Road Traffic Statistics Review: An Introduction

Department for Transport

Introduction

The Department for Transport's Road Traffic Statistics Team are conducting a review of the traffic estimates for Great Britain. This note gives an introduction to the review, and progress on the initial stages of work. User feedback is welcomed at all stages of the review, and contact details are provided below.

The aim of the review is to ensure robust traffic estimates continue to be produced, seeking innovation and greater value for money in their production while protecting user needs. The focus of the review is two-fold:

- A detailed review of the current methodology for producing quarterly and annual traffic estimates, which includes both 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.

CONTENTS:

An update to this note is planned for 2015 to detail the results of the next steps identified within this document, as well as the work on the other projects of the Methodology Review that will begin in 2015.

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.gsi.gov.uk. Any proposals for changes to the Road Traffic Statistics publication and datasets will be subject to a consultation process at the end of this project.

Acknowledgments

DfT is grateful to Charles Lound and Jim O'Donoghue for their input and helpful advice, and to the Office for National Statistics for providing this support for the project via the Quality Improvement Fund.

Current Methodology

Road traffic estimates are currently published for Great Britain on an annual and quarterly basis (https://www.gov.uk/government/collections/road-traffic-statistics), as well as an annual publication of street-level traffic data via the traffic counts website (http://www.dft.gov.uk/traffic-counts/).

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). In addition to counting flows, the ATCs record some of the physical properties of passing vehicles which are used to classify by vehicle type.

Annual estimates are currently based on around 7,000 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. For major roads (motorways and 'A' roads) a rolling-Census approach is taken; for minor roads ('B', 'C', and unclassified roads) a panel sample approach is taken.

Links to more detailed documents on the methodology currently used can be found via the road traffic section of this guidance note:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/270083/contents-page.pdf

Contact details: Tel: 020 7944 3095
Email: roadtraff.stats@dft.gsi.gov.uk

Overview of the Review

The traffic estimates methodology was last reviewed in a 2007 National Statistics Quality Review. Since then, new data sources have become available (such as mileage data from MOT tests), 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 greater value for money 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.

Scope

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

Methodology:

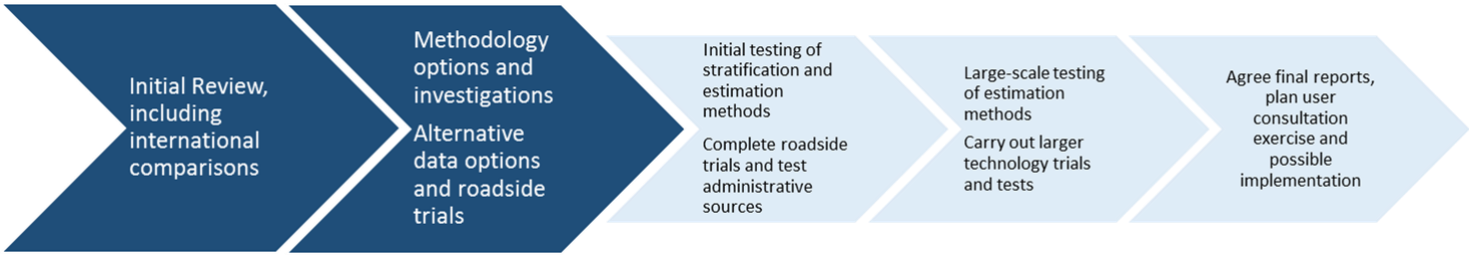
- How the samples are stratified.
- Imputation methods for dealing with missing and invalid data.
- Quarterly traffic estimation methods.
- Annual traffic estimation methods.

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 data sources could be integrated with DfT's data sources.

Timescales

The review will run until end March 2016, and is split into stages as set out below. As at end 2013-14 the review has covered the first section and was part way through the second section, as set out in dark blue.



Progress on initial projects

The rest of this report gives information on the initial investigations of the review, carried out up to end 2013-14. Specifically, this work has explored:

- How the Automatic Traffic Counter sample is stratified
- The quarterly traffic estimation methods
- Identification of alternative data sources. This includes methods of performing roadside counts, as well as potential administrative data sources.

In each of these projects, further work is needed to determine the feasibility of implementing the potential improvements identified. This work set out in each section, with the timescales for carrying out the next steps. An update to this note is planned for 2015 to detail the results of these next steps, as well as the work on the other projects of the Methodology Review that will begin in 2015.

1. Automatic Traffic Counters: sample stratification project

A fundamental component of the traffic estimation process is the stratification of the sample of Automatic Traffic Counters. The aim of stratification is minimise sampling error by grouping together roads with similar traffic patterns.

The main aspect of the ATC sample stratification project is to explore whether the current stratification groups satisfactorily reflect the traffic patterns on different types of roads.

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. As ATCs had already been deployed based on the previous stratification, the existing ATCs within these EFCats are not a random sample. In particular, some of the EFCats have quite a small set of ATCs within them meaning that estimates for these categories can show a lot of variation. Over the latter part of the 2000s, an effort was made to remedy this issue, and so the Methodology Review is a timely opportunity to assess the effect of this work and identify any further improvements needed.

Various exploratory analyses have been conducted to determine what road characteristics are important with regards to grouping roads and the traffic on them. This has included investigating the level of variation in traffic observed at each ATC, as well as variations between day and night, variations between weekday and weekend, and seasonal variations.

This work is ongoing, but the initial analyses suggest that the existing stratification structure can be improved upon. Further analyses are planned for 2014-15 to determine what improvements are appropriate, and to estimate what the impact of these would be on the traffic estimates.

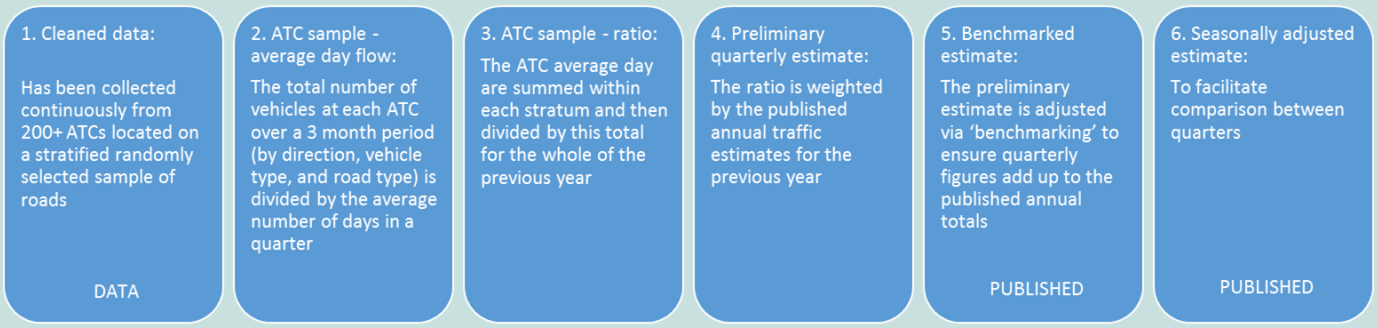
2. Quarterly estimation project

During 2013, the road statistics team moved the quarterly processing from Excel into SQL, which has provided the opportunity to make changes more easily to the methodology applied to quarterly estimation.

Current methodology

The current quarterly methodology uses the ATC data collected each quarter. The main stages to estimating quarterly road traffic are outlined in the diagram below, and in technical detail in Annex A.

Figure 1: Current process for quarterly traffic estimates



All aspects of the above process were reviewed by DfT Statisticians and an external statistical expert. Areas for potential further improvement were identified. They are:

A. The potential move away from a static base year to estimate the growth in traffic at ATC sites (step 3)

B. Updating the weights used to produce overall traffic estimates (step 4)

C. The technique used to constrain the quarterly results to the annual totals (benchmarking - step 5)

D. The application of seasonal adjustments (step 6)

The progress on these areas, and the next steps, are set out below.

A. Sample ratio

The sample ratio (step 3 in Figure 1) is the first clear initial indication of the change in traffic that has occurred 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. 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.

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 would allow for greater flexibility in the ATC sample with new ATCs coming into use after two or three years in place and non-functioning ATCs being removed immediately from the calculations either on a temporary or permanent basis depending on the severity of fault with the equipment.

Initial investigations tested this change alongside updating the weights, and the result of these combined changes suggests that they produce more accurate preliminary estimates, as they require smaller adjustments when being constrained to the finalised quarterly estimates (see technical annex B).

B. Weights

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).

These weights are currently 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 is to bring traffic estimates in line with best practice by moving to using weights calculated from the latest year only for which final annual estimates are available.

The initial investigation into using single-year weights alongside the updated ratio calculation (as in A above) suggests that they produce more accurate preliminary estimates, as they require smaller adjustments when being constrained to the finalised quarterly estimates (see technical annex B).

C. Benchmarking

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).

The current system uses a piece of software to do this called ‘Inter’, which was developed by the ONS in the late-1990s and uses interpolation. This software has been superseded, and a new approach is being developed by the ONS. Two alternative methods are being considered for the road traffic process: a regression method known as the Cholette-Dagum method (tested using the software package R); 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 (this can be implemented in any database/spreadsheet software package).

Inter, the Cholette-Dagum method, and the simple proportional method produced near identical results when the three systems were compared in the initial tests (see technical annex C).

A-C: Next Steps

The next steps are to investigate the effects of these improvements on the full quarterly series before determining if they are suitable for future implementation. This large-scale test of the possible changes to the quarterly methodology will occur during 2014-15.

D. Seasonal Adjustment

The quarterly traffic time series has a seasonal adjustment process applied to it to enable comparisons between quarters. This is a standard statistical technique, used in many other statistical series, such as GDP. Road traffic statistics use the standard software that is used across government, which was developed by the US Census Bureau. However, the team was using an older version of the software (called X-11 ARIMA) with a model that had not be reviewed for some time.

The package now recommended by the ONS is Win X-13 ARIMA, which offers a number of improvements including a Windows-based graphical user interface. A new seasonal adjustment model using Win X-13 ARIMA was developed with support from the ONS Time Series Analysis Branch. The impact on the overall numbers was very small.

The new model and updated software were adopted and implemented for the 2014 quarter one publication in May 2014.

3. Alternative data sources project

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, to meet stakeholder needs, and to deliver value for money.

Current data sources

Currently, road traffic estimates are based on

- counts performed by trained enumerators standing by the side of the road for a 12 hour period, known as manual counts
- a small number of counts performed by trained enumerators watching a video recording of traffic, known as manual video counts
- automatic traffic counter (ATC) data.

The mobility and flexibility of manual and video counts enables the Department's road traffic estimates to have the breadth of geographical coverage it has and to produce accurate and detailed vehicle classification.

The large number of manual and video counts on the major road network (motorways and 'A' roads - currently around 4,000 locations per year) 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. Since 2012, the use of this road-level data has grown. Over 200,000 customers accessed this data in 2013 via our traffic counts website (http://www.dft.gov.uk/traffic-counts/) or data.gov.uk (http://data.gov.uk/dataset/gb-road-traffic-counts). 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.

For minor road ('B', 'C' and Unclassified roads) traffic estimates, the same roads across Great Britain are counted each year (over 4,000 locations). This enables robust national level traffic estimates to be produced.

The ATC data provides continuous traffic count data for over 200 locations so that, at a national level, in-year traffic variations can be observed by road type and vehicle type.

The focus of the Alternative Data Sources project in the Methodology Review is 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 first stage of the project has been to investigate what options exist.

An extensive long list of possible count methods, and combinations of these, was drawn up. This was based on the expertise and experience within the Department, previous investigations into the issue by the team, and a review of UK and international research. It can be split into two groups: roadside and non-roadside data collection methods.

Alternative roadside methods

The options for roadside methods of collecting short-term traffic counts include: manual counts, video counts with/without some element of automatic processing, radar, rubber tubes, automatic number plate recognition (ANPR) cameras, and short-term counts using ATC-type equipment. Trials of several of the above options are planned to enable a comparison between these and existing methods.

Non-roadside traffic data

Non-roadside options for obtaining or synthesising traffic data have developed in recent years, and include:

- **Synthesised estimates,** which would use data on traffic flows from sources such as in-vehicle GPS (or, in the longer-term, from mobile phone data) to model how traffic flows around the road network.
- **Administrative data sources** for traffic, such as the mileage data from MOT tests, could be used to estimate overall traffic levels.

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. However, careful consideration will be needed as to the effect that fewer actual counts could have on the volatility of the overall estimates.

From this list, key candidates have been selected for further investigation during 2014-15. This selection was determined through discussions with internal experts, other authorities collecting traffic data (such as local authorities and the Highways Agency), and industry experts.

4. Next Steps

This note has provided an introduction to the Department for Transport's review of the traffic estimates for Great Britain, with an overview of the scope of the review and its timescales. The next steps for the review have been set out in the relevant sections of the document.

An update to this note will be provided in due course, and will cover the results from the further work identified throughout this document. If this review results in any proposals that change the Road Traffic Statistics publication and its datasets, they will be subject to a consultation process. Plans for implementation will then be made based on the results of the work and the consultation responses. If you wish to discuss any aspect of this work in the interim, please contact the team using the details below.

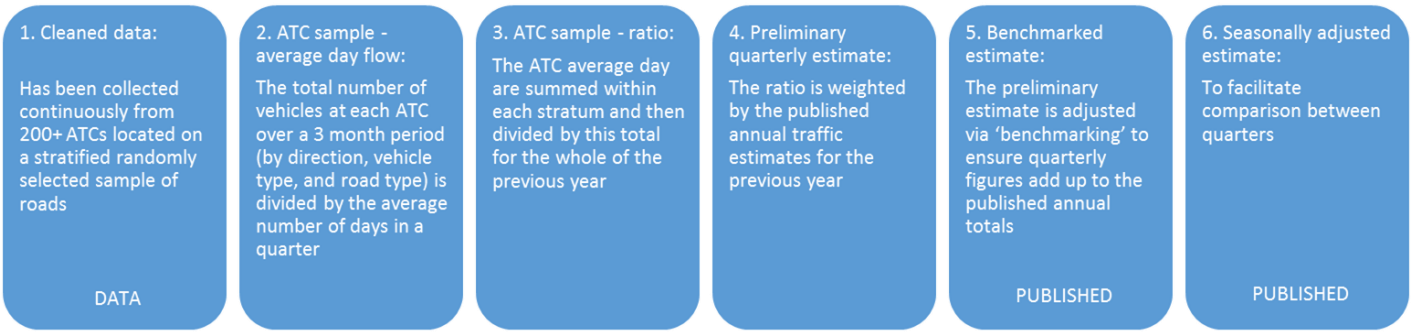
Contact details: Tel: 020 7944 3095
Email: roadtraff.stats@dft.gsi.gov.uk

Annexes

Annex A: Current quarterly method in detail

This annex sets out in detail the current methodology for estimating quarterly traffic, using data from a sample of automatic traffic counters (ATCs).

Figure 1: Current process for quarterly traffic estimates



For each vehicle type:

Step 2: If, for each ATC site i , we have ATC counts z_i , the first step is to calculate the average daily flow in the quarter q by dividing the ATC count by the average number of days in a quarter (91.25). These are then totalled up for all the ATCs within each EFCat e.

$$x_{eq} = \sum_{e} \frac{\sum_q z_i}{91.25}$$

Step 3: The next step is to create a ratio of the ATC daily totals for the current quarter, for EFCat e, to that for the annual total for 1999, multiplied by 4 to scale the quarter up to a year so that it can be compared to the annual total for 1999.

$$r_{eq}^{y,99} = 4 \times x_{eq} / \sum_{1999} x_{eq}$$

Step 4: These ratios r are then multiplied by the 1999 EFCat weight, added up for each road type (for all road types there is more than one EFCat), and then divided by the sum of the 1999 EFCat weights for that road type to produce a weighted ratio for each road category c.

$$r_{cq}^{w} = \frac{\sum_c W_e^{99} r_{eq}}{\sum_c W_e^{99}}$$

These are then used to create a 'Preliminary scaled up estimates' for this road category using the existing annual totals A_{cy}:

$$Q_{cq}^{P} = r_{cq}^{w} \frac{\sum_{y'=1993}^{y'=y(q)-1} A_{cy'}}{\sum_{q=1993q1}^{q=(y(q)-1)q4} r_{cq}^{w}}$$

The above formula is taking the total traffic for all years from 1993 to the previous year and scaling that by the sum of the weighted ratios for all the quarters of those years multiplied by the weighted ratio for the new quarter.

Step 5: This is then modified to produce the unseasonally adjusted publication estimates:

$$Q_{cq}^{F} = Q_{cq}^{P} + I_{cq} + \frac{1}{4} B_{cy(q)}$$

where I_{cq} is the adjustment output from the benchmarking program Inter, and

$$B_{cy(q)} = A_{cy} - \sum_{y} (Q_{cq}^{P} - I_{cq})$$

which is the extent to which the benchmarked quarterly preliminary estimates differ from the existing annual totals. For the latest year, where the final annual benchmark is not yet available, the Inter adjustment, I_{cq} is predicted by extrapolation from recent change. The latter adjustment, $B_{cy(q)}$ is zero for the current year.

Annex B: Ratio calculations and weighting

Calculating the ratios in the quarterly traffic estimates process is the first clear initial indication of the change in traffic that has occurred before these estimates are weighted, benchmarked, and seasonally adjusted. The weights are then applied to the sample ratios to ensure that the traffic estimates are representative of traffic in Great Britain across all vehicle and road types.

As part of the Methodology Review, the calculation/application of these steps in the quarterly process were reviewed.

Current method

The growth in traffic is based on a set of ratio calculations using the traffic counts from the ATC sites. Currently, this is calculated on a fixed base year, 1999, and all quarters are estimated based on a comparison to this base year. Once this ratio is calculated, it is then aggregated up to road type (motorway, rural 'A' roads, urban 'A' roads, rural minor roads, urban minor roads) using a weighted average, where the weights are based on final annual traffic estimates for 1999. To produce representative traffic estimates, the road type figures are weighted by the sum of the final annual estimates for 1993 to the latest year for which annual estimates are available.
These are the third and fourth steps in the calculation stages, as set out in Annex A.

The use of 1999 as a base year when calculating the ratios is potentially problematic as we get further away from that year. In addition, the ratio has to be based on common ATCs, otherwise it would include change simply due to change in the available ATCs rather than actual change in traffic. As a result, some missing ATC data is imputed when an ATC is no longer available or is introduced into the sample. The application of weights based on data for all years means that preliminary estimates can change each year, albeit by very small amounts.

Alternative Methods

The proposed changes are that:

a) The ratios are created for the current quarter over the ATC totals for a reference year just one or two years ago, i.e. for two years ago this would be:

$$r_{eq} = x_{eq} / \Sigma_{y-2} x_{eq}$$

The main benefit of this would be that each ratio would be likely to be based on a larger number of common ATCs.

b) Then weight this from annual estimates based on the most recent year. This may not be the previous year, as that would not necessarily be finalised in time, but it might be possible to work with two years back, so the preliminary quarterly estimate would then be:

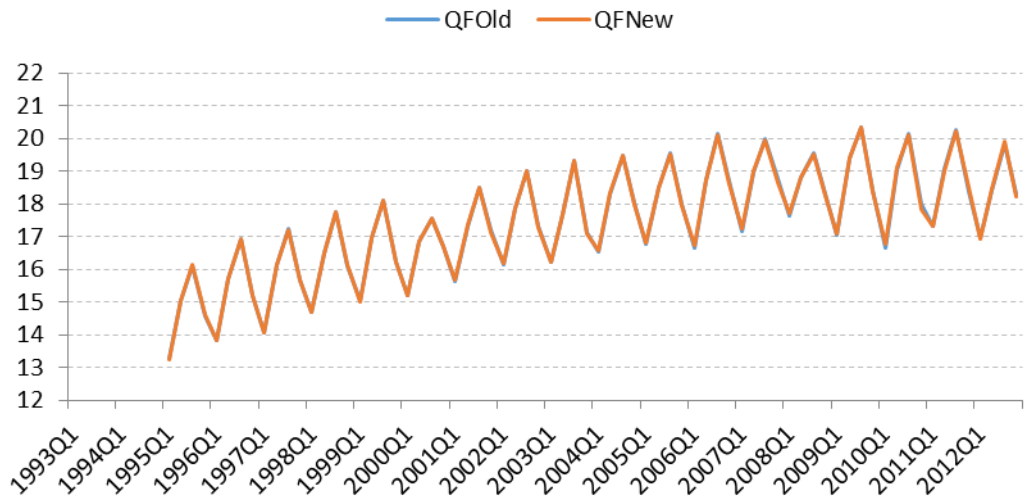
$$Q^{P}_{cq} = \sum_{e} W_{e,y-2} r_{eq}$$

Testing the new calculations

The new ratio calculations and the updated weights were compared to the current methodology to see what the impact was on the quarterly estimates. For this initial test, car traffic on motorways was the selected strand of traffic data for the comparison. This comparison will need to be expanded to all strands of traffic before implementation.

Chart B1 shows that the final quarterly estimates are hardly changed by this process (although note that this is after benchmarking). QFNew uses the new ratio calculation method based on a recent year and recent weights, while QFOld uses the old ratio calculation method using the 1999 base year and the old weighting method.

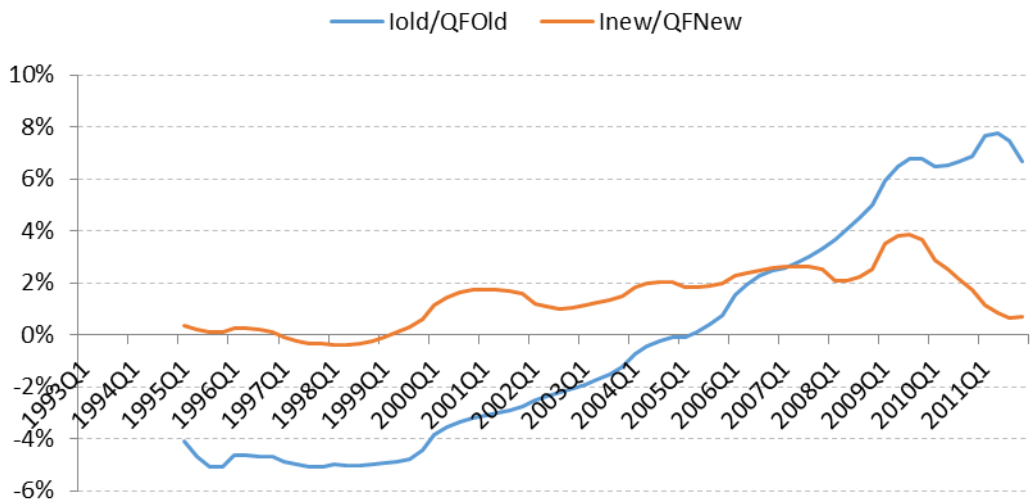
CHART B1: Comparing estimates produced by new and old calculation methods (Billion vehicle kilometres travelled by cars on motorways)



In the final stage of the traffic estimate process (step 5 in Annex A), the preliminary quarterly estimates are benchmarked to the annual estimate. The adjustment required has been getting larger as the quartiles drift further away from the annual[1], as can be seen from the blue line in Chart B2. This trend of the adjustments getting larger over time may partly be caused by the use of an old base year and/or by using the long-term total annual traffic to weight the sample estimate.

Chart B2 shows that the revised method does have an impact on the pattern of benchmark adjustments. It is now possible to see that the new adjustments, expressed as a proportion of the estimate (Inew/QFNew) no longer have a steady upward trend we saw with the previous adjustments (Iold/QFOld). The new adjustment is under 4% of the quarterly total. This would appear to be attractive in terms of reducing drift and, therefore, the extent to which quarterly estimates have to be revised on benchmarking to annual estimates.

CHART B2: Benchmark adjustment as proportion of traffic, car traffic on motorways



Conclusion

The new methods of calculating the ratios and weights in the quarterly has been discussed in this annex and give similar final estimates results compared to the current method already being used. Despite this, the new methods are more methodologically sound and gave a benchmark adjustment that was no longer uniformly increasing.

It might be expected that the benchmark adjustment would on average be zero. However, this is still not the case and possible reasons for its persistence will need to be investigated as part of the next stage to this project when we repeat the test for the other strands of road traffic beyond car traffic on motorways.

[1] Extrapolated adjustments are applied to the provisional quarterly estimates to allow for this and avoid increasing large adjustments at the benchmarking stage.

Annex C: Benchmarking preliminary quarterly estimates

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).

Current method

The current method used is described in *Interpolating Annual Data into Monthly or Quarterly Data* (http://www.ons.gov.uk/ons/guide-method/method-quality/specific/gss-methodology-series/gss-methodology-series--6--interpolating-annual-data-into-monthly-or-quarterly-data.pdf). The paper describes interpolation using a cubic spline method used in the software. The existing method for achieving this benchmark adjustment is implemented using a piece of software called Inter that was developed by ONS and formed part of the Central Shared Database (CSDB) system. The adjustments produced by the software are applied as described in Annex A where I_{cq} is the adjustment output from the benchmarking program Inter.

Alternative methods

Two alternatives were considered, the ONS recommended Cholette-Dagum method and an ad hoc method based on proportionally splitting the final annual estimate:

a) The Cholette-Dagum method

ONS now recommend the Cholette-Dagum method which is being implemented in their updated systems. The recommended approach is described in the paper *Benchmarking Time Series with Autocorrelated Survey Errors*[2]. The paper describes a regression model that takes into account bias and autocorrelated and heteroscedastic errors in the original data. In the paper, the benchmarking problem is written in matrix terms as a linear equation and the General Least Squares solution for the coefficients of that model, the bias and the underlying series, are derived.

b) The ad hoc approach

This approach applies a very simple ratio for the quarters in each year equal to the annual sum over the sum of the quarterly totals. Again in terms of previously used notation and given annual totals A_{cy}, the simpler unadjusted final quarterly estimate is given by:

$$Q_{cq}^{F\prime} = Q_{cq}^{P} \frac{A_{cy}}{\sum_{q=yq1}^{q=yq4} Q_{cq}^{P}}$$

The next sections in this annex deal with a comparison of these two methods against the current methods and themselves. Each method was applied to the preliminary quarterly traffic estimates produced using the current methodology, with none of the changes applied as set out in Annex B.

Testing the Cholette-Dagum method

Chart C1 summarises the results of benchmarking the preliminary quarterly estimates using the different methods:

- A series benchmarked using the Cholette-Dagum approach, using a proportional adjustment
- A series benchmarked using the Cholette-Dagum approach, using an additive adjustment

Compared against

- The benchmarked series provided by Inter[3]
- The annual benchmarks, divided by four, are shown as a series of horizontal black lines.

We can see that the overall qualitative impact on the estimates from these alternative approaches is negligible, with all three benchmarked series being virtually indistinguishable.

[2] Cholette, Pierre A., and Estela Bee Dagum, *Benchmarking Time Series with Autocorrelated Survey Errors*, International Statistical Review, 62, 3, pp. 365-377, 1994.

[3] Note that the series adjusted by *Inter* includes the small extra adjustment to account for the benchmark accuracy lost in rounding the adjustment.

CHART C1: Comparison of various benchmarking methods (Billion vehicle kilometres travelled by cars on motorways)

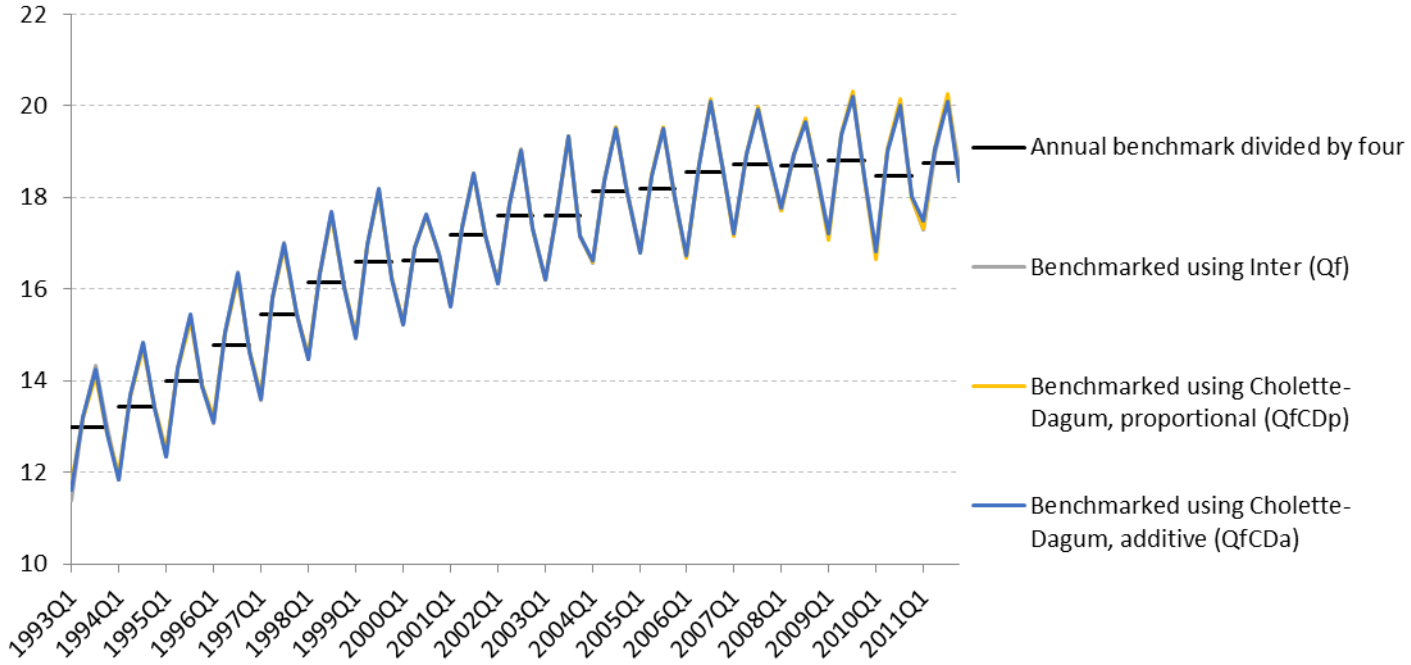
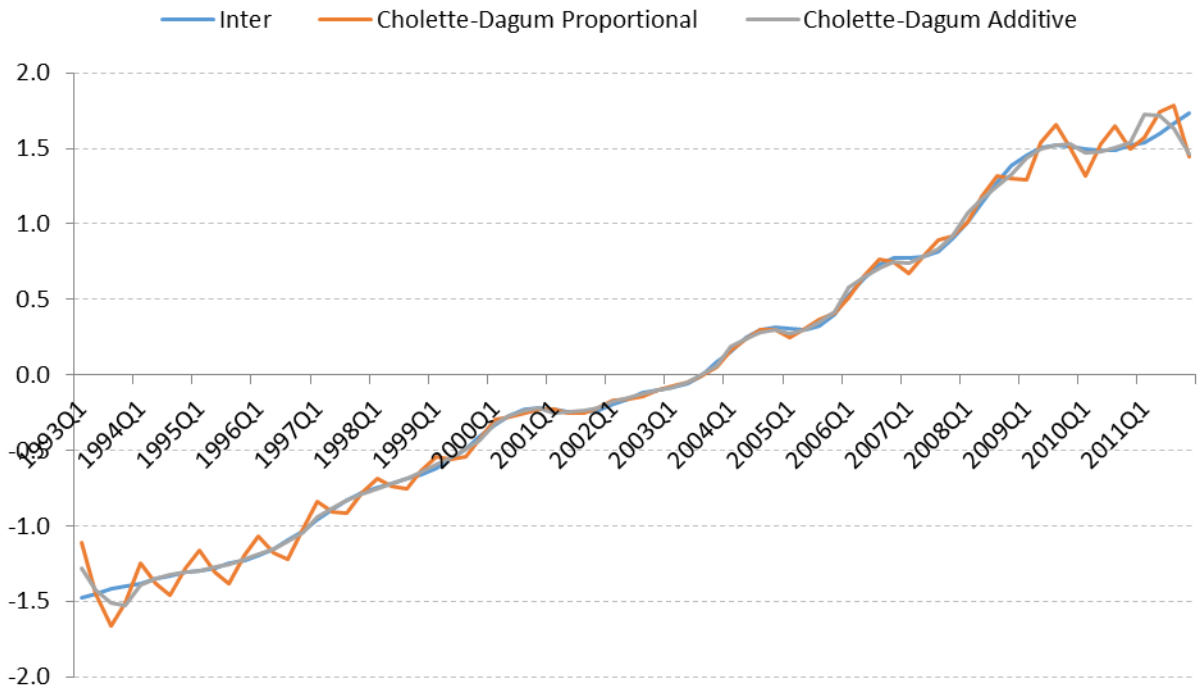


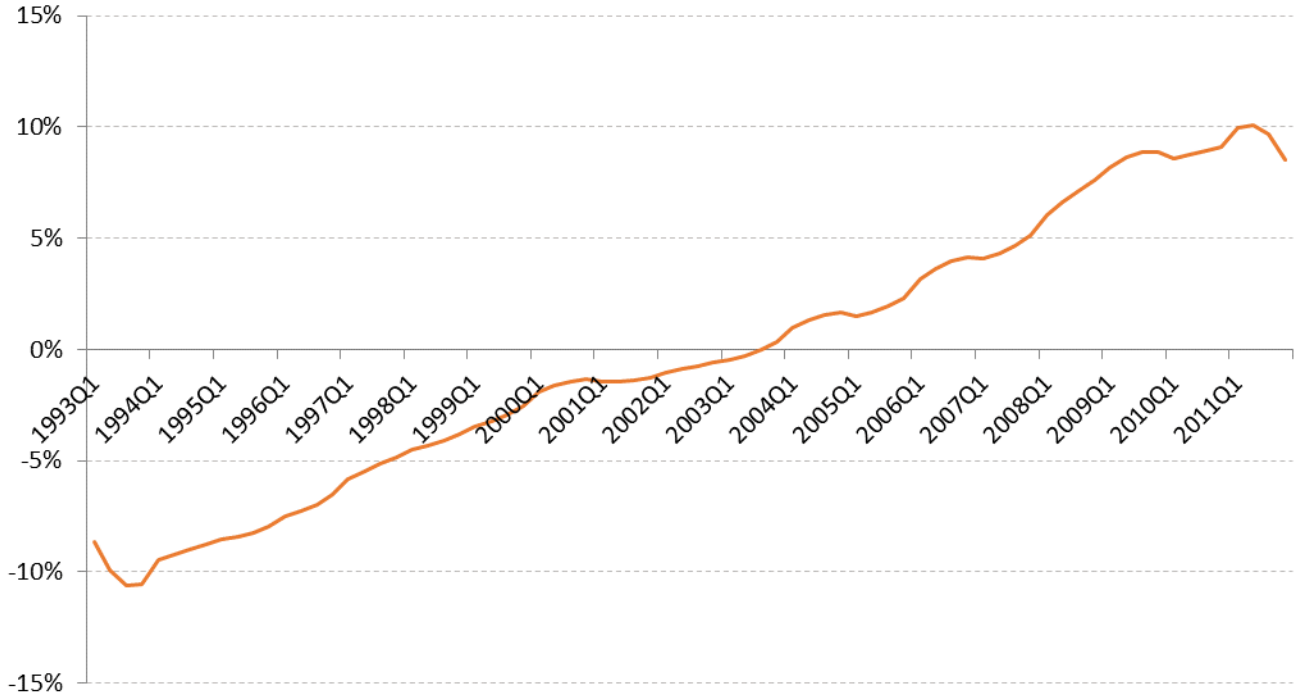
Chart C2 shows the absolute benchmarking adjustments made under the three methods, that is, the difference between the preliminary and final estimates. The Cholette-Dagum additive method and *Inter* method are both smooth and very similar for the majority of the series, except near the ends. The proportional method is clearly seasonal in the adjustment, which is to be expected as the series itself is seasonal. This has an implication for the quarterly processing since the current method extrapolates these adjustments to predict for the current year where no annual estimates are yet available.

CHART C2: Comparing the absolute adjustments made by Inter and Cholette-Dagum (Billion vehicle kilometres travelled by cars on motorways)



If we look at the Cholette-Dagum adjustment, as a proportion of the preliminary estimate, as presented in Chart C3, the line is smooth. Therefore, to replicate the current method of extrapolating the benchmark adjustment, it would appear to be more sensible to use this proportional adjustment, rather than the additive adjustment.

CHART C3: Cholette-Dagum adjustment as a proportion of traffic, car traffic on motorways



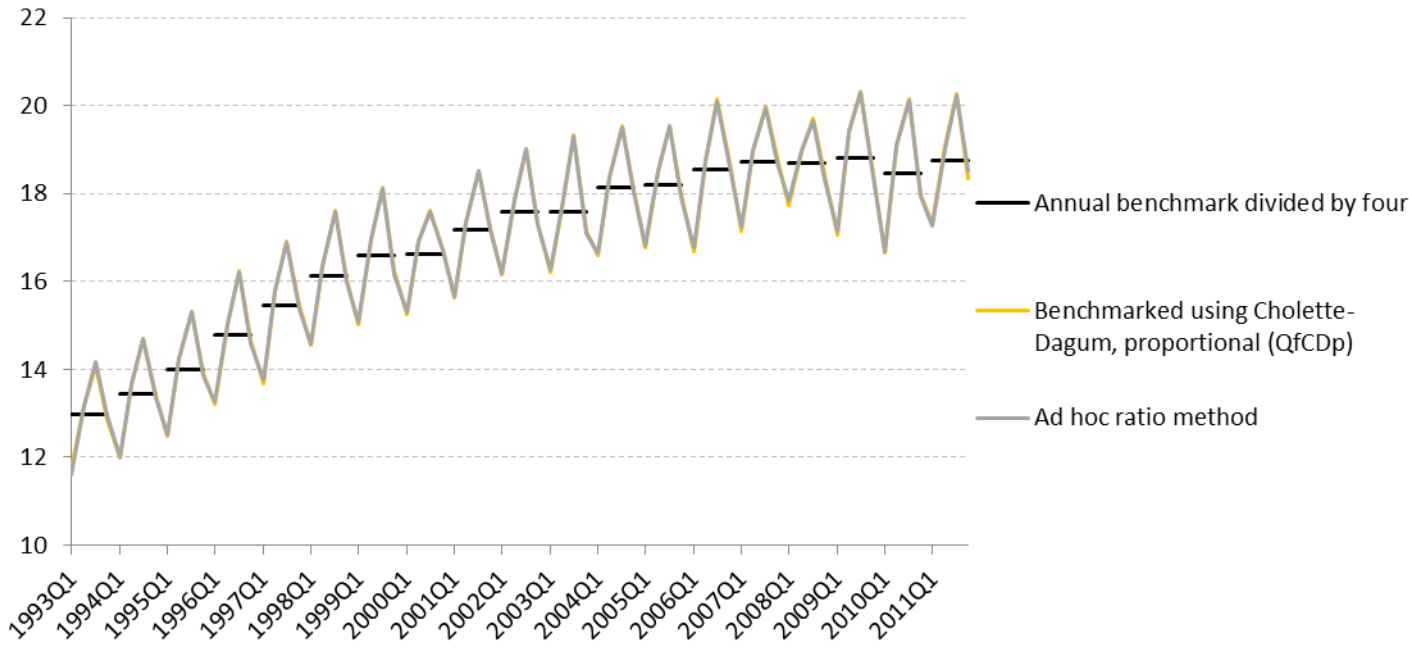
Both of the Cholette-Dagum methods produce a distinctive curve away from the trend at both ends of the series. This arises because the method implicitly forecasts the relationship between the path and the benchmarks, assuming no systematic relationship between the two other than an average difference in level which is modelled by the bias. It would, in principle, be possible to add an extra term to the underlying model to improve the behaviour at the ends of the series. However, in our alternative approaches to deriving the preliminary quarterly estimates, we have proposed an approach where the benchmark adjustment does not have this distinct relationship over time and so the assumptions are better met. Noting also the findings later about a simpler approach, we do not consider this additional complexity to be required at present.

Testing the ad hoc approach

The complexity of the formulation of the Cholette-Dagum approaches and the need to integrate an approach within the existing processing software has prompted the question of what could be achieved with a simple *ad hoc* approach.

The results are shown in Chart C4, together with the annual benchmark divided by four, and the Cholette-Dagum proportional results. The two benchmarked methods are again virtually indistinguishable, with the ad hoc results differing by between (-1.4%) and (+0.9%) of the Cholette-Dagum result across the whole series[4].

CHART C4: Comparing the complex and simple benchmarking methods (Billion vehicle kilometres travelled by cars on motorways)



Conclusion

The ONS-recommended Cholette-Dagum benchmarking method, applied using *R* software, is a potential substitute for the existing *Inter* software. This could be implemented as a standalone procedure by taking the preliminary quarterly results from the SQL system and reading back in the benchmarked outputs from the *R* program. It may be possible to better integrate the benchmarking through an *R* call from within the SQL programs, but this would require some further system development.

We noted that the results from the more complex Cholette-Dagum procedure, with the two alternative approaches, and the *Inter* results were very similar. We then went further to compare these more complex approaches with a very simple ratio approach that could be implemented in SQL directly, and again found that the results were very similar. From these results, we conclude that the issues of autoregression and heteroscedacity that the more complex Cholette-Dagum procedure is designed to cope with are not an issue with this series.

Our next steps are, therefore, to repeat the Cholette-Dagum benchmarking and simple ratio method adjustment for the quarterly series for the other road types and vehicle types. If the same conclusion is reached - that the two produce near identical results - then the simple method will be used in the production system. It seems reasonable to assume that this relationship will hold at least until any substantial change is made to the methodology for calculating the quarterly or annual estimates, at which point the comparison could be revisited.

[4] The largest differences are at the ends of the series and if the first and last years are omitted, the differences are between (-0.7%) and (+0.8%).