Transportation Research Part B 95 (2017) 196–213

Contents lists available at ScienceDirect

Transportation Research Part B

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

A statistical method for estimating predictable differences between daily traffic flow profiles

F. Crawford*, D.P. Watling, R.D. Connors

Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK

ARTICLE INFO

Article history:
Received 22 February 2016
Revised 6 October 2016
Accepted 5 November 2016
Available online 12 November 2016

Keywords:
Functional Data Analysis
Functional Linear Models
Systematic variations
Traffic flow profiles
Day-to-day variability

ABSTRACT

It is well known that traffic flows in road networks may vary not only within the day but also between days. Existing models including day-to-day variability usually represent all variability as *unpredictable* fluctuations. In reality, however, some of the differences in flows on a road may be *predictable* for transport planners with access to historical data. For example, flow profiles may be systematically different on Mondays compared to Fridays due to predictable differences in underlying activity patterns. By identifying days of the week or times of year where flows are predictably different, models can be developed or model inputs can be amended (in the case of day-to-day dynamical models) to test the robustness of proposed policies or to inform the development of policies which vary according to these predictably different day types. Such policies could include time-of-day varying congestion charges that themselves vary by day of the week or season, or targeting public transport provision so that timetables are more responsive to the day of the week and seasonal needs of travellers. A statistical approach is presented for identifying systematic variations in daily traffic flow profiles based on known explanatory factors such as the day of the week and the season. In order to examine day-to-day variability whilst also considering within-day dynamics, the distribution of flows throughout a day are analysed using Functional Linear Models. F-type tests for functional data are then used to compare alternative model specifications for the predictable variability. The output of the method is an average flow profile for each predictably different day type, which could include day of the week or time of year. An application to real-life traffic flow data for a two-year period is provided. The shape of the daily profile was found to be significantly different for each day of the week, including differences in the timing and width of peak flows and also the relationship between peak and inter-peak flows. Seasonal differences in flow profiles were also identified for each day of the week.

1. Introduction

There is now extensive literature analysing and modelling the extent to which traffic flows systematically vary within a day, due to time-of-day variations in demand, time-of-day variations in capacity (e.g. due to traffic signals), and the temporal and spatial interactions of congestion (Ukkusuri et al., 2012; Du et al., 2015; Han et al., 2015; Long et al., 2016; Ngoduy et al., 2016; Wang and Du, 2016). A corresponding body of work has additionally sought to address the considerable variation

* Corresponding author.
E-mail address: ts12fc@leeds.ac.uk (F. Crawford).

http://dx.doi.org/10.1016/j.trb.2016.11.004
0191-2615/

observed in traffic flows between days, known as day-to-day variability (Watling and Cantarella, 2013a, 2013b; Guo et al., 2015; Hazelton and Parry, 2016; Kumar and Peeta, 2015; Xiao et al., 2016). This twin focus, on within-day and day-to-day variation, is the topic of the present paper. Existing models including day-to-day variability usually represent variability by a single probability distribution for each randomly varying component[1], for example in research on demand (Watling, 2002; Clark and Watling, 2005; Sumalee et al., 2006; Shao et al., 2006; Nakayama and Watling, 2014), capacity (Lo and Tung, 2003; Siu and Lo, 2008; Sumalee et al., 2011b) or travel times (Noland and Small, 1995; Clark and Watling, 2005; Pu, 2011; Guo et al., 2012). In contrast to existing models, this research proposes separating the predictable and unpredictable components of variability so that variability could be represented by a set of probability distributions alongside a set of rules specifying which distribution relates to which day type. In this paper 'day type' relates to an exhaustive classification based on combining characteristics which would be known far in advance such as the day of the week or season. This classification could be anything from a simple weekday/weekend day split to a complex combination of days of the week and months. This research develops a stochastic model of within-day profiles that includes day types as explanatory factors to identify predictably different day types in a dataset. The functional day type coefficients are estimated, but the development of full probability distributions for models with both within-day and day-to-day dynamics is left for future work. The outputs of the method presented below are also useful in their own right as they can be used by practitioners to better understand travel patterns and perhaps inform day type specific policies in order to better utilise resources. The flow profiles for each day type can also be used to test the robustness of policies and, more importantly, as the day types are known far in advance plans can be made to mitigate potential problems.

The seminal work of Hanson and Huff (1988) provided evidence that cyclical patterns exist in individual travel behaviour. These patterns or cycles are often based around the days of the week or seasons, as evidenced by multi-day surveys (Kitamura and Van Der Hoorn, 1987; Schlich and Axhausen, 2003; Habib and Miller, 2008). More recently, data from emerging data sources have been used to examine activity patterns over a longer period of time, for example Järv et al. (2014) who used mobile phone data to examine monthly variations in activity spaces. Whilst in many cases such patterns will disappear once data has been aggregated, in some cases exogenous factors can cause systematic patterns in travel behaviour which translates into predictably different travel conditions. This could include the widely accepted weekday patterns of peak and off-peak traffic, but also patterns which are more likely to be overlooked, for example lower network demand in winter months or variations in daily flow profiles due to differences in shop opening hours between weekdays. Researchers often try to avoid the component of variability which is predictable by considering "some nominally 'typical' conditions" (Clark and Watling, 2005, p. 119), such as the 'peak' period of the day on 'non-holiday weekdays' only. Although some researchers have explored the impact of various types of predictable variability when undertaking analysis of flow or travel time data (Rakha and Van Aerde, 1995; Stathopoulos and Karlaftis, 2001; Zhang et al., 2007; Yazici et al., 2012), few have built on this to develop predictive models.

Two exceptions are Kamga and Yazıcı (2014) and Guardiola et al. (2014). Kamga and Yazıcı (2014) used GPS data from taxis to classify average travel times per unit distance across the city for each hour of the day and day of the week using regression trees. Guardiola et al. (2014) used Functional Data Analysis on traffic flow profiles from a detector set on a freeway for the purpose of classification and outlier detection. They used Functional Principal Component Analysis to identify the three Principal Components. The first appears to separate working from non-working days, the second may relate to the year and the third may be a seasonal factor.

The present paper builds on the work of Guardiola et al. (2014) as it also considers variability in daily traffic flow profiles, but it differs by considering day types which would be known in advance, such as the day of the week, rather than using data-driven category selection. Traffic flow data is perhaps the most widely collected data on road networks and therefore there are vast amounts of data to analyse, even when considering just one day type. Flow data informs us about what is occurring on the road network at any given point in time and therefore is particularly of interest to practitioners and those calibrating models.

This research proposes a method for identifying scenarios where traffic flow profiles differ predictably, as a result of characteristics which would be known in advance, such as the day of the week. Day type explanatory models will be estimated for both the magnitude and shape of the daily flow profiles. Data from one loop detector will be considered. The flow at a single location could vary due to many factors, including demand, route choice, departure time and the traffic conditions on other parts of the network. In this research the aim is to identify the day types or seasons where flows at this location are systematically different to those on other days, regardless of the cause, to inform scenario testing. For a single location, this is relevant for modelling localised policies such as capacity reductions (for example due to road works or changes in parking regulations) or congestion charging boundaries. The methodology could also be applied independently to many detectors in an area in order to identify days of the week or times of year where particular problems arise so that policy solutions to target the causes of these problems can be devised. Examples of such targeted policies might include: time-of-day varying congestion charges that themselves vary by day of the week or season; incentives to influence employers or shopping centres to adjust their opening times by day of the week or season; targeting public transport provision so that timetables are more responsive to the day of the week and seasonal needs of travellers.

[1] Whilst Ettema et al. (2005) did develop a day to day dynamical model where previously experienced travel times were stored in separate categories within memory, they did not discuss how the categories could be formed or used in practice.