AN ABSTRACT OF THE THESIS OF

Sean M. Mayes for the degree of Honors Baccalaureate of Science in Electrical and Computer Engineering presented on May 31, 2012. Title: Triple GPS Localization.

Abstract approved: ____

Huaping Liu

The purpose of this thesis is to increase the positional accuracy of Global Position System (GPS) modules using an artificial intelligence algorithm. Three basic and identical GPS modules were setup in an equilateral triangle formation with side lengths of one meter. The triangle was placed in four separate locations where approximately 700 GPS data points were collected per GPS module. The data was then analyzed and tested using a Gaussian distribution. In the best-case scenario, the Triple GPS Localization algorithm was able to improve the localization accuracy by 82%. This shows that utilizing a machine-learning algorithm can improve the positional accuracy of any GPS module.

Key Words: Triple GPS, Localization, TGPSL, DGPS, GPS, Artificial Intelligence, AI, Machine Learning, Normal, Gaussian, Distribution, Global Positioning System, Accuracy, Precision, Algorithm, Relative.

Corresponding e-mail address: SeanMayes7@gmail.com or Mayess@onid.orst.edu

Triple GPS Localization

By

Sean M. Mayes

A PROJECT

Submitted to

Oregon State University

University Honors College

in partial fulfillment of

the requirements for the

degree of

Honors Baccalaureate of Science in Electrical and Computer Engineering

Presented May 31, 2012

Commencement June 2012

Honors Baccalaureate of Science in Electrical and Computer Engineering of Sean M. Mayes presented on May 31, 2012.

APPROVED:

____

Mentor, representing Electrical and Computer Engineering

____

Committee Member, representing Computer Science

____

Committee Member, representing Electrical and Computer Engineering

____

Chair, Department of Electrical and Computer Engineering

____

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

____

Sean M. Mayes, Author

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF EQUATIONS & TABLES

Chapter 1:

Introduction

Accurate and precise positioning is needed in many fields of research. Currently many scientists use expensive surveying equipment such as Differential GPS (DGPS) modules to accurately find a position down to 10's of centimeters.[1] However, DGPS modules are not only expensive, they require communication with a pre-established base station. If however, a pre-established base station is not available, it can be very expensive and time consuming to set one up. An alternative is using an expensive stand-alone GPS module that is capable of locating your position to within a meter. Military grade GPS modules such as the Defense Advanced GPS Receiver (DAGR) cost around $2,000 and is only accurate to within a few meters using the DGPS capabilities[2].

Triple GPS Localization (TGPSL) is a quickly setup independent system that is easy to use, and low cost. In addition, this method will improve the accuracy using any set of three GPS modules. Triple GPS Localization consists of mathematically calculating the probability that the GPS reading will be within a particular radius from the actual position of the GPS module, utilizing three separate GPS modules setup in an equilateral triangle formation.

$$P(\,Radius < distance \mid three\ GPS\ modules, setup\ in\ an\ equilateral\ triangle\ formation\,)$$

Equation 1.1: Desired probability

The purpose of TGPSL is to provide an alternate method of increasing GPS positional accuracy. The work performed consisted of collecting real-time GPS data, applying a Gaussian machine-learned algorithm to this data, and creating a virtual triangle that adjusts itself according to a pre-determined utility.

The project utilizes three standard GPS modules arranged so that the modules form the corners of an equilateral triangle with side lengths of one meter. By utilizing the method described above, it is possible to increase the localization accuracy by approximately 82%.

In the next few chapters, the overall concept will be described as well as the general workings of current GPS satellites. The background covers how GPS accuracy is measured and introduces some typical problems associated with the GPS system. The theory chapter describes what the Gaussian distribution is and how it relates to GPS modules. In addition, the theory chapter shows how utilizing three GPS modules alone can increase the probability of your positional accuracy. The final sections reviews in depth detail of how Triple GPS Localization (TGPSL) works.

Chapter 2:

Background

Global Positioning System satellites continuously transmit location and current time signals to earth. GPS receivers do not transmit any information themselves, as they are simply receivers. GPS receivers require a direct line of sight to the satellite. Therefore, they do not work well indoors, near tall buildings, or under a dense forest canopy. Since the whole GPS operation depends on extremely accurate timing, every GPS satellite is equipped with an atomic clock.[3]

All GPS satellites are synchronized in order to transmit the signal at the exact same instant. The GPS receiver can then determine its relative position based on the difference in timing from at least 3 satellites (4 satellites for 3 dimensions). Because GPS receivers do not have a synchronized atomic clock, their positional information can be slightly off.[3]

When the whole GPS system was initially created, timing errors were inserted into GPS transmissions to limit the accuracy of non-military GPS receivers. The timing errors reduce the GPS accuracy to about 100 meters.[3] These timing errors were part of the operation known as Selective Availability. On May 1st, 2000, President Bill Clinton ordered the Selective Availability to be shut down.[4] Since then, GPS modules have been continually improving their accuracy with no hindrance from the government.

Current GPS modules range in accuracy from 20 meters radius down to only a few centimeters.[5] Today, it is relatively cheap to get GPS modules with accuracy between 3 and 10 meters radius. The GPS modules used for testing the TGPSL algorithm have an accuracy of 2.5m Circular Error Probability (CEP).

GPS accuracy is usually given as the modules Circular Error Probability. CEP means that the GPS readings will be within a radius of 2.5 meters of the GPS antenna 50% of the time. It also means that the GPS readings will be within a radius of 5 meters (twice the CEP radius)

95% of the time. CEP is not the same measurement as Root-Mean-Square however, through some calculation you can convert from one to the other.[6] The CEP has a similar concept to one standard deviation in a Gaussian distribution and follows the three-sigma rule.

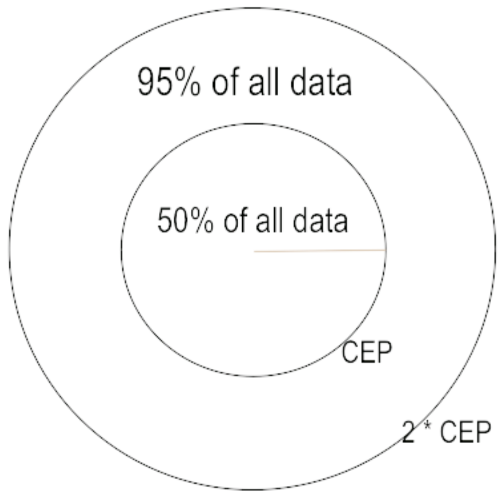

Figure 2.1: CEP (Circular Error Probability)

The three-sigma rule states that for a standard Gaussian distribution, 99.73% of the values will lie within 3 standard deviations. 68.27% of the values will lie within one standard deviation of the mean; 95.45% of the values will lie within 2 standard deviations of the mean.[7]

The most prominent and reliable GPS systems today require the use of a base station. A base station knows its exact location. Therefore, it is able to determine satellite signal errors, which are described later. The base station calculates the range from each satellite, and compares the results to the actual range calculated from its known location. The base station is then able to send the corrected range information to GPS receivers over FM broadcast band, by satellites, or by beacon transmitters that are maintained by the U.S. Coast Guard.[8]

Differential GPS systems are a well-established way to improve the accuracy of GPS modules. Depending on the distance from the base station, DGPS modules can process the correction data in real time or by post-processing. Most GPS modules sold today have the ability to utilize DGPS, WAAS (Wide Area Augmentation System), or MSAS (Multi-Functional Satellite Augmentation System). Although, all these systems require the use of a base station, they can greatly increase your positional accuracy.

Despite using the methods mentioned above or the TGPSL system, all GPS modules have several sources of error. The use of a base station helps eliminate the Ionosphere and Troposphere errors, which slow down the satellites signal as it passes through the atmosphere. Signal multi-path errors occur when a signal is reflected off of objects such as a tall building, or large mountains, which causes the signal travel time to increase. To add to these errors, a low number of visible satellites caused by large surrounding objects or bad satellite geometry when the satellites are located in a straight line or are closely grouped together can significantly reduce the accuracy.[9]

Chapter 3:

Theory

The Gaussian (or Normal) Probability Density Function (PDF) describes the relative likelihood for a GPS reading to occur in a given location. The Gaussian distribution will be used to determine if the data points collected tend to cluster around a single point. Ideally, the GPS data points will gather around the actual location of the GPS module.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Equation 3.1: Gaussian distribution

In addition to the Gaussian distribution, it is interesting to note that in the method that will be implemented, the farther the GPS modules are apart, the more accurate the actual position may be. As well as the farther apart the GPS readings, the more accurate the actual position will be.

It makes sense that the farther the GPS modules are apart the greater the accuracy of the actual position. However, it is an interesting concept that the farther the GPS readings are from the actual position, the more accurate the TGPSL actual position will be. (See figure 3.1) This is due to the CEP of GPS modules. 95% of the GPS reading will fall within 2*CEP. Therefore, if the GPS reading is far away from the actual GPS module, and the CEP radius around each GPS reading is drawn, the actual position (where all three circles intersect) has a higher probability than if the GPS readings were close together. When the GPS readings are close together, the CEP circles have too much overlap.

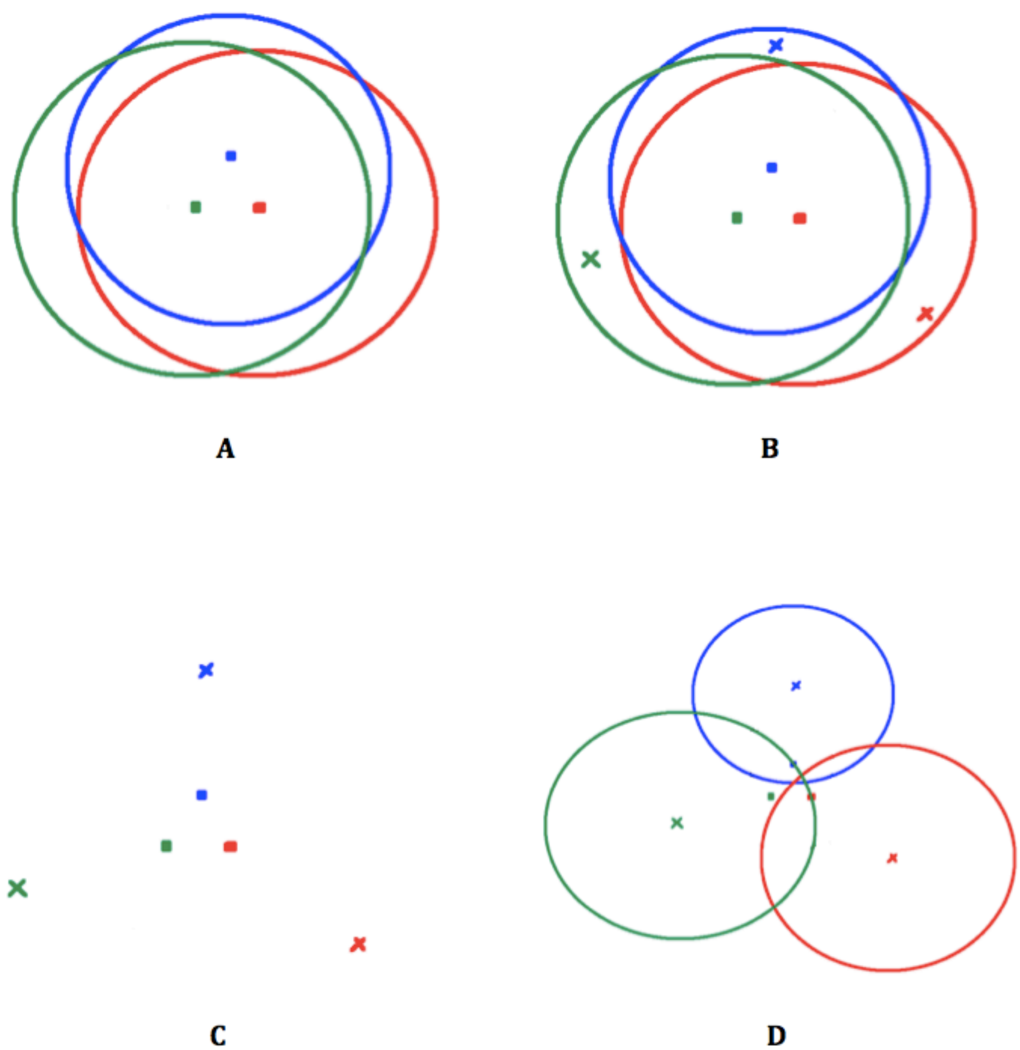

Figure 3.1: **A)** Three GPS modules mounted in a triangular formation. The 3 circles represent the CEP (Circular Error Probability). **B)** Each "x" represents one possible GPS reading (longitude and latitude). **C)** The only information known for the TGPSL algorithm. **D)** The CEP circle is drawn around each GPS reading. When the GPS readings are farther apart (or around 2 to 3 standard deviations away); the actual position (intersection of the circles) becomes more accurate.

The Venus GPS modules were used for this proposed system. They are the most accurate GPS modules found for a relatively low price ($49.95). These GPS modules have a Circular Error Probability (CEP) of 2.5 meters. Utilizing the CEP and the fact that TGPSL requires three GPS modules spaced one meter apart, we can see in figure 3.2 that by combining all three GPS module outputs and using basic rules of probability, TGPSL will obtain a location that is within a 4-meter radius, greater than 95% of the time.

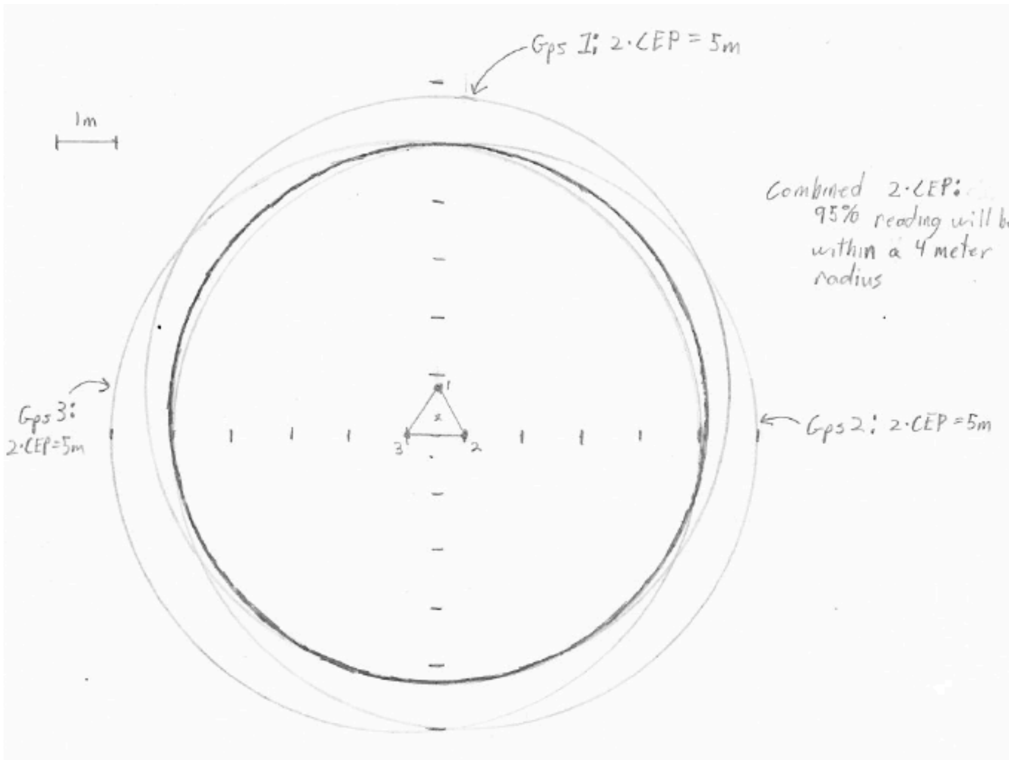

Figure 3.2: This figure represents the location of the three GPS modules mounted on the corners of an equilateral triangle with side lengths of one meter. The GPS circles represent a distance of 2*CEP (which is 5 meters for the Venus GPS modules). The darker circle represents the combined GPS module probability – The probability for this area (within a radius of 4 meters) is greater than 95% when simply taking the average of the three readings.

The goal is to decrease the radius and increase the accuracy. To accomplish this, a supervised-learning technique is used as well as artificial intelligence and machine-learning algorithms. In order to accomplish this task, a great amount of GPS data needed to be collected.

Chapter 4:

Process

4.1 Collection Process

In order to improve the accuracy of GPS modules, a great amount of GPS data needed to be collected. This data was used to find both the true location of the GPS modules and to test the algorithm. How the true location was obtained is described in section 4.4 The Algorithm.

The platform used for data collection was arranged such that the three GPS modules formed the corners of an equilateral triangle with side lengths of one meter. There were a couple reasons for choosing a one-meter equilateral triangle as tje foundation. First off, in order to achieve minimum CEP overlap the GPS modules cannot be on top of each other (figure 3.2 shows CEP overlap). The initial idea was to space the GPS modules out according to their CEP (2.5m for the Venus GPS modules). However, this thesis was apart of a greater project known as the Platform for Autonomous Multi-Vehicle Communication and Coordination (PAM-VCC). Essentially, the triangle platform needed to fit on top of a standard size car. Therefore, a one-meter equilateral triangle was the largest separation allowed that met the criteria. All GPS modules were connected directly to one micro-controller that stored all the necessary GPS data.

The collection process consisted of moving this platform to four separate locations; Alpha, Bravo, Charlie, and Delta. The four locations formed a square with side length of approximately 25 meters. The reason for collecting data from four separate locations was to verify that the algorithm in question works by applying it to all four locations independently. In addition, the 25m square allowed for a quick qualitative verification that the data collect was indeed correct. The spacing of 25 meters gave the GPS modules enough room so that their readings would not overlap. In order to orientate the platform with its actual location on earth and relative to the other locations, it was important that the square be as exact as possible and the orientation of the platform remain the same at all four corners. Maintaining the same orientation for all four locations both assisted the algorithm (discussed in section 4.4) in knowing which direction the virtual triangle should point and assisted in the qualitative verification of the graphs produced. The orientation of the platform was arbitrarily set to magnetic north with the center of the equilateral triangle placed at each corner of the square (see figure 4.1.1).

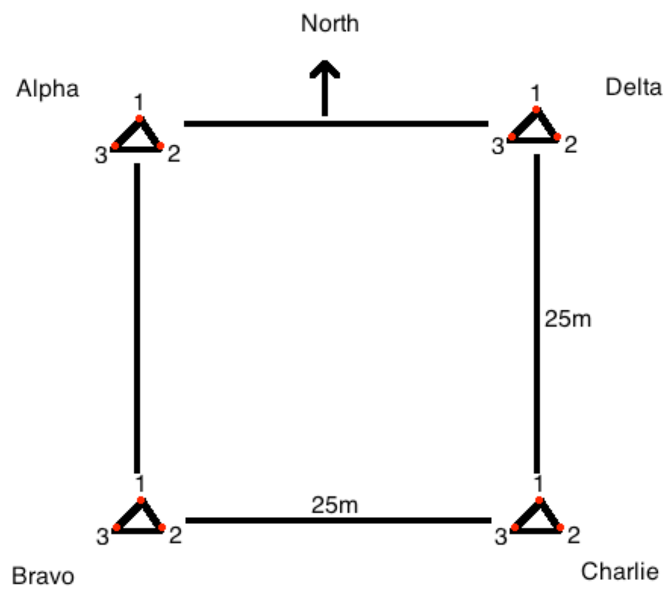

Figure 4.1.1: This figure represents the physical orientation of the GPS modules and GPS platform during the data collection process.

At each location, all three GPS module collected approximately 670 data points each – totaling about 2,000 data point per location. Each GPS module's sampling rate was set to 1Hz, giving the GPS modules one full second between collected samples. The GPS collected longitude, latitude, and time data. With this information, the artificial intelligence agent can get one coordinate from each GPS module with roughly the same time stamp. The TGPSL algorithm will then determine an accurate position based on only three GPS readings (one from each GPS module) and [for testing] compare this location to the actual location of the platform.

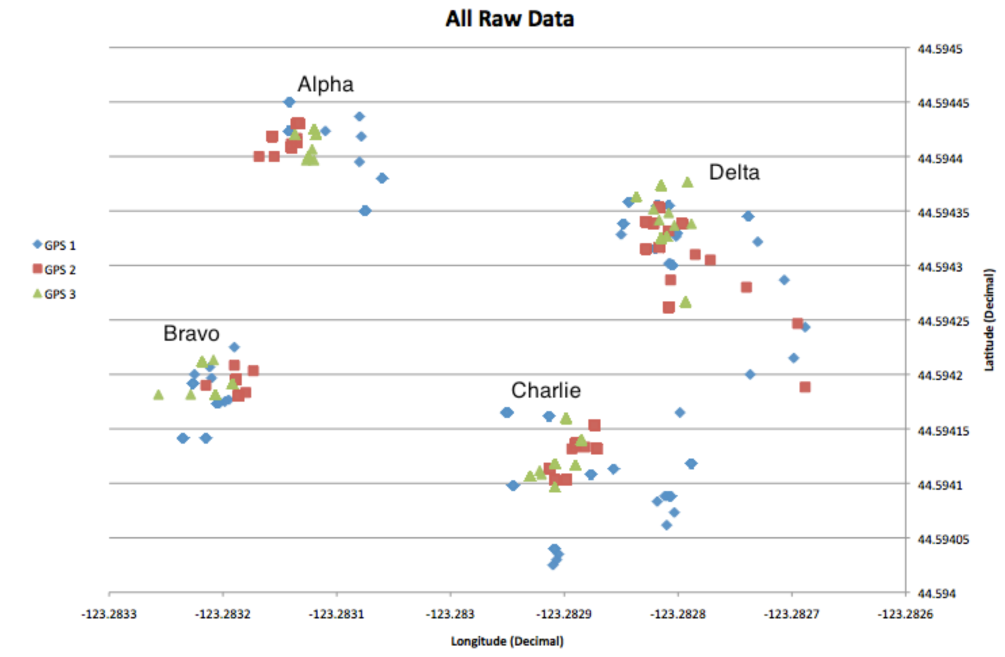

Figure 4.1.2: This figure represents all the data collected. The 25m square was oriented magnetic north, which has a declination of approximately 15 degrees in Corvallis Oregon. True North is perpendicular to the latitude lines.

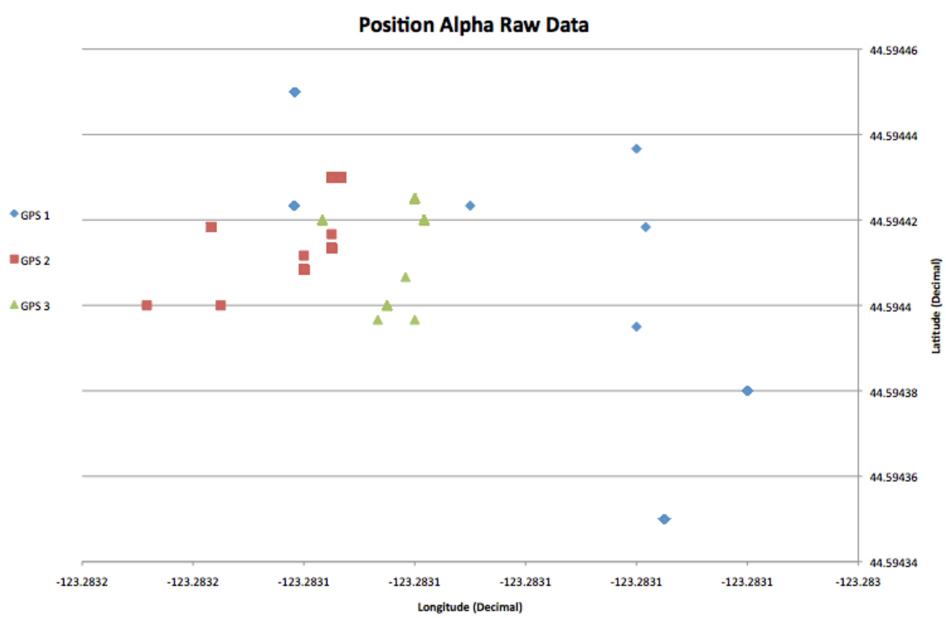

Figure 4.1.3: Alpha: GPS 1 had 670 samples. GPS 2 & 3 had 672 samples.

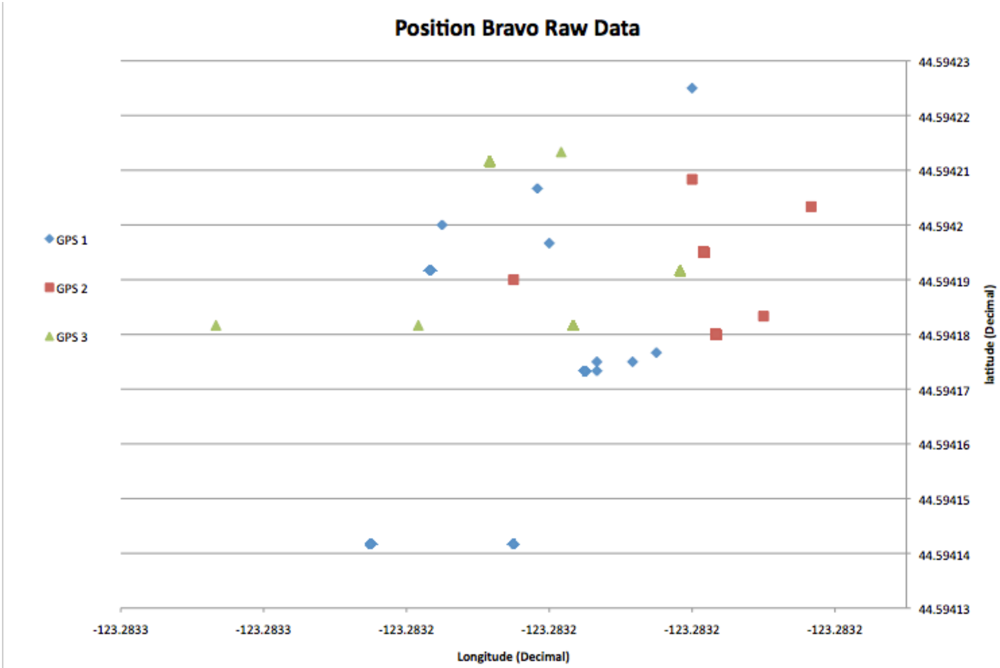

Figure 4.1.4: Bravo: There were 648 samples from each GPS module.

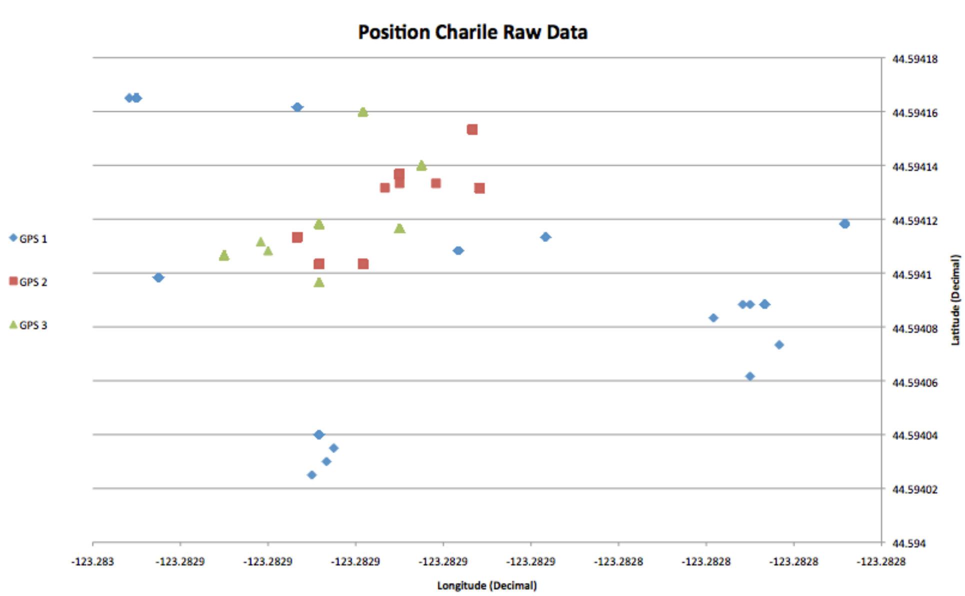

Figure 4.1.5: Charlie: GPS 1 had 726 samples. GPS 2 & 3 had 728 samples

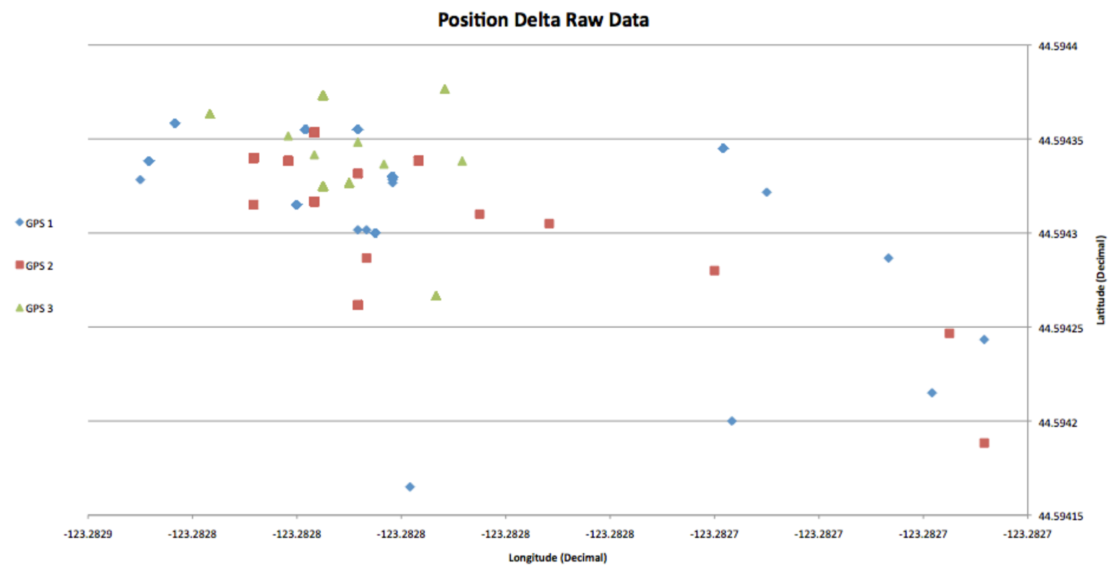

Figure 4.1.6: Delta: There were 696 Sample from each GPS module.

Process

4.2 Proof of Gaussian Distribution

In order to use the Gaussian (or Normal) distribution equations, the data collected needed to fit a Gaussian distribution. However, the lack of dispersion between GPS readings made this very difficult. Each GPS module collected about 672 data points. Although, the GPS modules were set to a 1Hz-sampling rate, the micro-controller is not multi-threaded; therefore, the micro-controller sampled every third reading – about 0.33Hz. A slower sampling rate turned out to be better than faster sampling because the majority of the GPS readings were exactly the same. Out of 672 data points, there were only about 5 to 12 different positions. This made proving that the GPS modules produce a Gaussian distribution much more challenging.

With such little variation in data points, the normality test chosen is known as a Q-Q Plot. The Q-Q plot is a way to graphically compare the distribution of a given data set to the Gaussian distribution. Along one axis is the normalized perfect standard Gaussian distribution based off the number of GPS readings collected. This was accomplished using Microsoft Excel's "NORMSINV" function. Along the other axis the actual GPS readings are sorted from minimum to maximum values and then normalized.

Ideally, if the data collected fits a perfect Gaussian distribution, the graph will produce a straight line that is equal to the $x = y$ line. Therefore, if the data falls along the $x = y$ line, it is a good indicator that the data is Gaussian distributed.

However, the data consisted of multiples of the same point. In order to determine if the data fit the Gaussian distribution, a linear trend-line was applied to the data. It is likely that the more data points collected, the trend-line would not be necessary. Therefore, the trend-line essentially represents the line produced if approximately 50,000 or more GPS data points were collected per GPS module. 50,000 data points would give a much smoother line as opposed to the 672 data points that give a step-like graph.

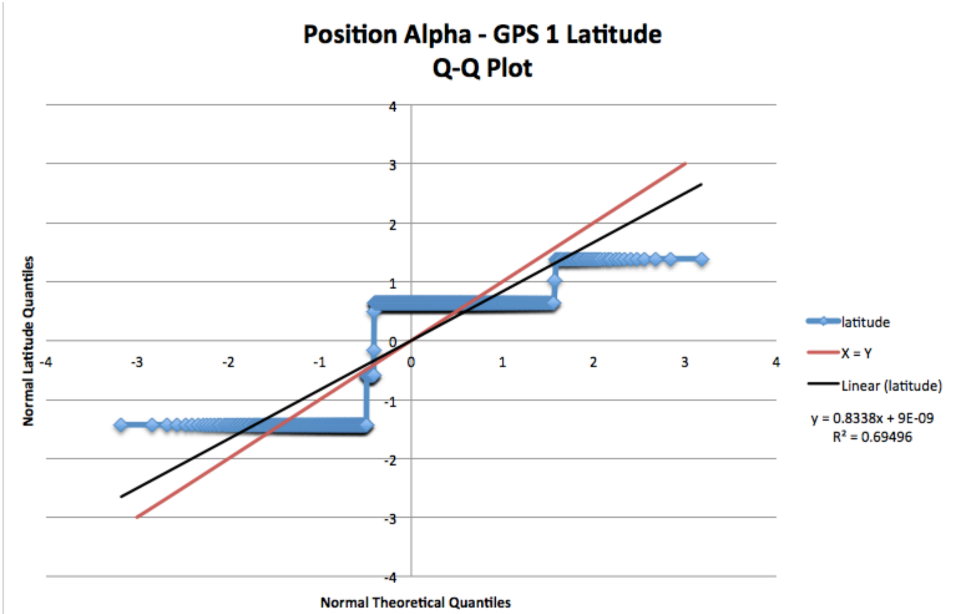

Figure 4.2.1: Q-Q Plot for data collection position Alpha, GPS module 1's latitude.

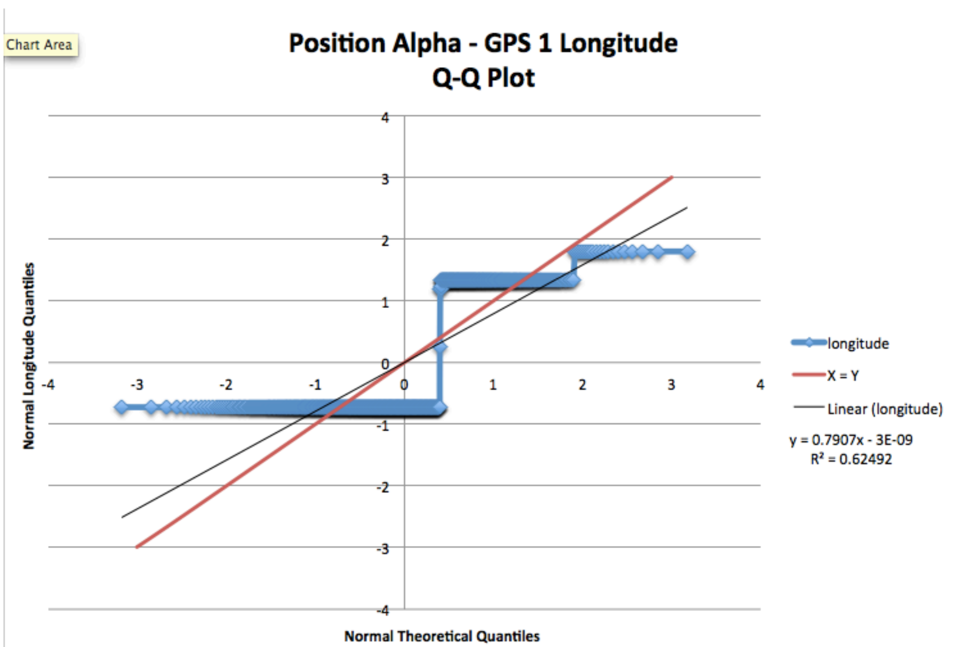

Figure 4.2.2: Q-Q Plot for data collection position Alpha, GPS module 1's longitude.

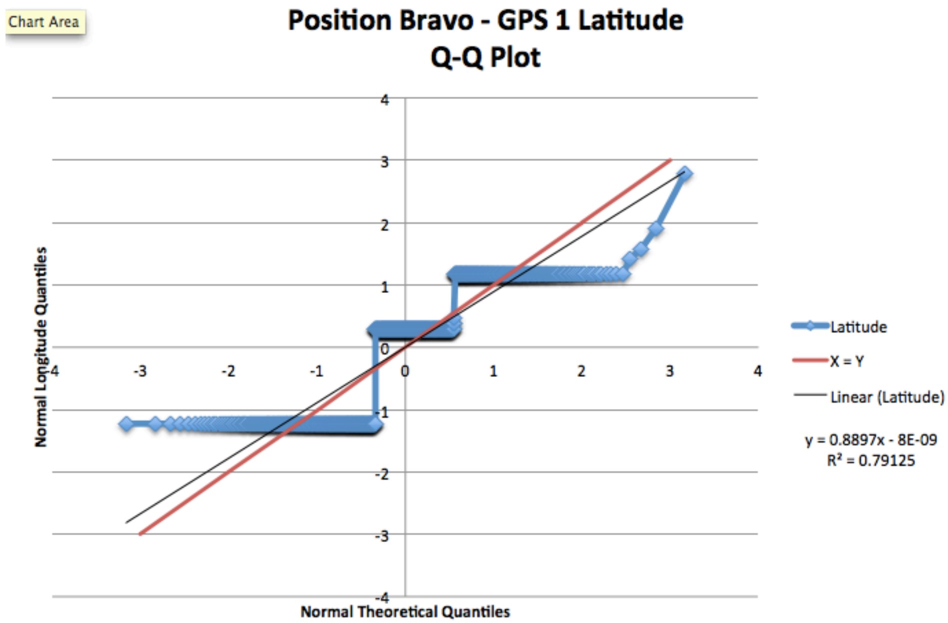

Figure 4.2.3: Q-Q Plot for data collection position Bravo, GPS module 1's latitude.

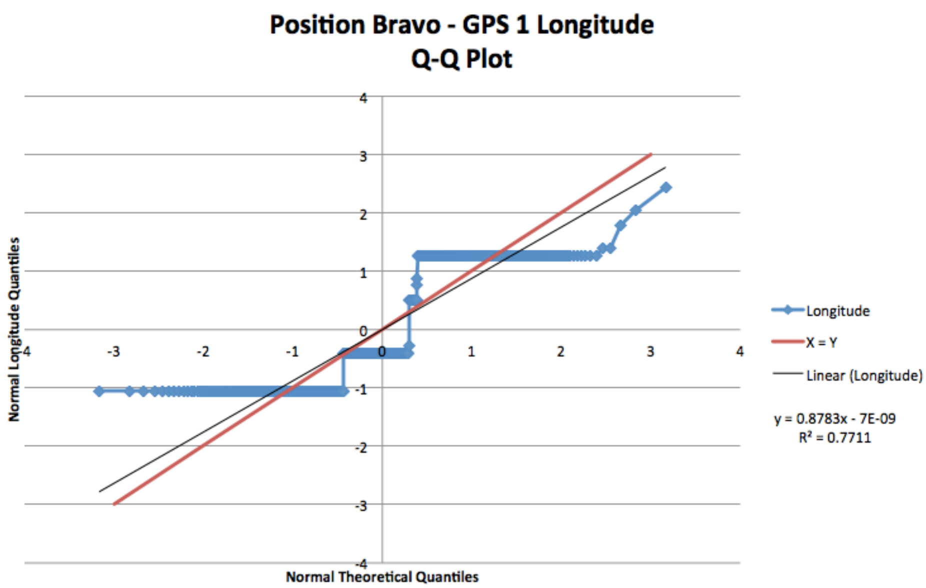

Figure 4.2.4: Q-Q Plot for data collection position Bravo, GPS module 1's Longitude.

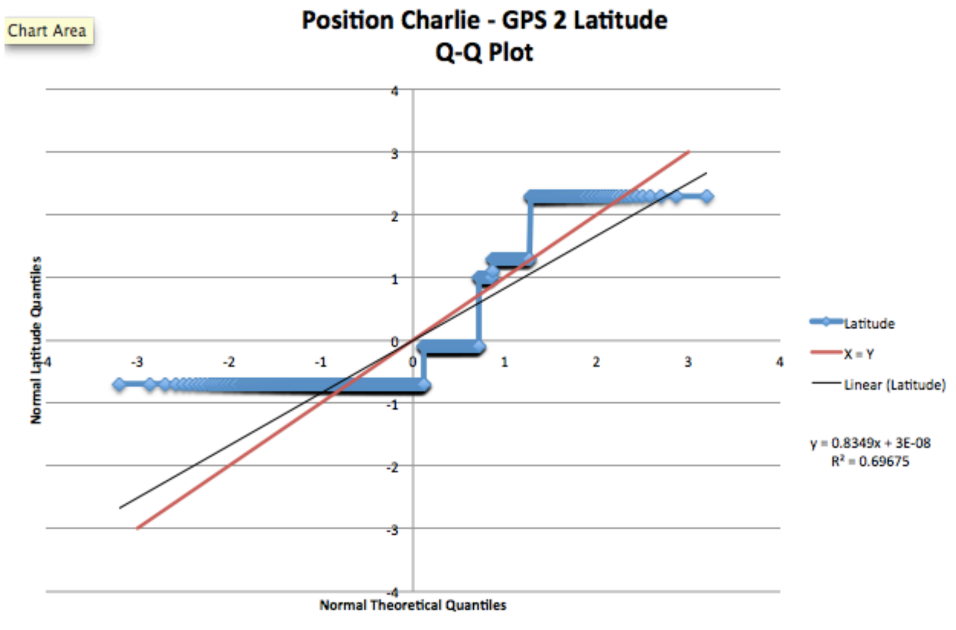

Figure 4.2.5: Q-Q Plot for data collection position Charlie, GPS module 2's latitude.

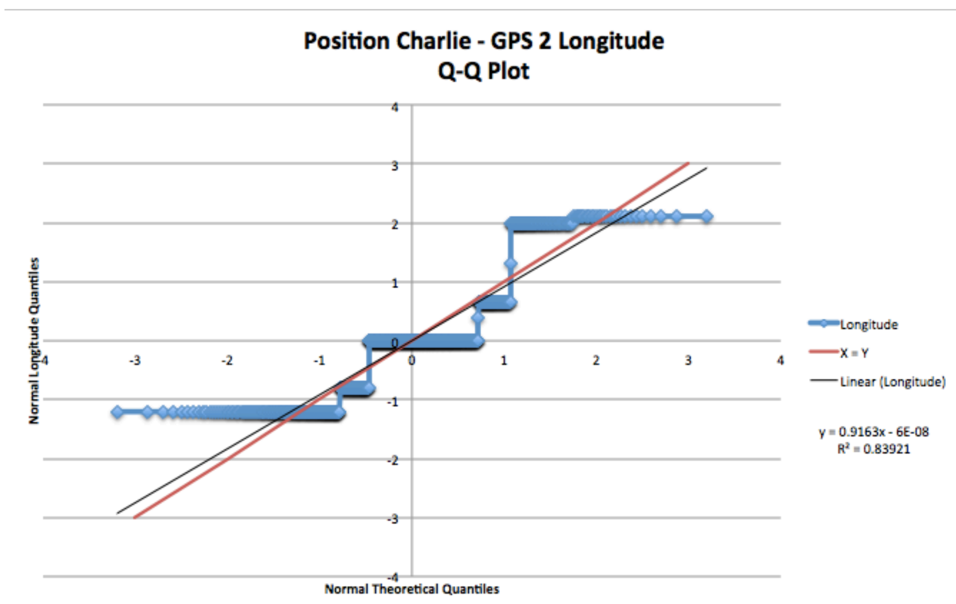

Figure 4.2.6: Q-Q Plot for data collection position Charlie, GPS module 2's longitude.

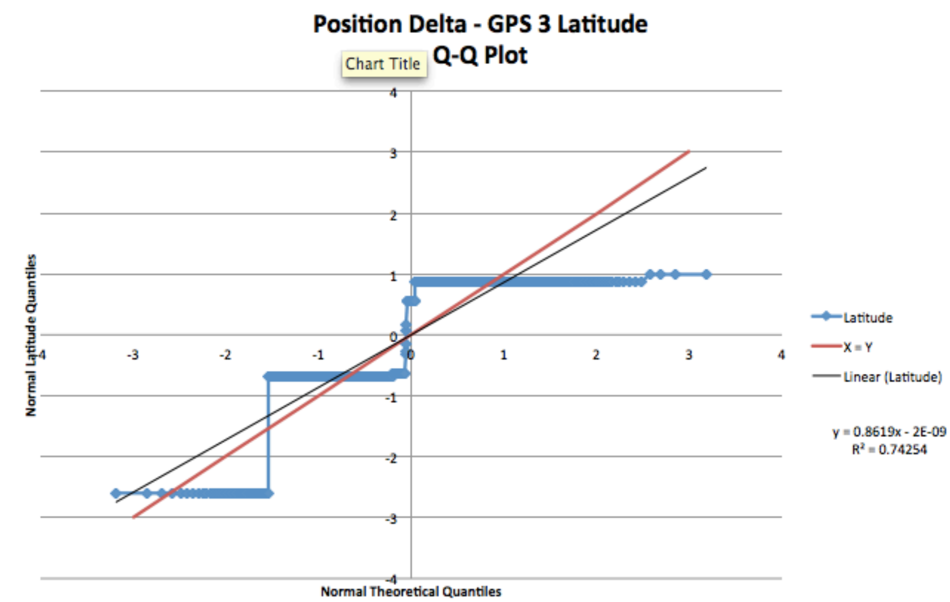

Figure 4.2.7: Q-Q Plot for data collection position Delta, GPS module 3's latitude.

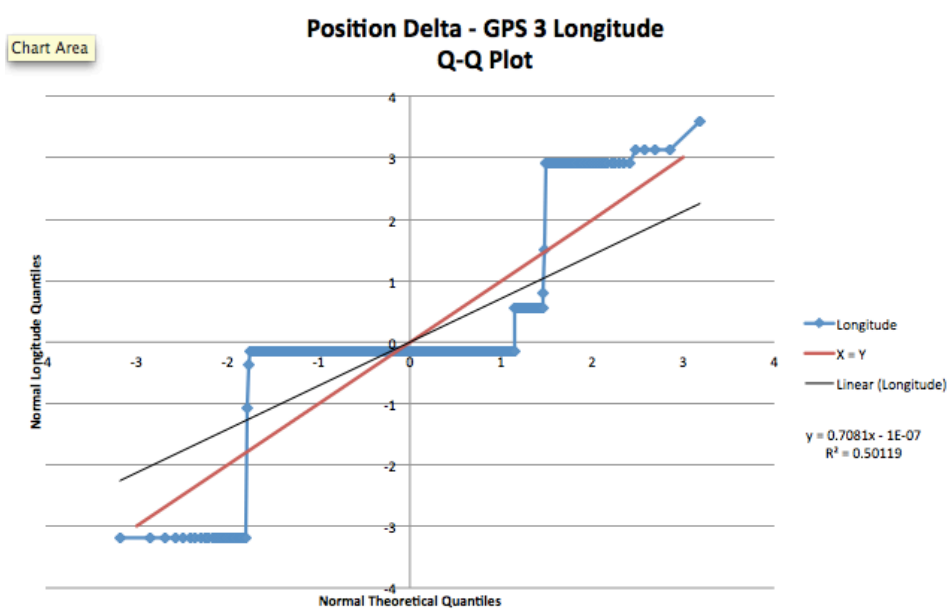

Figure 4.2.8: Q-Q Plot for data collection position Delta, GPS module 3's longitude.

In order to achieve a good spread of the data, I created one Q-Q Plot (latitude & longitude) for each location. At each location, I selected a different GPS module. Although, the trend lines do not exactly match the perfect Gaussian distribution (X = Y line), they are very consistent and are close enough to be approximated as a Gaussian distribution.

Process

4.3 Initial Conditions

The information below is the only information that was known at the time the algorithm was written.

- The entire set of collected GPS data points.
- The platform that was used to mount the GPS modules was an equilateral triangle with side length of one meter.
- At each position, the platform was oriented towards magnetic north.
- The four positions, in which the GPS data was collected, created a square with side length of approximately 25 meters.

Process

4.4 Algorithm

In order to test the algorithm, the actual or "true" location of the equilateral triangle (in relation with the earth) was needed. The output of the TGPSL algorithm will then be compared to the to the "true" center location. Using the TGPSL algorithm and the collection of all the GPS data points, I was able to determine the best approximation for the "true" center. This center was used to test my algorithm.

The following steps are applied to each location (Alpha, Bravo, Charlie, Delta) independent of each other. Each corner of the 25m square is used to test the algorithm individually. Therefore, achieving four separate TGPSL tests in one run-through of the algorithm. The first step simply consisted of finding the average points for each GPS module at each location.

1. Calculate the mean location for each GPS modules data.

In order to increase the accuracy, the mean of each GPS modules collection of data points is calculated. This mean, should ideally form the equilateral triangle of the collection platform. However, based on the collected data, a perfect equilateral triangle was never actually created **(see figures 4.4.1 – 4.4.4 below)**.

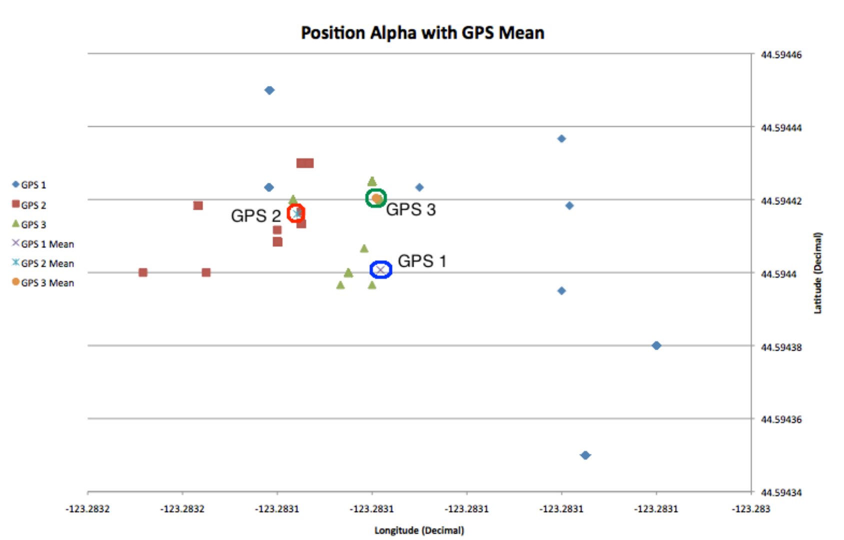

Figure 4.4.1: Position Alpha with each GPS Modules mean.

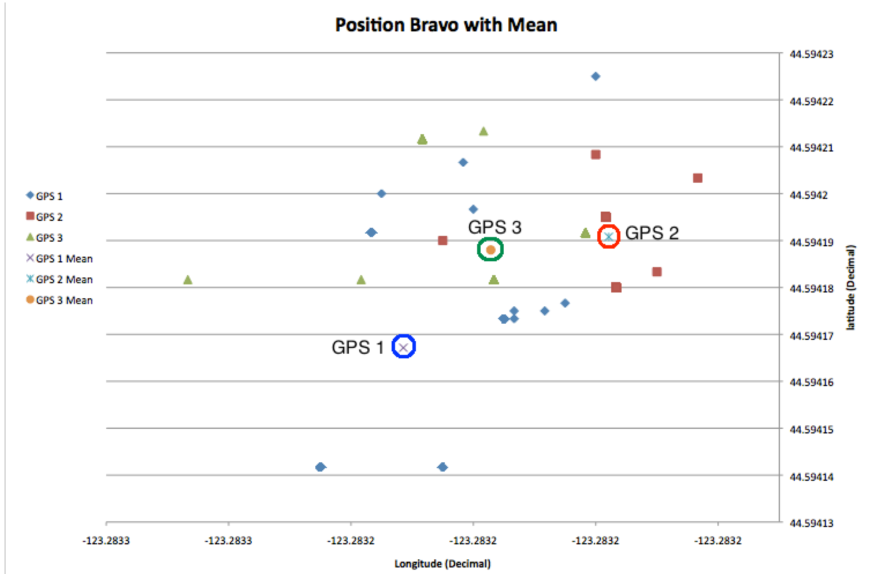

Figure 4.4.2: Position Bravo with each GPS Modules mean.

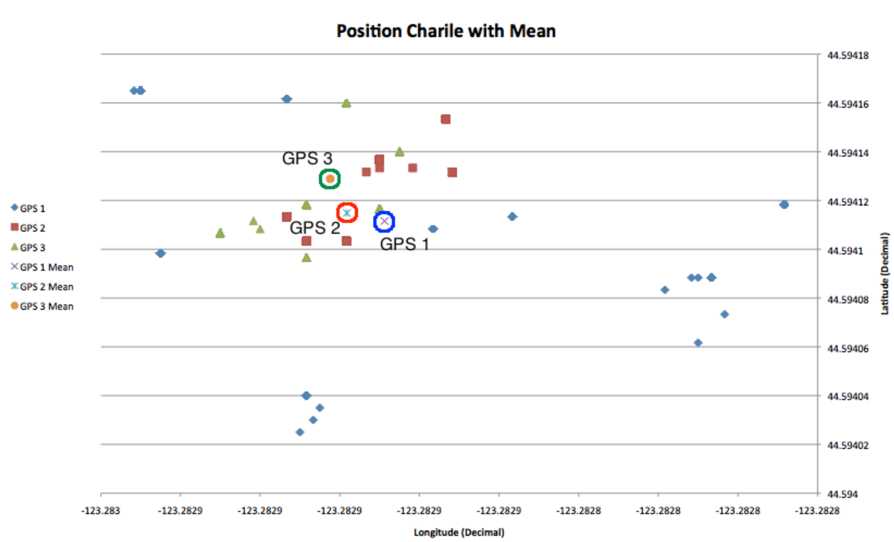

Figure 4.4.3: Position Charlie with each GPS Modules mean.

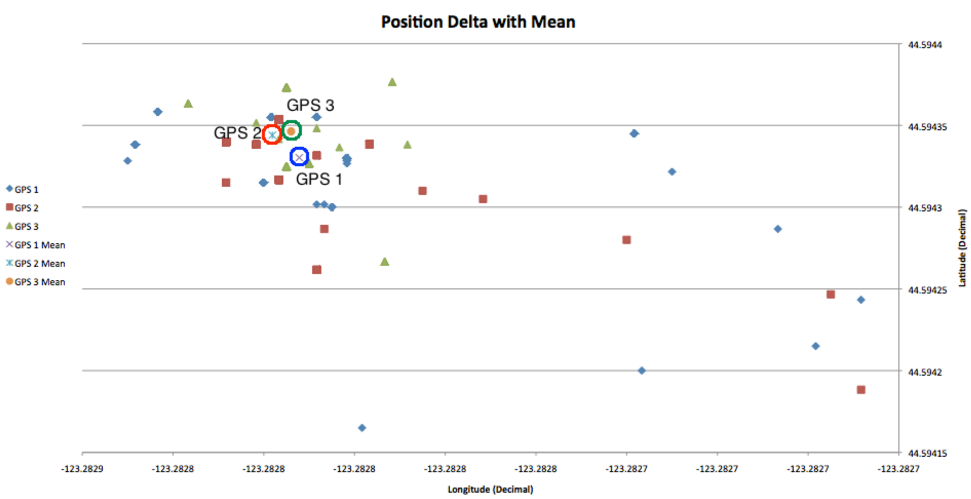

Figure 4.4.4: Position Delta with each GPS Modules mean.

2. Calculate the variance for each GPS module.

In order to calculate the variance for each GPS module, it was assumed that each GPS module is independent from the other. However, this is not actually true. If you know the coordinates of one GPS module, the other two GPS modules must be within a certain range since they are physically connected on the platform – therefore, this implies a dependent relationship. However, since it was my goal to determine the variance of each individual GPS module separately, an independent relationship was assumed.

Once each GPS module was assumed independent, then the GPS module's coordinates (longitude and latitude) could also be assumed independent. Even if the longitude is known, it says nothing about the latitude. This allowed the algorithm to use the standard one-dimensional Gaussian distribution **(see Equation 3.1)**

The Gaussian (or Normal) Probability Density Function (PDF) is a function that describes the relative likelihood for the GPS reading to occur at a given location. The Gaussian distribution will be used to determine if the data points collected tend to cluster around a single point. Ideally, the GPS data points will gather around the actual location of the GPS modules.

The variance is the sum of the squared distances of each term in the distribution from the mean, divided by the number of terms in the distribution. Unfortunately, the variance for each

GPS module is not the same. Below shows the variance for both the longitude and the latitude at each position.

Position Alpha

Variance GPS 1: Longitude = 12.92118590573728 Latitude = 15.58989594378628

Variance GPS 2: Longitude = 0.159456642737209 Latitude = 0.636391484753125

Variance GPS 3: Longitude = 0.096389847402263 Latitude = 0.170877091082693

Position Bravo

Variance GPS 1: Longitude = 2.063833878979651 Latitude = 5.331634287563483

Variance GPS 2: Longitude = 0.026265364539368 Latitude = 0.570495362042847

Variance GPS 3: Longitude = 0.649480152932984 Latitude = 1.582561775642153

Position Charlie

Variance GPS 1: Longitude = 37.02348165872709 Latitude = 24.26851977525434

Variance GPS 2: Longitude = 1.950915808490471 Latitude = 3.453247326641385

Variance GPS 3: Longitude = 1.758550343436165 Latitude = 4.797770405724188

Position Delta

Variance GPS 1: Longitude = 4.544172635965491 Latitude = 4.648228681300369

Variance GPS 2: Longitude = 1.530861933199642 Latitude = 3.597259393044629

Variance GPS 3: Longitude = 0.626205711364665 Latitude = 11.62076541008877

3. Compute the maximum probability.

The maximum probability is calculated by finding the highest probability (which is located at the mean of the of the probability density function). The maximum probability is the area under the highest point on the Gaussian distribution bell curve. This PDF will be used to calculate the maximum probability of each GPS module. **(See figure 4.4.5)**

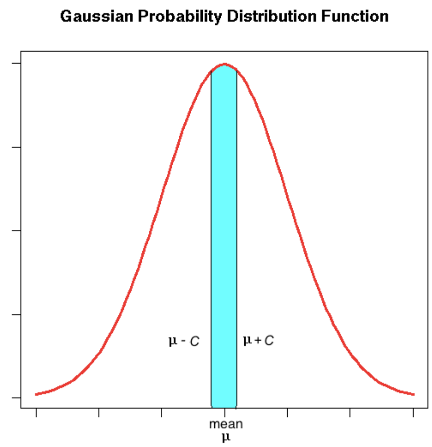

Figure 4.4.5: Maximum probability is the area under the curve surrounding the mean.

4. Create a virtual triangle based around the average location of all GPS points

Since it is known that the GPS modules were placed in an equilateral triangle formation with side length of one meter, the next step was to create this virtual equilateral triangle with those specifications. The triangle was arbitrarily placed with its center at the mean of all the data points collected and facing magnetic north.

5. Optimize the virtual triangle based on the Sum of Squared Error.

Once the equilateral triangle was place with its center at the overall mean location, it was optimized for best fit based on the Sum of Squared Error (SSE). Rotating and translating the virtual triangle to reduce the distance between the calculated mean of each GPS module and the corresponding virtual point on the equilateral triangle accomplished optimization. If the actual orientation of the triangle used to collect the data was not known, a random restart should be implemented. The random restart would reduce the chance of the virtual triangle getting stuck in a local maximum. However, the triangle used to collect the data was oriented magnetic north. Therefore, the initial orientation of true north for the virtual triangle worked as a close approximation.

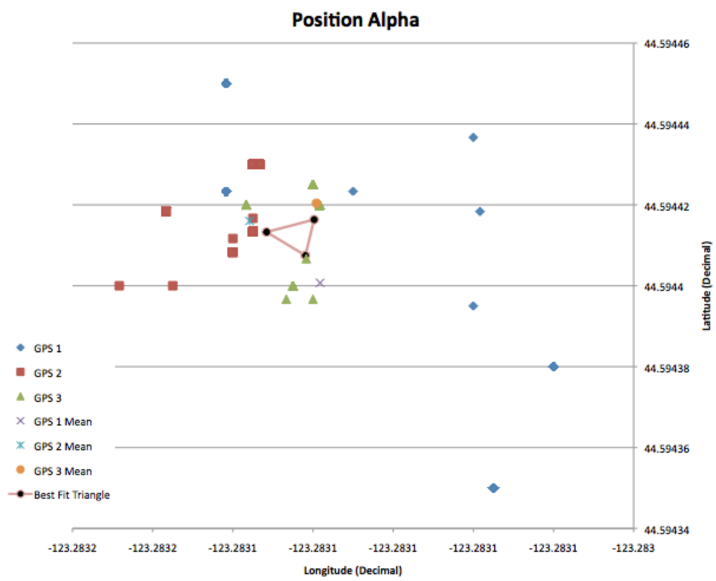

Figure 4.4.6: Position Alpha with the best-fit triangle based on the SSE.

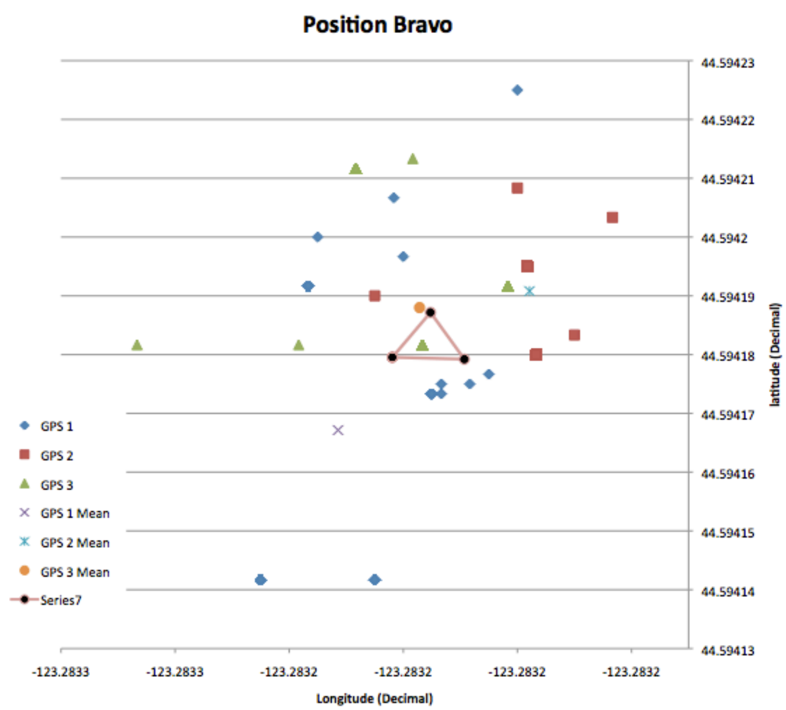

Figure 4.4.7: Position Bravo with the best-fit triangle based on the SSE.

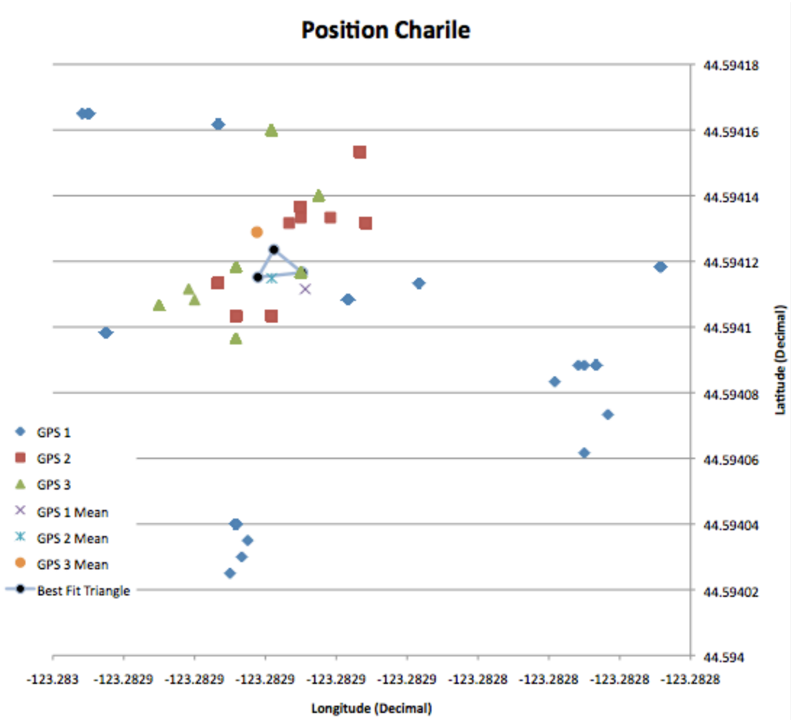

Figure 4.4.8: Position Charlie with the best-fit triangle based on the SSE.

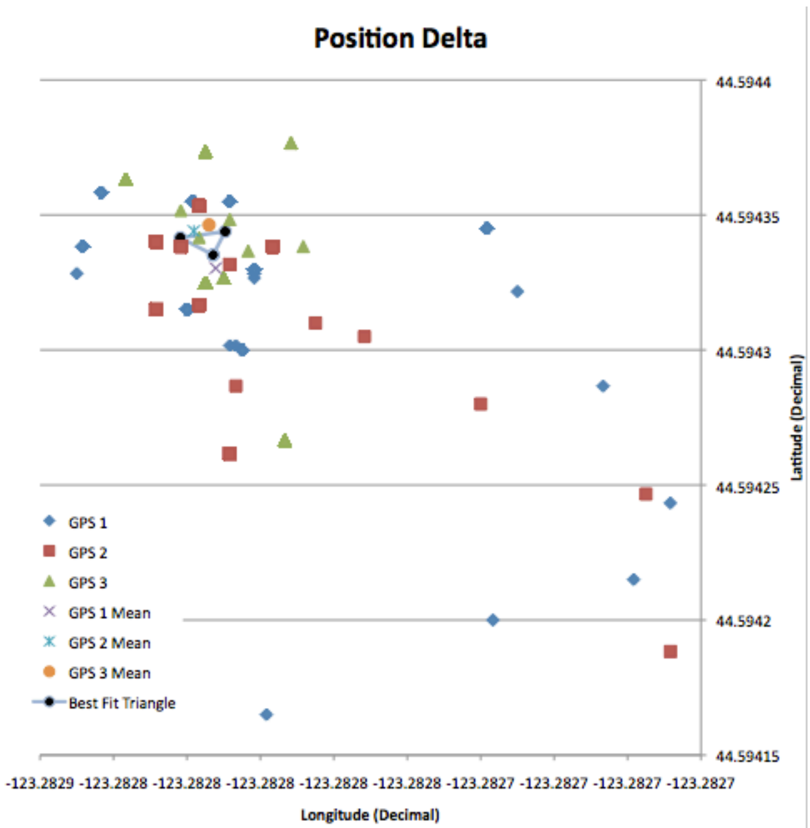

Figure 4.4.9: Position Delta with the best-fit triangle based on the SSE.

To verify the accuracy of the virtual triangle and its center, I compared the resulting locations against the actual arrangement that the data was collected in – a square with side length of 25 meters. As you can see in figure 4.4.10 & 4.4.11, the results were good. Setting up a perfect square of that size was challenging. There was most certainly human related error when laying out this square. I estimated this error to be around one meter at various locations. It turns out that I did better than I thought with exception to the distance between Alpha and Bravo (26.41m).

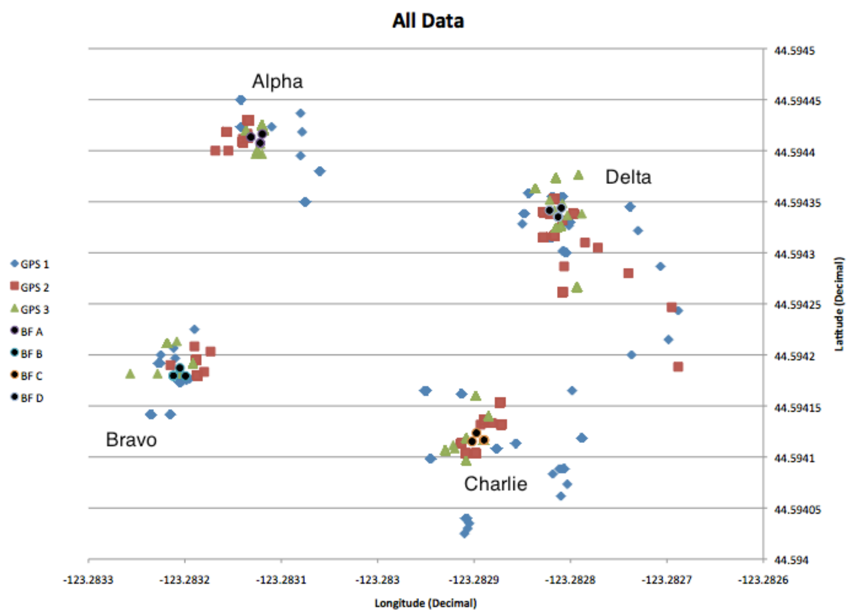

Figure 4.4.10: All of the GPS data collected with completely optimized virtual GPS positions.

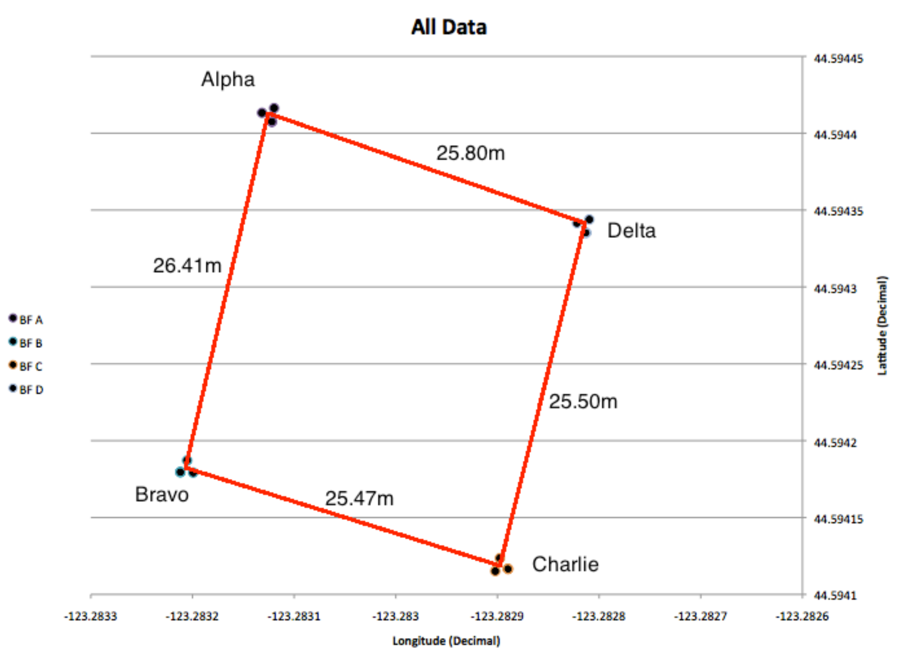

Figure 4.4.11: Completely optimized virtual GPS positions with the distance between collection positions marked.

6. Optimize the triangle based on the maximum probability for each GPS module.

The maximum probability of each GPS module was previously calculated in step 5. The larger the maximum probability is, the steeper the Gaussian bell curve, and the smaller the variance. Hence, the higher the maximum probability of a GPS module, the more accurate and reliable that GPS module's readings are.

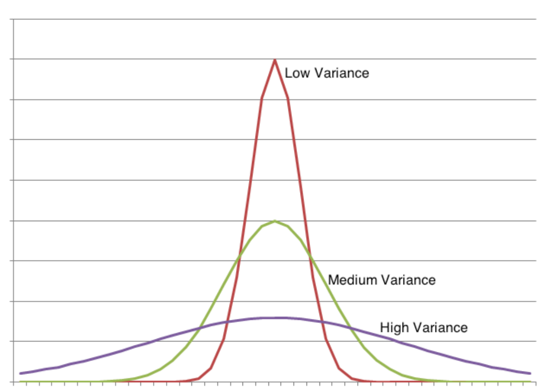

Figure 4.4.12: The smaller the variance the more accurate the GPS readings

Multiplying the maximum probability of the actual GPS mean location by the distance between the actual mean location and the corresponding virtual point on the virtual triangle will produce a score. **(See figure 4.4.11 & equation 4.1 below)** The idea is to equalize the playing field using the score as the utility. This will move the virtual triangle towards the more accurate (smaller variance) GPS module.

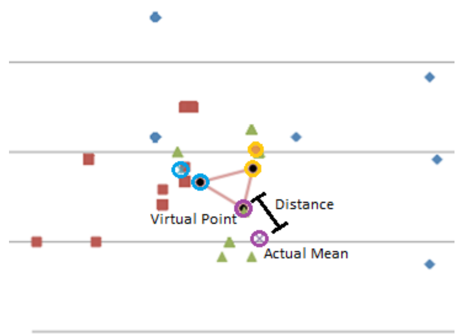

Figure 4.4.13: Graphical representation of the variables used to determine the score.

$$\underset{Probability}{\text{Max}} \times Distance = Score$$

Equation 4.1: Utility Calculation

For example, once the virtual point [GPS 1] moves close enough to the actual point [GPS 1], the score becomes less for GPS 1 and begins to increase for GPS 2 & 3. The scores for GPS 2 & 3 increase because the virtual triangle moves closer to GPS 1, increasing the distance between GPS 2 & 3's actual point and virtual point.

7. Use the center of the triangle as the actual center.

By simply equaling out the scores for each GPS module in the previous step, the virtual triangle moved closer to the more accurate/reliable GPS module and further away from the points that are considered to be less accurate. Once the virtual triangle is optimized, the center of the virtual triangle is considered to be the "true" center location. This center was used to test the algorithm.

Process

4.5 Testing the Algorithm

1. From the collected data, get one GPS point from each GPS module with [relatively] the same time stamp. The "star" represents the true center found based on the entire collection of GPS data points.

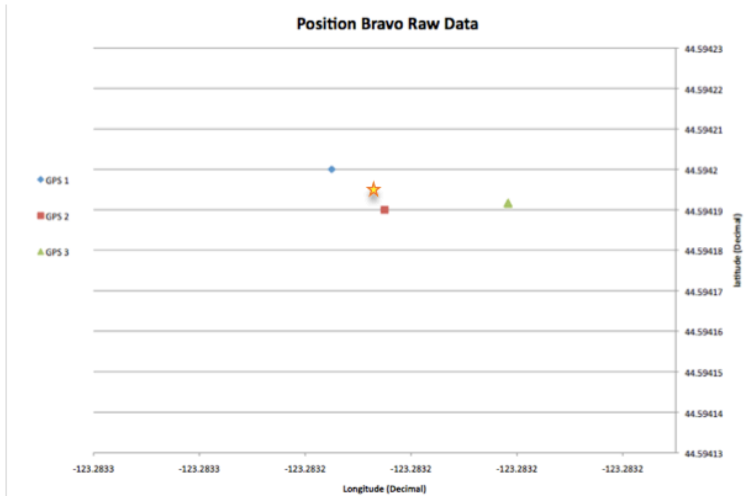

Figure 4.5.1: Graphical representation of the testing process. Snap-shot of three GPS readings.

2. Create a virtual triangle based around the average location of these three points.

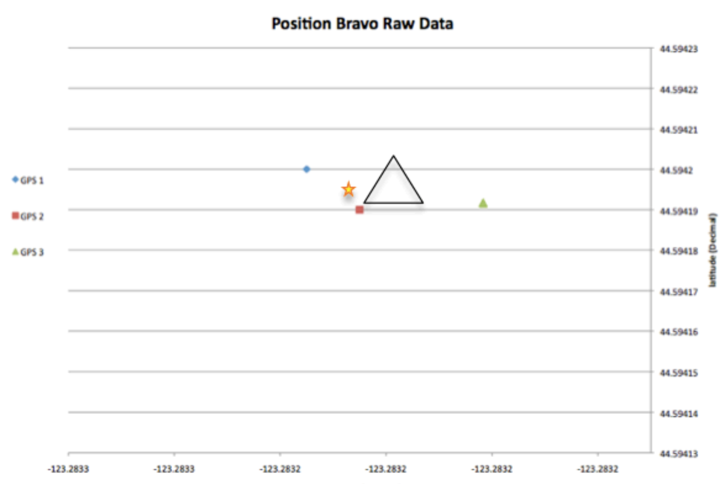

Figure 4.5.2: Graphical representation of the testing process. Virtual triangle arbitrarily placed at average location of the three GPS readings.

3. Optimize the virtual triangle based on the Sum of Squared Error.

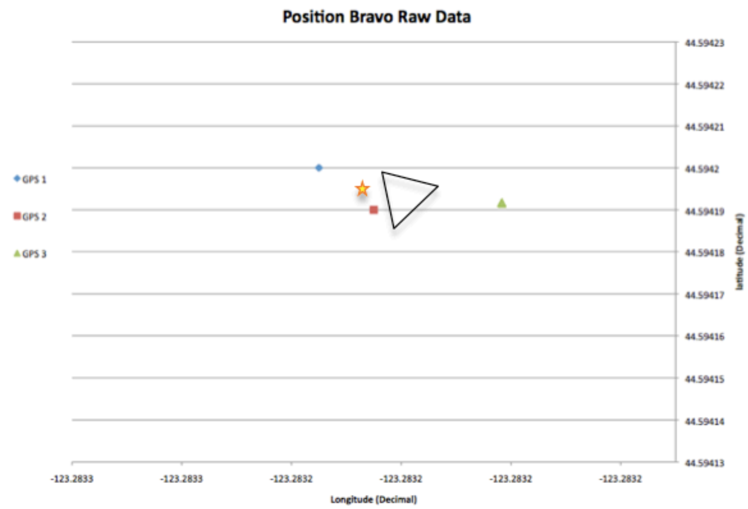

Figure 4.5.3: Graphical representation of the testing process. Virtual triangle optimized based on sum of squared error.

4. Optimize the triangle based on the maximum probability for each GPS module.

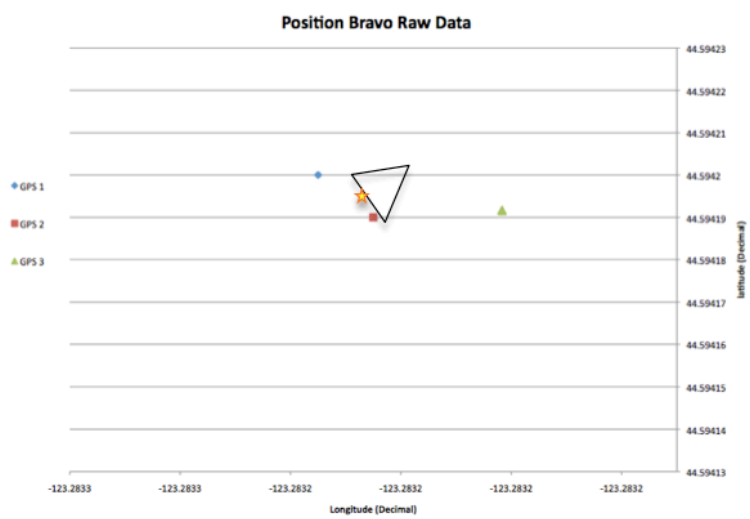

Figure 4.5.4: Graphical representation of the testing process. Virtual triangle optimized based on maximum probability.

5. Compare the center of the optimized virtual triangle with actual center computed above.

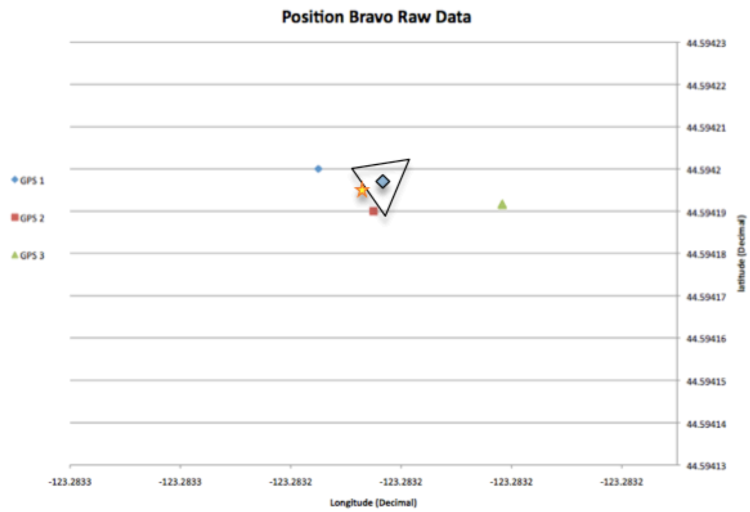

Figure 4.5.5: Graphical representation of the testing process. Virtual triangle center compared to actual location.

When the algorithm converges, the center of the optimized virtual triangle based off three GPS points (one from each GPS module) and true center previously calculated from the entire data set are compared. The performance was measured by the error of the prediction according to the distance from the actual center.

Chapter 5:

Results & Analysis

Predicting an accurate location can be achieved with the use of a few GPS modules and a machine-learning algorithm. The results show a remarkable improvement from the original accuracy. Originally, the Venus GPS modules were capable of determining a position within a radius of 2.5 meters 50% of the time and a radius of 5 meters 95% of the time. Using the algorithm described, the GPS modules can determine their position within a radius of 0.45 meters 50% of the time and a radius of 0.85 meters 95% of the time. That is an 82% increase in accuracy from the original 2.5m CEP and an 83% increase in accuracy from the original 5m (2*CEP).

	CEP (50%)	**2*CEP (95%)**
Single GPS	2.5m	5m
Alpha (TGPSL)	0.56m	1.24m
Bravo (TGPSL)	0.45m	0.85m
Charlie (TGPSL)	1.42m	1.94m
Delta (TGPSL)	2.5m	>5m
Delta Average	1.63m	3.72m

Table 5.1: Comparison of the resulting Circular Error Probability for each position.

It is interesting to note that positions Alpha and Bravo did a little better than position Charlie and significantly better than positions Delta. This is more likely due to the fact that position delta was about 10 to 15 meters away from a three story apartment building. Position delta is a great example of how nearby buildings can affect GPS accuracy. In addition, position Delta is the only position in which the average position found the correct spot more often than the TGPSL algorithm. In all other cases, the TGPSL algorithm out performed simply taking the average of the three GPS modules.

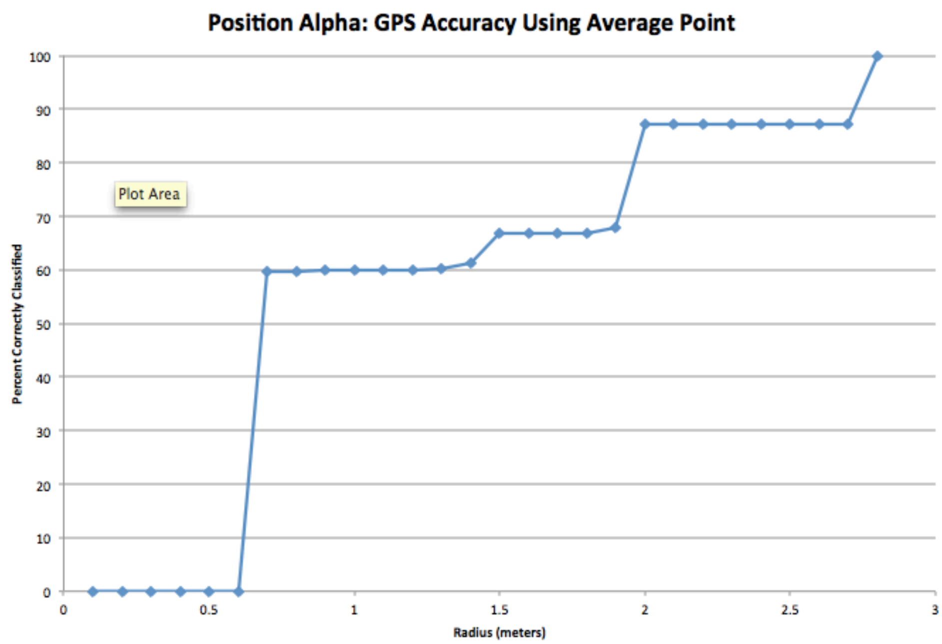

Figure 5.1: Position Alpha – The probability of correctly predicting the actual location within a particular radius using the average of the three GPS modules readings.

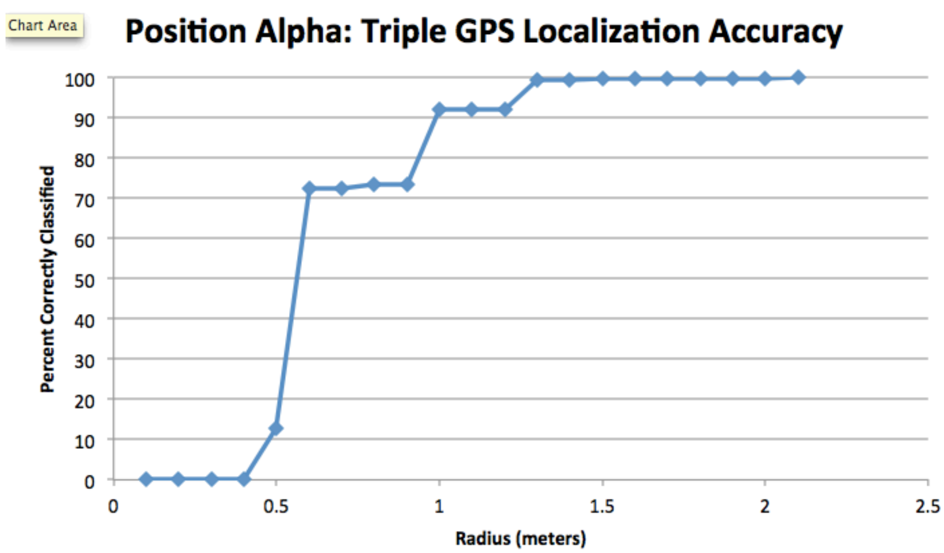

Figure 5.2: Position Alpha – The probability of correctly predicting the actual location within a particular radius using TGPSL.

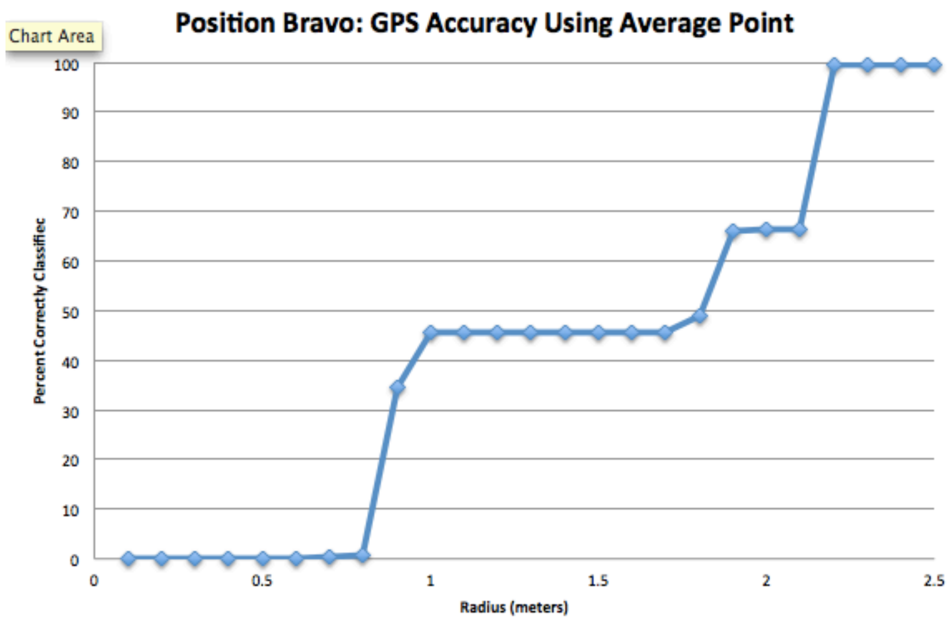

Figure 5.3: Position Bravo – The probability of correctly predicting the actual location within a particular radius using the average of the three GPS modules readings.

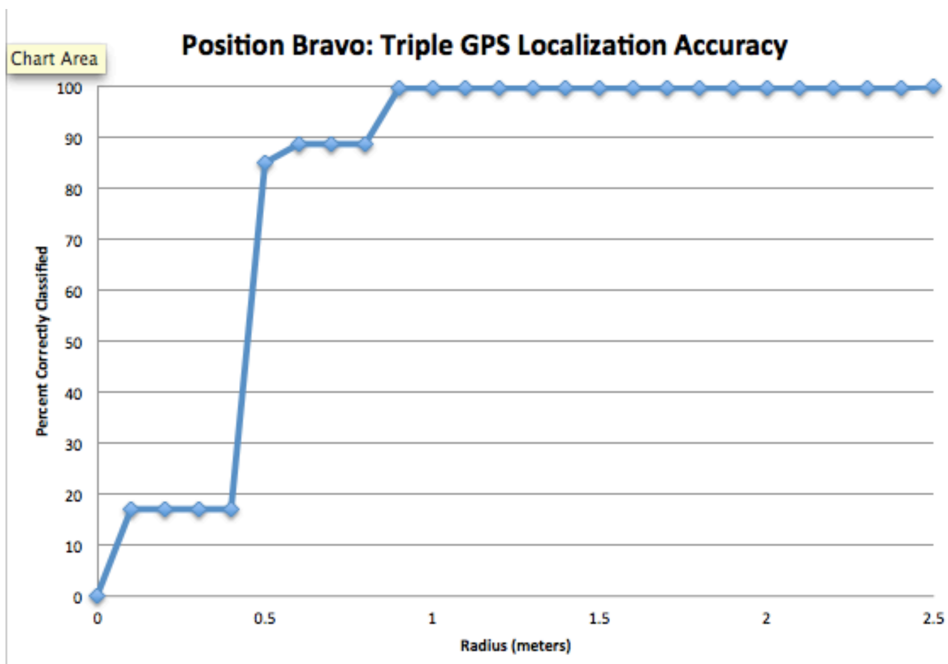

Figure 5.4: Position Bravo – The probability of correctly predicting the actual location within a particular radius using TGPSL.

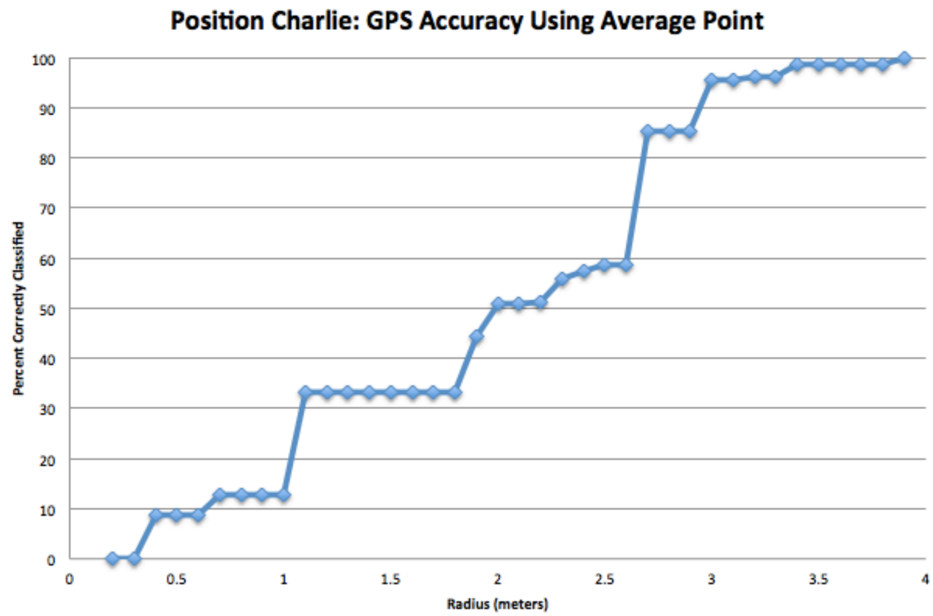

Figure 5.5: Position Charlie – The probability of correctly predicting the actual location within a particular radius using the average of the three GPS modules readings.

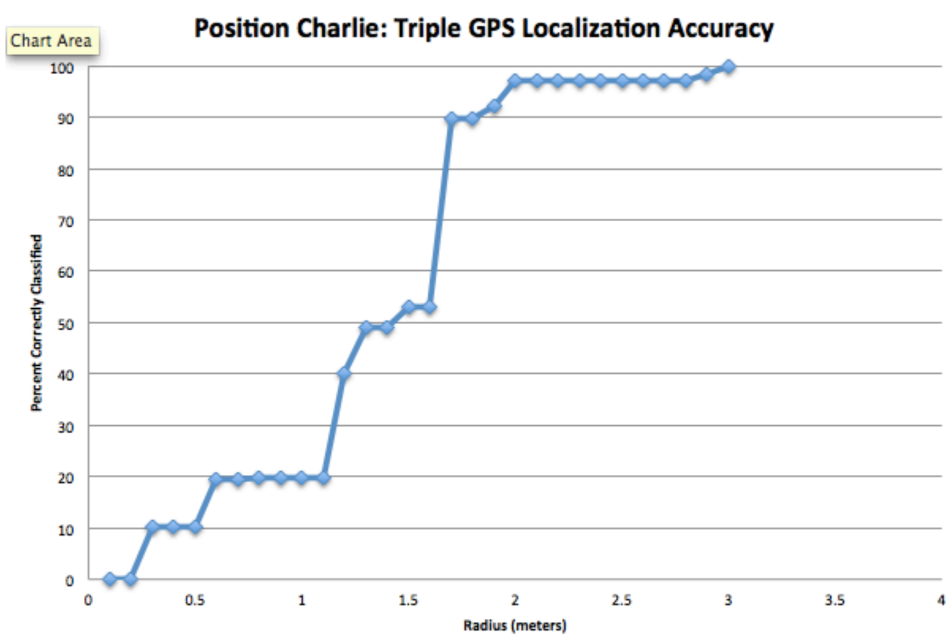

Figure 5.6: Position Charlie –The probability of correctly predicting the actual location within a particular radius using TGPSL.

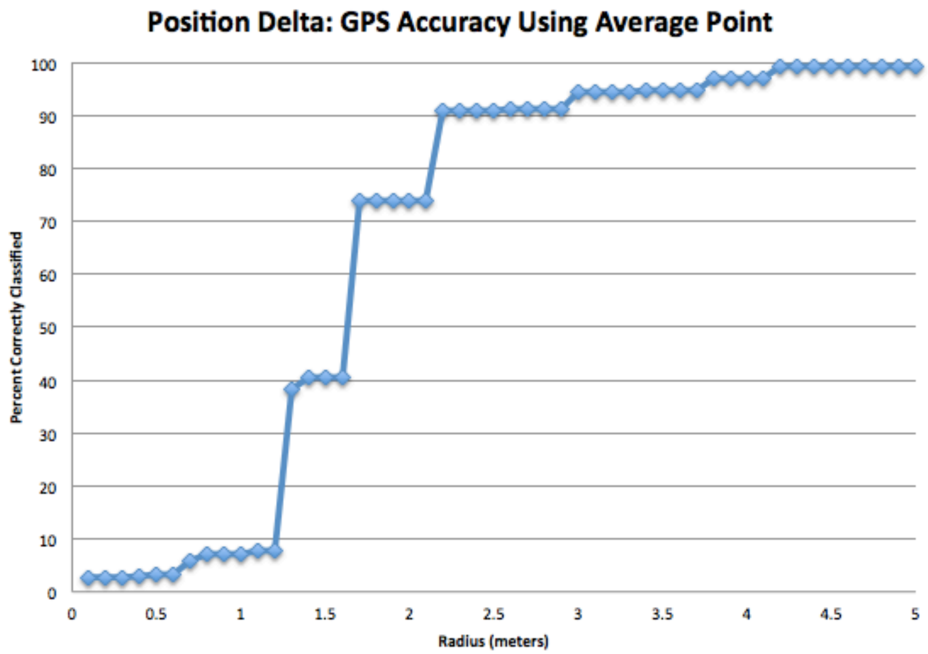

Figure 5.7: Position Delta – The probability of correctly predicting the actual location within a particular radius using the average of the three GPS modules readings.

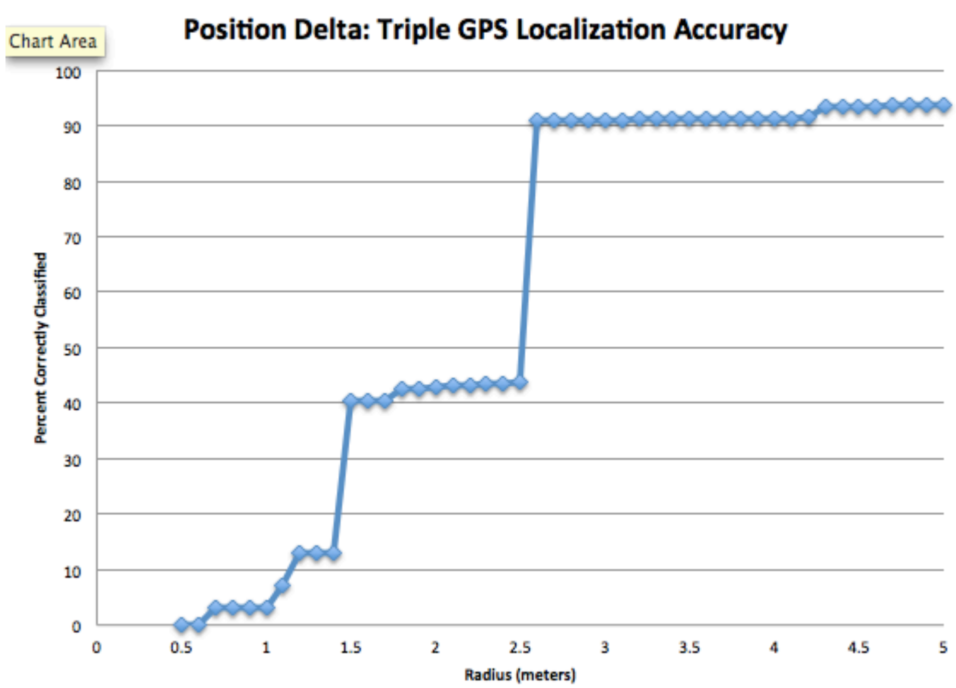

Figure 5.8: Position Delta – The probability of correctly predicting the actual location within a particular radius using TGPSL.

One Venus GPS module can calculate your location within a 5-meter radius 95% of the time. Combining an additional two Venus GPS modules in an equilateral triangle and taking the average location will increase your accuracy with minimum computation – No need to compute the Gaussian distribution, recursively optimize based on sum of squared error, nor recursively optimize based on maximum probability. However, by applying this machine-learning algorithm, TGPSL, to the three GPS modules can potentially increase your accuracy by approximately 82% the original rated accuracy. In addition, this algorithm can be applied to any three GPS modules of the same kind in the equilateral triangle formation. Therefore, depending on the accuracy desired, you could potentially buy any three GPS modules you like, and improve their overall accuracy.

GPS Receiver	CEP (m)	Price ($)
GS407 Helical	2	89.95
Venus	2.5	49.95
Copernicus II	2.5	44.95
LS20031	3	59.95
San Jose Navigation GPS	3.3	99.95
EM-406A SiRF III	5	59.95

Table 5.2: Accurate stand-alone GPS receivers under $100.[10]

The GPS modules listed in table 5.2 are accurate and relatively low cost. More GPS receivers can be found at the website below. These GPS receivers range from $100 to $250 and have accuracy between 1 and 5 meters using DGPS.

http://home-2.worldonline.nl/samsvl/oemtable.htm

Chapter 6:

Conclusion & Future Work

Triple GPS Localization (TGPSL) consists of mathematically calculating the probability that you will be within a particular radius, utilizing three separate GPS modules setup in an equilateral triangle formation.

$$P(Radius < 2.5m \mid Three\ GPS\ modules, setup\ in\ an\ equilateral\ triangle\ orientation) = 100\%$$

Equation 6.1: Achieved Probability

Although the Venus GPS modules used have DGPS capabilities, it was not enabled during any of the data collection. TGPSL does not have to operate independently of other positional correction methods such as DGPS or WAAS. It would be interesting to see how much the accuracy would improve if both TGPSL and DGPS methods were used.

Increasing the accuracy using TGPSL is very promising. For under $200, a localization system was built that is accurate within a radius of 0.45 meters 50% and 0.85 meters 95% of the time.

There are many more improvements that can be made using the TGPSL concept. In addition, expanding on the current concept could help answer a few questions.

- Find another optimization or method that could be applied to further increase the accuracy of TGPSL versus using a Gaussian distribution.
- Determine the accuracy improvement if you use four GPS modules arranged such that each GPS module lies at the corners of a square. Determine the accuracy improvement using five GPS modules?

- Create an equation that inputs accuracy desired and outputs required number of GPS modules needed.
- Based off the bell curve of the Gaussian distribution, can you create a probability that tells you how accurate your position actually is? How "good" is the localization information? For example, if all three GPS readings are far apart, how accurate is your position compared to if the GPS readings were close together?

REFERENCES

[1] Clynch, James. "A Short Overview of Differential GPS."*Fundamentals of Geospatial Information and Services*. Naval Postgraduate School, Dec 2001. Web. 10 May 2012. <http://www.oc.nps.edu/oc2902w/gps/dgpsnote.html>

[2] "Rockwell Collins." *Defense Advanced GPS Receiver*. Rockwell Collins, 2012. Web. 10 May 2012. <http://www.rockwellcollins.com/sitecore/content/Data/Products/Navigation_and_Guidance/GPS_Devices/Defense_Advanced_GPS_Receiver_-DAGR.asp&xgt>

[3] Smithsonian, . "How Does GPS Work." *GPS: A New Constellation*. National Air and Space Museum, 2012. Web. 10 May 2012. <http://airandspace.si.edu/gps/work.html>.

[4] "Selective Availability." *GPS Modernization*. National Coordination Office for Space-Based Positioning, Navigation, and Timing, 2011. Web. 10 May 2012. <http://www.gps.gov/systems/gps/modernization/sa/>.

[5] Clynch, James. "GPS Accuracy Levels." *Fundamentals of Geospatial Information and Services*. Naval Postgraduate School, 2001. Web. 10 May 2012. <http://www.oc.nps.edu/oc2902w/gps/gpsacc.html>.

[6] Diggelen, Frank. "GPS Accuracy: Lies, Damn Lies, and Statistics." *GPS World*. Questex Media Group LLC, 29 Nov 1998. Web. 10 May 2012. <http://www.gpsworld.com/gps/gps-accuracy-lies-damn-lies-and-statistics-1127>.

[7] "Normal distribution." *Absolute Astronomy*. Absolute Astronomy, 2011. Web. 10 May 2012. <http://www.absoluteastronomy.com/topics/Normal_distribution>.

[8] Chivers, Morag. "Differential GPS Explained." *ESRI*. N.p., 2012. Web. 10 May 2012. <http://www.esri.com/news/arcuser/0103/differential1of2.html>.

[9] "Sources of Errors in GPS." *The GPS System*. kowoma.de, 19 Apr 2009. Web. 10 May 2012. <http://www.kowoma.de/en/gps/errors.htm>.

[10] "GPS Receivers." *Sparkfun*. SparkFun Electronics, n.d. Web. 17 May 2012. <http://www.sparkfun.com/search/results?term=gps&what=products>.