to the Average Travel Time (ATT) of every train. There are altogether 19 EMUs and ATT equals to 4.16 h, 2 express trains with ATT equals to 18.98 h and 22 local trains with ATT equals to 22.54 h. ATT of all the trains $T_t$ equals to 14.25 h. $W_t$ is hard to investigate, we assume $W_t$ to be 0.5 h for the convenience of simulation.

Ticket price is also an important factor that affects PFV. Ticket price varies from travel distance to travel distance. Assuming that there are $\chi$ stations from origin to stop last stop, the distance between the $i$ station and the $i+1$ station is $P_{i, i+1}$, the ticket price between the $i$ station and the $i+1$ station is $\xi_t$, so the average price from station to station is shown as follows:

\[ \xi_{\text{avg}} = \frac{\sum_{i=1}^{\chi} (\chi - 1) \xi_t}{\sum_{i=1}^{\chi} (\chi - 1) \xi_t} \]  

\[ \xi_{\text{avg}} = \frac{\sum_{i=1}^{\chi} (\chi - 1) \xi_t}{\sum_{i=1}^{\chi} (\chi - 1) \xi_t} \]  

$\xi_{\text{avg}}$ is a power function on $t$.

Safety is one of the important factors that affect the passengers' choice of train. It is mostly affected by drivers' work time and the train's speed (Kecklund and Ingre, 2001). $S$ means safety coefficient. $W_t$ is work time, $\mu$ and $\rho$ are variable parameters, $T_t$ is the speed of the train. Safety can be described as:

\[ S = \mu \text{sgn} (W_t - 8) W_t + \rho \]  

\[ S = \mu \text{sgn} (W_t - 8) W_t + \rho \]  

**PFV PREDICTION MODEL AND ERROR CORRECTION**

We can view nonlinear factors into some simple units, therefore, factors affect PFV can be expressed as factors linear superposition. Take Qingdao's PFV as experimental data from 2005 to 2009 (Table 6).

First, $e_{\text{PV}}$, assuming that is the percentage of the influence from regional economic conditions to PFV, therefore:

\[ e_{\text{PV}} = \frac{e_{\text{PV}}}{e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}}} \times 100\% \]  

other factor's influence percentages can be figured out in the same way. PFV therefore, can be expressed as:

\[ e_{\text{PV}} = e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}} + e_{\text{PV}} \times 100\% \]  

Assume that means years and months to be predicted, for example 2020.04 and 2030.12 means April of 2020 and December of 2030. $t$ means months, $t = (T - [T]) \times 100$. Prediction model is a combining form of power function and exponential function on $t$. Substitute every into PFV prediction model, we get PFV of every month of every year, compared them to experimental data, we can see the comparison in Fig. 2 and 3. Figure 2 illustrates that we use the previous prediction model to predict the PFV, then we compared it with the actual PFV, we can see that there are obvious different between these.
two groups of data and in Fig. 3, we use prediction model expressed in expression 17 and we find the prediction data is very similar with the actual data, it means our latter prediction model is much better in prediction.

This simulation is generally in accordance with actual trend but below expectation. When analyze the monthly PFV line chart from 2005 to 2009, we will find PFV tending to ascending trend year by year totally (The x-coordinate represent 60 months of the 5 years). Table 6 in specific, PFV grow rapidly during vacation and drop rapidly after vacation, then come to gradual advance trend. This trend can be used to describe for other year. It means we should deal with the prediction of PFV in vacation and in peacetime separately, including fluctuation. The modified prediction model is shown as follows:

\[ p_{PFV} = (c'^{C}_{30} c^*_0 + c'^{C}_{0} Y_0 + c'^{T}_{m} T_{m} + c'^{I}_{m} I_{m} + c'^{S}_{m} S_{m}) f(t) + \sigma(t) \]  

\[ f(t) \] is a power function on \( t \). According to curve trend of experimental data, we assume that \( \sigma(t) \) equals to:

\[ \sigma(t) = \begin{cases} 
1.25 \times 10^5 \alpha^{T - 2010} \left( \frac{t}{10^3} \right)^4 & (t = 4, 7, 12) \\
4.0 \times 10^5 \beta^{T - 2010} \left( \frac{t}{10^3} \right)^{1.5} & (t = 1, 5, 8, 9, 10) \\
1.0 \times 10^5 \gamma^{T - 2010} \left( \frac{t}{10^3} \right)^{1.0} & (t = 2, 3, 6, 11)
\end{cases} \]

\( \alpha, \beta \) and \( \gamma \) are constant which are used to assure \( \sigma(t) \) increase regularly year by year. According to (Lewis, 1982), if MAPE is lower than 20%, the model can produce a fine prediction.
\[
\text{MAPE} = \frac{1}{M} \sum_{i=1}^{M} \left| \frac{\text{av}(i) - \text{pv}(i)}{\text{av}(i)} \right| \times 100
\]

(15)

\(\text{av}(i)\) and \(\text{pv}(i)\) means actual value and predictive vale, \(M\) means month, that is how many months' experimental data it will use. The over striking values in Table 7 are numbers that fail to make MAPE under 20%. Since all kinds of emergency may affect PFV greatly and values which fail to meet the requirement only take a little part of total values will be deemed to be within tolerance. In summary, the PFV prediction model proposed in this study is effective.

**CONCLUSION**

In this study, a PFV prediction model is proposed by using the method of factor-response simulation and the factorial design method on the factors which affect PFV mostly. We make a test on prediction model by using the experimental data of Qingdao railway station and we find the model's MAPE is under 20% totally. The precise and practical prediction model can help the National Railway Administration to draw up an effective railway operation plan to achieve the greatest economic benefit.

**ACKNOWLEDGMENTS**

Project supported by the State Key Laboratory of Rail Traffic Control and Safety (Contract No. RCS2008K001), Beijing Jiaotong University and National High-Tech Research and Development Program of China (863 Program) grants 2009AA062703, China.

**REFERENCES**


