Traffic Volume Prediction With Segment-Based Regression Kriging and its Implementation in Assessing the Impact of Heavy Vehicles

Yongze Song, Xiangyu Wang, Graeme Wright, Dominique Thatcher, Peng Wu, and Pascal Felix

***Abstract*— Geostatistical methods have been widely used for spatial prediction and the assessment of traffic issues. Most previous studies use point-based interpolation, but they ignore the critical information of the road segment itself. This can lead to inaccurate predictions, which will negatively affect decision making of road agencies. To address this problem, segment-based regression kriging (SRK) is proposed for traffic volume prediction with differentiation between heavy and light vehicles in the Wheatbelt region of Western Australia. Cross validations reveal that the prediction accuracy for heavy vehicles is significantly improved by SRK ($R^2 = 0.677$). Specifically, 78% of spatial variance and 53% of estimated uncertainty are improved by SRK for heavy vehicles compared with regression kriging, a best performing point-based geostatistical model. This improvement shows that SRK can provide new insights into the spatial characteristics and spatial homogeneity of a road segment. Implementation results of SRK-based predictions show that the impact of heavy vehicles on road maintenance is much larger than that of light vehicles and it varies across space, and the total impacts of heavy vehicles account for more than 82% of the road maintenance burden even though its volume only accounts for 21% of traffic.**

***Index Terms*— Geostatistics, segment-based regression kriging, traffic prediction, road maintenance.**

I. Introduction

ACCURATE spatial prediction of traffic volumes is of great importance for transportation analysis, planning and decision making. A wide range of methods have been applied for spatial prediction issues in the fields of road and traffic planning utilizing uncounted traffic data of road segments [1]–[3]. These studies can be divided into two categories: statistical approaches, such as multiple regression [4], [5], time-series analysis [6], [7], the U.S. Federal Highway Administration (FHWA) procedure [8], [9], machine learning algorithms [10] and geostatistical methods [11], and predictions with the help of image data such as light detection and ranging (LiDAR) data [12] and high-resolution remote sensing data [13], [14].

Among these approaches, geostatistical or kriging-based methods are very advantageous in providing insights into traffic behaviors across large spatial scales by exploring the spatial local correlations. These methods are a series of the best linear unbiased estimators for spatial data with expected bias of zero and minimized expected interpolation error, and could provide both prediction and its estimated uncertainty. Thus, compared with other models, their predictions tend to be more accurate. Comparison studies reveal that universal kriging (UK) performs much better than non-spatial multiple regression [1], [15] and has a small improvement over geographically weighted regression [11] in the spatial prediction of traffic volumes. In addition, traffic data, such as traffic speed and volume, are spatially continuous and autocorrelated, which means that traffic conditions at adjacent road segments are usually identical or similar. Traffic congestions may appear at road intersections especially in regions with dense human activities, and diffuse spatially leading to regional congestions [1]. This phenomenon is explained as spatial autocorrelation that could be explored by geostatistical methods [16]. For instance, geostatistical methods have been applied on traffic prediction issues including traffic count estimation [11], [15], speed prediction of the traffic system [17], traveling time estimation [18], congestion analysis [19] and incidents assessment [20].

However, all these studies utilize point-based spatial interpolation and they are not straightforward in traffic prediction. Actually, point-based methods are dominant for the interpolation of spatially continuous data over the areas of interest, since spatial distribution data are often collected from point sources [21]–[24]. But these methods simplify lines of road segments with various shapes and lengths as single points. This simplification process enables simple traffic modelling, but ignores the spatial characteristics and spatial homogeneity of lines of road segments. Ignoring these elements leads to

Manuscript received May 16, 2017; revised October 2, 2017 and February 2, 2018; accepted February 11, 2018. Date of publication March 7, 2018; date of current version December 21, 2018. The Associate Editor for this paper was P. Wang. This work was supported in part by the Australia Research Council Discovery Early Career Researcher Award under Project DE170101502 and in part by the Australian Research Council Linkage Project under Project LP140100873. *(Corresponding author: Yongze Song.)*

Y. Song and X. Wang are with the Australasian Joint Research Centre for Building Information Modeling, School of Built Environment, Curtin University, Perth WA 6845, Australia (e-mail: yongze.song@postgrad.curtin.edu.au; xiangyu.wang@curtin.edu.au).

G. Wright is with the Office of Research and Development, Curtin University, Perth WA 6845, Australia (e-mail: g.l.wright@curtin.edu.au).

D. Thatcher is with Heavy Vehicle Services, Main Roads Western Australia, Perth WA 6986, Australia (e-mail: dom.thatcher@mainroads.wa.gov.au).

P. Wu is with the Department of Construction Management, School of Built Environment, Curtin University, Perth WA 6845, Australia (e-mail: peng.wu@curtin.edu.au).

P. Felix is with Heavy Vehicle Services, Main Roads Western Australia, Perth WA 6986, Australia, and also with the WSP Global Inc., Perth WA 6000, Australia (e-mail: pfelix@pb.com.au).

Digital Object Identifier 10.1109/TITS.2018.2805817

inaccurate predictions which can negatively affect decision making by road agencies. Thus, a segment-based interpolation method is necessary for predicting uncounted traffic data.

Previous studies have attempted to develop geostatistical methods for data with a variable spatial support instead of traditional point support or a regular pixel support of remotely sensed images. These methods include area-to-point kriging (ATPK) for estimation from area data to point data [25], area-to-area kriging (ATAK) for estimation of different areas [26], top-kriging for estimation of data with variable spatial characteristics [27], and their relevant development such as area-to-area regression kriging (ATARK) [28] and area-to-point poison kriging [29]. They have been utilized for addressing interpolation problems with irregular spatial support from runoff of river networks [27], [30], geographical upscaling and downscaling [28], [31], [32], population estimation [33] and mapping disease and health data [29], [34], [35]. However, there are few research attempts to develop a geostatistical interpolation model for traffic prediction with the integration of the irregular shape of road segments.

As an important continuous and regular construction task, cost effective road maintenance greatly relies on accurate traffic volume predictions. Road maintenance becomes increasingly critical for social and economic development, especially for safe, accessible and serviceable travel and freight transportation. The primary objectives of maintenance decisions of a road network include determining the road segments to be repaired, the repair periods and the treatment strategies [36]. Among these works, road segment based maintenance burden analysis is a key evidence and foundation for engineering studies on road damage [37], evaluation of environmental impacts [38] and construction management of road maintenance [39]. The gradual cracking and pulverization of road surfaces is usually caused by the cumulative vehicle masses, especially trucks or heavy vehicles [40], [41], and the penetration of the road surface by saline groundwater [42]. In addition, road damage varies by locations, since the degree and speed of the impacts of cumulative vehicle masses vary on different roads that primarily serve passenger or freight transportation, and the impacts of groundwater salinization are highly variable across space depending on the local volume of water and the age and composition of road materials. Previous studies have stressed the importance of burden estimation of road maintenance due to its link with maintenance cost, but few of them accurately and geographically predict the burden distributions at segment level.

Therefore, this study employs segment-based ordinary kriging (SOK) and segment-based regression kriging (SRK), which are developed from point-based ordinary kriging (OK) and regression kriging (RK), for the traffic data with variable road segment support in the Wheatbelt region of Western Australia (WA). By borrowing ideas from ATAK and top-kriging, SOK and SRK integrate the spatial characteristics of road segments and the spatial homogeneity of each single segment, consider their spatial autocorrelation and enable segment-based data to compute the best linear unbiased estimation. Regression kriging (RK or SRK) is an effective supplement of ordinary kriging (OK or SOK) because it considers the information of covariates to deal with the non-stationary of random functions [28]. Both segment-based and point-based models are applied on the estimation of diverse types of traffic volumes, including heavy, light or total vehicles. A comparative study of both kinds of models, together with point-based inverse distance weighting (IDW) and universal kriging (UK), and non-spatial linear regression (LR), will give an insight of different traffic behaviors and provide road agencies with proper models that have the best prediction performance. The traffic volumes predicted by segment-based geostatistical models are applied on the assessment of road maintenance burden in the Wheatbelt, WA, which can help provide quantitative and accurate evidence for road asset management. Road maintenance burden is determined by the integration of predicted traffic volumes and restricted access vehicles (RAVs) network based vehicle mass estimation.

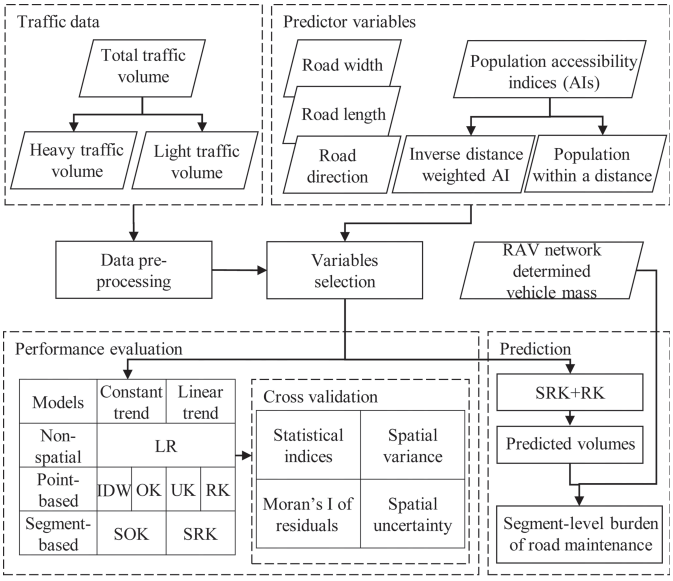


Fig. 1. Schematic overview of predicting the distribution of road maintenance burden.

II. Methodology

Fig. 1 shows the schematic overview of predicting traffic volumes with a comparison of segment-based and point-based models. The study process includes data pre-processing, variables selection, performance evaluation with cross validation for different categories of models and traffic volume prediction. As an implementation, the prediction results are applied to evaluation of the distribution of road maintenance burden in the Wheatbelt, WA.

A. Data Pre-Processing and Transformation

For both segment-based and point-based spatial prediction models, the dependent variable is a transformed normally distributed traffic volume of heavy vehicles, light vehicles or total vehicles due to the skewed distributed raw data. Box-Cox transformation is commonly used to remove skewed distributions and ensure stabilizing variations of traffic data [5], [11].

Observed traffic data Y is transformed by the maximum estimation of Box-Cox likelihood function $g_\lambda(Y)$ over a power variable λ [43], [44] and the transformed data Y_t is:

$$Y_t = g_\lambda(Y) = \begin{cases} \frac{1}{\lambda}\left(Y^\lambda - 1\right) & \lambda \neq 0 \\ \ln(Y) & \lambda = 0 \end{cases} \quad (1)$$

The maximum likelihood estimations of λ are -0.061, 0.101 and 0.141 for heavy, light and total vehicles respectively. Box-Cox transformation is performed by R*FitAR* package [45].

B. Variables Generation and Selection

Widths, lengths and directions of road segments are used as a predictor variable for both point-based and segment-based models. Population accessibility indices (AIs) are utilized to depict the accessing population within certain nearby buffer regions of count points or road segments. AIs used in this paper include the sum of inverse distance weighted population (WAI) within a given distance of count locations or along road segments, and the population within a given distance (DAI) that determines a maximum correlation with traffic volumes. WAI at location or segment u is computed by:

$$WAI_u = \sum\nolimits_k \frac{Population_k}{d_{uk}^\theta} \quad \forall k, d_{uk} \leq d_{max} \quad (2)$$

where $Population_k$ is the population within buffer region k, d_{uk} is the distance from count location or segment to the buffer region, d_{max} is the maximum band distance (50 km in this paper) and θ is a power parameter that ensures the maximum correlation between WAI and traffic volumes. Repeated computation of WAI and its correlation with traffic volumes derives that θ equals to 1.4, 1.6 and 1.6 for heavy, light and total vehicles respectively. To determine the distances of DAIs, the correlations between population within 5 km to 50 km (in increments of 5 km) and traffic volumes are computed for both point and segment observations of heavy, light and total vehicles. Results show that maximum correlation for heavy vehicle appears at 50 km buffer regions and that for light and total vehicles is at 15 km buffer regions. Step-wise linear regression is used to select predictor variables with significant correlations with traffic volumes and remove insignificantly correlated variables and variables with multicollinearity. Variables selected for predicting heavy vehicle volumes include road segment width, WAI (50 km and $\theta = 1.4$) and DAI (50 km), and those for the volume prediction of light vehicles and total vehicles are road segment width and DAI (15 km).

C. Segment-Based Geostatistical Modelling

SOK and SRK are proposed to predict traffic volumes at uncounted road segments that are characterized as linear road surfaces with various shapes and lengths. Segment-based geostatistical interpolations are primarily based on **three assumptions.** The **first** is the traffic data are assumed to be spatially homogeneous within a segment, but they are spatially heterogenous and autocorrelated for different segments. In general, a road segment is the specific representation of a portion of a road with uniform characteristics between two junctions or intersections [46]. It is usually defined by a series of rules such as same surface, width, number of lanes, pavement age and traffic conditions [46], [47]. Since transport authorities and researchers prefer to monitor traffic data (e.g. annual average daily traffic (AADT)) with the spatial unit of road segment, this assumption ensures their convenient analysis, management and decision making based on the spatial predictions. **Second**, observed traffic data are regarded as the output of a continuous traffic process across space. **Finally**, the spatial stationarity, a general geostatistical assumption, is also assumed for segment-based models that the expected variance between observations (or residuals for SRK) is a function of separation distance. SRK is a supplement of SOK by considering the information of covariates to deal with the nonstationary of random functions.

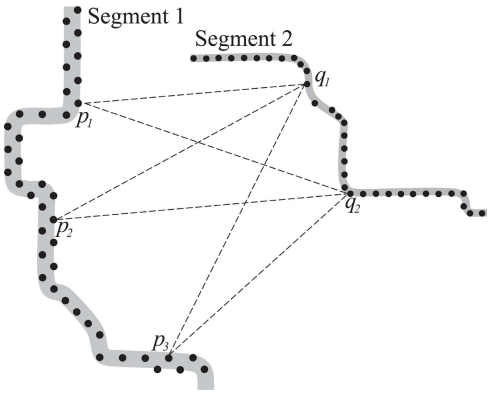


Fig. 2. Discretization of two road segments and their covariance with discretized points.

Using SRK to estimate traffic volumes includes the following **five steps**. The **first step** is modelling trends of traffic data using a linear regression where the dependent variables are transformed by Box-Cox function and the selected predictor variables.

The **second step** is to estimate residuals by removing trends in step 1. Based on the assumption of the second-order stationarity of residuals, the segment-based variogram is assessed by maximum likelihood estimation with a great many point values discretized from segments. The covariance calculation between any two segments with the discretized points at each segment is shown in the sketch map Fig. 2. The discretized points are linearly distributed along road segments which are generally several kilometers long and a few meters wide. A variogram is then determined by the parameters of nugget, sill, range and shape. A proper shape is selected from exponential, spherical and Gaussian functions by the comparison of the sum of squared errors (SSE) of model fitness. Thus, the segment covariance $C_s()$ between any two segments s_i and s_j is calculated with the equation

$$C_s(s_i, s_j) = \frac{1}{N(s_i)} \frac{1}{N(s_j)} \sum\nolimits_{r=1}^{N(s_i)} \sum\nolimits_{t=1}^{N(s_j)} C(p_r, p_t) \quad p_r \in s_i, p_t \in s_j \quad (3)$$

where $N()$ is the number of points derived from the discretisation of a segment and p_r and p_t are the discretised points within s_i and s_j.

The **third step** is to calculate the estimation and error variance of segment-based residuals. The SRK value $\hat{z}()$ at segment s_0 where traffic volumes are not counted is estimated by a linear combination of m neighbouring segments under the assumption of the second-order stationary of residuals by the equation

$$\hat{z}(s_0) = \sum_{i=1}^{m} \omega_i(s_0) z(s_i) \tag{4}$$

where s_i is the road segment that traffic volumes are counted, and $\omega_i(s_0)$ is the weight for $z(s_i)$ at segment s_0. The weights are estimated by

$$\begin{cases} \sum_{j=1}^{m} \omega_j(s_0) C_s(s_i, s_j) \\ \quad + \mu(s_0) = C_s(s_0, s_i), \quad i = 1, 2, \ldots, m \\ \sum_{j=1}^{m} \omega_j(s_0) = 1 \end{cases} \tag{5}$$

and the corresponding error variance of SRK estimation at line s_0 is calculated with the equation

$$\hat{\sigma}^2(s_0) = C_s(s_0, s_0) - \sum_{i=1}^{m} \omega_i(s_0) C(s_0, s_i) - \mu(s_0) \tag{6}$$

The **fourth step** is to generate the estimated traffic volumes by adding SRK estimations to their trends. The **last step** is using cross validation to validate the model performance of SRK by comparing with SOK and point-based interpolation methods, IDW, OK, UK and RK. Point-based interpolations are modelled by R *automap* and *gstat* packages, and segment-based models are done by R *rtop* package.

D. Integration of Segment-Based and Point-Based Predictions

The inverse-variance weighting method is used for the integrations of both SRK and RK derived results due to their respective contributions on diverse traffic behaviors. The inverse-variance weighted average traffic volume $\hat{Y}_{integ}$ and its least estimation variance $\hat{\sigma}^2_{integ}$ are

$$\hat{Y}_{integ} = \frac{\sum_{\tau=1}^{T} \frac{\hat{Y}_\tau}{\sigma_\tau^2}}{\sum_{\tau=1}^{T} \frac{1}{\sigma_\tau^2}} \tag{7}$$

$$\hat{\sigma}^2_{integ} = \frac{1}{\sum_{\tau=1}^{T} \frac{1}{\sigma_\tau^2}} \tag{8}$$

where $\hat{Y}_\tau$ and σ_τ^2 are estimation and variance of model τ ($\tau = 1, \ldots, T$). In this case, predictions and their kriging estimation variance of SRK and RK models are integrated for their combined predictions.

E. Performance Comparison Between Segment-Based and Point-Based Models

Statistical indices include mean error (ME), mean absolute percentage error (MAPE) and the coefficient of determination (R^2). Their respective equations are:

$$\mathrm{ME} = \frac{1}{n} \sum_{\kappa=1}^{n} (O_\kappa - P_\kappa) \tag{9}$$

$$\mathrm{MAPE} = \frac{100}{n} \sum_{\kappa=1}^{n} \left| \frac{O_\kappa - P_\kappa}{O_\kappa} \right| \tag{10}$$

$$R^2 = \frac{\left(\sum_{\kappa=1}^{n} \left(O_\kappa - \bar{O}\right)\left(P_\kappa - \bar{P}\right)\right)^2}{\sum_{\kappa=1}^{n} \left(O_\kappa - \bar{O}\right)^2 \sum_{\kappa=1}^{n} \left(P_\kappa - \bar{P}\right)^2} \tag{11}$$

The mean estimation variance and estimated uncertainty of geostatistical models (OK, SOK, UK, RK and SRK) are also computed to compare the improvement of segment-based models with that of point-based models. Wherein, the inverse Box-Cox transformation of traffic volume $\hat{Y}$ is:

$$\hat{Y} = G_\lambda\left(\hat{Y}_t\right) = \begin{cases} \frac{\left(\hat{Y}_t \lambda + 1\right)^{\frac{1}{\lambda}}}{e^{\hat{Y}_t}} & \lambda \neq 0 \\ \frac{}{} & \lambda = 0 \end{cases} \tag{12}$$

where $\hat{Y}_t$ is the predicted transformed data and $G_\lambda\left(\hat{Y}_t\right)$ is the inverse transformation function. Thus, its estimated standard deviation $\sigma_{\hat{Y}}$ is:

$$\sigma_{\hat{Y}} = \sqrt{\left(\frac{\partial G_\lambda(\hat{Y}_t)}{\partial \hat{Y}_t}\right)^2 \sigma^2_{\hat{Y}_t}} = \frac{\partial G_\lambda(\hat{Y}_t)}{\partial \hat{Y}_t}$$

$$\sigma_{\hat{Y}_t} = \begin{cases} \frac{\left(\hat{Y}_t \lambda + 1\right)^{\frac{1-\lambda}{\lambda}} \sigma_{\hat{Y}_t}}{e^{\hat{Y}_t} \sigma_{\hat{Y}_t}} & \lambda \neq 0 \\ \frac{}{} & \lambda = 0 \end{cases} \tag{13}$$

where $\sigma_{\hat{Y}_t}$ is the standard deviation estimation of $\hat{Y}_t$, and corresponding kriging estimation uncertainty is computed by:

$$\mu_{\hat{Y}} = \frac{\sigma_{\hat{Y}}}{\hat{Y}} \tag{14}$$

Further, the estimated volume of total vehicles $\hat{Y}_{total}$ is that of heavy vehicles $\hat{Y}_{heavy}$ plus that of light ones $\hat{Y}_{light}$, so the estimated standard deviation $\sigma_{\hat{Y}_{total}}$ and estimation uncertainty $\mu_{\hat{Y}_{total}}$ of total vehicle are:

$$\sigma_{\hat{Y}_{total}} = \sqrt{\sigma^2_{\hat{Y}_{heavy}} + \sigma^2_{\hat{Y}_{light}}} \tag{15}$$

$$\mu_{\hat{Y}_{total}} = \frac{\sigma_{\hat{Y}_{total}}}{\hat{Y}_{total}} \tag{16}$$

F. Estimation of Road Maintenance Burden

Segment-based prediction of traffic volumes is implemented on the estimation of road maintenance burden with the integration of the RAV network based vehicle mass estimation in the Wheatbelt, WA.

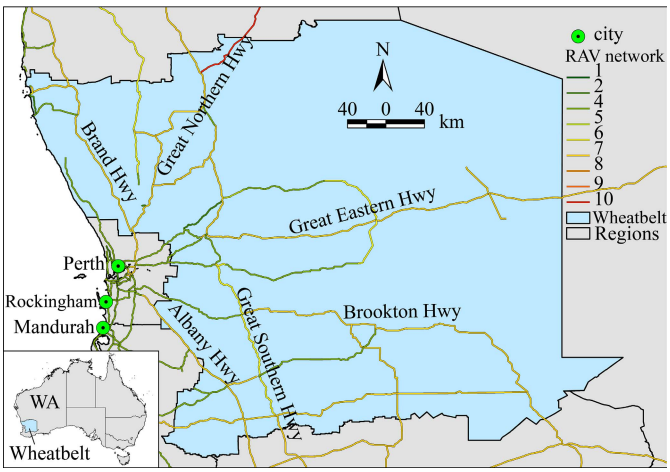


Fig. 3. Main roads and categories of RAV network in Wheatbelt, WA, Australia.

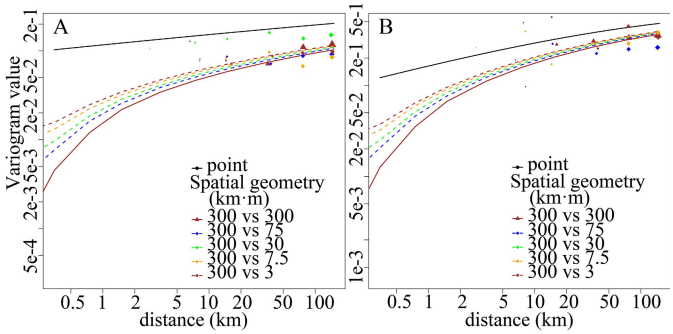


Fig. 4. Sample and fitted variograms of SRK model for heavy vehicle (A) and light vehicle (B) predictions.

III. Experiments

A. Study Area and Data

The Wheatbelt is one of the most important grain production regions in WA, Australia (Fig. 3). Heavy vehicles are a primary tool for the transportation of grain and industrial production. In the Wheatbelt, there are 280 primary segments of main roads distributed within the road networks by RAV network. Remote areas and Perth, the capital of WA, are linked by six major roads running through the Wheatbelt, including Brand Highway, Great Northern Highway, Great Eastern Highway, Brookton Highway, Great Southern Highway and Albany Highway. The classification of RAV networks is based on the axles numbers of heavy vehicles, and lists the mass of heavy vehicles in each category [48]. The total number of heavy vehicle accounts for about twenty percent of all vehicles, however, its impact on road damage is much greater than light vehicles. Heavy vehicles are primarily used for freight transportation and are characterized by large maximum permitted mass ranging from 42.5 t to 147.5 t [48].On the other hand, the weight of a standard light vehicle is only 1.65 t and the gross vehicle mass (GVM) is not allowed to exceed 4.5 t in Australia [49]. Volumes of heavy vehicle and light vehicle have been collected annually by Main Roads Western Australia, from fiscal years of 2008/2009 to 2013/2014 at 627 counting locations on 148 road segments distributed in the Wheatbelt [50]. This means volumes on about half of the road segments on main roads are counted, and the other half of road segments are not counted, even though they are also very important RAV networks. The counted volumes are summarized at segment level for spatial and temporal consistency. The mean summarized segment-level observations of heavy, light and total vehicle volumes are 194.2, 809.2 and 1003.4 vehicles/(km·day), and their maximum volumes are 1211.3, 4757.7 and 5105.0 vehicles/(km·day) respectively.

In addition, predictor variables are used to model trends for kriging-based models with regression. The predictor variables we collected include width (m), length (km) and direction of road segment (from east-west direction (0) to north-south direction (1)) and AIs, including WAI and DAI. Population is an effective indicator for traffic prediction since a dense traffic volume usually links with densely accelerated people's interactions [51]. Population data of raster type (2010) with a spatial resolution of 1 km is obtained from NASA Socioeconomic Data and Applications Center (SEDAC) [52].

B. Segment-Based Models for Traffic Prediction

The SRK variograms for heavy and light vehicle interpolations are computed respectively (Fig. 4). It illustrates the sample variograms and semivariogram functions of distance and spatial geometry, where solid lines represent semivariogram functions of equally sized segments and dotted ones are functions of the combination of various segments. The spatial characteristics of road segments are described by the multiplication of segment length and width. The segment-based variograms are a series of functions of distance and combined spatial geometry, and their shapes are significantly different with point-based fitted variograms (dark line). Shape distinction in the variograms of SRK for heavy vehicles is larger than that of SRK for light vehicles.

C. Performance Comparison for Segment-Based, Point-Based and Non-Spatial Models

Ten-fold cross validations are utilized to assess model performance with the statistical indices ME, MAPE and R^2. Statistical summary of cross validation is listed in Table 1. For interpolation models without regression, prediction errors (ME and MAPE) are significantly reduced by SOK compared with IDW and OK. The performance of models with regressions is generally improved compared with pure spatial interpolations such OK [53]. SRK performs better for predicting heavy vehicle volumes (ME = 8.5; MAPE = 25.0%) than non-spatial and point-based models, including UK (ME = 16.1; MAPE = 28.0%), RK (ME = 17.1; MAPE = 27.4%) and LR (ME = 17.3; MAPE = 27.4%). However, SRK with ME of 21.1 and MAPE of 22.2% is weaker than RK (ME = 13.2; MAPE = 20.8%) for the prediction of light vehicle volumes. The integration of SRK for heavy vehicles and RK for light vehicles (SRK + RK) predicts total vehicle volumes with the smallest prediction errors of ME (21.7) and MAPE (18.8%) compared with those of pure point-based RK (ME = 30.4; MAPE =19.2%), pure segment-based model

TABLE I

STATISTICAL SUMMARY FOR CROSS VALIDATION OF SEGMENT-BASED, POINT-BASED, AND NON-SPATIAL MODELS

Vehicle	Model	Mean	ME	MAPE	R2	Mean variance
Heavy vehicles	IDW	154.5	39.712	0.392	0.538	/
	OK	153.5	40.780	0.367	0.579	5.224E+05
	SOK	169.4	**24.803**	**0.366**	**0.564**	1.301E+05
	LR	178.1	16.117	0.280	0.634	/
	UK	176.6	17.658	0.274	0.666	2.224E+05
	RK	177.0	17.279	0.274	0.654	2.265E+05
	SRK	185.7	**8.499**	**0.250**	**0.677**	4.941E+04
Light vehicles	IDW	751.4	57.756	0.458	0.424	/
	OK	661.3	147.859	0.423	0.274	1.812E+08
	SOK	767.9	**41.253**	**0.382**	**0.431**	2.603E+07
	LR	778.7	30.438	0.212	0.758	/
	UK	778.0	31.193	0.208	0.753	8.493E+07
	RK	796.0	**13.151**	**0.208**	**0.763**	8.205E+07
	SRK	788.1	21.109	0.222	0.606	5.312E+06
Total vehicles (heavy plus light vehicles)	IDW	905.9	97.468	0.423	0.452	/
	OK	814.8	188.638	0.396	0.350	1.818E+08
	SOK	937.4	**66.055**	**0.360**	**0.470**	2.624E+07
	LR	956.9	46.555	0.193	0.805	/
	UK	954.6	48.851	0.193	0.801	8.516E+07
	RK	973.0	30.430	0.192	0.811	8.232E+07
	SRK	973.8	29.608	0.201	0.685	5.381E+06
	SRK+RK	981.8	**21.650**	**0.188**	**0.805**	8.209E+07
	SRK(all)	978.9	24.525	0.200	0.686	/

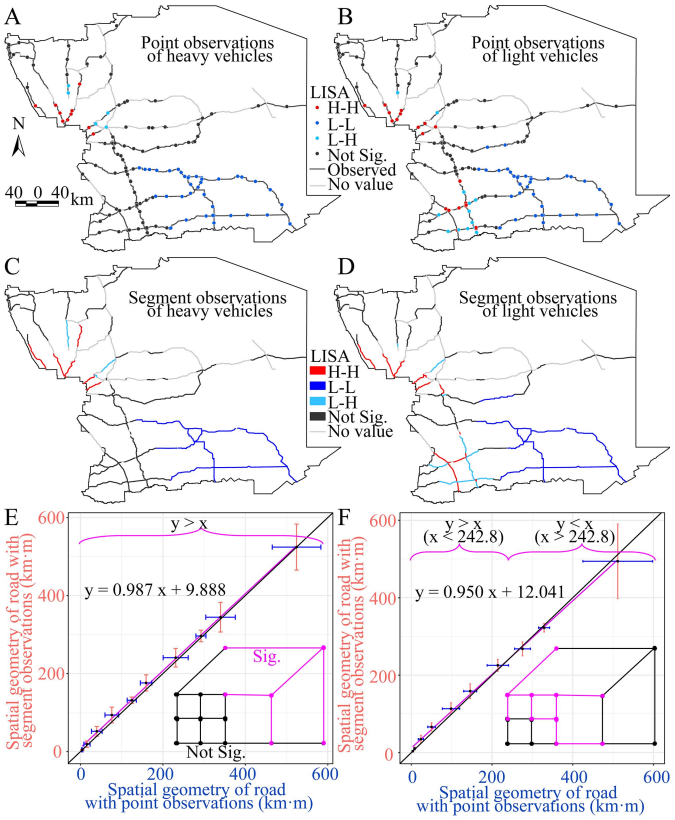


Fig. 5. Spatial geometry comparison (E: heavy vehicles; F: light vehicles) between roads with significantly local autocorrelated point-based (A: heavy vehicles; B: light vehicles) and roads with segment-based observations (C: heavy vehicles; D: light vehicles).

SRK (ME = 29.6; MAPE = 20.1%), and the model without the differentiation of two types of vehicles (SRK(all)) (ME = 24.5; MAPE = 20.0%). In addition, the coefficient of determination R^2 also proves that SRK with $R^2 = 0.677$ fits better than other models for heavy vehicles, but RK with $R^2 = 0.763$ fits better than SRK with $R^2 = 0.606$ for light vehicles. In total, cross validation R^2 of SRK + RK is 0.805, which is higher than pure SRK but slightly weaker than the pure RK model due to the imbalance of vehicle volumes (i.e. the average number of light vehicles is four times as many as that of heavy vehicles).

Cross validation results indicate that the prediction performance of SRK is associated with the relationship between the spatial heterogeneity and spatial geometry of spatial segment-based data. According to the characteristics of road segments mentioned above, short roads tend to be distributed in urban and densely populated areas, such as counties and towns, but long roads are primarily distributed in rural and outer areas, where the heavy vehicle freight transportation is more frequent than that in populated areas. This phenomenon enables SRK to predict segment-based traffic volumes of heavy vehicles better than those of light vehicles, since SRK characterizes the spatial geometry of road segments to predict traffic volumes, and the spatial geometry of long roads is more distinct than short roads compared with point-based observations. Due to the consideration of spatial geometry of road segments, spatial autocorrelations will be different between point-based and segment-based observations. In this paper, the spatial local autocorrelations are presented by local indicators of spatial association (LISA) [54]. The spatially significant autocorrelated roads are expressed by four groups, including high-volume roads and neighbors (H-H), low-volume roads and neighbors (L-L), low-volume roads and high-volume neighbors (L-H), and high-volume roads and low-volume neighbors (H-L) [55]. Fig. 5(A-D) shows the spatial autocorrelations of point-based and segment-based observations of heavy vehicles and light vehicles respectively. Fig. 5(E and F) compares the spatial geometry of roads with significant spatial autocorrelations between point-based and segment-based traffic volumes. Since the numbers of roads with significant spatial autocorrelations are different for point-based and segment-based data, the spatial geometry is summarized with mean and standard deviation between ten-quantile values. Schematic diagrams are added to explain the results. Results show that for the spatially significant autocorrelated volumes of all heavy vehicles, spatial geometry of roads with segment-based observations is larger than that with point-based observations, which means relatively long roads tend to be locally clustered and serve the heavy vehicles. While, when involving spatial geometry, volumes of light vehicles on short roads are significantly clustered, but the local autocorrelations of light vehicles on relatively long roads will be reduced. This result reveals that SRK can improve prediction accuracy for segment-based spatial data whose spatial autocorrelation is significant in segments with large spatial geometry than that in the segments with small spatial geometry. Thus, SRK model relies on both significant spatial heterogeneity of segment-based data and relatively large spatial geometry of segments.

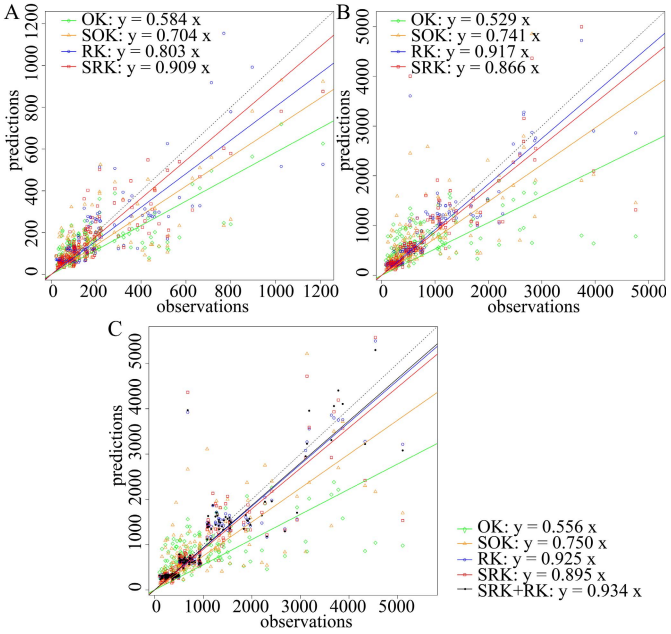


Fig. 6. Comparison of observations and cross-validation predictions from geostatistical segment-based and point-based models for heavy traffic volume (A), light vehicle volume (B) and total volume (C).

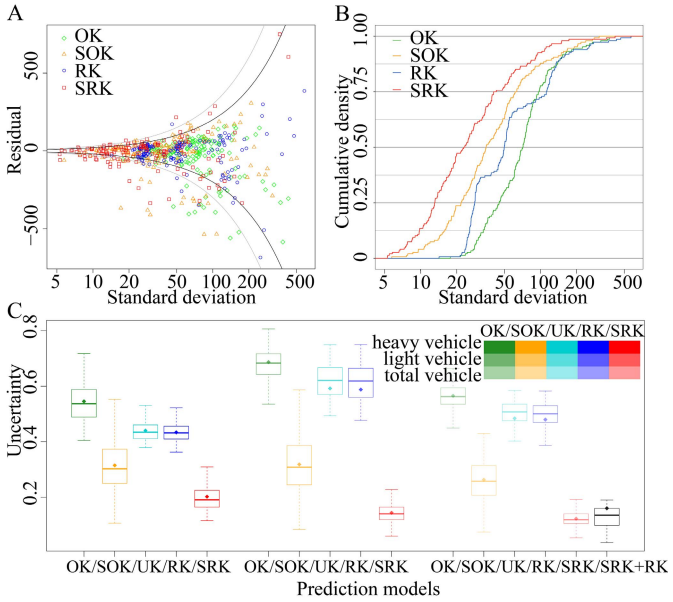


Fig. 7. Standard deviation and uncertainty summaries of predicted errors for geostatistical segment-based and point-based models. Relationship between residuals and standard deviation (A), cumulative possibility density of standard deviation (B) and estimated uncertainty summary (C).

Fig. 6 presents the results from the ten-fold cross validations performed for the predictions of heavy, light and total vehicle volumes at road segments respectively. Scatter plots and simple linear regressions are used to present the relationships between traffic volume observations and predictions from point-based models OK and RK, and segment-based models SOK and SRK, together with the combination of SRK for heavy vehicles and RK for light vehicles for the prediction of total vehicles. Results show that SRK and RK can better predict heavy and light vehicle volumes respectively, and SRK + RK is the best prediction model for total vehicle volumes.

Spatial autocorrelation of residuals reveals the fitting improvement of models. Moran's I is used to describe spatial autocorrection with a more significant spatial autocorrelation showing large absolutes of I and Z. For models without regression, spatial autocorrelation of SOK fitted residuals is Moran's I = 0.007, pseudo p = 0.331 and Z value = 0.383 for heavy vehicles and Moran's I = 0.041, pseudo p = 0.074 and Z value = 1.472 for light vehicles, and that of OK fitted residuals is Moran's I = −0.044, pseudo p = 0.104 and Z value = −1.116 for heavy vehicles and Moran's I = 0.044, pseudo p = 0.070 and Z value = 1.557 for light vehicles. This result shows that SOK reduces more spatial autocorrelations compared with OK for heavy vehicles but the reduction for light vehicles is slight. For kriging models with regression, spatial autocorrelation is reduced by SRK (Moran's I = −0.034, pseudo p = 0.183, Z value = −0.767) compared with RK (Moran's I = 0.098, pseudo p = 0.012, Z value = 3.109) for the prediction of heavy vehicle. In contrast, it is not reduced by SRK (Moran's I = −0.041, pseudo p = 0.098, Z value = −1.090) for predicting light vehicle volumes comparing to RK (Moran's I = −0.026, pseudo p = 0.234, Z value = −0.624). This result demonstrates that the traveling behaviors of heavy vehicles are more related with the spatial characteristics of road segments and segment-based models are more advantageous for explaining these characteristics. On the other hand, the traveling behaviors of light vehicles are different as they tend to have limited relationship with spatial geometry of segment and it is better to directly use a point-based model for their prediction.

Table 1 and Fig. 7 also illustrate the comparison of model performance from the perspective of kriging standard deviation and estimation uncertainty. Standard deviation is the squared root of spatial variance, and the estimated uncertainty is computed as kriging standard deviation divided by the prediction value [30]. Results show that the improvements of standard deviation of segment-based models, SOK and SRK, over point-based models, OK, UK and RK, are apparent. For the prediction of heavy vehicles, 78.19% of spatial variance and 69.03% of standard deviation are improved by SRK on average compared with RK, the best performing point-based model. It is also demonstrated that the fitting improvement of segment-based models primarily comes from the reduction of estimated standard deviation. Fig. 7(A) shows the relationships between residuals and standard deviation. It presents that most of the fitting residuals locate within three times estimated standard deviation (grey curve) even two times one (black curve), and segments with small standard deviations generally have relatively small fitting residuals. Segment-based models have both lower residuals and standard deviations than point-based models. Fig. 7(B) illustrates a comparison of the standard deviations among segment-based and point-based models. Segment-based models, especially SRK, have much higher cumulative density at low standard deviations.

Fig. 7(C) shows the statistical summary of estimated uncertainty. It shows that regression and segment-based models both run well on reducing estimated uncertainty and the

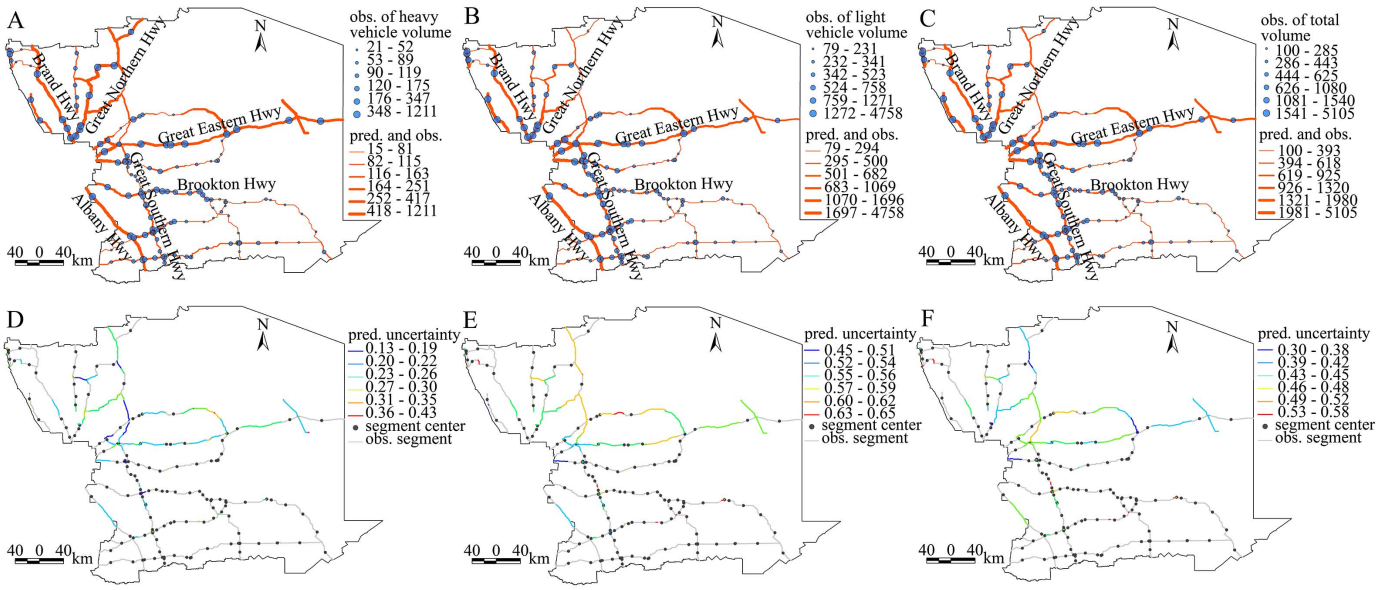


Fig. 8. Predictions of traffic volumes (heavy vehicle volume: A; light vehicle volume: B; total volume C) and corresponding prediction uncertainty (heavy vehicle volume: D; light vehicle volume: E; total volume F).

reduction from segment-based models is apparent. On average, 53.36% of estimated uncertainty is improved by SRK for heavy vehicles compared with RK. The uncertainty of the SRK + RK model with an average value of 0.160 is slightly higher than the pure SRK model due to the integration of RK-based prediction for light vehicles whose estimated uncertainty is high, but it is still much lower than the uncertainty of purely point-based models, including UK and RK with average values of 0.484 and 0.470. Thus, the improvement of estimated uncertainty from the SRK + RK model is 65.96% compared with the point-based RK model.

D. Spatial Prediction of Heavy and Light Vehicle Volumes

Fig. 8 shows the maps of predicted traffic volumes for heavy, light and total vehicles with the observations plotted at the middle points of road segments, and spatial distributions of respective prediction uncertainty. The mean predictions of heavy, light and total vehicle volumes are 277.4, 1006.0 and 1283.4 vehicles/(km·day), and the maximum volumes are 1172.0, 3547.0 and 4719.0 vehicles/(km·day) respectively. The corresponding distributions of kriging prediction uncertainty calculated by equation (14) and (16) are also displayed on the maps. Uncertainty of SRK-based heavy vehicle prediction model ranges from 0.13 to 0.43. Uncertainty of RK-based light vehicle prediction model ranges from 0.45 to 0.65, and uncertainty of SRK + RK-based total vehicle prediction model ranges from 0.30 to 0.58.

IV. Discussion

This study proposes segment-based models, SOK and SRK, for integrating spatial characteristics and spatial homogeneity of road segments with geostatistical interpolation methods to more accurately predict traffic volumes. Results show that segment-based models can reduce prediction errors compared

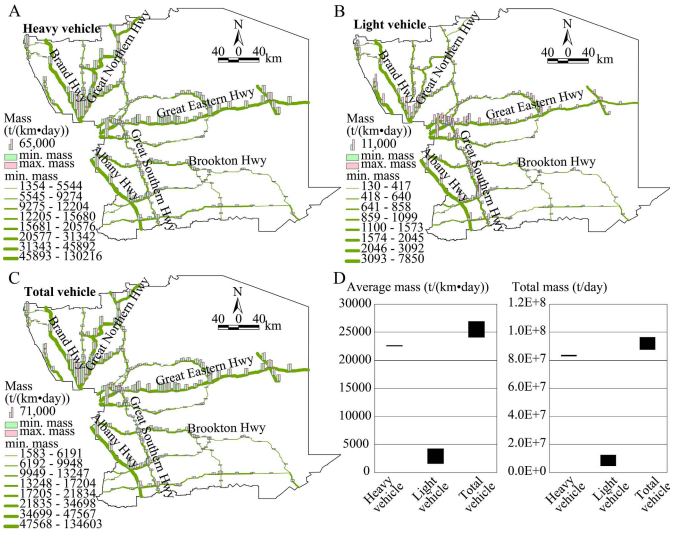


Fig. 9. Mass distributions for total vehicles (A), light vehicles (B) and total vehicles (C), and their statistical summary (D).

with point-based models by the consideration of spatial geometry of road segment, and reduce prediction standard deviation and uncertainty. Segment-based models also can better explain spatial autocorrelation of residuals through involving spatial geometry information of segments in kriging models with segment-based covariance and semivariogram functions. SRK performs much better than point-based models for the prediction of heavy vehicles, because the traveling behaviors of heave vehicles are more relevant when spatial characteristics of road segment are included. In the Wheatbelt, heavy vehicles primarily run on main roads and are used for long-distance freight transportation between grain production and remote industrial areas and densely populated regions located in the west of the Wheatbelt, such as Perth (capital city) and major

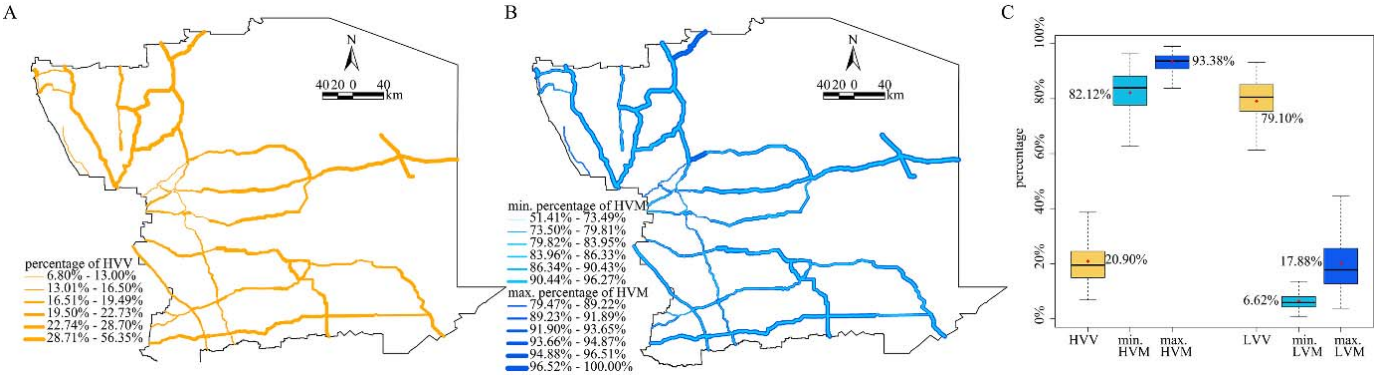


Fig. 10. Distributions of percentages of heavy vehicle volume (A) and mass (B), and their statistical summary (C).

centers of Rockingham and Mandurah. In contrast, RK is a better model for the prediction of light vehicle volumes than SRK. This phenomenon can be explained by the characteristics of SRK model that spatial geometry of segments are considered for prediction. SRK can improve prediction accuracy for segment-based spatial data whose spatial autocorrelation is higher in segments with large spatial geometry than that in the segments with small spatial geometry. Thus, we recommend that the relationship between spatial geometry of segments and the spatial heterogeneity of segment-based data should be examined before SRK modelling. A combination of SRK for heavy vehicles and RK for light vehicles is adopted for traffic volume prediction. Results reveal that the combined SRK + RK method has advantages on the prediction of both heavy and light vehicles with high prediction accuracy.

The SRK + RK-based predictions of traffic volumes are applied on the estimation of road maintenance burden in the Wheatbelt, WA. Road maintenance burden at segment level is estimated by the integration of the traffic volume predictions and the RAV network determined vehicle mass limitations. The distributions of the estimated ranges of vehicle masses for heavy, light and total vehicles are calculated at all segments, together with the statistical summaries shown in Fig. 9. The minimum masses of vehicles shown in the maps are conservative estimates, which assume that loads of all vehicles are restricted within the minimum limitations of the RAV network. On main roads in the Wheatbelt, the average masses of heavy, light and total vehicles are [22.49, 22.69], [1.56, 4.26], and [24.05, 26.95] thousand tons per kilometer per day, and their total masses are [82.78, 83.88], [4.64, 12.66], and [87.42, 96.54] million tons per day. The road maintenance burden from the mass of heavy vehicles is much bigger than that of light vehicles.

To explain the distinct impacts of heavy and light vehicles on road maintenance, percentages of heavy vehicle volumes and masses among all vehicles are mapped and summarized in Fig. 10. Results show that the percentage of heavy vehicles on road segments has a diverse range from 6.8% to 56.4% with a mean percentage of 20.9%, but its impact on road maintenance ranges from 51.4% to 96.3%, and the mean impact for all road segments accounts for 82.1% of the road maintenance burden. In contrast, 17.9% of the mean impact on road maintenance burden is related to light vehicles, although its volume is around 79.1%, which is about four times higher than that of heavy vehicle.

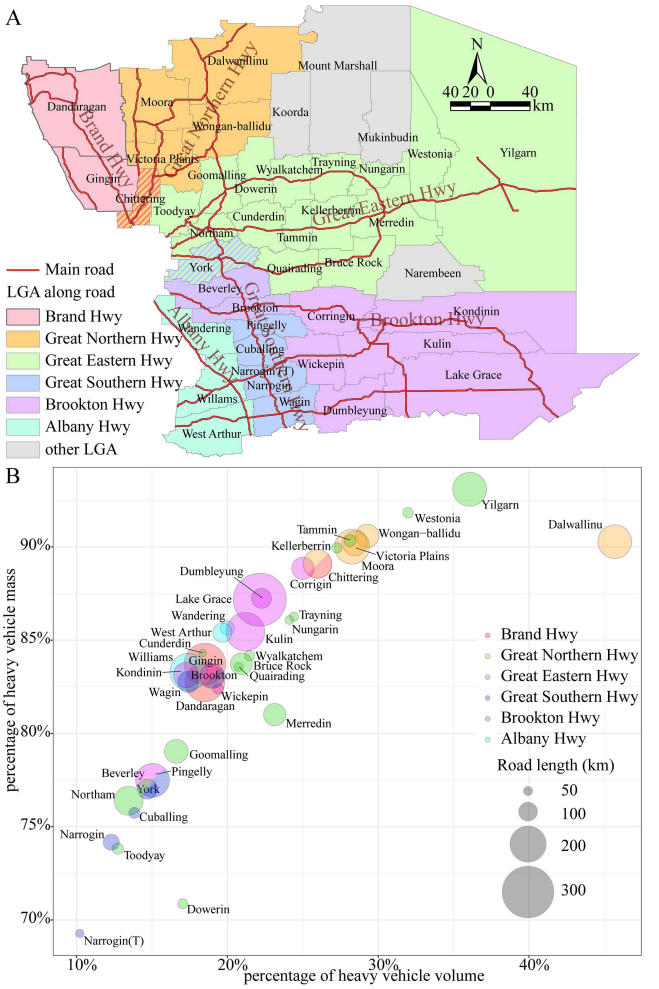


Fig. 11. Impacts of heavy vehicle on road maintenance burden for local government areas (LGAs).

Practically, grain production and remote industrial areas link densely populated cities with six primary main roads through the Wheatbelt, including Brand Highway, Great Northern Highway, Great Eastern Highway, Brookton Highway, Great Southern Highway and Albany Highway. To further analyses the impacts of heavy vehicles on the road maintenance burden,

road segments are grouped by local government area (LGA) groups that are spatially determined by the aforementioned six primary main roads. Fig. 11 shows the distributions of LGA groups, and the distribution of the impact of heavy vehicle on road maintenance burden in these groups. A general trend is that the percentage of heavy vehicle mass increases together with the percentage of heavy vehicle volume, but the relationships vary in different LGAs. The road maintenance burden from heavy vehicles is relatively low at LGAs along Great Southern Highway, and high in LGAs along Great Northern Highway. The impact of heavy vehicles varies at LGAs along Great Eastern Highway with high impact recorded at Dowerin and Toodyay, and a low impact at Yilgarn.

V. Conclusion

This study reveals that by involving the spatial geometry information of segments, the segment-based spatial interpolation method can more accurately estimate the segment-based spatial data with significant spatial heterogeneity and large spatial geometry compared with point-based interpolations. Due to this characteristic of segment-based methods, they have strong capability in predicting traffic volumes of heavy vehicle freight transportation especially in rural and remote areas, where the roads are longer than urban road segments, the monitoring data are sparsely distributed, and it is difficult to collect global positioning (GPS) data of vehicles with high spatial and temporal resolutions. In addition, segment-based spatial interpolation methods can be used in the spatial estimations for geographical objects with the spatial geometry of long segments. Besides the above academic contributions of this study, the methods and results also contribute to the industrial practice. The application of SRK on predicting the distributions of heavy vehicle volumes indicates that SRK can significantly improve prediction accuracy by considering the spatial geometry of road segments. The integration of accurately predicted traffic volumes at road segment level and the limitations of the RAV network is practical for assessing the impact of vehicles on the road maintenance burden in various places. Segment-based spatial interpolation methods are also a useful approach for the management of heavy and light vehicles, and can inform wise decision making of road maintenance strategies.

Acknowledgment

The authors would like to thank three reviewers and Editor Prof. Petros Ioannou for their insightful comments on the paper.

References

[1] H. Zou, Y. Yue, Q. Li, and A. G. O. Yeh, "An improved distance metric for the interpolation of link-based traffic data using Kriging: A case study of a large-scale urban road network," *Int. J. Geograph. Inf. Sci.*, vol. 26, no. 4, pp. 667–689, 2012.

[2] X. Wang and K. Kockelman, "Forecasting network data: Spatial interpolation of traffic counts from texas data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2009, pp. 100–108, Oct. 2105.

[3] H.-X. Zou, Y. Yue, Q.-Q. Li, and A. G.-O. Yeh, "Traffic data interpolation method of non-detection road link based on Kriging interpolation," *Jiaotong Yunshu Gongcheng Xuebao*, vol. 11, no. 3, pp. 118–126, 2011.

[4] B. Yang, S.-G. Wang, and Y. Bao, "Efficient local AADT estimation via SCAD variable selection based on regression models," in *Proc. Control Decision Conf. (CCDC)*, May 2011, pp. 1898–1902.

[5] M. Lowry, "Spatial interpolation of traffic counts based on origin–destination centrality," *J. Transp. Geogr.*, vol. 36, pp. 98–105, Apr. 2014.

[6] J. Tang, G. Zhang, Y. Wang, H. Wang, and F. Liu, "A hybrid approach to integrate fuzzy C-means based imputation method with genetic algorithm for missing traffic volume data estimation," *Transp. Res. C, Emerg. Technol.*, vol. 51, pp. 29–40, Feb. 2015.

[7] H. Tan, G. Feng, J. Feng, W. Wang, Y.-J. Zhang, and F. Li, "A tensor-based method for missing traffic data completion," *Transp. Res. C, Emerg. Technol.*, vol. 28, pp. 15–27, Mar. 2013.

[8] R. Rossi, M. Gastaldi, and G. Gecchele, "Comparison of clustering methods for road group identification in FHWA traffic monitoring approach: Effects on AADT estimates," *J. Transp. Eng.*, vol. 140, no. 7, p. 04014025, 2014.

[9] T. M. Guide, "Federal highway administration (FHWA)," US Dept. Transp., Washington, DC, USA, Tech. Rep. FHWA-PL-13-015, 2012.

[10] S. Islam, "Estimation of annual average daily traffic (AADT) and missing hourly volume using artificial intelligence," M.S. thesis, Clemson Univ., Clemson, SC, USA, 2016.

[11] B. Selby and K. M. Kockelman, "Spatial prediction of traffic levels in unmeasured locations: Applications of universal Kriging and geographically weighted regression," *J. Transp. Geogr.*, vol. 29, pp. 24–32, May 2013.

[12] C. K. Toth, A. Barsi, and T. Lovas, "Vehicle recognition from LiDAR data," in *Proc. Int. Arch. Photogramm. Remote Sens.*, 2003, pp. 1–5.

[13] Z. Jiang, M. R. McCord, and P. K. Goel, "Improved AADT estimation by combining information in image- and ground-based traffic data," *J. Transp. Eng.*, vol. 132, no. 7, pp. 523–530, 2006.

[14] M. McCord, P. Goel, Z. Jiang, B. Coifman, Y. Yang, and C. Merry, "Improving AADT and VDT estimation with high-resolution satellite imagery," in *Proc. Pecora 15/Land Satellite Inf. IV Symp.*, 2002, pp. 10–15.

[15] B. Selby and K. Kockelman, "Spatial prediction of AADT in unmeasured locations by universal Kriging," in *Proc. 90th Annu. Meeting Transp. Res. Board*, 2011, p. 21.

[16] V. Prasannakumar, H. Vijith, R. Charutha, and N. Geetha, "Spatio-temporal clustering of road accidents: GIS based analysis and assessment," *Proc.-Soc. Behavioral Sci.*, vol. 21, pp. 317–325, Jan. 2011.

[17] J. K. Hackney, M. Bernard, S. Bindra, and K. W. Axhausen, "Predicting road system speeds using spatial structure variables and network characteristics," *J. Geograph. Syst.*, vol. 9, no. 4, pp. 397–417, 2007.

[18] H. Miura, "A study of travel time prediction using universal Kriging," *Top*, vol. 18, no. 1, pp. 257–270, 2010.

[19] S. S. Prasetiyowati, M. Imrona, I. Ummah, and Y. Sibaroni, "Prediction of public transportation occupation based on several crowd spots using ordinary Kriging method," *J. Innov. Technol. Edu.*, vol. 3, no. 1, pp. 93–104, 2016.

[20] M. M. Molla, M. L. Stone, and E. Lee, "Geostatistical approach to detect traffic accident hot spots and clusters in north dakota," Dept. Civil Environ. Eng., North Dakota State Univ. Fargo, ND, USA, Tech. Rep. 276, 2014.

[21] J. Li and A. D. Heap, "Spatial interpolation methods applied in the environmental sciences: A review," *Environ. Model. Softw.*, vol. 53, pp. 173–189, Mar. 2014.

[22] Y. Song, Y. Ge, J. Wang, Z. Ren, Y. Liao, and J. Peng, "Spatial distribution estimation of malaria in northern China and its scenarios in 2020, 2030, 2040 and 2050," *Malaria J.*, vol. 15, no. 1, p. 345, 2016.

[23] Y.-Z. Song, H.-L. Yang, J.-H. Peng, Y.-R. Song, Q. Sun, and Y. Li, "Estimating $PM_{2.5}$ concentrations in Xi'an City using a generalized additive model with multi-source monitoring data," *PLoS ONE*, vol. 10, no. 11, p. e0142149, 2015.

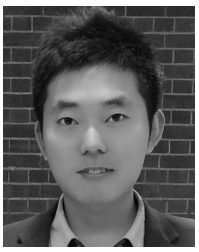
[24] J. Wang, Y. Ge, Y. Song, and X. Li, "A geostatistical approach to upscale soil moisture with unequal precision observations," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2125–2129, Dec. 2014.

[25] P. Goovaerts, "Combining area-based and individual-level data in the geostatistical mapping of late-stage cancer incidence," *Spatial Spatio-Temporal Epidemiol.*, vol. 1, no. 1, pp. 61–71, 2009.

[26] E. Pardo-Iguzquiza, V. F. Rodríguez-Galiano, M. Chica-Olmo, and P. M. Atkinson, "Image fusion by spatially adaptive filtering using downscaling cokriging," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 337–346, 2011.

[27] J. O. Skøien, R. Merz, and G. Blöschl, "Top-Kriging? Geostatistics on stream networks," *Hydrol. Earth Syst. Sci. Discussions*, vol. 2, no. 6, pp. 2253–2286, 2005.

[28] Y. Ge, Y. Liang, J. Wang, Q. Zhao, and S. Liu, "Upscaling sensible heat fluxes with area-to-area regression Kriging," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 656–660, Mar. 2015.
[29] P. Goovaerts, "Geostatistical analysis of disease data: Accounting for spatial support and population density in the isopleth mapping of cancer mortality risk using area-to-point Poisson Kriging," *Int. J. Health Geograph.*, vol. 5, no. 1, p. 52, 2006.
[30] J. O. Skøien, "Rtop: An R package for interpolation of data with a variable spatial support, with an example from river networks," *Comput. Geosci.*, vol. 67, pp. 180–190, Jun. 2014.
[31] Wang, Q., "Downscaling MODIS images with area-to-point regression Kriging," *Remote Sens. Environ.*, vol. 166, pp. 191–204, Sep. 2015.
[32] Y. Zhang *et al.*, "Spectral–Spatial adaptive area-to-point regression Kriging for MODIS image downscaling," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1883–1896, May 2017.
[33] X. Liu, P. C. Kyriakidis, and M. F. Goodchild, "Population-density estimation using regression and area-to-point residual Kriging," *Int. J. Geograph. Inf. Sci.*, vol. 22, no. 4, pp. 431–447, 2008.
[34] N. S. Asmarian, A. Kavousi, M. Salehi, and B. Mahaki, "Mapping of stomach cancer rate in Iran using area-to-area poisson Kriging," *J. Health Syst. Res.*, vol. 8, no. 4, pp. 681–687, 2012.
[35] N. S. Asmarian, A. Ruzitalab, K. Amir, S. Masoud, and B. Mahaki, "Area-to-area poisson Kriging analysis of mapping of county-level esophageal cancer incidence rates in Iran," *Asian Pacific J. Cancer Prevention*, vol. 14, no. 1, pp. 11–13, 2013.
[36] W. Chan, T. F. Fwa, and C. Y. Tan, "Road-maintenance planning using genetic algorithms. I: Formulation," *J. Transp. Eng.*, vol. 120, no. 5, pp. 693–709, 1994.
[37] A. M. M. Fakhar and A. Asmaniza, "Road maintenance experience using polyurethane (PU) foam injection system and geocrete soil stabilization as ground rehabilitation," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 136, no. 1, p. 012004, 2016.
[38] J. Min, L. Fu, T. Usman, and Z. Tan, "Does winter road maintenance help reduce air emissions and fuel consumption?" *Transp. Res. D, Transp. Environ.*, vol. 48, pp. 85–95, Oct. 2016.
[39] H. Gao and X. Y. Zhang Li, "Developing a weighted reward criterion for the Markov-based decision of road maintenance," *SpringerPlus*, vol. 5, no. 1, pp. 1–14, 2016.
[40] J.-P. Bilodeau, L. Gagnon, and G. Doré, "Assessment of the relationship between the international roughness index and dynamic loading of heavy vehicles," *Int. J. Pavement Eng.*, vol. 18, no. 8, pp. 693–701, 2017.
[41] B. S. Underwood, Z. Guido, P. Gudipudi, and Y. Feinberg, "Increased costs to US pavement infrastructure from future temperature rise," *Nature Climate Change*, vol. 7, pp. 704–707, Sep. 2017.
[42] S. Dasgupta, M. Hossain, M. Huq, and D. Wheeler, "Climate change, groundwater salinization and road maintenance costs in coastal Bangladesh," World Bank Policy Res., Washington, DC, USA, Working Paper 7147, 2014.
[43] S. Collins, "Prediction techniques for Box–Cox regression models," *J. Bus. Econ. Statist.*, vol. 9, no. 3, pp. 267–277, 1991.
[44] R. M. Sakia, "The Box–Cox transformation technique: A review," *Statistician*, vol. 42, no. 2, pp. 169–178, 1992.
[45] A. I. McLeod, Y. Zhang, and C. Xu, "Package 'FitAR,"' Dept. Statist. Actuarial Sci., Univ. Western Ontario, London, ON, Canada, Tech. Rep., 2013. [Online]. Available: https://cran.r-project.org/web/packages/FitAR/
[46] *Data Standard for Road Management and Investment in Australia and New Zealand*. Austroads, Sydney, NSW, Australia, 2016.
[47] S. Barua, K. El-Basyouny, and M. T. Islam, "Factors influencing the safety of urban residential collector roads," *J. Transp. Safety Secur.*, vol. 8, no. 3, pp. 230–246, 2016.
[48] *Prime Mover, Trailer Combinations—Vehicle Categories*, Main Roads Western Australia, Kingsford WA, Australia, 2016.
[49] *Driver and Vehicle Services, Licensing Fees and Charges*. Department Transport, Washington, DC, USA, 2016.
[50] *Main Roads Western Australia—Annual Report*, Main Roads Western Australia, Kingsford WA, Australia, 2015.
[51] L. Dong, R. Li, J. Zhang, and Z. Di, "Population-weighted efficiency in transportation networks," *Sci. Rep.*, vol. 6, May 2016, Art. no. 26377.
[52] *Center for International Earth Science Information Network-CIESIN—Columbia University, Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals*, NASA Socioeconomic Data Appl. Center (SEDAC), Palisades, NY, USA, 2016.
[53] B. Zou, Y. Luo, N. Wan, Z. Zheng, T. Sternberg, and Y. Liao. "Performance comparison of LUR and OK in $PM_{2.5}$ concentration mapping: A multidimensional perspective," *Sci. Rep.*, vol. 5, Mar. 2015, Art. no. 8698.
[54] L. Anselin, "Local indicators of spatial association—LISA," *Geograph. Anal.*, vol. 27, no. 2, pp. 93–115, 1995.
[55] Y. Ge *et al.*, "Geographically weighted regression-based determinants of malaria incidences in northern China," *Trans. GIS*, vol. 21, no. 5, pp. 934–953, 2017.



Yongze Song received the B.S. and M.S. degrees in surveying and mapping from China University of Geosciences, Beijing, in 2012 and 2015, respectively, and had four years of experience of cooperation research with the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. He is currently pursuing the Ph.D. degree in construction management with Curtin University, Perth, WA, Australia.

His current research interests include geographic-information-system theories, spatial statistical modeling, and geovisualisation-based road and traffic decision making.

Xiangyu Wang received the B.S. degree in transportation engineering from Tongji University, Shanghai, China, in 2000; the M.S. degree in engineering from University of Washington, USA, in 2002; and the Ph.D. degree in construction management from Purdue University, USA, in 2005.

He is currently a Professor and the Director with the Australasian Joint Research Centre for Building Information Modeling (BIM), School of Built Environment, Curtin University, Australia. He is also an internationally recognized Leading Researcher in the field of construction IT, BIM, lean, visualization technologies, and project management, and received about AUD $12 million in research funds and authored or co-authored over 300 peer-reviewed technical papers. He is also a Curtin-Woodside Chair Professor for Oil, Gas, and LNG Construction and Project Management, and the Founder and the Chair for the Curtin Advanced Technology Research and Innovation Alliance.

Dr. Wang has been involved in the organization of several international conferences. He is the Honorary Chair of the Conference on Innovative Production and Construction, the Editor-in-Chief of *Visualization in Engineering* (SpringerOpen).

Graeme Wright received the Ph.D. degree in spatial sciences from Curtin University, Australia.

He was the Deputy Vice-Chancellor for Research with Curtin University from 2011 to 2016. He received the title of Emeritus Professor in recognition of his contribution to the University. He has extensive knowledge and experience in education and research, including engagement with higher education policy at a strategic level. He has previously held appointments in the vocational educational sector and more recently across higher education at executive level, and has a profound understanding of the university research environment and its application to industry and the community.

Dr. Wright has extensive experience on Boards and Committees of research centers and CRCs, liaison with industry and negotiation of funding agreements, and broad research knowledge in spatial information sciences. He has been closely involved with the CRC for Spatial Information since 2003, then as a member of Research and Education Advisory Committee and the CRCSI Board. He was the inaugural Chair of the Research and Education College and currently chairs the CRCSI Research and Investment Committee, which is a CRCSI Board Committee. He led the engagement of Curtin with the CRCSI during its inception and his research back-ground is in remote sensing.

Dominique Thatcher is currently an accomplished Management Professional with over 25 years of Senior and Executive Management, specializing in end to end supply chain logistics and leading and implementing organizational cultural change.

He is currently involved in project and technology research and development with Heavy Vehicle Services, Main Roads Western Australia, Australia, including alternative road infrastructure funding model, road freight productivity and trade facilitation, digital transformation, connected autonomous vehicles, and artificial intelligence.

Mr. Thatcher is the member of the Australian Institute of Company Directors and the Associate Fellow of the Australian Institute of Management.

Peng Wu received the B.S. degree in project management from Tsinghua University, China, in 2006; the M.S. degree in construction management from Loughborough University, U.K., in 2007; and the Ph.D. degree in project management from National University of Singapore, Singapore, in 2012.

He is currently an Associate Professor with the Department of Construction Management, and an Associate Director with the Australasian Joint Research Centre for Building Information Modeling, Curtin University. His research areas include sustainable construction, lean production and construction, production and operations management, and life cycle assessment. In 2016, he received the Discovery Early Career Research Award from the Australian Research Council, which is a prestigious award to support excellent basic and applied research by early career researchers.

Pascal Felix is an experienced Strategic Leader with over 20 years of experience with the Western Australia public sector.

He is currently the Executive Director of Heavy Vehicle Services (HVS), Main Roads Western Australia, Australia. He has an interest in strategy and organizational development, technology and customer service. He is overseeing significant changes to the management of Heavy Vehicle Operations in Western Australia, including a refocusing of business operations for freight movements and improved service delivery. He established the Over Size Over Mass Unit—the first heavy vehicle—one stop shop for operators, who previously had to negotiate with up to four separate agencies. He also oversees the regulatory function of heavy vehicle and commercial operations, and works with the transport industry and to establish HVS as Leaders in researching and harnessing technology to improve freight productivity.