CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Freeway Performance Measurement System (PeMS), Version 3 Phase II

Pravin Varaiya
University of California, Berkeley

California PATH Research Report
UCB-ITS-PRR-2004-9

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation; and the United States Department of Transportation, Federal Highway Administration.

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

Final Report for Task Order 4127

April 2004

ISSN 1055-1425

Freeway Performance Measurement System (PeMS), Version 3, Phase II:

Final Report of MOU 4127

Pravin Varaiya

University of California, Berkeley CA 94720
Department of Electrical Engineering and Computer Science
Tel: (510) 642-5270, Fax: (510) 643-2356
varaiya@eecs.berkeley.edu

July 11, 2003

Executive summary

This continuing PeMS project had four tasks:

1. Quantify the reduction in delay from optimum freeway operations;
2. Release travel time estimates to the public;
3. Training Caltrans staff in the use of PeMS;
4. Improve system robustness/maintenance.

For task 1, we developed and implemented an algorithm that quantifes potential travel time reduction. The algorithm takes historical travel demand data from PeMS. It simulates the resulting traffic flow, based on an idealized ramp-metering algorithm, and calculates the resulting travel times on the freeway and waiting time at the ramps. As an example, the algorithm estimates that the annual congestion delay of 75 million vehicle-hours in Los Angeles, District 7, could be reduced to 25 million vehicle-hours.

For task 2, we developed an algorithm for travel time predictions for Los Angeles, and made it available on the PeMS website. Users could select any two points on the freeway network, and state a departure (or desired arrival) time. The algorithm finds the 15 shortest alternative travel routes and estimates a travel time for each. The user could select any of these routes.

For task 3, we assisted Booz Allen & Hamilton in the preparation and presentation of PeMS training material. One training session was held in District 7. The feedback received from the participants was incorporated in PeMS v. 4.

Berkeley Transportation Systems accomplished task 4 under a subcontract to U.C. Berkeley. The result was a clearly defined maintenance procedure and a very stable system.

1. Causes and cures of congestion

People believe congestion occurs because demand exceeds capacity, so they support initiatives to build additional highway capacity or curtail highway travel demand. Politicians work to bring highway construction projects into their districts, and environmentalists support proposals to make transit more attractive or automobile use more costly. Our analysis of the facts does not support the belief that congestion occurs mainly because demand exceeds capacity.

When traffic density reaches a critical value, traffic flow breaks down, resulting in a major reduction in speed and flow. Once the breakdown occurs, it takes a long time before traffic flows freely. In the meantime, travelers incur a large congestion delay, measured in vehicle-hours as the extra time spent driving below the free flow speed, taken to be 60 mph.

An ideal ramp-metering system would sense the traffic density, and maintain the on-ramp flows in such a way that the density is just below the critical (breakdown) density. This ideal scheme has two major effects:

1. Traffic on the freeway is always maintained at the free flow speed of 60 mph;
2. Vehicles spend some time queued up at the ramps.

The net benefit of such an ideal system is the reduction in travel time on the freeway minus the queuing delay at the ramps. We developed an ideal ramp-metering algorithm. The algorithm takes a travel demand profile from historical data, available in PeMS. It then simulates the freeway operation under this ideal ramp-metering algorithm. It calculates the resulting freeway travel time and the time spent by vehicles in queue at the ramps.

We applied this algorithm to data for District 7 for the week of October 3-9, 2000. The results were then ‘blown up’ for one year, assuming the same travel patterns. The estimated congestion delay for the year was 75 million vehicle-hours. With the ideal ramp-metering scheme the delay is reduced to 25 million vehicle-hours—a saving of 50 million vehicle-hours. At an opportunity cost of $20 per vehicle-hour, this amounts to an annual savings of $1 billion.

This exercise suggests that ramp-metering, properly applied, can lead to a huge reduction in congestion delay—far exceeding the gains from other means of congestion reduction. Because of limitations in ramp space, it may not be practicable to implement such a perfect ramp-metering system everywhere. Nonetheless the algorithm can be used in planning studies to determine locations where ramp-metering would provide the maximum benefit. Details of the algorithm and its application are in [1], [2].

2. Travel time predictions

PeMS provides 5-minute average speeds at every loop detector station. Given the speeds throughout the freeway network at the current time, *t*, a simple way to predict the travel time over a route from an origin O to a destination D in the network is given by the formula,

$$(1)\ T^*(O,D,t)=\sum_i \frac{L_i}{V_i(t)}\,.$$

In this formula $T^*(O,D,t)$ is the predicted travel time, the sum is over all segments i along the route from O to D, L_i is the length (in miles) of segment i, and $V_i(t)$ is the speed (miles per hour) on this segment at time t.

Most travel time estimation algorithms use this *naïve* prediction formula. We call it naïve because it is based on the assumption that current speeds will be maintained in the future. This assumption is false, particularly at the beginning and the end of the congestion period.

Our algorithm is more sophisticated and much more accurate. Suppose the current time is t, and we want to predict the travel time at future time $t + d$. (Here $d \geq 0$.) The PeMS prediction is given by the formula,

$$(2)\ T_P(O,D,t+d)= a(t,d)T^*(O,D,t)+b(t,d)T_h(O,D,t+d)\,.$$

In this formula $T_P(O,D,t+d)$ is the travel time prediction, $T^*(O,D,t)$ is the naïve current travel time estimate (evaluated according to formula (1)), and $T_h(O,D,t+d)$ is the *historical* travel time at time $t+d$. The regression coefficients $a(t, d)$ and $b(t, d)$ are calculated from historical data.

Calculation of the regression coefficients and storing them for every segment and time is cumbersome, but the calculation needs to be done only once. The algorithm and its application are described in [3].

This algorithm is used to give route guidance. The user selects an origin O and destination D. The algorithm first finds 15 shortest routes from O to D, using a version of Dijkstra's algorithm. Travel times for each of the 15 routes is estimated using (2). The results are returned to the user.

This 'route guidance' service was available on the PeMS website. However, because of a threat of an alleged patent violation, this service is no longer available. The route guidance design is described in [4].

3. Training

Booz Allen & Hamilton was contracted to develop a functional specification for PeMS, to document the system, and to hold one or more training sessions for Caltrans staff. The purpose of the training was to familiarize participants with PeMS and to obtain their feedback to improve PeMS usability.

We assisted Booz Allen & Hamilton in the preparation of the material and we attended the training. The result of this were extremely valuable. The suggestions have been incorporated in PeMS version 4.

4. Maintenance

Berkeley Transportation Systems (BTS) is responsible for code development and for maintenance of PeMS under a subcontract to the University of California. Under the current project, BTS made many changes that have dramatically improved the robustness and maintainability of PeMS. The most notable changes are:

1. A system that automatically backs up the data on tape;
2. An 'alarm' system that automatically notifies any disruption in data from the various districts;
3. Administrative aids to monitor PeMS user traffic and to determine the level of access that different users can obtain.

As a result, the system maintenance is carried out at very low cost, and the system manager receives good online reports on system usage and he or she has very good control over who can access the system.

5. Acknowledgements

The research reported here is the joint work of the PeMS Development Group, in particular, Chao Chen, Jaimyoung Kwon, Alexander Skabardonis and Pravin Varaiya of U.C. Berkeley and Bill Morris and Karl Petty of BTS. We have benefited greatly from advice, comments and interest of Tom Choe, Fred Dial, Joe Palen and John Wolf of Caltrans, and Tarek Hatata of System Metrics Group.

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views of or policy of the California Department of Transportation. This report does not constitute a standard, specification or regulation.

References

[1] C. Chen, Z. Jia and P. Varaiya, Causes and cures of highway congestion, *IEEE Control Systems Magazine*, vol 21 (6), 26 –32, Dec. 2001.
[2] P. Varaiya, California's performance measurement system: Improving freeway efficiency through transportation intelligence, *TRNews*, No.218,18-24, Jan-Feb2002., No.218, 18-24, Jan-Feb2002.
[3] J. Rice and E. van Zwet, A simple and effective method for predicting travel times on freeways, submitted to *IEEE Transactions on Intelligent Transportation Systems*, August 2002.
[4] C. Chen, Freeway Performance Measurement System (PeMS), PATH report, UCB-ITS-PRR-2003-22, July 2003.
[5] T. Choe, A. Skabardonis and P. Varaiya, Freeway performance measurement system (PeMS): an operational analysis tool, Tranportation Research Record, 1811, 2002.

Appendix

The appendix consists of versions of references [1], [2], [3] and [5]. The PhD thesis [4] is a PATH report. It contains a detailed description of the PeMS system.

Causes and Cures of Highway Congestion

Chao Chen, Zhanfeng Jia and Pravin Varaiya
{chaos,jia,varaiya}@eecs.berkeley.edu
Electrical Engineering & Computer Science
University of California, Berkeley

August 18, 2001

Introduction

People believe congestion occurs because demand exceeds capacity, so they support initiatives to build additional highway capacity or curtail highway travel demand. Politicians work to bring highway construction projects into their districts; environmentalists support proposals to make transit more attractive or automobile use more costly. This article argues that the facts do not support the belief that congestion occurs because demand exceeds capacity.

On the contrary, the major cause of congestion is the inefficient operation of highways during periods of high demand. Analysis of data shows that congestion reduces highway efficiency by 20 to 50 %; that is, vehicles take between 20 and 50 % more time to traverse sections that are congested than they would if congestion were prevented. Compensation of this efficiency loss through, say, a 20% capacity expansion is financially impossible; compensation through a 20% demand curtailment is practically impossible. The best way to combat congestion is through increases in operational efficiency. To increase efficiency, however, it is necessary to intelligently control access to highways through ramp metering.

We estimate that for Los Angeles, the annual congestion delay is 70 million vehicle-hours. If the highways were to be operated at 100 % efficiency, this delay would be reduced by 50 million vehicle-hours.

The paper is organized as follows. We first show that vehicles travel at 60 mph when there is maximum flow on a highway section (i.e., when the section is operating most efficiently). Thus congestion delay should be measured as the extra time vehicles spend on the highway traveling below 60 mph: a vehicle taking 20 minutes to travel 10 miles at 30 mph suffers a congestion delay of 10 minutes.

The efficiency η of any highway section may then be defined as

$$\eta = \frac{VMT/60}{VHT} ,$$

where *VMT* is the total number of vehicle-miles traveled and *VHT* is the total number of vehicle-hours traveled over the section over some time interval, such as the morning commute period. We will see that η overestimates efficiency; we use this definition anyway because it is easy to calculate from data.

We then present evidence to support the following model of traffic: if the occupancy on a highway section is kept below a certain critical level, its efficiency will be 100 %, congestion will not occur, and traffic will flow at 60 mph; attempts to increase occupancy above the critical level will cause congestion and a rapid drop in efficiency.

This model leads to an idealized ramp metering control policy (IMP), which holds vehicles back at the on-ramps so that the occupancy on each highway section is maintained at its critical level. The total travel time under IMP, VHT_{imp}, is the sum of highway travel time (at 100 % efficiency, 60 mph) and the delay at the ramps imposed by IMP:

$$VHT_{imp} = \frac{VMT}{60} + Delay_{ramp} .$$

Therefore, the travel time savings from IMP is

$$\begin{aligned} VHT_{saved} &= VHT - VHT_{imp} \\ &= VHT - \frac{VMT}{60} - Delay_{ramp}. \end{aligned}$$

As $VHT - VMT/60$ is, by definition, the congestion delay, this gives

$$Congestion\ delay = VHT_{saved} + Delay_{ramp}.$$

Observe that $Delay_{ramp}$ may be attributed to excess demand (i.e., demand that exceeds the maximum flow supported by the highway operating at 100 % efficiency). For Los Angeles, our estimates are $Congestion\ delay = 70$ million, $VHT_{saved} = 50$ million, and $Delay_{ramp} = 20$ million vehicle-hours per year.

In contrast to the belief that attributes *all* congestion delay to demand exceeding capacity, we find that the congestion delay consists of a (large) part that can be eliminated by IMP and a residual that can be reduced only by shifting demand during peak periods. Demand may be shifted to other modes, such as public transit, or over time to nonpeak periods.

The penultimate section compares the problems of highway congestion and strategies to relieve it by ramp metering with similar problems and proposed solutions in communication networks and power systems.

What Is Congestion?

Measures of congestion delay compare the actual travel time to some standard. There are two defensible standards: one is travel time under free flow conditions (nominally 60 mph), and the other is travel time under maximum flow. Drivers understand the first standard, transportation professionals approve the second. We analyze data to show that the two standards coincide at least for Los Angeles highways, where the maximum flow in most highway sections occurs near 60 mph.

California's Department of Transportation divides the state into 12 districts. The largest district, Los Angeles, comprises Los Angeles and Ventura counties. We obtained Los Angeles data from the PeMS (Performance Measurement System) database [1],[2].

PeMS receives real-time data from several districts. The data are produced by loop detectors buried in the pavement in each lane of the highway and spaced one-third to one-half mile apart. Every 30 s, the detectors report two numbers: flow and occupancy. *Flow* (often called count) is the number of vehicles that crossed the detector in the previous 30 s. We report flows in vehicles per hour, or *VPH*. *Occupancy* is the fraction of the previous 30 s that a vehicle was present over the detector.

A useful identity relates the three fundamental quantities of highway traffic:

$$Occupancy = \frac{Flow \times VehicleLength}{Speed} \ ,$$

where *VehicleLength* is the vehicle length (in miles) and *Speed* is the speed in mph. When occupancy exceeds a critical value, congestion sets in and speed drops, as is shown later. The critical occupancy varies with the section.

There are 4,199 detectors at 1,324 locations in Los Angeles highways. PeMS processes data from these detectors in real time and calculates 5-min averages of speed (mph) and flow (*VPH*). We analyze these averages for a 12-hr period beginning midnight of September 1, 2000, and bracketing the morning commute period. We limit the study to the 3,363 functioning detectors.

Detectors are located in all lanes. A section is a portion of a highway associated with a set of detectors (one per lane) and may contain one on- or off-ramp, as depicted in Fig. 1. We attribute a detector's data to the section in which it is located. So, for example, a recorded flow of 1,000 *VPH* for a half-mile section leads to a calculation of 500 *VMT* over that section during 1 hr.

For each detector, we find the 5-min interval in which it reported the maximum flow over the 12-hr study period. We then calculate the average speed reported by this detector over a 25-min interval surrounding this 5-min interval of maximum flow. That is, if the detector reported maximum flow in interval t, we calculate the average speed over the intervals $t-2, t-1, t, t+1, t+2$. This 25-min average is, therefore, a *sustained* speed. (The speed

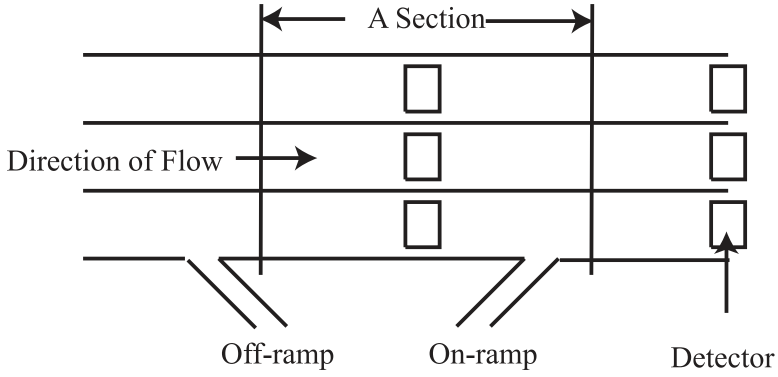

Figure 1: A section is a portion of highway at a detector location and may contain one on- or off-ramp.

at t is usually larger.) Fig. 2 gives the distribution of these speeds. The figure warrants the conclusion that the sustained speed at the time of maximum flow is 60 mph.

Fig. 3, which disaggregates by lane the data in Fig. 2, reinforces the conclusion, since the speed at maximum flow ranges from 65 mph in lane 1 (the innermost, fast lane) to 55 mph in lane 4 (the outermost, slow lane). Traffic on car-pool or HoV lanes is not included in the study.

Efficiency

We view a highway section as capital equipment that takes vehicle-hours traveled, *VHT*, as input and produces vehicle-miles traveled, *VMT*, as output. This is analogous to any other capital equipment that consumes certain variable inputs, such as labor, to produce some good or service.

According to this view, the output "produced" in one hour by a section of highway of length *SectionLength* miles is

$$VMT = Flow \times SectionLength.$$

The corresponding input is

$$VHT = \frac{VMT}{Speed}.$$

The ratio of output to input, *VMT/VHT*, is a measure of the *productivity* of this section (during this hour). Its unit is mph.

The maximum value of output produced is

$$MaxVMT = MaxFlow \times SectionLength,$$

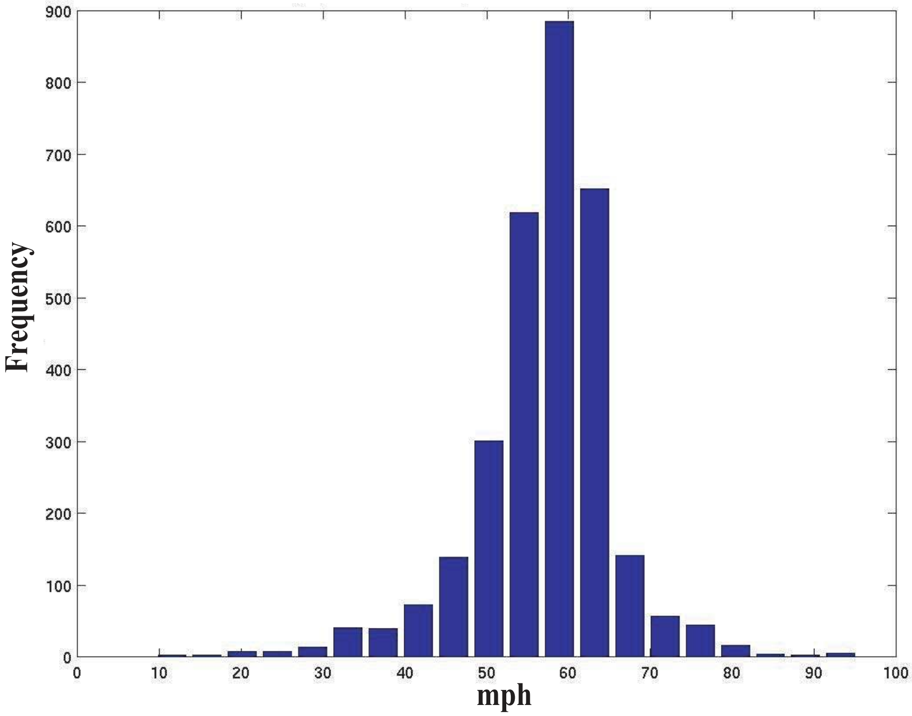

Figure 2: Distribution of average detector speed over a 25-min interval surrounding the time when the detector records maximum flow.

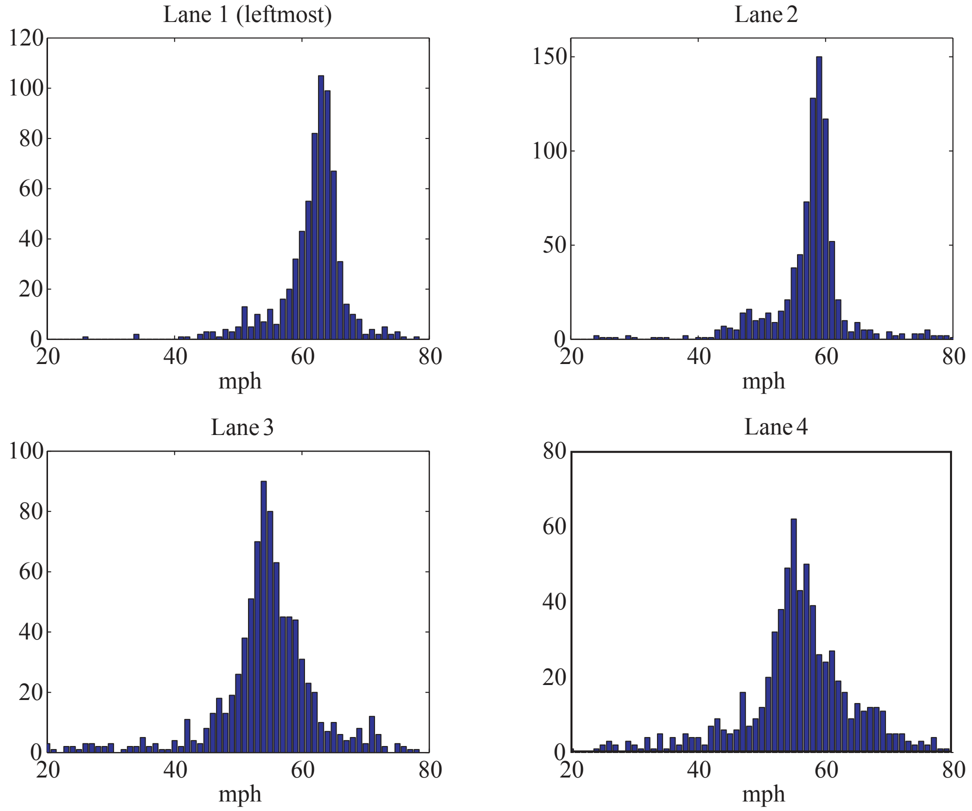

Figure 3: Distribution by lane of average detector speed over a 25-min interval surrounding the time when the detector records maximum flow.

where *MaxFlow* is the maximum flow that is observed on that section during the study period; the speed then is 60 mph. Thus, the maximum productivity is 60 mph. We define the *efficiency* index of a highway section as the ratio of actual to maximum productivity,

$$\eta = \frac{\mathit{VMT}/\mathit{VHT}}{60} \ . \tag{1}$$

This formula permits the following interpretation. Suppose the observed input over a section is $\mathit{VHT} = 10,000$ vehicle-hours and $\eta = 0.8$ or 80 %. Then for the *same* trips, the travel time would be reduced by 2,000 vehicle-hours were the section to operate at 100 % efficiency.

The formula also serves to calculate the efficiency not just for one section and one hour but for a highway network and any duration:

$$\eta_{network} = \frac{\sum_t \sum_k \mathit{VMT}_k(t)/\sum_t \sum_k \mathit{VHT}_k(t)}{60} = \frac{\mathit{VMT}_{network}/\mathit{VHT}_{network}}{60} \ , \tag{2}$$

where k ranges over all sections in the network and t ranges over the appropriate 5-min intervals; $\mathit{VMT}_{network}$, $\mathit{VHT}_{network}$ are simply the sum of the *VMT* and *VHT* over all the sections and time intervals.

We contrast our approach with the standard practice in traffic engineering. The observed *MaxFlow* depends upon physical characteristics, such as the section's grade and curvature, how it is connected to other sections, the location of on- and off-ramps, etc. It also depends on the pattern of traffic and how well the highway is operated. In the standard approach, the section is modeled in isolation to derive a theoretical maximum flow or *capacity* (which should be larger than the observed *MaxFlow*). The canonical model for deriving capacity is given in [3]. To prevent confusion with this notion of capacity, we do not use this term, and retain the empirically defined maximum throughput, *MaxFlow*. See [4] for a more detailed treatment of the distinction.

We estimate the congestion delay in Los Angeles using (2). PeMS provides 5-min averages of *VMT* and *VHT* for each section. For each day during the week of October 3-9, 2000, and the period midnight to noon, we calculate *VMT* and *VHT* for the network consisting of highways I-5, I-10, US 101, I-110, and I-405. For example, the calculation of *VMT* for I-5 for October 4 is

$$\mathit{VMT} = \sum_k \sum_t \mathit{VMT}_k(t),$$

where k ranges over all sections k in I-5 North and I-5 South and t ranges over all 5-min intervals on October 4 from midnight to noon. The results for the network are displayed in Table 1.

Table 1 deserves some comment. *VMT* is simply the sum of *VMT* over the five highways and the seven days in the week. *VHT* is obtained similarly. The % efficiency, η, is calculated using (2). The congestion delay, *VHT* - *VMT*/60, is the additional vehicle-hours spent

Table 1: Congestion delay and potential savings on five highways in Los Angeles during the week of October 3-9, 2000, midnight to noon.

Vehicle-miles traveled, *VMT*	= 86 million
Vehicle-hours traveled, *VHT*	= 1.85 million
Average efficiency, η	= 77 %
Vehicle-hours of congestion delay	= 404,000
Congestion delay saved by IMP	= 280,000
Delay due to excess demand	= 124,000

driving under 60 mph. The congestion delay saved by IMP is the potential reduction in congestion delay under the IMP ramp-metering policy, described later. Thus, the congestion delay is reduced by 280/404, or 70 %. The remaining delay is due to excess demand: it is the vehicle-hours spent behind ramps under IMP.

The period midnight to noon includes not only the morning congestion period but also periods when there is no congestion. The week of October 3-9 includes the weekend, when there is no congestion. The *VHT*, *VMT* include both highway directions, only one of which is congested during the morning commute. When there is no congestion, traffic is moving at 60 mph and efficiency is 100 %. The estimate η of 77 % average efficiency includes these non-congested periods and directions, so if we were to limit attention to the morning commute hours, the efficiency estimate would drop significantly. Fig. 4 shows the daily variation in congestion.

Evening traffic is more congested than morning traffic, so the congestion delay over the entire week is at least 808,000 vehicle-hours. For a 50-week year this amounts to 40 million vehicle-hours. The remaining highway network in Los Angeles is 75 % longer, so we estimate the annual congestion delay in all Los Angeles highways to be 70 million vehicle-hours. Assuming that 70 % of this delay can be saved by IMP, this amounts to 50 million vehicle-hours each year. Valuing the opportunity cost of time at $20 per vehicle-hour, this gives an annual savings of $1 billion.

From a more inclusive perspective, the efficiency index (1) is an underestimate, since it only accounts for changes in speed and not in flow. As a hypothetical example, consider a section with a maximum flow of 2,000 VPH at 60 mph, but which during congestion has a flow of 1,800 VPH at 30 mph. The efficiency according to (1) is 30/60 = 0.5, reflecting the drop in speed, but it does not reflect the 10 % reduction in flow. A better measure of the potential efficiency appears to be

$$\hat{\eta} = \frac{\mathit{Flow} \times \mathit{Speed}}{\mathit{MaxFlow} \times SpeedAtMaxFlow}. \tag{3}$$

For the hypothetical example, $\hat{\eta} = 0.45$ instead of 0.5. (The product $\mathit{Flow} \times Speed$ was proposed as a measure of performance in [5].)

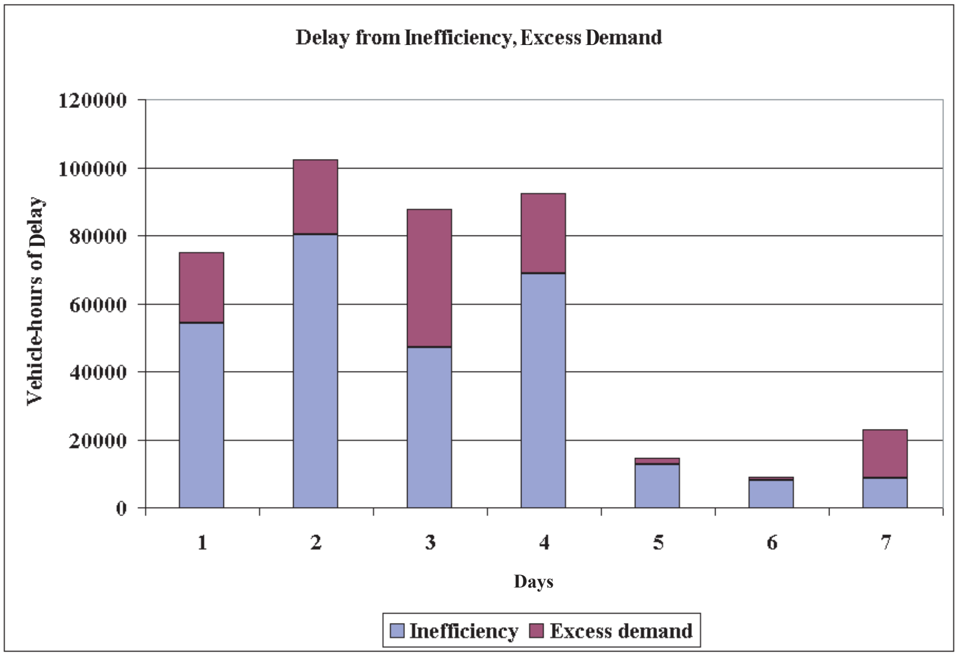

Figure 4: Variation in congestion during week of October 3-9, midnight to noon. Day 1, October 3, is Tuesday; Day 7, October 9, is Monday.

Viewing the highway section as a queuing system sheds light on the formula (3). The section provides a service to its customers (vehicles): the transport of a vehicle across the section. The service time of a vehicle is

$$\frac{SectionLength}{Speed}.$$

The system serves *Flow* vehicles in parallel, so its throughput is

$$\frac{Speed}{SectionLength} \times Flow.$$

The maximum throughput is

$$\frac{60}{SectionLength} \times MaxFlow.$$

$\hat{\eta}$ is the ratio of actual to maximum throughput.

We estimate $\hat{\eta}$ for all sections of I-10W during the morning congestion period on October 1, 2000, as follows. For each section we determine the 5-min interval between midnight and noon when its detector recorded the maximum occupancy. This is the time of worst congestion, and we find the speed and flow at that time. The efficiency during congestion for this section is

$$\hat{\eta} = \frac{FlowAtMaxOcc \times SpeedAtMaxOcc}{MaxFlow \times 60}.$$

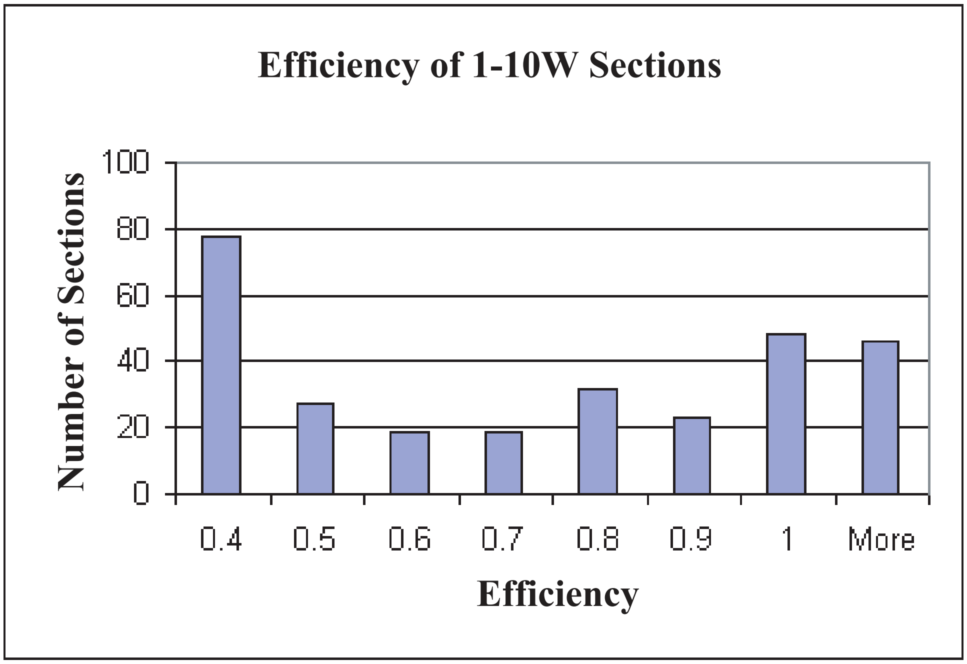

Figure 5: Variation in efficiency $\hat{\eta}$ during congestion along sections of I-10W, midnight to noon, October 1, 2000.

The distribution of $\hat{\eta}$ across 291 sections with functioning detectors on I-10W is shown in Fig. 5; 78 sections had an efficiency less that 40 %, 65 had an efficiency between 40 and 80 %, 71 had an efficiency between 80 and 100 %, and 46 had an efficiency above 100 % (recording speeds above 60 mph).

Behavior During Congestion

Fig. 6 is a plot of *Flow* vs *Occupancy* on one section of I-10W from midnight to noon on October 3, 2000. Each point is a 5-minute average of flow and occupancy. Successive points are connected by straight lines. Initially, vehicles travel at 60 mph and flow and occupancy increase in proportion. At 5:30 am, occupancy reaches a critical level, and flow reaches its maximum, 2400 VPH. Demand exceeds this maximum, congestion sets in, speed and flow decline while occupancy increases. At the depth of congestion, speed is 20 mph and flow has dropped to 1400 VPH. Demand then drops, and speed gradually recovers to 60 mph by 9:00 am. For this section, the critical occupancy level is 0.11.

This behavior suggests the model of congestion depicted in Fig. 7. Notice the three regimes in the “phase” portrait of the figure: free flow, then congestion, followed by recovery. The recovery phase is different from the congestion phase, reminiscent of hysteresis. Standard hydrodynamic models of fluid flow don’t exhibit such hysterisis. It is a challenge to invent

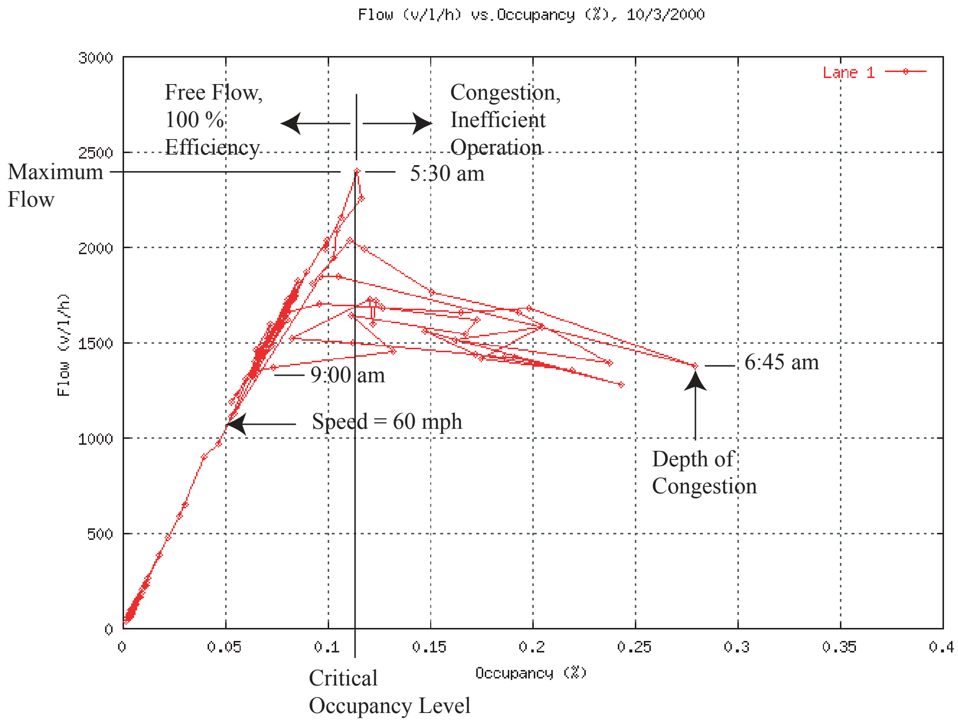

Figure 6: Flow vs. occupancy on a section at postmile 37.18 on I-10W, midnight to noon on October 3, 2000.

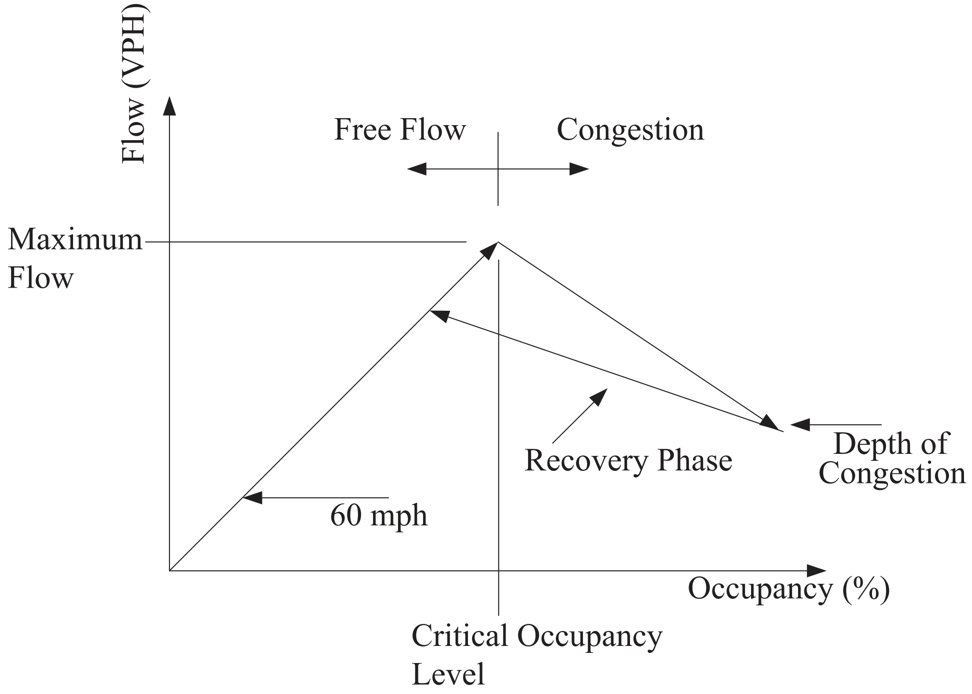

Figure 7: Model of congestion. If occupancy is maintained below critical level, section operates at 100 % efficiency and speed is at 60 mph.

a dynamic model that exhibits such transitions. Measurements of the recovery phase show erratic fluctuations as in Fig. 6. The model in turn supports the following hypothesis about traffic behavior:

> *If a metering policy keeps occupancy below its critical level in every section, efficiency will be 100 %, speed will be maintained at 60 mph, and highway congestion will be prevented. A consequence of the metering is that vehicles will be stopped at the ramps for some time.*

We call this the *ideal ramp metering (IMP) principle*. The IMP feedback strategy is to monitor the occupancy downstream of each on-ramp and to throttle the flow from the on-ramp whenever the occupancy exceeds its critical value. For a control-theoretic discussion of this policy, see [6].

The figures in Table 1 are computed as follows. For each highway section we calculate the critical occupancy level from PeMS data, as in Fig. 6. We assume that the pattern of demand is unchanged. We now simulate the traffic flow using the model of Fig. 7 and the IMP feedback strategy. In the simulation, vehicles will be held back at some ramps. We calculate the total time spent by the vehicles at the ramps. The ramp delay is 124,000 vehicle-hours. There is a net savings of 280,000 vehicle-hours.

Other Networks

As with highway networks, the movements of data and electric power are also organized in networks with high-capacity transmission links to take advantage of scale economies. Since they accommodate the demand from uncontrolled or unpredictable users, however, all three networks can experience congestion.

We explore similarities and differences in the congestion these three types of networks experience along the dimensions of demand, routing, and control.

The demand patterns in transportation and data networks are similar: users want to move commodities (vehicles, data packets) from some nodes to others. In power networks, users impose loads at some nodes that are supplied by power from other nodes. A power system is a single commodity network; the others are multicommodity networks.

Until recently, power system planners, like transportation planners, concerned themselves with finding effective ways to expand generation and transmission capacity to meet the forecast demand, assumed to be exogenous. Today there is a realization that curtailing demand for power is essential. One way to curtail demand is through real-time congestion pricing [7]-[9]. This is similar to the suggestion of transportation economists. There is also a growing literature exploring the use of pricing to shape demand in data networks [10]-[13]. Congestion pricing is not used in practice in data networks: it is used to an increasing extent in bulk power markets and to a very limited extent in highway transportation. It is practiced by airlines under "yield management." (We are ignoring the significant differences between spot prices for bulk power, time-of-day tolls in highways, and pricing of airplane seats for market segmentation. Only one component of these prices concerns congestion.)

In transportation networks, demand can be directly controlled by ramp metering. The counterpart in data networks is called admission control: at the entrance to the network certain flows or connections may not be admitted into the network. (Admission control is rare in data networks, but it is common in telephony: if the telephone network cannot route your call, it is blocked. Admission control can be based on the origin, destination, and other attributes of the data traffic. Ramp metering usually cannot distinguish between vehicles by their destination.) Power systems, too, employ admission control during emergencies by shutting down voluntary "interruptible" loads or by forced "rotating blackouts."

Routing of data packets is fully controlled by the routers located at the nodes. Power flows along routes determined by the laws of physics, given the patterns of sources (generation) and sinks (loads). Thus, the power flow routes cannot be directly controlled. (A limited amount of control can be exercised using expensive FACTS devices.) Similar to power, the routes chosen by drivers cannot be controlled. They prefer routes with shorter travel times, but the travel time on a congested highway section depends on the traffic flow, which in turn depends (through driver choice) on the travel time. Thus, routes and travel times are jointly determined in a simultaneous equation system, similar to the load flow equations that determine power flows and phase angles.

The differences in demand and routing, and the controls that can be exercised, affect the nature of congestion in the three networks.

A transmission link in a power network is congested if the power flowing through it is close to its thermal capacity. Additional power through the link carries the risk of a line fault, endangering the transfer of power in other links. Since power flows cannot be controlled, the *only* way to reduce the flow through an overloaded link is to change the pattern of power generation and consumption.

In data networks, transmission links do *not* get congested. They transmit at a fixed line rate. (An exception is congestion due to contention for access in shared Ethernet and wireless links.) Instead, congestion occurs at a node or router when the rate at which data to be forwarded over a particular outgoing link exceeds that link's line rate. The router's buffer then overflows and the router is forced to drop the packet.

Conclusions

For many years, the increase in travel demand has outstripped additions to California's highway infrastructure. Congestion is worse each year. Rising housing costs in high-employment regions force people to live further away, lengthening commutes, and increasing congestion. The resulting low-density housing makes current transit options (rail and buses) costly and less effective.

We have argued that a large portion of highway congestion can be attributed to inefficient operation. The inefficiency is greatest when demand is greatest. Empirical analysis indicates potentially large gains in efficiency, with dramatic reductions in congestion. Intelligent ramp metering control strategies can realize these gains.

One reason these strategies are not implemented is the widely held belief that congestion is determined by demand, and ramp metering merely transfers delay that would occur on the highway to delay at the ramps. But our analysis concludes that intelligent ramp metering transfers only a *fraction* of the highway delay to the ramps; the rest of the delay is eliminated. Of course, further empirical studies that test this conclusion are needed.

Transportation economists have long recognized that congestion is a "negative externality" and proposed congestion tolls to limit highway access during periods of high demand [14], [15]. But equity and engineering considerations suggest that in most places, ramp metering is easier to deploy than congestion tolls.

Transportation, power, and data networks face congestion. Congestion in data networks has to date been contained by expanding capacity ahead of demand growth. Until recently, that was also the strategy followed by transportation and power network operators; but that option today is frequently not available. The only option is to put in place efficiency-enhancing control strategies. But that poses challenges of control strategy design and the development and deployment of sensors and controller technologies to implement these strategies. Those challenges are just beginning to be addressed.

Acknowledgments

The paper builds on work of the PeMS Develpoment Group. Markos Papageorgiou and Alex Skabardonis helped us with their comments and criticisms. The PeMS project is supported by the State of California Department of Transportation, the EPRI/DoD Complex Interactive Networks Initiative under Contract WO8333-04, and the National Science Foundation under Grant CMS-0085739. The authors alone are responsible for the opinions expressed here and for any errors.

References

[1] transacct.eecs.berkeley.edu

[2] C. Chen, K. Petty, A. Skabardonis, P. Varaiya and Z. Jia, "Freeway performance measurement system: mining loop detector data," 80th Annual meeting of the Transportation Research Board, Washington, D.C., January 2001.

[3] Transportation Research Board. *Highway Capapcity Manual 2000*, Chapter 23. Washington, D.C., National Research Council, 1998.

[4] Z. Jia, P. Varaiya, K. Petty, and A. Skabardonis, "Congestion, excess demand, and effective capacity in California freeways," submitted to *Transportation Research*, December 2000.

[5] D.R. Drew and C.J. Keese, "Freeway level of service as influenced by volume and capacity characteristics," *Highway Research Record*, no. 99, pp. 1-47, 1965.

[6] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: an overview," *Proc. IEEE Intelligent Transportation Systems Conference*, Dearborn, MI, October 2000,

[7] C-W. Tan and P. Varaiya, "A model for pricing interruptible electric power service," in G.B. DiMasi, A. Gombani, and A.B. Kurzhanski (Eds.), *Modelling, Estimation and Control of Systems with Uncertainty*, pp. 423-444. Cambridge, MA: Birkhauser Boston, 1991.

[8] H-P. Chao, G. Huntington (Eds.), *Designing Competitive Electricity Markets*. Boston, MA: Kluwer, 1998.

[9] F.F. Wu and P. Varaiya, "Coordinated multilateral trades for electric power markets: theory and implmentation," *Electric Power and Energy Systems*, vol. 21, pp. 75-102, 1999.

[10] J. Walrand and P. Varaiya, *High-performance communication networks*, 2nd ed. San Francisco, CA: Morgan Kaufmann, 2000.

[11] R. Edell and P. Varaiya, "Providing Internet access: What we learn from INDEX," *IEEE Network*, vol. 13, no. 4, pp. 18-25, 1999.

[12] J.K. Mackie-Mason and H.R. Varian, "Pricing congestible network resources," *IEEE J. on Selected Areas in Communications*, vol. 13, no. 7, pp. 1141-9, 1995.

[13] S.H. Low and D.E. Lapsley, "Optimization flow control, I: basic algorithm and convergence," *ACM/IEEE Transactions on Networking*, vol. 7, no. 6, pp. 861-75, 1999.

[14] A. Waters, "Theory and measurement of private and social cost of highway congestion," *Econometrica*, vol. 29, pp. 676-99, 1961.

[15] T.E. Keeler, *The full costs of urban transport*, University of California, Inst. Urban and Regional Development, 1975.

Creating transportation system intelligence

Pravin Varaiya
University of California, Berkeley

Information technology (IT) provides the means to store, manipulate and disseminate massive amounts of data. The integration of IT at all levels of the transportation system creates the "intelligence" in Intelligent Transportation Systems (ITS). But this integration is a long and difficult process of the search for, and the exploitation of, the numerous opportunities in the interconnected set of operations, planning and investment procedures constituting today's transportation systems.

This paper gives a glimpse into the opportunities for enhancing freeway systems productivity. The discussion summarizes three years of experience with the Performance Measurement System or PeMS—a system that collects and stores data from California loop detectors, and converts these data into useful information. Examples from Los Angeles illustrate how this information can improve system management, assist travelers, and challenge current understanding of freeway traffic behavior.

With PeMS, freeway operations and planning can be based on information that previously was unavailable or too costly to gather. Routine reports, like California's congestion monitoring report that today consume appreciable resources to produce, can be generated at no cost. Significantly, engineers and planners can quickly isolate problem areas and focus on potential solutions as seen in two examples: identification of bottlenecks, and location of freeway segments where intelligent ramp-metering can significantly reduce congestion.

Travelers face a large variation in travel time during peak hours. Less appreciated is the fact that knowledge of the current state of the traffic can drastically reduce this variation, so travel times can be accurately predicted. PeMS makes these predictions available, and uses them to suggest optimum routes.

Our understanding of traffic behavior is embodied in models that inform decisions and professional training. Typically in traffic engineering, these models are insufficiently validated. However, PeMS data can be easily used to test these models, as seen in two surprising findings. First, maximum throughput occurs at the free flow speed of 60 mph, and not between 35 and 50 mph. Second, a large fraction of freeway congestion delay is due to inefficient operation rather than to excess demand.

ITS is often simply associated with a set of technologies (AVL, CMS, ETC, VIPS, etc), to be deployed one at a time in return for corresponding incremental benefits. But the paper suggests that much larger productivity gains in the freeway system can be obtained when operations, planning, and investment processes are transformed to make intelligent use of these technologies guided by the kinds of information that PeMS provides.

The paper begins with a brief overview of PeMS and then gives examples of PeMS applications. The reader will gain a better appreciation of PeMS from its website http://transact.eecs.Berkeley.EDU.

System overview

PeMS collects and stores data from loop detectors operated by the California Department of Transportation (Caltrans). PeMS applications convert these data into useful information accessed through the Internet by Caltrans personnel, value-added resellers (VARs), the public, and the research community. PeMS is a functioning prototype. It will be deployed statewide in July 2002. It is a low-cost system, built from commercial, off-the-shelf components. PeMS can be deployed incrementally, with no disruption of existing procedures. (See reference 1.) Its software is open: other data sources will be incorporated in PeMS as they become available; furthermore, PeMS is designed to include electronic data from other transportation modes, such as transit.

The PeMS database computer is located in the University of California, Berkeley. The computer has 4 GB of main memory, and 4 terabytes of disk—enough to store several years of California data online.

The software is organized in three layers. At the bottom is database administration. The work is standard but highly specialized: disk management, crash recovery, table configuration. The middle layer comprises software that works on the data as they arrive in real time. It

- Aggregates 30-second flow and occupancy values into lane-by-lane, 5-minute values;
- Calculates the *g*-factor for each loop, and then the speed for each lane. (Most detectors in California are single loop, and only report flow and occupancy. PeMS adaptively estimates the *g*-factor for each loop and time interval. The algorithm and tests of its validity are reported in reference 2.);
- Aggregates lane-by-lane values of flow, occupancy, and speed across all lanes at each detector station. At this point, PeMS has flow, occupancy, and speed for each 5-minute interval for each detector station (one station typically serves the detectors in all the lanes at one location);
- Computes basic performance measures such as congestion delay, vehicle-miles-traveled, vehicle-hours-traveled, and travel times.

The top software layer comprises applications some of which are described below.

Routine reports

Systems at Caltrans depend on monthly or annual reports and programs that provide high-level information to policy makers. These include the Traffic Operations Strategies report, the Highway Congestion Monitoring Program (HICOMP), and System Performance Measures Initiative. PeMS can benefit these reports and programs, as seen in the case of the annual HICOMP report.

The HICOMP report presents the location, magnitude, and duration of congestion for California freeways. Caltrans uses this information to identify problem areas and to establish priorities for operational and air quality improvement projects. Data for the report are obtained from "floating" cars driven through 5-7 mile freeway sections twice a year during congested periods. Figure 1 is one of PeMS' "average plots." It gives the maximum, average, and minimum vehicle-hours of congestion delay—the extra time spent driving below 35 mph—on US-101N for each day of the week, averaged over a 16-week period beginning February 4, 2001.

During these 16 Wednesdays (day 4 in Figure 1), the delay ranged from a minimum of 10,000 to a maximum of 60,000 vehicle-hours. This 600 percent variation implies that the twice-a-year samples of the HICOMP report are unreliable. With PeMS it is possible to track congestion accurately to determine trends and departures from the trend to produce (at no cost) a report that gives meaningful statistical measures of congestion. More significantly, PeMS forces the recognition that congestion delay is a random quantity so its measurement and report must take the statistical fluctuations into account.

The preceding remark applies equally to many reports (such as census counts) based on one-shot samples of randomly fluctuating quantities. Travel time is an important component "mobility" measures. PeMS computes the travel time for each freeway segment starting every five minutes. As shown later, travel times fluctuate widely from day to day. A meaningful summary of travel times must reflect these fluctuations.

PeMS also collects statewide incident data reported by the California Highway Patrol, and locates these incidents on the same geographical basis as the loop detector data. So hypotheses relating incidents to traffic variables such as vehicle-miles traveled or congestion delay or freeway geometry can be formulated and tested. Today, the relation between incidents and congestion delay is based on folklore. PeMS permits the discovery and accurate specification of such a relation, if one exists.

Finding bottlenecks

The PeMS application, called "plots across space," can assist in identifying bottleneck locations for more detailed investigation. To use the application, the engineer selects a section of freeway, a time, and a performance variable such as speed, flow, or delay. PeMS returns a plot of the variable across space. The plot in Figure 2, for example, displays speed averaged across all lanes in the figure—PeMS also provides lane-specific plots—over a 30-mile stretch of I10-W, beginning at post mile 20, at 7.30 am on September 14, 2000.

The precipitous drop in speed from 60 to 20 mph near post mile 23 indicates a potential bottleneck. There is another potential bottleneck near post mile 32. The PeMS contour plot of Figure 3 confirms the existence of both bottlenecks. Further confirmation may be obtained by examining the same plots for other days. Without PeMS this would be a very time-consuming analysis. Having quickly determined the existence of these bottlenecks, the engineer can go on to determine their cause, such as the location of interchanges, the highway geometry, large flows at ramps, etc, and propose potential solutions to alleviate the bottleneck.

Observe how PeMS can dramatically shift the use of the engineer's time from the drudgery of collecting data and making contour plots, to the creative understanding of the causes of bottlenecks and finding opportunities for relieving them. Furthermore, any scheme implemented to relieve a bottleneck can be rigorously evaluated by a thorough before-and-after comparison. Different schemes can be compared. Over time, there will be an accumulation of experience statewide so that schemes can be implemented with some degree of confidence in their effectiveness.

Maximum flow occurs at 60 mph

The speed-flow relationship is fundamental to traffic theory. In the Highway Capacity Manual this relation is given by a smooth curve, which yields a maximum flow at a speed between 35 and 50 mph. We use PeMS to test this hypothesis using cross-sectional data from all 3,363 functioning loop detectors (out of a total of 4,199 detectors) at 1,324 locations in Los Angeles, for a 12-hour period beginning midnight of September 1, 2000, bracketing the morning commute hours.

For each detector we find the 5-minute interval during which the flow reaches its maximum value. We then calculate the average speed during a 25-minute interval surrounding this maximum-flow interval. This is a measure of the *sustained* speed at the time of maximum flow. Figure 4 displays the per-lane distribution of this speed: in (the innermost) lane 1 this speed is between 60 and 70 mph, in lane 2 it is between 55 and 60, in lanes 3 and 4 between 50 and 60. The test rejects the hypothesis that maximum flow occurs between 35 and 50 mph.

The finding has some important implications. First, congestion delay should be measured as the time spent driving below 60 mph, both because it is the most efficient speed and because drivers experience congestion below this speed. (Caltrans today measures congestion as the time spent driving under 35 mph continuously for 15 minutes.)

Second, a ramp-metering strategy will only be effective if it maintains free flow speed. Lower speeds, such as say 45 mph, are simply not sustainable. Figure 5 illustrates this. (It is an instance of a PeMS "x-y plot" application in which 5-minute averages of any two variables at a detector are plotted against one another.) It gives the speed-flow relationship on lane 1 between 4.00 and 8.00 am on September 14, 2000, at post mile 32.87 on I-10W, near the second bottleneck in Figure 2. Clearly, once occupancy increases to cause the speed to drop below 60 mph at 5.10 am, the flow becomes unstable, dropping to 30 mph by 5.30, and 15 mph by 7.00 am. It seems unlikely that traffic flow at any speed below 55 mph can be sustained. This conclusion is confirmed by examining hundreds of similar plots.

Potential gains from ramp-metering

A fairly complex PeMS application calculates the potential reduction in congestion from an ideal ramp-metering policy (IMP). We study a freeway section that experiences recurrent congestion during the morning rush hour from 6.00 to 10.00 am, on a particular day. In the example this is the 16-mile segment of I-210W, starting at post mile 22, on January 11, 2001. The time period spans the rush hour, say 4.00 am to noon. So traffic is free flowing at the beginning and end of the study period.

The freeway section comprises several PeMS segments. (A segment is the freeway surrounding a detector station half way to the two neighboring stations). Some segments have on-ramps, some have off-ramps, and some have neither. PeMS gives, for each 30-second interval, the inflows of vehicles into the study section from each on-ramp and from upstream of the section, as well as the outflows at each off-ramp and downstream of the section. PeMS does not have origin-destination data, so the application calculates the constant turning ratio that matches total inflow and outflow in each segment.

The application next calculates the maximum throughput in each segment. This is simply the maximum flow that was in fact observed in that segment during the study time period, 4.00 am to noon, January 11, 2001. The maximum flow is an empirical quantity, which varies slightly from day to day and from segment to segment, see reference 4.

The hypothesis underlying the application is this: if the flow on each segment is *always* maintained below this maximum (say, by 3 percent), then vehicles on that segment will travel at 60 mph. The preceding section strongly supports this hypothesis, although a true test would require field experiments.

IMP imposes the policy that at each on-ramp (and upstream of the study section) vehicles are admitted so long as the flow in every section does not exceed the maximum flow, less 3 percent. (This is not the way to implement IMP; that should be based on measuring occupancy downstream of each on-ramp.) With this policy, under the hypothesis above, a vehicle may be held back at an on-ramp, but once it enters the freeway it will travel at 60 mph.

The result of IMP is displayed in the three plots of Figure 6. The top curve plots the *actual* vehicle-hours spent in the study section during each 5-minute slice from 4.00 am to noon. (The units are normalized to vehicle-hours/hour. This is a simple calculation since flows and speeds are known.) So the area under the top curve is the total vehicle-hours actually spent in the section during that period. The bottom curve gives the vehicle-hours those vehicles would have been spent if they experienced no delay at the ramps and traveled at 60 mph. So the area under the bottom curve is the *free flow* vehicle-hours that would have been spent by the same traffic demand. The difference in the area under the top and bottom curves is the vehicle-hours of delay suffered by traveling less than 60 mph. The two curves coincide outside the 6.00 to 10.00 am congestion period, as expected.

The middle curve plots the vehicle-hours that would have been spent under the IMP metering policy. Recall that under the hypothesis above, a vehicle is either queued up at an on-ramp or traveling at 60 mph on the freeway. So the area between the middle and bottom curves is the vehicle-hours spent queued up at the on-ramps, and the area between the top and middle curves is the net reduction in congestion delay. In the example of Figure 4, there is about 3,000 vehicle-hours of total congestion delay, of which 2,400 is eliminated by IMP with 600 vehicle-hours of queuing delay at on-ramps.

This application can be used by planners to locate potential sites where ramp-metering may be advantageous. Ramp-metering is a contentious local public policy issue in California, and most discussion is based on unfounded allegations about its impact. PeMS can provide an empirical basis for estimating the cost and benefit of a proposed ramp-metering scheme. The application also calculates the queue lengths formed at all the ramps and upstream of the study section. That information can be used to determine whether there is sufficient capacity in the ramps. As is known, the queue at one on-ramp can be traded off against another. So the application can stimulate a study of alternative coordinated ramp-metering strategies, and coordinated arterial signaling.

We use this application for five freeways in Los Angeles for the morning commute periods of the week of October 3-9, 2000. We then “blow up” the results to all LA freeways. This leads to the estimate that travelers spend an extra 70 million vehicle-hours each year driving below 60 mph, of which 50 million-hours can be eliminated by intelligent ramp-metering. At (say) $20 per vehicle-hour, there is a potential annual savings of $1 billion, see reference 3.

Efficiency of freeway operations

The freeway segment of Figure 5 can support a flow of 2,100 vehicles/lane/hour at 60 mph. But at 7.00 am, when congestion is worst, it serves only 1,300 vehicles/lane/hour at 15 mph, indicating a drop in operating efficiency. We propose a measure of efficiency, η, given by the formula

$$\eta = \frac{Flow \times Speed}{MaxFlow \times SpeedAtMaxFlow(60)} \quad . \tag{1}$$

According to this formula, the efficiency of this segment at the time of worst congestion was $\eta = \frac{1300 \times 15}{2100 \times 60} = 13$ percent.

We justify formula (1) by viewing the freeway segment as a queuing system. The queuing system provides a service to each customer (vehicle)—the transport of the vehicle across the segment. The vehicle’s service time is $\frac{SegmentLength}{Speed}$. The system serves $Flow$ vehicles in parallel. The *throughput* of this queuing system at any time is the number of vehicles served per hour at that time, namely $\frac{Speed}{SegmentLength} \times Flow$. The maximum throughput is $\frac{SpeedAtMaxFlow(60)}{SegmentLength} \times MaxFlow$. Formula (1) defines efficiency as the ratio of actual throughput to maximum throughput.

We use PeMS to estimate the efficiency of all 291 segments of I-10W with functioning detectors during the morning congestion period on October 1, 2000. For each segment we find the 5-minute interval between midnight and noon when its detector recorded the maximum occupancy. This is the time of worst congestion, and we find the speed and flow at this time. As before, we also find the maximum flow during the 12-hour interval. Using these in formula (1) gives the efficiency of the segment during worst congestion.

Figure 7 gives the distribution of efficiency for the 291 segments at the time of worst congestion: 78 segments had efficiency under 40 percent, 65 had efficiency between 40 and 80 percent, 71 had efficiency between 80 and 100 percent, and 46 had efficiency above 100 percent (they recorded a speed above 60 mph at time of maximum occupancy).

California's freeway system cost $1 trillion dollars. The calculation above shows that this capital stock is operated at a very low efficiency precisely at times of greatest demand (worst congestion). The potential gain from restoring efficiency will far exceed any practically conceivable increases in capacity through new construction. Any program to build "intelligence" in the transportation system must surely take as its main objective the recovery of this efficiency loss.

Travel times

Figure 8 gives the travel times for a 48-mile trip over I-10E, beginning at post mile 1.3 at any time between 5 am and 8 pm, for 20 working days in October 2000. The data are obtained from PeMS' travel time calculations.

Anyone planning this trip faces the statistical distribution implicit in this figure. If you intend to leave at say 5 pm, the "vertical slice" through the figure at 5 pm gives the prior distribution you face. So your trip may take between 45 and 130 minutes, with a 70 percent chance it will take between 60 and 100 minutes and a 10 percent chance it will take more than100 minutes. If you want to place a 90 percent confidence interval around your travel time, the best you can do is between 55 and 110 minutes—a 200 percent variation. But this variation can be drastically reduced if you know current conditions.

There are 20 curves in the figure, one for each day. The curve for a particular day is obtained from the travel times for that day starting every 5 minutes between 5 am and 8 pm. It is evident that if the trip starting at 5 pm takes more than 100 minutes, then that trip belongs to curves (days or random draws) for which the trip starting at 4 pm takes more than 90 minutes. What this means is that if you know the *current* travel time for a particular trip, you can predict the *future* travel time quite well. If at 4 pm the travel time is 90 minutes, then you can be 90 percent confident that at 5 pm the travel time will be between 85 and 110 minutes—a 25 percent variation. On the other hand, if at 4 pm the travel time is 60 minutes, the 90 percent confidence interval for a trip staring at 5 pm is between 60 and 80 minutes. In statistical terms this means that the unconditional variance in the travel time distribution is large, but the variance on the distribution *conditioned* on present and past values is much smaller. Of course, the further into the future you want to predict, the less important is knowledge of the current state of traffic: if you know the travel time at 5 am, and you want to predict the travel time at 5 pm, you can't do better than the unconditional distribution.

A PeMS application makes point estimates of future travel times for each freeway segment based on current and past travel times. (The past travel times are stratified by day of week and time of day to extract trends. See reference 5.) Through her browser, a user first indicates a proposed trip by clicking on the origin and the destination on the freeway map. The user then selects a start time (or the arrival time) and PeMS calculates 15 routes and their travel time estimates, including the routes with the shortest travel time and the shortest distance. The other routes differ by one link from these.

Other applications

The preceding sections indicate some uses of PeMS. We quickly list some others. Freeway lanes are often closed in response to requests for scheduled maintenance. With PeMS, one can compute the likely delay caused by a proposed lane closure, by comparing traffic demand for similar time intervals in the past with the reduction in throughput from the lane closure. So proposed lane closures may be shifted to time intervals to minimize the impact. A more ambitious scheme might involve including incentives in maintenance contracts that reflect these delays.

PeMS collects data on HoV lanes. These data can be used to determine the shift to car-pooling as a function of the congestion in the mainline lanes. If ramp-metering eliminates mainline congestion as suggested above, HoV lanes offer no advantage, and can be converted to mainline lanes, thereby increasing capacity. HoV bypass lanes at on-ramps can nonetheless encourage car-pooling.

General-purpose simulation models like Corsim and Paramics are typically used to answer a large number of "if-then" questions ranging from the effectiveness of ramp-metering schemes to the impact of traveler information. These models have many parameters and PeMS data may be used to calibrate those models. A more useful direction of research is the use of PeMS data to create a number of special-purpose statistical models. Statistical models of the impact of lane closures and HoV effectiveness are examples, as are the models for travel time prediction. Other models might estimate impact of incidents, weather, and special events. These special-purpose statistical models are easier to calibrate and maintain, and, where applicable, more reliable than the general-purpose simulation models.

Concluding remarks

Over the past 20 years, there has been a dramatic increase in productivity in the manufacturing sector. Much of this increase can be attributed to the integration of Information Technology (IT) into manufacturing. This has been a long process of learning, trial and error, seeking out opportunities for improvement. The process has been painful as Darwinian selection weeded out firms that did not successfully exploit the opportunities opened up by IT.

ITS initially was greeted as a set of technologies (AVL, CMS, ETC, VIPS, etc) that promised a quick path to productivity gains in transportation. A realistic assessment of actual productivity gains from the deployment of these technologies has not been made (there are several *un*realistic assessments), but an informed guess is that those gains are marginal. In retrospect this is hardly surprising. Like manufacturing, the production of transportation services is a highly complex operation, involving the orchestration of numerous interdependent activities, conventionally classified as operations, planning, and investment. Such highly complex systems do not admit quick technological fixes.

A wise traffic engineer remarked 40 years ago, "If you don't know how your system performed yesterday, you cannot expect to manage it today." A prerequisite to intelligent transportation systems is intelligence—the knowledge of what is happening to the system, an understanding of what decisions are effective, what key opportunities there are in the "value chain" that produces transportation services, and what technologies can help exploit them. Most transportation agencies operate without this intelligence.

It is instructive to think of the freeway system as an agency that consumes resources in order to produce a useful service. The service is transportation—the movement of vehicles carrying people and goods from one place to another. The service produced can be easily measured in vehicle-miles traveled or VMT. The resources consumed by the system are also easily measured. There are the fixed costs: depreciation of the freeway system's capital stock and the (essentially) fixed cost of the workforce that runs the system. There are also the variable costs: the time and money spent by people as they drive the vehicles over the freeways to go from one place to another. This is easily measured, too, as the vehicle-hours traveled or VHT. The ratio, Q = VMT/VHT, measures the freeway system's (marginal) productivity.

It is the common experience of drivers in California's urban areas that this productivity is declining. The decline would be much worse were it not for urban sprawl—people and businesses leave areas with very low productivity (this is an explanation of, not an argument for, sprawl); and because people adapt—they change their time of work, use their cell phones, listen to CDs, and give in to "road rage." Agencies that operate freeways usually form a monopoly (excepting a few private toll roads and alternative transit options), their frustrated customers have nowhere else to go, and so productivity continues to decline.

But the status quo can be changed. The first step is to equip system operators and their customers with intelligence about their system. PeMS shows that this step is easy. The next steps are difficult. They require carefully examining all the activities that affect the productivity measure Q, finding the changes in the activities that can lead to the greatest increase in Q, and implementing and monitoring those changes. Making routine the discovery and exploitation of opportunities needs organizational changes within the transportation agency. In the absence of any Darwinian mechanism that rewards those who carry out the required changes, these changes require inspired leadership.

Acknowledgements

This paper reports work of the PeMS Development Group, particularly Chao Chen, Zhanfeng Jia, Karl Petty, Jun Shu, Alex Skabardonis, and Erik van Zwet. I am also grateful to Hamed Benouar, Tom Choe, Ali Grabell, Joe Hecker, Randell Iwasaki, and John Wolf of Caltrans for their interest, support, and criticism of the work of PeMS. Elizabeth Stoltzfus and Chris Willige s of Booz.Allen and Hamilton have consistently provided thoughtful comment. Patrick Conroy and Steve Shladover of California PATH program have helped with their encouragement. The research summarized here is supported by grants from Caltrans and the National Science Foundation. The opinions expressed here and the errors of interpretation are those of the author alone.

References

1. C. Chan, K. Petty, A. Skabardonis, P. Varaiya, Z. Jia, "Freeway Performance Measurement System: Mining Loop Detector Data," 80th Annual Meeting of the Transportation Research Board, Washington, D.C., Jan 2001.

2. Z. Jia, C. Chen, B. Coifman, P. Varaiya, "The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors," IEEE 4th International ITS Conference, Oakland CA, August 2001.

3. Z. Jia, P.Varaiya, C. Chao, K. Petty, A. Skabardonis, "Maximum throughput in LA freeways occurs at 60 mph v.4," http://transacct.EECS.Berkeley.EDU/info.phtml, January 2001.

4. Z. Jia, P. Varaiya, C. Chen, K. Petty, A. Skabardonis, "Congestion, excess demand, and effective capacity in California freeways," http://transacct.EECS.Berkeley.EDU/info.phtml, December 2000.

5. J. Rice and E. van Zwet, "A simple and effective method for predicting travel times on freeways," IEEE 4th International ITS Conference, Oakland CA, August 2001.

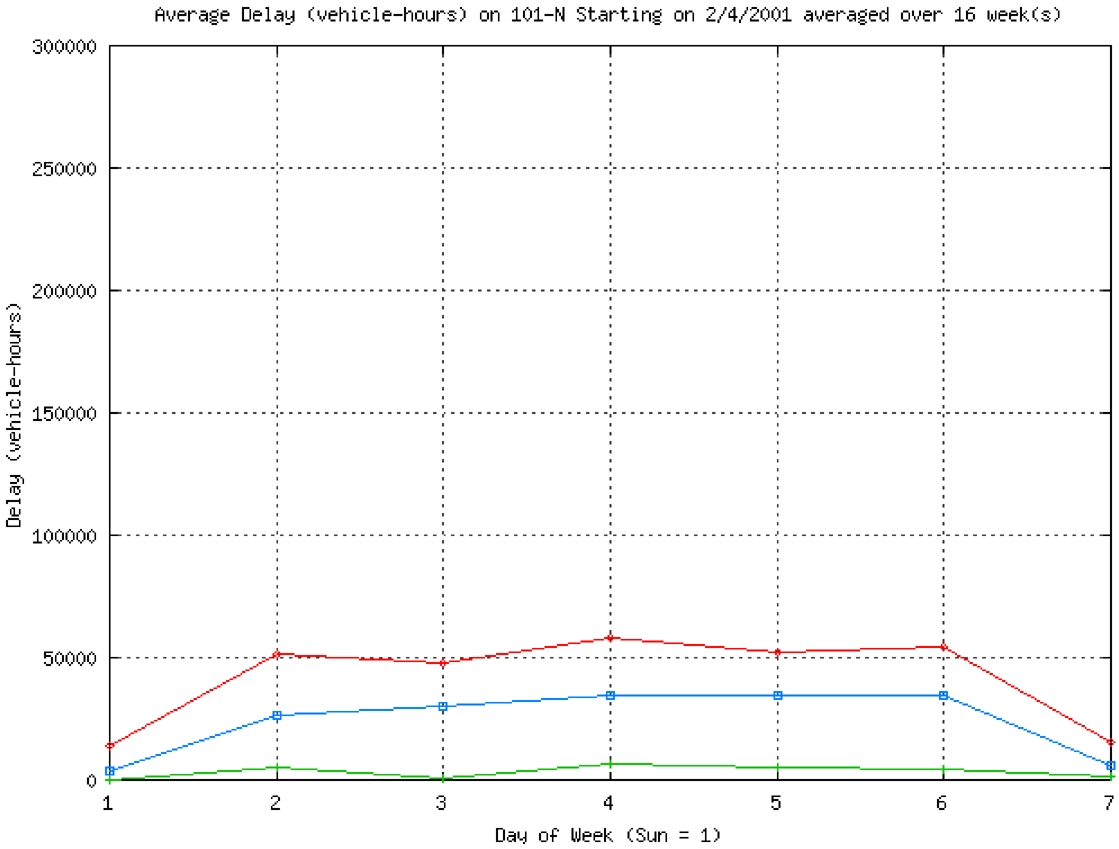

Figure 1 Maximum, minimum, and average delay over US 101-N by day of week

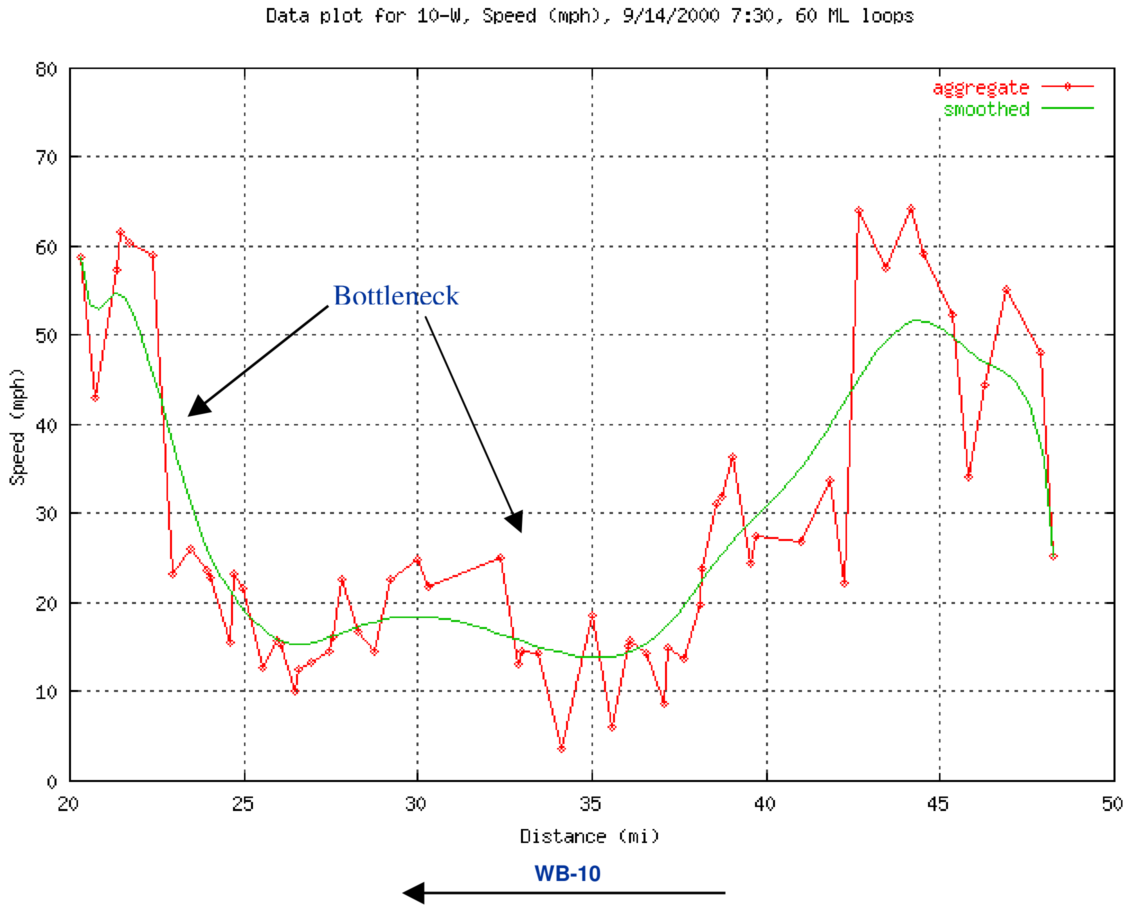

Figure 2 Speed across a 30-mile stretch of I-10W on September 14, 2000 at 7.30 am

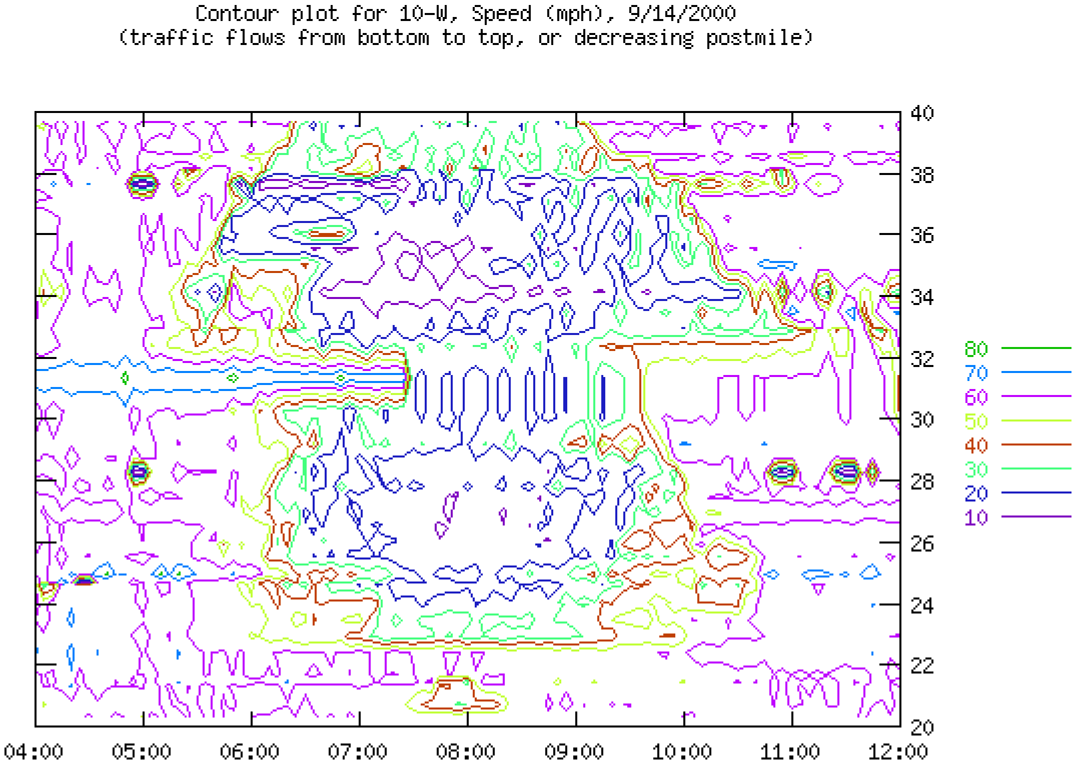

Figure 3 Contour plot of speed on I-10W from 4.00 am to noon, September 14, 2000

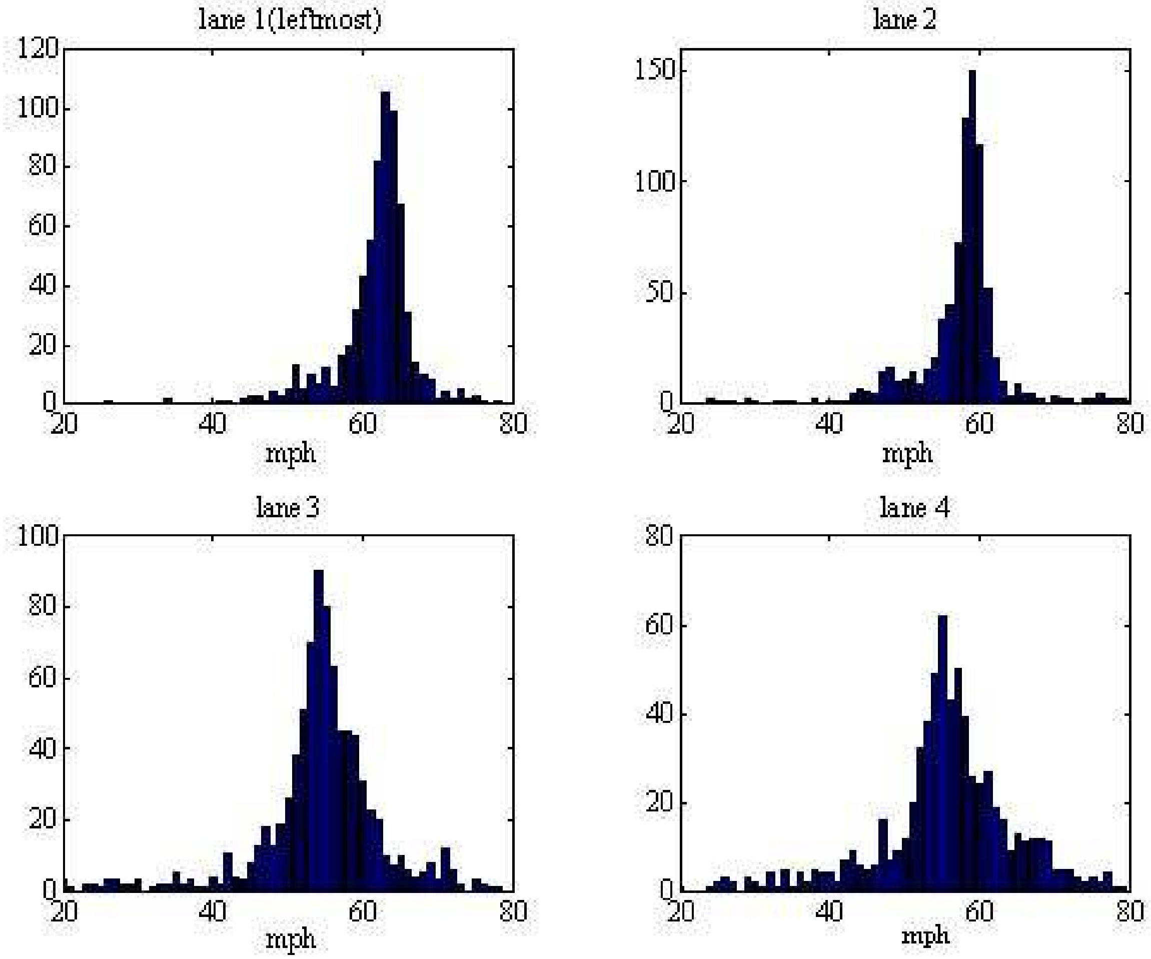

Figure 4 Distribution by lane of detector speed at time of maximum flow

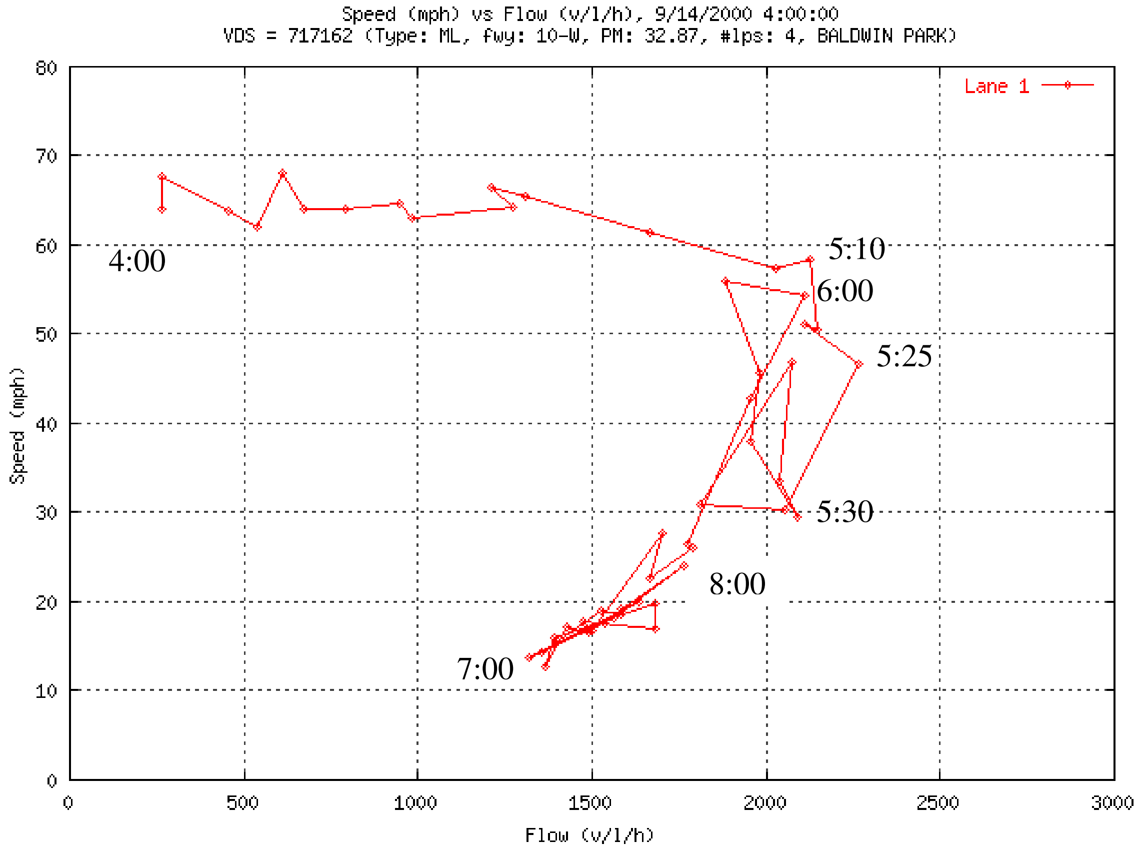

Figure 5 Speed-flow relationship between 4.00 and 7.00 am

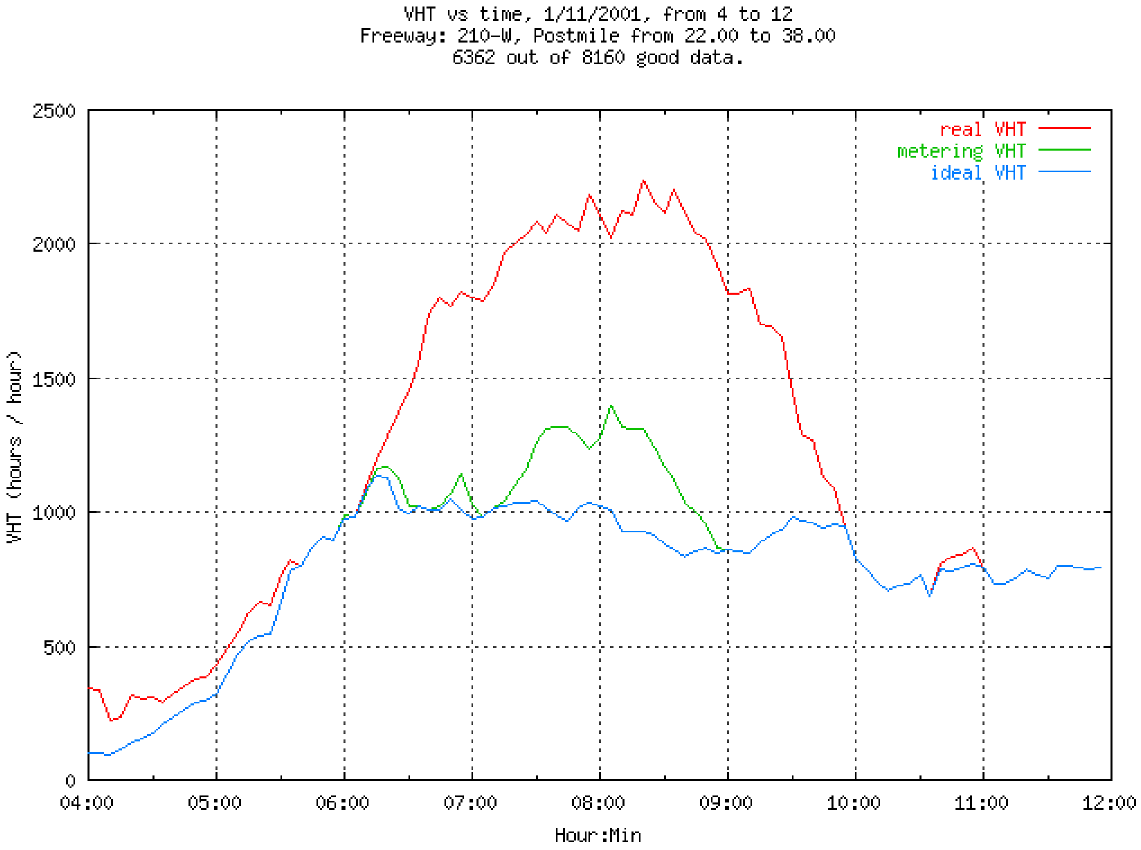

Figure 6 Calculation of potential reduction in congestion delay from ramp-metering

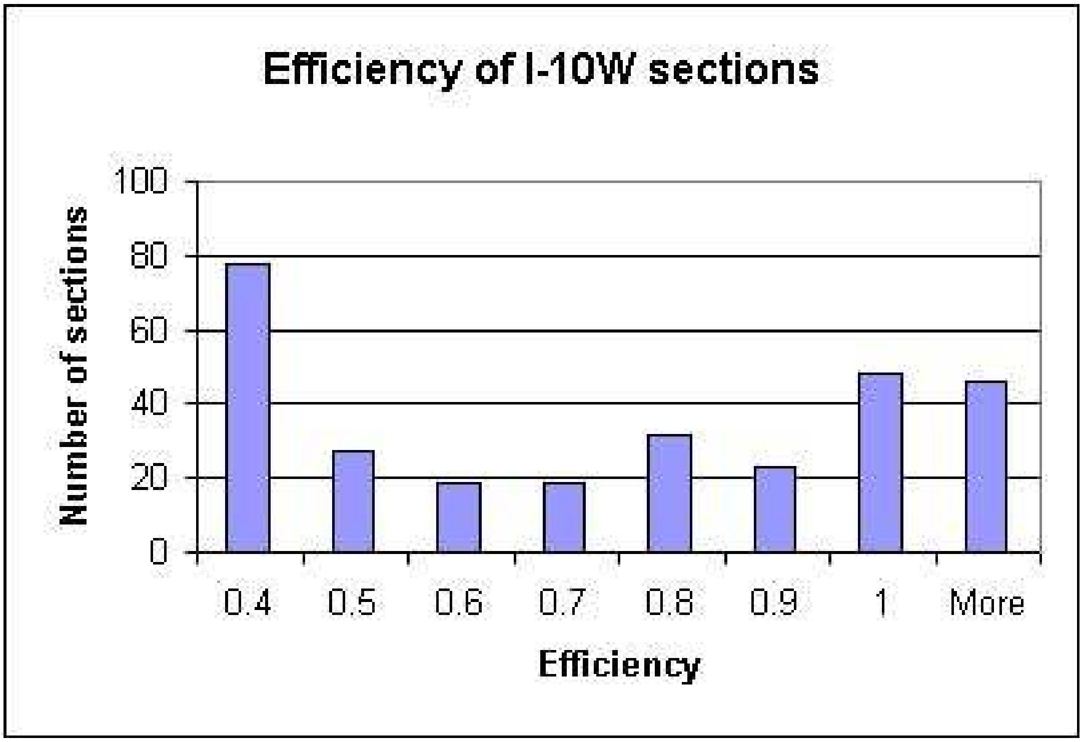

Figure 7 Variation in efficiency during worst congestion along segments of I-10W, midnight-noon, October 1, 2000

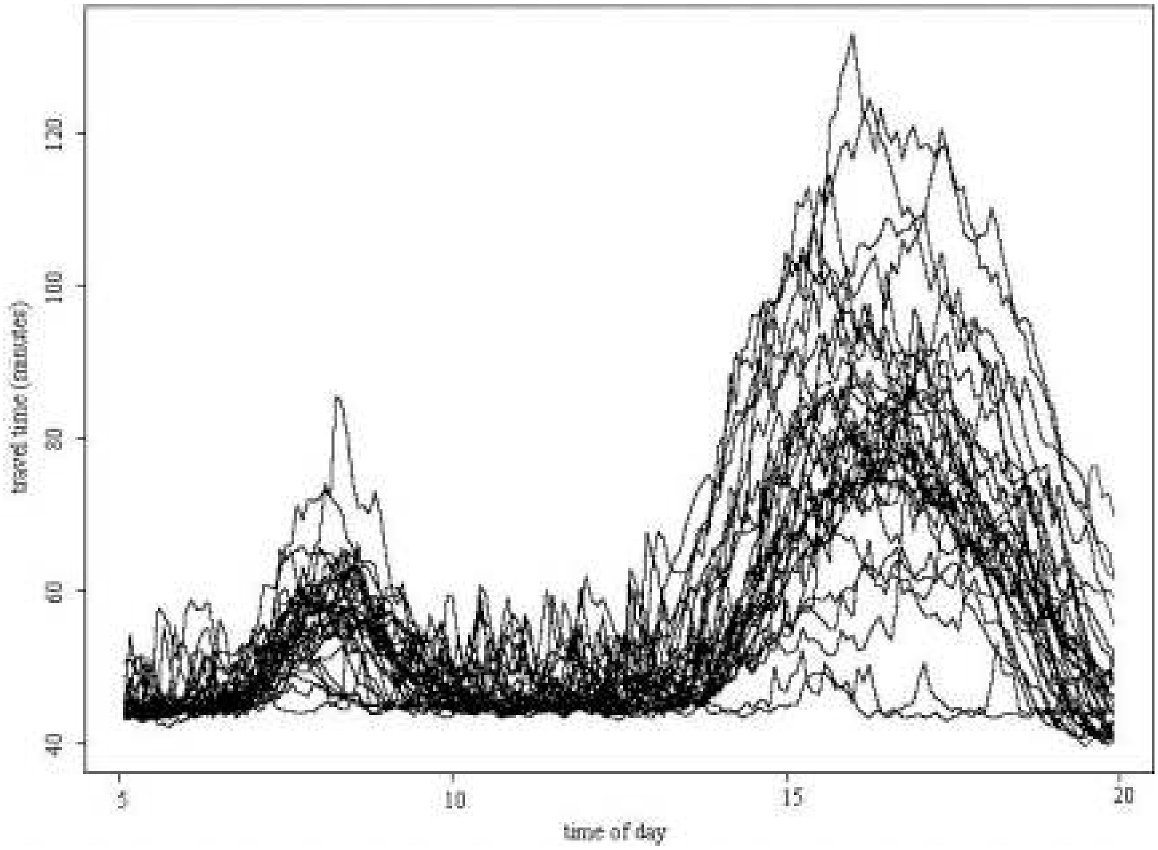

Figure 8 Travel time for I-10E between post miles 1.3 and 48.5 for 20 working days in October, 2000 for different starting times between 5 am and 8 pm

A Simple and Effective Method for Predicting Travel Times on Freeways

John Rice, Erik van Zwet

Abstract— **We present a method to predict the time that will be needed to traverse a certain stretch of freeway when departure is at a certain time in the future. The prediction is done on the basis of the current traffic situation in combination with historical data.**

We argue that, for our purpose, the current situation of a stretch of freeway is well summarized by the 'current status travel time'. This is the travel time that would result if one were to depart immediately and no significant changes in the traffic would occur. This current status travel time can be estimated from single or double loop detectors, video data, probe vehicles or by any other means.

Our prediction method arises from the empirical fact that there exists a linear relationship between any future travel time and the current status travel time. The slope and intercept of this relationship is observed to change subject to the time of day and the time until departure. This naturally leads to a prediction scheme by means of linear regression with time varying coefficients.

Keywords— **Prediction, Travel Time, Linear Regression, Varying Coefficients.**

I. Introduction

THE Performance Measurement System (PeMS) [1] is a large-scale freeway data collection, storage and analysis project. It involves the departments of EECS and Statistics and the Institute of Transportation Studies at the University of California, Berkeley in cooperation with the California Department of Transportation (Caltrans). PeMS' goal is to facilitate traffic research and to assist Caltrans by quantifying the performance of California's freeways. Useful information in various forms is to be distributed among traffic managers, planners and engineers, freeway users and researchers. In real time, PeMS obtains loop detector data on flow (count) and occupancy, aggregated over 30 second intervals. For all of California this amounts to 2 gigabytes per day. In its raw form, this data is of little use.

In this paper we focus our attention on travel time prediction between any two points of a freeway network for any future departure time. Besides being useful per se, travel time prediction serves as input to dynamic route guidance, congestion management, optimal routing and incident detection.

We are currently developing an Internet application which will give the commuters of Caltrans District 7 (Los Angeles) the opportunity to query the prediction algorithm which is described in this paper. The user will access our Internet site and state origin, destination and time of departure (or desired time of arrival). He or she will then receive a prediction of the travel time and the best (fastest) route to take.

In section II we state the exact nature of our prediction problem. We then describe our new prediction method ("linear regression") and two alternative methods which will be used for comparison. This comparison is made in section III with a collection of 34 days of traffic data from a 48 mile stretch of I-10 East in Los Angeles, CA. Finally, in section IV, we summarize our conclusions, point out some practical observations and briefly discuss several extensions of our new method.

II. Methods of Prediction

Consider a matrix V with entries $V(d,l,t)$ ($d \in D,\ l \in L,\ t \in T$) denoting the velocity that was measured on day d at loop l at time t. From V we can compute travel times $TT_d(a,b,t)$, for all $d \in D$, $a,b \in L$ and $t \in T$. This travel time is to approximate the time it took to travel from loop a to loop b starting at time t on day d. We can also compute a proxy for these travel times which is defined by

$$T^*_d(a,b,t) = \sum_{i=a}^{b-1} \frac{2d_i}{V(d,i,t)+V(d,i+1,t)}, \qquad (1)$$

where d_i denotes the distance from loop i to loop $(i+1)$. We call T^* the current status travel time (a.k.a. the snap-shot or frozen field travel time). It is the travel time that would have resulted from departure from loop a at time t on day d when no significant changes in traffic occurred until loop b was reached. It is important to notice that the computation of $T^*_d(a,b,t)$ only requires information available at time t, whereas computation of $TT_d(a,b,t)$ requires information of later times.

We fix an origin and destination of our travels and drop the arguments a and b from our notation. We now formally state the problem that is addressed in this paper.

Suppose we have observed $V(d,l,t)$ for a number of days $d \in D$ in the past. Suppose a new day e has begun and we have observed $V(e,l,t)$ at times $t \leq \tau$. We call τ the 'current time'. Or aim is to predict $TT_e(\tau+\delta)$ for a given (nonnegative) 'lag' δ. This is the time a trip that departs from a at time $\tau+\delta$ will take to reach b. Note that even for $\delta = 0$ this is not trivial.

Define the historical mean travel time as

$$\mu_{TT}(t) = \frac{1}{|D|} \sum_{d \in D} TT_d(t). \qquad (2)$$

Two naive predictors of $TT_e(\tau+\delta)$ are $T^*_e(\tau)$ and $\mu_{TT}(\tau+\delta)$. We expect—and indeed this is confirmed

Both authors are in the Department of Statistics of the University of California at Berkeley. Email: rice@stat.berkeley.edu and vanzwet@stat.berkeley.edu.

by experiment—that $T^*_e(\tau)$ predicts well for small δ and $\mu_{TT}(\tau+\delta)$ predicts better for large δ. We aim to improve on both these predictors for all δ.

A. Linear Regression

The main result of this paper is our discovery of an empirical fact: that there exist linear relationships between $T^*(t)$ and $TT(t+\delta)$ for all t and δ. This empirical finding has held up in all of numerous freeway segments in California that we have examined. This relation is illustrated by Figures 1 and 2, which are scatter plots of $T^*(t)$ versus $TT(t+\delta)$ for a 48 mile stretch of I-10 East in Los Angeles. Note that the relation varies with the choice of t and δ. With this in mind we propose the following model

$$TT(t+\delta) = \alpha(t,\delta)T^*(t) + \beta(t,\delta) + \varepsilon. \tag{3}$$

where ε is a zero mean random variable modeling random fluctuations and measurement errors. Note that the parameters α and β are allowed to vary with t and δ. Linear models with varying parameters are discussed by Hastie and Tibshirani in [2].

Fitting the model to our data is a familiar linear regression problem which we solve by weighted least squares. Define the pair $(\hat{\alpha}(t,\delta), (\hat{\beta}(t,\delta))$ to minimize

$$\sum_{\substack{d\in D\\ s\in T}} (TT_d(s) - \alpha(t,\delta) + \beta(t,\delta)T^*_d(t))^2 K(t+\delta-s), \tag{4}$$

where K denotes the Gaussian density with mean zero and a certain variance which the user needs to specify. The purpose of this weight function is to impose smoothness on α and β as functions of t and δ. We assume that α and β are smooth in t and δ because we expect that average properties of the traffic do not change abruptly. The actual prediction of $TT_e(\tau+\delta)$ becomes

$$\widehat{TT}_e^{\alpha\beta}(\tau+\delta) = \hat{\alpha}(\tau,\delta)T^*_e(\tau) + \hat{\beta}(\tau,\delta). \tag{5}$$

Writing $\beta(t,\delta) = \beta'(t,\delta)\mu_{TT}(t+\delta)$ we see that (3) expresses a future travel time as a linear combination of the historical mean and the current status travel time—our two naive predictors. Hence our new predictor may be interpreted as the best linear combination of our naive predictors. From this point of view, we can expect our predictor to do better than both. In fact, it does, as is demonstrated in section III.

Another way to think about (3) is by remembering that the word "regression" arose from the phrase "regression to the mean." In our context, we would expect that if T^* is much larger than average—signifying severe congestion—then congestion will probably ease during the course of the trip. On the other hand, if T^* is much smaller than average, congestion is unusually light and the situation will probably worsen during the journey.

Besides comparing our predictor to the historical mean and the current status travel time, we subject it to a more competitive test. We consider two other predictors that may be expected to do well. One resulting from Principal Component analysis and one from the nearest neighbors principle. Next, we describe these two methods.

B. Principal Components

Our predictor $\widehat{TT}^{\alpha\beta}$ only uses information at one time point; the 'current time' τ. However, we do have information prior to that time. The following method attempts to exploit this by using the entire trajectories of TT_e and T^*_e which are known at time τ.

Formally, let us assume that the travel times on different days are independently and identically distributed and that for a given day d, $\{TT_d(t) : t \in T\}$ and $\{T^*_d(t) : t \in T\}$ are multivariate normal. We estimate the covariance of this multivariate normal distribution by retaining only a few of the largest eigenvalues in the singular value decomposition of the empirical covariance of $\{(TT_d(t), T^*_d(t)) : d \in D,\ t \in T\}$. How many of the eigenvalues are retained must be specified by the user. Define τ' to be the largest t such that $t + TT_e(t) \le \tau$. That is, τ' is the latest trip that we have seen completed before time τ. With the estimated covariance we can now compute the conditional expectation of $TT_e(\tau+\delta)$ given $\{TT_e(t) : t \le \tau'\}$ and $\{T^*_e(t) : t \le \tau\}$. This is a standard computation which is described, for instance, in [3]. The resulting predictor is $\widehat{TT}_e^{\mathrm{PC}}(\tau+\delta)$.

C. Nearest Neighbors

As an alternative to Principal Components, we now consider nearest neighbors, which is also an attempt to use information prior to the current time τ. Similar to Principal Components, it is a non–parametric method, but it makes fewer assumptions (such as joint normality) on the relation between T^* and TT.

Nearest neighbors aims to find that day in the past which is most similar to the present day in some appropriate sense. The remainder of that past day beyond time τ is then taken as a predictor of the remainder of the present day.

The trick with nearest neighbors is in finding a suitable distance m between days. We suggest two possible distances:

$$m(e,d) = \sum_{l\in L,\ t\le\tau} |V(e,l,t) - V(d,l,t)| \tag{6}$$

and

$$m(e,d) = \left(\sum_{t\le\tau} (T^*_e(t) - T^*_d(t))^2\right)^{1/2}. \tag{7}$$

Now, if day d' minimizes the distance to e among all $d \in D$, our prediction is

$$\widehat{TT}_e^{NN}(\tau+\delta) = TT_{d'}(\tau+\delta). \tag{8}$$

Sensible modifications of the method are 'windowed' nearest neighbors and k-nearest neighbors. Windowed-NN recognizes that not all information prior to τ is equally relevant. Choosing a 'window size' w it takes the above summation to range over all t between $\tau - w$ and τ. So-called

k-NN is basically a smoothing method, aimed at using more information than is present in just the single closest match. For some value of k, it finds the k closest days in D and bases a prediction on a (possibly weighted) combination of these. Alas, neither of these variants appear to significantly improve on the 'vanilla' $\widehat{TT}^{NN}$.

III. Results

We have gathered flow and occupancy data from 116 single loop detectors along 48 miles of I-10 East in Los Angeles (between postmiles 1.28 and 48.525). Measurements were done at 5 minute aggregation at times t ranging from 5 am to 9 pm for 34 weekdays between June 16 and September 8 2000. We have used the so-called g factor method to convert flow and occupancy to velocity using the well-known formula

$$\text{velocity} = g \times \frac{\text{flow}}{\text{occupancy}}.$$

Here g is the unknown average length of vehicles, which has to be estimated. There are two problems with this method. First, the sensitivity of loop detectors differ and this difference is incorporated into g. Secondly, the average length of vehicles varies during a day from trucks late at night to compacts during rush hour. To try to counter these problems, we have used separate g factors for different detectors and allowed them to vary with the time of day. We estimate these g factors during times of freeflow when the occupancy is below 15% and known freeflow speeds occur. During congestion, when occupancy is above 15%, we keep the g constant.

Another problem with computing the velocity field is that loop detectors often do not report correct values or do not report at all. Fortunately, the quality of our I-10 data is quite good and we have used simple interpolation to impute wrong or missing values. The resulting velocity field $V(d, l, t)$ is shown in figure 3 where day d is June 16. The horizontal streaks typically indicate detector malfunction.

From the velocities we computed travel times for trips starting between 5 am and 8 pm. Figure 4 shows these $TT_d(t)$ where time of day t is on the horizontal axis. Note the distinctive morning and afternoon congestions and the huge variability of travel times, especially during those periods. During afternoon rush hour we find travel times of 45 minutes to up to two hours. Included in the data are holidays July 3 and 4 which may readily be recognized by their very fast travel times.

We have estimated the root mean squared error of our various prediction methods for a number of 'current times' τ (τ =6am, 7am,...,7pm) and lags δ (δ =0 and sixty minutes). We did the estimation by leaving out one day at a time, performing the prediction for that day on the basis of the remaining other days and averaging the squared prediction errors.

The prediction methods all have parameters that must be specified by the user. For the regression method we have chosen the standard deviation of the Gaussian kernel K to be 10 minutes. For the Principal Components method we have chosen the number of eigenvalues retained to be 4. For the nearest neighbors method we have chosen distance function (7), a window w of 20 minutes and the number k of nearest neighbors to be 2.

Figures 5 and 7 show the estimated root mean squared (RMS) prediction error of the historical mean $\mu_{TT}(\tau + \delta)$, the current status predictor $T_e^*(\tau)$ and our regression predictor (5) for lag δ equal to 0 and 60 minutes, respectively. Note how $T_e^*(\tau)$ performs well for small δ ($\delta = 0$) and how the historical mean does not become worse as δ increases. Most importantly, however, notice how the regression predictor beats both hands down.

Figures 6 and 8 again show the RMS prediction error of the regression estimator. This time, it is compared to the Principal Components predictor and the nearest neighbors predictor (8). Again, the regression predictor comes out on top, although the nearest neighbors predictor shows comparable performance.

The RMS error of the regression predictor stays below 10 minutes even when predicting an hour ahead. We feel that this is impressive for a trip of 48 miles right through the heart of L.A. during rush hour.

IV. Conclusions and loose ends

We stated that the main contribution of this paper is the discovery of a linear relation between $T^*(t)$ and $TT(t + \delta)$. But there is more. Comparison of the regression predictor to the Principal Components and nearest neighbors predictors unearthed another surprise. Given $T^*(\tau)$, there is not much information left in the earlier $T^*(t)$ ($t < \tau$) that is useful for predicting $TT(\tau + \delta)$. In fact, we have come to believe that for the purpose of predicting travel times all the information in $\{V(l, t),\ l \in L,\ t \leq \tau\}$ is well summarized by one single number: $T^*(\tau)$.

It is of practical importance to note that our prediction can be performed in real time. Computation of the parameters $\hat{\alpha}$ and $\hat{\beta}$ is time consuming but it can be done off-line in reasonable time. The actual prediction is trivial. As this paper is submitted, we are in the process of making our travel time predictions and associated optimal routings available through the Internet for the network of freeways of California District 7 (Los Angeles). It would also be possible to make our service available for users of cellular telephones—and in fact we plan to do so in the near future.

It is also important to notice that our method does not rely on any particular form of data. In this paper we have used single loop detectors, but probe vehicles or video data can be used in place of loops, since all the method requires is current measurements of T^* and historical measurements of TT and T^*.

We conclude this paper by briefly pointing out two extensions of our prediction method.

1. For trips from a to c *via* b we have

$$T_d(a, c, t) = T_d(a, b, t) + T_d(b, c, T_d(a, b, t)). \qquad (9)$$

We have found that it is sometimes more practical or advantageous to predict the terms on the right hand side than to predict $T_d(a, c, t)$ directly. For instance, when predicting travel times across networks (graphs), we need only predict travel times for the edges and then use (9) to piece these together to obtain predictions for arbitrary routes.

2. We regressed the travel time $T_d(t+\delta)$ on the current status $T^*_d(t)$, where $T_d(t+\delta)$ is the travel time departing at time $t+\delta$. Now, define $S_d(t)$ to be the travel time *arriving* at time t on day d. Regressing $S_d(t+\delta)$ on $T^*_d(t)$ will allow us to make predictions on the travel time subject to *arrival* at time $t+\delta$. The user can thus ask what time he or she should depart in order to reach his or her intended destination at a desired time.

Acknowledgments

The PeMS project is supported by grants from the National Science Foundation and from Caltrans through the California PATH program. The authors would like to acknowledge the help from (alphabetically) Peter Bickel, Chao Chen, Jaimyoung Kwon and Xiaoyan Zhang and the other members of the PeMS project. Also, we express our thanks to the engineers of Caltrans Headquarters and districts 7 and 12.

References

[1] `http://transacct.EECS.Berkeley.EDU/Public/`

[2] T. Hastie and R. Tibshirani (1993). Varying Coefficient Models. *Journal of the Royal Statistical Society Series B*, **55(4)** pp. 757–796.

[3] K.V. Mardia, J.T. Kent and S.M. Bibby (1979). *Multivariate Analysis*, Academic Press, London.

John A. Rice was born in New York on June 14, 1944. He received his B.A. degree in mathematics from the University of North Carolina at Chapel Hill in 1966 and his Ph.D. in statistics in 1972 from the University of California, Berkeley. He was in the Department of Mathematics at the University of California, San Diego from 1973–1991 and since 1991 he has been at the University of California, Berkeley where is is professor of statistics. His research interests include applied and theoretical statistics. He is a member of the Institute of Mathematical Statistics, the American Statistical Association, and the International Statistical Institute.

Erik W. van Zwet was born in The Hague, The Netherlands on November 10, 1970. He received his *doctorandus* (M.Sc.) degree in mathematics from the University of Leiden in 1995 and his Ph.D. degree, also in mathematics, from the University of Utrecht in 1999. He is currently a post doctoral visitor at the Department of Statistics of the University of California at Berkeley.

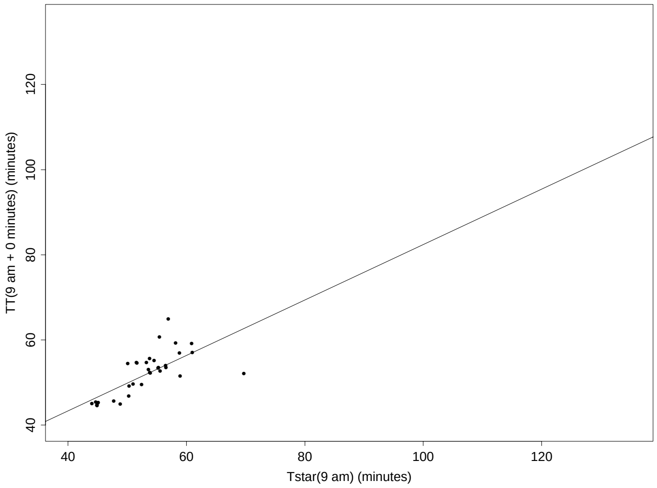

Fig. 1. $T^*(9\text{ am})$ vs. $TT(9\text{ am} + 0\text{ min's})$. Also shown is the regression line with slope $\alpha(9\text{ am}, 0\text{ min's})$=0.65 and intercept $\beta(9\text{ am}, 0\text{ min's})$=17.3.

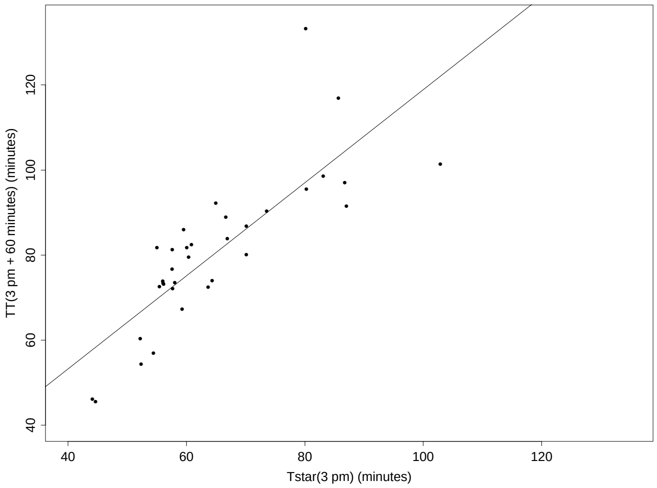

Fig. 2. $T^*(3\text{ pm})$ vs. $TT(3\text{ pm} + 60\text{ min's})$. Also shown is the regression line with slope $\alpha(3\text{ pm}, 60\text{ min's})$=1.1 and intercept $\beta(3\text{ pm}, 60\text{ min's})$=9.5.

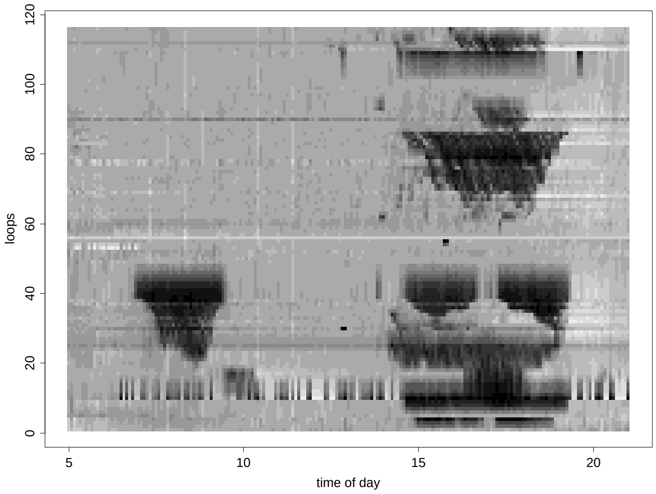

Fig. 3. Velocity field $V(d, l, t)$ where day $d =$ June 16, 2000. Darker shades refer to lower speeds. Note the typical triangular shapes indicating the morning and afternoon congestions building and easing. The horizontal streaks are most likely due to detector malfunction.

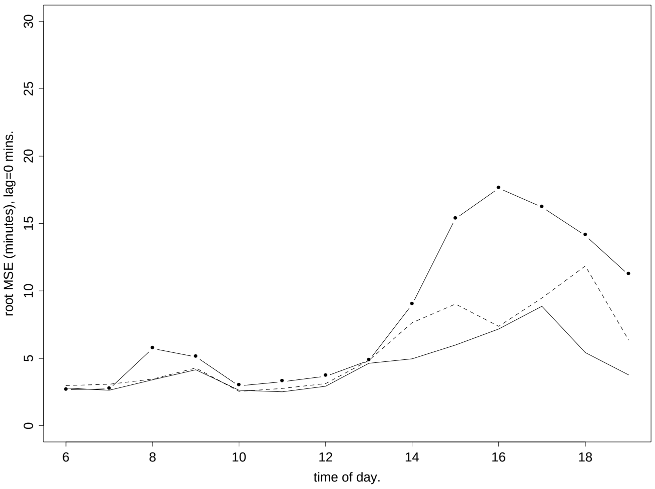

Fig. 5. Estimated RMSE, lag=0 minutes. Historical mean (– · –), current status (- - -) and linear regression (—).

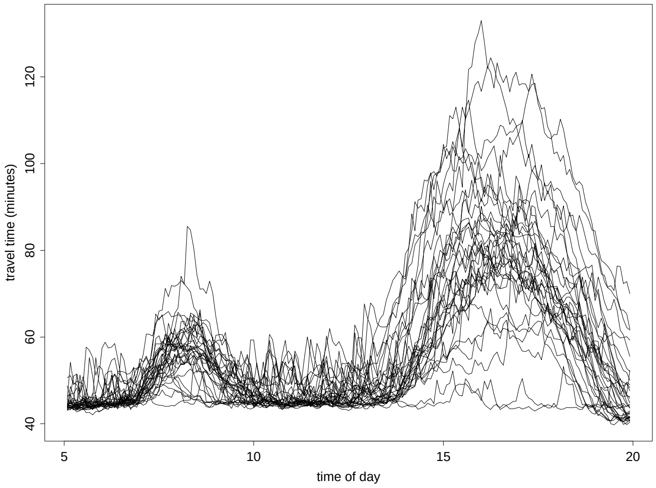

Fig. 4. Travel Times $TT_d(\cdot)$ for 34 days on a 48 mile stretch of I-10 East.

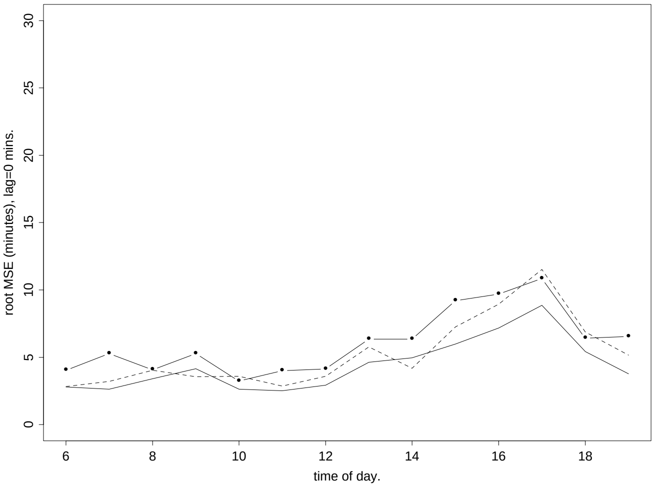

Fig. 6. Estimated RMSE, lag=0 minutes. Principal Components (– · –), nearest neighbors (- - -) and linear regression (—).

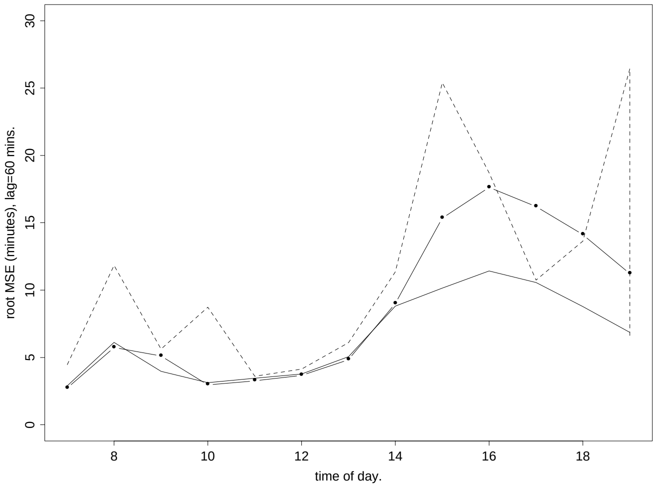

Fig. 7. Estimated RMSE, lag=60 minutes. Historical mean (– · –), current status (- - -) and linear regression (—).

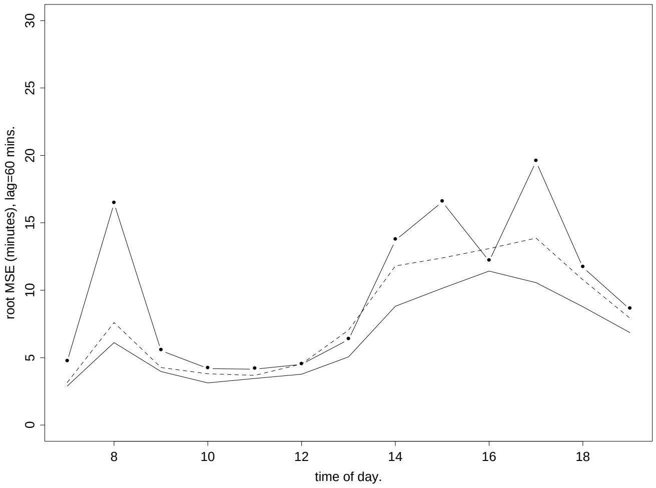

Fig. 8. Estimated RMSE, lag=60 minutes. Principal Components (– · –), nearest neighbors (- - -) and linear regression (—).

FREEWAY PERFORMANCE MEASUREMENT SYSTEM (PeMS):
AN OPERATIONAL ANALYSIS TOOL

Tom Choe
Office of Freeway Operations
California Department of Transportation District 7
120 S Spring Street
Los Angeles, CA 90012
Tel: (213) 897-0266, Fax: (213) 897-0894
tchoe@dot.ca.gov

Alexander Skabardonis*
Institute of Transportation Studies
University of California, Berkeley CA 94720-1720
Tel: (510) 642-9166, Fax: (510) 642-1246
dromeas@uclink4.berkeley.edu

Pravin Varaiya
Department of Electrical Engineering and Computer Science
University of California, Berkeley CA 94720
Tel: (510) 642-5270, Fax: (510) 642-6330
varaiya@eecs.berkeley.edu

For Presentation and Publication
81st Annual Meeting
Transportation Research Board
January 2002
Washington, D.C.

July 30, 2001

No WORDS: 3537
Plus 8 Figures (2000)
TOTAL: 5537
**Corresponding Author. Also in Transportation Research Record, 1811, 2002*

ABSTRACT

PeMS is a freeway performance measurement system for all of California. It processes 2 GB/day of 30-second loop detector data in real time to produce useful information. Managers at any time can have a uniform, and comprehensive assessment of freeway performance. Traffic engineers can base their operational decisions on knowledge of the current state of the freeway network. Planners can determine whether congestion bottlenecks can be alleviated by improving operations or by minor capital improvements. Travelers can obtain the current shortest route and travel time estimates. Researchers can validate their theory and calibrate simulation models.

The paper describes the use of PeMS in conducting operational analysis, planning and research studies. The advantages of PeMS over conventional study approaches is demonstrated from case studies on conducting freeway operational analyses, bottleneck identification, Level of Service determination, assessment of incident impacts, and evaluation of advanced control strategies

INTRODUCTION

Caltrans (California Department of Transportation) needs a freeway performance measurement system that extracts information from real time and historical data. ***PeMS*** (***Pe***rformance ***M***easurement ***S***ystem) is such a system. It presents information in various forms to assist managers, traffic engineers, planners, freeway users, researchers, and traveler information service providers (value added resellers or VARs).

Caltrans managers can instantaneously obtain a uniform, and comprehensive assessment of the performance of their freeways. Traffic engineers can base their operational decisions on knowledge of the current state of the freeway network. Planners can determine whether congestion bottlenecks can be alleviated by improving operations or by minor capital improvements. Traffic control equipment (ramp-metering and changeable message signs) can be optimally placed and evaluated. Travelers can obtain the current shortest route and travel time estimates. PeMS can serve to guide and assess deployment of intelligent transportation systems (ITS).

The purpose of this paper is to present the use of the PeMS as a tool to perform operations studies. The PeMS database and built-in applications offer several advantages in understanding the system performance and analyzing options compared to traditional approaches that are based on limited data due to the high effort and cost involved in field data collection. The use of PeMS maximizes the utility of data from loop detector surveillance systems that often are archived off-line without any processing and analysis.

The paper first gives a brief overview of the PeMS system. The following sections present the process and the findings from the application of PeMS by practicing engineers and researchers in conducting freeway operational analyses, bottleneck identification, determining the Level of Service, assessment of incident impacts, and evaluation of advanced control strategies. The last section summarizes the major study findings, and discusses ongoing and future work.

PeMS OVERVIEW

PeMS obtains 30 second loop detector data in real time from each Caltrans District Transportation Management Center (TMC). The data are transferred through the Caltrans wide area network (WAN) to which all districts are connected. Users can access PeMS over the Internet through a Web browser. The PeMS software architecture is modular and open. It uses commercial-of-the-shelf products for communication and computation. A brief overview of the system components is given below. These are described in detail elsewhere (*1*).

Data Processing

The 30 second data received by PeMS consist of counts (number of vehicles crossing the loop), and occupancy (the average fraction of time a vehicle is present over the loop). The software processes the data in real time and

- Aggregates 30-second values of counts and occupancy to lane-by-lane, 5-minute values;
- Calculates the *g*-factor of each loop;
- Uses the g-factor to calculate the speed for each lane;
- Aggregates the lane-by-lane value of flow, occupancy, and speed across all lanes at each detector station (one station typically serves the detectors in all the lanes at one location);

Most detectors in California have single loops. The g-factor (effective vehicle length) is used to calculate the average vehicle speeds from the flow and occupancy data. Typically, a constant value for the g-factor is used which leads to inaccurate speeds because the g factor varies by lane, time-of-day, as well as the loop sensitivity. PeMS uses an adaptive algorithm to compute the g-factor per each loop to provide accurate speed estimates. The algorithm has been tested and validated against "ground truth" data from double loop detectors and floating cars (*2*).

Calculation of Link Performance Measures

A link is defined as a freeway segment that contains a single loop detector (typical detector spacing is one-third to one-half mile). PeMS uses the 5 minute average values of flow and speed to compute the following performance measures: VMT (vehicle-miles traveled), VHT (vehicle-hours traveled), delay and travel time. Details of the computations are presented elsewhere (*3*).

Travel Time Estimation and Prediction

PeMS provides trip travel time estimates and shortest routes. You bring up the district freeway map on your Web browser, and select an origin and destination. PeMS displays 15 shortest routes, along with the estimates of the corresponding travel times.

PeMS also provides travel time predictions, for example, what will be the travel time 30 minutes from now. The travel time prediction algorithm combines historical and real time data (*4*). In addition, PeMS displays on the freeway system map the location and details about accidents and other incidents based on information retrieved in real-time from the California Highway Patrol (CHP) website. The PeMS data also can be transmitted to VARs, who provide traveler information services.

PEMS APPLICATION 1: FREEWAY OPERATIONAL ANALYSIS

PeMS was used by Caltrans staff to analyze existing operating conditions in the westbound direction of I-10 freeway during the am peak period. Figure 1 shows the study area. It extends from the Los Angeles/San Bernardino County line to Downtown Los Angeles, for a total length of 30 miles.

Figure 2 shows the volume and average speed for the test section at 7:30 am. A major bottleneck exists at the I-10/I-710 interchange. Figure 3 shows a three dimensional contour plot of speeds showing that congestion begins at about 6:30 am lasting until 11:00 am and extends to most of the study area.

The traditional approach to obtain performance data involves conducting floating car studies to obtain speed and delay data. This requires a minimum of two days field data collection with four person teams/segment driving instrumented vehicles in three 10-mile segments. This translates into 120 person hours. Additional field data collection to obtain statistically valid results is prohibitive due to the time and cost requirements. Further, several additional data need to be assembled including geometrics (aerial photos, as built plans), and traffic volumes (manual counts, historical data).

The use of PeMS brings several benefits. PeMS provides both the input (volumes) and performance data (speed, delay, VMT, VHT) for the study area. Contour and across space plots assist in determining problem locations and their impacts. More importantly, the data can be analyzed over several typical days. The entire analysis can be performed in less than one person day.

PEMS APPLICATION 2: BOTTLENECK IDENTIFICATION AND ANALYSIS

This example application involves the direct interaction with the PeMS database for customized applications. The objective is to identify where freeway bottlenecks are located, and assess their impacts. An important objective of this analysis is to determine if the bottleneck capacity could be preserved through traffic control measures (ramp metering).

The northbound direction of I-5 freeway in Los Angeles was analyzed. First, the PeMS built-in speed and occupancy contour plots were used to pinpoint bottleneck locations along the study section. Observations were performed for several weekdays. This preliminary analyses indicated that a potential bottleneck exists at postmile 29 (a weaving section). Figure 4 shows the average five-minute freeway occupancy at three loop detector locations for a four hour time period (2:00 to 6:00 pm). The loop occupancy at the bottleneck location (loop 716974) is about 11 percent. The downstream loop occupancy (loop 716978) is about 7 percent indicating free flow conditions. In contrast, the occupancy at the upstream loop (loop 71673) increases with time to about 25 percent from 4:00 to 6:00 pm, indicating congested conditions due to the presence of a downstream bottleneck.

Next, the 30 sec count and occupancy data for each detector were downloaded from the PeMS database and the results were analyzed in detail using cumulative count and occupancy plots (*5*). Figure 5 shows the plots for the upstream and downstream loop from the bottleneck. The values of counts and occupancy are appropriately scaled to remove stochastic fluctuations and reveal changes in traffic states. The plot for the downstream detector shows that the cumulative counts and occupancy track each other throughout the analysis period indicating free flow conditions. The opposite is true for the upstream loop.

At about 15:30 pm, the cumulative occupancy increases and the cumulative count decreases indicating congested conditions.

A broad based approach, another way to study bottleneck locations is by analyzing the speed contour maps. From Figure 3, we can easily identify potential bottleneck locations at postmiles 22 and 32 along the westbound I-10 corridor as the speeds are reduced from free flow conditions to virtually stop and go. Looking at a time slice between 7:30 am to 9:30 am, the second bottleneck at postmile 32 might not have been identified as it would be "buried" amongst the contour of congestion. With the speed contour maps, we can see the lengths of peak hours, formation and duration of bottlenecks, and indications of hidden bottlenecks. A key benefit of the PeMS is that this speed contour map is available for any time period, for any length of corridor 24 hours a day, 365 days a year. This allows engineers to study mid-day congestion periods, weekend peaks, holiday congestion, and alterations of traffic flow patterns due to extended construction road closures.

PeMS APPLICATION 3: LEVEL OF SERVICE (LOS) CHARACTERIZATION

The objective of this PeMS application was to determine the Level of Service (LOS) at several freeway segments per the Highway Capacity Manual--HCM2000 (*6*). Caltrans and the California Air Resources Board (ARB) are conducting chase car studies to derive speed correction factors to be incorporated into their emission factors for air quality analysis. This process involves recording vehicle speeds at selected freeway segments along with the prevailing operating conditions (LOS) as perceived by the observers. However, it is required to obtain LOS designations based on the actual operating conditions at each test segment.

The database included over 37 hours of chase car speed data collected in 250 segments in Los Angeles. The determination of the segments LOS using PeMS was done as follows:

- Match test segments with PeMS database
- Extract loop data (counts, speed, occupancies) from the loop detectors per segment
- Aggregate the data per segment per 15 minutes and compute the segment density (vpm/l)
- Determine the segment LOS per HCM2000 (basic freeway sections)

Currently, the effort is continuing using PeMS to identify and select test freeway sections with specific LOS to collect additional chase car data.

PeMS APPLICATION 4: INCIDENT IMPACTS

The PeMS database was used to analyze the impacts of a major incident in the eastbound direction of I-210 freeway (Figure 6). By utilizing the PeMS plots of speeds and volumes across space it was possible to determine the spatial and temporal impacts of the incident on the freeway, and the time for recovery to normal operating conditions.

Figure 6A shows the average speed vs. distance of 10 miles of freeway at 11:00 am. Traffic is free flowing at an average speed of about 60 mph. There are five through lanes on the freeway mainline until postmile 29, where they are reduced to four lanes. The traffic volume is about 6,000 vph (or 1500 vph/lane through the four lane section).

At 11:20 am a multi-vehicle collision occurred blocking three out of four travel lanes on the freeway. Figure 6B shows that the average speed drops to about 5 mph at the incident location. The incident lasted about 2.5 hours. Figure 6C shows the vehicle speeds at 2:00 pm shortly after the incident was cleared. The congestion has reached five miles upstream of the incident location. Normal operating conditions on the freeway resumed at 3:10 pm, 1.5 hrs following the incident removal (Figure 6D).

Further analysis of the PeMS data revealed the following regarding the incident impacts:

> **Remaining capacity**: The discharge rate of vehicles past the incident location on the single travel lane was very low (about 300 vph) the first 10 minutes of the incident. The discharge rate then increased to about 1,400 vph the rest of the incident duration. Assuming a typical capacity range of 8,000-8400 vph for the four lane section, the remaining capacity due to the incident is 17 percent of the capacity under normal conditions. This is higher than the suggested value of 13 percent reported in the HCM2000 (Chapter 25: Freeway Systems).
>
> **Discharge ("getaway" flow):** Following the incident clearance, it was observed that the queued vehicles discharged past the incident location at a rate of 7,400 vph, which is lower than the nominal capacity of the freeway section (8,000-8,400 vph).

PeMS APPLICATION 5: ASSESSMENT OF ATMIS STRATEGIES

Caltrans and other agencies nationwide have started to deploy Advanced Traffic Management and Information Systems (ATMIS) to manage freeway congestion. Examples include ramp metering, changeable message signs, and incident detection. The most important question to answer is: By how much can ATMIS reduce congestion? PeMS can help answer this question.

Congestion may be measured by Caltrans' definition of *delay (when freeway speeds fall below 35 mph)*, or by using VHT and VMT. We can use PeMS to analyze delay for any section of freeway, and the effectiveness of ramp metering. Figure 7 shows the results for a 6.3-mile section of I-405 from 5.00 to 10.00 am on 6/1/98.

The top curve in Figure 7 shows the actual VHT per 5 minutes on the study section from 5 to 10 am. The middle curve is the estimated VHT the *same* vehicles would spend if ideal ramp metering maintained throughput at capacity. This implies that a certain number of vehicles have to be stored at the freeway entrance ramps (excess demand). The lowest curve is the VHT that would result if demand-shift eliminated queues at ramps. The area between the top and middle curves is the delay that can be eliminated with ideal ramp metering (about 500 veh-hrs in this study section). The area between the middle and lowest curve is the delay due to the excess demand (about 200 veh-hrs). The total delay is then the area between the topmost and lowest curves. The delay due to the excess demand can only be reduced through temporal, spatial or modal demand-shifting. One way to shift demand is to use PeMS to inform travelers that they will face this delay. Travelers that are better off changing their trip departure time, route or travel mode would then do so.

DISCUSSION

Significant investments in ITS infrastructure are underway in California and in most metropolitan areas in the US to manage traffic congestion. Central to this infrastructure is a surveillance system that gathers real-time information from detectors on the state of the system and transmits to TMCs. However, in many cases the surveillance data are simply being archived and they are not analyzed to assist in operational analyses or calculate performance measures. PeMS is a unique data archival, processing and analysis system that allows access to the data and calculates performance measures.

PeMS is based on a modular and open software architecture. Currently, it stores data from over 4500 loop detectors in California. All the data are available on line. Users access PeMS through the Internet. PeMS is easy to use; built-in applications are accessed through a Web browser. Custom applications can work directly with the database. There is no need for special access to the agency's TMC data storage, writing scripts to access the database, or requesting off-site archived historical data.

The paper presented a number of case studies to demonstrate how PeMS can be used in operations, planning and research studies. PeMS brings large benefits. It maximizes the utility of the information from surveillance systems, and minimizes the effort and costs of performing studies through traditional data collection methods. Managers can instantaneously obtain a uniform, and comprehensive assessment of the performance of their freeways. Traffic engineers can base their operational decisions on knowledge of the current state of the freeway network, and determine whether bottlenecks can be alleviated by design or operational improvements. PeMS can serve to guide and assess deployment of ITS.

Caltrans engineers using PeMS are performing several other operations studies. These include analysis of peaking characteristics (midday and weekend peak), traffic patterns during special events, historical comparisons, planned lane closures, and CHP special programs on congestion management. Some of the ongoing collaborative work between Caltrans and the research team on PeMS includes:

Recurrent vs. non-recurrent congestion: Previous widely cited studies suggest that over 50 percent of freeway congestion is incident related (*7,8*). These studies are based on limited empirical data on freeway operating conditions and several simplifying assumptions on incident frequency, severity and impacts. Previous extensive field studies by the research team (*9,10*) developed a comprehensive data base on incident characteristics, and showed that only a fraction of the reported incidents causes delay. The example PeMS application described in the previous section demonstrated on how to evaluate the incident impacts based on real life data. Work is underway to estimate the amount of incident related vs. recurring congestion using the PeMS database and the incident data collected from the CHP.

Simulation models application, calibration and validation: simulation models are increasingly being used to analyze existing operations and evaluate the effectiveness of alternative scenarios. The critical questions regarding the practical application of simulation models are:

- Does the model accurately predict the existing bottleneck location(s) and impacts in the study section?
- Are the benefits from the simulated design/control improvements significant?

First, accurate input data are needed on design, demand and control characteristics, and proper calibration of the model parameters. The PeMS database provides the input data (traffic volumes, link lengths, number of lanes, detector locations) and the performance measures to assess the model accuracy in replicating existing conditions. Contour plots of speed and occupancy can be compared with the model results to verify that the model correctly identifies bottleneck locations. The PeMS flow/occupancy or speed/flow plots at selected locations can be used to adjust the model parameters (free flow speed, capacity or minimum headways) to better match field conditions. Other performance measures produced by PeMS (average speeds, delays, VMT or VHT) can be compared with the simulation outputs to assess the models' accuracy in replicating observed operating conditions.

The second question relates on how the variability in operating conditions affects the confidence of the predicted benefits from the testing of alternative scenarios (e.g., is the predicted 2 percent improvement significant?). This has not been investigated in past simulation studies, because the data were collected during a single day. PeMS can be used to readily assess the variability in the input data and performance measures for the study section over a long period of time. Figure 8 shows the distribution of travel times for a 22 mile section of I-405 in Orange County, California, during the am peak for a typical incident-free weekday during 1998. In this example, it would be really hard to determine that small predicted improvements are significant.

Currently, we are analyzing a section of a 10 mile section of I-210 freeway using the FREQ macroscopic simulation model (*11*), and a 15 mile section of I-10 freeway using the PARAMICS microscopic simulator (*12*). A key objective of this work is to automate the input/output data between PeMS and simulation models in order to minimize the effort for the models' application.

Performance Measures: It is widely recognized that planning and operational decisions should be based on the transportation system performance measures. PeMS' premise has been to produce performance measures that are based on real data that can be easily understood by the system manager (VMT, VHT) and the system user (travel time). PeMS also provides travel time reliability measures (Figure 8) which is perceived to be of high importance by travelers. Work is continuing to expand the built-in PeMS applications to produce performance measures for planning applications including areawide performance indicators and growth trends.

The usefulness of the PeMS system depends on the loop detector data accuracy. Detector failures result in lost and unreliable data. PeMS includes a set of diagnostics to check the incoming loop data for accuracy and reliability. It also provides information on the number of data samples received daily and spatial/temporal data dropouts. In addition, PeMS includes procedures to "fix" data holes at a location based on information from adjacent detector stations as appropriate. However, there is no substitute for accurate data and any agency installing and operating freeway surveillance systems which are primarily designed for real-time operating strategies must have a plan for intensive maintenance of the field and communications equipment.

ACKNOWLEDGEMENTS

The PeMS project is supported by grants from Caltrans to the California PATH Program, National Science Foundation Grant CMS-0095739 and EPRI/DoD Complex Interactive Networks Initiative Contract under WO8333-04. We are very grateful to engineers from Caltrans Districts 3, 4, 7, 8 and 12 and Headquarters for their encouragement, understanding, and patience. They continue to shape the evolution of PeMS vision, to monitor its progress, to champion PeMS within Caltrans. Without their generous support this project would not have reached such a mature stage.

The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views of or policy of the California Department of Transportation. This paper does not constitute a standard, specification or regulation.

REFERENCES

1. Chen, C., K. Petty, A. Skabardonis, P. Varaiya, and Z. Jia, "Freeway Performance Measurement System: Mining Loop Detector Data," paper 01-2354, presented at the 80th TRB Annual Meeting, Washington DC., January 2001 (forthcoming Transportation Research Record).

2. Jia, Z., et al, "The PeMS Algorithms for Accurate Real-Time Estimates of g-Factors and Speeds from Single Loop Detectors," 4th IEEE ITSC Conference, Oakland, CA, 2001.

3. PeMS Development Group, "PeMS: Calculations with Loop Detectors," PeMS Website http://transacct.eecs.berkeley.edu, 2001.

4. Rice, J., and E. Van Swet, "A Simple and Effective Method for Predicting Travel Times on Freeways, " PATH Working Paper, Institute of Transportation Studies, University of California, Berkeley, June 2001.

5. Cassidy, M.J., and R.L. Bertini, "Some Traffic Features at Freeway Bottlenecks," Transportation Research B, Vol 33B, pp.25-42, 1999.

6. Transportation Research Board, "Highway Capacity Manual," Washington, DC., 2000.

7. Lindley, J.A., 1986, "Qualification of Urban Freeway Congestion and Analysis of Remedial Measures, FHWA Report RD/87-052, Washington., D.C.

8. HICOMP Report, 1992, "Statewide Highway Congestion Monitoring Program," Caltrans, Division of Traffic Operations, Sacramento, CA.

9. Skabardonis, A., et al, "The I-880 Field Experiment: Database Development and Incident Delay Estimation Procedures," Transportation Research Record No, 1554, 1996.

10. Skabardonis, A., K. Petty, and P. Varaiya, "Evaluation of Freeway Service Patrol at a Los Angeles Freeway Site," PATH Research Report UCB-ITS-PRR-98-13, Institute of Transportation Studies, University of California, Berkeley, 1998.

11. May, A.D., et al, "Integrated system of freeway corridor simulation models," Transportation Research Record No. 1320, 1991.

12. Quadstone Ltd. *PARAMICS ModellerV3.0 User Guide and Reference Manual*. Edinburgh, February 2000

LIST OF FIGURES

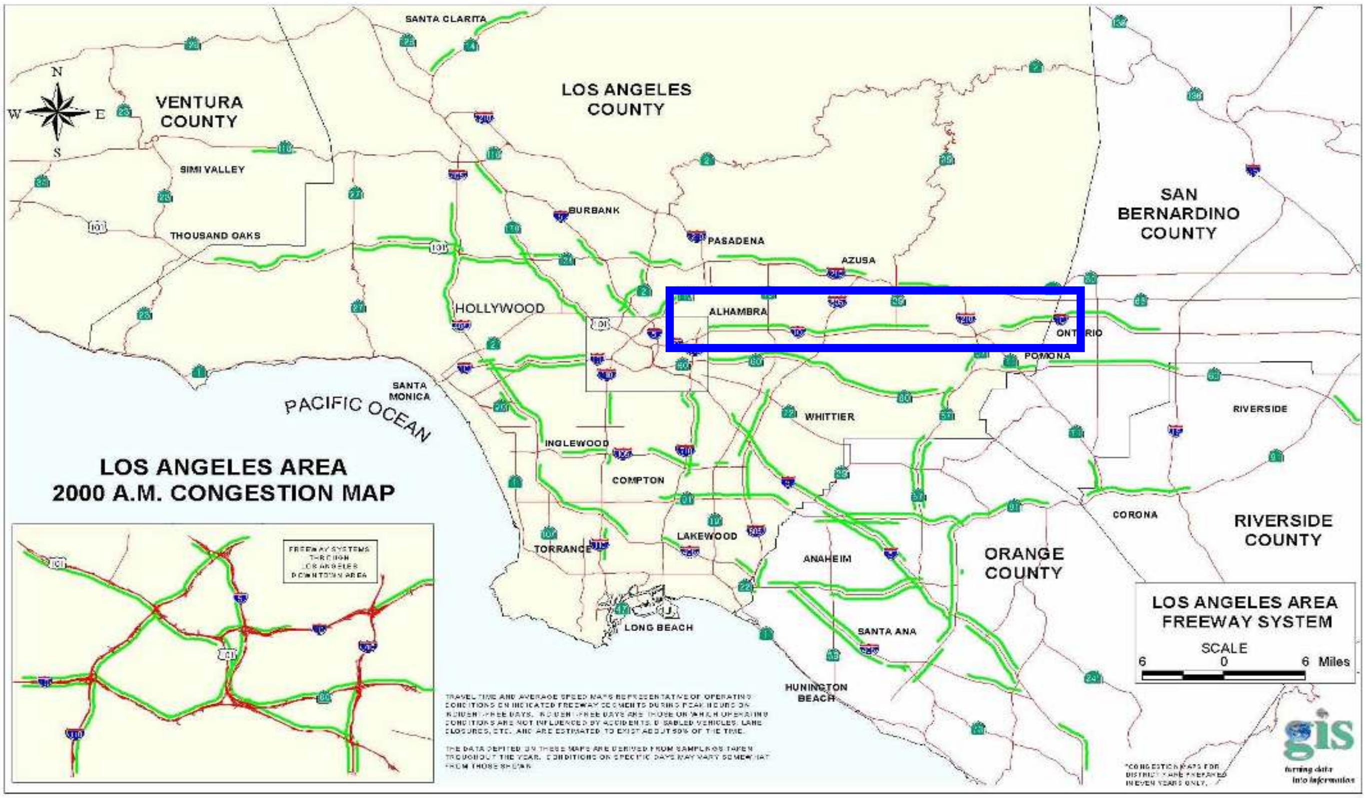
SANTA CLARITA
N
W
E
S
VENTURA COUNTY
LOS ANGELES COUNTY
SIMI VALLEY
THOUSAND OAKS
BURBANK
PASADENA
AZUSA
SAN BERNARDINO COUNTY
HOLLYWOOD
ALHAMBRA
ONTARIO
POMONA
SANTA MONICA
PACIFIC OCEAN
WHITTIER
RIVERSIDE
INGLEWOOD
COMPTON
LOS ANGELES AREA
2000 A.M. CONGESTION MAP
CORONA
RIVERSIDE COUNTY
LAKEWOOD
TORRANCE
ANAHEIM
ORANGE COUNTY
FREEWAY SYSTEMS THROUGH LOS ANGELES DOWNTOWN AREA
LONG BEACH
SANTA ANA
LOS ANGELES AREA FREEWAY SYSTEM
SCALE
6
0
6 Miles
HUNINGTON BEACH
TRAVEL TIME AND AVERAGE SPEED MAPS REPRESENTATIVE OF OPERATING CONDITIONS ON INDICATED FREEWAY SEGMENTS DURING PEAK HOURS ON INCIDENT-FREE DAYS. INCIDENT-FREE DAYS ARE THOSE ON WHICH OPERATING CONDITIONS ARE NOT INFLUENCED BY ACCIDENTS, DISABLED VEHICLES, LANE CLOSURES, ETC. AND ARE ESTIMATED TO EXIST ABOUT 50% OF THE TIME.
THE DATA DEPITED ON THESE MAPS ARE DERIVED FROM SAMPLINGS TAKEN THROUGHOUT THE YEAR. CONDITIONS ON SPECIFIC DAYS MAY VARY SOMEWHAT FROM THOSE SHOWN.
*CONGESTION MAPS FOR DISTRICT 7 ARE PREPARED IN EVEN YEARS ONLY.
gis
turning data into information

Figure 1. The Test Section (WB I-10)

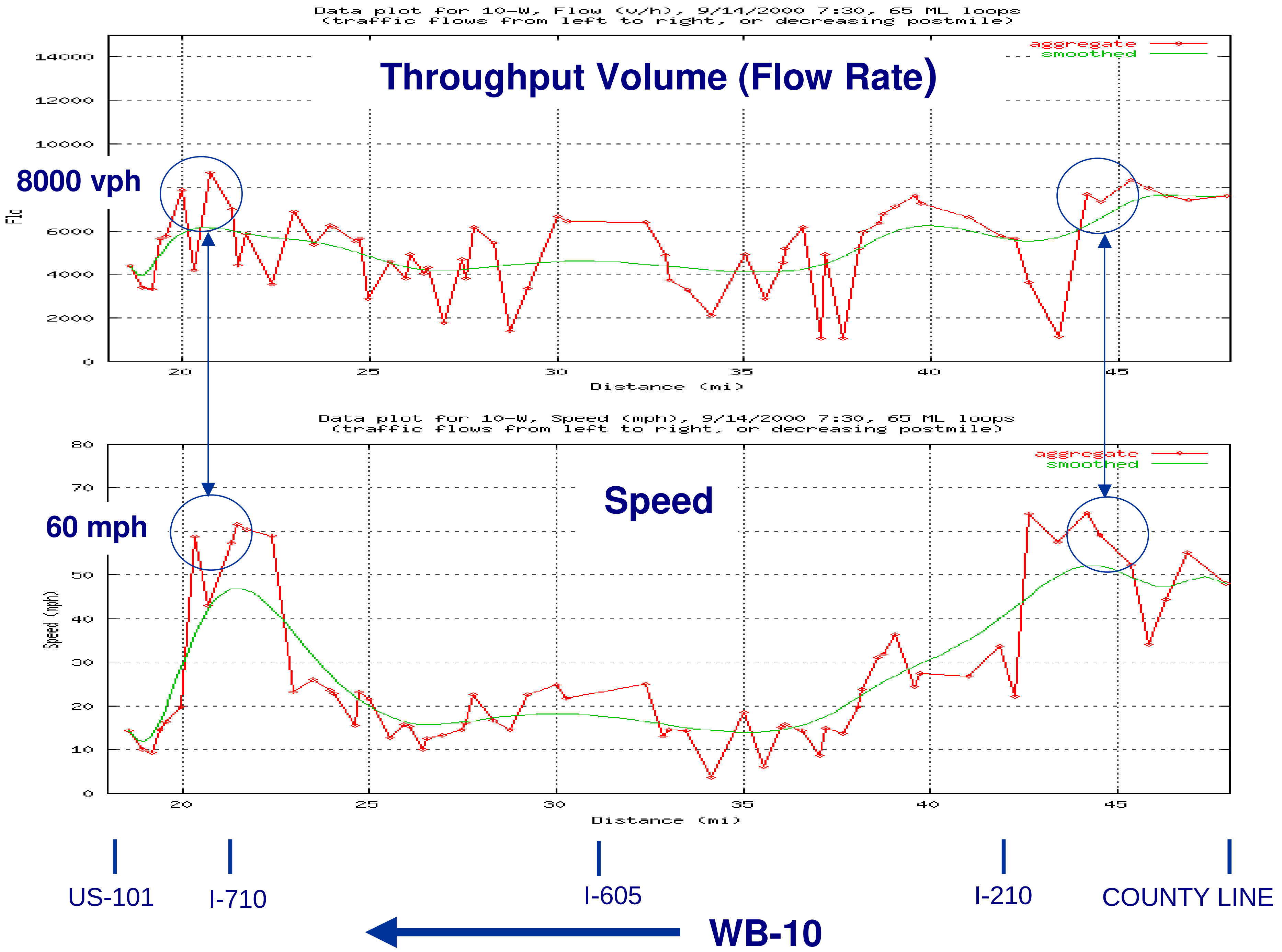

Figure 2. Flow and Speed along WB I-10, 7:30 am, 9/14/2000

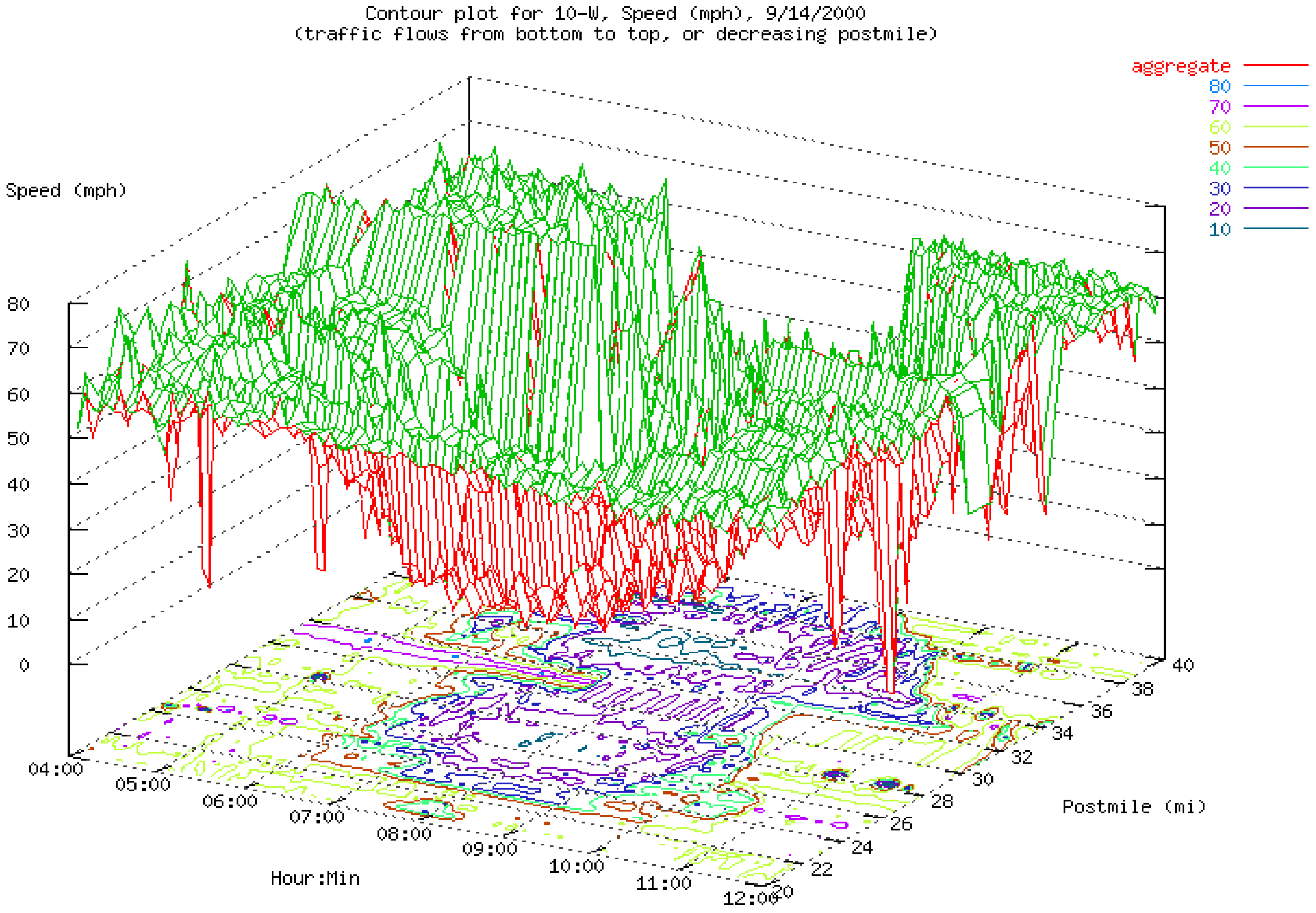

Figure 3. Speed Contour Plot (WB I-10)

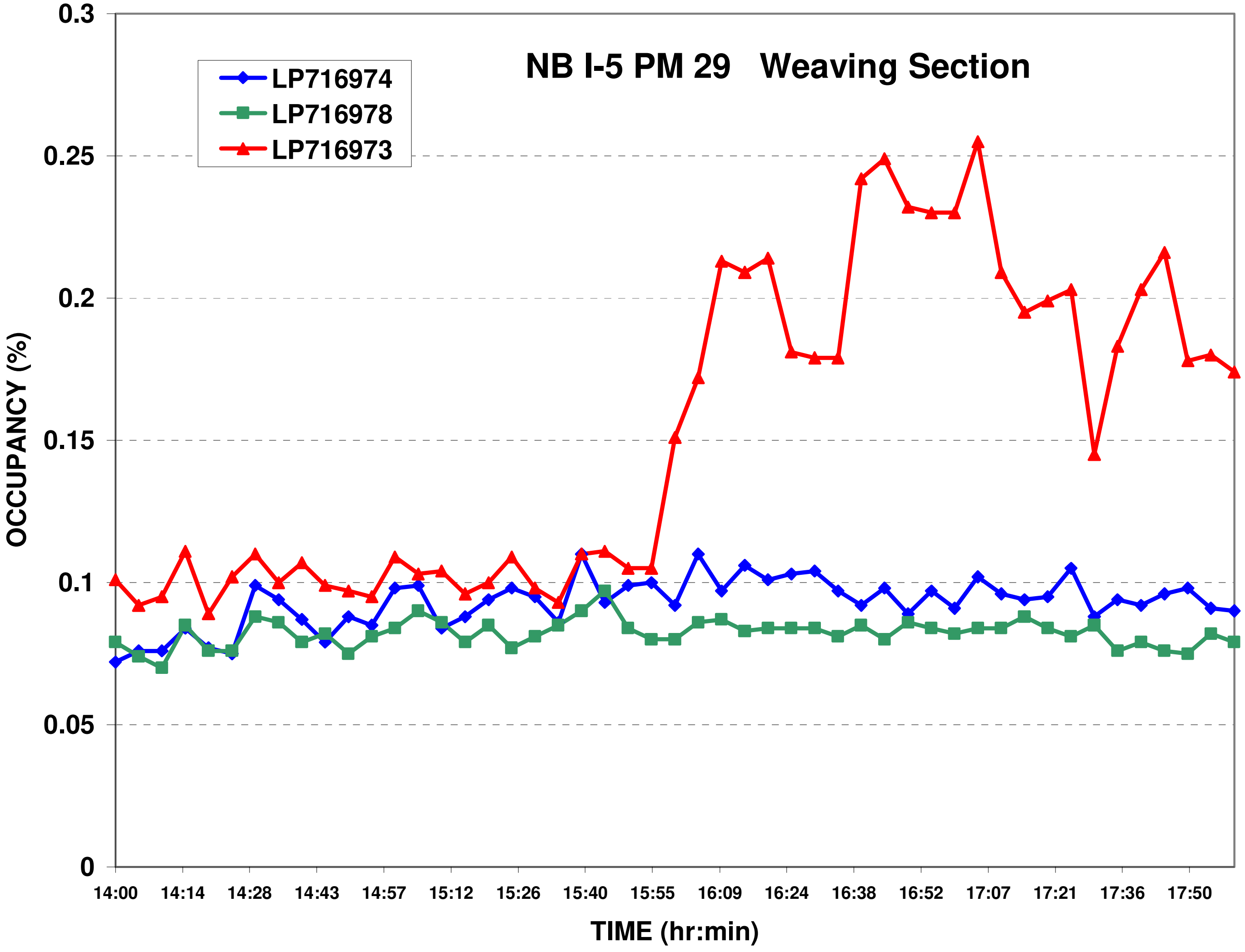

Figure 4. Detector Occupancy vs. Time

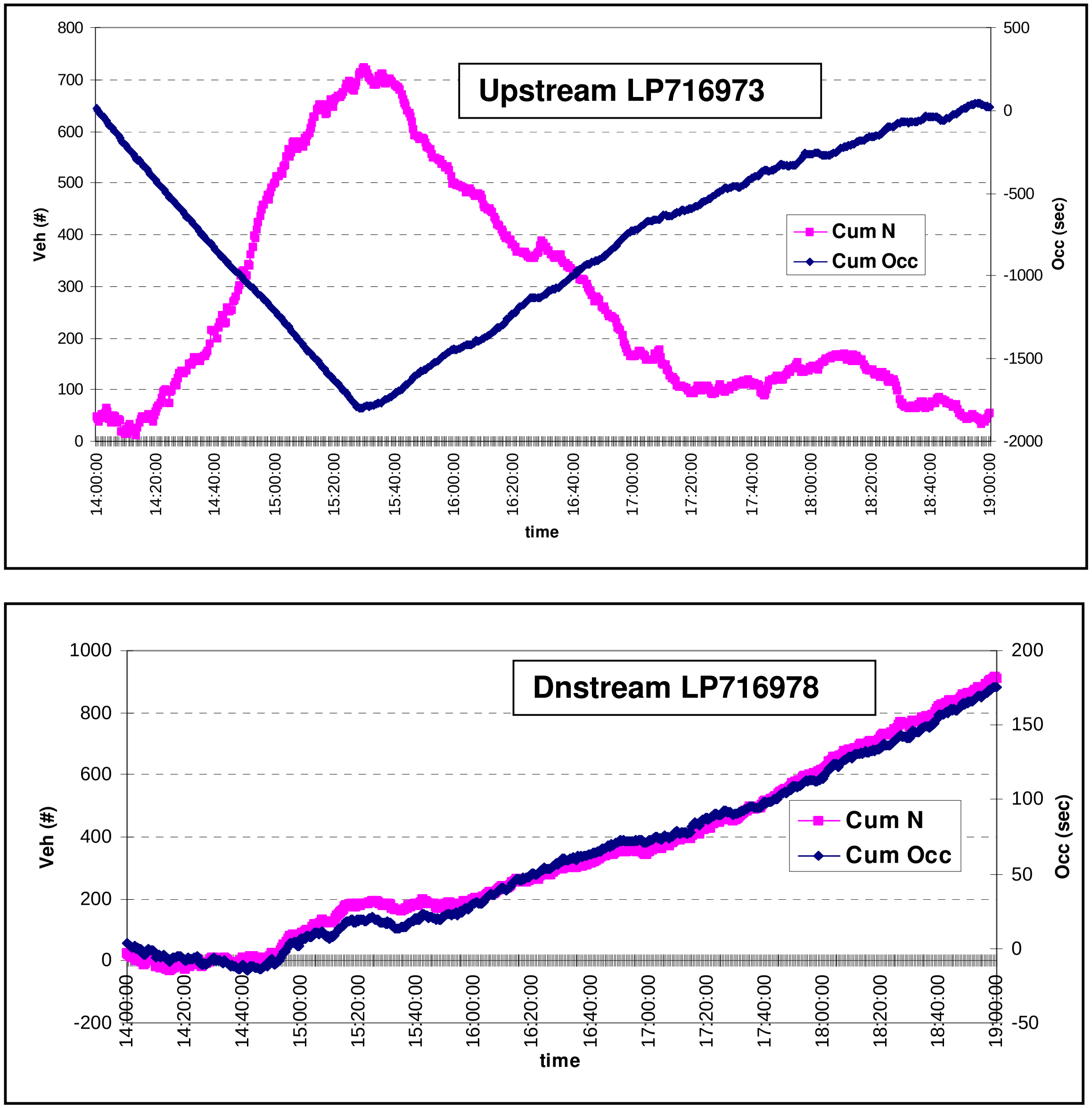

Figure 5. Cumulative Count and Occupancy Plots

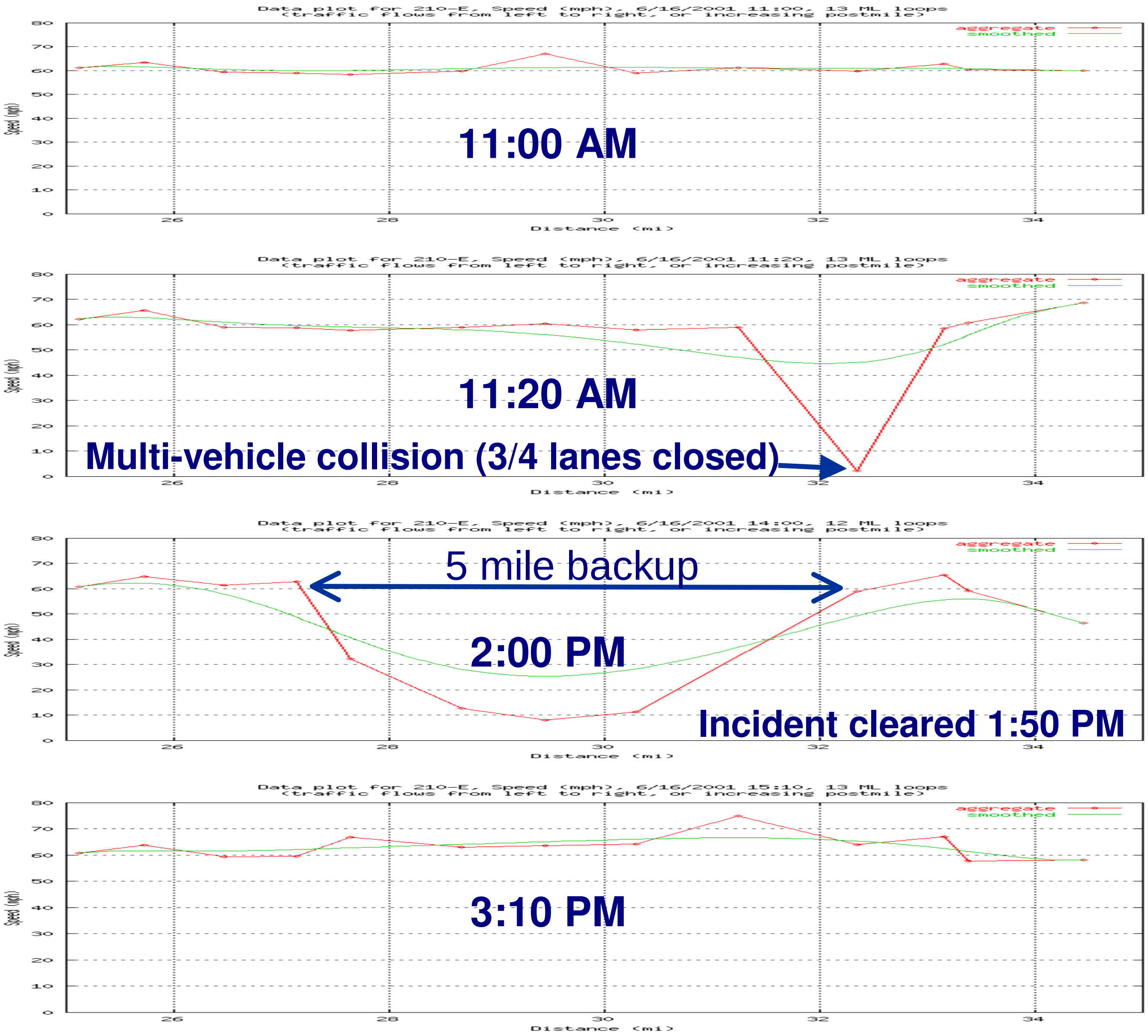

Figure 6. Incident Impacts: EB I-210 Saturday 6/16/2001

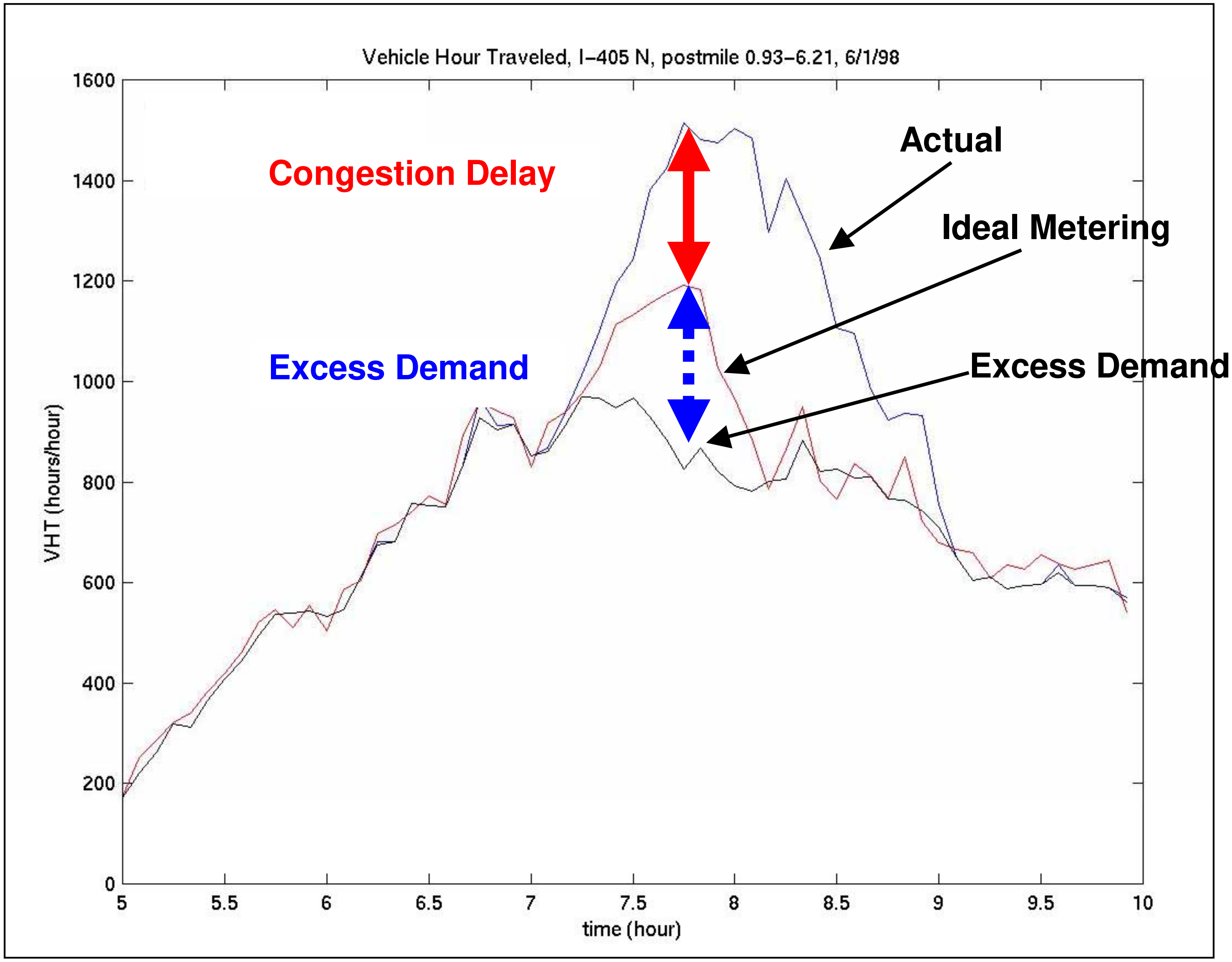

Figure 7. Veh-Hours Traveled (VHT) Under Different Scenarios

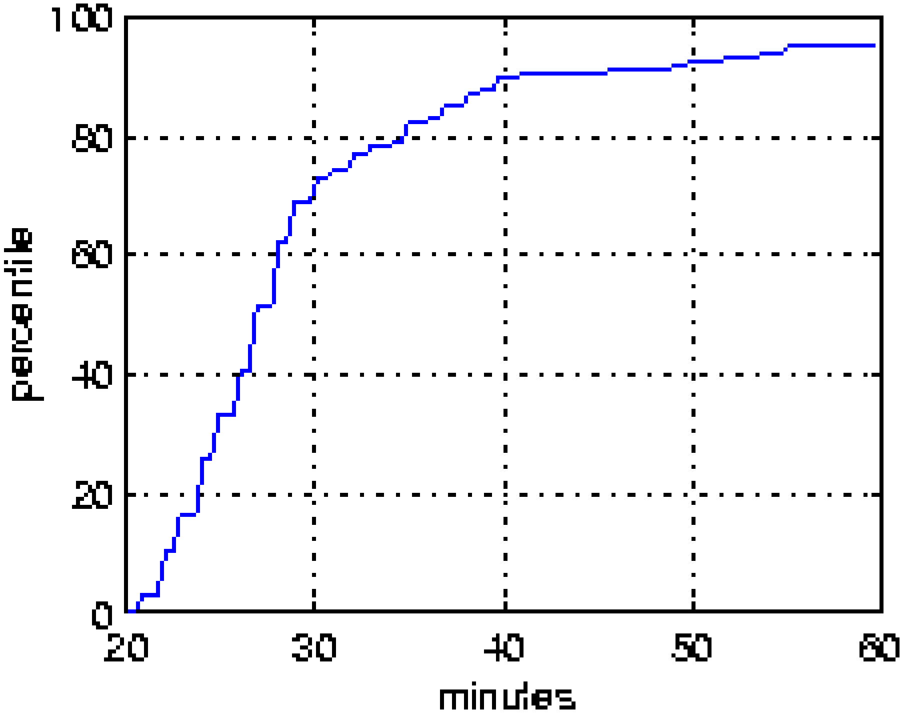

Figure 8. Distribution of travel times on I-405N in Orange County, CA. 8-9 am on Tuesdays in 1998. Mean Travel Time=31.5 min.