6. ANSWERS TO SPECIFIC QUESTIONS

The data analysis portion of this project was intended, in part, to answer eleven specific questions. These questions will be addressed specifically in this section of the report with references to related section of the report for additional information.

Question 1. How Accurately do the various types of vehicle classifiers sort the traffic stream into each of the 13 FHWA vehicle types?

Table XXXIX depicts the accuracy of each of the classifiers to sort the vehicle stream into the 13 FHWA vehicle types. This table lists the percent of the vehicle stream correctly classified and the percent sensor errors (axle miscounts) for each vendor and each 48-hour test. Note that the number in parenthesis in the "% Correctly Classified" column is the percentage of correctly classified vehicles with classes 2 and 3 combined. Errors between classes 2 and 3 are by far the most common errors since often time small pickup trucks have shorter lengths and wheelbases than large cars. Therefore, it is often impossible to correctly sort vehicles between class 2 and 3 based on axle spacing and length.

The accuracy of the classification appears to be inversely proportional to the ability of the classifiers to count axles. Missed axles (referred to as sensor errors) seem to directly relate to sensor errors and can be used as an indicator of overall classifier accuracy. Figure 72 in Section 5.3 plots the percentage of correctly classified vehicles as a function of the percent sensor errors (miscounted axles). As the number of sensor errors increased the accuracy of the classifiers declined almost linearly.

Question 2. For those vehicles that may be incorrectly classified by a device, into what class did the devices place the vehicle?

The classification matrices for each of the vendors are shown is Section 5.2. The matrices show the count and percentage of vehicles that were correctly and incorrectly classified into each of the categories. The classes into which an incorrectly classified vehicle was placed varied somewhat between the individual classifiers devices. Some general trends can however be noted.

The primary misclassification noted in the test results was the classification of class 3 vehicles as class 2 vehicles. Since these two classes of vehicles often have similar overall and wheelbase lengths (the primary measures for sorting the vehicles), this error is easily understood. Classifications of class 2 vehicles as class 3 occurred much less often than class 3 to class 2, but did occur in significant numbers.

Classification errors between all of the 2-axle classes (classes 2, 3, 4, and 5) were also common. Again, there are a number of small buses and 5-tire trucks, and large cars and pickup trucks with very near the same lengths and wheelbases. Therefore, differentiation of the classes based entirely on axle spacing and length is nearly impossible.

Table XXXIX. Classification Accuracy

Equipment Vendor	Model Number	Sensor Config.	48-Hour Test #	% Sensor Errors	% Correctly Classified
Mikros Systems	TEL-2CM	L-P-L (Philips)	1	13.91	70.3 (82.5)
			2	4.81	63.5 (78.8)
Peek Traffic, Inc.	TrafiCOMP III	P-L-P (Philips)	1	3.71	75.3 (93.0)
			2	3.53	74.8 (93.2)
	GK-6000	P-P (Philips)	1	6.29	73.7 (90.1)
			2	1.19	79.0 (96.2)
		P-P (Philips)	1	4.02	77.1 (92.4)
			2	1.07	79.1 (96.2)
PAT Equipment Corporation, Inc.	AVC-100	P-L-P (Atochem)	1	N/A	N/A
			2	N/A	N/A
	AVC-100	L-P-L (Philips)	1	1.49	76.6 (95.0)
			2	0.51	76.6 (95.1)
MITRON Systems Corporation	MSC-3000 DCP	P-P (Autologger)	1	N/A	N/A
			2	N/A	N/A
Electronic Control Measure	HESTIA	P-L-P (ECM)	1	10.05	67.5 (86.4)
			2	6.70	72.0 (94.3)
TimeMark, Inc.	Delta II	P-P (Philips)	1	2.83	78.9 (94.4)
			2	1.97	77.3 (94.6)
International Road Dynamics, Inc.	TC/C 530-4D/4P/4L	PR-L-PR (Dynax)	1	8.89	69.3 (88.4)
			2	5.35	72.6 (92.9)
		P-L-P (Philips)	1	7.25	70.8 (89.7)
			2	4.37	73.8 (93.9)
Golden River Traffic Ltd.	Marksman 660	P-L-P (Traffic)	1	N/A	N/A
			2	7.91	63.9 (82.3)
Diamond Traffic Products	TT-2001	P-L-P (Autologger)	1	3.17	73.1 (92.0)
			2	11.65 *	11.0 (11.8)*
		P-L-P (Philips)	1	3.47	75.5 (93.9)
			2	3.40 *	13.0 (16.2)*

* - Software During Test 2 Had Bug   N/A - Not Available for This Test

Classification of large trucks, especially classes 8 and 9, was very good. Classification errors were relatively uncommon. Common classification error that did occur included splitting a class 9 into a class 6 and a class 2, and classification of a class 9 as a class 8 (miscounted axles).

Question 3. For those vehicle classification devices that capture individual vehicle records of number of axles and distances between axles, how accurately monitored was this axle and distance information?

The accuracies of individual classifiers to measure axle spacings for both 48-hour tests are summarized in Table XL. The magnitude of the mean of the error in measuring the axle spacings ranged from 0.003 to 0.441 feet (0.4 to 5.3 inches), and the standard deviation ranged from 0.335 to 2.791 feet. Note that these ranges do not include those values for the Marksman 660 model, which had considerable errors throughout its data.

Question 4. For those vehicle classification devices that allow the operator to set the dimensional threshold between various vehicle classes, how accurate were these devices as defined by preceding points 1 through 3?

Virtually all of the equipment tested had the ability in some way or another to allow the operator to adjust the dimensional thresholds between vehicle classes. All of the equipment tested used the default dimensional thresholds of the classifier except the Mikros classifier. The Mikros classifier was programmed with the dimensional thresholds known as Schedule F or the Streeter-Amet Classification Categories ("Field Evaluation of FHWA Vehicle Classification Categories", Final Report, Technical Paper 84-5, Maine Department of Transportation, January 1985).

Question 5. How is the equipment accuracy as defined in the preceding points 1 through 3 affected by vehicle speed?

The accuracy of the classification equipment as a function of speed is discussed in the individual classifier results in Section 5.2. Most of the devices reported vehicle speed in their output files. The speed calculated by the vendor was used to separate the axle spacing, wheelbase and overall length measurements into separate speed bins. The mean and standard deviation of the measurements were calculated for each 5 MPH wide speed bin over the range of speeds from 30 to 80 MPH. Very few vehicles were recorded with speeds greater than 80 MPH or less than 30 MPH during any of the tests.

The effect of speed on the accuracy of the measurements varied from classifier to classifier. Typically, the mean error and standard deviation of the error slowly changed as a function of speed. The direction (towards positive or negative errors) and the magnitude of the change varied depending on the classifier. The absolute magnitude of the change was, however, fairly small for all of the classifiers.

Table XL. Accuracy of Axle Spacing Measurements
(All measurements in feet)

Equipment Vendor	Model Number	Sensor Config.	48-Hour Test #	Mean Error	Standard Deviation
Mikros Systems	TEL-2CM	L-P-L (Philips)	1	0.099	1.113
			2	-0.022	0.988
Peek Traffic, Inc.	TrafiCOMP III	P-L-P (Philips)	1	0.172	0.399
			2	0.064	0.356
	GK-6000	P-P (Philips)	1	0.110	1.449
			2	0.004	0.363
		P-P (Philips)	1	0.098	2.722
			2	0.003	0.335
PAT Equipment Corporation, Inc.	AVC-100	P-L-P (Atochem)	1	N/A	N/A
			2	N/A	N/A
	AVC-100	L-P-L (Philips)	1	0.078	2.055
			2	-0.252	0.420
MITRON Systems Corporation	MSC-3000 DCP	P-P (Autologger)	1	N/A	N/A
			2	N/A	N/A
Electronic Control Measure	HESTIA	P-L-P (ECM)	1	0.120	0.732
			2	0.441	1.519
TimeMark, Inc.	Delta II	P-P (Philips)	1	0.124	2.791
			2	-0.018	0.419
International Road Dynamics, Inc.	TC/C 530-4D/4P/4L	PR-L-PR (Dynax)	1	0.159	0.578
			2	0.099	0.515
		P-L-P (Philips)	1	0.139	0.726
			2	-0.014	0.610
Golden River Traffic Ltd.	Marksman 660	P-L-P (Traffic)	1	N/A	N/A
			2	-4.625	45.606
Diamond Traffic Products	TT-2001	P-L-P (Autologger)	1	0.011	2.340
			2	-0.166	0.978
		P-L-P (Philips)	1	0.133	2.124
			2	-0.203	0.498

N/A - Data Not Available

Question 6. How is equipment accuracy as defined in the preceding points 1 through 3 affected by the percentage of vehicles with more than 2 axles in the traffic stream being measured?

The equipment accuracy as a function of the number of trucks (vehicles with greater than two axles) for each classifier is presented in tables in Section 5.2. In that section, it shows that the classification accuracy as a function of the percentage of trucks does tend to show an improvement as the percentage of trucks increases. However, the percentage of trucks generally rose dramatically in the late evening and early morning hours when the traffic volume was very low. The effects of the lower traffic volume is likely to have had at least as great an impact on the classification accuracy as the percentage of trucks in the vehicle stream.

The general patterns of incorrect classifications showed no appreciable change as a function of the number of trucks in the traffic stream. The total number of class 3 vehicles classified as class 2 (the most common classification error) decreased as the traffic volume and number of class 3 vehicles decreased, but the tendencies and percentage of class 3 vehicles incorrectly classified stayed nearly the same.

Axle spacing measurements tended to show little or no appreciable variation with the percentage of vehicles with more than two axles. The response of individual classifiers are shown in tables in Section 5.2.

Question 7. What is the impact of pavement and/or air temperature on equipment accuracy?

The impact of pavement and air temperature on the classification and axle spacing measurement accuracy for each classifier and test is presented in tables in Section 5.2. Very few conclusions can be drawn from the data collected from the tests conducted in this project. The range of temperatures experienced during the test was relatively limited with no near or below freezing air temperatures experienced. Over the range of air and pavement temperatures experienced during the two 48-hour tests, the classification accuracy appeared to vary very little. At very low pavement temperatures (60-69 degrees), the mean error did appear to rise slightly, but at these times (early mornings) the traffic was also very light and consisted mainly of large trucks. Therefore, the measurement of axle spacings appeared to be affected very little by the temperature ranges experienced in these tests.

Question 8. What is the impact of precipitation on equipment accuracy?

During the tests performed for this project, no appreciable precipitation was recorded. Therefore, no information of sensor or classifier accuracy can be derived from the test results. All of the classifiers tested used combinations of piezoelectric sensors and magnetic loop detectors. While there is the possibility of some sensor degradation due to hydroplaning of vehicle tires or interference with the magnetic loops, the degradation is expected to be relatively minor. However, there is no test information which can be used to validate this assumption.

Question 9. How accurately can devices monitor overall vehicle length as a function of vehicle speed, traffic volume, temperature, or vehicle mix?

The accuracy of the classification devices which measured overall vehicle lengths is listed in Tables XLI and XLII. In Table XLI, the mean and standard deviation of the measurement errors (ground truth (true) length minus measured length) are listed for each vendor, classifier model and test. In Table XLII, the mean and standard deviation of the percentage length measurements are listed. The length percentage is calculated by multiplying the true length by 100 and dividing by the length measured by the classification equipment. These tables are summaries of the results of the individual classifiers listed in Section 5.2.

The effects of traffic speed on the overall length measurements of the individual classifiers can be examined by reviewing the figures in Section 5.2. The effects of traffic speed on the accuracy of the overall length measurements varied considerable from classifier to classifier. The Electronic Control Measure HESTIA had increasing mean errors at higher speeds (approaching 70 mph). The Mikros TEL-2CM showed a trend towards higher errors at lower speeds. The International Road Dynamics TC/C 530-4D/4P/4L and the Diamond TT-2001 units tended to have the greatest accuracies at speeds near 50-55 mph, and tended towards higher errors at higher and lower speeds (though there were very few vehicles with much lower speeds).

The accuracy of overall length measurements as a function of the air and pavement temperature tended to be nearly constant. The accuracy varied very little as the temperature changed for the classifiers tested. This is reasonable since the overall length measurements are made primarily using the loop sensor inputs (and possibly speed measured by the axle sensors). Magnetic loop sensors do not exhibit large variations as a function of temperature.

The overall length measurement accuracy as a function of the percentage of trucks (vehicles with greater than two axles) tended to degrade as the percentage of trucks increased. The magnitude of the mean error in overall length measurement increased by as much as 40% when the percentage of trucks increased from 0-20% to 40-60% (HESTIA during Test 2).

The accuracy of overall length measurements as a function of the traffic density is nearly impossible to separate from the accuracy versus temperature and percent of vehicles with greater than two axles. The high volume traffic occurred during the daytime hours where temperature was highest and the percentage of trucks was lowest. The traffic at the

Table XLI. Overall Length Measurement Errors
(All measurements in feet)

Equipment Vendor	Model Number	Sensor Config.	48-Hour Test #	Mean Error	Standard Deviation
Mikros Systems	TEL-2CM	L-P-L (Philips)	1 2	2.093 -2.167	2.549 4.031
Peek Traffic, Inc.	TrafiCOMP III	P-L-P (Philips)	1 2	N/A N/A	N/A N/A
	GK-6000	P-P (Philips)	1 2	N/A N/A	N/A N/A
		P-P (Philips)	1 2	N/A N/A	N/A N/A
PAT Equipment Corporation, Inc.	AVC-100	P-L-P (Atochem)	1 2	N/A N/A	N/A N/A
	AVC-100	L-P-L (Philips)	1 2	N/A N/A	N/A N/A
MITRON Systems Corporation	MSC-3000 DCP	P-P (Autologger)	1 2	N/A N/A	N/A N/A
Electronic Control Measure	HESTIA	P-L-P (ECM)	1 2	0.598 2.232	2.157 3.073
TimeMark, Inc.	Delta II	P-P (Philips)	1 2	N/A N/A	N/A N/A
International Road Dynamics, Inc.	TC/C 530-4D/4P/4L	PR-L-PR (Dynax)	1 2	-1.321 -1.739	11.302 10.952
		P-L-P (Philips)	1 2	-1.154 -2.152	11.423 10.493
Golden River Traffic Ltd.	Marksman 660	P-L-P (Traffic)	1 2	N/A N/A	N/A N/A
Diamond Traffic Products	TT-2001	P-L-P (Autologger)	1 2	-1.271 -3.717	8.747 4.613
		P-L-P (Philips)	1 2	-1.196 -2.382	5.845 1.600

N/A - Data Not Available

Table XLII. Overall Length Measurement Percentiles
(All measurements in percent of ground truth)

Equipment Vendor	Model Number	Sensor Config.	48-Hour Test #	Mean Percent	Standard Deviation
Mikros Systems	TEL-2CM	L-P-L (Philips)	1 2	108.43 91.29	26.62 28.32
Peek Traffic, Inc.	TrafiCOMP III	P-L-P (Philips)	1 2	N/A N/A	N/A N/A
	GK-6000	P-P (Philips)	1 2	N/A N/A	N/A N/A
		P-P (Philips)	1 2	N/A N/A	N/A N/A
PAT Equipment Corporation, Inc.	AVC-100	P-L-P (Atochem)	1 2	N/A N/A	N/A N/A
	AVC-100	L-P-L (Philips)	1 2	N/A N/A	N/A N/A
MITRON Systems Corporation	MSC-3000 DCP	P-P (Autologger)	1 2	N/A N/A	N/A N/A
Electronic Control Measure	HESTIA	P-L-P (ECM)	1 2	102.95 114.30	8.25 16.07
TimeMark, Inc.	Delta II	P-P (Philips)	1 2	N/A N/A	N/A N/A
International Road Dynamics, Inc.	TC/C 530-4D/4P/4L	PR-L-PR (Dynax)	1 2	98.35 96.21	51.48 52.28
		P-L-P (Philips)	1 2	99.38 94.21	52.81 47.94
Golden River Traffic Ltd.	Marksman 660	P-L-P (Traffic)	1 2	N/A N/A	N/A N/A
Diamond Traffic Products	TT-2001	P-L-P (Autologger)	1 2	97.62 94.04	45.35 14.75
		P-L-P (Philips)	1 2	95.06 90.82	14.46 6.56

N/A - Data Not Available

test site did tend to increase in the mornings and evenings, but there was no heavy rush hour traffic that could be used to assess performance under extreme traffic conditions. Therefore, the dependence on traffic volume is not expressly calculated for this study.

Question 10. How is the device accuracy affected by the sensing device used?

All of the classification systems tested for this project used some combination of piezoelectric axle sensors and inductive loops as sensing devices. The combination and configuration of the sensing devices can possibly have a significant impact on the accuracy of the classifiers. Two of the vendors (PAT Equipment Corp. and International Road Dynamics) used two identical classification models with two different sensor configurations. These provided good points for comparisons of the accuracy as a function of the sensing device configuration.

The primary difference between the sensing devices used by the individual classifiers was the number of piezoelectric axle sensors and inductive loops used. Most used 2 axle sensors and a single loop as the sensing devices. The Peek GK-6000, the TimeMark Delta II and the Mitron MSC 3000 DCP classifiers used no inductive loops. The Mikros TEL-2CM and one configuration of the PAT AVC-100 used a single axle sensor and two inductive loops. The remaining classifiers used two axle sensors and one inductive loop for their sensor configuration.

There was no appreciable difference in the classification accuracy between the equipment using different sensor configurations. Within the set of classifiers using each particular sensor configuration there was considerable variation in classification accuracy, but there was no trend or pattern between sensor configurations.

The magnitude of the mean error of the axle spacing measurements for classifiers using a single piezoelectric axle sensor averaged 0.113 feet, and the standard deviation averaged 1.144 feet. The magnitude of the mean error for classifiers using two piezoelectric axle sensors averaged 0.115 feet, and the standard deviation averaged 1.081 feet. The difference between one and two axle sensor configurations showed no appreciable difference in these tests. Again, due to considerable errors, the data for the Marksman 660 model was not included in these calculations.

All but one of the classifiers which measured overall vehicle length used the P-L-P sensor configuration. Only the Mikros TEL-2CM used the L-P-L configuration. The Mikros classifier's overall length measurement errors were near the average for the units using the P-L-P configuration. Therefore, there was not appreciable difference in overall length measurement accuracy noted in this project.

Question 11. Does the accuracy change over time? Is the accuracy the same when the equipment is first installed and when it is picked up, or does the length of time in operation affect the equipment?

The two detailed 48-hour tests were conducted approximately 4 months apart. Between the tests, the sensing devices remained in the roadway and were subjected to normal traffic volume. Comparing the classification accuracy of the each vendor classifier in the first and second 48-hour test can be accomplished by reviewing Table XXXIX (of Question 1. above). No Piezoelectric sensors were replaced between the two tests, but some piezoelectric sensors did fail. Those classifiers which used the failed sensors were not included in the second test and their results are not included in the table.

Nearly all of the classifiers had nearly equal to or better vehicle classification accuracy during the second 48-hour test than during the first test. The only notable exception to this pattern was the Diamond TT-2001 which had a software upgrade between tests that included a data recording bug. Using lessons learned during the first test, greater care was taken during the second test to prepare the piezoelectric sensors. Where axle sensors were installed low in the pavement, or where piezoelectric output was low, bituthane tape was placed over the sensor to increase its output voltage levels. This resulted in a lower sensor error (miscounted axle) rate during the second test and a corresponding higher percentage of correctly classified vehicles. The axle sensor accuracy appeared to have a greater effect on classification accuracy than time, and the sensor accuracy was dependent primarily on installation and preparation prior to each test.

Axle spacing and overall length measurement accuracy revealed no noticeable trend between the first and second 48-hour test. Almost all had mean errors which remained nearly constant between the tests.

An assessment of the accuracy over time during a continuous test can be made by reviewing the results in Section 5.2 of the 7-day test results for each vendor. No specific trends were noted in the accuracy of the classifiers over the 7-day test. Some classifiers displayed some accuracy improvement from the first to the last day of the test, but others showed some degradation. Therefore, the classification accuracy as a function of time over a seven day period depended on the particular classifier and no overall trends were noted.

7. AUGMENTED PNEUMATIC TUBE TEST

An augmented pneumatic tube test was conducted during the second 48-hour test. A Peek 241 (TrafiCOMP III) was configured to monitor both westbound lanes of traffic using pneumatic tubes. One pair of tubes covered the near lane (same lane as used for all other tests), and one pair of tubes were placed across both lanes of traffic.

The goal of the test was to determine how accurately the classifier could count vehicles and axles in two lanes at the same time. Obviously, the second set of tubes would detect axles in both lanes of traffic simultaneously. The goal of the test was to determine if the lanes of traffic could be separated and counted accurately.

The Peek 241 was intended to be configured to record and bin the separate lanes of traffic. Unfortunately, the classifier was configured to sum both lanes of traffic into one bin. The resulting data file had all vehicles in the near lane counted twice and the far lane counted once. This format adds to the ambiguity of the counting if one lane or the other has an error and thus would not be a fair test of the pneumatic tubes to count vehicles and axles.

It is recommended that this test be repeated in the future. The test should be conducted either with two separate classifiers, or with one classifier configured to store the results of each pneumatic tube pair in separate bins.