Cooperative agreement

No. DTFH61-99-X-00104

Intelligent Vehicle Initiative (IVI) Field Operational Test Program

Volume 4 of 4: Road Geometry Report

Freightliner Trucks Field Operational Test: The Freightliner/Meritor Wabco Roll Stability Advisor and Control at Praxair

By

Seth Rogers
DamilerChrysler Research and Technology North America

For

U.S. Department of Transportation
Federal Highway Administration
Washington, D.C.

September , 2002

Introduction

The IVI-RSA (Intelligent Vehicles Initiative-Rollover Stability Advisor) project is designed to evaluate and extend measures to reduce truck rollover. Current technology includes a box that measures a "Rollover" score while a truck rounds a curve, and communicates that score to the driver. The nearer the score to 100, the closer the truck came to tipping over. The intention is that the driver will learn to correct his own behavior when he sees examples of dangerous driving.

In this project, we collected data from many trips to test this hypothesis, and to improve the technology. One improvement would warn the driver ahead of the curve if the situation is dangerous, and possibly automatically slow the truck. This improvement requires a prediction of the rollover score without intervention, which in turn requires an accurate estimate of the radius of curvature. At the DaimlerChrysler Palo Alto research lab, we have an active research program in creating highly accurate maps with curvature from large collections of less accurate positioning traces.

PART I. Statistical Analysis

In this part, we give an overview of the data and processing results on all data, without attention to individual regions.

Raw Data

Positioning hardware was a differential GPS receiver. Positions were recorded twice a second. At each position, the onboard computer recorded time, longitude, latitude, height, dilution of precision, heading, speed, and number of satellites. The platform was a fleet of six liquid nitrogen delivery trucks. The trucks made daily runs through Indiana, Michigan, and surrounding states. Data were collected over a period of about 10 months, resulting in about 5000 usable vehicle traces. The traces covered about 10,000 hours of driving, or 773,000 kilometers. GPS requires at least 4 visible satellites to make a position fix. More is helpful because the geometry is likely to be better. The histogram in Figure 1 describes the satellite availability.

96% of the data reflect differential corrections. There are about eight differential beacons in range of at least part of the test area. Although the test data did not indicate which beacon(s) were in range for differential corrections, we can use this data to make a rough map of differential availability. The map in Figure 2 depicts a sampling of points with and without differential corrections.

The accuracy of the raw data is a key issue for our processing, but similarly important is a good accuracy estimate. We can use such an accuracy estimate to eliminate or deweight poor quality data. Most DGPS errors come from 3 sources: driving error (the difference between the driver's path and the center of the lane), satellite errors (few satellites or poor geometry), and differential errors (corrections too old or base station too far). Studies have shown that driving error is typically 10-30 centimeters. We can estimate satellite errors with the dilution of precision measure, available from the receiver. We also receive differential age from the receiver, and we can look up the location base station.

We propose to estimate the error of a single position as a function of these measurable factors. We base our estimate on ground truth data that we have recorded at our lab. Using a carrier-phase receiver synchronized with the same model DGPS receiver as was used in the data collection, we have the actual error of the position to within a few centimeters. Figure 3 shows the position error versus time for one of our data runs. Finally, we correlated the error with the available measurements, to see which measurement is most predictive of the actual error. Table 1 summarizes these results. Based on this study,

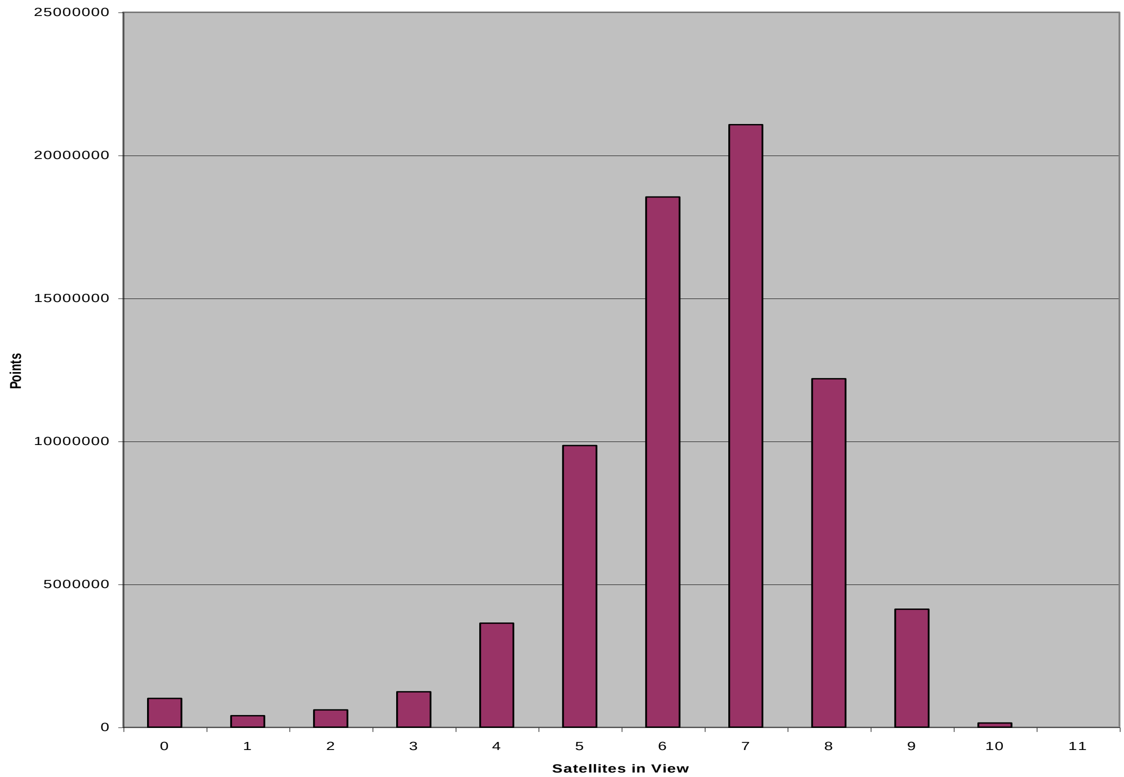

Figure 1. Satellite Visibility

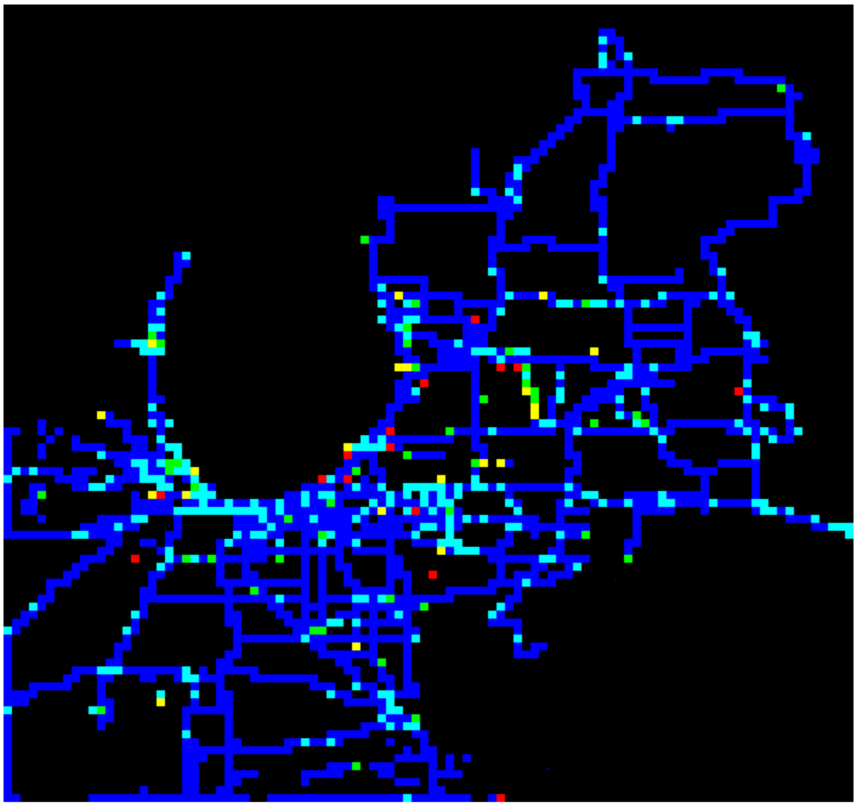

Figure 2. Differential availability. Blue areas are 90% or better availability, light blue 80%, green 70%, yellow 50%, red less than 50%. Black areas are not sampled.

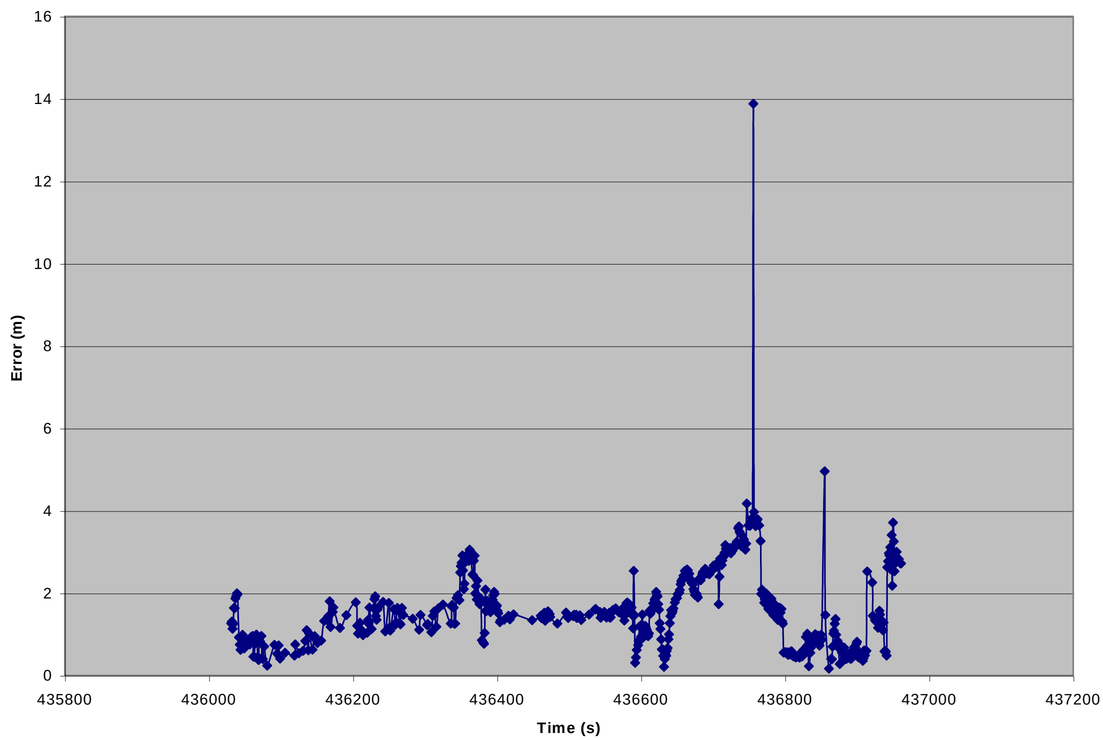

Figure 3. DGPS error, in comparison with a carrier-phase receiver.

Data Set	**Differential Range (km)**	**DOP-error correlation**
1	25	0.4682
2	150	0.3722
3	30	0.2178
4	150	-0.0295

Table 1. Error correlation. Since the correlation does not appear to systematically vary with differential range, DOP alone is taken to be the best predictor of true error, with a mean correlation of 0.27.

we found that the horizontal dilution of precision is a usable error estimate, so further processing weights data on this error.

Map Matching

The first stage in our processing is to separate the traces into map segment traversals, according to our baseline digital map. The baseline digital map is a commercial product that represents some of the roads in an area. The segments are the pieces of road between two intersections, or an on- or off-ramp on a highway. All of our later processing is based on collecting all portions of traces that traverse the same segment, so this is a crucial step. We only attempted to refine the road segments included in the digital map, so we did not cover some rural roads. We used the digital map developed by Navigation Technologies with region code DCA5. This map covers most of Michigan, Indiana, and northern Illinois, and part of Ohio and Wisconsin. About 1000 of the largest cities are covered in full detail, while the rest of the region contains just interstates and major roads.

The map matching process takes an entire trace and finds the sequence of segments that minimizes the distance between the trace and the sequence, using a Dijkstra shortest path-style algorithm. The map matcher produces a table of segment traversals, each containing a segment identifier, the time of entry, the duration of the traversal, the mean distance from the map (which itself has an error of up to 15 meters from

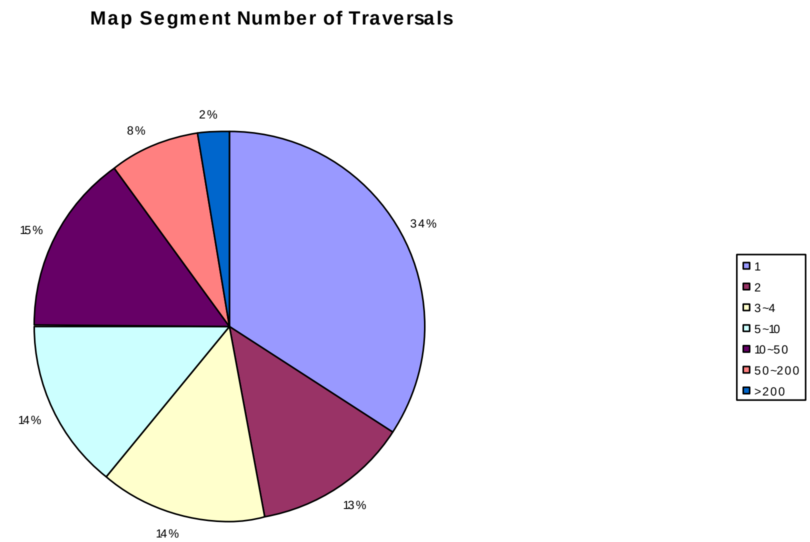

Figure 4. Number of traversals**.**

the true road centerline), and some general segment attributes such as road class, road name, and estimated transit time. The map matcher is not perfect, because it is dependent on the accuracy of the GPS data, the accuracy of the baseline digital map, and the assumption that the vehicle is on a segment in the map. 87% of the data matched some segment, for a total of 567,000 segment traversals. Some segments were visited much more often than others, as the pie chart in Figure 4 shows. Figure 5 shows the distribution of errors in map matching. Errors of more than 20 meters probably indicate a mapmatching error and those traversals were not used.

A byproduct of this processing step is some insights into the fleet's travel patterns. The main transit routes are evident in the color-coded route map in Figure 6. Also, it is interesting to analyze the types of roads normally driven. Figure 7 shows the distribution of road classes. It is also possible to refine attributes of the digital map besides geometry. The NavTech transit time estimates are very crude and do not reflect actual driving behavior. With our data, we can evaluate the accuracy of the estimates by comparing them with the actual traversal time. The distribution of the difference between actual and estimated transit times is in Figure 8. The agreement is generally good, but there are many more longer actual times (70%) than shorter (30%), possibly because the estimate is calibrated for passenger cars.

Segment Centerlines

The next step in processing involves creating a new, more accurate, road centerline than the one in the commercial base map. This need not be the geometrical center of the road; the centerline only needs to be parallel to the lanes for later processing. We generate the centerline by fitting a spline curve to the GPS points on the segment. The centerline fit normally functions well, but the endpoints of the segment need to be constrained to be continuous with the adjoining segments. The plot in Figure 9 illustrates the distribution of number of GPS points per meter. Many segments are not very well covered, with only 0.15 data points per meter, or 6.66 meters between points.

The higher the point density is, the higher we expect the map accuracy. However, some GPS points are more accurate than others. The spline fitting algorithm weights points by their inverse horizontal dilution of precision. Figure 10 shows the distribution of total weight per meter. Only 10% of the segments have a weight of more than 3.5. We estimate that this weight is the minimum for highly accurate maps.

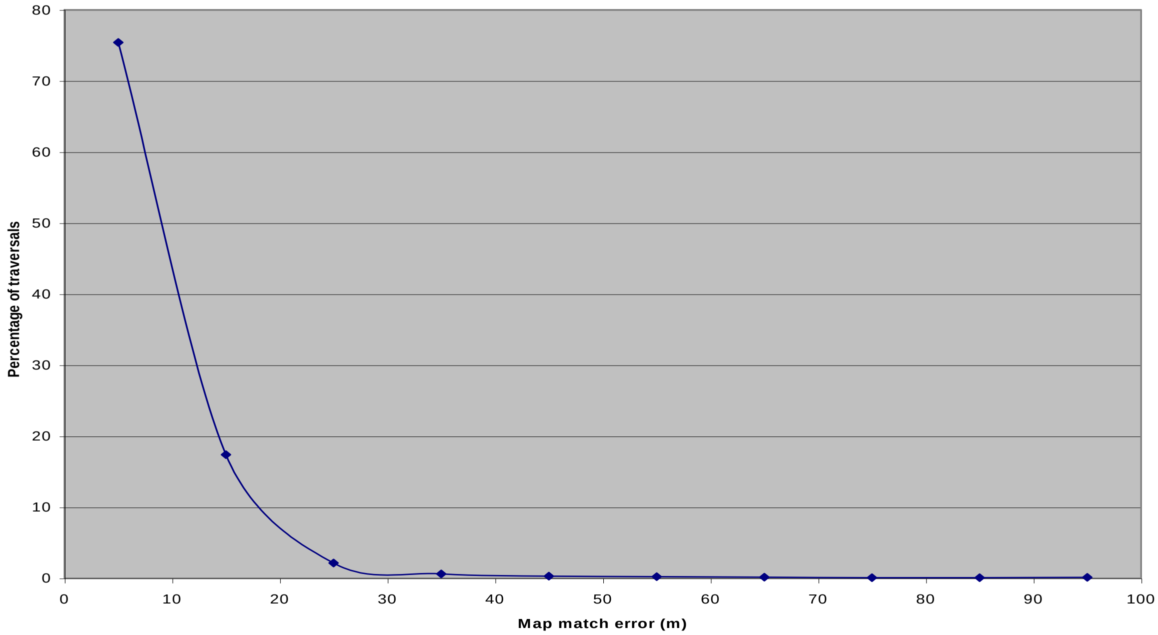

Figure 5. Map match error.

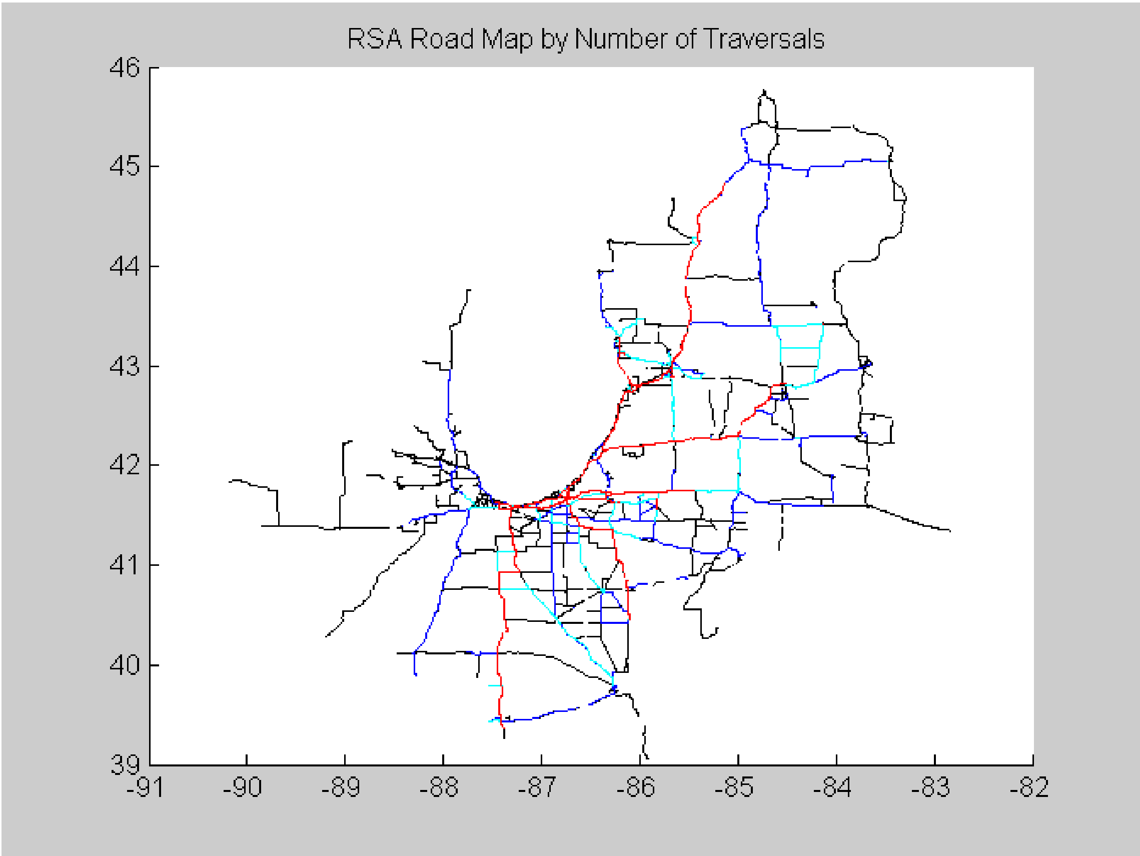

Figure 6. Coverage map. black 0-5 passes, blue 5-15, cyan 15-50, red > 50

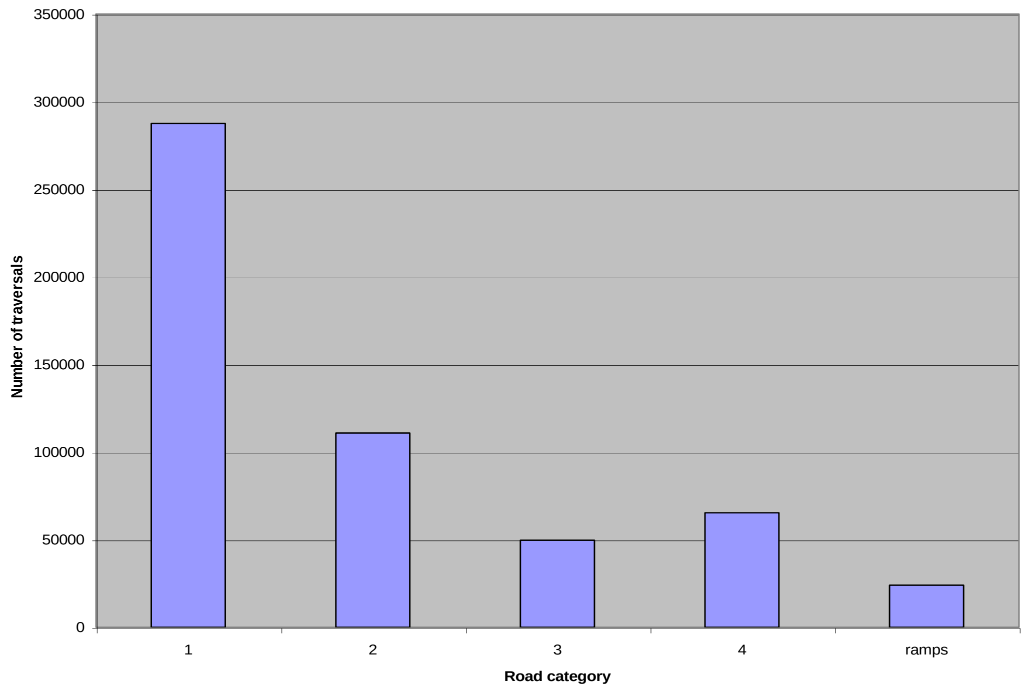

Figure 7. Road categories. 1=interstate, 2=highway, 3=major road, 4=local road

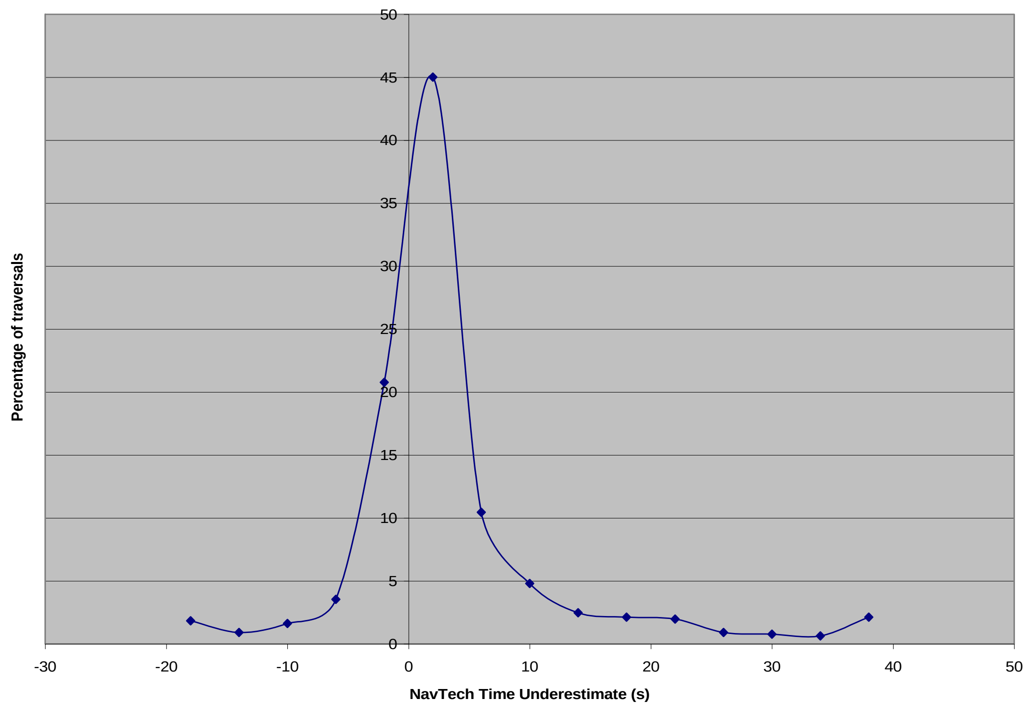

Figure 8. NavTech transit time estimate errors. About 70% of the traversals were longer than the NavTech estimate.

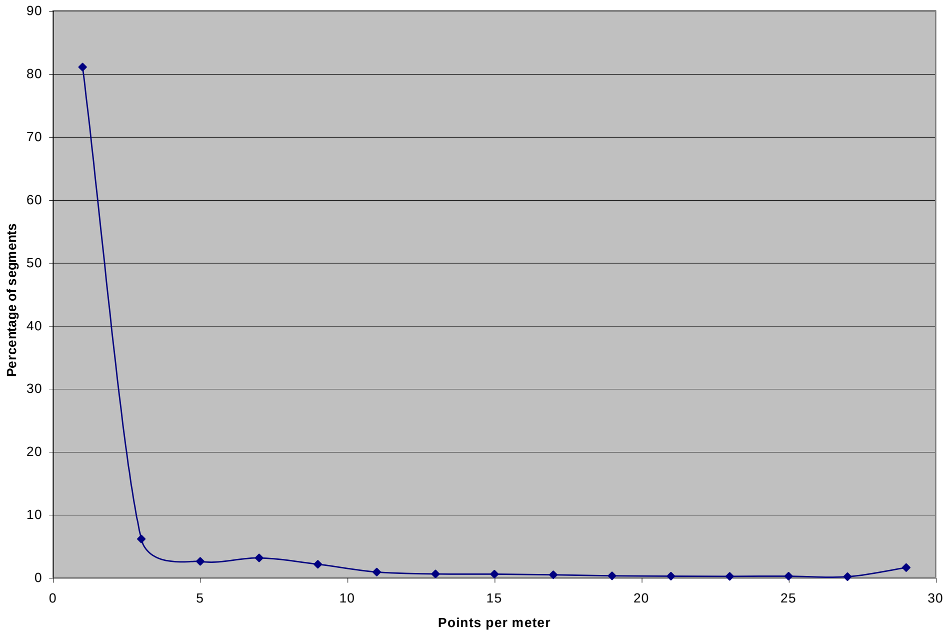

Figure 9. Point density of the segments.

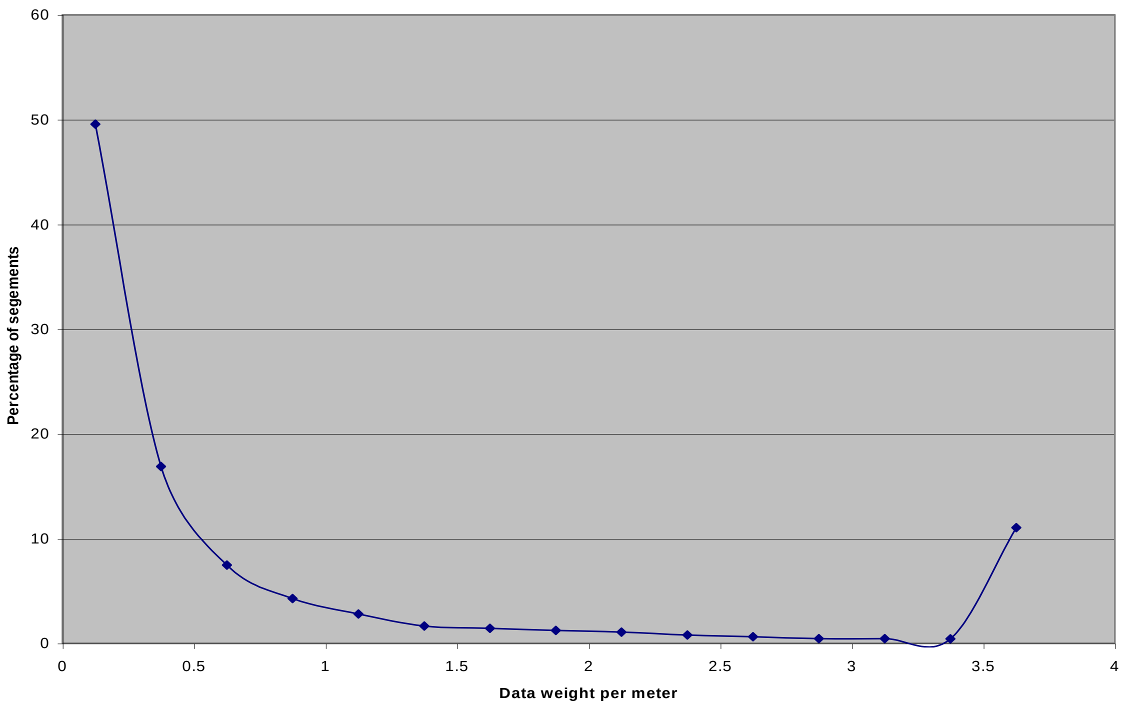

Figure 10. Data weight density.

To characterize the accuracy of points on the map we use bootstrap, as described in Part II.

It is impossible to directly determine the width of the road, but we assume that the distribution of the offset of vehicles from the road centerline should tail off at about +/- 8 meters for a 4-lane highway (of course, the truck fleet may not sample all the lanes, making some of them essentially invisible to us), and the standard deviation of the offsets should be about 4 meters. The distribution of actual standard deviations is graphed in Figure 11. Here we see that much of the data is within 1 meter of the road centerline, indicating that only one lane has been sampled.

Since each segment centerline is computed independently of its neighbors, we constrain the endpoints to match one of the connecting segments (matching all of the connecting segments would force a distortion in the shape. Future work will introduce short "connecting paths" to continuously connect segments.). About 60% of segments align perfectly with their neighbors, but there are instances of segment overlap or gaps between some segments. Figure 12 graphs the distribution of longitudinal gaps between neighboring segments. The gap is generally small but there are a few large overlaps. Figure 13 graphs the distribution of lateral gaps (misalignments) between segments. Again, the gap is generally zero but there are a few exceptions.

Sections of road with high curvature are most dangerous for truck rollovers. The curvature of the spline at a particular point is a function of the derivatives of the spline. Its geometrical interpretation is that, for a point with curvature κ, the curve follows a circle with radius $1/\kappa$ at that point. For highways, curvature of more than 0.001 for right turns and –0.001 for left turns is dangerous. Figure 14 shows that most roads in our data set are straight, while about 20% have a high left or right curvature.

Lanes

As the eventual aim is to find the exact curvature of the truck's current lane, the next step is to find the lane centerlines, informally defined as the invisible line which drivers in a lane are trying to follow. If the road centerline is parallel to all the lane centerlines, the lane centerlines are a constant offset from the road

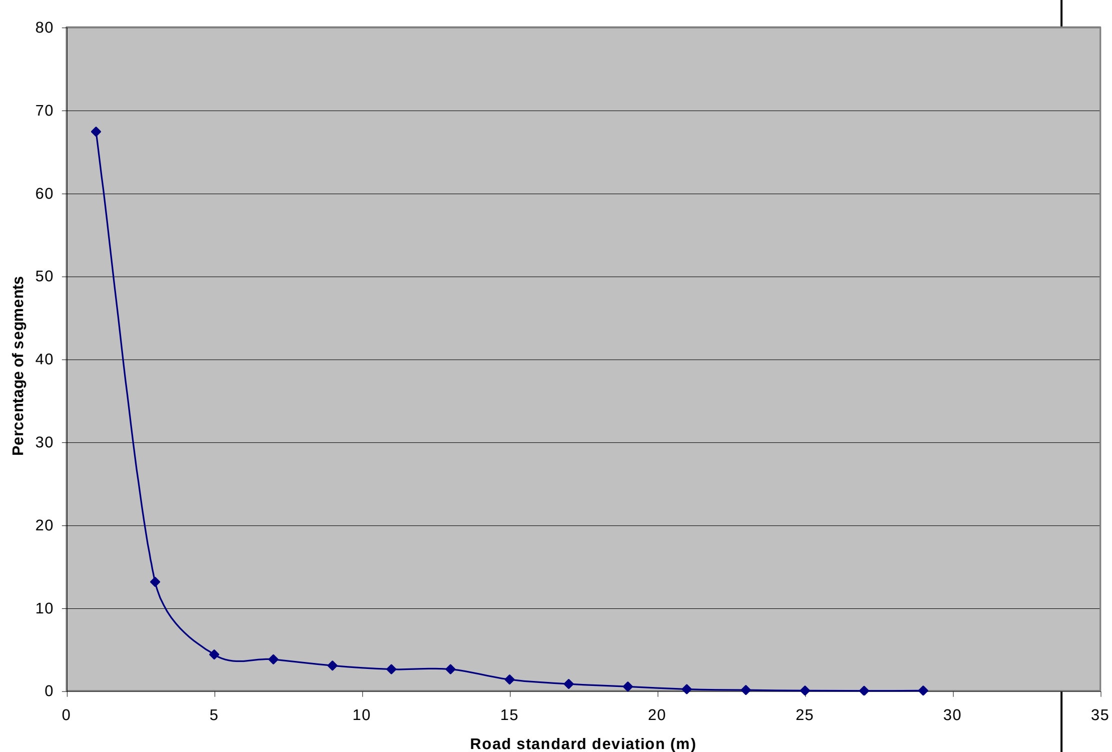

Figure 11. Distribution of road widths

centerline. If drivers are following a lane most of the time, most of the data points should cluster into these lane offsets. We have implemented a clustering technique to find the centers of these high-density regions that define the lane. To allow for lane merges and splits within the lane, we divide each segment into windows and assume the lane structure is constant within each window.

From the road width data in Figure 11, we expect most segments to only have one lane. Figure 15 displays the distribution of the number of lanes detected in each window. Although it is not as predominately single-lane as hypothesized, all segments seem to have a reasonable number of lanes. However, segments rarely change lane structure more than once or twice in the real world. But Figure 16 shows that 18% of the segments change lane structure at least 3 times. This indicates either a problem in the algorithms, or a lack of data in all the lanes for some lane windows.

As Figure 17 illustrates, further investigation into the number of points in each lane window shows that almost 30% of lane windows have less than 20 points. Since the lane finding algorithm discards lanes with too few representative points, this is likely the cause of rapid structure changes in some segments. We can repair this problem by “borrowing” evidence for a lane from neighboring windows, instead of processing each window in isolation.

Depending on the accuracy of the lane, the GPS and the driver’s lane following accuracy, the standard deviation of the distance to the lane center should be less than 1 meter. According to Figure 18, the standard deviation peaks at about 1.4 meters, probably due to GPS noise. However, there are some outliers up to 36 meters from the lane centerline that should be investigated. It is interesting to contrast the overall standard deviation with the standard deviation for each pass. Since we have observed that GPS error is slowly varying, each pass should be fairly internally consistent: the standard deviation of the offsets should be lower. In fact, if the trace is travelling parallel to the lane/road centerline, the standard deviation should be 0. Figure 19 indicates that this is the case, namely that most passes are travelling parallel to the centerline, and the overall standard deviation comes from differences amongst the traces.

If the error sources fit a normal distribution, the total offset from the lane centerline should be a gaussian. Figure 20 shows that many of the lane windows have a good Gaussian fit (low gaussian deviation), but several of them are quite poor and need a close look. Finally, the lane width is a good reality check for the lanefinder. We expect most lanes to be 3-4 meters wide, but surprisingly Figure 21 shows that over 15% of the lane windows are 5 meters wide or more.

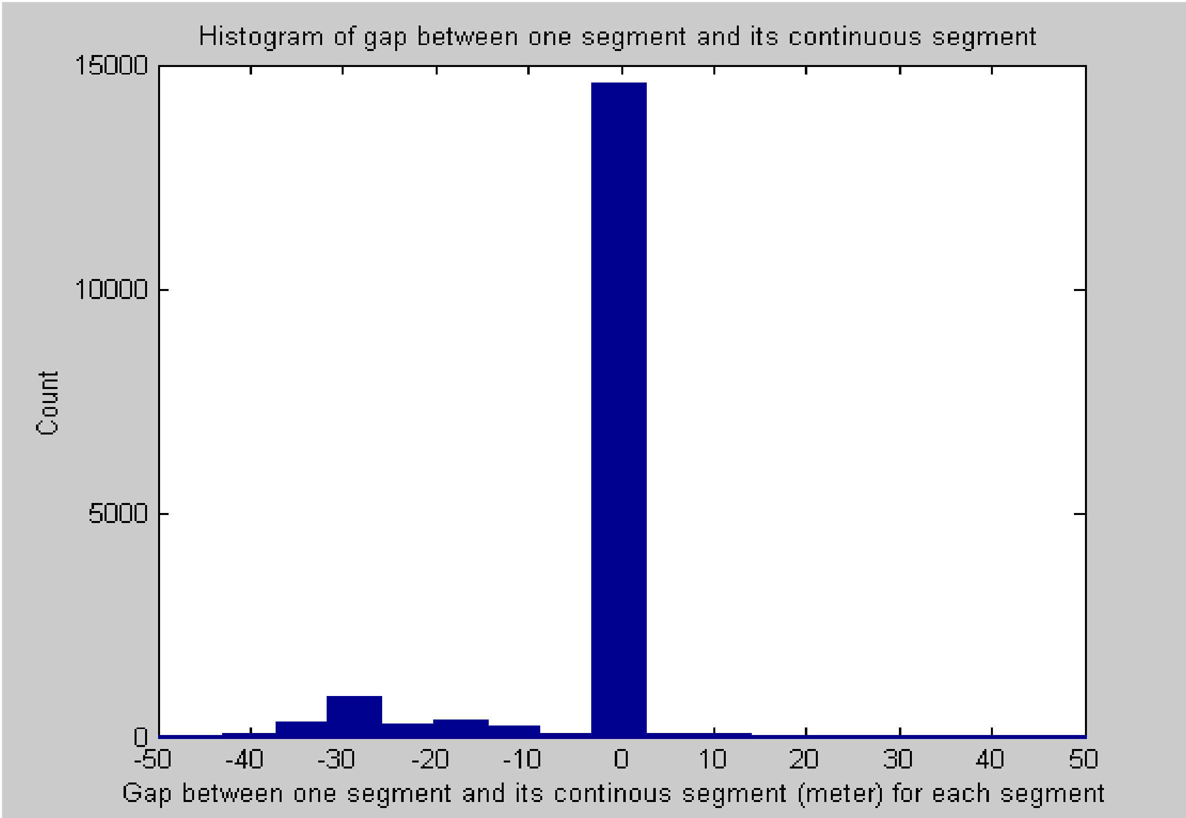

Figure 12. Longitudinal overlap/gap of each segment with its successor segment. Negative values indicate overlap while positive indicate gap.

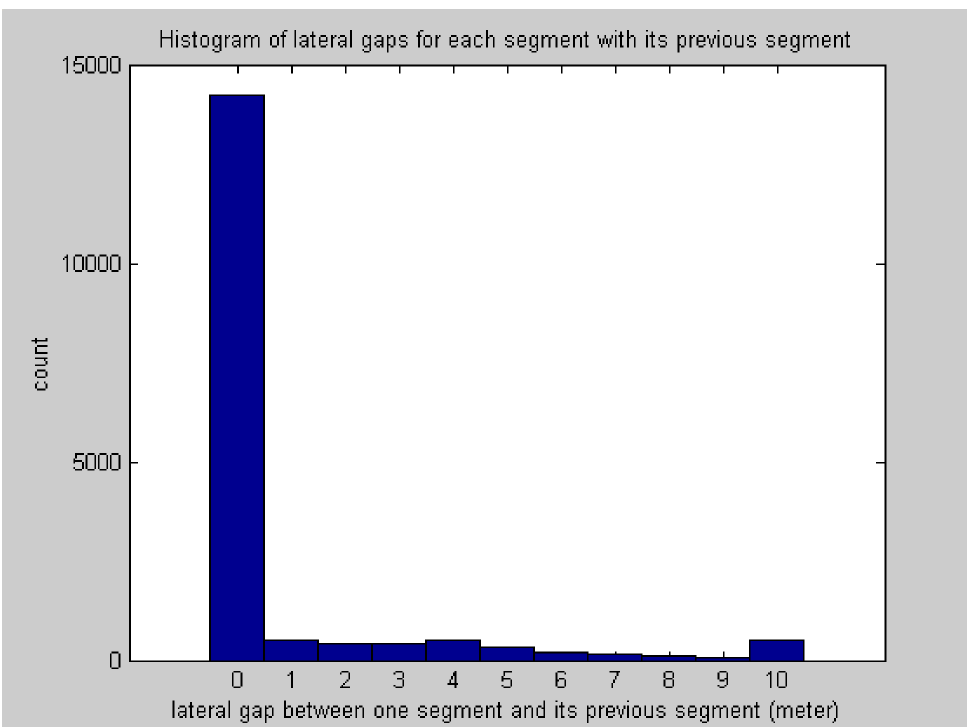

Figure 13. Lateral gap (misalignment).

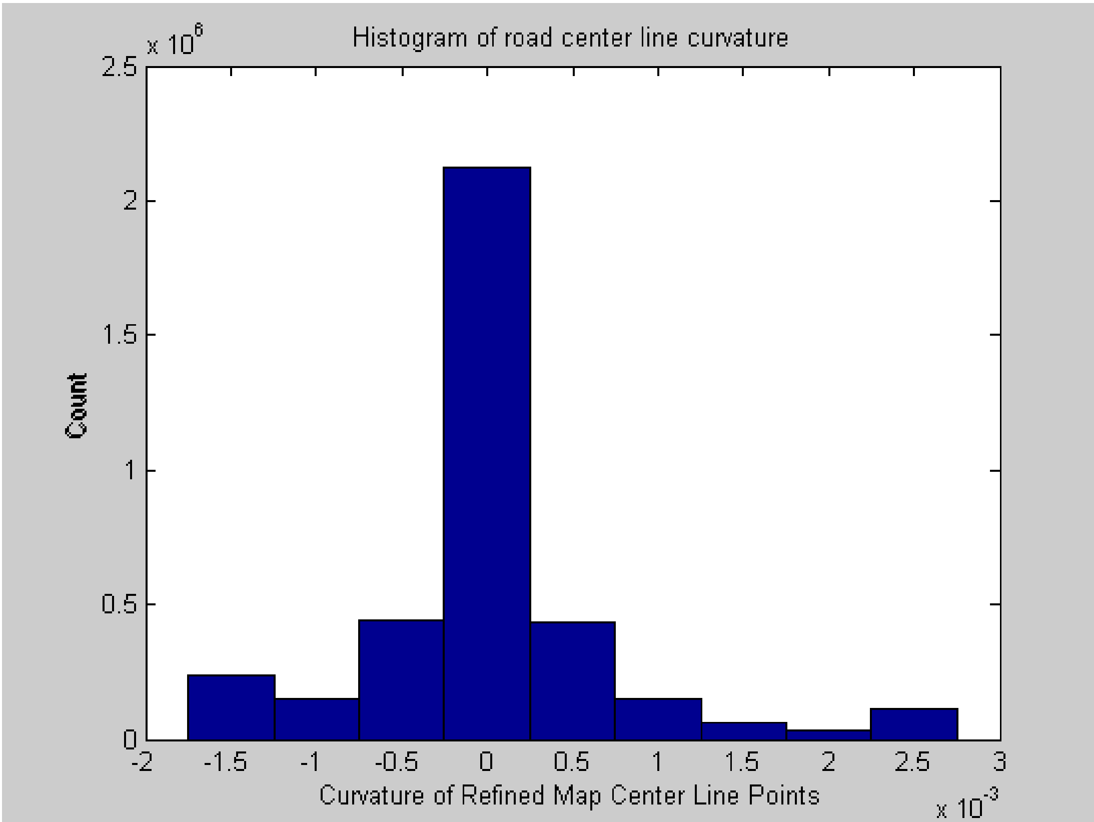

Figure 14. Histogram of road curvature. Most segments are straight (0 curvature), but some are curvy enough to pose a rollover hazard (curvature > ±0.001)

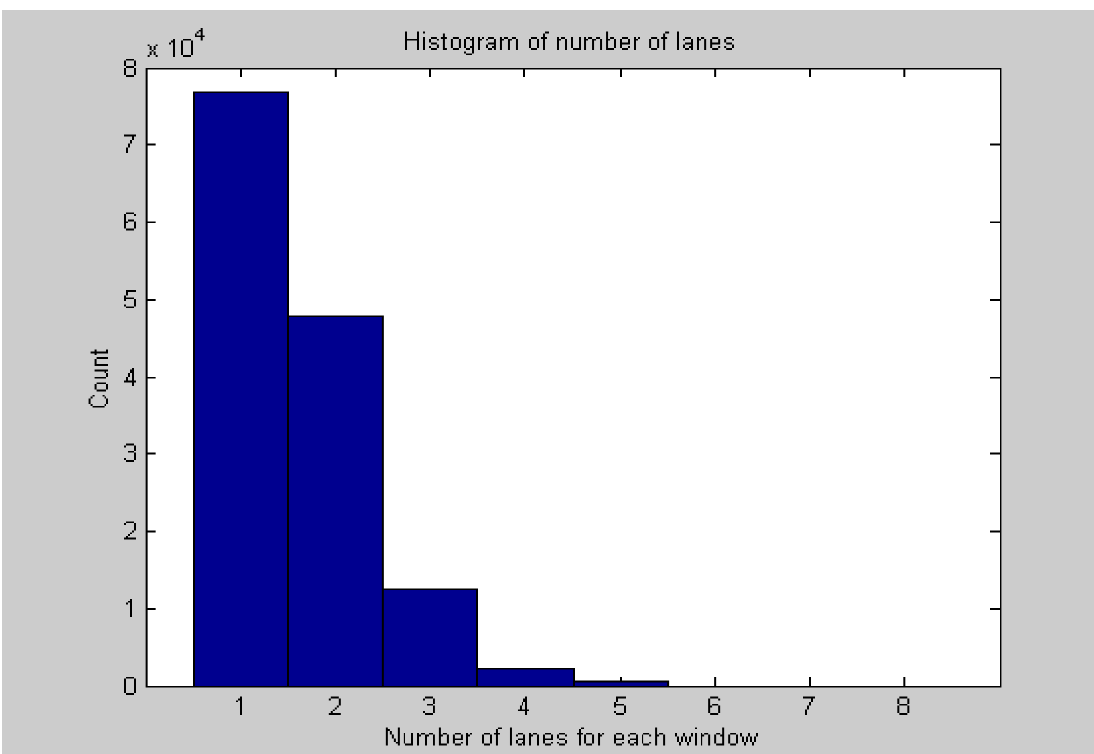

Figure 15. Number of lanes.

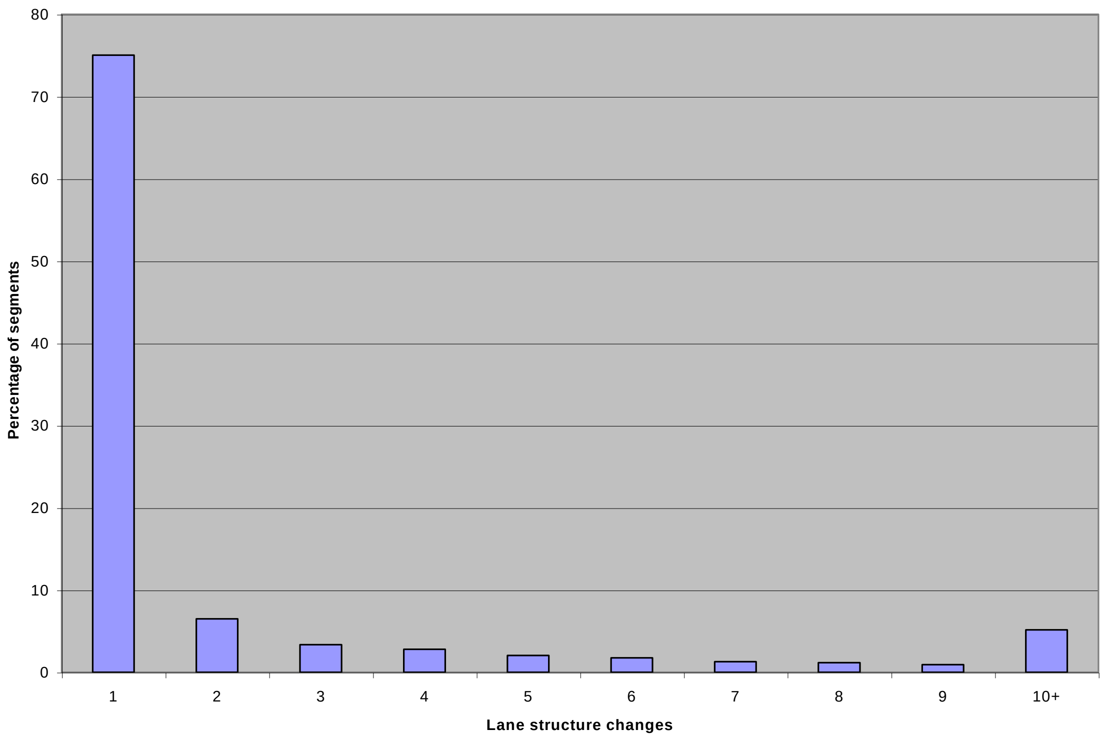

Figure 16. Number of changes in the lane structure over the length of the segment. Normally segments should not have more than 2 changes, so this may indicate a problem with the algorithms.

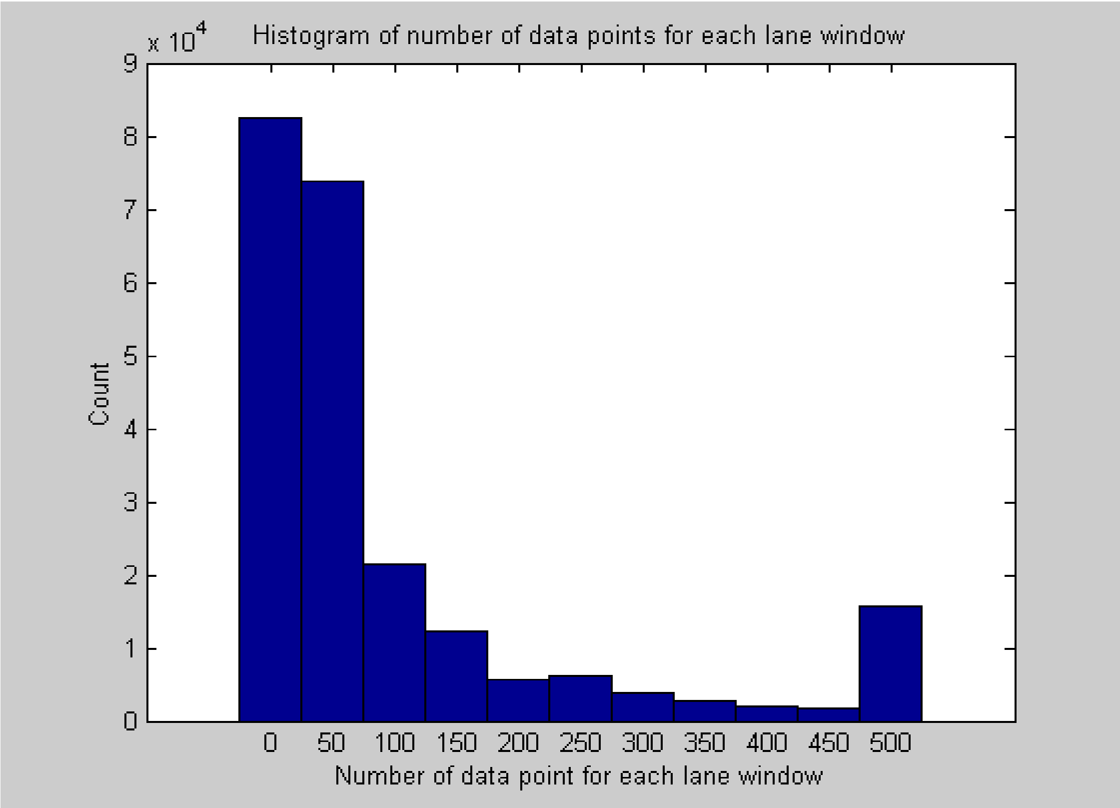

Figure 17. Number of points in each lane in each window. Less than 20 points is probably too little.

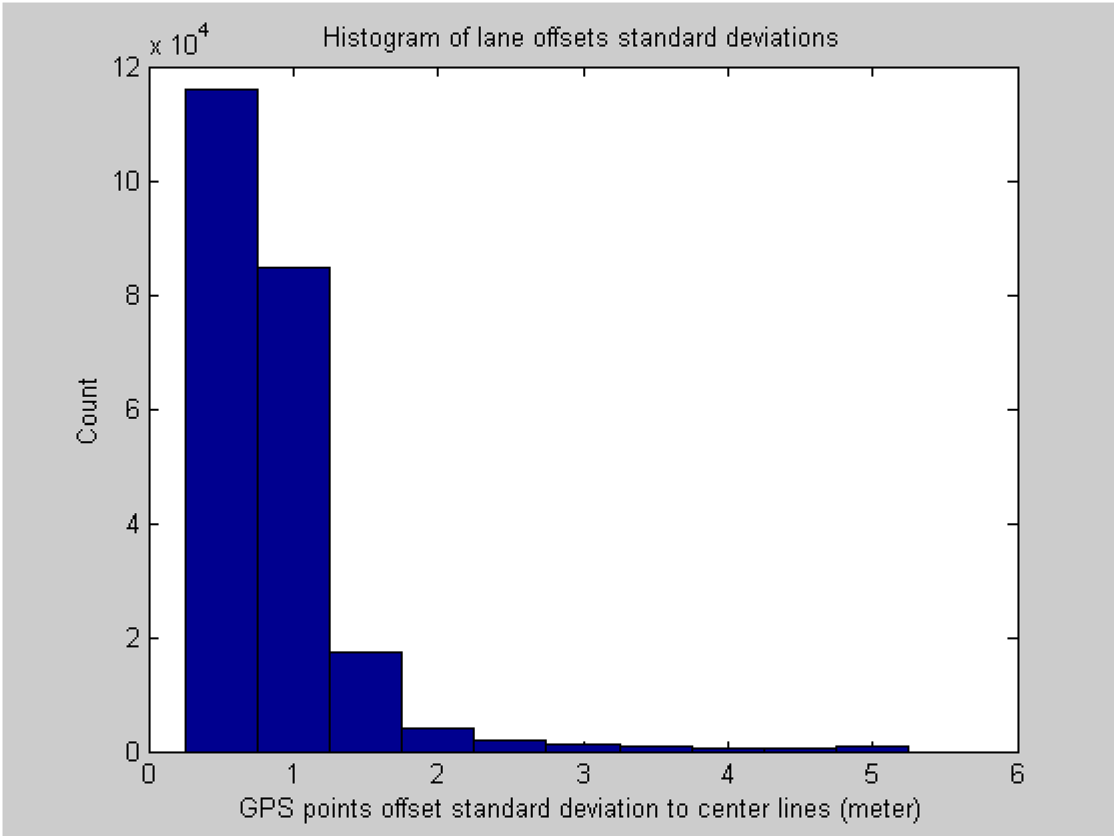

Figure 18. Standard deviation of offsets of the points in a lane from the lane centerline for each window.

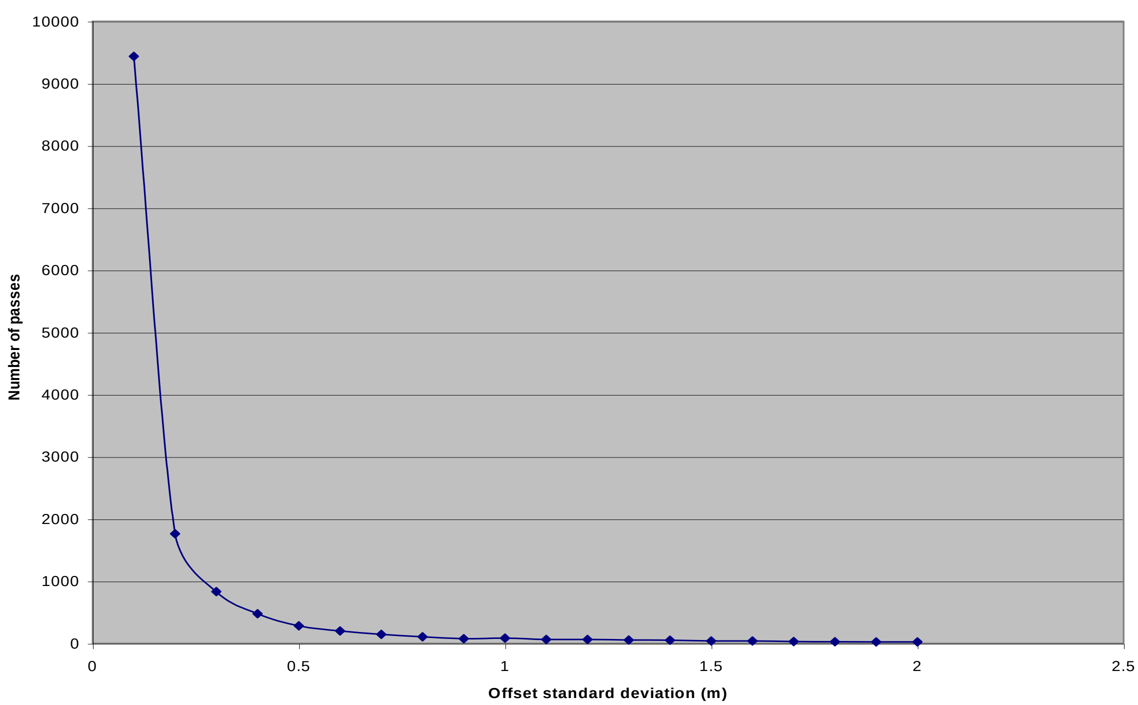

Figure 19. Standard deviation of the offsets for each pass. This score is much lower than the overall standard deviation.

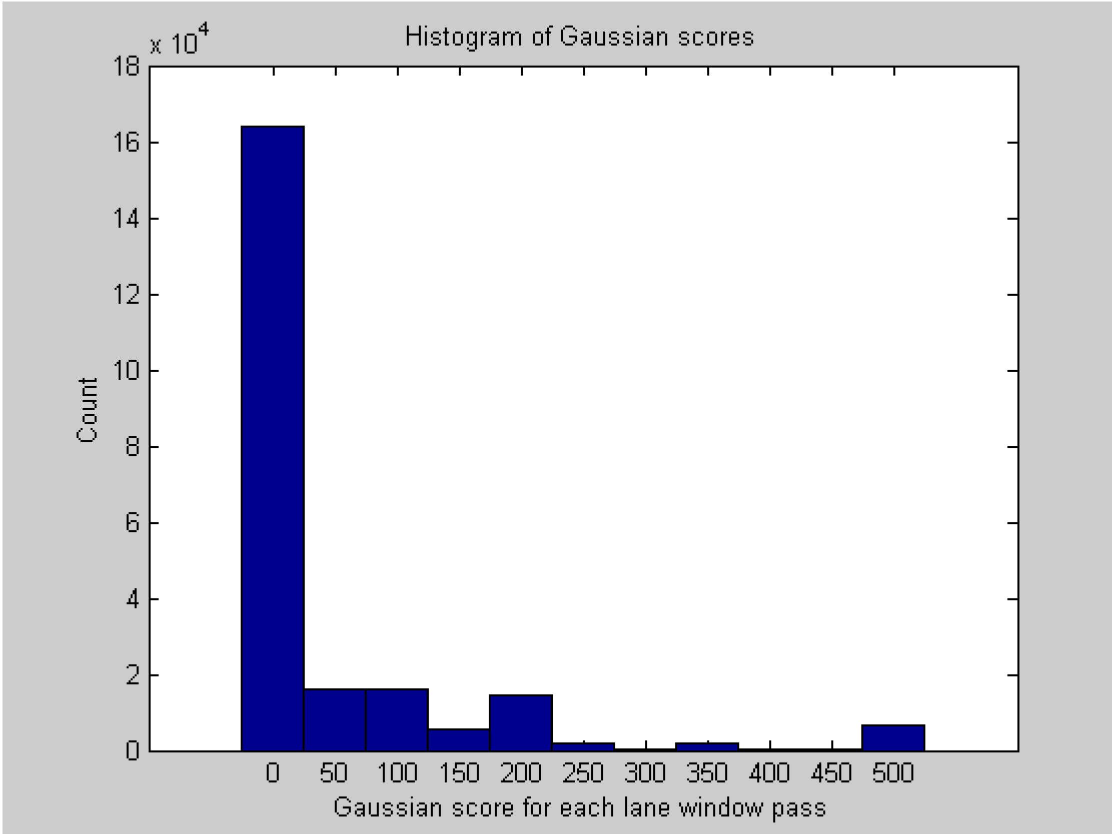

Figure 20. Gaussian deviation score. Lanes whose distribution does not sufficiently resemble a Gaussian (high Gaussian score) may be inaccurate.

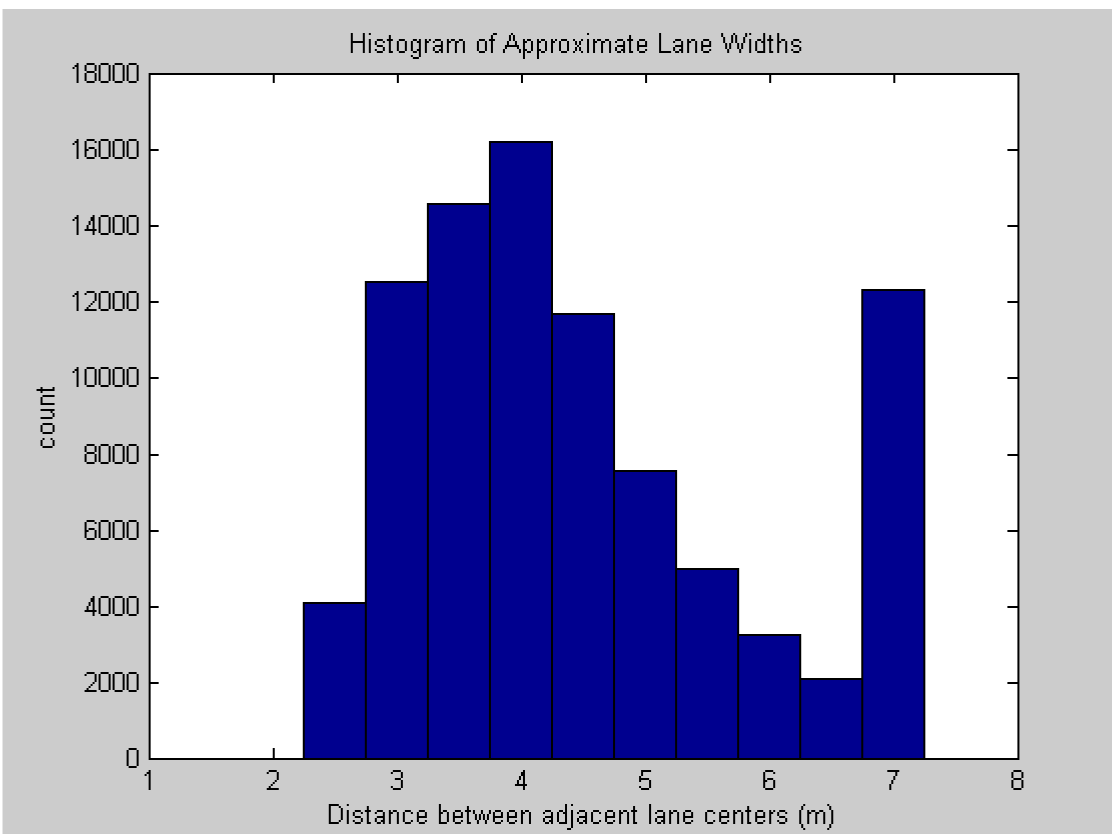

Figure 21. Lane width distribution.

PART II. Detailed Analysis

In this part, we give detailed treatment to particularly dangerous road segments, called "Hotspots."

Hotspot 1

The largest cluster of RSA warnings occurs on a highway onramp near Praxair's headquarters in Indiana. The onramp makes a 270 degree turn from a state highway to an interstate, and many trucks take the final portion of the curve too quickly as they accelerate in preparation for merging onto the highway. Figure 1 shows an aerial view of hotspot 1, with the road centerline in red and a circle approximating the spiral part of the curve in black. This centerline was calculated by fitting a spline to the position data, roughly 226 passes with a total of 19,000 points. Since the segment is roughly 600 meters long, the data density is 32 points per meter, one of the highest in the data set.

The major factors impacting the rollover score are road geometry, driver behavior, and truck parameters. Road geometry parameters include road curvature and superelevation. We calculate curvature from the derivatives of the road centerline spline. The curvature along this hotspot is in Figure 2, with an estimate of the true curvature according to the Indiana Department of Transportation. The curvature is mostly

Figure 1. Hot spot 1. The red line indicates the road centerline, as calculated by a spline fit to all passes on this segment. The black circle is an approximation to the spiral part of the onramp.

accurate, with the exception of some problems near the beginning due to poor positioning data. The superelevation, or bank, of the curve lets the truck drive faster around the curve without increasing its lateral acceleration. We calculate the bank from the measured lateral acceleration, speed, and curvature,

$$E = v^2\kappa/g - f,$$

where v is velocity in m/s, κ is curvature in m^{-1}, g is 9.81 m/s^2, and f is the lateral acceleration in g. Figure 3 shows the bank along the curve, with an estimate of the true bank according to the Indiana Department of Transportation.

Hotspot 2

The second-largest site of RSA warnings occurs on an S-curve. Figure 4 shows an aerial view of the curve with the computed centerline in red and circles for the two curves in yellow. This segment had 151 passes and about 34,000 points. This segment is somewhat longer than hotspot 1, about 930 m in length, giving an average of 36 points per meter. As in hotspot 1, Figures 5 and 6 show the curvature and bank angle, respectively. The results are good in the interior of the segment, but the endpoints are noisy because of low data density and distinct populations of traces entering and leaving the segment from different segments.

Map Accuracy

We directly estimate the accuracy of the centerline for these hotspots using a procedure called bootstrapping. Bootstrap is a computer-based method for assigning measures of accuracy to statistical estimates. It is particularly useful where standard statistical theory is useless because the situation is too complicated to be mathematically tractable or too ill understood to make justifiable assumptions about the nature of the stochastic phenomenon to arrive at a reasonable model.

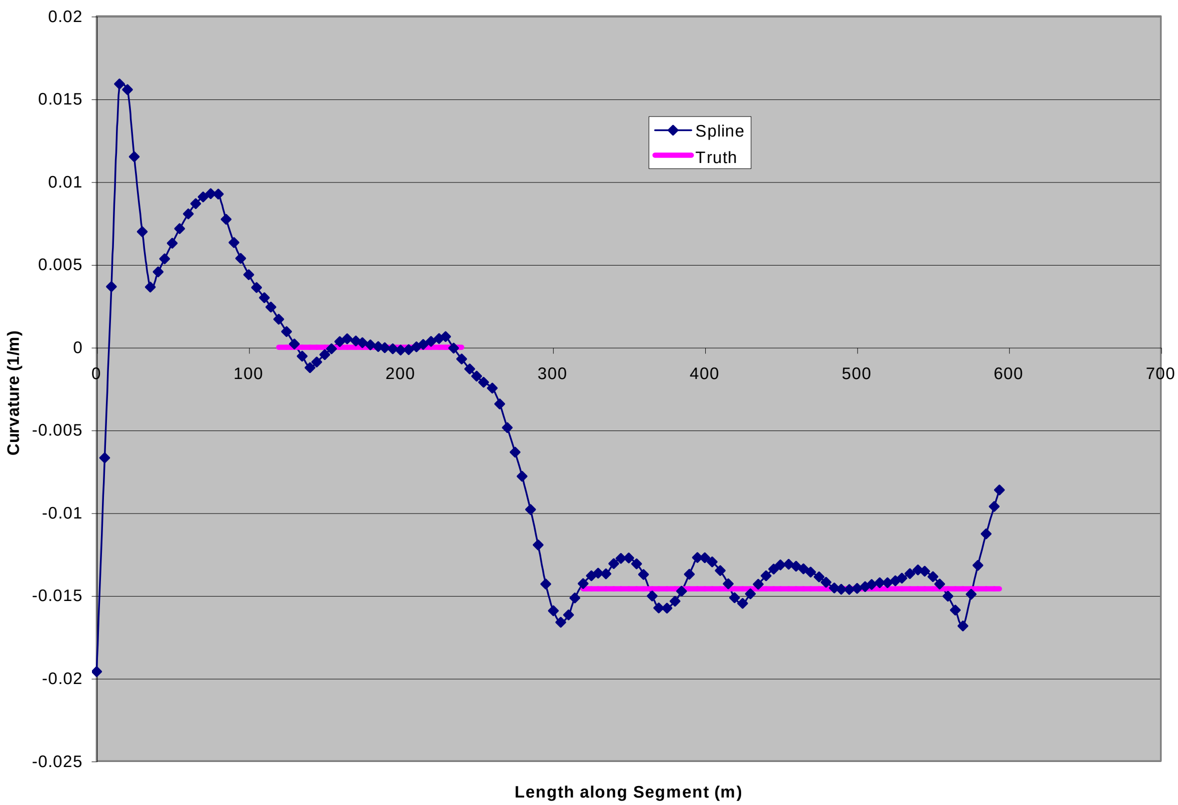

Figure 2. Hot spot 1. The blue line is the curvature computed from the spline fit. The magenta line is the ground truth according to the Indiana Department of Transportation.

Table 1. Map accuracy, in meters, for different areas and data volumes.

	10 traces	**200 traces**
Hotspot 1	0.041m	0.011 m
Hotspot 2	0.219 m	0.015 m

The latter is the case in our situation since the random nature of the sources of GPS and driving errors is not well understood, and it seems incorrect to make simplifying assumptions such as “error in each GPS point is an independent Gaussian random variable with zero mean”. It is best then to use a non-parametric statistical technique to attach a measure of accuracy to the map points. Bootstrap is the one used because of its simplicity, its universal applicability, and its reliable behavior in situations where its results can be compared with those from standard techniques.

The idea behind bootstrap is simple: in absence of any assumptions all we can know about the distribution of the population is present in the distribution of the data. So take the “empirical distribution” in place of the original distribution whatever that might be, and apply the usual statistical procedure, i.e., sample the data with replacement to create new data sets, compute the desired statistic for each of these, and look at the distribution of the statistic and compute its desired moments.

In our case the statistic of interest is the fitted spline. So we take the original data set of GPS points, call it $\boldsymbol{x}$, and create data sets $\boldsymbol{x^1, x^2, \ldots, x^b}$ by randomly choosing data points with replacement from $\boldsymbol{x}$. Each new data set is of the same size as the original. We fit splines to each of $\boldsymbol{x^1, x^2, \ldots, x^b}$. The collection of these splines reflects the distribution of the “spline” statistic. We can now calculate any measure of accuracy we choose for this statistic. We choose to take points on the splines at regular intervals and calculate the standard errors of these points. For example we take the points on all splines at parameter value 0.5; these are the mid-points of the splines (according to arc-length). We calculate the standard error of these points from the standard formula for standard error. Bootstrap theory guarantees that this standard error is close to the actual standard deviation of the spline mid-points, and gets closer as the number of bootstrap samples, $\boldsymbol{b}$, is increased.

Table 1 shows the map accuracy for Hotspots 1 and 2 using the bootstrap method with the number of samples $\boldsymbol{b}$ set to 200. We did not perform bootstrapping on the entire database because of its computational complexity. In addition to calculating the accuracy for the all traces, we also evaluated the accuracy with a partial data set of only 10 traces, to see how much accuracy is gained with more data. To arrive at a single number for each condition, we calculated the standard distribution of the error distribution for all points along the centerline and took the mean. Both hotspots are very accurate with complete data, but Hotspot 2 is significantly lower quality with only 10 traces. The 10 traces for Hotspot 2 are probably low-quality, illustrating that making maps from higher volumes of data reduces uncertainty over the final map quality, as well as improving the overall map quality. The practical effect of low quality maps on rollover warnings is evaluated in Task 20, Theoretical Rollover Warning Effectiveness.

Conclusion

This report has described techniques and results for creating precision maps of roadways from uncoordinated data collection vehicles. Precision maps are required for many advanced driver assistance systems, in order to provide detailed insight on current and upcoming situations. The upcoming curvature is particularly important for rollover warning, as detailed in the report for Task 20, Theoretical Rollover Warning Effectiveness.

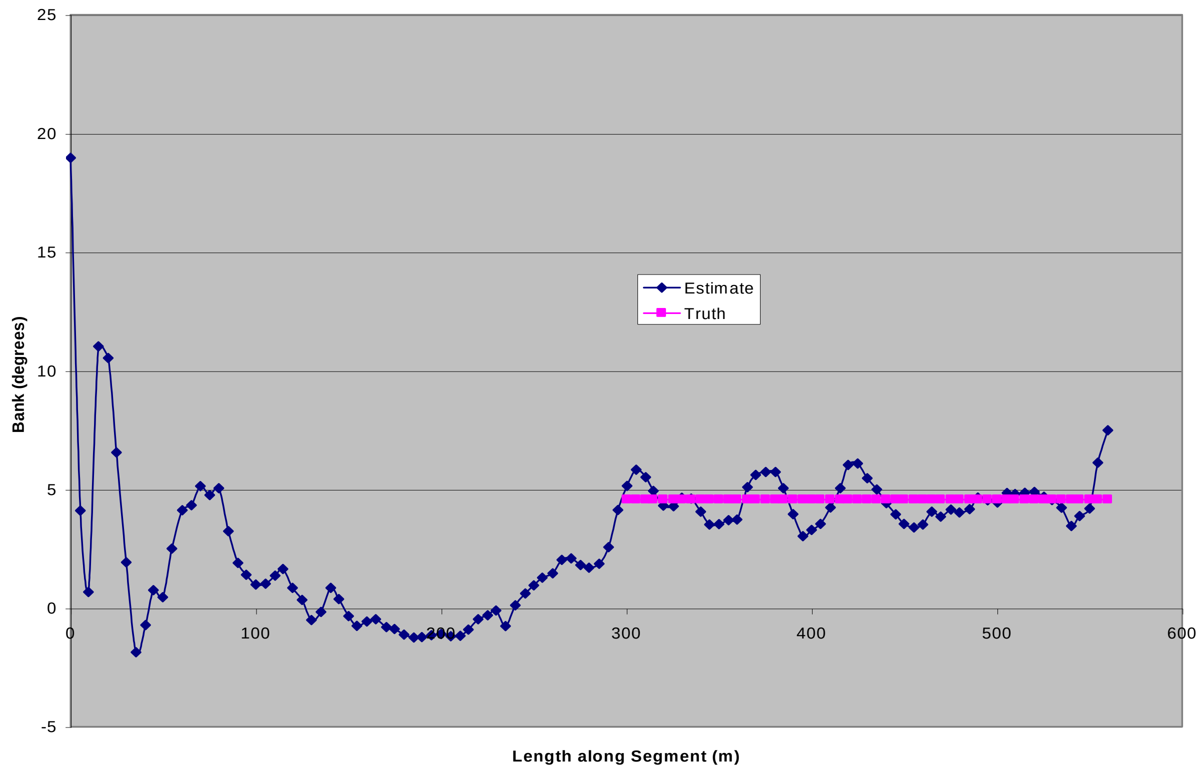

Figure 3. Hot spot 1. The blue line is the bank estimate computed from the formula. The magenta line is the ground truth according to the Indiana Department of Transportation.

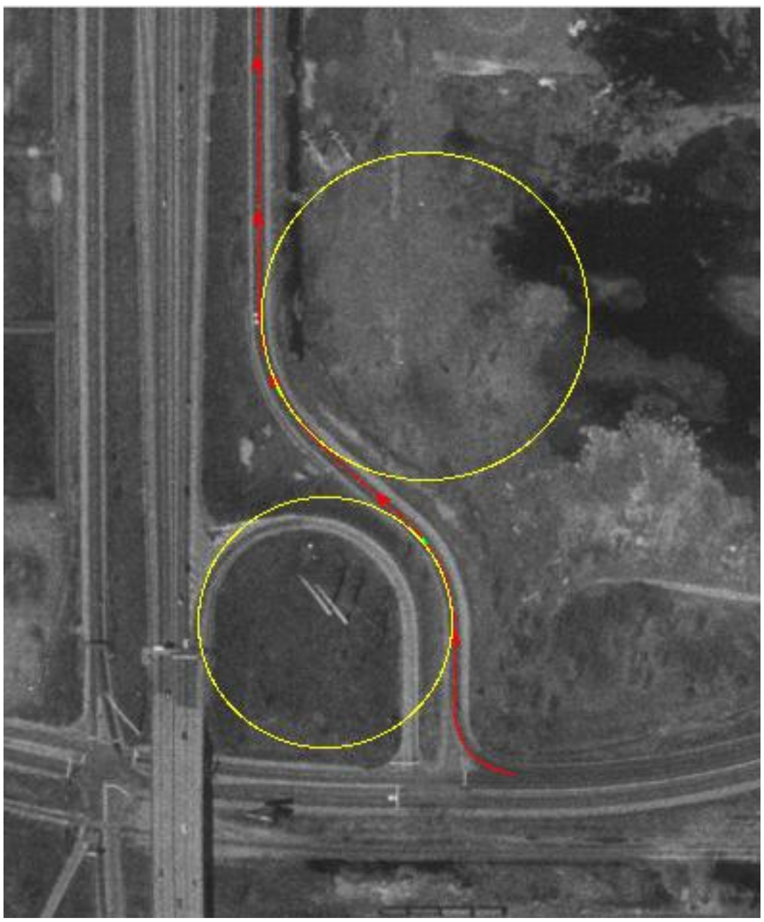

Figure 4. Hot spot 2. The red line indicates the road centerline, as calculated by a spline fit to all passes on this segment. The yellow circles are an approximation to the spiral parts of the onramp.

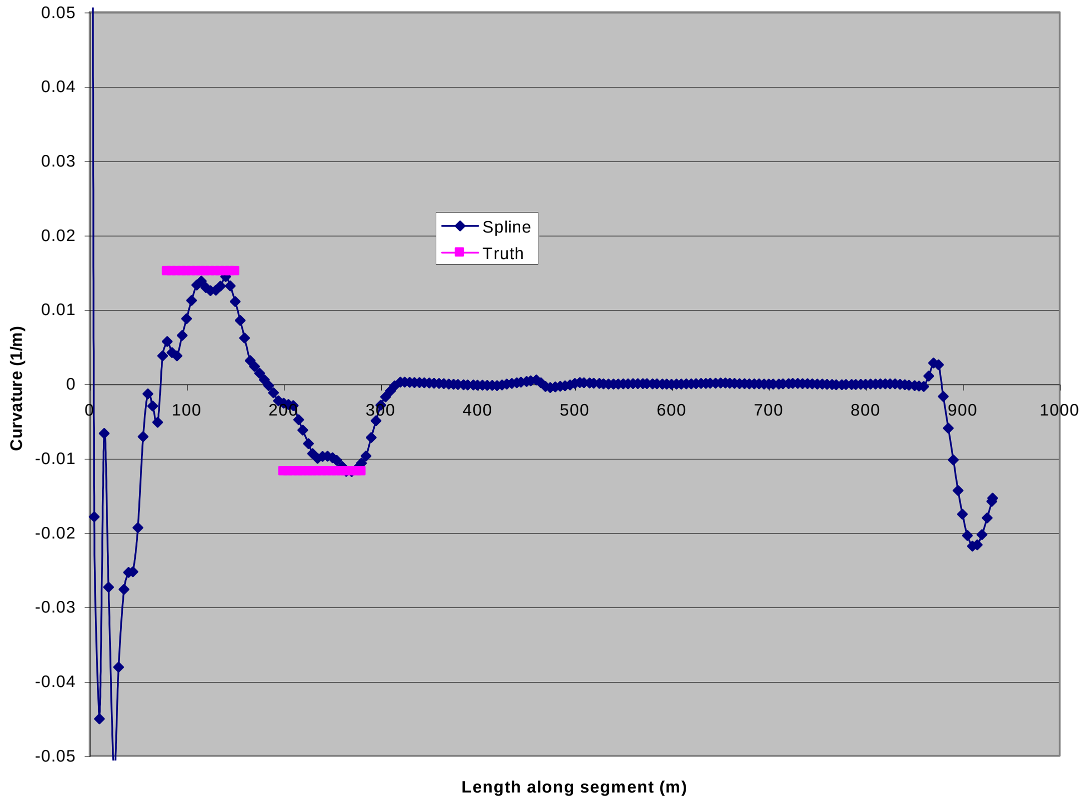

Figure 5. Hot spot 2. The blue line is the curvature computed from the spline fit. The magenta line is the ground truth according to the circular fits. The curvature shows some noise in the beginning and ending because trucks enter and leave the segment from different connecting segments, causing a poor fit.

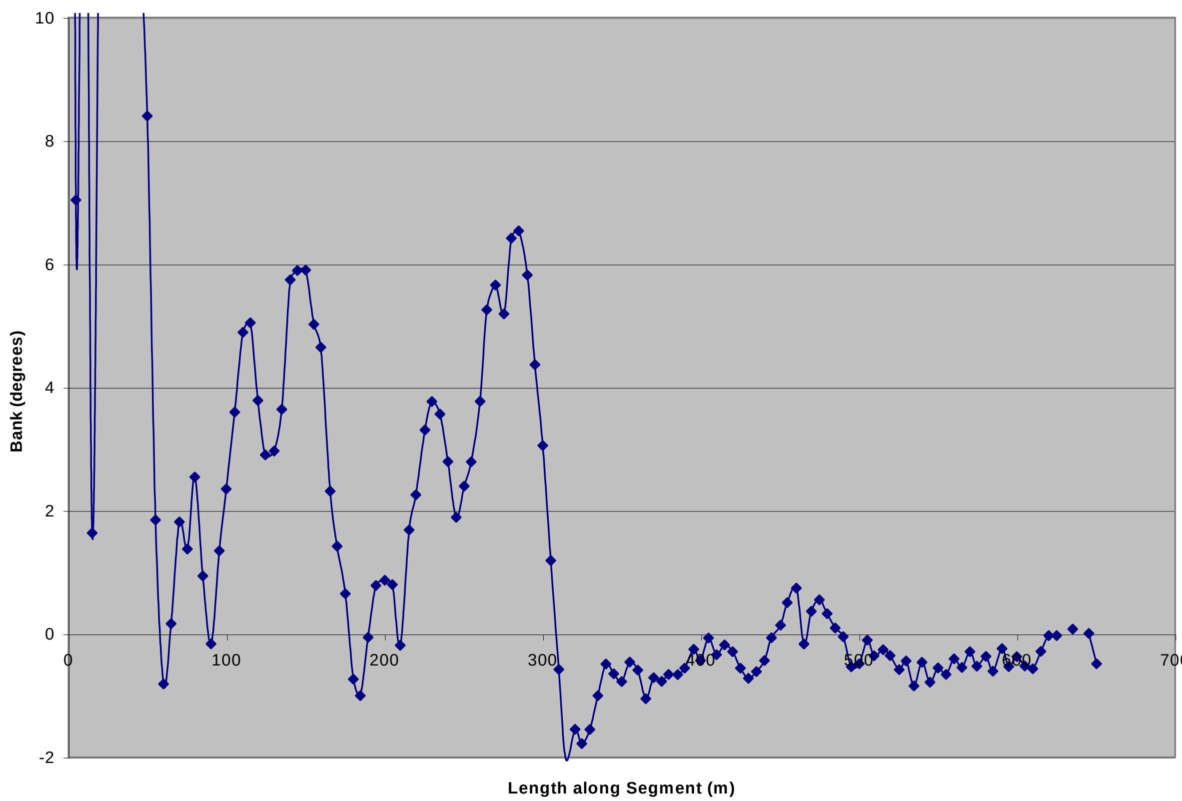

Figure 6. Hot spot 2. The blue line is the bank estimate computed from the formula. Ground truth is unavailable for the banking on this segment.