Modeling Operating Speeds on Residential Streets with a 30 km/h Speed Limit: Regression versus Neural Networks Approach

Do Duy DINH [a], Aya KOJIMA [b], Hisashi KUBOTA [c]

[a,b,c] *Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Saitama City, Saitama Prefecture, 338-8570, Japan*
[a] *E-mail: dinh@dp.civil.saitama-u.ac.jp*
[b] *E-mail: kojima@dp.civil.saitama-u.ac.jp*
[c] *E-mail: hisashi@dp.civil.saitama-u.ac.jp*

Abstract: This paper models operating speeds on residential streets with a 30 km/h speed limit by using: (i) regression methods including Single Equation Regression (SER) and Simultaneous Equation Approach (SEA), and (ii) Neural Networks (NN) modeling technique. Free-flow profile-speed data were recorded on 99 street sections with varying characteristics which were then used to develop and validate speed models for estimating maximum speeds obtaining within a section and speeds at the entrance to the next un-signalised intersection. The results suggest that the models developed by SEA performed better than those by the conventional regression (i.e., SER). Compared to regression models, NN models showed better performance especially regarding model fitness although the resultant models are quite complicated. Based on the developed models, various street features were found as determinants of driving speeds that provided helpful information for addressing speeding issues on neighborhood streets.

Keywords: Operating Speed; Residential Street; 30 km/h Speed Limit; Regression; Simultaneous Equations; Neural Networks.

1. INTRODUCTION

Urban residential streets with a 30 km/h speed limit are very popular in many countries such as Japan. The primary purpose of setting the speed limit is to ensure the safety of vulnerable street users (i.e., pedestrians, cyclists) on those streets and make residential neighborhoods as a safe living environment. The safety benefit of the 30 km/h speed limit was highlighted by the fact that 90% of pedestrians hit by a car travelling at 50 km/h survived, while only 20% of pedestrians hit by a car travelling at 50 km/h survived (OECD/ECMT, 2006). However, despite efforts to slow down vehicles by setting the speed limit, in Japan, excessive speeds on those streets were very common (Dinh and Kubota, 2013) making the streets dangerous to all users. To tackle the speeding issues on residential areas, it is necessary to find out speed influencing factors from roadway and roadside characteristics.

While abundant research has been dedicated to figure out the influences of roadway and roadside factors on vehicular speeds on rural highways (Gong and Stamatiadis, 2008; Wang *et al.*, 2006), a relatively limited number of studies have been completed for urban conditions. Amongst those studies focusing on urban streets, operating speeds were often studied at either curve locations or straight sections. As respect to speeds at curve locations, Fitzpatrick *et al.* (1997) showed that curve radius and approach density were variables of speeds at horizontal curves while the inferred design speed significantly affected speeds on vertical curves on suburban roadways. For four-lane suburban arterials, posted speed limit, deflection angle, and access density class were found as significant factors of speeds at horizontal curves while the presence of median and type of roadside developments became significant variables in a speed model without considering posted speed limit (Fitzpatrick *et al.*, 2001). By using ordinary regression, Tarris *et al.* (1996) developed speed models for estimating speeds at horizontal curves on urban collector streets with only one significant variable that was the degree of curve. With the same dataset, Peo and Manson (2000) used a mixed-model approach to build speed models with several speed predictors including: degree of curvature, longitudinal grade, lane width, and roadside characteristics. Bonneson (1999) studied speeds at 55 horizontal curves on urban/rural roadways and turning roadways. Based on the collected data, the researcher constructed a relationship between curve speed and influencing factors including approach speed, radius, and super-elevation. Up to date, there have been still a few operating speeds models available for estimating speeds on urban tangent streets. Previous studies (Fizpatrick *et al.*, 2003; Fizpatrick *et al.*, 2001; Ali *et al.*, 2007) identified posted speed limit as the only significant variable or the most significant predictor of operating speeds on the straight sections. Other street features such as roadside density, driveway density, availability of sidewalk, presence of on-street parking, number of lane, curb presence, and type of residential land uses were found as speed influencing factors on urban tangent streets (Wang *et al.*, 2006). There was only one previous study (Dinh and Kubota, 2013) that developed speed models for residential streets with a 30 km/h speed limit. The authors found various roadway and roadside factors as determinants of speed choice on those streets.

With regards to the methodologies used to develop operating speed models, most previous studies relied on a conventional regression approach namely a Single Equation Regression (SER). When using the SER, if data were available for modeling speeds at multiple locations along a street, speed models were separately developed for each point. An underlying assumption of the SER approach is that there is no endogenous relationship between dependent variables. However, when the locations under study are close, as is the case of residential streets with a 30 km/h speed limit, the conventional SER is not able to account for the potential endogenous relationship between speeds at the two locations that may lead to a bias on modeling results. Dinh and Kubota (2013) introduced for the first time a Simultaneous Equation Approach (SEA) to address the endogenous relationship issue on modeling operating speeds at multiple locations on a street section. Although, the developed model in that study had a reasonable fit and integrated a number of roadway and roadside factors, the same as most existing studies, the validation of the developed models based on an independent dataset have not been conducted to test its predictive ability. Up to date, there were no studies made a comparison regarding the performance between the SEA models and those developed by SER that makes it difficult to fully understand the robustness of the SEA method.

Beside regression approaches, Neural Networks (NN) modeling technique is an alternative methodology for modeling operating speeds. Different with statistical methods, NN models are not subject to distributional or other constraints inherent to regression (Taylor *et al.*, 2007). Specifically, the strength of the NN method lies on that no assumptions are needed regarding the model form; and that NN models are able to effectively deal with nonlinearity and multicollinearity that often ruin the linear regression models (Karlaftis and Vlahogianni, 2011). Transportation Research Board (2011) highlighted certain limitations of regression-based speed models and made a recommendation that the advance modeling methods such as NN should be used to overcome those limitations. So far, several NN-based operating speed models have been developed with promising results. Zaman *et al.* (2000) successfully developed NN models for predicting the 85^{th} percentile speeds for two-lane rural highways. In a study by McFadden *et al.* (2001), NN-based models were developed for estimating speeds at horizontal curves. The results showed that the performance of the NN models was comparable to those by regression approach. Recently, Signgh *et al.* (2012) used NN method to model 85th percentile speed for two-lane rural highways. The final models in that study showed a reasonable accuracy and integrated a number of input variables.

Although NN method has been demonstrated as an effective tool for modeling operating speeds, several limitations inherent to the method should be noted. The inference mechanism of a NN model is hidden itself causing difficulties on understanding the influence of each input variable on the model output. In addition, researchers often arbitrarily selected input variables for NN models and, very frequently, the implicit assumptions made regarding the data were disregarded. This implementation may lead to redundant input variables entering on the NN models. The redundant variables often require a large dataset and probably results to an overfitting issue which make NN models unable to generalize to a new dataset. Other limitations of NN models were also listed on a work by Karlaftis and Vlahogianni (2011).

Because as aforementioned, both the regression approach and the NN method possesses each own merits and limitations, it is necessary to carefully select the modeling technique suitable for each specific research question. Because there is no specific guideline on selecting the appropriate modeling approach (either a regression based or a NN based model), a comparison on the performance of these models on the same study context is needed to deeply understand about the strength and weakness inherent to each methodology. However, such knowledge up to date is still very limited and it deserves for more considerations (Karlaftis and Vlahogianni, 2011).

The primary objectives of the present study, therefore, is to evaluate the performance of operating speed models developed by using (i) regression methods (i.e., SER and SEA), and (ii) NN modeling technique. Different with previous studies that often used point-speed data; the model parameters in this study were calibrated based on the free-flow profile-speed data recorded on a number of residential streets with a 30 km/h speed limit. Speed models were then validated/tested with different dataset to confirm their predictive ability. From the developed models, the speed-influencing factors from roadway and roadside characteristics were also explored to provide helpful information for speeding intervention.

2. DATA COLLECTION

A free-flow speed survey was conducted on 99 street sections located in the areas of the city of Saitama, Kawaguchi, and Warabi in the Saitama Prefecture, Japan. The survey was divided into two periods of time. The first period was taken place from August 20th, 2011 to November 10th, 2011 with 85 selected street sections. The second survey was conducted to 14 street sections from November 15th, 2012 to December 10th, 2012. It was intended that, the data collected in the first period would be used to develop speed models which were then validated/tested with the data from the second survey period. In the present study, all selected street sections were located in residential areas. Only straight sections with a 30 km/h speed limit were selected. General street characteristics of the selected sections are summarised in Table 1.

This study used STALKER ATS radar guns connected to a laptop to record free-flow profile-speeds of individual vehicles travelling on each street section. A study section or site was defined as a segment between two intersections with a specified direction. The entering intersection must be a 4-leg or 3-leg intersection with a stop line for the entering approach. The exiting intersection must be un-signalised with a stop line for the study direction. It should be noted that in these sections drivers do not have to stop before the “stop line” of the

Table 1. Summary of selected street section characteristics

Characteristics	Measured value	
	The first survey period	The second survey period
Length of street section (m)	86.70 to 268.10; mean: 140.76	86.90 to 217.20; mean: 124.90
Number of lanes	1 to 2; mean: 1.64	1 to 2; mean: 1.71
Lane width (m)	2.35 to 5.70; mean: 3.51	2.45 to 5.30; mean: 3.28
Carriageway width (m)	3.40 to 7.10; mean: 5.30	3.60 to 7.00; mean: 5.32
Roadway width (m)	4.70 to 8.90; mean: 6.58	3.70 to 8.00; mean: 5.55
Left safety strip width (m)[a]	0 to 1.70; mean: 0.42	0 to 1.10; mean: 0.44
Right safety strip width (m)[b]	0 to 4.45; mean: 2.41	0.6 to 4.00; mean: 2.62
Presence of sidewalk	no sidewalk: 43 sites; sidewalk on one side: 24 sites; sidewalk on both sides: 18 sites	no sidewalk: 7 sites; sidewalk on one side: 3 sites; sidewalk on both sides: 4 sites
Sidewalk width (m)	0 to 5.10; mean: 1.32	0 to 2.50; mean: 1.07
Roadside object density (per 100 m)[c]	0 to 7.06; mean: 2.08	0 to 8.58; mean: 2.97
Driveway density (per 100 m)	0 to 3.44; mean: 0.90	0 to 1.91; mean: 0.65
Street marking	Centreline marking: 54 sites; edge marking only: 20 sites; no marking: 11 sites	Centreline marking: 10 sites; edge marking only: 2 sites; no marking: 2 sites
Land use development	Private houses are dominant: 35 sites; apartment/tall buildings are dominant: 15 sites; mixing development: 20 sites; near schools/ parks: 15 sites	Private houses are dominant: 8 sites; apartment/tall buildings are dominant: 3 sites; mixing development: 1 site; near schools/ parks: 2 sites
Type of entering intersection	Signalised intersection: 58 sites; 4-leg non-signalised intersection: 18 sites; 3-leg non-signalised intersection: 9 sites	Signalised intersection: 12 sites; 4-leg non-signalised intersection: 2 sites; 3-leg non-signalised intersection: 0 site
Type of exiting intersection	4-leg non-signalised intersection: 63 sites 3-leg non-signalised intersection and similar: 22 sites	4-leg non-signalised intersection: 13 sites 3-leg non-signalised intersection and similar: 1 site
Distance from the entrance (stop line) of exiting intersection to the nearest control point (m)[d]	43.30 to 339.40; mean: 139.8	82.50 to 265.00; mean: 163.47
Distance from the entrance (stop line) to the centre point of exiting intersection (m)	3.20 to 15.10; mean: 9.10	2.50 to 12.40; mean: 6.20
Distance from the entrance (stop line) to the nearest pedestrian crossing strip of exiting intersection (m)	2.40 to 14.10; mean: 4.79	4.80 to 12.70; mean: 9.33
Presence of pedestrian crossing strip at exiting intersection	Both before and after centre point: 35 sites; only before centre point: 25 sites; only after centre point: 25 sites	Both before and after centre point: 7 sites; only before centre point: 1 site; only after centre point: 6 sites
Width of crossing street (m)	2.40 to 13.30; mean: 5.60	3.90 to 6.70; mean: 5.22
Roadway width ratio between crossing street and study street	0.36 to 1.40; mean: 0.75	0.59 to 1.27; mean: 0.82

Notes: [a] Left safety strip width was measured from the edge of a study lane to the curb on the left. [b]Right safety strip width was measured from the edge of a study lane to the curb on the right. [c]Only rigid objects (such as utility poles) within 0.5 m from the edge of roadway were counted; and only objects on the left were counted if centreline marking is available; [d] The nearest control point is the nearest signalised intersection or the nearest location where drivers have to reduce speeds substantially.

exiting intersection except the case when they yield to pedestrians/cyclists crossing the roadway at the intersection. The radar gun was started to trigger when only one target vehicle entered a study section and at the same time there were no pedestrians or cyclists on the roadway. This was to ensure free-flow conditions for the selected vehicles as well as to reduce the interference from other moving objects that may have spoiled the recorded profile-speed data. The gun was kept operating until the vehicle reached an identified point. The point, which was predetermined and located after the entrance of the exiting intersections, was then used to match profile speeds with the street layout.

A lot of effort was made to enhance the accuracy of the speed data such as carefully eliminating the presence of the survey devices and surveyors to drivers or setting up the radar gun in the same side as the study lane. For each street section, at least 70 profile speeds were recorded. Speed data were collected only in good weather during the daytime. Only passenger cars and light trucks that did not turn at the exiting intersections were included in the study. A more detail about the street selection and technical issues of the survey were shown in Dinh and Kubota (2013).

3. DATA ANALYSIS

3.1 Data Reduction

Speed data were initially processed using the software program that accompanied the STALKER ATS radar guns to obtain a relationship between speed and distance for each vehicle. Vehicle speeds were matched exactly with street layouts by using the location of the identified points, i.e., the points where the radar guns were released as described in *Section 2*. Figure 1 provides typical profile-speed data for all vehicles in one street section.

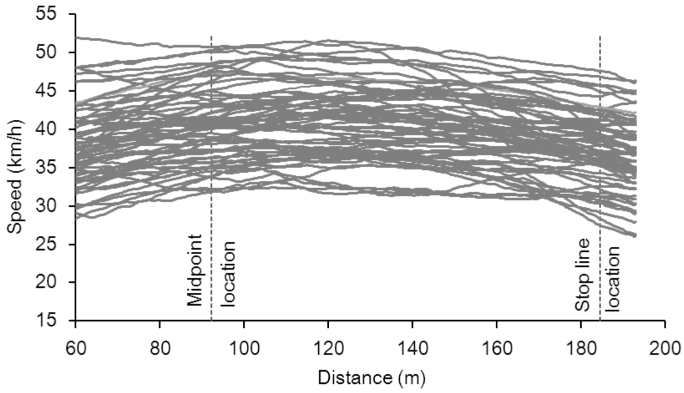

Figure 1. Typical profile-speed data for one street section (Dinh and Kubota, 2013)

By examining the speed data, it was found that after entering the sections, most drivers accelerated up to a maximum speed then decelerated, with the deceleration possibly because they were approaching the exiting intersection. Also, drivers likely reached their maximum speeds in the second half of the street sections because statistics indicated that more than 85% of the drivers in the first survey period and more than 93% drivers in the second period reached their maximum speeds after passing the midpoint of the street sections. Most drivers who had their highest speeds before reaching the midpoint location travelled at a very high speed when passing the entering intersection. As a result of a field observation in this study, these vehicles were highly affected by favourable driving conditions when entering the study sections such as passing the intersection during the green signal time. These maximum speeds, arguably, do not fully reflect the influence of roadway and roadside characteristics on driving speeds on the study street sections. Therefore, this study assumed the maximum speeds obtaining within the second half of a street section as the maximum speed influenced by the characteristics of the street. This is referred to hereafter as the “maximum speed”, “speed of tangent” or “tangent speed”.

In the next data analysis step, only speed profiles that covered the full second half of street sections were used. Speed profiles with abnormal driving patterns were excluded. Maximum speeds then were calculated for all individual vehicles. For each street section, individual speed profiles were excluded if their maximum speeds differed from the section mean by more than two standard deviations. The speed at the entrance to the exiting intersection (i.e., the stop line location), also hereafter called the “speed at intersection”, was determined for every vehicle in all sections. Speed profiles with speed at intersection less than 10 km/h were excluded because they were likely associated with unfavourable driving conditions at the exiting intersections that might not have been observed during the survey.

After data reduction, 5359 individual speed profiles for 85 street sections in the first survey period and 864 speed profiles for 14 sections in the second survey period remained for further analysis. The minimum number of individual speeds for one street section is 53, and the maximum number is 75. As noted by Dinh and Kubota (2013), the speeding problem is very serious because few people drove at the speed limit and nearly half of the drivers exceeded 40 km/h on streets with a 30 km/h speed limit.

3.2 Dependent Variables

The maximum speed obtained within a tangent section (speed of tangent) and the speed at the entrance to the next un-signalised intersection (speed at intersection) are both potentially related to traffic safety issues for urban streets with a 30 km/h speed limit, therefore both these speeds were examined in this research. To be consistent with previous studies, 85th percentile speeds were used to represent operating speeds. Mean speeds were also examined since they are likely to be considered as an indicator of speeding level on those street sections. Table 2 presents descriptive statistics of the dependent variables.

Table 2. Descriptive statistics of dependent variables

Variable description	Variable code	The first survey period					The second survey period				
		N	Min	Max	Mean	SD	N	Min	Max	Mean	SD
85^{th} percentile speed of tangent (km/h)	$V_{max,85}$	85	38.73	50.55	44.56	2.81	14	38.43	50.76	43.37	3.61
Mean speed of tangent (km/h)	$V_{max,av}$	85	33.46	45.04	39.50	2.35	14	33.96	44.04	38.80	3.00
85^{th} percentile speed at intersection (km/h)	$V_{in,85}$	85	22.81	49.01	40.54	4.58	14	32.07	47.03	40.29	4.00
Mean speed at intersection (km/h)	$V_{in,av}$	85	21.68	41.84	35.35	4.16	14	28.96	40.76	35.73	3.42

3.3 Modeling Approaches and Results

This study used regression methods (i.e., SER and SEA), as well as NN modelling to build models for estimating operating speeds. The following is a brief description of the model form for each respective approach and the corresponding modelling results.

3.3.1 Single equation regression (SER)

By this conventional approach, tangent speeds and speeds at intersection were separately modelled by using a multiple linear regression method. The model forms are as below:

$$\ln V_1 = \alpha_1 + \beta_1 X_1 + \varepsilon_1 \quad (1)$$
$$\ln V_2 = \alpha_2 + \beta_2 X_2 + \varepsilon_2 \quad (2)$$

where,

V_1: tangent speed (either $V_{max,85}$ or $V_{max,av}$),
V_2: speed at intersection (either $V_{in,85}$ or $V_{in,av}$),
X_1, X_2 : vectors of independent variables representing street characteristics
α_1, α_2 : estimable parameters, and
ε_1, ε_2 : disturbance terms.

To obtain the best model for predicting each target speed, a standard procedure for developing a multiple linear regression model was used. First, possible relationships between independent variables and each dependent variable were identified by using scatter plots and a simple regression method. The possible combinations of selected independent variables were then used to develop regression models. For each candidate model, a test of multicollinearity was performed by using the variance inflation factor (VIF) and extreme data were eliminated through casewise diagnostics. Other assumptions of linear regression such as homoscedasticity, normally distributed errors, and error independence were also tested. The finally selected models were those with a high adjusted coefficient of determination R^2_{adj} and passing all tested assumptions. All variables included in the final models have a significant level of 95%.

The data from 85 street sections in the first survey period were used to develop speed models. These models were then validated based on the data from 14 street sections in the second survey period. A number of indicators were used to evaluate the performance of the developed models. Table 3 provides the final selected models, while the performance of these models is shown in Table 4.

3.3.2 Simultaneous equation approach (SEA)

An important assumption of the SER is that there must be no relationship between dependent variables. However, the assumption seems to be violated because the data for this study showed a high correlation between tangent speeds and speeds at intersection with correlation coefficients (calculated based on the data from the first survey period) of 0.734 and 0.782 for the 85th percentile speeds and the mean speeds, respectively. That suggested that the SER may be inappropriate and it was necessary to consider the potential endogenous relationship while modelling speeds of tangent and speeds at intersection.

An effective way to address the afore-mentioned issues is using a SEA as discussed earlier. By this method, both tangent speeds and speeds at intersection were simultaneously modelled in each respective equation system. To account for the possible influences of speeds at intersections on tangent speeds and vice versa, the model specification in a general form based on SEA technique (see Green (2003) for a detailed description of the SEA technique) is expressed as:

$$\ln V_1 = \alpha_1 + \beta_1 X_1 + \gamma_1 \ln V_2 + \varepsilon_1 \quad (3)$$

$$\ln V_2 = \alpha_2 + \beta_2 X_2 + \gamma_2 \ln V_1 + \varepsilon_2 \tag{4}$$

where,

V_1: tangent speed (either $V_{max,85}$ or $V_{max,av}$),
V_2: speed at intersection (either $V_{in,85}$ or $V_{in,av}$),
X_1, X_2 : vectors of exogenous variables representing street characteristics,
α_1, α_2 : estimable parameters,
γ_1, γ_2 : estimable scalars, and
ε_1, ε_2 : disturbance terms.

Table 3. Operating speed models developed by SER

Variable	Estimated coefficient	S.E.	t-ratio	p-value	$\bar{X}$
Dependent variable: logarithm of 85th percentile speed of tangent ($LnV_{max,85}$) (km/h)					
Constant	3.534	0.043	82.678	0.000	
Number of lanes	0.034	0.014	2.381	0.020	1.635
Length of street section (m)	0.00050	0.00014	3.640	0.000	140.758
Sidewalk indicator (1 if sidewalks are available on both sides; 0 otherwise)	0.025	0.012	2.065	0.042	0.212
Roadside object density (per 100 m)	-0.0079	0.0026	-3.041	0.003	2.833
Carriageway width (m)	0.028	0.010	2.961	0.004	5.305
R^2 = 0.560; Adjusted R^2 = 0.532					
Dependent variable: logarithm of 85th percentile speed at intersection ($LnV_{in,85}$) (km/h)					
Constant	3.331	0.097	34.295	0.000	
Right safety strip width (m)	0.030	0.010	3.024	0.000	2.412
Sidewalk indicator (1 if sidewalks are available on both sides; 0 otherwise)	0. 079	0.026	3.055	0.003	0.212
Type indicator of exiting intersection (1 if 3-leg intersection or similar; 0 otherwise)	0.085	0.024	3.591	0.001	0.259
Carriageway width (m)	0.048	0.021	2.280	0.025	5.306
R^2 = 0.471; Adjusted R^2 = 0.445					
Dependent variable: logarithm of mean speed of tangent ($LnV_{max,av}$) (km/h)					
Constant	3.447	0.042	81.373	0.000	
Length of street section (m)	0.00039	0.00013	3.096	0.003	140.758
Right safety strip width (m)	0.012	0.004	2.648	0.010	2.412
Roadside object density (per 100 m)	-0.0091	0.0023	-4.041	0.000	2.833
Carriageway width (m)	0.032	0.009	3.680	0.000	5.306
R^2 = 0.577; Adjusted R^2 = 0.556					
Dependent variable: logarithm of mean speed at intersection ($LnV_{in,av}$) (km/h)					
Constant	3.162	0.095	33.341	0.0000	
Right safety strip width (m)	0.031	0.010	3.040	0.0000	2.412
Roadside object density (per 100 m)	-0.0017	0.0053	-3.228	0.0010	2.833
Distance from the entrance of exiting intersection to the nearest control point (m)	0.00034	0.00015	2.195	0.0134	139.782
Carriageway width (m)	0.058	0.020	2.890	0.0134	5.305
Type indicator of exiting intersection (1 if 3-leg intersection or similar; 0 otherwise)	0.059	0.024	2.434	0.0439	0.259
R^2 = 0.540; Adjusted R^2 = 0.511					

It should be noted that, The $ln V_2$ and $ln V_1$ on the right-hand-sides in Equation 3 and Equation 4 are endogenous variables that account for the endogenous relationship between the two speeds, as V_2 affects V_1 and vice versa. Different with the SER, the SEA is able to account for a correlation between the two error terms ε_1 and ε_2 in Equation 3 and Equation 4 respectively. The model parameters were estimated by using a 3-stage-least-square (3SLS) estimator. A detail description of the model development procedure was showed in Dinh and Kubota (2013). A further technical discussion of the technique used could be seen in Greene (2003).

As similar to the SER approach in *Section 3.3.1*, the data from the first survey period were used to develop speed models while those from the second survey period were used for validation. Table 5 provides the final selected models, while the performance of these models is shown in Table 6.

Table 4. Performance of operating speed models developed by SER

Indicator	$V_{in,av}$	$V_{in,85}$	$V_{max,av}$	$V_{max,85}$
With dataset for model development				
Mean squared error (MSE), $(km/h)^2$	8.35	10.98	2.39	3.61
Root mean square error (RMSE), km/h	2.89	3.31	1.55	1.90
Mean absolute error (MAE), km/h	2.21	2.45	1.23	1.53
% sites with absolute error (AE) ≤ 2.5	69.41	67.06	88.24	81.18
% sites with 2.5 < AE ≤ 5	23.53	20.00	11.76	18.82
% sites with 5 < AE ≤ 10	7.06	11.76	0.00	0.00
% sites with 10 < AE ≤ 12	0.00	1.18	0.00	0.00
With dataset for model validation				
Mean squared error (MSE), $(km/h)^2$	3.08	9.82	2.40	4.62
Root mean square error (RMSE), km/h	1.76	3.13	1.55	2.15
Mean absolute error (MAE), km/h	1.40	2.58	1.37	1.98
% sites with absolute error (AE) ≤ 2.5	71.43	57.14	92.86	85.71
% sites with 2.5 < AE ≤ 5	28.57	28.57	7.14	14.29
% sites with 5 < AE ≤ 10	0.00	14.29	0.00	0.00
% sites with 10 < AE ≤ 12	0.00	0.00	0.00	0.00

3.3.3 Neural networks (NN) modeling

NN is a modelling technique that was used to perform mapping of an input vector in to an output vector. A simple neural networks used in this study was a feed-forward neural networks consisting of three layers, namely input layer, hidden layers, and output layer as shown in Figure 2. The input layer included a number of input nodes that were selected variables representing street characteristics. The number of hidden layers and the number of nodes in each hidden layer were chosen in order that the best performance of the models was archived. The output layer consisted of one node representing the model outcome that was either $V_{max,85}$, $V_{max,av}$, $V_{in,85}$, or $V_{in,av}$. The relationship between the input variables and the output of the NN models was expressed by Equation 5. This study used a hyperbolic tangent or ‘tangsig’ function as an transfer function for the hidden layer (Equation 6), while the output layer used a pure linear function (called ‘purelin’) as an activation function (Equation 7). Training the network is a necessary step to estimate the model parameters (i.e., weights factors and bias factors).

Table 5. Operating speed models developed by SEA (adapted from Dinh and Kubota, 2013)

Variable	Estimated coefficient	S.E.	t-ratio	p-value	$\bar{X}$
Dependent variable: logarithm of 85th percentile speed of tangent ($LnV_{max,85}$) (km/h)					
Constant	3.5383	0.0392	90.248	0.0000	
Number of lanes	0.0323	0.0127	2.547	0.0109	1.635
Length of street section (m)	0.00050	0.00013	3.757	0.0002	140.758
Sidewalk indicator (1 if sidewalks are available on both sides; 0 otherwise)	0.0269	0.0111	2.415	0.0157	0.212
Roadside object density (per 100 m)	-0.0082	0.0023	-3.566	0.0004	2.833
Carriageway width (m)	0.0285	0.0085	3.345	0.0008	5.305
R^2 = 0.559; Adjusted R^2 = 0.532					
Dependent variable: logarithm of 85th percentile speed at intersection ($LnV_{in,85}$) (km/h)					
Constant	-3.4625	0.7480	-4.629	0.0000	
Logarithm of 85th percentile speed of tangent (km/h)	1.941	0.2001	9.700	0.0000	3.795
Length of street section (m)	-0.00123	0.00027	-4.528	0.0000	140.758
Distance from the entrance of exiting intersection to the nearest control point (m)	0.00024	0.00012	1.960	0.0500	139.782
Roadway width ratio between crossing street and study street	0.0902	0.0276	-3.274	0.0011	0.746
R^2 = 0.569; Adjusted R^2 = 0.547. The variable in italic is endogenous.					
Dependent variable: logarithm of mean speed of tangent ($LnV_{max,av}$) (km/h)					
Constant	3.4484	0.0396	87.109	0.0000	
Length of street section (m)	0.00039	0.00012	3.220	0.0022	140.758
Right safety strip width (m)	0.0125	0.0041	3.058	0.0013	2.412
Roadside object density (per 100 m)	-0.0089	0.0021	-4.279	0.0000	2.833
Carriageway width (m)	0.0313	0.0080	3.902	0.0001	5.306
R^2 = 0.577; Adjusted R^2 = 0.556					
Dependent variable: logarithm of mean speed at intersection ($LnV_{in,av}$) (km/h)					
Constant	-3.9230	0.6688	-5.866	0.0000	
Logarithm of mean speed of tangent (km/h)	2.0953	0.1842	11.373	0.0000	3.674
Length of street section (m)	-0.00139	0.00023	-5.921	0.0000	140.758
Width of crossing street (m)	-0.0152	0.0046	-3.299	0.0010	5.601
Type indicator of exiting intersection (1 if 3-leg intersection or similar; 0 otherwise)	0.0442	0.0179	2.473	0.0134	0.259
Distance from the entrance to the centre point of exiting intersection (m)	0.0056	0.0027	2.015	0.0439	9.106
R^2 = 0.697; Adjusted R^2 = 0.677. The variable in italic is endogenous.					

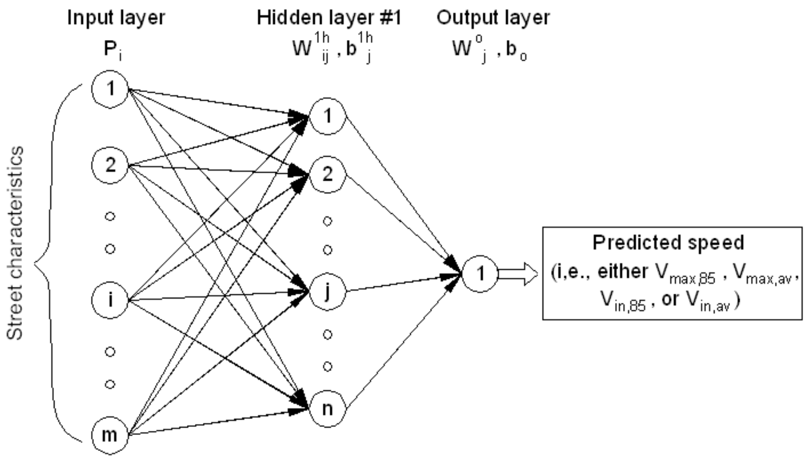

Figure 2. Architecture diagram of a NN model

Table 6. Performance of operating speed models developed by SEA

Performance indicator	$V_{in,av}$	$V_{in,85}$	$V_{max,av}$	$V_{max,85}$
With dataset for model development				
Mean squared error (MSE), $(km/h)^2$	7.95	10.33	2.39	3.58
Root mean square error (RMSE), km/h	2.82	3.21	1.55	1.89
Mean absolute error (MAE), km/h	2.06	2.38	1.23	1.54
% sites with absolute error (AE) ≤ 2.5	70.59	58.82	88.24	82.35
% sites with 2.5 < AE ≤ 5	20.00	30.59	11.76	17.65
% sites with 5 < AE ≤ 10	9.41	9.41	0.00	0.00
% sites with 10 < AE ≤ 12	0.00	1.18	0.00	0.00
With dataset for model validation				
Mean squared error (MSE), $(km/h)^2$	4.49	5.43	2.43	4.80
Root mean square error (RMSE), km/h	2.12	2.33	1.56	2.19
Mean absolute error (MAE), km/h	1.76	1.94	1.38	2.03
% sites with absolute error (AE) ≤ 2.5	71.43	71.43	85.71	71.43
% sites with 2.5 < AE ≤ 5	28.57	28.57	14.29	28.57
% sites with 5 < AE ≤ 10	0.00	0.00	0.00	0.00
% sites with 10 < AE ≤ 12	0.00	0.00	0.00	0.00

$$Y(P_1,\ldots,P_m) = f_0\left[b_0 + \sum_{j=1}^{n}\left\{W_j^0 f_h\left(b_j^{1h} + \sum_{i=1}^{m} W_{ij}^{1h} P_i\right)\right\}\right] \tag{5}$$

$$f_h(T) = \frac{2}{1+e^{-2T}} - 1 \tag{6}$$

$$f_0(T) = T \tag{7}$$

where,

Y : output value (i.e., either $V_{max,85}$, $V_{max,av}$, $V_{in,85}$, or $V_{in,av}$),
P_i : input variable i^{th},
i : subscript for input layer,
j : subscript for hidden layer,
m : number of input parameters,
n : number of nodes in hidden layer,
f_h : transfer function for hidden layer,
f_0 : transfer function for hidden layer,
W^{1h}_{ij} : weight factors for hidden layer,
W^0_j : weight factors for output layer,
b^{1h}_j : bias factors for hidden layer,
b_0 : bias factor for output layer.

To facilitate the training process, the dataset should be divided into two groups - one for training the network and the other for testing the network. In this study, the data from 85 street sections collected on the first survey period were used as a training dataset. Meanwhile, the data from 14 street sections on the second survey period were used only for testing. It should be noted that almost all investigated variables have the range of values in the training data seen in the testing dataset. The network was trained using the Bayesian regularization in combination with Levenberg-Marquardt training algorithm. Because to make the training algorithm generally works best, the network inputs and outputs should be scaled to fall approximately in the range [-1,1] (Beale *et al.*, 2012). In this study, the scaling, therefore, was performed for each variable to make its value ranged from -1 to +1. The model performance was determined by mean square error (MSE) calculated for both training dataset and testing dataset. The MATLAB software was used as a computational tool for developing the NN models.

The selection of input variables entered in the NN models should be carefully considered. A larger number of input variables leads to a more complex model and requires a bigger size of the training dataset to prevent model overfitting. Therefore, it is necessary to set criteria to judge whether a variable should be entered in the NN models. In the present study, the eligible variables were purposely determined as those that were included in at least one of the final models developed by the regression approaches (see Table 3 and Table 5), and those that were significant variables in the simple regression models with one of the output variables. Consequently, only 11 variables representing street characteristics each included in one or more models in Table 3 and Table 5 were satisfied these criteria. These variables, therefore, were selected as input variables with the orders when entering in the NN models as shown in Table 7.

Table 7. List of input variables for NN models

No	Variable	Code
1	Carriageway width (m)	P_1
2	Right safety strip width (m)	P_2
3	Length of street section (m)	P_3
4	Distance from the entrance of exiting intersection to the nearest control point (m)	P_4
5	Roadside object density (per 100 m)	P_5
6	Roadway width ratio between crossing street and study street	P_6
7	Number of lanes	P_7
8	Sidewalk indicator (1 if sidewalks are available on both sides; 0 otherwise)	P_8
9	Width of crossing street (m)	P_9
10	Distance from the entrance to the centre point of exiting intersection (m)	P_{10}
11	Type indicator of exiting intersection (1 if 3-leg intersection or similar; 0 otherwise)	P_{11}

The network architecture indicated by the number of hidden layers and the number of nodes in each hidden layer affects to the model accuracy and the model complexity. Because it is difficult to appropriately select the number of hidden layers and nodes in advance, this study conducted numerous trials to find the architecture that has the best model performance. For each model specification, network was repeatedly trained a number of times with different initial weights to enhance the probability of archiving the global minima. The results showed that the models for predicting $V_{max,av}$ and $V_{max,85}$ works best with the architecture with one hidden layer and three nodes in the layer, while those for predicting $V_{in,av}$ and $V_{in,85}$ has the best architecture with one hidden layer and two nodes in the layer. The parameters of the final NN models are shown below.

<u>Model 1:</u> Output $V_{in,av}$

$$W_{ij} = \begin{bmatrix} -0.2101 & -0.0513 & 0.0053 & -0.2093 & -0.0462 & -0.0560 & 0.1547 & -0.0280 & -0.0642 & -0.0066 & -0.1294 \\ 0.3350 & 0.2950 & -0.2512 & -0.0779 & -0.2943 & -0.3538 & 0.1726 & 0.2425 & -0.4030 & 0.1572 & 0.0703 \end{bmatrix} \quad (8)$$

$$b_j = \begin{bmatrix} -0.0871 & 0.5449 \end{bmatrix}^T \tag{9}$$

$$W_i = \begin{bmatrix} -0.4062 & 0.7732 \end{bmatrix} \tag{10}$$

$$b_0 = 0.0950 \tag{11}$$

<u>Model 2:</u> Output $V_{in,85}$

$$W_{ij} = \begin{bmatrix} 0.2014 & 0.0362 & 0.0394 & 0.1379 & 0.0235 & 0.0124 & -0.2775 & 0.0128 & 0.0121 & -0.0519 & 0.1602 \\ -0.3976 & -0.2900 & 0.2082 & 0.0520 & 0.2472 & 0.1963 & -0.1763 & -0.2558 & 0.2360 & -0.1094 & 0.0070 \end{bmatrix} \tag{12}$$

$$b_j = \begin{bmatrix} 0.0886 & -0.5847 \end{bmatrix}^T \tag{13}$$

$$W_i = \begin{bmatrix} 0.4307 & -0.7204 \end{bmatrix} \tag{14}$$

$$b_0 = 0.1592 \tag{15}$$

<u>Model 3:</u> Output $V_{max,av}$

$$W_{ij} = \begin{bmatrix} 0.1420 & 0.3355 & 0.1944 & -0.1626 & 0.1432 & 0.2103 & 0.4913 & 0.5815 & -0.1910 & 0.5423 & 0.0317 \\ 0.0764 & 0.2815 & -0.3483 & 0.3039 & 0.1054 & -0.4027 & 0.2750 & 0.1993 & -0.3927 & -0.1358 & -0.2814 \\ 0.5096 & 0.0477 & 0.0372 & 0.4709 & -0.1966 & -0.5664 & -0.2151 & -0.1172 & -0.1046 & -0.6993 & -0.1336 \end{bmatrix} \tag{16}$$

$$b_j = \begin{bmatrix} 0.4593 & -0.1549 & 0.3500 \end{bmatrix}^T \tag{17}$$

$$W_i = \begin{bmatrix} 0.9582 & -0.6514 & 0.7955 \end{bmatrix} \tag{18}$$

$$b_0 = -0.1789 \tag{19}$$

<u>Model 4:</u> Output $V_{max,85}$

$$W_{ij} = \begin{bmatrix} 0.2370 & -0.0480 & 0.2917 & -0.0121 & -0.4526 & -0.1181 & -0.1519 & -0.3494 & 0.1208 & -0.4391 & -0.1248 \\ -0.0075 & -0.1694 & 0.2064 & -0.0110 & 0.3586 & 0.0741 & -0.2155 & -0.1495 & 0.0189 & -0.1358 & 0.0338 \\ -0.4184 & -0.2332 & -0.1964 & 0.0665 & 0.0426 & -0.2577 & -0.3984 & -0.8143 & 0.1898 & -0.3105 & -0.1951 \end{bmatrix} \tag{20}$$

$$b_j = \begin{bmatrix} 0.1060 & 0.3388 & -0.5270 \end{bmatrix}^T \tag{21}$$

$$W_i = \begin{bmatrix} 0.6419 & 0.5914 & -1.0108 \end{bmatrix} \tag{22}$$

$$b_0 = -0.2298 \tag{23}$$

The same indicators used to evaluate the performance of operating speed models developed by regression approaches were used for NN models. These indicators are provided in Table 8.

Table 8. Performance of NN models

Performance indicator	$V_{in,av}$	$V_{in,85}$	$V_{max,av}$	$V_{max,85}$
With dataset for model development				
Mean squared error (MSE), $(km/h)^2$	5.46	7.70	1.23	2.21
Root mean square error (RMSE), km/h	2.34	2.78	1.11	1.49
Mean absolute error (MAE), km/h	1.78	2.11	0.90	1.18
% sites with absolute error (AE) ≤ 2.5	74.12	74.12	97.65	90.59
% sites with 2.5 < AE ≤ 5	21.18	17.65	2.35	9.41
% sites with 5 < AE ≤ 10	4.71	8.24	0.00	0.00
% sites with 10 < AE ≤ 12	0.00	0.00	0.00	0.00
With testing dataset				
Mean squared error (MSE), $(km/h)^2$	3.17	6.07	2.23	4.71
Root mean square error (RMSE), km/h	1.78	2.46	1.49	2.17
Mean absolute error (MAE), km/h	1.53	2.00	1.24	1.80
% sites with absolute error (AE) ≤ 2.5	92.86	64.29	85.71	78.57
% sites with 2.5 < AE ≤ 5	7.14	35.71	14.29	21.43
% sites with 5 < AE ≤ 10	0.00	0.00	0.00	0.00
% sites with 10 < AE ≤ 12	0.00	0.00	0.00	0.00

The validity of the NN models also was evaluated through conducting a parametric analysis to obtain the relationship between the output variables and each of the selected input variables. To examine the influence of a given input variable on a specific output variable, the input variable was varied while the other input variables were kept to constant values. For continuous variables, these constant values were determined as their average values of the respective variables calculated separately for groups of streets with one-lane carriageway and those with two-lane carriageway. The groups of one-lane streets and the two-lane streets were investigated separately because these two groups are significantly different in some aspects such as the range of carriageway width. For the dummy variables P_8 and P_{11}, the constant values were both set at 0 (i.e., streets with no sidewalk or with sidewalk on only one side and the existing intersection is a 4-leg intersection). The results of the parametric analysis are shown in Figure 3 and Figure 4.

4. DISCUSSION

The operating speed models developed by regression approaches had a coefficient of determination (R^2) ranged from 0.47 to 0.58 for SER models and from 0.56 to 0.70 for SEA models. This research is one of the rare studies that developed speed models by one dataset then validated them by another one. The developed models demonstrated a good performance on the validation dataset. For example, as shown in Table 6, the mean absolute error of the models developed by SEA ranged from 1.38 km/h to 2.03 km/h for the validation dataset and all of validated sites had an absolute error of less than 5.0 km/h in which from more than 70% to 85% of those sites had an absolute error of less than 2.5 km/h. In addition, the signs of all estimated coefficients of the developed models could be expected as a brief interpretation shown in Table 9. Collectively, the regression models owned a reasonable fit and it is reliable to use these models for estimating operating speeds on residential streets with a 30 km/h speed limit. It can be seen from these models, various roadway and roadside features were found as determinants of driving speeds. From the road design perspective, the results suggest that attention should be paid on the selection of street section length, the allocation of cross-section elements, and the characteristics of intersections to obtain desired driving speeds. The developed models also can be used to assess the speeding issues in existing streets, and/or to evaluate street designs regarding the intended operating speed goals.

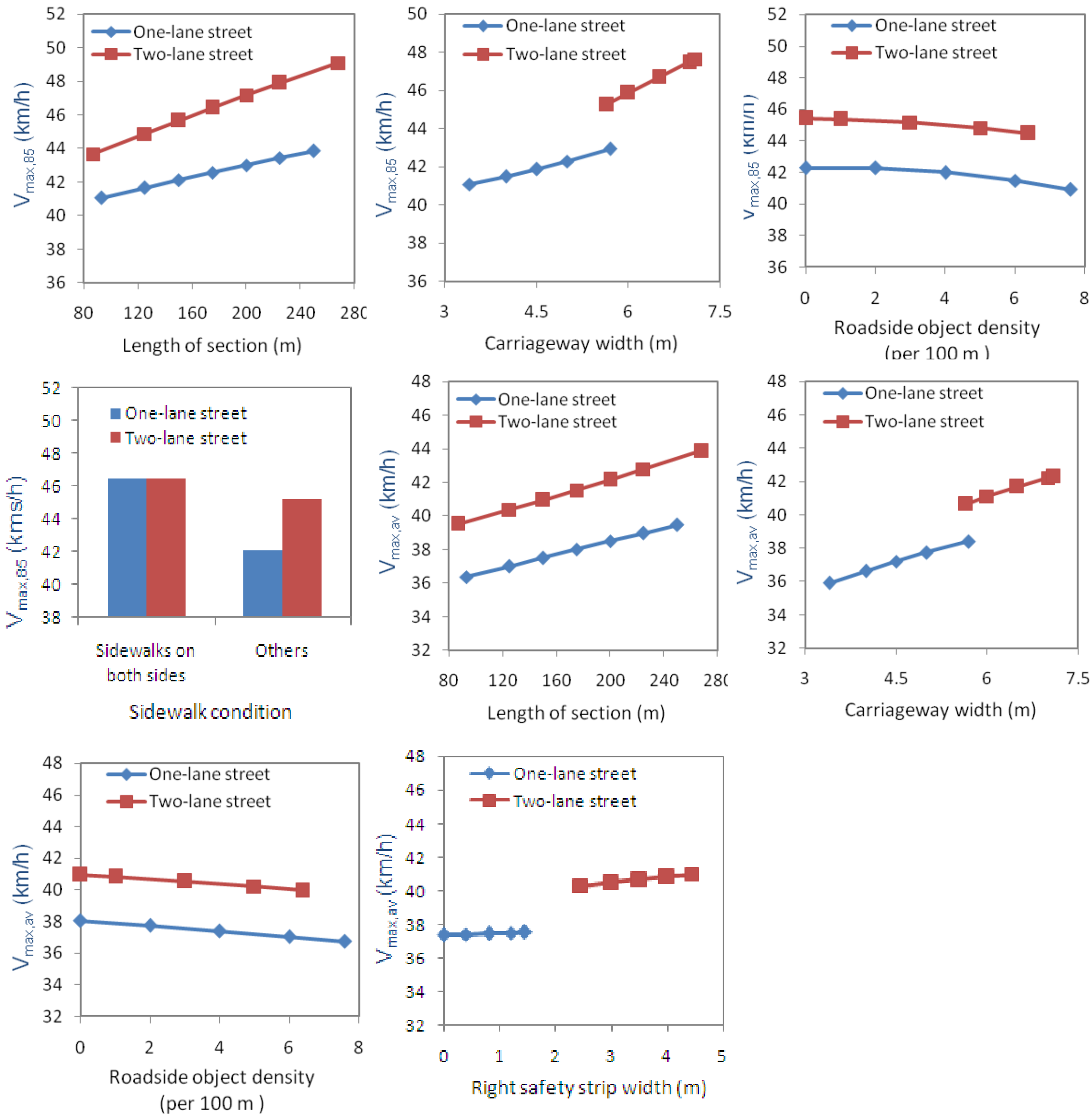

Figure 3. Variation of maximum speeds with selected street characteristics based on NN models

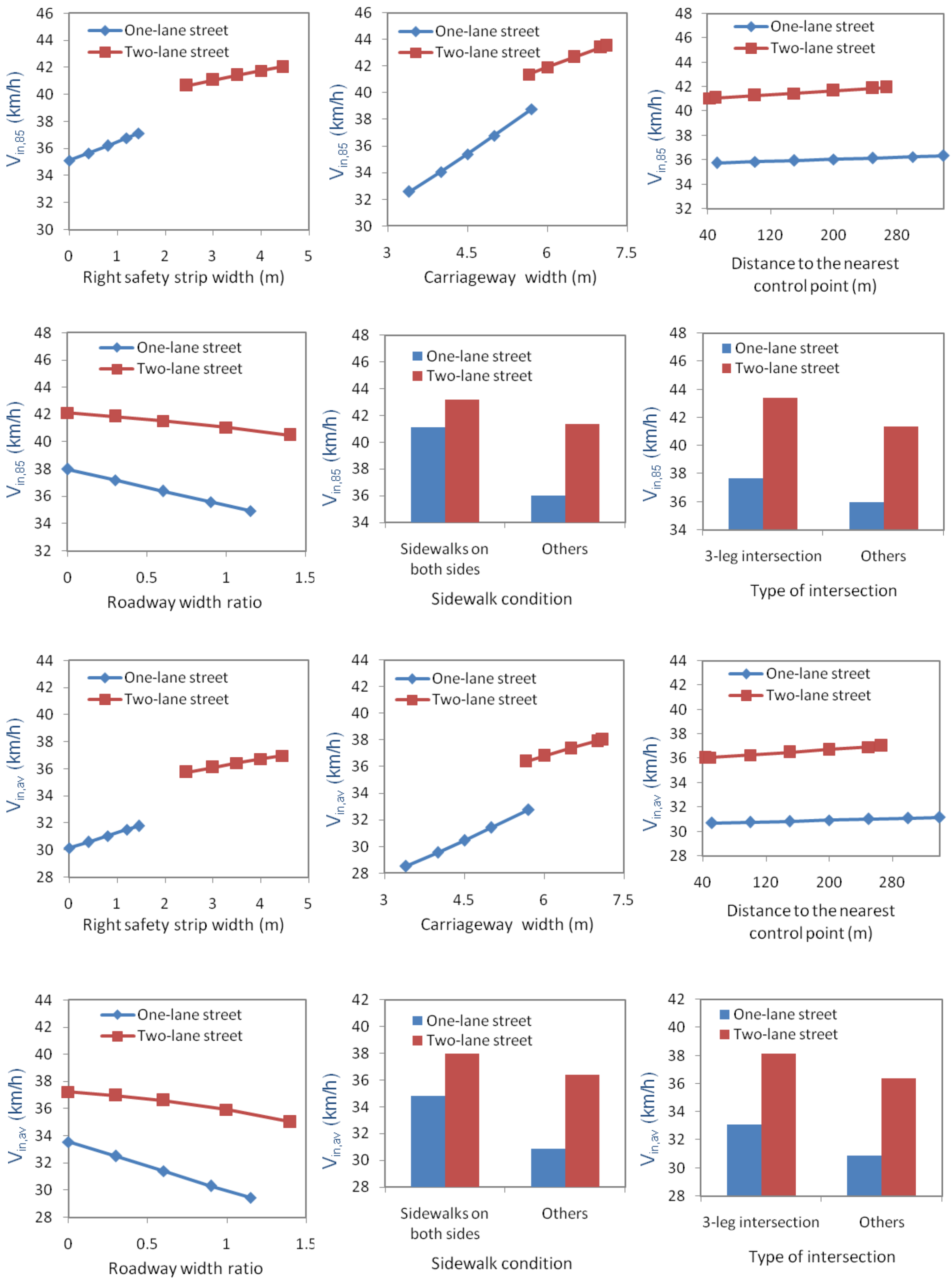

Figure 4. Variation of speeds at intersection with selected street characteristics based on NN models

Comparing the two regression approaches in this study, it was found that both methodologies yielded comparable results regarding models for estimating maximum speeds obtaining within a street section. The dependent variables in these models developed by SER are exactly the same as those in respective models developed by SEA. In addition, similar values of coefficients of determination could be found in the respective models by the two methods (see Table 3 and Table 5). Other model performance indicators in Table 4 and Table 6 also have comparable figures. One possible explanation for these comparable results is that, maximum speeds obtaining within a street section may not depend on speeds at other locations on the street. The argument is reinforced by the fact that speeds at intersection were not significant variables in the SEA models for predicting maximum speeds within a section.

With regards to models for estimating speeds at intersection, as could be seen in Table 3 and Table 5, more street characteristics were included in the models by SEA compared to those by SER. This fact demonstrates that the SEA is a better way to detect the speed influencing factors from street features. Regarding model performance, the coefficients of determination (R^2) of the models developed by SEA are significantly higher than those by SER. All performance indicators of the SEA models calculated based on the dataset for model development (see Table 4 and Table 6) are more favorable than those of the respective models by SER method. The performance on the validation dataset of the model developed by SEA for predicting 85th percentile speed at intersection is significantly better than that built by SER. As an illustration, the mean squared error of the model by SEA is 5.43 well lower than that of 9.82 for the model by SER. While the SEA model produced no site with absolute error larger than 5.0, the figures of 14.29% of sites could be seen in the model by SER. Although the model by SER for estimating mean speed at intersection performed on validation dataset slightly better than that by SEA, the performance indicators of the latter model still indicated a good predictive ability because the mean absolute error is only 1.76 km/h and no site has the absolute error over 5.0 km/h, while the absolute errors of 71.4% of sites are under 2.5 km/h. Furthermore, it should be noted that the high correlations between maximum speeds and speeds at intersection found in this study means that ignoring the endogenous relationship between the two speeds may consequently lead to bias model parameters and do not truly represent the mechanism regarding the influences of street features on driving speeds. In summary, it could be concluded that, in general SEA produces better modeling results than the conventional regression. This result suggests that SEA should be used instead of conventional regression for modeling speeds at several locations on short-length street sections similar as those in the current study.

This study also developed operating speed models by using NN approach. The NN models showed a very good predictive ability as demonstrating by the model performance indicators in Table 8. For example, on the testing dataset, the NN models had a mean absolute error ranged from 1.24 (km/h) to 2.00 (km/h) and there is no site with absolute error exceeding 5.0 (km/h). Compared to models by regression approaches, only the model for estimating 85th percentile speed at intersection by the SEA showed a slightly better performance on the testing dataset than that by NN method. The performance of other NN models on both model-development dataset and model-testing dataset outreaches that of regression models. In general, it could be concluded that with regards to predictive ability, the NN models have a trend to perform better than those by regression methods. This result is consistent with most previous studies on transportation field as noted by Karlaftis and Vlahogianni (2011).

The NN models in this study were also validated by conducting an input sensitivity analysis to investigate the influence of variables in the models on the model outputs. The results in Figure 3 and Figure 4 clearly showed the similar trends regarding to the effects of

Table 9. Speed-influencing variables from regression models

No	Variable	SER model	SEA model	Interpretation and/or possible explanation
1	Carriageway width (m)	$V_{max,85}$ $V_{in,85}$ $V_{max,av}$	$V_{max,85}$ $V_{max,av}$	An increase carriageway width led to an increase of speed because more room is provided for manoeuvring.
2	Right safety strip width (m)	$V_{max,av}$	$V_{max,av}$	A wider right safety strip width resulted in a higher speed as expected.
3	Length of street section (m)	$V_{max,85}$ $V_{max,av}$	$V_{max,85}$ $V_{max,av}$	A longer length resulted in higher speeds because a longer length provides more space for acceleration before reaching the maximum speed.
4	Length of street section (m)		$V_{in,85}$ $V_{in,av}$	A longer length resulted in lower speeds at intersection possibly because a longer length leads to a longer deceleration distance after reaching the maximum speed that consequently leads to a lower speed at intersection.
5	Distance from the entrance of exiting intersection to the nearest control point (m)	$V_{in,av}$	$V_{in,85}$	Drivers selected a higher speed with a higher value of the distance possibly because more information about the road ahead is provided.
6	Roadside object density (per 100 m)	$V_{max,85}$ $V_{max,av}$ $V_{in,av}$	$V_{max,85}$ $V_{max,av}$	A higher density of roadside object associated with a lower driving speed because the presence of such objects decreases the effective carriageway width and creates potential hazards to drivers.
7	Roadway width ratio between crossing street and study street		$V_{in,85}$	A smaller of the ratio resulted in higher speeds because it suggests that the crossing street is more minor compared to the study street.
8	Number of lanes	$V_{max,85}$	$V_{max,85}$	Streets with two lanes had higher speeds than one-lane streets because more room is provided for manoeuvring.
9	Sidewalk indicator (1 if sidewalks are available on both sides; 0 otherwise)	$V_{max,85}$ $V_{in,85}$	$V_{max,85}$	Sidewalk available on both sides resulted to higher speeds because pedestrians/cyclists activities are excluded from the roadway.
10	Width of crossing street (m)		$V_{in,av}$	A wider width of crossing street associated with a lower speed at intersection as expected.
11	Distance from the entrance to the centre point of exiting intersection (m)		$V_{in,av}$	A longer of the distance resulted in a higher speed because people are likely to drive slower when approaching the centre area of an intersection.
12	Type indicator of exiting intersection (1 if 3-leg intersection or similar; 0 otherwise)	$V_{in,85}$ $V_{in,av}$	$V_{in,av}$	As expected drivers selected a higher speed at a 3-leg intersection compared to a 4-leg intersection.
13	85th percentile speed of tangent		$V_{in,85}$	As expected a higher maximum speed resulted in a higher speed at intersection
14	Mean speed of tangent		$V_{in,av}$	As expected a higher maximum speed resulted in a higher speed at intersection

street characteristics on driving speeds as compared to those from regression models illustrated in Table 9. This fact proves a good generalization of the NN models and leads to a conclusion that it is reliable to use these models for predicting operating speeds on residential streets with a 30 km/h speed limit. However, it should be noted that the number of variables in the NN models outnumbers those in the regression models that means that the NN models are more complicated and requires more efforts to collect input data for prediction.

The good results of the NN models could be partly explained by the procedure for selecting input variables. Different with previous studies that often arbitrarily select input variables for NN models, this research used only variables those significantly affected driving speeds as detected by regression models. This approach is likely an effective way to reduce the redundant input variables while still remain the core information from the used dataset. It should be noted that a larger number of input variables required a larger size of the dataset and it may lead to an overfitting issue. The good performance of the NN models in the present study, therefore, suggests that it should be used regression methods complimentarily in order to select input variables for achieving desirable results for NN models especially when the sample size is rather small.

5. CONCLUSIONS

An attempt has been made in this paper to compare the performance of regression methods and Neural Networks (NN) approach to model operating speeds on residential streets with a 30 km/h speed limit. The regression models were developed and then validated/tested with a different dataset to confirm their predictive ability. The current research has proved that the Simultaneous Equation Approach (SEA) produced a better result compared to the conventional regression on modeling simultaneously maximum speeds obtaining within a street section and speeds at the entrance to the next un-signalised intersection. The issue of endogenous relationships between the two speeds found in this study could not be solved by the SER and it may invalid the resultant models. However, the issue could be well addressed by using the SEA. This finding suggests that compared to the conventional regression method, SEA is a more effective way and should be used for modeling speeds at different locations especially on short-length streets similar as those in this study.

Meanwhile, the NN models developed in this study showed a very good ability for prediction and outreached those developed by regression methods although the NN models are more complicated and require more input data. The good results on the performance of the NN models probably because in this study the selection of input variables for NN models was implemented by using a combination between regression methods and a NN technique in which input variables were selected based on significant variables found in regression equations. The implication, therefore, is that it is possible to achieve better results on modeling operating speeds by using complementarily an appropriate regression technique and a NN technique following a procedure as described in this paper.

Finally, while the speed data in the current study indicated that speed regulation is seriously violated on residential streets, the developed models can be applied as a useful tool for tacking this speeding problem. These applications may include assessing the extent of speeding violation in existing streets, re-designing streets to make them calmer, and planning and designing new urban streets to meet the intended operating speed goals.

REFERENCES

Ali, A.T., Flannery, A., Venigalla, M.M. (2007) Prediction models for free flow speed on urban streets. Presented in the Transportation Research Board 86th Annual Meeting, Washington, D. C.

Beale, M.H., Hagan, M.T, Demuth, H.B. (2012) *Neural Network ToolboxTM User's Guide R2012b*. The MathWorks, Inc.

Bonneson, J.A. (1999) Side friction and speed as controls for horizontal curve design.

Journal of Transportation Engineering, American Society of Civil Engineers, 125 (6) 473-480.

Dinh, D.D., Kubota, H. (2013) Profile-speed data-based models to estimate operating speeds for urban residential streets with a 30 km/h speed limit. *IATSS Research*, 36, 115-122.

Fitzpatrick, K., Carlson, P., Brewer, M., Wooldridge, M. (2001) Design factors that affect driver speed on suburban streets. *Transportation Research Record: Journal of Transportation Research Board*, 1751, 18-25.

Fitzpatrick, K., Carlson, P., Brewer, M., Wooldridge, M., Miaou, S.-P. (2003) *NCHRP Report 504: Design Speed, Operating Speed, and Posted Speed Practices*. Transportation Research Board of the National Academies, Washington, D.C.

Fitzpatrick, K., Shamburger, C.B., Krammes, R.A., Fambro, D.B (1997) Operating speed on suburban arterial curves. *Transportation Research Record: Journal of Transportation Research Board*, 1579, 89-96.

Gong, H., Stamatiadis, N. (2008) Operating speed prediction models for horizontal curves on rural four-lane highways. *Transportation Research Record: Journal of Transportation Research Board*, 2075, 1-7.

Greene, W. (2003) *Econometric Analysis, 5th edition*. Prentice Hall.

Karlaftis, M.G., Vlahogianni, E.I. (2011) Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C*, 19, 387-399.

McFadden, J., Yang, W., Durrans, S.R. (2001) Application of Artificial Neural Networks to Predict Speeds on Two-Lane Rural Highways. *Transportation Research Record: Journal of Transportation Research Board*, 1751, 9-17.

OECD/ECMT (2006) *Speed management*. Joint Transport Research Centre of the Organization for Economic Co-operation and Development and the European Conference of Ministers of Transport.

Poe, C.M., Mason, J.M. (2000) Analyzing influence of geometric design on operating speeds along low-speed urban streets: mixed-model approach. *Transportation Research Record: Journal of Transportation Research Board*, 1737, 18-25.

Singh, D., Zaman, M., While, L. (2012) Neural network modeling of 85th percentile speed for two-lane rural highways. *Transportation Research Record: Journal of Transportation Research Board*, 2301, 17-27.

Tarris, J.P., Poe, C.M., Mason, J.M., Goulias, K.G. (1996) Predicting operating speeds on low-speed urban streets: regression and panel analysis approaches. *Transportation Research Record: Journal of Transportation Research Board*, 1523, 46-54.

Taylor, D.R., Muthiah, S., Kulakowski, B.T., Mahoney, K.M., Porter, R.J. (2007) Artificial neural network speed profile model for construction work zones on high-speed highways. *Journal of Transportation Engineering*, 133(3), 198-204.

Transportation Research Board (2011) *“Modeling operating speed” Transportation Circular EC151*. Operational Effects of Geometrics Committee, Transportation Research Board, Washington, D.C.

Wang, J., Dixon, K.K., Li, H., Hunter, M.P. (2006) Operating-speed model for low-speed urban tangent streets based on in-vehicle global positioning system data. *Transportation Research Record: Journal of Transportation Research Board*, 1961, 24-33.

Zaman, M., Issa, R., Najjar, Y. (2000) Modeling the 85th percentile speed on Oklahoma two-lane rural highways via neural network approach. *Computational Intelligence Applications in Pavement and Geotechnical Systems*, 99-108.