Transportation Research Part C 19 (2011) 206–217

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Analysis and modelling of traffic flow under variable speed limits

B.G. Heydecker *, J.D. Addison

Centre for Transport Studies, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Keywords:
Traffic flow
Traffic modelling
Motorway control
Variable speed limit
Speed–density

ABSTRACT

We develop the relationship between speed and density to analyse the flow of traffic under the operation of variable speed limits. By statistical analysis of traffic data from the UK motorway network, we find that the functional form preferred for this does not have an explicit jam density that will induce zero speed in traffic. We deduce that it is zero speed that induces jam density in traffic rather than vice versa, so that the direction of causality between speed and density differs according to circumstances. We develop an approach to modelling traffic in light of this. We apply this to analyse speed control as a traffic management measure and show how it can be used to estimate the effect of speed management on road capacity.

1. Introduction

Variable speed limits are used on motorway roads to manage traffic flows with the intention of improving capacity and hence throughput of traffic. In order for this control of traffic speeds to affect the quality of traffic flow, speed should have a causal influence on the traffic flow. This is distinct from the more usual approach to modelling of traffic flow in which density is taken to determine speed, and hence flow; there is some accordance with models that use different relationships between speed and density according to circumstance. In this paper, we investigate the relationships among speed, flow and density that are apparent in motorway traffic data under each of normal and speed-controlled operation. We discover evidence for circumstances in which speed is the influential quantity – typically under speed-controlled conditions – and others under which density is the influential quantity – typically in low-density conditions.

Descriptions of traffic flow at a point that are based upon the three quantities of speed (v), flow (q) and density (k) are well established. Analysis of their definitions leads in the case of homogeneous traffic directly to the formula that has become known as the *fundamental equation* (Greenshields, 1935):

$$q = kv, \tag{1}$$

where q is flow, measured in vehicles per unit of time, k is density, measured in vehicles per length of road, and v is space-mean speed (Wardrop, 1952), measured in length per unit of time.

This fundamental equation can be supplemented by adoption of a model relationship to describe the association between speed and density of the form

$$v = f(k). \tag{2}$$

Traffic flow can be analysed effectively following on from these relationships (see, for example, Wardrop, 1952; Lighthill and Whitham, 1955; Leutzbach, 1988; Daganzo, 1997). When Eq. (2) is populated with an explicit form, any one of the three variables (speed v, flow q, and density k) can be used to calculate values for the other two, subject to the ambiguity between

* Corresponding author. Tel.: +44 20 7679 1553; fax: +44 20 7679 1567.
E-mail addresses: ben@transport.ucl.ac.uk (B.G. Heydecker), puff@transport.ucl.ac.uk (J.D. Addison).

0968-090X/$ - see front matter
doi:10.1016/j.trc.2010.05.008

free-flow and congested values of flow q. Higher order models of traffic flow have been developed that use (2) as an equilibrium relationship for steady-state flows, towards which traffic speeds will relax over time (Zhang et al., 2001). Several authors have suggested that different relationships should apply according to the circumstances (for example, Hall et al., 1992; Zhang, 1999); typical variations in circumstances are according to accelerating or decelerating traffic, whilst in the present paper we consider the circumstances according to the prevailing speed limit. Whichever of these model developments is pursued, the purpose is to describe traffic flow, to project from current traffic states to future ones, and to develop a deeper understanding of the behaviour of traffic as a medium and its gross response to management measures.

Whilst Eq. (1) follows inexorably from the definitions of the quantities that are involved, equations of the form (2) represent descriptive models of traffic behaviour. Two consequences follow from this in respect of Eq. (2): first that the choice of the function $f(\cdot)$ is *ad hoc*, and second that because this is used to represent an association between speeds and densities – irrespective of the model development within which it is adopted – no inference can be drawn from it concerning causality. A key element of this modelling, which we investigate here, is the identification of causal relationships between the variables: this becomes crucial when variable speed limits are applied as traffic management measures.

In this paper, we investigate the relationship between speed and density (for which detector occupancy is used as a proxy). We consider a range of different speed–density functions, and establish clear preference for one that does not have a finite density that causes speed to become zero. However, we note that stationary traffic does attain a certain maximum density, which is know as *jam density* k_j. We deduce that under some conditions such as this, speed determines density. We use this observation to explore the effects of variable speed limits as a traffic management measure as is used on some busy sections of motorway roads in the UK. This leads us to estimate speed–density relationships differently according to the presence or absence of speed control, with consequences for estimates of the capacity of the road. Another consequence of this preference for a speed–density relationship without a jam density is that it is consistent with the presence of stable start waves in dense traffic, as have been observed (Zhang, 1999).

2. Modelling traffic flow

2.1. Introduction

In this section, we raise various issues that arise in modelling traffic flow. First, we consider the way in which quantities that can be measured practically relate to those used in the fundamental equation (1). We then consider a range of different functional forms that could be used to represent the relationship that is expressed formally in (2) between speed and density: for each of these candidate relationships, we investigate certain properties. We then address the statistical estimation of these models from observed data, and identify a suitable approach for this together with associated statistical criteria.

2.2. Properties of traffic models

Although the fundamental equation (1) is expressed in terms of the density k of traffic, direct measurement of this quantity can be difficult in practice as it requires instantaneous observation of a substantial length of road. A more practical approach that is in widespread current use (Hall, 2001) is based upon the proportion of time ω for which a fixed detector is occupied, which is known as the detector *occupancy*. Thus on a homogeneous section of road, the spatial proportion ω_s of the length that is occupied by vehicles will be

$$\omega_s = kL, \tag{3}$$

where L is the mean effective length of a vehicle. Provided that the speed of traffic is constant along the length of this section so that $\omega = \omega_s$, the occupancy ω of a detector on it can be used to estimate density as

$$k = \omega/L. \tag{4}$$

From this relationship, density is proportional to mean occupancy; it is therefore sometimes more convenient to express speed directly as a function of occupancy rather than as one of density, thus $v = f(\omega)$. This then leads to the relationship between flow and occupancy as

$$q = f(\omega)\omega/L. \tag{5}$$

We note, however, that the mean effective length of vehicles can differ between traffic in different lanes of a motorway, and between traffic control states (Heydecker and Addison, 2008), and this will influence estimates of flow and capacity that are made using (5).

This form of the fundamental relationship (5) between flow and occupancy can be used to determine several properties of the traffic model. In the present context we note that if this relationship is concave throughout its domain (i.e. with negative curvature given by the second derivative q'' of flow with respect to occupancy) then stable shock waves can only occur as transitions from low to high occupancy; start waves, where the spatial gradient of traffic density in the direction of travel is negative, will lead to rarefaction waves (Lighthill and Whitham, 1955) so that jammed traffic will disperse from the downstream boundary. However, if the fundamental relationship has a subdomain within which it has positive curvature (i.e.

where $q'' > 0$) then stable start waves can arise when traffic accelerates from a region with occupancy in this subdomain to a region of lower occupancy (see, for example, Morgan, 2002). Regions of this kind occur in speed–density functions that do not have a value of density that gives zero speed: models that have this positive curvature are consistent with the observed phenomenon of stable start waves, but without necessarily providing a mechanism for their occurrence.

As the traffic density, and hence occupancy, diminishes towards 0, the opportunities for interactions between vehicles in the traffic will diminish and so vehicles will tend to travel at some free-flow speed denoted here as v_0. On the other hand, when traffic is forced to stop as, say, at a red traffic signal or an obstruction on the road, it will do so at a certain *jam density* with an associated expected value of jam occupancy, here denoted as ω_j. Throughout the present paper, we suppose these two quantities of free-flow speed and jam occupancy to be well determined, as seems reasonable from the arguments presented above. However, we note here that the direction of causality differs between the arguments presented in these two cases. There seems to be no reasonable argument that travelling at free-flow speed will itself induce low density and occupancy of traffic. However, the counterpart question of whether or not traffic at jam density could move at non-zero speed is open at this stage and is one to which we will return.

2.3. Modelling approaches

A speed–occupancy model relationship $v = f(\omega)$ can be estimated from paired observations of speed and occupancy as

$$v_i = f(\omega_i) + \varepsilon_i, \tag{6}$$

where ε_i is an error term for the estimated speed. Statistical procedures can be applied to estimate parameters of a given model form for $f(\cdot)$, and these generally depend on the statistical distribution that is adopted for the error term ε. This formulation supports the hypothesis that traffic speed is determined by traffic density (as represented here by occupancy), whilst recognising the possibility that individual observations of speed deviate from the typical value that is associated with the prevailing occupancy.

An alternative view is that the traffic assumes a certain density and hence occupancy depending on the prevailing speed. Thus, for example, in slow traffic conditions, drivers could choose to follow at a headway (and hence density and occupancy) that is comfortable for them given the prevailing speed. This leads to a statistical model formulation as

$$\omega_i = g(v_i) + \eta_i \tag{7}$$

where η_i is an error term for the estimated occupancy. In this case, the form of the function $g(\cdot)$ is inverse to that of $f(\cdot)$.

In fitting either of models (6), (7), the criterion adopted in the present study is maximum likelihood according to the statistical distribution of the error term. The log-normal distribution adopted for this is appropriate for continuous quantities such as speed and occupancy that are of definite sign, and provides parameters to represent the mean and the variance of the data. The variate x has log-normal distribution with parameters μ and σ if $\log_e(x)$ has the normal distribution with mean μ and variance σ^2. Thus the probability density function of a lognormally distributed variate x is

$$p(x|\mu,\sigma) = \frac{1}{x\,\sigma\sqrt{2\pi}} \exp\left(\frac{-(\log_e(x) - \mu)^2}{2\sigma^2}\right). \tag{8}$$

This function also quantifies the likelihood of the parameters (μ, σ) given the observation x, and its logarithm (the *log-likelihood*) can be summed over observations to provide the joint log-likelihood, which provides a useful measure of performance of a statistical model.

The parameters of this distribution can be estimated to maximise the logarithm of their likelihood by solving for $\hat{\mu}$ the quadratic minimisation:

$$\underset{\mu}{\text{Minimise}} \sum_{i=1}^{n} (\log_e(x_i) - \mu)^2 \tag{9}$$

after which the parameter σ can be estimated as

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\log_e(x_i) - \hat{\mu})^2}. \tag{10}$$

We use the likelihood ratio test (Washington et al., 2003) to test for support for variations to a model. This test statistic for an extension to a model by introduction of v parameters from a set $\boldsymbol{\theta}$ to an extended one $\boldsymbol{\theta}^+$ is

$$\lambda = -2(\ell^*(\boldsymbol{\theta}, \mathbf{x}) - \ell^*(\boldsymbol{\theta}^+, \mathbf{x})). \tag{11}$$

Under the null hypothesis that the distribution is specified fully by the parameters $\boldsymbol{\theta}$, this test statistic asymptotically in n has the χ^2 distribution on v degrees of freedom.

2.4. Traffic models

Several forms have been proposed in the literature for the function $f(\cdot)$ between occupancy and speed, and hence for its inverse $g(\cdot)$. Drake et al. (1967) investigated a range of possibilities. These included Greenshields' (1935) linear relationship

$$f(\omega) = v_0(1 - \omega/\omega_j), \tag{12}$$

which confers the analytical advantage of giving convenient closed formulae for each of the five other bivariate relationships among speed, flow and occupancy. Drake et al. extended this to piecewise linear relationships, considered Underwood's and Greenberg's relationships described below, and also Drake's relationship $f(\omega) = v_0 \exp[-0.5(\omega/\omega_m)^2]$, which has subdomains of each of negative and positive curvature f''.

Underwood's (1961) model relationship between speed and density was intended primarily for use in free-flow conditions. This can be expressed in terms of occupancy as

$$f(\omega) = v_0 \exp(-\omega/\omega_m), \tag{13}$$

where ω_m is the occupancy that gives rise to the maximum flow of $q^* = v_0\omega_m/eL$, with corresponding speed $v = v_0/e$. This model has an explicit free-flow speed v_0, but there is no value of occupancy that will cause the speed to be 0 so that it is consistent with the possibility that traffic can travel at a non-zero speed even when it is at jam density. According to this model, the second derivative of flow with respect to occupancy is

$$q'' = \frac{v_0}{\omega_m L}\left(\frac{\omega}{\omega_m} - 2\right) \exp\left(\frac{-\omega}{\omega_m}\right) \tag{14}$$

so that the fundamental relationship between flow q and occupancy ω has negative curvature in the subdomain $[0, 2\omega_m)$ and positive curvature for greater values of occupancy. Thus according to this model, stable start waves can arise when traffic accelerates from a region in which the occupancy exceeds $2\omega_m$ into a region of lower occupancy. The range of possible speeds for this start wave is determined by the first derivative q' to be between $-v_0e^{-2}$ when $\omega = 2\omega_m$ to $-v_0(1 - \omega_m^{-1})\exp(-\omega_m^{-1})$ when $\omega = 1$.

Greenberg's (1959) model relationship between speed and occupancy was intended for use at high densities. This can be expressed in terms of occupancy as

$$f(\omega) = v_m \log_e\left(\frac{\omega_j}{\omega}\right). \tag{15}$$

In this model, v_m is the speed that gives rise to the maximum flow of $q^* = v_m\,\omega_j/eL$, which occurs when $\omega = \omega_j/e$. This model has an explicit jam occupancy ω_j, but there is no bound on the speed as occupancy approaches 0. Because of this, although this model might be suitable to represent congested conditions, it is not suitable for free-flow conditions. According to this model, the second derivative of flow with respect to occupancy is

$$q'' = \frac{-v_m}{\omega L} < 0 \tag{16}$$

so that the fundamental relationship is concave throughout its domain. According to this model, stable shock waves can only occur when traffic passes from a region into another of greater occupancy, thus undergoing an abrupt decrease in speed.

Edie (1961) developed a speed–density model that conforms to Underwood's (13) at low density and to Greenberg's (15) at high density. In the standard form of this model, the transition between these components is made smoothly (i.e. with continuity of q and of q') at the point of maximum flow. In order to achieve this, the parameters of the two elements are related as $v_m = v_0/e$, and $\omega_m = \omega_j/e$ so that the capacity is $q^* = v_0\omega_j/e^2$. In the present study, we view this model as providing an alternative to Underwood's (13) in which the high occupancy relationship follows Greenberg's form (15), and hence has a finite jam density and is concave throughout.

Other models have been developed that can be used to relate speed to occupancy; notable amongst these are the parametric family considered by Del Castillo and Bentez (1995) that varied between the forms

$$f(\omega) = v_0[1 - \exp(1 - \exp(\alpha(\omega_j/\omega - 1)))]$$

and

$$f(\omega) = v_0[1 - \exp(\alpha(1 - \omega_j/\omega))].$$

This family provides smooth approximations to the form

$$f(\omega) = v_0 \mathrm{Min}\left\{1, \frac{\omega_m}{\omega}\left(\frac{\omega_j - \omega}{\omega_j - \omega_m}\right)\right\},$$

which gives a two-part linear fundamental diagram with triangular profile. All of these models have a finite jam density ω_j/L that gives rise to zero speed and fundamental diagrams that are concave.

The two relationships between speed and occupancy that we consider in detail here are continuous but non-linear. This has the disadvantage of analytical inconvenience in that closed formulae are not generally available to calculate either speed

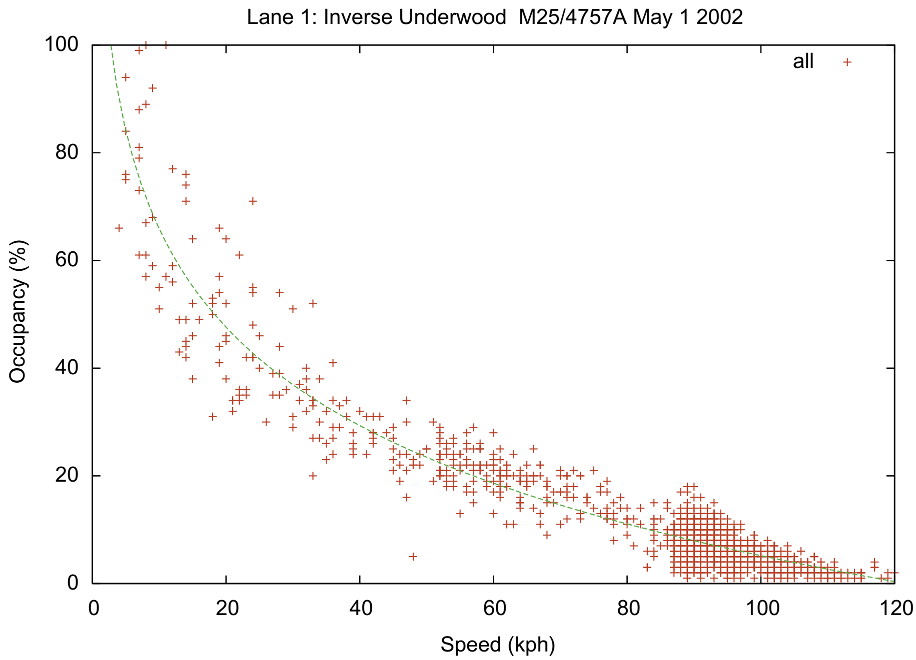

Fig. 1. Observations of occupancy and speed in lane 1.

or occupancy from specified values of flow. Inspection of the observed relationship (see Fig. 1) between speed and occupancy in the dataset that we use in the present analysis shows positive curvature but little evidence for discontinuity in speed as occupancy varies. The two speed–occupancy relationships that we investigate have positive curvature $f'' > 0$ throughout the range of occupancy, which makes them suitable to represent the data used in this study. This analysis can be applied readily to other relationships between speed and occupancy, but the focus of the present work is on the direction of causality between occupancy and speed rather than on the particular form of the relationship.

3. Traffic data

3.1. Introduction

The data used for the present investigation are derived from the MIDAS system (Highways Agency, 1994) and the associated Halogen data logging systems. They represent measurements made between junctions 10 and 16 on the 188 km M25 London Orbital Motorway in England. In this section, the road has four running lanes in each direction and carries up to 85,000 vehicles per day in each direction. The running lanes are numbered consecutively from 1 at the nearside (shoulder) to 4 at the offside (median). The rules of the road include that overtaking vehicles should pass to the offside of those being overtaken, and vehicles should return to the available lane that is furthest to the nearside as soon as safely possible (Department for Transport, 2007). Under most circumstances, heavy goods vehicles and those drawing trailers are prohibited from using the most offside lane of a motorway (i.e. lane 4 in the present case).

Under normal operating conditions, the national speed limit of 70 mph applies. However, mandatory speed limits can be imposed on this section of the M25 motorway at each of the reduced values of 60 mph and 50 mph in order to manage traffic flow under busy conditions, and the value of 40 mph can be used to slow traffic in advance of its approach to queues: these reduced speed limits are enforced by radar speed detectors supported by automatic cameras.

The MIDAS system has arrays of detectors spaced at approximately 500 m intervals along the carriageway, with a separate detector in each lane. This system records data that are aggregated over 1 minute intervals, each record containing:

flow the number of vehicles crossing the detector during that minute,
speed the mean speed in km/h of vehicles crossing the detector, and
occupancy the proportion of time that the detector is occupied.

The data used in the present analysis represent all of the available records for Wednesday 1 May 2002, at location code 4757 on the clockwise (A) carriageway. Analysis of corresponding data from the same site on Wednesday 15 May 2002, and from another site 4747 between the same junctions on Wednesday 1 May 2002 were found to yield similar results to those reported here and support identical findings.

Table 1
Number of observations available for analysis.

Speed limit (mph)	Lane			
	1	2	3	4
70	1015	893	733	541
60	18	18	18	18
50	97	97	97	97
40	239	239	240	241

The distribution of observations available for analysis is reported in Table 1. There were over 1000 min of the day during which the speed limit defaulted to the national one of 70 mph. By contrast, only 18 min of data were available for the 60 mph speed control condition. The number of available observations in each controls state varied substantially among lanes in the case of the 70 mph speed limit, with about half the number available in lane 4 compared with lane 1: this arises because in light traffic conditions such as occur late at night, the high availability of low numbered lanes leads to lane 4 being empty for extended periods of time.

Preliminary analysis of the data showed that the relationship between speed and occupancy was in all cases non-linear, and follows a trend that has positive curvature. A typical example is given in Fig. 1, which shows all observations of occupancy and speed made in lane 1 and the clearly non-linear trendal relationship between them. In view of this, only the non-linear models given by Underwood (13) (fitted line in Fig. 1) and Edie, as described in Section 2.4, were investigated further. Comparison of the results of fitting a single model to data from all four lanes taken together with corresponding ones from four separate models, one for each lane, showed that there is a substantial difference among the relationships for the different lanes. In light of this, all models were developed separately for traffic in each of the four lanes of the motorway.

4. Analysis and results

4.1. Introduction

We now consider the application of the analytic techniques described in Section 2 to the data introduced in Section 3. The approach to this that we pursue is that of generalised linear statistical modelling (see, for example, McCullagh and Nelder, 1989). This provides objective criteria for selection between model forms, and within each form for the specification and parameter values. The preferred traffic model can then be used together with the parameter values that are estimated for it to describe traffic properties such as capacity in various different circumstances.

According to the generalised linear modelling approach, the dependent variable that is being investigated is considered to be causally dependent upon a set of other variables (the *explanators*), with expected value calculated as a monotonic function (the inverse of the *link function*) of a linear combination of their values. Once the link function and statistical distribution of the error have been decided, values for the coefficients associated with each of the explanators can be calculated so as to maximise the likelihood of the model given the observations.

On the basis of the preliminary investigations, the values for the parameters of each model were taken to differ between lanes and so were fitted separately. We now address in turn the choice of model form for the speed–occupancy function, the effect of speed control status on the parameters of the functions, and the matter of causality in the modelling. Once each model had been refined so far as possible, the resulting forms were compared with each other according to the value of the maximised likelihood that was achieved so as to determine which can be justified on the basis of the data.

4.2. Choice of model form

The first step in modelling was to investigate the choice of model form as between Underwood's form (13) and Edie's form, which differs in that the high occupancy region is modelled using Greenberg's form (15). These two models have the same form at low occupancy, and have finite free-flow speeds v_0. The present investigation focuses on the high occupancy region as this is of greatest interest in management of busy roads and hence for speed control.

According to the generalised linear model framework, Underwood's model (13) can be expressed in the log-linear form as

$$v_i = \exp(\alpha + \beta\omega_i) + \varepsilon_i, \tag{17}$$

where $v_0 = \exp(\alpha)$ and $\omega_m = -\beta^{-1}$. When the log-normal distribution is postulated for the error term ε, this model can be fitted conveniently in the form

$$\log_e(v_i) = \alpha + \beta\omega_i + \xi_i, \tag{18}$$

where ξ has the normal distribution Nor$(0, \sigma^2)$: in this form, the model can be fitted by ordinary least squares regression to minimise σ^2.

In the subdomain within which it applies in Edie's model, Greenberg's model (15) can be estimated by fitting observed values of v to the logarithm of ω according to the regression

$$v_i = \alpha + \beta \log_e(\omega_i) + \varepsilon_i \quad (\omega_m < \omega_i) \tag{19}$$

In this case, the fitted parameters can be interpreted as $v_0 = -\beta$ and $\omega_j = \exp(-\alpha/\beta)$. In view of the form of this model, values for the parameters were estimated so as to maximise their likelihood using a purpose-specific optimiser applied to the joint likelihood calculated according to the log-normal distribution for the error term ε.

Values of the maximised log-likelihoods of the Underwood models for each lane, calculated according to the log-normal distribution (8) are given in Table 2. These are presented for 4 model variants in which the degree of specialisation of the model to the speed control increases progressively; in each case the mean is estimated using two parameters in Eq. (18) whilst the variance is estimated as a separate parameter. The first column of results is for a single model (with two parameters for the mean and one for the variance) fitted to all of the data. The second column of results is for two separate 3-parameter models, one fitted to each of the groups of data corresponding to no control (denoted as {70}) and control of some kind (denoted as {60, 50, 40}). The third column distinguishes further between the cases of control for speed management (denoted as {60, 50}) and queue protection control (denoted as {40}). Finally, the last column has a separate model for each speed control state.

The last row of Table 2 shows the values of the maximised log-likelihood for each of these models when traffic is modelled in all lanes together. The increase in log-likelihood when separate models are used for each lane is greater than 700 for each level of distinction in speeds giving likelihood ratio test values (11) in excess of 1400; by comparison, the critical value of the test statistics according to the likelihood ratio test at the 0.05 level of significance with 9 degrees of freedom is 16.9. We therefore find that a separate model is required for each of the 4 lanes of traffic.

The corresponding results for the Edie model are given in Table 3. This model differs from the Underwood model at high occupancy by adopting the logarithmic relationship of Eq. (15) in place of the exponential relationship of Eq. (13). This model therefore has an intrinsic jam occupancy that will cause traffic to stop. In this case, the improvement in log-likelihood associated with use of separate models for each lane is greater than 500 for each combination of speed control categories giving likelihood ratio test statistics (11) in excess of 1000, which is sufficient to justify the use of the additional 9 model parameters. Because of the smoothness conditions at the spline point $\omega = \omega_j/e$ between the two parts of the relationship, the Edie model is determined by the same number of parameters as the corresponding Underwood one.

Comparison of the likelihood values in Tables 2 and 3 for these two models shows that for each combination of lane and speed group, the Underwood model has the greater likelihood value. The difference in values is most pronounced when no allowance is made for variations speed limits, but exceeds 140 in all cases. From this, we find that the Underwood model provides a better description of the present M25 dataset than does the Edie model. This shows that in the high occupancy region $\omega > \omega_m$, Underwood's exponential relationship (13), which has no jam occupancy, provides a description of the data that is preferable to Greenberg's logarithmic relationship (15), which does have an explicit jam occupancy. Based on the empirical preference for Edie's model over Underwood's, we can draw a series of conclusions. First, even at the highest occupancies recorded, traffic can be expected to flow. In light of this, we conclude that jam occupancy arises as the occupancy

Table 2
Log-likelihood values for fitted Underwood models.

Lane	Model formulation (speed groups, mph)			
	{70, 60, 50, 40}	{70}, {60, 50, 40}	{70}, {60, 50}, {40}	{70}, {60}, {50}, {40}
1	633.74	775.09	838.23	841.92
2	507.29	617.23	725.68	731.62
3	305.69	529.16	628.68	638.57
4	132.07	292.13	388.79	399.28
Total	1578.79	2213.61	2581.38	2611.39
All	867.55	1286.04	1527.15	1537.29

Table 3
Log-likelihood values for fitted Edie models.

Lane	Model formulation (speed groups)			
	{70, 60, 50, 40}	{70}, {60, 50, 40}	{70}, {60, 50}, {40}	{70}, {60}, {50}, {40}
1	186.67	344.74	505.58	506.56
2	95.16	288.76	534.37	536.36
3	−174.39	119.49	391.21	394.51
4	−243.45	21.81	241.03	246.34
Total	−136.01	774.80	1672.19	1683.77
All	−722.37	131.97	918.63	926.42

Table 4
Estimates of capacity q^* for of each lane from Underwood's model (13).

Lane	v_0 (km/h)	ω_m	L (m)	$q^* = v_0\,\omega_m/e\,L$ (vehicles/h)
1	116.60	0.2606	10.93	1022.6
2	133.52	0.2234	7.58	1448.4
3	154.71	0.2114	6.32	1903.8
4	160.49	0.2910	6.21	2765.0

that is attained by traffic when the speed is forced to be zero by some blockage, and that occupancy this high will not in itself cause traffic to halt. This in turn supports the idea that in circumstances when the speed of traffic is controlled, traffic will adopt that speed and assume an occupancy according to that speed.

The estimates of fitted parameters for Underwood's model (13) for each lane, corresponding to all observations for that lane, are given in Table 4 together with the associated estimates of capacity q^* for that lane according to the formula $q^* = v_0\omega_m/eL$ for this model. The values for the mean effective vehicle length L in each lane given in this table were calculated in Heydecker and Addison (2008). This shows increasing estimates of free-flow speed and apart from lane 4 decreasing estimates of occupancy at capacity ω_m for increasing lane number. Together with the estimates of mean effective vehicle length, they lead to increasing estimates of capacity, as measured in vehicles per unit time, with increasing lane number.

4.3. Effect of speed control

We now consider the effect of speed control on the speed–occupancy relationship that is fitted to the data. In light of the preference for Underwood's speed–occupancy relationship that is established in Section 4.2, we focus on Table 2, which shows the results for that model. The log-likelihoods that result from fitting Underwood's model (13) that are given in Table 2 show that for each lane, the most substantial improvement in log-likelihood occurred when a separate model was fitted to data from conditions when speed was limited to one or other of the values 60, 50 or 40 mph as opposed to having the national value of 70 mph. For each lane, the improvement in log-likelihood associated with this refinement of the model ranged from a little over 30 in the case of distinction between 60 mph and 50 mph, to more than 600 for the separation of the national speed limit case from the speed restricted cases. Although not central to development of the argument, we note that the improvements in fit of Edie's model are similar in pattern and magnitude. Even the smallest of these improvements is sufficient to justify the associated model refinement that introduces 3 further parameters, for which the critical improvement in log-likelihood according to the likelihood ratio test (11) is 3.91 (=7.82/2).

4.4. Causality and the direction of influence

We now turn to the implications of the choice of Underwood's exponential model (13) in preference to one that has an explicit jam occupancy, and consider them in combination with the observation that the greatest improvement in model fit was achieved with the distinction between national and controlled speed operation of the motorway. The argument associated with the nature of jam occupancy and the occurrence of 0 speed in Underwood's model shows that in this case at least, speed is determined exogenously because it cannot result from high occupancy. Consideration of the improvements in log-likelihood associated with separation of the various different speed control categories shows that the greatest improvement in model fit occurred when the speed-controlled cases (60, 50 and 40 mph speed limits) were treated separately from the national speed limit of 70 mph. We therefore investigate for the speed-controlled cases estimation of the inverse Underwood model

$$\omega = g(v) = \omega_m \log_e\left(\frac{v_0}{v}\right), \tag{20}$$

which treats speed v as the explanatory variable. This is fitted by minimising with respect to the parameters (v_0, ω_m) the sum of squares of the difference between observed and estimated values in (20). The fitted model was evaluated according to the log-likelihood of the log-normal distribution (8).

The results of estimating this model for all data in lane 1 is shown in Fig. 2 on the scale of occupancy against logarithm of speed. This shows good agreement with the trend between occupancy and speed, though we note that the dispersion around the model estimate is greater at lower speeds than at higher ones. Comparisons between estimation of a single model for all data against separate estimation for the cases of national speed limit (70 mph) and all controlled speeds (60, 50 and 40 mph) taken together are given in Table 5. The likelihood ratio test statistics, which are all substantial, show that the distinction between national speed limit and the controlled states is fully justified. The smallest of these values, 37.18 for lane 1, is greatly in excess of the critical value at the 5% level of significance on 3 degrees of freedom of 7.82, whilst each of the values for the other lanes is greater than 400. The standard errors of estimation s_ω in the range 0.019–0.07 show that this model achieves a reasonably good fit to the observed occupancy ω, as can be seen in Fig. 2.

Both of the usual and the inverse Underwood models between speed and occupancy, with results presented in Tables 2 and 5 respectively, support use of separate models for the cases of national speed limit (70 mph) and controlled speeds.

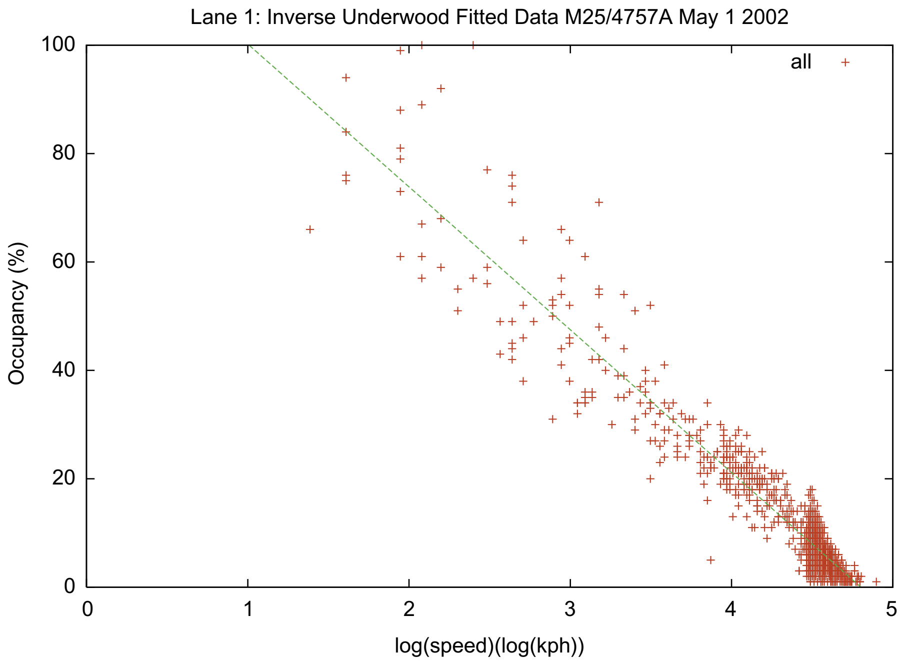

Fig. 2. Inverse Underwood model fitted to all data for lane 1.

Table 5
Comparison of inverse Underwood models for all flows.

Lane	Speed group	v_0 (km/h)	ω_m	Standard error s_ω	Log-likelihood	Likelihood ratio λ (2 df)
1	All	121.71	0.2636	0.0422	−871.88	37.18
	70 mph	120.00	0.2707	0.0265	−879.27	
	Control	130.71	0.2523	0.0690	25.98	
2	All	139.09	0.2405	0.0403	−954.79	603.16
	70 mph	128.36	0.3360	0.0226	−768.72	
	Control	163.03	0.2194	0.0633	115.51	
3	All	163.47	0.2048	0.0391	−852.83	737.08
	70 mph	138.48	0.4506	0.0191	−616.29	
	Control	157.07	0.2082	0.0576	132	
4	All	172.58	0.1910	0.0420	−590.48	492.82
	70 mph	143.89	0.4540	0.0196	−329.06	
	Control	167.76	0.1932	0.0584	−15.01	

Argument concerning traffic behaviour at low density and hence low detector occupancy supports use of the usual Underwood model (13) to estimate speed from occupancy when the national speed limit applies, whereas corresponding argument supports use of the inverse Underwood model (20) for speed controlled conditions.

We therefore investigated further use of the inverse Underwood model (20) by considering data only from controlled speeds. The results of this are given in Table 6, which shows the likelihood ratios for the test of separate models for each speed within the controlled range against a single model for all controlled speeds. The likelihood ratio test statistics all exceed 20, compared to the critical value at the 5% level of significance on 6 degrees of freedom, which is 12.6 : we therefore find that further distinction between models according to the controlled speed is justified. The results of estimating the inverse Underwood model for data observed in lane 4 under 40 mph speed limit control is shown in Fig. 3 on the scale of occupancy against logarithm of speed. This shows good agreement with the trend between occupancy and speed, though we note that the dispersion around the model estimate, which is accommodated where this occurs between speed control cases, remains greater at lower speeds within the 40 mph control regime than at higher ones.

The fitted values of the parameters for the inverse Underwood model for the speed controlled conditions are shown in Table 7, together with the maximum modelled value of flow at an occupancy that was observed for each one. For each lane, the maximum flow under 60 mph speed control occurred at the greatest observed occupancy; under other speed controls, it occurred at an occupancy within the range observed. In lanes 1 and 2, the greatest flow occurred under speed controls of 50 mph, whilst in lanes 3 and 4 it occurred under speed control of 60 mph. The estimates of maximum flow for lanes 1–3

Table 6
Comparison of inverse Underwood models for speed–controlled flows.

Lane	Speed (mph)	v_0 (km/h)	ω_m	Log-likelihood	Likelihood ratio λ (4 df)
1	Controlled	130.71	0.2523	25.98	23.8
	60	111.44	0.3950	−7.19	
	50	127.60	0.2768	16.2	
	40	126.16	0.2566	28.87	
2	Controlled	163.03	0.2194	115.51	49.46
	60	110.87	0.6282	0.35	
	50	177.56	0.2169	37.95	
	40	145.99	0.2307	101.94	
3	Controlled	157.07	0.2082	132	30.98
	60	123.83	0.4771	−4.41	
	50	179.50	0.1923	57.49	
	40	138.83	0.2212	94.41	
4	Controlled	167.76	0.1932	−15.01	182.96
	60	121.34	0.5599	−8.3	
	50	153.02	0.2339	14.57	
	40	151.21	0.2022	70.2	

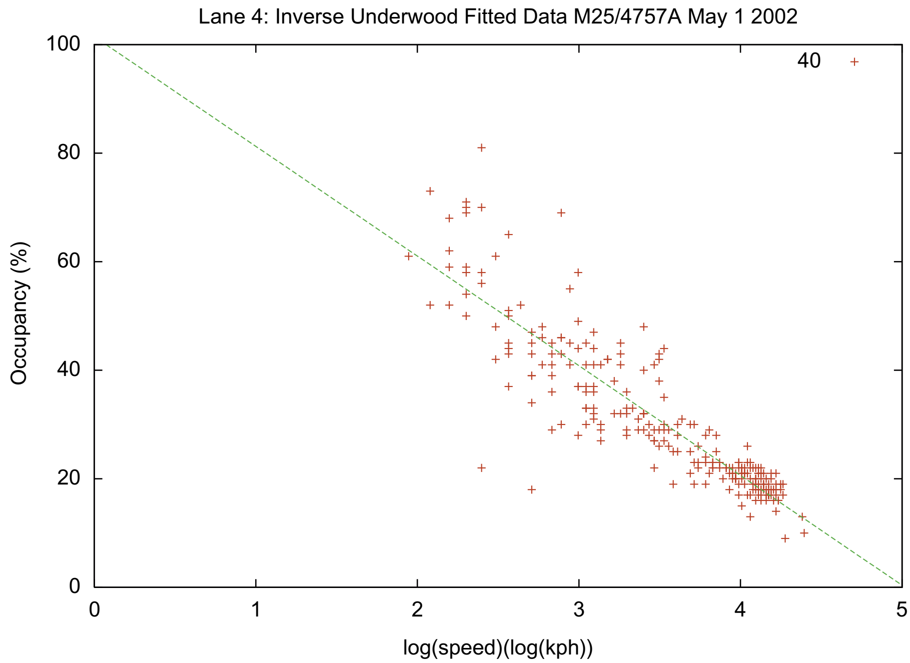

Fig. 3. Inverse Underwood model fitted to data for lane 4 under 40 mph speed control.

exceed those from the usual Underwood model applied to all speed limits that are shown in Table 4, though it is slightly less for lane 4. The estimated values of maximum flow are lower, and more realistic, in each lane where the different speed control levels are modelled separately. This indicates that different relationships apply under each level of speed control and hence emphasises the importance of modelling traffic separate under each of them.

5. Discussion and conclusions

The work reported in this paper considers the relationship between speed and occupancy on motorway roads, and the way in which this is influenced by speed controls. Understanding and modelling this is important in developing and refining traffic management measures. Analysis of observations of speed and occupancy in each of the 4 lanes of the M25 motorway shows that the relationship between them differs significantly from lane to lanes, so that each lane should be modelled separately.

Table 7
Estimates of capacity q^* for of each lane from the inverse Underwood model in controlled speeds.

Lane	Speed (mph)	v_0 (km/h)	ω_m	L (m)	Maximum flow at observed occupancy (vehicles/h)	Speed at maximum flow (km/h)
1	Controlled	111.44	0.3950	8.24	1965.12	48.09
	60	127.90	0.2468	10.61	821.53	84.00
	50	115.18	0.3274	7.91	1753.82	46.94
	40	151.99	0.2168	9.02	1343.92	46.41
2	Controlled	110.87	0.6282	8.75	2927.46	59.98
	60	127.05	0.3786	8.83	1490.02	84.00
	50	161.04	0.2380	8.68	1624.41	65.32
	40	163.52	0.2131	9.40	1363.74	53.71
3	Controlled	123.83	0.4771	6.82	3189.08	57.78
	60	133.73	0.3735	6.57	2363.85	79.00
	50	164.92	0.2214	6.57	2044.52	66.03
	40	176.05	0.1849	7.20	1663.21	51.07
4	Controlled	121.34	0.5599	6.48	3858.43	61.72
	60	135.73	0.3856	6.19	2663.62	79.00
	50	162.25	0.2242	6.30	2124.15	56.29
	40	159.61	0.1991	6.80	1719.21	55.63

Investigation of the form of this relationship shows support for use of Underwood's exponential relationship, which does not have a jam density. This suggests that even at the highest densities recorded, traffic could continue flow. In light of this, we conclude that jam density arises as the density attained by traffic when the speed is forced to be zero by some blockage, and that density this high will not in itself cause traffic to halt. This supports the idea that in some circumstances when the speed is controlled, traffic will assume a density according to that speed.

A consequence of the adoption of Underwood's relationship between speed and occupancy is that the fundamental relationship between flow and occupancy has positive curvature at occupancy values that are more than twice that at which flow is maximised. This positive curvature leads to the possibility of transitions from low speed and high density to higher speed and lower densities that are stable rather than dispersive. Stable start waves of this kind are observed in practice: together with ordinary shock waves (transitions from low density to high) that have similar speed they form persistent regions of high density that travel upstream along the road in the form of stop–start waves. This then provides a basis for modelling stop–go traffic conditions that are commonly found on congested motorway roads.

Statistical estimation of Underwood's model relationship between speed and occupancy using observations separately in each lane showed that different relationships obtain according to the speed control status. Thus we find that driver behaviour, and hence traffic performance, varies according to this control status.

Drawing on these findings, we conclude that use of a relationship in which speed is the explanatory variable upon which occupancy and flow are taken to depend is justified when traffic is subject to speed control. This prompts the development of a modelling approach in which the form and directions of influence as well as the parameters of the relationships differ according to the speed control status of the road. The lines of argument that are established in the present paper and hence the resulting modelling approach are independent of the road location and functional forms, and so will have general applicability.

Estimation of the capacity in vehicles per unit time of each of the lanes of the road using occupancy data depends upon the mean effective length of vehicles at the detectors, which is known to differ between combinations of lane and speed control state. Use of appropriate values together with the estimated models of speed, flow and occupancy leads to estimates of capacity of each lane in different control states. According to this, the estimated capacity of each of lanes 1–3 increases when speed control is introduced, with the greatest values that were supported by the data occurring when the control speed is 50–60 mph depending on the lane. This increase in capacity can lead to reduced congestion and hence reduced travel times. The present modelling approach therefore provides a description of how imposing a reduced speed limit can induce greater road occupancy and hence increase flow on motorway roads.

The results of this research relate to the flow of traffic on busy motorway roads in three distinct ways. These are the inherent nature of traffic flow in these conditions, the way in which models for this can be formulated and used, and consequent findings for management of this traffic. Whilst these findings are based upon analysis of traffic on the controlled section of the M25 motorway in England, their nature informs issues of traffic modelling in support of development of traffic management in general.

Acknowledgments

The work reported here was undertaken as part of a project funded by the UK Engineering and Physical Science Research Council under research Grant EP/D 506840, and was supported by the UK Highways Agency. The authors are grateful to anonymous referees for their valuable comments on an earlier draft of this paper.

References

Daganzo, C.F., 1997. Fundamentals of Transportation and Traffic Operations. Pergamon, Oxford.

Department for Transport, 2007. The Highway Code. The Stationary Office, London.

Del Castillo, J.M., Bentez, F.G., 1995. On the functional form of the speed–density relationship – I: general theory. Transportation Research Part B 29B (5), 373–389.

Drake, J., Schofer, J., May Jr., A.D., 1967. A statistical analysis of speed-density hypotheses. In: Edie, L.C., Herman, R., Rothery, R. (Eds.), Vehicular Traffic Science. Elsevier, New York, pp. 112–117.

Edie, L.C., 1961. Car following and steady state theory for non-congested traffic. Operations Research 9, 66–76.

Greenberg, H., 1959. An analysis of traffic flow. Operations Research 7, 79–85.

Greenshields, B.D., 1935. A study of traffic capacity. Highway Research Board Proceedings 14, 448–477.

Hall, F.L., 2001. Traffic stream characteristics. In: Gartner, N.H., Messer, C.J., Rathi, A. (Eds.), Traffic Flow Theory: A State-of-the-art Report. Transportation Research Board, Washington, DC (Chapter 2).

Hall, F.L., Hurdle, V.F., Banks, J.H., 1992. Synthesis of recent work on the nature of speed–flow and flow–occupancy (or density) relationships on freeways. Transportation Research Record 1365, 12–18 (Washington DC: TRB).

Heydecker, B.G., Addison, J.D., 2008. Measuring traffic flow using real-time data. In: Proceedings of the TRB Traffic Flow Theory and Characteristics Committee (AHB45) Summer Meeting 2008: The Fundamental Diagram: 75 Years (Greenshields75 Symposium). 8–10 July 2008, Woods Hole, Massachusetts.

Highways Agency, 1994. Motorway Incident Detection and Automatic Signalling (MIDAS) Design Standard TD 45/94. In: Design Manual for Roads and Bridges, 9(1)2. Department for Transport, London.

Leutzbach, W., 1988. Introduction to the Theory of Traffic Flow. Springer, London.

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves: II. A theory of traffic flow on long crowded roads. Proceedings of the Royal Society 229A, 317–345.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models, second ed. Chapman and Hall, London.

Morgan, J.V., 2002. Numerical methods for macroscopic traffic models, Thesis for PhD, University of Reading.

Underwood, R.T., 1961. Speed, volume and density relationships, Quality and Theory of Traffic Flow, Yale Bureau of Highway Traffic, pp. 141–188.

Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. In: Proceedings of the Institution of Civil Engineers Part II, pp. 325–378.

Washington, P.S., Karlaftis, M.G., Mannering, F.L., 2003. Statistical and Econometric Methods for Transportation Data Analysis. Chapman and Hall/CRC, London.

Zhang, H.M., 1999. A mathematical theory of traffic hysteresis. Transportation Research Part B 33B (1), 1–23.

Zhang, H.M., Kühne, R., Michalopoulos, P., 2001. Continuum flow models. In: Gartner, N.H., Messer, C.J., Rathi, A. (Eds.), Traffic Flow Theory: A State-of-the-art Report. Transportation Research Board, Washington, DC (Chapter 5).