Sensors and Materials, Vol. 32, No. 11 (2020) 3893–3893
MYU Tokyo

S & M 2383

A Multiscale Sequential Data Assimilation System and Its Application to Short-term Traffic Flow Prediction

Wenzhong Shi[1] and Runjie Wang[1,2*]

[1]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong, P. R. China
[2]College of Surveying and Geo-Informatics, Tongji University,
No. 1239 Siping Road, Shanghai 200092, P. R. China

(Received June 28, 2020; accepted November 4, 2020)

Keywords: sequential data assimilation, assimilation models, historical measurements, multiscale noise separation, short-term traffic flow prediction

We present a multiscale sequential data assimilation (M-SDA) system and apply it to short-term traffic flow prediction. Assimilation models in traditional sequential data assimilation (T-SDA) systems, which are usually constructed using historical measurements, are always disturbed by local noises. Simultaneously, the accuracy of assimilation results is also affected. To reduce the effects of these noises on assimilation models and the accuracy of results, an M-SDA system combining a T-SDA system and noise separation methods is constructed. This paper comprises four main parts: (1) a T-SDA system for short-term traffic flow prediction and multiscale noise separation methods are briefly discussed, and an example of denoised measurements with separated multiscale noises is given; (2) an M-SDA system for short-term traffic flow prediction is established; (3) the impacts of different noise separation scales on the accuracy of assimilation results are analyzed; and (4) applications of the M-SDA system to short-term traffic flow prediction are presented and compared with those of a T-SDA system. Experimental results were acquired from traffic flow measurements collected from a sub-area of a highway near Liverpool and Manchester, UK. The gap between the true and predicted values was evaluated by the root mean square error (RMSE) and mean absolute percent error (MAPE). By comparison with the prediction results from the T-SDA system, it was experimentally shown that the M-SDA system can successfully reduce the effects of noises in historical measurements on assimilation model construction and improve the accuracy of short-term traffic flow prediction results.

1. Introduction

Traffic flow prediction, including long-term and short-term predictions, is a crucial component in many intelligent transportation systems. For the purpose of dynamic traffic management or providing advance information to travelers, short-term traffic flow prediction that reflects fast-changing local temporal and spatial fluctuations in traffic flow values is necessary.(1) Data used in short-term traffic flow predictions are aggregated over seconds to

[*]Corresponding author: e-mail: runjie.wang@connect.polyu.hk

https://doi.org/10.18494/SAM.2020.2969

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

hours. Most studies commonly used intervals of a number of minutes, such as 1,(2) 5,(1) 10,(1) 15,(3) or 30 min.(4)

Owing to the stochastic nature of traffic flow values, prediction algorithms with both accuracy and robustness have become increasingly important. Much attention has focused on taking advantage of different measurements and models to make predictions.(5) Data assimilation (DA) is a technique that can estimate state vectors by integrating physical model information and measurements with the consideration of the data distribution in time and space, as well as measurements and background field errors. It plays a significant role in meteorology,(6) oceanography,(7) hydrology,(8) and other fields. Related DA studies based on Bayesian theory have been applied in some short-term traffic state predictions.(9,10)

There are three components in DA systems: assimilation models (dynamic state and observation models), measurements, and assimilation methods. Using the discrete time index k, DA can be mathematically expressed as(11)

$$\begin{cases} X_k = \boldsymbol{A}_{k,k-1} X_{k-1} + \boldsymbol{G}_{k,k-1} w_{k-1} \\ y_k = \boldsymbol{H}_k X_k + v_k \end{cases}, \tag{1}$$

where the first expression is the dynamic equation and the second expression is the observation equation. $\boldsymbol{A}$ is the dynamic state model. $\boldsymbol{H}$ is the time-dependent measurement operator, which connects the statements X and measurements y. w and v are assumed to be zero-mean Gaussian random noises with the covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively. $\boldsymbol{G}$ is a coefficient matrix.

Sequential assimilation methods, as one of the implementation classes of DA, posteriorly estimate the state vectors on the basis of status updates using the weights of the model and measurement errors when measurements are available.(12) Each of such methods contains two steps: predict the state using the previously analyzed one and update it using Bayes' formula.(13) The Kalman filter (KF) method is a basic algorithm in sequential DA systems. It can update variable states using real-time measurements and adapt to changes in traffic flows. The KF method has excellent performance in many traffic flow prediction applications.(9,14) It also has a low computational cost and low storage requirements. Owing to these advantages, the KF method will be used later in the sequential DA system for traffic flow prediction research.

Although traditional sequential data assimilation (T-SDA) systems have already been used for short-term traffic flow prediction, the following issues need to be further investigated. (i) Most researchers have focused on finding appropriate traffic flow prediction models but have ignored the quality of the models.(15) Variation patterns in historical measurements are similar on the same day of consecutive weeks or months, and they are always used to construct forecast models such as the vector autoregressive (VAR) model used in Ref. 16. However, noises in historical measurements usually make it difficult to abstract patterns of traffic flow data and further affect the accuracy of models for prediction. (ii) There are some disadvantages in methods that have been proposed for dealing with noises. One type of method directly smooths the errors in the time domain through a mean filtering algorithm,(17) a nonlocal mean filter algorithm,(18) a median filter algorithm,(19) and so forth. However, in these methods, the window size used for filtering is often difficult to determine, which has a major impact on the precision of noise processing. Moreover, mean filtering and nonlocal mean filter algorithms are only suitable for processing Gaussian random noises. Owing to these limitations, these methods are often used for image denoising.(20) The other type of noise processing method is performed in the frequency domain. Noises in measurement series usually have a high frequency. The discrete wavelet transform (DWT)(21) and empirical mode decomposition (EMD)(22) methods have been hot topics in applications that process measurement noises in recent years. These methods can be used to separate noises from measurements.(23) However, different degrees of noise separation (denoted as the noise separation scale later in this paper) will have different effects on model construction and assimilation results. Detailed analyses of these different noise separation methods with different noise separation scales in traffic flow prediction are required.

This study mainly focuses on the problem of assimilation model inaccuracy caused by noises in historical measurements, which are used to construct models in T-SDA systems. The purpose of this study is to build a multiscale sequential data assimilation (M-SDA) system and apply it to short-term traffic flow prediction. The M-SDA system is a combination of a T-SDA system and noise separation methods under multiple scales. Three critical issues are investigated as follows. (i) We establish an M-SDA system combining a T-SDA system with two noise separation methods, that is, a DWT method and an EMD method. (ii) We analyze the impacts of different noise separation scales on assimilation forecasting results. (iii) We apply the proposed M-SDA system to short-term traffic flow prediction and verify its effectiveness by comparison with the T-SDA system.

The remainder of the paper is organized as follows. The theoretical background is introduced in Sect. 2. The construction of the M-SDA system is described in Sect. 3. Section 4 presents detailed application experiments utilizing the M-SDA system and analyzes the accuracy of the results. Finally, conclusions are made in Sect. 5.

2. Theoretical Background

2.1 T-SDA system for short-term traffic flow prediction

Short-term traffic flow prediction models are an important part of a T-SDA system. A VAR model that considers the effects of downstream and upstream location information on the traffic flow of a specific location is used in this study.(16) The model is expressed as

$$\begin{aligned} r(k+1) = {} & para\_0(k) \times Z(k) + para\_1(k) \times Z(k-1) + \\ & para\_2(k) \times Z(k-2) + \ldots + para\_n(k) \times Z(k-n) + v(k), \end{aligned} \tag{2}$$

with

$$\begin{cases} r(k+1) = \dfrac{q_S(k+1)}{\overline{q}_S(k+1)} \\ Z(k) = \begin{bmatrix} \dfrac{q_S(k)}{\overline{q}_S(k)} & \dfrac{q_{a_i}(k)}{\overline{q}_{a_i}(k)} \end{bmatrix}^T (i = 1, 2, 3, \ldots, m), \end{cases} \tag{3}$$

where [*para*_0(*k*), *para*_1(*k*), ..., *para*_*n*(*k*)] are the unknown parameters of the state model. $q_s(k + 1)$ is the traffic flow value of a specific path that needs to be predicted through the T-SDA system. $\overline{q}_s(k+1)$ denotes the average value of the specific path, which can be calculated by historical flow measurements of the same day of the week in previous weeks at the time interval [*kT*, (*k* + 1)*T*]. $q_s(k)$ is the historical traffic flow value of the specific path at the time interval [(*k* − 1)*T*, *kT*], and $\overline{q}_s(k)$ is the corresponding average value. $q_{a_i}(k)$ denotes the traffic flow values of the downstream and upstream paths, which are also the adjacent paths at the time interval [(*k* − 1)*T*, *kT*]. $\overline{q}_{a_i}(k)$ is the historical average value.

In Eqs. (2) and (3), the unknown parameters should be calculated before acquiring the traffic flow forecasting results. Thus, for the calculation in the T-SDA system, the form in Eq. (3) can be transferred into Eq. (1) by setting

$$\begin{cases} A_{k,k-1} = I, G_{k,k-1} = 0 \\ X_k = \left[para\_0(k), para\_1(k), ..., para\_n(k)\right]^T \\ y_k = r(k+1) \\ H_k = \left[Z^T(k), Z^T(k-1), Z^T(k-2), ..., Z^T(k-n)\right]. \end{cases} \tag{4}$$

The measurements are traffic flow values. The standard KF method(24) is used as the assimilation method in the T-SDA system(24) as shown in Fig. 1.

It can be seen from Eq. (4) and Fig. 1 that the measurement operator $\boldsymbol{H}_k$ is built using historical measurements. It plays an important role in the calculation of the Kalman gain matrix $\boldsymbol{K}$ in the KF method, which is the key to balancing the weight between the state estimates and the new measurements. Noises in historical traffic flow measurements disturb the specification of the measurement operator $\boldsymbol{H}_k$ and then further reduce the accuracy of assimilation results through the effect on the Kalman gain matrix $\boldsymbol{K}$. Denoising processing for the measurement operator $\boldsymbol{H}_k$ is essential before short-term traffic flow forecasting.

2.2 Multiscale noise separation methods

Denoising processing is a major application of the DWT(21) and EMD(22) methods. Typical DWT and EMD methods applied to a signal $f(k) \in S^2(R)$ [where $S^2(R)$ denotes the square integrable function space] for multiscale noise separation are briefly introduced as

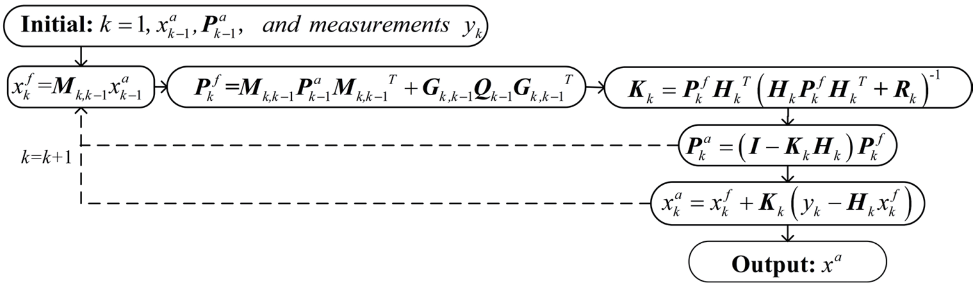
Initial: $k=1, x^a_{k-1}, \boldsymbol{P}^a_{k-1}$, *and measurements* y_k
$x^f_k = \boldsymbol{M}_{k,k-1}x^a_{k-1}$
$\boldsymbol{P}^f_k = \boldsymbol{M}_{k,k-1}\boldsymbol{P}^a_{k-1}\boldsymbol{M}_{k,k-1}{}^T + \boldsymbol{G}_{k,k-1}\boldsymbol{Q}_{k-1}\boldsymbol{G}_{k,k-1}{}^T$
$\boldsymbol{K}_k = \boldsymbol{P}^f_k\boldsymbol{H}_k{}^T\left(\boldsymbol{H}_k\boldsymbol{P}^f_k\boldsymbol{H}_k{}^T + \boldsymbol{R}_k\right)^{-1}$
$\boldsymbol{P}^a_k = \left(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k\right)\boldsymbol{P}^f_k$
$x^a_k = x^f_k + \boldsymbol{K}_k\left(y_k - \boldsymbol{H}_k x^f_k\right)$
k=k+1
Output: x^a

Fig. 1. KF method.

$$\text{DWT:} \quad \begin{cases} (W_\psi f)(a,b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} f(k)\overline{\psi\left(\frac{k-b}{a}\right)} dt \\ f(k) = C_\psi^{-1} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (W_\psi f)(a,b)\psi_{a,b}(k)\frac{da}{a^2} db, \end{cases} \tag{5}$$

where $W_\psi f$ denotes the wavelet transform. $\psi_{a,b}(k)$ is the wavelet function and a and b are the scale and shift factors, respectively. $|a|^{-1/2}$ is a normalizing factor. The Haar wavelet is appointed as the mother wavelet as it is commonly used in the DWT method. By using the DWT method, the original signal is decomposed into several levels of purer and noisy series, expressed as A_i and D_i, respectively, with i mean levels, as described in detail in Ref. 21. Using different levels of decomposed pure data A_i as measurements means different noise separation scales. The maximum level in this study was taken to be 2. This is because after many experiments, decomposed noisy series will contain excessively pure information beyond decomposed level 2. Excessive noise separation will make the left purer series lose the changing profile of the original data series.

$$\text{EMD:} \quad f(k) = \sum_{i=1}^{g} c_i(k) + r(k) \tag{6}$$

The original signal $f(k)$ can be decomposed into several intrinsic mode functions (IMFs). $c_i(t)$ denotes the ith IMF_i and g is the number of decomposed IMFs. $r(t)$ denotes the separated noises. Details of the algorithm are given in Ref. 22. In this study, we define different numbers of *IMF* combinations for rebuilt data, such as the sum of IMF_2–IMF_g and the sum of IMF_3–IMF_g, as different noise separation scales. The maximum level in this study was taken to be the sum of IMF_3–IMF_g to avoid the excessive separation of noises, which would adversely affect the accuracy of pure series on the basis of the results of many experiments.

2.3 Example of multiscale noise separation

In this section, an example of multiscale noise separation using the above two methods is given. The original data are 15 min traffic flow measurements of path 3339 (LM91) belonging to Highways England. The original traffic flow data on three Mondays in consecutive weeks, which have similar variations but are affected by observation noises, are shown in Fig. 2(a). These original data series were decomposed into several purer and noisy series under four noise separation scales. Details of the four noise separation scales are given in Table 1. After noise separation, different purer series [*P* in Figs. 2(b)–2(e)] and noise series [*N* in Figs. 2(f)–(i)] were acquired. The purer series are smoother than the original series but still keep their variation trends. They can be treated as denoised data for further multiscale model building. The noisy series were regarded as noises and removed. Figure 2 indicates that different separation scales can give different noise separation results.

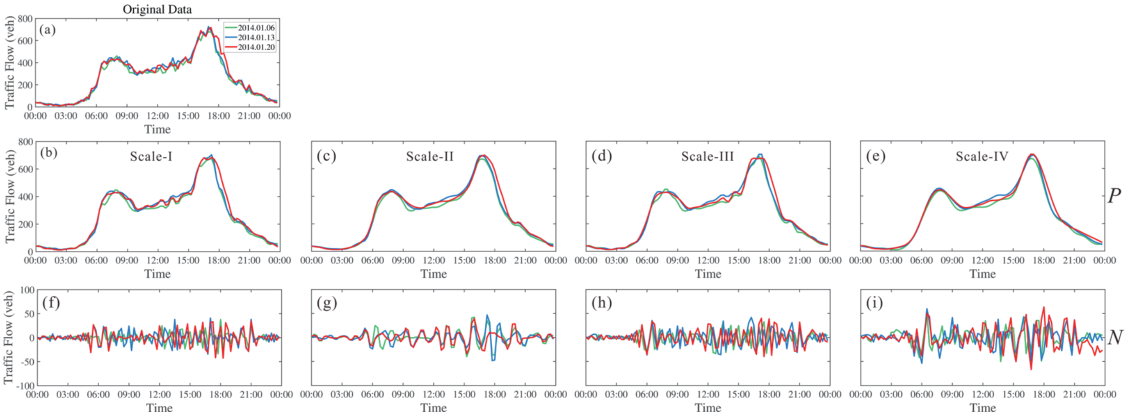

Fig. 2. (Color online) (a) Original data, (b)–(e) purer series acquired under four scales, and (f)–(i) noise series acquired under four scales.

Table 1
Details of four noise separation scales.

Scale I	Scale II	Scale III	Scale IV
A_1	A_2	Sum (IMF_2–MF_g)	Sum (IMF_3–IMF_g)

3. M-SDA System for Short-term Traffic Flow Prediction

As reported in this section, an M-SDA system was built combining T-SDA and multiscale measurement noise separation methods, that is, the DWT and EMD methods. For the M-SDA system in this study, multiseries historical measurements with various degrees of noise removal are used to build multiple measurement operators $\boldsymbol{H}_k$ in the sequential DA system to satisfy different situations of modeling data. After noises are separated at multiple scales using the above methods, more precise historical measurements remain. Furthermore, different models can be constructed using multiseries historical measurements with multiscale noises separated. "Multiscale models" is shorthand for these different models. Multiscale models will be used for different traffic flow forecasting situations with the purpose of achieving highly precise assimilation results. A schematic of the proposed M-SDA system based on the above discussion for short-term traffic flow prediction is presented in Fig. 3. The M-SDA system was built in three steps: (1) multiseries historical measurements were acquired with multiscale noises separated using the DWT and EMD methods; (2) multiscale assimilation models in the M-SDA system were constructed under multiseries historical measurements after the noises were separated; and (3) a schematic of the M-SDA system for short-term traffic flow prediction was established. Figure 3 shows the entire technological framework of the study.

4. Empirical Study

In this section, real-world data are used in short-term traffic flow prediction. Our experiment is designed to investigate the impact of different noise separation scales on the prediction results and performance of the M-SDA system for short-term traffic flow prediction.

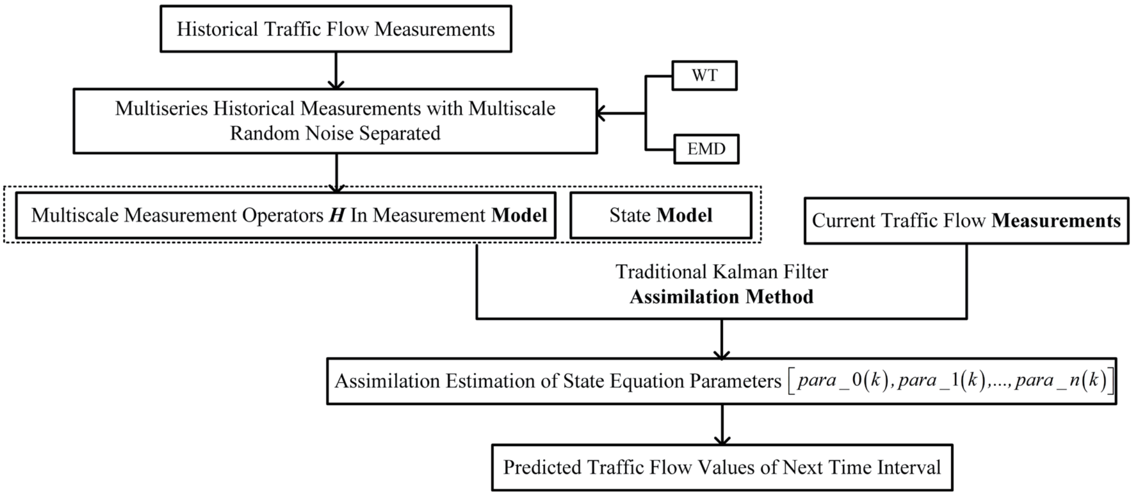
Historical Traffic Flow Measurements
Multiseries Historical Measurements with Multiscale Random Noise Separated
WT
EMD
Multiscale Measurement Operators *H* In Measurement **Model**
State **Model**
Current Traffic Flow **Measurements**
Traditional Kalman Filter **Assimilation Method**
Assimilation Estimation of State Equation Parameters $[para\_0(k), para\_1(k),..., para\_n(k)]$
Predicted Traffic Flow Values of Next Time Interval

Fig. 3. Schematic of M-SDA system for short-term traffic flow prediction.

4.1 Study area and material description

The datasets used in this paper are downloadable from the website of Highways England (*highwaysengland.co.uk*). The data were collected from a sub-area of a highway between Liverpool and Manchester, UK, as shown in Fig. 4(a). The time interval for the data is 15 min. The traffic flow forecasting results from Monday to Sunday were acquired and analyzed. The data of each path contain eight days from consecutive weeks. As the mean traffic flow values are required in the assimilation models shown in Eq. (4), datasets of the first seven days were used for model construction in the S-DA system, and the data from the eighth day were employed to test the effectiveness of the proposed approach. Also, because the traffic flow early in the morning and late at night was small and of little concern to traffic management, prediction results from 6:00 a.m. to 9:00 p.m. were used.

4.2 Impacts of different noise separation scales on prediction results

The short-term traffic flow prediction of path 3339 (LM91) shown in Fig. 4(b) was first taken as a detailed example to illustrate the impacts of different noise separation scales on assimilation prediction results. There are four paths adjacent to path 3339 (LM91): path 3338 (LM89), path 3339 (LM93), path 6200 (LM844A), and path 6296 (LM87); thus, $m = 4$ in Eq. (3). We then set $n = 2$ in Eq. (4). Without the loss of generality, the prediction results obtained on the workday Wednesday and the non-workday Sunday were analyzed in detail. For a comparative analysis of experimental results, prediction results from the T-SDA system obtained using the raw data without the noises separated were also obtained. In addition to the four noise separation scales already shown in Table 1, the five noise separation scales used in later tests are redefined in Table 2. Five assimilation models were also built. Model 1 was the assimilation model in the T-SDA

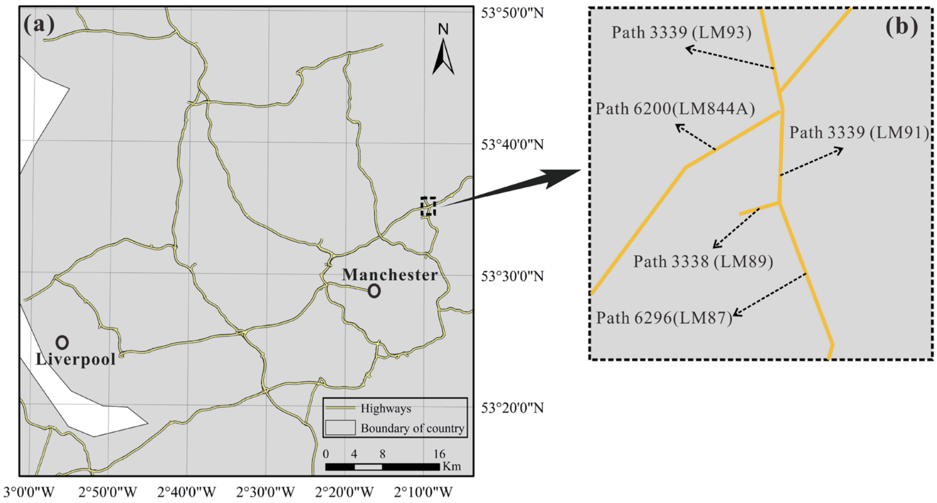

Fig. 4. (Color online) (a) Study area and (b) part of the paths.

Table 2
Details of five noise separation scales and five models.

Scale-I	Scale-II	Scale-III	Scale-IV	Scale-V
Raw data	A_1	A_2	Sum (IMF_2–IMF_g)	Sum (IMF_3–IMF_g)
Model 1	Model 2	Model 3	Model 4	Model 5

system. Models 2 through 5 were built using historical measurements with the noises separated under the four scales and are the models in the M-SDA system.

Root mean square error (RMSE)$^{(1)}$ and mean absolute percent error (MAPE)$^{(1)}$ values were employed to evaluate the prediction performance under different noise separation scales. *RMSE* and *MAPE* are defined as

$$RMSE = \sqrt{\frac{1}{tn}\sum_{k=1}^{tn}(\hat{x}(k) - x(k))^2}\,, \tag{7}$$

$$MAPE = \frac{1}{tn}\sum_{k=1}^{tn}\frac{|\hat{x}(k) - x(k)|}{x(k)} \times 100\%\,, \tag{8}$$

where $\hat{x}(k)$ denotes the prediction value, $x(k)$ is the true observed value, and tn is the total number of time intervals. The smaller the *RMSE* and *MAPE* values, the higher the prediction performance.

Figure 5 shows the prediction results on the workday Wednesday for the five models. Figure 6 shows the prediction results on the non-workday Sunday. The prediction results from Model 1, which are also the results from the T-SDA system, were taken as reference to verify

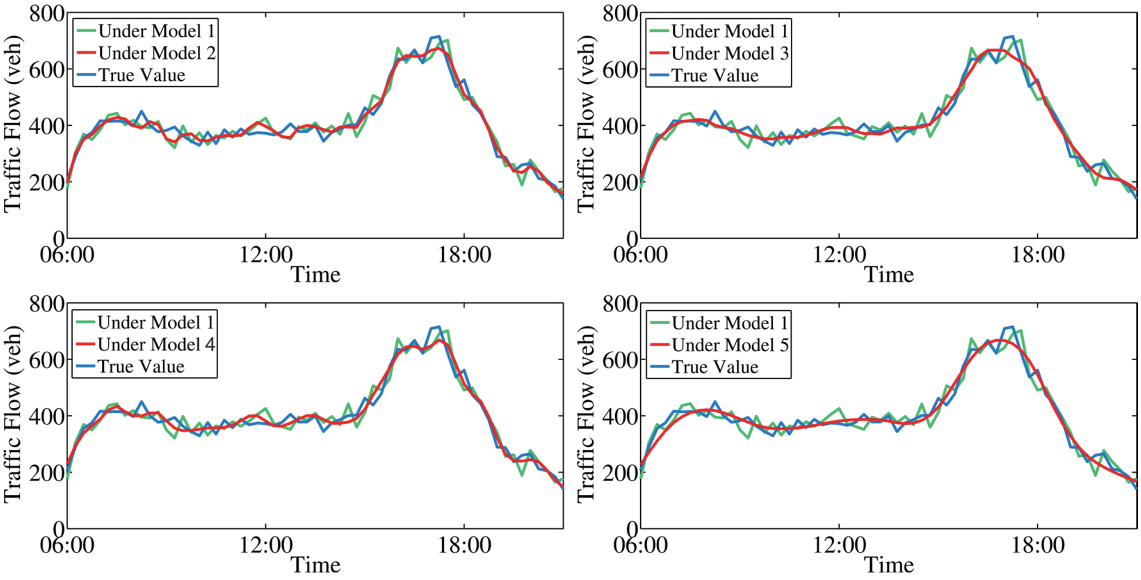

Fig. 5. (Color online) Different traffic flow prediction performance characteristics of path 3339 (LM91) on Wednesday for five models.

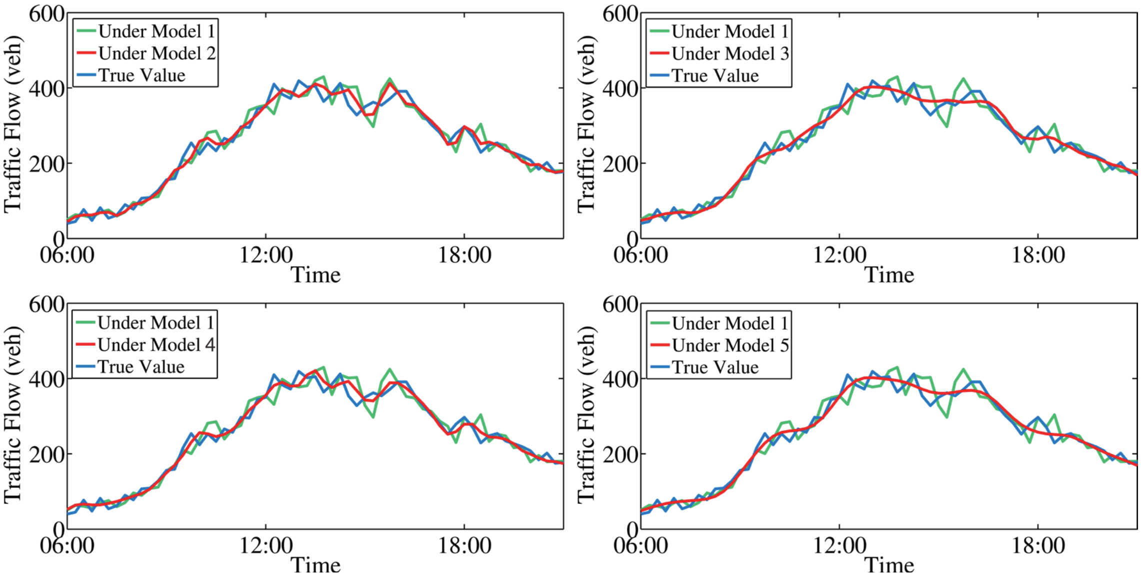

Fig. 6. (Color online) Different traffic flow prediction performance characteristics of path 3339 (LM91) on Sunday for five models.

the effectiveness of the models in the M-SDA system. The true values were also added. Table 3 shows the corresponding *RMSE* and *MAPE* values of the short-term traffic flow prediction results on Wednesday and Sunday, respectively.

Table 3
RMSE and *MAPE* values for five models in M-SDA system on Wednesday and Sunday.

	Wednesday		Sunday	
	RMSE	*MAPE* (%)	*RMSE*	*MAPE* (%)
Model 1	35.33	7.91	30.29	10.47
Model 2	24.04*	5.03	19.28	7.62
Model 3	25.25	5.03	17.71	7.40
Model 4	24.40	4.99*	16.20*	7.38*
Model 5	24.32	5.29	16.75	7.61

*Best performance.

It can be seen from the results shown in Table 3 that the prediction results acquired from Models 2–5 in the M-SDA system using historical measurements with noises separated were superior to those obtained from the T-SDA system (Model 1). Also, different good performance characteristics were produced by the different models in the M-SDA system. For example, *RMSE* and *MAPE* are 24.04 and 5.03% for Model 2 on Wednesday, and *RMSE* and *MAPE* were reduced by 10.93 (from 35.33 to 24.40) and 2.92% (from 7.91 to 4.99%) for Model 4 compared with the results for Model 1. Similar prediction results were obtained using the data collected on Sunday.

It is concluded from Figs. 5 and 6 and Table 3 that different noise separation scales have different impacts on assimilation models and assimilation prediction results. Also, from the lower *RMSE* and *MAPE* values shown in Table 3 and the better distributions displayed in Figs. 5 and 6, good results were obtained when using the models in the M-SDA system. This indicated that the M-SDA system built in this study is effective in improving prediction accuracy and that different noise separation scales have different impacts on prediction results.

4.3 Performance of M-SDA system for short-term traffic flow prediction

In this section, to further verify the effectiveness of the M-SDA system, it was applied to short-term traffic flow prediction of all the paths shown in Fig. 4(a) on Monday to Sunday. There are four models in the M-SDA system, which were constructed using historical measurements with noises separated under the four scales shown in Table 1. As an example of a detailed analysis, the prediction performance characteristics of the five paths shown in Fig. 4(b) are listed first. *RMSE* and *MAPE* values of the prediction results acquired from the models in the T-SDA system (Model 1) and the M-SDA system (Models 2–5) from Monday to Sunday are given in Fig. 7, and the average *RMSE* and *MAPE* values of the five paths are shown in Tables 4 and 5, respectively.

The prediction performance characteristics were different for the four models in the M-SDA system. As shown in Tables 4 and 5, compared with prediction results from the T-SDA system (Model 1), the average *RMSE* and *MAPE* were improved by various degrees when using the models from the M-SDA system. By taking path 3339 (LM91) as an example, the average *RMSE* and *MAPE* acquired from the T-SDA system (Model 1) were 37.12 and 8.46%, respectively, and the best prediction performance was acquired from the M-SDA system for Model 4, with the average *RMSE* reduced by 14.15 (from 37.12 to 22.97) and the relative

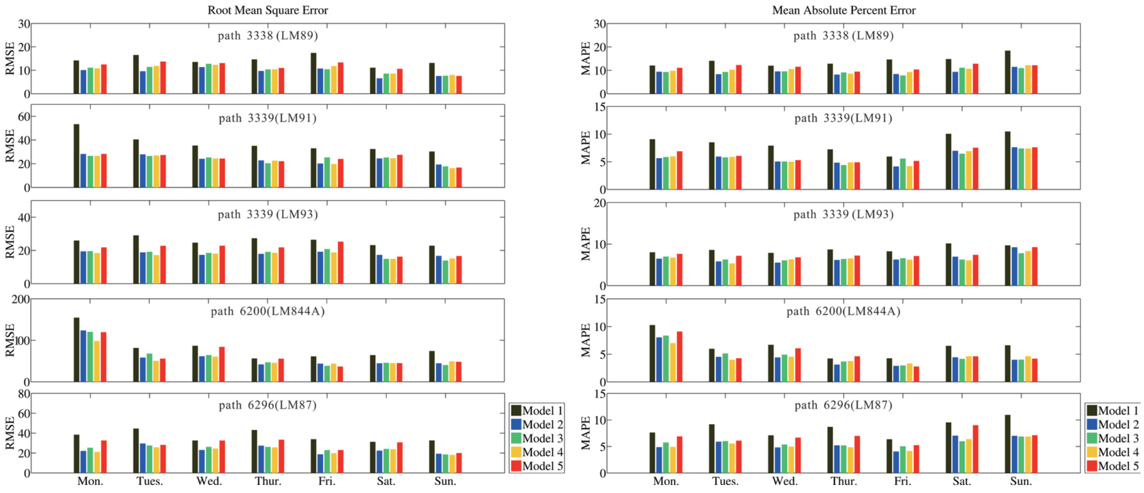

Fig. 7. (Color online) *RMSE* and *MAPE* values of five paths for five models from Monday to Sunday.

Table 4
Average *RMSE* values of five paths for five models.

Path	Model 1	Model 2	Model 3	Model 4	Model 5
3338 (LM89)	14.36	9.34*	10.31	10.51	13.20
3339 (LM91)	37.12	23.79	23.82	22.97*	24.30
3339 (LM93)	25.65	17.97	18.07	17.27*	24.26
6200 (LM844A)	82.72	59.65	60.52	56.08*	65.92
6296 (LM87)	36.67	23.17	24.37	22.63*	31.84

*Best performance.

Table 5
Average *MAPE* values of five paths for five models.

Path	Model 1	Model 2	Model 3	Model 4	Model 5
3338 (LM89)	14.10	9.22*	9.55	10.16	12.64
3339 (LM91)	8.46	5.74	5.78	5.74*	6.20
3339 (LM93)	8.76	6.62	6.63	6.50*	8.67
6200 (LM844A)	6.37	4.48*	4.74	4.55	5.39
6296 (LM87)	8.50	5.55	5.73	5.38*	7.49

*Best performance.

accuracy improved by 38.12%. The corresponding average *MAPE* was reduced by 2.72% (from 8.46 to 5.74%) and the relative accuracy was improved by 32.15%. Similar results were also obtained for the other four paths. The results of this test indicate that the multiscale models built using historical measurements with multiscale noise separation in the M-SDA system can achieve a higher prediction accuracy.

For further verification, the T-SDA and M-SDA systems built in this study were used for all the paths shown in Fig. 4(a). The corresponding average *RMSE* and *MAPE* values from Monday to Sunday are given in Tables 6 and 7, respectively. Results show that the average *RMSE* and

Table 6
Average *RMSE* values of all paths for five models.

	Model 1	Model 2	Model 3	Model 4	Model 5
Mon.	79.69	71.05*	73.46	72.11	77.16
Tues.	83.04	74.79*	81.64	75.32	79.07
Wed.	83.89	74.84*	77.54	77.91	82.37
Thur.	83.12	72.79*	75.01	74.51	75.64
Fri.	84.47	73.17*	74.71	74.95	81.90
Sat.	50.84	41.33	40.25*	41.04	46.44
Sun.	47.25	38.44	37.74*	38.72	42.40

*Best performance.

Table 7
Average *MAPE* values of all paths for five models.

	Model 1	Model 2	Model 3	Model 4	Model 5
Mon.	9.50	8.16*	8.22	8.28	9.25
Tues.	9.63	8.31*	8.94	8.37	8.98
Wed.	10.17	8.77*	9.03	8.98	9.84
Thur.	9.35	7.98*	8.15	8.22	8.14
Fri.	9.40	7.91*	8.10	8.17	8.31
Sat.	9.32	7.68*	7.75	7.80	8.23
Sun.	9.76	7.85	7.82*	7.99	8.91

*Best performance.

MAPE values from the M-SDA system (Models 2–5) are all smaller than those from the T-SDA system (Model 1). Taking the workday Monday and non-workday Sunday as examples, the average *RMSE* and *MAPE* from the T-SDA system (Model 1) on Monday were 79.69 and 9.50%, compared with 71.05 and 8.16% from the M-SDA system for Model 2, respectively. On the non-workday Sunday, the *RMSE* and *MAPE* values from the T-SDA system were 47.25 and 9.76%, compared with 37.74 and 7.82% from the M-SDA system for Model 3, respectively.

Overall, the results in Fig. 7 and Tables 4–7 suggest that the performance of the M-SDA system is higher than that of the T-SDA system, although different models in the M-SDA system have different effects on improving the prediction accuracy. The built M-SDA system can be effectively applied to predict short-term traffic flows.

5. Conclusions

The T-SDA system has been shown to be effective in short-term traffic flow prediction. However, short-term traffic flow data are always corrupted by local noises. To improve the accuracy of assimilation models and prediction results, this work sheds a new light on the impacts of noises in historical measurements on the construction of assimilation models and the accuracy of the prediction results. The M-SDA system combining the T-SDA system and noise separation methods was built to overcome the problems of noises in historical measurements. It was found to perform well in short-term traffic flow prediction applications. The main conclusions from the analysis results are as follows:

(1) Noises can be successfully separated from historical measurements under different scales using the DWT and EMD methods. Also, different noise separation scales give different noise separation results.

(2) Different noise separation scales have different impacts on assimilation prediction results. The prediction results of path 3339 (LM91) on Wednesday and Sunday were taken as examples for a detailed analysis. *RMSE* and *MAPE* were 24.04 and 5.03% for Model 2 on Wednesday, and *RMSE* and *MAPE* were reduced by 11.01% (from 35.33 to 24.32%) and 2.62% (from 7.91 to 5.29%) for Model 5 compared with the results from Model 1, respectively. Similar improvements in the prediction results were obtained using the data collected on Sunday.

(3) The built M-SDA system was successfully applied to short-term traffic flow prediction. The prediction results acquired from the M-SDA system outperformed those from the T-SDA system. After removing the noises existing in historical measurements used to build the measurement model in the T-SDA system, the M-SDA system is effective for predicting short-term traffic flow with a higher accuracy.

However, the conclusions of this study are based on an experiment with a time interval of 15 min. In future work, we will consider the application of the M-SDA system to shorter-term traffic flow prediction, such as to traffic flow prediction in urban areas, as traffic conditions in urban areas with a shorter time interval are much more complicated. Moreover, an adaptive scale model in the M-SDA system may be studied in the future to dynamically obtain the highest prediction performance of the M-SDA system.

Acknowledgments

We would like to thank the open data of Highways England for providing the traffic flow data used in the study.

References

1 Y. Xu, H. Chen, Q. J. Kong, X. Zhai, and Y. Liu: J. Adv. Transport. **50** (2016) 489. https://doi.org/10.1002/atr.1356

2 S. Dunne and B. Ghosh: J. Transp. Eng. **138** (2012) 455. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000337

3 S. Sun, R. Huang, and Y. Gao: J. Transp. Eng. **138** (2012) 1358. https://doi.org/10.1061/(asce)te.1943-5436.0000435

4 M. V. D. Voort, M. Dougherty, and S. Watson: Transport. Res. C-Emer. **4** (2012) 307. https://doi.org/10.1016/s0968-090x(97)82903-8

5 E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis: Transport. Rev. **24** (2004) 533. https://doi.org/10.1080/0144164042000195072

6 D. F. Parrish and J. C. Derber: Mon. Weather. Rev. **120** (1992) 747. https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2

7 I. Hoteit, D. T. Pham, G. Triantafyllou, and G. Korres: Mon. Weather Rev. **136** (2008) 317. https://doi.org/10.1175/2007mwr1927.1

8 R. Lei and M. Hartnett: Comput. Geosci. **99** (2017) 81. https://doi.org/10.1016/j.cageo.2016.10.012

9 Y. Xie, Y. Zhang, and Z. Ye: Comput-Aided. Civ. Inf. **22** (2007) 326. https://doi.org/10.1111/j.1467-8667.2007.00489.x

10 J. Guo, W. Huang, and B. M. Williams: Transport. Res. C-Emer. **43** (2014) 50. https://doi.org/10.1016/j.trc.2014.02.006

11 P. J. Smith, G. D. Thornhill, S. L. Dance, A. S. Lawless, D. C. Mason, and N. K. Nichols: Q. J. R. Meteor. Soc. **139** (2013) 314. https://doi.org/10.1002/qj.1944

12 G. Evensen: J. Geophys. Res-Oceans. **99** (1994) 10143. https://doi.org/10.1029/94jc00572

13 A. Fournier, G. Hulot, D. Jault, W. Kuang, A. Tangborn, and N. Gilet: Space Sci. Rev. **155** (2010) 247. https://doi.org/10.1007/s11214-010-9669-4

14 M. J. Deng and S. R. Qu: Comput. Intel. Neurosc. **2015** (2015) 1. https://doi.org/10.1155/2015/875243

15 E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias: Transport. Res. C-Emer. **43** (2014) 3. https://doi.org/10.1016/j.trc.2014.01.005

16 G. J. Shen, X. J. Kong, and X. Chen: Control. Eng. Appl. Inf. **13** (2011) 65. https://www.researchgate.net/publication/266522351_A_Shortterm_Traffic_Flow_Intelligent_Hybrid_Forecasting_Model_and_Its_Application

17 S. Rakshit, A. Ghosh, and B. U. Shankar: Pattern Recogn. **40** (2007) 890. https://doi.org/10.1016/j.patcog.2006.02.008

18 S. Kindermann, S. Osher, and P. W. Jones: Multiscale Model Sim. **4** (2005) 1091. https://doi.org/10.1137/050622249

19 V. Gupta, V. Chaurasia, and M. Shandilya: J. Vis. Commun. Image Represent. **26** (2015) 296. https://doi.org/10.1016/j.jvcir.2014.10.004

20 B. H. Tracey and E. L. Miller: IEEE Trans. Biomed. Eng. **59** (2012) 2383. https://doi.org/10.1109/tbme.2012.2208964

21 R. Yan, R. X. Gao, and X. Chen: Signal Process. **96** (2014) 1. https://doi.org/10.1016/j.sigpro.2013.04.015

22 N. E. Huang, M. L. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, and P. Gloersen: Proceedings A **459** (2003) 2317. https://doi.org/10.1098/rspa.2003.1123

23 M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells: IEEE Signal Proc. Lett. **3** (2002) 10. https://doi.org/10.1109/97.475823

24 R. E. Kalman: J. Basic Eng. Trans. **82** (1960) 35. https://doi.org/10.1115/1.3662552

About the Authors

Wenzhong Shi received his B.S. and M.S. degrees from Wuhan University, China, in 1985 and 1988, respectively, and his Ph.D. degree from Osnabruck University, Germany, in 1994. He is the Otto Poon Charitable Foundation Professor in Urban Informatics, the Chair Professor in GISci and Remote Sensing, the Director of PolyU-Shenzhen Technology and Innovation Research Institute (Futian), the Director of Smart Cities Research Institute, and the Head of Department of Land Surveying and Geo-Informatics. His research interests are in GISci, remote sensing, and urban informatics, focusing on analytics and quality control for spatial big data, object extraction and change detection from satellite images and LiDAR data, integrated indoor mapping technology, 3D and dynamic GISci modeling, and smart city applications. (john.wz.shi@polyu.edu.hk)

Runjie Wang received her B.S. degree from Chang'an University, China, in 2012 and her Ph.D. degree from Hong Kong Polytechnic University, China, in 2020. Her research interests are in analytics and quality control for sequential data assimilation and for traffic flow forecasting applications.
(runjie.wang@connect.polyu.hk)