Chris Kennett
August 2015

Merging Traffic at Signalled Junctions

Introduction

Back in 2012, at the JCT Symposium, I presented a paper 'Modelling Merges at Signalled Junctions'. In that paper I showed that lanes merging after a junction affects driver lane choice approaching the junction. Using data from three different sites, I showed that the share of traffic between merging lanes was biased towards the nearside lane, was predictable and results were statistically significant. This in turn, made it possible to model the effect using Linsig, although some changes were needed; in particular, a modified version of the Linsig give-way model was used.

This approach provided a breakthrough in predicting and modelling traffic flows where lanes merge after a junction, however the paper did not explore the details of the sites, or try to understand the driver behaviour.

Scope

This paper looks into the detail behind the original statistical observations and predictions; we try to explain how and why drivers are reacting to merging lanes. We will also examine the differences between the data from the sites, using a subjective analysis to explore differences and similarities, to build on confidence in the original paper. Finally, more data, from a wider range of sites is presented in addition to the original results.

Methodologies

To discover and explore potential differences between data from sites, and to examine the reliability of the original data, a wide range of data collection methods have been used. Video surveys and on site observations have provided key qualitative information, while MOVA flows have been collected to provide raw data. From MOVA flow logs, both the cumulative lane flows and the X detector flows have been collected. Although this does not provide a firm base for statistical analysis between sites, it does allow us to confirm the same trends can be seen, regardless of the method of data collection. This is helpful in establishing that a single methodology does not overly influence the result.

In total, data was collected from seven new sites not originally included in the original paper. Amongst these seven sites, there were 12 approaches that merged. All were examined and appeared to follow similar trends, consistent with the original findings. Eight of the approaches were discounted from the study though, as the count data could not be verified to relate to a specific merge on an exit, as opposed to other exit lanes.

Key Questions

Several key questions arose as a result of the original paper. These predominantly relate to driver behaviour in differing circumstances, and therefore the general application and relevance of the results. Broken down into simple questions, they are:

1. Why do drivers appear to prefer the nearside lane, when lanes merge after a junction?
2. Although all sites show a statistically consistent trend, there are small differences between datasets collected by different methodologies. What explains these differences?
3. How do different geometries and interventions affect driver behaviour?

Why do driver prefer the nearside lane?

The original paper hypothesised that some drivers will be willing to take higher risks and use the offside lane. As the perceived benefit of the offside lane increases (I.e. delay increases in the nearside lane), proportionally more people will accept the higher risks and use the offside lane.

Figure 1. A5/A47 Longshoot Junction, Nuneaton

This does seem to support and be borne out by the results established, however there is little evidence to demonstrate that drivers do perceive an increased risk in the offside lane. Standard methods for testing this hypothesis would include interviewing drivers, however that is beyond the resources available for this research. Instead, clues as to the perceived risk can be inferred from watching driver behaviour.

Using a video survey of a site in Plymouth, driver behaviour could be examined in detail. Observing the video over several time periods and several days, the site was first of all confirmed to follow the originally established trend. A difficulty in measurement became apparent though.

The video showed the approach to the signals and the exit, looking forward in the direction of travel. All drivers using the offside lane to go ahead clearly must merge, however it is clear that drivers change lanes from Lane 2 to Lane 1 at many different points. Focusing on drivers in Lane 2, during quiet periods, many drivers tend to merge to Lane 1 earlier, often before the stop line. In some cases, vehicles are even moving between lanes as they come into the frame, some 50 metres before the stop line.

At busier times though, drivers tend to stay in Lane 2 longer, crossing the stop line and either merging in the junction, or using the merge on the exit to re-join Lane 1. In mid and high flow situations, several drivers can even be seen moving out of Lane 1 late, or starting to change lanes before moving back. This late-weaving behaviour can be seen to be attributable to perceived delay – drivers in a free flowing Lane 1 rarely move across late. Likewise, drivers in Lane 2 rarely join Lane 1 early if Lane 1 is not free-flowing.

Figure 2. A46/A607, Hobby Horse Roundabout, Leicester

Figure 3. A1101 / Weasenham Lane, Wisbech

It is quite clear therefore that delay is one of the factors affecting driver choice, as Lane choice can be seen to be influenced by perception of delay in both lanes. Inferring perception of risk is more difficult. Despite this however, there is clearly a bias towards Lane 1 – drivers sometimes weave from Lane 1 towards 2 and back to Lane 1, but the opposite is never true. Drivers using Lane 2 also appear to tend to drive slightly faster and during low traffic conditions, the later they merge, the faster they tend to drive. This does not prove that risk perception is a second primary factor, but it does support the idea.

Figure 4. A46/A607, Hobby Horse Roundabout, Leicester.

What explains the differences in the data?

The unexpected behaviour observed above, quickly became apparent as a likely cause of differences between recording methodologies.

When modelling traffic using a program such as Linsig, we do not tend to consider the location of vehicles in a lane; they are either in one lane, or another. Sometimes an allowance may be made to allow flows to be automatically balanced by allowing traffic to 'move' between lanes in between junctions, however the exact location of this is never considered in detail.

What appears to be happening in practice though, is that the measured lane flow changes depending on three factors: overall flow, the longitudinal position on the carriageway at which flow is being measured, and the geometry of the road itself.

Leaving the geometry for the last 'question', the flow and point of measurement together pose a significant problem for standardised data collection. If drivers do tend to change lanes (both to Lane 2 and back to Lane 1) later as flow increases, the measured flow in each lane will vary depending on where the measurement is taken. Measurements taken 40 - 50 metres away from the stop line (MOVA X loops for example) may under-represent the number of vehicles actually crossing the stop line in Lane 2 in busy periods.

Measuring at the stop line though, (such as in a typical manual survey) may be likely to miss the drivers who merge into Lane 1 in advance of the stop line, particularly during mid – low flow periods. This may also depend on the exact methods and consistency of each individual person conducting a manual survey – some may choose to count vehicles merging shortly before the stop line, whereas others may not. Both count positions are therefore likely to give slightly varying trends.

The alternative to a manual survey of course is an automatic vehicle count. Automated surveys, collecting data constantly will collect a much larger sample of low-flow data. They may be affected by loop-clipping though, as vehicles straddling or changing lanes may be counted in both. Manual counts may be more accurate (subject to caveats above) but due to cost, tend to under-sample low flow conditions. This may exacerbate discrepancies and errors in the trend during low flow conditions.

As a result, there is no clear “best method” for taking a representative measurement of lane flow. Generally speaking though, there is no clear benefit (or loss) to capacity from a vehicle changing lanes shortly before the stop line, compared to after it. An automated lane flow measurement between 12m and 50m from the stop line over a 24 hour period is likely to give the most reliable measurement.

On balance, it may be necessary to tacitly accept these differences between methodologies. By compiling data from these different sources, the trend will be an aggregate of both and results of modelling based upon it should be suitable for all circumstances, within reasonable error bands.

How does the geometry effect the result?

At first glance, the results obtained from a wide range of sites, with substantial differences in geometry, follow a remarkably consistent trend. The following geometric traits are all included within the cumulative data samples:

- T-Junction, cross-roads and gyratory layouts
- Rural minor, inter-urban strategic and urban road networks
- Geographic spread of sites including Plymouth, Cambridgeshire, Norfolk, Peterborough, Nottingham, Leicestershire and Derbyshire.
- Lane 2 approach lane lengths from infinite (i.e. dual carriageway) to 60m flared lane
- Merge lengths between 60m and 100m beyond junction
- Various signage and road markings to indicate drivers should merge.
- Various radii (ahead and right turns) through junction.

There is insufficient sample data available to specifically analyse the wide range of differing geometric traits. Since all the sites fit with the general trend, it is fair to assume that no one of these traits makes a significance difference to the lane flows. What does seem more likely from the observations made, is that they do play an important part in establishing the point at which drivers are likely to merge.

A short flared approach lane will only allow a late movement to Lane 2, whereas a short merge on the exit is likely to encourage drivers to merge earlier. This may have an effect on the measurable traffic flow, depending on how the data is collected.

From this it could be expected that a site with short flare and short merge would not fit the general trend of lane use. The short distance available would minimise the benefit in delay but increase the risk. If all this holds true, then fewer drivers should use Lane 2 in almost all circumstances. Unfortunately no data from sites of this design was available.

Other traits that may have an effect could include signage of various types. From the observations made, primarily of a site in Norfolk, signage appears to be potentially effective in encouraging drivers to merge slightly later.

Figure 5. Dereham Road / Longwater Road, Norwich.

Theoretical Capacity of Merges

Drivers merging later is not the same as an increase in capacity. As was mentioned in the introduction of my original paper, any exit lane after the merge has a finite capacity – which can be calculated by well established methods. The maximum capacity achieved by the two merging lanes logically cannot be higher than the capacity of the exit lane.

Of course the traffic flow at the point of merge is not constant, as the traffic signals disrupt the flow between Stages. Even if both merging lanes are saturated at the point of the merge, lost time in the junction *may* provide sufficient time for the traffic to merge and dissipate, without the resulting queue from the slowing/merging traffic impacting upon the junction. In this sense, with traffic at or near capacity, and a relatively high amount of lost time, a merge may be effective in creating an increase in capacity from a dual carriageway or flared approach.

Where the flow exceeds the capacity of the exit lane, the result is each vehicle slowing down, with (normally) the nearside lane allowing the offside to merge by leaving larger gaps between vehicles. This initially sacrifices nearside lane capacity due to the increase in headway between vehicles. From this point of view, regardless of where traffic merges, once the exit lanes become saturated, there is no longer a capacity benefit from the merge - flare. Encouraging late merging may initially help to keep a queue clear of the junction, but it gains little or no capacity advantage in these conditions.

Figure 6. Dereham Road / Longwater Road, Norwich. Traffic merging shortly after start of green.

In the above photo, traffic can be seen merging shortly after the start of green. At this point, there is a much higher number of drivers in the offside lane than the nearside, and drivers in the nearside can be seen to be leaving large gaps. From this viewpoint, for a short while after the start of the stage, the flow in the offside lane exceeds the nearside.

As we have already established though, regardless of the point at which traffic actually merges, the proportion of vehicles in each lane remains the same. Looking at the same lanes of traffic, but as the lanes approach the signals, a different flow composition can be seen.

Figure 7. Dereham Road / Longwater Road, Norwich. Queue shortly after start of green.

From this viewpoint, it is clear that the demand for the offside lane is much lower than the view at the merge would suggest.

Accuracy and Variations in Trend

Using the ideas and theories above, we can go back and look again at the data we collected from various sites. As previously discussed above, there are differences in the trend lines between different sites. Looking at the data presented in graphs, on some, the trend line does not fit as well. There is no polynomial trend line that fits consistently across all sites. Instead, the data points appear to be linear (with a higher coefficient) up to a point, after which variation increases and coefficient decreases.

In this example, from The Long Shoot Junction near Nuneaton, nearside lane flow can clearly be seen to drop compared to total flow after about 700 vehicles (in a one hour period). During periods of lower flow though, the trend appears very consistent and slightly higher. One immediate reaction would be to assume that the trend line is a cubic curve, arcing as flow increases. This does not fit though, and no cubic trend line can be found that describes the flow better than the linear expression.

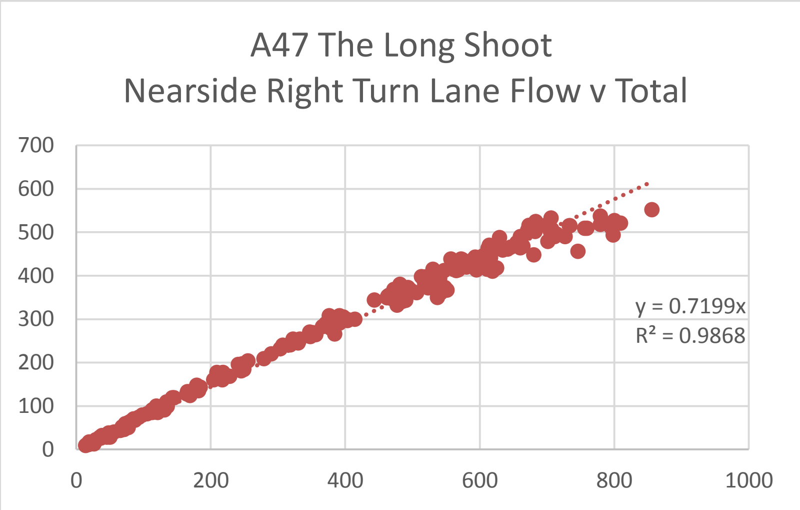

Figure 8. A5/A47 Longshoot Junction, All Data

Instead, if we hypothesise that the maximum saturation flow of the exit merge may be reached between 600 and 700 vehicles, the pattern becomes easier to understand. As the exit becomes saturated, traffic in the nearside lane slows down, to allow the offside lane to merge. The effect of this slowing is to create a longer, slower queue in the nearside lane. At this point though, the offside lane may still be relatively free-flowing, as was seen in the Norfolk videos. The important point to note is that the data above only shows vehicles that were able to pass over the relevant detector – it does not represent the demand in each lane, once a queue has formed.

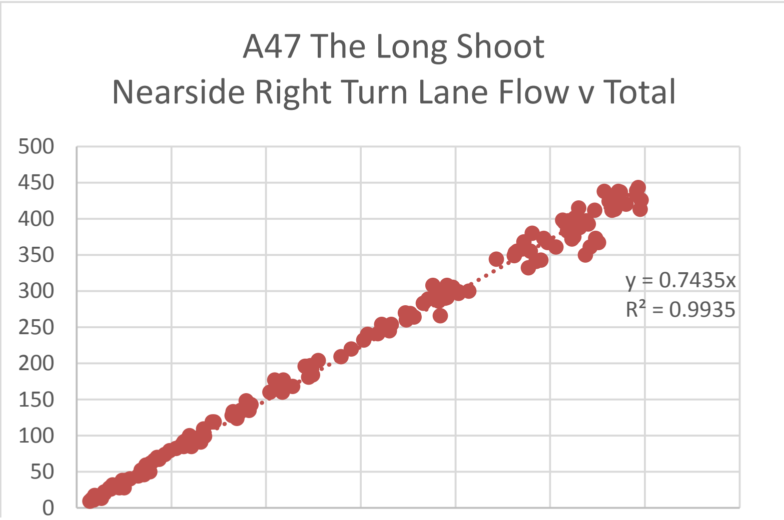

Figure 9. A5/A47 Longshoot Junction, Excluding saturated flows

If we exclude the data relating to flows at or near our assumed exit saturation, the trend line becomes a very close fit.

Correcting the data as I did in the original paper, we can once again aggregate the data to compare, and to look at the overall fit of our trend as a general predictor of traffic distribution between lanes. This does not exclude any high-flow data points, however the significance of the highest flow data points reduces, as there are comparatively fewer of them.

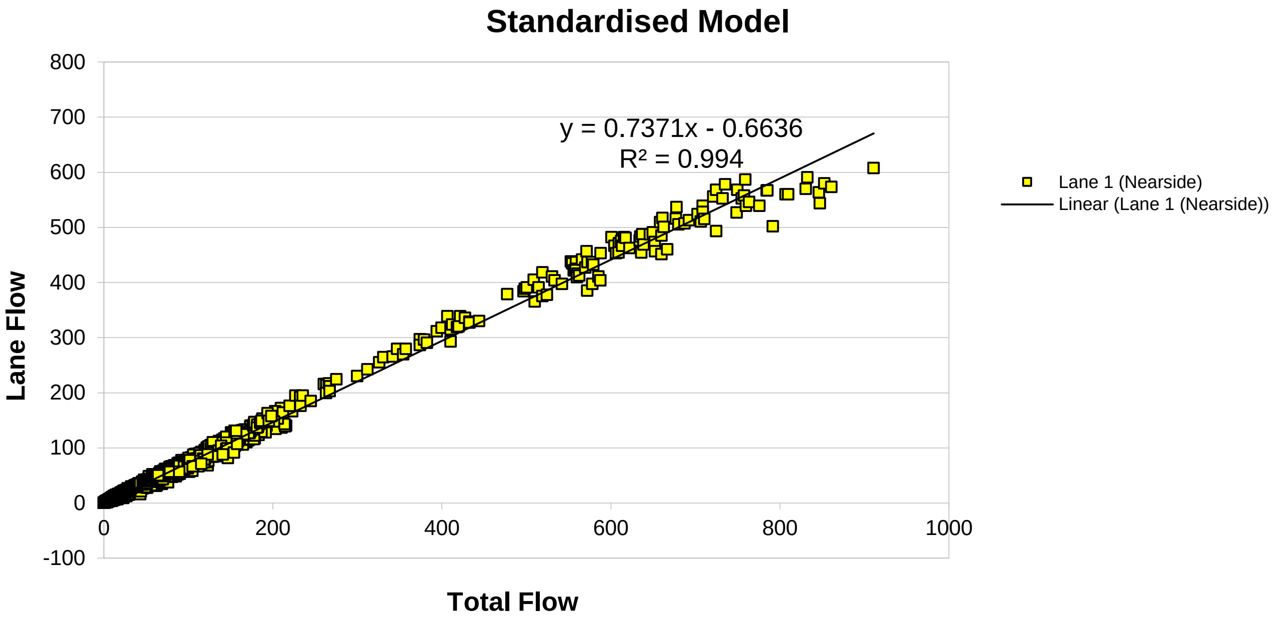

Figure 10. Corrected and Standardised Data, All Sites

Previously, we have simply examined the data to check for a trend and to measure the fit of the trend line to the data points. With this amount of data, we can look in more detail to establish the level of precision, and therefore how good the trend may be for prediction through modelling. Using basic statistical analysis techniques, the trend line (the "linear regression"), can be examined.

Input

The data series above contains aggregated data from 8 different merging approaches, sampled from 6 sites across Cambridgeshire, Peterborough, Norfolk, Nottinghamshire and Plymouth. There is variation in the number of samples from each site. There are a total of 1696 measurements.

Data Analysis

The measured coefficient is 0.7370 with an intercept of -0.6658, which is close to the original findings. The flow can obviously never be below 0 in practice, although there could be a number of different reasons that explain this small difference. The R Square value of the linear regression (how well the trend matches the data) is 0.9939. The standard deviation of the error of the regression is 8.4935 and the "t Statistic" value is 528.1711. The "P-value", is 0.0000. This indicates that the linear regression (the trend) is a very strong fit, with a very high probability of predicting the traffic distribution in the nearside lane. The actual confidence boundaries of the regression are: Lower 99.0%, 0.7343; Upper 99.0%, 0.7398; Lower 95.0%, 0.7343; Upper 95.0%, 0.7398. For the Intercept, the t-Stat value of -2.63 and P-value of 0.00859 indicate that there is less certainty in the result, although it is likely to be close (possibly 0, as hypothesised).

If we set the Y-intercept at 0, there is a slight variation in the coefficient to 0.735, with no appreciable decrease in any of the above indicators of consistency for the regression.

Limitations

Although the fundamental premise is supported with a significant amount of data (old and new), this paper has deliberately concentrated on discussing and analysing the theory as to why this may be true. Although this feature is reasonably widespread, there is simply insufficient data for the wide range of geometric variables to prove the impact of each feature, and therefore explain the slight differences between sites. That does not cast any doubt on the underlying findings though, which have remained strongly consistent, regardless of the variables.

The original paper did note that even with the merge modelled, the results did not follow the measured flows accurately. This now appears likely to be the result of the point of merge not being accurately reflected. It is possible that this can be improved on, but more work will be needed with Linsig to incorporate a model and refine it.

Although the ultimate capacity of the exit is constrained by the saturation flow of the lane after the merge, the junction still benefits from an increase in capacity from the dual lane approach or flared lanes. The capacity can only be modelled with reasonable accuracy if the proportional division of flow between the two lanes is accurately reflected.

We must remember in interpreting the result, that there are many other factors influencing the results of a traffic count at any given time. Short sample periods will always tend to have a greater divergence from any model of the same period, however long sample periods may become less relevant. This is true of all modelling. All other features such as bus stops, lane alignment and the upstream road network should be taken fully into account, as well as considering any merge.

Conclusion

The new data collected reinforces the findings of the earlier paper; that traffic flow in lanes approaching a merge is predictable and can be modelled. Furthermore, there is a consistent bias towards the nearside lane.

This gives us confidence that traffic flow in the nearside lane can be reasonably predicted by the expression:

$$\mathbf{F_n = 0.735 * F_T}$$

Where F_n is the nearside lane flow and F_T is the total flow.

The new observations and analysis go some way to explaining why this happens and how it works in practice. Although the bias in the proportions of traffic flows between lanes is consistent, the point at which drivers choose to use a lane, or merge back is widely variable, depending on conditions and geometry. Indeed many of the features previously considered to ‘increase capacity’ actually just change the point of merge, with little or no difference to actual capacity.

Some sites do still seem to show trends of slightly different proportional flows in the nearside lane. These tend to be sites from which a smaller sample size was gathered and have a higher variance, indicating these results are less sound.

Acknowledgements

With thanks to the many people who have helped by providing data or access to sites, including colleagues at Norfolk and Cambridgeshire County Councils, Peterborough and Plymouth City Councils, the Highways Agency, A-One+ and Amey.

Photos and video from Google Street View and from CCTV and direct recording on sites.