A Comparison of Freeway Work Zone Capacity Prediction Models

Nan Zheng\textsuperscript{a}1*, Andreas Hegyi\textsuperscript{b}, Serge Hoogendoorn\textsuperscript{b}, Henk van Zuylen\textsuperscript{b} and David Peters\textsuperscript{c}

\textsuperscript{a}: École Polytechnique Fédérale de Lausanne, Station 18-GC C2 406, CH-1015 Lausanne, Switzerland. (nan.zheng@epfl.ch)
\textsuperscript{b}: Delft University of Technology, Stevinweg 1, 2628CN, Delft, the Netherlands (a.hegyi, s.p.hoogendoorn, h.j.vanzuylen@tudelft.nl)
\textsuperscript{c}: BAM Infraconsult B.V., Breda, Postbus 9559, 4801LN, the Netherlands (david.peters@home.nl).

Abstract

To keep the freeway networks in a good condition, road works such as maintenance and reconstruction are carried out regularly. The resulting work zones including the related traffic management measures, give different traffic capacities of the infrastructures, which determines the travel time for road users. A work zone capacity prediction model therefore is highly needed to evaluate mobility. Considering the work zone capacity as a function of work zone configurations, different prediction models have been developed in the past. The conventional models assume a linear relationship between the capacity of a work zone and its configuration variables. Recent artificial intelligence models are more flexible in constructing nonlinear relationships, but the accuracy of the models is not sufficiently tested. This research gives a comparison study of the existing models. Firstly, a selection of the critical work zone configuration variables is shortly discussed. Then three currently used prediction models are introduced, namely the model in the Highway Capacity Manual (2000), two multi-linear regression models, and a fuzzy logic based artificial neural network model. These models are tested for Dutch cases. Results show that comparing to the widely-applied linear regression models, the neuro-fuzzy model has the highest average accuracy and the prediction error can be reduced as large as 20%. The neuro-fuzzy model is recommended to serve in practice, as the choice of work zone configuration and the corresponding traffic measures can be made based on the capacity calculation.

© 2011 Published by Elsevier Ltd.

Keywords: work zone capacity; work zone configuration; capacity prediction; traffic management

1. Introduction

In order to keep the freeway network in a good state, regular road works such as maintenance and extension are necessary.

These road works create physical changes on freeway and result capacity reduction. If the capacity can be predicted, a systematic planning of traffic management can be executed for maintaining certain capacity. In practice, predicting capacity is not easy. Firstly, a number of variables, e.g. the composition of vehicles, road geometric design, and traffic measures, affect the capacity [1]. Those variables should be thoroughly considered, especially for
work zones since more infrastructural variables can influence the capacity. Secondly, the relationship between the capacity and work zone variables should be established by a mathematical model that has a great flexibility in constructing the relationship, as well as a good accuracy in prediction. The current work zone capacity prediction models do not sufficiently meet these demands. Previous studies and guidelines take limited number of variables into account respectively [2] [3] [4]. In most studies, the relationship between the variables and the capacity is assumed to be linear [4] [5]. Furthermore, the recent non-linear models are not sufficiently proved to have better prediction accuracy, although there were a few comparison studies [6] [7]. Therefore this research aims to compare the existing prediction models for freeway work zone capacity, in term of the assumptions, the construction and the accuracy of the models. Note that since the behavioral-based models require a large amount of data for model calibration for all possible work zone configurations, which is out of the available data source for this research, the focus of this study is only on data-driven models.

Section 2 gives a literature summary of the existing freeway work zone capacity prediction models. While section 3 gives a brief discussion on input variable selection. In Section 4, three prediction models are calibrated and a cross-validation method is designed to test the accuracy of the models. In Section 5, the test results for a Dutch case are presented. Section 6 will draw conclusions based on the discussion and the test results in Section 5.

2. Literature review

Two types of models are generally distinguished for capacity prediction, namely the data-driven models in which the capacity is related to the infrastructural and external variables, such as linear regression models, and the behavioral based model in which the capacity is the result of driving behavior in response to various conditions, such as the car-following models. The classic data driven model is the linear regression model. The Highway Capacity Manual (2000) gives a linear prediction model, considering the capacity in relation to the following four variables: intensity of the work activities, proximity to ramps, number of available lanes, and percentage of heavy vehicles. While Al-Kaisy et al. [4] and Kim et al. [5] also proposed linear models and added the following variables into their linear models: location of the closed lane, driver population, work zone gradient, lateral distance, work duration, weather condition, and work time. Recent studies focus on applying artificial intelligence (AI) models. Karim et al. [6] proposed a radial basis neural network model with eleven variables while Jiang and Adeli [7] proposed a neuro-fuzzy logic model considering seventeen variables, six of which were added to [6]: work zone location, work zone duration, weather condition, work day, work time, and work zone length. Compared to linear regression models, the AI models are more flexible in constructing the relationship between the variables and the capacity. Both Karim [6] and Adeli [7] showed that the AI models have better accuracy than the linear regression models. Although the results are promising, they are not sufficiently convincing. This is because in most researches the available data were split into construction data and test data, and the split was fixed. There could be different results by different choices of data. The other type of model is the behavioral-based model however not the focus of this research. Studies on applying behavioral-based models for work zone conditions can be found in [8] [9] [10].

3. Variable selection

The considered variables in the abovementioned models are summarized. Based on another study by the authors [11], the following eighteen variables should be considered and the corresponding definitions of the variables are given below: Composition of heavy vehicles: the percentage of heavy vehicles to the whole traffic population. Driver population: it distinguishes travel purposes (e.g. commuters and visitors) or the aggressiveness of drivers by age differences. Lane width: the total/average width of the available lanes/lane. Lateral distance: the distance from the edge of a lane to the work zone or to the physical separations (e.g. barrier). Number of available/ closed lanes: the number of lanes. Distance to ramps: the distance between a work zone and ramps in its vicinity. Month factor: different months. Presence of traffic signs: warning signs for infrastructure change, speed limit regulation, route information and etc. Presence of signal control: a control strategy that is used for demand management, especially at the areas that have reduction of available lanes. Road gradient: the geometric gradient of the lanes/temporary infrastructure. Road curve radius: the geometric radius of the lanes/temporary infrastructure. Sight deprivation: the absence or presence of sight proof, which is to prevent drivers from being distracted by work activities. Separation measures: to separate freeway from work zones. “Open” separation measures and the “closed” separation measures
can be distinguished. The open measures are those that only mark the division between the freeway and the work zones, such as cones. The closed measures are those that have continuous and complete separations, e.g. guard rail. Work zone intensity: the intensity depends on the type of the project, the number of workers, the size of the equipment and etc. Most manuals give their own classifications based on those figures. Work zone layout: three types of layout are generally recognized: lane reduction, lane split without lane reduction and lane-shift without lane reduction. Work zone length: the length of work zone activity. Work zone transition/buffer length: the distance between the late warning sign when approaching to a work zone and the actual start of the work zone.

4. Approach

In this section, the three considered models, the HCM 2000 model, the multi-linear regression model and the neuro-fuzzy model, are introduced and discussed. Section 4.1 introduces the three models respectively. For consistency of the notation, similar symbols for each capacity model are used. Section 4.2 describes how the parameters of each model are estimated. Section 4.3 shows the design of accuracy test. Section 4.4 gives the quantification for each variable.

4.1 Models

The proposed model in HCM 2000 is based on a simple linear regression [13] that considers four variables, as shown in (1):

\[ C = (1600 + v_i - v_r)HN_0 \]  

where \( C \) is the predicted traffic capacity of a cross-section; 1600 is the capacity of a single lane under ideal condition; \( v_i \) is the adjustment factor for work zone intensity (level 0~5); \( v_r \) is the adjustment factor for the distance from the cross-section to ramps; \( H \) is the adjustment factor for heavy vehicle (function of heavy vehicle proportion and passenger-car equivalent) and \( N_0 \) is the number of open lanes in work zone. To predict the capacity, the recommended values should be found for \( v_i \) and \( v_r \) from the HCM 2000, according to the specific work zone condition. While \( H \) is calculated by another equation provided in HCM 2000. An HCM model in a general form will be estimated:

\[ C = (C_0 + \beta_1 v_i - \beta_2 v_r)HN_0 \]  

where \( \beta \) is the estimated coefficients for the variables and \( C_0 \) is a constant. In more complicated linear regression model (Kim et al.,2000), seven variables were included.

\[ C = 1857 - 168.1v_{numcl} - 37v_{ocl} - 9v_{hv} + 92.7v_{ld} - 34.3v_{wl} - 106.1v_{w} - 2.3v_{wg} \]  

where \( C \) is the predicted capacity; \( v_{numcl} \) is the number of closed lane; \( v_{ocl} \) is the location of closed lane (right=1, otherwise=0); \( v_{hv} \) is the heavy vehicle percentage; \( v_{ld} \) is the lateral distance; \( v_{wl} \) is the work zone length; \( v_{w} \) is the work zone intensity and \( v_{wg} \) is the work zone gradient. In this research, a general multi-linear model will be estimated with the eighteen proposed variables, which are shown in (4):

\[ C = C_0 + \beta_1 x_1 + ... + \beta_{18} x_{18} \]  

where \( \beta \) is the estimated coefficients for the variables and \( C_0 \) is a constant. Jiang and Adeli (2003) introduced a fuzzy rule based neural network model for predicting freeway work zone capacity. Based on given data, the model generates rules
that represent the relationship between the input variables and the output. When given new inputs data, a consequent output \( C \) is obtained by rule \( i \) based on the membership degree of the new data to rule \( i \). This process is described by (5):

\[
C_i = \sum_{j=1}^{J} \exp \left( \frac{(q_j - c_j)^2}{2\sigma_j} \right) q_j.
\]  

In which the exponential function is the membership function [14] of rule \( i \); \( C_i \) is the rule-based value by rule \( i \); \( q_j \) are input variables; \( c_j, \sigma_j \) : the parameters of the membership function for variable \( j \) by rule \( i \). Given the new data, the strength of rule \( i \) in determining the final output is evaluated by (6) and then used to determine the final prediction result by (7).

\[
W_i = \prod_{j=1}^{J} \exp \left( \frac{(q_j - c_j)^2}{2\sigma_j} \right)
\]

\[
C = \sum_{i=1}^{N} C_i \frac{W_i}{\sum_{i=1}^{N} W_i}
\]

4.2 Model calibration

For the regression model, the coefficients are normally estimated by the least squares method [15]. When more variables are considered, the significance of including any variable on the prediction is examined and the variables that are not linearly significant to the prediction result are excluded, known as stepwise regression method [15]. The method begins with an initial model with several variables. Then it compares the explanatory power of incrementally larger or smaller models. This regression model will also be used in this research.

In the neuro-fuzzy model, the estimated parameters of the membership functions require adjustment. Due to the neural network structure, the backpropagation algorithm, a fundamental training algorithm for a neural network [15], is applied to do the adjustment. The backpropagation algorithm has two functions: an error calculation function and a learning function. In the error function as given by (8), a mean squared error is defined as the average of the squared differences between the actual values and the estimated values of training errors. The training error is the average difference between the estimated capacities by each consequent equation (equation (5)) and the corresponding measured actual capacities.

\[
E(c_{ij}, \sigma_{ij}) = \frac{1}{M} \sum_{k=1}^{M} \left| C_n^k - \hat{C}_n^k \right|^2
\]

where \( C_n^k \) is the \( k \)th normalized measured actual work zone capacity and \( \hat{C}_n^k \) is the \( k \)th normalized estimated work zone capacity. In the learning function, shown by (9), the new values of the parameters are a function of the old values of the parameters, the current error and a “learning ratio” which determines the training speed.

\[
W(c_{ij}, \sigma_{ij})_{new} = W(c_{ij}, \sigma_{ij})_{old} + \sum_{k=1}^{M} \eta (C_n^k - \hat{C}_n^k)
\]

where \( W(c_{ij}, \sigma_{ij}) \) represents the parameter set of the membership function and \( \eta \) is the learning ratio. This calibration process will be stopped when the error function is minimized by the current values of the parameters.

To ensure the generalization capability of the resulting model and avoid the overfitting problem in training neural network [15], the calibration data are split into training data and checking data, and are used to calculate the training error and the checking error. The training error is calculated by equation (8) as described above, while the checking
error is calculated by the same equation. The overfitting is supervised by optimizing both the error of training and of checking. This is illustrated in Figure 1, in which the upper line is the curve of checking error and the lower line is the curve of training error. It can be seen that although a better model can be always achieved with more training, the ability to fit in new data which are the checking data, the general performance becomes worse after a certain number of training. In this case, the optimal training is obtained 16 or 17 epochs. For other data sets, the optimal number of epochs may be different, but the method to determine it is the same.

Figure 1 an illustration of the value of training and checking errors versus the number of training times (y-axis, normalized values of training error; x-axis, number of iteration epochs)

4.3 Model accuracy and design of accuracy test

The accuracy of the considered models will be tested. One critical objective of designing the accuracy test is to ensure the test will not be influenced by the data itself. For instance, one calibrated model may show the highest accuracy in prediction when tested by a certain group of datasets, while when tested by a different group of datasets the model does not perform well. In such case, the conclusion on which model has the highest accuracy is rather unreliable. In previous studies, there were efforts made on avoiding data bias is, for example filtering the outliers [7]. However, the conclusions drawn by a single-round test are still not reliable. To overcome this difficulty, a cross-validation method [15] is employed here. All of the available datasets will be randomly split into construction (training data and checking data for the neuro-fuzzy model) datasets and test datasets. For each split, the models are fitted and their predictive accuracy is assessed. The process of data division is illustrated by Figure 2.

Figure 2 the process of an accuracy test based on cross-validation method (NF: neuro fuzzy, MR: multi regression)

4.4 Data collection

Collecting the capacity: the capacity can be defined a flow rate, below which traffic breakdowns will happen. A breakdown can be defined as a speed drops below a properly pre-defined critical speed. Furthermore, as
demonstrated in many studies that the capacity is not a fixed value rather follows a certain distribution, a distribution of the capacity values should be estimated rather than a single value. In these senses, the Product-limit method used by Brilon et al. [1] is suitable to estimate the capacity. By using this method capacity distributions are obtained, an example is displayed in Figure 3. The 50% percentile value is chosen to represent the average capacity of the data collection road section.

![Distribution of capacity estimated by Product limit model (2007-06, A2 160,5km)](image_url)

Figure 3 the cumulative probability distribution of the capacity values, A2 160,5km, 2007-06

Collecting the input variable: for the measurable variables, the values are the actual figures that are displayed on the blueprints of work zones or they are measured from the blueprints. For the linguistic variables, each is classified into levels based on preview studies and practical experiences. Integer numbers are assigned to distinguish the levels. Measurable variables: Driver population, Heavy vehicles percentage, Lane width, Lateral distance, Distance to ramps, Number of available lanes, Number of closed lanes, Road curve radius, Road Gradient, Temporary speed limit, Work zone length, Work zone transition length. Linguistic variables: Open or closed separations of traffic flow: 0 for no measures, 1 for open measures (open barriers, markers), 2 for closed measures (closed barriers, fence, cordon), Presence of signal control: 0 for yes, 1 for no, Sight deprivation: 0 for high level of deprivation (close to the lanes, without cover for the work zone, draw driver’s attention greatly), 1 for low level of deprivation (with good covers for the work zone, sometimes draw driver’s attention), Month factor: from 1 to 12, each number represents the month of the year, Work zone phase: 0 for the start phase (1 month after start), 1 for the mid-term phase (longer than 2 months), 2 for the end phase (less than 1 month before finish), Work zone intensity: 0 for high level of intensity (bridge renovation, road expansion), 1 for medium level of intensity (resurfacing), 2 for low level of intensity (pavement, median barrier repair or installation, pavement marking), Work zone layout: 0 for the layout of lane merge, 1 for the layout of narrowed lane only, 2 for bypass/shift

5. Case study

A case study is carried out for Dutch freeway work zones. Section 5.1 shortly describes data collection locations and data availability. In Section 5.2 the parameters of the models are estimated and the resulting models are discussed. The results of the designed accuracy test are given in Section 5.3.

5.1 Data collection

Traffic data and work zone data are collected on Dutch freeways. The datasets cover one finished bridge renovation project on Dutch freeway A16 that lasted nine months, and two ongoing freeway reconstruction projects on Dutch freeway A2, A58 and A67 that have been going on for two years. Due to the unavailability of data, the following four variables are not able to be considered in the datasets: road gradient, road curve radius, presence of signal control and driver population. Besides, presence of traffic signs is considered together with work zone
transition distance, since the latter one is the distance between the location of the last warning sign before the entrance of a work zone and the entrance. The capacities are estimated monthly. For each project, monthly capacity is estimated for each month of the whole project period if the data are available. In total seventy-one datasets, consisting of the remaining fourteen variables and the corresponding capacities, are collected. The datasets cover eight different locations within the three project areas.

5.2 Model calibration

In this section, the parameterized models estimated by one of the data divisions, are presented. And the rationale of the resulting models is discussed. The resulting HCM model is:

$$c = (1347.1 - 344.8v_r + 1357.2v_i)HN_o$$

Equation 10 gives a reasonable estimation of the parameters. The work zone capacity is positive coefficient to distance to ramps, indicating that the further a work zone is located to a ramp the better less the capacity it can have. The capacity is negative proportion to work zone intensity, indicating that a more intensive work zone causes a relatively larger capacity drop. A neuro-fuzzy model is estimated using the same four variables. The result fuzzy logic based model has 10 rules. The values of the parameters of the first five membership functions are given in Table 1.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Values of parameters $(c_{ij}, \sigma_{ij})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule 1</td>
<td>0.047, 0.875</td>
</tr>
<tr>
<td>Rule 2</td>
<td>0.047, 0.500</td>
</tr>
<tr>
<td>Rule 3</td>
<td>0.047, 0.875</td>
</tr>
<tr>
<td>Rule 4</td>
<td>0.047, 0.875</td>
</tr>
<tr>
<td>Rule 5</td>
<td>0.047, 0.500</td>
</tr>
</tbody>
</table>

In Figure 4, the fuzzy rules between the capacity and the considered four variables are displayed. The capacity decreases as composition of heavy vehicles increases, which is not linear but first slowly decrease and then fast. The capacity increases as the distance to ramps increases. If work zone intensity is not large, the capacity increases first and then decreases as the number of available lanes increases. When work zone intensity is large, the capacity increases as the number of available lanes increases. This could be explained as when multiple lanes are available, the interaction between the traffic flows is larger and causes more capacity reduction.

Another linear model is calibrated by regression. Several variables are excluded based on a correlation test. The resulting function is shown in equation (11). It can be seen that some estimated coefficients in function (equation (11)) are not consistent with common sense. For example the capacity is not in a direct proportion to lane. 

---

Figure 4 the relation viewers for the four variables (Left figure: z-axis capacity, x-axis heavy vehicles, y-axis distance to ramps; Right figure: z-axis capacity, x-axis work zone intensity, y-axis available lanes)
width nor lateral distance. Therefore we perform a stepwise multi-regression, to examine if it is the inclusion of redundant variables that confuse the estimation of coefficients. The resulting function after a stepwise regression is shown in equation (12). Four variables, sight deprivation, work phase and month factor, are excluded. One can notice that these four variables are linguistic and dummy variables and they have to be skipped in order to make good linear-regression models. It shows the lack of capability of the linear regression models in handling the linguistic and dummy variables.

\[
C = 1069 + 3.3v_{\text{heavy vehicle}} - 62v_{\text{lane width}} - 315v_{\text{lateral distance}} - 184\text{separation measure} + 744\text{distance ramps} + 25v_{\text{lanes available}} - 72v_{\text{lanes closed}} - 152v_{\text{speed limit}} + 6.6v_{\text{work phase}} + 49v_{\text{wz length}} - 184v_{\text{wz transition}} + v_{\text{month}} \tag{11}
\]

\[
C = 1690 - 300v_{\text{heavy vehicle}} - 62v_{\text{lane width}} - 3v_{\text{lateral distance}} + 60\text{separation measure} + 230\text{distance ramps} + 6v_{\text{lanes available}} - 44v_{\text{lanes closed}} + 9v_{\text{speed limit}} - 100v_{\text{wz length}} - 184v_{\text{wz transition}} \tag{12}
\]

Another neuro-fuzzy model with all the proposed variables is trained. The result model has 33 rules. The result of training errors and checking errors are illustrated in Figure 5. The overfitting problem is believed to happen after the 25th training. The trained parameters at the 20th training are chosen as the desired model parameters. The number of the pairs of the membership functions parameter is 495 (33 rules x 15 variables). Example membership functions for variable work zone length are illustrated in Figure 6. The physical meaning of the membership functions, for instance the most left membership function can be explained as representing a work zone that has the shortest length while the most right membership function can be explained as representing a work zone that has the longest length.

The predicted relation between the capacity and two proposed variables when keeping values of the other variables constant, are shown in Figure 7. Capacity is plotted with lane width and traffic separation measures. It is shown that a closed traffic separation measure gives more capacity. This could be because drivers feel safer when
work zones are completely separated from the freeway so that they drive faster than given open separation measure. The capacity will decrease as lane width increases, which is consistent with the result by the regression models. The right graph in Figure 8 shows that the capacity increases as the number lanes increase, when the lateral distance is no more than 0.5 meters. This could be because if the total available width for work zone freeway is fixed, given more lateral distance means less horizontal freedom between vehicles. Drivers slow down for safety concerns and cause the capacity reduction. Figure 8(left) indicates the capacity is not influenced by number of closed lanes much when work zone transition distance is very small or large enough. It also indicates that the capacity increases when enough transition length is provided. However, an unusual observation is that when work zone transition distance is between 300 and 800 meters, the capacity increases as closed lanes increases. In the right graph, when the number of available lanes is low, the capacity decreases just first and increases back after the work zone activity is finished. When more lanes are available, the capacity does not decrease or only a little.

![Figure 7](image1.png)  
**Figure 7** the relation viewers for four variables (left: z-axis capacity, x-axis lane width, y-axis traffic separation measure, right: z-axis capacity, x-axis lateral distance, y-axis available lanes)

![Figure 8](image2.png)  
**Figure 8** the relation viewers for four variables (left: z-axis capacity, x-axis work zone transition length, y-axis available lanes, right: z-axis capacity, x-axis work phase, y-axis available lanes)

Based on the discussion above, it can be concluded that the neuro-fuzzy model is more flexible in constructing and representing the relationship between the variables and the capacity. And the model gives better explanation of the relationship between the capacity and the linguistic variables than the multi-regression models.

5.3 Model accuracy test

The designed cross-validation technique was executed. In this case study, the datasets were randomly divided for a hundred times. The prediction accuracy was compared firstly between the HCM 2000 model and the neuro-fuzzy model that has the same variables. The root mean squared error (RMSE) for the accuracy set was calculated a one-hundred times. The result is illustrated by the first frequency histogram in Figure 10. From the figure we can see that although there is risk to result extreme large error (near 1000veh/h) by the neuro-fuzzy model (“Fuzzy 1”), in 90% cases the value of errors by the neuro-fuzzy model is lower than the HCM model. This result indicates that when considering the same variables while the number of considered variables is not large, the neuro-fuzzy model is more accurate than the HCM model.

Then the proposed fourteen variables were included. The prediction accuracy was compared between the regression models and the neuro-fuzzy model. The results are illustrated by the second frequency histogram in Figure 9. From the histogram, we see that the stepwise multi-regression model in general results smaller prediction
errors than the multi-linear regression model. By stepwise regression, variables that are statistically insignificant to the output capacity are removed from the model until the explanatory effect of the variables on the response cannot be improved anymore. This means that adding those variables may decrease the explanatory power of the resulting model, therefore result a less accurate model. According to the results of one hundred stepwise regressions variables such as month factor and work phase are excluded frequently.

Back to the histogram, 50% of cases the errors by the neuro-fuzzy model (“Fuzzy 2”) is smaller than the errors by the regression models. However there is still risk for the neuro-fuzzy model to result high value of error, indicating that it is possible to obtain a bad-trained model. The observation in the second histogram also indicates that when the same variables are considered, the neuro-fuzzy model is more accurate than the regression models. Furthermore, by comparing the HCM model and the regression models, we can see that the value of errors is reduced when including more variables in the model. The same result can be observed by comparing the error of “Fuzzy 1” and of “Fuzzy 2”.

6. Conclusion

In this research, a study is performed on comparing the prediction models for freeway work zone capacity. Firstly, previous studies on the topic are reviewed. Secondly the selection of variables is shortly explained. Then, the three currently-developed models namely the HCM 2000 model, multi-linear regression models and a neuro-fuzzy model are discussed, in term of input variables, model assumptions, model construction and model accuracy. The models are modified and tested in Dutch freeway cases.

The proposed neuro-fuzzy model has the average highest prediction accuracy. This is because it constructs a non-linear relationship between the variables and the capacity, and it has a better capability of handling linguistic and dummy variables than the traditional capacity prediction models. The proposed neuro-fuzzy model with the proposed variables can be implemented as a work zone capacity prediction in practice.

Due to the limited amount of data, some possibly important variables such as road gradient and road curve radius were not included in the proposed method. In future studies, those variables may be studied in order to, for instance, explain the uncommon relation between the lane width and the capacity, and etc. Besides, studies are also suggested.
on applying behavior-based model to predict the capacity, whose results can be used to test the conclusions from this research.

Acknowledgement

This project was supported by Dutch Ministry of Transport and Dutch company Heijmans Techniek & Mobiliteit.

References