Computational Statistics and Data Analysis 73 (2014) 189–204

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

# Modelling trends in road accident frequency—Bayesian inference for rates with uncertain exposure

Louise K. Lloyd [a,*], Jonathan J. Forster [b]

[a] Transport Research Laboratory, Crowthorne House, Nine Mile Ride, Wokingham, RG40 3GA, United Kingdom
[b] School of Mathematics, University of Southampton, United Kingdom

ARTICLE INFO

*Article history:*
Received 4 July 2012
Received in revised form 17 October 2013
Accepted 20 October 2013
Available online 26 October 2013

*Keywords:*
Bayesian analysis
Disaggregation
Exposure
MCMC
Road accidents
Uncertainty

ABSTRACT

Traffic flow data are primarily used to monitor road use and to compute road accident rates in Great Britain. The main traffic flow data used for these purposes measure annual traffic flow in vehicle kilometres, however this dataset is limited in its disaggregation. In particular, it is not possible to quantify traffic flow by different types of cars using just these flow data. Two additional sources of data are introduced (the number of cars registered each year and an induced exposure dataset containing information about different road use by different car types) and a model combines the three datasets in order to produce a disaggregation of traffic flow by car type and road type on the commonly used annual traffic data. MCMC algorithms are used to simulate from the posterior distributions and produce estimates of the traffic by three road types and six car types across 12 years. These flow estimates are then used in models for accident rates.



## 1. Introduction

Annual traffic flow, defined as the number of vehicle kilometres travelled in a year in Great Britain, is used to monitor trends in road travel across the country on different road types, at different times of day and year, and by different vehicle types. These trends help to define areas of congestion and inform government expenditure for road construction, improvement and structural maintenance.

In addition, traffic flows are also used to estimate the risk of being involved in an accident. Typically accident risk is defined as the rate of accidents per vehicle kilometre travelled each year which can be disaggregated as much as the traffic flow data allows: that is, by region, road type, time of day and year, and vehicle type. Flow information about different types of cars, from small saloons and minis to 4 × 4s and people carriers, is not available and is estimated from other sources. In general, traffic flow data are replaced by information on the number of registered cars each year in order to compute accident rates in these situations.

Notably, it is clear from registered vehicle data that the number of 4 × 4 and people carrier type cars has increased dramatically over the last 10 years from 4% in 1999 to 12% of the car population in 2010, and it has been shown by Knowles et al. (2007) that 4 × 4 style cars have different types and consequences of accidents to smaller cars. In previous similar research (Knowles et al., 2007; Starnes and Longthorne, 2003; Broughton, 2007; Keall and Newstead, 2007), accident rates are quoted in relation to the number of registered vehicles as the most appropriate flow data available at the time, however the appropriateness of this measure is disputed in each of these papers. af Wahlburg and Dorn (2007) suggest that traffic flows cannot be directly related to the distribution of the different car types registered, as drivers of different car types generally have different driving habits, including use of different road types and varying annual mileage. Therefore a method

* Corresponding author.
*E-mail address:* lklloyd@trl.co.uk (L.K. Lloyd).

0167-9473/$ – see front matter © 2014 Transport Research Laboratory. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2013.10.020

**Table 1**
Annual car traffic flow in Great Britain (billion vehicle kilometres).

| Year | Motorway | A roads | Minor roads |
|---|---|---|---|
| 1999 | 63 | 174 | 143 |
| 2000 | 71 | 165 | 143 |
| 2001 | 72 | 168 | 144 |
| 2002 | 70 | 177 | 145 |
| 2003 | 70 | 179 | 144 |
| 2004 | 73 | 181 | 145 |
| 2005 | 73 | 180 | 145 |
| 2006 | 74 | 181 | 142 |
| 2007 | 75 | 178 | 145 |
| 2008 | 75 | 177 | 143 |
| 2009 | 75 | 178 | 141 |
| 2010 | 74 | 175 | 137 |

**Table 2**
Number of registered vehicles in Great Britain by car type and year (millions).

| Year | Minis and superminis | Small saloons | Medium saloons | Large saloons | 4 × 4 and people carriers | Sports cars |
|---|---|---|---|---|---|---|
| 1999 | 6.5 | 7.8 | 5.8 | 1.9 | 1.0 | 0.6 |
| 2000 | 6.7 | 7.8 | 5.8 | 1.8 | 1.1 | 0.6 |
| 2001 | 7.0 | 7.9 | 5.9 | 1.8 | 1.3 | 0.7 |
| 2002 | 7.3 | 8.0 | 5.9 | 1.7 | 1.6 | 0.8 |
| 2003 | 7.6 | 7.9 | 5.9 | 1.7 | 1.8 | 0.9 |
| 2004 | 7.9 | 7.9 | 5.8 | 1.7 | 2.1 | 0.9 |
| 2005 | 8.1 | 7.9 | 5.7 | 1.7 | 2.4 | 1.0 |
| 2006 | 8.2 | 7.9 | 5.5 | 1.7 | 2.6 | 1.0 |
| 2007 | 8.4 | 7.9 | 5.4 | 1.8 | 2.9 | 1.0 |
| 2008 | 8.6 | 7.9 | 5.2 | 1.8 | 3.0 | 1.0 |
| 2009 | 8.7 | 7.8 | 4.9 | 1.8 | 3.2 | 1.0 |
| 2010 | 8.9 | 7.8 | 4.7 | 1.8 | 3.3 | 1.0 |

of estimating the traffic flow distribution of different car types is required. This paper discusses a possible method for combining all available data to derive such estimates, and applies these new estimates to some recent accident data.

We have chosen a Bayesian approach to modelling flow data in order to coherently quantify the uncertainty in traffic flow, and allow this uncertainty to be propagated into models for accident data. Combining three sources of data has also resulted in a complex multivariate posterior distribution which requires MCMC methods to solve.

## 2. Data

### 2.1. Traffic data

The Department for Transport collects and analyses traffic count data on a large selection of roads in Great Britain. These counts are combined with road network lengths in order to estimate the total number of vehicle kilometres travelled each year. Overall, there are approximately 50,000 km of Motorway and A roads and 338,000 km of Minor (B, C and unclassified) roads (Department for Transport, 2010). Traffic densities vary considerably over different types of roads, different areas and different vehicle types, so detailed data are collected automatically and manually at a large number of sites across Great Britain. Traffic flow, measured in vehicle kilometres (where one vehicle-kilometre is one vehicle travelling one kilometre), is the product of the average daily flow (measured in vehicles and calculated from the traffic count) and the length of the road on which the daily flow was based. Due to the nature of the counting mechanisms, traffic flow can be approximately disaggregated by time, month, road type, region and vehicle type, however a disaggregation by different car types is not available.

The data used in this research are car traffic flow from 1999–2010 on different road types, as shown in Table 1.

In order to gain a useful flow measure which can be disaggregated by car size, additional information is required.

### 2.2. Registered vehicle data

The Driver and Vehicle Licensing Agency (DVLA) holds information on each registered vehicle in Great Britain, including the make and model of the vehicle. This make and model information can be used to categorise the British car fleet into six subgroups by car size: minis and superminis; small saloons; medium saloons; large saloons; 4 × 4s and people carriers; and sports cars. The data that have been used in this research are the number of registered vehicles by car type and year from 1999–2010, shown in Table 2.