Transportation Research Part B 76 (2015) 27–47

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Nonlinear multivariate time–space threshold vector error correction model for short term traffic state prediction

Tao Ma [a,*], Zhou Zhou [b], Baher Abdulhai [a]

[a] Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario M5S 1A4, Canada
[b] Department of Statistical Sciences, University of Toronto, 100 St. George Street, Toronto, Ontario M5S 3G3, Canada

ARTICLE INFO

Article history:
Received 8 August 2014
Received in revised form 18 January 2015
Accepted 22 February 2015
Available online 19 March 2015

Keywords:
Cointegration
Vector error correction
Threshold regime switching
Short term traffic state prediction
Neural Network

ABSTRACT

We propose Time–Space Threshold Vector Error Correction (TS-TVEC) model for short term (hourly) traffic state prediction. The theory and method of cointegration with error correction mechanism is employed in the general design of the new statistical model TS-TVEC. An inherent connection between mathematical form of error correction model and traffic flow theory is revealed through the transformation of the well-known Fundamental Traffic Diagrams. A threshold regime switching framework is implemented to overcome any unknown structural changes in traffic time series. Spatial cross correlated information is incorporated with a piecewise linear vector error correction model. A Neural Network model is also constructed in parallel to comparatively test the effectiveness and robustness of the new statistical model. Our empirical study shows that the TS-TVEC model is an effective tool that is capable of modeling the complexity of stochastic traffic flow processes and potentially applicable to real time traffic state prediction.

1. Introduction

With an accurate prediction of traffic state, we are able to use Advanced Traffic Management System (ATMS) and Advanced Traveler Information System (ATIS) to effectively manage traffic flow with proactive dynamic traffic control and traveler's route guidance. Our ultimate goal is to bridge the gap between infrastructure supply and traffic demand, maximize utilization of capacity of existing infrastructure, and timely formulate traffic management solutions to solve traffic congestion problems as well as provide safe and convenient ways for travelers to reduce their travel time in a complex urban environment. Therefore, this research focusing on hourly traffic state prediction is an essential component of Intelligent Transportation Systems and applications.

1.1. Challenges of traffic state modeling and forecasting

The challenge of modeling traffic state lies in the intricate characteristics of dynamic and stochastic traffic processes. The time series of traffic volume, speed, and occupancy collected from different locations usually exhibit various characteristics. From a statistical viewpoint, all these characteristics can be summarized with the following properties: multi-seasonality, non-stationarity, temporal and spatial correlations, and dynamics between traffic variables that lead to the interaction of multivariate traffic time series. From a viewpoint of macroscopic traffic flow theory, traffic flow is generally classified as a

* Corresponding author.
E-mail address: tao.ma@utoronto.ca (T. Ma).

http://dx.doi.org/10.1016/j.trb.2015.02.008
0191-2615/

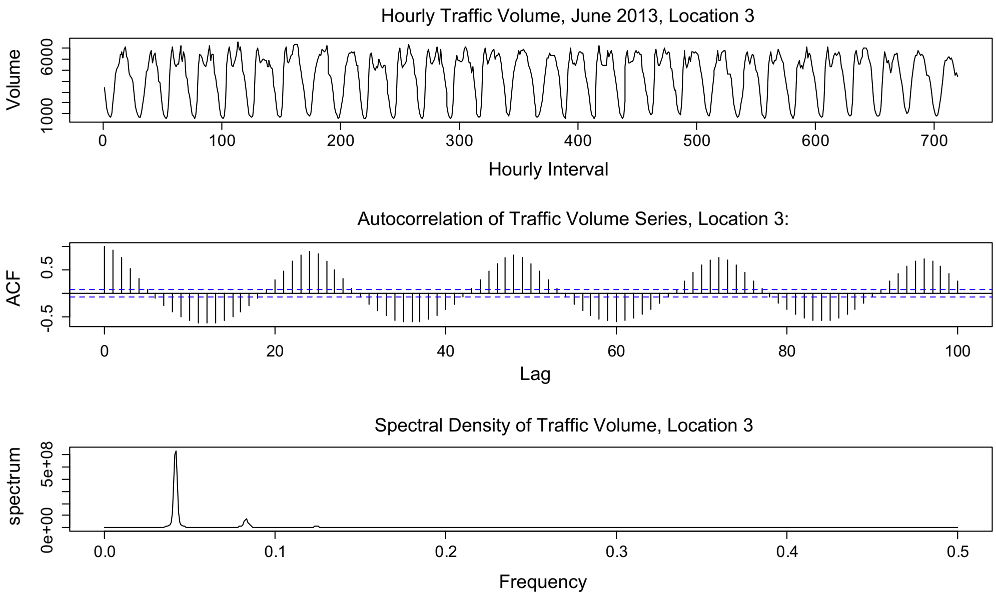

Fig. 1. Plot of traffic volume, ACF, and spectrum from sample data.

free flow or congested state. In congested state, abrupt declination in traffic volume and speed due to instability and/or incidents can be defined as outliers in the context of conventional time series modeling. However, they cannot be simply removed in traffic modeling because those outliers imply traffic congestion, which is of primary interest in traffic management. In order to model and forecast traffic more accurately using time series, all these properties and factors have to be taken into account and captured during the course of modeling and forecasting. There is no single time series model available in the literature that can incorporate these factors all at once.

Fig. 1 shows the plot of traffic volume series from location 3,[1] its autocorrelogram (ACF), and spectral density. Both the ACF and spectral density indicate the existence of seasonality in the time series.

Table 1 shows the results of Augmented Dickey–Fuller (ADF) (Dickey and Fuller, 1979, 1981) and KPSS (Kwiatkowski et al., 1992) unit root test to check the degree of non-stationarity in the traffic time series. These statistical tests indicate that the null hypothesis with unit roots cannot be rejected such that the traffic time series from the three locations are non-stationary. Fig. 2 shows the cross-correlograms of two pairs of traffic time series to investigate the dynamics between traffic variables as well as their spatial cross-correlation. These cross-correlograms reveal the existence of cross-correlation between the traffic volume and speed at the same site and also indicate that the spatial cross-correlation exists between the upstream traffic volume and the volume at the site of prediction.

1.2. Existing methods

Over the past two decades, the attentions and efforts of modeling and forecasting macroscopic traffic states evolved from univariate temporal correlation to multivariate temporal-spatial correlation and from linear to nonlinear form. Those models may be loosely classified as statistical and non-statistical methods. Some representatives include the class of Time Series models, e.g. seasonal ARIMA (Autoregressive Integrated Moving Average) (Williams and Hoel, 2003) and STARIMA (Space–Time ARIMA) (Kamarianakis and Prastacos, 2003), Kalman Filter State-Space model (Antoniou et al., 2005), Neural Network (Qiao et al., 2001; Abdulhai et al., 2002; Zheng et al., 2006), Nonparametric Regression (Smith et al., 2002), and Support Vector Regression (Wu et al., 2004), Bayesian Network (Castillo et al., 2008).

Another class of works is based on macroscopic traffic flow theory to estimate the internal traffic state for any intermediate point on a freeway or arterial segment from the boundary conditions. Many researchers use the cell-based or link-based modeling approach. Representative works include the Cell Transmission Model (CTM) (Daganzo, 1994, 1995), variational kinematic waves (Daganzo, 2005), CTM-based traffic state estimation models (Muñoz et al., 2003), second-order traffic flow model with Kalman filter (Nanthawichit et al., 2003,Wang and Papageorgiou, 2005), CTM-based second-order traffic flow model with particle filtering (Mihaylova et al., 2007), the Lighthill–Whitham–Richards partial differential equation (LWR PDE) (Lighthill and Whitham, 1955; Richards, 1956) with the Lagrangian measurements (Herrera and Bayen, 2010), Newell's simplified kinematic wave model (Newell, 1993), stochastic Newell's three-detector method (Laval et al., 2012; Deng et al., 2013), etc.

[1] See data map in Fig. 5 for information on the locations of data collection.