On the reasonableness of traffic forecasts

Robert Bain, RBconsult Ltd

It is difficult to approach the topic of this short paper without commenting on overconfidence. Range estimation is popular in many disciplines where professionals, facing uncertainty, estimate possible outcome ranges. Where estimates are biased by overconfidence, narrower ranges result. The psychology literature is rich with examples of individual and professional overconfidence. This form of behavioural bias has been tested by many academic studies and has been documented for – among others – investment bankers, financial analysts and economic forecasters (van der Venter & Michayluk, 2008). Are traffic forecasters any different?

A literature review of behavioural research material reveals some consistent findings which could have relevance here:

- **Overconfidence increases with the prediction difficulty level (traffic forecasting is not easy);**
- **The phenomenon occurs more frequently among males than females (this is a male-dominated profession);**
- **. . . and increases with age and experience (the job titles of the survey respondents – described later – suggest relatively high levels of seniority).**

Overconfidence is not the central issue here. The paper focuses on a small survey and the survey results. However, reflecting on these results, some comments are made about overconfidence towards the end.

INTRODUCTION

On the 22nd March, 2011 a message was posted to two popular transport modelling-related email lists[1] asking for subscribers' views about the 'reasonableness' of traffic forecasts. The aim was to conduct a short, simple survey about predictive capability.

List subscribers were presented with four scenarios:

- – an existing regular (toll-free) road
- – an existing toll road
- – a new-build toll-free road
- – a new-build toll road

...and four forecast horizons:

- – the following (ie next) day
- – one-year ahead
- – five-years ahead
- – 20-years ahead

Subscribers were asked to provide their answers in the form of ranges, such as ± 5% or ± 25%. The intention was for practitioners (such as traffic modellers) and those otherwise involved professionally with transport models (department of transport officials or academics) to provide their estimates of the error ranges – or notional confidence intervals – that would likely apply to state-of-the-practice traffic forecasts of different planning horizons.

The survey was, through necessity, somewhat crude and sacrificed sophistication for simplicity (and speed of completion) to encourage participation. A follow-up email message was sent on the 28th of March with a 'final call' being posted on the 30th March.

RESPONDENTS AND RESPONSES

48 replies to the survey were received, however, as two appeared to have submitted counter-intuitive responses, only 46 were carried forward for analysis.

Respondents represent consultants (22), state and other government officials (13) and academics/researchers (11). 21 are based in the US, 12 in the UK, three in New Zealand, two (each) in Australia and Canada, and one (each) in Bangladesh, Brazil, Chile, Hong Kong, Ireland and Sweden.

Although the response rate was low, the calibre of respondent was high. Many of the consultants hold senior positions (President, Managing Director, Director of Transport Planning) as do the government officials (Transport Modelling Manager, Senior Transport & Economics Advisor, Traffic & Toll Modelling Manager). Four of the academics are professors (one of whom is a leading author in the field of traffic modelling), two are senior lecturers and one is the deputy director of a centre for transport studies. The quality of the respondents thus part-compensated for the quantity of responses received.

RESULTS

Early responses to the poll demonstrated that the majority of respondents drew little (if any) distinction between what they regarded as 'forecast reasonableness' for a toll road and a toll-free road. As such, the follow-up email omitted that distinction and simply asked respondents to self-define reasonableness under two scenarios: for an existing and a new road.

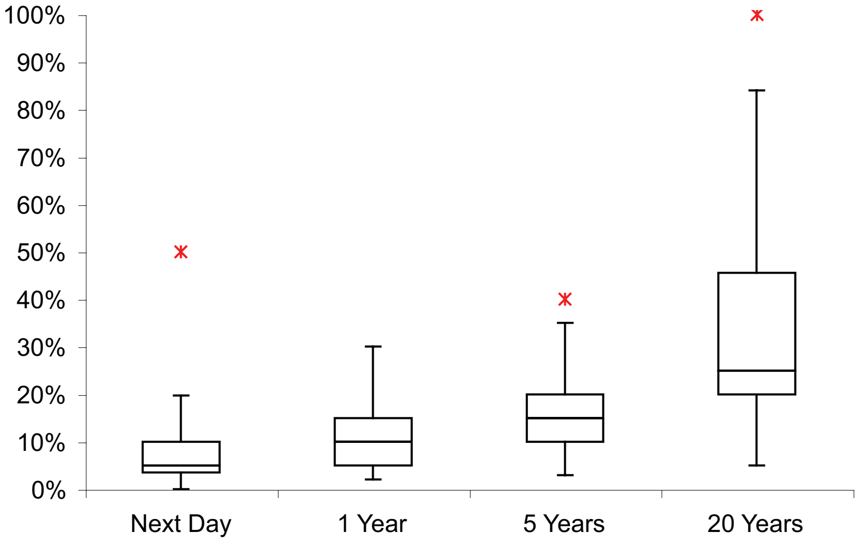
The average (mean) responses for each forecast horizon for each scenario are presented in Table 1.

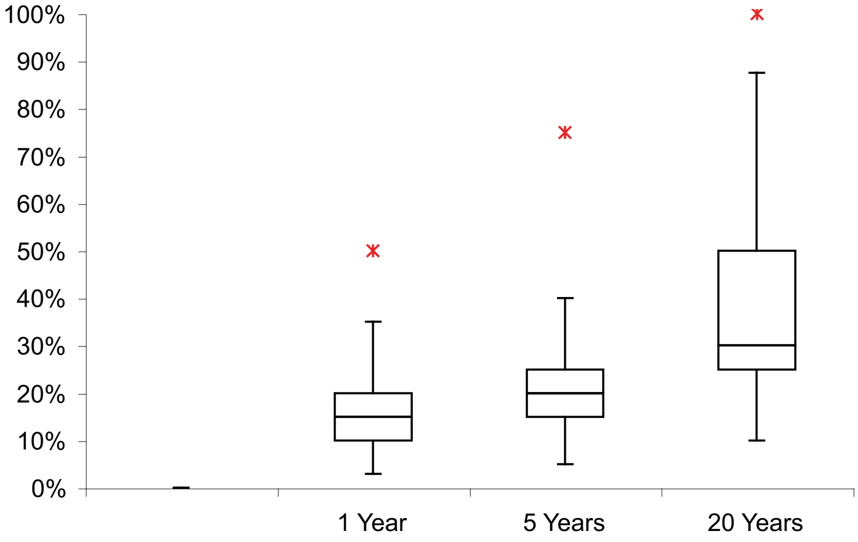
Forecast Horizon	Existing Road	New Road
Next Day	± 7.5%	
1 Year	± 10%	± 15%
5 Years	± 15%	± 25%
20 Years	± 32.5%	± 42.5%

Note: Percentages have been rounded.

Table 1: Forecast Reasonableness (means: all responses)

ABOUT THE AUTHOR
Robert Bain is a chartered civil engineer. For a number of years Robert worked for the rating agency Standard & Poor's where he was a Director in the firm's Infrastructure Finance Ratings practice with responsibility for transportation projects and public-private partnerships. Today he runs his own consultancy conducting infrastructure investment analysis. Much of his work involves reviewing transport models and demand forecasts for banks, infrastructure funds and institutional investors. For more information see www.robbain.com. He can be contacted by email at info@robbain.com





Note: In Figures 1 and 2, the ends of the 'whiskers' represent the lowest datum still within 1.5 times the interquartile range (IQR) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile. Outliers are identified in red. The vertical axis represents the error range (± %) associated with each of the forecast horizons. Respondents were not asked about next day forecasting performance for new-build roads.

Top, Figure 1: Boxplot – Forecast Reasonableness (existing road)
Bottom, Figure 2: Boxplot – Forecast Reasonableness (new road)

Setting aside the reported ranges for now, the pattern of responses accords with intuition. Under both scenarios – reflecting increasing uncertainty associated with deeper futures – the prediction intervals grow as the forecast horizon extends. Also, in both cases, predictive capability was felt to be stronger for existing (known) facilities than for new-builds – which introduce uncertainties of their own.

The results are depicted graphically in separate box-and-whisker plots (Figures 1 and 2).

DISCUSSION

Do the survey results suggest that traffic forecasters are overconfident in terms of their predictive abilities? With a sample of 46 it is impossible to draw definitive (statistically significant) conclusions. However between a quarter and a third of respondents reported the following predictive ranges:

- ± 8% or less for one-year forecasts (30%) [sub-sample mean = less than 5%]
- ± 12.5% or less for five-year forecasts (28%) [sub-sample mean = less than 9%]
- ± 22% or less for 20-year forecasts (32%) [sub-sample mean = less than 17%]

These do seem to be narrow ranges. These respondents are labelled 'low-ballers' and are referred to later.

To explore the issue of overconfidence further two approaches were developed; an input analysis and an outcome analysis.

Input Analysis

The majority of traffic forecasts incorporate growth and it would appear to be difficult to argue that the uncertainty associated with this growth could be less than the uncertainty associated with its determinants ('drivers')[2]. The growth drivers typically include projections of population, GDP, car ownership, households, employment, fuel price (and/or efficiency) or some combination thereof. Take possibly one of the more predictable of those drivers; population (certainly in relation to the predictability of GDP or fuel price – for example).

In terms of accuracy, an initial review of the literature on population projections appears positive (see – for example – Shaw, 2007). However two common trends emerge. Although accurate at the aggregate (state or national) level, forecasting performance deteriorates rapidly (a) as the study area shrinks – towards the zone sizes typically used in transport modelling – and (b) as the forecasting horizon expands. Smith and Shahidullah (1995) calculate errors for 20-year small-area population projections lying between 25% and 35%. Yet 13% of respondents to the survey suggested that 20-year traffic forecasts would have an associated predictive range of ± 15% or less, and only one-third of respondents reported possible ranges in excess of ± 30%. And census tract analysis by Smith, Tayman and Swanson (2001) suggest average errors of 45% and 54% for 25-year and 30-year population projections respectively. These are horizons frequently used in traffic forecasting and are wide intervals for a variable often used to part-explain traffic growth.

In forecasting reports, although the determinants of traffic growth are frequently described, it is rare to find a discussion of the future-year uncertainties associated with these determinants, how these uncertainties combine and the resulting implications for the traffic forecasts themselves. This would be helpful. Indeed, more research focussed here might serve to usefully frame debates about predictive capability more generally.

Outcome Analysis

This section turns to actual comparisons of traffic forecasts with outturn performance. This topic has been covered before – see, for example, Flyvbjerg et al (2006), Bain (2009) and Welde & Odeck (2011) – although, overall, it continues to receive surprisingly little attention in the literature.

As part of its Post-Opening Project Evaluation (POPE) initiative, in recent years the UK Highways Agency (HA)[3] has started to publish comparisons of traffic forecasts with outturn figures for its 'major schemes' (road improvements costing more than £5m)[4]. The comparison (for early-period, ie opening year) for 55 schemes is presented in Figure 3. In this figure, predictive performance is presented along the horizontal axis in terms of percentage error: (forecast – outturn)/outturn.

The fitted distribution (red line) in Figure 3 is centred almost on zero, suggesting an absence of bias (no system-

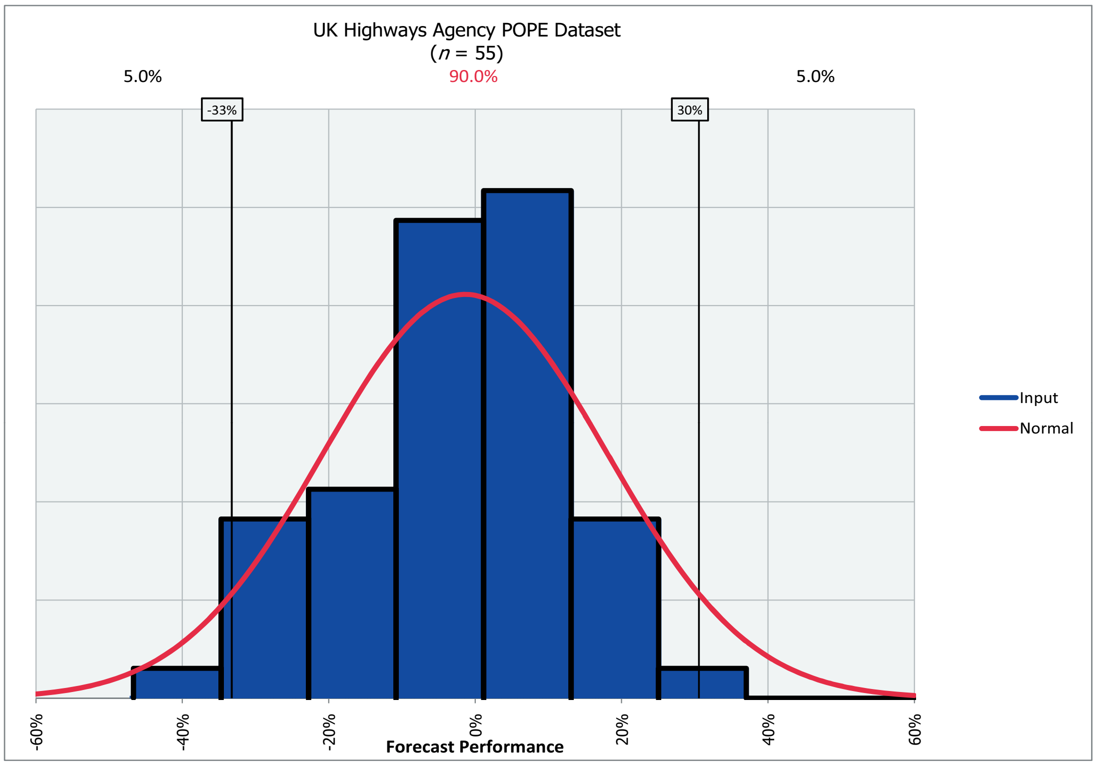


Figure 3: Traffic Forecasting Performance (UK Highways Agency, 2010)

atic tendency for over- or under-prediction). In terms of performance, 90% of the opening year traffic volumes fall between -33% and +30% of their respective forecasts. The HA provides both the source year for the forecasts and the opening year for each scheme. On average, the time lapsed between the two is just under 5 years (but say 5 years to be consistent with the survey). In contrast with the observations, over 70% of survey respondents provided ranges of ± 20% or less for 5-year forecasts.

In truth, the Highways Agency's forecasting performance has actually been quite strong in the past (an issue previously discussed by the author with the Agency). Other studies of traffic forecasting performance have cast predictive capability in a less favourable light. A comparative analysis of opening year traffic on toll roads versus forecasts (from Bain, 2009) is presented in Figure 4. This figure shows the data after adjustments were made for optimism bias, which has been a consistent finding in toll road-related research – see, for example, J P Morgan (1997), Vassallo (2007) and Li & Hensher (2010). The distribution shows 90% of outturn traffic volumes lying between ± 43% of their respective forecasts. This observed early-period range is almost identical to the range reported by survey respondents for 20-year forecasts (see Table 1)!

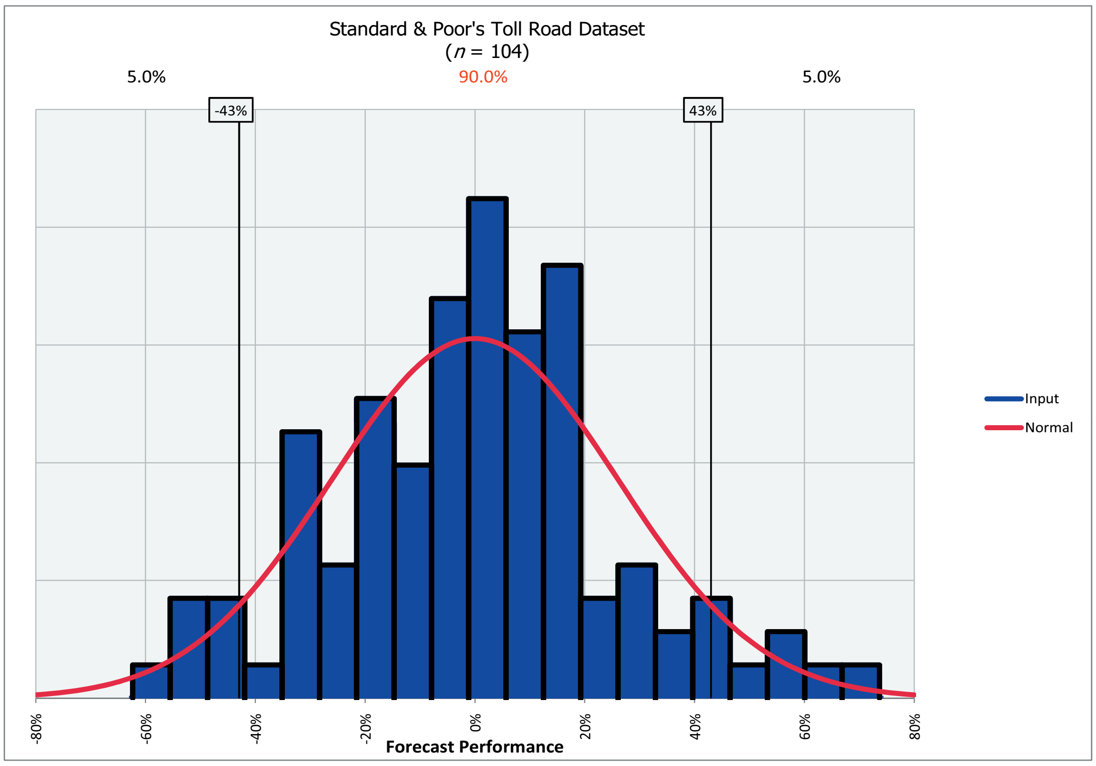
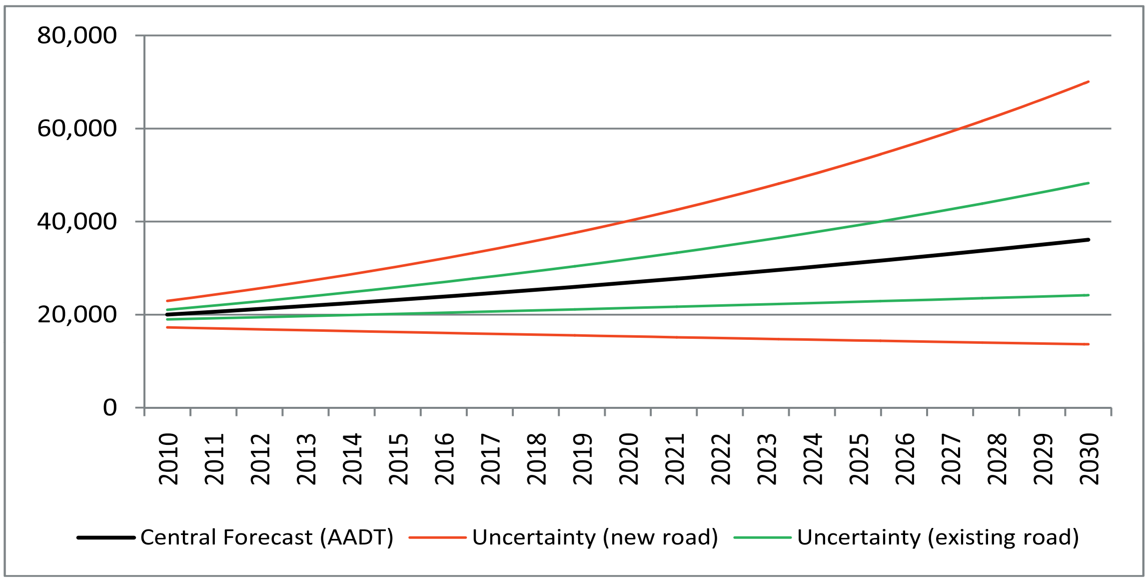


Figure 4: Traffic Forecasting Performance (S&P, 2009)

Figure 5:
Traffic Forecast Showing 'Uncertainty Envelopes' (all respondents)



CONCLUSIONS

- The survey results were reported earlier (Table 1 and Figures 1 & 2). To provide additional insight, the reported (mean) ranges are overlaid on a simple forecast in Figure 5. The forecast takes a (hypothetical) 2010 traffic volume of 20,000 vehicles/day and applies a 3% per annum growth rate over a 20-year horizon. The forecast range for an existing road is shown in green. The (wider) range for a new-build project is shown in red. This illustrates the forecasting 'uncertainty envelopes' as reported by survey respondents.
- Although the survey did not set out to examine the issue of overconfidence in terms of predictive capability (forecasting ability), clear signs of this behavioural bias are evident from the responses. To give some examples:
  - 5 respondents gave ranges for next-day forecasts of ± 0%. Other respondents pointed out that, as day-to-day traffic varies by around ± 10%, no forecasting range could possibly lie below that.
  - In terms of 1-year forecasts, 8 respondents suggested ranges of ± 3% or less.
  - In terms of 5-year forecasts, 10 respondents suggested ranges of ± 8% or less.
  - In terms of 20-year forecasts, 6 respondents suggest ranges of ± 10% or less.
- The concept of 'low-ballers' was introduced earlier. These are respondents who appear to have reported unfeasibly narrow ranges associated with the four forecasting horizons. Their responses place downward pressure on the average values reported earlier (Table 1). If the low-ballers are omitted from the sample, the resulting means are those shown in Table 2.

Table 2:
Forecast Reasonableness (omitting low-ballers)

Forecast Horizon	Existing Road	New Road
Next Day	± 7.5%	
1 Year	± 12.5%	± 17.5%
5 Years	± 20%	± 27.5%
20 Years	± 42.5%	± 47.5%

Note: Percentages have been rounded.

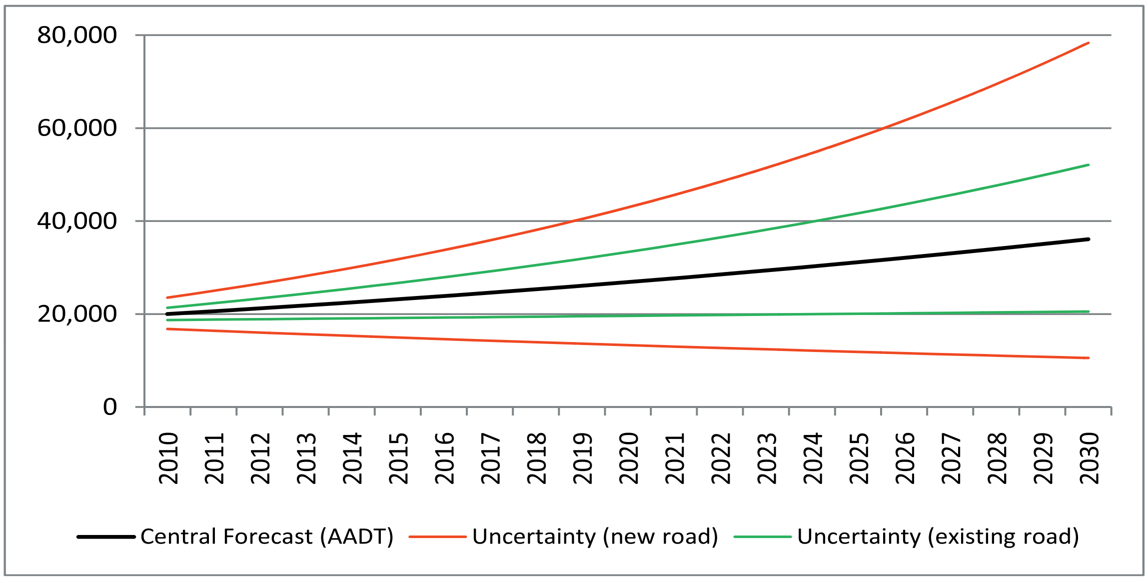


Figure 6:
Traffic Forecast Showing 'Uncertainty Envelopes' (omitting low-ballers)

- Applying these revised ranges to the hypothetical example given earlier produces the uncertainty envelopes shown in Figure 6 (as before, the ranges for an existing and a new road are presented in green and red respectively).
- The observed Highways Agency data discussed earlier appears to support at least one of the ranges reported in Table 2. 89.1% (nearly 90%) of the outturn traffic volumes fall within ± 27.5% of the Agency's forecasts, and these forecasts had an average 'age' of around 5 years. This suggests that, at a 90% confidence level, 5-year traffic forecasts for new-builds are likely to have an accuracy of ± 27.5% (although substantially more data would be needed to draw firm conclusions).
- This brings us to the clearest conclusion from the whole exercise. More data and more research are required. The author is the first to acknowledge the limitations of the survey reported here. When reviewing the population projection accuracy literature described earlier, it was obvious that demographers spent some time reflecting on the accuracy of their forecasts and reporting their findings back to their profession – so that they could assign empirically-derived confidence intervals and learn lessons that might guide future forecasting exercises. Given the extensive use made of – and reliance placed on – traffic projections internationally (by planners, policy-makers, economists and so forth) surely we should be doing the same?

REFERENCES:

- Bain, R (2009), Error and Optimism Bias in Toll Road Traffic Forecasts, Transportation, 36 (5), pp. 469-482, available at: http://www.robbain.com/Bain_Transportation_2009.pdf
- Flyvbjerg, B et al (2006), Inaccuracy in Traffic Forecasts, Transport Reviews, 26 (1), pp. 1-24.
- FSUTMS-Cube Framework Phase II, Model Calibration and Validation Standards, Final Report, Florida Department of Transportation Systems Planning Office (prepared by Cambridge Systematics Inc.), October 2008, Tallahassee FL.
- Li, Z and Hensher, D (2010), Toll Roads in Australia: An Overview of Characteristics and Accuracy of Demand Forecasts, Transport Reviews, 30: 5, pp. 541-569.
- Morgan, J P (1997), Examining Toll Road Feasibility Studies, Municipal Finance Journal, Vol. 18, No. 1, Spring 1997.
- Shaw, C (2007), Fifty Years of United Kingdom National Population Projections: How Accurate Have They Been?, Population Trends, Summer 2007, available at: http://www.gad.gov.uk/Documents/Demography/Population%20Trends/Population_Trends_128.pdf
- Smith, K and Shahidullah, M (1995), An Evaluation of Population Projection Errors for Census Tracts, Journal of the American Statistical Association, March 1995, Vol. 90, No. 429, Applications and Case Studies.
- Smith, K et al (2001), State and Local Population Projections: Methodology and Analysis, European Journal of Population, Volume 18, Number 3, 303-305, Springerlink.
- Travel Model Validation and Reasonableness Checking Manual, Second Edition, Federal Highway Administration (prepared by Cambridge Systematics Inc. as part of the Travel Model Improvement Program), September 2010, Washington DC.
- Van de Venter, G and Michayluk, D (2008), An Insight into Overconfidence in the Forecasting Abilities of Financial Advisors, Australian Journal of Management, Vol. 32, No. 3, Special Issue, March 2008, University of New South Wales.
- Vassallo, J M (2007), Why Traffic Forecasts in PPP Contracts are Often Overestimated?, EIB University Research Sponsorship Programme, EIB, Luxembourg.
- Welde, M and Odeck, J (2011), Do Planners Get it Right? The Accuracy of Travel Demand Forecasting in Norway, European Journal of Transport and Infrastructure Research, Issue 11(1), pp. 80-95.

[1] The Transport Model Improvement Program (TMIP) list – supported by the US Federal Highway Administration – and the Universities' Transport Study Group (UTSG) list – established to promote transport teaching and research in the UK and Ireland.
[2] The argument presented here assumes that all predictive uncertainty stems from assumptions about growth and that, otherwise, traffic models are perfect (introduce zero forecasting uncertainty of their own).
[3] The HA is responsible for operating, maintaining and improving the strategic road network in England.
[4] See http://www.highways.gov.uk/roads/18386.aspx