Role model

The latest Mott MacDonald Masterclass sees transport modelling expert Tom van Vuren ask: is it model error or forecasting inaccuracy?

WORDS BY TOM VAN VUREN

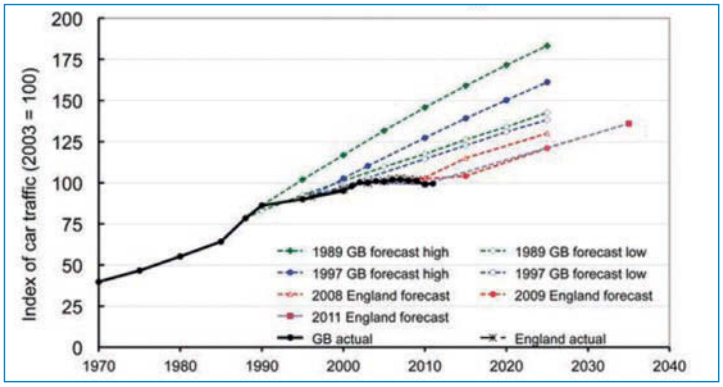
For a recent tram extension project in the English city of Birmingham for which we used the PRISM model, the modelled passenger forecasts in the scheme's business case were within one per cent of the outturn values, one year after opening. Surely that is cause for a pat on the back, but it led to another question: how good are our models in general? Of course, the model is only a simplification of reality, but expensive decisions with long term impacts are based on its forecasts. So I wonder, can we quantify the reliability of forecasts made by our existing models?

Surprisingly, the systematic evaluation of transport models used for infrastructure project appraisal is relatively young. The earliest reference I have been able to identify is a Rand Europe report from 2005, which compared model outcomes from the Dutch National and Regional Models with observed total number of tours and kilometres (by mode) in the study areas as well as vehicle flows on selected links. Their conclusion was that "the contribution of input uncertainty (eg in future incomes, car ownership levels) to these errors is generally much larger than that of model uncertainty". Colleagues here in the UK have queried the reliability of the Department for Transport's road traffic forecasts. Actual traffic volumes hover persistently at the lowest end of or below their predictions, going back as far as 1989 – is that because the model is wrong or (as in the Dutch case) because input variables (economic growth, population figures, fuel costs) have been consistently overestimated? And if so, does it really matter what the reason is? In the end it's the forecast figures that affect policy, not the model itself.

Confidence

Currently, confidence in the accuracy of our transport models tends to rely on how well the base model represents observed conditions. A lot of time, money and data resources are spent to get the base year right. Unfortunately, a good base year fit may give confidence in the model's structure and parameters but doesn't guarantee that any forecasts made using the model are reliable. Many of the drivers of change into the future are totally unrelated to transport, such as economic and population growth. Luis Willumsen has explained this well, separating model error from errors in base year as well as future year data, and finally in future scenarios.

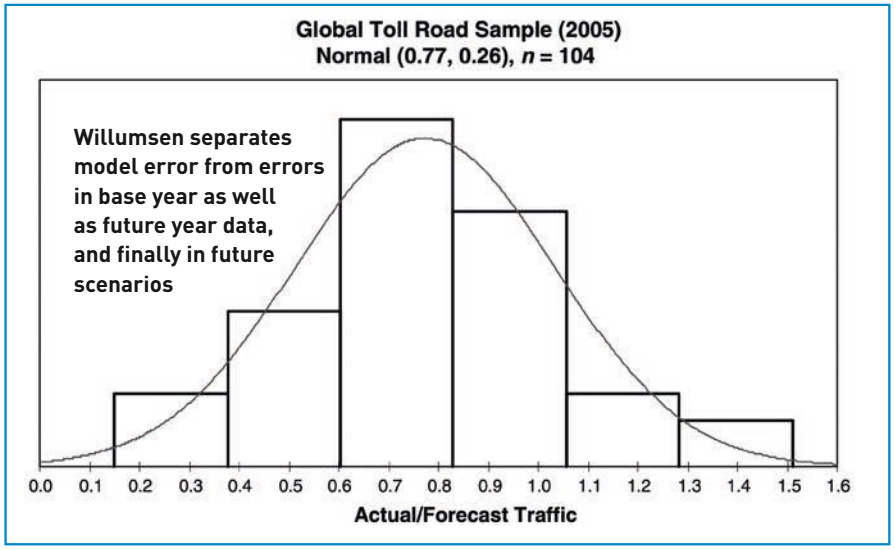
Source: https://transportist.org/2014/12/03/extrapolations-in-traffic-vs-reality/



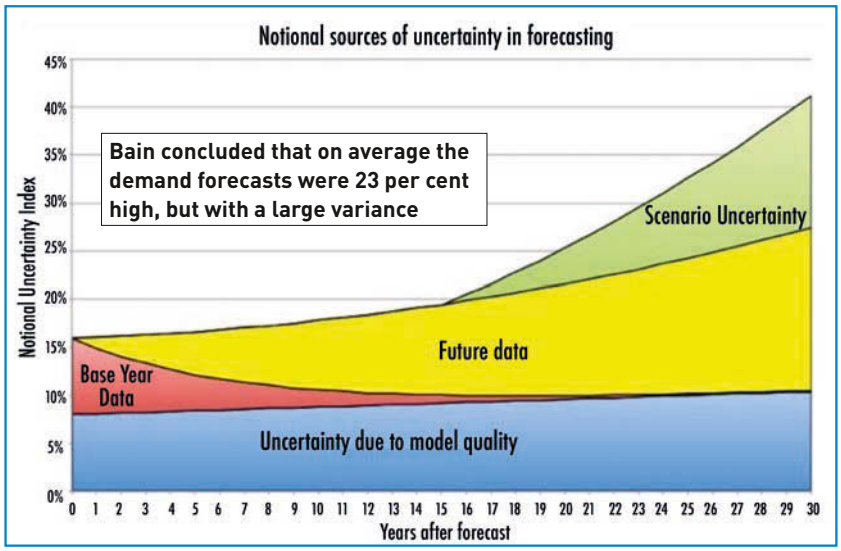
Actual traffic volumes hover persistently at the lowest end of or below their DfT predictions

And then there is the additional problem that a model fitted to data in, say 2018, is used to forecast behavioural responses way into the future, assuming that the observed and embedded travel behaviour doesn't change. Partly covering model error, partially covering scenario uncertainty, this temporal stability is a notion that is increasingly challenged – think of the recent findings of the All Change report (by Greg Marsden et al, 2018); and the National Travel Survey, indicating falling trip numbers (down nine per cent), travel distances (also down nine per cent) and time spent travelling (a three per cent reduction) since 2002.There is a quite substantial volume of literature on the accuracy of travel forecasts, but almost all of these do not and cannot separate the contribution of model inaccuracies from input errors. A few examples are:

- Seminal work by Flyvbjerg et al in 2005 identified from a sample of almost 200 projects that for road schemes forecasts were



Source: www.robbain.com/Bain_Transportation_2009.pdf



Source: https://goo.gl/mn9dT9

on average nine per cent high, and the figures for rail demand forecasting were considerably worse (realised demand was on average only half the forecast). But Flyvbjerg and his team could not identify the contribution of model inaccuracies – this would have required model re-runs to which they did not have access.

- Rob Bain had similar findings in his analysis of around 100 toll road projects. Bain concluded that on average the demand forecasts were 23 per cent high, but with a large variance. As in the Flyvbjerg study, he did not distinguish between uncertainty factors, but like the Rand study, he identified external input factors (and in the case of toll roads in particular, assumed values of time) as important contributors to uncertainty and error.
- The findings by Parthasanti and Levinson (2010) are quite different. As opposed to Flyvbjerg and Bain, they found that traffic forecasts have on average been lower than outturn flows, particularly for high-volume roads. But like the other studies, there is a substantial variance in the figures with both overestimates and underestimates occurring.
- Finally, a meta-analysis of post-opening project evaluations by Highways England (2015) found that two thirds of tested schemes accurately forecast traffic flows (to within plus-or-minus 15 per cent), but again there is much variability between schemes. The report also claims that the accuracy of traffic forecasts has improved over time, which they attribute to variable demand modelling, as is now expected for most projects in WebTAG.

Virtually all of the comparisons that are in the public domain look at the accuracy of the forecasts (conflating model error and input error), or more generally the extent to which the infrastructure project's objectives have been achieved.

There is another complication: the timeframe within which modelled responses occur differs. It is much easier and quicker to change departure time or route, than mode, let alone destination. For all effects in the model to have occurred in reality, a substantial amount of time (years!) needs to pass after implementation – and in that time the errors in future data and scenarios will also have grown.

What to do? It seems obvious that we need to deal with all aspects of uncertainty: model precision will almost certainly improve over time, with new techniques emerging such as agent-based models, and with big data sources providing much richer sources for estimation and validation. And novel data sources such as mobile phone traces and GPS-derived travel times and speeds will allow much more detailed before and after studies, building an evidence base for assessing the reliability of existing tools, or pointing towards necessary improvements.

So here's my suggestion for dealing with model and forecasting uncertainty:

- Develop alternative scenarios for medium to long term futures, say 5-50 years ahead. Even if these only respond to what we currently consider as the key drivers of demand (population, economy, housing and other land use development), the uncertainty fan will help understand the risks associated with the decision;
- Sensitivity-test the models that we use to forecast the impacts of policy and interventions in those alternative futures. Allow for errors in model estimates, as well as different assumptions on how currently observed behaviour (such as trip rates) may change over time. Where tolls, fees or fares play a role, a range of values of time should definitely be tested;
- Finally, monitor not just whether the project's objectives are being met, but also the accuracy of the travel demand forecasts that fed into their appraisal in the first instance, and at different times after implementation or opening. This can be done much easier now with big and passively collected data.

The greatest challenge (and opportunity) is to accept and start using models and model results without expecting or even desiring a single version of the truth. With hindsight it is amazing that we ever felt comfortable doing so in the past!

Further reading

- **Rob Bain (2009)** Error and Optimism Bias in Toll Road Traffic Forecasts. Transportation DOI 10.1007/s11116-009-9199-7 http://www.robbain.com/Bain_Transportation_2009.pdf
- **Flyvbjerg, B., Holm, M., Buhl, S.:** How (in)accurate are demand forecasts in public works projects? J. Am. Plan. Assoc. 71(2), 131–146 (2005). https://goo.gl/ceJwFG
- **Gerard de Jong et al (20015)** Uncertainty in Traffic Forecasts: Literature Review and New Results for the Netherlands. Working Paper WR-268-AW, March 2005. https://goo.gl/wZFy3Q
- **Highways England (2015)** Post Opening Project Evaluation (POPE) of Major Schemes - Executive Summary. https://goo.gl/rsj91i
- **Pavithra Parthasarathi and David Levinson (2010)** Post-construction evaluation of traffic forecast accuracy Transport Policy, (2010), https://goo.gl/yC8oso
- **Greg Marsden et al (2018)** All Change - The future of travel demand and the implications for policy and planning. https://goo.gl/xMRWPb
- **Bert van Wee (2007)** Large infrastructure projects: A review of the quality of demand forecasts and cost estimations, Environment and Planning B, Planning and Design 34(4): 611-625

TOM VAN VUREN is Divisional Director at Mott MacDonald
mottmac.com
Tom.vanVuren@mottmac.com