

# **DRAFT Whitepaper**

## **Travel Demand Modeling for Oregon Highway 217 Using Archived Inductive Loop Detector Data**

Robert L. Bertini  
Civil & Environmental Engineering  
Portland State University  
P.O. Box 751  
Portland, OR 97207-0751  
Phone: 503-725-4249  
Fax: 503-725-5950  
Email: bertini@pdx.edu

Andrew Byrd  
Computer Science / Intelligent Transportation Systems Lab  
Portland State University  
Phone: 503-725-4285  
Fax: 503-725-5950  
Email: abyrd@cs.pdx.edu

Thareth Yin  
Civil Engineering / Intelligent Transportation Systems Lab  
Portland State University  
Phone: 503-725-4285  
Fax: 503-725-5950  
Email: thareth@thareth.com

Kyle Hauger  
Metro  
600 NE Grand Ave.  
Portland, Oregon 97232-2736  
Phone: 503-797-1813  
Fax: 503-797-1949  
Email: haugerk@metro.dst.or.us

February 2004

## ABSTRACT

[\[Add abstract\]](#)

## INTRODUCTION

Researchers at Portland State University's Intelligent Transportation Systems (ITS) Lab routinely analyze large amounts of traffic data collected from nearly 400 inductive loop detectors embedded in Portland area highways. These detectors were originally installed by the Oregon Department of Transportation (ODOT) to facilitate day to day management of the region's transportation facilities by providing vehicle count, speed, and detector occupancy information at twenty-second intervals. As part of the region's Automated Transportation Management System (ATMS), they serve primarily to inform transportation management facilities of congestion and incidents on the freeways as they happen. Through cooperation with ODOT, large amounts of this sensor data have been archived for alternate uses in studies of traffic phenomena.

In the fall of 2003, staff members at Metro, Portland's regional planning organization, were faced with a transportation planning project in need of supporting data. While ATMS data are frequently used in transportation research, its use is currently much less common in urban planning. The Metro staff were aware of PSU's access to and experience analyzing ATMS data and recognized the potential of this rich data source to assist their work. A partnership was formed PSU and Metro to determine the feasibility of using this archived ATMS data in the urban planning process.

## PROBLEM

For various reasons it had become necessary to consider the addition of new lanes to Oregon Highway 217. As a metropolitan planning organization (MPO), Metro is responsible for demonstrating evidence of future demand for increased road capacity in order to secure federal funds for such a transportation project. The demand for these improvements is estimated through a travel demand forecasting process, in which planners model the performance of the transportation network and predict what changes will be necessary to accommodate the needs of residents as a result of expected future growth.

Travel demand forecasting is a complex process and is usually carried out with the assistance of specialized computer software. In Portland, Metro uses the [EMME/2](#) system, which was originally developed at the Center for Research on Transportation at the University of Montreal. EMME/2 represents the entire regional transportation network as nodes connected by links which have associated capacities and performance characteristics.

The demand model for this network is typically created in four major steps: trip generation, trip assignment, mode choice, and trip distribution. In the first step, planners determine how many trips on the transportation network originate from each location within the region. Next, each trip is assigned a destination and a mode of transportation such as car, bus, rail, or bicycle. Once the number, destination, and mode of the trips within a region have been predicted, the actual simulation of roadway conditions occurs in the step called trip distribution. Under the assumption that travelers will favor the quickest path from source to destination, the trips are iteratively redistributed over the available transportation links until equilibrium is established. This distribution process requires that a relationship be specified between the volume of vehicles on a particular link and the expected amount of time required to traverse that link.

While the first three steps draw on statistical data from extensive travel surveys (the 1994/1995 Regional Household Activity Survey in the case of Portland), many of the volume-travel time relationships used in the trip distribution stage are theoretical or default values. Metro has collected data on the characteristics of Portland area arterial streets using probe vehicles, but this process had not been carried out on area highways due to lack of funds. Because it is the only step not based on empirical data, trip distribution can be viewed as a weak point in Portland's travel demand forecasting process.

The volume-travel time relationships used for highways at the time of this study were standard United States Bureau of Public Roads (BPR) functions taken from literature published as long ago as the 1960s. [\[Add a reference\]](#) The BPR, which later became the Federal Highway Administration, used these functions historically to make transportation decisions for the entire country. Regional variations in highway performance may exist which are not accounted for, and it is reasonable to assume that the characteristics of automobiles and the behavior of drivers have changed significantly since these functions were created.

Perhaps more importantly, recent developments in transportation such as freeway ramp metering may have deeply altered the performance characteristics of our highways. These original BPR functions model a steady degradation of highway speeds as the volume of traffic approaches capacity. Recent observations indicate that our current more efficient use of transportation facilities may cause speeds to remain much more consistent as volume increases, and that highway capacities may have actually increased. The hope in this study was that ATMS data could serve the same role for highways that probe vehicle tests had for arterial roads and provide some model of road performance without relying on these theoretical relationships.

## **DATA SOURCES**

At the time of this study, the ITS Lab did not yet have access to a continuous stream of ATMS data from ODOT. It was therefore necessary to choose a dataset from among those previously requested from ODOT for other research projects. A set of loop detector data from April 14, 2002 to April 20, 2002 was chosen because it was the most recent large dataset available, providing a full week of observations from all Portland area loop detectors. These data were loaded into a database table as shown in Table 1.

**TABLE 1 Database table containing a full week of loop detector data.**

detectorid	samplestart	volume	speed	occupancy
1181	4/20/2002 11:57:00 PM	4	55	5
1181	4/20/2002 11:57:20 PM	3	63	3
1181	4/20/2002 11:57:40 PM	4	54	4
1181	4/20/2002 11:58:00 PM	4	74	3
1181	4/20/2002 11:58:20 PM	2	79	1
1181	4/20/2002 11:58:40 PM	7	64	6
1181	4/20/2002 11:59:00 PM	2	63	2
1181	4/20/2002 11:59:20 PM	3	63	4
1181	4/20/2002 11:59:40 PM	2	64	1
1182	4/14/2002	3	62	3
1182	4/14/2002 12:00:20 AM	2	81	1
1182	4/14/2002 12:00:40 AM	1	83	0
1182	4/14/2002 12:01:00 AM	1	48	0
1182	4/14/2002 12:01:20 AM	1	65	0
1182	4/14/2002 12:01:40 AM	2	65	1
1182	4/14/2002 12:02:00 AM	1	65	0

Record: 2691361 of 11672640

Each row in the database represents an observation from one detector during one 20-second time period, including the number of vehicles passing the detector (count), the average speed of those vehicles, and the percentage of this 20 seconds when a vehicle was over the detector (occupancy). This table, originally containing well over 11 million records, was reduced in size by instructing the database engine to keep only observations from detectors located on highway 217, leaving information from 39 detectors at 9 locations (stations) in each direction. This substantial reduction made later analysis by the database engine much faster and ensured that results would focus on the study area. Figure 1 contains a map of Highway 217 and the detector stations that remained in the working dataset.

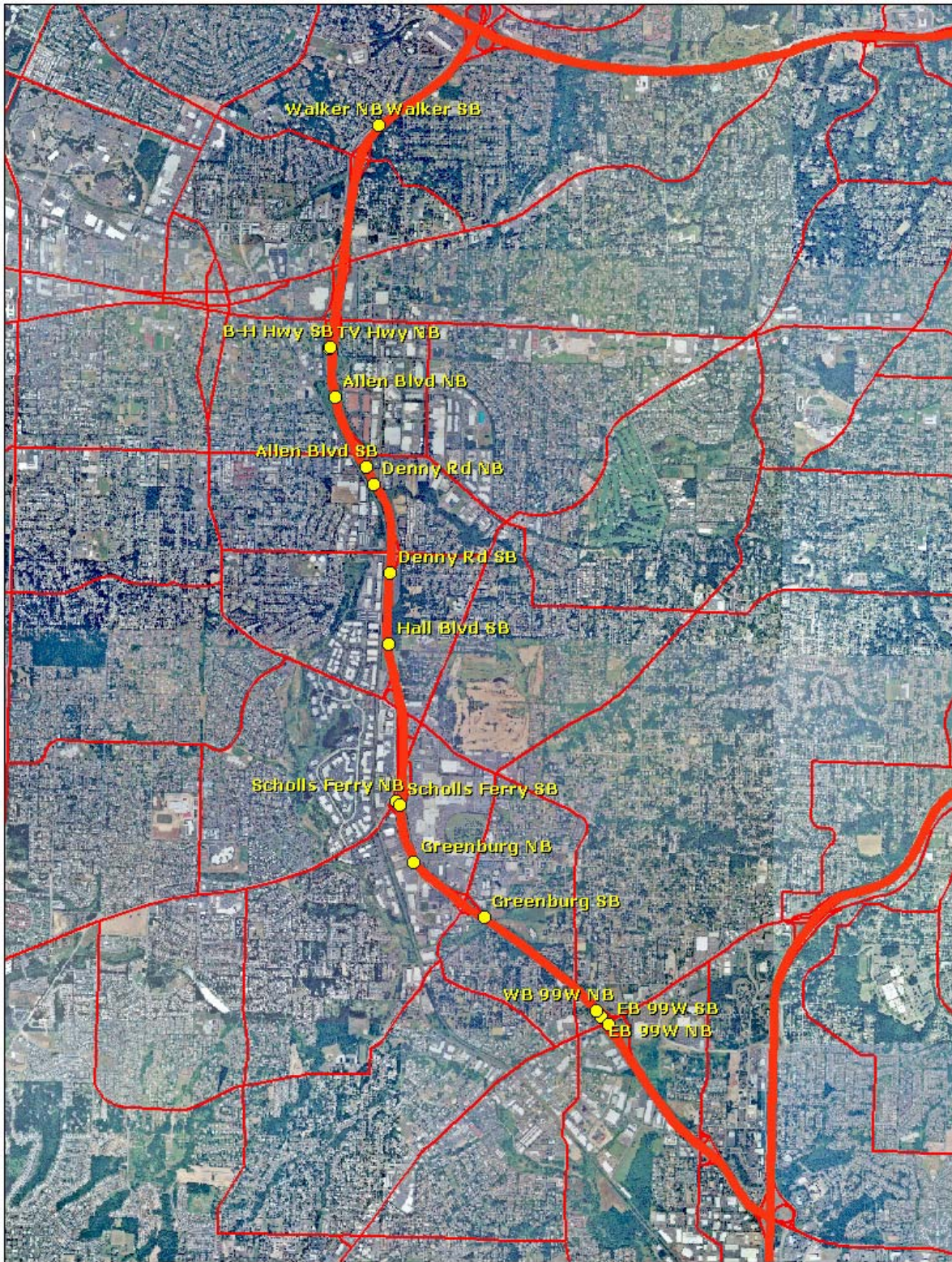


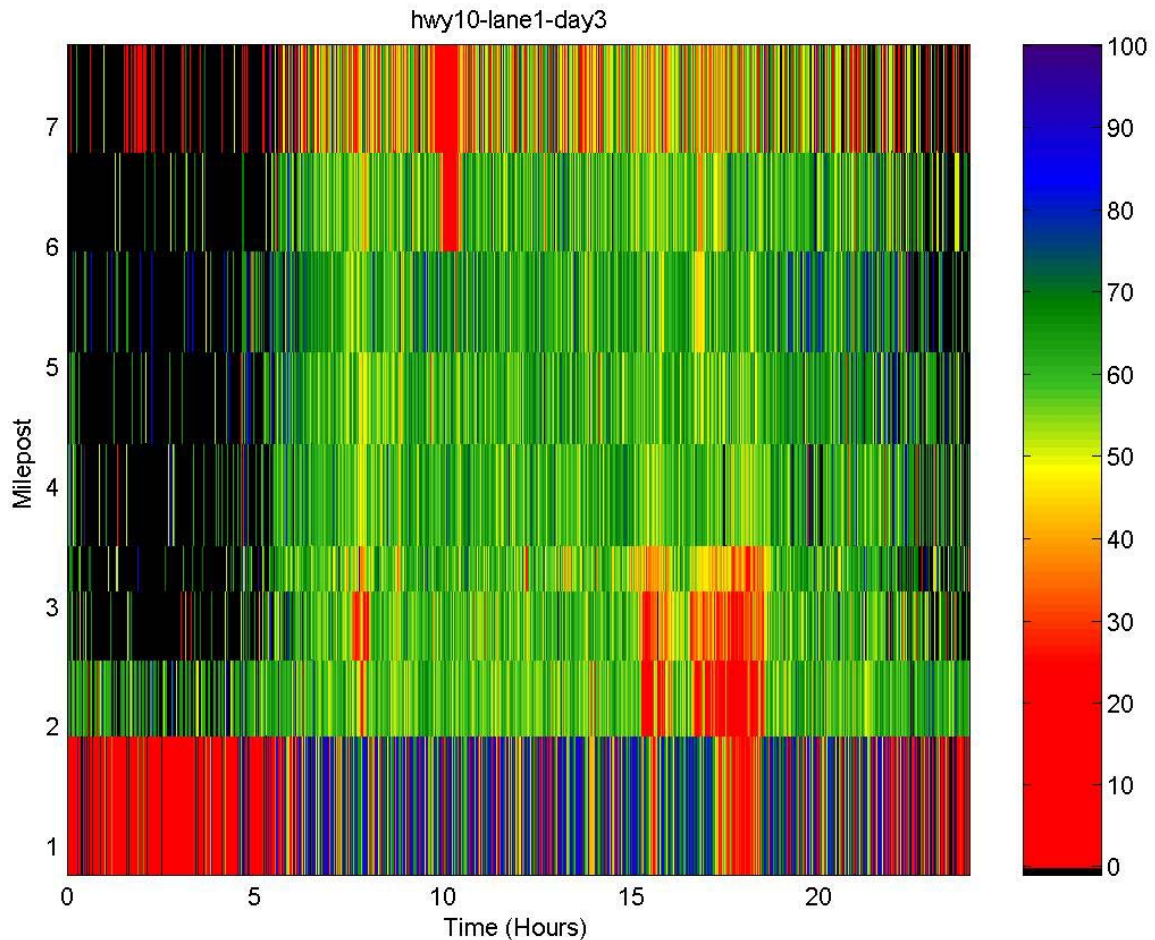
FIGURE 1 Aerial photo of study area showing arterials, highways, and loop detectors

## REMOVAL OF ERRORS AND DATA QUALITY VERIFICATION

There were a number of known issues with loop detector data from the Portland area that had been uncovered by previous studies. [Add a reference] Corrective measures were applied to the data set before analysis to minimize the impact of these errors on the results of this study.

When no vehicles pass a detector during a 20 second sample period, the detector currently reports an error code rather than a zero volume (ODOT intends to correct this in the near future). Fortunately, information about these zero-volume time periods was not needed to complete the required analysis. An update query was issued which removed all records containing these error codes from the working data set.

Malfunctioning or disabled detectors will often continue to report erroneous values, which are then recorded in ODOT's database alongside valid data. Ideally, the database should contain information about the health of individual detectors, but this functionality is not currently available.



**FIGURE 2** Speed contour plot showing malfunctioning detectors.

Reported speed or occupancy values from these unreliable detectors will often be very low or very high, and are frequently inconsistent with information from neighboring detectors. Many of these inconsistencies can be spotted using contour plots in which speed or occupancy is plotted on a plane defined by time on one axis and distance along the highway on the other, as shown in

Figure 2. In reviewing contour plots generated from the entire working data set, it was discovered that the first station at each end of the highway was reporting error codes or invalid data during most of the study period. Because of these errors, these stations were eliminated from the study.

## DETERMINING INFLUENCE AREAS

The resulting database table contained a week of high resolution information about traffic on Highway 217, sampled at discrete points along the length of the highway. For the purpose of travel demand forecasting, these point measurements need to be extrapolated to lengths of highway that can be represented as links in the model. These representative segments are often referred to as the influence areas of the detectors. The lengths of these influence areas are important values because the loop detectors cannot directly provide travel time information; travel times must be estimated by dividing segment length by speed.

In deciding which road segment a particular detector represents, several options were available. Data from each detector station could be associated with a region downstream of that station, upstream of the station, or extending in both directions around the station. As shown in Figure 3, the third approach was used for consistency with previous studies and station data from ODOT.

In this approach, each detector was assumed to represent the traffic conditions on both the segment of road extending halfway to the next upstream station and the segment extending halfway to the next downstream station. Where no next station existed, the detector's influence extended to the end of the highway. The position and length of influence areas can be easily calculated because the database provides a table specifying the position along the road (milepost) of each detector.



**FIGURE 3** Determining loop detector influence area lengths

## PASSENGER CAR EQUIVALENCE ADJUSTMENT

Volumes in the travel demand model are expressed in passenger car equivalents (PCE). Real highway traffic consists of a mix of vehicle types, so the standard vehicle is taken to be a typical passenger car and all heavier, larger, or slower vehicles are represented as the number of passenger cars displaced by their vehicle type given the local terrain and road conditions. These larger vehicles tend to both occupy more roadway space and, particularly in hilly terrain, create downstream gaps in traffic which are left unfilled.

It is technically possible to classify vehicles by length using inductive loop detectors. Unfortunately, Portland area detectors are not yet programmed to perform this function, so vehicle classification information was not available with the same resolution as the rest of our data. Instead, ODOT [State Highway Inventory Reports](#) from December 31, 2001 were used to estimate vehicle mix for this highway. These reports must be interpreted using a separate table of vehicle [classification codes](#). Counts were not available covering a full range of locations in both directions on the highway, so a single centrally located location was chosen on Southbound 217. These data are reproduced in Table 2.

**TABLE 2 Vehicle classifications for Oregon Highway 217**

Effective Date	Beginning Mile Point	Dir Cd	ADT Volume	Manl Cnt	Portbl Cnt	ATR Cnt	Vehicle Class Code	Vehicle Class Pct	Vehicle Class Volume
12/31/2001	3.02	S	115300	N	Y	N	1	0.37	427
12/31/2001	3.02	S	115300	N	Y	N	2	81.02	93416
12/31/2001	3.02	S	115300	N	Y	N	3	10.78	12429
12/31/2001	3.02	S	115300	N	Y	N	4	0.39	450
12/31/2001	3.02	S	115300	N	Y	N	5	3.56	4105
12/31/2001	3.02	S	115300	N	Y	N	6	1.11	1280
12/31/2001	3.02	S	115300	N	Y	N	7	0.09	104
12/31/2001	3.02	S	115300	N	Y	N	8	1.48	1706
12/31/2001	3.02	S	115300	N	Y	N	9	0.59	680
12/31/2001	3.02	S	115300	N	Y	N	10	0.18	208
12/31/2001	3.02	S	115300	N	Y	N	11	0.15	173
12/31/2001	3.02	S	115300	N	Y	N	12	0.05	58
12/31/2001	3.02	S	115300	N	Y	N	13	0.23	265
							sum 4-13	7.83	

According to information taken from the Highway Capacity Manual [specify year and page], the PCE for trucks traveling on level terrain is 1.2. The overall heavy vehicle correction factor can then be determined as follows:

$$f_{HV} = \frac{1}{1 + P_T(E_T - 1)}$$

Where:

$$P_T = \text{proportion of trucks (combinations plus single units) from inventory, expressed as a decimal}$$

$$E_T = \text{passenger car equivalents}$$

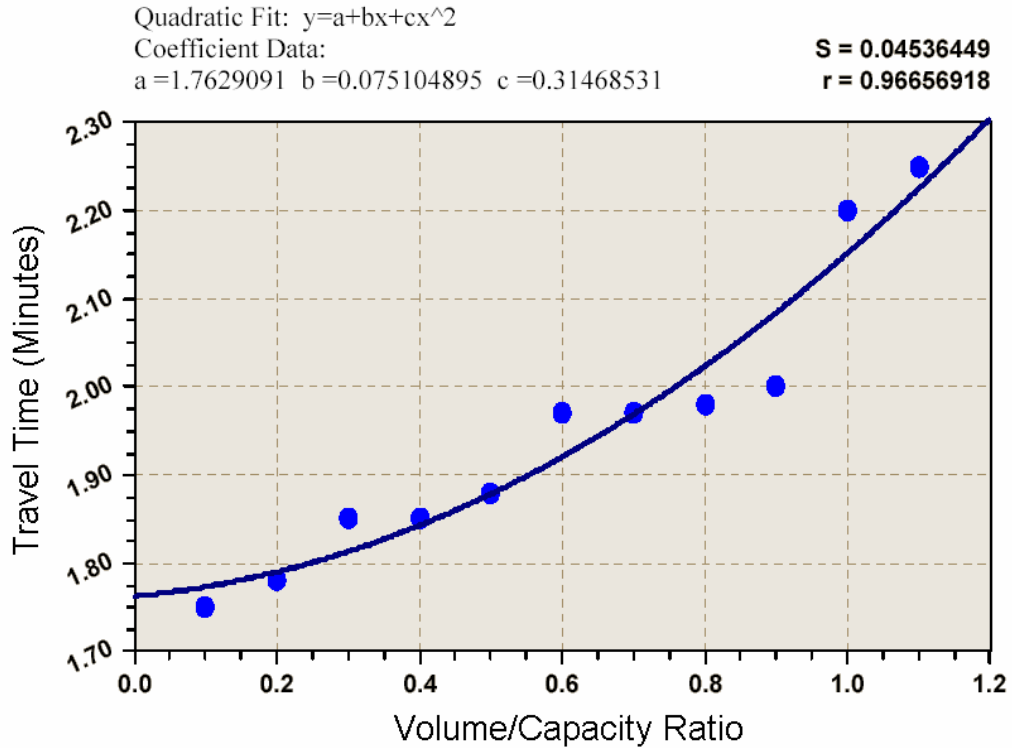
Summing the percentages for vehicle classifications 4 through 13 gives a total of 7.83 percent heavy vehicles, and using this value in the above equation gives a correction factor of 1.016. All vehicle counts in the data set were multiplied by this PCE correction factor for consistency with other portions of the demand model.

## ANALYSIS

The trip distribution process in EMME/2 specifically requires functions relating the volume/capacity (V/C) ratio on a link to the travel time over that link. The first step to reaching this end goal of our analysis was to develop a relationship between PCE flow and travel time.

For each detector individually, all reported volume-speed pairs were grouped together regardless of time of observation. All volumes were then multiplied by a factor of 180 to convert the reported 20-second counts to equivalent hourly flows. A travel time was determined for each data point by dividing the segment length that the detector represented by the reported speed.

Because the 20-second volume figures are integer values, the equivalent hourly rates are naturally quantized to multiples of 180. The resulting flow – travel time pairs can therefore be sorted into about sixteen ‘bins’ based on these different hourly flows. For each of these bins, the average and standard deviation of the reported travel times was calculated and plotted. At this point in the study, flows were divided by an assumed capacity of 2400 vehicles per hour to allow presentation of the data in a form similar to that required by the demand model. The first of these plots is shown in Figure 4, including a regression curve derived from the data.



**FIGURE 4 Preliminary plot of V/C to travel time relationship at one detector assuming capacity of 2400 vehicles per hour.**

The use of true travel times based on the varying road segment lengths at different detector stations produced functions which were difficult to validate or compare with one another. To facilitate comparison, a normalized form of travel time was later used in which the inverse of the reported speeds was multiplied by 60, giving a result with units of minutes per mile. The resulting functions were more readily comparable, but could be converted to output true travel times by multiplying the normalized travel time by the segment length in miles.

Most of this analysis process was expressed in structured query language (SQL) and submitted to the database engine as a single query. The database engine and query optimizer determined the most efficient way to extract the requested information and returned it in tabular form, ready for plotting and regression analysis. This query is shown in Table 3 with a sample of the resulting data as it would appear in a database table in Table 4.

**TABLE 3 Example SQL query issued to perform analysis on loop detector data.**

```

SELECT C.detectorId, H.highwayName, H.longDirection, D.milePost, D.locationText, D.laneNumber,
       C.pce_flow, C.pce_flow / D.capacity AS pce_vc_ratio, count(*) AS num_datapoints,
       stdev((S.stationlength/C.speed)*60) AS stdev_time, avg((S.stationlength/C.speed)*60) AS avg_time,
       avg(C.occupancy) AS avg_occupancy
FROM data217 AS C, detectors AS D, stations AS S, highways AS H
WHERE C.detectorid = D.detectorid
AND D.stationid = S.stationid
AND D.highwayId = H.highwayId
AND (D.highwayId = 9 or D.highwayId = 10)
GROUP BY C.detectorid, H.highwayName, H.longDirection, D.milepost, D.locationtext, D.lanenumner,
         C.flow
HAVING count(*) > 100;
    
```

**TABLE 4 Resulting database table from the above query.**

detector	highway	longDirection	milepost	locationtext	lane	pce_flow	pce_vc_ratio	num_datapoints	stdev_time	avg_time	avg_occ
1488	217	NORTH	6.61	72nd Ave NB t	1	184	0.07666667	274	19.129068	28.0962	3.9197
1488	217	NORTH	6.61	72nd Ave NB t	1	367	0.15291667	788	13.273118	16.19562	3.5495
1488	217	NORTH	6.61	72nd Ave NB t	1	551	0.22958333	1393	10.901135	11.14845	3.649
1488	217	NORTH	6.61	72nd Ave NB t	1	734	0.30583333	1915	8.2327717	7.706607	3.8366
1488	217	NORTH	6.61	72nd Ave NB t	1	918	0.3825	2196	6.0736122	5.334191	4.3201
1488	217	NORTH	6.61	72nd Ave NB t	1	1102	0.45916667	2328	5.9975480	4.435000	5.0949
1488	217	NORTH	6.61	72nd Ave NB t	1	1285	0.53541667	2328	4.6836066	3.237047	6.0301
1488	217	NORTH	6.61	72nd Ave NB t	1	1469	0.61208333	2198	2.6246818	2.349118	7.6733
1488	217	NORTH	6.61	72nd Ave NB t	1	1652	0.68833333	1991	2.0139127	2.064485	9.0090
1488	217	NORTH	6.61	72nd Ave NB t	1	1836	0.765	1821	0.9531100	1.770364	9.8523
1488	217	NORTH	6.61	72nd Ave NB t	1	2020	0.84166667	1405	0.9162783	1.648790	10.671
1488	217	NORTH	6.61	72nd Ave NB t	1	2203	0.91791667	1070	0.4893894	1.506646	10.979
1488	217	NORTH	6.61	72nd Ave NB t	1	2387	0.99458333	683	0.4030402	1.421709	10.943
1488	217	NORTH	6.61	72nd Ave NB t	1	2570	1.07083333	466	0.2928134	1.341658	11.279
1488	217	NORTH	6.61	72nd Ave NB t	1	2754	1.1475	236	0.2357972	1.286756	11.301
1488	217	NORTH	6.61	72nd Ave NB t	1	2938	1.22416667	117	0.1774824	1.242546	12.043
1489	217	NORTH	6.61	72nd Ave NB t	2	367	0.15291667	172	13.381014	13.65847	10.698
1489	217	NORTH	6.61	72nd Ave NB t	2	551	0.22958333	463	8.7157974	9.255895	7.8164
1489	217	NORTH	6.61	72nd Ave NB t	2	734	0.30583333	1037	8.2393877	7.227688	8.3539
1489	217	NORTH	6.61	72nd Ave NB t	2	918	0.3825	1603	5.7875835	5.118584	7.8422
1489	217	NORTH	6.61	72nd Ave NB t	2	1102	0.45916667	2339	6.0804025	4.502112	8.1672
1489	217	NORTH	6.61	72nd Ave NB t	2	1285	0.53541667	2880	4.4856085	3.589240	8.5997
1489	217	NORTH	6.61	72nd Ave NB t	2	1469	0.61208333	3099	4.4726683	3.171985	9.0397

Because of the large size of the dataset to be analyzed, the use of a relational database had clear advantages in this research. Taking the time to structure the original data allows the rapid, repeatable use of many different analysis techniques by issuing saved queries. Relational databases in ITS research are an essential tool that seems to underused.

**ESTIMATING CAPACITIES**

At this point, an assumed capacity of 2400 vehicles/hour was being used to convert all reported volumes to V/C ratios. The final step in the analysis was to replace this assumed value with an empirically estimated capacity of the particular facility. Theoretical capacity values could be

calculated for each highway segment, but in keeping with the concept in this study of seeking accuracy from ATMS data, it was desirable to determine capacities directly from loop detector data. Several techniques were discussed, and it was eventually decided that the maximum 15-minute flow on each segment during the study period would be regarded as its capacity. [Show sample cumulative plot of flow and indicated “capacity” and also mention that only one day’s data were used for this step] For each detector, the observations were sorted by time and the sum of the next 15 minutes’ volumes was computed at each 20-second interval.

Demonstrating that the maximum 15-minute flows often coincided with the onset of congestion increased confidence that these figures were close to the true capacities being sought. The occupancy figure reported by loop detectors is known to be related to congestion, so the 15-minute flows were plotted together with occupancy versus time. [Show sample] The majority of the maxima in the 15-minute flows were indeed located just before notable increases in occupancy, and therefore presumably just before the onset of congestion.

There is some question as to whether these values should be referred to as ‘capacities’ since this word has a strong preestablished meaning. However, the results of this process appear to be the maximum sustainable volume of traffic on a particular segment, beyond which a breakdown in flow occurs. This meaning of ‘capacity’ was determined to be acceptable for the purposes of this study by Metro staff. Results which did not appear to meet this description were discarded.

## RESULTS

Because some uncertainty exists about the validity of these capacity figures, the final result set is presented in Table 5 as the average and standard deviation of segment travel time by detector and flow bin. These flows may be divided by the reported capacity figure for each detector if a V/C to travel time relationship is desired. [can we present this graphically somehow, or compare it to default values?]

These results were presented to Metro, who performed regression analysis on the data for each detector, producing the necessary functions for EMME/2 trip distribution. [request from Kyle Hauger: 1 or 2 example travel time plots with regression curves]



## RECOMMENDATIONS / CONCLUSIONS

The results of this study indicate that it is indeed feasible to use ATMS data in the travel demand forecasting process. [[request from Kyle Hauger: qualitative discussion of effectiveness in demand model](#)]

If these results prove to be useful and accurate, this process could be refined and repeated for other highways in the Portland region. It may even be possible to incorporate the production of trip distribution functions into the new regional data archiving and evaluation system under development at PSU, allowing this analysis to be repeated automatically using larger or more reliable data sets.

In this study, only a week of loop detector data was analyzed. Future studies should attempt to use larger, more carefully chosen data sets in which all detectors are providing reliable information. The use of more formal methods for determining data quality would increase confidence in the results. Techniques for automatically determining the validity of loop detector data have been documented as part of the California PEMS system. [[add reference](#)]

It is recommended that future results be presented as normalized (1-mile) travel times, organized by hourly flow rather than V/C ratio as some readers may question the method used in calculating capacities. Segment lengths and estimated capacities should then be reported for each detector.

Lastly, this study clearly demonstrates the utility of relational databases in analyzing these large, structured sets of transportation data. Desired results can often be clearly expressed as a database query, with the majority of sorting and analysis being optimized and carried out by fast, well-tested software.

## REFERENCES

## ACKNOWLEDGMENTS

## LIST OF TABLES

Table 1	Database Table Containing A Full Week Of Loop Detector Data
Table 2	Vehicle Classifications For Oregon Highway 217
Table 3	Example SQL Query Issued To Perform Analysis On Loop Detector Data.
Table 4	Resulting Database Table From The Above Query.
Table 5	Summary Table Showing Average (Grey) And Standard Deviation (Blue) Of Travel Times For Each Flow Range At Each Detector.

## LIST OF FIGURES

Figure 1	Aerial Photograph Of The Study Area With Arterials, Highways, And Loop Detector Locations Marked.
Figure 2	Speed Contour Plot Showing Malfunctioning Detectors.
Figure 3	Determining Loop Detector Influence Area Lengths
Figure 4	Preliminary Plot Of V/C To Travel Time Relationship At One Detector Assuming Capacity Of 2400 Vehicles Per Hour.