SHORT TERM FORECASTING OF URBAN TRAFFIC CONGESTION

Nick HOUNSELL, Saeed ISHTIAQ and Mike MCDONALD
Transportation Research Group
University of Southampton
SOUTHAMPTON - U.K.

1. INTRODUCTION

Traffic congestion is becoming increasingly common in many cities of the world as the demand for road travel continues to grow. This has led to the development of a number of traffic management and control systems aimed at making best use of existing roadspace. Traffic responsive Urban Traffic Control Systems, such as SCOOT (Hunt et al, 1981) are one such example.

Traffic congestion is a **dynamic** phenomenon which varies both within and between days according to regular and random variability in link/network demand and capacity. Dynamic methods are therefore required for the effective management and control of congestion. While many of the existing dynamic methods are **responsive** to congestion, new techniques involving signal control and/or vehicle routing require a **predictive** element involving short term forecasting of traffic conditions.

This paper describes research which is underway at Southampton University concerning short term forecasting methods for traffic applications. The fundamental research is sponsored by the Science and Engineering Research Council, while the applications are related to projects within the European Commission's DRIVE programme in which the University of Southampton is involved, particularly SCOPE and LLAMD (see section 7).

2. REQUIREMENTS

Requirements for short term forecasting can be identified in many new dynamic traffic management and control systems, such as:

* Traffic control systems, in which a forecast of congestion could trigger a 'remedial' control strategy, such as 'gating' where the rate of traffic inflow to a region is metered using appropriate signal timings.
* Dynamic route guidance systems, where the calculation of optimum routes requires a forecast of traffic conditions on links for the time at which vehicles will arrive on those links.

* Parking guidance systems, where a forecast of car park occupancy is required for efficient guidance.

In its 'simplest' form, requirements are for a forecast of traffic conditions on a link-by-link basis for periods of up to about 1 hour ahead depending on the application. For traffic control, a much shorter range forecast will typically be most appropriate, while for routeing systems, the required forecast horizon depends on the journey time from the point of routeing advice to the destination.

The parameter to be forecast also depends on the application. Again taking the above two examples, traffic control systems may require forecast of traffic demand and queue length, while route guidance systems require link journey time/cost forecasts. It is also likely that key descriptors of congestion will vary depending on the 'user'. For example, in network traffic control, the existence of queues blocking upstream junctions may be a key factor, whereas for a driver, congestion may be perceived as low averge speed, high delay, etc.

In the following sections, discussion is centred on the forecasting of average delay per vehicle. This parameter is of relevance for traffic control (e.g. being directly related to queue length) and is the key component of link journey times (for route guidance systems) which describes the level of congestion. Delay per vehicle is also relatively easily interpreted.

3. TIME-DEPENDENT VARIABILITY

The development of appropriate short-term forecasting techniques requires an initial analysis of the underlying time dependent variability in the parameter to be forecast: This is important in forecasting since the forecast model selected to generate the forecasts should be compatible with the demand pattern of the data. Clearly, if the parameter values are relatively stable within and between days, the forecasting process is greatly simplified.

Considering the average delay per vehicle on a link as a parameter reflecting the level of congestion, a number of sources of time dependent variability can be identified:

(i) **Cyclic variability**: where a link is under signal control, queues and delays will vary in a cyclic manner according to the cycle time of the signals. This very short term variability is usually not of interest in forecasting, except for particular signal control applications, and a time aggregation for data of one signal cycle or more is therefore usually necessary.

(ii) **Variability by time of day** Most urban networks show pronounced and regular variations in congestion by time of day, such as peak (morning and evening) and off peak hours.

(iii) **Variability by day of week** Variability here is largely due to the variations in activities which occur on different days, such as working and non-working days.

(iv) **Variability by month** This 'seasonal' variability may be related to environmental changes, changes in 'work' practices (e.g. vacation periods) and so on.

(v) **Long term variability** This describes the underlying trend in the parameter, such as an underlying increase in congestion due to a growth in traffic demand which is outpacing the provision of additional capacity.

The sources of variability described above are 'predictable' and can be quantified to a reasonable extent by analysis of historic data patterns, and incorporated into the forecasting process. In addition, there are unpredictable short-term effects such as congestion caused by traffic incidents (accidents, breakdowns, roadworks etc) which are much more difficult to forecast but which are particularly important in the context of congestion and its control.

4. FORECASTING METHODS

A number of 'time-series' forecasting models are available which could be applied to the traffic situation. After taking consideration of different sources of variability, as described earlier, two methods have been selected for detailed evaluation, as described here.

1: Box-Jenkin ARIMA modelling. (Pankratz, 1983)

2: Horizontal-seasonal model. (Thomopoulos, 1980)

These are univariate forecasting methods; forecasts are based only on past patterns in the series being forecast, although updating is possible as new data is received.

4.1 Box-Jenkin ARIMA modelling

In term series analysis, the principal data that is used in generating forecasts is the history of the past demand entries (Z_1 , Z_2 , Z_2 , - - - , Z_t). The forecasting process begins by developing a statistical type of fit through the demands of the past. The fit that is sought may vary, depending on the flow pattern of the history of demands. The fit is then projected forward to estimate the demands for the time periods of the future. These estimates are essentially the forecasts of future demands.

Box and Jenkins propose a family of Algebraic models called ARIMA models (Auto Regressive Integrated Moving Average) from which we select one that seems appropriate for forecasting a given data series.

The term AutoRegressive means that Z_t, the current value of the time series, is "regressed" or expressed as a function of Z_{t-1}, Z_{t-2}, - - -, Z_{t-p} , which are the previous values of the same time series, and to an unknown noise a_t, in a linear manner by the relation

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + - - - + \phi_p Z_{t-p} + a_t$$

The parameters of Autoregressive Integrated Moving Average models are:

$$\text{ARIMA } (p,d,q)\ (P,Q,D)_s$$

where

p = Order of Non-Seasonal Autoregressive operator
d = Order of Non-Seasonal Differencing
q = Order of Non-Seasonal Moving-Average operator
P = Order of Seasonal Autoregressive operator
D = Order of Seasonal Differencing
Q = Order of Seasonal Moving-Average operator
s = Length of Seasonality

Here season refers to a period after which the pattern of the series repeats itself, e.g: traffic flows are expected to be higher during peak hours on every day, here day is one season and on different days the pattern of traffic flows are more or less the same.

The fundamental idea in ARIMA models is that the variable Z_t is related to its own past (Z_{t-1}, Z_{t-2}, Z_{t-3}, - - -) and (Z_{t-s}, Z_{t-2s}, Z_{t-3s}, - - -).

Where

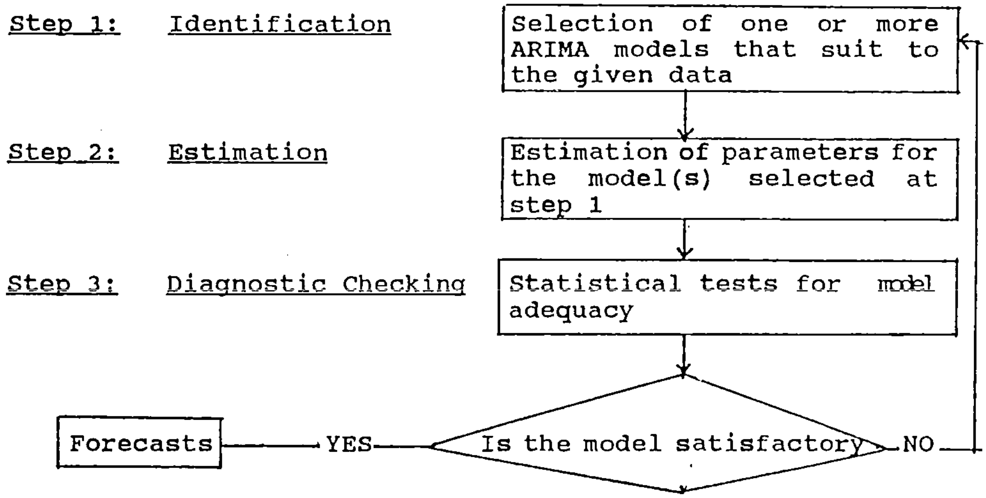
(Z_{t-1}, Z_{t-2}, Z_{t-3}, - - -) are the values of the series in the same season at different time periods, and
(Z_{t-s}, Z_{t-2s}, Z_{t-3s}, - - -) are the values of the series at same time period in different seasons.

Consider an example where traffic delays on different days at different time periods are denoted by

	Day1	Day2	Day3	Day4
t_1	$Delay_{11}$	$Delay_{12}$	$Delay_{13}$	$Delay_{14}$
t_2	$Delay_{21}$	$Delay_{22}$	$Delay_{23}$	$Delay_{24}$
t_3	$Delay_{31}$	$Delay_{32}$	$Delay_{33}$	$Delay_{34}$
- - - - - -	- - - - - -	- - - - - -	- - - - - -	- - - - - -
- - - - - -	- - - - - -	- - - - - -	- - - - - -	- - - - - -
t_n	$Delay_{n1}$	$Delay_{n2}$	$Delay_{n3}$	$Delay_{n4}$

$Delay_{34}$ is a function of $Delay_{24}$, $Delay_{14}$ and/or $Delay_{33}$ $Delay_{32}$, $Delay_{31}$

The Box-Jenkin method for finding a good ARIMA model involves three steps; Identification, Estimation and Diagnostic checking. The schematic representation of these three steps is as follow.



4.2 Horizontal-seasonal model

Most of the traffic parameters (flow, delay etc) show periodic behaviour within a day, i,e there are higher flows and delays during the peak period, this periodic behaviour of the parameters can be explained by so called seasonal models, here season refers to the period.

In these models smoothing of the past demand entries is used and in this way higher weights are assigned to the more current entries. The feature allows the forecasts to react quicker to more current shifts in the level or seasonal influences of the demands.

The model applies when the expected demand at time t is

$$\mu_t = \mu \rho_t$$

where

μ represents the average demand per day and

ρ_t represents the seasonal ratio at time t.

The seasonal ratio for time period t is found from the relation between μ_t and μ. The seasonal ratios are always greater or equal to zero and over a day their average value is 1, where ρ_t=1, the expected demand at time period t is the same as the average daily demand, where $\rho_t < 1$, then μ_t is less than μ and when $\rho_t > 1$, then μ_t is greater than μ.

Two phases are necessary in order to implement the model. The first is concerned with initialising the system and second is with updating the forecasts.

In the initialising phase the past demand entries (z_1, z_2, - - -, z_T) are used to find estimates of μ and ρ_t. The estimate of μ as of time t is $\hat{a}_t$ and estimates of the seasonal ratios ρ_{t+i} are r_{t+i} for i=1, 2, 3, - - -. These estimates can be obtained by the following recursive relations:

$$\hat{a}_t = \alpha(z_t/\hat{r}_t) + (1-\alpha)\ \hat{a}_{t-1}$$

$$\hat{r}_{t+n} = \gamma(z_t/\hat{a}_t) + (1-\gamma)\ \hat{r}_t$$

Here at time t, the ratio $(z_t/\hat{r}_t)$ represents the current seasonally adjusted demand, the ratio $(z_t/\hat{a}_t)$ gives the seasonal ratio for time period t.

α and γ are smoothing constants, their values should b between 0 and 1, higher values of α and γ gives more weight to the current data.

The most current n seasonal ratios

$$\hat{r}_{T+1},\ \hat{r}_{T+2},\ - - -,\ \hat{r}_{T+n}$$

are normalised so that their average is 1. This is performed by first finding the average

$$\bar{r} = (\hat{r}_{T+1} + \hat{r}_{T+2} + - - - + \hat{r}_{T+n})/n$$

and then adjusting the ratios by

$$r_{T+i} = \hat{r}_{T+i}/r \qquad \text{for } i = 1, 2, - - -, n)$$

Having carried out these steps, the initialisation phase is complete.

At this time the first set of forecasts can be generated. The forecast for the ith future time period is

$$\hat{z}_T(i) = \hat{a}_T r_{T+i}$$

<u>Updating</u>

As each new demand entry becomes available, an updating scheme is carried forward to yield the current estimates of the mean demand level and the seasonal ratios.

Calling the current time period T, the new observation is z_T and the updating relations are the following:

$$\hat{a}_T = \alpha\ (z_T/r_T) + (1-\alpha)\ \hat{a}_{T-1}$$

$$r_{T+n} = \gamma(z_T/\hat{a}_T) + (1-\gamma) r_T$$

As before, the seasonal ratios are normalised so that their average is 1.

With updating of the estimates completed, the forecasts for the ith future time period is generated by:

$$\hat{z}_T(i) = \hat{a}_T r_{T+i}$$

4.3 Examples

Examples of the application of the above models to data on two links in Southampton with different traffic characteristics are given in Figures 1 to 4. The data was obtained from the SCOOT UTC system using the ASTRID database facility (McLeod et al, 1989). Figures 1 and 2 show the application of the two models to a link with relatively low and regular congestion, while Figures 3 and 4 apply to a link with considerably higher variability. In each case, forecasts are being made for data for a particular peak period, and are compared with actual measurements. Tables (1) and (2) show the Forecast-Error statistics. In both cases the one-step ahead (or updated) forecasts show improvement over 36-step ahead (not updated) forecasts. Forecast-horizons were also studied, which show that forecasting errors become greater with the increase in forecasting horizon.

From the above examples and application of these models to some other links, it is seen that though Box-Jenkin modelling requires more rigorous analysis and computer time, it handles a wider variety of situations and provide more accurate short-term forecasts. Also, the concepts associated with Box-Jenkin models are derived from a solid foundation of classical probability theory and mathematical statistics, once a model is selected for the given data set, the estimation of the parameters are done by certain algorithms.

On the other hand, the Horizontal-Seasonal model is simple to apply and can be implemented on computer by any high level programming language, this can reduce the cost of forecasting process. However the choice of the smoothing parameters is critical as the values of the smoothing parameters can affect the forecasts but there is no available method which can select the optimum values of these smoothing parameters.

5. NETWORKS

The forecasting methods described above apply to individual links and do not describe the interaction between links and the build up of congestion which can occur on a network basis. This analysis of links on an individual basis is relevant, because it reflects the operation of many traffic control systems, such as isolated signal controlled junctions and dynamic route guidance. It also has the great merit that a forecast is generated for the specific parameter of interest (e.g. delay). However, congestion in urban networks is often associated with one or more "pinchpoints" from which queues spread to affect a number of upstream links. Queues on adjacent links are then interrelated both in time and space, and the forecasting of congestion on an independent link by link basis becomes less relevant.

A useful method for studying the time varying growth and decay of network congestion is the use of a suitable network traffic model. At the University of Southampton, work is centring on the CONTRAM dynamic traffic assignment model (Leonard et al, 1989). The objective is to study the time varying congestion phenomenon in a variety of network, traffic and control scenarios to try and then develop short term congestion forecasting procedures for networks. It may be possible, for example, to relate statistically the rate of growth of congestion with its key controlling parameters, such as the incident characteristics (severity, location, duration), the underlying levels of flow and capacity and so on.

The work is particularly concerned with incident-induced congestion, and a new logic has been written to better reflect drivers' routes and queueing characteristics for incident conditions than is available in "normal" traffic assignment models. Early work is illustrating the number of links affected by an incident and the effects on each link in terms of the increase in journey time per time interval. "Congestion trees" are also being studied, illustrating the growth, spread and decay of queues.

Short term forecasting of network congestion for use in control systems ideally requires the use of an on-line dynamic traffic assignment model to produce the forecast itself. Such models are also being recommended for other applications, such as dynamic route guidance. Research and development in this area is therefore likely to be of particular importance over the next few years, perhaps incorporating parallel processing techniques to enable rapid predictions and control recommendations to be obtained.

6. CONCLUDING COMMENTS

* Short term forecasting of urban traffic congestion will become an increasingly important requirement in traffic control and informatic systems. The current research has revealed suitable methods for such forecasting on a link-by-link basis and where congestion is reasonably recurrent. These methods could be further tested and refined on-line by adopting methods for automatic validation within systems where such forecasts are used (e.g. for route guidance). An expert system approach could also be adopted in which the performance of a range of forecasting methods is automatically monitored and the method chosen appropriate to the traffic situation.
* The current requirement is to develop forecasting procedures appropriate to unpredictable

congestion, such as that caused by traffic incidents and where congestion is a network phenomenon. This will inevitably require a dynamic network modelling approach incorporating traffic assignment, for the formulation of appropriate procedures and, perhaps, for on-line applications.

7. REFERENCES

- Pankratz, Alan. Forecasting with univariate Box-Jenkins models. John Wiley and Sons, 1983.

- Thomopoulos, Nick T. Applied forecasting methods. Prentice-Hall, 1980.

- SCOPE: Project in DRIVE II dealing with Integrated Urban Traffic Management in Southampton, Cologne and Piraeus.

- LLAMD: Project in DRIVE II dealing with Integrated Urban Traffic Management in London, Lyon, Amsterdam, Munich and Dublin.

- University of Southampton. The forecasting and control of congestion in urban networks. Study in progress for the U.K. Science and Engineering Research Council.

- CARGOES: Project in DRIVE I dealing with the integration of dynamic route guidance and traffic control systems.

- Hounsell, N. and McLeod, F. ASTRID: Automatic SCOOT traffic information database. Transport and Road Research Laboratory Contractors Report CR235, Crowthorne, 1990.

- Leonard, D.R. et al. CONTRAM: Structure of the model. Transport and Road Research Laboratory Report RR178, Crowthorne, 1989.

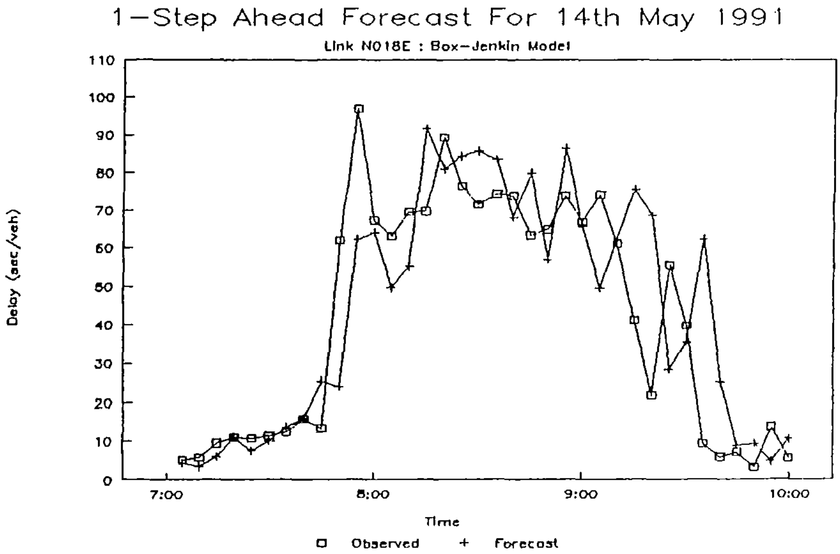


Figure 1

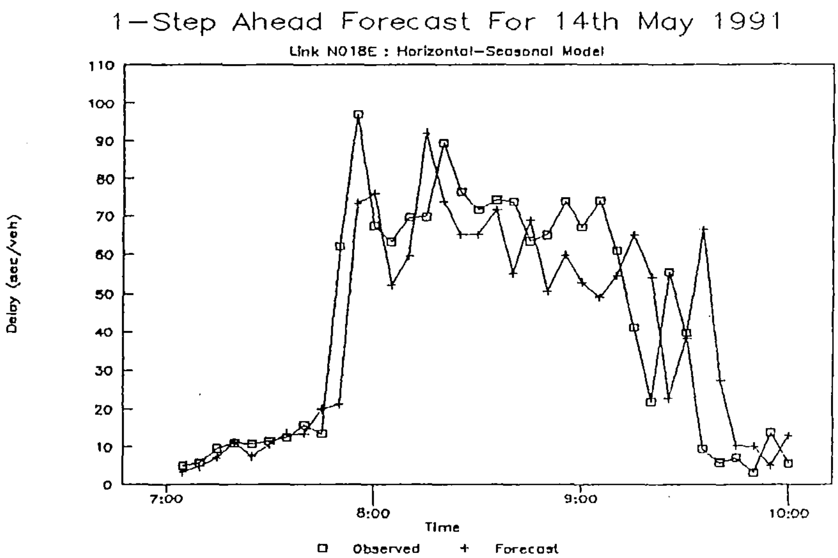


Figure 2

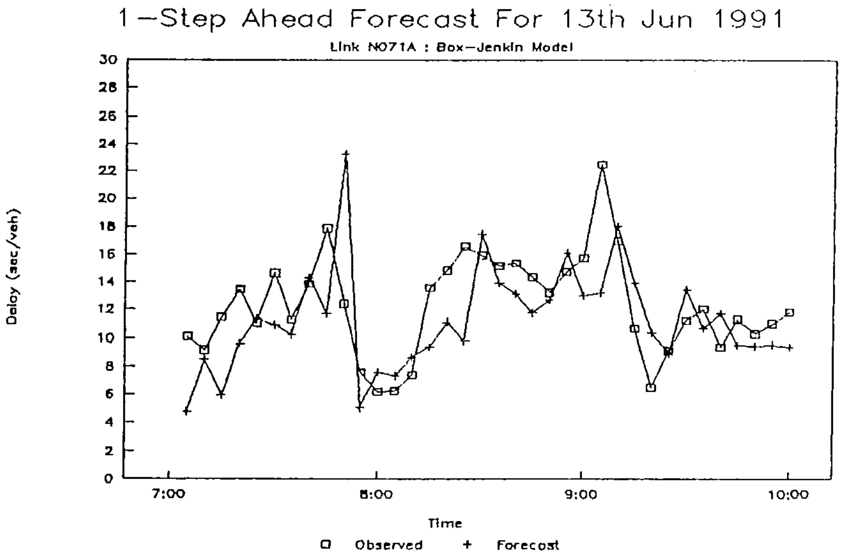


Figure 3

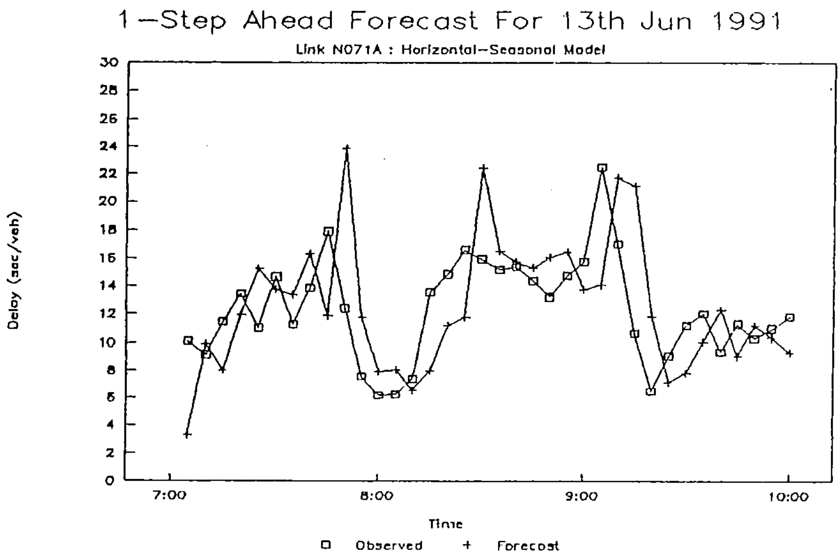


Figure 4

Link N071A : Forecast-Error Statistics

	Box-Jenkin		Horizontal-Seasonal	
	36-Step Ahead	1-Step Ahead	36-Step Ahead	1-Step Ahead
ME	2.22	1.09	5.66	-0.28
MAE	3.81	2.82	5.79	3.44
MSE	20.69	13.82	44.55	18.96
MPE	12.30	5.58	41.47	-5.78
MAPE	30.37	22.84	43.49	29.16

TABLE(1)

Link N018E : Forecast-Error Statistics

	Box-Jenkin		Horizontal-Seasonal	
	36-Step Ahead	1-Step Ahead	36-Step Ahead	1-Step Ahead
ME	28.58	-1.68	36.20	2.01
MAE	32.29	12.97	36.21	12.89
MSE	1640.44	352.51	2113.40	325.44
MPE	16.45	-35.24	72.78	-31.89
MAPE	88.45	61.40	73.11	62.62

TABLE(2)

ME Mean Error
MAE Mean Absolute Error
MSE Mean Square Error
MPE Mean Percentage Error
MAPE Mean Absolute Percentage Error