NISS

Real-Time Prediction of Incipient Congestion on Freeways from Detector Data

Todd L. Graves, Alan F. Karr, Nagui M. Rouphail,
and Piyushimita Takuriah

Technical Report Number 79
March, 1998

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

Real-Time Prediction of Incipient Congestion on Freeways from Detector Data

Todd L. Graves*
National Institute of Statistical Sciences and Bell Laboratories

Alan F. Karr
National Institute of Statistical Sciences

Nagui M. Rouphail
North Carolina State University

Piyushimita Thakuriah
University of Illinois at Chicago

graves@research.bell-labs.com; karr@niss.org;
rouphail@eos.ncsu.edu; thakuria@voyager.utc.uic.edu

Abstract

This paper is devoted to identifying precursors of *freeway breakdown*: decreased speed and flow as the result of excess demand. We do so by formulating prediction problems wherein we try to predict whether a breakdown will begin in the next five or ten minutes. We use tree-structured statistical models to select conditions that are likely to precede breakdowns. The data we use are from single loop detectors on Interstate 5 in the Seattle area, and were

*Correspondence contact: Room 1Q-331, Bell Laboratories, 1000 East Warrenville Road, Naperville, IL 60566. Research supported in part by NSF grants DMS-9313013 and DMS-9208758 to the National Institute of Statistical Sciences.

available essentially in real time, so results such as ours could be useful to indicate that anti-breakdown countermeasures are called for. We also discuss techniques for visualizing the breakdown process which we found critical in our research, and the extensive process by which researchers must clean up detector data of this form before performing data analyses.

1 Introduction

Freeway breakdown — decreased speed and flow as the result of excess demand, rather than decreased capacity (as from an incident) — is increasingly common. Its consequences range from economic (e.g., lost time)[1] to environmental (increased pollution). At the same time, countermeasures such as ramp metering or lowering the speed limit have been proposed, and are feasible *provided they can be implemented in time*, that is, before breakdown becomes widespread and severe.

This paper is devoted to understanding the processes by which breakdown occurs. In particular, we identify precursors that can alert motorists to imminent breakdown and allow managers to implement countermeasures. We use data from single-loop detectors, which are readily available in many locales. In our specific setting (Interstate Highway 5 in Seattle, WA and its suburbs), the data are available in near-real-time over the Internet, and the computations needed for our methods can be done sufficiently rapidly that countermeasures could be implemented in time to make a difference.

1.1 What is Breakdown?

Critical to any empirical study of breakdown is a quantitative definition of breakdown. Figure 1 depicts a relationship fundamental to the understanding of freeway dynamics. For each flow rate, or number of vehicles that pass per unit time, there are two possible speeds, which correspond to high speed and relatively light traffic, and low speed and very heavy traffic. The points in Figure 1 (left) were taken at one minute intervals from a single detector location on a single day (June 4, 1996) of the data collection. We obtained the ideal relationship by fitting a quadratic curve to these data.[2] The quadratic curve predicts a free-flow speed of 86 MPH, and a maximum flow rate of 1980 vehicles per hour per lane, attainable when traffic moves at 42 MPH.

Heuristically, breakdown *at a single location* is the shift of traffic conditions from the upper segment of the curves in Figure 1 to the lower. Reflecting this, we define breakdown

[1] The December 6, 1997, issue of *The Economist* quotes the Transportation Research Board and the Federal Highway Administration in reporting that increased congestion on US freeways will cost an additional $41 billion in 2010.

[2] Fitting a smooth curve to the data using a smoothing spline did not indicate that it was possible to improve on a quadratic fit.

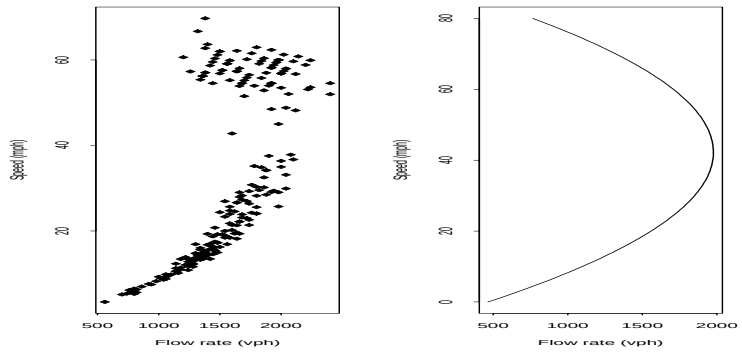

Figure 1: Left: Speed-flow diagram, using data from one location on one day of our data collection. Right: Speed-flow diagram, ideal relationship, fit to the data

conditions at a given location as follows:

> The start of a breakdown at a given location is the first of five consecutive one-minute intervals in which the estimated speed over the detector at that location is less than 30 MPH. Similarly, a breakdown ends with the first of five consecutive one-minute intervals in which speeds exceed 30 MPH.

Thus, breakdowns, if they exist at all, last at least five minutes. Shorter periods below the 30 MPH threshold are of less interest. Exactly what we mean by detector speed (recall that data are from single-loop detectors) appears in §2.3. Figure 2 illustrates this definition.

Our analyses indicate that the conditions that precede breakdown conditions at a given location are slow speeds, or decreases in speeds, either at that location or further downstream. The relationship of these findings to the "phase transition" view of breakdown propounded by physicists, as in [Browne, 1997], is unclear.

1.2 Scope of Work

The remainder of the paper is structured as follows. In §2, we describe the process by which we obtained the data on which the study is based: this involved downloading data from the Internet, a substantial amount of preprocessing to weed out the unreliable data, as well as structuring the data in such a way that we could address the question of what conditions precede breakdown. In §3 we present the visualization techniques we developed, which were particularly critical in identifying the patterns which breakdowns propagate and dissipate. §4 details how we were able to formulate the problem as a prediction problem and describes the models we obtained for conditions that precede breakdowns.

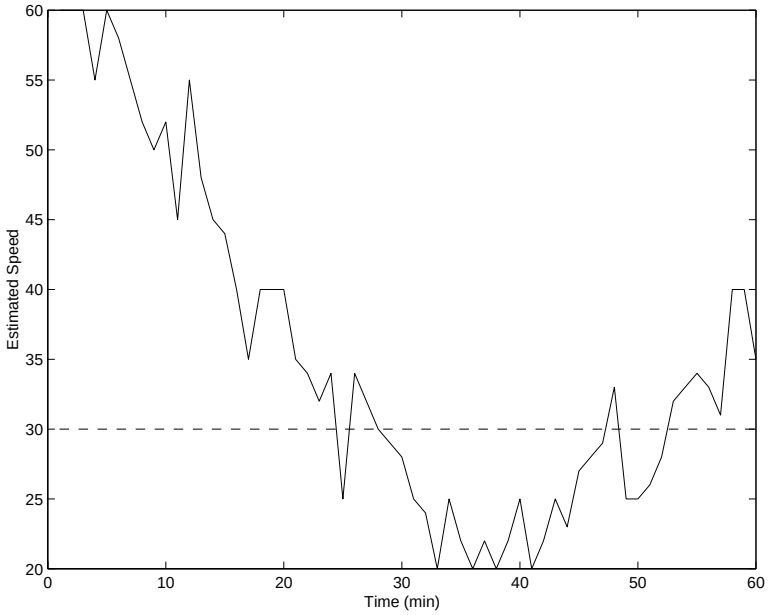

Figure 2: The Start and End of Breakdown at a Detector Location. The thirty MPH threshold is shown, along with the time trace of speeds. Despite the 25 MPH speed at time 25, breakdown does not begin until time 28. Similarly, it ends at time 53 even though the speed at 48 is 33 MPH.

1.3 Comparison with Incident Detection

The freeway breakdown prediction problem has some similarities to the automatic incident detection problem [Chen and Cheng, 1992, Incident Detection Project, 1990, Cheng *et al.*, 1993]. There are similarities in the form of the algorithm (e.g., pattern recognition, statistical approaches, catastrophe theory), in the selection of variables (e.g., detector counts, speeds and occupancies) and in the assessment of the algorithm utility (e.g., detection rates and false alarm rates). For example, the pattern recognition approach, exemplified by the California Algorithm Series [Payne and Tignor, 1976], uses a decision tree structure to identify incidents. The variables of interest are the discontinuities over time, of speed and occupancy over a local detector, and discontinuities over space, of speeds and occupancies across multiple detectors.

An important distinction between the two problems is the time dimension. Incident detection algorithms are reactive, i.e., they detect (or miss) an incident only after it occurs. On the other hand, the freeway breakdown prediction algorithm attempts to anticipate the occurrence of congestion in the near future. The reader should keep that distinction in mind when the results presented in this paper are compared to those available from the incident

detection literature.

2 Data collection

The data we analyzed were collected during the summer of 1996 from an 11.68 mile-long section of Interstate 5 in the downtown and northern suburbs of Seattle, with the detectors located between 220th Street SW and State Road 520, between milepost markers 168.02 and 179.90. See Figure 3 and Figure 4 for maps of the study area. In this region, I-5 has three–four all-purpose lanes; in some places there is also a high-occupancy vehicle (HOV) lane. The northbound direction has detectors in 24 locations; there are 23 detector locations in the southbound direction. Each lane has a detector at each location. We collected data only from mainline detectors, and not from HOV lanes or on- or off-ramps. (Northbound, there are eleven on-ramps and nine off-ramps, while southbound has twelve on-ramps and ten off-ramps.)

Our main goal was to study the morning and evening peak periods in their appropriate directions. Initially, southbound data were between 6:00 AM and 10:00 AM, and northbound data between 3:00 PM and 7:00 PM on weekdays. Ultimately, we increased the afternoon collection time to between 2:00 PM and 8:00 PM because many events seemed to begin before 3:00 PM and last beyond 7:00 PM.

The detectors are of the single-loop variety with twenty-second time resolution. They record both *count*, the number of vehicles that pass over the detector in that period of time (typically reported in units of vehicles per hour per lane) and *occupancy* — the fraction of time that some vehicle is "detected." How speed was imputed from these data is described in §2.3.

2.1 Obtaining data from the Internet

As shown in Figure 5, the "raw" detector data are collected in computers operated by the Washington State Department of Transportation (WSDOT). These data were accessed over the Internet using software developed by researchers in the University of Washington (UW) at Seattle,[3] and stored on computers at the National Institute of Statistical Sciences (NISS) in Research Triangle Park, NC. The Internet-based data access is essential: the detector data are not retained by WSDOT.

Also as shown in Figure 5, UW computers transform the data by appending time stamps and concatenating data from all detectors for a given 20-second time period into a single vector. The time stamps are the clock times at which the data are received in UW, which differ from the time of actual collection times by at most 2-3 milliseconds.

[3]We thank Professor Dan Dailey, Department of Electrical Engineering, for making the software available to us.

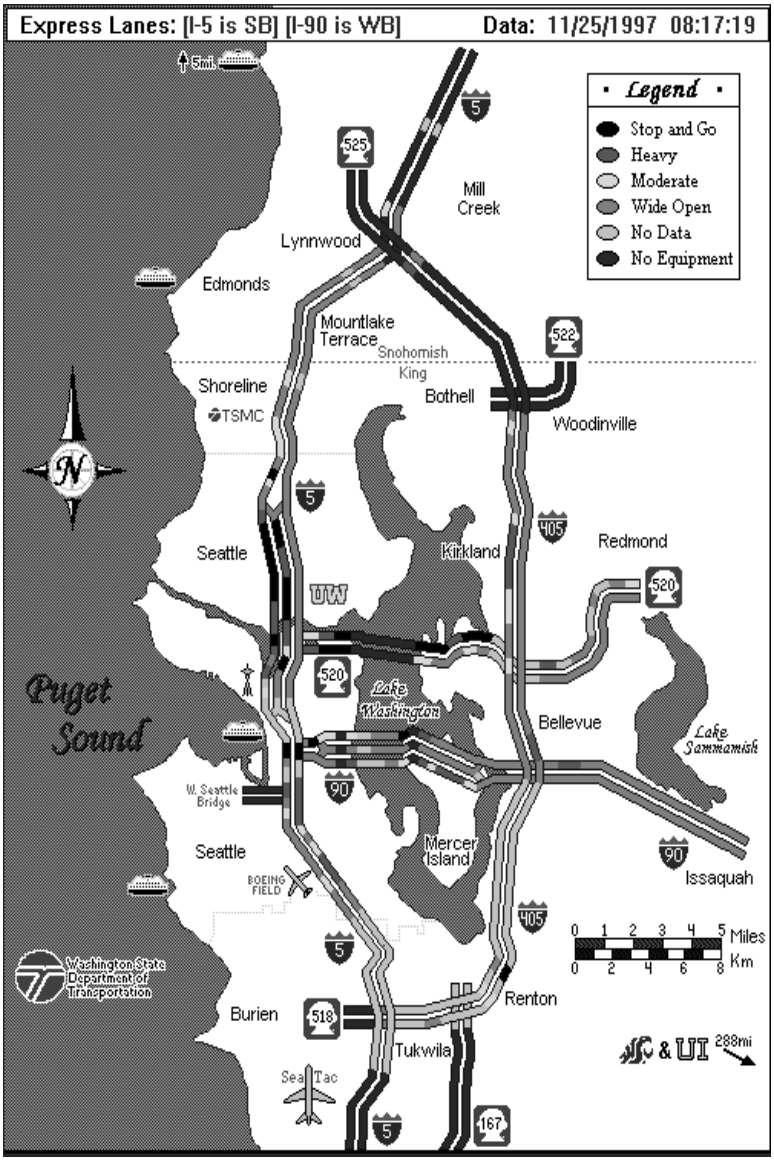

Figure 3: Map of the study area and its surroundings. The northern edge of the study area is Interstate 5's 220th Street SW exit, near the bend in I-5 near the "Mountlake Terrace" label. The southern tip of the study area is close to the State Road 520 exit. Thanks to the Washington State Department of Transportation, and their web site at `http://www.wsdot.wa.gov/`, for providing this map.

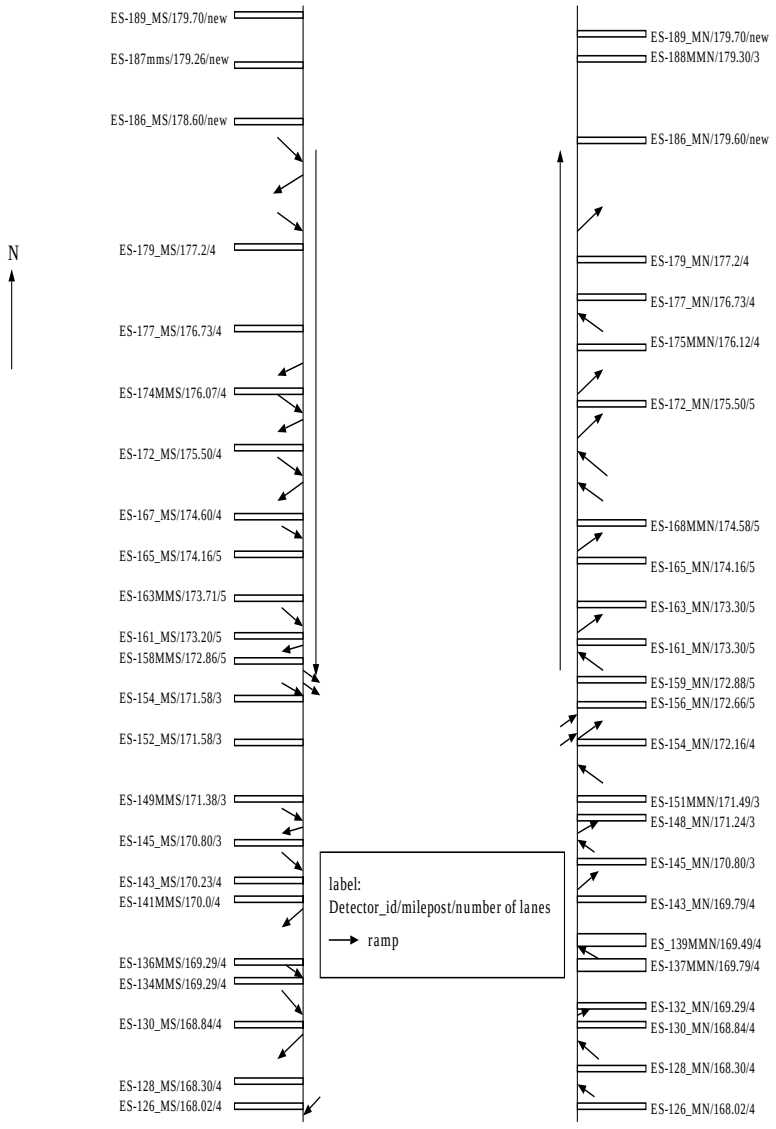

Figure 4: Schematic map of the two directions of traffic in the study area. The label for each detector location contains three pieces of information: the detector ID, its milepost, and the number of lanes (high occupancy lanes included) at that location. Arrows leading onto or off the freeway denote on- and off-ramps respectively.

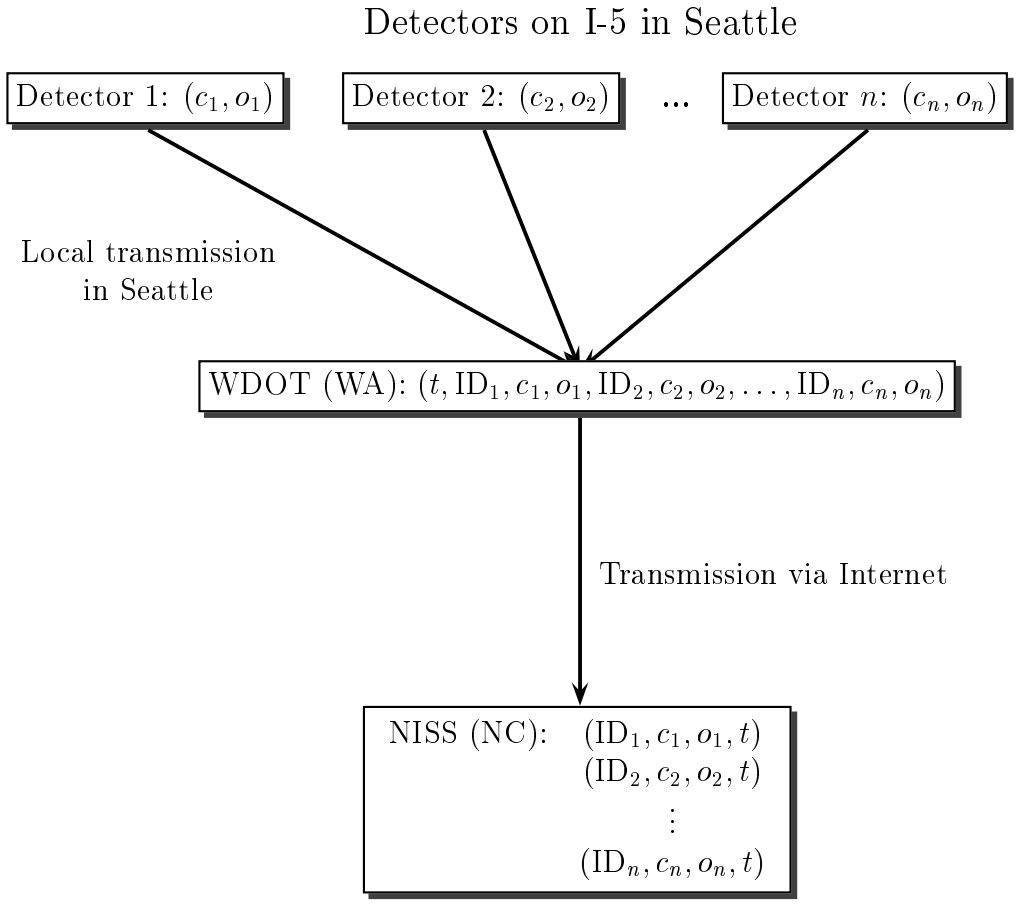

Figure 5: Collection and Transformation of the Count (c) and Occupancy (o) Data. Detector IDs were added by WDOT and all readings for a given time interval were concatenated into a single vector. At NISS the data were disaggregated by detector. Only data for the detectors of interest were retained.

At NISS, these data were transformed once more by separating data into one stream for each detector, which enabled us to save data only for the detectors of interest (91 northbound and 81 southbound, out of more than 3000). Separate files were created for the morning and evening of each day of data collection.

One day's data (AM and PM) amounted to approximately 4 MB; the roughly fifty days of data we collected between May 28 and August 16, 1996, gave a total of 200MB of data (about one-half of which, however, consists of detector IDs).

2.2 Data Preparation

The raw data downloaded from the Internet were not yet suitable for analysis. Both substantial cleanup — eliminating detectors that appeared to be giving nonsensical data, and exploring aggregation (over detectors at the same location and over time) of the data were necessary. We employed the S-Plus language not only to prepare the data for use, but also to conduct the analyses.

2.2.1 Data Screening

Initial inspections of the data revealed many clearly flawed entries. Interactive analyses, with graphical output exemplified in Figure 6, identified many problems. This figure contains one plot of count and occupancy over time from three detectors (in different lanes) at a single location. The right lane detector (whose data are shown in the top row of plots) we accept without modification. The middle lane (second row) detector seems to be generally acceptable, except for a short period of bizarre behavior, so we would use these data, but omit the suspect period from the calculations of average speeds and occupancies. The left lane detector (bottom row) is clearly malfunctioning, and we would disregard its data.

For the southbound (AM) data, one of the 23 locations we examined had no functioning detectors, while five other locations had at least one completely malfunctioning detector, in the sense of the third row of Fig 6. Locations with one demonstrably faulty detector also tended to have counts that seemed inconsistent with the output of adjacent detectors: for instance, the hourly average counts at the downstream detector might have been larger than those at the upstream detector by more vehicles than could plausibly have entered on the on-ramp between the two detectors. To be conservative, we removed from our analyses all locations with at least one malfunctioning detector. This left fifteen northbound and thirteen southbound locations for which we were willing to use the data.

Further purging of data on the basis of implausible imputed speeds is described in §2.3. Ultimately, as a result of the entire cleanup process, approximately 40% of the data were discarded. This is a significant insight into the quality of detector data.

A usable real-time implementation of our techniques would, of course, require that the data cleanup process be automated. (All of the remaining steps are already performed in

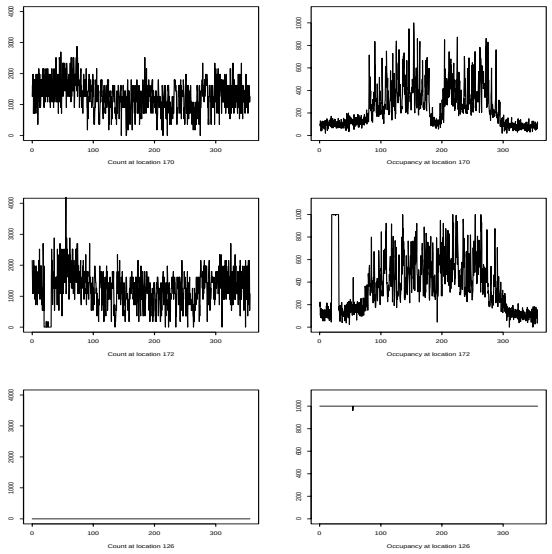

Figure 6: Example plots used for interactive identification of faulty detectors. Upper row: count and occupancy from a healthy detector. Middle row: count and occupancy from a detector with suspect behavior early in the data collection. Lower row: count and occupancy from a failed detector.

real time.)

2.2.2 Aggregation

To remove excessive noise and improve computational aspects of our analyses, the detector data were aggregated in two ways. First, 20-second readings were aggregated to one minute averages, which smoothed the data with no loss of essential information.

Second, we took averages over all detectors in the same location. The justification for this was concurrent research [Click *et al.*, 1996] on data taken from a freeway in North Carolina. In these data, the correspondences between lanes were remarkable (the rightmost lane tended to have slower speeds than the other lanes, but slowdowns happened in all lanes at the same time, to the nearest minute); Figure 7 demonstrates this phenomenon.

After all aggregation was complete, we were still left with some 6MB of usable data.

2.3 Speed imputation

Single-loop detectors record count and occupancy measurements, which are customarily used to derive estimates of the prevailing speed by dividing count by occupancy. More precisely,

suppose that $C(T)$ vehicles pass over the detector during a period of time of length T, and these vehicles have lengths $\ell_1, \ldots, \ell_{C(T)}$ and speeds $v_1, \ldots, v_{C(T)}$. Vehicle i spends $(\ell_i + \ell_D)/v_i$ time over the detector, where ℓ_D is the length of the detection zone, so that

$$\text{Occupancy} = \frac{1}{T} \sum_{i=1}^{C(T)} \frac{\ell_i + \ell_D}{v_i}.$$

Then if, for instance, all the speeds v_i are equal to v, one may recover this speed by computing

$$v = \frac{1}{T \times \text{Occupancy}} \sum_{i=1}^{C(T)} (\ell_i + \ell_D) = \frac{C(T)\bar{\ell}}{T \times \text{Occupancy}}, \tag{1}$$

where $\bar{\ell}$ is the average vehicle length, plus the length of the detection zone. More generally if speeds vary but are independent of vehicle lengths (length is uncorrelated with 1/speed), then by computing `Count/(Occupancy/Length)` one obtains the harmonic mean of these speeds.

In this study we have used 24 feet as the value of $\bar{\ell}$. This assumption is not critical: if, for instance, 22 feet were a more reasonable value of $\bar{\ell}$, all of our conclusions could be modified by multiplying our reported speeds by 12/11. We did, however, (on the advice of WSDOT) assume that vehicle mix does not depend on location, time of day, day, or direction.

Using (1), we removed from the data set all detectors reporting speeds in excess of 100 MPH.

Double–loop detectors would not require imputation of speed. Our techniques would not perform worse in this case than for single–loop detectors, but it is not clear that they would perform significantly better.

3 Visualization of Congestion

Visualizations of the data were the essential step in developing insight about breakdown. Most of these depict speeds, calculated using (1). Recall from §1.1 that breakdown conditions begin at a given location on the first of five consecutive one-minute speed measurements that are slower than 30 MPH, and end with the first of five consecutive speed measurements greater than 30 MPH.

3.1 Breakdown in Time

Figure 7 (Cited in §2.2.2 as justification for aggregating over detectors in different lanes at one location.) shows breakdown during a typical morning peak at a single location on I-40 in Research Triangle Park, NC. The three graphs correspond to the three lanes at a single location, and the x-axis measures minutes after 6:00 AM. Each column of blocks displays a

color histogram of the distribution of speeds of cars passing that detector during that minute: light colors mean a high proportion of cars traveling at that speed. Early in the day, all three detectors saw most cars traveling at the free flow speed of about 70 mph. Traffic first hit very slow speeds at about 7:25 (time 85 on the x-axis), but there was a rapid partial recovery before another brief slowdown at about 7:35. From 7:50 to the end of the data collection, the location was in extended breakdown.

While useful, visualizations such as Figure 7 do not reveal the spatial structure of breakdown, which is what we consider next.

3.2 Breakdown in Time and Space

A visualization of the space-time properties of breakdown on a typical afternoon (June 20, 1996) for northbound traffic on I-5 in Seattle is given in the left panel of Figure 8. There, the x-axis displays time, beginning at 2 PM and going to 8 PM (time zero is 3 PM, and time is measured in minutes). The y-axis displays space (measured in mileposts); traffic flows downward.

Breakdown is shown as follows: a horizontal line is drawn through each one-minute interval and at each location during which the imputed (using (1)) speed is less than 30 MPH. For instance, at milepost 176.73 (the bottom line), breakdown conditions predominate between times 0 and 260 (3:00 and 7:20 PM).

Figure 8 shows clearly that farther downstream, breakdown conditions begin earlier and last longer. This pattern appeared on virtually all serious breakdowns, and was particularly obvious in the evenings up to Friday, July 12, which without exception had long-lasting breakdowns at the same locations. (Intriguingly, breakdowns at these locations became much rarer beginning Monday, July 15.) The triangular pattern is also present, albeit less dramatically, in morning breakdowns, which were generally less severe.

The triangular pattern is characteristic of congestion caused by an excessively high demand — that is, of breakdown. Thus, evening breakdowns start at or downstream of milepost 176.73, and propagate upstream, eventually stretching all the way back to milepost 169.79. Ultimately, cars are no longer joining the congested area as rapidly as they are leaving it, and breakdown conditions dissipate in the downstream direction. In this case the "cause" of the breakdown seems apparent: downstream of milepost 176.73, the freeway changes from five lanes including an HOV lane, to three lanes with no HOV lane.

Also of interest (for example, for purposes of prediction) is the speed of propagation of breakdown, which one may estimate using the slope of the line fitted to the starts of the breakdowns. Figure 8 (left) indicates that this speed is on the order of 3 MPH: of 31 breakdowns appearing in afternoon data, 25 of them had propagated at speeds between 2.5 and 5 MPH.

By contrast with breakdown, Figure 8 (right) contains evidence of a temporary capacity reduction (perhaps an accident) on June 12. Congestion appears at mileposts 169.79–168.3,

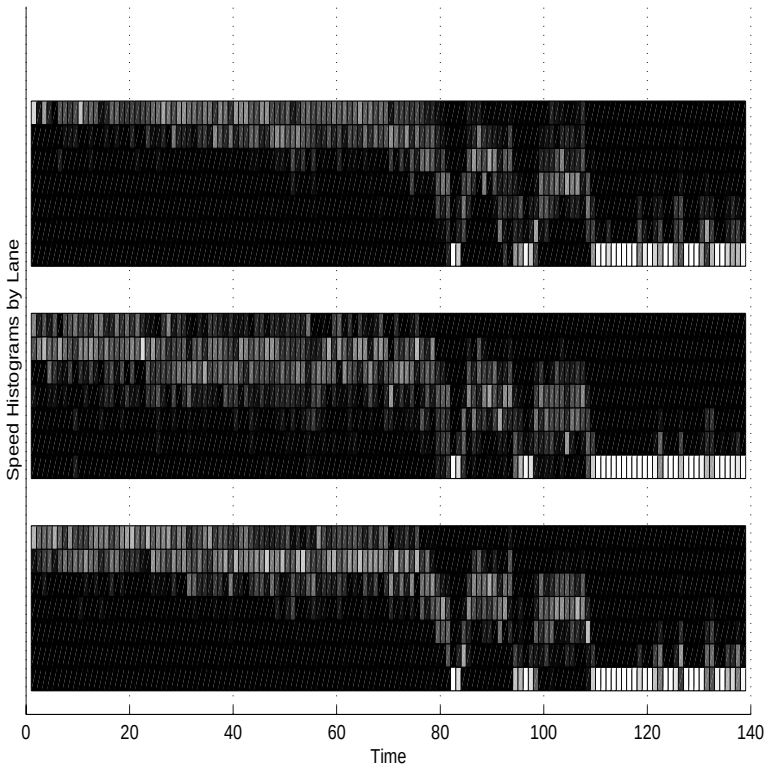

Figure 7: Three detectors at the same location on I-40 in North Carolina on one breakdown-ridden morning. The x-axis represents time in minutes after 6:00 AM. Within each of the three graphs, the y-axis displays speeds, and a vertical column in a single graph gives the speed distribution of the cars passing over the detector in that minute: a light color reflects a large count of vehicles traveling at that speed during that minute. Note that while the lanes have different speeds, they enter breakdown at the same time, to the nearest minute. Speed bins are < 35 MPH, 35–40 MPH, ..., 60–65 MPH, > 65 MPH.

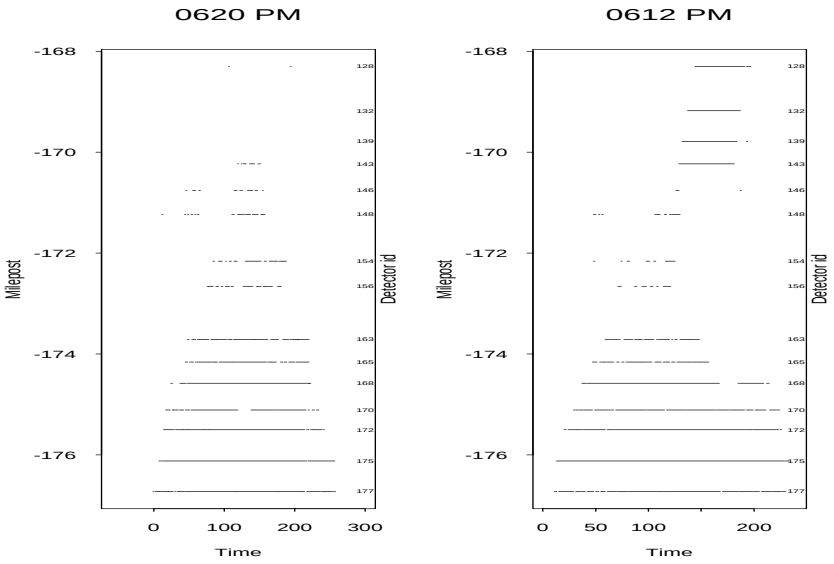

Figure 8: Left: recurring congestion on June 10. Right: recurring congestion (the triangle) and nonrecurring congestion (the parallelogram) on June 12.

which is much less common than further downstream, and the recovery pattern is different: the breakdown shape is a parallelogram instead of a triangle. Moreover, the congestion propagates more rapidly than recurring congestion— from milepost 169.79 through milepost 168.3 at about 7.5 MPH.

4 Development of Predictive Models for Breakdown

In this section we describe our algorithms for predicting when breakdown conditions are likely to appear in the near future. After formulating the prediction problem (§4.1), we discuss classification trees, the methodology we used to build prediction rules, in §4.2, and present our results in §4.3.

4.1 The Prediction Problem

We first predict breakdowns at a single location using only measurements at that location, and then using spatial prediction techniques, which consider measurements from downstream detectors.

The initial question was, "Is a breakdown going to begin here in the next five minutes?", which makes sense only when a breakdown has not yet started. We divided time into fixed

five-minute intervals (whose last digit was five or zero). For each interval that did not contain a start or any other part of a breakdown, (call these potential alarm intervals), the problem is to predict, using the five count, occupancy, and speed measurements *from the five-minute interval*, whether there is (the start of) a breakdown.

As predictor variables we used not only the count, occupancy, and speed measurements themselves, but also several functions of them, including mean, maximum, and minimum and slope (computed using least squares). To capture variability in the five consecutive measurements of a given variable, we used both the standard deviation and mean absolute deviation $\sum_{i=1}^{4} |a_{i+1} - a_i|$, where $a_1, \ldots, a_5$ are the five measurements.

Slope and variability functions were meant to capture volatile conditions in traffic, which have been proposed [NCHRP, 1995] to lead to breakdown. However, at an early stage we found that the variability measurements were not benefiting the predictions, so they were not included in the analyses described in §4.3.

It was also of interest to evaluate how important the aggregation interval is in making prediction. Since many detectors report five-minute rather than 20-second data, we also conducted analyses using only the data these detectors would have available, namely mean count, mean occupancy, and a speed measurement derived from the ratio of mean count to mean occupancy. We also tried to look farther into the future, predicting whether a breakdown will begin in the next 6–10 minutes.

Since we saw in §3.2 that recurring congestion tends to propagate downstream at speeds between 2.5 and 5 MPH, and since detector locations are on the average one-half mile apart, breakdown conditions should require 6–12 minutes to propagate from the downstream detector to the detector of interest. Therefore, the downstream detector should be at its most useful for prediction of breakdowns starting in the second five minute interval from now (i.e. six to ten minutes in the future).

4.2 The Prediction Methodology: CART

To identify values of count, occupancy, and speed in the potential alarm interval particularly likely to precede breakdowns, we used classification tree methodology (see [Breiman *et al.*, 1984] and [Clark and Pregibon, 1992]). Tree-structured classifications divide the set of possible values of predictor variables (in this case, functions of counts, occupancies, and speeds) into subsets, and predict a binary result (in this case, breakdown or no breakdown) depending on which subset the predictor variables fall into. The subsets allowed by the S-Plus implementation (and most others) of classification trees are hyperrectangles, as shown in Figure 9.

An hypothetical example of a tree-based classifier is shown in Figure 9. Here the x-axis denotes the mean count in the potential alarm interval, and the y-axis is the mean occupancy. Each dot corresponds to a single potential alarm interval in the data set, and the dot is colored black if a breakdown occurred, and white if one did not. The dotted lines

describe the classifier: the four subsets are

$$\texttt{Occupancy} > 40\%,$$

$$\texttt{Occupancy} <= 40\% \text{ and } \texttt{Count} <= 1000,$$

$$\texttt{Occupancy} <= 40\% \text{ and } \texttt{Count} > 1000 \text{ and } \texttt{Count} <= 2000,$$

$$\texttt{Occupancy} <= 40\% \text{ and } \texttt{Count} > 2000.$$

If count and occupancy fall into the third subset, then the classifier predicts no breakdown; otherwise it predicts a breakdown.

Another way of depicting the classifier is the tree shown in the bottom view in Figure 9.

A critical part of a tree algorithm is an "impurity measure," which measures how much variability there is in the response variable (here, breakdown or no breakdown) within a node. One begins with a tree which consists of only one node and all the data. Then one looks for ways of splitting the data into two subsets, whose impurity measures — when summed — are as much as possible less than the impurity measure of the entire data set. Allowable "splits" are those into high and low values of one of the predictor variables, such as `Occupancy` <= 40%. After splitting into two subsets, the algorithm turns to making further refinements of these subsets. The impurity measure used here was binomial deviance, computed as

$$-2\{b\log(\frac{b}{b+n}) + n\log(\frac{n}{b+n})\},$$

with $0\log 0$ interpreted as 0, when the data consist of b breakdowns and n non-breakdowns. This measure is a constant multiple of the log of the maximized likelihood function

$$\frac{(b+n)!}{b!n!}p_b^b p_n^n,$$

maximized when $p_b = 1 - p_n = \frac{b}{b+n}$, up to the factorial constant which does not depend on p_b and p_n. It equals zero when all the data are identical (here, all breakdowns or no breakdowns), and increases as the data become less "pure."

A troublesome question to tree "growers" is how many splits should be allowed and how leafy the tree should grow. A popular choice is to grow a tree with a clearly excessive number of splits and then "prune" back using a method such as cross-validation, for which see either [Breiman *et al.*, 1984] or [Clark and Pregibon, 1992].

4.3 Results

To assess prediction methods, the data were divided into two halves, the first used to construct trees and the second to evaluate the trees (with data not used to construct them). The "training set" for the AM data was simply the first half of the days, while that for the

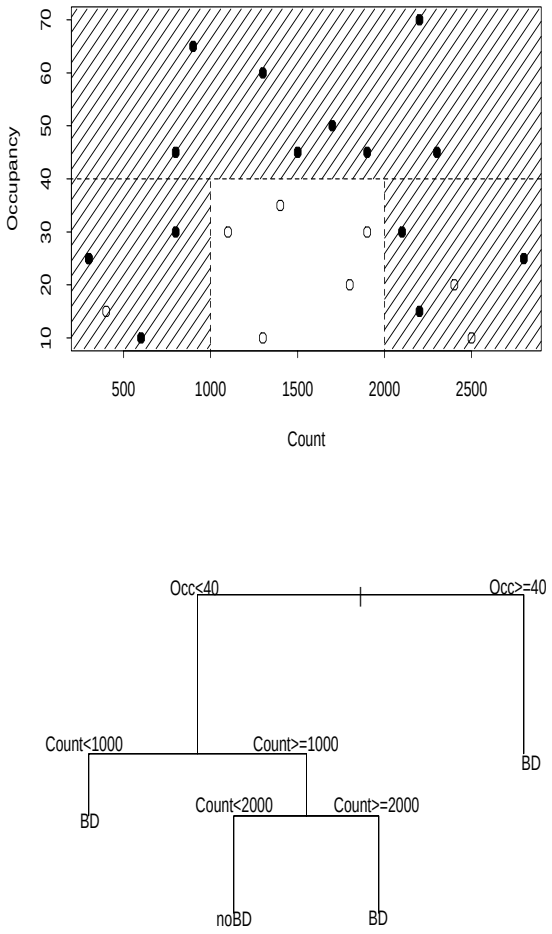

Figure 9: Top: graph view of tree-based classifier. The shaded subset consists of those conditions which the tree predicts that a breakdown will follow. The hollow icons correspond to (synthetic) observations where breakdowns did not follow the count-occupancy conditions shown, so that hollow icons inside the shaded area are false alarms. The filled-in icons correspond to conditions that preceded breakdown, so that if there were any filled-in icons in the unshaded area, they would represent failed detections. Bottom: tree view of same classifier. Nodes that terminate in a "BD" are those in which we predict breakdown.

PM data was more complicated: the first one-half of the days up to July 12, and the one-first half of the days after July 15. Each tree was constructed using data from roughly 30,000 five-minute intervals, 300 of which preceded breakdowns, and then tested on a data set of similar size.

We use three measures of the success of a tree model. First, *detection rate* is the fraction of breakdown starts that we predict successfully. We also use two different *false alarm rates*: the fraction of alarms which are not followed by breakdowns, and the fraction of potential alarm intervals not followed by breakdowns but in which we set off alarms. Since starts of breakdown follow only about one percent of the potential alarm intervals in the data set, the first false alarm measure is typically large while the second is small.

To compare different trees as informatively as possible, we strive to set the second false alarm rate to the arbitrarily chosen value 5%, and we introduce another element to tree-based prediction: if the conditions are in the hyperrectangle which would bring the false alarm rate over 5%, execute an auxiliary randomization so that an alarm is set off with some probability, the probability chosen to set the false alarm rate to exactly 5%.

Results from eight trees are shown in Table 1. The eight trees represent all combinations of three two-level factors: the length of time into the future that we are trying to predict breakdown, the aggregation of the predictor variables and which detectors are used. The entries in the Table are the *case identifier*, showing which detectors are used (LD = local and downstream; L = local only), aggregation of predictor variables (1 = one minute; 5 = five minutes) and the prediction interval (1–5 = next five minutes; 6–10 = six to ten minutes in the future); the *training set detection rate*; the *test set detection rate*; and the *test set false alarm rate*. Details of the associated trees are in Table 2, in the Appendix.

The false alarm rates we found for the test data sets were very close indeed to the 0.05 of the training set: all but one (0.058) of the eight were within 0.02 of 0.5. The other measure of false alarm rate, the fraction of alarms which do not precede breakdowns, is also not shown, because these measures tended to be very similar, ranging from 86 to 90 percent.

Our results show that breakdown speeds are normally preceded by speeds that are slow but possibly not quite breakdown level, both at the current detector and downstream. High occupancies have similar effects to low speeds. Low counts (the bottom segments of the curves in Figure 1) are more predictive of breakdown than high counts; exceptions are in the first and third cases of Table 1, when large downstream counts can increase the likelihood of breakdown beginning six to ten minutes from now, given that no breakdown begins in one to five minutes.

In terms of time aggregation of predictor variables, comparing relevant entries shows that the loss of predictive power from aggregating one-minute data to the five-minute level is not large.

Comparing the first two cases (or the third and fourth) in Table 1 demonstrates that readings from the downstream detector add little predictive ability to readings from the current detector in terms of predicting breakdowns beginning in only one to five minutes.

Case	Training Set Detection Rate	Test Set Detection Rate	Test Set False Alarm Rate
(LD, 1, 1–5)	.68	.64	.048
(L, 1, 1–5)	.58	.61	.051
(LD, 5, 1–5)	.66	.60	.058
(L, 5, 1–5)	.57	.57	.050
(LD, 1, 6–10)	.72	.61	.049
(L, 1, 6–10)	.51	.52	.050
(LD, 5, 6–10)	.65	.59	.049
(L, 5, 6–10)	.48	.47	.048

Table 1: Summary of Prediction Results. The case identifier shows which detectors are used (LD = local and downstream; L = local only), aggregation of predictor variables (1 = one minute; 5 = five minutes) and the prediction interval (1–5 = next five minutes; 6–10 = six to ten minutes in the future.)

However, for predictions 6–10 minutes into the future (the last four cases), data from the downstream detector do improve predictions; in fact, predictions six to ten minutes in the future are very nearly as successful as predictions one to five minutes into the future, provided we have downstream detector information available. This finding is also consistent with our observations in §4.1 about the speed of propagation of breakdown.

A number of other variables proposed in the literature as indicative of incipient congestion failed to enter our classification trees. These include `Count` × `Occupancy` and `Count` × `Speed` and all measures of variability.

It should be noted that the false alarm rates and detection rates are actually underestimates of the success of the predictions. Many of the false alarms occurred when breakdowns did not begin in the immediately following interval, but near-breakdown conditions persisted for several minutes (often as long as half an hour) before a breakdown began. Arguably such events should be evaluated as greater successes than the events we classified as successful prediction, because they predict breakdowns farther in the future and are thus more useful for implementing countermeasures.

5 Conclusions

This research has a number of implications. First, it made use of the Internet to obtain data from a remote location. Since the data were available in real time, it would have been possible to observe from North Carolina the conditions on Seattle roadways, and even to apply control measures if we had the capability.

Included in this research is a study of the quality of detector data. We found it necessary to discard roughly 40% of our data. Future researchers should be under no illusions about the amount of effort that will be required to sort out poorly functioning detectors.

This research provided significant insights regarding differences between recurring and non-recurring congestion, including visualization tools for identifying both types. Both types of congestion propagate from downstream to upstream, and while recurring congestion clears upstream first, the opposite is true for non-recurring congestion. This leads to triangular patterns for recurring congestion in graphics such as Figure 8, while non-recurring congestion appears as a parallelogram. Our data also indicate that non-recurring congestion propagates more rapidly than recurring congestion.

We showed, further, that CART is an effective technique for constructing models to predict breakdown. Key predictors are speed and occupancy. No evidence was found that measures of variability or volatility in traffic lead to improved predictions.

The results from the trees demonstrate that using more than one detector station can lead to large benefits. In particular, by observing detectors downstream from the detector of interest, one may predict the onset of breakdown farther into the future.

We also compared the usefulness of detector data aggregated to five minute averages to finer, one minute data. The gains from using one minute data were small, but the data strongly suggested that there is a gain: in Table 1, there are four pairs of trees that differ only in the degree of aggregation, and in all four pairs, the one minute data beat the five minute data in both training set and test set detection rate. The margin was only about 3% on the average, though.

Appendix

Here we present the specific classification trees associated with the results presented in §4.3, using the case identifiers in Table 1. The variable names appearing here typically consist of three parts: they end with either `Count`, `Occ`, or `Speed`; this identifier is preceded by the function applied to the five count, occupancy, or speed observations (the possibilities are `Max`, `Min`, `Mean`, and `Slope`); and if the name begins with `Down`, the readings from the downstream detector are required instead of those from the current detector. The italicized lines incorporate the randomization described in §4.3.

(LD, 1, 1–5) (MinSpeed $\leq$ 49) AND [(DownMaxOcc>24) OR (DownSlopeCount $\leq$ -109) OR (MeanSpeed $\leq$ 29)] OR [(49 < MinSpeed $\leq$ 60) AND (DownSlopeOcc>1.7)] *OR [(MinSpeed $\leq$ 49) AND (DownMaxOcc $\leq$ 24) AND (DownSlopeCount > -109) AND (MeanSpeed > 29)]*
(L, 1, 1–5) (MinSpeed $\leq$ 49) AND (MeanSpeed $\leq$ 33) OR [(MinSpeed $\leq$ 49) AND (MeanSpeed > 33) AND (SlopeOcc > 1.3) AND (MeanCount $\leq$ 1889)] *OR [(MinSpeed $\leq$ 49) AND (MeanSpeed > 33) AND (SlopeOcc $\leq$ 1.3) AND (MinCount $\leq$ 1757)]*
(LD, 5, 1–5) (Speed $\leq$ 55) AND (DownSpeed $\leq$ 34) OR [(Speed $\leq$ 55) AND (34 < DownSpeed $\leq$ 44) AND (Count $\leq$ 1529)] OR [(Speed $\leq$ 29) AND (DownSpeed >44)] OR [(Speed > 55) AND (DownSpeed $\leq$ 59) AND (Occ > 14) AND (Count $\leq$ 1910)] *OR [(Speed $\leq$ 55) AND (34 < DownSpeed $\leq$ 44) AND (Count > 1529)]*
(L, 5, 1–5) (Speed $\leq$ 37) *OR [(37 < Speed $\leq$ 55) AND (Count $\leq$ 1669)]*
(LD, 1, 6–10) (DownMinSpeed $\leq$ 19) AND (MaxSpeed $\leq$ 62) OR [(19 < DownMinSpeed $\leq$ 36) AND (MaxSpeed $\leq$ 62) AND (DownSlopeSpeed $\leq$ 1.7)] OR [(DownMinSpeed $\leq$ 36) AND (MaxSpeed > 62) AND (DownSlopeSpeed $\leq$ -7.8) AND (DownMeanOcc $\leq$ 24)] OR [(DownMinSpeed > 36) AND (MeanOcc > 24) AND (DownMinCount $\leq$ 1830) AND (MeanSpeed $\leq$ 29)] OR [(DownMinSpeed > 36) AND (MeanOcc > 24) AND (DownMinCount > 1830)] *OR [(DownMinSpeed > 36) AND (MeanOcc $\leq$ 24) AND (MaxSpeed $\leq$ 64) AND (DownSlopeSpeed $\leq$ -2.7)]*
(L, 1, 6–10) (MeanSpeed $\leq$ 33) OR [(33 < MeanSpeed $\leq$ 55) AND (SlopeSpeed $\leq$ -1.8) AND (MaxCount $\leq$ 2120)] OR [(55 < MeanSpeed $\leq$ 62) AND (MinCount $\leq$ 1250) AND (SlopeSpeed $\leq$ -2.5)] *OR [(33 < MeanSpeed $\leq$ 55) AND (SlopeSpeed > -1.8) AND (MinCount $\leq$ 1750)]*
(LD, 5, 6–10) (DownSpeed $\leq$ 27) AND (Speed $\leq$ 55) OR [(27 < DownSpeed $\leq$ 45) AND (Speed $\leq$ 55) AND (Count $\leq$ 1950)] OR [(DownSpeed $\leq$ 45) AND (Speed > 55) AND (DownCount $\leq$ 960)] OR [(DownSpeed > 45) AND (Speed $\leq$ 58) AND (Occ > 24) AND (DownCount > 2030)] *OR [(DownSpeed $\leq$ 45) AND (55 < Speed $\leq$ 62) AND (DownCount > 960)]*
(L, 5, 6–10) (Speed $\leq$ 37) *OR [(37 < Speed $\leq$ 55) AND (Count $\leq$ 1920) AND (Occ > 11.7)]*

Table 2: Classification trees corresponding to Table 1. Italicized nodes invoke the randomization described in §4.3.

References

[Breiman *et al.*, 1984] Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees, Wadsworth & Brooks/Cole, Monterey, CA (1984).

[Browne, 1997] Browne, M. W., "Stuck in Traffic? Consult a Physicist," *New York Times*, November 25, 1997, F1-3.

[Chen and Cheng, 1992] Chen, C.-H. and G.-L. Cheng, "A Review of Recent Freeway Incident Detection Algorithms," University of Maryland, 1992.

[Cheng *et al.*, 1993] Cheng, G.-L., Y. L. Li, H. J. Payne, and S. G. Ritchie, "Incident Detection Issues, Task A, Draft Interim Report: 'Automatic Freeway Incident Detection'," FHWA Contract DTFH61-93-C-00015, October 1993.

[Clark and Pregibon, 1992] Clark, L. A. and D. Pregibon, "Tree-based models," in Statistical Models in S (J. M. Chambers and T. J. Hastie, eds.), Wadsworth & Brooks/Cole, Pacific Grove, CA, pp. 377–419, (1992).

[Click *et al.*, 1996] Click, S. M., N. M. Rouphail, R. G. Hughes, and T. L. Graves, "Using Advanced Vehicle Monitoring Systems to Extend System Capacity along North Carolina Freeways," Final Report, North Carolina Department of Transportation (ITRE) (1996).

[Incident Detection Project, 1990] Incident Detection Project, "Incident Detection Algorithms," University of Minnesota, 1990.

[NCHRP, 1995] North Carolina Highway Research Project, "Speed-Flow Relationships for Basic Freeway Segments," JHK and Associates and Texas Transportation Institute, May 1995.

[Payne and Tignor, 1976] Payne, H., and S. Tignor, "Freeway Incident Detection Algorithms based on Decision Trees and States," *Transportation Research Record* 682, TRB, Washington, D.C., April 1976.

[Petty *et al.*, 1998] Petty, K. F., P. J. Bickel, J. Jiang, M. Ostland, J. A. Rice, Y. Ritov, and F. Schoenberg, "Accurate estimation of travel times from single-loop detectors," *Transportation Research A* 32A, 1–17 (1998).