



City Research Online

City St George's, University of London

Citation: Hu, J., Kaparias, I. & Bell, M. G. H. (2009). Spatial econometrics models for congestion prediction with in-vehicle route guidance. IET Intelligent Transport Systems, 3(2), pp. 159-167. doi: 10.1049/iet-its:20070062

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/6526/>

Link to published version: <https://doi.org/10.1049/iet-its:20070062>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Hu, J, Kaparias, I & Bell, MGH (2009). Spatial econometrics models for congestion prediction with in-vehicle route guidance. IET Intelligent Transport Systems, 3(2), pp. 159-167. doi: 10.1049/iet-its:20070062 <<http://dx.doi.org/10.1049/iet-its:20070062>>



**CITY UNIVERSITY
LONDON**

[City Research Online](#)

Original citation: Hu, J, Kaparias, I & Bell, MGH (2009). Spatial econometrics models for congestion prediction with in-vehicle route guidance. IET Intelligent Transport Systems, 3(2), pp. 159-167. doi: 10.1049/iet-its:20070062 <<http://dx.doi.org/10.1049/iet-its:20070062>>

Permanent City Research Online URL: <http://openaccess.city.ac.uk/6526/>

Copyright & reuse

City University London has developed City Research Online so that its users may access the research outputs of City University London's staff. Copyright © and Moral Rights for this paper are retained by the individual author(s) and/ or other copyright holders. All material in City Research Online is checked for eligibility for copyright before being made available in the live archive. URLs from City Research Online may be freely distributed and linked to from other web pages.

Versions of research

The version in City Research Online may differ from the final published version. Users are advised to check the Permanent City Research Online URL above for the status of the paper.

Enquiries

If you have any enquiries about any aspect of City Research Online, or if you wish to make contact with the author(s) of this paper, please email the team at publications@city.ac.uk.

This paper is a postprint of a paper submitted to and accepted for publication in IET Intelligent Transport Systems and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at IET Digital Library

SPATIAL ECONOMETRICS MODELS FOR CONGESTION PREDICTION WITH IN-VEHICLE ROUTE GUIDANCE

Jun Hu, Ioannis Kaparias and Michael G.H. Bell

Centre for Transport Studies, Imperial College London, SW7 2AZ, UK

Telephone: +44(0)20-7594-5952 Fax: +44(0)20-7594-6102 Email: jh205@imperial.ac.uk

Keywords: congestion propagation, congestion prediction, reliability, in-vehicle route guidance, panel data model.

ABSTRACT

This study explores the congestion dependence relationship among links using microsimulation, based on data from a real road network. The work is motivated by recent innovations to improve the reliability of Dynamic Route Guidance (DRG) systems. The reliability of DRG systems can be significantly enhanced by adding a function to predict the congestion in the road network. This paper also talks about the application of spatial econometrics modelling to congestion prediction, by using historical Traffic Message Channel (TMC) data stored in the vehicle navigation unit. The nature of TMC data is in the form of a time series of geo-referenced congestion warning messages which is generally collected from various traffic sources. The prediction of future congestion could be based on the previous year of TMC data. Synthetic TMC data generated by microscopic traffic simulation for the network of Coventry are used in this study. The feasibility of using spatial

econometrics modelling techniques to predict congestion is explored. Results are presented at the end.

1. INTRODUCTION

As Dynamic Route Guidance systems are a rapidly growing market, more sophisticated systems are continuously being developed, including more advanced functions. In recent research work [1], [2], [3] and [4], a new Dynamic Route Guidance algorithm was developed, which, besides travel time, also takes travel time reliability into account. By defining reliability as the probability that a link will be uncongested, the algorithm uses a penalty procedure to, where possible, exclude from the route search links classified as unreliable, under the assumption that a link's state is a binary variable, reliable or unreliable. The level of link congestion is chosen to measure reliability.

Congestion data are required in many new dynamic route guidance systems. At present, car navigation systems from Garmin, Tomtom and other suppliers can be supplied with congestion warning messages via the Traffic Message Channel (TMC), provided that the driver has subscribed for the service. These warnings can then be displayed on the in-vehicle map in the form of triangles (see Fig. 1) and can cause the route recommended to the driver to be modified. However, this data relates to congestion that has already happened and has been detected by traffic monitoring systems. The quality of guidance could be significantly improved if accurate short-term predictions of congestion were also available. Then the route could be sought which would take encountered, rather than current, network conditions into account. [The congestion prediction in the context of this paper means the prediction of congestion events on certain links of a route at \$n\$ unit of time ahead.](#)

As current vehicle navigation systems are limited to detect incidences of congestion, there is no guarantee that the recommended route will turn out to be congestion-free or indeed avoid sites of undetected congestion. Since congestion often propagates in predictable ways and is recurrent, reasonably accurate short-term forecasts of congestion should be possible.



Fig. 1: In-vehicle display of a BMW

The aim of this paper is first to explore the congestion dependence relationships among links and then investigate the potential for using spatial econometrics models to predict the congestion propagation and dissemination on the road network for eventual use in car navigation systems. Link reliability, failure dependence relationships and congestion propagation, which form the basis of the study presented here, are first reviewed. By using a case study of a real road network for the City of Coventry in the UK, the impact of congestion on neighbouring links as well as the propagation and dissipation of congestion are studied. A congestion forecasting framework based on link speed data is established. Results are presented at the end.

2. BACKGROUND

2.1. Failure dependence relationship

The concept of link failure is derived from the definition of link reliability and represents the state of being unacceptably congested, which in the case of a road network means that the travel time on the link is longer than a preset threshold. Links in a real road network, however, do not usually fail independently. When a link fails or is degraded, the adjacent links or the links in the same area are also likely to be affected by the same degradation, although with a lag. In order to propose a more reliable path for the travellers, link failure dependence relationships should be taken into account in a robust route guidance algorithm.

Only a few studies on link failure dependence have been carried out so far. [1] introduced the idea of link failure dependence by defining three possible types of failure dependence between two links. Considering two links i and j , if the performance of link i deteriorates, then link j is positively failure dependent on i if its performance also deteriorates. Alternatively, j is negatively failure dependent on i if a deterioration of the performance of i results in an improvement of the performance of j , and if a deterioration of the performance of i leaves the performance of j unaffected, then i and j are failure independent. In order to quantify the degree of failure dependence, [1] introduce a failure dependence coefficient μ_{ij} , where $-1 < \mu_{ij} < 1$.

The calculation of μ_{ij} is carried out using a so-called topological failure dependence approach, according to which it is assumed that links are failure dependent because of their location. For instance, if link j is located directly upstream to the incident link i or is an alternative to it, it is very likely that its performance will have a high positive correlation with the performance of link i and j will therefore be positively failure dependent on i . Conversely,

if j is located directly downstream of the incident link i , it is very likely that it will be negatively failure dependent on i , because it will benefit from the bottleneck effect of the incident. Thus, depending on the location of the link in question, the failure dependence coefficient between them varies accordingly.

2.2. Congestion propagation

Through the study of link failure dependence relationship, we further find one interesting aspect of congestion, which is its propagation. It tends to propagate in a direction opposite to the flow of traffic. Its propagation has been studied in previous research [5]. The results shed light on the phenomenon of progressive link failure which characterises the propagation of congestion, where failure is characterised by a collapse in link speed. Fig. 2 illustrate the situation at the time of the initial link failure (circled on left) and 30 minutes later on the right (red indicates link failure). So if we study the pattern of congestion generation, propagation and dissipation, in other words the dependent relationship among these link speeds, we could predict when congestion is likely to arise on a certain link. However, this dependent relationship is complicated, since links speeds are usually spatially and temporally correlated with each other.

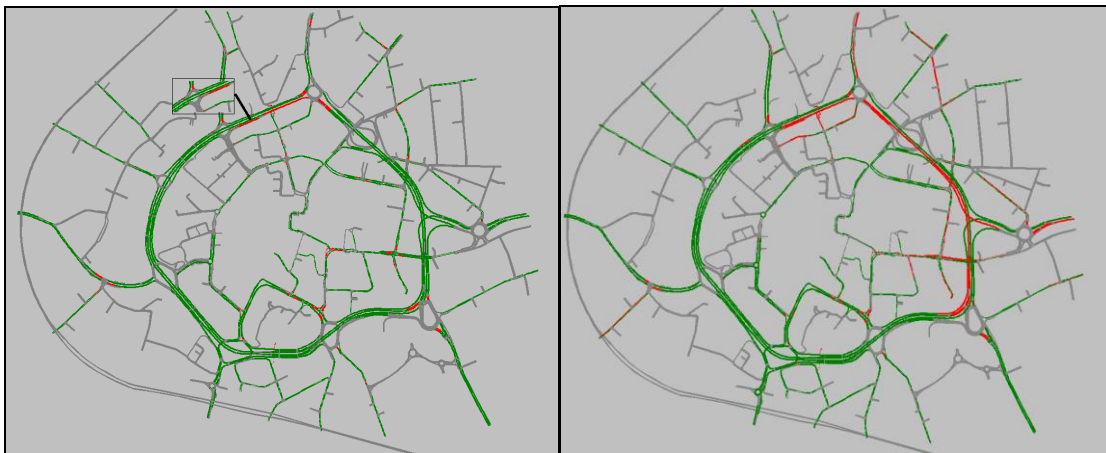


Fig. 2(a): Initial link failure

Fig. 2(b): Link failure after 30 minutes

Inter-dependencies in link speeds depend on network topology. The incidence of congestion on a certain link depends on its location and distance from the initial congested link. The blockage of the initially congested link usually results in the congestion of the upstream links. The distance to that initial congested link and the extent of shared flow determine the degree of impact. On the other hand, downstream links usually appear to be less congested when the traffic is blocked upstream. In previous link reliability studies, [1] categorise this geometrical relationship into three types: positively dependent, negatively dependent and independent. This can be quantified by the so-called *failure dependency coefficient*, which has mentioned above. This is elsewhere used to find reliable routes for dynamic route guidance.

Link speeds on the network are also autocorrelated. The speed on a link at a particular period t is determined by the state of the link at previous periods, $t-1$, $t-2$, etc. For example, if we know a link is congested in the previous period, there is a high probability that the link will be congested in next period.

2.3. Congestion incidence prediction models

As can be found in the literature, a large number of traffic prediction models exist. Broadly speaking, there are two types of prediction models in terms of their methodological approaches: simulation-based models and data-driven models.

Simulation-based approaches make use of traffic flow simulation models, derived from traffic flow theory, to predict the traffic conditions on the route of interest. Many simulation software tools have been developed, the most notable of which are DynaMIT [6] and SAVaNT [7]. DynaMIT was developed to predict the inductive travel time on a route of

interest, its underlying model being based on driver behaviour assumptions such as car following, gap acceptance and risk avoidance. SAVaNT, on the other hand, is used to predict link travel times in decentralized route guidance architectures, its underlying model being based on driver behaviour.

Data-driven approaches refer to a broader category of congestion prediction models. These approaches include general statistical models [8], neural network models [9] and machine learning models [10]. General statistical methods have been used by many researchers [8] to develop journey time forecasting models, on a link-by-link basis. These methods have been proven to be relatively simple and robust, and as they usually do not require many input parameters, they are efficient in their operation. In the model proposed by Hounsell and Ishtiaq [8], the required parameters are only incident severity and location in the network, from which the forecasted journey time on the incident link and on affected links can be generated. Other statistical models also include time series models [11], ARIMA (AutoRegressive Integrated Moving Average) models for predicting traffic flow [12], and Bayesian travel time prediction models [13].

In this paper, we adopt the spatial econometrics approach. Spatial econometrics is a subfield of econometrics that deals with the treatment of spatial interaction (spatial autocorrelation) and spatial structure (spatial heterogeneity) in regression models for cross-sectional and panel data [14]. Spatial dependence can be defined as the existence of a functional relationship between what happens at one point in space and what happens elsewhere.

There are mainly two causes of spatial dependence: measurement errors and spatial interaction phenomena. Regarding the former, spatial dependence is a by-product of

measurement errors from observations in contiguous spatial units. In most empirical studies, data is only collected at an aggregate scale and there may be little correspondence between the spatial scopes of the phenomenon and the spatial units of observations. As a consequence, measurement errors are likely. Regarding the latter cause, spatial dependence occurs due to the existence of a variety of spatial interaction phenomena. As a result, what is observed at one point is determined by what happens elsewhere in the system. This can formally be expressed in the following equation:

$$y_i = f(y_1, y_2, y_3 \dots y_N)$$

i ∈ *S*, where *S* as the set containing all spatial units of observations

Every observation of a variable *y* at *i* is related formally through function *f* to the magnitudes of the variable in other spatial units of the system. In this study, we are more concerned about the second case of spatial dependence, as this type of dependence is more likely to appear in a road network.

Spatial heterogeneity can be seen as a lack of structural stability of the various phenomena over space. In road networks, for example, according to the road hierarchy, accidents occurring on different roads will have different effects on the network, such that major roads tend to have a higher impact than minor roads. Therefore, different functional forms should be considered in the model. The easiest way to illustrate spatial heterogeneity is to express it by the following expression:

$$y_{it} = f_{it}(x_{it}, \beta_{it}, \varepsilon_{it})$$

where *i* = spatial unit of observation,

t = time period

The f_{it} is a time-space specific functional relationship, which explains the dependent variable y_{it} in terms of a vector of independent variables x_{it} , a vector of parameters β_{it} , and an error term ε_{it} .

Spatial econometrics models saw widespread application after Anselin's description of the method [14]. The broad application of spatial analysis in the transportation and urban planning fields is reviewed in [15]. Mainly, one branch of study is concerned with the explicit accounting for the spatial structure assumptions inherent in aggregation, such as defining travel zones. Another is for correlations in statistical models of spatially interacting processes like network flows and land uses. Recent applications in the field of transportation have been made in long-term land use models by Zhou et al [16] and in short-term location choice by Zhao & Bhat, Guo & Bhat [17], [18]. The work of Aerts et al [19] is an example of spatial autocorrelation analysis of incidents on networks. A spatial correlation model is used by Bernard et al [20] to establish the correlation of link travel speed.

3. CASE STUDY OF COVENTRY CITY MODEL

3.1. Description of the model

In order to generate useful database for this study, we used the Coventry city traffic simulation model. The model represents the existing traffic situation during a weekday afternoon peak hour (17:00 – 18:00) in 2004 [21]. A sudden road closure caused by an incident or vehicle breakdown is artificially created on the city ring road, in order to generate the congestion data for the analysis. The incidence link is denoted as the 'failed link' on the map. Then the following 8 links are selected for analysis. In Fig. 3 colour is used to identify

the links analysed. The traffic flows in an anticlockwise direction on the selected links, so these links are all upstream links of the incidence link.

The road closed for duration of 30 minutes and reopened at 17:30. The whole simulation period is from 16:30 to 18:30, which is 2 hours. Therefore, we could be able to look at the three stages (incident-free, congestion propagation and congestion dissipation).

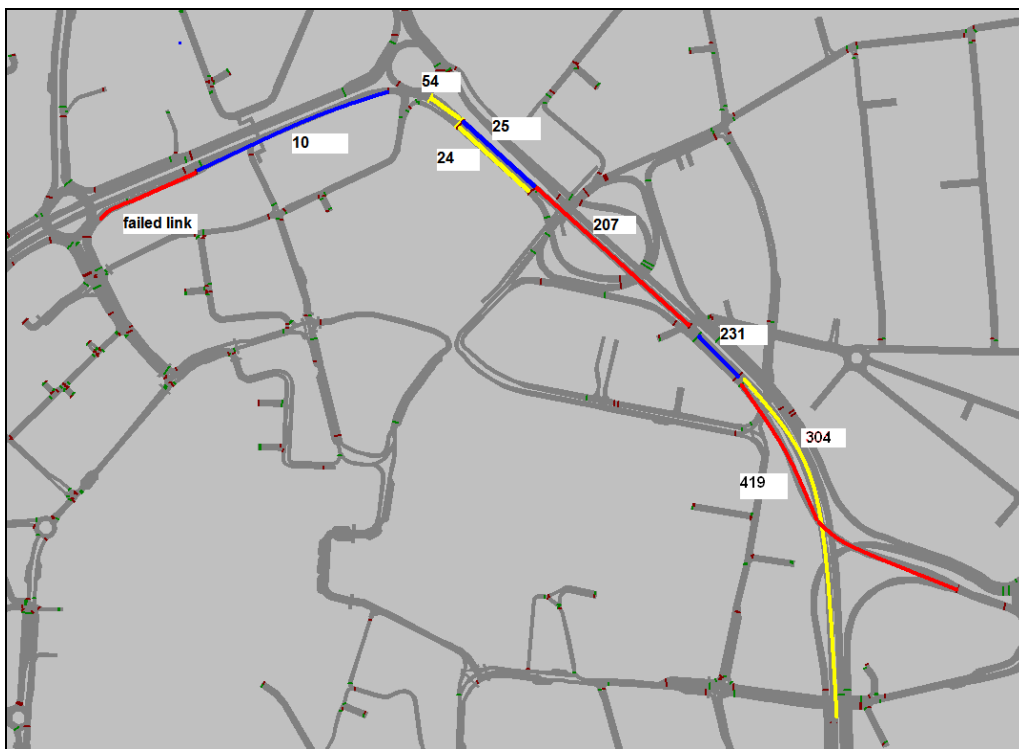


Figure 3: location of links

3.2. Database

Analogue speed data is used in this study to investigate the propagation of congestion and therefore predict the congestion. The state of the link can be very well characterised by the speed on that link. The lower the speed, the more severe the state of congestion is on that link.

Speeds on the 8 selected links are collected. The average speed of vehicles travelled on the links is recorded for 3 minutes interval over 2 hours, giving 40 observations for each link. That leads to a cross-sectional ($N=8$) time series ($T=40$) panel data set with a total of 320 ($N \times T$) observations.

However, there are 3 observations of speed missing on link 25 during the highly congested period, as the simulation software could not estimate the travel time on that link since the speed was zero or nearly so. Other data missing is due to the nature of the data. For example, our independent variable $V_{i,t-1}$, speed lag on the link i at time $t-1$. The first value of this variable will naturally be missing. An example of the structure of the data is shown in Table 1.

Table 1: Example of panel dataset

Link id	t	$V_{i,t}$	S (speed on downstream link at time t)	$V_{i,t-1}$	$V_{i,t-2}$	$V_{j,t-1}$
10	1	37.0	30.3			
10	2	37.3	30.3	37.0		30.3
10	3	35.9	28.9	37.3	37.0	30.3
10	4	35.0	28.5	35.9	37.3	28.9
10	5	36.5	30.3	35.0	35.9	28.5
10	6	34.7	31.3	36.5	35.0	30.3
10	7	35.9	29.4	34.7	36.5	31.3
10	8	32.7	25.1	35.9	34.7	29.4
10	9	36.5	24.4	32.7	35.9	25.1
10	10	35.9	26.8	36.5	32.7	24.4

3.3. Method of analysis

Here we use the simplified network diagram to illustrate the way that the analysis was carried out.

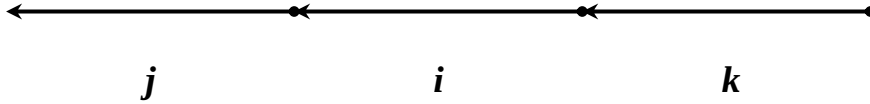


Figure 4: network diagram

The arrow in the diagram indicates the direction of traffic flow. Suppose link i is the link in question, we wish to predict the average speed on link i at time t . This is determined by various factors, most importantly: the speed on downstream link j in the previous time interval ($V_{j,t-1}$) as well as the speed on link i in the previous time intervals $t-1$ and $t-2$ ($V_{i,t-1}$, $V_{i,t-2}$). We used a panel data model to establish this relationship. This is explained in the next chapter.

4. RESULTS

4.1. Introduction

Panel data are observations on a cross-section of units, links in our case, on the network that are observed over a period. The fundamental advantage of panel data models over cross-section models is that they take into account autocorrelation as well as correlation between links [22]. Two types of panel data model are frequently encountered in the literature. One is the fixed effects model while the other is the random effects model. The main difference

between these two models is whether or not the unobserved link effects are correlated with the regressors in the model. The basic forms of two models are as follows:

The fixed effects model is:

$$y_{it} = \mathbf{x}_{it}^T \boldsymbol{\beta} + \alpha_i + \varepsilon_{it}$$

where α_i is a link-specific constant term and does not vary over time. The corresponding random effects model is:

$$y_{it} = \mathbf{x}_{it}^T \boldsymbol{\beta} + u_i + \varepsilon_{it}$$

where u_i is a link-specific random element. Vector $\mathbf{x}$ represents the regressors and vector $\boldsymbol{\beta}$ represents the parameters.

The regressand and regressors of the model are listed below:

$V_{i,t}$	= speed on link i at time t
$V_{j,t-1}$	= speed on the most immediate downstream link at $t-1$
$V_{i,t-1}$	= speed on the link i at time $t-1$
$V_{i,t-2}$	= speed on the link i at time $t-2$

The model captures both spatial and temporal autocorrelation. However, because the panel data in our case has a small number of cross-sectional unit N and large T , if we use the

dynamic panel data model, its estimator, called the Arellano and Bond estimator, will not be efficient for our problem [23]. It is thought that the general panel data model is more appropriate to fit a dataset with small N and large T .

In our case, both the random and fixed effects models will be fitted and then we use the Hausman test for random effects model to decide which model is more appropriate.

Panel data model

Both the fixed and random effects models are fitted to the data. The results are compared and Hausman test is carried out for the random effects model. Table 2 shows that the chi-square test statistic is 24.42. From Chi-square table, we find $Prob(\chi^2 > 24.42) = 0.0000$ with three degrees of freedom, showing that the difference is statistically significant. Therefore the null hypothesis that the individual effects are uncorrelated with the other regressors in the model should be rejected. So we would conclude that of the two alternatives we have considered, the fixed effects model is the better choice according to the Hausman test result.

Table 2: Hausman test for random effects model

Hausman test	(b)	(B)	(b-B)	sqrt(diag(V_b- V_B))
	FE	RE	Difference	S.E.
$V_{j,t-1}$	0.26	0.12	0.14	0.03
$V_{i,t-1}$	0.88	0.98	-0.10	0.017
$V_{i,t-2}$	-0.16	-0.14	-0.02	.
Test: Ho: difference in coefficients not systematic				
chi2(3)=(b-B)'[(V_b-V_B)^(-1)](b-B)= 24.42				
prob>chi2= 0.0000				

This fixed effect model is therefore defined as:

$$V_{it} = a_i + \beta V_{j,t-1} + \gamma V_{i,t-1} + \phi V_{i,t-2} + \varepsilon_{it} \quad (1)$$

$$i = 1, \dots, N; j = 1, \dots, N; t = 1, \dots, T_i$$

Where as defined earlier V_{it} is the speed for an observation unit (link) i in a given period t (3-minute intervals), α_i is a link-specific constant term in the model, which does not change over time, and ε_{it} is the usual residual.

In case of random effect model, the form can be defined as:

$$V_{it} = \alpha_i + \beta V_{j,t-1} + \gamma V_{i,t-1} + \phi V_{i,t-2} + v_i + \varepsilon_{it} \quad (2)$$

where v_i is a link-specific random element.

4.2. Model estimation results

The fixed effects model is estimated by the OLS regression, the results of the estimation is shown in the Table 3:

Table 3: Fixed effects model estimation results

Model: Fixed effects	Coef.	t-value	p-value	95% CI	
				Lower limit	Upper limit
$V_{j,t-1}$	0.26	5.31	0.00	0.16	0.35
$V_{i,t-1}$	0.88	13.94	0.00	0.75	1.00
$V_{i,t-2}$	-0.16	-2.93	0.00	-0.27	-0.05
Constant	0.94	1.72	0.09	-0.14	2.03

Observations	292				
Overall R-square	0.87				

From the estimation result, the p-value indicates that all coefficients are statistical significant from zero. The coefficient of variable $V_{j,t-1}$ from estimation is 0.26 with a positive sign, indicating that the speed on link i is positively correlated with speed on the downstream link. This indicates that increasing speed by one unit on the downstream link will result in an increment of 0.26m/h speed on the current link i in the same period. In the congestion situation, when the downstream link starts congested, the speed will be gradually reducing, consequently the speed on the following link will also be reduced. The intuition is consistent with the estimated coefficient sign.

Comparing the coefficients of $V_{i,t-1}$ and $V_{i,t-2}$, it appears the coefficient of $V_{i,t-1}$ is much higher than $V_{i,t-2}$ and has the opposite sign. This indicates that the influence of $V_{i,t-1}$ is much higher than $V_{i,t-2}$ and that speeds tend to oscillate when perturbed.

In order to inspect the goodness of fit of the model, we make a comparative plot between of the observed & fitted speeds on links 10 and 304.

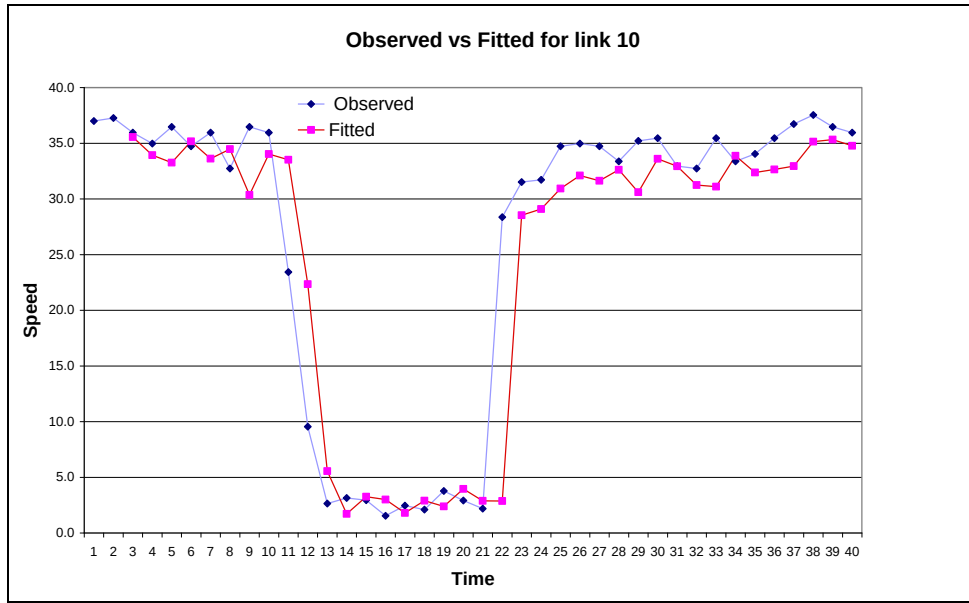


Figure 5: Observed vs fitted speed on link 10

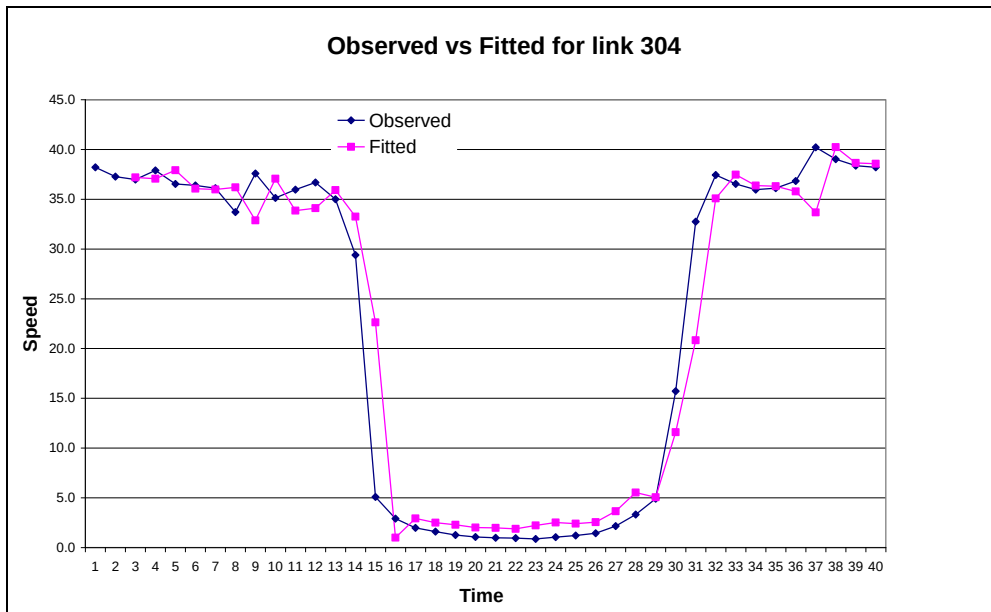


Figure 6: Observed vs fitted speed on link 304

The estimated speeds are consistent with the observed speeds. The estimated values follow the trend of actual speed changes. However, there is one lag delay between fitted data and observed data. This is due to prediction of our model largely relied on the value of one time lagged speed on the same link. It is evident from the value of estimated coefficient imposing

upon the $V_{i,t-1}$, which is the largest among other coefficients. This causes the result of fitted value highly correlated to the one time lagged value. It is suggested that reducing the time interval may improve the model accuracy.

It also needs to be noted that the estimated values tend to underestimate the actual speed values, especially in periods before and after incident. A possible explanation for this is because the model only has a few explanatory variables, it cannot capture all effects on the dependent variable. It may also be the case that the model form is misspecified with respect to the variables currently included. Hence, other models are being considered.

4.3. Other models

Other possible forms of model have also been explored in this study. Potential explanatory variable $V_{j,t}$, the speed on the downstream link at time t , is included in these models. The form of model B is as follows:

$$V_{it} = a_i + \beta V_{j,t} + \gamma V_{i,t-1} + \phi V_{i,t-2} + \varepsilon_{it} \quad (3)$$

$$i = 1, \dots, N; j = 1, \dots, N; t = 1, \dots, T_i$$

The model presumes the speed on link i depends on the speed on the downstream link simultaneously (that is V_{jt}) instead of speed of downstream link at time $t-1$ (that is $V_{j,t-1}$).

The model C will be:

$$V_{it} = a_i + \beta V_{j,t} + \beta' V_{j,t-1} + \gamma V_{i,t-1} + \phi V_{i,t-2} + \varepsilon_{it} \quad (4)$$

$$i = 1, \dots, N; j = 1, \dots, N; t = 1, \dots, T_i$$

This model assumes both the $V_{j,t}$ and $V_{j,t-1}$ affect the speed on link i at time t .

The results of the two model estimations are as follows:

Table 4: Other models estimation results

Model: Fixed Effects	Model B			Model C		
	Coef.	t-value	p-value	Coef.	t-value	p-value
$V_{j,t}$	0.27	7.65	0.00	0.24	5.51	0.00
$V_{j,t-1}$	.	.	.	0.09	1.53	0.13
$V_{i,t-1}$	0.85	14.52	0.00	0.82	13.30	0.00
$V_{i,t-2}$	-0.12	-2.23	0.03	-0.13	-2.44	0.02
Constant	0.40	0.75	0.45	0.39	0.73	0.47
Observations	289			289		
Overall R-square	0.87			0.88		

From the results in Table 4, the coefficients in model B are statistically significant as suggested by the p -values, except the constant. The value of the coefficients emphasize the importance of each explanatory variables on the speed on link i at time t . The coefficient of $V_{i,t-1}$ is the highest among others. It indicates the speed on link i at time t is largely dependent on the speed in the most previous time interval $t-1$ on the link i . The finding is consistent with our basic model A.

Model C includes one more variable than models A and B. The estimation results show that the coefficient of $V_{j,t-1}$ and constant term are statistically insignificantly different from zero. Again the coefficient of $V_{i,t-1}$ is the highest among others.

4.4. Comments on models

All three models show that the speed V_{it} is largely dependent on the speed in the previous time interval on the same link. This is shown by the value of the coefficient for $V_{i,t-1}$. Model A's $V_{i,t-1}$ coefficient is 0.88, models B and C are 0.85 and 0.82 respectively, which are reassuringly consistent with each other.

In terms of the prediction, only model A could be used to predict the speed on a link at future time period, because the model only requires information from past time periods, $V_{j,t-1}$, $V_{i,t-1}$ and $V_{i,t-2}$. Both models B and C involve the variable V_{jt} . That means we need to know the speed on the downstream link for the prediction period. However, models B and C are still valuable as a comparison for model A and provide an insight into the problem.

5. APPLICATION

In in-vehicle route guidance, in order to provide the driver with a fast and reliable route, current and anticipated congestion have to be taken into account. A recently developed route guidance approach, ARIAdNE [24], incorporates link travel time reliability in the search for a set of efficient, reliable, maximally disjoint paths. The main concept behind the operation of

ARIAdNE is the fact that specific links, classified as unreliable, are avoided as much as possible by the route finding algorithm, such that the probability of their inclusion in a route is low. However, ARIAdNE currently lacks the ability to forecast future congestion. The proposed spatial econometrics model in this study is expected to be used in conjunction with ARIAdNE to identify unreliable links, i.e. links on which congestion is likely to occur. The model shall provide ARIAdNE with information on future traffic congestion, which will enable it to not only take into account the current traffic situation, but also to consider future traffic conditions, thus making it more robust.

The model presented in this paper is designed to run in parallel with ARIAdNE in the vehicle, such that a historical database of TMC incident messages will be stored in the system. The model will first be estimated and calibrated off-line against the historical TMC data, and the calibrated model will then be taken on-board. The model will then recursively update itself with the real-time TMC messages, as soon as they are received. The predicted congestion information will be used to inform ARIAdNE which links are unreliable, such that they can be avoided by the route finding procedure.

6. CONCLUSION

The aim of this study was to set up econometric models to predict the congestion propagation and dissemination on the road network for eventual use in car navigation systems. Before the modelling was carried out, the nature and causes of the problem are illustrated using a VISSIM model for Coventry. The spatial and temporal nature of link congestion pointed to a spatial econometric approach. Speed data for eight links was simulated. Simple linear panel data models were fitted, with a view to gaining a model that could be used for forecasting.

The results showed that the fitted value of speed from model is fairly consistent with actual speed. However, only a limited number of explanatory variables are considered in the model, so we could not explain all the variation of speed on the selected links.

For the future work, other forms of model possibly involving more explanatory variables will be investigated. Currently, the regression model has a simple linear form with spatial and autocorrelation. The support for alternative non-linear regression models will be explored. Special treatment for autocorrelation among variables should be further considered in the model, the suggested technique being co-integration. Finally, we intend to build a model to forecast congestion along a route based on real TMC data.

ACKNOWLEDGEMENT

Authors would like to thank Mott Macdonald Ltd and Coventry city council for their permission to use the Coventry model for this study.

REFERENCES:

1. Chen, Y. Kaparias, I. Bell, M. G. H. & Bogenberger, K. (2005), '*Reliable autonomous route guidance by a constrained A* search considering intersection delays*', The Reliability of Travelling and the Robustness of Transport Systems, TRAIL Research School, Delft, The Netherlands, pp. 69-92.
2. Chen, Y. Bell, M. G. H. & Bogenberger, K. (2006), '*Reliable pre-trip multi-path planning and dynamic adaptation for a centralized road navigation system*', ITSC 2005 - 8th International IEEE Conference on Intelligent Transportation Systems, pp. 257-262.
3. Kaparias, I., Bell M. G. H., Bogenberger, K. & Chen, Y. (2007), '*An approach to time-dependence and reliability in dynamic route guidance*'. Transportation Research Record 2039, pp. 32-41.
4. Kaparias, I., Bell, M. G. H., Chen, Y., & Bogenberger, K. (2007), '*ICNavS: A tool of reliable dynamic route guidance*', IET Intelligent Transport Systems 1 (4), pp. 225-233.
5. Hu, J. (2006) '*Reliability and Link Failure Dependence Relationships in Dynamic Route Guidance*', MSc thesis. Imperial College London.
6. Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N., & Mishalani, R. (1998), '*DynaMIT: a simulation-based system for traffic prediction*', Delft, The Netherlands.
7. Wunderlich, K. E., Kaufman, D. E., & Smith, R. L. (1998), '*Link travel time prediction for decentralized route guidance architectures*', The university of Michigan ITS research centre.
8. Hounsell, N. Ishtiaq, S. (1997), '*Journey time forecasting for dynamic route dynamic systems in incident conditions*', International journal of forecasting 13.
9. van Lint, H. (2004), '*Reliable travel time prediction for freeways*', Delft University of Technology, Netherland.

10. Horvitz, E., Apacible, J., Sarin, R. & Liao, L. (2005), '*Prediction, Expectation, and Surprise: Methods, Designs, and Study of a Deployed Traffic Forecasting Service*', Twenty-First Conference on Uncertainty in Artificial Intelligence, UAI-2005.
11. Nihan, N. L., and Holmesland, K. O. (1980), '*Use of the Box and Jenkins time series technique in traffic forecasting*', Transportation 9, pp. 125-143.
12. Kamarianakis, Y., and Prastacos, P. (2005), '*Space-time modelling of traffic flow*', Computer & Geosciences 31, pp. 119-133.
13. van Hinsbergen, C. & van Lint, H., '*Bayesian combination of travel time prediction models*', in 87th TRB meeting, Washington D.C.
14. Anselin, L. (1988), '*spatial econometrics: methods and models*', Kluwer Academic Publishers.
15. Paez, A., and Scott, D. M. (2004), '*Spatial statistics for urban analysis: a review of techniques with examples*', Journal of Geographic System 61, pp. 53-67.
16. Zhou, B. and Kockelman, K., '*Neighborhood impacts on land use change: a multinomial logit model of spatial relationships*', 52nd North American Regional Science Association International Conference, Las Vegas, Nevada.
17. Zhao, H. and Bhat, C. (2002), '*the spatial analysis of activity stop generation*', Transportation Research B 36, pp. 557-575.
18. Guo, J.Y. and Bhat, C. (2007), '*operationalizing the concept of neighbourhood: application to residential location choice analysis*', Journal of Transport Geography 15, pp. 31-45
19. Aerts, K., Lathuy, C., Steenberghen, T., & Thomas, I. (2005), '*spatial clustering of traffic accidents distances along the network*'.
20. Bernard, M., Hackney, J. K., & Axhausen, K. W. '*Correlation of link travel speeds*', Monte Verita, Switzerland.

21. Cusworth, R. Riat, J. Storey, B. & Dreher, F. (2005) '*Coventry ring road modelling report Friday and Saturday*', Technical report of Mott Macdonald.
22. Greene, W. (2003), '*Econometric analysis*', Fifth edition. Prentice hall.
23. Baltagi, B. (1995), '*Econometric analysis of panel data*', John Wiley & Sons. New York.
24. Kaparias, I., Bell, M. G. H. & Belzner, H. (2008), '*Field testing of a reliable in-vehicle navigation algorithm*', 10th International Conference on Applications of Advanced Technologies in Transportation, Athens, Greece.