Prediction of Traffic Flow by an Adaptive Prediction System

JIAN LU

In a dynamic (real-time) traffic control system, the accuracy of the prediction of traffic characteristics such as flow, speed, and headway is one of the key factors affecting the performance of the system. Because the traffic characteristics can be described by stochastic processes, nonlinear and time-variate types of prediction models could be more adequate than linear or time-invariate prediction models. A traffic control system model for traffic flow is described, and the importance of the accuracy of the prediction model is emphasized. Then, the concept of adaptive prediction of traffic flow is introduced, and its mathematical derivation and the least-mean-square algorithm are described. As an experiment to validate the adaptive prediction system, a sine function is used to simulate traffic flow as input to the adaptive prediction system. Finally, the adaptive prediction system is applied to actual traffic flow data collected from a highway network. The predicted traffic flow is then compared with the real traffic flow. The performance of the model as to its dynamic response to a step function, convergence of the adaptive prediction system, and related matters are also discussed.

Since World War II, the scope of research involving vehicular traffic has widened exponentially. The dramatic development of the traffic system network has opened the door of traffic engineering research to modern system theory, control theory, and information theory. Classic traffic engineering theories are now combined with other applied theories such as optimum control (*1*) and kinetics (*2*).

As understood by researchers and users of traffic systems, some traffic characteristics such as flow, speed, and headway affect the performance of a traffic system. Certainly, the performance of a traffic system can be improved by measuring and controlling these characteristics through modern detection and control methods. Macroscopically, a traffic system equipped with a detection system and a control system could be mathematically abstracted into a dynamic multivariable system. The actual behavior of an abstracted dynamic traffic control system depends largely on the control rule and also on the accuracy of the detection system and the prediction of traffic characteristics.

Consider the real-time freeway traffic control system shown in Figure 1. In this system, $S_1(k)$, $S_2(k)$, . . ., $S_j(k)$ are the traffic characteristics measured at Time Step k by sensors. To optimize the traffic characteristics of the freeway network, the states of traffic characteristics at Step $k + 1$, or $S_1(k + 1)$, $S_2(k + 1)$, . . ., $S_j(k + 1)$, must be predicted. According to the feedback signal predicted, the real-time control system can adjust the timing of all related traffic lights or some type of control equipment so that the traffic variables in the freeway system are at the optimal state. In this case, the predicting algorithm should be an important part in the control system. The performance of the freeway network depends not only on the control equipment itself but also on the accuracy of the feedback signal predicted. The importance of the accuracy of traffic prediction to the performance of the traffic control system has been emphasized by Stephanedes et al. (*3*) and by Kreer (*4*).

Traditionally, linear regression and extrapolation models have been used as predicting models. However, these kinds of predicting models are nonparameter estimation models, and are usually used to predict the processes with obvious trends. For stochastic processes such as traffic flow, speed, and headway for a limited time interval, say from 5 to 9 a.m., the traffic characteristics of a system change with obvious trends. In a macroscopic time interval, 1 day, 1 week, or 1 month, traffic characteristics are more appropriately described as the approximation of stochastic processes without the obvious constant trends. In this case, the linear regression and extrapolation models have limitations for the purpose of prediction.

Traffic characteristics are random variables. Strictly speaking, these types of variables are not stationary in terms of mean value, variance, or covariance in a relatively long time period. To apply modern control theories and signal processing theories (e.g., stochastic system control and signal prediction theories) to traffic problems, the traffic variables can be treated as asymptomatically stationary processes with certain constraining conditions. These constraining conditions could be the time frame considered, the sampling rate, given time periods, locations, and environmental conditions. One of the reasons to assume that the traffic variables are asymptomatically stationary with certain constraining conditions is that most of the prediction models could be guaranteed to converge if the signal or noise is stationary. Examples of these prediction models are the Kalman filtering model (*5*), the Wiener filtering model (*6*), the autoregressive-moving average (ARMA) filtering model (*7*), and the adaptive filtering model (*8*).

In the last two decades, several important parameter estimation models have been applied to traffic flow prediction (*9*), intersection traffic distribution prediction (*10*), origin-destination matrix estimation (*11,12*), freeway incident detection (*13*), and traffic density estimation (*14*). Several typical models have been used to predict traffic flow. Among these models are the Kalman filtering model (*9*), the time series prediction model (*15*), the spectral analysis model (*16*), and the urban traffic control system (UTCS) (*17*).

Engineering Research and Development Bureau, New York State Department of Transportation, 1220 Washington Ave., Albany, N.Y. 12232.

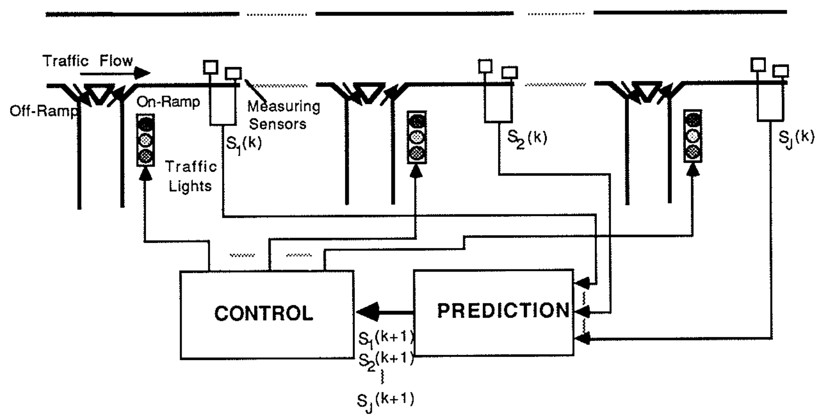

FIGURE 1 A freeway traffic control system.

Recent studies have used an adaptive filtering model to predict traffic flow. An adaptive traffic prediction system or model can be considered as a dynamic (real time) parameter estimation model; that is, the traffic characteristics predicted at Time Step k are a function of past traffic characteristics at Time Steps $k - 1, k - 2, \ldots, k - M$, where $M < k$, and M and k are positive integers, and of a set of parameters estimated by the adaptive prediction model.

Mathematically, the objective function of the adaptive prediction model—minimizing the resulting mean square error—might be similar to that of the Kalman filtering, time-series, and UTCS prediction models. However, the adaptive prediction model uses a simplified least-mean-square (LMS) algorithm to search for the optimal filter weights or states. This difference means that the dynamic response of the adaptive prediction model could be better and the data storage space could be less than the earlier prediction models. Consequently, the adaptive model could have more potential to be used in an on-line traffic flow prediction and control system. Intuitively speaking, if an adaptive prediction model is viewed as a system, then this system is one whose structure is adjustable in such a way that its performance improves through contact with its environment.

BASIC PRINCIPLE OF ADAPTIVE PREDICTION SYSTEM

Figure 2 shows an adaptive prediction system. In this system, Z^{-1} is a one-step delay factor, and Z^{-S} is an s-step delay factor (where s is a positive integer). Mathematically, $q(k)Z^{-1} = q(k - 1)$, and $q(k)Z^{-S} = q(k - s)$. As can be seen from Figure 2, the core of the system is the adaptive processor, the structure of which is shown in Figure 3. In the adaptive processor, all of the parameters (weights) at Step k are adjustable. The error of prediction $e(k)$ controls the adjustment of the system. From Figures 2 and 3, the following equation can be derived:

$$\hat{q}(k) = \sum_{j=0}^{N} W_{jk} q(k - s - j) \qquad k = s + 1, s + 2, \ldots \tag{1}$$

Equation 1 indicates that $\hat{q}(k)$ is a linear weighted combination of $q(k - s), q(k - s - 1), \ldots, q(k - s - N)$. The weights are $W_{0k}, W_{1k}, \ldots, W_{Nk}$, and the Index k denotes the time step. If $\hat{q}(k)$ is used to predict $q(k)$, then the error of prediction at Step k is

$$e(k) = q(k) - \hat{q}(k) \qquad k = s + 1, s + 2, \ldots \tag{2}$$

The purpose of using an adaptive processor is to adjust the weights at each Step k so that the mean square error $E[e^2(k)]$ is minimized. The vectors $\mathbf{W}_k$ and $\mathbf{Q}_{k-s}$ are defined as follows:

$$\mathbf{W}_k = (W_{0k}, W_{1k}, \ldots, W_{Nk})^T$$

$$\mathbf{Q}_{k-s} = [q(k - s), q(k - s - 1), \ldots, q(k - s - N)]^T$$

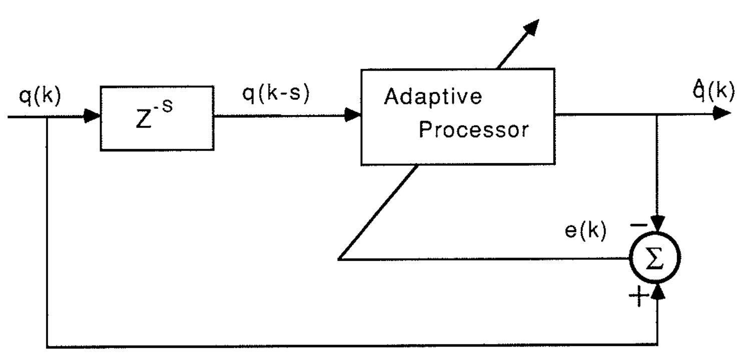

FIGURE 2 Adaptive prediction system.

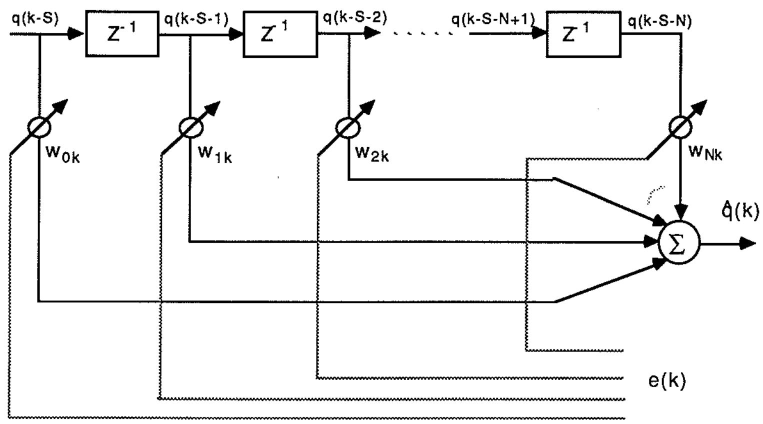

FIGURE 3 Adaptive processor [W_{ik} are adjusted by $e(k)$, i = 0 to N].

With these definitions, Equation 1 can be expressed using vector notation as follows:

$$\hat{q}(k) = \mathbf{Q}_{k-s}^{T}\mathbf{W}_k = \mathbf{W}_k^{T}\mathbf{Q}_{k-s} \tag{3}$$

Equation 4 describes how the adaptive processor adapts, that is, how the vector $\mathbf{W}_k$ is adjusted as the Time Step Index k changes. From Equations 2 and 3,

$$e(k) = q(k) - \mathbf{W}_k^{T}\mathbf{Q}_{k-s} = q(k) - \mathbf{Q}_{k-s}^{T}\mathbf{W}_k \tag{4}$$

By squaring Equation 4, the instantaneous squared error can be obtained as follows:

$$\begin{aligned} e^2(k) &= [q(k) - \mathbf{W}_k^{T}\mathbf{Q}_{k-s}][q(k - \mathbf{Q}_{k-s}^{T}\mathbf{W}_k] \\ &= q^2(k) + \mathbf{W}_k^{T}\mathbf{Q}_{k-s}\mathbf{Q}_{k-s}^{T}\mathbf{W}_k - 2q(k)\mathbf{Q}_{k-s}^{T}\mathbf{W}_k \end{aligned} \tag{5}$$

In order to find the expected value of Equation 5 over k, $e(k)$ and $q(k)$ are assumed to be statistically stationary. This assumption can usually be approximately satisfied for the particular traffic characteristics. Then the expectation of $e^2(k)$ is

$$\begin{aligned} E[e^2(k)] &= E[q^2(k)] \\ &\quad + \mathbf{W}_k^{T}E[\mathbf{Q}_{k-s}\mathbf{Q}_{k-s}^{T}]\mathbf{W}_k - 2E[q(k)\mathbf{Q}_{k-s}^{T}]\mathbf{W}_k \end{aligned} \tag{6}$$

Let $\mathbf{R}$ be defined as the square matrix

$$\mathbf{R} = E[\mathbf{Q}_{k-s}\mathbf{Q}_{k-s}^{T}] \tag{7}$$

Therefore, $\mathbf{R}$ is the correlation matrix of $q(k - s)$ with dimension $N \times N$. Let $\mathbf{P}$ be defined as the column vector

$$\begin{aligned} \mathbf{P} &= E[q(k)\mathbf{Q}_{k-s}^{T}] \\ &= E[q(k)q(k - s), q(k)q(k - s - 1), \ldots, \\ &\quad q(k)q(k - s - N)]^{T} \end{aligned} \tag{8}$$

This vector is the set of autocorrelation of $q(k)$. Therefore, $\mathbf{R}$ and $\mathbf{P}$ are the second-order statistics of the random variable $q(k - s)$ at Step k. By the definitions of $\mathbf{R}$ and $\mathbf{P}$, Equation 6 can be expressed as,

$$E[e^2(k)] = E[q^2(k)] + \mathbf{W}_k^{T}\mathbf{R}\mathbf{W}_k - 2\mathbf{P}^{T}\mathbf{W}_k \tag{9}$$

According to the assumption that $q(k)$ is statistically stationary, $\mathbf{R}$ and $\mathbf{P}$ are a constant matrix and vector, respectively. In this case, $E[e^2(k)]$ is a quadratic function of the weight vector $\mathbf{W}_k$. If the adaptive processor has the ability of self-study to seek the minimum $E[e^2(k)]$ by adjusting $\mathbf{W}_k$, and $E[e^2(k)]$ tends to be minimal when $\mathbf{W}_k$ tends to be optimal solution $\mathbf{W}_k^*$, then the prediction of the processor will be optimal. The question is how to find the optimal solution $\mathbf{W}_k^*$ of $\mathbf{W}_k$ so that $E[e^2(k)]$ is minimized at each Step k. This question can be solved by the gradient method. The gradient of the mean square error $E[e^2(k)]$ can be expressed by

$$\nabla_k = 2\mathbf{R}\mathbf{W}_k - 2\mathbf{P}$$

In order to obtain the optimal solution $\mathbf{W}_k^*$ so that $E[e^2(k)]$ is minimized, let

$$\nabla_k = 0 = 2\mathbf{R}\mathbf{W}_k^* - 2\mathbf{P}$$

or

$$\mathbf{W}_k^* = \mathbf{R}^{-1}\mathbf{P} \tag{10}$$

Equation 10 is the optimal solution for $\mathbf{W}_k$. By substituting Equation 10 into Equation 9 and noting that the correlation matrix is symmetrical, then

$$\begin{aligned} E[e^2(k)]_{\min} &= E[q^2(k)] + [\mathbf{R}^{-1}\mathbf{P}]^{T}\mathbf{R}\mathbf{R}^{-1}\mathbf{P} - 2\mathbf{P}^{T}\mathbf{R}^{-1}\mathbf{P} \\ &= E[q^2(k)] + \mathbf{P}^{T}\mathbf{R}^{-1}\mathbf{P} - 2\mathbf{P}^{T}\mathbf{R}^{-1}\mathbf{P} \\ &= E[q^2(k)] - \mathbf{P}^{T}\mathbf{R}^{-1}\mathbf{P} \end{aligned} \tag{11}$$

Although Equation 10 is the optimal solution of $\mathbf{W}_k$, in a practical sense $\mathbf{W}_k^*$ is not estimated by Equation 10. In the next section, the algorithm for estimating $\mathbf{W}_k^*$ is discussed.

LMS ALGORITHM

Recall in Equation 10,

$$\nabla_k = 2\mathbf{R}\mathbf{W}_k - 2\mathbf{P}$$

or

$$\mathbf{W}_k^* = \mathbf{R}^{-1}\mathbf{P}$$

Combining these two equations,

$$\mathbf{W}_k^* = \mathbf{W}_k - 0.5\mathbf{R}^{-1}\boldsymbol{\nabla}_k \quad (12)$$

Equation 12 can be changed into an adaptive algorithm as follows:

$$\mathbf{W}_{k+1} = \mathbf{W}_k - 0.5\mathbf{R}^{-1}\boldsymbol{\nabla}_k \quad (13)$$

If the vector of weight ($\mathbf{W}_k$) is adjusted in the direction of the gradient at each Step k, and a constant μ ($0 < \mu < 1$) is defined, then Equation 13 can be simplified as follows:

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \mu\boldsymbol{\nabla}_k \quad (14)$$

where μ regulates the step size (from k to $k + 1$) and has dimensions of reciprocal signal power.

In order to develop the LMS algorithm, $e^2(k)$ itself can be taken as an estimate of $E[e^2(k)]$; then the estimate of the gradient $\boldsymbol{\nabla}_k$ can be expressed by

$$\hat{\boldsymbol{\nabla}} = \begin{bmatrix} \frac{\partial e^2(k)}{\partial W_{0k}} \\ - \\ - \\ - \\ \frac{\partial e^2(k)}{\partial W_{Nk}} \end{bmatrix} = 2e(k) \begin{bmatrix} \frac{\partial e(k)}{\partial W_{0k}} \\ - \\ - \\ - \\ \frac{\partial e(k)}{\partial W_{Nk}} \end{bmatrix}$$

$$= -2e(k)\mathbf{Q}_{k-s} \quad (15)$$

With this simple estimate of the gradient, the LMS algorithm can be specified by Equations 14 and 15:

$$\begin{aligned} \mathbf{W}_{k+1} &= \mathbf{W}_k - \mu\boldsymbol{\nabla}_k \\ &= \mathbf{W}_k + 2\mu e(k)\mathbf{Q}_{k-s} \end{aligned} \quad (16)$$

Thus, Equations 3 and 16 constitute the adaptive prediction model. Equation 16 indicates that the LMS algorithm can be implemented in a practical system without squaring, averaging, or differentiation and is elegant in its simplicity and efficiency.

STABILITY AND CONVERGENCE

The stability and response performance need to be evaluated for a system model. In the adaptive prediction system, the stability and response performance depend mainly on the step size (μ) and the order (N) of the adaptive processor. In order to simplify the discussion for the response performance of the adaptive prediction system, only the number (CS) of convergence steps is of concern. CS is defined as follows. If $q(k)$, a step function, changes from a constant level to another constant level, then CS is the number of steps for $\hat{q}(k)$ to converge to $q(k)$ for a given criterion (C) after $q(k)$ changes. That is,

$$q(k) = A \qquad \text{if } k < L$$

or

$$q(k) = B \qquad \text{if } k \geq L$$

where $A \neq B$ and L is a positive integer.

CS is defined by

$$|\hat{q}(k) - q(k)| \leq \mathrm{C} \qquad \text{if } k \geq L + \mathrm{CS} \quad (17)$$

where C is a given criterion and CS is the minimum number of steps for Equation 17 to exist.

Conceptually, small values for μ and N result in good stability but slow convergence speed or large CS. This means that there is a trade-off between stability and response performance. In this research effort, another parameter (AL1) was defined as follows:

$$\mathrm{AL1} = \frac{1}{2\mu}$$

For this discussion, the average vehicle speed is defined as the mean speed of n vehicles at Time Step k, where n is the number of vehicles in a unit length of highway section. Therefore, the average vehicle speed is also a spatial average speed. If $q(k)$ is considered the average vehicle speed and is a step function, then

$$q(k) = 30 \text{ (mph)} \qquad \text{if } k < 20$$

or

$$q(k) = 50 \text{ (mph)} \qquad \text{if } k \geq 20 \quad (18)$$

Figure 4 shows the results of observed average speed versus predicted average speed for $N = 10$, $\mathrm{AL1} = 125$, $S = 1$, and $C = 0.1$. Because the average speed changes from 30 to 50 mph after several steps, there is a significant difference between $q(k)$ and $\hat{q}(k)$. The predicted average speed $\hat{q}(k)$ returns to almost the same as the observed average speed $q(k)$ after eight steps.

A step function is usually used as a reference input in the performance evaluation of a system because of its simplicity. To check the effects of AL1 ($1/\mu$) and N on the stability and convergence step of the adaptive prediction system, the step function (Equation 18) is considered as the reference input. Figure 5a shows the curve of AL1 versus Convergence Step CS. As AL1 gets larger (or μ gets smaller), the adaptive prediction system takes more steps to converge. Therefore, CS becomes larger. Figure 5b shows the curve of N versus Convergence Step CS. As N gets larger, the adaptive prediction system takes fewer steps to converge. Therefore, CS becomes smaller.

As previously discussed, the stability of the system depends mainly on AL1 and N. Then, in the plane of (AL1, N) a zone should exist in which the adaptive prediction system should be stable. The predicted value would then converge to the observed value for a given criterion C. Otherwise, the system

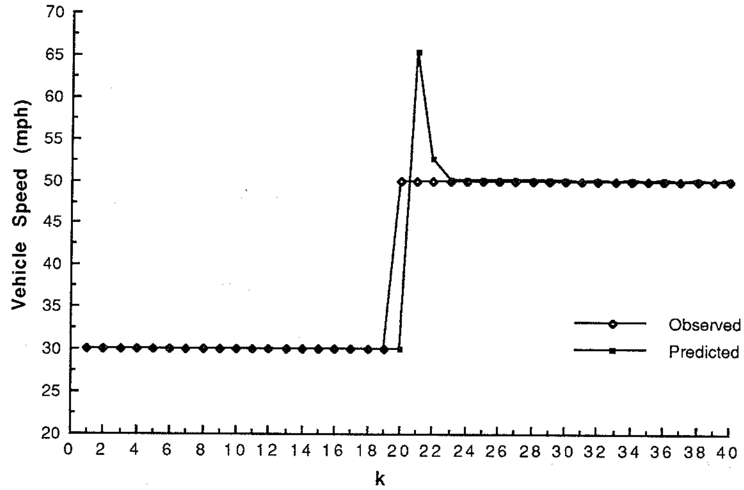

FIGURE 4 Response of adaptive prediction system [$q(k)$-step function].

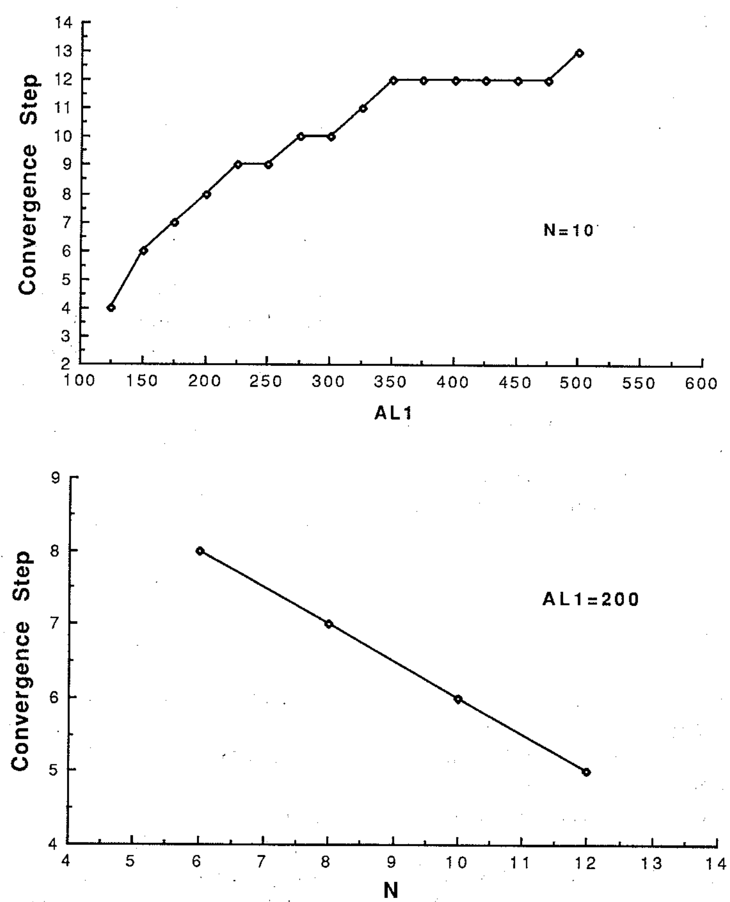

FIGURE 5 The effect of step size on the convergence step (AL1 = $1/2\mu$) (*top*) and the effect of N on convergence step (*bottom*).

would be unstable. Figure 6 shows the stable and unstable zones. The reference input is the same as for Equation 18. AL1 changes from 75 to 500, and N changes from 6 to 30. If (AL1, N) values belong to the area above the straight line, the system is stable; otherwise it is unstable.

EXPERIMENTS WITH THE ADAPTIVE PREDICTION SYSTEM

Two types of experiments were conducted to evaluate the performance of the adaptive prediction system. In the first experiment, a sine function is used to simulate the traffic flow $[q(k)]$:

$$q(k) = A \sin (kT) + B \quad \text{(vehicles/hour, } k = 0, 1, 2, \ldots) \tag{19}$$

where A and B are positive constants and T is the sample interval. The experimental results are shown in Figure 7, with AL1 = 125, N = 10, and S = 1. The predicted traffic flow $[\hat{q}(k)]$ follows the true (observed) traffic flow $[q(k)]$. The error of prediction is shown in Figure 8. For this kind of deterministic traffic flow, the adaptive prediction system has the ability to predict precisely the future characteristics of the traffic flow by understanding the past process of $q(k)$. This ability could be the result of the cyclic nature of the sine function input.

In the second experiment, real traffic data were collected from northbound lanes of the Mopac Highway and West 35th Street intersection in Austin, Texas. The data collection period covered three consecutive days with a 1-hr time interval between data sampling. The unit of the traffic flow data is vehicles per hour. Because the experiment was conducted to test how well the adaptive prediction system worked and because predicting traffic flow by historical traffic data is more difficult with a large sampling interval than with a small sampling interval, a large sampling interval (1 hr) was used in the experiment instead of a small sampling interval (3 to 5 min). Although the past characteristics of traffic flow were observed, precisely predicting the future values of the traffic flow was impossible. This result was different from that of the first experiment. The adaptive prediction system can figure out the statistical characteristics of the traffic flow using the adaptive processor and the past values to optimally predict the future traffic flow. Statistically, the adaptive prediction system's performance is optimal. Figure 9 shows the curves of observed traffic flow and predicted traffic flow with the given structure (AL1 = 10^8, N = 23, and S = 1). Figure 10 shows the same prediction

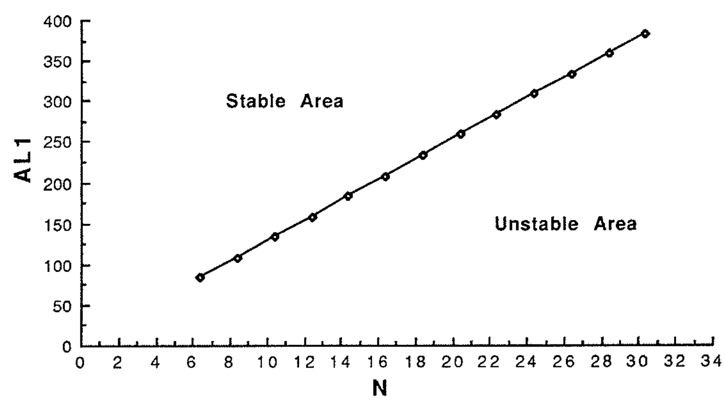

FIGURE 6 Stable and unstable zones in (AL1, *N*) plane.

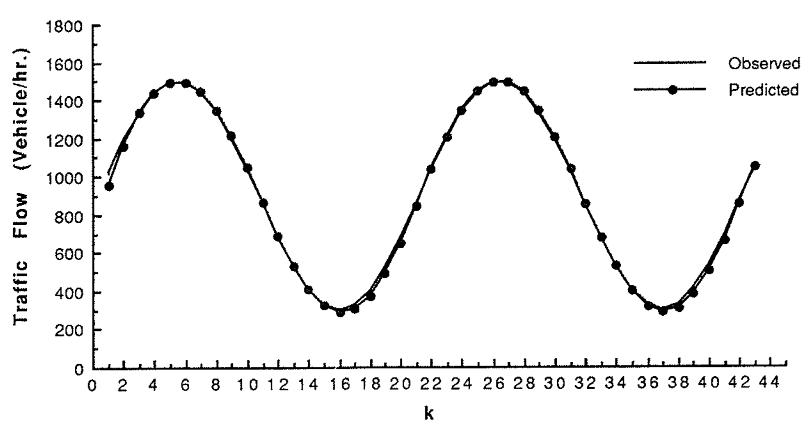

FIGURE 7 Prediction of traffic flow [$q(k) = A \sin (kT) + B$, T-sample interval].

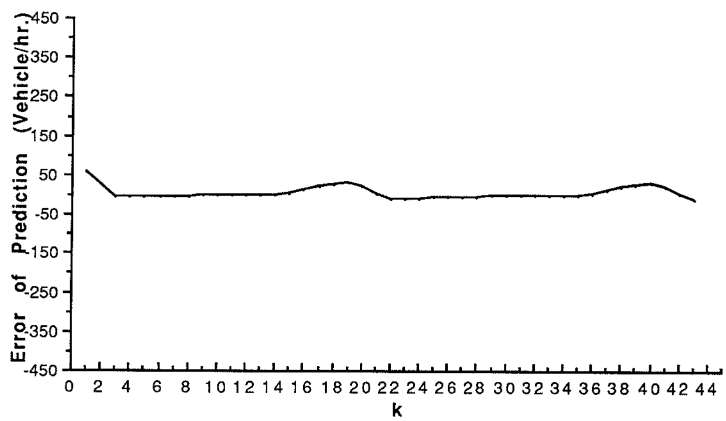

FIGURE 8 Error of prediction [$q(k) = A \sin (kT) + B$].

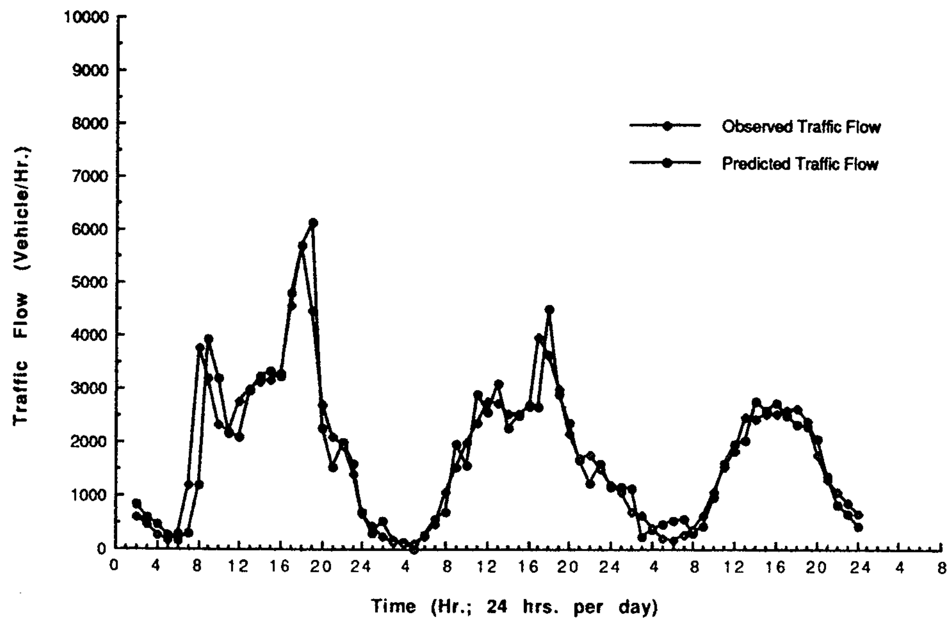

FIGURE 9 Prediction of traffic flow (traffic flow data collected from Mopac Highway, Austin, Texas) with $N = 23$, AL1 = 10^8.

process but with a different model structure (AL1 $= 9 \times 10^7$, $N = 6$, and $S = 1$). The figures indicate that the model structure (AL1 and N) affects the prediction error and performance of the prediction system.

CONCLUSIONS

The adaptive prediction system can be used as a real time predictor. The performance of the predictor depends on the characteristics of the traffic variables and on the structure of the predictors AL1 and N. In the process of choosing AL1 and N, consideration should be given to the stability and convergence step. For the prediction of a specific traffic variable, an adequate number of tests should be run to obtain the optimal AL1 and N values.

As discussed, the LMS algorithm is so simple that the adaptive prediction system can be conducted by hardware. The control system would then have better performance (less delay and faster response).

An adaptive prediction system is only one of the applications for the adaptive processor. The adaptive processor can be used widely in other traffic control areas such as filtering, signal detection, and signal processing. The adaptive prediction system, like the other prediction models cited, has some limitations for practical applications. One of the most critical problems seems to be the convergence. Although in certain situations, the adaptive prediction system could converge to the optimal states with given model structures (i.e., AL1 and N), in some other situations, the adaptive prediction system might not converge with the same model structures.

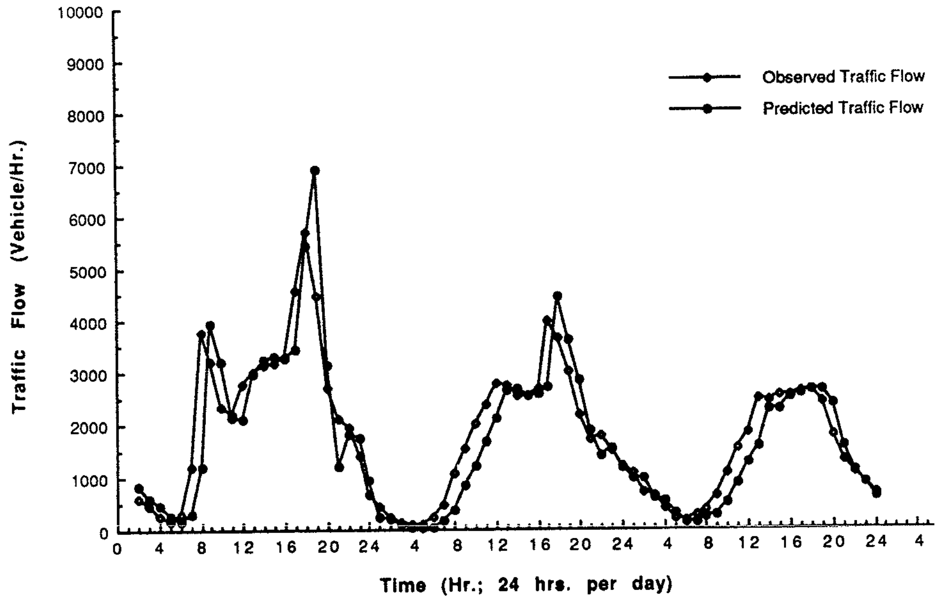

FIGURE 10 Prediction of traffic flow (traffic flow data collected from Mopac Highway, Austin, Texas) with $N = 6$, AL1 = $9 * 10^7$.

ACKNOWLEDGMENTS

The author would like to thank Clyde E. Lee, Professor of Transportation Engineering, in the Department of Civil Engineering of the University of Texas at Austin, for providing data and valuable suggestions. The author would like also to thank Carl Bertrand, Instrumentation Engineer, in the Center for Transportation Research of the University of Texas at Austin, for significant editing work and valuable suggestions.

REFERENCES

1. M. Papageorgiou. *Applications of Automatic Control Concepts to Traffic Flow Modeling and Control.* Springer-Verlag, New York, 1983.
2. I. Prigogine and R. Herman. *Kinetic Theory of Vehicular Traffic.* American Elsevier, New York, 1971.
3. Y. J. Stephanedes, P. Michalopoulos, and R. Plum. Improved Estimation of Traffic Flow for Real-Time Control. In *Transportation Research Record 795*, TRB, National Research Council, Washington, D.C., 1981, pp. 28–39.
4. J. B. Kreer. Factors Affecting the Relative Performance of Traffic Responsive and Time-of-Day Traffic Signal Control. *Transportation Research*, Vol. 10, 1976, pp. 75–81.
5. R. Kalman. A New Approach to Linear Filtering and Prediction Problem. *Journal of Basic Engineering*, Vol. 82D, ASME, 1960, pp. 34–35.
6. N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Time Series.* MIT Press, Cambridge, Mass., 1949 and 1964.
7. G. E. P. Box and G. M. Jenkins. *Time Series Analysis, Forecasting, and Control.* Holden Day, San Francisco, 1970.
8. B. Widrow and S. D. Stearns. *Adaptive Signal Processing.* Prentice-Hall, Englewoods Cliffs, N.J., 1985.
9. I. Okutani and Y. J. Stephanedes. Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory. *Transportation Research*, Vol. 18B, No. 1, 1984, pp. 1–11.
10. M. Cremer. Determining the Time-Dependent Trip Distribution in a Complex Intersection for Traffic Responsive Control. *Proc., 4th International Conference on Control in Transportation Systems*, IFAC/IFFIP/IFORS, 1983, pp. 145–151.
11. N. L. Nihan and G. A. Davis. Application of Prediction-Error Minimization and Maximum Likelihood to Estimate Intersection O-D Matrices from Traffic Counts. *Transportation Science*, Vol. 23, No. 2, 1989, pp. 77–90.
12. N. L. Nihan and G. Davis. Recursive Estimation of OD Matrices Using Inflow/Outflow Counts. *Transportation Research*, Vol. 21B, No. 2, 1987, pp. 149–163.
13. A. S. Willsky et al. Dynamic Model-Based Techniques for the Detection of Incidents on Freeways. *IEEE Transactions on Automatic Control*, AC-25, 1980.
14. D. C. Cazis and C. H. Knapp. On-Line Estimation of Traffic Densities from Time-Series of Traffic and Speed Data. *Transportation Science*, Vol. 5, No. 3, 1971, pp. 283–301.
15. M. S. Ahmed and A. R. Cook. Analysis of Freeway Traffic Time-Series Data by Using Box-Jenkins Techniques. In *Transportation Research Record 722*, TRB, National Research Council, Washington, D.C., 1979, pp. 1–9.
16. H. Nicholson and C. D. Swann. The Prediction of Traffic Flow Volumes Based on Spectral Analysis. *Transportation Research*, Vol. 8, 1974, pp. 533–538.
17. *Urban Traffic Control System and Bus Priority System Traffic Adaptive Network Signal Timing Program: Software Description.* FHWA, U.S. Department of Transportation, 1973.

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