Evaluation of Car-Following Input Variables and Development of Three-Vehicle Car-Following Models with Artificial Neural Networks

Mitsuru TANAKA a, Nobuhiro UNO b

a Johnson, Mirmiran & Thompson, Herndon, VA, 20171, USA
b Graduate School of Management & Department of Urban Management, Kyoto University, Kyoto, 6158540, Japan
a E-mail: mitsurude@msn.com
b E-mail: uno@trans.kuciv.kyoto-u.ac.jp

Abstract: A four-layer artificial neural network (ANN) structure was set up in the models and a genetic algorithm (GA) and back-propagation methodology were utilized to customize individual driver's behavior. A number of combinations of the input variables was examined with the R2 values representing the model fitting. This paper concluded that there are significant differences in degrees of contribution to the models among the several input variables. The most important finding was that the leading vehicle's (LV) acceleration had stronger relationship to the following vehicle's (FV) acceleration rate rather than the relative speed between the leading vehicle and the following vehicle. Additionally, the input variables related to the preceding vehicle of the leading vehicle (PLV) were added in the model. It was also found the variables of the preceding vehicle of the leading vehicle help car-following models slightly, but they were not as much as expected.

Keywords: Car-Following Model, Artificial Neural Networks (ANN), Back-Propagation, Genetic Algorithm (GA), Various Input Variables, Three-Vehicle Model

1. INTRODUCTION

Various microscopic car-following models have been developed to represent vehicular movements in the microscopic level of traffic flow in the past half century. The car-following models have generally been modeled based on experiments observing variables of two vehicles called a leading vehicle, LV, and a following vehicle, FV. It started with a very simple mathematical model that utilized a linear relationship between spacing between a leader and a follower, and a desirable speed of the follower. Since then, the car-following models have become increasingly complex, with extra input variables and specific predefined rules in newer proposed models. Developing such a complicated car-following model normally requires us to spend enormous effort and sometimes there is difficulty in calibration with the predefined rules. The models and rules are predetermined by the modelers, however they may not accurately represent human driving behavior as there may be some key factors missing from the natural driving behavior in the models. In order to better present driving behavior, it may be more effective to have a car-following model learn the human driving behavior through the car-following traffic variables from the observed data sets. The artificial neural network (ANN) model has such a structure for building an input-output system throughout the learning process from training data sets.

Mathematical models of ANN have been developed from the structure and functional aspects of biological neural networks. ANNs are adaptive systems that change structures based on external or internal information that flows through the network during the learning phase.
Therefore, ANNs have been often used for non-linear statistical data modeling. Car-following behavior is a human behavior and a driver learns from experience utilizing the external information surrounding the driving vehicle. It is expected that the ANN structure has great potential to model car-following behavior. There are some shortcomings of using ANNs and one of them is the long computation time during the training process due to its heavy repeating process in the data learning algorithms. However, in the past several decades, the computer technologies have been significantly improved and computation time for processing complicated computer algorithms has significantly shortened. Thus, the computation time of heavy repeating processes for learning data sets in ANN has significantly reduced as the computer technology has improved. Therefore, ANN model structure was applied for building car-following models in this paper. Multiple ANN car-following models were developed with different input variables and they were compared each other for evaluating the importance of the model input variables among them. The proposed ANN models were also compared to a well-known existing car-following model in order to evaluate how the ANN structures are valid and feasible to apply as car-following models in this study.

2. BACKGROUND

2.1 Car-Following Models

Microscopic car-following models began with a simple linear relationship between driving speed and distance headway with the vehicle running in front of the subject vehicle introduced by Pipes (1953). Forbes and Zagorsk (1958) introduced the same car-following model, but used speed and time headway instead of space headway. After these speed-distance models, researchers at General Motors Company (GM) established a series of car-following models called stimulus-response system. In this car-following system, stimulus is represented by relative speed between a leader and a follower and response is represented by acceleration of the follower as shown in equation (1).

\[
\ddot{x}_n(t+T) = \lambda [\dot{x}_{n-1}(t) - \dot{x}_n(t)]
\]  

where,
\[
\begin{align*}
    t & : \text{time}, \\
    T & : \text{reaction time}, \\
    \ddot{x}_n(t+T) & : \text{acceleration of the } n^{th} \text{ vehicle at time } t+T, \\
    \dot{x}_n(t) & : \text{speed of the } n^{th} \text{ vehicle at time } t, \\
    \lambda & : \text{sensitivity factor}
\end{align*}
\]

GM researchers also developed generalized car-following models by considering that the sensitivity factor in equation (1) can be as simple as a constant value, but also can be a complicated function of headway and speed of the vehicles. This generalized car-following equation is called the GM model shown as equation (2).

\[
\ddot{x}_n(t+T) = \alpha \left[ \frac{\ddot{x}_n(t+T)}{[\dot{x}_{n-1}(t) - x_n(t)]} \right] [\dot{x}_{n-1}(t) - \dot{x}_n(t)]
\]  

where,
\( x_n(t) \) : position of the \( n^{th} \) vehicle at time \( t \),
\( \alpha, l, m \) : model parameters

It is known that equation (2) can represent many of car-following models introduced in past. For example, the equation becomes Pipes’ and Forbes’ models when \( l = 0 \) and \( m = 0 \). In the case where \( l = 1 \) and \( m = 0 \), the equation becomes the GM’s third model, and it becomes the GM’s fourth model with \( l = 1 \) and \( m = 1 \). This well-known GM’s car-following model is also famous for making its connection to the macroscopic traffic stream models.

The GM model has been well-known as representing the stimulus-response system in car-following behavior. Other than the GM models, there are many other existing car-following models with various modeling aspects. Brackstone and McDonald (1999) classified the existing car-following models into several groups; GM Model, Collision Avoidance Model, Linear Model, Psychophysical Action Point Model and Fuzzy Logic Based Model. Other than these model groups, there are widely-used models such as Optimal Velocity Model (Bando et al., 1995) and Intelligent Driver Model (Treiver et al, 2000). In addition to these models, several researchers recently started developing neural network driver models (Chong and Medina, 2011; Matthew and Ravishankar, 2012; Panwai and Dia, 2007; Colombaroni and Fusco, 2011). They are all modeled with representing unique and specific driving behavior aspects. Although there are several candidates for model comparison, the GM model was selected for model comparison in this study due to its long history and name recognition.

3. CAR-FOLLOWING DATA SETS

The car-following data was recorded in a test track at the Civil Engineering Research Institute of Hokkaido, Japan in October 2000. The test track consisted of two 1.2 km straight sections with two 0.3 km semicircular curves. To eliminate any geometry effects from the curves, only the data on the straight sections were used. The location and speed of each vehicle were recorded at 0.1 second intervals throughout the experiment. RTK GPS (Real-Time Kinematic GPS) was equipped on each of the ten test vehicles and the trajectories of the vehicles were recorded. Gurusinghe et al. (2002) examined the accuracy of the experimental data and confirmed these data sets are sufficiently accurate.

Two different types of data were recorded in the experiment: one for car-following conditions without stops and the other for the vehicle movements from start to stop in short distances. The car-following data portion has been used for the studies of several researchers such as Gurusinghe et al. (2002), Ranjitkar et al. (2003; 2004; 2005), and Tanaka et al. (2008). There were a total of ten drivers in the car-following platoon. In this study, the ten drivers were identified as D01 through D10 from head to tail of the platoon. The driver of the first vehicle in the platoon, D01, initiated several speed patterns in the experiment. The data sets for D02 through D10 were used for fitting car-following models. There were a total of 47 data sets consisting of 21 car-following data sets and 26 start & stop data sets used in this analysis. The car-following data sets were set to start at 40-50 km/h, have some speed fluctuations, and end at the same initial speed within the straight sections. The start & stop speed data sets were designed to accelerate up to an expected optimum speed of 40-50 km/h from a stop condition, travel with speed fluctuations, and decelerate to complete stop. The length of the data collection period was 4342.3 seconds with 43,423 data points for each of ten vehicles in these data sets. The reaction time for each driver’s response had to be preset in training for use in calibrating the car-following models in this study. The reaction times for the individual drivers were
referred to the previous study by Ranjitkar et al. (2003). The reaction time approximately ranged between 1.3 to 1.6 seconds among the drivers D02 through D10.

4. METHODOLOGY

As mentioned in the previous paragraph, the raw car-following data included more than forty-two thousand data points for each vehicle. In the back-propagation training process, the widely-spread, enormous data can sometimes cause overlearning and/or divergence problems. Therefore, data sampling and data conversion to fewer points were conducted for a smooth learning process during back-propagation training. Assuming the data points are not fluctuated in a short time interval, simple data sampling was conducted at every 20 points, which corresponds to 2.0 second intervals.

Five types of ANN car-following models were developed with various input variables in this study. Initial synaptic weights in the proposed ANN models were first calibrated with a GA model. After the synaptic weight initialization by GA, the back-propagation process was performed for training the synaptic weights further by learning the sampled data sets at every 2.0 second. The number of input data variables varied from one to five depending on the ANN model.

In order to compare the performance of the ANN car-following models, the GM’s car-following model was calibrated to fit to the same data sets. The GA model was again utilized for finding the best constants in calibrating the GM models. The ANN and GM models were evaluated with the $R^2$ value in order to measure how they fit to the raw data sets.

4.1 Genetic Algorithm (GA)

The Genetic Algorithm (GA) program was used for two purposes in this study: 1) to calibrate the initial weights of the ANN car-following models, and 2) to find the best fit coefficients of GM’s car-following model shown in equation (2). The GA program used in this study was the Genecop III program proposed by Z. Michalewicz (1992). It is characterized by the operation by floating-point number, the flexibility in dealing with constraints and boundaries, and the systematic applications of mutations and crossovers. The GA’s objective function was set to minimize the average of the sum squared error between estimated and observed data values as seen in Equation (3). In determining the initial weights of the ANN models, the population size was set at 200 and the generation was set at 100 in the GA program. These GA parameters were determined after investigating the results of several calibration examples.

$$J(e) = \text{Minimize} \frac{1}{N} \sum_{i=1}^{N} (y_{i,\text{data}} - y_{i,\text{est}})^2$$

where,

- $J(e)$ : objective function,
- $y_{i,\text{data}}$ : measured value for the output variable in $i$th input data,
- $y_{i,\text{est}}$ : estimated value for the output variable in $i$th input data,
- $N$ : number of data points
Figure 1. The Mutilayer ANN Car-Following Model Structures
Table 1. Input and Output Variables for the ANN and GM Car-Following Models

<table>
<thead>
<tr>
<th>Model</th>
<th># of Input Variables</th>
<th>Input Variables</th>
<th>Output Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN Model-1</td>
<td>Any 1 Input</td>
<td>Relative Speed (t), ( \dot{x}_{n-1}(t) - \dot{x}_n(t) )</td>
<td>FV Acceleration (t+T), ( \ddot{x}_n(t+T) )</td>
</tr>
<tr>
<td>ANN Model-2</td>
<td>Combination of any 2 Inputs</td>
<td>Spacing (t), ( x_{n-1}(t) - x_n(t) - L_n )</td>
<td></td>
</tr>
<tr>
<td>ANN Model-3</td>
<td>Combination of any 3 Inputs</td>
<td>FV Speed (t), ( \dot{x}_n(t) )</td>
<td></td>
</tr>
<tr>
<td>ANN Model-4</td>
<td>Combination of any 4 Inputs</td>
<td>FV Acceleration (t), ( \ddot{x}_n(t) )</td>
<td></td>
</tr>
<tr>
<td>ANN Model 5</td>
<td>All 5 Inputs</td>
<td>LV Acceleration (t), ( \dot{x}_{n-1}(t) )</td>
<td></td>
</tr>
<tr>
<td>GM Model</td>
<td>3 Inputs</td>
<td>Relative Speed (t), ( \dot{x}_{n-1}(t) - \dot{x}_n(t) )</td>
<td>FV Acceleration (t+T), ( \ddot{x}_n(t+T) )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Headway (t), ( x_{n-1}(t) - x_n(t) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FV Speed (t+T), ( \dot{x}_n(t+T) )</td>
<td></td>
</tr>
</tbody>
</table>

Note: FV denotes a following vehicle (\( n^{th} \) vehicle), LV denotes a leading vehicle (\( (n-1)^{th} \) vehicle), \( x_n(t) \) is the position of the \( n^{th} \) vehicle at time \( t \), \( T \) is a reaction time of the \( n^{th} \) vehicle, and \( L_n \) is the vehicle length of the \( n^{th} \) vehicle.

4.2 Multilayer Artificial Neural Network (ANN) Models

The main structure of the ANN models was developed for training the car-following data sets with back-propagation methodology. Figure 1 shows the structures of the multilayer ANN car-following models used in this study. All of the multilayer ANN models consist of four layers: an input layer, two intermediate hidden layers, and an output layer. Each layer has one or more neurons which are connected to the neurons on adjacent layers with some connection strengths called synaptic weights. The normalized input variable, between 0 and 1, was entered into the input layer. The input signals were transmitted in sequence from the input layer to the output layer while neural operations were repeated. The output layer produces the normalized objective variable. The forward signal process is completed with this computation sequence from the input layer to the output layer.

After the forward signal process, the synaptic weights were adjusted so that the error between the output signals and the target signals was minimized. The back-propagation method (Wasserman, 1989) produces the adjustments of synaptic weights in each layer. The synaptic weights are adjusted by the momentum method to smooth the adjustment and urge the convergence. The forward and back-propagation processes were repeated until the error between the target and the output was reduced within a tolerance limit or until a preset number of iterations was reached. The preset number of iterations was set at 5000 in the calibrations of all ANN models. The learning coefficient, one of the key parameters of back-propagation process, was set at 0.5 in all ANN models.

Five ANN models with different input variables were developed in this study. They are identified as ANN Model-1, ANN Model-2, ANN Model-3, ANN Model-4, and ANN Model-5 throughout this work. The descriptions of the layer and node structure as well as the input and output variables for these ANN car-following models are listed in Table 1.
5. ANALYSIS RESULTS

Since the back-propagation network sometimes causes a problem in the learning of numerous, widely-spread data points, the effort of reducing the number of training data points was first attempted with the data set for one driver. The data was simply sampled at every 20 raw data points. Assuming the car-following data points are continuous and that there are no sudden changes in a short time period, it was expected that the sampled points with this simple sampling methodology can represent the raw data very. The ANN Models were trained through the back-propagation methodology with the synaptic weight initialization by GA. In the training process, the average error percentage was significantly reduced by GA at the beginning, then the back-propagation process takes over and starts reducing the error further for fine tuning through many iterations. In this combined error reduction methodology, often a bump occurs in the learning curve and the error increases temporarily before it decreases when the back-propagation process takes over. This is perhaps because the vector of reducing the error in the back-propagation process is not always the same as GA’s error reduction vector.

5.1 The Five ANN Model Results

Five ANN models were trained with the sampled data sets. The ANN models were made for all individual drivers, D02 through D10, except the very first driver, D01, in the car-following platoon. The results depended on each drive’s behavior, so that the minimum, the maximum, and average values among different drivers are shown throughout the analysis.

The first result shown is regarding the ANN Model-1 with a single input variable in the model. Figure 2(a) summarizes the $R^2$ values of the ANN Model-1 for each driver with a different input variable. There are five different input variable for Model-1. They are Spacing between FV and LV, Speed of FV, Acceleration of FV, Relative Speed between FV and LV, and Acceleration of LV. As illustrated in Figure 3, the ANN Model-1 with Spacing between FV and LV and Speed of FV show almost no relationship to the output, Acceleration of FV at time (t+T). The $R^2$ values of these inputs are 0.008 and 0.0025. The input of Acceleration at time (t) shows better relationship to the model output, Acceleration at (t+T). This is because Acceleration is continuous between time t and (t+T) and could not jump significantly within the reaction time T. The minimum and maximum $R^2$ values show 0.484 and 0.697 which spread widely. Compared to the Acceleration input, Relative Speed between FV and LV shows better relationship to the output with a narrower minimum and maximum range as expected. This result reconfirms that Relative Speed has a strong relationship to the model output, Acceleration at time (t+T). However, one surprise result here is that Acceleration Rate of LV at time (t) even shows better $R^2$ value than the one with Relative Speed. Although the difference in the $R^2$ values between these models is a little, this result shows another significantly important variable for a car-following modeling other than Relative Speed. The $R^2$ values for all the nine drivers’ models are steadily high.
Figure 2(a). The $R^2$ Values after Training ANN Model-1

Figure 2(b). The $R^2$ Values after Training ANN Model-2

Note: Spacing denotes spacing between the following vehicle and the leading vehicle, Speed denotes speed of the following vehicle, RS denotes the relative speed between the following vehicle and the leading vehicle, Accel denotes the acceleration rate of the following vehicle, and LV Accel denotes the acceleration rate of the leading vehicle.
Figure 2(c). The $R^2$ Values after Training ANN Model-3

Figure 2(d). The $R^2$ Values after Training ANN Model-4 & 5

Note: Spacing denotes spacing between the following vehicle and the leading vehicle, Speed denotes speed of the following vehicle, RS denotes the relative speed between the following vehicle and the leading vehicle, Accel denotes the acceleration between the following vehicle and the leading vehicle, and LV Accel denotes the acceleration rate of the leading vehicle.
Figure 2(b) summarizes the $R^2$ values of the ANN Model-2. The ANN Model-2 are with two input variables and all combinations of two out of the five input variables shown in Table 1 are used as the model inputs. As expected from the Model-1 results, the combination of Spacing and Speed show very low $R^2$ values and they implicate that they have almost no direct relationship to the output value by themselves. However, once one of other input variables of Acceleration, Relative Speed, and LV Acceleration is involved, Spacing and Speed contribute to improve the $R^2$ values. It looks that LV Acceleration is becoming the base of generating high $R^2$ values among the input combinations as well as Relative Speed. The highest average $R^2$ value seen in the ANN Model-2 was 0.805 with a combination of Relative Speed and LV Acceleration. The $R^2$ value range among different drivers is also very narrow between 0.778 and 0.832. This result implies that Relative Speed and LV Acceleration are significantly important inputs as supported by the ANN Model-1 results.

The $R^2$ value results of the ANN Model-3 are shown in Figure 2(c). The ANN Model-3 are with three input variables and all combinations of three out of the five input variables shown in Table 1 are used as the model inputs. Similar to the ANN Model-2, as expected from the Model-1 results, the combination of Relative Speed, Acceleration, and LV Acceleration generated the best $R^2$ values among any three input combinations. The highest $R^2$ value for the ANN Model-3 was 0.853 with Relative Speed, Acceleration, and LV Acceleration improving 0.048 from the ANN Model-2. It is also seen in the figure that the worst three combinations of the Model-3 include the combination of Spacing and Speed.

<table>
<thead>
<tr>
<th>Model</th>
<th># of Input Variables</th>
<th>Input Variables</th>
<th>Output Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN Model-6</td>
<td>Common 5 Inputs</td>
<td>Relative Speed (t), $\dot{x}_{n-1}(t) - \dot{x}_n(t)$</td>
<td>FV Acceleration (t+T), $\dot{x}_n(t+T)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spacing (t), $x_{n-1}(t) - x_n(t) - L_n$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FV Speed (t), $\dot{x}_n(t)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FV Acceleration (t), $\ddot{x}_n(t)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LV Acceleration (t), $\dddot{x}_{n-1}(t)$</td>
<td></td>
</tr>
<tr>
<td>ANN Model-6 With PLV Information</td>
<td></td>
<td>PLV Speed (t), $\dot{x}_{n-2}(t)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Common 5 Inputs +</td>
<td>PLV Spacing (t), $x_{n-2}(t) - x_n(t) - L_n$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Any of 1 PLV Input</td>
<td>PLV Relative Speed (t), $\dot{x}_{n-2}(t) - \dot{x}_n(t)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLV Acceleration (t), $\ddot{x}_{n-2}(t)$</td>
<td></td>
</tr>
<tr>
<td>ANN Model-7 With PLV Information</td>
<td></td>
<td>PLV Relative Speed (t), $\dot{x}_{n-3}(t) - \dot{x}_n(t)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Common 5 Inputs +</td>
<td>PLV Acceleration (t), $\dddot{x}_{n-2}(t)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 PLV Inputs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Input and Output Variables for the ANN Models with the Preceding Vehicle Information (Extra 2 ANN Models)
The $R^2$ value results of the ANN Model-4 and Model-5 were shown together in Figure 2(d). The ANN Model-4 have four input variables and only four combinations available using the five input variables in Table 1. The ANN Model-4 show that, as far as the combination of Relative Speed, Accel, and LV Accel is in, the model shows the best $R^2$ results, which are seen as 0.865 in the two combinations. The ANN Model 5 has only one combination with all the five input variables in Table 1. The $R^2$ value of the ANN Model-5 resulted as 0.866 which is almost the same as the best of ANN Model-4. This implies that the effects of Speed and Spacing may depend on an individual driver. Some may react with Speed only, and some may react with Spacing only, and some may react for both of them. Comparing the $R^2$ results of the best ANN Model-4 to the bests of ANN Model-3, it can be said that at least one of Speed and Spacing affects a driver’s behavior very slightly.

### 5.2 Extra Two ANN Models with the Preceding Vehicle of the Leading Vehicle

Some drivers may be having some effects by not only the leading vehicle but also a vehicle in front of the leading vehicle. The preceding vehicle in front of the leading vehicle is defined as the preceding vehicle of the leading vehicle, PLV, in this paper. Since these ANN models do not require any pre-settings of a formula for the car-following models, extra information of the PLV was plugged in a couple of enhanced ANN models in order to investigate the effect of the preceding vehicle. Table 3 and Figure 3 show the enhanced ANN models with the extra information inputs on the PLV. The models are named as ANN Model-6 and 7. One of the ANN Model-6 actually did not include the information of PLV, but Speed of LV was used as the sixth input variable. This Model-6 was made only for the comparison purpose and the effect
of an extra input was investigated. The leftmost results in Table 3 and Figure 4 show the results of the ANN Model-6 without PLV information. The results show that the sixth input, LV Speed, did not affect the model results as comparing them to the ANN Model-5 results seen in Figure 2(d). Since LV Speed should have been in the model as the sum of Relative Speed and Speed of FV, the effect was expected.

Table 3. The $R^2$ Values after Training ANN Model-6 & 7

<table>
<thead>
<tr>
<th>Driver ID</th>
<th>Model-6 (6 Input Variables)</th>
<th>Model-6 (6 Input Variables) With LLV Information</th>
<th>Model-7 (7 Input Variables) With LLV Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RS Spacing Speed</td>
<td>RS Spacing Speed</td>
<td>RS Spacing Speed</td>
</tr>
<tr>
<td></td>
<td>Accel</td>
<td>Accel</td>
<td>Accel</td>
</tr>
<tr>
<td></td>
<td>LV Accel</td>
<td>LV Accel</td>
<td>LV Accel</td>
</tr>
<tr>
<td></td>
<td>LV Speed</td>
<td>PLV Speed</td>
<td>PLV Speed</td>
</tr>
<tr>
<td>D03</td>
<td>0.881</td>
<td>0.885</td>
<td>0.876</td>
</tr>
<tr>
<td>D04</td>
<td>0.836</td>
<td>0.842</td>
<td>0.832</td>
</tr>
<tr>
<td>D05</td>
<td>0.888</td>
<td>0.813</td>
<td>0.828</td>
</tr>
<tr>
<td>D06</td>
<td>0.898</td>
<td>0.912</td>
<td>0.894</td>
</tr>
<tr>
<td>D07</td>
<td>0.856</td>
<td>0.881</td>
<td>0.854</td>
</tr>
<tr>
<td>D08</td>
<td>0.894</td>
<td>0.855</td>
<td>0.892</td>
</tr>
<tr>
<td>D09</td>
<td>0.850</td>
<td>0.850</td>
<td>0.847</td>
</tr>
<tr>
<td>D10</td>
<td>0.848</td>
<td>0.880</td>
<td>0.867</td>
</tr>
<tr>
<td>Low</td>
<td>0.836</td>
<td>0.813</td>
<td>0.832</td>
</tr>
<tr>
<td>High</td>
<td>0.898</td>
<td>0.912</td>
<td>0.894</td>
</tr>
<tr>
<td>Average</td>
<td>0.869</td>
<td>0.865</td>
<td>0.868</td>
</tr>
</tbody>
</table>

Note: Spacing denotes spacing between the following vehicle and the leading vehicle, Speed denotes speed of the following vehicle, RS denotes the relative speed between the following vehicle and the leading vehicle, Accel denotes the acceleration rate of the following vehicle, LV Accel denotes the acceleration rate of the leading vehicle, LV Speed denotes speed of the leading vehicle, PLV Spacing denotes spacing between the following vehicle and the preceding vehicle of the leading vehicle, PLV Speed denotes speed of the preceding vehicle of the leading vehicle, PLV RS denotes the relative speed between the following vehicle and the preceding vehicle of the leading vehicle, and PLV Accel denotes the acceleration rate of the preceding vehicle of the leading vehicle.

The ANN Model-6 with PLV information includes one of four different inputs related to PLV. The information used related PLV were, PLV Speed, Speed between the following vehicle and PLV, Relative Speed between the following vehicle and the PLV, and PLV Acceleration. The middle columns in Table 3 and Figure 4 show the results of the ANN Model-6 with PLV information. It is hard to say if there is an effect by the PLV information, however, it can be said that the models with PLV RS ad PLV Accel, 0.872 and 0.885 respectively, were somewhat slightly better than the ANN Model-6 without PLV information which was 0.869.
Therefore, one more ANN Model, ANN Model-7, was created including two PLV inputs, PLV RS and PLV Accel. The results of the ANN Model-7 ended with almost no changes from the ANN Model-6 with PLV Accel input. Therefore, this implies that the acceleration information of the PLV is somewhat related to the driving behavior, even though the effect is very small.

Finally, all the ANN models were compared in one chart as seen in Figure 5. GM models were also calibrated by GA for comparison purpose. The average $R^2$ value of the GM models was calculated as 0.707 for all drivers. Among all the ANN Models examined in this paper, it is concluded that the greatest ANN model is either the ANN Model-6 with PLV Accel Input, which $R^2$ value was 0.885, or the ANN-Model-7, which $R^2$ value was 0.883.

Figure 4. The $R^2$ Values after Training ANN Model-6 & 7
6. CONCLUSION

In this study, the total of seven ANN car-following models were developed and input variables were evaluated with the benchmark $R^2$ values. The ANN models were also compared to the existing GM car-following model. All the ANN models were successfully trained using GA and back-propagation methodologies with the sampled data sets.

Some key facts related to the input variables for car-following models were found in this paper. The most important finding was that LV Acceleration had stronger relationship to the model output, the following vehicle’s acceleration, than the relative speed between the following vehicle and the leading vehicle, which has been traditionally said as the most important component of a car-following model. This paper also verified importance of the relative speed to the car-following models. It was also found that the acceleration rate of the following vehicle at time (t) can also contribute to estimate the output, acceleration rate of time (t+T), in the car-following models.

In the comparison among all the examined ANN models, the models with three inputs could have a good car-following model if the three key input variables, Relative Speed, Acceleration, and LV Acceleration, are correctly selected. Extra inputs of Speed and Spacing can improve the model accuracy, but the effect is somewhat limited.

Finally, extra inputs related to PLV were evaluated. The results showed that somewhat slightly positive effect if the PLV’s Acceleration rate is used as an input. The acceleration information of the PLV might somewhat help driver’s acceleration decision, but the influence may be much smaller than expected.
The GM model was also calibrated to the same data sets with the GA technique and the results were compared to the several ANN models. The ANN models were generally resulted in better than GM models if the number of input variables is three or more. This concludes that the ANN car-following models were successfully developed to represent car-following conditions by human drivers.

REFERENCES


