Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 198 (2022) 43–50

Procedia Computer Science

www.elsevier.com/locate/procedia

The 12th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2021)
November 1-4, 2021, Leuven, Belgium

Modeling of road traffic flows in the neighboring regions

Henrik Fredriksson[a,*], Mattias Dahl[a], Benny Lövström[a], Johan Holmgren[b], Håkan Lennerstad[a]

[a]Department of Mathematics and Natural Sciences, Blekinge Institute of Technology, Karlskrona, 37179 Sweden
[b]Department of Computer Science and Media Technology, Malmö University, Malmö 20506, Sweden

Abstract

Traffic flows play a very important role in transportation engineering. In particular, link flows are a source of information about the traffic state, which is usually available from the authorities that manage road networks. Link flows are commonly used in both short-term and long-term planning models for operation and maintenance, and to forecast the future needs of transportation infrastructure. In this paper, we propose a model to study how traffic flow in one location can be expected to reflect the traffic flow in a nearby region. The statistical basis of the model is derived from link flows to find estimates of the distribution of traffic flows in junctions. The model is evaluated in a numerical study, which uses real link flow data from a transportation network in southern Sweden. The results indicate that the model may be useful for studying how large departing flows from a node reflect the link flows in a neighboring geographic region.

Keywords: link flows; traffic volumes; flow distribution; flow estimation; transportation network

1. Introduction

Road traffic flows are a key component in the analysis of the movement of people and goods, both in urban areas and on a regional level. The generation of road traffic flows stems mainly from individuals' decisions to make a trip and which routes they decide to use. The number of vehicles traveling on each route over the course of a day will induce flow on each of the individual links in the network. In practice, it is more straightforward and more cost-effective to obtain link flow data than route flow data. Link flow data can be obtained by traditional sensor technologies such as pneumatic tubes and inductive loops, which are placed in the network [1]. To be able to collect route flow data, information about route choices is required, and it may be difficult to obtain due to data protection and privacy [2, 3].

* Corresponding author. Tel.: +46-455-385548
E-mail address: henrik.fredriksson@bth.se

10.1016/j.procs.2021.12.209

However, link flows have low informative value on the travel patterns and travel demands (e.g., route choices and the origin–destination of the vehicles). Therefore, a large body of the literature is devoted to the reconstruction or estimation of route flows or origin–destination flows from link flows, for instance, using license plate recognition technologies or by making assumptions about travelers' route choices [4, 5]. However, link flows can be used to study the distribution of vehicles in the network from a broader perspective, including how locations in the network are connected in terms of flows.

The connection between locations and the distribution of aggregate traffic flows play an important role in the study of travel demand and the prediction of traffic flows in the transportation network. For instance, the connection between locations in terms of flows may be used to detect links, which in case of accidents or any disruptive event may affect traffic flows negatively [6, 7]. The study of the distribution of traffic flows also plays an important role in planning the maintenance of roads, studying how congestion spreads, and maintaining road traffic fluidity in general [8, 9, 10].

In order to analyze the movement of vehicles, it is necessary to know how the aggregate vehicle flows disperse in junctions such as intersections and roundabouts, or which on-ramps or off-ramps vehicles are using when they enter and leave freeways. Knowledge about which links vehicles travel on along their journey is essential information in transportation planning and engineering and is used as input data for traffic modeling, analysis, and forecasting [11]. The travel demand through a node is generally estimated using historical data about travel demand along with traffic counts [12, 13]. From a broader perspective, the distribution of traffic flow in junctions provides a statistical basis for studying how the traffic flows propagate in a larger geographic region, rather than in a single isolated intersection or roundabout [14, 15].

The distribution of traffic flows is also determined by which routes travelers use. In reality, it is reasonable to assume that the majority of travelers select a route that will allow them to reach their destination as efficiently as possible. Travel time is one of the most important criteria for selecting a certain route; however, several studies show that the preferred routes may not be optimal with respect to the minimization of travel time [16]. Thus, in order to study how aggregated traffic flows propagate in the network and how locations are connected in terms of flows, it is relevant to consider the possible ways vehicles can travel between locations.

In the current paper, we present a model to investigate how link flows can be used to study how the traffic flows in one location reflect the traffic flows in a nearby region. Such a study may be helpful in examining how sudden changes in traffic volumes can be expected to affect the traffic flow in neighboring links or links in the surrounding area. The traffic scenario considered in this paper concerns traffic flows without congestion, with the aim of studying the stochastic behavior of aggregated traffic flows. Thus, we assume that route choices are independent of traffic flow volumes.

The remainder of the paper is organized as follows. In Section 2 we describe the proposed model. To evaluate the model, we present a numerical study using real link flow data from a region in the southern part of Sweden in Section 3. Discussions, conclusions and future research are presented in Section 4 and Section 5.

2. The proposed model

In this section we describe the proposed model and establish notations and terminology used in the remainder of this paper. A transportation network is represented by a directed graph $G = (N, A)$, where N is the set of nodes and A is the set of links. For each link $a \in A$, we let $v_a(t)$ denote the hourly flow on that link during time period $[t, t+1]$, where $t = 0, 1, \ldots, 23$. The set of departing links from the node n is denoted by $\delta^{\text{out}}(n)$, i.e.,

$$\delta^{\text{out}}(n) = \{a \in A \mid \exists x \in N : a = (n, x)\}. \tag{1}$$

To study how the traffic flow in one location reflects the flow in a location in the surrounding area, the idea is to study the routes connecting the two locations and the traffic flow that may occur on these routes. Here, we emphasize that by a route, we mean any sequence of unique links connecting two nodes in the network. The link flows are used as the statistical basis to determine the probability that a particular route between the locations is used. We assume that the traffic flows on the routes are Poisson distributed random variables. Let π denote the probability that a vehicle selects a particular route and suppose that there are m candidate vehicles, which may travel on the route. The route probabilities and the number of candidate vehicles may have sources of variability [2]. The measurements of link flows may have errors, and the travel demand is expected to vary from day to day. Therefore, we assume that the number of vehicles that actually select a particular route is a random variable X from the distribution $\text{Bin}(m, \pi)$ with expected value $E[X] = m\pi$. Suppose there are k different routes between two fixed locations and let X_i be the random variable of vehicles that select alternative i with the corresponding route probability $\pi^{(i)}$, where $i = 1, \ldots, k$. As vehicles cannot travel on multiple routes simultaneously, the X_i's are independent random variables. Let $X = X_1 + X_2 + \ldots + X_k$ be the number of vehicles which selects one of the route alternatives $1, \ldots, k$. It follows that the number of vehicles X that actually travel between the two locations is a random variable from the distribution $\text{Bin}(m, \pi^{(1)} + \ldots + \pi^{(k)})$, with expected value $E[X] = m(\pi^{(1)} + \ldots + \pi^{(k)})$, where m is the number of candidate vehicles that may travel on any of the routes.

As the link flows are the only parameters of the model, we determine the route probabilities by the product of transition probabilities from node to node along the routes. For a route $p^{(i)} = (1, 2, \ldots, N)$, we compute the route probability by the formula

$$\pi^{(i)} = \prod_{1 \le j \le N-1} \pi^{(i)}_{j,j+1}, \tag{2}$$

where $\pi^{(i)}_{j,j+1}$ is the probability of moving from node j to the next node $j + 1$ along route alternative i. The superscript (i) indicates the turning movement probability with respect to route alternative i. For instance, the turning movement probabilities may take into account that vehicles do not return to an already visited node.

3. Numerical study

In this section, we present a numerical study to evaluate the proposed model using real traffic data from a region in southern Sweden.

3.1. Scenario description

The network in our numerical study consists of 30 nodes and 92 links, and it is depicted on top of the traffic flow map in Fig. 1. The selected nodes captures the major traffic flows in the region, which explains why some nodes are located near one other. As the nodes are located near cities and other communities, the node locations mainly represent highway junctions or on-ramps and off-ramps for vehicles to enter and leave freeways.

The link flow data used in the numerical study consist of hourly link flows of trucks over a 24-hour period. Using real measurements calls for calibration to ensure that the inflow is equal to the outflow for each node and each hour. By introducing virtual links of the form (n, n) at each node n acting as ”parking“, it is always possible to find flows on the virtual links such that the transportation system is balanced for each node and for each time period, even though the flow on the links varies throughout the day.

The purpose of the numerical study is to evaluate how the departing flow from a node reflect the flow on some link in a region near the node. Therefore, we selected a set of ”origin nodes“ and a ”destination link“ for each one of the origin nodes. Our conjecture is that the model is appropriate for studying how large flows propagate, we mainly selected nodes with large departing flows or where we hypothesized that the departing flow reflect the flow on some link in the nearby region. The selection of a destination link for each origin node was based on the underlying traffic flow map. We selected a destination link that was not directly connected to the origin node, and we selected for travel times of less than one hour between the origin node and the destination link. For each origin node, we identified and selected one or more routes between the node and the destination link, and for each time period we computed the route probabilities according to (2), where we assumed that no vehicle returned to an already visited node. For a route

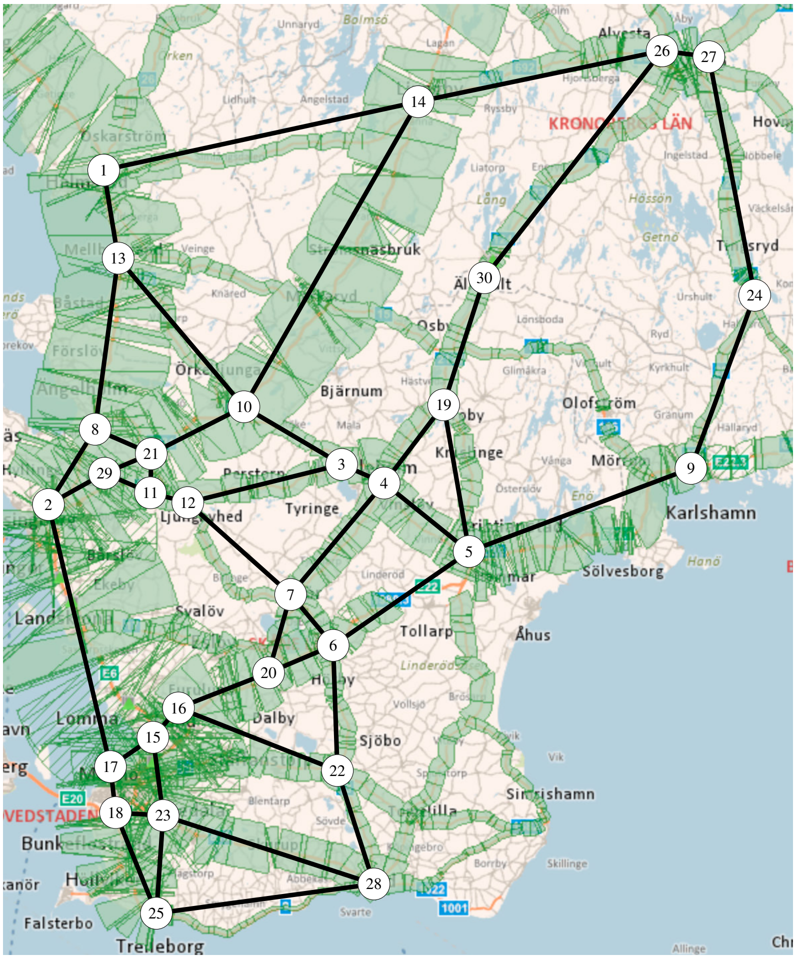

Fig. 1. Graph representation of the considered network showing nodes and links on top of the traffic flow map.

$p^{(i)} = (1, 2, \ldots, N)$, we used the formula

$$\pi^{(i)}_{j,j+1} = \frac{v_a}{\sum_{a' \in \delta^{\text{out}}(n) - \{b\}} v_{a'}}, \quad 1 \le j \le N - 1, \tag{3}$$

where $a = (j, j+1)$ and where $b = \emptyset$ if $j = 1$, and $b = (j, j-1)$ otherwise, to determine the transition probabilities in the nodes.

The number of candidate vehicles m for each origin node n and hour is determined by

$$m(t) = \sum_{a \in \delta^{\text{out}}(n)} v_a(t), \quad t = 0, 1, \ldots, 23, \tag{4}$$

which is the total number of departing vehicles from node n during time period $[t, t+1]$.

3.2. Results

To evaluate how the departing vehicle flows from the origin nodes reflect the flow on the destination links, we first computed the difference between the expected link flows and the measured link flow by the root mean square error (RMSE) and the mean absolute error (MAE) for each hour. We emphasize that the aim is not to reproduce the true link flows but to study the impact of the departing flows from the origin node on the destination link. For the expected link flows $\widetilde{v_a}$ and the true link flows v_a, the RMSE is computed by

$$\mathrm{RMSE} = \sqrt{\frac{1}{24}\sum_{t=0}^{23}(v_a(t) - \widetilde{v_a}(t))^2}, \tag{5}$$

and the MAE is computed by

$$\mathrm{MAE} = \frac{1}{24}\sum_{t=0}^{23}|v_a(t) - \widetilde{v_a}(t)|\,. \tag{6}$$

We also computed Pearson's correlation coefficient (PCC) between the resulting link flow time series and the true link flow time series. Table 1 presents the RMSE, MAE and PCC and which routes we included in the computations, for each pairing of a selected origin node and destination link. The number of truck departures per hour for the origin nodes 2, 5 and 17 are in the ranges $[72, 702]$, $[20, 321]$ and $[84, 754]$, with mean values 418, 173 and 401 of departures per hour, respectively. Fig. 2–7 illustrate the expected link flow and the measured link flow for each hour of the day. In total, the numerical study includes $24 \times 6 = 144$ observations, i.e., one observation per hour on each destination link. For each observation, we computed the relative error η according to

$$\eta = \frac{|v_a(t) - \widetilde{v_a}(t)|}{v_a(t)}, \quad t = 0, 1 \ldots, 23. \tag{7}$$

For 82 of the 144 observations, the relative error is less than 25%, and for 46 of the 144 observations, the relative error is less than 15%.

Table 1. Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Pearson's correlation coefficient (PCC) between the expected link flow and the measured link flow over a 24-hour period. The selection of origin nodes was mainly based on the total number of departing vehicles throughout the day.

Origin node	Destination link	RMSE	MAE	PCC	Route(s)
2	(29, 11)	9.22	6.55	0.94	2–29–11
2	(29, 21)	22.86	17.91	0.85	2–29–21
5	(19, 30)	10.26	7.67	0.89	5–19–30, 5–4–19–30
17	(16, 20)	24.11	14.30	0.90	17–15–16–20
17	(16, 22)	10.20	6.02	0.91	17–15–16–22
17	(23, 28)	7.21	5.54	0.94	17–15–23–28, 17–18–23–28

4. Discussion

Based on the results presented in the previous section, we find it reasonable to assume that the model can be used to estimate how the departing flow from an origin node is reflected in the flow on its destination links. Thus, we

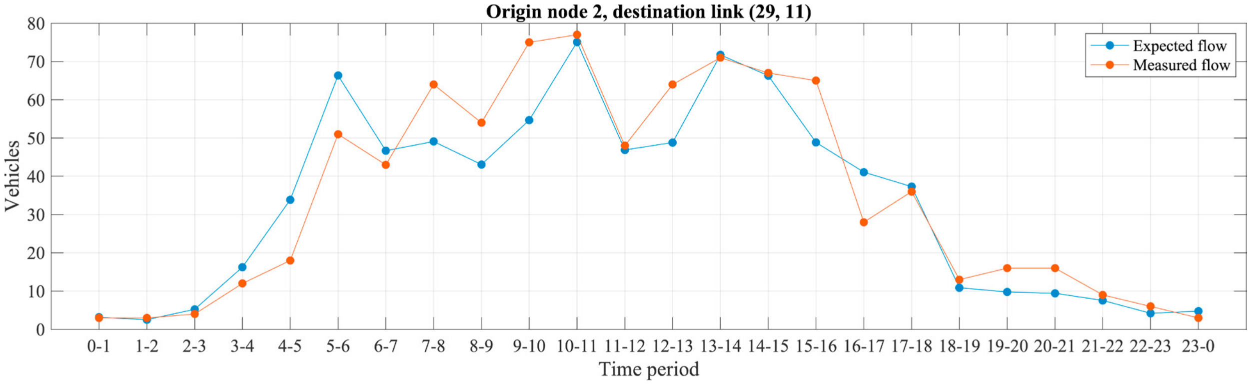

Fig. 2. Expected flow on link (29, 11) from origin node 2 and the measured flow over a 24-hour period.

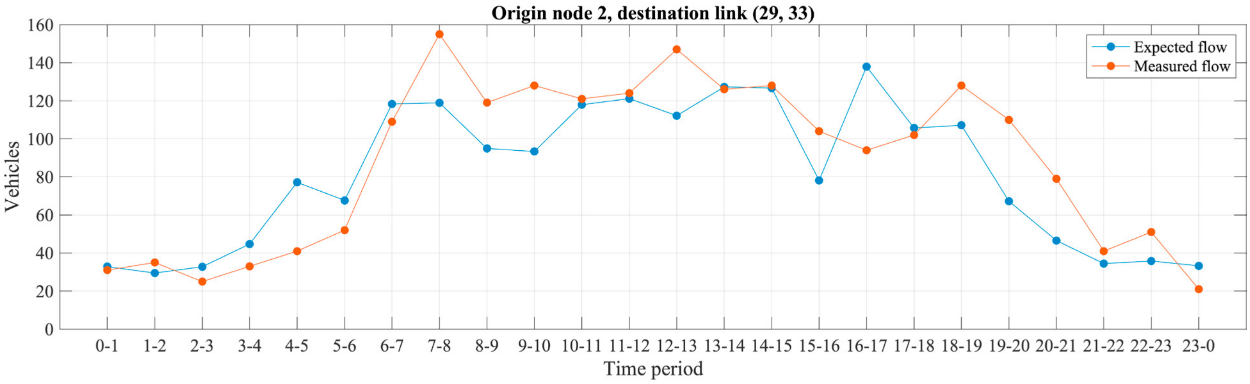

Fig. 3. Expected flow on link (29, 33) from origin node 2 and the measured flow over a 24-hour period.

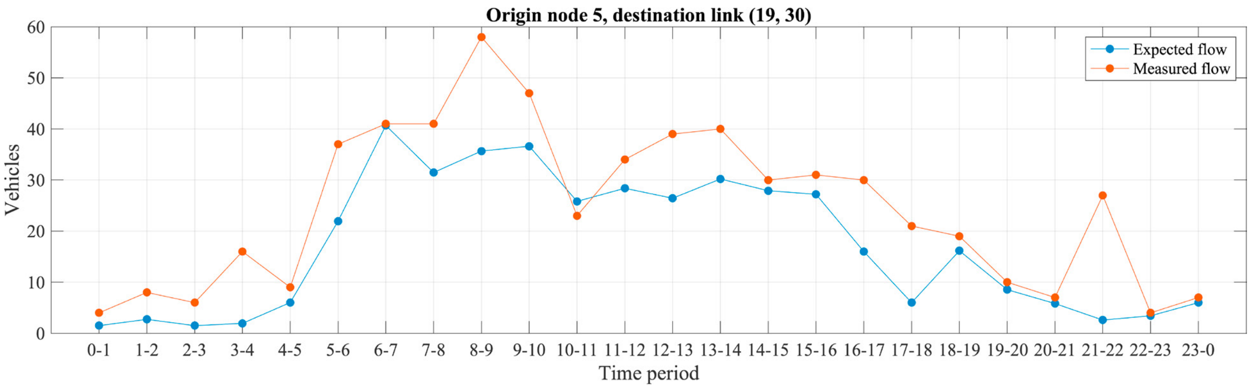

Fig. 4. Expected flow on link (19, 30) from origin node 5 and the measured flow over a 24-hour period.

conjecture that if there are m vehicles that depart from an origin node, the number of vehicles that will travel on the destination link can be estimated by a random variable X from distribution $\mathrm{Bin}(m, \pi)$, where π is the sum of route probabilities of the routes connecting the origin node and the destination link. The selected nodes include the major roads in the network, where congestion rarely occurs. Thus, we claim that the route probabilities in our study depend on the time period and not on the traffic volumes, which means that the expected flow on the destination link can be computed by $m\pi$ for arbitrary number of candidate vehicles m. As seen in Fig. 2-7, there are time periods when the expected flow differs from the measured flow. For hours when the expected link flow is higher than the measured link flow, we conjecture that some share of the vehicles may have reached their destination before reaching the selected

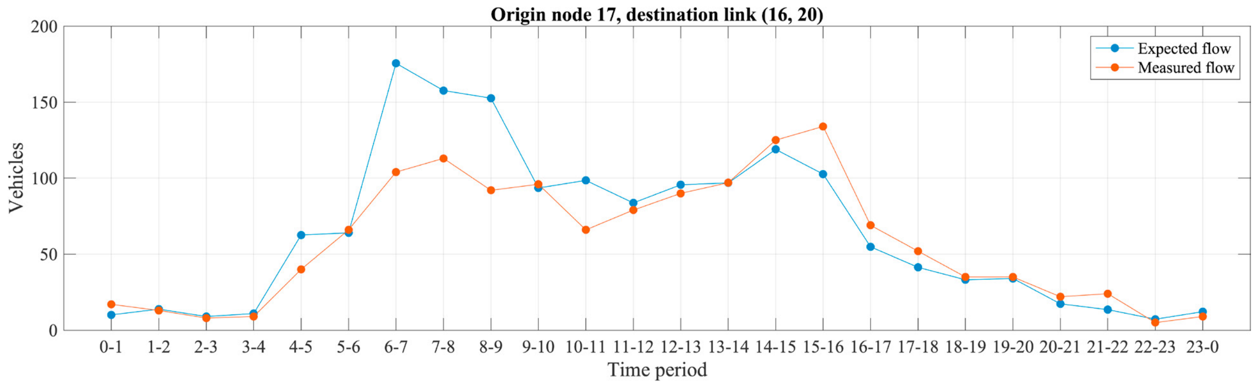

Fig. 5. Expected flow on link (16, 20) from origin node 17 and the measured flow over a 24-hour period.

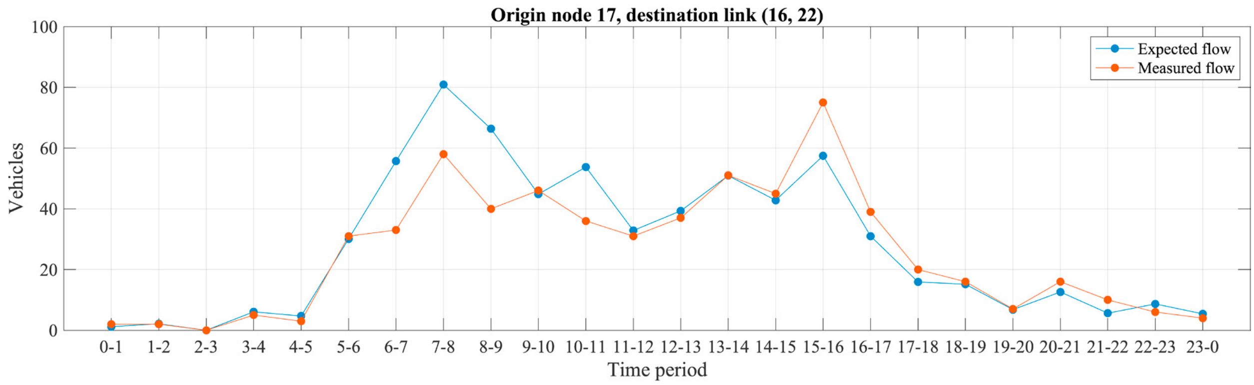

Fig. 6. Expected flow on link (16, 22) from origin node 17 and the measured flow over a 24-hour period.

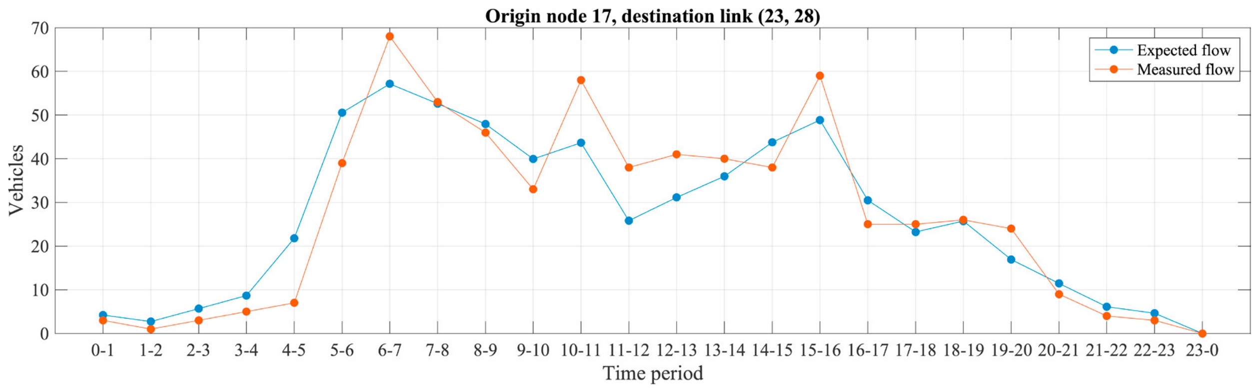

Fig. 7. Expected flow on link (23, 28) from origin node 17 and the measured flow over a 24-hour period.

destination link. Conversely, if the expected link flow is lower than the measured link flow during some hour of the day, vehicles may originate from some other neighboring node during that hour.

Since the main component of the statistical basis of the model is the link flows, which inherently do not contain any information about the travel behavior (such as travelers' route choices), the accuracy of the proposed model may vary. If the alternative ways to travel between two locations are parts of routes that are actually used, then it is expected that the difference between the expected flow and the measured flow will be small.

A potential limitation of using link flows to determine the route probabilities is that there are situations in which the model may produce less accurate results. For instance, if a route carrying a significantly larger flow crosses the route from the origin node to the destination, a link-flow–based turning movement probability may ”redirect“ a large portion of the vehicles in the wrong direction. This is why we chose to study nodes with large outflows. However, it should be emphasized that knowledge about the actual routes is hard to derive from the link flows unless additional information is provided. Nonetheless, there are a few scenarios in which the model may be particularly applicable, and in which it suffices to use link flows to determine route probabilities [17]. Large flows may travel through a particular node, and some links may carry significantly larger traffic volumes and characterize the traffic flows in the surrounding area. Another scenario in which the model is applicable is when the studied network has a tree structure, e.g., when studying the traffic flows in a bidirected acyclic graph.

5. Conclusions and Future work

In this paper, we have presented a model that aims to study how traffic flows in one location reflect the traffic flows in a nearby region. The proposed model uses link flows, which are relatively inexpensive to obtain, as its statistical basis. Using real traffic flow data in a network of southern Sweden, the numerical study indicates that the proposed model may be used to estimate how changes in traffic flow volumes in one location affects the flow on a nearby link. Future work includes the study of how the hourly link flows can be analyzed from one hour to the next hour, to gain knowledge about the movement of vehicles over a larger geographic region.

References

[1] G. Leduc, et al., Road traffic data: Collection methods and applications, Working Papers on Energy, Transport and Climate Change 1 (2008) 1–55.
[2] M. L. Hazelton, Estimation of origin–destination matrices from link flows on uncongested networks, Transportation Research Part B: Methodological 34 (2000) 549–566.
[3] B. Hoh, M. Gruteser, H. Xiong, A. Alrabady, Enhancing Security and Privacy in Traffic-Monitoring Systems, IEEE Pervasive Computing 5 (2006) 38–46.
[4] E. Castillo, Z. Grande, A. Calviño, W. Y. Szeto, H. K. Lo, A State-of-the-Art Review of the Sensor Location, Flow Observability, Estimation, and Prediction Problems in Traffic Networks, Journal of Sensors 2015 (2015).
[5] T. Larsson, J. T. Lundgren, A. Peterson, Allocation of Link Flow Detectors for Origin-Destination Matrix Estimation-A Comparative Study, Computer-Aided Civil and Infrastructure Engineering 25 (2010) 116–131.
[6] J. L. Sullivan, D. C. Novak, L. Aultman-Hall, D. M. Scott, Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach, Transportation Research Part A: Policy and Practice 44 (2010) 323–336.
[7] E. Jenelius, T. Petersen, L.-G. Mattsson, Importance and exposure in road network vulnerability analysis, Transportation Research Part A: Policy and Practice 40 (2006) 537–560.
[8] X. Mao, C. Yuan, J. Gan, Incorporating Dynamic Traffic Distribution into Pavement Maintenance Optimization Model, Sustainability 11 (2019) 2488.
[9] X.-s. Li, X.-l. Zheng, W.-w. Guo, Y.-y. Ren, Y.-n. Wang, M. Yang, Traffic Accident Propagation Properties and Control Measures for Urban Links Based on Cellular Automata, Advances in Mechanical Engineering 5 (2013) 905640.
[10] M. Guettiche, H. Kheddouci, Critical links detection in stochastic networks: application to the transport networks, International Journal of Intelligent Computing and Cybernetics (2019).
[11] M. S. Ghanim, K. Shaaban, Estimating Turning Movements at Signalized Intersections Using Artificial Neural Networks, IEEE Transactions on Intelligent Transportation Systems 20 (2018) 1828–1836.
[12] C.-J. Lan, G. A. Davis, Real-time estimation of turning movement proportions from partial counts on urban networks, Transportation Research Part C: Emerging Technologies 7 (1999) 305–327.
[13] H.-T. Liou, S.-R. Hu, S. Peeta, Estimation of Time-Dependent Intersection Turning Proportions for Adaptive Traffic Signal Control under Limited Link Traffic Counts from Heterogeneous Sensors, Technical Report, School of Civil Engineering, Purdue University, 2017.
[14] R. Besenczi, N. Bátfai, P. Jeszenszky, R. Major, F. Monori, M. Ispány, Large-scale simulation of traffic flow using Markov model, PLoS ONE 16 (2021) e0246062.
[15] B. Priambodo, A. Ahmad, R. A. Kadir, Predicting Traffic Flow Propagation Based on Congestion at Neighbouring Roads Using Hidden Markov Model, IEEE Access (2021).
[16] J. Vreeswijk, T. Thomas, E. van Berkum, B. van Arem, Drivers' Perception of Route Alternatives as Indicator for the Indifference Band, Transportation Research Record 2383 (2013) 10–17.
[17] H. Bar-Gera, P. B. Mirchandani, F. Wu, Evaluating the assumption of independent turning probabilities, Transportation Research Part B: Methodological 40 (2006) 903–916.