External factors affecting motorway capacity

B Morris, K Boddington, S Notley and T Rees
External factors affecting motorway capacity

B Morris, K Boddington, S Notley, T Rees
Disclaimer

This report has been produced by the Transport Research Laboratory under a contract with TRL Academy. Any views expressed in this report are not necessarily those of TRL Academy.

The information contained herein is the property of TRL Limited and does not necessarily reflect the views or policies of the customer for whom this report was prepared. Whilst every effort has been made to ensure that the matter presented in this report is relevant, accurate and up-to-date, TRL Limited cannot accept any liability for any error or omission, or reliance on part or all of the content in another context.

When purchased in hard copy, this publication is printed on paper that is FSC (Forest Stewardship Council) and TCF ( Totally Chlorine Free) registered.

Contents amendment record

This report has been amended and issued as follows:

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Description</th>
<th>Editor</th>
<th>Technical Referee</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25/10/2010</td>
<td>First Draft for Review</td>
<td>B Morris</td>
<td>A Stevens</td>
</tr>
<tr>
<td>3</td>
<td>04/11/2010</td>
<td>Final draft for technical review</td>
<td>B Morris</td>
<td>A Stevens</td>
</tr>
<tr>
<td>4</td>
<td>04/01/2011</td>
<td>Final version for publishing</td>
<td>B Morris</td>
<td>A Stevens</td>
</tr>
</tbody>
</table>
Contents

1 Introduction 2

2 Literature Review 3
   2.1 Defining Highway Capacity 3
   2.2 Calculating Highway Capacity 4
   2.3 The Impact of External Factors on Capacity 6

3 Methodology 8
   3.1 Selecting Factors to Assess 8
   3.2 From Factors to Parameters 9
   3.3 Site Selection 10
   3.4 Raw Data Descriptions 12
   3.5 Building the Data Sets 13

4 Results 15
   4.1 M4 Eastbound J19 – J20 16
   4.2 M25 Anti-Clockwise J11 – J12 17

5 Discussion 20

6 Conclusion 22
Executive Summary

This project was commissioned by the TRL Academy to investigate external factors affecting motorway capacity. The work was carried out at TRL during 2010.

Motorway capacity has long been considered fixed. Though extensive research related to concepts such as Controlled Motorways, which aim to reduce the occurrence of congestion, it has been observed that motorway capacity is in fact variable and not fixed.

This research has been carried out to investigate factors that affect capacity, and in doing so, build a simple model that parameterises these factors and describes the variable nature of capacity.

The objectives of this research were:

1. Construct an operational definition of capacity
2. Investigate factors that affect capacity using readily available data

Two motorway sites were investigated (J19 – J20 of the M4 and J11 – J12 of the M25) where congestion occurs regularly. MIDAS (Motorway Incident Detection and Automatic Signalling) data and data from Wunderground, a web based weather provider, was used.

The factors investigated were:

- Merge / Diverge percentage
- HGV percentage
- Precipitation
- Lighting conditions (day / night / twilight)

A statistical analysis showed that 46% and 27% of the observed variance in capacity was explained by the factors investigated for the M4 and the M25 site respectively.

These results form a promising basis for further research to be carried out and suggestions for potential improvements the methods developed here have been made, including:

- The inclusion of more factors, including geometric factors;
- The use of higher resolution, better quality data; and
- More sites where recurrent congestion forms
1 Introduction

This project was commissioned by the TRL Academy to investigate external factors affecting motorway capacity. The work was carried out at TRL during 2010.

TRL has been involved in transport research for over 75 years, and during that time has made significant contributions to the field of traffic, and, more specifically, research into the causes and dynamics of congestion on motorways in the UK. It is from this basis that the work detailed in this report has been carried out.

Motorway capacity, a concept which is defined in section 2, has long been considered fixed. Though extensive research related to concepts such as Controlled Motorways, which aim to reduce the occurrence of congestion, it has been observed that motorway capacity is in fact variable and not fixed.

This research has been carried out to investigate factors that affect capacity, and in doing so, build a simple model that parameterises these factors and describes the variable nature of capacity.

The objectives of this research were:

1. Construct an operational definition of capacity
2. Investigate factors that affect capacity using readily available data

This report covers the following:

Section 2 includes a literature review of relevant research in the field, where different methodologies for estimating capacity are discussed and results compared.

Section 3 includes information on the methodology that was developed to test the hypothesis that capacity is variable and that the variability can be explained by observable factors.

Section 4 includes the results of the study.

Section 5 includes a discussion of the results and a discussion on how this research could be extended and improved.
2 Literature Review

The concept and derivation of highway capacity originated with Greenshields seminal paper in 1935 (Greenshields, 1935), where the correlation between traffic speed and flow was first modelled. Subsequently, a sizeable body of research has sought to better understand the complexities of the ‘speed-flow’ relationship and to develop theoretical and mathematical methodologies for determining highway capacity that more accurately reflect the real-world dynamics of traffic flows.

Whilst such research continues to contribute to effective highway design and engineering, it is importantly enabling more efficient and responsive management of traffic flows for the benefit of road users.

2.1 Defining Highway Capacity

Traditionally, the capacity of a highway is given as a fixed value based on speed-flow relationships and geometric factors. Widely accepted design guidance, such as the Highway Capacity Manual (HCM) (2000), defines capacity as the maximum flow rate that can be reasonably expected to traverse a highway link under prevailing roadway, traffic and control conditions. The HCM provides, for defined types of highway link with different geometric layouts, a set of capacity values derived from observed data that vary only according to free-flow speeds. Importantly, the capacity value is not the absolute maximum flow rate ever observed but the reasonably expected and repeatedly achieved flow rate under ‘ideal’ conditions.

In recent years, there has been a growing recognition that capacity should not be considered as a fixed value. Empirical studies have demonstrated that capacity is dependent on external factors that vary in time and space. These external factors arise from changes to driver, vehicle, road and environmental conditions (including weather) (Ponzlet, 1996).

Furthermore, even under constant external conditions, observed flow breakdowns on a particular highway can be preceded by highly variable flow rates (Brillon et al, 2005), that do not always occur at the maximum ‘capacity’ flow. It is understood, therefore, that the breakdown state affects capacity values and, as such, capacity varies for different degrees of congestion. The concept of ‘capacity randomness’ assumes that the flow rate at which breakdown occurs has the properties of a random variable, since it is dependent on the behaviours and interactions of multiple drivers in the context of a specific local constellation of a highway (Brillon et al, 2005).

In summary:

\[
\text{Expected capacity} = \text{‘ideal’ capacity} - \text{‘capacity bubble’}
\]

Capacity is more generally thought to be, for a given time and location, the traffic flow below which conditions are acceptable (free-flow) and above which conditions are unacceptable (congestion).

Current research therefore focuses on modelling the dynamic and stochastic nature of capacity and, in particular, applying methodologies that determine capacity relative to
flow breakdown. This has also necessitated wider understanding of the external factors affecting capacity and the probabilistic nature of flow breakdown.

2.2 Calculating Highway Capacity

At a simplistic level, highway capacity has been calculated using a fitted distribution model so that it is equivalent to a percentage on the statistical distribution of observed flow rates. Over time, studies have refined this approach to improve accuracy of estimations, such as through the removal of long-headways data (Chang & Kim, 2000) or choosing a percentile within a specified range of observed maximum flow rates (Washburn et al, 2010). Such methods lack a theoretical basis and, without incorporation of breakdown events or congested data points, the accuracy of the estimated capacity values is unknown.

The aforementioned HCM uses a deterministic approach to calculating capacity based on analysis of speed-flow diagrams. A mathematical function is fitted to speed-flow data points, from which the apex of the function is determined as the capacity. Through subsequent publications of the HCM, capacity values have increased as a result of improved understandings of the influence of geometric factors on capacity (especially those such as lane width, which can be ‘masked’ by driver behaviour).

In the approaches above, the analyses tend to exclude data for ‘non-ideal’ conditions e.g. unstable flow. (Hwang et al, 2005) highlights that by only considering ‘ideal’ conditions and also excluding the effects of external factors, methodologies such as that used in the HCM result in double excluded errors.

Van Arde (1995) introduced a now well-known and frequently adopted methodology for calculating capacity which is also based on the fundamental speed-flow diagram. Subsequent comparisons with other approaches have found this method to be useful in providing general capacity estimates in terms of providing a good fit with the empirical data and consistency with traffic flow theory (Washburn et al, 2010). This method incorporates congested data points which increases accuracy of estimations; however, the capacity values are not tied directly to breakdown events. Therefore, it is still not possible to determine whether more traffic than the highest observed flow rates could be served.

Efforts have subsequently been made to develop a theoretical concept for incorporating breakdown events into capacity estimations, so that capacity value is based on the flows which cause a breakdown event. The Product Limit Method (PLM) estimates a breakdown probability based on flow rates preceding breakdown events, from which capacity is then determined as a certain percentile value of the breakdown probability distribution (Brillon et al, 2005).

The PLM uses the statistics of life-time analysis, and requires the application of a suitable algorithm for identifying breakdown events and thus a clear definition of what constitutes a breakdown event. Models using this approach are considered as ‘stochastic models’ since they capture the random variability of capacity as observed in empirical data. Therefore, for a certain traffic speed-flow relationship, there is a distribution of capacities (capacity distribution function) as opposed to a single capacity from traditional deterministic models (Wang et al, 2009).

This method was only applied to a highway with a bottleneck where the lane “drops” (Brillon et al, 2005), but later studies validated that PLM can be applied more widely to
basic highway links for estimation of capacity values (Washburn et al, 2010). It should be noted that the determination of appropriate breakdown probabilities is complex, and further studies (Washburn et al, 2010) suggested that estimates were only reliable for sites with more than 0.5 breakdown events in one day.

Nevertheless, this technique has improved the methodology for investigating the individual effects of various external conditions such as rainfall, visibility and operation of traffic control systems, as discussed later. It has also been used to investigate capacity variability under dynamic traffic conditions; for example, when traffic moves between the congested and free-flow state (Brillon et al, 2005).

The research cited above, through empirical analysis of traffic dynamics, has demonstrated that recovery from synchronised to free-flowing traffic always involves lower volumes than those observed during breakdowns. This difference between breakdown volume and recovery has been termed a ‘capacity drop’, and is thought to be dependent on factors such as driver behaviour and vehicle conditions (e.g. restricted acceleration capabilities of some vehicles when emerging from a congested area). The findings of Brillon et al (2005) are supported by other studies, whereby capacity has been seen to ‘drop’ by up to 6% (Ponzlet, 1996) and even 24% (Brilon and Zurlinden, 2003).

Geistefeldt (2008) has identified the breakdown probability as an important measure of effectiveness in highway design, since it represents the reliability of traffic operation. The authors therefore propose that defining a maximum acceptable breakdown probability could be considered as an alternative to derive design capacities, and the results of their study suggest that appropriate 1-hour breakdown probabilities are in a range of 25-50%.

The same study investigated the empirical relationship between deterministic and stochastic capacities by determining, for 27 data samples from German highways, the percentile of the capacity distribution function that corresponds to the conventional capacity estimated in the speed-flow diagram. The conventional capacity in 1-hour intervals was found to roughly imply a 40% breakdown probability (Geistefeldt, 2008).

Alongside the development of stochastic and deterministic models, efforts have also been made to develop dynamic estimation methods that model the aggregate effects of external factors, including driver, vehicle and highway conditions (Hwang et al, 2005). One such method has calculated capacity by restating an equation assuming that driver and vehicle conditions follow certain distributions and can be set as an error term. Dynamic capacities can then be estimated depending on speed.

The research proposes that when traditional ‘fixed’ estimates of capacity are used, such as that provided by the HCM, in some circumstances the volume per capacity of highway changes to oversaturated condition (e.g. flow rate can exceed capacity). However, when a dynamic estimation is calculated, the conditions where traffic flow rate exceeds the capacity of a highway do not occur as the capacity is dynamically changing.

Whilst the dynamic highway estimation method covered by Hwang et al (2005) more accurately reflects traffic characteristics then previous studies, it is still based on the statistical distribution of traffic speeds and as such is not linked to the occurrence of flow breakdown. Also, due to the generalisation of ‘driver condition’ and ‘vehicle condition’, it also does not contribute to any improved understanding of the significance of individual external factors on capacity.
In the studies referenced above, it is commonly noted that estimated capacity rates are lower than those predicted by the HCM. It is thought that this is explained in part by the increasing use of more disaggregated data values, as well as the understanding that external factors can contribute to flow breakdown well below the ‘ideal capacity’. Additional guidelines (NCHRP, 2008) provide equations to enable improved application of HCM methods in planning applications that take local factors into account.

A number of studies have explicitly looked into the relationship between external factors and capacity, as will now be discussed in more detail.

### 2.3 The Impact of External Factors on Capacity

Weather, and in particular rainfall, has been a focus of research seeking to quantify the relationships between external factors and highways capacity. Early studies produced generalised findings regarding correlation between weather conditions and traffic flow and volume, including consideration of the impact on driver behaviour and demand. It is only relatively recently that more research has been undertaken on the impact of weather on highway capacity, since capacity can only be measured on relatively limited sections of motorway with high traffic flow and as such require more localised weather information (Chung et al, 2006).

As explained, traditional methodologies, including that used by the HCM, assume good weather conditions (e.g. clear, dry, daytime), which is not an accurate representation of reality. Within the HCM, weather conditions are only taken into account with respect to free-flow speed, through the display of different speed-flow curves for different weather conditions. However, this does not extend to the consideration of capacity.

Prevedouros & Kongsil (2003) provide a synthesis of 26 previous studies investigating the effects of wet conditions on highway speed and capacity. A key finding was that the impedance factors caused by rainfall have a significant impact on road, vehicle and driver conditions. For example:

- Rain causes a reduction in friction and skid resistance which affects vehicle stopping distance;
- Rain causes light scattering, which impacts impact on visibility and object recognition; and
- These factors can trigger driver behavioural change, such as decreased speeds to account for longer perception / reaction time and stopping distance.

It is observed that this results in fundamental changes to the speed-flow diagram. The results from 11 studies (post-1980), where original data from highways are averaged assuming equal weights, produced an average speed reduction of 4.7mph in light rain. This is significantly less than the 12.0mph decrease suggested in the HCM (2000). The average capacity reduction is 8.4% in light rain (7 studies), which is comparable with the authors’ own research (Prevedouros & Kongsil, 2003). However, the study highlighted that there is considerable variance in previous research and that it is difficult to compare findings given the type of data, time periods and sites used.

A more recent study (Chung et al, 2006) was undertaken in Japan, and used high resolution weather and traffic data to investigate the impact of rain on motorway capacity and speed, and the effect of daylight on capacity. Findings demonstrated that
rain decreases capacity ranging from 4 – 7% for light rain and 14% for heavy rain. This is accompanied by a noticeable decrease in speed (between 4.5 and 8.2% in light and heavy rain conditions respectively). Capacity during daylight was found to be 12.8% greater than capacity during daybreak. Due to the limited sample size, as only 1 site was used, it was not possible to draw firm conclusions. However, the results regarding capacity reduction roughly correspond with previous studies (Prevedouros & Kongsil, 2003).

Other factors that have been considered include the proportion of HGVs and merge and diverge driver behaviour (Washburn et al, 2010). These factors are increasingly considered within ‘fixed’ capacity estimations; for example the NCHRP (NCHRP, 2008) provides adjustment factors for HCM values.

It is clear that due to a lack of studies, or in some cases due to a lack of consistency in findings between studies, there is still extremely limited understanding of the effect and level of impact of external factors on highway capacity. The majority of methodologies rely on a calculation of capacity as a percentile of observed flow rates although, as mentioned, the PLM method using a stochastic estimation model does allow investigation of external factors whilst also considering the ‘random’ element of capacity variance. Research has also highlighted that cultural differences in driver behaviour can play an important part in accounting for variance, making it difficult to compare international findings and highlighting the importance of conducting UK studies to inform applications for the UK network.
3 Methodology

The focus of this study had been to investigate and describe variability in motorway capacity. This section outlines the methodology developed to gather and analyse data in the form of a multiple linear regression, where the output of the study will be an equation of the form:

\[ y = a_n x_n \]

Where \( a_n \) and \( x_n \) are the constants and associated parameters respectively.

To investigate factors that affect capacity, potential factors were listed and then categorised into three logical areas; those that related to road geometry, those that related to driver behaviour and any other external factors.

The underlying principle behind this grouping is that flow breakdown is an observed aggregate behaviour of traffic and as such will be affected by the interactions between these three broad categories. The list of initial factors is included in Table 1 below.

<table>
<thead>
<tr>
<th>Road Geometry</th>
<th>Driver Behaviour</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane width</td>
<td>Gap acceptance</td>
<td>Rainfall</td>
</tr>
<tr>
<td>Link length</td>
<td>Headway adjustment</td>
<td>Visibility</td>
</tr>
<tr>
<td>Gradient</td>
<td>Speed variation</td>
<td>Time of day</td>
</tr>
<tr>
<td>Curvature</td>
<td>Use of brakes</td>
<td>Day type</td>
</tr>
<tr>
<td></td>
<td>Traffic composition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Traffic demand</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Commuter, Commercial or Leisure Traffic</td>
<td></td>
</tr>
</tbody>
</table>

3.1 Selecting Factors to Assess

The parameterisation and integration of all these factors into a unified model was beyond the scope of this project. A subset of these factors was therefore shortlisted for investigation. The criteria and process for selecting which factors to assess is outlined below:

- Factors should be easy and simple to parameterise and measure.
- Factors should have readily available data sources to allow analysis to be performed on the associated parameters.
- Geometric factors were not considered for this project as they have been covered extensively in the Highway Capacity Manual (HCM, 2000).

The shortlist based on the criteria above is included in Table 2 below.
Table 2 – Factors chosen for investigation

<table>
<thead>
<tr>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane changing</td>
<td>Rainfall</td>
</tr>
<tr>
<td>Traffic composition</td>
<td>Time of day</td>
</tr>
</tbody>
</table>

These factors were selected because they were the easiest to parameterise and had data that was readily available. Further modification was made to these factors based on the type of data that was available. This is discussed in more detail in the next section of this report.

3.2 From Factors to Parameters

The factors listed in Table 2 needed transforming into measureable parameters. Data sources relating to each of the factors were identified and their properties considered before parameters were defined.

On the UK Strategic Road Network, the Highways Agency collect extensive data from inductive loops embedded in the road surface. This data, collected from the Motorway Incident Detection and Automatic Signalling network, is referred to as MIDAS data for short. MIDAS data does not include lane changing information, as the loops are static in space and are spatially too infrequent to infer lane changing from in-lane data. It was therefore decided that the parameters of merge percentage and diverge percentage be considered as a proxy for lane changing. This is based on the evidence that most flow-breakdown occurs in close proximity to merge and diverge junctions, and very rarely mid-link.

MIDAS data also includes length classification information. This meant that traffic composition could be easily parameterised. Therefore HGV flow as a percentage of main carriageway flow was also considered so that an assessment could be made of how the proportion of HGVs might affect capacity.

Wunderground, an online weather information provider, catalogues historical weather data from its weather station network. Precipitation information, recorded by various weather stations was referred to during the analysis that was performed in this project.

The parameters investigated are listed in Table 3 below.

Table 3 – Parameters chosen for analysis

<table>
<thead>
<tr>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merge Percentage</td>
<td>Rainfall</td>
</tr>
<tr>
<td>Diverge Percentage</td>
<td>Light, Twilight, or Night</td>
</tr>
<tr>
<td>HGV Percentage</td>
<td></td>
</tr>
</tbody>
</table>
3.3 Site Selection

As described in section 2, the HCM defines capacity as the maximum flow rate that can be reasonably expected to traverse a highway link under prevailing roadway, traffic and control conditions. For the purpose of this study, capacity was defined as the peak 15-minute throughput prior to flow-breakdown occurring. The premise of this study is that in the time period before flow breakdown, deterministic factors interact to cause the onset of congestion.

The criteria for selecting sites, therefore, were as follows:

- Congestion must be recurrent, demonstrating flow-breakdown on at least 3 days a week on average.
- MIDAS data must be available for at least a year, both at the immediate location of the congestion seedpoint, as well as upstream and downstream of the surrounding junctions.
- MIDAS data must also be available on the surrounding slip roads.
- Weather information must be available from a representative site within a 5 mile radius of the congestion seedpoint.

Based on these criteria, congestion seedpoints on the UK Strategic Road Network (SRN) were catalogued and categorised. From this list, the most suitable sites were then selected. These sites were:

1. The M4 from junctions 19 to 20 in the eastbound direction.
2. The M25 from junction 11 to 12 in the anti-clockwise direction.
Figure 1 – Map of the M4 J19 – J20 (Image courtesy of Open Street Map)
3.4 Raw Data Descriptions

The raw data used in this study consists of:

- MIDAS Data (from the Highways Agency)
- Historical weather data (from Wunderground)

MIDAS data contains the following:

- 1-minute averaged speeds by lane
- 1-minute averaged occupancy by lane
- 1-minute traffic counts by lane (normalised to hourly flows)
- 1-minute classified vehicle counts by carriageway

Wunderground weather data simply recorded the normalised hourly equivalent rainfall, where the measurement period varied from station to station and also within station.
3.5 Building the Data Sets

Once two appropriate sites were selected, data for each site was then collected. To build data sets that would then be used for statistical analysis, MIDAS data for 2008 and 2009 was collected for each site. From this data two operations were performed:

1. Visual identification of congestion seedpoints and shockwaves using space-time plots generated from TRL’s Motorway Traffic Viewer (MTV) software.
2. Data processing to extract the relevant data for the 15 minute period prior to the onset of flow breakdown.

MTV is a tool that has been widely adopted in the UK and the concepts it uses in terms of traffic data visualisation have been even more widely adopted in the academic community. An example of an MTV space-time plot is shown in Figure 3.

![MTV plot showing the standing wave (seedpoint) and resulting shockwaves](image)

**Figure 3 – An MTV plot showing the standing wave (seedpoint) and resulting shockwaves**

Using this tool, each day for 2008 and 2009 was examined. The onset of flow breakdown can be identified as the left most part of the standing wave of low speed vehicles. The seedpoint is also identifiable through traffic immediately downstream travelling away from the standing wave at free-flow speeds. Through this method of seedpoint identification, the time at the start of flow breakdown was recorded to the nearest 5 minutes.
Once this initial data set was captured, a fully automated set of analysis was then performed using Microsoft SQL Server, which hosted all the MIDAS data that had been collected. This analysis is outlined below.

Table 4 – Description of how the data was handled

<table>
<thead>
<tr>
<th>Operation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finding the peak 15-minute throughput</td>
<td>The times at which flow breakdown occurred were derived from visual inspection. Inherent in this process is a lack of fineness. To improve the granularity, data from plus and minus 15 minutes of the initial times identified was analysed, where the maximum 15-minute throughput was extracted on a 1-minute rolling sum.</td>
</tr>
<tr>
<td>Calculating merge and diverge flow</td>
<td>As described in the site selection criteria above, sites had MIDAS data available on the slip roads as well as upstream and downstream of the adjacent merge and diverge slip roads. These data were used to derive the merge and diverge flows based on the difference between the flows upstream and downstream of the merge and diverge slips. These flows were then validated against the slip road flows for completeness. These flows were then described as a percentage of the 15-minute throughput prior to flow breakdown.</td>
</tr>
<tr>
<td>Calculating HGV flows</td>
<td>MIDAS data stores classified vehicle counts. The longest category (&gt;11.6m) was used and vehicles in the class were considered as HGVs for the purpose of this study. These counts were then described as a percentage of the 15-minute throughput prior to flow breakdown.</td>
</tr>
<tr>
<td>Precipitation</td>
<td>Normalised hourly precipitation values were looked up for the 15-minute throughput prior to flow breakdown.</td>
</tr>
<tr>
<td>Night, Twilight or Daylight</td>
<td>These values were looked up for each day that flow breakdown had occurred. The relevant state was then recorded based on the category in which the greatest period of the 15-minutes prior to flow breakdown fell.</td>
</tr>
</tbody>
</table>

Once the full data sets for each location were compiled, statistical analysis was performed. This analysis is outlined in the next section, along with the results.
4 Results

The data sets generated as described in the section above were subjected to statistical analysis in the form of multiple linear regression. The multiple linear regression performed considered all continuous variables only initially. Subsequently, categorical variables were then also considered, where precipitation was also transformed from a continuous variable (measured in mm of rainfall per hour), to a categorical variable (had it rained in the last 15 minutes or not).

The initial multiple linear regression included the following continuous variables:

- Hourly precipitation rate for the 15-minutes prior to the onset of flow breakdown
- Merge flow as a percentage of main carriageway flow
- Diverge flow as a percentage of main carriageway flow
- HGV flow as a percentage of main carriageway flow

The subsequent multiple regression, where categorical variables were also considered, included the following variables:

- Merge flow as a percentage of main carriageway flow
- Diverge flow as a percentage of main carriageway flow
- HGV flow as a percentage of main carriageway flow
- Precipitation as a categorical variable (wet or dry)
- Lighting conditions as a categorical variable (darkness, morning twilight, daylight, evening twilight)

Standard statistical tests were performed, where:

- $R^2$ is the proportion of the variance in the dependant variable accounted for by the model, which is a number between 0 and 1
- Sig. is the statistical significance of the individual variables and the model overall. A variable or model is considered significant if the significance is below 0.05 (conversely, 95% confidence interval), where the number refers to the probability that the variable explains none of the observed variance (the null hypothesis).
- Partial Eta Squared describes the size of the effect that the independent variables have on the dependant variable.
- F describes the ratio of explained variance to unexplained variance in the data set, where a high value indicates that the variable or model contributes to explaining the observed variance.
- Beta (for the linear regression) indicates the standardised coefficients, allowing their effects to be compared with each other.
- t is similar to F but is only relevant to linear regression analysis.

The data points collected for each site are shown in Table 5.
Table 5: Summary of data points

<table>
<thead>
<tr>
<th>Category</th>
<th>M4</th>
<th>M25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Days Investigated</td>
<td>231</td>
<td>286</td>
</tr>
<tr>
<td>Dry</td>
<td>195</td>
<td>279</td>
</tr>
<tr>
<td>Wet</td>
<td>36</td>
<td>7</td>
</tr>
<tr>
<td>Night</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td>Morning Twilight</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Light</td>
<td>150</td>
<td>251</td>
</tr>
<tr>
<td>Evening Twilight</td>
<td>46</td>
<td>25</td>
</tr>
</tbody>
</table>

The results for each site are discussed below.

### 4.1 M4 Eastbound J19 – J20

Table 6 includes the results from the multiple linear regression performed with the continuous variables shown in the first column. The coefficients reveal that HGVs, merging traffic and precipitation all reduce capacity while diverging traffic contributes to an increase in capacity. This is an unusual result as the expectation at the outset was that all lane changing, induced by merging or diverging traffic, would reduce capacity. This supposed increase in capacity as more traffic diverges could be related to the specific road geometry at this site.

HGVs contribute the most to the observed variability in capacity, followed by merge percentage. All factors are statistically significant, except diverge percentage, which is on the borderline of the 95% confidence interval.

Table 6: Multiple linear regression for M4 J19-J20

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Co-efficient</th>
<th>Std. Error</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1964.49</td>
<td>126.70</td>
<td>-</td>
<td>15.51</td>
<td>0</td>
</tr>
<tr>
<td>HGV %</td>
<td>-54.47</td>
<td>6.41</td>
<td>-0.493</td>
<td>-8.50</td>
<td>0</td>
</tr>
<tr>
<td>Diverge %</td>
<td>3.59</td>
<td>1.82</td>
<td>0.111</td>
<td>1.97</td>
<td>0.05</td>
</tr>
<tr>
<td>Merge %</td>
<td>-7.26</td>
<td>1.49</td>
<td>-0.280</td>
<td>-4.88</td>
<td>0</td>
</tr>
<tr>
<td>Hourly Precipitation (mm)</td>
<td>-43.71</td>
<td>12.19</td>
<td>-0.202</td>
<td>-3.59</td>
<td>0</td>
</tr>
</tbody>
</table>
This analysis was then extended to include the lighting conditions, darkness, twilight and daylight. Precipitation was also transformed from a continuous variable to a categorical variable which represented whether it had rained or not in the 15 minutes prior to flow breakdown. The results of this analysis yielded an $R^2$ of 0.463. The full results can be seen in Table 7 below.

### Table 7: Analysis of variance for M4 J19-J20

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>Significance</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>9</td>
<td>21.188</td>
<td>0.000</td>
<td>0.463</td>
</tr>
<tr>
<td>Intercept</td>
<td>1</td>
<td>255.263</td>
<td>0.000</td>
<td>0.536</td>
</tr>
<tr>
<td>HGV percentage</td>
<td>1</td>
<td>74.502</td>
<td>0.000</td>
<td>0.252</td>
</tr>
<tr>
<td>Diverge percentage</td>
<td>1</td>
<td>6.728</td>
<td>0.010</td>
<td>0.030</td>
</tr>
<tr>
<td>Merge percentage</td>
<td>1</td>
<td>36.122</td>
<td>0.000</td>
<td>0.140</td>
</tr>
<tr>
<td>Precipitation</td>
<td>1</td>
<td>9.816</td>
<td>0.002</td>
<td>0.043</td>
</tr>
<tr>
<td>Lighting</td>
<td>3</td>
<td>5.794</td>
<td>0.001</td>
<td>0.073</td>
</tr>
<tr>
<td>Precipitation * Lighting</td>
<td>2</td>
<td>0.127</td>
<td>0.881</td>
<td>0.001</td>
</tr>
</tbody>
</table>

The important outputs of the UNIvariate ANalysis Of Variance (UNIANOVA) are the F test, the significance and the partial eta squared values. The inclusion of lighting conditions and the transformation of the precipitation variable from continuous to categorical significantly improve the fit of the model. All factors are significant and there is no interaction between lighting conditions and precipitation (WetorDry*Light) which would reduce the appropriateness of the analysis.

In comparison to the multiple linear regression, the HGV percentage and the merge percentage still have the largest bearing on the observed capacity. Interestingly, lighting conditions appear to impact capacity rather more than precipitation and diverge percentage.

### 4.2 M25 Anti-Clockwise J11 – J12

The same set of analysis was performed for this site as for the M4 Eastbound J19 - J20. The initial multiple linear regression yielded an $R^2$ of 0.227 and the coefficients are described in Table 8 below.
Table 8: Multiple linear regression for M25 J11-J12

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Co-efficient</th>
<th>Std. Error</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1427.36</td>
<td>75.96</td>
<td>-</td>
<td>18.79</td>
<td>0</td>
</tr>
<tr>
<td>HGV %</td>
<td>12.02</td>
<td>3.68</td>
<td>0.18</td>
<td>3.27</td>
<td>0.001</td>
</tr>
<tr>
<td>Diverge %</td>
<td>13.67</td>
<td>1.83</td>
<td>0.40</td>
<td>7.46</td>
<td>0</td>
</tr>
<tr>
<td>Merge %</td>
<td>-0.54</td>
<td>2.20</td>
<td>-0.014</td>
<td>-0.243</td>
<td>0.81</td>
</tr>
<tr>
<td>Hourly Precipitation (mm)</td>
<td>-81.17</td>
<td>26.03</td>
<td>-0.16</td>
<td>-3.12</td>
<td>0.002</td>
</tr>
</tbody>
</table>

The results from the M25 site are less convincing than the M4 site. This is potentially due to there being significantly less data points where precipitation or lighting conditions may have been a factor.

A result of note from the M25 multiple linear regression analysis is that increases in HGV percentage appear to increase capacity. This is a highly unexpected result as prevailing thinking, as well as the results from the M4 site, suggests that HGVs reduce capacity, due to occupying more road space than cars and they also travel at reduced speeds compared to normal traffic, due to legislative restrictions.

Another point worth noting from these results is that merge percentage plays no part in the variability in capacity, which is the converse to the M4 site where merge percentage was a statistically significant factor.
This analysis was then extended in the same way as for the M4 Eastbound J19 – J20, yielding an $R^2$ of 0.274. The full results can be seen in Table 8 below.

**Table 9: Analysis of variance for M25 J11-J12**

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>Significance</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>9</td>
<td>11.547</td>
<td>0.000</td>
<td>0.274</td>
</tr>
<tr>
<td>Intercept</td>
<td>1</td>
<td>244.480</td>
<td>0.000</td>
<td>0.470</td>
</tr>
<tr>
<td>HGV percentage</td>
<td>1</td>
<td>10.457</td>
<td>0.001</td>
<td>0.037</td>
</tr>
<tr>
<td>Diverge percentage</td>
<td>1</td>
<td>71.370</td>
<td>0.000</td>
<td>0.205</td>
</tr>
<tr>
<td>Merge percentage</td>
<td>1</td>
<td>0.032</td>
<td>0.858</td>
<td>0.000</td>
</tr>
<tr>
<td>Precipitation</td>
<td>1</td>
<td>1.700</td>
<td>0.193</td>
<td>0.006</td>
</tr>
<tr>
<td>Lighting</td>
<td>3</td>
<td>3.424</td>
<td>0.018</td>
<td>0.036</td>
</tr>
<tr>
<td>Precipitation * Lighting</td>
<td>2</td>
<td>0.115</td>
<td>0.891</td>
<td>0.001</td>
</tr>
</tbody>
</table>

The analysis above shows an improvement of the fit of the model over multiple linear regression for the M25. This is in part due to the lack of data points for each independent variable compared to the M4 site. However, there are some interesting points that should be drawn out of the analysis above:

- Precipitation becomes statistically insignificant in terms of its impact on capacity.
- There is no interaction between precipitation and lighting conditions (Wet or Dry * Light) as with the M4 site.
- Diverge percentage appears to have the greatest impact on capacity at this site.
5 Discussion

The original hypothesis at the outset of this study was that a linear equation could be constructed that could describe the variability in capacity that is observed at many points on the UK SRN, using readily available traffic and weather data.

The results have demonstrated the following:

1. There is some merit in the novel approach that has been developed in this study.
2. The interaction of factors is site specific in a way that is not yet clearly understood.
3. The number of factors and/or the resolution of data are not sufficient to fully describe the variance in capacity.

The results have shown some interesting details which are summarised below:

- At each location and for each type of analysis, only the merge or diverge flows are significant, with the respective parameter showing a lesser degree of significance. This would indicate that flow breakdown is contributed to from lane changing due to either merging traffic, or diverging traffic, not both.
- There is a noticeable difference in the amount of capacity variance explained by the analysis at each site, where less variance is explained for the M25 site than the M4 site. This indicates that the parameters investigated do not account for any site specific factors. It is not known whether these can be purely explained through geometric differences or if there are other factors that have not been taken account of.
- The relative significance of HGVs is different for both sites and even contributes to an increase in capacity on the M25. This is an counter intuitive effect and could form the basis of further study.

The novelty in the approach that was taken for this study is twofold. Firstly, capacity has long been considered constant and that constant is defined somewhat loosely as the flow which can be sustained under prevailing conditions. Secondly, the approach was developed to test if readily available data could be used to test the original hypothesis above, thereby avoiding costly instrumentation and data collection regimes. This study has re-cast capacity as a variable property, which is determined by the interaction of factors prior to the onset of congestion.

The values of $R^2$ obtained for each site demonstrate that, to a varying degree, a significant proportion of the variability in capacity can be explained using the simple methods developed in this study. However, there are significant shortcomings in the methods and data that were employed in this study, due to the constraints that were imposed at the outset, namely, the requirement that only readily available data was to be used.
Listed below in are relationships that cannot be verified or tested using this methodology. This list is followed by a further list describing potential improvements and next steps that could be taken to improve on the results obtained in this study.

Relationships and concepts not testable or verifiable by using this technique:

- The use of merge and diverge flows as proxies for lane changing
- How factors interact to vary the observed capacity
- The physical mechanisms that induce congestion at a microscopic level
- Discrimination between types of driver (commuter or leisure traffic for instance) and any effect they might have on the variability of capacity
- How geometric factors affect capacity

Further enhancements that could improve and refine the method and results generated in this study could be:

- Perform analysis on more sites
- Collect higher resolution data
- Collect better quality data
- Introduce more parameters
- Improve the rigor in defining congestion seedpoints
6 Conclusion

This study has investigated factors that affect motorway capacity. Motorway capacity has long been considered as a fixed value. However, through extensive experience working with traffic data, as well as recent literature exploring the issue, capacity is now considered to be variable.

Some studies have focused on refining the traditional speed-flow curve, through better approximations of the curve’s functional structure as well as calibrating the curve to observed data. Other studies have considered the variability in capacity to be purely stochastic.

In this study, a different approach has been taken where capacity is defined as the 15-minute throughput prior to flow breakdown. Factors that could affect capacity have also been investigated over the same 15-minute period. Statistical analysis was performed and partial correlations between the factors investigated were observed.

The outcomes of this study demonstrate that estimating motorway capacity is still not a trivial task. The methods developed here demonstrate promise and further research could utilise higher resolution data and investigate more motorway sites.

The potential application of the findings of this study are still some way off, but with further work it may be possible to construct an algorithm that evaluates real-time data and provides a so called “on-line” estimate of the capacity of the motorway at the desired location.

Further enhancements could also be made by introducing a predictive element into the algorithm. This would be extremely useful for pro-actively managing the motorway network through schemes such as managed motorways, where the hard shoulder could be opened in due time to avoid flow breakdown, or variable mandatory speed limits could be set more effectively so as to not slow drivers unnecessarily.

This study has also highlighted the need for a rigorous and systematic methodology for identifying and characterising flow breakdown.

In conclusion, this study has achieved the objectives that were originally set out, namely:

1. Construct an operational definition of capacity
2. Investigate factors that affect capacity using readily available data

While the results were not as conclusive as was originally hoped, a basis for further study has been produced and suggestions for how this method could be improved have been made.
References


This report outlines a novel approach taken to characterise the variability in traffic throughput prior to the onset of congestion by investigating factors that could potentially affect capacity. Two motorway sites experiencing recurrent congestion in the UK were selected for investigation, the M4 eastbound between junctions 19 and 20, and the M25 anti-clockwise between junctions 11 and 12. The factors assessed include merge and diverge flows upstream and downstream of the congestion seedpoint, HGV flows, precipitation and lighting conditions. Capacity was defined as the 15 minute throughput prior to the onset of congestion, and all other factors were measured for the same 15 minute period. Multiple linear regression analysis revealed an R2 value of 0.46 at the M4 site and 0.27 at the M25 site. This demonstrates that the methodology developed shows some merit and that further investigation is warranted where higher resolution data should be used and more parameters should be considered. The potential application of this research lies in the operation of traffic management schemes such controlled motorways, where speed limits could be set based on a number of factors, such as those investigated above.

Other titles from this subject area

PPR193 | Channel corridor and community transport study – final report. A Davies, C Burke, K. Townley and S Reid. 2007
PPR196 | Modelling congestion caused by abnormal loads. N Taylor. 2007
PPR106 | In-depth interviews with van drivers and managers of van drivers. B Lang. 2006
PPR113 | Literature review on van use in the UK. B Lang and L Rehm. 2006