

UC Berkeley

Research Reports

Title

Loop Detector Data Collection and Travel Time Measurement in the Berkeley Highway Laboratory

Permalink

<https://escholarship.org/uc/item/0248v7w8>

Authors

May, Adolf D.
Cayford, Randall
Coifman, Ben
et al.

Publication Date

2003-04-01

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

Loop Detector Data Collection and Travel Time Measurement in the Berkeley Highway Laboratory

**Adolf D. May, Randall Cayford,
Ben Coifman, Greg Merritt**

**California PATH Research Report
UCB-ITS-PRR-2003-17**

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 4134

April 2003

ISSN 1055-1425

Loop Detector Data Collection and Travel Time Measurement in the Berkeley Highway Laboratory

Adolf D. May University of California
Randall Cayford University of California
Ben Coifman Ohio State University
Greg Merritt University of California

ABSTRACT

This document is the final report for the 2001-2002 Berkeley Highway Laboratory (BHL) Project which is a part of the University of California's PATH program and supported by the California Department of Transportation (Caltrans). The primary objective of this project has been to maintain, improve, and conduct research on the Berkeley Highway Laboratory (BHL) detector system.

- This report contains ten chapters that describe the work undertaken and the results of each task of the project. The first chapter introduces the project, provides a project background, and a site description. The next nine chapters describe each of the project's nine tasks including the task objective, process followed, and results. The last chapter contains a summary of major accomplishments and identifies future research directions.

KEY WORDS: Data Communication, Freeways, Real Time Information, Sensors, System Engineering, Traffic Surveillance, Vehicle Detectors

EXECUTIVE SUMMARY

The BHL detector system has been continuously maintained. Its major accomplishments include improved system reliability and more efficient recoveries from failures.

The BHL detector data has been continuously collected and archived. Its major accomplishments include daily summary data and vehicle-level data.

The BHL detector data has been provided to researchers upon request. This has resulted in the use of BHL data for additional traffic research beyond the scope of this project.

The BHL detector system link to District 04 has been maintained. This has allowed the BHL loop detector stations to function as standard Caltrans District 04 detectors while providing much richer data and research possibilities.

Research of alternative detector diagnostic tests has been undertaken. Its major short-term and longer-term accomplishments include the development and refinement of several single- and dual-loop detector data validation tests.

An initial set of detector diagnostics has been implemented into the BHL detector system. This has resulted in a set of four single-loop and one dual-loop tests that are run continuously against the incoming data stream from every station.

Further research effort has been given to vehicle re-identification. This has resulted in higher frequency of vehicle identification while simultaneously reducing the error rate.

Alternative communication technologies have been evaluated. This has resulted in the development of a new communication system using off-the-shelf components which is cheaper, simpler, and more robust.

Efforts have been expended on moving toward larger-scale implementation. This effort has resulted in modifications to the BHL system that make it more adaptable to different cabinet, station, and loop configurations. The system has been modified to more closely match Caltrans standards for station identification and layout. The cost of such systems, in both time and hardware, has been reduced by the use of same standard field components that are currently used for data collection.

The efforts on the 2001-2002 BHL project have resulted in the identification of future research directions with particular focus on next year's BHL project. A strong working relationship has developed between the BHL research team with both Caltrans District 04 and Headquarters staff.

Table of Contents

1. INTRODUCTION	1
2. MAINTAIN BHL SURVEILLANCE SYSTEM	6
2.1 SYSTEM COMPONENTS.....	6
2.2 MONITORING THE DATA SYSTEM.....	8
2.3 FAILURE MODES.....	11
2.4 FUTURE PLANS.....	14
3. GENERATE AND ARCHIVE SUMMARY DATA SETS	16
3.1 DATA GENERATION.....	16
3.1 DATA ARCHIVING.....	19
4. DISTRIBUTE SURVEILLANCE SYSTEM DATA	21
5. MAINTAIN DATA LINK TO DISTRICT 4	26
6. DEVELOP ONLINE DETECTOR DIAGNOSTICS	28
6.1 DEVELOPMENT OF SINGLE LOOP TESTS.....	29
6.2 DEVELOPMENT OF DUAL LOOP TESTS.....	38
6.3 TESTS IMPLEMENTED IN REAL TIME.....	44
6.4 FUTURE DIRECTIONS.....	44
7. IMPLEMENTING DETECTOR DIAGNOSTICS	46
7.1 THE ACTIVITY TEST.....	47
7.2 MINIMUM ON TIME TEST.....	47
7.3 MAXIMUM ON TIME TEST.....	47
7.4 MODE ON TIME TEST.....	47
7.5 DUAL LOOP ON TIME DIFFERENCE TEST.....	48
7.6 ARCHIVING TEST RESULTS.....	49
7.7 MONITORING THE LOOPS.....	50
7.8 FUTURE DIAGNOSTICS.....	50
8. UPDATE ON THE VEHICLE REIDENTIFICATION ALGORITHMS	54
8.1 REDUCING THE SEVERITY OF ERRORS.....	55
8.2 EXTENDING TO SINGLE LOOP DETECTORS.....	62
8.3 FUTURE DIRECTIONS.....	65
9. EXPERIENCE WITH ALTERNATIVE COMMUNICATIONS	65
10. MOVING TOWARDS A LARGER SCALE IMPLEMENTATION	72
11. CONCLUSION	76
11.1 SUMMARY OF MAJOR ACCOMPLISHMENTS.....	76
11.2 FUTURE RESEARCH DIRECTIONS.....	78

List of Figures

Figure 1-1	Location map of the BHL detector system	2
Figure 1-2	Flow chart of the BHL detector system and its connection to the District 4 Traffic Management Center and the state-wide performance monitoring system	3
Figure 2-1	Components of the BHL data collection and processing system	7
Figure 2-2	Typical density contour map	9
Figure 2-3	Typical monthly overview of density contour maps	10
Figure 2-4	Modem connectivity and data availability status	12
Figure 4-1	30 second aggregate data for each bi-directional station	22
Figure 4-2	Typical density chart for a directional station	23
Figure 4-3	Typical display of travel times between pairs of adjacent stations	24
Figure 6-1	Cumulative distribution of effective vehicle lengths on a single day at a dual loop detector	30
Figure 6-2	Distribution of on-times during free flow on a single day at a dual loop detector	31
Figure 6-3	Acceptable length and velocity combinations for min and max on-time tests	34
Figure 6-4	Acceptable length and velocity combinations for mode on-time test	36
Figure 6-5	On-times from a good dual-loop detector	40
Figure 6-6	On-times from a bad dual-loop detector	41
Figure 7.1	Summary Status Page	51
Figure 7.2	Detail Status Page	52
Figure 8-1	Vehicle Match Matrix	57
Figure 8-2	Vehicle matches with filter test	58
Figure 8-3	One vehicle passing over a dual loop detector	60
Figure 8-4	One vehicle traversing an extended link between two detector stations, illustrating the free flow travel time range.	63
Figure 8-5	The resulting travel times for matched vehicles in one lane over 20 hours	64
Figure 9.1	Controller to PC to Ricochet to Server	67
Figure 9.2	Controller to CDPD to Server	69

1. Introduction

This document is the final report for the 2001-2002 Berkeley Highway Laboratory (BHL) Project which is part of the University of California's PATH program and supported by the California Department of Transportation (Caltrans). The primary objective of this project has been to maintain, improve, and conduct research on the Berkeley Highway Laboratory detector system.

The BHL detector system is located along I-80 in the East Bay portion of the San Francisco Bay Area and lies between the Powell Street interchange and the Gilman Street interchange. Figure 1 is a location map of the BHL detector system which consists of eight detector stations in each direction over this 2.7 mile study section of I-80. The detector stations are approximately one-third of a mile apart and each station includes a pair of detectors in each of its 10 or 12 lanes. Each of the 164 detectors in the system is interrogated every 1/60 of a second as to whether a vehicle is occupying the loop detector and this microscopic information has been archived continuously during the duration of this project.

The BHL detector system is part of the Caltran's District 4 detector system which in turn is included in the state-wide performance monitoring system (PEMS). Figure 2 is a flow chart that includes the elements of the BHL detector system (on the left) and its connection to the District 4 Traffic Management Center and the state-wide performance monitoring system (on the right). The remaining portions of this report will deal with the UCB activities as shown on the left side of this figure and its important connection to District 4 Traffic Management Center.

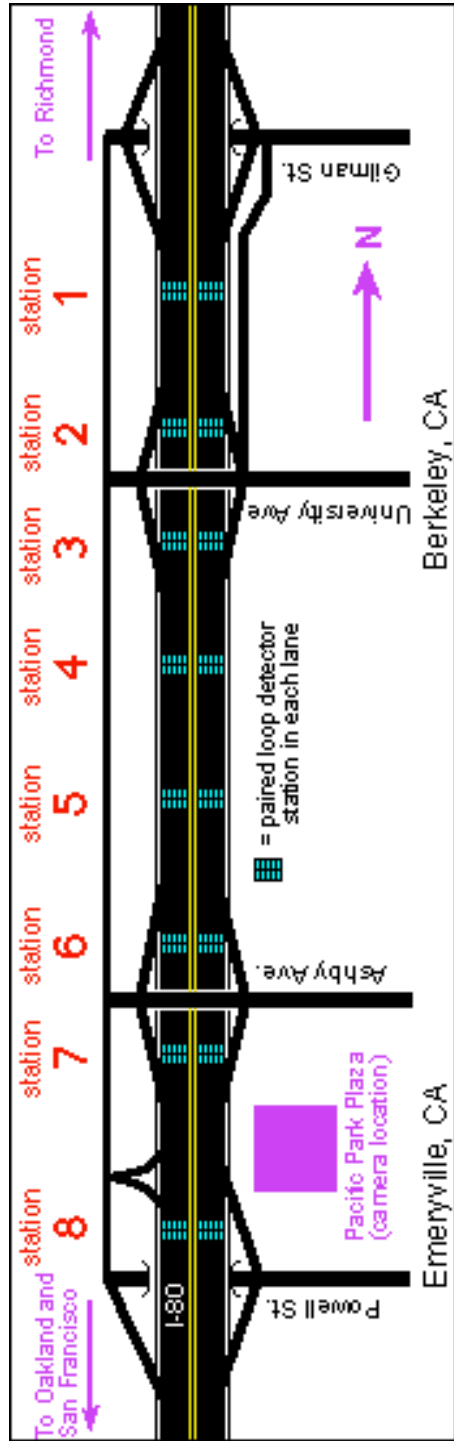


Figure 1-1 Location map of the BHL detector system

RELATIONSHIP BETWEEN DISTRICT 04 DATA COLLECTION, PEMS SYSTEM, AND BERKELEY HIGHWAY LABORATORY

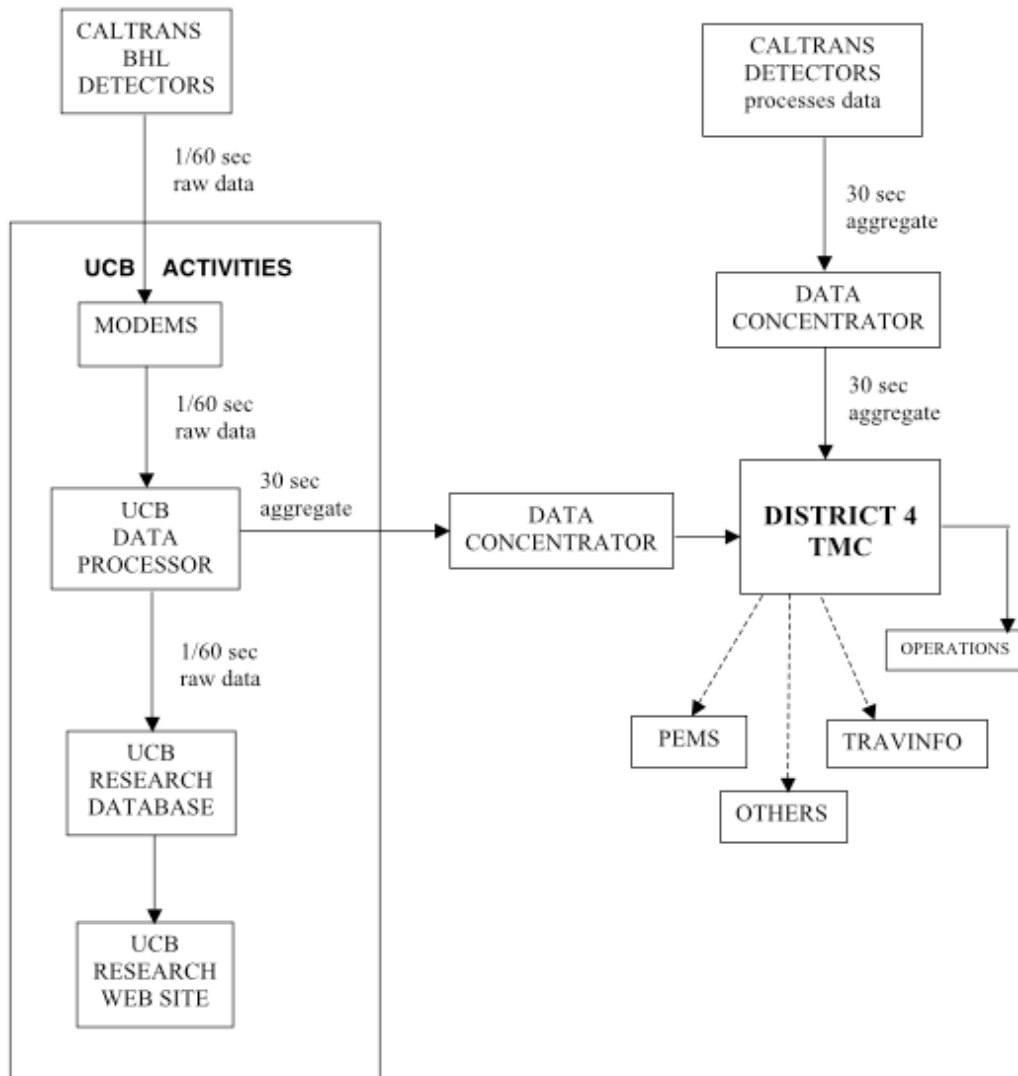


Figure 1-2 Flow chart of the BHL detector system and its connection to the District 4 Traffic Management Center and the state-wide performance monitoring system

An important element of this BHL project has been its coordinated team effort, its close working relationship with Caltrans, and quarterly reporting to the PATH program. Several days were set aside at the beginning of the project and at the end of each of the four quarters during the project for exchanging team member accomplishments, preparing for meeting with Caltrans, inspecting field sites, meeting with Caltrans, and developing plans for the next quarter's efforts. Additional meetings were held as required between team members and with Caltrans staff, and Caltrans staff was informed whenever significant detector problems were encountered.

The work on this project was divided into nine interrelated tasks and the next nine chapters of this report will address these nine tasks. The first set of tasks of the project were to continuously maintain the BHL detector system (Chapter 2), generate and archive data (Chapter 3), distribute BHL detector data to researchers (Chapter 4), and maintain the data link to District 4 (Chapter 5). The successful completion of these tasks insured that the BHL detector system was maintained, data stored, distributed to researchers as requested, and supported District 4 requirements.

The second set of tasks of the project were more research-oriented and included developing detector diagnostics (Chapter 6), implementing detector diagnostics (Chapter 7), furthering research on vehicle re-identification (Chapter 8), evaluating alternative communication technologies (Chapter 9), and moving toward larger scale implementation (Chapter 10). The successful completion of these tasks resulted in an improved BHL detector system that included detector diagnostics, system monitoring diagnostics, and improved communications. In addition, these research efforts provided the framework for future enhancements and potential larger scale implementations.

The final chapter in this report is divided into two major parts. The first part summarizes the major accomplishments of this year's project both in terms of immediate and longer term consequences. The second part describes future research directions with particular focus on next year's project.

2. Maintain BHL Surveillance System

The Berkeley Highway Lab consists of many physical components, including traffic detection loops, network communications infrastructure, and computers. To capture as much data as possible, it is important to maintain full functionality of all of these components. As part of this project, a number of tools and procedures have been used to monitor the status of system components and diagnose failures and outages.

2.1 System Components

The components of the BHL data collection and processing system are shown in Figure 2-1. Data is collected from the loop controllers every second by CDPD (Cellular Digital Packet Data) modems and sent via TCP/IP to the Data Collection Server on the UC Berkeley campus. Software running on this computer records the raw CDPD data in local files and also transmits the data to a database running on a separate computer. Most of the BHL data computation and display tools extract data directly from this database.

As 1/60 second field data is recorded in the database, software running on a dedicated computer computes 30 second averages. The results of these calculations are stored back in the database.

Also running continuously, the Vehicle Re-identification and Travel Time Processor extracts raw data from the database and feeds the computational results back into the database. This data processing computer also sends 30 second data to Caltrans District 4 via a dedicated frame relay data line (Chapter 5).

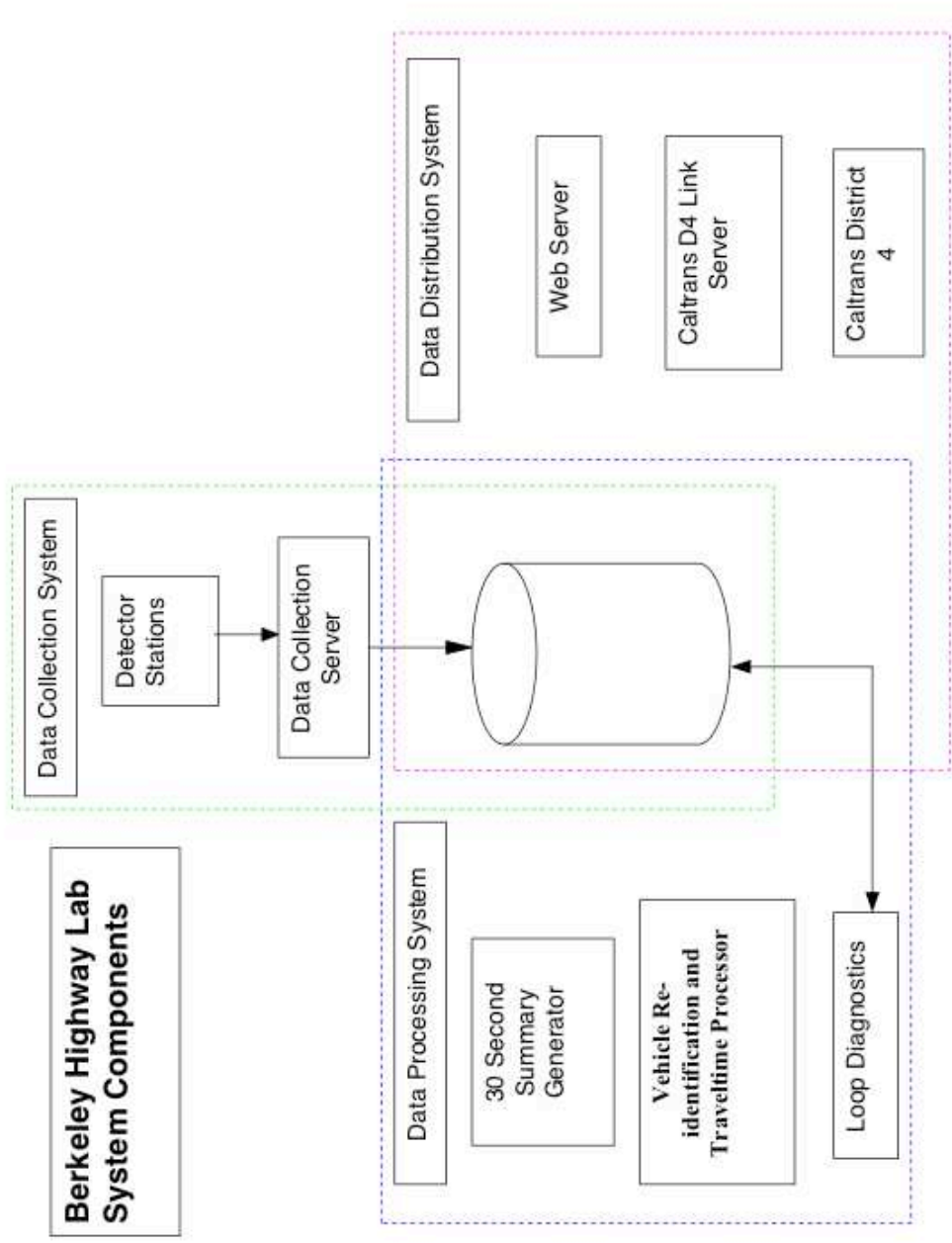


Figure 2-1 Components of the BHL data collection and processing system

The UC Berkeley Institute of Transportation Studies Web server runs software that can query the database and present BHL data numerically and graphically. These features are publicly accessible.

2.2 Monitoring the data system

BHL team members monitor the status of the data collection and processing system on a daily basis. Several of the publicly accessible Web-based presentations of live data are useful for a quick overview of system status, as the successful operation of these displays requires many system components to be functioning. If these displays are not functioning or are missing data that should be available, additional tools are accessed to determine the source of the problem. Most of these status display tools are passive, displaying information only on demand. However, several components are more active, generating an e-mail notification to BHL team members in the event of a problem. BHL team members correct problems directly whenever possible, or report the problem to the appropriate service provider.

The Web-based I-80 Density Contour Maps provide a quick overview of recent system status. Figure 2-2 shows a typical map. This map shows vehicle density in vehicles per mile per lane, for stations 1-7, calculated every 5 minutes. The time period shown is the current calendar day; maps for previous days can also be displayed. If current data is displayed for all stations, it is certain that the corresponding controllers and modems are functioning, the Data Collection Server is functioning, the database is receiving raw data, and the Web server is functioning. Monthly composites of these daily plots are created to give an overview of data acquisition history (Figure 2-3).

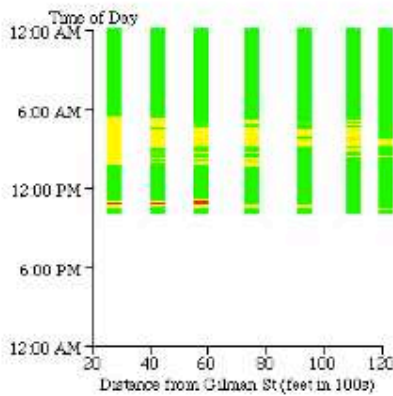


Density Contour Map for I-80

- Education
- Research
- Publications
- The Library
- People
- Centers
- Tech Transfer
- Computing
- Labs
- Contact Us
- Home

Congestion between Gilman and Powell, westbound, for Today

To Oakland/SF >>>>>



This chart shows the congestion level at all 7 stations over the course of one day for one direction. A blank chart indicates we have no data for that day.

Times are listed from top to bottom.

Green denotes free-flow conditions - less than 40 vehicles per hour per lane

Yellow denotes densities from 40 to 70 vehicles per mile per lane

Red denotes congestion - 70 or more vehicles per mile per lane

White denotes missing data. ([Why is it missing?](#))

Select a new day or side below to see another chart.

Direction: North or East South or West

Day: (enter today, yesterday or a date in the form mm/dd/yy - we have data for most days since May 1, 2000)

[Back](#) to the chart selection page

[Back](#) to the I-80 Freeway Data page

Please send any questions, comments, suggestions on this or any of the I-80 freeway data pages to Randall Cayford (cayford@its.berkeley.edu).

© Copyright 2000 Regents of University of California. Last updated on 9/19/00.

Figure 2-2 Typical density contour map



Figure 2-3 Typical monthly overview of density contour maps

The main system components not confirmed to be operational by a positive result of the test above include the 30s generator and the data feed to Caltrans. The computer which processes the 30s data and makes it available to Caltrans responds to status queries; a desktop workstation polls this computer several times per minute and archives the current 30s data feed status.

A dynamic Web-based status page is also used to confirm modem connectivity and data availability status on a station-by-station basis (Figure 2-4).

2.3 Failure Modes

Over the course of the project, there have been several different types of system component failures that have been detected by the monitoring procedures. These are described below.

Controller failure

If a CDPD modem is seen to be on-line, but no data is coming from that station, the cause is typically a controller failure. When the loop controllers at the highway stop functioning, they require a cycling of power to return to normal operation. Also, on several occasions, there have been errors related to cabinet maintenance. For example, controllers have been switched off and left off. On another occasion, the detector cards had been removed from a controller but were not reinstalled. These controller-related difficulties require a physical visit to the cabinet at the highway.

Institute of transportation studies



BHL Status

search go

- [Education](#)
- [Research](#)
- [Publications](#)
- [The Library](#)
- [People](#)
- [Centers](#)
- [Tech Transfer](#)
- [Computing](#)
- [Labs](#)
- [Contact Us](#)
- [Home](#)

	Station_1	Station_2	Station_3	Station_4	Station_5	Station_6	Station_7	Station_8	Station_9
Network status at Tue Jul 23 12:45:00 2002									
Receiving data	(0 seconds)	(0 seconds)	(0 seconds)	(0 seconds)	(0 seconds)	(0 seconds)	(0 seconds)	(0 seconds)	(0 seconds)

Vehicle reidentification algorithms developed by [Benjamin Coifman](#) and [implemented](#) in real time by [Randall Cayford](#). For more information on this work, [click here](#).

Please send any questions, comments, suggestions on this or any of the I-80 freeway data pages to [Randall Cayford \(cayford@its.berkeley.edu\)](mailto:Randall.Cayford@its.berkeley.edu).

© Copyright 2000 Regents of University of California. Last updated on 6/09/00.

Figure 2-4 Modem connectivity and data availability status

Late in the project, the controller at station 3 began to require power cycling after running for several days or several weeks. A lamp timer, designed to automatically cycle the power once per day, was installed at this station in an effort to ensure that this controller remain operational without requiring frequent visits to the highway.

Power outages

Power outages have occurred both at the highway and at the U.C. Berkeley campus. These have included scheduled outages (e.g., highway maintenance) as well as unplanned failures. Most components of the data system come back on-line automatically after a power failure. However, since computer hard drives can be damaged when shut down unexpectedly, there is always concern that one of the data collection or processing machines may have difficulty after a power failure. When power is restored, all computers are checked to ensure proper functioning.

Software-related computer failures

During the course of the project, most of the data collection and processing computers have “crashed” at least once. An outage associated with any particular computer will have a specific and unique impact on the data collection and processing system. BHL team members use the system status tools to identify the location of such an outage, and physically visit the machine to address the outage. In most cases a software or system restart returns normal service. (In the case of a database outage, other data system components must be restarted after the database returns to service so that database connections can be re-established.)

Network outage

Each main component of the BHL data collection and processing system is connected to other components via Internet connections. The first step in the data flow (the CDPD modems)

is wireless; the remaining connections are via physical networks. During the course of the project, there have been several brief network outages on the U.C. Berkeley campus. Given the particular network topology at our data processing site, these outages isolate the Data Collection Server from the network, isolate the data storage/processing machines from the network, or both. (For further discussion of network availability, see Chapter 9.)

2.4 Future Plans

Our work to date suggests a number of possible future directions. Future work may include the construction of a flow chart or decision tree that allows a system operator to evaluate the symptoms of a data loss and arrive quickly to an indication of which system component is likely to have failed. A problem with a specific component of the BHL data system will have specific (and usually predictable) results; this situation lends the BHL monitoring system to being interfaced with an automated “self diagnosis” routine. For example, an outage of the Data Collection Server would cause a loss of all incoming data. If the Web server becomes unavailable, the data collection, processing, and archiving all continue to operate, but views of current and historical data are temporarily unavailable. A database outage interferes with data processing, but raw data continues to be stored by the Data Collection Server. If combined with computer-based system status tools, this automated decision tree could be coded into system monitoring software, which could alert system operators of specific system problems and likely solutions.

Most of the BHL system status tools report the status of specific system components upon request. It would be useful to record a history of the status of all system components so that transient problems or problem patterns could be more easily identified and addressed. System status data to be archived might include daily data packet losses for each modem, interruptions in data requests from Caltrans District 4, network outages between system components, and server availability.

3. Generate and Archive Summary Data Sets

3.1 Data generation

At the start of this project the BHL server collected and archived five data sets:

- Raw, bitpacked detector data
- Sorted, bitpacked detector data
- Measured travel times
- 30 second aggregate by lane
- 5 minute aggregate by station

The raw data are stored in six hour long intervals and include all stations. The following is a sample of seven lines from this data set.

```
983693022000 983693017120 1351 40584 3 270B3D65D174C86D8AAB0DB4ABE9E4937
983693022000 983693026730 1009 38223 1 270BD0FAE7DA0D87E7C39
983693021000 983693027110 281 75085 5 2702D103DC3FF7DF70
983693023000 983693026620 1256 -80516 2 2713B632305505B18D14A
983693022000 983693016400 710 148 7 270004A117962A34B61CEFDB5E14E8235
983693023000 983693017010 2278 -3749 4 2700F8425D0C
```

The first two columns are the time in the field (1/1000 sec) corrected for clock drift since restart and uncorrected, respectively. These values were originally relative to the PC clock in the field, which was reset once a day when the computer underwent a forced hard restart. The two columns were used because the PC clocks would drift significantly throughout the day. Now, however, the PC has been eliminated from the loop and CDPD modems are being used. So the times are generated in the server based on when the packet arrives. The next column shows the cumulative number of PC restarts and was used to determine if the station has restarted. The fourth column shows the offset (sec) of the controller clock, i.e., the value to add to the controller clock to generate the correct time. The controller clocks do not drift significantly from one day to the next but may show significant drift after power outages and other unusual events. By

tracking the offset, the need to visit the cabinets and reset the clocks has been eliminated. The fifth column indicates the station number and the sixth column contains the encoded controller data. The encoded data are sent each second and include the controller time and information about each observed transition (detector number, state and time).

The sorted data followed the same format as the raw data; however, there are several differences in the content. The data from each station are stored in different files. Unlike the raw data, within each file the data are sorted by time and any duplicate packets are removed.

The measured travel times include the link number, lane, direction, travel time and time that the reidentified vehicle exited the link [1-2]. The aggregate data include station, time, lane, flow (veh/hr), occupancy (percent), and velocity (mph) averaged every 30 sec.

Over the course of this project we have added three new data sets to the archives.

- Daily summary images
- Diagnostic data
- Individual vehicle data

The daily summary images provide an efficient overview of the BHL throughout the day. Figure 2-3 shows a sample of these images. Distance along the facility is shown on by the abscissa and time of day by the ordinate. Since each detector station is at a fixed location, the data fall in discrete columns. In this figure green indicates less than 40 vehicles per mile per lane, yellow is 40 to 70 vehicles per mile per lane, and red represents 70 or more vehicles per mile per lane. Any period without data is shown in white. Inspection of these plots shows congestion growing

from downstream and then receding in the opposite direction, both from recurring bottlenecks and incidents.

Several diagnostic tools were implemented to assess the performance of the surveillance and communication system, as discussed in section 2-2. These tools alert the BHL team to any new hardware problems, allowing for a rapid response. Researchers can benefit from this information to assess the completeness of collected data and the results of detector diagnostics are now being archived.

Figure 8-3A shows a time-space diagram depicting a vehicle passing over a dual loop detector. The controller normally records four transitions, i.e., the turn-on and turn-off times at each of the loops, as shown in Figure 8-3B. These measurements are used to calculate: dual loop traversal time via the rising edges, TT_r , dual loop traversal time via the falling edges, TT_f , total on-time at the first loop, OT_1 , and total on-time at the second loop, OT_2 , as indicated in the figure.

Previously, these measurements are discarded after being used for vehicle reidentification. Now, all of the vehicle transitions are retained after being extracted from the controller packets. It is worth noting that occasionally a vehicle will change lanes over the dual loop detector, one of the loop detectors will malfunction or a data packet is not received. All of these events will impact subsequent analysis. Currently these unmatched transitions are discarded. This individual vehicle data consists of:

- Station
- Lane (0-9)
- Upstream loop on time
- Upstream loop off time
- Downstream loop on time
- Downstream loop off time

All times are in number of 1/60 second intervals since midnight; if a vehicle enters a dual loop detector before midnight and exits after midnight, the entry time will be expressed as a negative number to represent the time interval before midnight.

A planned enhancement to the individual vehicle data set will include all transitions seen even if they don't match a complete vehicle. A status code will be added to differentiate complete vehicle records for incomplete records. The status data allows a user to decide whether they want to only use the data that have been grouped into vehicles or if they want to include the small amount of data that was excluded from the grouping. This fact is important since these data will be used to generate conclusions about traffic flow theory. Any cleaning done by BHL team members without telling users can lead to erroneous conclusions, e.g., a partial packet might fall in the middle of an otherwise long headway, indicating that a second vehicle may have passed but full information on that vehicle is unavailable. Of course many users will simply be interested in the complete data and can easily filter out the partial data using the status flag.

3.1 Data archiving

The rate of data generation by the BHL data processing system began to exceed 4 Gigabytes per month at the conclusion of this project. This data must be removed from the data

servers so that their local storage capacity is not consumed by excessive amounts of data, but it must be stored for future analysis.

At the beginning of each calendar month, data is transferred from the data collection and data processing computers to a desktop workstation. These data files are compressed, and the compressed files are copied to CDR media. Duplicate copies are created so that data loss is less likely. After duplicate data CDs have been produced for data from a given month, the original copies of the data are removed from the servers so that there is sufficient storage space for new data. A list of archived data appears below.

Raw, bitpacked detector data	July 1999 – present
Sorted, bitpacked detector data	July 1999 – present
Individual vehicle travel times	July 1999 – present
30 second summary data	May 2000 – present
5 minute summary data	May 2000 – present
Individual vehicle data	January 2002 - present
Daily summary images	generated December 2001 - present
Monitor log data	October 2001 - present

4. Distribute Surveillance System Data

Sharing the unique Berkeley Highway Lab data with other researchers allows others to conduct additional traffic research beyond the scope of this project. Some Berkeley Highway Lab data is publicly available via the BHL Web site; additional data is made available to other researchers upon request, and with the understanding that Caltrans will be apprised of all data sharing arrangements.

The Berkeley Highway Lab Web site offers current and historical data on demand. This includes:

- 30 second aggregate data for each bi-directional station (Figure 4-1)
- Density contour maps showing stations 1-7 (Figure 2-2)
- Performance charts for each directional station, including:
 1. Speed
 2. Flow
 3. Occupancy
 4. Density (Figure 4-2)
 5. Density-Flow
- Travel times between pairs of adjacent stations (Figure 4-3)

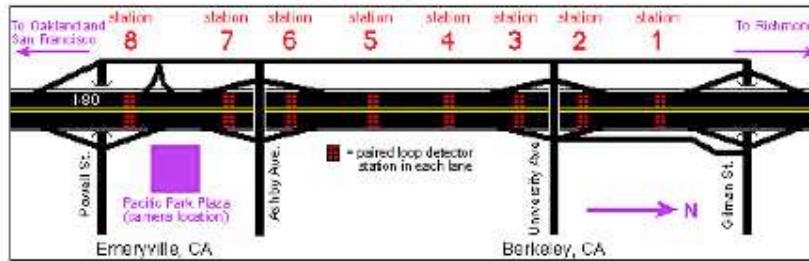
In addition to data presented on the Web, raw and processed data sets are archived (Chapter 3) so that a history of I80 traffic is available to researchers. The BHL Web pages invite traffic engineers to contact BHL regarding data sharing.

BHL data has been used by a number of researchers at several institutions. These include:



30 Second Aggregate Data for Station 2

- Education
- Research
- Publications
- The Library
- People
- Centers
- Tech Transfer
- Computing
- Labs
- Contact Us
- Home



Other Stations: [8](#) [7](#) [6](#) [5](#) [4](#) [3](#) [2](#) [1](#) [Home](#)

Station 2	Westbound to San Francisco/ Oakland					Eastbound to Richmond				
	W5	W4	W3	W2	HOV W1	HOV E1	E2	E3	E4	E5
Flow (vehs/hr)	1,320	1,320	1,140	1,740	720	300	1,680	1,020	1,320	420
% Occupancy	23.5	16	10.5	20	5	3	9	11	8	2.5
Speed (mph)	38	38	38	30	55	76	69	64	59	63

Current Date: 7/31/2002
Current Time: 9:29:00

[Vehicle reidentification algorithms](#) developed by [Benjamin Coifman](#) and [implemented](#) in real time by [Randall Cayford](#). For more information on this work, [click here](#).

Please send any questions, comments, suggestions on this or any of the I-80 freeway data pages to Randall Cayford (cayford@its.berkeley.edu).

© Copyright 2000 Regents of University of California. Last updated on 6/09/00.

Figure 4-1 30 second aggregate data for each bi-directional station



Performance Chart for I-80: Density

- Education
- Research
- Publications
- The Library
- People
- Centers
- Tech Transfer
- Computing
- Labs
- Contact Us
- Home

Density for Station 2, North or East bound

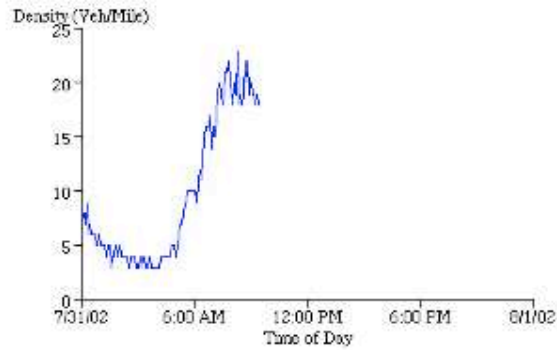


Chart Type:

Station:

Direction: North or East South or West

Day:

Use Cache?:

[Back](#) to the chart selection page

[Back](#) to the I-80 Freeway Data page

Please send any questions, comments, suggestions on this or any of the I-80 freeway data pages to Randall Cayford (cayford@its.berkeley.edu).

© Copyright 2000 Regents of University of California. Last updated on 9/19/00.

Figure 4-2 Typical density chart for a directional station

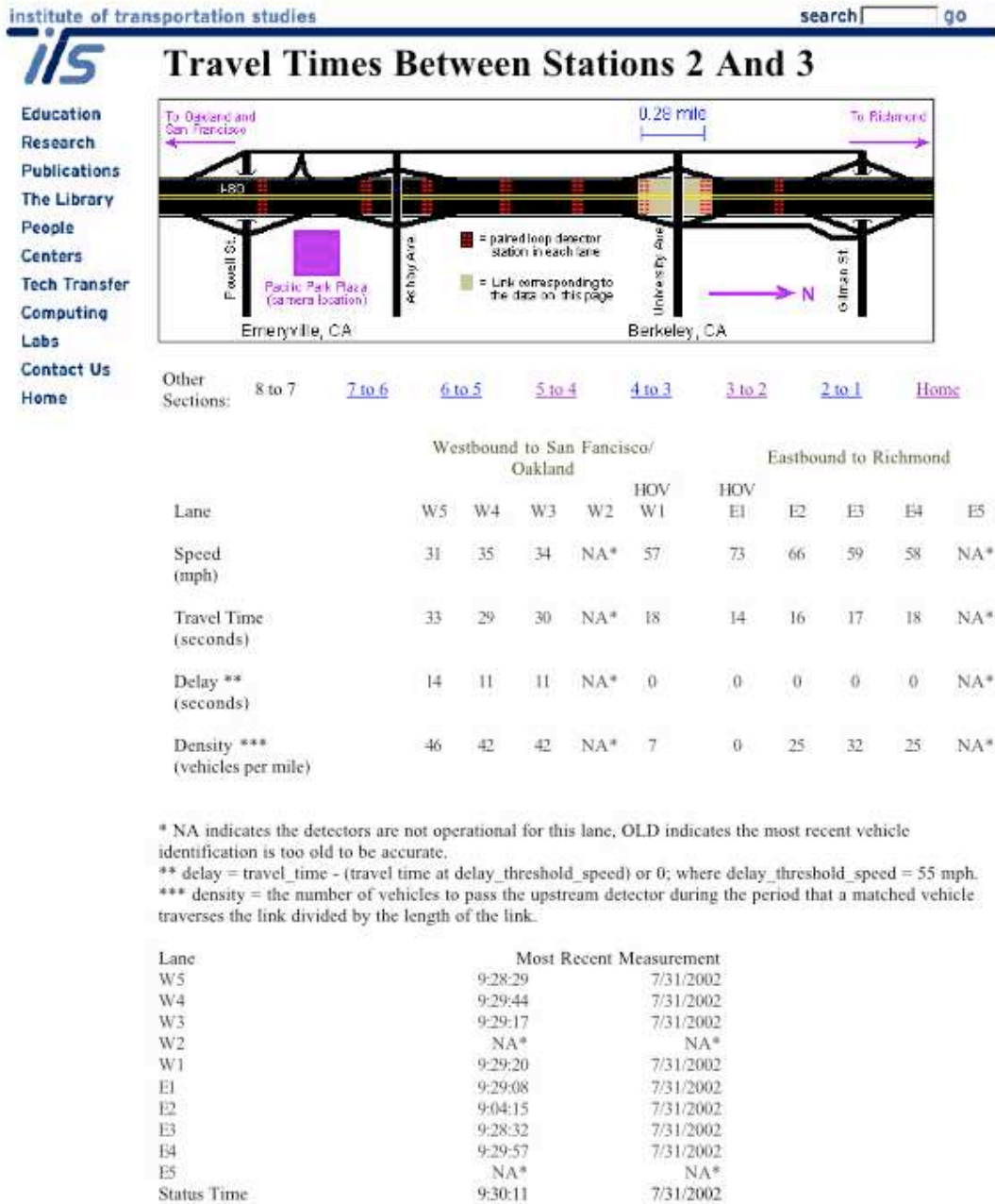


Figure 4-3 Typical display of travel times between pairs of adjacent stations

- Joy Dahlgren, Wei Lin, Carlos Daganzo, and Jorge Laval (U.C. Berkeley) are using BHL vehicle headway data for a project on traffic dynamics and traffic management strategies. To offset some of the costs of BHL personnel time, Dr. Dahlgren has purchased hard disk storage for the BHL servers which enables us to store and archive data more efficiently.
- Taewan Kim and Michael Zhang (U.C. Davis) are using BHL data for modeling microscopic traffic flows, with a focus on vehicle headways under different traffic conditions. They use BHL detector event data to validate their models.
- Oliver Gao (U.C. Davis), with Joy Dahlgren (U.C. Berkeley), is studying individual vehicle length distributions.
- Sue Ahn and Mike Cassidy (U.C. Berkeley) are using 30 second data for a new project. They're also interested in individual vehicle data.
- Edgar Ergueta (U.C. Berkeley), with Ben Coifman, is using BHL data to develop improved vehicle travel time calculation algorithms.

As long as the lab continues to operate with a scope similar to the current project, BHL data available will be made available to other researchers on a case-by-case basis and with the approval of Caltrans. The BHL research team will continue to evaluate with Caltrans the importance of the project and announcing the availability of data sets.

A goal of future work might include user-accessible data archived presented on the Web, and software designed to extract specific subsets of data from raw data files.

5. Maintain Data Link To District 4

Caltrans collects 30 second data from highway loop detectors statewide. The Berkeley Highway Lab, in addition to collecting 1/60 second data, generates 30 second data which is relayed to Caltrans and integrated into the statewide Caltrans IFS2 data collection system. For this data to reach Caltrans, the dedicated data communication line between the Berkeley campus and Caltrans District 4 must function, and 30 second data must be made continuously available over this link.

The data link that allows 30 second data to travel from the BHL system to Caltrans District 4 is supported by a frame relay connection. This is a virtual point-to-point link with robust error correction, and has performed very reliably during the course of the project.

A Caltrans server at the remote end of the frame relay connection continuously sends data requests to the BHL Caltrans D4 Link Server, which provides the 30 second data in response to each request. The BHL team monitors the Caltrans data requests and can inform Caltrans in the event that data is not being pulled from the Berkeley Highway Lab.

The frame relay data line between the Berkeley campus and District 4 is highly reliable; several outages of approximately several hours each have occurred in the past year, and these have been due to electrical power failures in the campus building where the BHL data collection system resides. Service on this line is estimated to exceed 99% availability.

As part of this project, a special status reporting function has been built into the BHL data server that makes the 30 second data available to Caltrans. This reporter generates the following output when a status request is sent to it:

Last Caltrans read was: 1027595390000
Last database read was: 1027725288493
Last clock time from database was: 1027725210000

These time values represent the last time that the Caltrans server requested data, the last time that the BHL data feed grabbed current data from the BHL database, and the time of the data measurements themselves. The times are shown in elapsed 1/1000 seconds of “epoch” time, and can easily be compared with the current time to generate the age of the most recent event of each type.

Additional code running on another computer has been written to run continuously, polling this reporter several times each minute and record the status in a log file. This log file can be scanned for historical outages and other transient problems.

This frame relay data link performs well, and there are no plans to replace or modify this portion of the BHL infrastructure. Future work may see this data link status built into an automated system trouble warning system as described in Chapter 2.

6. Develop Online Detector Diagnostics

The accuracy of detector data is paramount for traffic control and information systems, yet little has been published on identifying or correcting errors in the data. Many practitioners and some researchers have worked to automate tests of aggregate data by asking the question, "Are the time series 30 second average flow and occupancy within statistical tolerance?" [3-5] These systems often go undocumented in the literature because they are either designed in-house by an operating agency or were developed by a consulting firm using proprietary information.

Chen and May [6] developed a new approach for verifying detector data using event data, i.e., individual vehicle actuations. Their methodology examines the distribution of vehicles' on-time, i.e., the time the detector is occupied by a vehicle. Unlike conventional aggregate measures, their approach is sensitive to errors such as "pulse breakups", where a single vehicle registers multiple actuations because the sensor output flickers off and back on.

Coifman [7] went a step further and compared the measured on-times from each loop in a dual loop detector on a vehicle by vehicle basis. At free flow velocities the on-times from the two loops should be virtually identical regardless of vehicle length, even allowing for hard decelerations. Many hardware and software errors will cause the two on-times to differ. At lower velocities, vehicle acceleration can cause the two on-times to differ even though both loops are functioning properly and thus, congested periods were excluded from the earlier analysis. Coifman and Dhoorjaty [8] developed several additional tests for single and dual loops.

To verify the performance of the surveillance system, many microscopic tests have been developed on this project to assess the quality of the data from the loop detectors. Ultimately, five of these tests were deployed, as discussed in Chapter 7, four of which are compatible with

single loop detectors and they are applied in real time to the 164 loops (each loop of the 82 dual loop pairs) in the BHL. While one test compares information for the paired loops in a dual loop detector, it is applied to all of the dual loop pairs in the BHL.

6.1 Development of single loop tests

The single loop tests are based on the fact that very little information is available on any given vehicle. But there are some bounds on the data that is available. A single loop detector records each vehicle's on-time, i.e., the time that the given vehicle occupies the detector. This on-time, OT , is defined by the following relationship,

$$OT = L / v \tag{6-1}$$

where on is the on-time, L is the vehicle's effective length¹ and v is the vehicle's velocity as it passes the detector. In other words, $OT(L,v)$ depends on physical characteristics of the vehicle and traffic conditions. Unlike single loops, dual loop detectors can measure v by taking the quotient of loop separation and TT and thus, allowing for direct measurement of L . Figure 6-1 shows the cumulative distribution of vehicle lengths as measured at a dual loop detector over an entire day. During free flow conditions, v is relatively constant and on-time is roughly proportional to L , as shown in Figure 6-2. The longest on-times correspond to 80 ft trucks traveling at free flow velocities. Congested data are excluded from this figure because the

¹ Effective vehicle length is the length as seen by the detector and may include the detection zone in addition to the physical vehicle length. Throughout this section *vehicle length* or simply *length* are sometimes used as shorthand for *effective vehicle length*.

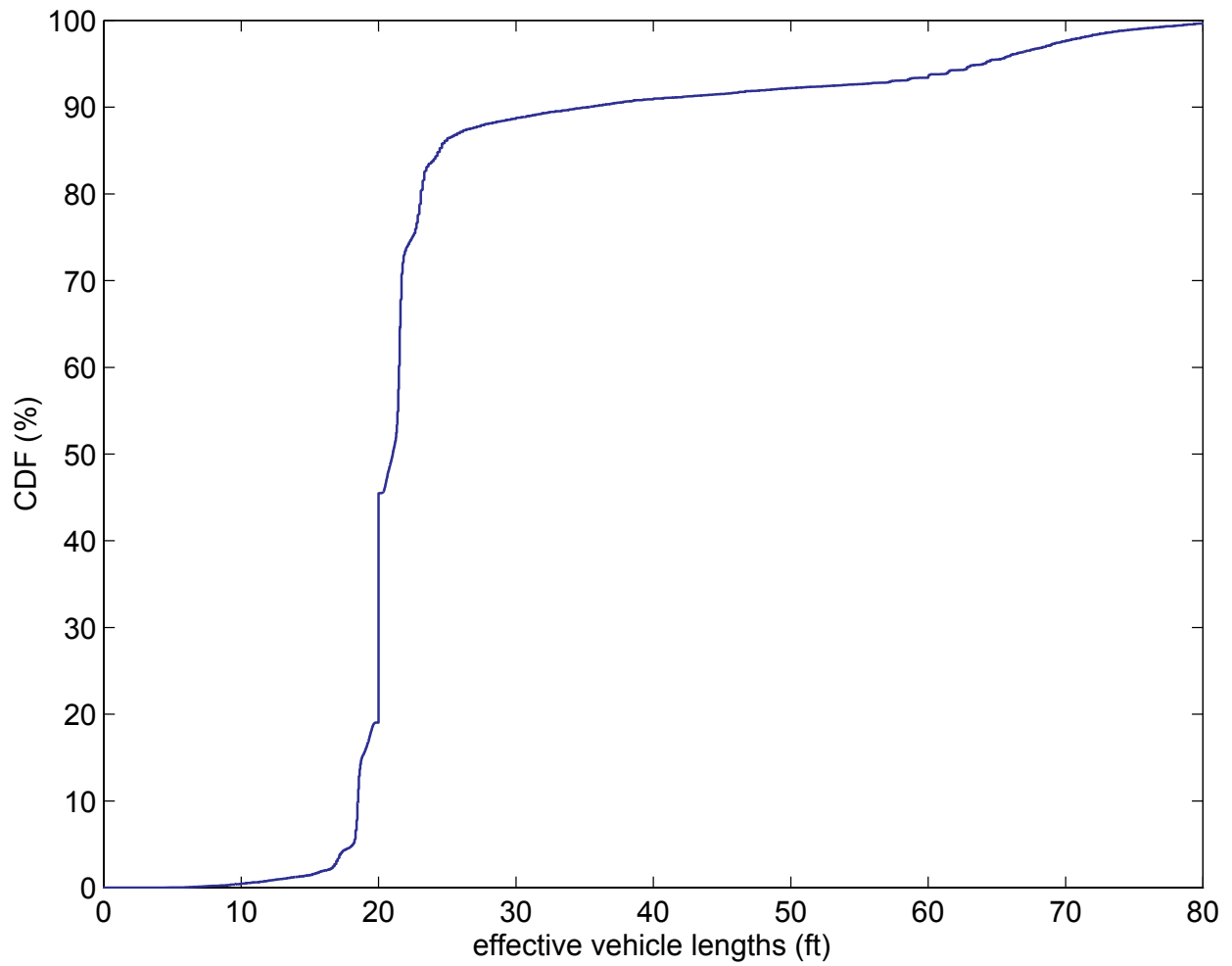


Figure 6-1 Cumulative distribution of effective vehicle lengths on a single day at a dual loop detector

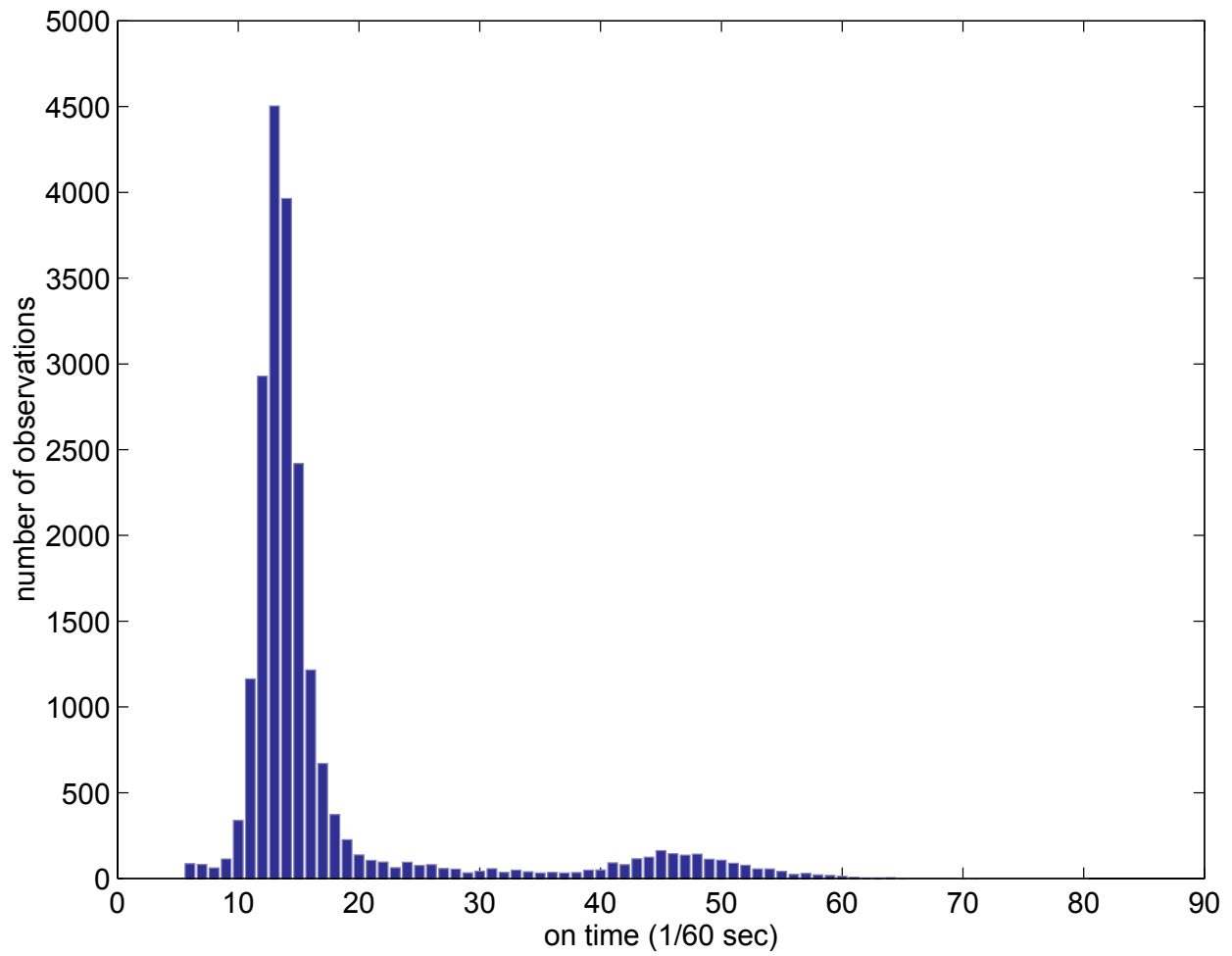


Figure 6-2 Distribution of on-times during free flow on a single day at a dual loop detector

simple relationship breaks down, e.g., some of the observed on-times were up to 10 sec long simply because the vehicles slowed or stopped over the detector. Without the redundancy of a dual loop detector, it is more difficult to separate the influence of L and v in the on-time measurement. Conventional velocity estimation from single loops uses the equation,

$$v_{est_conventional} = L_A * \text{flow} / \text{occupancy} \quad (6-2)$$

where L_A is the assumed average effective vehicle length at the detector. This equation is inversely proportional to average on-time. Work by our group [9] has demonstrated that,

$$v_{est} = L_A / \text{median}(OT) \quad (6-3)$$

is less sensitive to outlying OT measurements. Provided the sample size is large enough, the accuracy of Equation 6-3 approaches that of the average velocity measured from a dual loop detector. For example, [9] found that for the subject data, a median of 11 OT measurements centered on the given vehicle yields an average absolute error under 3 mph. With this equation, one can also estimate vehicle lengths, L_{est} , as follows,

$$L_{est} = L_A * OT / \text{median}(OT) \quad (6-4)$$

Obviously, the lower bound of Equation 6-1 is defined by the shortest feasible L passing at the highest possible v . There is no upper bound during congestion, but if one determines the

presence of free flow traffic using Equation 6-3, then the upper bound is defined by the longest L and the mode is defined by the most common L . These observations lead to the following tests:

Minimum On-time: One cannot expect any speeds greater than certain maximum speed (which corresponds to an on-time). The percentage of on-times that are greater than the cut-off on-time is calculated. A significant percentage of low on-times would therefore indicate some error in the loop. An on-time measurement is accepted as valid in the min on-time test at single loops if it is greater than 7/60 sec. The acceptable velocity and length region is shown in Figure 6-3A and the threshold selection was based on measured length distributions similar to Figure 6-1. Two alternative thresholds are shown with dashed lines for reference. Of course loop detectors are not perfect and the occasional error is expected even from perfectly functioning loops. So this test fails if more than 3.5 percent of the pulses in a sample fall below the threshold on-time. Currently the test is applied to contiguous samples of 100 pulses.

Maximum On-time: Similarly, at the other extreme, almost all the on-times during free flow are expected to be less than the maximum on-time. An on-time measurement is accepted as valid in the max on-time test at single loops if it is less than 75/60 sec and the current velocity estimate from Equation 6-3 is above 45 mph. The acceptable velocity and length region is shown in Figure 6-3B (note scale change from part A). Any periods in which the velocity estimate falls below 45 mph should be excluded from this test. We have not yet implemented Equation 6-3 to estimate velocity in real time, so a larger threshold of 700/60 sec is currently used, as shown in Figure 6-3C. This larger threshold will accept even 80 ft vehicles traveling at

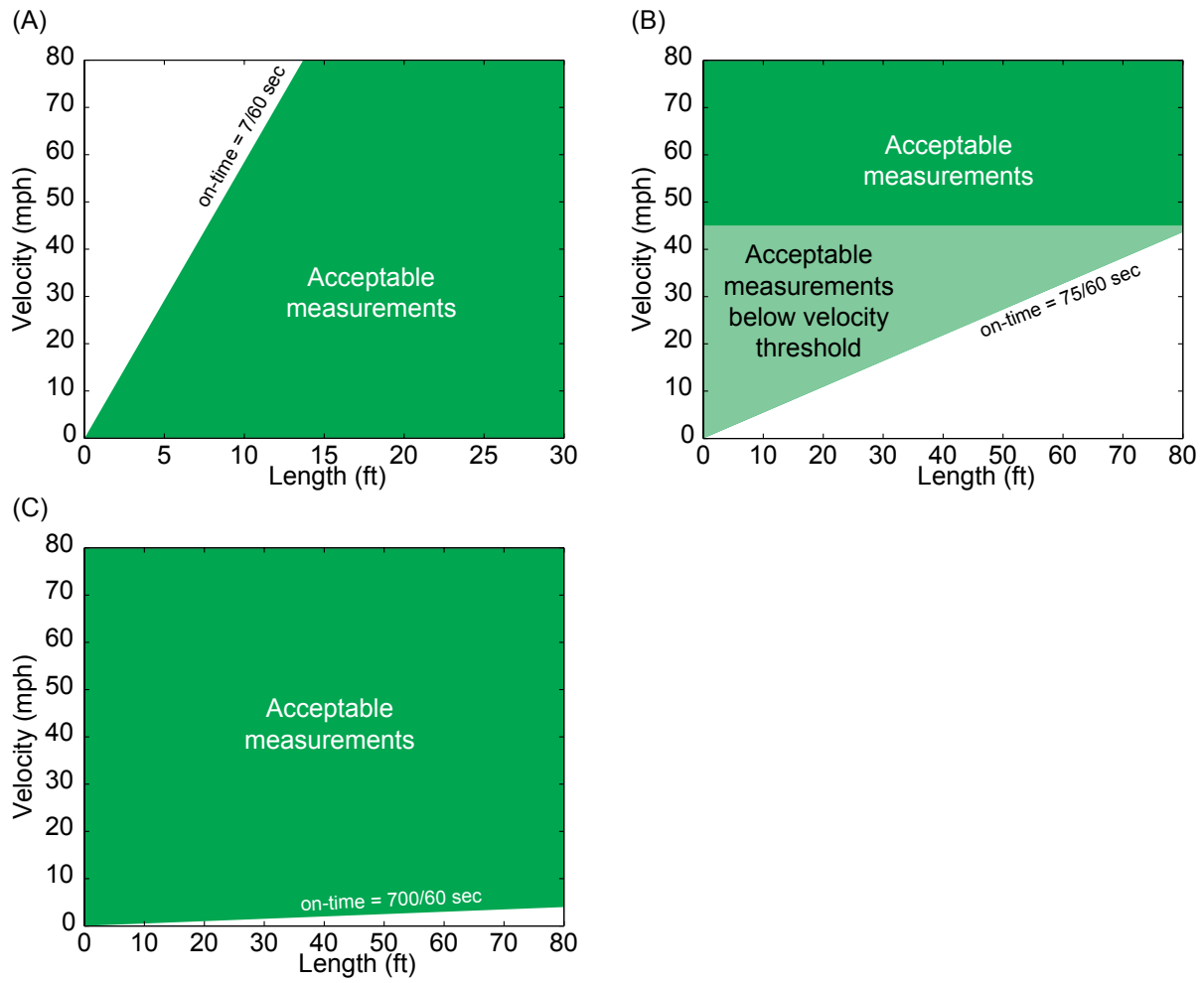


Figure 6-3 Acceptable length and velocity combinations for min and max on-time tests
 (A) Min on-time test (B) Max on-time test (C) max on-time without velocity check

5 mph. The test fails if more than 3.5 percent of the pulses in a sample fall above the threshold on-time. Once more, currently the test is applied to contiguous samples of 100 pulses.

Mode On-time: For the mode on-time test at single loops, one must first generate the distribution on-times, e.g., Figure 6-2. This test collects on-time measurements from 1,000 consecutive free flow pulses (discarding any intermediate pulses labeled as being from congested conditions via Equation 6-3). The larger sample size is used to ensure that the distribution will not be skewed from any transient measurements. If any intervening pulses correspond to a velocity estimate below 50 mph, they are ignored and the sampling is resumed when free flow conditions return. Using the current acceptable values of OT , greater than 10/60 sec and less than 16/60 sec, the acceptable velocity and length region is shown in Figure 6-4. Note that the threshold selection was based on measured length distributions similar to Figure 6-1. Once 1,000 pulses have been observed, test passes if the mode of the distribution falls in the acceptable measurement region. The mode effective vehicle length (unobserved) is expected to be near 20 ft, given the free flow constraint, the mode on-time (observed) should correspond to one of these vehicles. The on-time constraints do not explicitly account for velocity, so the region above 50 mph is shown darker than the acceptable region below the threshold. Few true measurements should fall in the lighter region and those that do should normally be discarded by the velocity test.

Once more, we have not yet deployed the velocity estimation, but one can not work around the fact that velocities are unknown in this test. So, the test is expected to fail if a significant portion of the observed vehicles come from congested conditions without being

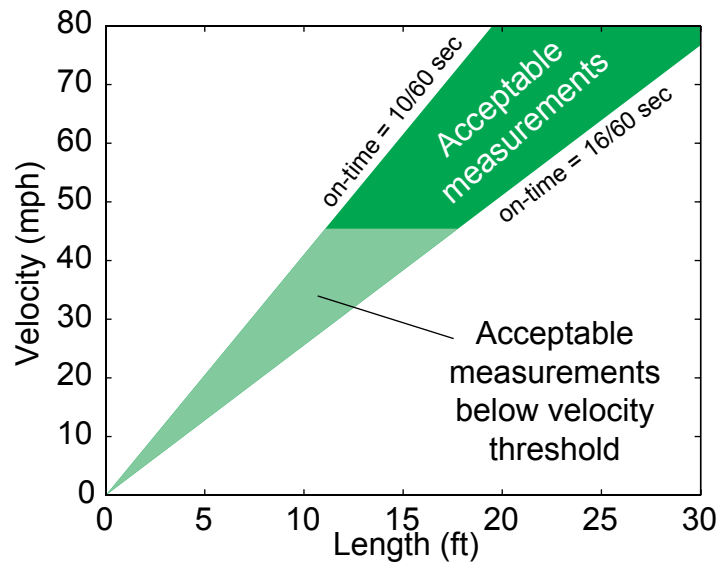


Figure 6-4 Acceptable length and velocity combinations for mode on-time test

eliminated with the velocity test. It should, however, yield correct results when most of the observed vehicles come from free flow periods even without the velocity test.

If deployed on a large scale, there is a need for a second test to make sure that *hanging on* errors do not fool this test into thinking that it never sees free flow. The simplest solution would be to keep track of how many pulses were discarded due to the velocity threshold. Alternatively, one could calculate the mode both for all vehicles and just for the free flow vehicles.

Minimum Off-time: Drivers maintain certain "safe" gap between vehicles and one can find a critical gap (off-time) such that all drivers exceed this value. Any gaps below this critical value should not be feasible. A loop with a significant value of low off-time percentage will be a suspect. Typically, low off-times can be attributed to pulse breakups, tailgating or flicker. Preliminary investigations suggest the use of a threshold value on the order of 10/60 sec. This test has not yet been implemented in real time.

Mode Off-times: During moderate to high flows, the mode on-time should fall within a reasonable window of feasible values. We suspect that the diagnostic power of this test will be surpassed by the other single loop tests. This test has not yet been implemented in real time.

Estimated Minimum Length: Following logic similar to the on-time tests, the effective vehicle length estimated from Equation 6-4 should not be less than a minimum value. The percentage of vehicles with estimated lengths under a given threshold can be calculated and a large percentage

of low (infeasible) vehicle lengths would indicate a problem in the loop. This test has not yet been implemented in real time.

Estimated Maximum Length: This is the analog to the previous test. Unlike on-time, though, the estimated vehicle length is less sensitive to vehicle velocity. So it is possible to specify the upper threshold for all traffic conditions and then track the percentage of measurements that correspond to infeasible vehicle lengths. This test has not yet been implemented in real time.

Single Loop Operation: Because all of the on-time tests are event driven, we have to account for the fact that the tests will not report an error if the detector stops sending data. A single loop detector passes this test if a pulse is detected in the preceding time period, currently set to 15 min. This relatively simple test verifies that a detector is sending data. It also reduces the need for a maximum on-time test or maximum off-time test to catch stuck detectors. A stuck detector would likely result in a single long observation, but long off times are to be expected during early morning hours. As an alternative to this test, the on-time tests could also be applied over fixed periods, e.g., a day, rather than a specific number of actuations, but that would increase the need to track long on or off-times.

6.2 Development of dual loop tests

Of course the BHL consists of dual loop detectors and the redundancy between the loops can be used to compare the performance of one loop against the other since each vehicle is observed twice, once at each loop. The two observations are normally used to measure velocity,

but the redundancy can also be used to assess the performance of the dual loop detector and identify detector errors.

On-time Difference: At free flow velocities, the on-time should be virtually identical between the two loops, regardless of the vehicle length. Many hardware errors will cause the two on-times to differ, but below free flow velocity, acceleration can cause the two on-times to differ even though the detectors are working properly. Exploiting these facts, [7] developed a formal methodology for testing dual loop detectors off-line. In this project we have updated the test and implemented it in real time through out the BHL.

After matching pulses from the same vehicle between the two loops we measure v . Note that the simplest pulse matching is recommended, so that the matching does not accidentally eliminate any detector errors. Instead, these errors should be preserved so that they will appear in the test set. Furthermore, to include any transient errors that may impact individual measurements of v , we take a moving median of 11 samples centered on the subject measurement. If the median is greater than 50 mph, then we calculate the difference between the two on-times and update the observed distribution. After 1,000 free flow vehicles are counted, the number of observations that have a difference between $-3.5/60$ sec and $+3.5/60$ sec inclusive is counted. This sample size is necessary to ensure that transient but expected events, such as the occasional unmatched vehicle, will not cause the test to fail. If fewer than 5 percent of the observations fall outside this region then the loops pass the test. Figures 6-5 and 6-6 illustrate the test applied to a good and bad dual loop detector, respectively.

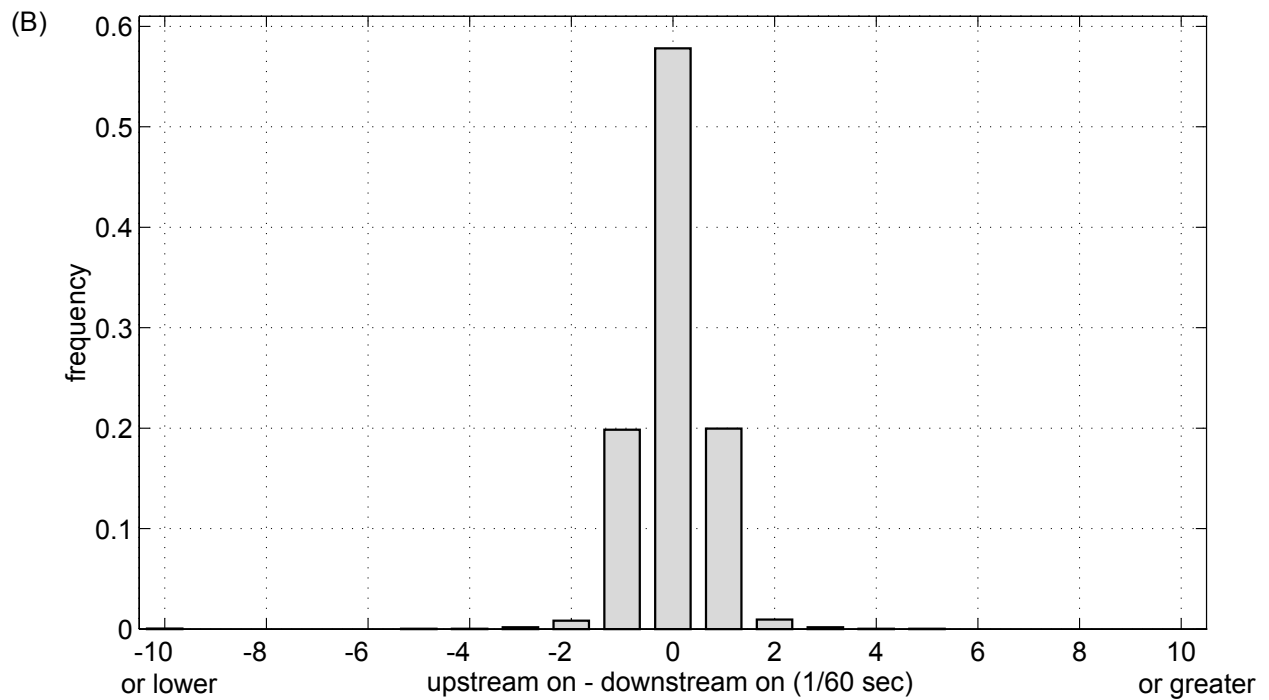
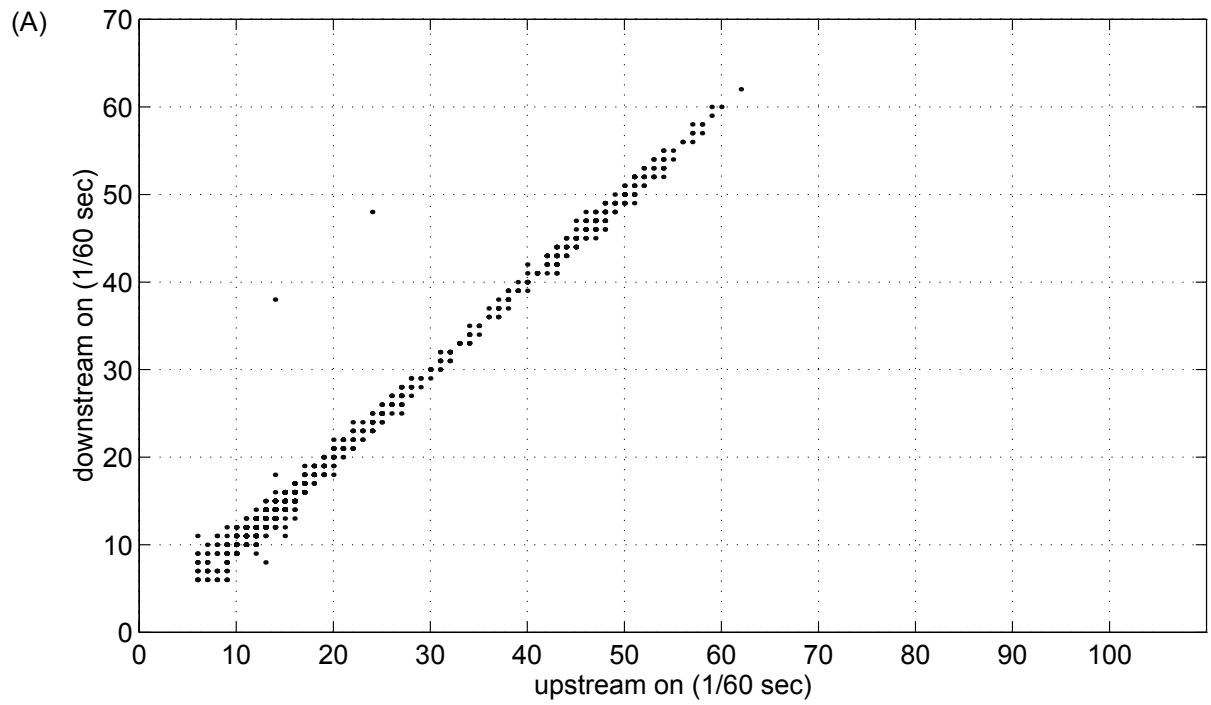


Figure 6-5 On-times from a good dual-loop detector

(A) Scatter plot of on-time for 3,067 matched pulses from a good dual loop detector (B) Distribution of $OT_1 - OT_2$ for the same pulses.

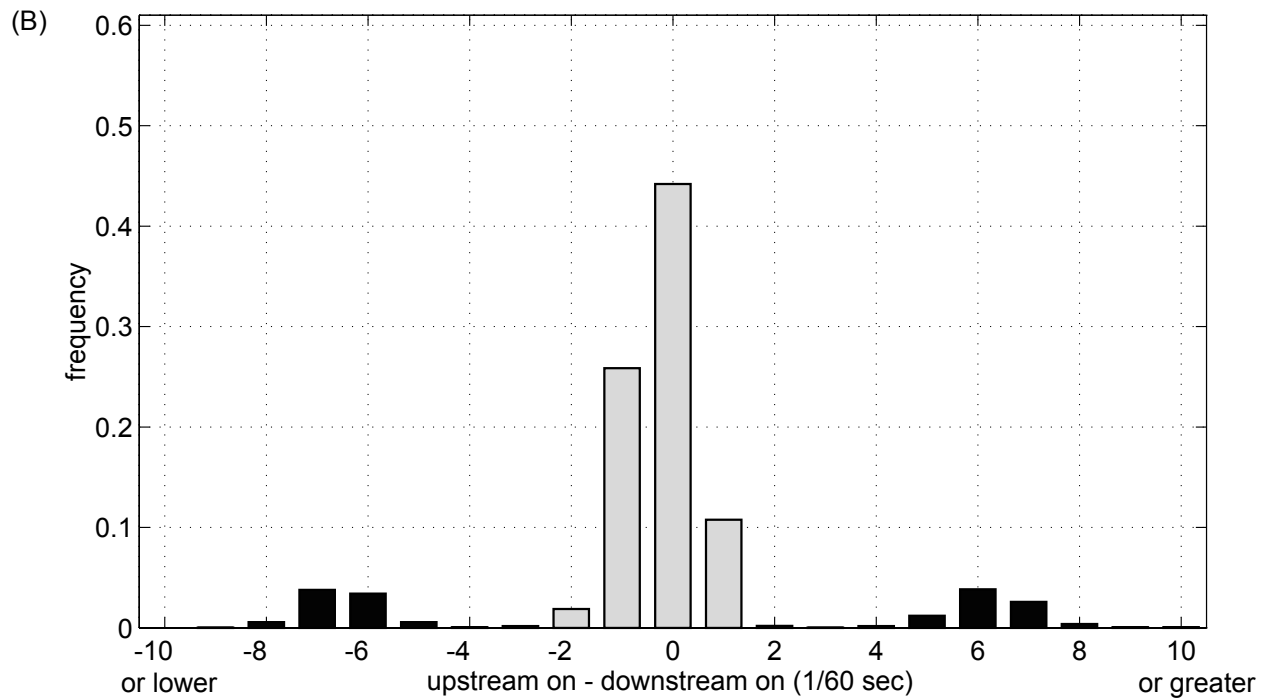
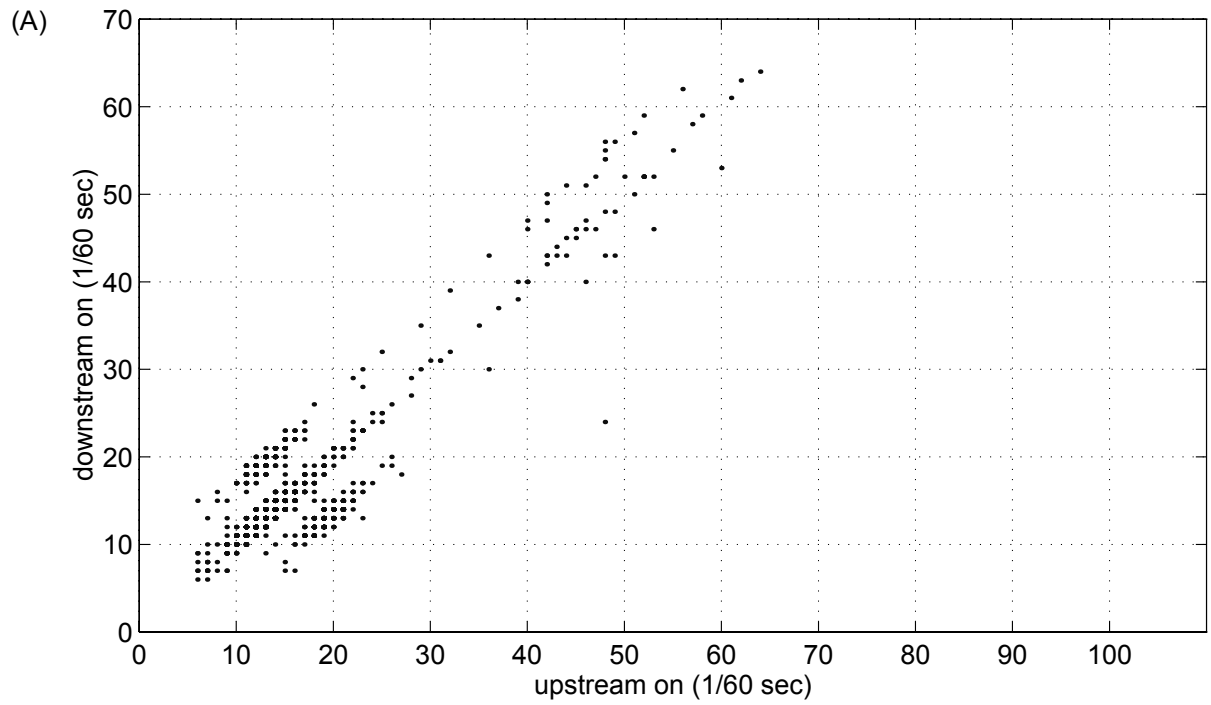


Figure 6-6 On-times from a bad dual-loop detector
 (A) Scatter plot of on-time for 2,803 matched pulses from a bad dual loop detector (B) Distribution of $OT_1 - OT_2$ for the same pulses.

Off-time Difference: This is similar to the above test but uses off-times instead of on-times. The one key difference is the fact that in the presence of a detector error, two vehicles may yield the same on-time and pass the previous test if they were incorrectly matched, but they would have significantly different off-times. This test has not yet been implemented in real time.

Minimum Distance: This test examines the measured distance between vehicles, i.e., the product of off-time and velocity. As with the minimum off-time test, drivers maintain certain "safe" distance between vehicles and one can find a critical distance such that all drivers exceed this value. Any gaps below this critical value should not be feasible. A dual loop with a significant percentage of low off-times will be a suspect. This test complements the single loop Minimum Off-Time test by accounting for measured vehicle velocity. Thus, the passing criteria can be tighter than the single loop test and in the presence of detector errors, the measured velocity may lead to a greater number of vehicles failing this test. This test has not yet been implemented in real time.

Measured Minimum Length: This test is identical to the single loop Estimated Minimum Length test, except the measured L are used in place of L_{est} . This test has not yet been implemented in real time.

Measured Maximum Length: This test is identical to the single loop Estimated Maximum Length test, except the measured L are used in place of L_{est} . This test has not yet been implemented in real time.

Moving Median Velocity Difference: This test compares the individual vehicle velocity against the median of 11 velocity measurements centered on the given vehicle. If the difference between the two exceeds a preset threshold, the velocity measurement is considered erroneous. This test can be used to eliminate transient errors in velocity and tracking the frequency of errors can be used to assess the performance of the dual loops. This test has not yet been implemented in real time.

Mode On-time Difference: The single loop Mode On-time test considers the performance of each loop individually, and the dual loop On-time Difference test compare the performance of each loop on a vehicle by vehicle basis. After observing a large number of free flow vehicles, the mode on-time at each loop in a dual loop should be identical. Many detector errors would cause the two measurements to differ. This test has not yet been implemented in real time.

Mode Off-time Difference: This is similar to the above test except that it uses off-times. This test has not yet been implemented in real time.

Median On-time Difference and Median Off-time Difference: These tests are similar to the respective mode difference tests, except for the fact that these tests should also apply during congested conditions. These tests have not yet been implemented in real time.

6.3 Tests implemented in real time

Roughly half of the tests were designed to be compatible with either single loops or dual loops, while the remainder of the tests require the redundancy of dual loops. Out of the 18 tests, the following five were selected for real-time implementation,

- Single loop operation
- Min on-time test at single loops
- Max on-time test at single loops
- Mode on-time test at single loops
- On-time difference test at dual loops

Min and Max on-time assure that the single loop measurements are within the feasible range of measurements. The mode on-time test further verifies that the center of the distribution is in the expected range. The on-time difference exploits the redundancy of the two loops. Finally the single loop operation test assures that the loops are sending data. Refer to chapter 7 for implementation.

6.4 Future directions

One of the three major tasks in next year's project deals with detector diagnostics. The first step will be assessing the experience gained with the detector diagnostics that have been implemented this year. The planned work for next year consists of improvements in some of the

implemented detector diagnostics and the implementation of additional detector diagnostics. The following paragraphs describe some of the detector diagnostics that will be considered, although it may not be feasible to examine all of them thoroughly within the scope of the forthcoming project. Future work, either next year or beyond will include the diagnosing detector problems and proscribing possible solutions.

We would like to implement the single loop velocity estimation and improve the fidelity of the Maximum On-time and Mode On-time tests, since the large tolerance currently used will admit large errors. Related to this improvement, we would need to verify that hanging-on errors do not fool these tests into thinking that they never sees free flow. We would also like to deploy more of the tests to provide a more complete picture of what is happening at the loop detectors.

Through extended experience with the tests, we would like to refine the thresholds used in the tests. Although the BHL includes 164 loop detectors, they are on a single facility; we would like to gather data from other locations to balance the trade off between precise error detection and required tolerance for normal variance. When the tests do fail, we would like to work with Caltrans to follow up on the problems and diagnose what went wrong in the field.

Finally, we hope to examine various means of removing the impact of errors in the collected data, either by substituting available data for missing data or some other means. Such work would have to include the accuracy of such substitutions and their impact on the traffic surveillance system, i.e., determining when is some information better than none.

7. Implementing Detector Diagnostics

The work done on developing detector diagnostics produced a large number of potentially valuable and interesting tests which could be run against the high resolution data collected in the Berkeley Highway Lab. Of the set of possible tests, five were chosen to be implemented in software and run continuously in real time on the incoming data stream.

Another software module, the Diagnostic Processor, was developed and added to the BHL system architecture (Figure 2.1). The Diagnostic Processor pulls the detector packets from the central database, passes the loop transitions to a set of diagnostic routines, and stores the results back into a diagnostic status table in the central database. Each diagnostic routine is implemented separately and operates independently on the incoming data. This allows new diagnostic tests to be easily integrated into the system as they are developed. The Diagnostic Processor provides a flexible framework within which many tests of different kinds can be incorporated.

Five diagnostic tests were implemented in the BHL system; activity test, minimum on time test, maximum on time test, mode on time test, and dual loop mode on time difference test. Four tests run on a specified number of vehicles, one runs on a time interval basis. Three tests operate over all traffic conditions while two operate in free-flow conditions only. Four operate on single loops, one operates on paired loops. These tests were chosen as providing the most immediate feedback on the correct operation of the loops.

7.1 The activity test

This test provides basic information on loop operation. The test looks for activity on a single loop in a 15 minute interval. If any activity, meaning at least one loop transition is seen, the loop passes the activity test. The BHL has several broken loops, including one at station 5 (the Eastbound upstream loop in lane 1) and two at station 3 (the Eastbound upstream loop in lane 5 and the Westbound downstream loop in lane 2). All other loops pass the activity test except during periods when the communications system is not working correctly.

7.2 Minimum on time test

The second test is a test for minimum on time. This test looks at the length of the on times for the last 100 vehicles to cross the loop. If more than 3.5% of the vehicles generate on times shorter than $7/60$ ths of a second, the loop fails the minimum on time test. All the active loops in the BHL pass the minimum on time test.

7.3 Maximum on time test

The maximum on time test looks at the length of the on times for the last 100 vehicles to cross the loop. If more than 3.5% of the vehicles generate on times greater than 700 ticks (700 $1/60$ s time intervals, or approximately 12 seconds) the loop fails the maximum on time test. All the active loops in the BHL pass the minimum on time test.

7.4 Mode on time test

The mode on time sorts the lengths of the on times for the last 1,000 vehicles to pass the loop. The times are sorted into bins of $1/60$ th of a second width and the mode of the distribution

is found. If the mode of the distribution is greater than 16/60ths of a second or less than 10/60ths of a second, the loop fails the mode on time test.

This test is valid only for vehicles traveling at free flow speeds. During congestion, the mode will quite often be much higher than the cutoff value used in the test. To adjust for this, vehicles are only counted in the mode on time test if they are likely traveling at free flow speeds. For every vehicle, its velocity is estimated based on the median on times seen for this vehicle and the previous ten vehicles. An average vehicle length of 20 feet is divided by the median on time of these 11 vehicles. The result is compared against a free flow speed of 50 mph. If the result is greater than or equal to 50 mph, the vehicle stream is likely moving at free flow speeds and the current vehicle is counted. Once 1,000 free flow vehicles have been seen, the mode is calculated and tested against the thresholds.

The mode on time is very sensitive. Many of the loops in the BHL will pass and then fail this test. At this time, more work needs to be done to determine whether the changes in test results are caused by changes in the loops or caused by thresholds set too tight for normal traffic variation.

7.5 Dual loop on time difference test

The dual loop on time difference test is the only test implemented which uses paired loops. This test compares the on times seen at the upstream loop with the on times seen at the downstream loop for the last 1,000 vehicles crossing the pair of loops. For each vehicle, the difference between the on times for the pair of loops is calculated. In general, this value should be zero. If more than 5% of the vehicles have a difference between the two on times greater than or equal to 3.5/60ths of a second, the pair of loops fails the on time difference test.

Like the mode on time test, the dual loop test is only valid during free flow conditions. A vehicle is only counted in the test if the median velocity of the it and the last 10 vehicles is greater than or equal to 50 mph. Since this test uses pairs of loops, the velocity calculation for the vehicle is done differently than for the mode on time test. A velocity is calculated by dividing the distance between the loops by the difference between the on time at the downstream loop minus the on time at the upstream loop. A similar calculation is done based on the off times at the loops. The velocity of a vehicle is calculated as the average of the two velocities. This produces a much more accurate velocity measurement than the calculation used for a single loop.

7.6 Archiving Test Results

Once the diagnostic tests have been run, the results are stored. Test results are stored in two database tables in the central database server. The first table, diagnostic status, holds the most current test results. This is used for real time display of the loop status. The second table, diagnostic history, holds past test results. Each diagnostic test stores the results from its previous run. When the results of the current test are different from the last run, the current result is stored in the diagnostic history table.

Archiving test results allows the cross-referencing of collected data sets with the diagnostics. Periods where a particular loop has failed a particular test can be identified and the data collected from that loop can be marked as bad. Users of the data can check the loop status before working with the data.

7.7 Monitoring the Loops

The status of all loop tests can be monitored via a set of web pages. The first page displays a summary status light for each of the 16 stations (each station is one side of the freeway only) in the BHL. Figure 7.1 shows the summary status page. Lights for each station indicate how many of the tests have been passed at that station. A green light for a station means that all tests for all loops in all lanes have passed. A yellow light indicates 70% of the tests have been passed. A red light means that more than 30% of the tests have failed for that particular station.

Each light is linked to a detail page with test results for each of the individual tests performed on the loops at that station. Figure 7.2 shows the detail page. Each column in the detail page represents a lane. Each row represents a particular test. Each test result is shown in two cells of the table. The left cell shows the result of the test, green for passed, red for failed. The right cell shows the number of minutes since the test was last run. Since most of the tests are based on the number of vehicles passing the loop, the time required to run the test will vary for each loop and by time of day. The minutes since the last run are shown to give an indication of whether the result has just been calculated or is from a while ago.

7.8 Future Diagnostics

The five diagnostic tests are just a start on the large set of tests that could be implemented to run on the BHL loops. Other tests were developed as part of this project and will be implemented in the future. The diagnostic history archive will provide a view of the operation of



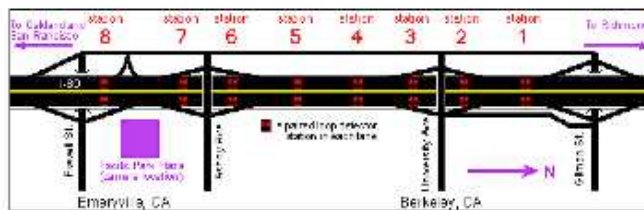
I-80 Freeway Status Page

- Education
- Research
- Publications
- The Library
- People
- Centers
- Tech Transfer
- Computing
- Labs
- Contact Us
- Home

Station and Loop Diagnostics

Station:	8	7	6	5	4	3	2	1
EastBound								
WestBound								

Click the lights below the station number to see detailed data for that station.
 Green means that station passed all tests for all lanes for all loops
 Yellow means that station passed 2/3 of all tests for all lanes for all loops
 Red means that station failed at least 1/3 of all tests for all lanes for all loops



Vehicle re-identification algorithms developed by Benjamin Coffman and implemented in real time by Randall Cuyford. For more information on this work, [click here](#).

Please send any questions, comments, suggestions on this or any of the I-80 freeway data pages to Randall Cuyford (cuyford@its.berkeley.edu).

© Copyright 2002 Regents of University of California. Last updated on 7/17/2002.

Figure 7.1 Summary Status Page

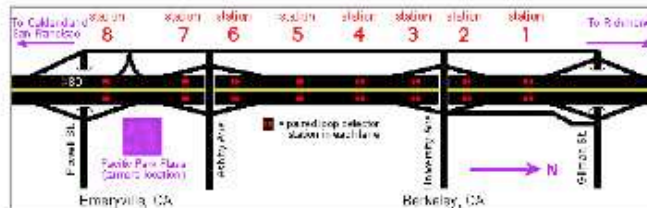


Loop Diagnostics for 3, WestBound

- Education
- Research
- Publications
- The Library
- People
- Centers
- Tech Transfer
- Computing
- Labs
- Contact Us
- Home

Other Stations: 8 [East](#) [West](#) 7 [East](#) [West](#) 6 [East](#) [West](#) 5 [East](#) [West](#) 4 [East](#) [West](#) 3 [East](#) [West](#) 2 [East](#) [West](#) 1 [East](#) [West](#) [Home](#)

Loop	Test	Lane									
		1		2		3		4		5	
		Status	Last Run	Status	Last Run	Status	Last Run	Status	Last Run	Status	Last Run
Upstream	Activity		12 min		12 min		12 min		12 min		12 min
Upstream	Min On Time		0 min		3 min		1 min		2 min		4 min
Upstream	Max On Time		0 min		3 min		1 min		2 min		4 min
Upstream	Mode On Time		4 min		29 min		6 min		18 min		33 min
Downstream	Activity		12 min		12 min		12 min		12 min		12 min
Downstream	Min On Time		3 min		NA min		1 min		3 min		3 min
Downstream	Max On Time		3 min		NA min		1 min		3 min		3 min
Downstream	Mode On Time		43 min		NA min		30 min		3 min		20 min
Dual	On Time Differential		30 min		NA min		39 min		45 min		31 min



Vehicle reidentification algorithms developed by [Benjamin Coifman](#) and implemented in real time by [Randall Cayford](#). For more information on this work, [click here](#).

Please send any questions, comments, suggestions on this or any of the I-80 freeway data pages to [Randall Cayford \(cayford@its.berkeley.edu\)](mailto:cayford@its.berkeley.edu).

Figure 7.2 Detail Status Page

the loops over time. Currently, there are no convenient viewing tools for examining the test history of a particular loop. The development of viewing tools will allow the examination of each loop over long periods. Analysis of the test history will also allow an investigation into the causes of loop failures, such as correlation with weather conditions.

8. Update on the Vehicle Reidentification Algorithms

The fundamental benefit of the BHL is the provision of higher resolution data for the development of new applications. Among these applications, vehicle reidentification has been a driving force behind the creation of the BHL.² Previous PATH projects deployed real-time vehicle reidentification tools that are compatible with Caltrans' existing dual loop detector hardware and communications infrastructure, e.g., [1-2]. Once a vehicle has been reidentified, its travel time can be measured by taking the difference between the two arrival times. This real-time travel time information can be used to assess the benefits of emerging technologies that promise to match a larger percentage of the passing vehicle fleet.

Although this study did not explicitly support the development of vehicle reidentification algorithms, the BHL was originally created to test such algorithms, verify their feasibility and demonstrate their benefit. This history is evident in the real-time travel time measurement presented on the BHL web page. A related study, PATH TO 4107 has continued to refine the vehicle reidentification algorithms. Access to the high resolution data is still uncommon since it is normally aggregated in the controller and discarded. Any large scale deployment of the vehicle reidentification algorithms would likely require further refinement of the algorithms to make them robust to phenomena that are not observed in the BHL.

Before this study, the vehicle reidentification algorithms matched between seven and 70 percent of the passing vehicles, depending on traffic conditions. These matches had an accuracy

² Vehicle reidentification is the process of matching individual vehicle measurements at a freeway detector station with the vehicles' corresponding measurements taken at another detector station located upstream.

over 95 percent, but the occasional errors may be severe. During the past year our attention focused on reducing the severity of the errors when they occur. We also examined the feasibility of adapting the algorithms to single loop detectors. Both of these efforts will be summarized here and discussed at greater length in the final report of TO 4107.

8.1 Reducing the severity of errors

Although the performance of the vehicle reidentification algorithms is good, it remains a difficult task to differentiate between two length measurements. To summarize the steps in the Basic Reidentification algorithm, first, vehicles are assigned successive arrival numbers as they pass each detector station and these numbers are assigned independently at each station. Next, a set of *feasible upstream measurements* is identified for each downstream measurement, bounded by minimum travel time and maximum density. The algorithm finds all vehicles in this set with a length range that intersects that of the downstream vehicle. The results of this resolution test are stored as one row in a matrix termed the vehicle match matrix (VMM). Note that in the following discussion, the columns of the VMM are indexed by the difference between the arrival numbers assigned to vehicles at the upstream and downstream detector stations (upstream offset). Therefore, each diagonal³ in the VMM corresponds to all possible matches for a given upstream vehicle.

The false positives are randomly distributed over the VMM, but if vehicles usually maintain their order between stations (i.e., assuming lane change maneuvers are relatively

³ Throughout this paper, a diagonal in the VMM will refer to a diagonal going from the upper right to the lower left corner of this matrix.

infrequent), the true (but unknown) matches should manifest themselves as sequences of possible matches in the same column⁴. In other words, false positives will typically form short sequences in the VMM while true matches will usually form longer sequences. Figure 8-1A shows a sample of the VMM for 100 vehicles and Figure 8-1B shows the matches found by the algorithm. In general the matches such a small number of downstream vehicles should usually all fall near a single column. In this case the true matches fall near column 80 while errors are seen as short sequences elsewhere in the matrix. The preceding algorithm would catch most of these errors but occasionally some would slip through. When one of the accepted errors is several columns off, the resulting error in travel time measurement could be severe.

During the previous year we developed several tests to exploit the fact that the true matches should not drift rapidly in the VMM. We use one of these tests, the so-called Filter Test, to illustrate the benefits. As previously discussed, true matches tend to form longer sequences than false positives do. This test will favor areas of high density of matches in the VMM. The Filter Test uses aggregate trends to narrow the range of *feasible upstream matches* to a small number of columns in the VMM that are most likely to contain the true matches. So the Filter Test does not look for final matches, but for a "final region."

The Filter Test starts with the VMM and selects all matches that are members of the three longest vertical sequences for each row and for each diagonal, i.e., the most promising matches for a downstream vehicle and an upstream vehicle, respectively. Figure 8-2A shows an example

⁴ For simplicity, the term sequence is used to refer to a sequence of possible matches in the same VMM column.

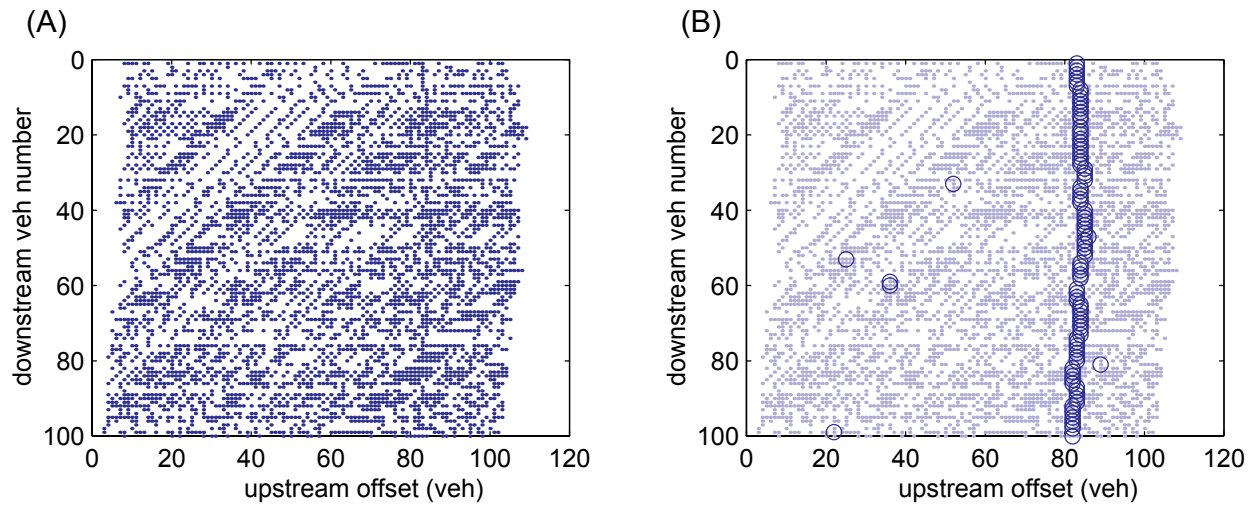


Figure 8-1 Vehicle Match Matrix

(A) A sample VMM for 100 downstream vehicles, (B) matches (circles) after accounting for lane change maneuvers and the initial VMM (light points) from part A.

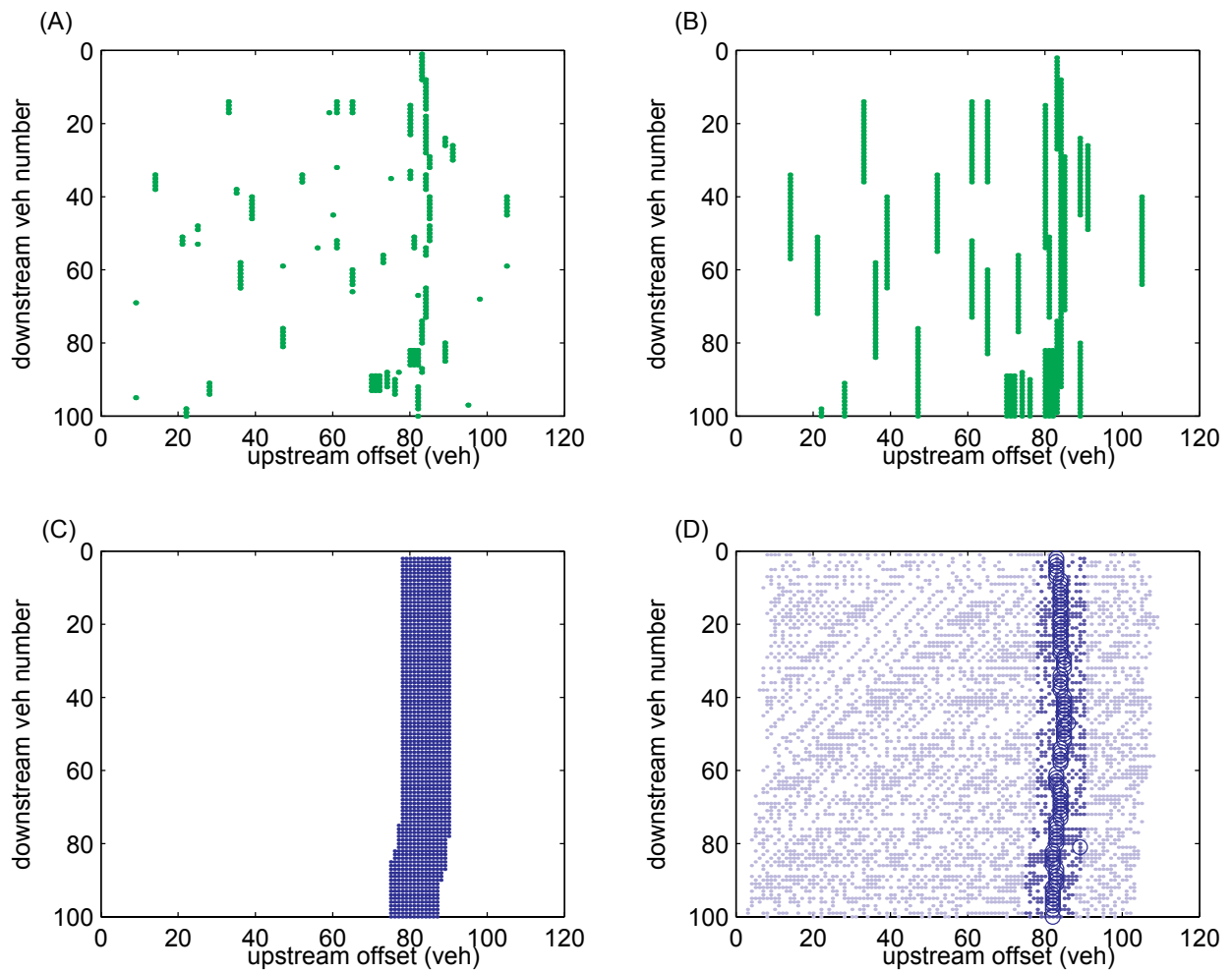


Figure 8-2 Vehicle matches with filter test

(A) Selection of best 3 matches per row and diagonal before lane changes, (B) extended sequences after vertical moving averages, (C) good region, (D) final matches (circles), possible matches within good region (dark points), and initial VMM (light points).

of this *pre-selection* applied to the VMM of Figure 8-1A. The Filter Test will ignore matches if they fall in sequences shorter than five vehicles.

The remaining matches are assigned a weighting of "1", and all other elements in the matrix are assigned a weighting of "0" (blank spaces in the figure). Most vehicle lengths fall within a small range; the remaining vehicles (about ten percent for the subject location) are distinct and have fewer false positives associated to them. Exploiting this fact, if a downstream vehicle has a possible match for less than 10 percent of its feasible matches, the vehicle's weighting will be doubled. The Filter Test also favors matches that fall in longer sequences (i.e., situations with no detected lane change maneuvers). Specifically, if a match belongs to a sequence that is 1.25 times longer than the median sequence length passing through the row of the VMM, the match's weighting is also doubled. Thus, a distinctive vehicle in a long sequence may ultimately receive a weighting of "4".

Each element of the filtered VMM now has a weighting of 0, 1, 2 or 4. Next, the Filter Test takes a moving average of 20 rows within each column, i.e., for each element in the filtered VMM, it calculates the mean weighting over the previous 20 elements in the same column. Once more, assuming that true matches usually form longer sequences than false positives, this process should favor the former. Comparing Figures 8-3A-B, this averaging has extended the sequences. Although it is not shown in Figure 8-2B, each of the matches has a different weighting associated with it. Also note that the small sequences (shorter than five vehicles long) that are shown in Figure 8-2A have been discarded in a previous step, and therefore do not appear in Figure 8-2B.

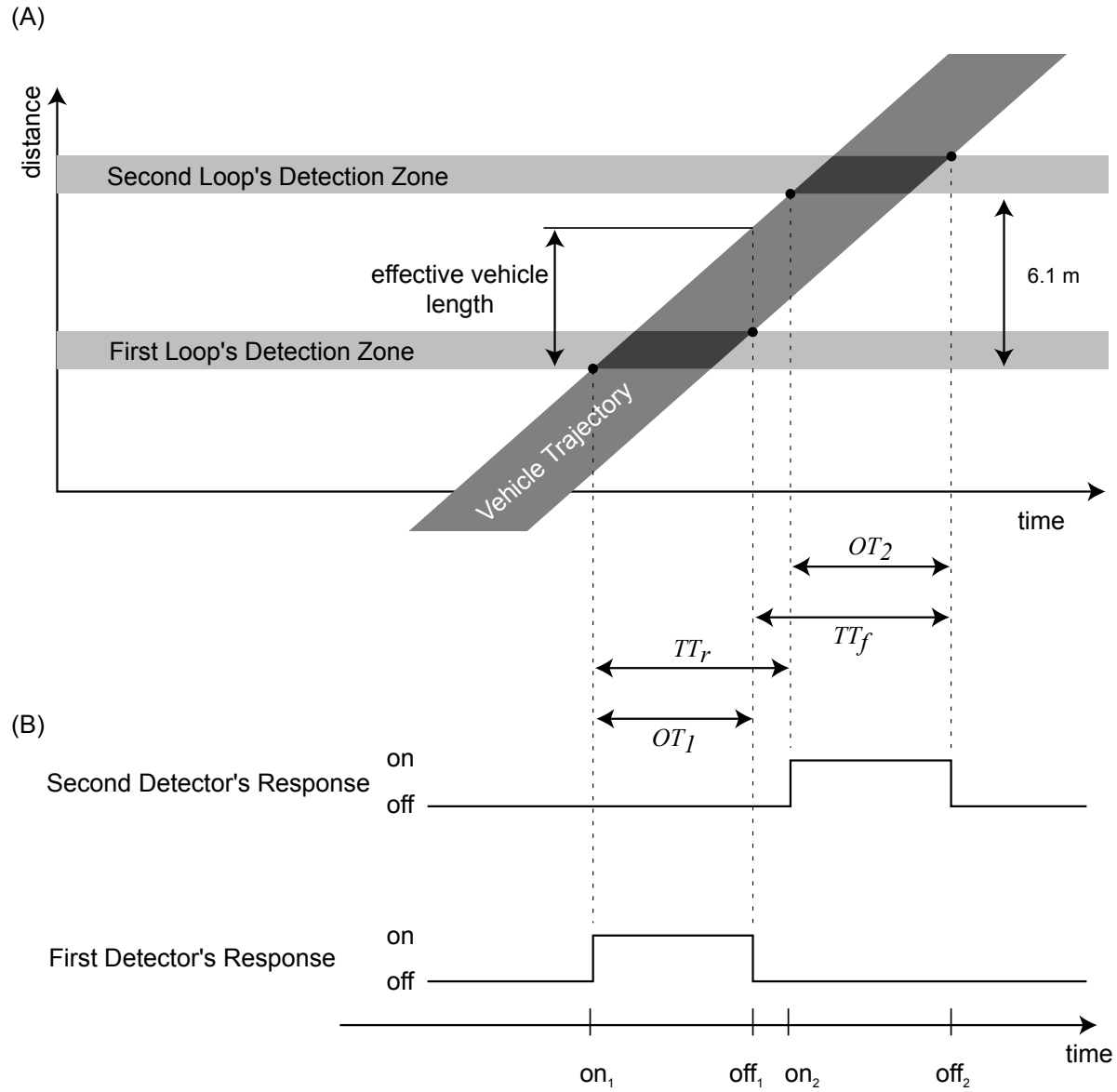


Figure 8-3 One vehicle passing over a dual loop detector

(A) the two detection zones and the vehicle trajectory as shown in the time space plane. The height of the vehicle's trajectory reflects the non-zero vehicle length.
 (B) The associated turn-on and turn-off transitions at each detector.

Under most traffic conditions true matches usually will not change columns rapidly because each column shift represents the relatively infrequent lane change maneuvers. This property is evident in the higher density of matches shown around column 82 in Figure 8-2A. Taking the results from the vertical moving average (Figure 8-2B), the Filter Test then calculates a moving average of five columns within each row in order to "connect" sequences from adjacent columns of the filtered VMM and account for lane change maneuvers. The results of this horizontal moving average represent the final weightings for each element in the filtered VMM.

Next the Filter Test selects all matches that have a weighting greater than a threshold. In the present study, to determine this threshold we calculated the mean over all non-zero final weights from the filtered VMM. The threshold was then set to 5 times this mean weighting. The averaging and thresholding process is repeated a second time to further narrow the set of possible matches. Everything that remains defines the *good region*, e.g., Figure 8-2C.

Even though determining the *good region* is the ultimate goal of the Filter Test, in order to appreciate its performance, we apply the Basic Algorithm to the same VMM independently of this test. Finally, only the intersection of the best matches from the Basic Algorithm (Figure 8-1B) and the *good region* from the Filter Test (Figure 8-2C) are retained as final matches for this test, e.g., Figure 8-2D.

8.2 Extending to single loop detectors

During free flow periods, the measurement resolution from dual loop detectors degrades simply because the difference of 1/60 sec becomes more significant in TT and OT (see Figure 8-3). During these periods, it becomes impossible to differentiate between the short vehicles; however, the large range in the long vehicles allows for continued matching (see Figure 6-1). To address this problem, we have developed a reidentification algorithm that identifies relatively distinct vehicles, i.e., long vehicles, at the downstream detector station. For each of these vehicles it looks for a similar vehicle in the same lane at the upstream station within a time window of *reasonable* free flow travel times [2]. Thus, if traffic is free flowing over the link between detectors, this approach will usually find a match in the time window (Figure 8-4A). If the freeway is congested, vehicles will be delayed and the true match for a vehicle will not be found in the time window (Figure 8-4B). In the event a match is found the algorithm can also calculate that vehicle's travel time by taking the difference in the vehicle's arrival time at the two locations. In other words, the algorithm is capable of reporting free flow travel times or that "traffic is not free flowing". Over the past year we have extended the work to use estimated vehicle lengths from single loop detectors and then apply the vehicle reidentification methodology to these common detectors. The length estimation is simply the product of Equation 6-3 and the vehicle's measured on-time. Figure 8-5 compares results from single loop detectors against dual loop detectors over an 1,800 ft link. The algorithm performs well with single loop data, the results are almost as good as the dual loop

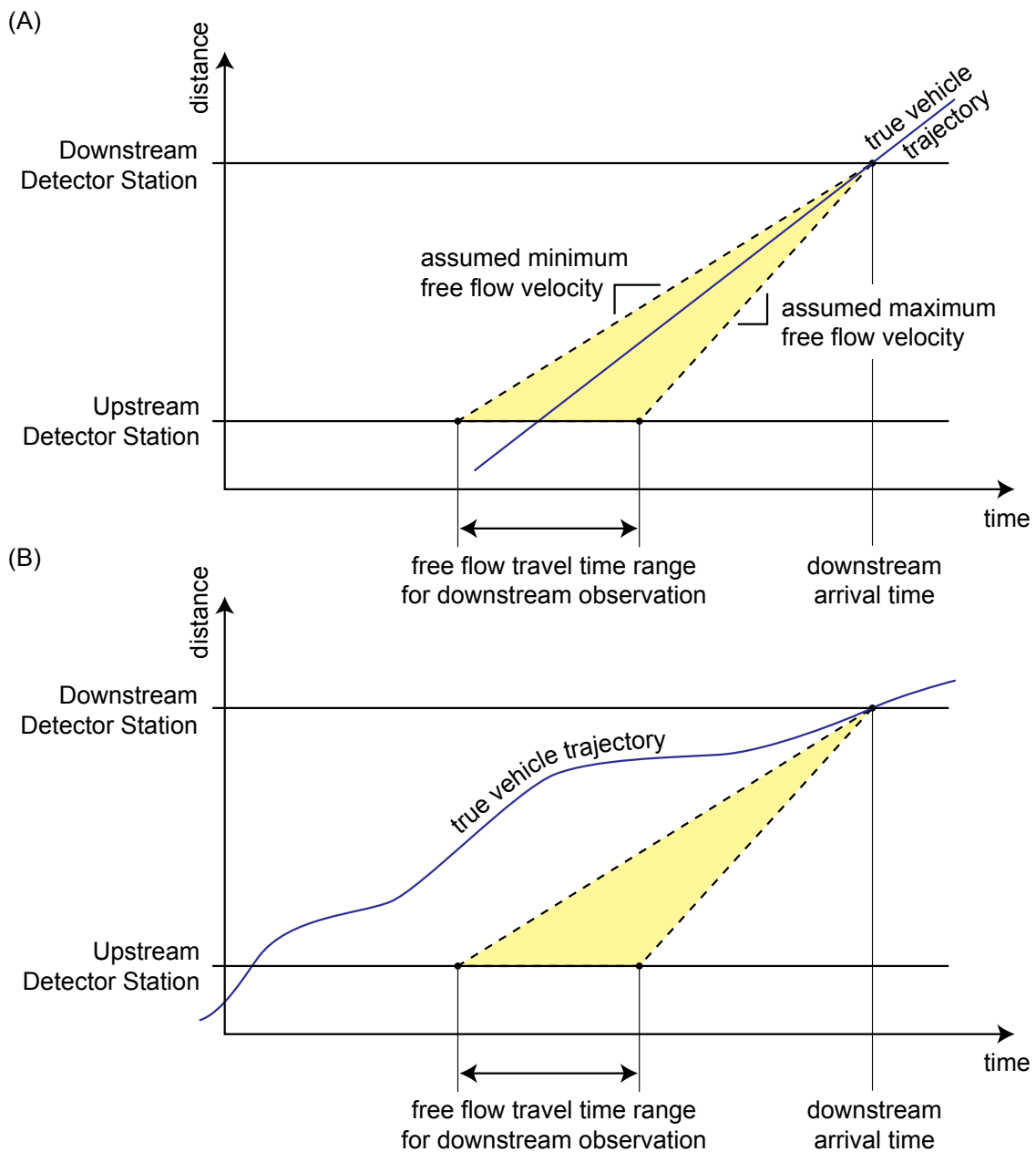


Figure 8-4 One vehicle traversing an extended link between two detector stations, illustrating the free flow travel time range.

(A) The vehicle travels at a free flow velocity and it was observed at the upstream station during the time range; (B) the vehicle traveled slower than the minimum free flow velocity and it passed the upstream station before the start of the time range.

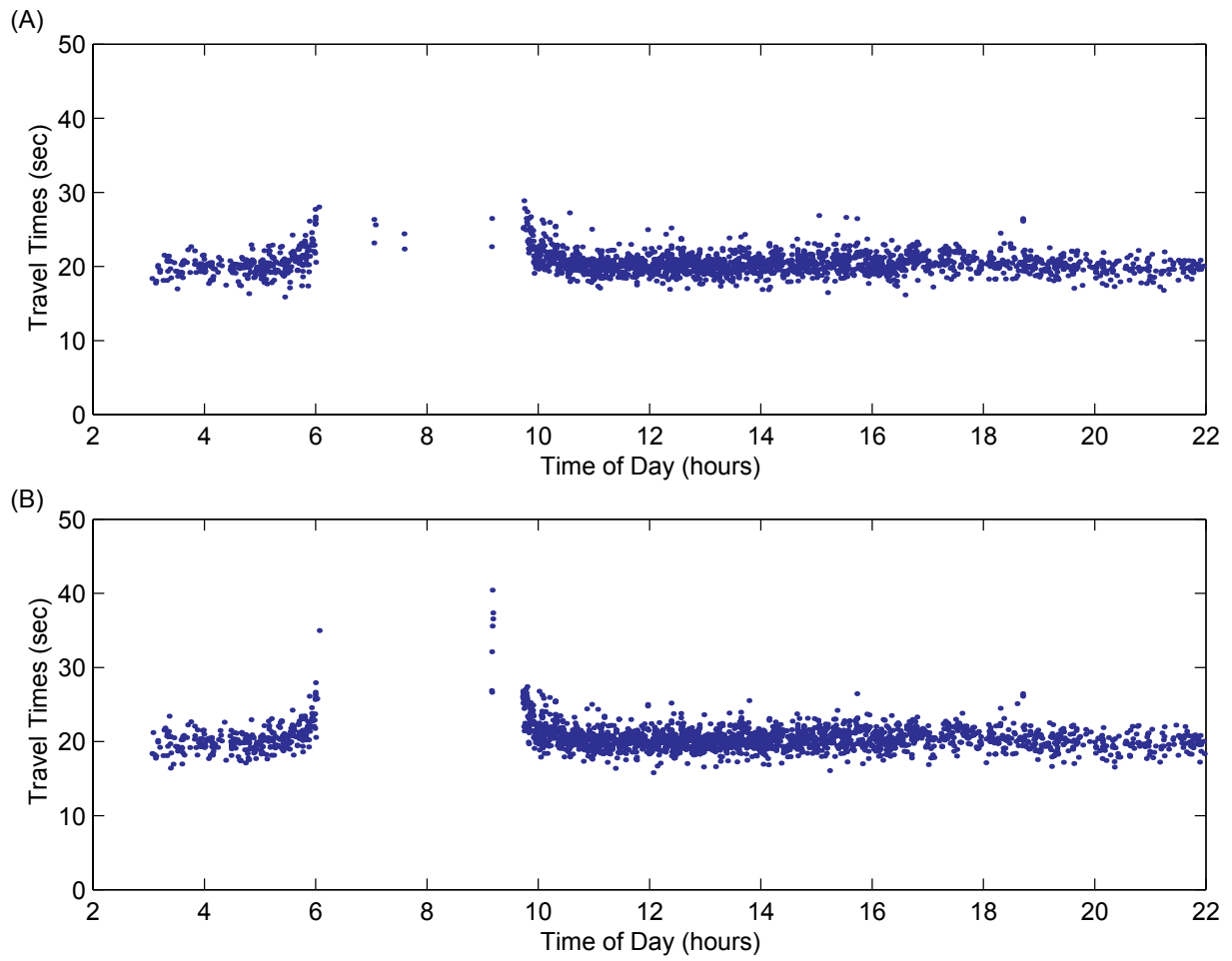


Figure 8-5 The resulting travel times for matched vehicles in one lane over 20 hours
(A) single loop data (B) dual loop data.

8.3 Future directions

Although the next round of PATH funding does not support development of vehicle reidentification, we still hope to further improve the performance of the algorithms. In addition to improving performance in the existing scenarios, we would like to extend the algorithms across lane adds/drops and freeway merges/diverges. These tasks will increase the number of freeway links that can be monitored in a larger scale deployment. We also hope to consider data from emerging detector systems that promise improved measurement resolution.

9. Experience with Alternative Communications

A key part of the BHL loop collection system is the communications link between the loop cabinets in the field and the UCB campus servers. Unlike most cabinet setups, data on every loop actuation seen by the controller is being sent to the central server. This entails sending one data packet every second from every cabinet.

At the start of this year's project, field communications were handled by a computer installed in the field cabinets. The field computer received data packets once each second from the controller. The data packets were time stamped and sent to the UCB server using a Ricochet wireless modem. A messaging protocol between the field computer and the UCB server handled acknowledging the receipt of data packets at the server and resending any packets lost in transmission. Figure 9.1 shows the primary communications pieces.

The field computer based communications system had several problems. The wireless modem established a standard internet connection for the computer. Wireless communications is much less reliable than most other connection technologies, however. The network connection failed often which caused serious problems with the operating system (Windows 98) installed on the field computer. Depending on what operation was being performed when the failure occurred, the computer either recovered the connection and continued, or attempted to reboot. Reboot was generally successful but occasionally would fail causing the computer to hang until the power was cycled. Lamp timers were added to the field system to power cycle the computers regularly, every morning at 2:00 AM. This prevented long term failures of the communication system by limiting the length of time during which a particular station was off-line to no more than a day. However, data losses of several hours were common.

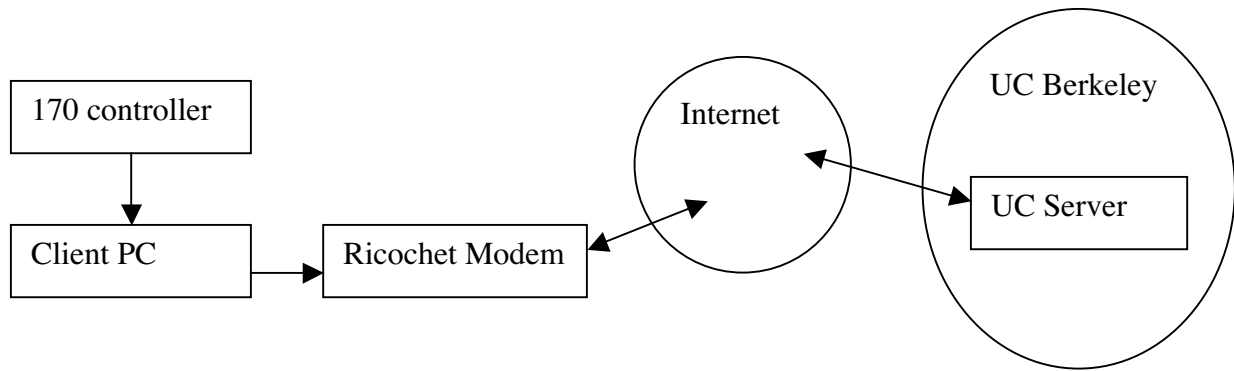


Figure 9.1 Controller to PC to Ricochet to Server

In August 2001, Metricom, the provider of the Ricochet wireless service, went into bankruptcy and all the BHL field modems stopped working. This presented an opportunity to try alternative communications systems.

A new communications system was developed which eliminated the field computer in favor of stand-alone CDPD wireless modems. The new system uses just a CDPD modem installed in the cabinet connected directly to the controller. Figure 9.2 shows the new communication system.

The CDPD modems accept serial data from the controller and forward it over a wireless internet to the UCB server. The modems use the UDP internet protocol. This protocol is very efficient but provides no acknowledgement for received packets and does not automatically resend lost packets. In normal operation, occasional packets will be lost in transmission. Two brands of CDPD modems were tried, one from Sierra Wireless and one from Airlink. They had different requirements on the server side and different problems with transmissions.

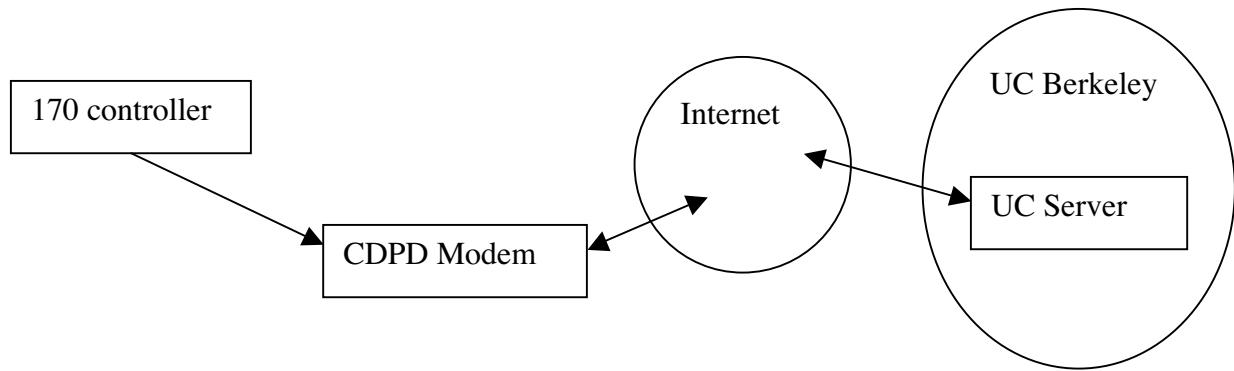


Figure 9.2 Controller to CDPD to Server

The first modems tried were ones from Sierra Wireless. These modems will automatically send data received over the serial port to a server once a communications channel has been opened. Opening a communications channel requires having the server contact the modem to initiate communication. While the Sierra Wireless modems generally worked well, were reliable and rarely lost network connection, they had an internal software problem which caused them to truncate the data for every 30th packet. This rendered the packet unusable at the server end. We were unable to resolve the problem with the vendor.

A second set of modems, from Airlink, were installed in August 2001. These modems automatically send serially received data to a server address preconfigured in the modem. Once turned on, they start sending data immediately. This simplified the server software considerably. The Airlink modems are smaller and considerably cheaper than the Sierra Wireless modems and have been extremely reliable.

The new communications system has proven to be much more reliable than the old communications system. Configuring a modem for installation takes only a few minutes, dramatically less time than configuring a field computer. There have been many fewer network related failures. While the CDPD modems still lose connection for short periods of time on rare occasions due to problems with the underlying wireless network, provided by Verizon, they recover automatically as soon as the network problems are resolved. Random connection loss caused by radio interference results in the loss of a few packets but rarely more than a few seconds at a time. The field computer based system lost several minutes of data whenever radio interference caused a lost network connection.

Overall, the data collection system captures the controller packets with a very high degree of reliability. The primary problem with the new communications system is the occasional

packet lost in transmission. On a typical day, the system collects 98.5%-99.5% of all controller packets. One or two seconds of data from a particular station will simply never arrive at the server. No analysis has been done matching packet loss times between stations to determine whether there is any pattern to the losses or any correlations either between the stations or to network activity on the UCB campus.

While the new communications system only loses approximately 1% of the packets, several possible ways to further reduce the number of lost packets have been explored. The simplest is to send each packet to two different servers, preferably on two widely separated network. Sending two packets, each of which has a 99% chance of arriving, should lower the overall lost packet rate significantly. A second proposed solution is to implement a acknowledgement/resend protocol in the modem. While more complicated, this would lower the loss rate the most. Changes to the server software would be necessary to check that all packets are received and to request a resend of packets that did not arrive. Much of the development for this enhanced server software has already been done, however, as the original communications system implemented such a scheme. Unfortunately, these techniques require modifying the software embedded in the CDPD modem. This will force the use of a vendor specific solution. While this may not be a problem for a small experimental installation, it does present problems for a long term and larger scale deployment where multiple suppliers would be desirable.

10. Moving Towards a Larger Scale Implementation

The loop collection system of the BHL has been in operation for over 4 years. The original design of the system was guided by the particularities of the location in which it was developed. The first seven stations originally part of the BHL all cover both sides of the freeway with five lanes in each direction. As part of this project, modifications to the system were made to support the easy implementation of the BHL data collection system in other locations.

The biggest barrier to implementing the BHL system at other locations was the variability in station configurations. All the original stations in the BHL had the same configuration and the software made many implicit assumptions about how stations were designed. The first step towards a system deployable at other locations was to redesign all the software modules to use a more general definition of stations.

Caltrans convention is to define a station as a set of loops on one side of the freeway only. The original BHL system defined a station as both sides of the freeway. The definition of station in the BHL system was changed to match the Caltrans definition. This was done by splitting the data packets by side of freeway. For seven of the physical cabinets in the BHL, data from both sides of the freeway are mixed in the same controller data packet. The server software running on the data collection server was modified to parse the controller packet into two separate packets for these stations. Each new data packet matches the format and structure of the original packet but has loop transitions only for loops on one side of the freeway. The new packets are stored in new tables in the central database identified by new station identifiers which are side specific.

Once stored, the new packets are processed in much the same way the old packets were. Each software module in the BHL system was modified to pull the new data packets from the database by side of freeway. Much of the software became simpler as there was no longer any need to split the data by side of freeway within the secondary processing modules. All modules were changed to use the new station identifiers.

The second change was to make the system flexible with respect to the number of lanes at each station. For every station —now defined as a single side of the freeway— the number of lanes became an input parameter for that station. The internal structures and processes were modified to use a variable number of lanes.

When it was first established, the BHL adopted its own identification scheme for labeling individual stations and cabinets. This made internal operations easier and more understandable but limits the ability to easily deploy the system elsewhere. It also causes confusion when communicating with agencies outside the BHL such as Caltrans. As part of the redefinition of stations to be side specific, the standard Caltrans naming convention for stations was adopted. Each station is identified by road, direction and milepost in a single unique identifier string. All software modules in the BHL were modified to use this identifier string to specify stations. The major advantage of this change is to easily allow the addition of other stations into the system. The old naming convention depended on numbering stations sequentially moving down the freeway. The new naming convention allows new stations can be added independently of the old stations without causing numbering problems. A new station may be located anywhere and will be identified and treated correctly within the system.

The software modules in the BHL system all depend on a certain amount of configuration information. Each module needs to know things like which stations to process,

what the loop to port mappings for those stations are, the number of lanes for those stations and other parameters. To simplify the system, all station specific information was moved from configuration files specific to each software module into a station data table in the central database. Each module still has a small configuration file, but many of the parameters are now read out of the station data table. This simplifies adding a new station to the system by centralizing the information in one place accessible to each software module.

Other changes were made to the BHL system to support its ability to be deployed in new locations. While the software changes make the system more flexible, changes in the hardware make it much simpler and cheaper to install in new areas. The field installed hardware was changed by switching from computers with wireless modems to stand-alone CDPD modems. The single modem solution is significantly cheaper than the combined computer –modem solution, both in terms of hardware cost and in terms of installation and configuration time. Configuring the CDPD modem can be done in minutes and requires no site specific information. More importantly, the field installation uses the same off the shelf components as those commonly installed in the cabinets to collect 30 second aggregate data. This makes the cost of field components identical to other data collection systems.

While the field components match are comparable to other solutions, much of the processing of the data has been moved from the controller to central servers. A larger scale implementation of the BHL would need to be able to economically scale to process increased amounts of data. The BHL system is designed to scale easily. Processes are independent and can be easily separated onto different machines. With the exception of the travel time calculations, each station can be processed independently. If there are too many stations to process on a single machine, a second machine can be added to process half the stations. The travel time calculation

deals with pairs of stations only so it, too, can be partitioned across multiple machines if necessary. The current BHL implementation runs on four low-end desktop computers. The partitioning across multiple computers was done for ease of updating and managing the separate software modules rather than because of load issues. Experience with the system indicates that none of the four machine is using more than approximately 10% of the available processing power to deal with the 16 individual stations in the current BHL. Switching to higher end equipment , still within the inexpensive desktop machine range, and optimizing the software a small amount would allow the current BHL implementation to process several thousand stations.

The new software for the BHL data collection system is now flexible enough to be easily deployed to new locations. The cost should be comparable to existing data collection systems and yields significant additional benefits such a accurate travel times between stations through vehicle re-identification and greatly enhanced detector diagnostics. A full cost-benefit study comparing the BHL system to other systems is planned for the coming year.

11. Conclusion

This final chapter in the report is divided into two major parts. The first part summarizes the major accomplishments of this year's project both in terms of immediate and longer term consequences. The second part describes future research directions with particular focus on next year's project.

11.1 Summary of Major Accomplishments

The major accomplishments of each of the nine tasks described in the previous nine chapters will be summarized for each chapter in the following paragraphs.

The BHL detector system has been continuously maintained during the current project (Chapter 2). Its major accomplishments include improved system reliability and more efficient recoveries from failures.

The BHL detector data has been continuously collected and archived during the current project (Chapter 3). Its major accomplishments include daily summary data and individual vehicle data.

The BHL detector data has been provided to researchers upon request during the current project (Chapter 4). This has resulted in the use of BHL data for additional traffic research beyond the scope of this project.

The BHL detector system link to District 4 has been maintained during the current project (Chapter 5). This has allowed the BHL loop detector stations to function as standard Caltrans

District 4 detectors while providing much richer data and research possibilities under the BHL project.

Research of alternative detector diagnostic tests has been undertaken as part of this current project (Chapter 6). Its major short-term and longer-term accomplishments include the development and refinement of several single and dual loop detector data validation tests.

An initial set of detector diagnostics has been implemented into the BHL detector system (Chapter 7). This has resulted in a set of tests which are run continuously against the incoming data stream from every station. There are four single loop detector tests and one dual loop test run for every loop of every lane of every station in the BHL.

Further research effort has been given to vehicle re-identification (Chapter 8). This has resulted in higher frequency of vehicle reidentification while simultaneously reducing the error rate.

Alternative communication technologies have been evaluated (Chapter 9). This has resulted in the development of a new communications system using off the shelf components which is cheaper, simpler and more robust.

Efforts have been expended on moving toward larger-scale implementation (Chapter 10). This effort has resulted in modifications to the BHL system which make it more adaptable to different cabinet, station and loop configurations. The system has been modified to more closely match Caltrans standards for station identification and layout. The cost of the system, in both time and hardware, has been reduced by the use of the same standard field components which are currently used for data collection.

11.2 Future Research Directions

The efforts on this current research project have resulted in the identification of future research directions with particular focus on next year's project, expected to begin at the beginning of 2003 and continue through 2003. A proposal for continued research on this project has been submitted and approval for funding for next year is expected to begin in the fall of 2002 and continue until mid-2003. Next year's BHL project includes eight major tasks which are briefly described in the following paragraphs. Research for next year and implications for research beyond next year are included in the discussion of each of next year's eight tasks.

The first task will be to continue to work toward an automated detector diagnostics and real-time data validation. The goal of next year's project will be to extend and enhance the current detector diagnostics, and to begin to diagnose causes and possible solutions of identified detector problems. Beyond next year's project, means of substituting available data for missing data due to detector difficulties should receive attention.

The second task will be to continue the development and implementation of system diagnostics. The goal of next year's project will be to develop additional software to monitor system components, pro-actively notify project personnel in the case of component failure, and configure system components to automatically recover from outages where feasible.

The third task will be to undertake a cost-effectiveness evaluation of the BHL-type detector system. The objective of this task is to provide guidance to Caltrans as to the cost-effectiveness of a BHL-type detector system as compared with the current Caltrans detector system approach when applied at another freeway site. An appropriate freeway site would be selected considering Caltrans suggestions.

The fourth task will be to continue to maintain the data link to Caltrans District 4 in order to provide the required conventional aggregate data and at the same time, permit the continued use of the BHL detector system for research without interfering with the day-to-day operations in the District. The detector and system diagnostics described in the two previous tasks and the associated pro-active responses to identified problems will insure a continuous flow of high-quality set of detector data to Caltrans.

The fifth task will be to continue to maintain and operate the BHL detector system data and to archive data. The BHL surveillance system will be continually updated as detector and system diagnostics are added and as additional data is archived.

The sixth task will be to enhance the benefit of the system by generating and archiving additional detector and system data. Consideration will be given to generating and archiving at least some of the following additional detector and system data:

- 5 minute aggregate data by lane
- enhanced individual vehicle data at each station to include unmatched transitions
- aggregate travel times between each station
- loop diagnostics results by appropriate time interval
- system errors by occurrence
- loop errors by occurrence

The seventh task will be to continue to distribute the loop detector data to researchers upon request. Requests will be carefully evaluated in terms of effort required and potential benefits. Initially this will simply consist of providing CDROMs to researchers in an organized manner. A desired goal beyond next year's project will be to develop an on-line web-based distribution method for some portion of the data base.

The eighth and final task will be to prepare a final project report. A draft final report will first be developed and distributed to Caltrans and PATH for their review and comment. Considering comments received, the final project report would be prepared and distributed as a PATH report. It is also anticipated that a research paper based upon the final report will be prepared and submitted for national and/or international presentation and publication.

(References)

- [1] Coifman, B. and Cassidy, M. (in press). "Vehicle Reidentification and Travel Time Measurement on Congested Freeways", *Transportation Research-A*. Draft available at: http://www-ceg.eng.ohio-state.edu/~coifman/documents/VRI_basic.pdf.
- [2] Coifman, B. (in press). "Identifying the Onset of Congestion Rapidly with Existing Traffic Detectors", *Transportation Research-A*. Draft available at: <http://www-ceg.eng.ohio-state.edu/~coifman/documents/Truck.pdf>.
- [3] Jacobson, L, Nihan, N., and Bender, J. (1990). Detecting Erroneous Loop Detector Data in a Freeway Traffic Management System, *Transportation Research Record 1287*, TRB, Washington, DC, pp 151-166.
- [4] Cleghorn, D., Hall, F., and Garbuio, D. (1991). Improved Data Screening Techniques for Freeway Traffic Management Systems, *Transportation Research Record 1320*, TRB, Washington, DC, pp 17-31.
- [5] Nihan, N. (1997). Aid to Determining Freeway Metering Rates and Detecting Loop Errors, *Journal of Transportation Engineering*, Vol 123, No 6, ASCE, pp 454-458.
- [6] Chen, L., and May, A. (1987). Traffic Detector Errors and Diagnostics *Transportation Research Record 1132*, TRB, Washington, DC, pp 82-93.
- [7] Coifman, B. (1999) Using Dual Loop Speed Traps to Identify Detector Errors, *Transportation Research Record no. 1683*, Transportation Research Board, pp 47-58.
- [8] Coifman, B., Dhoorjaty, S., "Event Data Based Traffic Detector Validation Tests", *ASCE Journal of Transportation Engineering*, [submitted for publication].
- [9] Coifman, B., Dhoorjaty, S., Lee, Z. (in press). Estimating Median Velocity Instead of Mean Velocity at Single Loop Detectors, *Transportation Research: Part C*. (draft available at: <http://www-ceg.eng.ohio-state.edu/~coifman/documents/MedianSingle.pdf>)