Adaptive Forecasting of Freeway Traffic Congestion

GARY A. DAVIS, NANCY L. NIHAN, MOHAMMAD M. HAMED, AND LESLIE N. JACOBSON

Problems of forecasting freeway traffic variables a few minutes in advance, particularly lane occupancy and the difference between inflowing and outflowing traffic for a short section of freeway, are explored. Methods based on linear time series analysis were found to do reasonably well at forecasting mean values but not so well for those extremes corresponding to the onset of congestion. Techniques based on statistical pattern recognition principles were found to be promising. The most promising of the pattern recognition algorithms was put into use on a section of I-5 and is being field tested.

Since 1981, the Washington State Department of Transportation (WSDOT) has used integrated traffic-responsive on-ramp control to cope with recurring traffic congestion on the Seattle region's portion of Interstate 5 (I-5). The algorithm used to calculate the on-ramp entry rates has two basic routines, local control and bottleneck control. In the local control routine, the algorithm calculates metering rates on the basis of the measured lane occupancy at the main-line station immediately upstream of the on-ramp merge point. This portion of the algorithm is effective only as long as the demand for use of a section of freeway does not greatly exceed the capacity of the section.

As is well known, both from traffic flow theory and from practical experience, when demand for access to a section of freeway is excessive, the operating speed of that section is reduced and vehicles queue at the points of congestion (bottlenecks). The bottleneck control routine of the WSDOT algorithm eliminates these speed reductions by restricting access to the freeway at one or more on-ramps upstream from the bottleneck point. A detailed description of the WSDOT control algorithm was provided by Jacobson et al. (*1*).

One limitation of the bottleneck control algorithm is that it is reactive rather than anticipatory. It does not take action until a bottleneck has formed, so speed reduction and instability already exist in the traffic stream. The bottleneck algorithm is oriented toward cleaning up messes rather than preventing them. If bottleneck formation could be forecast and bottleneck control could be used to prevent bottleneck formation, then at least in theory overall traffic volumes should increase and delay of the traveling public should decrease. The goal of recent research conducted at the University of Washington is to develop an algorithm that can reliably forecast bottleneck formation 1 min or more in advance of occurrence and to incorporate this new algorithm into the WSDOT ramp control algorithm.

G. A. Davis, Department of Civil and Mineral Engineering, University of Minnesota, Minneapolis, Minn. 55455. N. L. Nihan and M. M. Hamed, Department of Civil Engineering, University of Washington, Seattle, Wash. 98105. L. N. Jacobson, Washington State Department of Transportation, Seattle, Wash. 98105.

BOTTLENECK CONTROL ALGORITHM

The placement of main-line loop detector stations divides the I-5 main line into short (½-mi or less) sections. Each section is bounded upstream and downstream by a set of loop detectors, one in each lane. On-ramps merge and off-ramps diverge inside these section boundaries. The loop detectors provide the freeway controller with measurements of traffic volume and lane occupancy. The controller transmits the volume and occupancy data to the freeway management system computer.

The computer uses 1-min moving averages of main-line and ramp volume and occupancy to compute on-ramp metering rates every 20 sec. From the main-line and ramp volume measurements, the computer calculates the difference between the traffic entering and exiting a freeway section; this difference is called the storage rate for the freeway section. When the difference is positive, indicating that more vehicles have entered the section than have left it and when the measured lane occupancy at the section's downstream boundary exceeds 18 percent, indicating that the traffic flow is near capacity, the computer concludes that a bottleneck has formed. The computer reduces the entry rates at on-ramps upstream from the section's downstream boundary by the storage rate. If a straightforward anticipatory algorithm could be developed to forecast each section's downstream occupancy and storage rate, then when these forecast values indicate a bottleneck, the computer could use them in the bottleneck algorithm in place of actual measurements. The bottleneck forecasting algorithm would come into effect before the occurrence of a bottleneck, and metering rates would be calculated to prevent or minimize the bottleneck.

The research team used a bottleneck section of southbound I-5 approximately 12 mi north of downtown Seattle as a study site. On this section, congestion routinely begins on weekday mornings. The congestion from this section affects traffic flow for several miles upstream. This section was designated as Section 2, the section immediately downstream as Section 1, and the section immediately upstream as Section 3. Figure 1 shows the geometry of this study area. The total length of the three sections was about 1.3 mi. The research team's goal was to develop a routine that would forecast lane occupancy and storage rate for Section 2, given past measurements from Sections 1, 2, and 3.

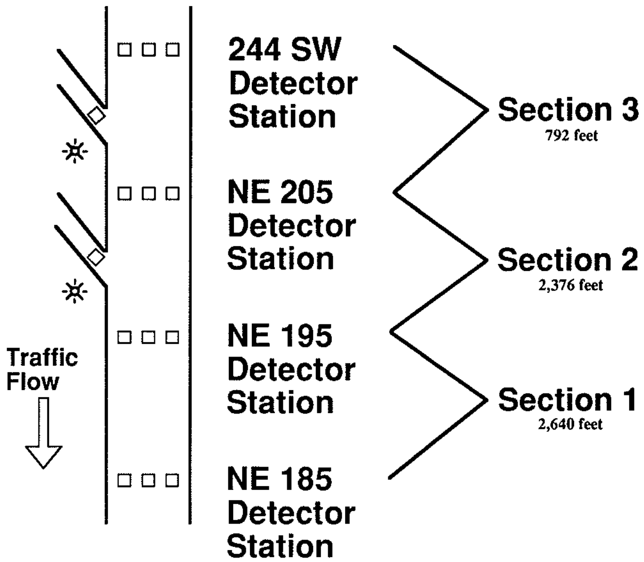

FIGURE 1 Test section of southbound I-5.

The researchers collected 1-min volume and occupancy data from the main-line stations defining these sections. One-minute volume data were collected from all the ramps in these sections as well. Data were collected from 6:00 to 8:00 a.m. on five weekdays in late November and early December of 1988. These data were used to develop and test potential forecasting routines.

TIME SERIES FORECASTING

Results presented by Ahmed and Cook (*2*) and by Kyte et al. (*3*) indicate that the time series methods described in Box and Jenkins (*4*) can be used to model the short-term dynamics of freeway traffic flow variables. In the Box and Jenkins (*4*) method, data are first collected in time series form, that is, as measurements indexed by time. The autocorrelation functions (the correlation of a variable with its own past) and the cross-correlation functions (the correlation of a variable with another variable's present and past) are then inspected to obtain candidate models. These models all express some variable's current value as a linear combination of its own and other variables' past values. The linear weights for each candidate model are then estimated and the models compared. The best model is then selected as the forecaster.

The researchers prepared time series of 1-min storage rate and average lane occupancy for Sections 1, 2, and 3 from the data. They fit univariate and transfer function models to the data from the 1st day using standard Box-Jenkins (*4*) methods and then used the fitted models to generate 1-min-ahead forecasts of Section 2 occupancy and storage rate for each value of the data from each of the 4 days. The detection rule was as follows:

If forecast occupancy > 18 percent and forecast storage rate > 0, then forecast a bottleneck.

The estimated auto- and cross-correlation functions had potential for predicting occupancy and storage rates from Section 2, both from past measurements at Section 2 and from past measurements from the downstream Section 1. This latter result appears to be caused by the propagation of shock waves upstream from Section 1 into Section 2. It has been identified in cross-correlation functions by Kyte et al. (*3*). Table 1 presents the detection rates for the time series forecasts on four of the study days. The Percent Correct column gives the unconditional percentage of correct forecasts, whereas the False Positive and False Negative columns give conditional percentages; that is, the False Positive column gives the percentage of nonqueuing (nonbottleneck) intervals that were falsely predicted to show queuing, whereas the False Negative column gives the percentage of queuing (bottleneck) intervals falsely predicted not to show queuing. These conditional rates are more informative than unconditional rates, such as the Percent Correct, because they correct for the high proportion of nonqueuing intervals in the available data sample. The rates presented in Table 1 are not encouraging.

Figures 2 and 3 show sequences of actual and 1-min-ahead forecasts of storage rate for Section 2 on a lightly congested day. The forecast values of storage rate tend to hover around the mean value for the time series and ignore extreme values.

TABLE 1 DETECTION RATES FOR TIME SERIES FORECASTS

Date	Percent Correct	False Positives	False Negatives
11/23/88	94	0	64
12/13/88	94	1	67
12/14/88	82	3	58
12/15/88	96	0	83

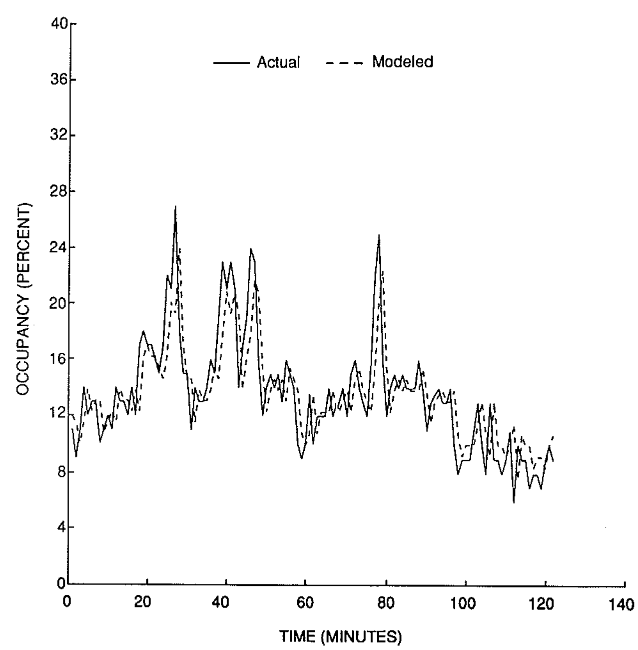

FIGURE 2 Actual and modeled occupancies for Section 2.

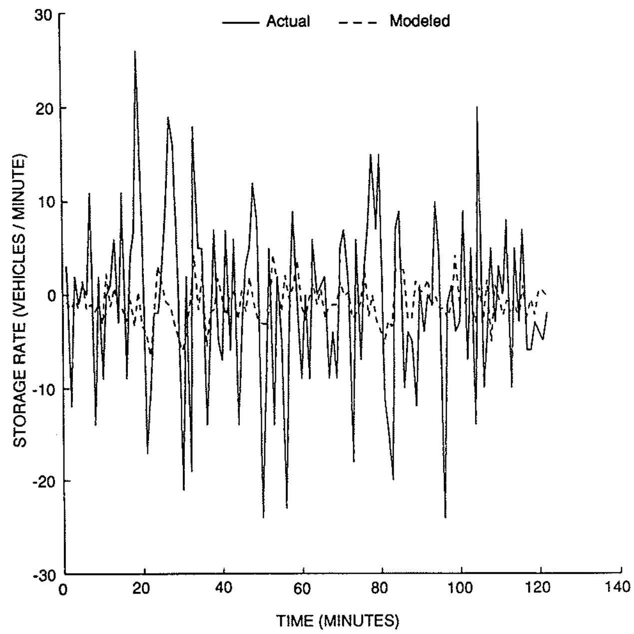

FIGURE 3 Actual and modeled storage rates for Section 2.

This tendency of time series forecasters to avoid extreme values is especially distressing because the extreme values are of primary interest. The forecast values of occupancy track the true values, but 1 min too late. Similar results were found for a number of univariate and multivariate time series models, and Davis and Nihan (*5*) have reported similar findings for the nonparametric nearest-neighbor forecaster. Least-squares-based approaches, although tending to minimize the average error between forecast and actuality, were not promising candidates to predict the formation of bottlenecks.

PATTERN RECOGNITION APPROACH

In statistical pattern recognition, the primary activity is to sort observations into two or more categories (*6*). Using the data, the 1-min intervals were sorted into queuing and nonqueuing intervals. Queuing intervals were those for which the data met the bottleneck criteria in Section 2 (positive storage rate, occupancy > 18 percent). Nonqueuing intervals did not meet those criteria. Measurements from previous (lagged) intervals for Sections 1, 2, and 3 could then be sorted into those that preceded bottleneck formation and those that did not. A rule based on these lagged measurements was developed that discriminated intervals with bottlenecks from intervals without bottlenecks. For example, suppose the storage rates from Section 1 measured 3 min in the past ($SR1_{t-3}$), preceding the bottleneck intervals in Section 2, are normally distributed with a mean of 4 veh/min and a standard deviation (SD) of 8 veh/min, whereas those $SR1_{t-3}$ values preceding nonbottleneck intervals are normal with a mean of −2 veh/min and SD of 9 veh/min. Obviously, larger positive values of $SR1_{t-3}$ are more likely to precede bottlenecks in Section 2. The rule

If $SR1_{t-3} > -2 + (2 * 9)$
$= 16$ veh/min, then forecast a bottleneck

would falsely anticipate a bottleneck about 5 percent of the times when bottlenecks did not occur. However, the value of 16 veh/min corresponds to the standard score of 1.5 on the distribution for bottleneck predecessors. From a table of normal probabilities, about 93 percent of the bottleneck intervals will be missed when they do occur. Therefore, a low false positive rate (5 percent) is traded for a high false negative rate (93 percent). Any attempt to lower the false negative rate by lowering the cutoff will produce a corresponding increase in the false positive rate. Obviously, distributions with large differences between their means and low standard deviations will produce data that are relatively easy to classify, whereas distributions with substantial overlap will produce ambiguous data.

In this pattern recognition approach, variables must be identified that show good separation between bottleneck and nonbottleneck conditions. Using the boxplot feature of Minitab (*7*), the researchers evaluated the storage rate and occupancy measurements at time intervals lagged 1, 2, and 3 min for Sections 1, 2, and 3 to determine which had the greatest ability to discriminate between bottleneck and nonbottleneck intervals. Figure 4 shows boxplots for the two variables that appear to give the best discrimination. In interpreting a boxplot, note that the interval between the *I* symbols corresponds to the interval between the first and third quartiles of the frequency distribution of the data. The + symbols indicate the location of the estimated median of the distribution, and the parentheses give a 95 percent confidence region around the estimate of the median. The single lines extending to the right and left of the quartile boundaries indicate the probable range of the distribution, while the symbols mark the location of possible outliers.

From the boxplots, one occupancy (OC) variable and one SR variable were selected to be used as forecasters for lightly congested data (November 23, 1988) and another set was selected to be used for heavily congested data (November 22, 1988). For lightly congested data, the OC values from Section 2 lagged one interval ($OC2_{t-1}$) and the SR values from Section 1 lagged two intervals ($SR1_{t-2}$) were selected. For heavily congested data, the same occupancy variable ($OC2_{t-1}$) and storage rate from Section 1 lagged three intervals ($SR1_{t-3}$)

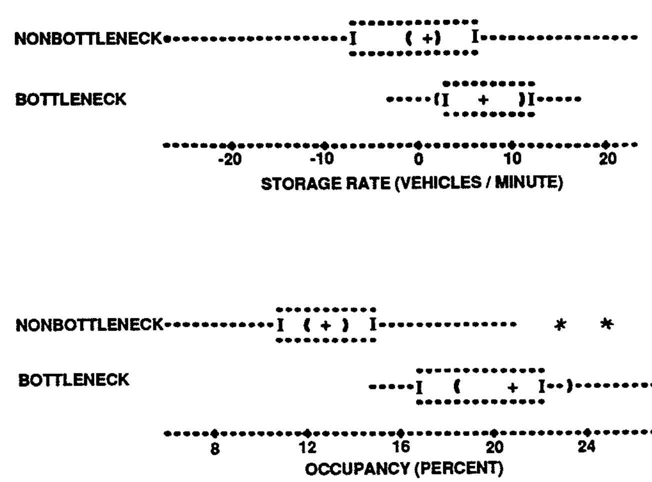

FIGURE 4 Boxplots for Section 1 storage rate, Lag 2 (*top*), and for Section 2 occupancy, Lag 1 (*bottom*).

were selected. The following simple forecasting rule was adopted:

If the measurement exceeds the 75th percentile of the non-bottleneck distribution, then forecast a bottleneck.

This rule resulted in a 25 percent false positive rate in all cases, and a false negative rate depending on the variable used. In addition to separate forecasts for the SR and OC variables, the following compound forecasts were constructed:

If OC indicates a bottleneck *and* SR indicates a bottleneck, forecast a bottleneck.

If OC indicates a bottleneck *or* SR indicates a bottleneck, forecast a bottleneck.

Table 2 presents the percent of correct forecasts, the false positive rate, and the false negative rate for each pattern recognition forecaster and for each of the 5 days for which data were collected. The traffic congestion on November 22 was considered heavy, whereas that on November 13, 15, and 23 was considered light. The congestion on November 14 was intermediate between light and heavy, and forecasts were made using the light congestion rules. Generally, for a given false positive rate of about 25 percent, the individual forecasts have lower accuracy rates and higher false negative rates as traffic congestion increases. However, the *anded* forecasts for lightly congested days appear to do well, with false positive rates ranging from 5 to 11 percent and false negative rates ranging from 33 to 36 percent. Apparently, the *anded* forecast could prove useful in forecasting those bottlenecks that characterize the transition from uncongested flow to congested flow, but after congestion sets in, forecasting bottlenecks becomes more difficult.

TABLE 2 DETECTION RATES FOR PATTERN RECOGNITION FORECASTS

Date	Forecaster	Percent Correct	False Positives	False Negatives
11/22/88	SR	63	28	49
	OC	68	24	47
	AND	68	7	73
	OR	63	46	22
11/23/88	SR	71	29	36
	OC	77	25	0
	AND	92	5	36
	OR	55	50	0
12/13/88	SR	76	23	33
	OC	57	46	0
	AND	88	11	33
	OR	45	60	0
12/14/88	SR	69	25	48
	OC	56	58	3
	AND	78	11	52
	OR	45	72	0
12/15/88	SR	67	33	33
	OC	82	23	0
	AND	89	10	33
	OR	55	47	0

IMPLEMENTATION

For practical purposes, the critical errors are false positives. In the WSDOT freeway management system, if a bottleneck is not forecast but one occurs, the situation is what currently exists—a responsive reaction to the bottleneck after it is detected. (In essence, the current system predicts bottlenecks with a 100 percent false negative rate.) However, a false positive forecast of a bottleneck would lead the system to calculate metering rates that would be more restrictive than necessary. This restriction could lead to excessive ramp queues and possibly a more congested freeway as the queues extend to the queue detectors and metering rates are set higher as a result. The false positive rates presented in Table 2 were within an acceptable range, while the false negative rates were low enough to indicate that an algorithm based on this forecaster would likely improve the existing bottleneck control algorithm. Therefore, WSDOT has incorporated this forecaster into its bottleneck control algorithm for field testing.

The forecaster's simplicity fits well into real time process control. The forecasting algorithm was programmed as a two-step process. First, the computer program checks occupancy at Section 2. If the occupancy is above the level that indicates a positive prediction (13 percent), the algorithm checks the storage rate for Section 1 from the previous minute. If the storage rate is greater than six vehicles, the algorithm predicts the formation of a bottleneck in Section 2 during the next minute and calls the normal bottleneck algorithm. The bottleneck algorithm uses an average storage rate for the metering rate reduction that it distributes over upstream ramps.

The test period is underway. Data are being collected that will be analyzed using time series analysis to determine if the forecasting and early metering intervention has a significant impact on traffic flow.

CONCLUSION

The onset of capacity-reducing freeway congestion should be forecast so that measures can be taken to prevent its formation. Box-Jenkins-type time series methods tend to concentrate on forecasting mean values and avoiding extremes. However, the occurrence of extreme values is what characterizes the transition from smoothly flowing traffic to the stop-and-go traffic that must be forecast. An approach based on statistical pattern recognition principles was aimed at identifying those variables with conditional distributions that show good discrimination between the presence and absence of traffic bottlenecks.

Although the results must be regarded as preliminary, they show enough promise to be field tested. Other combinations of variables, including upstream volume, may improve the algorithm's forecasting ability.

REFERENCES

1. L. Jacobson, K. Henry, and O. Mehyar. Real-Time Metering Algorithm for Centralized Control. In *Transportation Research Record 1232*, TRB, National Research Council, Washington, D.C., 1989.

2. S. Ahmed and A. Cook. Application of Time-Series Analysis Techniques to Freeway Incident Detection. In *Transportation Research Record 841*, TRB, National Research Council, Washington, D.C., 1982, pp. 19–21.
3. M. Kyte, J. Marek, and D. Frith. Using Multivariate Time-Series Analysis to Model Freeway Traffic Flow. Presented at 68th Annual Meeting of the Transportation Research Board, Washington, D.C., 1989.
4. G. Box and G. Jenkins. *Time-Series Analysis: Forecasting and Control.* Revised ed., Holden-Day, San Francisco, 1976.
5. G. Davis and N. Nihan. An Application of a Nonparametric Regression Technique to Short-Term Freeway Traffic Forecasting. *ASCE Journal of Transportation Engineering*, 1989.
6. C. Chen. *Statistical Pattern Recognition*. Spartan Books, Rochelle Park, N.J., 1973.
7. T. Ryan, L. Joiner, and B. Ryan. *Minitab 82 Reference Manual.* Document N434. Academic Computer Center, University of Washington, Seattle, 1982.

Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.