Transport Reviews

ISSN: 0144-1647 (Print) 1464-5327 (Online) Journal homepage: https://www.tandfonline.com/loi/ttrv20

Spatiotemporal traffic forecasting: review and proposed directions

Alireza Ermagun & David Levinson

To cite this article: Alireza Ermagun & David Levinson (2018) Spatiotemporal traffic forecasting: review and proposed directions, Transport Reviews, 38:6, 786-814, DOI: 10.1080/01441647.2018.1442887

To link to this article: https://doi.org/10.1080/01441647.2018.1442887

Published online: 06 Mar 2018.

TRANSPORT REVIEWS
2018, VOL. 38, NO. 6, 786–814
https://doi.org/10.1080/01441647.2018.1442887

Spatiotemporal traffic forecasting: review and proposed directions

Alireza Ermagun[a] and David Levinson[b]

[a]Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL, USA; [b]School of Civil Engineering, University of Sydney, Sydney, Australia

ABSTRACT
This paper systematically reviews studies that forecast short-term traffic conditions using spatial dependence between links. We extract and synthesise 130 research papers, considering two perspectives: (1) methodological framework and (2) methods for capturing spatial information. Spatial information boosts the accuracy of prediction, particularly in congested traffic regimes and for longer horizons. Machine learning methods, which have attracted more attention in recent years, outperform the naïve statistical methods such as historical average and exponential smoothing. However, there is no guarantee of superiority when machine learning methods are compared with advanced statistical methods such as spatiotemporal autoregressive integrated moving average. As for the spatial dependency detection, a large gulf exists between the realistic spatial dependence of traffic links on a real network and the studied networks as follows: (1) studies capture spatial dependency of either adjacent or distant upstream and downstream links with the study link, (2) the spatially relevant links are selected either by prejudgment or by correlation-coefficient analysis, and (3) studies develop forecasting methods in a corridor test sample, where all links are connected sequentially together, assume a similarity between the behaviour of both parallel and adjacent links, and overlook the competitive nature of traffic links.

ARTICLE HISTORY
Received 14 March 2017
Accepted 15 February 2018

KEYWORDS
Traffic forecasting; spatial correlation; systematic review; traffic network; short-term traffic; competitive links; network structure

1. Introduction

Short-term traffic forecasting aims to predict the number of vehicles on a link during a given time slice, typically less than an hour. With the growing need to develop more adaptive traffic management systems, short-term traffic forecasting interests traffic engineers. This is a fundamental objective of advanced traffic management systems and advanced traveller information systems. Approaches generally take advantage of the fact that many of the cars that will be on one link soon are already on the network upstream of the relevant location, and of typical patterns of flow.

CONTACT Alireza Ermagun alireza.ermagun@northwestern.edu Department of Civil and Environmental Engineering, Northwestern University, 2145 Sheridan Rd, Evanston, IL 602088, USA

Does spatial interdependence exist between traffic links? Is embedding this dependency in short-term traffic forecasting methods propitious? If so, how is this information captured? These questions have been confronting researchers who seek to maximise the performance of the network by anticipating traffic conditions. Two strands of research tackled these questions in two discrete time spans. One benefits from the information of upstream and downstream traffic links as an input of the system. The other predefines the spatial dependence structure between traffic links and embeds this structure in forecasting methods. Irrespective of which strand is chosen, the success of the method heavily relies on detecting the spatial dependence structure.

Embedding the spatial components in traffic forecasting methods has been the focus of countless research papers over the past few years. The related literature has compelling evidence to support the potential of spatial components to augment traffic forecasting. Nevertheless, coupling the spatial components with forecasting methods may act as either a catalyst or a hindrance. It behaves as a catalyst when actual spatial information feeds the system and behaves as a hindrance when misrepresented spatial information causes erroneous results. Ample of methods have emerged aiming to extract spatial dependency between traffic links as accurately as possible (Ermagun and Levinson, 2018b). However, little is known about whether and to what extent the methods represent the spatial interdependence realistically.

This paper reviews studies that fall into the aforementioned two strands of research. Particularly, we delve into the existing research through the lens of a comprehensive systematic framework. This approach comprehensively searches the literature, rather than just one part of it, and thereby lowers the chance of bias. Drilling down further, we seek to answer the following questions in this review:

- What are the spatial components and their role in traffic forecasting?
- To what extent does spatial dependency exist between traffic links?
- How is spatial dependence captured and embedded in forecasting methods?
- Is the current knowledge exhaustive or crude?
- Where are the lacunae in the current literature?
- What directions should research take?

Answering these questions enables us to uncover what and how much we know about the effectiveness of spatial information in traffic forecasting methods. It also sheds light on the consistencies and inconsistencies of the findings across multiple studies and leads to identifying gaps in our knowledge that require further research.

The remainder of the paper is set out as follows. First, we discuss the methodology of the review. Second, we summarise the statistics of 130 research papers extracted from the pool of studies with our systematic approach. Third, we review and synthesise the extracted research papers from two perspectives: (1) methodological framework of the models and (2) methods for capturing and incorporating spatial information in the models. Fourth, we conclude the paper with a broad discussion on gaps of the current literature, and propose future directions.

2. Review methodology: a systematic approach

There is a general agreement on reviewing the literature "systematically" to avoid representing islands without continents. Despite the emphasis on systematic literature review, researchers adopt the following recipe sporadically (Booth, Sutton, and Papaioannou, 2016):

> Take a simmering topic, extract the juice of an argument, add the essence of one filing cabinet, sprinkle liberally with your own publications and sift out the work of noted detractors or adversaries.

To avoid this pitfall, we follow five steps in conducting a systematic review proposed by Khan, Kunz, Kleijnen, and Antes (2003):

Step 1: Framing questions for a review
Step 2: Identifying relevant work
Step 3: Assessing the quality of studies
Step 4: Summarising the evidence
Step 5: Interpreting the findings

To capture the potential range of published articles in the field, we identified relevant articles by an electronic search of Google Scholar, IEEE Xplore, and Scopus academic search engines along with electronic library records. The limited coverage of electronic sources does not cause any bias in our case, as we trace back utilisation of spatial information in traffic forecasting methods to 1984. We hunted for studies while considering manifold and distinct search keys not just simply in titles, keywords, and abstracts, but in the text of articles. Although this necessitated double effort, it resulted in extracting a more comprehensive pool of research. The main search keys were "traffic forecast", "forecasting traffic", "forecasting of traffic", "spatial", and "space". We searched for both "spatial" and "space" terms, as they are interchangeably used to describe spatial components in the literature of traffic forecasting.

We summarised the study exclusion process in Figure 1. This process encompasses three steps. First, we searched the literature to extract all articles including the combination of selected keywords as shown in Figure 1. This search led to extracting articles from diverse disciplines. Second, we executed four distinct assessment criteria to not only exclude irrelevant disciplines, but to only include articles that are germane to using spatial components for traffic forecasting. Thus, we excluded literature about wireless local area networks, internet traffic, railways, and groundwater, to name but a few. We also dropped articles where our search keys appeared in the introduction, literature review, recommendation, and reference sections. Concretely speaking, we perused the pool of articles closely and excluded articles which lack implementation of spatial information in traffic forecasting methods. This resulted in 113 English language research articles. Third, we systematically reviewed the lists of references from excluded articles. We then added those research papers that met the inclusion criteria in accordance with the second step. The consequence of this systematic search resulted in 130 publications in peer-reviewed journals, conference proceedings, and dissertations. The extensive list of references for the 130 extracted articles is provided in the Appendix. We have limited the scope of this review to only English language publications.

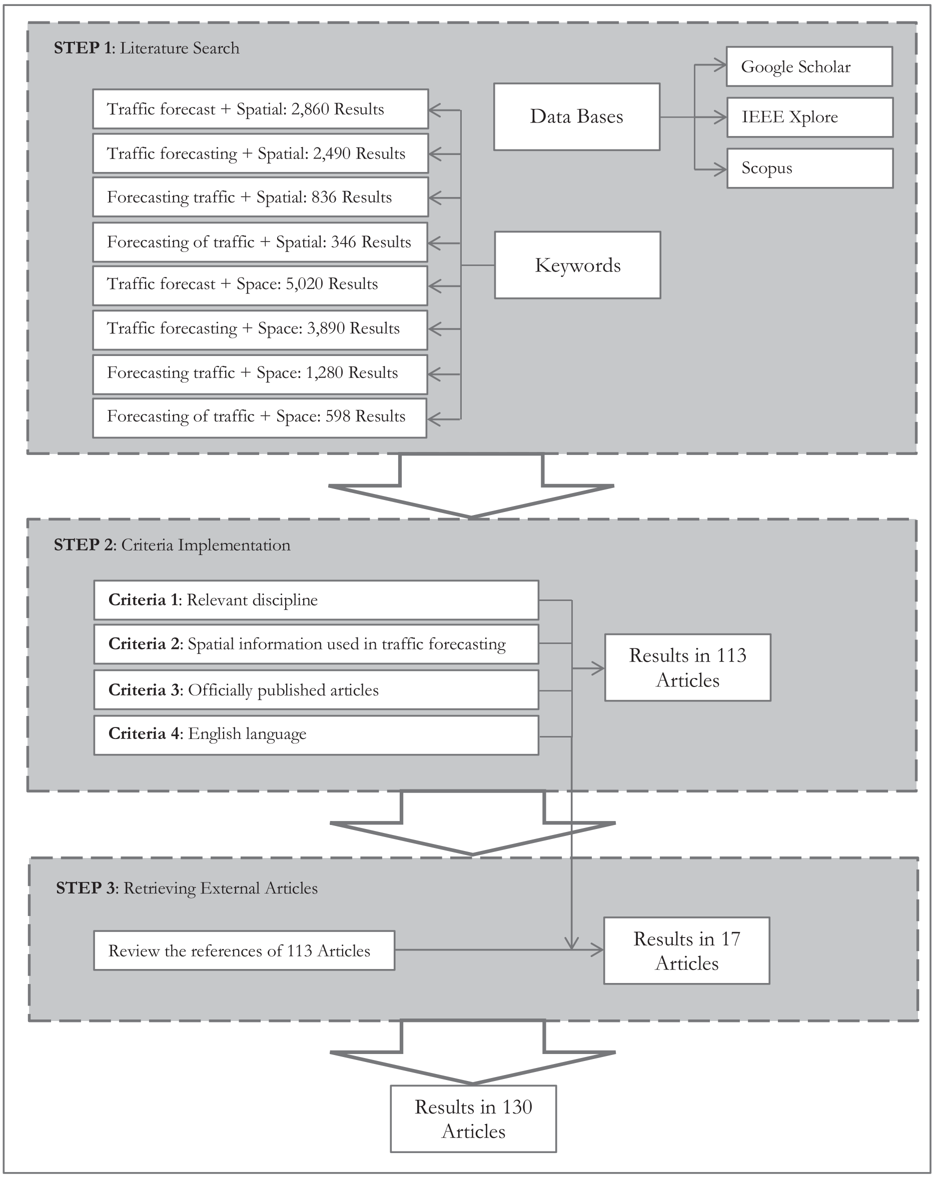


Figure 1. Flowchart diagram of study inclusion process.

3. Review statistics

This section provides a statistical overview of extracted articles. Table 1 classifies the source of the articles that embedded spatial components in traffic forecasting methods. As depicted in Table 1, 63.8% of articles were published in peer-reviewed journals. Almost 67.5% of the extracted articles appeared in *Transportation Research Part C*, *Transportation Research Record*, *IEEE Transactions on Intelligent Transportation Systems*, *Journal of Transportation Engineering*, *Computer-Aided Civil and Infrastructure Engineering*, *IET Intelligent Transport Systems*, *Journal of Intelligent Transportation Systems*, *Transportation*

Table 1. Distribution of publications by source.

Classification of sources	Number of retrievals	Percentage
Article division		
Scientific journals	83	63.8%
Dissertations	6	4.6%
Conference proceedings	41	31.6%
Journal source		
Transportation Research Part C: Emerging Technologies	12	14.5%
Transportation Research Record	11	13.3%
IEEE Transactions on Intelligent Transportation Systems	11	13.3%
Journal of Transportation Engineering	6	7.2%
Computer-Aided Civil and Infrastructure Engineering	4	4.8%
IET Intelligent Transport Systems	4	4.8%
Journal of Intelligent Transportation Systems	4	4.8%
Transportation Research Part B: Methodological	2	2.4%
Journal of Advanced Transportation	2	2.4%
Other	27	32.5%

Research Part B, and*Journal of Advanced Transportation*. This statistic reveals that articles on traffic forecasting using spatial components are concentrated in emerging technology journals. It is not surprising as traffic forecasting is an integral part of intelligent transportation systems. The other 32.5% of the articles appeared in 27 other journals.

To give the reader a sense of the temporal evolution of the field, we drew the life-cycle graph of publications in Figure 2. This figure shows the number of publications per year over the extracted articles in this review. As shown, utilising spatial components in traffic forecasting methods is an emerging research field. We designate 1984 as the historical starting point for earmarking spatial components as a potential input of forecasting methods. The growth phase of this field is laid as early as 2001. As portrayed in Figure 2, the number of publications had a significant jump in the past two years. The drop in the number of publications for 2016 is due to the time of search, which was 30 June 2016. We expect the field would continue its growth, and more research is needed to reach the apex of maturity as we discuss later in detail.

Figure 2. The life-cycle diagram of the research field.

We summarised the key characteristics of 130 extracted articles in Table 2, which include forecasting resolution, the type of data, number of traffic links incorporated in the study, and modelling framework. As alluded to in the preceding section, the extensive list of references for these articles is provided in the Appendix.

As for the type of output data, the larger part of spatiotemporal traffic forecasting concerns predicting traffic flow. Almost 65.3% of articles predicted traffic flow, while this share is around 19.2% for traffic speed and 16.1% for travel time. Three studies (Chandra, 2009; Chandra and Al-Deek, 2009; Van Grol, Inaudi, and Kroes, 2000) predicted both traffic flow and speed.

As for the area of implementation, two articles (Chen, Li, Tian, Chen, and Wang, 2012; Dong, Shao, and Li, 2009) did not indicate their study areas. Out of 128 remaining articles, 34 studies used urban street or urban road as their area of implementation. The target of the rest of the articles is expressway, freeway, highway, motorway, carriageway, or ring road, of which 23 studies and 5 studies conducted their research on interstate highways and U.S. highways, respectively. One challenge in short-term traffic forecasting is comparing the forecasting applications across studies and their areas of implementation. Although Vlahogianni, Golias, and Karlaftis (2004) noted the forecasting applications become more complex in urban areas, there is no concrete evidence demonstrating the weakness of prediction in urban areas. Comparing the forecasting applications across studies is almost impossible as studies:

(1) use different data resolution ranging from 0.5 min (Abdulhai, Porwal, and Recker, 1999, 2002; Chen et al., 2012; Liang and Wakahara, 2014) to 60 min (Ma, Zhou, and Abdulhai, 2015);
(2) employ various modelling techniques to forecast traffic conditions ranging from statistical methods (Clark, Dougherty, and Kirby, 1993) to machine learning techniques (Yu, Song, Guan, Yang, and Yao, 2016);
(3) measure distinct types of output data including traffic flow (Zhu, Peng, Xiong, and Zhang, 2016), speed (Fusco, Colombaroni, and Isaenko, 2016), travel time (Reza, Pulugurtha, and Duddu, 2015), and relative velocity (Kamarianakis and Prastacos, 2003);
(4) report the prediction accuracy at different time and link aggregations;
(5) report different measurement errors when calculating the prediction accuracy of models such as mean absolute percent error (Yu et al., 2016), root mean square error (Zhang and Zhang, 2016), and mean absolute error (Chen, Pao, and Lee, 2014).

As for the number of studied links, studies typically took into account a small number of traffic links for short-term traffic forecasting. Looking at Table 2, it is found that more than three-quarters of the articles analysed traffic networks encompassing less than 20 traffic links, of which 35% includes less than 5 traffic links. A handful of studies analysed traffic networks with more than 1000 traffic links. However, the spatial information of only a small part of the network neighbouring each study link is considered for short-term traffic forecasting. For example, Salamanis, Kehagias, Filelis-Papadopoulos, Tzovaras, and Gravvanis (2016) used a traffic network comprising 218,576 road segments in Berlin, Germany. However, only the spatial information of up to 13 neighbouring links was taken into consideration.

Table 2. Summary of literature.

No.	Study	Location	Road	Step (min)	Predictor	Data	Link	Method	
								ST	ML
1	Fusco et al. (2016)	Italy	Urban Highway	5	Speed	R	5	✓	✓
2	Zhu et al. (2016)	Germany	SUMO Traffic Network	5, 10, 15	Flow	S	19	✓	✓
3	Yu et al. (2016)	China	Urban Street	5	Flow	R	6		✓
4	Xia et al. (2016)	China	Urban Street	5	Flow	R	3		✓
5	Zhang and Zhang (2016)	US	U.S. Highway 290	5	Flow	R	6	✓	✓
6	Ko et al. (2016)	Korea	Expressway	1	Flow	R	5	✓	✓
7	Zhao and Sun (2016)	China	Urban Highway	15	Flow	R	10		✓
8	Jiang et al. (2016)	China	Ring Road	2	Speed	R	3	✓	✓
9	Polson and Sokolov (2016)	US	Interstate 55	5	Speed	R	21	✓	✓
10	Salamanis et al. (2016)	Germany	Urban Network	5	Travel time	R	218,576	✓	
11	Wu et al. (2016)	US	Urban Street	5	Flow	R	14	✓	✓
12	Xu et al. (2015)	China	Urban Street	10	Flow	R	17	✓	✓
13	Lv et al. (2015)	US	Statewide Freeway	5	Flow	R	–		✓
14	Zou et al. (2015)	US	Interstate 394	5	Speed	R	5	✓	
15	Ma et al. (2015)	Canada	Urban Highway	60	Flow	R	9	✓	✓
16	Schimbinschi et al. (2015)	Australia	Urban Network	15	Flow	R	4		✓
17	Dong et al. (2015)	China	Regional Freeway	2	Flow	R	12	✓	
18	Agafonov and Myasnikov (2015)	Russia	Urban Network	10	Travel time	R	3387		✓
19	Fusco et al. (2015)	Italy	Urban Highway	5	Speed	R	7	✓	✓
20	Zou et al. (2015)	China	Urban Street	5	Flow	R	3		✓
21	Reza et al. (2015)	US	Interstate 77	1	Travel time	R	28	✓	
22	Hou et al. (2015)	US	Interstate 270	15	Flow	R	8		✓
23	Shahsavari and Abbeel (2015)	US	Statewide Freeway	15	Flow	R	36	✓	✓
24	Xing et al. (2015)	China	Highway	15	Flow	R	120	✓	
25	Ahn et al. (2015)	Korea	Expressway	1	Flow	R	4		✓
26	Yang et al. (2015)	US	Urban Highway	10	Flow	R	3254	✓	✓
27	Yang (2015)	US	Statewide Highway	1	Speed	R	9	✓	
28	Dell'Acqua et al. (2015)	US	U.S. Highway 101	15	Flow	R	–	✓	✓
29	Ran et al. (2015)	US	Freeway	5	Speed	R	13		✓
30	Zhong and Ling (2015)	US	Interstate 35E	–	Flow	R	2		✓
31	Dong et al. (2015)	China	Urban Freeway	5	Flow	R	10	✓	✓
32	Wu et al. (2014)	China	Urban Street	5	Flow	R	5		✓
33	Cheng et al. (2014)	England	Urban Highway	5	Flow	R	22	✓	
34	Zhu et al. (2014)	China	Urban Street	15	Flow	R	3		✓
35	Chen et al. (2014)	US	Interstate 5	5	Speed	R	5		✓
36	Niu et al. (2014)	China	Urban Street	15	Flow	R	64		✓
37	Daraghmi et al. (2014)	Taiwan	Urban Street	2	Flow	R	13	✓	
38	Mohan et al. (2014)	Singapore	Expressway	5	Speed	R	12		✓
39	Yang et al. (2014)	US	Interstate 24	1	Speed	R	9	✓	
40	Dong et al. (2014)	China	Urban Freeway	2	Flow	R	12	✓	✓
41	Ratrout (2014)	US	Urban Street	15	Flow	R	4		✓
42	Zhao et al. (2014)	US	California State Route 85	5	Flow	R	2	✓	
43	Fabrizi and Ragona (2014)	Italy	Ring Road	3	Speed	R	3		✓
44	Qing et al. (2014)	China	Urban Street	5	Flow	S	11	✓	✓
45	Liang and Wakahara (2014)	Germany	SUMO Traffic Network	0.5	Flow	S	6	✓	
46	Zou et al. (2014)	US	U.S. Highway 290	5	Travel time	R	5	✓	
47	Haworth (2014)	England	Urban Freeway	5	Flow	R	22	✓	✓
48	Li et al. (2013)	US		5	Flow	R	3	✓	

(*Continued*)

Table 2. Continued.

No.	Study	Location	Road	Step (min)	Predictor	Data	Link	Method ST	Method ML
			California State Route 85						
49	Pan et al. (2013)	US	Freeway	5	Flow	R	7	✓	
50	Zeng and Zhang (2013)	US	Freeway	5	Travel time	R	3		✓
51	Fowe and Chan (2013)	US	Interstate 430	15	Flow	R	8	✓	
52	Han and Moutarde (2013)	France	Urban Network	15	Flow	S	13,627		✓
53	Huang and Wang (2013)	China	Ring Road	5	Flow	R	11	✓	
54	Zheng and Van Zuylen (2013)	Netherlands	Urban Road	1	Travel time	S	3		✓
55	Kamarianakis et al. (2012)	US	Freeway	5	Speed	R	–	✓	
56	Haworth and Cheng (2012)	England	Urban Freeway	5	Travel time	R	22		✓
57	Cheng et al. (2012)	England	Urban Freeway	5	Travel time	R	22	✓	
58	Guo et al. (2012)	England	Urban Street	15	Flow	R	2		✓
59	Guo et al. (2012)	England	Urban Street	15	Flow	R	–		✓
60	Pan (2012)	US	Expressway	5	Travel time	R	3	✓	
61	Ngan et al. (2012)	China	Urban Intersection	–	Flow	R	8		✓
62	Chen et al. (2012)	–	–	0.5	Flow	R	4		✓
63	Wu et al. (2012)	US	Urban Street	5	Flow	R	14	✓	
64	Yuan and Wang (2012)	China	Urban Street	5	Flow	R	–		✓
65	Sun et al. (2012)	China	Urban Freeway	15	Flow	R	31		✓
66	Pascale and Nicoli (2011)	US	U.S. Highway 101	15	Flow	R	11		✓
67	Djuric et al. (2011)	US	Interstate 35W	5	Speed	R	11	✓	✓
68	Han and Moutarde (2011)	France	Urban Network	15	Flow	S	13,627		✓
69	Samaranayake et al. (2011)	US	Interstate 880	2.5	Speed	R	–		✓
70	Cheng et al. (2011)	England	Urban Highway	5	Travel time	R	22	✓	
71	Deng and Jiang (2011)	–	Urban Highway	15	Flow	R	7	✓	
72	Min and Wynter (2011)	–	Highway	5	Flow	R	502	✓	
73	Khosravi et al. (2011)	Australia	Intercity Highway	15	Travel time	R	4		✓
74	Lippi et al. (2010)	US	Interstates 10, 605, 710	15	Flow	R	7		✓
75	Min et al. (2010)	China	Urban Street	5	Flow	R	50	✓	
76	Herring et al. (2010)	US	Urban Street	30	Travel time	S	322	✓	
77	Sun et al. (2010)	China	Expressway	5	Micro-LOS	R	18	✓	
78	Guo et al. (2010)	England	Urban Street	15	Flow	R	–		✓
79	Lee (2010)	Germany	Intercity Freeway	15	Flow	R	15	✓	
80	McCrea and Moutari (2010)	England	Country Highway	–	Flow	S	6		✓
81	Min et al. (2009)	China	Urban Street	5	Flow	R	10	✓	
82	Chandra and Al-Deek (2009)	US	Interstate 4	5	Flow and speed	R	5	✓	
83	Dong et al. (2009)	–	–	2	Flow	R	20		✓
84	Li and Lu (2009)	China	Ring Road	5	Flow	R	3		✓
85	Venkata and Chandra (2009)	US	Interstate 4	5	Flow and speed	R	5	✓	
86	Ghosh et al. (2009)	Ireland	Urban Street	15	Flow	R	10	✓	
87	van Hinsbergen et al. (2009)	Netherlands	Intercity Highway	5	Travel time	R	19		✓
88	Hu et al. (2009)	England	Urban Highway	3	Speed	S	8	✓	
89	Innamaa (2009)	Finland	Ring Road	5	Travel time	R	2		✓
90	Yue and Yeh (2008)	Hong Kong	Carriageway	–	Flow	R	7	✓	

(*Continued*)

Table 2. Continued.

No.	Study	Location	Road	Step (min)	Predictor	Data	Link	Method ST	Method ML
91	Stathopoulos et al. (2008)	Greece	Urban Street	3	Flow	R	2	✓	✓
92	Chandra and Al-Deek (2008)	US	Interstate 4	5	Speed	R	5	✓	
93	Dimitriou et al. (2008)	Greece	Urban Street	3	Flow	R	2	✓	✓
94	De Fabritiis et al. (2008)	Italy	Ring Road	3	Speed	R	–		✓
95	Wu et al. (2008)	US	Interstate I-5	5	Flow	R	2		✓
96	Hu et al. (2008)	US	Statewide Freeway	5	Flow	R	4		✓
97	Van Lint (2008)	Netherlands	Intercity Freeway	5	Travel time	R	14		✓
98	Ye et al. (2008)	China	Urban Street	5	Flow	R	8	✓	
99	Vlahogianni et al. (2007)	Greece	Urban Street	3	Flow	R	4	✓	✓
100	Yue et al. (2007)	Hong Kong	Carriageway	1	Flow	R	7	✓	
101	Sun and Zhang (2007)	China	Urban Highway	15	Flow	R	20		✓
102	Xie and Zhang (2006)	US	Interstate 80	5	Flow	R	4		✓
103	Van Lint (2006)	Netherlands	Intercity Freeway	1	Travel time	R	26		✓
104	Wang et al. (2006)	–	Hypothetical Freeway	1	Flow	S	23	✓	
105	Vlahogianni et al. (2005)	Greece	Urban Street	3	Flow	R	3	✓	
106	Sun et al. (2005)	China	Urban Highway	15	Flow	R	31		✓
107	Kamarianakis and Prastacos (2005)	Greece	Urban Street	7.5	Flow	R	25	✓	
108	Van Lint (2006)	Netherlands	Intercity Freeway	1	Speed	S	19		✓
109	Bajwa et al. (2005)	Japan	Expressway	5	Travel time	R	5		✓
110	Innamaa (2005)	Finland	Intercity Highway	1	Travel time	R	4		✓
111	Ishak and Alecsandru (2004)	US	Interstate 4	5	Speed	R	3		✓
112	Kamarianakis et al. (2004)	Greece	Urban Street	7.5	Flow	R	11	✓	
113	Alecsandru (2003)	US	Interstate 4	5	Speed	R	–		✓
114	Vlahogianni et al. (2003)	Greece	Urban Street	3	Flow	R	–		✓
115	Kamarianakis and Prastacos (2003)	Greece	Urban Street	7.5	Relative velocity	R	25	✓	
116	Stathopoulos and Karlaftis (2003)	Greece	Urban Street	3	Flow	R	5	✓	
117	Ishak et al. (2003)	US	Interstate 4	5	Speed	R	4		✓
118	Hu et al. (2003)	China	Urban Intersection	–	Speed	R	60		✓
119	Van Lint et al. (2003)	Netherlands	Intercity Highway	–	Travel time	S	13		✓
120	Van Lint et al. (2002)	Netherlands	Intercity Highway	–	Travel time	S	12		✓
121	Tebaldi et al. (2002)	US	Interstate 5	1	Flow	R	15	✓	
122	Abdulhai et al. (2002)	US	Interstate 5	0.5, 1, 2, 5, 15	Flow	R	3		✓
123	Williams (2001)	France	Motorway	30	Flow	R	–	✓	
124	Van Grol et al. (2000)	Netherlands	Motorway	1, 5, 10	Flow and speed	R	2	✓	
125	Park and Rilett (1999)	US	U.S. Highway 290	5	Travel time	R	6	✓	✓
126	Abdulhai et al. (1999)	US	Interstate 5	0.5	Flow	R	9		✓
127	Park et al. (1998)	US	Interstates 35, 10	5	Flow	R	4	✓	✓
128	Larry (1995)	US	Urban Street	5	Flow	R	4	✓	
129	Clark et al. (1993)	England	Urban Street	5	Flow	R	3	✓	
130	Okutani and Stephanedes (1984)	Japan	Urban Street	5	Flow	R	4	✓	

Note: R: real data; S: simulation data; ST: statistical; ML: machine learning.

4. A review of model selection and testing

Selection of an appropriate model is a major challenge in short-term traffic forecasting. In a traditional classification, models fall into either parametric or non-parametric techniques. Out of 130 extracted articles, 43 studies used parametric techniques, 72 studies used non-parametric techniques, and 9 studies employed both parametric and non-parametric techniques. The remaining 6 articles applied a combination of parametric and non-parametric techniques. We eschew comparing parametric and non-parametric techniques as previous reviews (Smith, Williams, and Oswald, 2002; Vlahogianni et al., 2004) allocated a comprehensive discussion about their advantages and disadvantages.

In a different classification, models fall into either statistical or machine learning methods. The two methods differ on philosophy, goals, model development process, and knowledge acquisition. Statistical methods concern inference and estimation, aim at providing a model that offers insights on the data, assume the functional form a priori, make a number of hypotheses and place restrictions, and are tough disciplines. However, machine learning methods concern implementation, aim at providing an efficient prediction, approximate the functional form via learning, are feasible without prior models or error distribution specifications, and are easy to apply using software packages.

Out of 130 extracted articles, 48 studies used statistical methods, 59 studies used machine learning methods, and 23 studies employed both statistical and machine learning methods. Figure 3 depicts the life-cycle graph of articles using statistical and machine learning methods. As shown, machine learning methods were employed later than statistical methods in the life-cycle of spatiotemporal short-term traffic forecasting. However, they received more attention over time than statistical models, particularly in recent years.

The remainder of this section takes a further look at studies that compared the prediction accuracy of statistical and machine learning methods.

Zhu et al. (2016) did not find a significant difference between the Bayesian network (BN) and Autoregressive Integrated Moving Average (ARIMA) model in predicting traffic flow. However, the proposed linear conditional Gaussian BN outperformed ARIMA model in

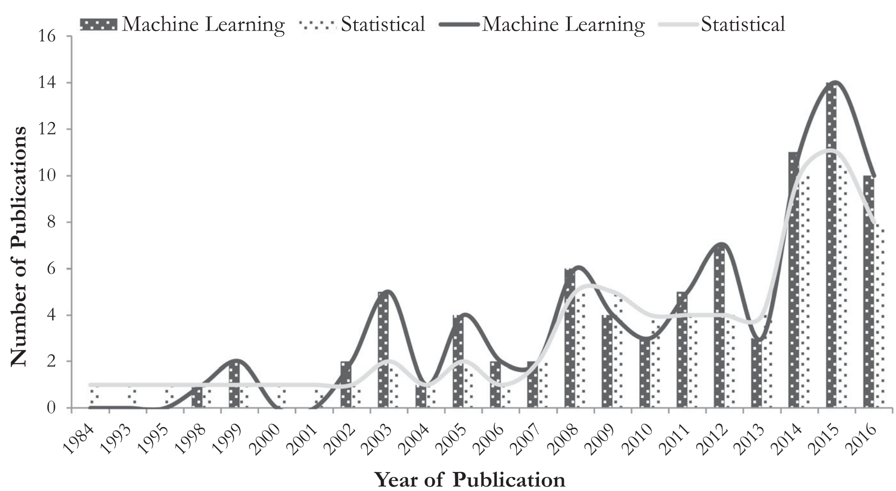


Figure 3. Comparing the distribution of statistical and machine learning methods over time.

all 5, 10, and 15 min horizons. Zhang and Zhang (2016) compared historical average (HA), vector autoregression (VAR), and general regression neural network forecasting models, while predicting traffic flow of six selected links extracted from U.S. Highway 290 in Houston, TX. They concluded that the VAR model has the highest forecasting accuracy when there are no or few missing data. They also noted that the HA model, which has already been applied to the urban traffic control systems and other traveller information systems because it is easy to implement and understand, has the worst prediction accuracy.

Jiang, Zou, Zhang, Tang, and Wang (2016) compared five machine learning methods: (1) back propagation neural network (BPNN), (2) nonlinear autoregressive model with exogenous inputs neural network (NARXNN), (3) support vector machine with radial basis function as kernel function (SVM-RBF), (4) support vector machine with linear function (SVM-LIN), and (5) multilinear regression with three statistical methods: autoregressive integrated moving average (ARIMA), VAR, and space-time (ST) model. Using the travel speed of three links collected from 4th ring road in Beijing, they concluded: (1) the BPNN, NARXNN, and SVM-RBF clearly outperform two traditional ARIMA and VAR statistical models and (2) as time step increases, the ST model consistently and VAR model in some cases provide the lowest errors.

Wu, Chen, Lu, and Yang (2016) compared three well-established traffic flow prediction models, namely autoregressive moving average (ARMA), spatiotemporal ARMA, and artificial neural network (ANN) in different traffic conditions. The results were mixed. The authors did not find any superiority of the machine learning method to statistical methods in short horizons. In longer horizons, however, the machine learning method outperformed the statistical methods. To predict traffic flow in Shanghai urban road network, Zhu et al. (2016) tested autoregression (AR), multivariate adaptive regression splines (MARS), support vector regression (SVR), seasonal autoregressive integrated moving average (SARIMA), spatiotemporal Bayesian MARS, and variable selection-based SVR. The results indicated the superiority of machine learning methods to statistical methods.

Shahsavari and Abbeel (2015) compared the simple univariate ARIMA and SARIMA with the Graph Neural Network (GNN) model, while predicting traffic flow. He concluded that ARIMA or SARIMA slightly outperforms the GNN model in one or at most two-step ahead prediction problems. However, the GNN performance is larger than statistical methods when it comes to a longer forecasting horizon. He then noted that GNN is more efficient in extracting long-term dependencies and learning the network dynamics, unlike simple time-series models. Earlier studies also showed the superiority of machine learning methods, including neural network model (Qing, Yongqin, Yongguo, and Qingming, 2014), Elman model (Dong, Shao, Richards, and Han, 2014), and adaptive hybrid fuzzy rule-based system (Dimitriou, Tsekeris, and Stathopoulos, 2008) to ARIMA.

This trajectory reveals that machine learning methods, which have attracted more attention in recent years, outperform the naïve statistical methods such as historical average, real-time profile, and exponential smoothing. However, there is not a certain superiority when machine learning methods are compared with advanced statistical methods such as space-time, spatiotemporal autoregressive integrated moving average, and VAR. To enable definite conclusions to be reached, we need more comprehensive studies comparing the accuracy of prevalent machine learning and statistical methods. We also alert researchers and practitioners to be cautious when deciding the superiority of the models across studies as each study considers networks of different topologies and sizes, distinct types of roads, and congested versus uncongested condition.

5. A review of spatial component testing

In 1984, Okutani and Stephanedes (1984) were the first to achieve a better traffic flow prediction on a link by taking into account the spatial information of its upstream feeder links. Twenty years later, Kamarianakis and Prastacos (2003) borrowed a model, the so-called space-time autoregressive integrated moving average (STARIMA), from the regional science literature to forecast relative velocity on major arterials of Athens, Greece. Although the fundamental of STARIMA is laid as early as 1975 by Cliff and Ord (1975), they were the first to test this model in a traffic forecasting framework. The STARIMA family model is considered a generic form of autoregressive linear models used in traffic forecasting. This model is quite distinct from the traditional autoregressive integrated moving average (ARIMA) model by capturing the spatial information of neighbouring links for traffic forecasting. The taxonomy of this family of models is depicted in Figure 4. Letting Y_t is the vector of observations at time t in this family, $\triangledown$ is a $l \times l$ difference operator matrix such that:

$$\triangledown Y_t = Y_t - Y_{t-1},$$
$$\triangledown^2 Y_t = \triangledown(\triangledown Y_t) Y_t - 2Y_{t-1} + Y_{t-2},$$

where g is the number of differences, p is the autoregressive order, m is the moving-average order, p_z is the spatial order of the zth autoregressive term, m_z is the spatial

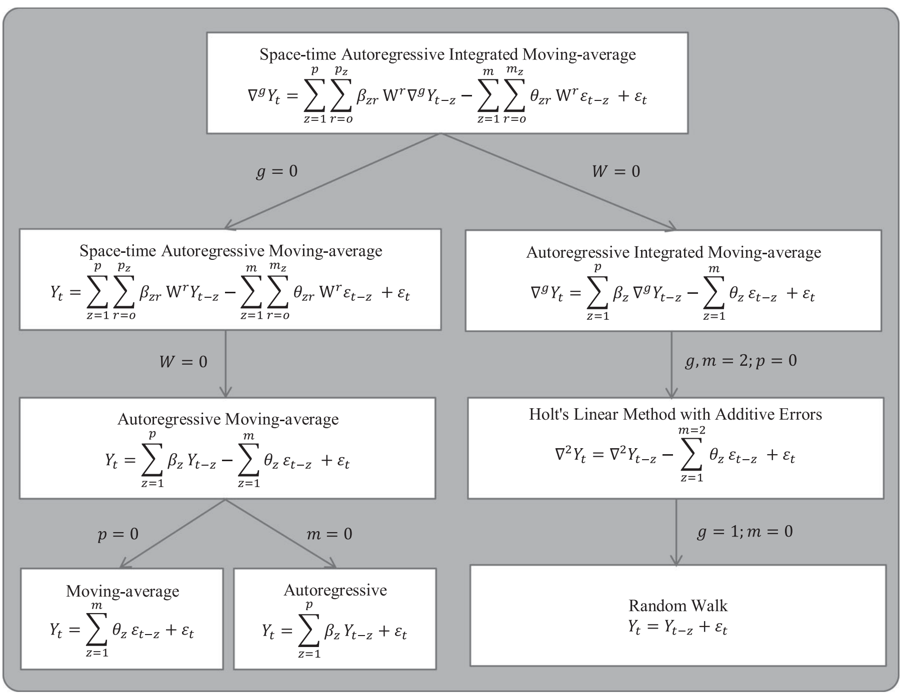


Figure 4. Taxonomy of spatiotemporal family models.

order of the zth moving-average term, β_{zr} is the autoregressive parameter at temporal lag z and rth spatial lag, θ_{zr} is the moving-average parameter at temporal lag z and rth spatial lag, W_r is the $I \times I$ matrix of spatial weights for rth-order neighbours, and ε_t is the random normally distributed error vector at time t. The components of the spatial weight matrix regularly satisfy three major rules:

(1) $w_{i,j} \geq 0$,
(2) $w_{i,i} = 0$, and
(3) $\sum_{j=1}^{I} w_{i,j} = 1$, for all $i = 1, 2, \ldots, I$ (Bavaud, 1998).

As shown in Figure 4, the space-time autoregressive integrated moving-average model branches into two leaves: (1) space-time autoregressive moving-average (STARMA) models when the number of temporal differences equals zero and (2) autoregressive integrated moving-average (ARIMA) models when the spatial term is eliminated from the equation. The STARMA family collapses to the autoregressive moving-average (ARMA) model when the spatial term is eliminated from the equation. This model is employed to understand and predict future traffic conditions as a statistical analysis tool and involves an AR and a moving-average (MA) part. The AR part regresses the vector of observations at time t against the vector of observations at time $t-1$. The MA part smooths out short-term fluctuations by modelling the error term as a linear combination of error terms occurring contemporaneously and at various times in the past. The ARIMA family collapses to Holt's Linear Method with Additive Errors when the autoregressive order equals zero, and the number of differences and the moving-average order equal two. To the best of the authors' knowledge, this model has not been used in short-term traffic forecasting. A unit reduction in the number of differences and a two units reduction in the moving average transform this model into the random walk model.

The studies of Okutani and Stephanedes (1984) and Kamarianakis and Prastacos (2003) formed the essence of a methodological strand of thinking of different points in time. They acknowledged embedding spatial information as the potential of enhancing traditional temporal models. These methods have burgeoned and developed in the literature. The spatiotemporal methods stand on the foundation of traditional temporal techniques. The only refinement is benefiting from the spatial information to advance the accuracy of predictions. The literature that discusses this subject has been prolific. For instance, Smith et al. (2002) classified temporal traffic forecasting models into parametric and non-parametric and discussed their pros and cons in detail. Vlahogianni et al. (2004) also broadly reviewed the short-term traffic forecasting methods and compared the proposed models in parametric and non-parametric framework. We hence eschew digging into the performance of models and their formulations. Rather, we elaborate the results of the studies through the lens of spatial components effectiveness and modelling performance. To achieve this, we review the studies in three separate classes.

5.1. Class 1: spatial effectiveness emphasis

In this class, the studies aim to examine the effectiveness of spatial components by comparing models with and without spatial information. Williams' (2001) ARIMAX model treats the upstream traffic flow series as transfer function inputs into the ARIMA model.

Embedding the spatial factor enhanced the accuracy of traffic flow forecasting by 15.6%. To forecast the traffic speed in five stations on I-4 in the downtown region of Orlando, FL, Chandra and Al-Deek (2008) developed a univariate ARIMA time-series model and a VAR model including the information of neighbouring links. Comparing both models, they found VAR significantly outperforms ARIMA. This is consistent with Chandra and Al-Deek (2009).

To investigate whether the inclusion of spatial information improves the accuracy of the ANN model, Zeng and Zhang (2013) compared the state-space neural network (SSNN) model with traditional ANN models. The SSNN model consistently outperforms other neural network models in both short and long horizons.

To investigate the effectiveness of spatial information, Dong, Xiong, Shao, and Zhang (2015) compared a spatiotemporal model with traditional ARIMA and a linear regression model encompassing only spatial information. The output of the models affirmed the superiority of the spatiotemporal model. It was also noted that the temporal input factor provides more accurate information than the spatial input factor in uncongested situations. In congested conditions, it reverses.

5.2. Class 2: modelling performance emphasis

Studies of this class compare the performance of sundry modelling techniques to introduce the most efficient method. Kamarianakis and Prastacos (2005) embedded the spatial information in the traditional ARIMA model and compared its performance with STARIMA, where the spatial components are captured with a spatial weight matrix. The performance of both models was found quite close. However, the STARIMA model included 7 parameters and a naïve spatial weight matrix (first- and second-order adjacent matrix), whereas the ARIMA model encompassed 75 different parameters. Sun, Zhang, and Zhang (2005) employed both spatial and temporal information and compared the accuracy of random walk, Markov chain, and BN methods by the RMSE. The findings stated that the BN performs better than Markov chain, and the latter outperforms the random walk model.

Min, Hu, Chen, Zhang, and Zhang (2009) compared the accuracy prediction power of Dynamic STARIMA with multivariate adaptive regression splines (MARS). The former and STARIMA are alike in structure, whereas the spatial weight matrix of the Dynamic STARIMA is derived from traffic flow information of links, and not simply adjacency. It enables the model to be updated dynamically in a real network. The latter is a non-parametric model. The comparison of two models indicated the superiority of Dynamic STARIMA. Interestingly, Ye, He, Hu, and Zhang (2008) found that MARS is more accurate than linear regression and neural network methods. This may result in the superiority of Dynamic STARIMA over ANN. Min, Hu, and Zhang (2010) generalised the STARIMA model and introduced GSTARIMA model, which relaxes the assumption that the autoregressive parameters and the moving-average parameters are the same for all traffic locations. They noted the performance of GSTARIMA model exceeded the STARIMA model.

5.3. Class 3: hybrid emphasis

Class 3 is a combination of Class 1 and Class 2. We thereupon labelled this class hybrid analysis emphasis, as the studies of this class both explore the potential of spatial information and compare the modelling techniques. One comprehensive study developed four different artificial neural network models and compared the accuracy of them with historical average, Kalman filtering, real-time profile, and exponential smoothing (Park and Rilett, 1999). The four artificial network models were distinct in input data, which ranged from no spatial information to spatiotemporal information of four upstream and downstream links. Comparing the mean average percentage error of the models, it is found that (1) adding spatial information of upstream and downstream traffic links augments the forecasting models for longer horizons, but not for short horizons and (2) Kalman filtering is superior to other models in predicting one or two steps ahead, while the neural network models outperform the other models in longer horizons.

Kamarianakis and Prastacos (2003) compared the forecasting performance of different statistical methods including historical average, ARIMA, VARMA, and STARIMA models. Comparing the RMSE of the models, it is inferred that (1) the HA method performs significantly worse than the other three models and (2) embedding the spatial component does not improve the prediction accuracy of the models as the ARIMA model outperforms STARIMA. The authors claimed the small number of loop detectors and their dispersed locations in the extracted network cause the unexpected superiority of ARIMA to STARIMA.

Vlahogianni, Karlaftis, and Golias (2007) developed modular predictor incorporating spatiotemporal patterns to predict the traffic volume of four traffic links in Greece. The proposed model was then compared with ARIMA and state-space models developed by Stathopoulos and Karlaftis (2003), and genetically optimised multilayer perceptrons (MLP) and statistic MLP developed by Vlahogianni, Karlaftis, and Golias (2005) using the same data. Comparing the mean relative percentage error of the models, it is concluded (1) adding the spatial component enhances the prediction accuracy (Stathopoulos and Karlaftis, 2003), (2) the neural network method outperforms statistical models, and (3) the accuracy of the neural network method depends on the prediction technique, where the proposed modular predictor performs better than statistic MLP and equally well to genetically optimised MLP.

Wu, Yang, Zhu, and Yu (2014) added the spatial information in k-nearest neighbour model to enhance the accuracy of traffic flow forecasting in urban roads of Guiyang, China. The results of comparing the developed model with historical average and neural network models indicated: (1) including both temporal and spatial information reduces the error significantly in comparison with the model with only temporal information and (2) the HA model performs the worst.

This trajectory leads us to the conclusion that irrespective of which method is selected, spatial information inclusion in short-term traffic forecasting models boosts the accuracy of prediction, particularly in congested traffic regimes and longer time horizons.

6. A review of methods for capturing spatial information

It has been over three decades since spatial information was first captured in a traffic corridor for the sake of traffic flow prediction (Okutani and Stephanedes, 1984). In this section, we discuss the evolution of techniques for dealing with capturing spatial information for traffic forecasting. We take a fairly narrow view of analysis and delve into the emerged approaches from two conceptual aspects. For each aspect, we elaborate on the nature of spatial components used in traffic forecasting and identify the notion behind an objective evaluation of approaches.

6.1. A naïve approach

Traffic conditions of a downstream section of a road are highly associated with traffic conditions upstream (as those vehicles will ultimately travel to the link in question). Thereupon, spatial information of upstream sections may capture the dynamics of traffic. Following the study of Okutani and Stephanedes (1984), Head (1995) utilised the traffic flow of detectors on the approach of each upstream intersection to predict future arrivals. He noticed that the longer horizons are achieved when spatial information is embedded in traffic forecasting methods. In another study, Park, Messer, and Urbanik (1998) found that the traffic flow of upstream links is highly correlated with the study link, and mentioned that spatial information is as informative as temporal information. Stathopoulos and Karlaftis (2003) predicted traffic flow in an urban corridor while using the spatial information of four consecutive loop detectors in the upstream of the study section. Although they acknowledged spatial information as a catalyst, they noticed farther links are correlated with the study link over a longer time lag.

Not only is a link affected by its upstream links, but downstream links also may affect traffic conditions of their upstream links. Abdulhai et al. (2002) benefited from both upstream and downstream flow information to take backward propagating shockwaves into account. Studies took a step forward by examining to what extent the downstream information is crucial in traffic forecasting. Van Lint, Hoogendoorn, and Van Zuylen (2002) highlighted that the downstream information plays a more critical role than upstream information in congested situations for travel time forecasting. However, no superiority was observed in uncongested conditions. In the congested regime, Djuric, Radosavljevic, Coric, and Vucetic (2011) also concluded that the current speed of the downstream link has a greater weight than the upstream link for speed forecasting. This is also confirmed by Daraghmi, Yi, and Chiang (2014), who made the same conclusion for traffic flow forecasting on an arterial road. Zou, He, Zhang, Du, and Gao (2015) used the information of two upstream and two downstream links. They developed distinct models to explore the role of downstream and upstream links in forecasting of traffic speed. No significant difference was found between using either downstream or upstream information for various prediction horizons in both congested and uncongested regimes.

6.2. A modest approach

Studies corroborated the hypothesis about enhancing the accuracy of forecasting methods by incorporating the information of neighbouring links. However, little information was known about which and how many links are needed to be included in forecasting methods. Two criteria were introduced to select the neighbouring links: (1) correlation-coefficient assessment and (2) distance adjustment. The former probes deeply into the data to explore whether and to what extent the information of neighbouring links is correlated with the study link. The highly correlated links are then selected as an input of forecasting methods. The latter borrows from regional science, and more specifically from the first law of geography. In accordance with this law, every link is related with every other links, but near links are more related than distant links. Despite the existence of many alternative methods to define the nearness and distance threshold in regional science, the traffic forecasting field has benefited mostly from spatial information of adjacent links.

From the correlation-coefficient assessment side, Sun et al. (2005) calculated the Pearson correlation coefficient to rank the input spatial and temporal exogenous variables. They then selected the four most correlated upstream and downstream links in different time regimes. Building on their experimental results, they concluded that not only near links, but also distant links in a traffic network have high correlation coefficients. Likewise, Chandra and Al-Deek (2008, 2009) utilised cross-correlation function and found that past values of an input series influence the future values of a response series. Hu, Xie, Song, and Wu (2008) also adopted the cross-correlation function to select the relative neighbouring links, rather than the selection of immediate upstream and downstream links. The results of the analyses showed the immediate upstream and downstream link as well as the eighth link located in the upstream are the most correlated links. They also found that the downstream link is more effective than upstream links and validated this by the existence of a ramp between the upstream link and the study link, which reduces the correlation.

From the distance adjustment side, most studies using these criteria fall into the spatio-temporal methodological category and prejudge the spatial dependency by creating a spatial weight matrix. As we mentioned, two methods are adopted to identify the components of a spatial weight matrix in traffic forecasting. One simply assumesthat just adjacent links have a spatial dependence with the study links. The other takes a step forward more comprehensively measuring the spatial dependency and states both adjacent and distant links are spatially correlated with the study link; however, the strength of the dependency is reduced by increasing the distance. In traffic forecasting, the ring of dependency is labelled by "order". For example, the first-order adjacency matrix shows the dependency between the study link and its immediate adjacent links. The second-order adjacency matrix, however, indicates the links that are connected to the study links indirectly and with having the first-order links in middle.

Kamarianakis, Prastacos, and Kotzinos (2004) used the first- and second-order adjacency matrix to capture spatial dependency. Studies using the distance adjusted approach simply expect all adjacent links have a similar effect on the study link. Thereupon, spatial weight matrices encompass binary elements, in which zero and one values stand for spatial independence and spatial dependence, respectively. These matrices are occasionally row normalised for statistical and prediction reasons that lead to not binary elements. Although this normalisation results in dissimilar spatial dependency, this dissimilarity does not stem from a conceptual traffic theory. To our best knowledge, only three studies considered dissimilar spatial dependency in creating spatial weight matrices. One studied the traffic flow forecasting of a link using the flow information of the upstream T-junction (Min et al., 2009). The weight of spatial dependence for each upstream links equals the traffic flow ratio of each link to the sum of the flow in the T-junction. This needs a dynamic update of spatial weights in real time. The second used the speed differentials over space formula and defined the spatial dependency between two links as the difference between the average speeds of links divided by their distance (Cheng, Wang, Haworth, Heydecker, and Chow, 2011). Likewise, the third employed the speed differentials over speed formula and defined the spatial dependency between two links as the difference between the average speeds of links divided by the speed of the target link (Cheng, Wang, Haworth, Heydecker, and Chow, 2014). The theoretical concept behind this calculation is a decrease in traffic speed on one link follows a relative decrease in traffic speed of its adjacent link.

7. Closing remarks and opportunities for future research

In this section, we intend to deal with the last two questions from the introduction:

- What is the lacunae in the current literature?
- What directions should research take?

To answer these questions, we need to discuss the types of traffic networks studied in the literature. However, a preliminary knowledge of graph theory is required, which drove us to provide a brief introduction here. For details the reader may refer to Rodrigue, Comtois, and Slack (2013).

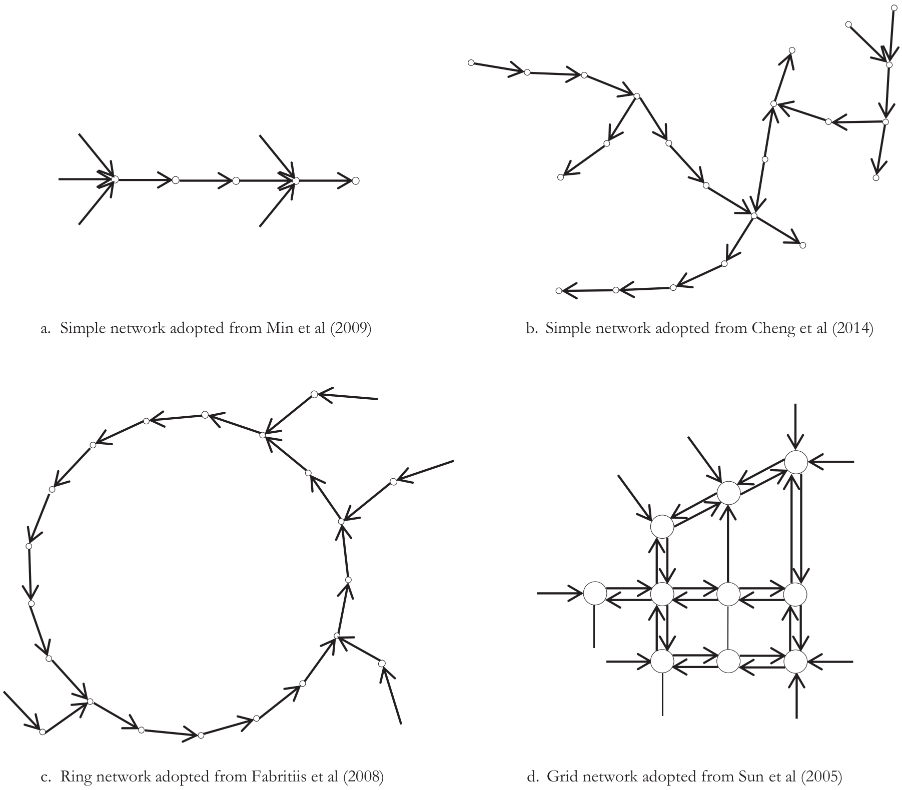
A graph is a collection of nodes that are connected by links. In accordance with graph theory, the following terminologies are drawn:

- Two links are parallel if they connect the same pair of nodes.
- Two links are adjacent if they share a common node.
- A link is loop if its two nodes are the same.
- A graph is simple if it has no parallel links or loops.
- A graph is directed if its links show direction.
- A graph is connected if at least one link exists between every pair of nodes.
- A ring network is a closed path where every node has exactly two links incident with it.

Having these terminologies, a traffic network is exemplified by a graph $G = (N,L)$ encompasses N nodes and L links, which is both directed and connected. Studies have explored the spatial dependency between traffic links in three distinct network topologies: (1) simple network, (2) grid network, and (3) ring network. The first topology is dominant in analysis, and where one of the other two topologies was analysed, the selected test subgraph collapsed the network to a simple network. Irrespective of which topology what number of links are chosen, all studies, except one (Yang, Shi, Hu, and Wang, 2015), have explored spatial dependency between traffic links for the sake of traffic forecasting in a simple graph including upstream and downstream links. We draw the schematic of networks used in the studies in Figure 5.

We are of the opinion that a large gulf exists between the realistic spatial dependence of traffic links on real networks and the typical sub-networks which have been studied in the research to date. We detected the following gaps in the literature, which signpost the way forward for further research.

(1) As alluded to previously, studies capture spatial dependency of either adjacent or distant upstream and downstream links with the study link. We hypothesise that the spatial correlation between traffic links follows a more sophisticated pattern,



a. Simple network adopted from Min et al (2009)
b. Simple network adopted from Cheng et al (2014)
c. Ring network adopted from Fabritiis et al (2008)
d. Grid network adopted from Sun et al (2005)

Figure 5. Typical network topology used in the literature.

which is not captured simply by distance rule. We now have new evidence to corroborate our hypothesis. For instance, Hu et al. (2008) revealed the first- and the eighth-order upstream links, but not other upstream links, are highly correlated with the link of interest in their specific example. A comprehensive recent paper (Yang et al., 2015) investigated the correlation between traffic links in the highway network of Twin Cities, MN. The results highlighted that the contributive links in forecasting models are widely distributed in the traffic network and are not a function of distance. This leads to this conclusion that the spatial dependency between traffic links is more complex in a whole network than what is presumed to exist in a corridor.

(2) The spatially relevant links are selected either by prejudgment or by correlation-coefficient analysis, each of which has drawbacks. In the former, researchers assume that neighbouring links are the most spatially correlated links with the study link and embed their information in forecasting methods as an input. This prejudgment increases error, if the adjacent link has not any spatial effect on the study link, as we discussed in preceding paragraph. The latter does not suffer from this shortcoming, as the input information is selected according to the most highly correlated links. However, a similar spatial effect is typically considered for all selected links, which may distort the accuracy of models.

(3) According to graph theory, two links are adjacent if they share a common node, while they are parallel if they connect the same pair of nodes. All studies, except one (Yang

et al., 2015), have developed forecasting methods in a corridor test sample, where all links are connected sequentially together. As a result, they studied the correlation of adjacent links and assume a similarity between the behaviour of both parallel and adjacent links in the hope that the reader will unquestioningly accept this assumption. We do not hold this assumption reasonable and present here the complementary and competitive nature of traffic link (Ermagun, Chatterjee, and Levinson, 2017) to shed light on the dissimilarity of spatial correlation between parallel and adjacent links. By our definition, two links are complementary, when an increase in the cost of one decreases the flow of both links. Two links are competitive, when an increase in the cost of one link decreases the flow of itself, but increases the flow of the other. We then expect a positive and a negative spatial dependency between complementary and competitive links, respectively (Ermagun and Levinson, 2018a). This nature, however, has not been captured in the literature.

This systematic review highlighted that the field is approaching its maturity, while it is still as crude as it is perplexing. It is perplexing in the conceptual methodology used, and it is crude in capturing spatial information. We further intend to bring the reader's attention to exploring the implementation of big data or smart system for spatiotemporal short-term traffic forecasting. As big data in transport coming from sensor networks, connected vehicles, and especially mobile phones continue to produce ever more measurements, these can be exploited to produce more spatially and temporally detailed short-term forecasts.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abdulhai, B., Porwal, H., & Recker, W. (1999). *Short term freeway traffic flow prediction using genetically-optimized time-delay-based neural networks*. California Partners for Advanced Transit and Highways (PATH), California PATH Working Paper, UCB-ITS-PWP-99-1, Report for MOU 360.

Abdulhai, B., Porwal, H., & Recker, W. (2002). Short-term traffic flow prediction using neuro-genetic algorithms. *ITS Journal-Intelligent Transportation Systems Journal*, *7*(1), 3–41.

Bavaud, F. (1998). Models for spatial weights: A systematic look. *Geographical Analysis*, *30*(2), 153–171.

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review*. London: Sage.

Chandra, S. R. (2009). *Spatio-temporal analyses for prediction of traffic flow, speed and occupancy on I-4* (PhD thesis). University of Central Florida Orlando, Florida.

Chandra, S., & Al-Deek, H. (2008). Cross-correlation analysis and multivariate prediction of spatial time series of freeway traffic speeds. *Transportation Research Record: Journal of the Transportation Research Board*, *2061*, 64–76.

Chandra, S. R., & Al-Deek, H. (2009). Predictions of freeway traffic speeds and volumes using vector autoregressive models. *Journal of Intelligent Transportation Systems*, *13*(2), 53–72.

Chen, L., Li, Q. R., Tian, X. Y., Chen, X. S., & Wang, R. X. (2012). Paratactic spatial-temporal two dimension data fusion based on support vector machines for traffic flow prediction of abnormal state. In S. Cai & Mingli Li (Eds.), *Advanced materials research* (Vol. 532, pp. 1225–1229). Stafa-Zuerich, Switzerland: Trans Tech Publications.

Chen, X.-Y., Pao, H.-K., & Lee, Y.-J. (2014). *Efficient traffic speed forecasting based on massive heterogenous historical data*. 2014 IEEE international conference on big data (Big Data) (pp. 10–17). Washington, DC: IEEE.

Cheng, T., Wang, J., Haworth, J., Heydecker, B., & Chow, A. (2011). *Modelling dynamic space-time autocorrelations of urban transport network*. Proceedings of the 11th international conference on geocomputation 2011, University College London, London, pp. 215–210.

Cheng, T., Wang, J., Haworth, J., Heydecker, B., & Chow, A. (2014). A dynamic spatial weight matrix and localized space-time autoregressive integrated moving average for network modeling. *Geographical Analysis*, *46*(1), 75–97.

Clark, S. D., Dougherty, M. S., & Kirby, H. R. (1993). *The use of neural networks and time series models for short term traffic forecasting: A comparative study*. Transportation planning methods. Proceedings of seminar D held at the PTRC European transport, highways and planning 21st summer annual meeting, September 13–17, UMIST (Vol. P363), University of Manchester Institute of Science and Technology, England.

Cliff, A. D., & Ord, J. K. (1975). Space-time modelling with an application to regional forecasting. *Transactions of the Institute of British Geographers*, *64*, 119–128.

Daraghmi, Y.-A., Yi, C.-W., & Chiang, T.-C. (2014). Negative binomial additive models for short-term traffic flow forecasting in urban areas. *IEEE Transactions on Intelligent Transportation Systems*, *15* (2), 784–793.

Dimitriou, L., Tsekeris, T., & Stathopoulos, A. (2008). Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow. *Transportation Research Part C: Emerging Technologies*, *16*(5), 554–573.

Djuric, N., Radosavljevic, V., Coric, V., & Vucetic, S. (2011). Travel speed forecasting by means of continuous conditional random fields. *Transportation Research Record: Journal of the Transportation Research Board*, *2263*, 131–139.

Dong, C., Shao, C., & Li, X. (2009). *Short-term traffic flow forecasting of road network based on spatial–temporal characteristics of traffic flow*. 2009 WRI world congress on computer science and information engineering (Vol. 5, pp. 645–650). Xiamen, China: IEEE.

Dong, C., Shao, C., Richards, S. H., & Han, L. D. (2014). Flow rate and time mean speed predictions for the urban freeway network using state space models. *Transportation Research Part C: Emerging Technologies*, *43*, 20–32.

Dong, C., Xiong, Z., Shao, C., & Zhang, H. (2015). A spatial–temporal-based state space approach for freeway network traffic flow modelling and prediction. *Transportmetrica A: Transport Science*, *11*(7), 547–560.

Ermagun, A., Chatterjee, S., & Levinson, D. (2017). Using temporal detrending to observe the spatial correlation of traffic. *PloS one*, *12*(5), 1–21. doi:10.1371/journal.pone.0176853

Ermagun, A., & Levinson, D. (2018a). An introduction to the network weight matrix. *Geographical Analysis*, *50*(1), 76–96. doi:10.1111/gean.2018.50.issue-1

Ermagun, A., & Levinson, D. (2018b). Development and application of the network weight matrix to predict traffic flow for congested and uncongested conditions. *Environment and Planning B: Urban Analytics and City Science*, *1*, 1–22. doi:10.1177/2399808318763368

Fusco, G., Colombaroni, C., & Isaenko, N. (2016). Comparative analysis of implicit models for real-time short-term traffic predictions. *IET Intelligent Transport Systems*, *10*(4), 270–278.

Head, L. (1995). Event—based short—term traffic flow prediction model. *Transportation Research Record*, *1510*, 125–143.

Hu, C., Xie, K., Song, G., & Wu, T. (2008). *Hybrid process neural network based on spatio-temporal similarities for short-term traffic flow prediction*. 2008 11th International IEEE conference on intelligent transportation systems (pp. 253–258). Beijing: IEEE.

Jiang, H., Zou, Y., Zhang, S., Tang, J., & Wang, Y. (2016). Short-term speed prediction using remote microwave sensor data: Machine learning versus statistical model. *Mathematical Problems in Engineering*, *2016*, 1–13.

Kamarianakis, Y., & Prastacos, P. (2003). Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches. *Transportation Research Record: Journal of the Transportation Research Board*, *1857*, 74–84.

Kamarianakis, Y., & Prastacos, P. (2005). Space-time modeling of traffic flow. *Computers & Geosciences*, *31*(2), 119–133.
Kamarianakis, Y., Prastacos, P., & Kotzinos, D. (2004). *Bivariate traffic relations: A space-time modeling approach*. AGILE proceedings, Crete, Greece, pp. 465–474.
Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, *96*(3), 118–121.
Liang, Z., & Wakahara, Y. (2014). Real-time urban traffic amount prediction models for dynamic route guidance systems. *EURASIP Journal on Wireless Communications and Networking*, *2014*(1), 1–13.
Ma, T., Zhou, Z., & Abdulhai, B. (2015). Nonlinear multivariate time–space threshold vector error correction model for short term traffic state prediction. *Transportation Research Part B: Methodological*, *76*, 27–47.
Min, X., Hu, J., Chen, Q., Zhang, T., & Zhang, Y. (2009). *Short-term traffic flow forecasting of urban network based on dynamic STARIMA model*. 2009 12th International IEEE conference on intelligent transportation systems (pp. 1–6). St. Louis, MO: IEEE.
Min, X., Hu, J., & Zhang, Z. (2010). *Urban traffic network modeling and short-term traffic flow forecasting based on GSTARIMA model*. 2010 13th International IEEE conference on intelligent transportation systems (ITSC) . Funchal, Madeira Island: IEEE. pp. 1535–1540.
Okutani, I., & Stephanedes, Y. J. (1984). Dynamic prediction of traffic volume through Kalman filtering theory. *Transportation Research Part B: Methodological*, *18*(1), 1–11.
Park, B., Messer, C., & Urbanik II, T. (1998). Short-term freeway traffic volume forecasting using radial basis function neural network. *Transportation Research Record: Journal of the Transportation Research Board*, *1651*, 39–47.
Park, D., & Rilett, L. R. (1999). Forecasting freeway link travel times with a multilayer feedforward neural network. *Computer-Aided Civil and Infrastructure Engineering*, *14*(5), 357–367.
Qing, L., Yongqin, T., Yongguo, H., & Qingming, Z. (2014). The forecast and the optimization control of the complex traffic flow based on the hybrid immune intelligent algorithm. *Open Electrical & Electronic Engineering Journal*, *8*, 245–251.
Reza, R., Pulugurtha, S. S., & Duddu, V. R. (2015). *ARIMA model for forecasting short-term travel time due to incidents in spatio-temporal context*. Transportation research board 94th annual meeting, number 15-5553, Washington, DC.
Rodrigue, J.-P., Comtois, C., & Slack, B. (2013). *The geography of transport systems* (4th ed.). New York: Routledge.
Salamanis, A., Kehagias, D. D., Filelis-Papadopoulos, C. K., Tzovaras, D., & G. A. Gravvanis (2016). Managing spatial graph dependencies in large volumes of traffic data for travel-time prediction. *IEEE Transactions on Intelligent Transportation Systems*, *17*(6), 1678–1687.
Shahsavari, B., & Abbeel, P. (2015). Short-term traffic forecasting: Modeling and learning spatio-temporal relations in transportation networks using graph neural networks. University of California at Berkeley, Technical Report No. UCB/EECS-2015-243. Retrieved from http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-243.html
Smith, B. L., Williams, B. M., & Oswald, R. K. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, *10*(4), 303–321.
Stathopoulos, A., & Karlaftis, M. G. (2003). A multivariate state space approach for urban traffic flow modeling and prediction. *Transportation Research Part C: Emerging Technologies*, *11*(2), 121–135.
Sun, S., Zhang, C., & Zhang, Y. (2005). *Traffic flow forecasting using a spatio-temporal Bayesian network predictor*. International conference on artificial neural networks (pp. 273–278). Berlin: Springer.
Van Grol, R., Inaudi, D., & Kroes, E. (2000). *On-line traffic condition forecasting using on-line measurements and a historical database*. 7th World congress on intelligent transport systems, Turin, Italy, pp. 6–9.
Van Lint, J., Hoogendoorn, S., & Van Zuylen, H. (2002). Freeway travel time prediction with state-space neural networks: Modeling state-space dynamics with recurrent neural networks. *Transportation Research Record: Journal of the Transportation Research Board*, *1811*, 30–39.
Vlahogianni, E. I., Golias, J. C., & Karlaftis, M. G. (2004). Short-term traffic forecasting: Overview of objectives and methods. *Transport Reviews*, *24*(5), 533–557.

Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach. *Transportation Research Part C: Emerging Technologies*, *13*(3), 211–234.
Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2007). Spatio-temporal short-term urban traffic volume forecasting using genetically optimized modular networks. *Computer-Aided Civil and Infrastructure Engineering*, *22*(5), 317–325.
Williams, B. (2001). Multivariate vehicular traffic flow prediction: Evaluation of ARIMAX modeling. *Transportation Research Record: Journal of the Transportation Research Board*, *1776*, 194–200.
Wu, Y.-J., Chen, F., Lu, C.-T., & Yang, S. (2016). Urban traffic flow prediction using a spatio-temporal random effects model. *Journal of Intelligent Transportation Systems*, *20*(3), 282–293.
Wu, S., Yang, Z., Zhu, X., & Yu, B. (2014). Improved k-NN for short-term traffic forecasting using temporal and spatial information. *Journal of Transportation Engineering*, *140*(7), 04014026.
Yang, S., Shi, S., Hu, X., & Wang, M. (2015). Spatiotemporal context awareness for urban traffic modeling and prediction: Sparse representation based variable selection. *PloS One*, *10*(10), e0141223.
Ye, S., He, Y., Hu, J., & Zhang, Z. (2008). *Short-term traffic flow forecasting based on mars*. Fifth international conference on fuzzy systems and knowledge discovery, 2008 (FSKD'08) (Vol. 5, pp. 669–675). Jinan, Shandong: IEEE.
Yu, B., Song, X., Guan, F., Yang, Z., & Yao, B. (2016). k-Nearest neighbor model for multiple-time-step prediction of short-term traffic condition. *Journal of Transportation Engineering*, *142*(6), 04016018.
Zeng, X., & Zhang, Y. (2013). Development of recurrent neural network considering temporal–spatial input dynamics for freeway travel time modeling. *Computer-Aided Civil and Infrastructure Engineering*, *28*(5), 359–371.
Zhang, Y., & Zhang, Y. (2016). A comparative study of three multivariate short-term freeway traffic flow forecasting methods with missing data. *Journal of Intelligent Transportation Systems*, *20*(3), 205–218.
Zhu, Z., Peng, B., Xiong, C., & Zhang, L. (2016). Short-term traffic flow prediction with linear conditional gaussian Bayesian network. *Journal of Advanced Transportation*, *50*(6), 1111–1123.
Zou, T., He, Y., Zhang, N., Du, R., & Gao, X. (2015). Short-time traffic flow forecasting based on the k-nearest neighbor model. In *ICTE 2015* (Vol. 1, pp. 1689–1695). Dalian, China.

Appendix

(1) Fusco, G., Colombaroni, C., & Isaenko, N. (2016). Comparative analysis of implicit models for real-time short-term traffic predictions. *IET Intelligent Transport Systems*, *10*(4), 270–278.
(2) Zhu, Z., Peng, B., Xiong, C., & Zhang, L. (2016). Short-term traffic flow prediction with linear conditional Gaussian Bayesian network. *Journal of Advanced Transportation*, *50*(6), 1111–1123.
(3) Yu, B., Song, X., Guan, F., Yang, Z., & Yao, B. (2016). k-Nearest neighbor model for multiple-time-step prediction of short-term traffic condition. *Journal of Transportation Engineering*, *142*(6), 04016018.
(4) Xia, D., Wang, B., Li, H., Li, Y., & Zhang, Z. (2016). A distributed spatial-temporal weighted model on MapReduce for short-term traffic flow forecasting. *Neurocomputing*, *179*, 246–263.
(5) Zhang, Y., & Zhang, Y. (2016). A comparative study of three multivariate short-term freeway traffic flow forecasting methods with missing data. *Journal of Intelligent Transportation Systems*, *20*(3), 205–218.
(6) Ko, E., Ahn, J., & Kim, E. Y. (2016). 3D Markov process for traffic flow prediction in real-time. *Sensors*, *16*(2), 147.
(7) Zhao, J., & Sun, S. (2016). High-order Gaussian process dynamical models for traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, *17*(7), 2014–2019.
(8) Jiang, H., Zou, Y., Zhang, S., Tang, J., & Wang, Y. (2016). Short-term speed prediction using remote microwave sensor data: Machine learning versus statistical model. Mathematical Problems in Engineering, 2016.

(9) Polson, N., & Sokolov, V. (2016). Deep learning predictors for traffic flows. arXiv preprint arXiv:1604.04527.
(10) Salamanis, A., Kehagias, D. D., Filelis-Papadopoulos, C. K., Tzovaras, D., & Gravvanis, G. A. (2016). Managing spatial graph dependencies in large volumes of traffic data for travel-time prediction. *IEEE Transactions on Intelligent Transportation Systems*, *17*(6), 1678–1687.
(11) Wu, Y. J., Chen, F., Lu, C. T., & Yang, S. (2016). Urban traffic flow prediction using a spatio-temporal random effects model. *Journal of Intelligent Transportation Systems*, *20*(3), 282–293.
(12) Xu, Y., Chen, H., Kong, Q. J., Zhai, X., & Liu, Y. (2016). Urban traffic flow prediction: A spatio-temporal variable selection-based approach. *Journal of Advanced Transportation*, *50*(4), 489–506.
(13) Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y. (2015). Traffic flow prediction with big data: A deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, *16*(2), 865–873.
(14) Zou, T., He, Y., Zhang, N., Du, R., & Gao, X. (2015). Short-time traffic flow forecasting based on the k-nearest neighbor model. *ICTE 2015* (pp. 1689–1695).
(15) Ma, T., Zhou, Z., & Abdulhai, B. (2015). Nonlinear multivariate time–space threshold vector error correction model for short term traffic state prediction. *Transportation Research Part B: Methodological*, *76*, 27–47.
(16) Schimbinschi, F., Nguyen, X. V., Bailey, J., Leckie, C., Vu, H., & Kotagiri, R. (2015, October). Traffic forecasting in complex urban networks: Leveraging big data and machine learning. *2015 IEEE international conference on Big data (big data)* (pp. 1019–1024). IEEE.
(17) Dong, C., Xiong, Z., Shao, C., & Zhang, H. (2015). A spatial–temporal-based state space approach for freeway network traffic flow modelling and prediction. *Transportmetrica A: Transport Science*, *11*(7), 547–560.
(18) Agafonov, A., & Myasnikov, V. (2015, April). Traffic flow forecasting algorithm based on combination of adaptive elementary predictors. *International conference on analysis of images, social networks and texts* (pp. 163–174). Cham: Springer.
(19) Fusco, G., Colombaroni, C., Comelli, L., & Isaenko, N. (2015, June). Short-term traffic predictions on large urban traffic networks: Applications of network-based machine learning models and dynamic traffic assignment models. *2015 International conference on models and technologies for intelligent transportation systems (MT-ITS)* (pp. 93–101). IEEE.
(20) Zou, Y., Hua, X., Zhang, Y., & Wang, Y. (2015). Hybrid short-term freeway speed prediction methods based on periodic analysis. *Canadian Journal of Civil Engineering*, *42*(8), 570–582.
(21) Reza, R. M., Pulugurtha, S. S., & Duddu, V. R. (2015). ARIMA model for forecasting short-term travel time due to incidents in spatio-temporal context (No. 15-5553).
(22) Hou, Y., Edara, P., & Sun, C. (2015). Traffic flow forecasting for urban work zones. *IEEE Transactions on Intelligent Transportation Systems*, *16*(4), 1761–1770.
(23) Shahsavari, B., & Abbeel, P. (2015). *Short-term traffic forecasting: Modeling and learning spatio-temporal relations in transportation networks using graph neural networks.* University of California, Berkeley.
(24) Xing, X., Zhou, X., Hong, H., Huang, W., Bian, K., & Xie, K. (2015, September). Traffic flow decomposition and prediction based on robust principal component analysis. *2015 IEEE 18th international conference on intelligent transportation systems (ITSC)* (pp. 2219–2224). IEEE.
(25) Ahn, J. Y., Ko, E., & Kim, E. (2015, August). Predicting spatiotemporal traffic flow based on support vector regression and bayesian classifier. *2015 IEEE fifth international conference on Big data and cloud computing (BDCloud)* (pp. 125–130). IEEE.
(26) Yang, S., Shi, S., Hu, X., & Wang, M. (2015). Spatiotemporal context awareness for urban traffic modeling and prediction: Sparse representation based variable selection. *PloS One*, *10*(10), e0141223.
(27) Yang, J. (2015). Spatio-temporal dynamics of short-term traffic.
(28) Dell'Acqua, P., Bellotti, F., Berta, R., & De Gloria, A. (2015). Time-aware multivariate nearest neighbor regression methods for traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, *16*(6), 3393–3402.
(29) Ran, B., Tan, H., Feng, J., Liu, Y., & Wang, W. (2015). Traffic speed data imputation method based on tensor completion. *Computational Intelligence and Neuroscience*, 2015, 22.

(30) Zhong, J. T., & Ling, S. (2015). Key factors of K-nearest neighbours nonparametric regression in short-time traffic flow forecasting. *Proceedings of the 21st international conference on industrial engineering and engineering management 2014* (pp. 9–12). Paris: Atlantis Press.

(31) Wu, S., Yang, Z., Zhu, X., & Yu, B. (2014). Improved k-NN for short-term traffic forecasting using temporal and spatial information. *Journal of Transportation Engineering*, *140*(7), 04014026.

(32) Cheng, T., Wang, J., Haworth, J., Heydecker, B., & Chow, A. (2014). A dynamic spatial weight matrix and localized space-time autoregressive integrated moving average for network modeling. *Geographical Analysis*, *46*(1), 75–97.

(33) Zhu, J. Z., Cao, J. X., & Zhu, Y. (2014). Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections. *Transportation Research Part C: Emerging Technologies*, *47*, 139–154.

(34) Chen, X. Y., Pao, H. K., & Lee, Y. J. (2014, October). Efficient traffic speed forecasting based on massive heterogenous historical data. *2014 IEEE international conference on Big Data (Big Data)* (pp. 10–17). IEEE.

(35) Niu, X., Zhu, Y., & Zhang, X. (2014, December). DeepSense: A novel learning mechanism for traffic prediction with taxi GPS traces. *2014 IEEE Global Communications Conference (GLOBECOM)* (pp. 2745–2750). IEEE.

(36) Daraghmi, Y. A., Yi, C. W., & Chiang, T. C. (2014). Negative binomial additive models for short-term traffic flow forecasting in urban areas. *IEEE Transactions on Intelligent Transportation Systems*, *15*(2), 784–793.

(37) Mohan, D. M., Asif, M. T., Mitrovic, N., Dauwels, J., & Jaillet, P. (2014, October). Wavelets on graphs with application to transportation networks. *2014 IEEE 17th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1707–1712). IEEE.

(38) Yang, J., Han, L., Freeze, P., Chin, S. M., & Hwang, H. L. (2014). Short-term freeway speed profiling based on longitudinal spatiotemporal dynamics. *Transportation Research Record: Journal of the Transportation Research Board*, *2467*, 62–72.

(39) Dong, C., Shao, C., Richards, S. H., & Han, L. D. (2014). Flow rate and time mean speed predictions for the urban freeway network using state space models. *Transportation Research Part C: Emerging Technologies*, *43*, 20–32.

(40) Ratrout, N. T. (2014). Short-term traffic flow prediction using group method data handling (GMDH)-based abductive networks. *Arabian Journal for Science and Engineering*, *39*(2), 631–646.

(41) Dong, C., Xiong, Z., Shao, C., & Zhang, H. (2015). A spatial–temporal-based state space approach for freeway network traffic flow modelling and prediction. *Transportmetrica A: Transport Science*, *11*(7), 547–560.

(42) Zhao, N., Li, Z., & Li, Y. (2014, October). Improving the traffic data imputation accuracy using temporal and spatial information. *2014 7th International conference on intelligent computation technology and automation (ICICTA)* (pp. 312–317). IEEE.

(43) Fabrizi, V., & Ragona, R. (2014). A pattern matching approach to speed forecasting of traffic networks. *European Transport Research Review*, *6*(3), 333-342.

(44) Qing, L., Yongqin, T., Yongguo, H., & Qingming, Z. (2014). The forecast and the optimization control of the complex traffic flow based on the hybrid immune intelligent algorithm. *Open Electrical & Electronic Engineering Journal*, *8*, 245–251.

(45) Liang, Z., & Wakahara, Y. (2014). Real-time urban traffic amount prediction models for dynamic route guidance systems. *EURASIP Journal on Wireless Communications and Networking*, *2014*(1), 85.

(46) Zou, Y., Zhu, X., Zhang, Y., & Zeng, X. (2014). A space-time diurnal method for short-term freeway travel time prediction. *Transportation Research Part C: Emerging Technologies*, *43*, 33–49.

(47) Haworth, J. (2014). *Spatio-temporal forecasting of network data* (Doctoral dissertation). UCL (University College London).

(48) Li, L., Li, Y., & Li, Z. (2013). Efficient missing data imputing for traffic flow by considering temporal and spatial dependence. *Transportation Research Part C: Emerging Technologies*, *34*, 108–120.

(49) Pan, T. L., Sumalee, A., Zhong, R. X., & Indra-Payoong, N. (2013). Short-term traffic state prediction based on temporal–spatial correlation. *IEEE Transactions on Intelligent Transportation Systems*, *14*(3), 1242–1254.

(50) Zeng, X., & Zhang, Y. (2013). Development of recurrent neural network considering temporal-spatial input dynamics for freeway travel time modeling. *Computer-Aided Civil and Infrastructure Engineering*, *28*(5), 359–371.
(51) Fowe, A. J., & Chan, Y. (2013). A microstate spatial-inference model for network-traffic estimation. *Transportation Research Part C: Emerging Technologies*, *36*, 245–260.
(52) Han, Y., & Moutarde, F. (2013). Statistical traffic state analysis in large-scale transportation networks using locality-preserving non-negative matrix factorisation. *IET Intelligent Transport Systems*, *7*(3), 283–295.
(53) Huang, X., & Wang, W. (2013). Traffic prediction based on correlation of road sections. *Indonesian Journal of Electrical Engineering and Computer Science*, *11*(10), 5523–5529.
(54) Zheng, F., & Van Zuylen, H. (2013). Urban link travel time estimation based on sparse probe vehicle data. *Transportation Research Part C: Emerging Technologies*, *31*, 145–157.
(55) Kamarianakis, Y., Shen, W., & Wynter, L. (2012). Real-time road traffic forecasting using regime-switching space-time models and adaptive LASSO. *Applied Stochastic Models in Business and Industry*, *28*(4), 297–315.
(56) Haworth, J., & Cheng, T. (2012). Non-parametric regression for space-time forecasting under missing data. *Computers, Environment and Urban Systems*, *36*(6), 538–550.
(57) Cheng, T., Haworth, J., & Wang, J. (2012). Spatio-temporal autocorrelation of road network data. *Journal of Geographical Systems*, *14*(4), 389–413.
(58) Guo, F., Krishnan, R., & Polak, J. W. (2012). Short-term traffic prediction under normal and incident conditions using singular spectrum analysis and the k-nearest neighbour method.
(59) Guo, F., Krishnan, R., & Polak, J. W. (2012). Short-term traffic prediction under normal and abnormal traffic conditions on urban roads. *Transportation Research Board 91st annual meeting* (No. 12-1627).
(60) Pan, T. (2012). *The stochastic dynamic journey time reliability analysis by considering the spatial and temporal correlation* (Doctoral dissertation). The Hong Kong Polytechnic University.
(61) Ngan, H. Y., Yung, N. H., & Yeh, A. G. (2012, August). Modeling of traffic data characteristics by Dirichlet process mixtures. *2012 IEEE international conference on automation science and engineering (CASE)* (pp. 224–229). IEEE.
(62) Chen, L., Li, Q. R., Tian, X. Y., Chen, X. S., & Wang, R. X. (2012). Paratactic spatial-temporal two dimension data fusion based on support vector machines for traffic flow prediction of abnormal state. *Advanced materials research* (Vol. 532, pp. 1225–1229). Trans Tech Publications.
(63) Wu, Y. J., Chen, F., Lu, C., Smith, B., & Chen, Y. (2012). Traffic flow prediction for urban network using spatio-temporal random effects model. *91st Annual meeting of the Transportation Research Board (TRB)*.
(64) Yuan, Z. W., & Wang, Y. H. (2012, August). Research on K nearest neighbor non-parametric regression algorithm based on KD-tree and clustering analysis. *2012 Fourth international conference on computational and information sciences (ICCIS)* (pp. 298–301). IEEE.
(65) Sun, S., Huang, R., & Gao, Y. (2012). Network-scale traffic modeling and forecasting with graphical lasso and neural networks. *Journal of Transportation Engineering*, *138*(11), 1358–1367.
(66) Pascale, A., & Nicoli, M. (2011, June). Adaptive Bayesian network for traffic flow prediction. *Statistical signal processing workshop (SSP), 2011 IEEE* (pp. 177–180). IEEE.
(67) Djuric, N., Radosavljevic, V., Coric, V., & Vucetic, S. (2011). Travel speed forecasting by means of continuous conditional random fields. *Transportation Research Record: Journal of the Transportation Research Board*, *2263*, 131–139.
(68) Han, Y., & Moutarde, F. (2011, October). Analysis of network-level traffic states using locality preservative non-negative matrix factorization. *2011 14th International IEEE conference on intelligent transportation systems (ITSC)* (pp. 501–506). IEEE.
(69) Samaranayake, S., Blandin, S., & Bayen, A. (2011, December). Learning the dependency structure of highway networks for traffic forecast. *2011 50th IEEE conference on decision and control and european control conference (CDC-ECC)* (pp. 5983–5988). IEEE.
(70) Cheng, T., Wang, J., Haworth, J., Heydecker, B. J., & Chow, A. H. F. (2011). Modelling dynamic space-time autocorrelations of urban transport network. *Proceedings of the 11th international conference on geocomputation 2011* (pp. 215–210).

(71) Deng, R., & Jiang, L. (2011). Traffic state forecast of road network based on spatial-temporal data mining. *ICTE 2011* (pp. 734–739).

(72) Min, W., & Wynter, L. (2011). Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies*, *19*(4), 606–616.

(73) Khosravi, A., Mazloumi, E., Nahavandi, S., Creighton, D., & Van Lint, J. W. C. (2011). Prediction intervals to account for uncertainties in travel time prediction. *IEEE Transactions on Intelligent Transportation Systems*, *12*(2), 537–547.

(74) Lippi, M., Bertini, M., & Frasconi, P. (2010, September). Collective traffic forecasting. *Joint European conference on machine learning and knowledge discovery in databases* (pp. 259–273). Berlin, Heidelberg: Springer.

(75) Min, X., Hu, J., & Zhang, Z. (2010, September). Urban traffic network modeling and short-term traffic flow forecasting based on GSTARIMA model. *2010 13th International IEEE conference on intelligent transportation systems (ITSC)* (pp. 1535–1540). IEEE.

(76) Herring, R., Hofleitner, A., Abbeel, P., & Bayen, A. (2010, September). Estimating arterial traffic conditions using sparse probe data. *2010 13th International IEEE conference on intelligent transportation systems (ITSC)* (pp. 929–936). IEEE.

(77) Sun, X., Jia, L., Dong, H., Qin, Y., & Guo, M. (2010). Urban expressway traffic state forecasting based on multimode maximum entropy model. *Science China Technological Sciences*, *53*(10), 2808–2816.

(78) Guo, F., Polak, J. W., & Krishnan, R. (2010, September). Comparison of modelling approaches for short term traffic prediction under normal and abnormal conditions. *2010 13th International IEEE conference on intelligent transportation systems (ITSC)* (pp. 1209–1214). IEEE.

(79) Lee, M. T. (2010). Short-term freeway traffic flow forecasting with ARIMAX modeling.

(80) McCrea, J., & Moutari, S. (2010). A hybrid macroscopic-based model for traffic flow in road networks. *European Journal of Operational Research*, *207*(2), 676–684.

(81) Min, X., Hu, J., Chen, Q., Zhang, T., & Zhang, Y. (2009, October). Short-term traffic flow forecasting of urban network based on dynamic STARIMA model. *12th International IEEE conference on intelligent transportation systems, 2009, ITSC'09* (pp. 1–6). IEEE.

(82) Chandra, S. R., & Al-Deek, H. (2009). Predictions of freeway traffic speeds and volumes using vector autoregressive models. *Journal of Intelligent Transportation Systems*, *13*(2), 53–72.

(83) Dong, C., Shao, C., & Li, X. (2009, March). Short-term traffic flow forecasting of road network based on spatial-temporal characteristics of traffic flow. *2009 WRI world congress on computer science and information engineering* (Vol. 5, pp. 645–650). IEEE.

(84) Li, R., & Lu, H. (2009, May). Combined neural network approach for short-term urban freeway traffic flow prediction. *International symposium on neural networks* (pp. 1017–1025). Berlin, Heidelberg: Springer.

(85) Venkata, C., & Chandra, S. R. (2009). Spatio-temporal analyses for prediction of traffic flow, speed and occupancy on I-4.

(86) Ghosh, B., Basu, B., & O'Mahony, M. (2009). Multivariate short-term traffic flow forecasting using time-series analysis. *IEEE Transactions on Intelligent Transportation Systems*, *10*(2), 246–254.

(87) Van Hinsbergen, C. P. I. J., Van Lint, J., & Van Zuylen, H. (2009). Bayesian training and committees of state-space neural networks for online travel time prediction. *Transportation Research Record: Journal of the Transportation Research Board*, *2105*, 118–126.

(88) Hu, J., Kaparias, I., & Bell, M. G. (2009). Spatial econometrics models for congestion prediction with in-vehicle route guidance. *IET Intelligent Transport Systems*, *3*(2), 159–167.

(89) Innamaa, S. (2009). Self-adapting traffic flow status forecasts using clustering. *IET Intelligent Transport Systems*, *3*(1), 67–76.

(90) Yue, Y., & Yeh, A. G. O. (2008). Spatiotemporal traffic-flow dependency and short-term traffic forecasting. *Environment and Planning B: Planning and Design*, *35*(5), 762–771.

(91) Stathopoulos, A., Dimitriou, L., & Tsekeris, T. (2008). Fuzzy modeling approach for combined forecasting of urban traffic flow. *Computer-Aided Civil and Infrastructure Engineering*, *23*(7), 521–535.

(92) Chandra, S., & Al-Deek, H. (2008). Cross-correlation analysis and multivariate prediction of spatial time series of freeway traffic speeds. *Transportation Research Record: Journal of the Transportation Research Board*, *2061*, 64–76.
(93) Dimitriou, L., Tsekeris, T., & Stathopoulos, A. (2008). Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow. *Transportation Research Part C: Emerging Technologies*, *16*(5), 554–573.
(94) De Fabritiis, C., Ragona, R., & Valenti, G. (2008, October). Traffic estimation and prediction based on real time floating car data. *11th International IEEE conference on intelligent transportation systems, 2008. ITSC 2008* (pp. 197–203). IEEE.
(95) Wu, T., Xie, K., Song, G., & Hu, C. (2008, October). A multiple svr approach with time lags for traffic flow prediction. *11th International IEEE conference on intelligent transportation systems, 2008. ITSC 2008* (pp. 228–233). IEEE.
(96) Hu, C., Xie, K., Song, G., & Wu, T. (2008, October). Hybrid process neural network based on spatio-temporal similarities for short-term traffic flow prediction. *11th International IEEE conference on intelligent transportation systems, 2008. ITSC 2008* (pp. 253–258). IEEE.
(97) Van Lint, J. W. C. (2008). Online learning solutions for freeway travel time prediction. *IEEE Transactions on Intelligent Transportation Systems*, *9*(1), 38–47.
(98) Ye, S., He, Y., Hu, J., & Zhang, Z. (2008, October). Short-term traffic flow forecasting based on mars. *Fifth international conference on fuzzy systems and knowledge discovery, 2008. FSKD'08* (Vol. 5, pp. 669–675). IEEE.
(99) Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2007). Spatio-temporal short-term urban traffic volume forecasting using genetically optimized modular networks. *Computer-Aided Civil and Infrastructure Engineering*, *22*(5), 317–325.
(100) Yue, Y., Yeh, A. G., & Zhuang, Y. (2007, September). Prediction time horizon and effectiveness of real-time data on short-term traffic predictability. *Intelligent transportation systems conference, 2007 (ITSC 2007)*. IEEE (pp. 962–967). IEEE.
(101) Sun, S., & Zhang, C. (2007). The selective random subspace predictor for traffic flow forecasting. *IEEE Transactions on Intelligent Transportation Systems*, *8*(2), 367–373.
(102) Xie, Y., & Zhang, Y. (2006). A wavelet network model for short-term traffic volume forecasting. *Journal of Intelligent Transportation Systems*, *10*(3), 141–150.
(103) Van Lint, J. W. (2006). Reliable real-time framework for short-term freeway travel time prediction. *Journal of Transportation Engineering*, *132*(12), 921–932.
(104) Wang, Y., Papageorgiou, M., & Messmer, A. (2006). RENAISSANCE–A unified macroscopic model-based approach to real-time freeway network traffic surveillance. *Transportation Research Part C: Emerging Technologies*, *14*(3), 190–212.
(105) Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach. *Transportation Research Part C: Emerging Technologies*, *13*(3), 211–234.
(106) Sun, S., Zhang, C., & Zhang, Y. (2005, September). Traffic flow forecasting using a spatio-temporal Bayesian network predictor. *International conference on artificial neural networks* (pp. 273–278). Berlin, Heidelberg: Springer.
(107) Kamarianakis, Y., & Prastacos, P. (2005). Space-time modeling of traffic flow. *Computers & Geosciences*, *31*(2), 119–133.
(108) Van Lint, J. W. (2006). Reliable real-time framework for short-term freeway travel time prediction. *Journal of Transportation Engineering*, *132*(12), 921–932.
(109) Bajwa, S. I., Chung, E., & Kuwahara, M. (2005, September). Performance evaluation of an adaptive travel time prediction model. *2005 IEEE proceedings of intelligent transportation systems* (pp. 1000-1005). IEEE.
(110) Innamaa, S. (2005). Short-term prediction of travel time using neural networks on an interurban highway. *Transportation*, *32*(6), 649–669.
(111) Ishak, S., & Alecsandru, C. (2004). Optimizing traffic prediction performance of neural networks under various topological, input, and traffic condition settings. *Journal of Transportation Engineering*, *130*(4), 452–465.

(112) Kamarianakis, Y., Prastacos, P., & Kotzinos, D. (2004). Bivariate traffic relations: A space-time modeling approach. *AGILE Proceedings* (pp. 465–474).
(113) Alecsandru, C. D. (2003). A hybrid model-based and memory-based short-term traffic prediction system .
(114) Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2003, August). A multivariate neural network predictor for short term traffic prediction in urban signalized arterial. *10th IFAC symposium on control in transportation systems*, Tokyo, Japan.
(115) Kamarianakis, Y., & Prastacos, P. (2003). Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches. *Transportation Research Record: Journal of the Transportation Research Board*, *1857*, 74–84.
(116) Stathopoulos, A., & Karlaftis, M. G. (2003). A multivariate state space approach for urban traffic flow modeling and prediction. *Transportation Research Part C: Emerging Technologies*, *11*(2), 121–135.
(117) Ishak, S., Kotha, P., & Alecsandru, C. (2003). Optimization of dynamic neural network performance for short-term traffic prediction. *Transportation Research Record: Journal of the Transportation Research Board*, *1836*, 45–56.
(118) Hu, J., Song, J., Yu, G., & Zhang, Y. (2003, October). A novel networked traffic parameter forecasting method based on Markov chain model. *IEEE international conference on systems, man and cybernetics, 2003* (Vol. 4, pp. 3595–3600). IEEE.
(119) Van Lint, J. W. C., Hoogendoorn, S. P., & van Zuylen, H. J. (2003). Toward a robust framework for freeway travel time prediction: Experiments with simple imputation and state-space neural networks. *TRB 82nd annual meeting (CD-ROM)* (Vol. 6).
(120) Van Lint, J., Hoogendoorn, S., & Van Zuylen, H. (2002). Freeway travel time prediction with state-space neural networks: Modeling state-space dynamics with recurrent neural networks. *Transportation Research Record: Journal of the Transportation Research Board*, *1811*, 30–39.
(121) Tebaldi, C., West, M., & Karr, A. F. (2002). Statistical analyses of freeway traffic flows. *Journal of Forecasting*, *21*(1), 39–68.
(122) Abdulhai, B., Porwal, H., & Recker, W. (2002). Short-term traffic flow prediction using neuro-genetic algorithms. *ITS Journal-Intelligent Transportation Systems Journal*, *7*(1), 3–41.
(123) Williams, B. (2001). Multivariate vehicular traffic flow prediction: Evaluation of ARIMAX modeling. *Transportation Research Record: Journal of the Transportation Research Board*, *1776*, 194–200.
(124) Van Grol, R., Inaudi, D., & Kroes, E. (2000, November). On-line traffic condition forecasting using on-line measurements and a historical database. *7th World congress on intelligent transport systems* (pp. 6–9).
(125) Park, D., & Rilett, L. R. (1999). Forecasting freeway link travel times with a multilayer feedforward neural network. *Computer-Aided Civil and Infrastructure Engineering*, *14*(5), 357–367.
(126) Abdulhai, B., Porwal, H., & Recker, W. (1999). Short term freeway traffic flow prediction using genetically-optimized time-delay-based neural networks .
(127) Park, B., Messer, C., & Urbanik II, T. (1998). Short-term freeway traffic volume forecasting using radial basis function neural network. *Transportation Research Record: Journal of the Transportation Research Board*, *1651*, 39–47.
(128) Larry, H. K. (1995). Event—based short—term traffic flow prediction model. *Transportation Research Record*, *1510*, 125–143.
(129) Clark, S. D., Dougherty, M. S., & Kirby, H. R. (1993). The use of neural networks and time series models for short term traffic forecasting: A comparative study. Transportation planning methods. *Proceedings of seminar held at the PTRC European transport, highways and planning 21st summer annual meeting* (September 13–17, 1993), Umist. Volume p363.
(130) Okutani, I., & Stephanedes, Y. J. (1984). Dynamic prediction of traffic volume through Kalman filtering theory. *Transportation Research Part B: Methodological*, *18*(1), 1–11.