

Study of a Spatial Structure of Urban Traffic Flows Using a Regime-Switching Vector Autoregressive Model*

Dmitry Pavlyuk^[0000–0003–3710–9678]

Transport and Telecommunication Institute, Riga, Latvia
Dmitry.Pavlyuk@tsi.lv

Abstract. This paper is devoted to multivariate regime-switching vector autoregressive models application taking into account a spatial structure in urban traffic flows. The spatial structure of traffic flows is estimated as statistical relationships between traffic characteristics (including lagged) at different road segments. A broad literature review is provided and complex spatial dependencies at road segments are illustrated. We execute empirical analyses of out-of-sample forecasting accuracy of several widely used traffic flow models: autoregressive integrated moving average model, vector autoregressive model, and Markov-switching vector autoregressive model. We provided empirical evidence for improving forecasting accuracy by including regime-dependent spatial dependencies of traffic flows at road segments into vector autoregressive models. Special attention is paid to selection of time resolution and its effects on the forecasting results.

Keywords: spatiotemporal modelling, traffic flow, vector autoregressive model, regime-switching model, forecasting

1 Introduction

During the last decades, models of urban traffic flows have become an integral part of intellectual transportation systems. A wide range of stakeholders use results of traffic forecasting on an everyday basis: many drivers use navigation software for routing and efficient determining departure time; public transport operators schedule bus routes on the base of expected traffic jams; emergency services use current and expected traffic data for solving vehicle routing problems; road network planning and maintenance operators try to predict effects of their road activities. Nowadays an intensive usage of urban traffic models is typical for large developed cities, but the rapid growth of modern traffic control technologies and inevitable spreading of autonomous vehicles will lead to the growing demand for new flexible approaches to traffic flow modelling.

* This work was financially supported by the post-doctoral research aid programme of the Republic of Latvia (project no. 1.1.1.2/VIAA/1/16/112, “Spatiotemporal urban traffic modelling using big data”), funded by the European Regional Development Fund.

At this moment the most popular approach to traffic flow forecasting is based on univariate models like the autoregressive integrated moving average model (ARIMA). Many scientific articles and practical applications include modelling of traffic flows at a selected (usually the most critical) road segment without any links to other parts of the road network. This non-spatial approach has several advantages, such as satisfactory forecasting accuracy, simplicity of models, limited data requirements, but ignore a fact of spatial dependencies between space points (crossroads, road segments, etc.) of traffic flows. As a result, the non-spatial approach becomes inconsistent in case of real-time changes of the road network like traffic congestion, accidents, and others. Simultaneous modelling of traffic flows at multiple space points and incorporation of spatial dependencies between space points should lead to higher forecasting accuracy. This fact, supported by high availability of real-time traffic data from multiple data sources (sensor loops, video recorders, probe cars' routes), raises attention to development of spatiotemporal models and their application to traffic flows.

This paper is our small contribution to the methodology of spatiotemporal traffic modelling. We applied a regime-switching vector autoregressive model where spatial dependencies between traffic flows at space points are directly included into the specification and may vary over the time. The suggested model is validated for real world traffic data. Our main contributions are the presented evidences of advantages of spatiotemporal regime-switching vector autoregressive models over widely used univariate time series models, and also several practical recommendations for such models applications.

2 Spatial dependencies in traffic flow modelling

Although the majority of previous studies utilized non-spatial forecasting models, recently several researchers extended the models to spatial settings. Space-time ARIMA (STARIMA) models seems the most widely used model with an explicitly included road network spatial structure. STARIMA represents a vector autoregressive model where dependencies between space points are provided in a form of a spatial contiguity matrix. The spatial contiguity matrix can be exogenously defined on the base of road network characteristics or identified using correlations. Kamarianakis and Prastacos [6] applied STARIMA with a fixed spatial contiguity matrix, based on inverted distances between space points, for traffic flow forecasting and demonstrated its advantages over non-spatial models; Min et al. [12] suggested to use dynamic spatial contiguity matrixes to reflect changes in road networks; Salamanis et al. [14] introduced a graph-theory-based technique for managing spatial dependencies in STARIMA models. However, an assumption about a fixed spatial contiguity matrix is rarely satisfied in practice, so several recent researches applied correlation analysis to dynamic identification of spatial relationships. Cheng et al. [3] executed exploratory space-time autocorrelation analysis for a road network and discovered the structure to be dynamic and heterogeneous in both space and time. Min and Wynter [11] utilised separate correlation-based contiguity matrixes for different time lags. Recently

Schimbinschi [15] suggested a topology-regularized vector autoregressive model, where potential relationships between space points are limited by distance, but remain flexible within this distance limit.

A popular least absolute shrinkage and selection operator (LASSO) and improved graphical LASSO (GLASSO) are frequently used for regularization and dependency selection to limit model complexity and prevent overfitting. Kamarianakis et al. [7] successfully applied the LASSO technique for real-time road traffic forecasting and later Haworth and Cheng [5] applied graphical LASSO for local spatiotemporal neighborhood selection. Li et al. [9] used Granger causality tests and further LASSO regression for similar purposes.

Classical autoregressive models reconstruct a time pattern of the traffic flow and use this pattern for forecasting. Although accuracy of such models is satisfactory for many practical purposes, it potentially could be improved by the utilizing the fact of different patterns of the traffic flow, which are switched over the time. For example, the traffic flow pattern could be different for free, stabilized or congested road conditions. Thus regime-switching models which allow the existence of multiple forms of spatial dependency, became a promising technique for traffic flow analysis (models are referred to as threshold models if time points of regime switching are predefined and as Markov-switching models if regime switching is a Markov process). Recently advanced regime-switching models have been utilized by several researchers. Yu and Zhang [17] proved that the Markov-switching ARIMA outperforms regular ARIMA specifications in terms of in-sample prediction accuracy. Cetin and Comert [1] also found significant advantages of threshold models in case of presence time intervals with different traffic regimes. Kamarianakis et al. [7] found the threshold ARIMA model useful for discovering spatial dependencies between values of the traffic flow in different locations. The latter paper, to the best of our knowledge, is the only research where spatial dependencies and the threshold regime-switching technique were successfully combined for traffic flow forecasting.

3 Research Methodology

Let define space-time point traffic information at a particular space point $i = 1, \dots, N$ and time point $t = 1, \dots, T$ as a vector X of K traffic flow characteristics:

$$X_{i,t} = (X_{1,i,t}, X_{2,i,t}, \dots, X_{K,i,t}) \quad (1)$$

Typically, a set of traffic flow characteristics includes traffic volume, average flow speed, flow density, and road occupancy. A one-step-ahead forecasting model for the traffic flow can be considered as a relationship between a space-time point traffic information and observed historical information:

$$X_{i,t+1} = f(X_{i,1}, \dots, X_{i,t}) + \varepsilon_{i,t+1} \quad (2)$$

where f is an unknown function, representing relationships between space-time points, and $\varepsilon_{i,t+1}$ is a random component.

A popular vector autoregressive model (VAR) model specification could be presented as [4]:

$$X_{i,t+1} = \sum_{h=0}^p A_h X_{i,t-h} + \varepsilon_{i,t+1} \quad (3)$$

where p is an order of autoregressive component, and A_h are matrices of unknown coefficients. The model could be reduced to the set of independent autoregressive models putting A_h to diagonal matrices: $A_h = \alpha_h I_N$. Moving average(MA) components, included into classical ARIMA models, are omitted in the presented VAR model specification. VAR models can be found as a good approximation for VARMA for sufficiency large lags p [10] (similarly to Wold's theorem), thus in this research we omitted MA components.

Thanks to the flexibility and ability to fit data, VAR models became an important forecasting tool in many applied areas. But this flexibility comes from a large number of model parameters and brings a risk of overfitting data and poor out-of-sample forecasting accuracy. Bayesian approach to estimation of model parameters partly helps to overcome the problem of overfitting [8]. In this research we use this approach to estimation and corresponding models will be referred to as BVAR (Bayesian VAR).

There are several widely used metrics for estimation of model's forecasting accuracy, and we use two of them – root-mean-square error (RMSE) and mean absolute error (MAE):

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (4)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (5)$$

See Chai and Draxler [2] for an extended discussion about advantages of RMSE and MAE as forecasting accuracy metrics.

To prevent effects of in-sample overfitting of VAR models, we use the out-of-sample values of RMSE and MAE for model comparison. In particular, we implemented a one-step-ahead model validation procedure, where parameters are estimated for every observation in the testing subsample and the resulting model is used to forecast one following observation.

One of the main assumptions of this research is presence of dynamical changes in spatial relationships between space points. An underlying mechanism of this dynamics is based on drivers' awareness of road conditions and selection of the optimal route. Nowadays many drivers use navigation software with real-time information about current traffic conditions and make their routing decision (or just follow navigation software instructions, which also based on real-time traffic information). As a result, traffic flows at two crossroads may be independent in case of free movement, but become dependent in case of congestion on one or both of them. We applied a regime-switching version of BVAR to reveal presence

of such regimes in data. We assume the Markovian nature of these regimes, so transitions between regimes (states) don't depend on process history.

We consider Markovian hidden states s that represent actual traffic flow regimes:

$$P(s_t | s_0, s_1, \dots, s_{t-1}, X_{i,1}, \dots, X_{i,t}) = P(s_t | s_{t-1}) \quad (6)$$

Each regime has its own estimates of the VAR model parameters, so the Markov-switching VAR model is formulated as MS-VAR(p):

$$X_{i,t+1 | s_t} = \sum_{h=0}^p A_h(s_{t-h}) X_{i,t-h} + \varepsilon_{i,t+1}^s \quad (7)$$

Given the current regime, the model is reduced to the regular BVAR model.

A list of additional techniques, utilized in this research, includes:

- Simple historical averages for separation a long-term pattern and disturbances in the traffic flow;
- Augmented Dickey-Fuller test for stationarity;
- Granger causality tests for identification of lagged relationships;
- Naïve forecasts as a base point for models' forecasting accuracy.

4 Experimental results

4.1 Data description

Similarly to Pavlyuk [13], we utilize data of traffic flows publicly available from the Minnesota Department of Transportation. Data is collected by two sensors located before bridges in Minneapolis city center and serving the same direction (Fig. 1).

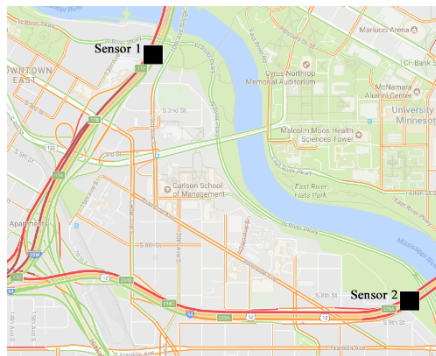


Fig. 1. Location of sensor loops

Traffic information includes average traffic flow density values, aggregated by three time resolutions: 10-minutes, 15-minutes, and 30-minutes time frames.

The research period is from the 1st of March till the 31st of May 2017 (92 days); the data sets are complete.

Typical daily patterns of analyzed traffic flow densities are presented on Fig. 2 and fairly usual for urban roads – low intensities during night hours, higher intensities during day hours and several congestion “spikes” in morning/evening rush hours.

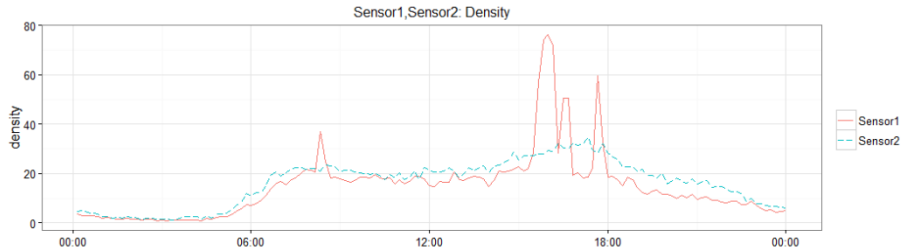


Fig. 2. Typical daily patterns of analyzed traffic flow densities

Traffic flow densities have clear historical profiles (for each time-of-day and day-of-the-week) which could be predicted by historical average values or seasonal exponential smoothing (with weekly “seasons”). The research period is relatively short and uniform, so the exponential smoothing does not significantly outperform simple historical averages, thus we continue with the latter option:

$$\text{Density}_t = \text{Density}_{avg,t} + \text{DensityStat}_t \quad (8)$$

where $\text{Density}_{avg,t}$ are historical average values for day-of-the-week time-of-day, and DensityStat_t are deviations from the historical averages and the main point of interest. We analyzed DensityStat_t using augmented Dickey–Fuller tests and found them stationary for both sensors.

4.2 Model specification and estimation

Density levels at both sensors become stationary after subtracting historical averages, so a direct application of VAR model is allowed. We apply Granger causality tests for initial identification of spatial relationships between traffic flow densities in two space points. The tests were implemented for density levels and for residuals of univariate ARIMA models (Table 1).

Granger tests support our hypothesis about presence of spatial relationships, but also give us information about potential asymmetry of this relationship: discovered effects of density at Sensor 2 point on density at Sensor 1 point have higher significance than in the opposite direction. Note that this conclusion has a preliminary nature, and VAR models will provide more consistent estimates of these spatial dependencies.

A stack of considered models includes:

Table 1. Results of Granger’s causality tests for independent ARIMA model residuals

Cause	Effect	Time resolution	F-statistic	p-value
Sensor 1	→ Sensor 2	10-mins	3.871	0.021
		15-mins	10.096	0.000
		30-mins	2.417	0.089
Sensor 2	→ Sensor 1	10-mins	5.973	0.003
		15-mins	16.752	0.000
		30-mins	38.246	0.000

- Naïve forecasts as a base point for models’ accuracy comparison.
- ARIMA($p, 0, q$) model, where optimal orders of autoregressive lags p and moving average component q , is selected for every new model on the base of the Akaike Information Criterion (AIC).
- Spatial BVAR model with 1 and 2 autoregressive lags (models with higher order of lags also were estimated, but didn’t outperform the presented models and excluded from this paper for brevity reasons)
- Markov-switching spatial BVAR model (MS-BVAR) with 1 and 2 autoregressive lags.

The main goal of this research is testing of forecasting accuracy of different models, so we exclude models’ estimation results from this paper and concentrate on their out-of-sample forecasts. To estimate out-of-sample forecasting accuracy we split the data set into training (91 days) and testing (1 day) subsamples (in the background we also executed similar procedures for a testing subsample of 7 days and obtained similar results). For every time point in the testing subsample (48, 96, and 144 points for 30-mins, 15-mins, and 10-mins time frames respectively) we estimated all models for previous observations and made one-step-ahead forecasts. These forecasts were used to estimate models’ out-of-sample forecasting accuracy using RMSE and MAE metrics. Estimation results are summarized in Table 2.

The obtained results are fairly consistent for RMSE and MAE metrics, so we arbitrary use MAE for further discussion.

5 Results and discussion

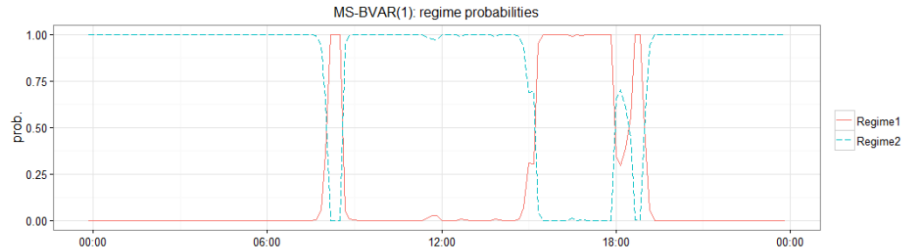
Firstly, we note significant difference in forecasting accuracy of Sensor 1 and Sensor 2 points (MAE for Sensor 1 forecasts are 2–3 times higher than for Sensor 2 forecasts). This conclusion corresponds to daily density patterns presented on Fig. 1: Sensor 1 point is more congested in rush hours and has obvious spikes, which significantly affect forecasting accuracy.

VAR models significantly outperform ARIMA and naïve forecasts for both space points; Markov-switching BVAR models are preferred for Sensor 1, and regular BVAR models are preferred for Sensor 2. This observation is consistent with the studies presented in Section 2: embedding of spatial relationships into

Table 2. Out-of-sample accuracy of models

		Time resolution	Naïve	ARIMA	BVAR ($p = 1$)	BVAR ($p = 2$)	MS-BVAR ($p = 1$)	MS-BVAR ($p = 2$)
Sensor 1	RMSE	10-mins	6.969	6.722	6.592	6.655	6.370	6.201
		15-mins	7.176	6.628	6.663	6.599	6.550	6.231
		30-mins	9.306	8.492	8.058	8.141	7.935	7.892
	MAE	10-mins	2.907	2.778	2.691	2.776	2.573	2.447
		15-mins	2.970	2.805	2.828	2.737	2.799	2.651
		30-mins	4.187	3.909	3.620	3.700	3.587	3.589
Sensor 2	RMSE	10-mins	1.699	1.515	1.556	1.513	1.757	1.805
		15-mins	1.220	1.165	1.137	1.121	1.311	1.524
		30-mins	1.374	1.301	1.297	1.307	1.293	1.302
	MAE	10-mins	1.263	1.136	1.167	1.118	1.247	1.318
		15-mins	0.946	0.912	0.846	0.828	0.987	1.059
		30-mins	1.005	0.936	0.893	0.903	0.910	0.889

traffic flow forecasting models leads to increased forecasting accuracy. An important observation is related to the preference of the regime-switching model for Sensor 1 and the regular (one regime) model for Sensor 2. Fig. 3 provides the clues for this fact.

**Fig. 3.** Estimated regime probabilities in the MS-BVAR(1) model

Note that the MS-BVAR model identified two regimes: rush hours (Regime 1) and normal traffic flow (Regime 2). On Fig. 1 we observed that Sensor 1 point is more affected by congested traffic during rush hours, so the regime-switching model provides more accurate forecasts. Congestion is less typical for Sensor 2 point, so one-regime BVAR model provides better out-of-sample forecasts, while MS-BVAR model suffers from over-parametrization and in-sample overfitting. Thus we can make a conclusion that even within a small road segment the best overall accuracy of forecasts is archived not from one general model, but from

a correct combination of models. The problem of forecasts combining is a hot topic in research publications on traffic forecasting; see Vlahogianni et al. [16] for an extended description of scientific challenges in this area.

Another important point of this research lays in effects of time resolution on models' forecasting accuracy. Note that forecasting results for different time frames are not directly comparable due to different forecasting horizons (one-step-ahead forecasts correspond to 10, 15, and 30 minutes respectively). Nevertheless, we can note that smaller time resolution requires a higher number of autoregressive lags as it was expected: for 30-mins time frames VAR models with one lags are preferred, for 15- and 10-mins time frames – VAR models with two lags (further increasing of the number of lags don't have significant effects on the models' forecasting accuracy for analyzed time resolutions). There is no solid approach to selection of the optimal time resolution for traffic flow modelling, and so the preferred time frames are a trade-off between detailed but noisy high resolution data, and low resolution data where some dependencies are smoothed by aggregation. An optimal resolution for discovering spatial relationships should correspond to the time of spatial effects (i.e. what time is required for drivers and their navigation software to identify a congestion and reroute). Impulse response functions, usually estimated with VAR models, could provide evidences for this spatial time lag, so are suggested as a tool for model identification.

6 Conclusions

This paper contributes to multivariate statistical modelling of traffic flows, where spatial dependencies and potential regimes are explicitly included. We applied modern Markov-switching vector autoregressive models for traffic flow forecasting and compared out-of-sample forecasting accuracy for several popular models. Also, we briefly discussed problems of optimal model and time resolution selection. Our main conclusions are:

- Vector autoregressive models with spatial dependencies between space points outperform classical univariate model like ARIMA.
- Spatial relationships between space points should be considered as asymmetric, i.e. effect of changes in traffic flow in one space point on another is not invertible.
- Multiparametric models (i.e. Markov-switching vector autoregressive models) suffer from sample data overfitting, which leads to poor out-of-sample forecasting accuracy. Selection of an appropriate model is space point-specific, so a complex forecast for road network with a large number of space points should be constructed from a set of models and combined forecasts.
- Selection of time resolution (an aggregation level) is an important step of modelling. Time resolution is a trade-off between detailed but noisy high resolution data, and low resolution data, where some dependencies are smoothed by aggregation. As lags, related to spatial dependencies, depend on road network and drivers' behavior, the optimal time resolution should be estimated for every specific data set.

References

1. Cetin, M., Comert, G.: Short-term traffic flow prediction with regime switching models. *Transportation research record. Journal of the Transportation Research Board* 1965, 23–31 (2006)
2. Chai, T., Draxler, R.R.: Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* 7(3), 1247–1250 (2014)
3. Cheng, T., Haworth, J., Wang, J.: Spatio-temporal autocorrelation of road network data. *Journal of Geographical Systems* 14, 389–413 (2012)
4. Hamilton, J.: *Time series analysis*. Princeton University Press, Princeton, N.J (1994)
5. Haworth, J., Cheng, T.: Graphical LASSO for local spatio-temporal neighbourhood selection. In: *Proceedings the GIS Research UK 22nd Annual Conference*. Presented at the GISRUK. pp. 425–433 (2014)
6. Kamarianakis, Y., Prastacos, P.: Space-time modeling of traffic flow. *Computers & Geosciences* 31(2), 119–133 (2005)
7. Kamarianakis, Y., Shen, W., Wynter, L.: Real-time road traffic forecasting using regime-switching space-time models and adaptive LASSO. *Applied Stochastic Models in Business and Industry* 28, 297–315 (2012)
8. Karlsson, S.: Forecasting with bayesian vector autoregression. In: *Handbook of Economic Forecasting*, vol. 2, chap. Chapter 15, pp. 791–897. Elsevier (2013)
9. Li, L., Su, X., Wang, Y., Lin, Y., Li, Z., Li, Y.: Robust causal dependence mining in big data network and its application to traffic flow predictions. *Transportation Research Part C: Emerging Technologies* 58(Part B), 292–307 (2015)
10. Lütkepohl, H.: *Forecasting Cointegrated VARMA Processes*, pp. 179–205. Blackwell Publishing Ltd (2007)
11. Min, W., Wynter, L.: Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies* 19, 606–616 (2011)
12. Min, X., Hu, J., Chen, Q., Zhang, T., Zhang, Y.: Short-term traffic flow forecasting of urban network based on dynamic STARIMA model. In: *Proc.12th Int. Conf. Intelligent Transportation Systems (ITSC)*. pp. 1–6 (01 2009)
13. Pavlyuk, D.: On application of regime-switching models for short-term traffic flow forecasting. In: Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., Kacprzyk, J. (eds.) *Advances in Dependability Engineering of Complex Systems: Proceedings of the Twelfth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX*, July 2–6, 2017, Brunów, Poland. pp. 340–349. Springer International Publishing, Cham (2018)
14. Salamanis, A., Kehagias, D., Filelis-Papadopoulos, C., Tzovaras, D., Gravvanis, G.A.: Managing spatial graph dependencies in large volumes of traffic data for travel-time prediction. *IEEE Transactions on Intelligent Transportation Systems* 17, 1678–1687 (2016)
15. Schimbinschi, F., Moreira-Matias, L., Nguyen, V., Bailey, J.: Topology-regularized universal vector autoregression for traffic forecasting in large urban areas. *Expert Systems with Applications* 82, 301–316 (2017)
16. Vlahogianni, E., Karlaftis, M., Golias, J.: Short-term traffic forecasting: Where we are and where we’re going. *Transportation Research Part C: Emerging Technologies* 43, 3–19 (2014)
17. Yu, G., Zhang, C.: Switching ARIMA model based forecasting for traffic flow. In: *Acoustics, Speech, and Signal Processing*, 2004. *Proceedings*. pp. ii–429 (2004)