A flexible approach to motorway control

Keith McCabe
Principal Consultant
Intelligent Transport Systems
Atkins

Andy Riley
Senior Engineer
Intelligent Transport Systems
Atkins

Abstract

This paper describes recent work undertaken by the UK Highways Agency to consolidate the experience gained from the implementation of Controlled Motorway and Ramp Metering Schemes, in addition to the innovative Active Traffic Management Pilot Scheme. A flexible approach to motorway control is described that can address a particular problem at a specific location at a given time via combined and coherent use of available traffic control measures based on Intelligent Transport Systems (ITS). The form of control will be based on the available physical infrastructure (e.g. gantry or post mounted signals) used in different ways at different times. The traffic problem may change over time, which requires flexibility in the control algorithms to cope with changing circumstances.

Introduction

The aim of this paper is to provide a summary of the progress on the development of variable speed limits in the UK. The paper provides a short review of the results from the Controlled Motorways (CM) pilot scheme on M25 and considers proposals for possible enhancements. Results of a theoretical study into combining variable speed limits with ramp metering are also presented.

There are three main reasons to control motorway traffic using mandatory or advisory variable speed limits and/or ramp metering, namely:

- To improve efficiency – reducing congestion and journey times
- To improve safety – reduce the number and severity of accidents
- To produce environmental benefits – reduce emissions.

A positive result in all three areas is the aim, however, under particular circumstances it may be sensible to prioritise one form of benefit over another (e.g. a reduction in speed to reduce accidents and emissions may result in increased journey times). The aim of this paper is to discuss the relationship between particular control interventions and the impact on efficiency, safety and environment.

Controlled Motorways on M25

The Controlled Motorways pilot scheme on the M25 began in 1995 and a business case for this system was completed in 2002[1] The original main aims of the scheme were to provide a smooth traffic flow, to improve journey times, journey time reliability and lane utilisation and to reduce the incidence of stop-start driving and the stress of driving. Subsequent notable achievements also included the development of proven technology and the reduction in environmental problems (i.e. noise and pollution).

During the monitoring of CM there was an improvement in journey times on the clockwise carriageway in the morning peak period. The observed behaviour was attributed to drivers becoming familiar with the CM and driving more smoothly to prevent flow breakdown. Other areas of the M25 did not show as much improvement in journey time performance because congestion had increased, which was attributed to yearly increases in flow levels (more road users).

The number of shockwaves decreased between 1995 and 2002, with a reduction from a typical seven shockwaves per morning rush hour down to a typical five. This observed behaviour could be attributed to the smooth driving behaviour or improvements to the control system over this period. It contributes to the improvements in journey time and queuing despite the increased throughput.

Evidence of safety improvements were demonstrated by the studies of injury accidents. The variable speed limit system on the M25, known as controlled motorways, has resulted in steadier and less stressful journeys and thus reduced the number and severity of accidents, "injury" accidents by 10% and "damage only" accidents by 30%.

Emissions have decreased overall by between 2% and 8%. The weekday traffic noise adjacent to the scheme has reduced by 0.7 decibels.

Lane utilisation and headway distribution had been improved. Lane utilisation became more balanced making better use of the road space, with a reduction in the number of very short headways (pre 1995).

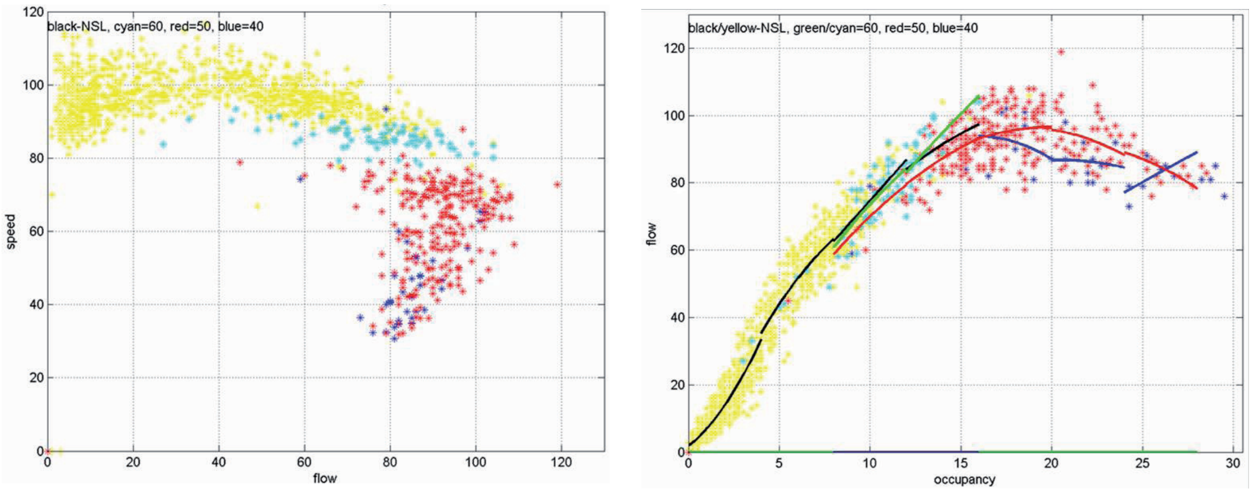

Figure 1 - M42 study mandatory speed limits - speed flow diagram and flow occupancy diagrams

The aim to reduce the stress of driving appears to have been achieved, with the primary road users indicating a positive response to the CM scheme. The driver opinion survey carried out in the first year of operation indicated that over half the respondents noticed an overall improvement; two thirds would like to see the scheme extended to other stretches of motorway. The survey indicated very clearly that the automatic speed cameras were essential to ensure drivers' compliance with speed limits.

Controlled Motorway on M42

A study was undertaken to investigate the effect of mandatory Variable Speed Limits (VSL) on motorway throughput and traffic flow efficiency using traffic data from the M42 motorway Active Traffic Management section (Junctions 3A to 7) as well as to determine possible proposals for improvement of the VSL control algorithm used on M42 as well as in other UK motorways. Speed limits are activated/modified/de-activated when flow and/or speed measurements cross pre-set thresholds - full details can be found in (2).

A mandatory 3 lane VSL has been in operation in the aforementioned stretch of M42 since late 2005. Threshold values have been fine tuned over a period since the mandatory speed limits went operational, to ensure safe operation and consistent signals to motorists. The major tuning activities were completed in January 2006.

Traffic data (i.e. flow/occupancy/speed measurements as well as speed limit records) were available for this study from two periods: 2005 period traffic data corresponding to the VSL control algorithm application before the major fine tuning and 2006 period traffic data when the control algorithm was operating using the fine-tuned threshold values.

Validation of the traffic data was performed to remove faulty data or data that corresponded to low traffic demand, incidents of significant time duration. Visual inspection as well as curve fitting methods were used for the purposes of the evaluation analysis.

The signals were activated significantly less frequently during the period before the major fine tuning, hence the data set contained many non-VSL data at dense and critical traffic conditions than those included in the 2006 period.

A main focus of this work was on verifying the impact on the shape of the flow-occupancy diagram, see Figure 1, (or, equivalently, of the speed-flow diagram); more specifically VSL are thought to:

(i) Lower the slope of the flow-occupancy diagram at under critical occupancies

(ii) Shift the critical occupancy (at which maximum flow occurs) to higher values and possibly increase the flow capacity.

Notice that, if (ii) is correct, then VSL control (if properly operated) can have the effect of allowing a higher throughput at critical or slightly overcritical traffic conditions as compared to the no control case, even if the VSL affected flow capacity does not actually increase. On the other hand, if (i) is correct, then VSL could increase journey times for less congested traffic conditions if applied with less appropriate threshold values.

Main findings

Speed limits - when applied at less congested occupancies - decrease the slope of the flow-occupancy diagram (see (i) above). Moreover, the smaller the imposed speed limit, the larger the decrease in the slope of the flow-occupancy diagram.

Importantly, the VSL-affected flow-occupancy curve indeed crosses (at least for some VSL) the non-VSL curve, shifting the critical occupancy to higher values in the flow-occupancy diagram (see (ii) above). The major cross points were found to lie on or slightly beyond the non-VSL critical occupancy.

Regarding the potential increase of flow capacity, the data analysis was rather inconclusive, as at some locations a slight increase is indeed visible while at other locations no increase can be observed.

On "rainy" days, the non-VSL flow capacity and the critical speed are reduced by some 10%.

The same quantities were observed to change quite substantially from day to day (even for the same location) without any obvious reason (stochastic effects). In contrast, the non-VSL critical occupancy was found to be rather insensitive to weather conditions and stochastic effects.

INTELLIGENT TRANSPORT SYSTEMS

The speed limit of 50mph was found to be the main contributor to modified aggregate traffic flow behaviour that could be exploited towards more efficient traffic flow. A speed limit of 60mph was found to have a rather moderate impact; while a speed limit of 40mph is used at high occupancies in the interest of traffic safety rather than traffic flow efficiency.

In the data period before the major fine tuning, the speed limits were activated either "too early" or "too late", leading to cases where either speed limits were imposed although traffic was under critical; or speed limits were imposed after congestion had occurred, hence having little effect on congestion avoidance. In contrast, the behaviour of the VSL control algorithm after the major fine tuning is more robust, preserving similar behaviour at different days.

Algorithm issues

The algorithm currently used for mandatory VSL is based on a rule-based method designed to harmonise traffic speeds and to reduce the severity of congestion. The control system responds to one minute traffic speed/flow data measured by the MIDAS loop detectors. Speed limits are reduced if the measured flow goes above a pre-defined Rising Flow Threshold or the measured speed goes below a pre-defined Falling Speed Threshold. Likewise speed limits are increased if the measured flow goes below a pre-defined Falling Flow Threshold and the measured speed goes above a pre-defined Rising Speed Threshold. The speed threshold is essentially used to ensure that the signals remain set when flow drop in "stop & go" driving conditions. On this basis the algorithm is based on a practical method.

Whether such an algorithm would be successful in preventing traffic breakdown largely depends on the methodology used to determine Rising/Falling Speed and Flow Thresholds.

The methodology originally used for the determination of speed/flow thresholds for the M25 CM was to apply speed limits at times when average traffic speeds were approaching these speeds anyway with a high likelihood of flow breakdown occurring. This was based on the assumption that the reduction in actual traffic speeds due to the speed limits is likely to be small; hence if there is an increase in delay to drivers, this is also going to be small. On this basis, the initial algorithm was based on a homogenisation approach.

The M25 CM has been subject to continuous monitoring and fine tuning over recent years. As a result, there have been a number of changes to the algorithm parameters which may have positively affected the impacts of the algorithm. The monitoring results suggest that the system has produced positive impacts on flow breakdown, journey time reliability, safety and air pollution but neutral or negative impacts on journey time and throughput.

The review of the algorithms indicated that limitations can be attributed to the current algorithm threshold setting constraints, location dependencies and general complexity of combined parameters. In addition, MIDAS capability is generally underused, but there are limitations at multiple MIDAS sites. There are also high speed limit settings in heavy congestion:

- The algorithm does not allow the use of separate thresholds for different time of day (e.g. am/pm), day of the week, different weather or lighting conditions, seasonal variations or bank holidays
- The thresholds apply to the total flow across the carriageway lanes - separate thresholds per lane are not available.

The algorithm parameters have to reflect a wide range of factors involved in forming different traffic situations e.g. road geometry (i.e. number of carriageway lanes, gradient, distance between the junctions) and traffic characteristics (e.g. traffic composition, lane utilisation, lane changing frequency, proportion of merging and diverging traffic).

It is very difficult to reflect the combined effects all these factors in a limited number of control parameters (e.g. speed/flow thresholds). As the tuning of the system is done manually, it is a great challenge to correlate all thresholds and algorithm parameters in this way.

No provisions are currently made for the use of lower than 40mph speed limits for incident and congestion management; severely congested areas are largely uncontrolled, therefore. Customers often get frustrated when speed limits of 40mph or higher are posted when they are actually doing much lower speeds, despite the fact the speed limits represent the maximum speed at which they should travel. Any improvement in this respect may contribute to improved driver acceptance of the system.

Although the VSL control algorithm behaves more robustly and efficiently after the major fine-tuning, it still faces some limitations mainly due to the sensitivity of the appropriate thresholds to weather and stochastic traffic conditions. Thus, the VSL control algorithm in several cases does not impose the appropriate speed limits although the traffic state has entered the unstable overcritical occupancy-flow region, thus leading to strong flow variations.

The remaining drawbacks of the VSL algorithm are mainly due to the use of absolute threshold values of flows or speeds for VSL activation (e.g. VSL set to 60mph if measured flow higher than a flow threshold); while in reality the appropriate critical thresholds may change from day-to-day even under the same weather and light conditions.

Possible modifications to the algorithm

There is significant potential for the improvement of the algorithms in the longer term by changing the MIDAS Site Data to allow the use of different thresholds for different weather and lighting conditions. This would also require linking the MIDAS Subsystem to the Meteorological (MET) Subsystem to send the alerts related to the changes in weather/lighting conditions to the MIDAS Outstations. For this purpose, the MIDAS Site Data should also be modified to allow for using different thresholds for different weather/lighting conditions.

In view of changing real absolute thresholds, a more robust and efficient algorithm might result if the decisions on VSL-activation are based on estimated slopes in the flow-occupancy diagram. In other words, while, e.g. the motorway's flow capacity value may vary from day to day, the fact that flow capacity values occur at zero slope does not change.

An adaptive algorithm - based on the algorithm for critical occupancy estimation developed in a previous study[3] - has been preliminarily tested for slope estimation using the M42 traffic data. The algorithm parameters were chosen to be the same for all motorway locations, and the preliminary tests indicate that the adaptive algorithm may be more robust than the current VSL control algorithm even after the major fine tuning. More precisely, the adaptive algorithm seems to avoid situations where speed limits are imposed "too early" or "too late". Quite importantly, the adaptive algorithm is expected to avoid the time and effort consuming process of fine tuning the location dependent VSL control algorithm thresholds.

As a longer term development, the Highways Agency is looking towards a network oriented traffic control model for the management of incidents and congestion on a motorway corridor or network. For control systems, this would offer features such as coordination, prediction, integration with other systems, control by other parameters (e.g. environmental) and macro level area traffic management. This model is called the Network ATM Supervisory Subsystem (NASS)[4].

Network oriented traffic control has several advantages compared to local control. For example, the capacity of a route is the same as the capacity of its most constrained point. Solving a single local traffic problem can only result, therefore, in the vehicles travelling faster to downstream congestion. The same number of vehicles have to pass the downstream bottleneck (with a given capacity). In such a case, the average travel time on the network level will be determined by the capacity of the downstream congestion point. A network oriented approach would take this into account and, if possible, would resolve both congested areas.

Using network oriented traffic control, it would be possible to integrate the control system with other motorway traffic management measures such as Ramp Metering (RM), Mainline Metering, Strategic Traffic Management, etc. The synergistic effects of the integrated systems would increase the overall benefits of the HA Traffic Management Systems.

Multiple objectives could be achieved using network oriented traffic control (for example in pollution and noise control as well as for incident and congestion management).

The inflow to a congested area could also be limited to a level that is less than outflow of the area using network oriented traffic control. The overall delays on the network could be maintained below the non-controlled situation. This is a fast and effective approach to the alleviation of severe congestion and to prevent secondary breakdowns.

Theoretical design for combining CM with RM

An examination of the theoretical impact of the two systems was carried out to determine the likely impacts of operating CM and RM on the same section of motorway. Research into two previously conducted studies was considered. These studies simulated the use of control strategies of Model Predictive Control (MPC)[5] and Stackelberg Game Theory[6]. Whilst it is not possible to directly compare the two strategies against each other (as they were tested under different conditions), the review showed definite theoretical benefits in terms of journey time.

Combining the systems was shown to be better than using either RM or CM singularly. Hence it is considered that linking the control systems could be beneficial.

Complementary RM and CM

The existing operation of RM and CM could act in a complementary way when deployed on the same section of motorway flow breakdown is delayed and platoon management can be achieved.

In the period when the flow is reaching capacity the operation of RM delays flow breakdown allowing the upstream speeds to remain at 50mph for longer. The benefits of increased throughput due to RM and higher harmonised upstream speeds due to CM are potentially enhanced by operating both control measures on the same section of motorway.

In the period after flow breakdown RM improves the merge behaviour by producing small platoons of traffic, at the same time, CM protects the back of any queue that has formed (setting 40mph using HIOCC), the resultant impact on safety of both control measures operating on the same section has the potential to provide enhanced safety benefits.

Interaction of RM and CM

There are two situations where the interaction between the two control systems on the same section of motorway needs to be carefully considered:

(1) When the speeds are rising as a result of RM holding back the slip road traffic, there is the potential to interact with a purely flow based signal setting (60 and 50mph) on CM. The interaction between the two systems in this area of operation using the existing control algorithms requires careful consideration as there is the potential either to delay the flow recovery process or have inappropriate speeds for the traffic conditions.

(2) The potential impact of setting mandatory 60/50/40mph speed limits on the value of critical occupancy also requires attention. Difficulties may arise from different values of critical occupancy when different speed limits are set. Potentially this could lead to inefficient operation of RM.

Support of CM by RM

RM could assist the current operation of CM when the CM system is setting 60/50mph in congested and heavily congested periods. RM could potentially control the flow to complement the setting of 50mph and 60mph signals on the main line. This could result in an improved speed harmonisation allowing a higher level of service (higher main line speed/ shorter journey time) to be achieved for longer, with only short delays to slip road traffic. This mode of operation would require sufficient acceleration distance from the stop line to the merge to enable the vehicle to reach merge speeds of between 50 and 60mph. In a number of circumstances this may require works to extend the length of the slip road. In such a circumstance a site specific business case would be required prior to undertaking the work.

Support of RM by CM

CM could assist the operation of RM when the slip road queue is reaching the slip road storage capacity. In theory setting a 30mph signal upstream of the merge area could reduce the mainline flow to a level that would allow more slip road traffic to join without exceeding the downstream capacity. This is currently a modelled theoretical concept that in theory produces benefits in terms of allowing the downstream flow to run at capacity for longer without the bottleneck causing flow breakdown. When this has been modelled a significant reduction in the number of main line shockwaves has been noted, with resulting accident and journey time reliability benefits.

Overview of combined RM/CM operation

Overall the benefits of CM and RM that were demonstrated in the pilots are achievable when the two control systems operate on the same section of motorway. The areas of concern that needs to be examined are during flow recovery and when inappropriate speeds are set on the main line. The first of these problems may be overcome with the enhancements to CM because the problem is most likely to occur with a CM algorithm based purely on flow level settings. The second could be addressed through the use of an adaptive system for critical occupancy for the RM (possibly AD-ALINEA). Careful selection of a test site may enable an early demonstration of the overall benefits of the two systems operating on the same section of road.

The RM and CM could be enhanced to improve combined operation. RM could improve the operation of CM at 60 and 50mph. This would require an enhancement to the RM control system and careful site selection (to allow vehicles to achieve merge speeds of greater than 50mph). A link from CM to RM to let the RM system know when the 60mph and 50mph limits are in operation would be a useful feature of this system. A carefully selected trial site where the benefits of the enhancement could be isolated would be required prior to a wider implementation of this enhancement. CM could improve RM operation by managing the flow at capacity and after flow breakdown by setting 30mph limits on gantry signals. This would require a fully integrated control system that could be a subset of the NASS control system using the principles in Stackleberg Game or Model Predictive Control systems. This type of system could be developed to "lock in the benefits of motorway widening" and possibly form part of a broader access management/ motorway control system.

Conclusions

Studies into the operation of controlled motorways have shown that potentially benefits can be achieved with reduced congestion, reduced accidents and reduced emissions. However, it should be noted that the current methodology of using speed and flow thresholds requires substantial fine tuning to achieve efficient results. A system that automatically fine tunes parameter settings potentially gives long term efficiency gains.

A fully integrated system using variable speed limits linked to ramp metering/access management potentially can manage a motorway to address a range of local priorities (congestion, safety or environmental). Reducing speeds to improve safety and produce environmental benefits can potentially increase journey times. However, the safety and environmental benefits can be enhanced by linking to ramp metering/access management to maintain constant speeds and improve journey time reliability.

Overall the principle of using variable speed limits and ramp metering to improve efficiency, safety and environmental impacts has been theoretically proven. To maximise the benefits of implementations the automation of parameter setting should result in more consistent outcomes.

Acknowledgements

The work described in this paper is sponsored by the Highways Agency. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Highway Agency.

The Technical University of Crete Dynamic Systems and Simulation Laboratory undertook the analysis of the data on variable speed limits on the M42 on behalf of the Highways Agency.

References

1. Rees, T; Harbord, B; Dixon, C; Abou-Rahme, N. Speed Control and Incident Detection on the M25 Controlled Motorway (Summary of Results 1995-2002), Report, Transport Research Laboratory, PPR033.
2. Markos Papageorgiou, Elias Kosmatopoulos, Matt Protopapas, Ioannis Papamichail (2006), "Evaluation of the effects of variable speed limits on motorway traffic using M42 traffic data, DSSL Technical University of Crete Internal Report 2005-25.
3. Kosmatopoulos, E., Papageorgiou, M., Monolis, D., Hayden, J., Higginson, R., McCabe, K., Rayman, N. (2006) Real Time estimation of critical occupancy for maximum motorway throughput, Transportation Research Record, in press
4. NASS fact sheets #1 - #3, Available from IPL Ltd.
5. Hegyi, A. (2004). Model Predictive Control for Integrating Traffic Control Measures. TRAIL Thesis Series T2004/2, The Netherlands, TRAIL Research School.
6. Zhenlong, L. (2005) Optimal Coordination of Variable Speed and Ramp Metering based on Stackelberg Game. 12th World Congress on Intelligent Transport Systems. Nov. 2005, San Francisco, USA.