



Department  
for Transport

# Traffic Statistics Methodology Review Overview

**Moving Britain Ahead**



July 2018

The Department for Transport has actively considered the needs of blind and partially sighted people in accessing this document. If you have other needs in this regard please contact the Department.

Department for Transport  
Great Minster House  
33 Horseferry Road  
London SW1P 4DR  
Telephone 0300 330 3000  
General enquiries <https://forms.dft.gov.uk>  
Website [www.gov.uk/dft](http://www.gov.uk/dft)



© Crown copyright 2018

Copyright in the typographical arrangement rests with the Crown.

You may re-use this information (not including logos or third-party material) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or e-mail: [psi@nationalarchives.gov.uk](mailto:psi@nationalarchives.gov.uk).

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned.

# Contents

Foreword	4
User Feedback	4
Acknowledgments	4
Executive summary	5
1. Current Methodology	6
2. Methodology Projects	7
A. Sample Stratification Project	7
B. Validation and Imputation methods for dealing with missing and invalid data	10
C. Quarterly Traffic estimation methods	12
D. Annual Traffic estimation methods	15
3. Alternative data sources	17
E. Alternative road side data sources	17
F. Alternative non-road-side data sources	18
G. Data sharing with other organisations	19
4. Implementation	21

# Foreword

The Department for Transport's Road Traffic Statistics Team have conducted a review of the traffic estimates for Great Britain. This document gives an introduction to the review, an overview of the projects contained within it, and their conclusions. It was first published in August 2015 and has been updated to include final outcomes in July 2018.

The aim of the review was to seek opportunities for innovation and efficiencies in the production of traffic statistics, without degrading their quality in terms of accuracy and reliability, timeliness and meeting user needs. The focus of the review was two-fold:

- A detailed review of the current methodology for producing quarterly and annual traffic estimates, which includes processing of raw data, methodological improvements, plus any efficiencies in sampling and analysis.
- Investigating alternative data sources and, if appropriate, determining robust methods for their use in producing traffic estimates.

Throughout this document the term "current methodology" refers the methodology that was in use until the publication of 2017 Road Traffic Estimates in July 2018.

## User Feedback

We are keen to receive user feedback on the issues covered in this document. This can be given via the Road Traffic Statistics Team inbox: [roadtraff.stats@dft.gov.uk](mailto:roadtraff.stats@dft.gov.uk).

## Acknowledgments

DfT is grateful to Charles Lound and Jim O'Donoghue from the Office for National Statistics (ONS) Methodology Advisory Service (MAS) for their input and helpful advice, and to the UK Statistics Authority for providing this support for the project via the Quality Improvement Fund.



# Executive summary

The Department for Transport publishes annual and quarterly estimates of road traffic on Great Britain's roads.

The methodology used to produce these estimates was last reviewed in a 2007 National Statistics Quality Review<sup>1</sup>. Since then, new data sources have become available, more use has been made of video traffic counting methods, DfT has conducted a review of uses and users of traffic statistics, and further IT improvements have been made to the statistical processing systems.

As a result, it is believed that there is potential to deliver innovation and efficiencies when estimating road traffic in Great Britain. However, the key challenge is to do this while ensuring robust estimates continue to be produced that meet user needs.

The review covered a number of areas, split into two topics, as set out below.

## Methodology:

- How the road samples are stratified.
- Imputation methods.
- Quarterly traffic estimation methods.
- Annual traffic estimation methods.

## Alternative data sources:

- Alternative methods of performing short-term road-side traffic counts.
- Use of non-road-side data sources, such as GPS data and MOT data.
- Exploring opportunities for entering data sharing agreements with organisations that collect traffic data and how these could be integrated with DfT's data sources.

This document provides an overview of the projects above and their conclusions. Comprehensive reports including detail on statistical methods and analyses have been published online alongside this document as the projects were completed<sup>2</sup>.

## Implementation

Section 5 provide a summary of the improvements that were implemented in the publications "*2017 Road Traffic Estimates for Great Britain*" and the "*Provisional road traffic estimates for the year ending March 2018*" in July 2018.

Feedback and questions are welcomed via the Road Traffic Statistics Team inbox: [roadtraff.stats@dft.gov.uk](mailto:roadtraff.stats@dft.gov.uk).

---

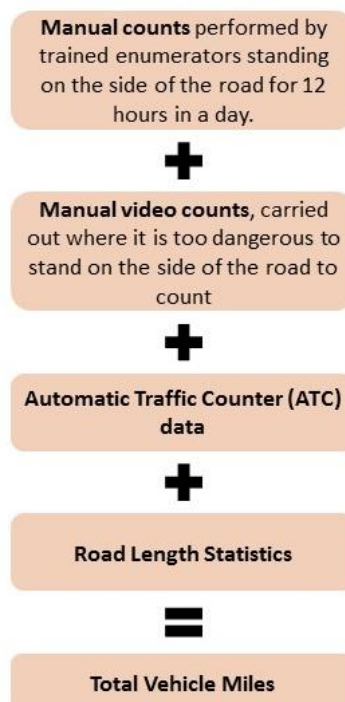
<sup>1</sup> Available via [www.ons.gov.uk/ons/guide-method/method-quality/quality/quality-reviews/theme/travtrans/nsqr-49/index.html](http://www.ons.gov.uk/ons/guide-method/method-quality/quality/quality-reviews/theme/travtrans/nsqr-49/index.html)

<sup>2</sup> Available at [www.gov.uk/government/statistics/road-traffic-statistics-methodology-review](http://www.gov.uk/government/statistics/road-traffic-statistics-methodology-review)

# 1. Current Methodology

- 1.1 Road traffic estimates are currently published for Great Britain on an annual and quarterly basis as well as an annual publication of street-level traffic data via the traffic counts website.
- 1.2 Quarterly estimates are calculated on a panel sample approach, with traffic data collected continuously from a national network of around 200 Automatic Traffic Counters (ATCs) which count flows and classify by vehicle type.
- 1.3 Annual estimates are currently based on around 8,000 manual counts, where trained enumerators count traffic by vehicle type over a 12 hour period. This data is combined with the ATC data and road lengths statistics to produce the number of vehicle miles travelled each year by vehicle type, road category, and region.
- 1.4 For major roads (motorways and 'A' roads) a rolling-Census approach is taken to manual counts, and the large number of counts enable detailed road-level traffic estimates to be produced for these road types. The 2012 review of uses and users of traffic statistics established that the existing traffic data for each junction-to-junction link of the major road network was essential for a number of customers of road traffic data, in particular for producing road traffic forecasts and road transport emission statistics.
- 1.5 For minor roads ('B', 'C', and unclassified roads) a panel sample approach is taken, whereby the same roads across Great Britain are counted each year (over 4,000 locations). This enables robust national level minor road traffic estimates to be produced.
- 1.6 More detailed explanations of the current methods used to produce traffic estimates, from the above data sources, are available via the websites below.

## Current data sources



## Links

Quarterly and annual statistics:

[www.gov.uk/government/collections/road-traffic-statistics](http://www.gov.uk/government/collections/road-traffic-statistics)

Street level traffic counts (traffic estimates for each link on the major road network):

[www.dft.gov.uk/traffic-counts/](http://www.dft.gov.uk/traffic-counts/)

More detailed guidance in the road traffic section of this note:

[www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/270083/contents-page.pdf](http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/270083/contents-page.pdf)

## 2. Methodology Projects

### A. Sample Stratification Project

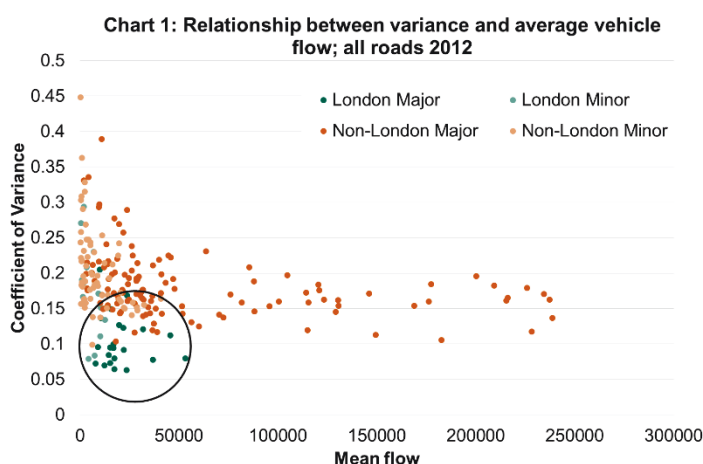
- 2.1 A fundamental component of the traffic estimation process is the stratification of the sample of Automatic Traffic Counters. The aim of stratification is to minimise sampling error by grouping together roads with similar traffic patterns.
- 2.2 The stratification is also used when producing both expansion factors and growth factors for calculating annual traffic estimates. Therefore, it needs to take into account both variation in traffic flows within the year and variation between years.
- 2.3 The sample stratification project explored whether the current stratification satisfactorily reflects the traffic patterns on different types of roads. A summary of the analyses, their findings, and the conclusion is given below<sup>3</sup>.

#### Current Methodology

The current sample is classified into 'Expansion Factor Categories' (EFCats), which split roads into 22 categories based on their characteristics such as road classification, area type, and traffic flow level. This stratification was introduced in the early 2000s.

#### Analyses: Within Year Approach

- 2.4 The project looked at the effects of certain road attributes (e.g. road classification) on fluctuations in traffic flow within the year. To do this, a measure of variation in flow was calculated that was independent of the size of traffic flow (a coefficient of variance). This meant that an analysis could be carried out to test for correlations and significant differences between the variation on different roads without being confounded by the size of the traffic flow. The analysis looked at different types of variation within the year: seasonal variation; 12 hour daytime to 24 hour variation; and weekend to weekday variation. In combination with statistical tests, graphical representations of trends were inspected for distinct patterns.
- 2.5 For example, the cluster of sites at the bottom left hand corner of Chart 1 are largely London sites with low levels of variation, suggesting that London is something of an



<sup>3</sup> A full report on this project is available at [www.gov.uk/government/statistics/road-traffic-statistics-methodology-review](http://www.gov.uk/government/statistics/road-traffic-statistics-methodology-review)

anomaly and, therefore, perhaps should have its own stratum. This was supported by a significant difference between expansion factors for roads inside and outside London.

- 2.6 As a result of the analyses of traffic flow variation within year, four factors were found to be significant;
- whether the road was inside or outside London;
  - road classification (motorway, 'A' road etc);
  - whether the road was in an urban or rural area (for 'A' roads and Minor roads);
  - the level of flow on the link (for 'A' roads only);
- 2.7 To review this analysis, the ONS Methodological Advisory Service (MAS) ran a regression with fixed effects that took into account the effect that day of the week and month have on the expansion factors. The proposed strata were supported by the model.

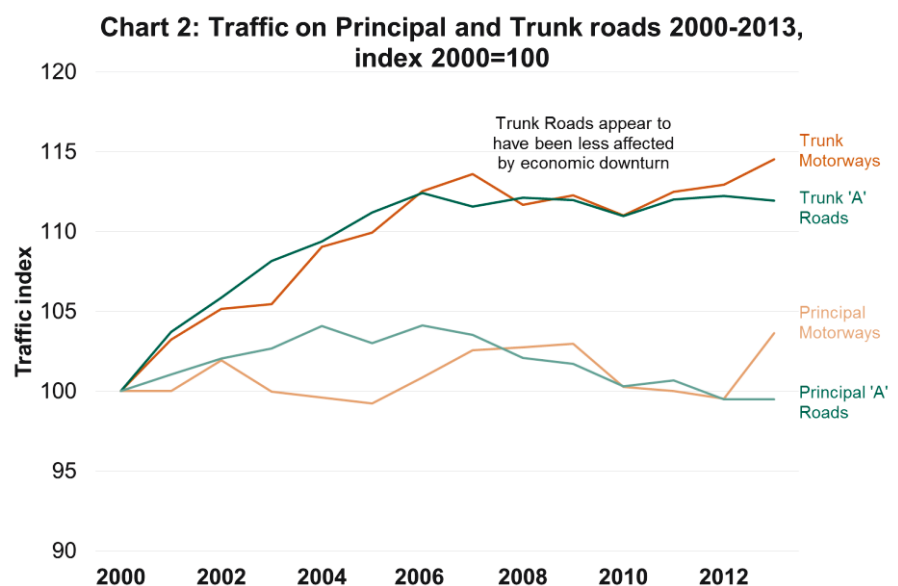
### Analyses: Between Year Approach

- 2.8 The next step in the analyses was to look at variation between years. An examination of historical trends established that there are clear differences in growth between different road types meaning that using one growth factor across the board would be unsuitable. Therefore, a similar approach to the within year analysis was taken to establish suitable categories based on between year variations, using correlation analysis, ANOVA tests, and regression models. In addition, visualisations based on historical traffic patterns for different road types were inspected for similarities in trends, as in Chart 2.

- 2.9 Different characteristics were found to affect variation between years than those for variation within the year. For example, level of flow on the link was not found to be a significant factor for 'A' roads, but road management status (whether a road was managed by Highways England or a Local Authority) was.

### Conclusion

- 2.10 A revised stratification was recommended (see Box 1). The new stratification consists of 18 categories that can be aggregated up into two different groups of strata: one for use in estimating within year variation in traffic flows (calculating expansion factors); and the second for change across years (calculating growth factors and provisional traffic estimates).





## Box 1: Road Traffic Statistics Stratification

- 01 Trunk Motorway
- 02 Principal Motorway
- 03 Trunk Urban 'A' roads <20,000
- 04 Trunk Urban 'A' roads >20,000
- 05 Trunk Rural 'A' roads <20,000
- 06 Trunk Rural 'A' roads >20,000
- 07 Principal Urban 'A' roads <20,000
- 08 Principal Urban 'A' roads >20,000
- 09 Principal Rural 'A' roads <20,000
- 10 Principal Rural 'A' roads >20,000
- 11 Outer London 'A' roads
- 12 Inner London 'A' roads
- 13 London 'B' roads
- 14 London 'C' and Unclassified roads
- 15 Urban 'B' roads
- 16 Urban 'C' and Unclassified roads
- 17 Rural 'B' roads
- 18 Rural 'C' and Unclassified roads

### Expansion Factor Stratification

- 01 Motorways
- 02 Urban 'A' roads with flows over 20 thousand vehicles a day
- 03 Urban 'A' roads with flows under 20 thousand a day
- 04 Rural 'A' roads with flows over 20 thousand vehicles a day
- 05 Rural 'A' roads with flows under 20 thousand a day
- 06 Outer London 'A' roads
- 07 Inner London 'A' roads
- 08 Urban Minor roads
- 09 Rural Minor roads
- 10 London Minor roads

### Growth Factor Stratification

- 01 Trunk Motorway
- 02 Trunk Urban 'A' roads
- 03 Trunk Rural 'A' roads
- 04 Principal Urban 'A' roads & Principal motorways
- 05 Principal Rural 'A' roads
- 06 London 'A' roads
- 07 London 'B' roads
- 08 London '*C' and unclassified* roads
- 09 'B' Roads outside London
- 10 'C' and unclassified roads outside London

## B. Validation and Imputation methods for dealing with missing and invalid data

2.11 DfT Statisticians and an external methodological expert from the ONS reviewed the current imputation methodology.

### Current imputation method

**Automatic Traffic Count (ATC)** data goes through an automatic and manual validation check. Hourly flows that fail the validation test are removed and replacement values are imputed by copying forward the previous year's flows on the same day and week, thus giving a "normal" flow for that site. There are two main reasons why ATC data is imputed;

- because the ATC is faulty and producing no data or incorrect data that needs to be imputed,
- the ATC is functioning correctly but because of road works, or an event, there are extreme traffic levels on the road being sampled.

Quarterly and annual growth rates are calculated using these flows, as are expansion factors. Therefore these events which cause anomalous traffic data are discarded and replaced with imputed data to prevent over-inflating the effect of a one-off incident on total traffic estimates. Validation tests are carried out for each vehicle type independently in order to reduce the amount of imputation necessary.

For the **manual count** data, automatic and manual checks are carried out to identify invalid counts. Abnormal counts are omitted from the calculation of annual traffic estimates. For these road links with missing or invalid data, an estimation of the flow is calculated by using the previous year's annual traffic flow data and a 'growth factor' calculated from ATC data.

Growth factors are also used for roads which were not counted in a given year. Some manual counts on major roads are done annually, others at 2, 4 and 8 year intervals. For those done less frequently than annually, the most recent manual count is converted to a traffic estimate for the year in question through the application of an expansion factor. The previous year's traffic estimate then has a growth factor applied, for the relevant growth category, to grow the estimate from the year it was counted to the year in question.

2.12 The review of this method identified no major changes for the validation and imputation of manual count data. However, for the validation and imputation of ATC data, three main areas were identified for further investigation. A summary of the improvements for each of these three areas is provided below.

### 1. Identification of invalid automatic traffic counter data

2.13 As part of the ongoing improvements, the traffic statistics team regularly review the validation checks that are used to identify invalid traffic count data. In addition to these regular refinements, a new database system has been implemented that includes new processes for identifying invalid data from the automatic traffic counters.

- 2.14 The new system includes checks based on error codes and other information provided by the counter, as well as checks highlighting data that is an outlier when compared to historic data from the site. Once these automated checks have run, a visual check is made of the data to identify any other anomalous results.
- 2.15 During 2017, the new validation system was run on the data from ATCs for 2012 onwards. The efficiencies from the improved system have resulted in a significant decrease in the time taken to validate the DfT ATC data.

## **2. What, if any, historically anomalous data should be imputed?**

- 2.16 The review considered whether all types of invalid ATC data should be removed and replaced as part of the imputation process. Data from faulty ATCs will continue to be imputed. However, traffic data for abnormal events, which in the current method are regarded as invalid data and consequently imputed, can be considered to be random events, and representative of similar situations occurring every day across the road network. Therefore, the review proposed dealing with these abnormal flows as follows:
- Abnormal events with over 48 hours duration (such as long-term road works) are not random; they are persistent and exist over more than one day. Their presence or absence in the relatively small number of sites with ATCs is unlikely to be typical of the wider road network. Therefore, it would not be right to take these flows into account and these flows will continue to be imputed.
  - Abnormal events of 48 hours or less duration, are proposed to have their flows retained and not impute them, with the exception of:
    - When calculating expansion factors (for the annual estimation process). An expansion factor value is the 'annual average 24 hour flow' divided by the '12 hour flow on a particular day'. If any of the hours within the '12 hour flow' part of this calculation are part of an abnormal event, then they should be imputed. In that way, an abnormal ATC flow on a particular day does not impact on every manual count taking place on that day. However, it is taken into account in the annual total to which the day's count is being scaled.
- 2.17 The proposed approach is that short term abnormal events such as road works of 48 hours or less duration and cycling events will not be imputed for grossing to annual totals, but will be imputed for the calculation of expansion factors. This is under consideration for future implementation.

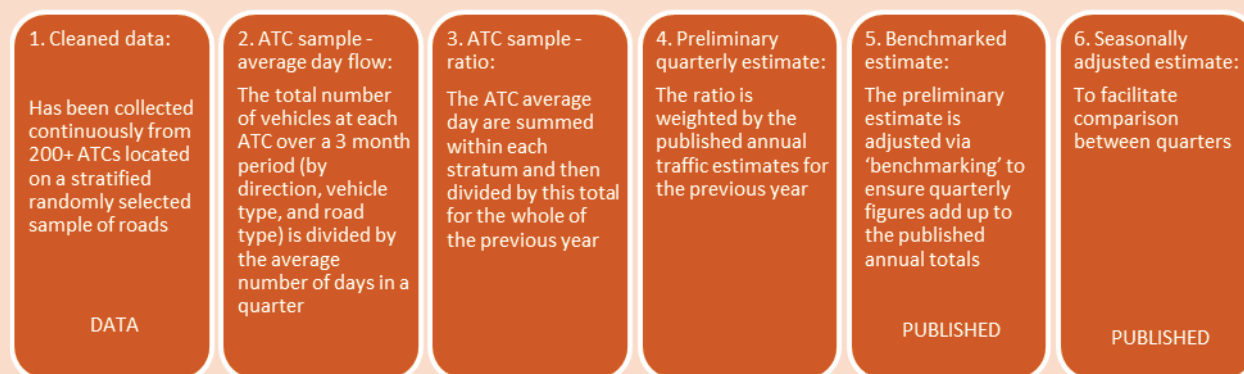
## **3. How to impute**

- 2.18 It was proposed that the imputation process, for each vehicle type, should be changed to the following three steps:
- 1 To use the average flow for each vehicle type recorded on the same weekday, site and hour for the preceding and following weeks, provided they do not fall on a bank holiday and the data is valid;
  - 2 If only one of the preceding or following weeks is suitable, use the flow from that week;
  - 3 If neither the week preceding nor following is suitable, use the flow on the same weekday, site, direction and hour in the preceding year, uprated by the corresponding change in average traffic flows for the same weekday and hour for other sites in the same stratum.
- 2.19 During 2017, this method was applied to the data from ATCs for 2012 onwards.

## C. Quarterly Traffic estimation methods

**Figure 1: Current methodology**

The quarterly methodology uses the ATC data collected each quarter. The main stages to estimating quarterly road traffic are outlined in the diagram below, and further technical detail is provided in the Quarterly Traffic estimation methods report<sup>4</sup>.



2.20 All aspects of the quarterly calculation process (as set out in Figure 1) were reviewed by DfT Statisticians and an external statistical expert and areas for potential further improvement were identified to the stages:

- sample ratio calculation
- weighting
- benchmarking

2.21 A summary of the changes and the main findings are set out below<sup>4</sup>.

### Sample Ratio

2.22 The sample ratio is the first clear initial indication of the change in traffic that has been observed at the sample ATC sites, before these estimates are weighted, benchmarked, and seasonally adjusted. The ratio is currently produced by using a static base year (1999) against which data for all other quarters are compared (step 3 in Figure 1 above). The use of a static base year introduces inflexibility to the sample, as it requires the same sample of locations to have traffic data in the static base year and in the current quarter. This is problematic as the ATC sample can change over longer periods, for example due to road layout changes. To date, any such changes to the ATC sample have been compensated for by using a combination of calculated adjustments and imputation.

2.23 A solution to this is to calculate the ratio between the current quarter and the data for the same set of ATCs for the whole of the previous year. This allows for greater flexibility in the ATC sample, via new ATCs to be brought into the sample to replace faulty equipment, either on a temporary or a permanent basis depending on the severity of the fault.

### Proposed Ratio

Current Quarter's Flow : Equivalent figures for all 4 quarters in the previous year

<sup>4</sup> A full report on this project is available at [www.gov.uk/government/statistics/road-traffic-statistics-methodology-review](http://www.gov.uk/government/statistics/road-traffic-statistics-methodology-review)



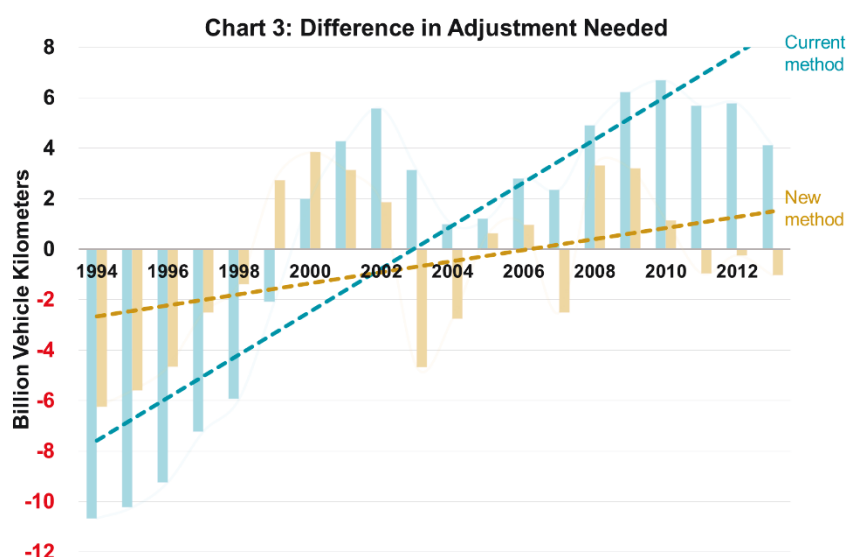
## Weights

- 2.24 Weights are applied to the ATC sample ratios to gross up the sample data collected by the ATCs and produce preliminary traffic estimates that are representative of all roads in Great Britain (step 4 in Figure 1).
- 2.25 These weights are based on the sum of the final annual estimates from 1993 to the latest year for which annual estimates are available. These are applied to all years of data, resulting in preliminary estimates that change each year and are never final. The proposed improvement brings traffic estimates in line with best practice by moving to the use of weights calculated from the latest year for which final annual estimates are available.

## Testing

- 2.26 Both of these proposed improvements were assessed by calculating new preliminary traffic estimates using the new methods, and then measuring how close these preliminary estimate figures are to the published final annual road traffic estimates. This measurement was compared with the equivalent figures for the current method.

Chart 3 shows the differences between the preliminary estimates and the final annual figures for both the current and proposed new methods, it clearly shows that the new method requires less adjustment than the current method. Therefore, the proposed new method was recommended for implementation.



## Benchmarking

- 2.27 The phrase benchmarking is used here to refer to the process by which the preliminary quarterly estimates (based on weighted ATC data) are adjusted so that they add up to the final annual estimate (which combine ATC and manual counts data).
- 2.28 The current system uses a piece of software to do this called 'Inter' to do this, which is no longer supported. Two alternative methods were considered for the road traffic process: a regression method known as Cholette-Dagum; and a simpler proportional, ad hoc approach, where the percentage of traffic that each quarter contributed in that year is applied to the final annual estimates to produce the final quarterly estimates.

## Testing

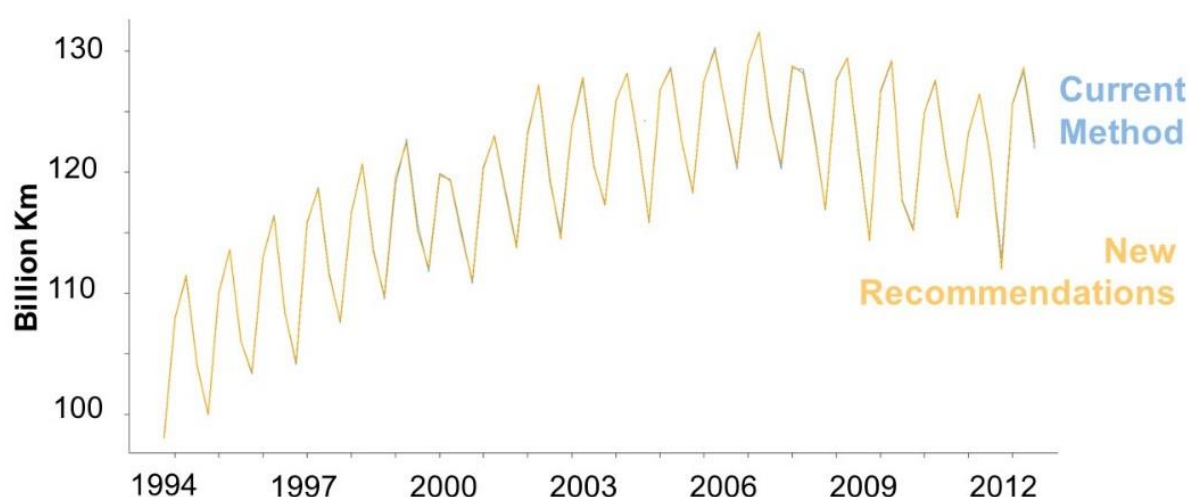
- 2.29 The investigation compared the two alternative methods. In general the differences were very small. For all motor vehicles, the largest difference was a fraction of a percent in any quarter, looking at both quarterly estimates and rolling annual totals. The largest differences (which were still marginal) were for buses and motorcycles on motorways, both of which have larger seasonal fluctuations, and make up a small proportion of total traffic.

	Ad-hoc proportional adjustment	Cholette-Dagum method
Advantages	<ul style="list-style-type: none"> <li>i. Can be implemented very easily and quickly in the same software as the rest of the process</li> <li>ii. Simple to explain and understand for traffic statistics users</li> <li>iii. Already in use elsewhere in DfT</li> </ul>	<ul style="list-style-type: none"> <li>i. Used by ONS for some time-series</li> <li>ii. Available as source code implemented in "R", so replicable and future proof</li> <li>iii. Takes into account autocorrelation between neighbouring quarterly deviations</li> </ul>
Disadvantages	<ul style="list-style-type: none"> <li>i. Less sophisticated than the Cholette-Dagum method.</li> <li>ii. Can produce unrealistic steps in the adjustment to be applied (see figure C)</li> </ul>	<ul style="list-style-type: none"> <li>i. Not easily explained, and not easily understood by most users of traffic statistics</li> <li>ii. Requires export and import of files between R and other programs</li> </ul>

## Conclusions

- 2.30 As a result of the testing, the new sample ratio and weighting methods were recommended. In addition, the ad hoc proportional benchmarking method produced similar traffic estimates to the Chollete-Dagum method. In light of the advantages of the ad-hoc proportional method in terms of the ease of integration into the rest of the process and simplicity, it was recommended that this benchmarking process be adopted.
- 2.31 A comparison was made of the quarterly estimates calculated via the current methodology and estimates calculated via the new recommendations. As can be seen in Chart 4, the differences between the final quarterly statistics were negligible.

**Chart 4: Impact of recommendations on traffic statistics**



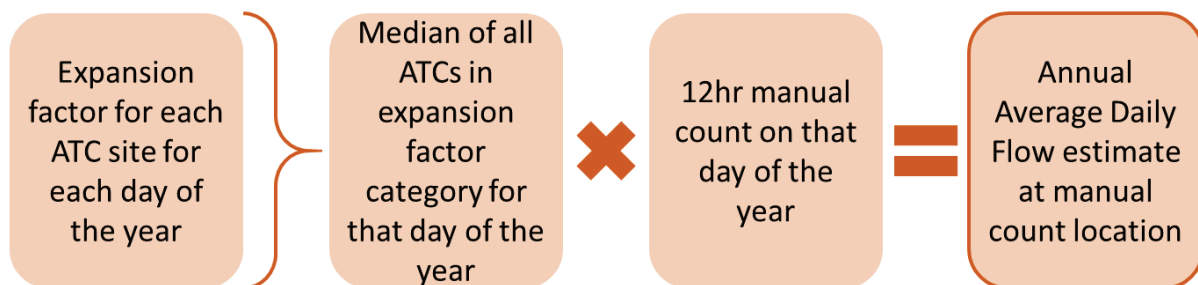
- 2.32 During 2017, the new recommendations were applied to the data from ATCs for 2013 onwards.

## D. Annual Traffic estimation methods

### Expansion Factor Methods

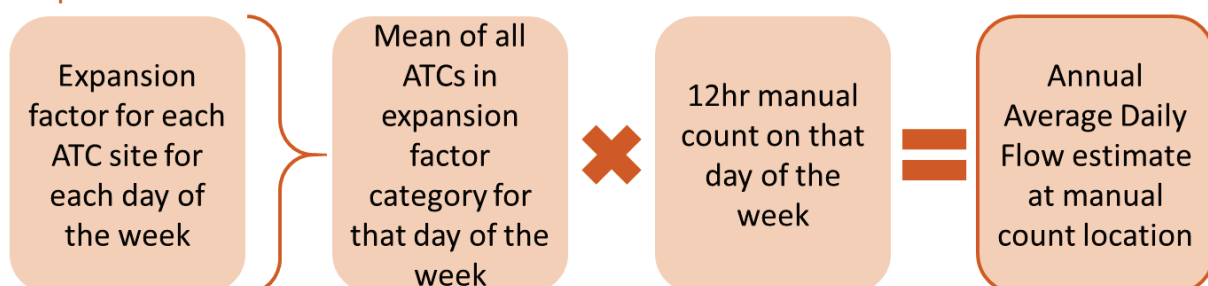
- 2.33 Manual counts are undertaken for a 12-hour period on one weekday in a neutral month during the course of the year. This is converted to an estimate of annual average daily flow at that location through the application of an expansion factor (for the expansion category to which the road location belongs). Currently the calculation of expansion factors is site weighted, that is, an expansion factor is calculated for each ATC site on a particular day and then the median of these factors is taken for the expansion factor category.

#### Current method



- 2.34 The previous National Statistics Quality Review looked at various methods for calculating expansion factors. As part of the Annual Traffic Estimation Methodology project, simple averages of the expansion factors using the different methods set out in this review were produced for one year. The discrepancies between them and the current method were inspected to establish how appropriate each method was.
- 2.35 From this analysis it was recommended that the site averaged day of the week expansion factors are used. This method smooths out some of the random fluctuations associated with the calculation of daily factors, particularly when there are very low flows. It also takes account of the fact that the expansion factors vary by day of week, tending to be lower on Fridays, and higher mid-week. It gives each ATC site equal weighting in the calculation of the expansion factors.
- 2.36 It was also recommended that the mean in each expansion factor category (rather than the median) be used. This is because the latter can be biased when dealing with a skewed distribution, such as traffic flows. Averaging over all observations for a particular day of the week will help minimise the influence of “unusual sites” or particular days.

#### Proposed new method



- 2.37 The next stage of this project is to run both the current method and proposed new expansion factor method with the same data and compare differences over time.

## Minor roads methods

- 2.38 The scale of the minor road network in Great Britain means it is not possible to count traffic on every stretch of road. Instead, a representative panel sample of minor road locations are counted each year. The change in traffic flows between two consecutive years is applied to the overall minor road traffic estimates for the previous year, to calculate regional and national estimates for the latest year. Every decade, a one-off benchmark review is carried out in order to recalibrate the traffic estimates on minor roads, to correct for errors in the sample that accumulate over time.
- 2.39 All aspects of the minor roads annual estimations methods were reviewed by DfT statisticians and an external statistical expert and areas for potential improvement were identified, including:
- the sampling method for benchmark and subsequent panel sampling,
  - using a different level of disaggregation for calculating the change in traffic levels between years,
  - pedal cycle traffic estimates on minor roads.

### Benchmark sample

- 2.40 All work on conducting the 2018 and 2019 benchmark of minor road traffic estimates will be subject to separate reports to be published at the completion of the benchmark. This is expected to be in 2020, and is outside the scope of this methodology review.

### Traffic growth estimation on minor roads

- 2.41 The review identified improvements to the method for calculating the change in traffic flows between two consecutive years.
- 2.42 The current methodology estimates, for each vehicle type, the median traffic flow for the panel sample, split road type. The change between these figures for the two years provides the traffic growth estimate for minor roads, split by vehicle type and road type.
- 2.43 The change identified was to add a regional split into the above calculation, and to use a mean rather than a median. These improve the robustness of the regional traffic estimates for minor roads.
- 2.44 During 2017-18, this method was implemented on the minor road data for 2016 and 2017.

### Pedal cycles

- 2.45 The work on improving the method for calculating the change in traffic flows between two consecutive years has been implemented for pedal cycle figures as well. Further work to improve the minor road traffic estimates for pedal cycles is under consideration.



## 3. Alternative data sources

- 3.1 A data source review is essential to ensure that the road traffic data collection continues to produce robust estimates of traffic, to keep up with technology developments, and to meet stakeholder needs.
- 3.2 The focus of the Alternative Data Sources project was to investigate and determine feasible options for collecting short-term count data, which would replace or supplement the current manual and video count sources. The 2012 review of uses and users of traffic statistics established that the existing traffic data for each junction-to-junction link of the major road network was essential for a number of customers of road traffic data, over 350,000 customers accessed this data online in 2017. Therefore, it is desirable that any alternative data sources will need to be able to produce this detailed geographic coverage either on their own or in combination with other sources.
- 3.3 Working with a range of colleagues, an extensive list of possible methods, and combinations of these, was drawn up. This was split into: feasibility trials of roadside technologies, desk-based analysis of non-roadside administrative data, and data sharing with other organisations<sup>5</sup>.

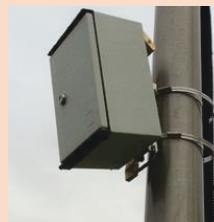
### E. Alternative road side data sources

#### Feasibility trials

- 3.4 The scope of the trials was to look at the feasibility of roadside technology options for traffic counting purposes. Technology that performed well would be considered for later in-depth trials. The aim was to determine how well each equipment type performs, including:
  - potential accuracy and reliability
  - the range of vehicles classified
  - the ease of deployment and operation, including health and safety issues
  - how well the equipment works in different environments such as heavy rain

#### Technologies types trialled

##### Radar



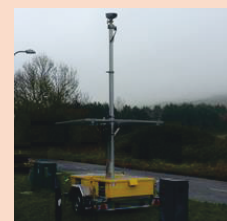
##### Rubber tubes



##### Video technology



##### Automatic Number Plate Recognition cameras



<sup>5</sup> A report on this project is available at [www.gov.uk/government/statistics/road-traffic-statistics-methodology-review](https://www.gov.uk/government/statistics/road-traffic-statistics-methodology-review)

### 3.5 The main findings from these feasibility trials were:

- Accuracy: Some of the technologies tested showed promising accuracy of total vehicle counts (within 1% deviation from the video count).
- Vehicle classification comparison: The analyses for different vehicle classifications found that none of the technologies provided detailed and accurate enough vehicle classification counts for DfT's statistical purposes.
- Sensitivity: all of the technologies tended to under count when compared to the manually enumerated video count. Some technologies were less sensitive than others but, even within the same type of technology, different brands had differing levels of sensitivity.
- Costs: some of the equipment could potentially be more expensive and less flexible than manual counts. Some types of equipment require local authority permissions for their temporary installations.
- Reliability: some of the technologies failed at some of the sites, but this was not always obvious until the analysis stage. Exploring boxplots by site and by technology revealed some set-up issues that had not been identified at the time of the trials.
- Security: some of the equipment was vandalised, some were more likely to be vandalised.

3.6 The analysis found that some of the technologies were within 1% deviation from the enumerator total, however, none provided the accuracy of vehicle classification required. Therefore, none were deemed suitable for the wholesale replacement of the existing automatic traffic counters and manual counts combination for DfT traffic statistics production.

3.7 The DfT team will explore the potential for the use of the total count data (i.e. without vehicle classification) from the technology types tested to estimate traffic statistics for all motor vehicles only.

3.8 The team will continue to investigate new technologies under development that may have applications for traffic counting in the future. Examples that have arisen since the completion of this project include new video analytics technologies, and drone or satellite images.

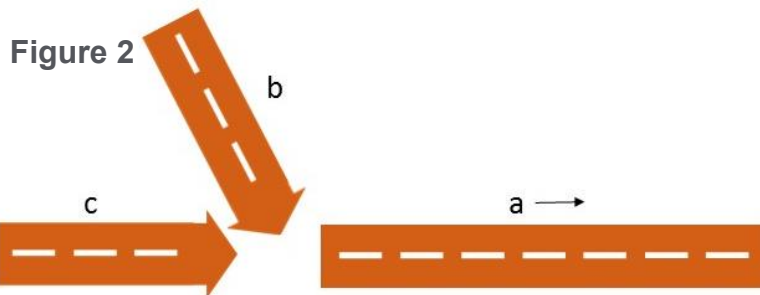
## F. Alternative non-road-side data sources

3.9 Non-roadside options for obtaining or synthesising traffic data have developed in recent years. It is likely that these options would still require observations of traffic from short- or long-term traffic counts, but they could reduce the sample size or frequency of these counts and therefore reduce costs. However, careful consideration will be needed as to the effect that fewer actual counts could have on the volatility of the overall estimates. That is, as you base estimates on fewer real counts the estimates become less reliable, so vigorous testing will be carried out before any methods are implemented.

### Synthesised traffic estimates

3.10 Synthesised estimates would use data on traffic flows from sources such as in-vehicle GPS (or, in the longer-term, possibly from mobile phone data) to model how traffic flows around the road network. From these types of data, the proportion of traffic that flows onto a given road link from its feeder roads can be obtained and, when combined with observed count data on those feeder roads, a traffic estimate for the road link can be generated.

3.11 For example, in Figure 2 the traffic flow for link a in the direction shown, would be derived using the percentages of vehicles that join from links b and c. These percentages would then be applied to counts on links b and c to get an estimate of traffic on link a.



3.12 This approach was trialled for a sample of motorway and 'A' road links, but the analyses determined that this methodology was not sufficiently reliable for large-scale implementation at the current time.

3.13 The team will continue to investigate if it could be used for individual road links, such as those where it is difficult or dangerous to conduct traffic counts.

### Administrative data

3.14 Administrative data sources for traffic, such as the mileage data from MOT tests, could be used to estimate overall traffic levels. Whilst MOT test data does not include all vehicle types (such as HGVs, buses, pedal cycles), nor does it provide information about where the vehicle is being driven, it does give a wealth of other information such as propulsion type, detailed vehicle types, age of vehicle, etc.

3.15 DfT statisticians have been working with the Driver & Vehicle Standards Agency to develop the MOT test dataset so that it can be robustly analysed for this purpose. That project is outside this Methodology Review.

3.16 The road traffic statistics team are working closely with the statisticians leading the project developing the MOT test data, with experimental statistics currently planned for 2019.

## G. Data sharing with other organisations

3.17 The annual traffic estimates currently use data from a subset of the ATCs owned by the Scottish Government to estimate traffic growth on certain motorway and 'A' roads in Scotland. The data sharing element of the Methodology Review reviewed this method and explored other opportunities with organisations that collect traffic data and how these data sources could be integrated with DfT's data sources.

3.18 A thorough investigation was conducted to determine whether integration of Highways England's ATC data is possible<sup>6</sup>, whilst ensuring that traffic statistics remain robust and the time series consistent.

<sup>6</sup> Highways England ATC data is available via: <http://webtris.highwaysengland.co.uk/>

- 3.19 The project's conclusion was that Highways England's automatic traffic counter data would be a useful additional data source, subject to obtaining the resources required to match the network of counters to the road network used for the DfT road traffic estimate calculation.
- 3.20 Since this project was completed, DfT have included Highways England, Transport Scotland, and Transport for London ATC data in the publication "*2017 Road Traffic Estimates for Great Britain*".



## 4. Implementation

- 4.1 The review was completed in 2016. It identified a number of proposed improvements to the data and methodology used to estimate road traffic in Great Britain, as set out in the previous chapters of this document<sup>7</sup>.
- 4.2 A summary of the implemented improvements is set out below. During 2017-18, these were implemented in the calculation of the publications *"2017 Road Traffic Estimates for Great Britain"* and the *"Provisional road traffic estimates for the year ending March 2018"*.

### Quarterly provisional estimates

- 4.3 The publication *"Provisional road traffic estimates for the year ending March 2018"* has been produced using:
- the new stratification of the ATC data (see page 7)
  - the improved validation and imputation of ATC data (see page 10)
  - the new quarterly estimation methods (see page 12)

### Final annual estimates

- 4.4 The publication *"2017 Road Traffic Estimates for Great Britain"* has been produced using:
- the newly analysed ATC data (as set out in paragraph 4.3 above)
  - the improved expansion factor calculation (see page 15)
  - the improved minor road traffic calculation (see page 16)
  - ATC data from Transport Scotland, Highways England, and Transport for London (see page 19)

---

<sup>7</sup> The technical reports on the review are available at [www.gov.uk/government/statistics/road-traffic-statistics-methodology-review](http://www.gov.uk/government/statistics/road-traffic-statistics-methodology-review)