This report investigates the following characteristics of relative positioning in detail:

1. Achievable relative positioning accuracy using VSC-A RTK SW
2. Assessment of accuracy benefits compared to the alternative method
3. Relative position solution availability with GPS outages

2.2 Test Setup

The test setup was a scaled down version of the full VSC-A test bed as shown in Figure 3. No applications were run in the test setup. Instead, the relative positioning system output was logged and post-mission analysis was conducted. Functions of the VSC-A test bed components Dedicated Short Range Communications (DSRC) Radio, Wireless Message Handler (WMH), and Sensor Data Handler (SDH) were handled using a Wireless Access in Vehicular Environments (WAVE) Radio Module (WRM) and a PC in the test setup. The Relative Positioning Module (VSC-A RTK SW or its non-Graphical User Interface (GUI) version) and the GPS receiver (NovAtel® OEMV®-1) in the test setup were identical to that of the full VSC-A implementation. The generic test setup with the Cooperative Intersection Collision Avoidance System for Violations (CICAS-V) system as an add-on is shown in Figure 4.

All tests included two or more vehicles with the same basic setup that included an OEMV-1 GPS receiver, test PC, and the WRM. A single vehicle with CICAS-V capability was included in the tests, and this vehicle was configured such that a CICAS-V-enabled GPS receiver is used in the test setup. Several vehicles were equipped with alternative GPS receivers, and these were used only in data logging mode. Standard data logging tools that come with these receivers were used for logging, and analysis was conducted only in the post-mission mode.

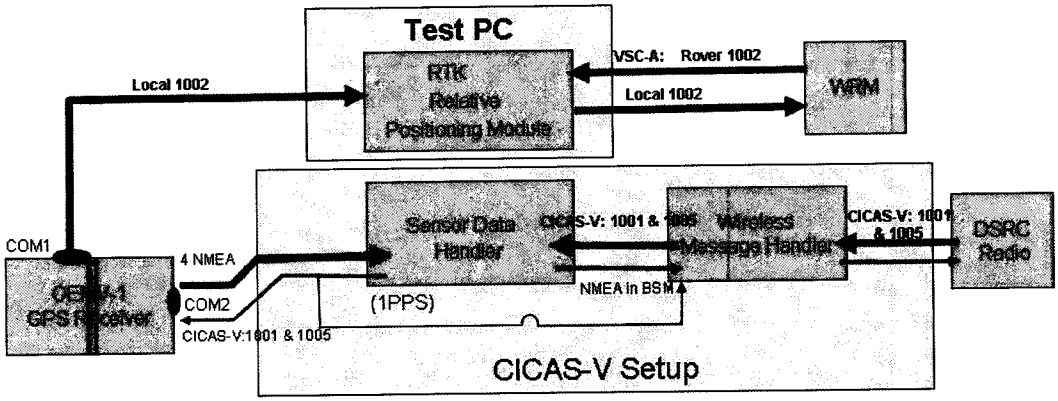

Figure 4: Vehicle Hardware Setup

2.3 Collected Data

The following data was collected from each vehicle:

- VSC-A RTK SW real-time relative position solutions for each vehicle
- Raw GPS data from host vehicle (HV) (RTCMV3.0 and device-specific binary)
- Raw GPS data as received from other vehicles (RTCMV3.0)
- HV position, speed, and heading (device-specific binary)

2.4 Test Scenarios

Test scenarios were designed such that a range of vehicle operating conditions and GPS visibility conditions were covered. The following variables were used to define the scenarios:

- Sky visibility: Open sky, short complete interruptions (i.e., overpasses), and tree cover
- Vehicle speed: slow speed (< 40 mph) and high speed (> 55 mph)
- Roadway: freeway and urban streets
- Local GPS measurement noise: mainly addressed by looking at high and low density traffic

3 Analysis Summary

The analysis is primarily focused on assessing the relative positioning accuracy benefits of the VSC-A relative positioning SW. The SP method and RTK method of relative positioning are compared in the analysis. The impact of short GPS outages on the RTK solution is also investigated.

3.1 Analysis Methodology

Between-vehicle across distance (D_{Across}) was used as the primary analysis variable. The selection was based on the fact that this is one of the most critical relative distance measures for VSC-A SW applications. The ease of measuring this variable for validation was also a factor in the analysis variable selection. The following two methods were used to validate D_{Across} in the tests.

1. Using a post-mission truth solution generated using raw GPS data gathered during the tests. A post-mission precise positioning version of VSC-A RTK SW was used to generate the truth solution as discussed below.
2. The tests were designed and executed such that the D_{Across} measure is always a multiple of lane width except for turns and situations where it is unsafe to do so. For instance, vehicles were always driven in the same lane or adjacent lane formations such that D_{Across} is either approximately zero or approximately a single lane width.

3.2 Data Decoding and Processing

For each scenario analysis, a vehicle was picked as the HV and the others were considered remote vehicles (RVs). Analysis is always done using a local coordinate frame, and the HV position is considered the origin of the local coordinate frame. It is noted that any given vehicle in a test can be considered the host and the others as remotes. The data decoding and processing involved the following steps:

1. Decode RTK output records from Host logs
2. Decode HV heading information
3. Time match (1) and (2) and estimate D_{Across} and D_{Along} using the coordinate transformation shown in Figure 5. It is noted that regardless of the relative positioning method used, GPS-based relative observations will always generate D_{East} and D_{North} with respect to the global coordinate frame. The same coordinate transformation routine will be implemented in the SDH in the full VSC-A implementation.
4. Decode vehicle position records (Latitude and Longitude) from Host and Remote logs
5. Derive D_{East} and D_{North} by first time matching host and remote position record (4) and then converting remote position into local coordinates with respect to the host position at the same time epoch. Note that these coordinate transformation functions are also identical to that implemented in the full VSC-A implementation.
6. Generate reference D_{Across} and D_{Along} using values using post-mission RTK SW
7. Plot D_{Across} and D_{Along} generated using RTK method (3) and SP method (5) along with reference data from (6), if generated

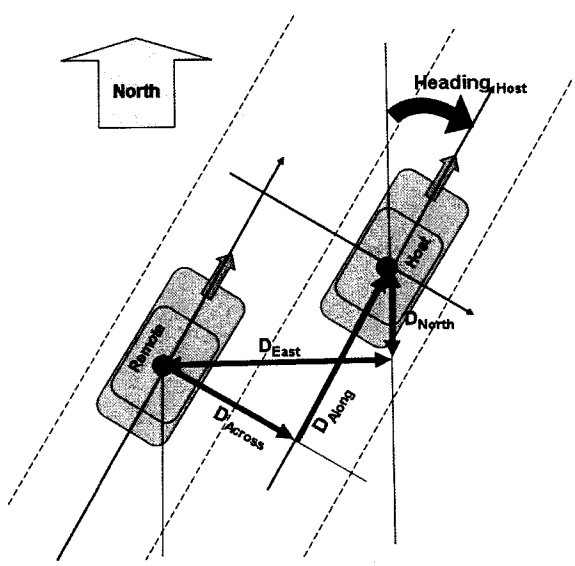

Figure 5: Global to Local Coordinate Transformation

3.3 Generating the Truth Solution

Two methods were used to validate the real-time between vehicle distances generated by the RTK and the same values estimated by the SP method discussed below:

1. Formation of vehicles during the test scenarios: Vehicles were driven in the same lane and in adjacent lanes during all possible times. Lane changes were kept to sharp changes that are short rather than a slow deviation from a lane to the next. Hence, the across distances between vehicles are expected to be either close to zero (same lane) or a multiple of a lane width (adjacent lane) for most of the test durations.
2. Post-mission truth trajectory using post-processing SW: Raw GPS data gathered from all test vehicles were individually post-processed using a SW tool to generate a reference vehicle trajectory solution D_{Across} and D_{Along}.

It is noted that reference solutions for the majority of tests were generated using only L1 GPS. This method does not guarantee a cm-level solution due to the convergence time required for such a solution to converge to cm-level. The accuracy of the reference solution can be indirectly estimated by using the forward-backward solution comparison method in the post-processing tool. In general, forward and backward L1 reference vehicle position solutions were found to agree within 0.4 m whereas the corresponding L1L2 solutions agree within a few cm. An example of the forward-backward solution discrepancy plot is shown in Figure 6. For these tests, only L1 reference solutions were available and may not be presented in the analysis.

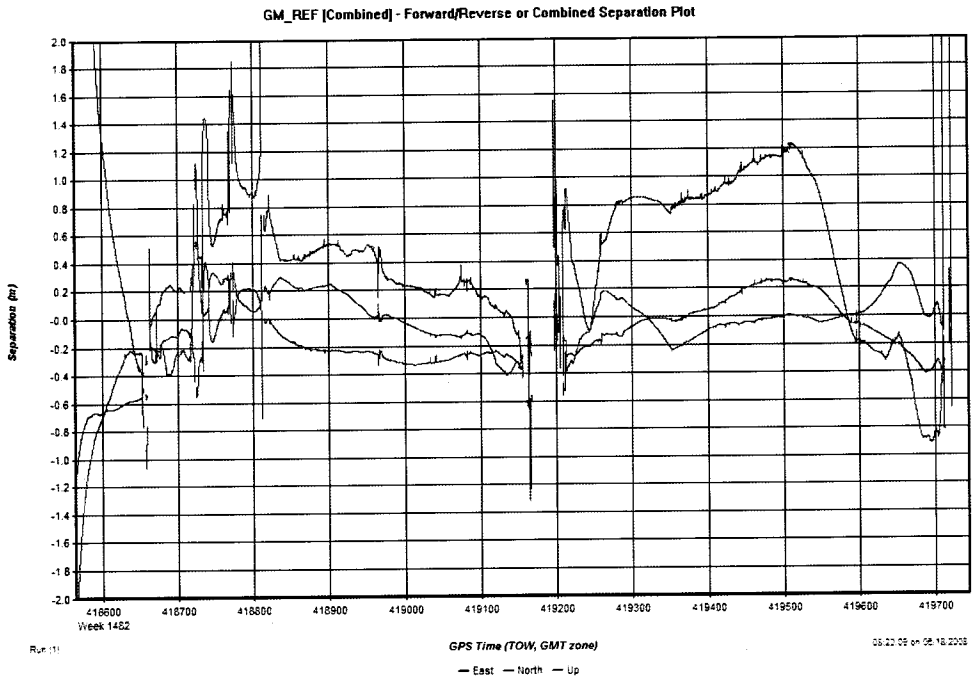

Figure 6: Post-Mission L1 Forward-Backward Solution Discrepancy

4 Results

An analysis summary is presented below under several scenarios.

4.1 Scenario 1: Impact of Differences in Satellite Visibility

This scenario looks at the impact of sky visibility differences between V2V (i.e., one vehicle using a GPS satellite or more that is not available to the other) and its impact on relative positioning using SP and RTK methods. As observed during these field tests, such difference may occur often depending on the sky visibility obstructions around the vehicles. A typical example is a situation where a vehicle is driven next to (i.e., adjacent lane) a semi-like vehicle that obstructs a part of the sky view of the vehicle and other vehicles in the same area may not have such obstructions.

An illustration of how this impacts the SP method is given in Figure 7. At time = 0, vehicles A and B are in adjacent lanes and both are operating in WAAS positioning mode. Given that certain conditions discussed in background are true, vehicle receivers output their locations as A_{WAAS} and B_{WAAS}. As shown in Figure 7, the reported positions (i.e., A_{WAAS} and B_{WAAS}) are offset by a common error vector from the actual positions of the vehicles A and B. However, at time = 5 sec, vehicle B may see additional GPS satellites that are not seen by vehicle A, and its error vector may change. This could result in a change of reported vehicle position B_{WAAS}. This error vector change is a function of many variables including quality of the additional/lost satellite(s) measurement(s) and the way each receiver translates the GPS measurements into position solutions. From a VSC-A point of view, if the SP method is used to estimate D_{Across} in this sequence of events, D_{Across} could potentially change from a single lane width to more than (as depicted in Figure 7) or less than a lane width with no actual vehicle orientation change.

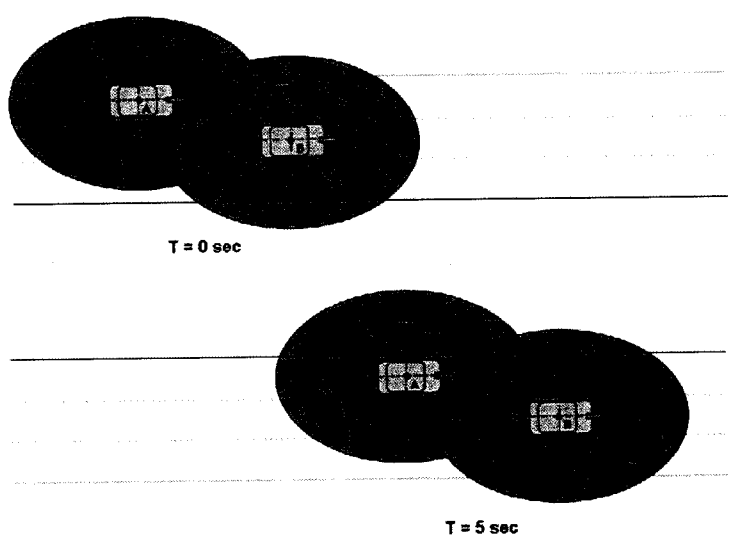

Figure 7: SP Relative Positioning Method Dependency on Satellite Visibility

It is noted that the sky visibility changes only affects the number of observations available to a given vehicle. If a certain subset of satellites is seen by both vehicles in

consideration, measurements from those satellites are not affected by the local obstruction. Therefore, this scenario is not expected to impact a relative positioning method that uses actual GPS measurements such as the RTK method. More information on the technical basis for this hypothesis can be found in the Background section (section 1) of this report and the references.

In order to verify the above hypothesis, D_{Across} estimated using SP method and RTK method of relative positioning was compared in situations in which GPS satellite count seen from each test vehicle were different. Figure 8 to Figure 12 show segments of data highlighting such time intervals. A description of the vehicle orientation and other parameters are given with each figure.

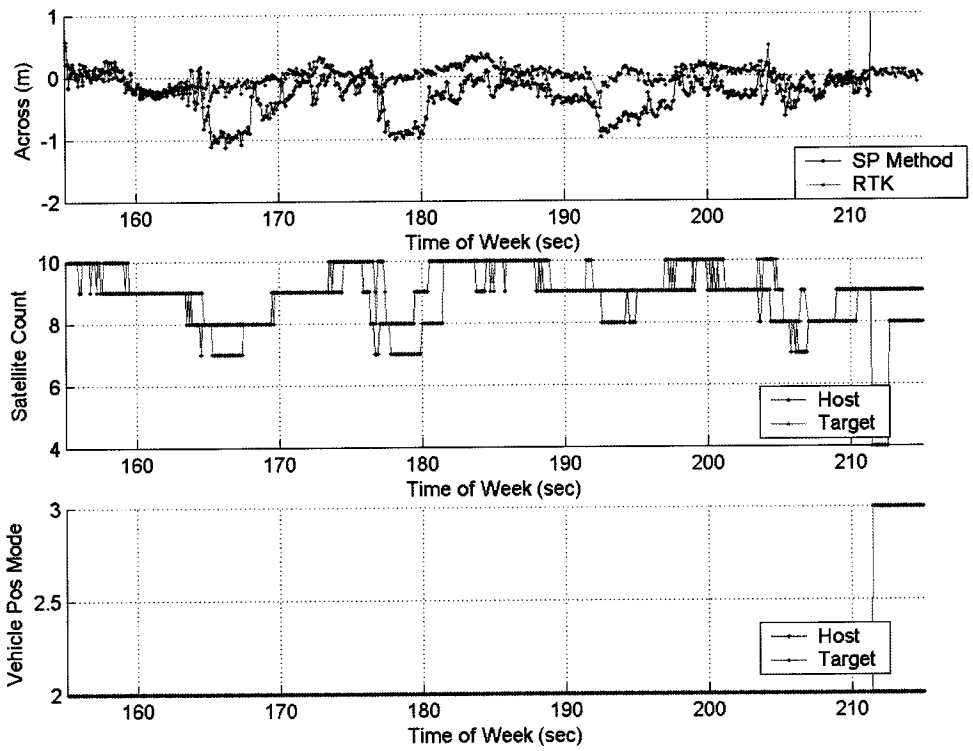

Figure 8: Visibility Difference Analysis Set 1 (WAAS)

Figure 8 shows a 60-second time window in which the host and the target (i.e., remote) vehicles were driven in the same lane on a straight-road segment. D_{Across} observed using RTK and WAAS methods is shown in the top plot of Figure 8. The plot in the middle shows the total number of satellites seen by each vehicle. The bottom plot shows the positioning modes of individual vehicles which indicates mode 2 (WAAS) for both vehicles for the majority of the duration shown.

As expected, the D_{Across} estimated by the RTK method shows values around zero for the whole duration, and the maximum deviation was found to be within ± 0.5 m from 0 m, the best estimate of the reference D_{Across} for the same lane formation. However, for the SP method based estimate, clear deviations of up to 1 m are seen. These deviations are directly related to the differences in total number of satellites seen by vehicles. For

instance, the deviation around 180 seconds is highly correlated with the target vehicle satellite count dropping one below that of the host. However, it is noted that every satellite count differences does not necessary result in a bias in SP method D_{Across} estimate as indicated by data shown around time 158 seconds and 210 seconds. The target vehicle positioning mode changes after 210 seconds to the run, and the impact of this is discussed in a later scenario. Importantly, D_{Across} observed by the RTK method remains unchanged throughout the whole time interval, including after the target vehicle positioning mode change.

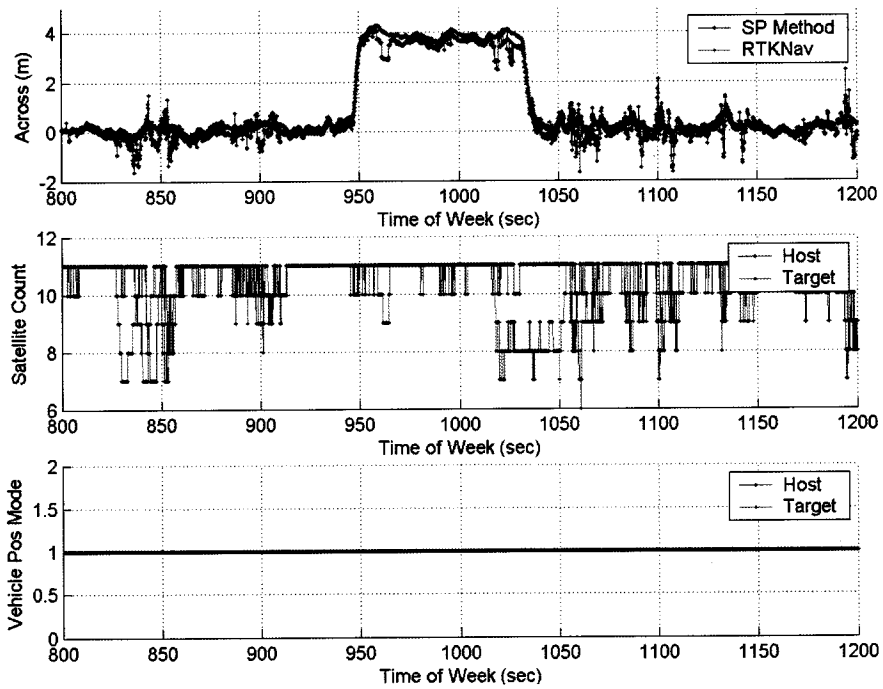

Figure 9: Visibility Difference Analysis Set 2 (GPS)

Figure 9 shows a dataset in which one of the vehicles changed lanes and came back to the original lane during the time period shown. As seen in the top plot, both the target and the HV remained in the same lane until time 950 seconds based on the near zero D_{Across}. The target vehicle subsequently changed lanes to the adjacent lane, as indicated by a D_{Across} close to a single lane width. After around 70 seconds, the target vehicle comes back to the same lane formation with the HV.

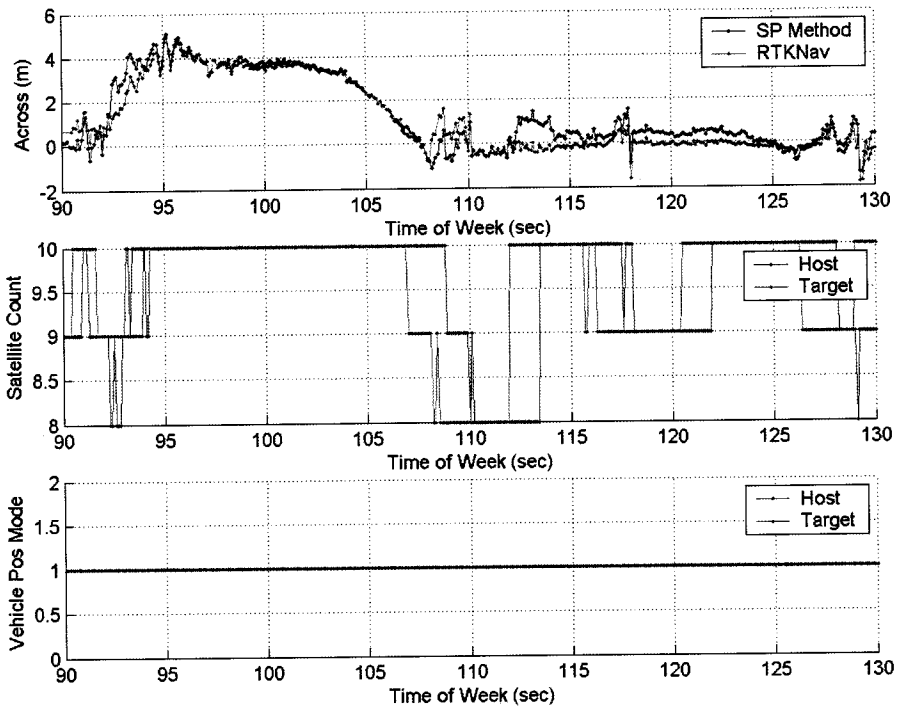

Figure 10: Visibility Difference Analysis Set 3

Figure 10 shows the time duration where the target vehicle performs a lane change maneuver similar to the one in Figure 9. Errors in excess of 1 m are seen when using the SP mode of relative positioning. Around 112 seconds, a deviation in SP D_{Across} is observed due to the HV seeing two satellites less than the target vehicle. However, it is noted that the visibility difference around 120 seconds, where the HV once again sees two satellites less than the target vehicle, does not introduce any significant changes to the SP method D_{Across}.

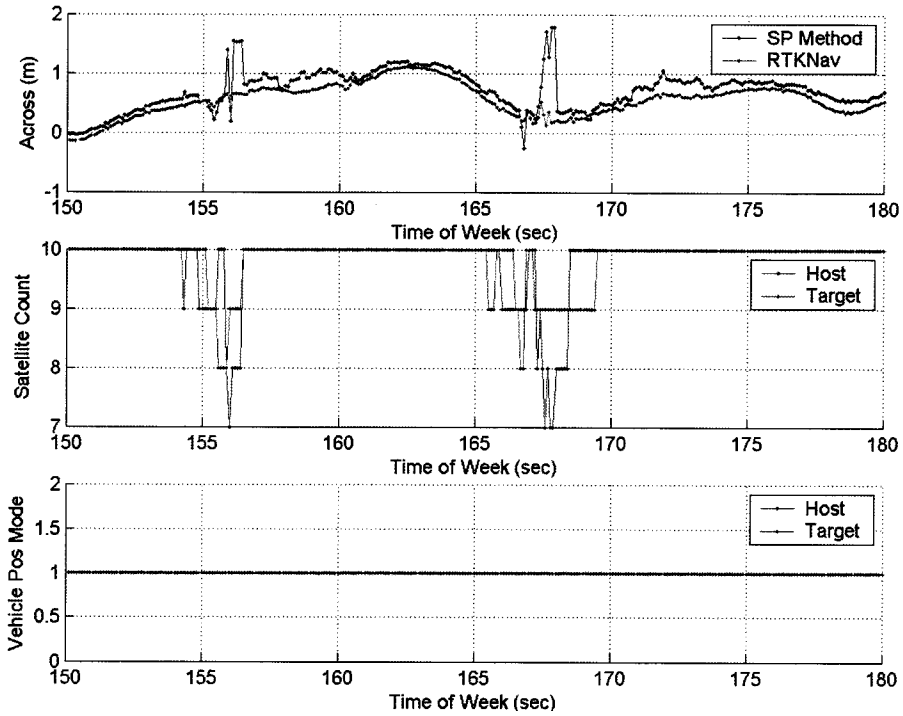

Figure 11: Visibility Difference Analysis Set 4

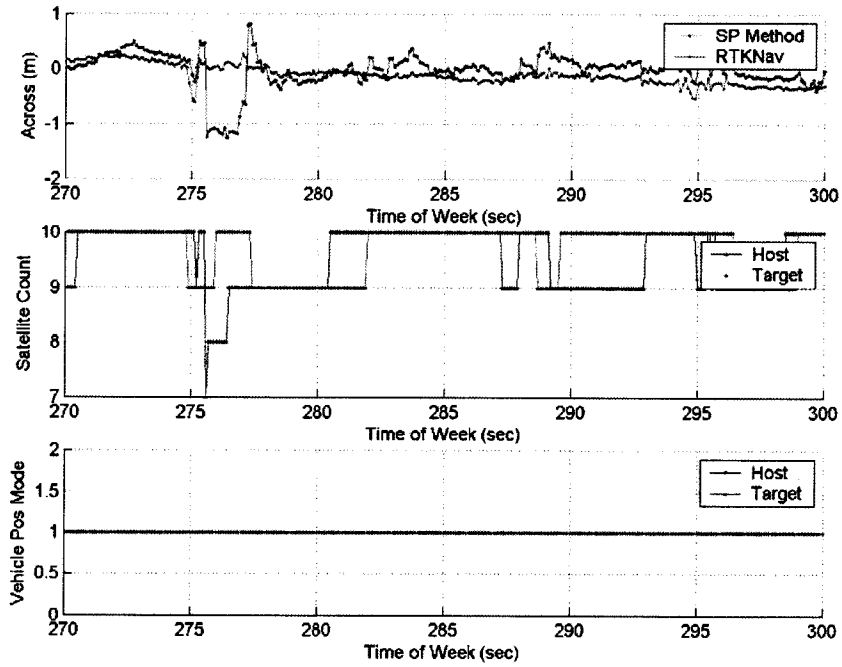

Figure 12: Visibility Difference Analysis Set 5

Figure 11 and Figure 12 also show two other data segments in which the SP method reported D_{Across} becomes erroneous due to satellite visibility changes.

4.2 Scenario 2: Impact of Positioning Mode on Relative Positioning

As outlined in the Background section, error characteristics of different positioning modes are significantly different. These differences cause problems when receivers change modes or vehicles operating in different modes use SP method for relative positioning. The root of these issues is the variations of vehicle position reported by receivers when positioning modes switch. In the case of vehicles operating in difference modes, positioning mode dependent errors may not cancel out in the SP mode. This scenario investigates the impact of vehicle positioning mode on the relative positioning when using RTK and SP methods.

This scenario is considered especially important for the deployment of VSC-A and CICAS-V-like applications as positioning mode changes are inevitable under these conditions. For instance as illustrated in Figure 13, a CICAS-V enabled vehicle is most likely to operate in WAAS mode when local RTK data is not available from a CICAS-V intersection. As the vehicle enters a CICAS-V coverage area (i.e., at t = 5 seconds in Figure 13), the vehicle positioning mode will change from WAAS to RTK and this in turn tightens the absolute accuracy of the vehicle position. As a result, the reported vehicle position may have an instantaneous change that could be as high as 2 m (i.e., WAAS has a root mean square (RMS) accuracy bound of ~2 m, and RTK typically has an accuracy bound better than 0.5 m). If multiple vehicles are considered that are traveling very close to each other, such mode transition may not take place at the same time. In Figure 13 illustration, only vehicle B has switched to the RTK mode and vehicle A remains in WAAS mode at t = 5 seconds. Hence, using the SP mode of relative position under these conditions is expected to create relative positioning issues. It is noted however that RTK method is not vulnerable to this mode of failure as GPS measurements used in RTK method are vehicle position mode independent.

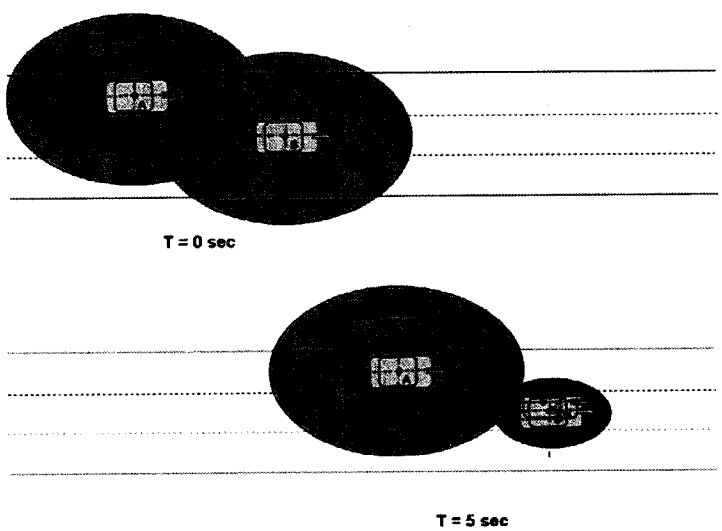

Figure 13: SP Relative Positioning Method Dependency on Positioning Mode

Figure 14 shows a segment of data from a driving scenario where the host and the target vehicles were in the same lane and were driving toward a CICAS-V intersection in a straight road. Both vehicles were configured to operate using WAAS and only one of them was configured to work with CICAS-V data (RTK mode) when available. From time 715 seconds onward, the vehicles were stopped for a traffic light in the same lane.

The target vehicle switched to RTK mode around 650 seconds as shown in the bottom plot of Figure 14. It is noted that the number of satellites used in the target vehicle drops to 5 at the same time as shown in the middle plot of Figure 14. This is a characteristic of the RTK engine which starts precise positioning by using the best 4 satellites and later adds additional satellites to the solution. Although not presented in this report, the absolute position accuracy of the target vehicle increases to better than 0.5 meters after the mode switch to RTK whereas that of the host remains the same.

The top plot shows the D_{Across} estimated by the SP and RTK modes, and the mode switch introduces a bias of ~1 meter due to the mode transition. It is noted that this is solely due to the error in Host vehicle position that is in WAAS mode. As expected, the RTK D_{Across} remains within ± 0.5 meters of the expected D_{Across} of 0 meters as the vehicles are in the same lane formation.

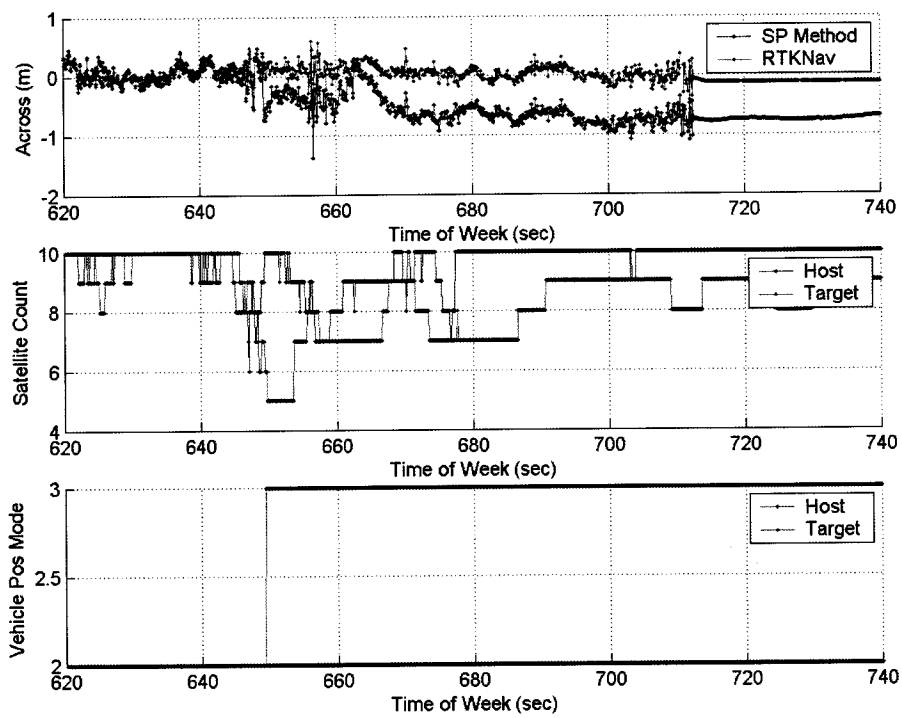

Figure 14: Positioning Mode Impact Analysis Set 1

Figure 15 shows a driving scenario where the host and target vehicles were driven in the same lane while driving through a CICAS-V intersection coverage area. The vehicles perform three lane changes at 1525, 1560, and 1580 seconds. Each of these lane changes was performed such that the target vehicle changes to the adjacent lane first followed by the host performing the same lane change. The vehicles were in the same lane formation

outside of the three lane changing maneuvers. As shown in the bottom plot of Figure 15, the target vehicle enters RTK mode around 1500 seconds and leaves the CICAS-V coverage area around 1585 seconds, thus transitioning back to WAAS mode.

As seen in the top plot of Figure 15, the HV RTK D_{Across} estimate correctly reflects the orientation of the target vehicle which shows D_{Across} values of the order of a lane width when the lane changes take place and close to 0 values when the vehicles are in the same lane formation. However, the SP method shows a D_{Across} error of approximately 2 meters for the whole duration the target vehicle is in RTK mode. As pointed out in the preceding discussion, this happens due to the presence of ~2 meters of absolute positioning error in the HV position as it operates in WAAS mode. This error is corrected by the RTK processing using the local data in the target vehicle. When both vehicles are in WAAS mode, this absolute error is almost removed in relative processing as they are similar in both vehicle positions. It is noted, however, that in WAAS mode, both vehicles are ~2 meters off from the true location of the vehicle which is not acceptable for CICAS-V operations.

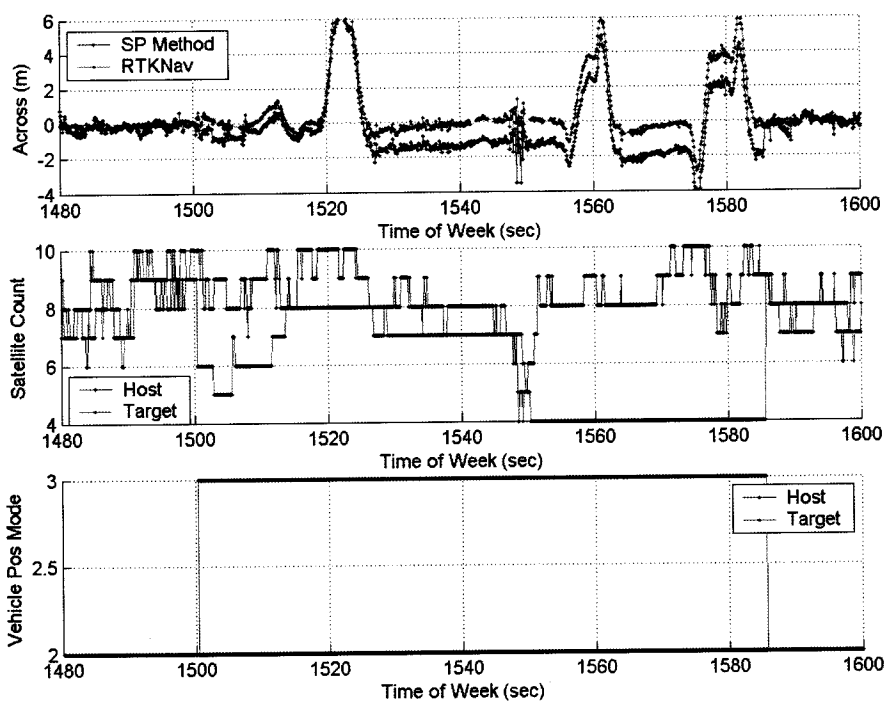

Figure 15: Positioning Mode Impact Analysis Set 2

4.3 Scenario 3: Impact of Using Other Positioning Mode Combinations

This section illustrates the impact of vehicles using positioning mode combinations other than WAAS/RTK on relative positioning. One of the vehicles in these tests was set to operate in standalone GPS mode at all times. Although VSC-A/CICAS-V systems are likely to operate in WAAS mode by default, there could be many instances where some

vehicles could operate in GPS mode due to unavailability of WAAS data (for instance, due to local visibility restrictions).

Figure 16 illustrates a driving scenario in which the target vehicle starts in the same lane as the host, performs a lane change around 230 seconds, drives in adjacent lane for around 20 seconds before changing back to the same lane formation around 250 seconds into the run. As shown in the top plot of Figure 16, the RTK D_{Across} estimate reflects these changes as expected and within the ± 0.5 meter error bound of reference D_{Across} for same lane and adjacent lane formations. However, the D_{Across} estimate from the SP method shows an error of 2 meters around 260 seconds that remains for the rest of the time duration shown. It is noted that the vehicles were brought to a stop in a same lane formation at the end of the run (i.e., after 280 seconds). In addition, the SP method D_{Across} shows errors in the order of 1 meter due to satellite visibility differences around 215 and 225 seconds prior to the transition of target vehicle to RTK mode.

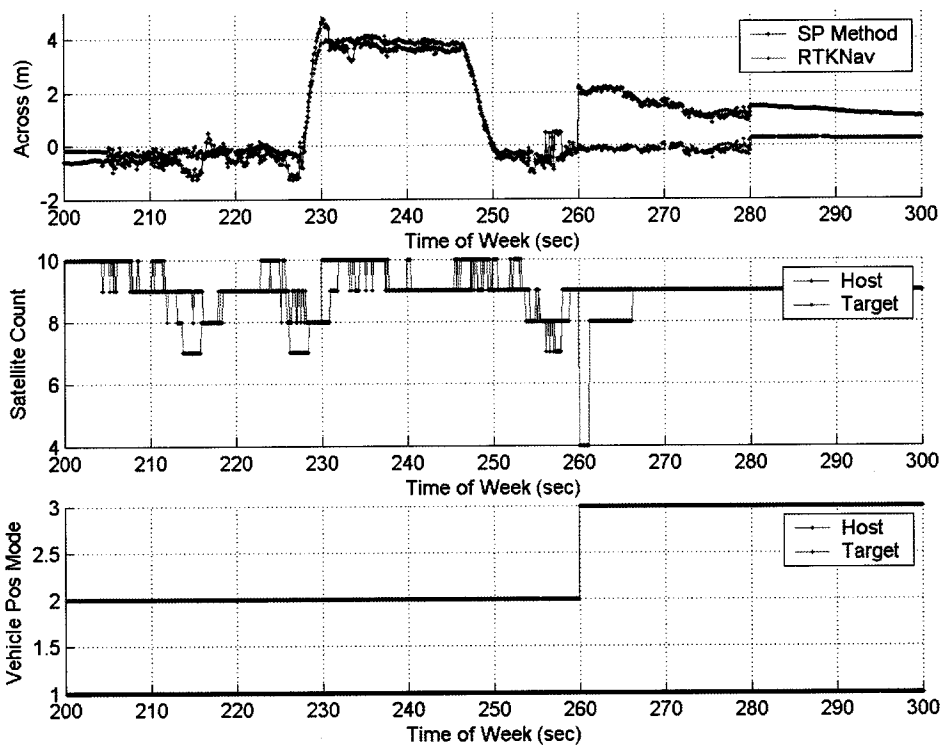

Figure 16: Positioning Mode Impact Analysis Set 3

Figure 17 illustrates a similar scenario to that shown in Figure 16 with two lane change maneuvers by the target vehicle and an instance where the satellite availability in both the target and host drops to 5 or less (i.e., at 1600 seconds). It is noted that the SP mode D_{Across} error in this case vary from almost zero to more than 2 meters when the target vehicle is in the RTK mode. Also the satellite visibility differences compound the SP mode D_{Across} estimate error, causing it to reach errors as high as ~5 meters at around 1600 seconds.

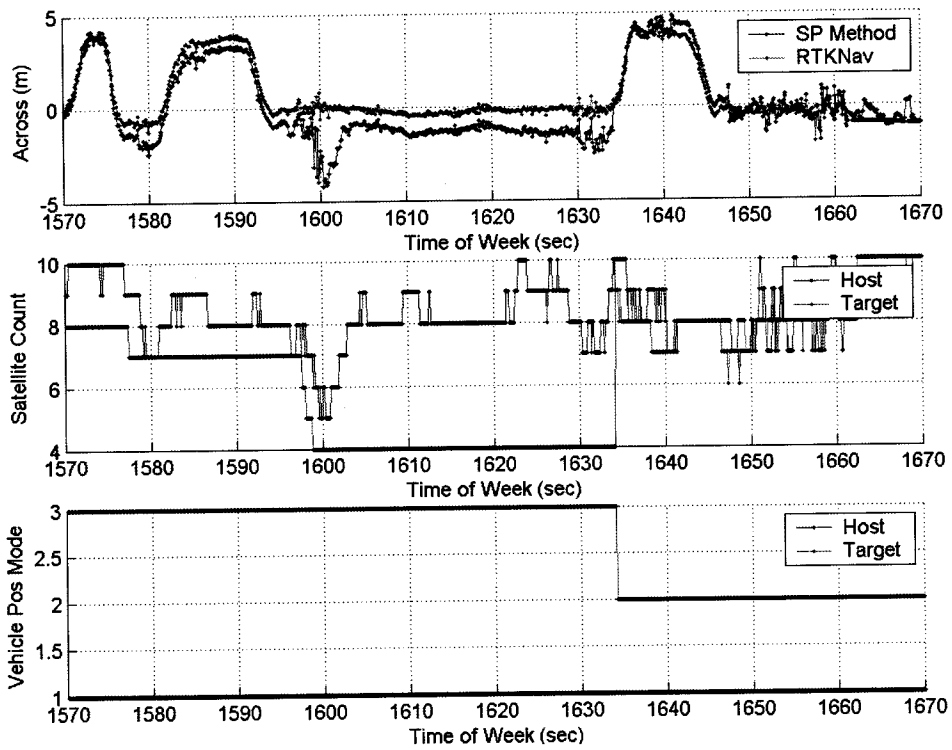

Figure 17: Positioning Mode Impact Analysis Set 4

4.4 Impact of Short Complete GPS Interruptions

The current implementation of the VSC-A test bed is designed for open-sky GPS conditions. However, this implementation is expected to perform under conditions that are commonly encountered but are not necessarily open sky. An example would be a situation where an overpass obstructs the view of the sky for a short duration on a freeway that is otherwise mostly open sky. Under these conditions, the VSC-A positioning system is expected not to provide misleading information to applications. Also the relative positioning system is expected to recover within a certain amount of time after normal GPS reception becomes available. This section shows some data excerpts that illustrate the behavior of the RTK and SP relative position solutions under short but complete GPS outages.

Figure 18 shows a data plot for a 30 second time duration within which the test vehicles were driven under two overpasses. As seen in the plot in the middle, two complete outages (i.e., satellite count dropping to zero) of approximately 3 seconds each were experienced by both vehicles. It is also noted that both vehicles were operating in GPS mode as shown by Vehicle Pos Mode 1 in the bottom plot. The VSC-A RTK SW stops outputting data as soon as common satellite visibility drops below 4. More importantly, VSC-A RTK SW starts outputting accurate estimates with 4-5 seconds of seeing more than 4 common satellites after the outage. This was considered typical under such conditions.

Since the GPS receivers continue to produce predicted vehicle position data during the outage, the SP method continues to generate estimates during the outage. As seen in the top plot of Figure 18, these estimates can be erroneous and highly unreliable. However, in contrast to the VSC-A RTK method, the SP method starts outputting position estimates near immediately within seeing 4 or more satellites.

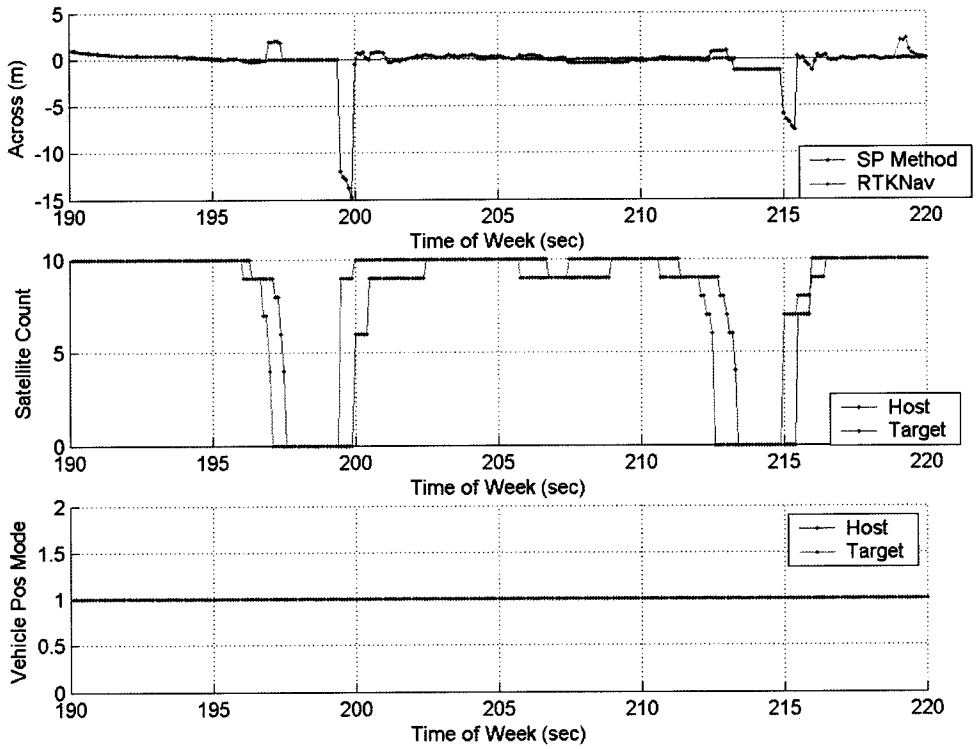

Figure 18: Impact of GPS Interruptions on Relative Positioning Set 1

Figure 19 shows another data duration that shows similar SP and VSC-A RTK SW relative positioning performance during and after GPS outages. The VSC-A RTK SW stops sending estimates as soon as the common satellite count drops below 4, and it resumes reliable output within 5 seconds of getting measurements from 4 or more satellites. Whereas the SP method continues providing a solution during the outage which should be considered unreliable, however, starts providing a solution near immediately after getting measurements from 4 or more satellites.

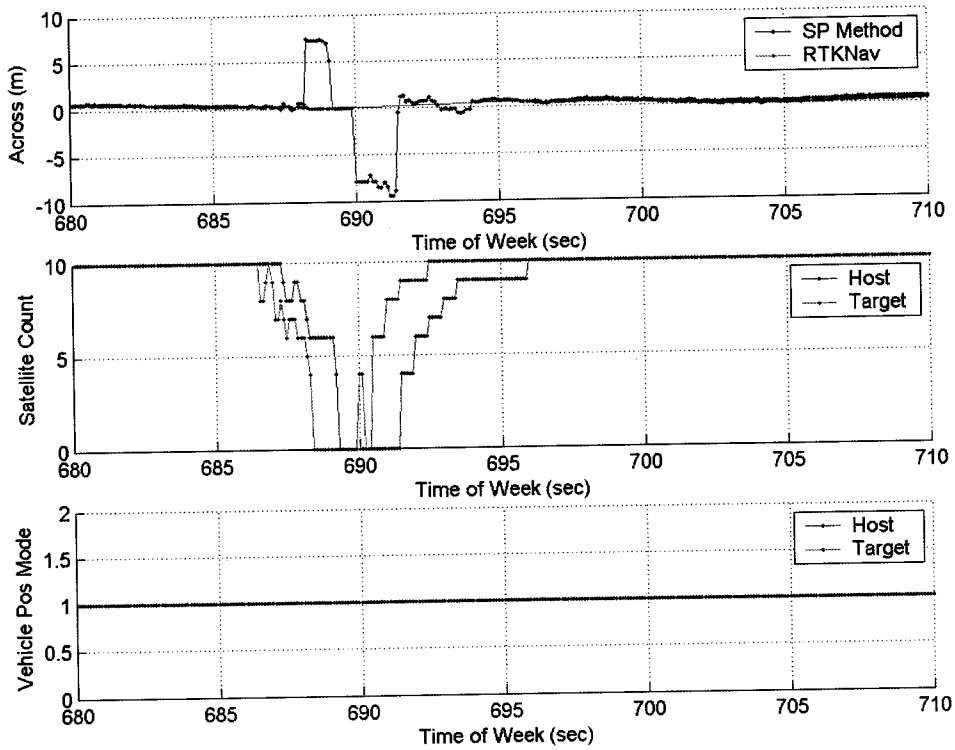

Figure 19: Impact of GPS Interruptions on Relative Positioning Set 2

5 Conclusions

Please refer to the main body of the final report for the positioning conclusions and recommendations.

6 References

[1] Misra, P. and Enge, P., "*Global Positioning System: Signals, Measurements, and Performance,*" Second Edition, Ganga-Jamuna Press, 2006.

VSC-A Final Report: Appendix E-2

GPS Service Availability Study Literature Review

Prepared by

PLAN Group
University of Calgary

List of Acronyms

CAMP	Crash Avoidance Metrics Partnership
CEP	Circular Error Probable
CODE	Center for Orbit Determination in Europe
CORS	Continuously Operating Reference Station
DD	Double-Differenced (measurement)
DGPS	Differential Global Positioning System
DOD	Department of Defense
DOP	Dilution of Precision
DRMS	Distance Root Mean Squared
EKF	Extended Kalman Filter
EU	European Union
FAA	Federal Aviation Administration
FDMA	Frequency Division Multiple Access
FKP	Flächen Korrektur Parameter
GAO	U.S. Government Accountability Office
GDOP	Geometric Dilution of Precision
GIM	Global Ionosphere Map
GLONASS	Global Navigation Satellite System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HDOP	Horizontal Dilution of Precision
HW	Hardware
I2V or I-V	Infrastructure-to-Vehicle
IGS	International GNSS Service
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
ITS	Intelligent Transportation Systems
JPO	Joint Program Office
LKF	Linearized Kalman Filter
MEMS	Micro-Electro-Mechanical System

MRSE	Mean Radial Spherical Error
NHC	Non-Holonomic Constraints
NHTSA	National Highway Traffic Safety Administration
PDOP	3d Position Dilution of Precision
PF	Particle Filter
PPM	Parts Per Million
PRN	Pseudo-Random Noise
RITA	Research and Innovative Technology Administration
RMS	Root Mean Squared
RTK	Real Time Kinematic
SEP	Spherical Error Probable
SP	Single Point
SPKF	Sigma Point Kalman Filter
TDOP	Time Dilution of Precision
UERE	User Equivalent Range Error
UKF	Unscented Kalman Filter
USAF	United States Air Force
URE	User Range Error
USDOT	United States Department of Transportation
UTC	Coordinated Universal Time
V2I	Vehicle-to-Infrastructure
VDOP	Vertical Dilution of Precision
VRS	Virtual Reference Station
VSC2	Vehicle Safety Communications 2
VSC-A	Vehicle Safety Communications – Applications
V-V or V2V	Vehicle-to-Vehicle
WASS	Wide Area Augmentation System
ZUPTs	Zero Velocity Updates

Table of Contents

List of Acronyms **E-2-ii**

1 Introduction **E-2-1**

2 GPS Overview **E-2-1**

2.1 Constellation Overview E-2-1

2.2 GPS Signal Structure E-2-1

2.3 GPS Modernization E-2-2

3 Outline of GPS Error Sources and Characteristics **E-2-3**

3.1 Spatially Correlated Errors E-2-3

3.1.1 Ionosphere Errors E-2-3

3.1.2 Troposphere Errors E-2-5

3.1.3 Orbit Errors E-2-6

3.1.4 Clock and Timing Errors E-2-6

3.2 Uncorrelated Errors E-2-7

3.2.1 Noise and Multipath E-2-7

3.2.2 Differences in Receiver Technology E-2-10

3.3 Measurement Errors in V2V Navigation E-2-12

4 Parameters that Impact GPS Performance **E-2-13**

4.1 Definitions of Performance Measures E-2-13

4.2 Factors that Affect Dilution of Precision E-2-15

4.2.1 Availability of Satellites and Constellation Size E-2-15

4.2.2 Time of Day E-2-16

4.2.3 Geographic Location E-2-16

4.2.4 Elevation Mask Angle and Obstructions E-2-20

5 Reference Stations and Reference Station Networks **E-2-22**

5.1 Differential GPS E-2-22

5.2 WAAS E-2-23

5.3 Local Base Stations E-2-24

5.4 Network-based RTK E-2-25

6 Review of Integration of GNSS with Inertial Navigation **E-2-29**

6.1 Introduction and Motivation E-2-29

6.2 Methodology E-2-30

6.2.1 Inertial Navigation Basics E-2-30

6.2.2 Equations of Motion and Error Equations E-2-31
6.2.3 Integration of GNSS and INS E-2-33
6.3 Review of Published Results E-2-36
7 Impact of Future GNSS and GPS Modernization E-2-45
7.1 Impact of New GPS Signals E-2-45
7.2 Potential Use of Galileo, GLONASS, and COMPASS E-2-45
7.2.1 Additional GNSS Constellations E-2-47
8 References E-2-52

List of Figures

Figure 1: L1 Ionosphere Error for January 1, 2004, at 0:00 Coordinated Universal Time (UTC) Derived from a Global Ionosphere Map from the CODE E-2-5
Figure 2: DD Code Residuals – Combined Effect of Noise and Multipath for L1 C/A Code for a High Quality L1 Receiver E-2-8
Figure 3: DD L1 Phase Residuals – Combined Effect of Noise and Multipath for L1 Carrier Phase for a High Quality L1 Receiver E-2-9
Figure 4: DD Code Residuals – Combined Effect of Noise and Multipath for L1 C/A Code for a High Sensitivity "Low Cost" Receiver E-2-9
Figure 5: DD L1 Phase Residuals – Combined Effect of Noise and Multipath for L1 Phase for a High Sensitivity "Low Cost" Receiver E-2-10
Figure 6: Zero Baseline L1 Double Difference Phase Errors for u-blox LEA-4T for June 31, 2006 (from Schleppe 2006) E-2-11
Figure 7: Zero Baseline L1 Double Difference Phase Errors for NovAtel OEM4 – April 20, 2006 (from Schleppe 2006) E-2-12
Figure 8: Median HDOP (50 Percentile Value) over the United States Computed with the Current 31-Satellite GPS Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask E-2-17
Figure 9: Median HDOP (50 Percentile Value) over the United States Computed with the "as Designed" 24-Satellite GPS Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask. E-2-18
Figure 10: 95 Percentile HDOP (95 Percent of the Time the Values are Less than This) over the United States Computed with the Current 31-Satellite GPS Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask E-2-19
Figure 11: 95 Percentile HDOP (95 Percent of the Time the Values Are Less than This) over the United States Computed the "As-Designed" 24-Satellite Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask E-2-20
Figure 12: 95th Percentile HDOP Value for a 24-Satellite GPS Constellation and a 10-Degree Elevation Mask E-2-21
Figure 13: 95th Percentile HDOP Value for a 24-Satellite GPS Constellation and a 20-Degree Elevation Mask E-2-22
Figure 14: Example of the Measurement Error of One Satellite Across a Region (red line) (This error is measured at two locations indicated by the green and blue reference stations.) E-2-26
Figure 15: The Estimated Residual Dispersive (Ionosphere) Error for a Single Reference Station User (Left) and a Network Reference Station User (Right) E-2-27
Figure 16: The Estimated Residual Non-Dispersive (Troposphere/Geometry) Error for a Single Reference Station User (Left) and a Network Reference Station User (Right) E-2-27
Figure 17: Concept of GNSS/INS Integration E-2-34
Figure 18: 95th Percentile HDOP Using a 24-Satellite GPS Constellation and a 20-Degree Elevation Mask E-2-48

Figure 19: 95th Percentile HDOP Using a 24-Satellite GPS constellation in Conjunction with a 30-Satellite Galileo Constellation and a 20-Degree Elevation Mask E-2-49
Figure 20: 95th Percentile HDOP Using GPS, Galileo, and GLONASS with a 20-Degree Elevation Mask E-2-50
Figure 21: 95th Percentile HDOP Using GPS, Galileo, GLONASS, and Compass with a 20-Degree Elevation Mask E-2-51

List of Tables

Table 1: Combined Code and Carrier Phase Noise and Multipath RMS Error Shown in Raquet (1998) E-2-7
Table 2: GPS Errors in Single Point and Differential Mode Assume a Baseline Length of the Order of 10s of Kilometers and a Baseline Length of 10s of Seconds (after Misra and Enge 2001) E-2-13
Table 3: Key Characteristics of GNSS and INS E-2-29
Table 4: General IMU Classifications (from Petovello 2003b) E-2-31
Table 5: Summary of Results with Navigation-Grade IMUs E-2-39
Table 6: Summary of Results with Tactical-Grade IMUs E-2-40
Table 7: Summary of Results with Automotive-Grade IMUs E-2-42
Table 8: Summary of Results with Reduced IMUs E-2-44
Table 9: Comparison of GPS and Galileo Ambiguity Resolution over a 1 KM Baseline (from Cao, et al. 2008b). E-2-47

1 Introduction

This is the first of two reports as part of the Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) testing and analysis project. This report contains the literature review only. The other report contains the results, analysis, and conclusions from the study.

2 GPS Overview

The Global Positioning System (GPS) is an all-weather satellite navigation system operated by the United States Air Force (USAF). The system consists of three segments: the Control Segment, consisting of ground-based tracking and control stations; the User Segment, which includes all user receivers; and the Space Segment, which consists of a constellation of satellites in medium Earth orbit that transmit synchronized ranging signals, and information about the satellite orbit, to users on or near the surface of the Earth. Users must track at least four satellites, decode the navigation message transmitted by each in order to determine the satellite positions, and then use this information and the four or more range observations to compute four unknowns, which are the user's three coordinates (latitude, longitude, and altitude) and the user-receiver clock offset. The system is well described in a number of textbooks. Two popular texts are Misra and Enge (Misra and Enge 2001) and Leick (2004). A more detailed treatment is given in the two volume GPS Blue Books (Parkinson and Spilker, Jr. 1996).

2.1 Constellation Overview

The system design calls for 24 satellites in near circular 12-hour orbits distributed in 6 orbital planes inclined at 55 degrees to the equator. The 24 satellites in the guaranteed constellation were originally described as 21 satellites plus 3 active spares; however, since at least 1995, the guaranteed minimal constellation is 24 operational satellites (DOD 1995). The constellation is designed to ensure at least four satellites are in view at all times at all locations on Earth. As of June 2009, there were 31 operational GPS satellites. Updates on constellation status, as well as almanac files describing the satellite orbits and notices to users, can be found at the United States Coast Guard Navigation Centre webpage. Although GPS has reliably provided continuous and consistent service for many years, there are concerns that it will not be possible to maintain the same level of service in the future. The U.S. Government Accountability Office (GAO) (2009) suggests that there is a small chance that the number of active GPS satellites could decrease below 24 over the next five years if efforts are not made by the USAF to correct management projects in its satellite procurements programs.

2.2 GPS Signal Structure

The GPS signal structure is complex and will not be reviewed in detail here. Instead, only the material relevant to this report is provided. For more details on the signal structure, refer to Ward (1996), Spilker (1996b), and Misra and Enge (2001).

The GPS signal is comprised of two frequencies, namely L1 (1575.42 MHz) and L2 (1227.60 MHz). Modulated on these carriers are the:

- Pseudo-Random Noise (PRN) codes used for ranging measurements
- Navigation data to communicate the satellite's position, time, health, etc., to users in real-time (50 bps modulation)

Currently, only two types of PRN codes are used, namely the Coarse/Acquisition code (C/A-code) on L1, and the Precise code (P-code) on L1 and L2. Exploiting the characteristics of the signal structure, the following three types of measurements can be obtained from most GPS receivers.

- Pseudorange (code) measurements - These are derived from the PRN codes and are, therefore, classified according to code and frequency as L1-C/A, L1-P. and L2-P
- Carrier phase (phase) measurements - By measuring the phase of the incoming carrier (L1 and/or L2), the range to a satellite can be measured; however, it is biased by an ambiguous number of cycles. This is due to the fact that it is only possible to measure a phase between 0 and 360 degrees and its subsequent change over time, and it is not possible to determine a pseudorange directly from a phase measurement.
- Doppler measurements - The derivative of the carrier phase measurement is the Doppler shift caused by the relative receiver-satellite motion

In terms of code measurements, the P-code theoretically provides better overall performance. Unfortunately, the P-code signal is currently encrypted in an attempt to limit its use to the military community (including the entire L2 signal). However, codeless and semi-codeless tracking techniques have been developed which allow the civil community access to these signals. Unfortunately, these techniques decrease the signal-to-noise ratio by 14 dB or more and, therefore, produce considerably noisier measurements than would otherwise be expected. These techniques are described as "unauthorized use" by the USAF and are generally limited to land surveying and scientific applications while consumer and civil aviation and maritime users use L1 C/A exclusively. Almost all currently commercially available dual frequency receivers use codeless or semi-codeless techniques and, therefore, are more expensive than a dual frequency receiver that is to track a second frequency directly such as L2C.

2.3 GPS Modernization

The GPS L1 and L2 signal structure was designed in the 1970's and has been used operationally for over 30 years. In parallel with ongoing replacement of satellites, several major initiatives are underway to modernize the GPS signal for both civil and military users. Only those aspects of GPS modernization relevant to civil users are discussed here. The first modernization is the addition of a civil code to the L2 signal. L2C (for L2 Civil) is, at the time of this writing, available on five recently deployed GPS satellites. The addition of this new ranging code allows for direct acquisition of L2 and allows civil use of L2 without the requirement for codeless techniques which will result in less expensive dual frequency GPS being available in the near future.

A third frequency, L5 1176.45 MHz, is planned; and one L5 capable satellite has been deployed. The fully modernized GPS, or GPS III, will consist of at least 24 satellites each broadcasting civil signals on three frequencies. The implications of this for users are discussed in Section 7 below as well as the other global navigation satellites systems that are being developed and deployed by the European Union, Russia, and China.

3 Outline of GPS Error Sources and Characteristics

The precision, accuracy and reliability of GPS positioning and navigation is dependent on the level of errors present in the observations. The properties of observation errors change over time and geographic region. This section describes the error sources relevant to single point (SP) and differential code and carrier phase positioning.

The error sources effecting GPS observables can be roughly divided into two categories based on whether or not the errors are correlated with the antenna location (i.e., spatially correlated). Errors that are spatially correlated can be reduced or eliminated by differencing between receivers that are located close to one another. Between receiver observation difference works because both receivers are affected by similar levels of systematic errors. If observations are affected by similar (spatially correlated) systematic errors, then the common part of these errors will be cancelled out in the difference leaving a smaller residual error. In the process, the position estimation changes from estimating the absolution position of one receiver to estimating the relative position or baseline between the two receivers. In the case that the coordinates of one receiver are already precisely known, then the coordinates of the other receiver can be determined. This is the basis for differential GPS, which is discussed in Section 3.3 below.

3.1 Spatially Correlated Errors

The GPS errors that can be reduced or eliminated by differencing between receivers include the ionosphere, troposphere, and satellite clock. The two atmospheric errors (troposphere and ionosphere) can be reduced by differencing between receivers. The shorter the distance between the receivers, the greater the reduction of correlated errors due to the signal traveling through more or less the same path in the atmosphere to reach the two receivers. The remaining errors after differencing are called differential errors and are usually expressed in relative terms (such as parts per million (PPM)) with respect to the baseline length. One PPM is equivalent to 1 mm of error over 1 km. In the case of the satellite clock error, it is cancelled completely using between-receiver differencing. Each error is discussed below.

3.1.1 Ionosphere Errors

The ionosphere is a region of the atmosphere which contains weakly ionized plasma (Klobuchar et al 1995). The ion content (free electrons) in this region has various effects on electromagnetic signals, such as GPS.

The ion content of the ionosphere is distributed from 60 to more than 1000 km above the surface of the Earth (Klobuchar et al 1995; Leva et al 1996). However the peak density is located around 300 to 450 km. The effect of the ionosphere on radio-navigation signals is

a function of the integration of the electron density along the signal's path. The effect of the ionosphere is a function of the frequency of the signal for L-band signals. Carrier phase measurements are advanced by the ionosphere by the same amount that the code is delayed. This property is commonly used to estimate or mitigate the ionospheric effect.

The ionosphere's variability is due to the number of free electrons, which is a function of solar radiation. As a result, there is a daily variation of the ionosphere such that it is relatively calm at night and is most active around 14:00 local time.

There are also regional effects due to the sun. These effects can be seen in Figure 1, which shows the estimated global ionospheric error on January 1, 2004, at 0:00 UTC. This is derived from a Global Ionosphere Map (GIM) produced by the Center for Orbit Determination in Europe (CODE). There is a significant ionosphere gradient at low geomagnetic latitudes which is amplified at approximately 14:00 local time. During ionospheric storms, there can also be significant ionospheric gradients at the poles.

Ionospheric storms can also cause localized pockets of charged particles. When the Global Navigation Satellite System (GNSS) signals pass through these pockets, they can change rapidly. These rapid changes can cause significant measurement biases and, in some cases, cause the receivers to lose tracking lock of the signals. This effect is called scintillation and only occurs under the most severe circumstances in high latitudes and in equatorial regions.

The ionosphere error is, at times, the largest error source for absolute GNSS positioning. It can vary from less than 5 meters to more than 150 meters during extreme conditions (Wells 1999) although is typically 2 to 10 meters (Wells 1999).

The ionosphere is also the largest error source for differential GPS ranging from less than 1 part per million (ppm) of the inter-antenna distance during low ionospheric periods at mid latitudes to greater than 10 ppm at low geomagnetic latitudes during midday.

For vehicle-to-vehicle (V2V) relative positioning, with distances less than 1 km, the magnitude of the residual differential ionosphere error typically ranges from 1 mm to 1 cm depending on the level of ionospheric activity.

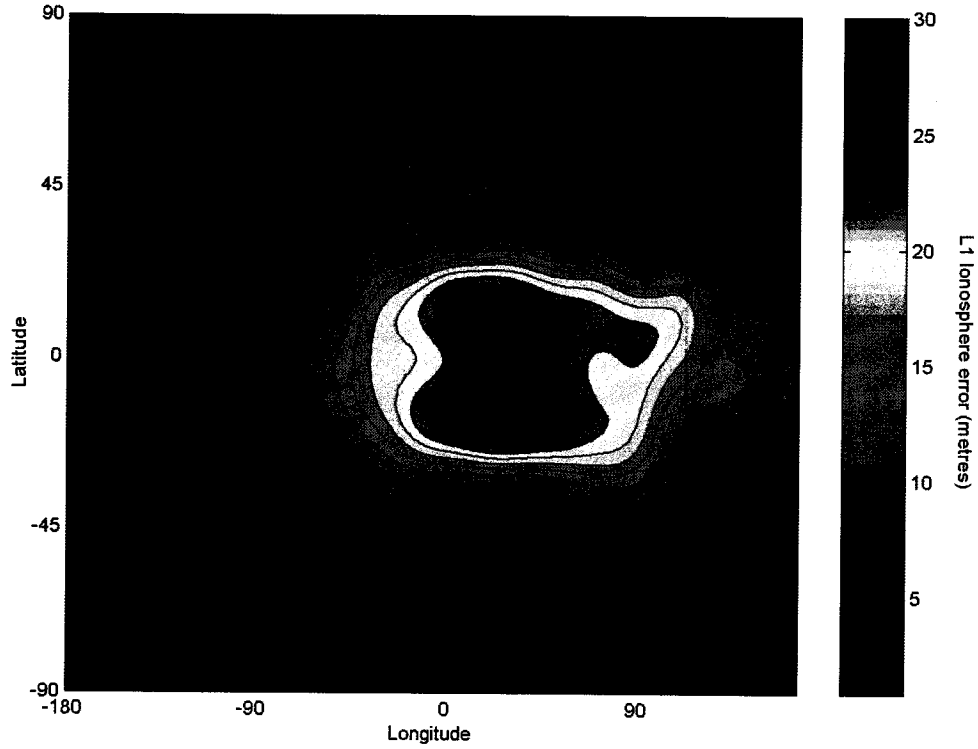

Figure 1: L1 Ionosphere Error for January 1, 2004, at 0:00 Coordinated Universal Time (UTC) Derived from a Global Ionosphere Map from the CODE

3.1.2 Troposphere Errors

The troposphere is a region of the atmosphere that spans from the Earth's surface from 12 to 14 km above the surface (Spilker, Jr. 1996b). The composition of the gases in this region has an impact on GPS signals. As the signal travels through these tropospheric gases, the signal refracts and slows the transmission speed of the signal, which both lengthens the measurement's path causing a delay in the time at which the signal is received by the user. The magnitude of the delay is relative to the atmospheric profile along the signal path.

The tropospheric delay is divided into two components, the dry and wet delays. The dry (or hydrostatic) part is due to the non-water content, and the wet delay is caused by the water content in the atmosphere. The hydrostatic part comprises approximately 90 percent of the delay but can be predicted with an accuracy of about 1 percent at the zenith using meteorological data. In contrast, the wet term makes up the remaining 10 percent of the error and can only be predicted with about 10–20 percent accuracy (de Jong, et al. 2002). The wet delay varies by 10 to 20 percent in a few hours (Spilker, Jr. 1996b).

Many models have been developed to reduce the effect of the troposphere on GPS measurements. Shrestha (2003) and Zhang (1999) give an overview of many common troposphere models.

The hydrostatic error is typically around 2.3 meters at the zenith and up to 10 times higher at low elevations; however, this effect is reduced to a few millimeters using any of the troposphere models (Zhang 1999; Shrestha 2003). The wet delay can be less than a few centimeters up to 35 centimeters depending on the amount of water vapor in the atmosphere. The differential residual troposphere error (after modeling) is on the order of 0.1 to 0.4 ppm of the inter-antenna distance (Alves 2005). This residual error over distances of less than 1 km (V2V navigation applications) is less than the magnitude of the measurement noise and is, therefore, negligible.

3.1.3 Orbit Errors

The orbit error is due to inaccuracies in the satellite position reported by the broadcast ephemeris. The effect of a satellite position error on the differential position is the projection of this error onto the direction of the observation vector (Parkinson 1996).

The magnitude of the absolute errors are about 3.5 m (50th percentile) (Ryan, 2002) and vary slowly with time (Olynik 2002). Orbit errors are also correlated as a function of the inter-antenna distance. Raquet (1998) shows that orbit error is usually less than 0.1 ppm of the inter-antenna distance. In terms of V2V navigation, the effect of orbit errors is less than 1 mm and is negligible.

3.1.4 Clock and Timing Errors

SP GPS positioning assumes that transmitting satellites are time synchronized to a time system called GPS Time. GPS Time is a realization of Coordinated Universal Time (UTC), and it is based on an average of the atomic clocks operated by the control and space segments of GPS. The time system is steered such that it follows UTC with the exception that it does not contain leap seconds. The GPS time scale is expressed in weeks and seconds from 0h UTC January 6, 1980. In reality, all of the satellite clocks differ from GPS time. This difference is called the satellite clock error, or offset. To mitigate this effect, each GPS satellite is monitored by the control segment; and a clock correction model is broadcast as part of the navigation message. The clock correction consists of a polynomial model representing the clock offset, drift, and rate of drift. After this correction is applied, a small residual satellite clock error term remains. 1 ns of satellite clock error corresponds to 30 cm of ranging error.

Separate from the satellite clock error, is a receiver clock error or offset. This arises from the fact the receiver clock is usually a low-cost crystal oscillator. The approach usually taken to remove this effect is to estimate it along with the unknown position of the receiver. Estimating the receiver clock offset will completely remove the effect of receiver clock biases provided that all of the measurements on the receiver are taken at precisely the same time. The satellite clock error contributes to 1-3 m of position error in SP mode (Kaplan 1996; Parkinson and Spilker 1996).

In differential mode, there is a small error created by the differences in measurement times of the two receivers. To minimize this effect, receivers steer their internal clocks to GPS time using the internal position and timing computations.

The estimated clock offset is correlated with the estimated position solution (since all 4 parameters, 3-D position, and receiver clock offset, are estimated together). The clock estimate is most highly correlated with the height component of the position estimation. This makes sense intuitively since if the receiver clock offset is estimated incorrectly and satellites are more or less uniformly distributed around but above the user, the average effect of the clock offset estimation error will be to either raise or lower the estimated position. This is the reason that GPS estimates height more poorly than the horizontal position components. The addition of a clock constraint (using a high quality clock) will improve the vertical solution and conversely the application of a height constraint will improve the clock offset estimation.

Another method for removing the effect of the clock is through differencing measurements from different satellites. The effect of receiver clock biases can be removed by differencing the measurements of two satellites that were observed at the same receiver and the effect of satellite clock biases can be removed by differencing the measurements of the same satellite that was observed at two different receivers. Differencing these two differences removes both the satellite and receiver clocks. This combination of measurements is called a double-differenced (DD) measurement.

3.2 Uncorrelated Errors

The uncorrelated errors, namely noise and multipath, are not a function of distance between antennae. These errors are described in more detail in this section.

3.2.1 Noise and Multipath

Multipath error is caused by the interference of a reflected signal mixing with the direct satellite signal. The level of multipath is a function of the receiver tracking technology, the antenna type, and the antenna environment.

The noise term consists of receiver measurement noise and the sum of all other unmodeled and second order effects. This is also a function of the receiver technology used. Raquet (1998) shows the code and carrier phase noise and multipath root mean squared (RMS) errors (Table 1) from sample data using a Trimble 4000 SSi receiver.

Table 1: Combined Code and Carrier Phase Noise and Multipath RMS Error Shown in Raquet (1998)

Measurement Type	RMS Error
L1 CA code	0.4 m
L2 P code	1.0 m
L1 phase	4.3 mm
L2 phase	6.2 mm

Figure 2, Figure 3, Figure 4, and Figure 5 show the combined effects of noise and multipath for L1 C/A code, and L1 carrier phase, respectively, for a high quality L1

receiver and a high sensitivity "low cost" receiver. The data was obtained in an open-road-kinematic-vehicular environment from the Final Field Study. The misclosure is the remaining error after the effect of receiver and satellite positions have been removed. See Petovello (2003a) for a detailed explanation of how the noise and multipath is calculated using GPS data.

Noise is not a spatially correlated effect; therefore, the magnitude of the errors will be the same regardless of the antenna separation. Multipath is generally considered not to be spatially correlated; however, there is some spatial correlation over extremely short (1 to 10 centimeter) inter-antenna distances (Ray 2000).

Multipath and noise are the largest error sources for short inter-antenna baselines since all of the other GPS error sources are spatially correlated and are thus removed during processing.

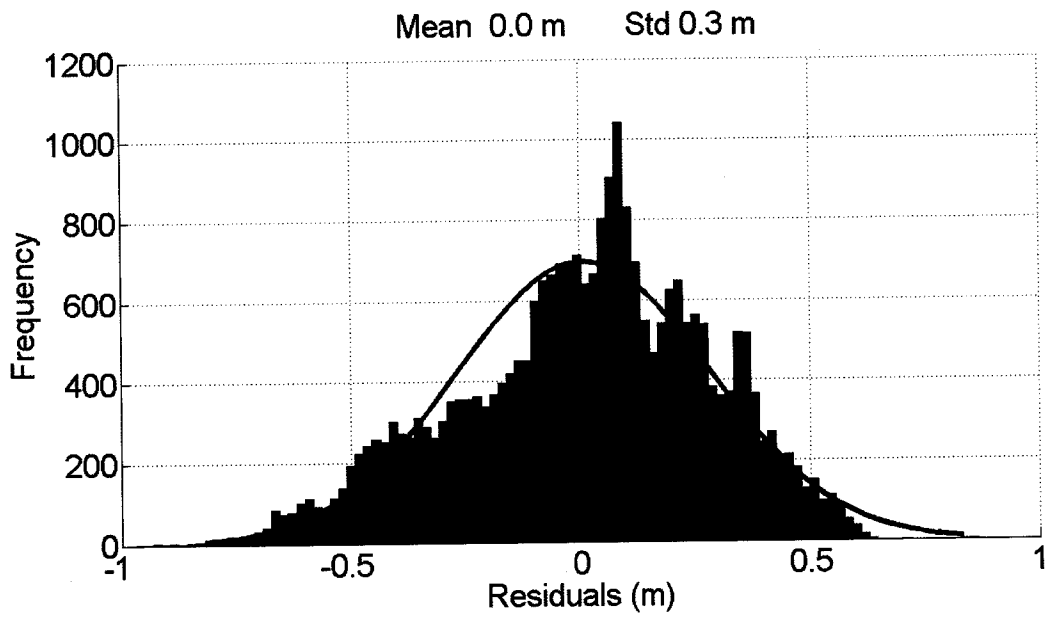

Figure 2: DD Code Residuals – Combined Effect of Noise and Multipath for L1 C/A Code for a High Quality L1 Receiver

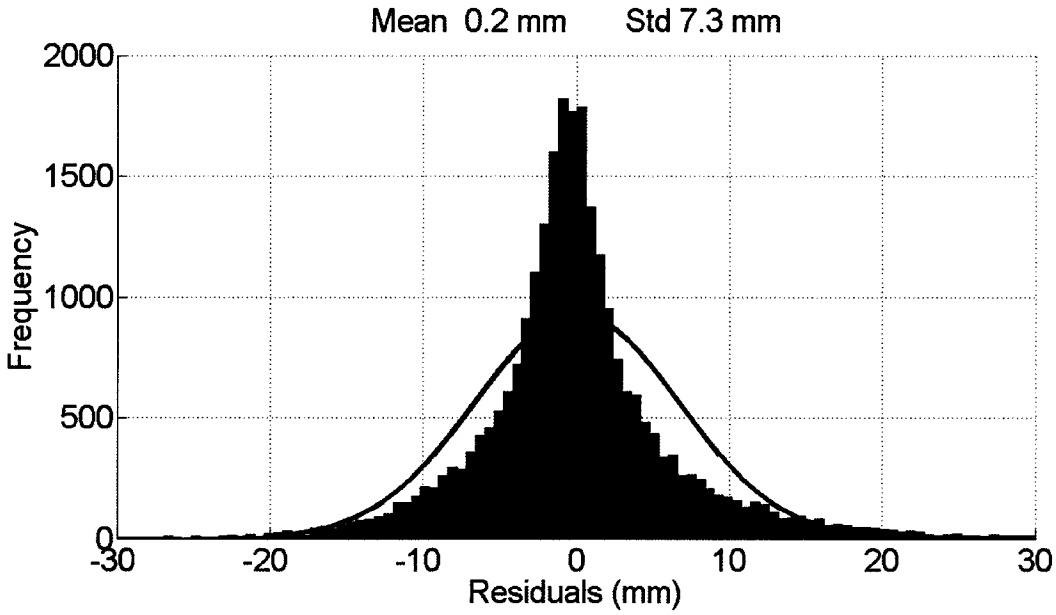

Figure 3: DD L1 Phase Residuals – Combined Effect of Noise and Multipath for L1 Carrier Phase for a High Quality L1 Receiver

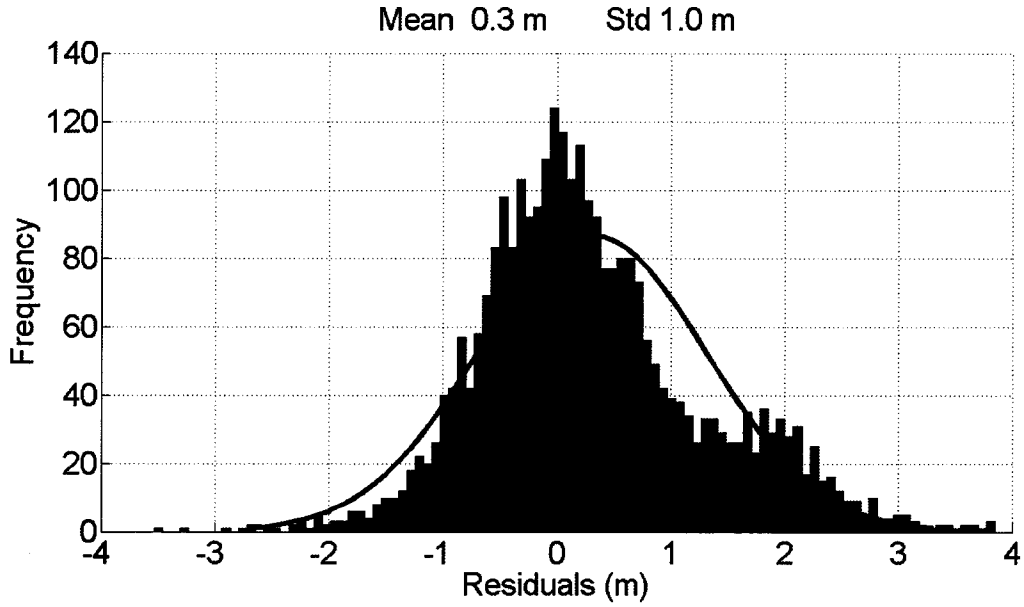

Figure 4: DD Code Residuals – Combined Effect of Noise and Multipath for L1 C/A Code for a High Sensitivity "Low Cost" Receiver

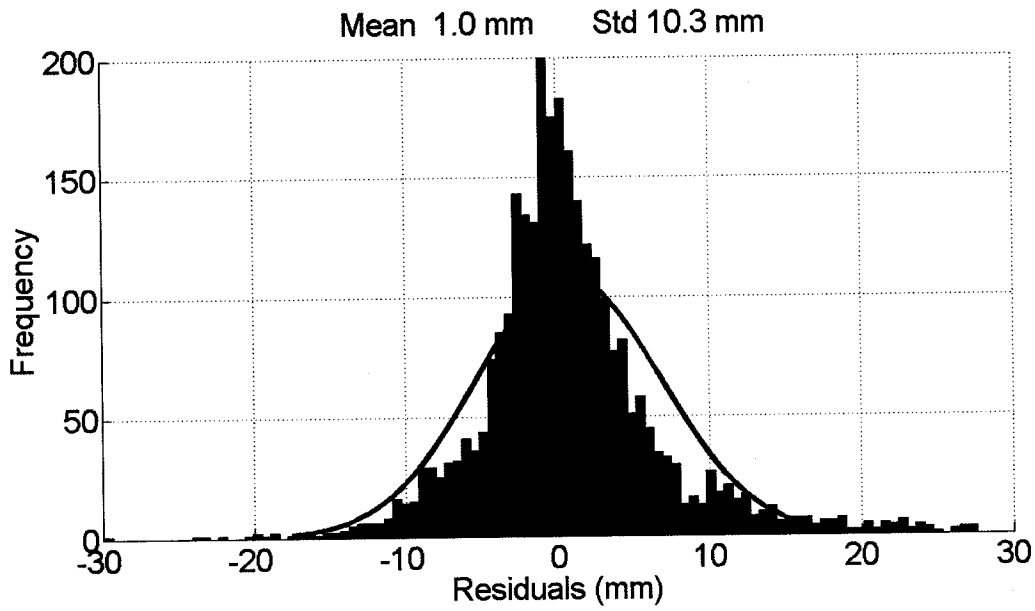

Figure 5: DD L1 Phase Residuals – Combined Effect of Noise and Multipath for L1 Phase for a High Sensitivity "Low Cost" Receiver

3.2.2 Differences in Receiver Technology

Receivers have an array of tuning parameters that can be adjusted depending on the intended use of the receiver. High sensitivity receivers are specially tuned to have high sensitivity so that they can maximize the number of satellites tracked. This can be especially advantageous in environments where there are obstructions between the receiver and satellite antennas, such as under foliage or in urban canyons. The disadvantage of high sensitivity receivers is that they are more susceptible to multipath and measurement noise.

Figure 6 and Figure 7 show the L1 phase double difference measurement noise for two u-blox LEA-4T receivers and for two NovAtel OEM4 receivers, respectively. The RMS of the two solutions (with outliers removed) is 6.6 mm for the u-blox LEA-4T receiver pairs and 4.8 mm for the NovAtel receiver pairs. These are consistent with the results shown in Table 1. The u-blox receivers have 38 percent more receiver noise as a consequence of the higher sensitivity. See MacGougan (2003) for more information about the properties and capabilities of high sensitivity receivers relative to surveying quality receivers.

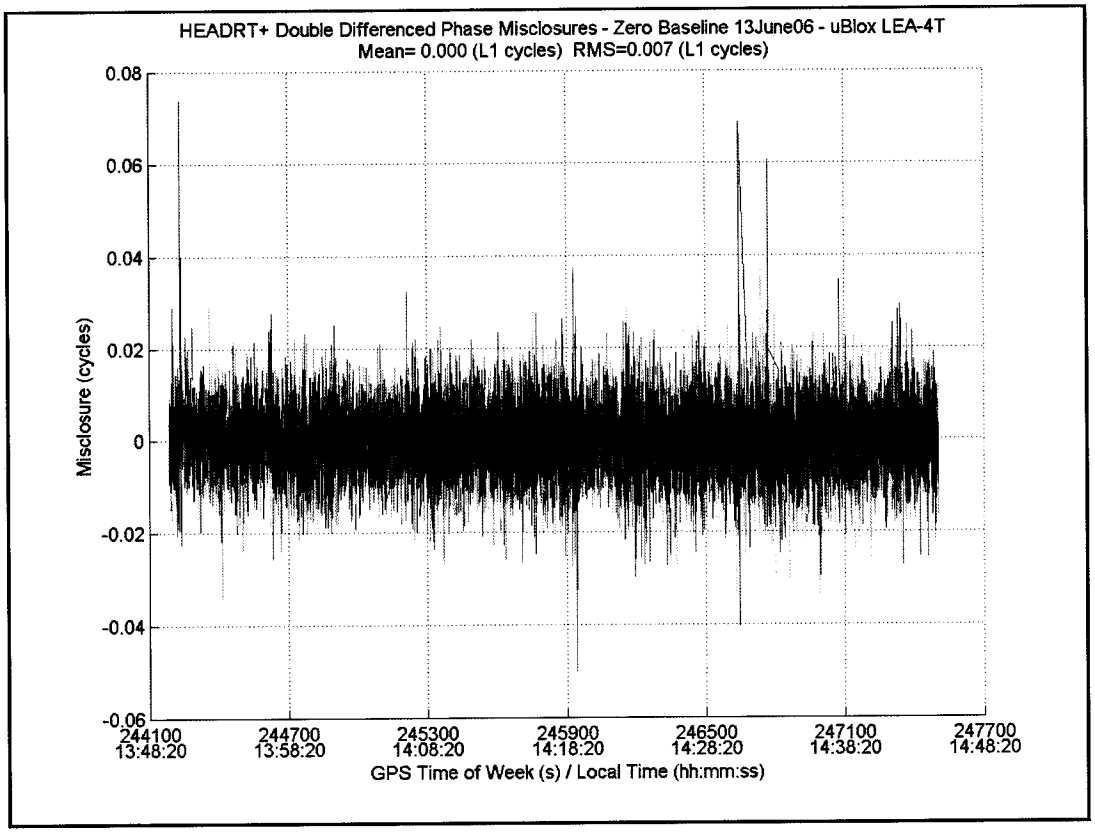

Figure 6: Zero Baseline L1 Double Difference Phase Errors for u-blox LEA-4T for June 31, 2006 (from Schleppe 2006)

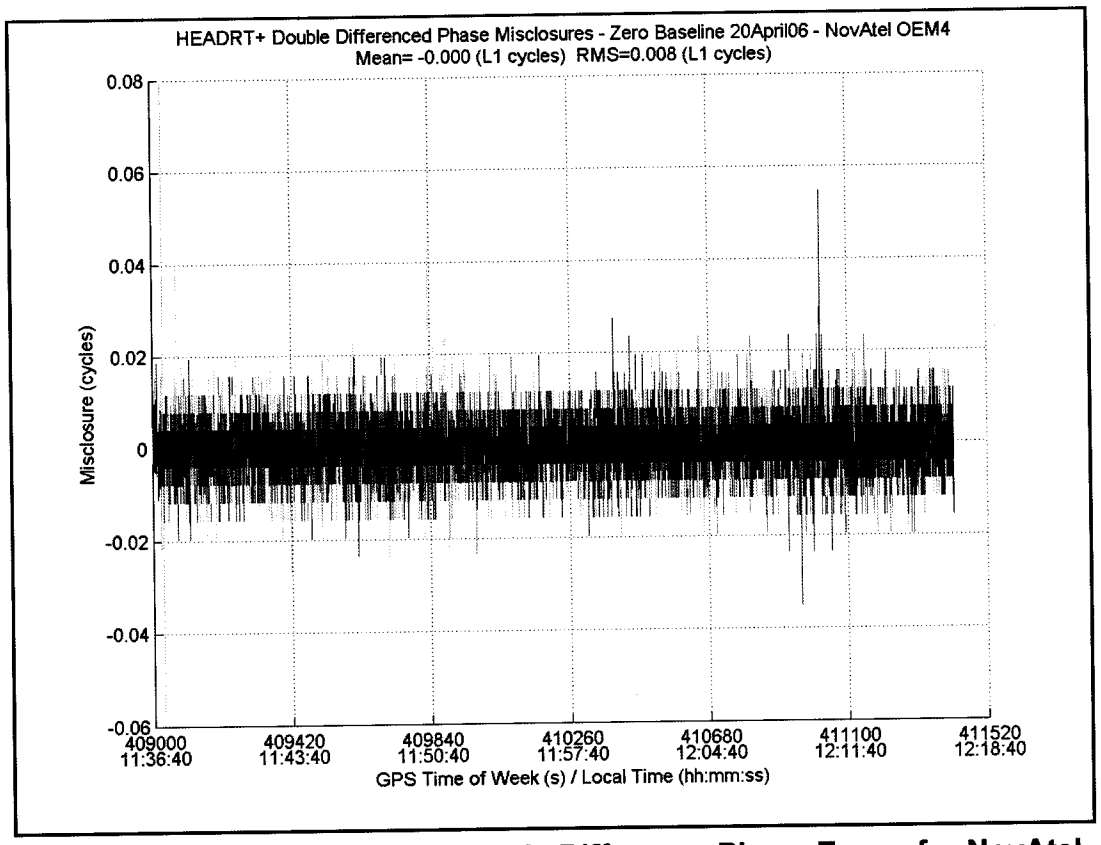

Figure 7: Zero Baseline L1 Double Difference Phase Errors for NovAtel OEM4 – April 20, 2006 (from Schleppe 2006)

3.3 Measurement Errors in V2V Navigation

Spatially correlated errors can be reduced or removed by differencing the measurements of the same satellite that is observed at two different receivers, provided the time between the observations (latency) to be differenced is short. Approximate error levels before and after between-receiver differencing are given in Table 2.

The effect of the errors discussed in this section can be applied to short distance (< 1 km) V2V-relative navigation. When the same satellite is used in the positioning calculation for each vehicle, then the differential errors shown in Table 2 are expected. Namely, decimeter level errors for correlated error sources. Multipath is the largest error source on the order of 0.5 to 1 m.

For all satellites that are not common between the two receivers, the correlated error sources (ionosphere, troposphere, satellite clock, and orbit error) will not be reduced and will, therefore, have the full effect. When one receiver tracks a satellite that the other vehicle's receiver cannot, this satellite is most likely a low elevation satellite; because low elevation satellites typically have the lowest signal strength. In addition, low elevation satellites have the greatest impact on horizontal position error. Consequently, using uncommon-view satellites in V2V applications decreases horizontal position accuracy.

Table 2: GPS Errors in Single Point and Differential Mode Assume a Baseline Length of the Order of 10s of Kilometers and a Baseline Length of 10s of Seconds (after Misra and Enge 2001)

Error Source	Error Size	Residual Differential Error
Satellite Clock	2 m	0 m
Orbit	2 m	0.1 m
Ionospheric Delay	2 – 10 m	Broadcast model 1-5 m Differential Error 0.2 m
Tropospheric Delay	2.3-2.5 m (zenith)	Models 0.1-1 m Differential Error 0.2 m
Multipath	in a "clean" environment Code 0.5 – 1 m Carrier 0.5 – 1 cm	Uncorrelated Mitigated by antenna and receiver design
Receiver Noise	Code 0.25-0.5 m Carrier 1-2 mm	Uncorrelated

4 Parameters that Impact GPS Performance

The purpose of this section is to document parameters that affect GPS performance other than the errors discussed in Section 3 above. In general, GPS performance is a function of the level of errors in the observations and the geometry of the satellites contributing to the position solution. First, these and other performance measures will be defined and then factors that affect the geometry of the satellites will be discussed in detail.

4.1 Definitions of Performance Measures

There is a wide range of performance measures used to assess GPS positioning. In general, they include system measures such as availability and reliability, solution quality indicators such as accuracy or precision, and time and statistical performance metrics such as "time to first fix." Each of these will be defined below.

Availability can have two meanings. The first is "is a position solution (meeting some performance measure) available or not." However, availability can also be used to indicate "the number of satellites available;" in other words, "tracked by the receiver" or "used in the solution." In this report a detailed description of what is being used will be described when the meaning of availability is ambiguous.

Reliability also has two meanings. The first is simply "can the system be trusted." The second is a more specific meaning in the field of high-precision positioning where reliability indicates the ability of a system to control gross errors or blunders through the detection and elimination of outlier observations. Reliability will not be discussed in this report.

Accuracy, broadly defined, is the closeness of an estimate or a group of estimates to the true value, while precision represents the closeness of a group of estimates to the mean value of the group. Several very specific measures of accuracy have been developed and are used to assess navigation systems including GPS. The most common include Circular Error Probable (CEP), Spherical Error Probable (SEP), Distance Root Mean Squared (DRMS), and Mean Radial Spherical Error (MRSE) (de Jong, et al. 2002). All of these are based on the probable density function of the position solution which can be represented by a 3x3 covariance matrix of the estimated position solution coordinates. Shown here expressed in east, north, height coordinates (E,N,h) but equally valid in any other coordinate system.

$$C_{\hat{x}} = \begin{bmatrix} \sigma_E^2 & \sigma_{EN} & \sigma_{Eh} \\ \sigma_{EN} & \sigma_N^2 & \sigma_{Nh} \\ \sigma_{Eh} & \sigma_{Nh} & \sigma_h^2 \end{bmatrix}$$

CEP is a 2-dimensional error defined as the radius of a circle containing 50 percent of the probability density of the 2-D solution. In other words, 50 percent of the time the true solution will be within a circle of radius CEP centered at the estimated solution. SEP is the 3-dimensional equivalent to this. CEP and SEP are difficult to determine theoretically and are generally obtained by assessing a large time series of solutions obtained at known points.

A more theoretical measure is DRMS, which is derived from the covariance matrix of the horizontal coordinates ([x,y] or [E,N] or [latitude, longitude]). It is the geometric mean of the estimated standard deviations of the horizontal coordinates and is obtained by taking the square-root of the diagonal elements of the covariance matrix.

$$DRMS = \sqrt{\sigma_E^2 + \sigma_N^2}\ .$$

MRSE is the 3-D analogue of DRMS and is defined as

$$DRMS = \sqrt{\sigma_E^2 + \sigma_N^2 + \sigma_h^2}\ .$$

All of these accuracy measures depend on the probability distribution of the position estimate which in turn depends on two things: The accuracy of the measurements, and the geometry of the satellites. The accuracy of the measurements depends on all of the error sources described in Section 3. The level of measurement error can be expressed by a single number called either User Equivalent Range Error (UERE) or just User Range Error (URE). If the assumption is made that all of the ranges used in a solution have the same UERE, then this value can be factored out of the error estimate. What remains is a dimensionless term that represents the satellite geometry which is called the Dilution of Precision (DOP). The definition of DOP depends also on what is being estimated leading to the definition of HDOP (horizontal) VDOP (vertical), PDOP (3d position), TDOP (time), and GDOP (Geometric, or the DOP associated with estimating 3d position and

receiver clock offset). Specifically HDOP and PDOP are related to DRMS and MRSE through UERE as follows:

$$HDOP = \frac{1}{UERE} DRMS = \frac{1}{UERE}\sqrt{\sigma_E^2 + \sigma_N^2}$$

and

$$PDOP = \frac{1}{UERE} MRSE = \frac{1}{UERE}\sqrt{\sigma_E^2 + \sigma_N^2 + \sigma_h^2}\,.$$

The advantage of using DOP to assess GPS performance is that it allows the satellite availability and geometry to be assessed independently of the ranging measurement accuracy. The DOP can then simply be multiplied by the ranging accuracy to obtain a positioning accuracy; so, for example, if a standalone or SP GPS has a ranging accuracy of 5 meters, code differential 1 meter, and fixed ambiguity carrier phase 2 cm, then in horizontal accuracy in a scenario where HDOP = 1.2 (a typical value, for example) can be assessed as 6 m, 1.2 m, and 2.4 cm without having to repeat an evaluation for each type of measurement.

4.2 Factors that Affect Dilution of Precision

Spilker, Jr., (1996a) presents an excellent overview of the design of the GPS constellation and factors that affect DOP. Many others have used DOP to assess proposed new constellations or changes to the GPS constellation. The main factors that affect DOP in open sky conditions are constellation size, time of day, and geographic location. DOP is clearly dependent on the number and location of satellites in view and will be further affected when the view of the sky is obstructed.

4.2.1 Availability of Satellites and Constellation Size

The GPS constellation, as designed, consists of 24 satellites in 6 orbital planes, with 4 satellites unequally spaced in each plane. According to Spilker, Jr., (1996a) this design was chosen to ensure continuous worldwide availability of at least 5 satellites and also to minimize the effect of a single satellite failure. The 24 satellite constellation represents a minimum level of service guaranteed by the United States government. In practice, the number of available satellites has been greater than 24 for at least the past decade. In 2000 there were 27 active satellites, and there are presently 31 active satellites. Clearly, increasing the constellation size improves performance (shown as better or smaller values of HDOP). Figure 8 and Figure 9 show the 50 percentile values of HDOP over the United States over the period of 1 day that can be obtained using the current 31 satellites and the "as-designed" 24 satellite constellation. These were computed by computing the locations of all satellites using the GPS almanac file for the week of June 21-27, 2009, and a second almanac file containing the parameters for the "as-designed" 24 satellite constellation as published in Spilker, Jr., (1996a). Looking at these figures, there appears to be very little difference, and this is true. The median HDOPs obtainable with a 24

satellite constellation and with a 31 satellite constellation are on the order of 1 to 1.5 with a slight improvement in the case of the 31 satellite constellation.

4.2.2 Time of Day

Because GPS satellites have an orbital period of 12 sidereal hours, the GPS ground tracks repeat (in theory) exactly once every sidereal day. In the time it takes the satellite to make two orbits, the Earth has rotated exactly once. Thus, the configuration of the constellation with respect to the Earth repeats once every sidereal day (approximately 23h56m). Based on this, the performance of the constellation can be evaluated by simulating a single day of operation. During a day the performance does vary. A useful way of presenting performance over a day is to evaluate a higher percentile value, for example, when HDOP is better (smaller) than a particular value 95 percent of the time. Figure 10 and Figure 11 show the 95th percentile value of HDOP over the United States for the same 2 simulations (31 and 24 satellite constellations). Clearly in this case, the 24 satellite constellation performs poorly compared to 31 satellites. What Figure 11 indicates is that with a 24 satellite constellation, on the particular day of simulation, 5 percent of the time the HDOP was greater than 10 over much of the eastern United States when using a 15 degree elevation mask. Contrast this with Figure 10 where the HDOP only exceeds 3 5 percent of the time and then only over a limited region around the Great Lakes. This indicates the importance of the additional satellites that currently form part of the GPS constellation and their role not in maintaining average performance but in insuring continuous reliable operation of the system.

4.2.3 Geographic Location

The above simulations were limited to the United States; however, numerous studies have shown that GPS coverage varies slightly with latitude and when averaged over a day but does not vary greatly with longitude. Both Spilker, Jr., (1996a) and O'Keefe, et al., (2002) show that GPS coverage, in terms of availability is best at low and high latitudes and is slightly degraded at mid-latitudes. DOP corresponds generally to availability and is generally better in the equatorial and high latitude regions; however, at high latitude locations, the HDOP is generally improved while the VDOP is degraded. This occurs because the 55-degree inclination of the orbits means that north of latitude of 55 degrees, satellites will no longer pass overhead. At the pole, the maximum elevation of a GPS satellite would be 44.7 degrees (Spilker, Jr., 1996a) making vertical position estimation difficult.

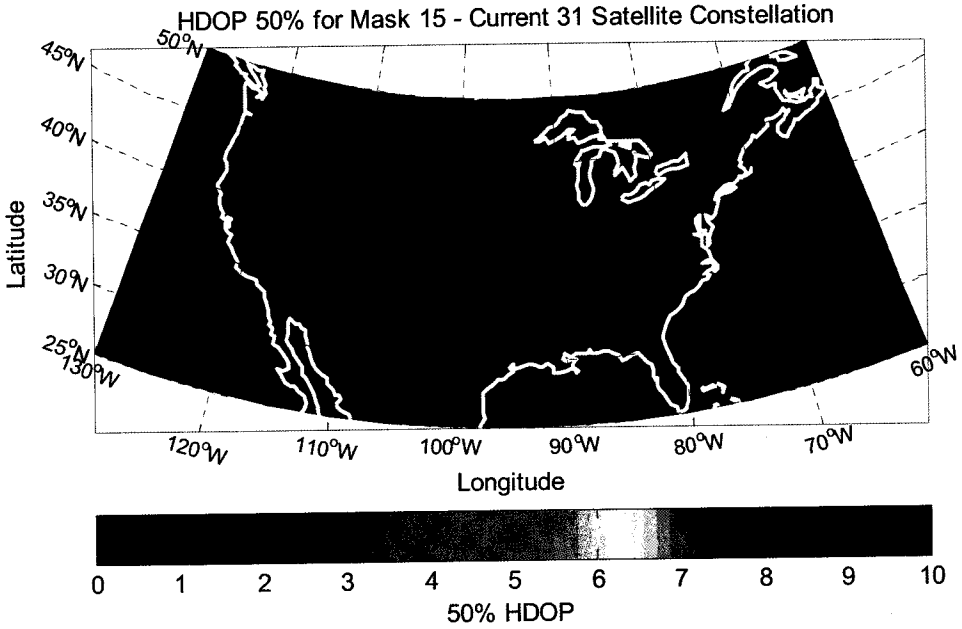

Figure 8: Median HDOP (50 Percentile Value) over the United States Computed with the Current 31-Satellite GPS Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask

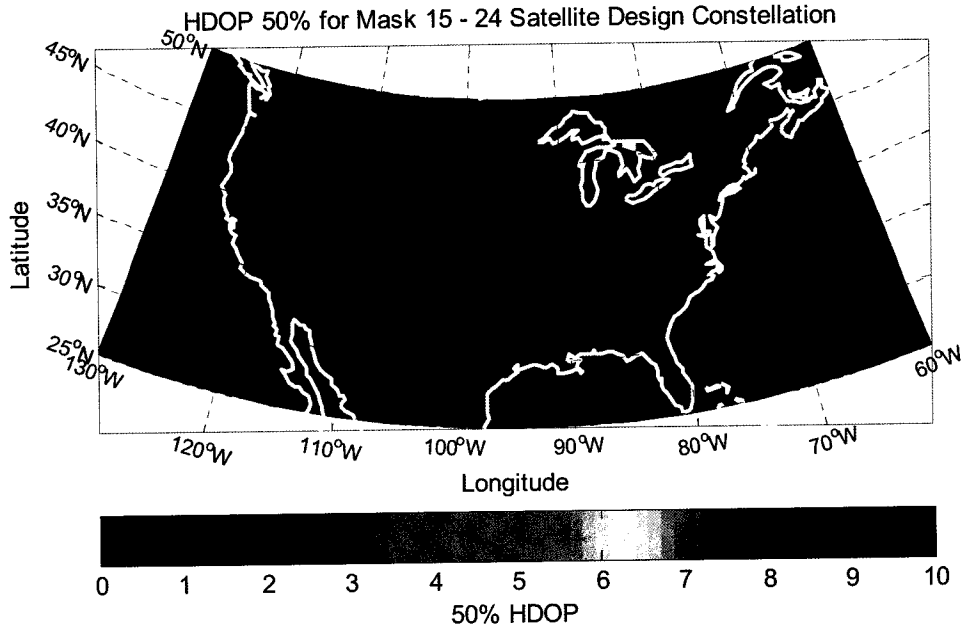

Figure 9: Median HDOP (50 Percentile Value) over the United States Computed with the "as Designed" 24-Satellite GPS Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask

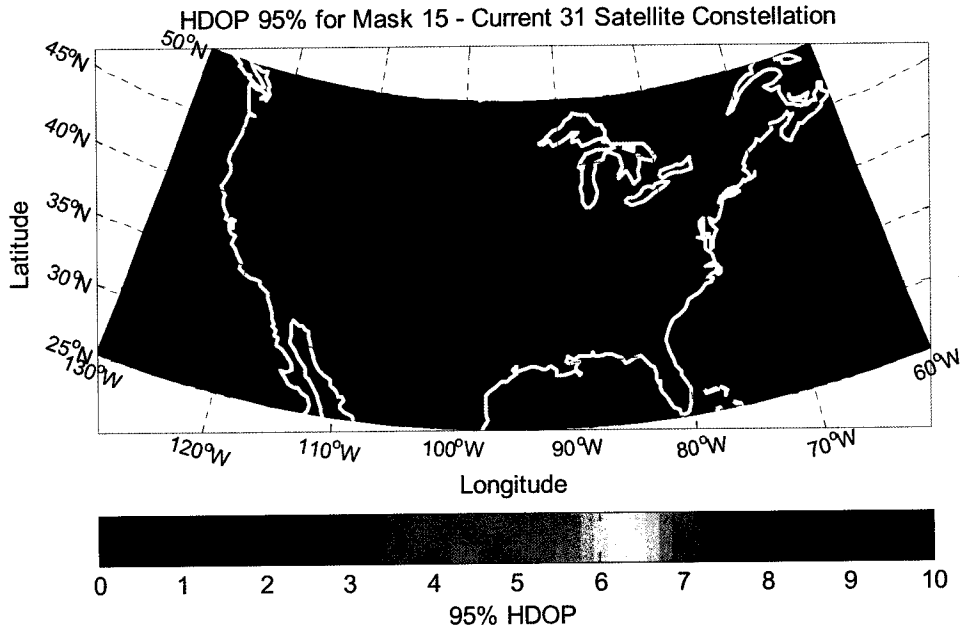

Figure 10: 95 Percentile HDOP (95 Percent of the Time the Values are Less than This) over the United States Computed with the Current 31-Satellite GPS Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask

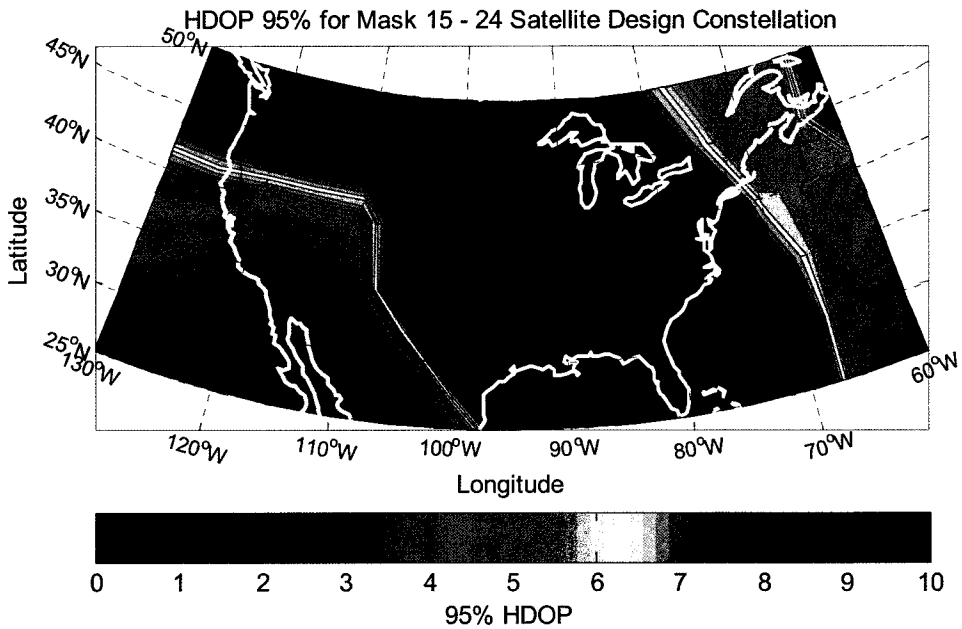

Figure 11: 95 Percentile HDOP (95 Percent of the Time the Values Are Less than This) over the United States Computed the "As-Designed" 24-Satellite Constellation Simulated on June 21, 2009, Using a 15-Degree Elevation Mask

4.2.4 Elevation Mask Angle and Obstructions

Figure 12 shows the 95th percentile value of HDOP for a 27-satellite GPS constellation with a 10 degree elevation mask. In contrast, Figure 13 shows the same when a 20-degree elevation mask is used. The degradation in the high-mid latitudes is evident in this case. In general as the elevation mask is increased, performance is decreased as satellites are excluded from the solution. Generally elevation mask values of 10 to 15 degrees are used in order to include as many satellites as possible while excluding extremely low elevation satellites that generally exhibit more multipath and ionospheric error.

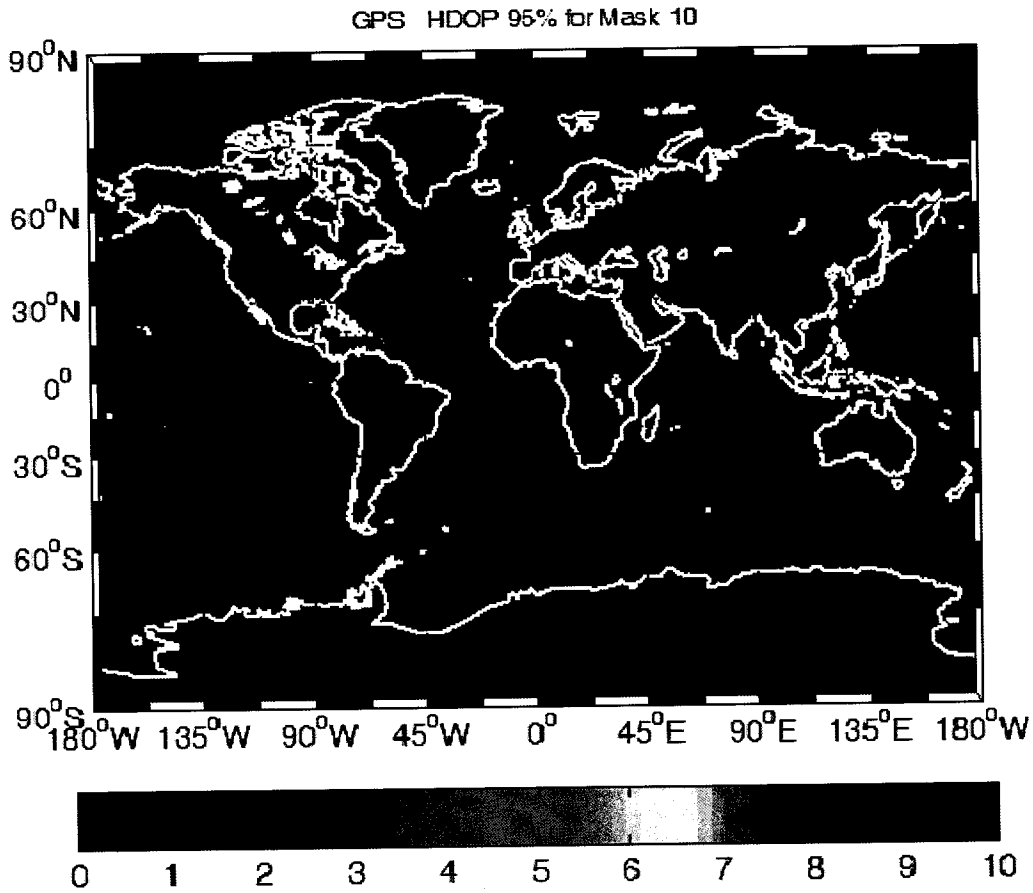

Figure 12: 95th Percentile HDOP Value for a 24-Satellite GPS Constellation and a 10-Degree Elevation Mask

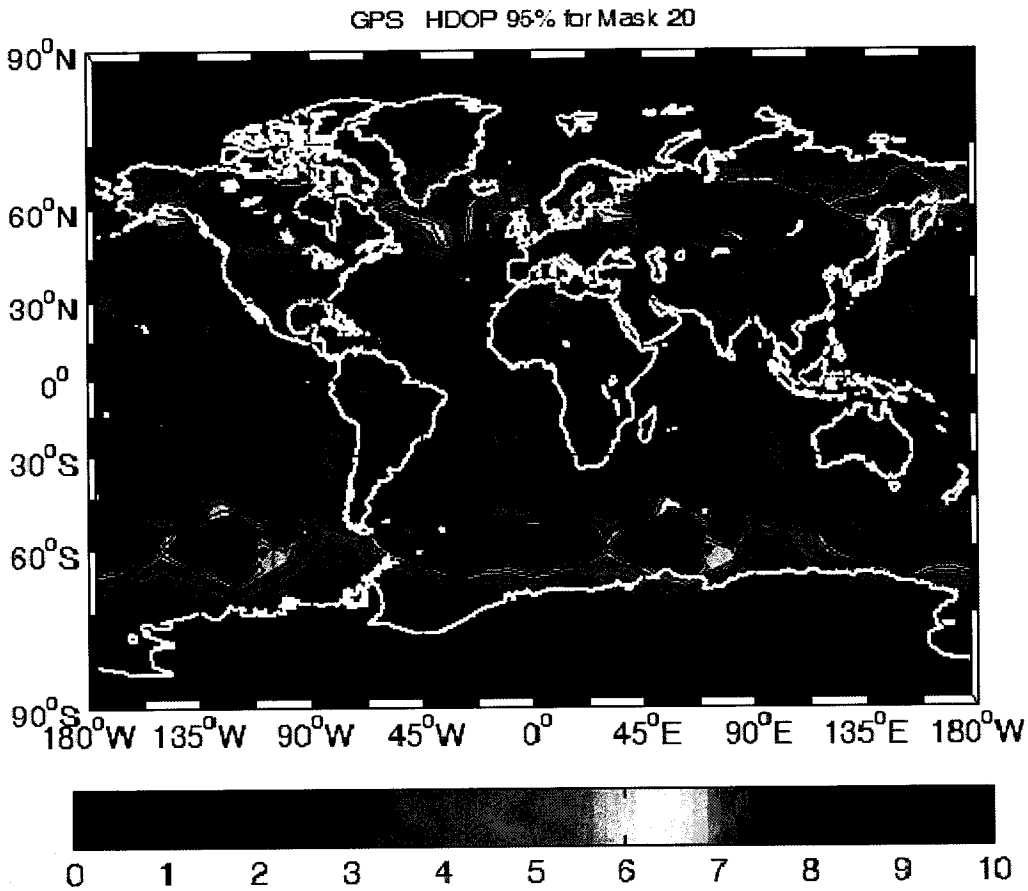

Figure 13: 95th Percentile HDOP Value for a 24-Satellite GPS Constellation and a 20-Degree Elevation Mask

5 Reference Stations and Reference Station Networks

This section describes the impact of reference stations and reference station networks on GNSS accuracy and reliability. Single reference station differential GPS (DGPS), Wide Area Augmentation System (WAAS), precise carrier phase positioning (RTK), and network-based RTK are discussed in the context of vehicle navigation.

5.1 Differential GPS

Combining the GPS error sources summarized in Table 2 a single, stand-alone GPS receiver is expected to achieve approximately 9 meters of position accuracy assuming low levels of atmospheric errors and a large number of GPS satellites well distributed throughout the sky. The global average position domain accuracy for GPS is less than 9 meters (95 percent) in the horizontal direction and 15 meters (95 percent) in the vertical (DOD 2008). The worst site position domain accuracy is 17 meters (95 percent) and 37 meters in horizontal and vertical, respectively (DOD 2008). The United States Department of Defense (DoD) is committed to maintaining these minimum levels of accuracy. These accuracies may be worse than what is commonly achievable under the

current constellation and receiver technology because these accuracies are the minimum commitment with a 24-satellite constellation.

The most significant error sources shown in Table 2 can be reduced or completely removed by using a nearby reference station to estimate the regional error sources. DGPS is the use of mainly code measurements for positioning using a reference station. DGPS position accuracy can be reduced to less than one meter (one sigma) if the reference station is 50 km from the user's antenna (Parkinson and Spilker, Jr., 1996) provided the base and user GPS devices are of sufficient quality. This accuracy decreases steadily as the distance between the reference station and the user increases.

5.2 WAAS

The WAAS is a GPS correction estimation and distribution system. The Federal Aviation Administration (FAA) administers the program, which has been available in initial operational capability mode for aviation use since July 10, 2003 (FAA 2007). GPS measurement errors are monitored for WAAS corrections using a reference station network distributed throughout the Continental U.S. with additional stations in Alaska, Canada, and Mexico. The correction messages for WAAS are distributed through geostationary satellites.

WAAS has many services including integrity monitoring and error modeling. The geostationary satellites are themselves additional ranging satellites which provide further satellite availability along with the other data services.

The integrity messages notify users within 6 to 8 seconds if a satellite's ranging data becomes unstable or unusable (FAA 2007). This is a critical feature for many applications because it allows for protection levels of GPS service. If the GPS ranging errors exceed the protection levels, then the integrity messages alert the applications that there may be a problem.

The WAAS network measures and monitors the clock, orbit, and ionosphere errors throughout the network and provides corrections for these error sources. The clock corrections are transmitted at a high rate and a slow rate for different applications. Yousuf (2005) shows that the RMS WAAS clock corrections are 31 to 44 percent better than the broadcast clock model, and the RMS error of the satellite positions is 13 to 35 percent better than the broadcast positions. This study did show significant outliers in the WAAS corrections due to satellites entering the view of the network. These outliers could exceed 50 meters in error for some cases.

Yousuf (2005) concludes that the accuracy of the WAAS ionosphere correction is 2 to 3 meters during low ionosphere periods; however, it can be much larger during ionosphere storms. The ionosphere errors after applying the WAAS ionosphere model can be significantly larger than the usual 2 to 3 meter error levels during ionosphere storms. The corresponding integrity information responds well to the event suggesting that users would be properly warned about the reduced accuracy of the corrections.

In terms of the position accuracy, with dual frequency International GNSS Service (IGS), Continuously Operating Reference Station (CORS), and WAAS reference stations as rovers, the RMS position error when using WAAS was 0.7 to 1.1 meters (95 percent) in

the horizontal and 1.0 to 1.7 meters (95 percent) in the vertical (FAA 2004) Yousuf 2005). This is roughly the same level of accuracy as shown for DGPS. FAA (2005) shows a WAAS-enabled 62 to 76 percent improvement in the horizontal position error and 79 to 88 percent improvement in the vertical position errors relative to standalone GPS using WAAS-certified user equipment.

5.3 Local Base Stations

When a reference station is within a few kilometers to a user, then carrier phase methods can be used to further improve the positioning performance. Carrier phase measurements are significantly more precise than code measurements, and they are much less susceptible to multipath. Unfortunately, carrier phase measurements have an integer wavelength biases (ambiguities) that must be estimated and removed. These biases are commonly referred to as carrier phase ambiguities.

Carrier phase ambiguities are tracking biases induced by the measurement process of the carrier phase of the GPS signal. These biases have integer values (in cycles). They are random values that change with each loss of lock of the tracked signal. In other words, each satellite is assigned a random integer ambiguity that will be different every time the signal is tracked.

These biases must be estimated before the signal can be used effectively, and if possible, the ambiguities should be constrained (fixed) to their true integer values. This fixed ambiguity case will give the best possible performance when using carrier phase measurements.

There are many methods to determine the correct integer ambiguities; however, they are not relevant to this discussion. For more information on ambiguity resolution methods see Erickson (1992), Chen and Lachapelle (1995), Teunissen (1994), and Jong and Tiberius (1996).

The ability to determine the integer values of the carrier phase ambiguities is limited by the errors in the carrier phase measurements. When the reference station and user are less than 5 kilometers apart, then the carrier phase ambiguities can be reliably determined. Under these conditions the precision of the positioning solution is mainly limited by the carrier phase noise and multipath. Positioning accuracies of a few centimeters is achievable at these short distances.

As the distance between the reference station and user increases, the spatially correlated errors also gradually increase. When the errors become too large to reliably determine the carrier phase ambiguities, a less precise float ambiguity can be estimated. This distance varies depending on the magnitude of the ionosphere and troposphere errors. Under typical conditions, this distance is on the order of 20 to 30 km.

When the distance between the user and reference station increase further, the level of ionosphere error increases. The increase in ionosphere error decreases the accuracy of the float carrier phase ambiguity estimates. At this distance, a combination between the two GPS measurement frequencies can be used to completely remove the effect of the ionosphere. This ionosphere-free measurement can be used to calculate the user's

position with decimeter-level accuracy. The level of accuracy continues to increase further as the distance between the user and the reference station increases.

The ionosphere-free measurement cannot be created with single frequency receivers. In general, all of the reference station distances discussed above are slightly shorter when using only single frequency receivers instead of full dual frequency receivers.

5.4 Network-based RTK

A multiple reference station RTK is a complex, yet natural extension of single reference station RTK. A single reference station RTK actively and dynamically measures GNSS measurement errors. These measurement errors can be characterized based on spatial correlation of the error sources. The following measurement error sources are spatially correlated over tens of kilometers or more and can be useful to reduce the measurements of other GNSS users: satellite clock, satellite orbit, troposphere, and ionosphere errors.

In a single reference station RTK, the errors are assumed to be constant everywhere around the reference station. In reality, however, the quality of these error estimates degrade as a function of distance and can reach an unacceptable level for ambiguity resolution after tens of kilometers. One approach to ensure an acceptable level of measurement error over a wide geographic region is to deploy many reference stations, each operating independently. Once this infrastructure is in place, users select the reference station that will provide them with the greatest reduction of measurement errors and use the corresponding corrections in the traditional single reference station RTK approach. Unfortunately, the decision as to which reference station to use can be problematic especially when the user is located between nearby and equally spaced reference stations. The estimated measurement errors at each of the reference stations may be different, but the user is forced to discretely choose one or the other.

The solution to this problem is a multiple reference station RTK. Instead of discretely choosing the solution from one reference station or another, the multiple reference station solution allows users to combine the estimated measurement errors at each of the reference stations and smoothly transition from the errors at one reference station to another.

The multiple reference station solution is not only better because of the ease of use when transitioning between reference stations but also because the smooth combined solution is more likely to represent the user-observed measurement errors providing an even further reduction of user-measurement errors relative to the single reference station case.

An example of multiple reference station processing is illustrated in Figure 14. The red line represents the changing errors as a function of receiver location. There are two reference stations in this example. A blue reference station is located at -2, and a green reference station is located at +2 while the user in this example is located at 0 in the middle of the network. If the user were to use only the green reference station, then the residual error experienced by the user is shown by the vertical green line. Alternatively, if the user were to use the blue reference station then the error experienced by the user is represented by the blue line. When both reference stations are used, then a combined interpolated solution is shown in black and the corresponding residual error is shown as the vertical black line.