Journal of Transport Geography 46 (2015) 244–247

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Temporal aggregation and spatio-temporal traffic modeling

Marco Percoco *

Department of Policy Analysis and Public Management, CERTeT, Università Bocconi, Italy

ARTICLE INFO

Article history:
Received 2 April 2014
Revised 22 June 2015
Accepted 1 July 2015
Available online 7 July 2015

Keywords:
Traffic forecasting
Uncertainty
Spatial–temporal correlation
Covariance separability

ABSTRACT

Traffic forecasting is crucial for policy making in the transport sector. Recently, Selby and Kockelman (2013) have proposed spatial interpolation techniques as suitable tools to forecast traffic at different locations. In this paper, we argue that an eventual source of uncertainty over those forecasts derives from temporal aggregation. However, we prove that the spatio-temporal correlation function is robust to temporal aggregations schemes when the covariance of traffic in different locations is separable in space and time. We prove empirically this result by conducting an extensive simulation study on the spatial structure of the Milan road network.

1. Introduction

Decision making about transport planning and investment projects depends on traffic forecasting and its reliability. Recently, Selby and Kockelman (2013) and Wang and Kockelman (2009) have proposed the use of spatial interpolation techniques to obtain spatio-temporal predictions of traffic, with specific reference to Texas in terms of Annual Average Daily Traffic (i.e., an aggregate measure of traffic intensity). In this paper, I argue that temporal aggregation of traffic data may pose severe problems of estimation, unless the covariance function is separable in space and time.

Data are, in fact, often aggregated with respect to the data-generating process (Marcellino, 1999). Many researchers have considered temporal aggregation in time-series models (Granger, 1993; Lutkepohl, 1987; Marcellino, 1999; Wei, 1982). As for spatio-temporal models, Giacomini and Granger (2004) have compared the relative forecasting efficiency of different methods after aggregating over space and by using small sample simulations. They found that better results can be obtained by imposing *a priori* constraints on the cross-section dependence of observation units in vector autoregressions. Similarly, Percoco (2007) has studied the forecasting performance of the space–time autoregressive model under temporal aggregation.

In this paper, some issues related to the approach proposed by Selby and Kockelman (2013) and Wang and Kockelman (2009) are proposed by studying the robustness of the spatio-temporal correlation function of traffic time series in different locations under temporal aggregation schemes. In particular, the paper presents a relatively simple implication of predicting Annual Average Daily Traffic (AADT) from aggregated data when the covariance is not separable. As such, the contribution of the paper should not be found in a new estimation procedure, but as a discussion of the implication of the intuition of Selby and Kockelman (2013). The research found that it does not depend on temporal aggregation only if it is separable. Results are confirmed by an extensive simulation study over the Milan, Italy, road network. The paper is organised as follows. In Section 2, basic definitions are presented, followed in Section 3 by the main results. A simulation study is offered in Section 4, and Section 5 presents conclusions.

2. Assumptions and related concepts

Let us assume a spatio-temporal stochastic process of traffic generation in the form, $\{Y(\mathbf{u},t); \mathbf{u} \in D_s, t \in D_t\}$ evolving through the spatiotemporal index $D_s \times D_t$. Let T be a time aggregation scheme,[1] then it is possible to define the integral process[2] as

$$X_T(\mathbf{u},t) = \int_{t-\frac{T}{2}}^{t+\frac{T}{2}} Y(\mathbf{u},t)\mathrm{d}\tau \tag{1}$$

[1] A temporal aggregation scheme is an aggregation rule over the time. In the case of AADT, it is the average daily traffic over a year period. In the case of the integral process in (1), the aggregation rule is the sum (integral).

[2] The results in this paper are derived by considering an integral process, however, they hold also for the mean intensity process, such as the Annual Average Daily Traffic, since one is a scaled version of the other.

* Address: PAM, Università Bocconi, via Rontgen 1, 20121 Milano, Italy.
E-mail address: marco.percoco@unibocconi.it

http://dx.doi.org/10.1016/j.jtrangeo.2015.07.001
0966-6923/

Let us also assume that process $\{Y(\mathbf{u},t)\}$ is stationary in space and time and that, without loss of generality, $D_s \subset R^2$. Under the assumption of temporal stationarity[3] (at least of the second order), the covariance of process (1) measured at locations u_1 and u_2 at time t_1 and t_2 can be expressed as:

$$\mathrm{cov}(u_1,u_2,|t_1-t_2|) = \sigma(u_1)\sigma(u_2)r(u_1,u_2,|t_1-t_2|) \tag{2}$$

where $\sigma(u_1), \sigma(u_2)$ are the standard deviation at locations u_1 and u_2 and $r(u_1,u_2,|t_1-t_2|)$ is the spatio-temporal correlation function. In accordance with Cressie and Wikle (2011), the following definition is introduced:

Definition 1. The correlation function in (2) is separable if

$$\mathrm{cov}(u_1,u_2,|t_1-t_2|) = \sigma(u_1)\sigma(u_2)r_d(|u_1-u_2|)r_t(|t_1-t_2|) \tag{3}$$

where r_t is the correlation function over time and r_d is the correlation function over space.

As primarily argued by Dalezios and Adamowski (1995), the identification of a spatio-temporal dynamic model in terms of spatial and temporal lags can be conducted by estimating a space–time autocorrelation function and a space–time partial autocorrelation function. Interestingly enough, Di Giacinto (2010) makes use of this approach to identify the lags in a space–time vector autoregressive model. In what follows, I show that the spatio-temporal correlation function is sensitive to aggregation schemes when it is not separable.

More importantly, Selby and Kockelman (2013) and Wang and Kockelman (2009) make use of kriging to predict traffic flows. This spatial interpolation procedure (or set of procedures) is highly dependent on the covariance function in (3) and the correlation function in both the trend and the drift part of the traffic generating process, as shown in Cressie and Wikle (2011). In what follows, I first prove that temporal aggregation may pose some problems in the correlation function, and then in kriging predictions.

3. Results

In this section I show that if the covariance function of traffic flows is separable, then it is robust to temporal aggregation schemes. Furthermore the analysis shows that if the covariance is not separable, then temporal aggregation imposes a bias in the correlation function.

Proposition 1. *If the covariance of the traffic generating process $\{Y(\mathbf{u},t)\}$ is separable, then the spatio-temporal correlation function is robust to temporal aggregation schemes.*

Proof. Let us consider the correlation function of traffic flows at two spatial locations, at two points in time t and $t+\tau$ and under the same time aggregation window T. Setting $\Gamma_T(0) =) \iint r_t(|\mu-\alpha|)\mathrm{d}\mu\mathrm{d}\alpha$ with $\tau = 0$, the covariance of the integral process at points u_1, u_2 can be expressed as

$$\mathrm{cov}_T(u_1,u_2,t,t) = \sigma(u_1)\sigma(u_2)r_d(|u_1-u_2|)\Gamma_T(0) \tag{4}$$

The corresponding correlation coefficient of the integral process $\{X_T(u,t)\}$ at aggregation level T is defined as:

$$r_T(u_1,u_2,t,t) = \frac{\mathrm{cov}_T(u_1,u_2,t,t)}{\sqrt{\mathrm{cov}_T(u_1,u_1,t,t)\mathrm{cov}_T(u_2,u_2,t,t)}}$$

and if the assumption of separability holds then:

$$r_T = \frac{\sigma(u_1)\sigma(u_2)r_d(|u_1-u_2|)\Gamma_T(0)}{\sqrt{\sigma^2(u_1)\Gamma_T(0)\sigma^2(u_2)\Gamma_T(0)}}$$

Finally, we have:

$$r_T = r_d(|u_1-u_2|) \qquad \square \tag{5}$$

The result in (5) implies that if the correlation function of the traffic-generating process $\{Y(\mathbf{u},t)\}$ can be separated into a product of the correlation function, depending only on inter-observation points distance $|u_1-u_2|$ and the correlation function depending only on time interval, $|t_1-t_2|$ then the correlation coefficient r_T of the integral process does not depend on the time aggregation scale T.[4] Let us now turn our attention to the proof that if the covariance is not separable then the correlation function of traffic flows depends on temporal aggregation.

Lemma 1. *If the covariance of the process $\{Y(\mathbf{u},t); \mathbf{u} \in D_s, t \in D_t\}$ is not separable, then the spatio-temporal correlation function depends on temporal aggregation schemes.*

Proof. Let us consider the correlation coefficients at two different spatial points u_1, u_2 of time t, and $t+\tau$, the covariance can be obtained as

$$\begin{aligned}\mathrm{cov}_T(u_1,u_2,t,t+\tau) &= E\left\{\int_{t-\frac{T}{2}}^{t+\frac{T}{2}}[X(u_1,\mu)-m(u_1)]\mathrm{d}\mu \right.\\ &\quad \left.\cdot \int_{t+\tau-\frac{T}{2}}^{t+\tau+\frac{T}{2}}[X(u_2,\mu)-m(u_2)]\mathrm{d}\alpha\right\}\\ &= \sigma(u_1)\sigma(u_2), \iint r(|u_1-u_2|,|\mu-\alpha|)\mathrm{d}\mu\mathrm{d}\alpha\end{aligned}$$

That is:

$$\mathrm{cov}_T(u_1,u_2,t,t+\tau) = \sigma(u_1)\sigma(u_2)\Gamma'_T(d;\tau)$$

where $\Gamma'_T(d;\tau) = \iint r(|u_1-u_2)|;|\mu-\alpha|)\mathrm{d}\mu\mathrm{d}\alpha$ m is the mean at locations u_1 and u_2 and $E[.]$ denotes the expected value. The corresponding correlation coefficient of the same process under aggregation scheme T can be written as:

$$r_T(d,\tau) = \frac{\mathrm{cov}_T(u_1,u_2,t,t+\tau)}{\sqrt{\mathrm{cov}_T(u_1,u_1,t,t)\mathrm{cov}_T(u_2,u_2,t+\tau,t+\tau)}}$$

If the correlation function is not separable, then:

$$r_T(d,\tau) = \frac{\sigma(u_1)\sigma(u_2)\Gamma'_T(d;\tau)}{\sqrt{\sigma^2(u_1)\Gamma_T(0;0)\sigma^2(u_2)\Gamma_T(0;0)}}$$

i.e., $r_T = \frac{\Gamma'_T(d,\tau)}{\Gamma_T(0,0)}$ where $\Gamma'_T(0,0) = \Gamma_T(0)$ and then $r_T = \frac{\Gamma'_T(d,\tau)}{\Gamma_T(0)}$. In other words, if the covariance of the process is not separable, then the correlation function depends on temporal aggregation. Furthermore, it can be shown that both spatial and temporal autocorrelation functions depend on temporal aggregation when the covariance is not separable. To this end, let us consider the covariance at point u and time $t, t+\tau$:

$$\mathrm{cov}_T(u,t,t+\tau) = \sigma^2(u)\Gamma_T(\tau)$$

In this case, we have $\Gamma'_T(0;\tau) = \Gamma_T(\tau)$ and the corresponding correlation coefficient is:

[3] A process $\{Y(\mathbf{u},t)\}$ is stationary if for any couple of points in space and time (u_1, u_2) and (t_1, t_2) its covariance function can be written as, $\mathrm{cov}(|u_1-u_2|,|t_1-t_2|)$ where the first argument stands for a distance measure.

[4] As it can be noted, the sole assumption needed for the derivation of the above results is the required stationarity in time. In fact, the relations developed for the covariance of the process $\{Y(\mathbf{u},t)\}$ and that of the integral process require the hypothesis of stationarity to hold at least on the interval $[t-\frac{T}{2}, t+\tau+\frac{T}{2}]$.