Dynamic prediction of traffic volume through Kalman filtering theory

Article *in* Transportation Research Part B Methodological · February 1984

DOI: 10.1016/0191-2615(84)90002-X · Source: RePEc

Transpn Res.-B Vol. 18B, No. 1, pp. 1 11, 1984
Printed in Great Britain

0191-2615/84 $3.00 + .00
Pergamon Press Ltd.

DYNAMIC PREDICTION OF TRAFFIC VOLUME THROUGH KALMAN FILTERING THEORY

IWAO OKUTANI
Shinshu University, Nagano, Japan

and

YORGOS J. STEPHANEDES
University of Minnesota, Minneapolis, MN 55455, U.S.A.

(*Received* 17 *September* 1982; *in revised form* 14 *March* 1983)

Abstract—Two models employing Kalman filtering theory are proposed for predicting short-term traffic volume. Prediction parameters are improved using the most recent prediction error and better volume prediction on a link is achieved by taking into account data from a number of links. Based on data collected from a street network in Nagoya City, average prediction error is found to be less than 9% and maximum error less than 30%. The new models perform substantially (up to 80%) better than UTCS-2.

INTRODUCTION

Predicting traffic flow or volume on a link of a heavily utilized road network in order to control traffic there is one of numerous problems associated with computerized signal control. If traffic congestion can be foreseen on a link in a central area of the street network, congestion could be avoided in several ways. For example, upstream vehicles could be instructed to take alternative route(s) or green time of traffic signals in the outer area could be reduced for vehicles feeding into the link. If it is foreseen that the traffic flow will exceed the predetermined safety level on a link in an urban expressway, this could be prevented by controlling the traffic volume flowing into the expressway at the toll gates of on-ramps. Optimal traffic signal settings could, of course, best be determined if the value of future traffic flow rather than present or past data were employed.

BACKGROUND

Existing demand prediction algorithms fall into two major categories, i.e. the second generation and the third generation. The former are designed for control intervals on the order of 5–15 min. They are older and typically require historical data as reference. They use current traffic measurements to correct for the traffic deviations from the average historical pattern. Second generation UTCS (UTCS-2) (FHWA, 1973), ASCOT (SRI, 1977), and ASCOT-RTOP (CMT, 1976) all belong to this category.

Third generation algorithms (FHWA, 1973; Mengert *et al.*, 1979), generally more recent than the second generation, were developed with the objective of making predictions based on current traffic measurement only. However, the third-generation UTCS (Lieberman *et al.*, 1974), the best known algorithm in this category, requires a "representative" data set for estimating prediction coefficients (Stephanedes *et al.*, 1981).

More recent studies have used Box-Jenkins-type analyses of time series data (Eldor, 1977; Gafarian *et al.*, 1977; Ahmed and Cook, 1979), spectral analysis (Nicholson and Swann, 1974), and Kalman filtering (Nahi, 1973; Chang and Gazis, 1975; Gazis and Knapp, 1971). Other filter-predictors have also been developed for use in traffic-responsive computer control of urban traffic (Baras and Levine, 1977; Baras *et al.*, 1978; Baras *et al.*, 1979); these, however, predict queue size at an intersection rather than demand.

Predictions using Box-Jenkins type analyses have at times resulted in unsatisfactory goodness of fit and high errors (Gafarian *et al.*, 1977); in certain cases they have not been more accurate than a simple moving average (Ahmed and Cook, 1979). While second generation predictor performance has not justified the substantial effort for historical data collection, 3rd generation predictors have been characterized by an intrinsic lag; both have resulted in prediction errors comparable to errors obtained via the corresponding UTCS predictor (Eldor, 1977).

Spectral analysis has resulted in acceptable prediction errors. The main disadvantage of this method is its inability to account for rapid unforeseen changes which are not reflected in the covariance matrix based on the characteristics of previous data (Nicholson and Swann, 1974).

Kalman filtering, one of the most advanced methods in modern control theory, is based on theory proposed by Kalman (1960). It may be applied to short term stationary or nonstationary stochastic phenomena. In traffic it has been applied successfully for demand prediction with great accuracy (Gazis and Knapp, 1971).

Most existing demand prediction methods make use of the time series of traffic flow data on the study link only. In this work an alternative method is developed for predicting 15-min volume during the day using not only the traffic flow on the study link but also the traffic flows on other links which may feed into the study link. Owing to its desirable characteristics Kalman filtering theory is employed for developing the prediction algorithms. The new algorithms are compared to UTCS-2, previously determined as the best of the most widely used algorithms (Stephanedes *et al.*, 1981) for this type of prediction.

UTCS-2

The second generation UTCS, UTCS-2, predicts the next-control-interval (on the order of 5–15 min) traffic volume at each detector location in real time based on the measurements from the same location only. The algorithm makes use of both smoothed historical traffic data and current traffic-volume measurements from the vehicle detector.

The UTCS-2 set of equations has been presented elsewhere (FHWA, 1973). The complexity of the solution, however, is usually not shown. By solving the difference equations of UTCS-2, it can be shown that UTCS-2 results in the following demand prediction equation (Stephanedes *et al.*, 1981):

$$\begin{aligned}\hat{v}_t = m_t + \gamma(m_{t-1} - \gamma f_{t-1}) + (1-\alpha)\sum_{s=0}^{t-1} \\ \times \alpha^s(f_{t-s-1} - m_{t-s-1}) + \gamma(1-\alpha)\sum_{s=0}^{t-2} \\ \times \alpha^s(f_{t-s-2} - m_{t-s-2})\end{aligned} \tag{1}$$

in which

$$m_t = a_0 + \sum_{i=1}^{M}[a_i \cos(2\pi it/N) + b_i \sin(2\pi it/N)] \tag{2}$$

and

$$\begin{aligned}\gamma = \Bigg\{(n-1)\sum_{s=1}^{n-1}\left[f_s - m_s - (1-\alpha)\sum_{p=0}^{s-1}\alpha^p(f_{s-p-1} - m_{s-p-1})\right] \\ \times\left[f_{s-1} - m_{s-1} - (1-\alpha)\sum_{p=0}^{s-2}\alpha^p(f_{s-p-2} - m_{s-p-2})\right]\Bigg\} \\ \div (n-2)\sum_{i=1}^{n}\left[f_s - m_s - (1-\alpha)\sum_{p=0}^{s-1}\alpha^p(f_{s-p-1} - m_{s-p-1})\right]^2\end{aligned} \tag{3}$$

where: $\hat{v}_t$ = predicted volume at time t; m_t = historical volume at time t; f_t = measured volume at time t; α = constant computed off-line from representative volume data of the location in question (e.g. for the UTCS system in Washington, D.C., α was 0.2); γ = smoothing coefficient (e.g. for the UTCS system in Washington, D.C., γ was 0.9); a_0, a_i, b_i = coefficients (computed off-line) of Fourier series approximation of historical traffic patterns for each measurement location; M = user input parameter determining the fidelity of Fourier series approximation, usually the result of a trade-off between Fourier series accuracy and storage space and computation effort (in general, for more rapidly varying functions, higher values of M should be used; M-values from 6 to 20 have been used in past applications); n = number of sample points of the representative data set; N = total number of time intervals in the representative data set (e.g. for 15-min intervals, the data for a 24-hr day will consist of 96 intervals).

For $k > 1$ the following prediction model was constructed based on UTCS-2:

$$\hat{v}_{t+k} = m_{t+k} + \gamma m_{t+k-1} - \gamma \hat{v}_{t+k-1} + (1-\alpha) \sum_{s=0}^{t+k-1} \alpha^s (f^*_{t+k-1-s} - m_{t+k-1-s}) \tag{4}$$

where

$$f^*_{t+k-1} = \begin{cases} f_{t+k-1-s} & \text{if } t+k-1-s \leqq t \\ \hat{v}_{t+k-1-s} & \text{if } t+k-1-s > t. \end{cases}$$

It is easily seen that eqn (4) can be rewritten as

$$\begin{aligned} \hat{v}_{t+k} = m_{t+k} + \gamma m_{t+k-1} - \gamma \hat{v}_{t+k-1} + (1-\alpha) \sum_{s=0}^{k-2} \\ \times \alpha^s (\hat{v}_{t+k-1-s} - m_{t+k-1-s} + (1-\alpha) \sum_{s=k-1}^{t+k-1} \\ \times \alpha^s (f_{t+k-1-s} - m_{t+k-1-s}). \end{aligned} \tag{5}$$

The UTCS-2 algorithms for predicting the next-control-interval (on the order of 5–15 min) traffic volume and the traffic volume at k ahead of time t as modified above are based on single-location traffic measurements. Both use the linear combination of residues (differences between traffic measurements and either historical data or smoothed traffic data) as the basic feature for prediction and require historical data as the reference.

Major UTCS-2 drawbacks are related to its high reliance on historical data: Traffic volume can vary substantially, depending on various external (with respect to algorithm) factors (e.g. weather conditions, special events, developments in other modes of transportation, and even the traffic-control change itself). UTCS-2 is not responsive to such changes. Because of this reliance on historical data, UTCS-2 is not readily transferable across systems and therefore may not be practical. A large data base is required for the historical data. This data base consumes computer storage space and must be updated periodically off-line. Furthermore, an analysis conducted early in the UTCS project in which "simulated" traffic data were used indicated that historical data were not always necessary to achieve good prediction (Stephanedes *et al.*, 1981).

Previous performance tests have indicated that UTCS-2 performs better than a number of existing algorithms. For example, it consistently performs better than UTCS-3, with a lower mean square and a lower mean absolute error (Kreer, 1975; Kreer, 1976; Stephanedes *et al.*, 1981), and a larger portion of small-magnitude errors (Kreer, 1975). It is not subject to an inherent time lag—as is UTCS-3—and, therefore, provides reasonably good values during a vehicle detector outage and is available as soon as detector operation is restored.

PREDICTION MODEL 1

Let $z(\tau)$ denote the traffic volume at time τ to be predicted. Although in practical applications this is a real number, for generality $z(\tau)$ will be assumed to be a vector of dimension m. For predicting $z(\tau + k)$, the traffic volume k time intervals ahead of time τ, the following linear prediction model is introduced:

$$^*z(\tau + k) = H_0(\tau)x(\tau) + H_1(\tau)x(\tau - 1) + \ldots + H_r(\tau)x(\tau - r) + w(\tau), \tag{6}$$

where $H_j(\tau)(j = 0, 1, 2, \ldots, r)$ is a parameter matrix of dimensions $m \times n$, $x(\tau)$ is a vector of dimension n composed of traffic characteristics judged to be useful for the prediction of $^*z(\tau + k)$ by the results of *a priori* investigation or by the experience of practicing traffic engineers and $w(\tau)$ is a noise vector of dimension m.

While $x(\tau)$ can be measured directly, the identification of $H_j(\tau)$ is carried out by making use of Kalman filtering theory as follows. Let $a_i^j(\tau)$ be the ith row of $H_j(\tau)$ and define an $m \cdot n \cdot (r+1)$ vector $h(\tau)$,

$$h(\tau) = (a_1^0(\tau), a_2^0(\tau), \ldots, a_m^0(\tau), a_1^1(\tau), a_2^1(\tau), \ldots, a_m^1(\tau), \ldots, a_1^r(\tau), a_2^r(\tau), \ldots, a_m^r(\tau))^T.$$

Also define an $m \cdot m \cdot n \cdot (r+1)$ matrix $\Lambda(\tau)$ and an m vector $z(\tau)$,

$$\Lambda(\tau) = \left[\begin{array}{ccc|ccc|c|ccc} x^T(\tau) & & \bigcirc & x^T(\tau-1) & & \bigcirc & \cdots & x^T(\tau-r) & & \bigcirc \\ & x^T(\tau) & & & x^T(\tau-1) & & \ddots & & x^T(\tau-r) & \\ \bigcirc & & x^T(\tau) & \bigcirc & & x^T(\tau-1) & \cdots & \bigcirc & & x^T(\tau-r) \end{array}\right], \tag{7}$$

$$z(\tau) = {}^*z(\tau+k).$$

Then eqn (6) may be rewritten as

$$z(\tau) = \Lambda(\tau)h(\tau) + w(\tau) \tag{8}$$

where $z(\tau)$, $\Lambda(\tau)$, $h(\tau)$ and $w(\tau)$ are regarded as the observation vector, transfer matrix, state vector and measurement error vector in the new system of parameter $h(\tau)$, respectively. If parameter $h(\tau)$ is stationary, the system equation for the new state vector $h(\tau)$ is

$$h(\tau) = h(\tau-1) + e(\tau-1) \tag{9}$$

where $e(\tau-1)$ is a noise vector of dimension $m \cdot n \cdot (r+1)$.

According to the Kalman filtering theory, the optimum state vector $\hat{h}(\tau)$—optimum in the sense that it minimizes the variance of an arbitrary weighted sum of estimation errors $\tilde{h}(\tau)$—estimated after $z(\tau)$ is observed, is given by

$$\hat{h}(\tau) = \hat{h}(\tau|\tau-1) + K(\tau)[z(\tau) - \hat{z}(\tau)], \tag{10}$$

where $K(\tau)$ denotes the so-called Kalman gain matrix of dimension $m \cdot n \cdot (r+1) \times m$ and $\hat{h}(\tau|\tau-1)$ and $\hat{z}(\tau)$ are the optimum estimates of $h(\tau)$ and $z(\tau)$ which can be derived before $z(\tau)$ is observed. If $h(\tau)$ is assumed stationary as expressed by eqn (9),

$$\hat{h}(\tau|\tau-1) = \hat{h}(\tau-1).$$

and

$$\hat{z}(\tau) = \Lambda(\tau)\hat{h}(\tau|\tau-1) = \Lambda(\tau)\hat{h}(\tau-1).$$

Thus eqn (10) is expressed as:

$$\hat{h}(\tau) = \hat{h}(\tau-1) + K(\tau)[z(\tau) - \Lambda(\tau)\hat{h}(\tau-1)]. \tag{11}$$

Since the value in the parentheses [.] describes the prediction error of $z(\tau)$, eqn (11) means that $\hat{h}(\tau-1)$ is updated by making use of this newest error information adjusted by the Kalman gain matrix $K(\tau)$.

The Kalman gain matrix is derived from the following successive operation equations:

$$K(\tau) = S(\tau)\Lambda^T(\tau)[R(\tau) + \Lambda(\tau)S(\tau)\Lambda^T(\tau)]^{-1} \tag{12}$$

$$S(\tau) = P(\tau-1) + Q(\tau-1) \tag{13}$$

$$P(\tau) = S(\tau) - K(\tau)\Lambda(\tau)S(\tau) \tag{14}$$

$$S(\tau_0) = D \tag{15}$$

where: $S(\tau)$ = covariance matrix of the estimation error $\tilde{h}(\tau|\tau-1)(=h(\tau)-\hat{h}(\tau|\tau-1))$; $P(\tau)$ = covariance matrix of the estimation error $\tilde{h}(\tau)$; $R(\tau)$ = covariance matrix of $w(\tau)$; $Q(\tau)$ = covariance matrix of $e(\tau)$; D = covariance matrix of $h(\tau_0)$; τ_0 = starting time.

Actual calculation follows the steps below, beginning at $\tau = \tau_0$.

Step 1. Let $\tau = \tau_0$ and $S(\tau_0) = D$ (eqn 15).

Step 2. Calculate $K(\tau)$ by eqn (12).

Step 3. Calculate $\hat{h}(\tau)$ by eqn (11) for $\tau \neq \tau_0$. When $\tau = \tau_0$, calculate $\hat{h}(\tau_0)$ by the following equation;

$$\begin{aligned}\hat{h}(\tau_0) = \hat{h}(\tau_0|\tau_0 - 1) + K(\tau_0)[z(\tau_0) - \Lambda(\tau_0)\\ \times \hat{h}(\tau_0|\tau_0 - 1)]\end{aligned}$$

where $\hat{h}(\tau_0|\tau_0 - 1)$ expresses an *a priori* estimate of $h(\tau_0)$ and is usually put as the mean of $h(\tau_0)$ designated by μ, i.e.

$$\hat{h}(\tau_0|\tau_0 - 1) = \mu.$$

From this fact $S(\tau_0)$ becomes equal to D as is described at Step 1.

Step 4. Calculate $P(\tau)$ by eqn (14).

Step 5. Calculate $S(\tau + 1)$ by eqn (13).

Step 6. Let $\tau = \tau + 1$ and return to Step 2. Since our original prediction model is given by eqn (6), it is evident that

$$\tau_0 = r.$$

If no value is available for $R(\tau)$, $Q(\tau)$ and D, it is customary to express them as diagonal matrices. Similarly, μ is set equal to the zero vector.

It should be noted that, if t is the present time, the newest possible observed data is $^*z(t)(=z(t-k))$. Hence the most updated parameter obtained using this observation is $\hat{h}(t-k)$. If $h(\tau)$ is stationary, parameter $\hat{h}(t-k)$ can be used for predicting $^*z(t+k)$ or $z(t)$, i.e.

$$^*\hat{z}(t+k) = \Lambda(t)\hat{h}(t-k), \tag{16}$$

where $^*\hat{z}(t+k)$ denotes the prediction of $^*z(t+k)$. However, according to eqns (7) and (8), $^*\hat{z}(t+k)$ must be of the form

$$^*\hat{z}(t+k) = \hat{z}(t) = \Lambda(t)\hat{h}(t),$$

which is different from eqn (16). Evidently it is impossible to obtain $\hat{h}(t)$ at time t since calculation of its value requires $^*z(t+k)$. Nevertheless, if the assumption of stationarity for $h(\tau)$ holds, $\hat{h}(t-k)$ can be used in place of $\hat{h}(t)$ in eqn (16).

In order to ensure the stationarity, it is necessary and sufficient to use the difference between the traffic volume on a day in question and that on the day a week before as variable $^*z(\tau)$ rather than the traffic volume on the day itself; a similar observation holds for the vector of traffic characteristics $x(\tau)$, i.e.

$$^*z(\tau) = {}^*Z(d,\tau) - {}^*Z(d-1,\tau) \tag{17}$$

$$x(\tau) = X(d,\tau) - X(d-1,\tau) \tag{18}$$

where: $^*Z(d, \tau)$ is the traffic volume at time τ on the dth day which is to be predicted; $X(d, \tau)$ is the traffic characteristic at time τ on the dth day which is used for the prediction of $^*Z(d, \tau)$; $^*Z(d-1, \tau)$, $X(d-1, \tau)$ are the traffic volume and traffic characteristic on the day a week before the dth day, which correspond to $^*Z(d, \tau)$ and $X(d, \tau)$, respectively.

With $^*\hat{z}(t+k)$, the predicted value of $^*z(t+k)$ known, $^*\hat{Z}(d, t+k)$, the prediction of $^*Z(d, t+k)$ is obtained from

$$^*\hat{Z}(d, t+k) = {}^*\hat{z}(t+k) + {}^*Z(d-1, t+k).$$

PREDICTION MODEL II

The similarity of the traffic flow pattern from day to day is taken into account in making up the second model. In this model the assumption of stationarity for parameter h with respect to time t is not required. The following prediction model is adopted:

$$^*Z(d, t+k) = H_0(d, t)X(d, t) + H_1(d, t)X(d, t-1)$$

$$+ \ldots\ldots + H_r(d, t)X(d, t-r) + w(d, t), \tag{19}$$

where: $H_j(d, t)$ = parameter matrix at time t on the dth day; $w(d, t)$ = noise vector at time t on the dth day.

As described in the previous section eqn (19) can be written in the form

$$^*Z(d, t+k) = \Lambda(d, t)h(d, t) + w(d, t) \tag{20}$$

where $\Lambda(d, t)$ and $h(d, t)$ correspond to $\Lambda(t)$ and $h(t)$ previously defined; eqn (20) is an observation equation for the new state vector $h(d, t)$. Since traffic flow fluctuates in a rather similar fashion from day to day, $h(d, t)$ may be assumed stationary with respect to day d and the system equation is

$$h(d, t) = h(d-1, t) + e(d-1, t) \tag{21}$$

where the $(d-1)$-th day is the day a week before the dth day.

For the new system $\hat{h}(d, t)$, the optimum estimate of $h(d, t)$ is represented by

$$\hat{h}(d, t) = \hat{h}(d-1, t) + K(d, t)[{}^*Z(d, t+k) - \Lambda(d, t)\hat{h}(d-1, t)]$$

where $K(d, t)$ denotes the Kalman gain matrix calculated through a number of steps as in Model I.

The prediction $^*\hat{Z}(d, t+k)$ is given by

$$^*\hat{Z}(d, t+k) = \Lambda(d, t)\hat{h}(d-1, t),$$

i.e. the optimum parameter estimate derived on the day a week before the day under analysis is used for the prediction of $^*Z(d, t+k)$.

The traffic volume difference between two consecutive weeks as expressed by eqns (17) and (18) could also be used as an input variable in place of the dth day traffic volume data, but this is not required in this case since stationarity has been assumed from the start.

APPLICATION

For testing the new prediction models data were collected from a network in Nagoya City, Japan. In this example, with four links in the network numbered as shown in Fig. 1, let $X_i(d, t)$ denote the traffic volume on the ith link and

$$^*Z(d, t) = X_1(d, t).$$

Since traffic data from three previous time intervals were used for prediction, the problem

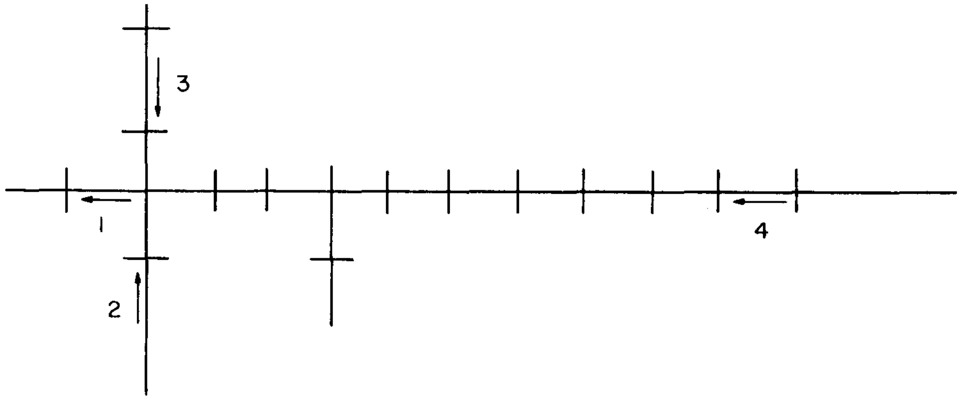

Fig. 1. Study network in Nagoya City.

examined is characterized by

$$m = 1,$$
$$n = 4,$$
$$r = 3.$$

The data are 15-min vehicle counts recorded at intervals of 5 min, the time unit in the models. Data obtained by an exponential smoothing technique were also used instead of raw data to examine possible improvements in prediction accuracy.

Three prediction error indices were computed: (a) mean relative error (ϵ_{mean}), which indicates the expected error as a fraction of the measurement, (b) root relative square error (ϵ_{rs}), which penalizes large prediction errors, and (c) maximum relative error, $\epsilon_{\max}$. These error indices are defined as follows:

$$\epsilon_{\text{mean}} = \frac{1}{N}\sum_t \left\{\frac{X_{\text{true}}(t) - X_{\text{pred}}(t)}{X_{\text{true}}(t)}\right\},$$

$$\epsilon_{rs} = \sqrt{\frac{1}{\Sigma_t X_{\text{true}}(t)}\sum_t \left\{\frac{X_{\text{true}}(t) - X_{\text{pred}}(t)}{X_{\text{true}}(t)}\right\}^2 X_{\text{true}}(t)},$$

$$\epsilon_{\max} = \max_t \left\{\frac{X_{\text{true}}(t) - X_{\text{pred}}(t)}{X_{\text{true}}(t)}\right\},$$

where: N = the number of samples (here, $N = 144$); $X_{\text{true}}(t)$ = measured value at time t; $X_{\text{pred}}(t)$ = predicted value at time t.

Four models were tested: UTCS-2 (modified here for making predictions k time intervals ahead of t, where $k > 1$), and new models $M1$, $M2$-a, and $M2$-b, where: $M1$ = prediction model I where the difference between the dth day and $(d-1)$-th day traffic data are used as the input data (see eqn 17 and eqn 18); $M2 - a$ = prediction model II where traffic data on the dth day are used as the input data; $M2 - b$ = prediction model II where data used are same as for $M1$.

The three errors, ϵ_{mean}, ϵ_{rs}, and $\epsilon_{\max}$ for the four models are presented in Tables 1 and 2 in two ways. The value for each relative error is given first; evidently, lower errors indicate better algorithm performance. In addition, each model is compared to modified UTCS-2, and the deviation of its error with respect to that of UTCS-2 is presented. A negative deviation implies that the model results in an error smaller than that of UTCS-2 and is, therefore, more desirable.

The following conclusions can be drawn from the test results and the relative per-

Table 1. Prediction errors relative to raw data for 15-min prediction

		ε_{mean}		ε_{rs}		ε_{max}	
k	Model	Value	Difference from UTCS-2 (%)	Value	Difference from UTCS-2 (%)	Value	Difference from UTCS-2 (%)
1	UTCS-2	0.1149	---	0.1357	---	0.4592	--
	M1	0.0488	-57.5	0.0609	-55.1	0.2132	-53.6
	M2-a	0.0712	-38.0	0.0868	-36.0	0.2291	-50.1
	M2-b	0.0810	-29.5	0.0963	-29.0	0.2720	-40.8
3	UTCS-2	0.1087	---	0.1284	---	0.3623	---
	M1	0.0749	-31.1	0.0919	-28.4	0.2717	-25.0
	M2-a	0.0760	-30.1	0.0937	-27.0	0.2225	-38.6
	M2-b	0.0867	-20.2	0.1025	-20.2	0.2425	-33.1
6	UTCS-2	0.1274	---	0.1459	---	0.3200	---
	M1	0.0601	-52.8	0.0813	-44.3	0.2821	-11.8
	M2-a	0.0818	-35.8	0.1015	-30.4	0.2519	-21.3
	M2-b	0.0853	-33.0	0.1068	-26.8	0.2921	- 8.7
9	UTCS	0.1523	---	0.1674	---	0.3216	---
	M1	0.0705	-53.7	0.0901	-46.2	0.2453	-23.7
	M2-a	0.0788	-48.3	0.1012	-39.5	0.2509	-22.0
	M2-b	0.0847	-44.4	0.1049	-37.3	0.2960	- 8.0

formance comparisons given in Tables 1 and 2:

(1) At all times all new prediction models perform substantially (up to 80%) better than UTCS-2.

(2) As a rule, prediction model $M1$, where the difference between the dth day and $(d-1)$-th day traffic data are used as input, outperforms prediction model $M2$. A possible reason is that, in $M2$ prediction on the day under study is accomplished using traffic data from the day a week before the study day.

(3) When prediction is performed vs smoothed—rather than raw—data, UTCS-2 performance worsens; by contrast, $M1$ achieves its best performance under these conditions.

Table 2. Prediction errors relative to smoothed data for 15-min prediction

		ε_{mean}		ε_{rs}		ε_{max}	
k	Model	Value	Difference from UTCS-2 (%)	Value	Difference from UTCS-2 (%)	Value	Difference from UTCS-2 (%)
1	UTCS-2	0.1269	---	0.1439	---	0.3832	---
	M1	0.0249	-80.4	0.0341	-76.3	0.1620	-57.7
	M2-a	0.0626	-50.7	0.0788	-45.2	0.2385	-37.8
	M2-b	0.0816	-35.7	0.0979	-32.0	0.2818	-26.5
3	UTCS-2	0.1261	---	0.1425	---	0.3698	---
	M1	0.0538	-57.3	0.0736	-48.4	0.2432	-34.2
	M2-a	0.0673	-46.6	0.0827	-42.0	0.2235	-39.6
	M2-b	0.0858	-32.0	0.0998	-30.0	0.2521	-31.8
6	UTCS-2	0.1503	---	0.1665	---	0.3769	---
	M1	0.0480	-68.1	0.0740	-55.6	0.2775	-26.4
	M2-a	0.0730	-51.4	0.0912	-45.2	0.2554	-32.2
	M2-b	0.0852	-43.3	0.1037	-37.7	0.2563	-32.0
9	UTCS-2	0.1736	---	0.1853	---	0.3463	---
	M1	0.0566	-67.4	0.0738	-60.2	0.2294	-33.8
	M2-a	0.0747	-57.0	0.0929	-49.9	0.2189	-36.8
	M2-b	0.0870	-49.9	0.1059	-42.8	0.2567	-25.9

(4) By increasing k (the number of 5-min time intervals ahead of current time for which prediction is performed) from one to nine in the proposed models, performance is not significantly affected. Such robust performance is highly desirable for longer-term prediction.

A sample of plots of measured 15-min. traffic volume data (smoothed value) versus predicted values (prediction of $k = 6$ or 30 min. ahead) is shown in Figs. 2–5. The conclusions cited above can also be drawn from these figures.

DISCUSSION AND CONCLUSIONS

The test results suggest that the proposed models perform better than UTCS-2 in 5-min prediction. When the difference between traffic data on the day under study and that on the same day one week before is used as input, performance improves, probably since

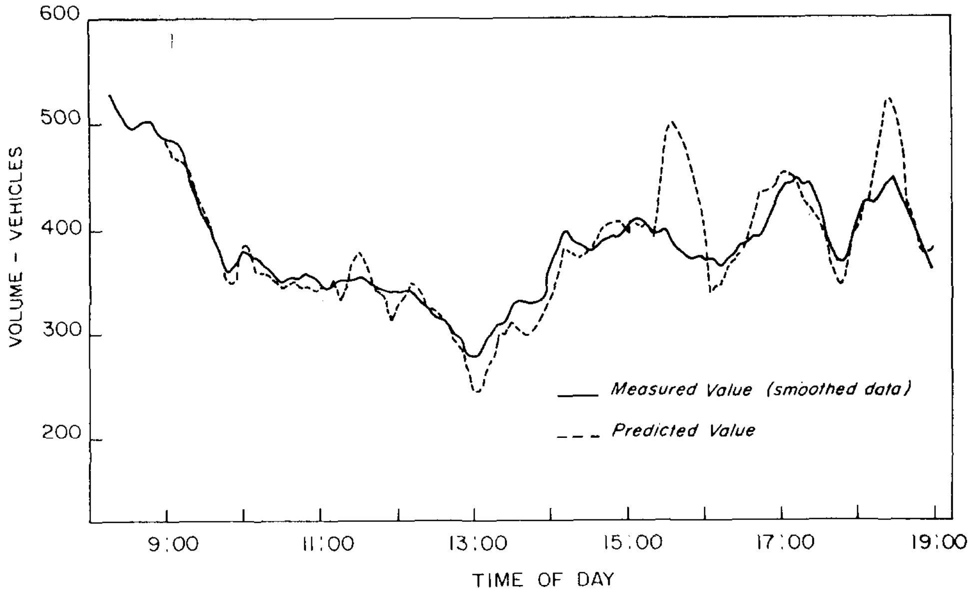

Fig. 2. Measured values and predicted values obtained by $M1(k = 6)$.

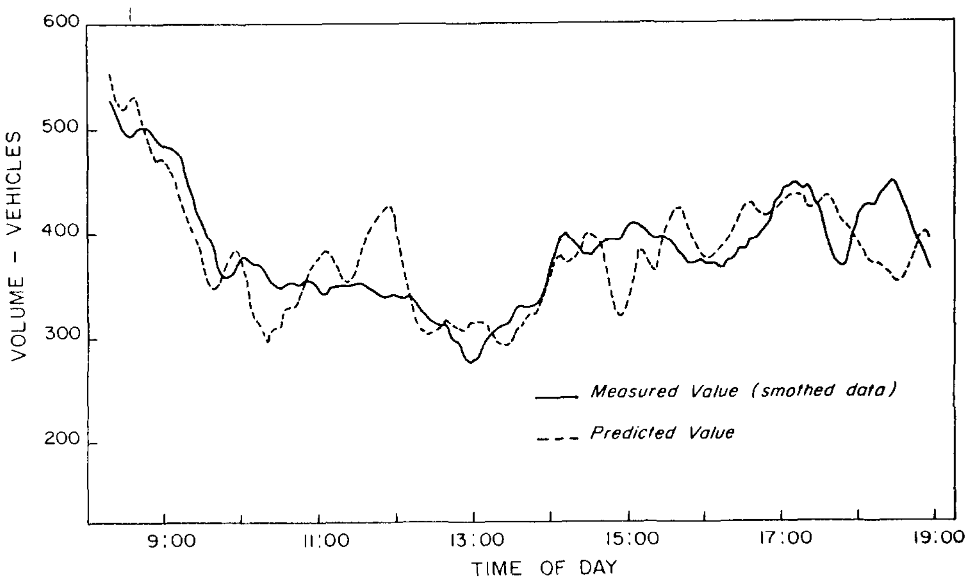

Fig. 3. Measured values and predicted values obtained by $M2 - a(k = 6)$.

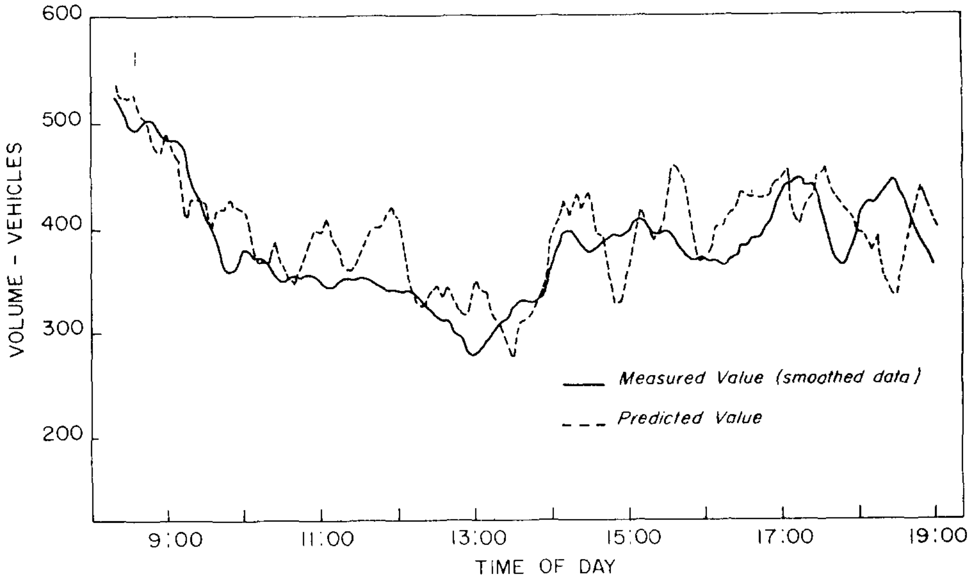

Fig. 4. Measured values and predicted values obtained by $M2-b(k=6)$.

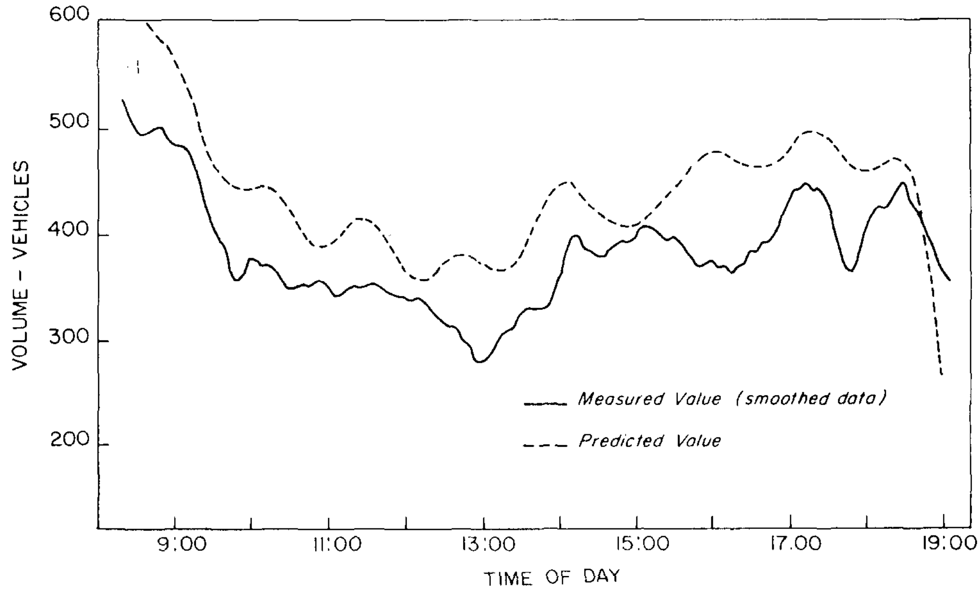

Fig. 5. Measured values and predicted values obtained by UTCS-2 ($k=6$).

stationarity in the parameters is better ensured. The tests also indicate a robustness in model performance as the prediction horizon increases, thus suggesting desirable behavior for longer-term prediction. Additional tests have indicated no appreciable change in performance for $r > 3$. These properties, however, should be more deeply analyzed before final conclusions are drawn.

Although comparison tests were performed against UTCS-2 only, previous work (Stephanedes *et al.*, 1981) has indicated the superiority of UTCS-2 over a number of algorithms. The superiority of the new models over these algorithms could also be inferred.

Four are the major advantages of the new models over the existing. First, they can be used for traffic volume prediction using data from links adjacent to the one under study. Second, they employ future data estimates to continuously update the error for better prediction. Further, the models can easily and quickly compute desired traffic characteristics, especially when $m = 1$. In addition, they can be employed for predicting traffic volume k time intervals ahead, for any k.

The new models are promising for practical applications. While prediction error has improved, computation time required is reasonable. For example, 4500 predictions can be obtained in approximately 10 sec. for model $M1$ and 25 sec. for $M2$ on a HITAC M-200H at the Computer Center of the University of Tokyo. For making 10 predictions in real time, as would often be the case, approximately 5 sec. would be required for $M1$ on a PDP-11 minicomputer (or less than 2 min. on an APPLE-II microcomputer) acting as a master or local controller. Storage requirements for the proposed models are not excessive as two-week data would be adequate for algorithm implementation. Finally, these models could be used for predicting a variety of traffic characteristics such as time occupancy and traffic density.

Acknowledgements—This study was partially supported by a Grant-in-Aid for Scientific Research funded by the Ministry of Education of Japan. The U.S. Department of Transportation has provided partial support through project No. DOT-RC-82008. The authors would also like to thank P. Michalopoulos for his cooperation and valuable suggestions.

REFERENCES

FHWA (1973) Urban traffic control system and bus priority system traffic adaptive network signal timing program: software description. Federal Highway Administration, U.S. Dept. of Transportation, Washington, D.C.

SRI (1977) Improved control logic for use with computer-controlled traffic. Stanford Research Institute, NCHRP Project 3–18(1), Final Rept., Menlo Park, CA.

CMT (1976) Improved operation of urban transportation systems: volume 3—the development and evaluation of a real-time computerized traffic control strategy. Canada Ministry of Transport, Toronto, Ontario.

Liberman E. B. *et al.* (1974) Variable cycle signal timing program: volume 4—prediction algorithms, software and hardware requirements, and logical flow diagrams. Federal Highway Administration, U.S. Dept. of Transportation, NTIS: PB 241 720, Washington, D.C.

Mengert P., Brown P. and Yuan L. (1979) Prediction algorithms for urban traffic control. Transportation Systems Center, U.S. Dept. of Transportation, Internal Project Memorandum, Cambridge, MA.

Stephanedes Y. J., Michalopoulos P. G. and Plum, R. A. (1981) Improved estimation of traffic flow for real time control. *Transpn Res. Rec.* No. 795, 28–39.

Eldor M. (1977) Demand predictors for computerized freeway control systems. Proceedings of the 7th International Symposium on Transportation and Traffic Theory, Kyoto, Japan, 14–17 August 1977, 341–358.

Gafarian A. V., Paul J. and Ward T. L. (1977) Discrete time series models of a freeway density process. Proceedings of the 7th International Symposium on Transportation and Traffic Theory, Kyoto, Japan, 14–17 August 1977, 387–411.

Ahmed S. A. and Cook A. R. (1979) Analysis of freeway traffic time-series data by using Box-Jenkins techniques. *Transpn Res. Rec.* No. 733, 1–9.

Nicholson H. and Swann C. D. (1974) The prediction of traffic flow volumes based on spectral analysis. *Transpn Res.* **8**, 533–538.

Nahi N. E. (1973) Freeway traffic data processing. Proceedings of the IEEE 61 No. 5, 537–541.

Chang M. F. and Gazis D. C. (1975) Traffic density estimation with consideration of lane changing. *Transpn Sci.* **9**(4), 308–320.

Gazis D. C. and Knapp C. H. (1971) On-line estimation of traffic densities from time-series of flow and speed data. *Transpn Sci.* **5**, 283–301.

Baras J. S. and Levine W. S. (1977) Estimation of traffic flow parameters in urban traffic networks. Proceedings of the IEEE Annual Conference on Decision and Control, San Francisco, 428–433.

Baras J. S., Levine W. S. and Lin T. L. (1978) Discrete time point processes in urban traffic queue estimation. Proceedings of the IEEE Conference on Decision and Control, San Diego, 1025–1031.

Baras J. S. (1979) *et al.* Advanced filtering and prediction software for urban traffic control systems. Transportation Studies Center, University of Maryland, Draft Final Rept., College Park, MD.

Kalman R. E. (1960) A new approach to linear filtering and prediction problems, *J. Basic Engng* **82D** (1), 35–45.

Kreer J. B. (1975) A comparison of predictor algorithms for computerized traffic control systems. *Traffic Engng* **45**(4) 51–56.

Kreer J. B. (1976) Factors affecting the relative performance of traffic responsive and time-of-day traffic signal control. *Transpn Res.* **10**, 75–81.