GPS Data Analyzation

Luke Davidson - Professor Singh - EECE 5554 - Lab 1

Abstract

With all measurements come errors, especially when trying to identify the pinpoint location of a 1-inch diameter GPS receiver on the surface of the Earth. Understanding and minimizing these errors is crucial to improving many of the localization challenges that face technology. In this experiment, we used the BU-353S4 GPS Receiver to receive GPS location data in both stationary and transient states, used that data to complete conversions to UTM coordinates, and published and recorded those resulting values to be analyzed for error.

Testing and Data

We created a way to acquire data by creating a ROS publisher node that reads serial data from the GPS device through a USB port, parses that incoming data for the $GPGGA format, calculates necessary values such as UTM coordinates, and publishes final values to a ROS topic in the format of a custom message containing the values: latitude, longitude, altitude, UTM Easting and Northing, and a header. This data was recorded for two states: 10 minutes of stationary readings, and roughly 700 m of transient walking data.

Stationary Data

Test setup: GPS Receiver ~5 ft off the ground in an open area. The results are plotted below:

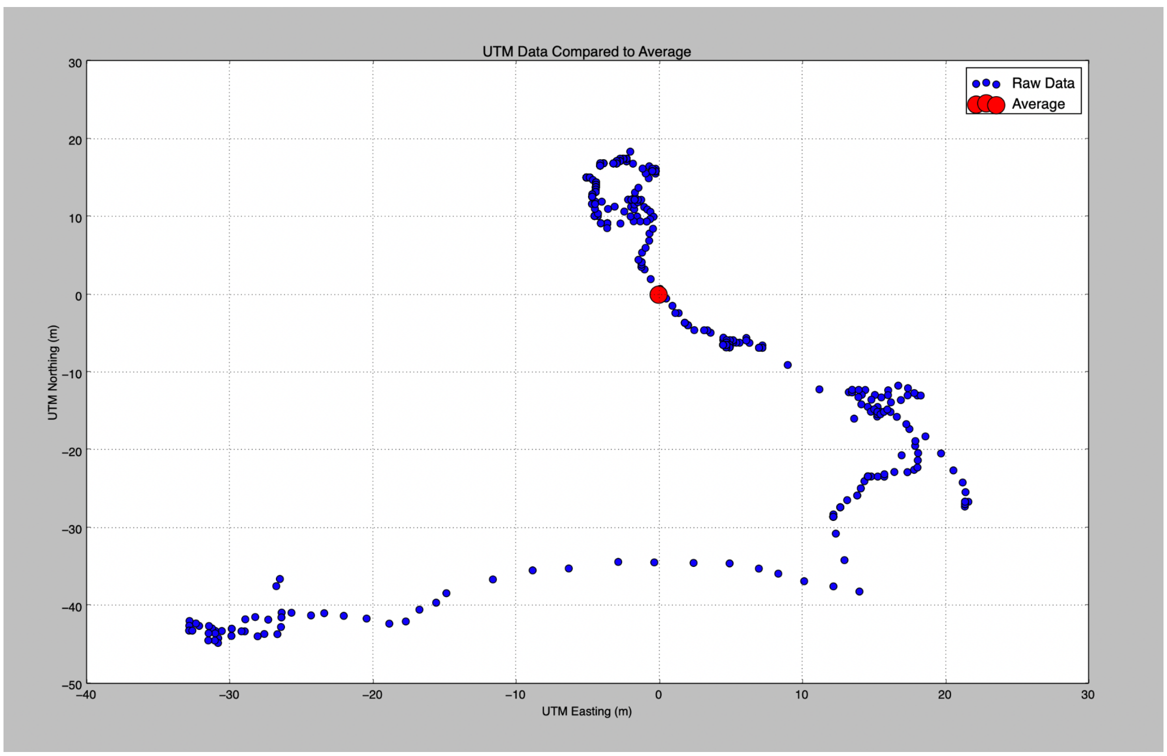
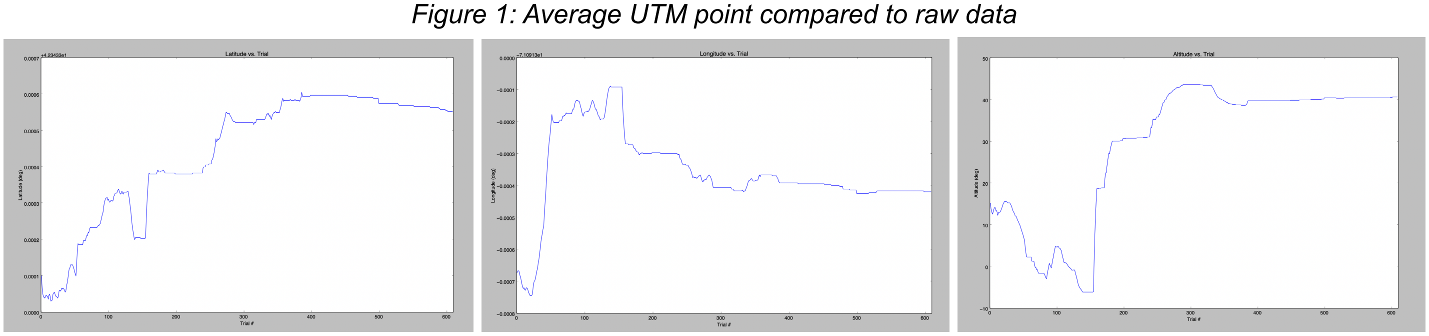


Figure 1: Average UTM point compared to raw data



Figures 2-4: Latitude, Longitude, and Altitude vs. Sequence Number

Transient

Test setup: Walking down Commonwealth Avenue in a straight line holding the receiver at ~6 ft. The resulting UTM Northing vs. UTM Easting data, including a linear line of best fit, is shown below:

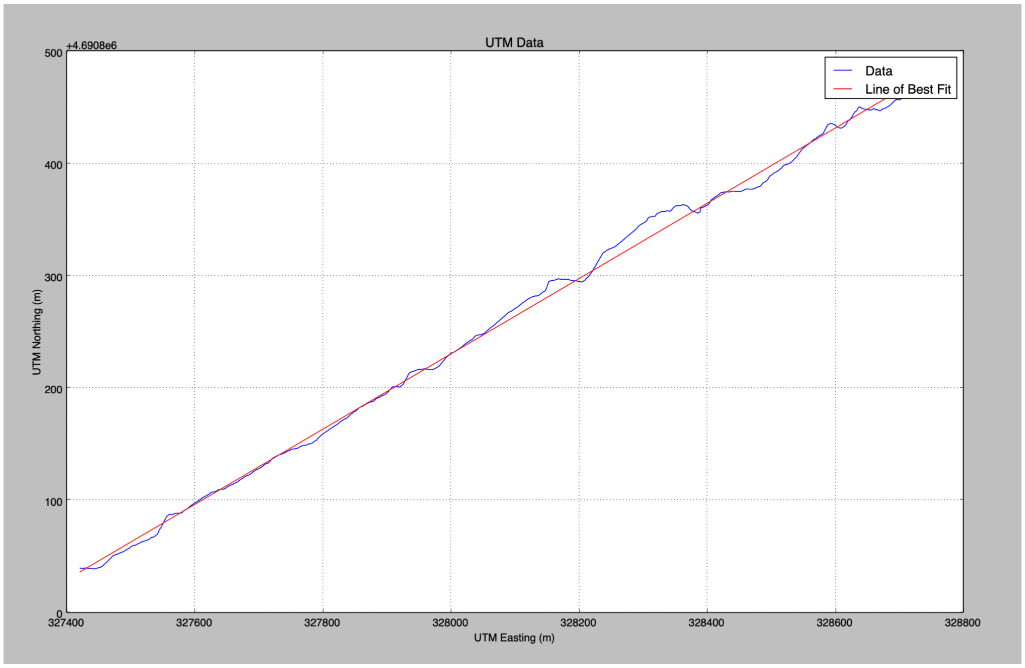


Figure 5: UTM Northing vs. UTM Easting Compared to Linear Line of Best Fit

Data and Error Analysis

To accurately analyze the errors seen in the above plots, I calculated the average distance of all of the data points to a theoretically correct location and compared the results. For the stationary data, I set this theoretically correct value as the average of UTM E and N. For the transient data, I fit a linear line of best fit to the data and calculated the distance of each data point to that line. The results are shown below:

Average Distance from Average (Stationary): 18.343 m
Average Distance from L.o.B.F (Transient): 4.587 m

As seen, the transient data showed less error than the stationary data. This was the expected behavior before conducting the study for multiple reasons. First, recorded GPS data of a moving body tends to follow a certain trajectory, the trajectory of that moving body, which can be modeled by fitting a line of best fit to that dataset. This line of best fit is a much better representation of the true expected location of the moving body than, for example, taking the average UTM E and N value of the stationary data because there is lots of noise associated with the stationary reading, and the true correct location is not accurately known and may not be the average. In Figures 2-4, we can see the data may have converged to a certain, more accurate, reading around sequence number 400.

Other sources of error in this study that caused the stationary data to have more error include the weather and location at the time of the data recording, and human error such as not walking in a perfect straight line. When I recorded the stationary data, it was cloudy and rainy, and I was situated relatively near two buildings, whereas it was sunny and I was in more of an open area when recording the transient data. The buildings can cause interferences with the communication between the GPS and satellites such as blocking or bouncing of signals. The cloudy weather can also cause communication issues as it acts as a physical barrier between the data sender and receiver.

Conclusion

Unfortunately for almost all forms of data, including GPS localization data, there will be some form of error that will need to be accounted for. One of the most important aspects of designing an efficient and effective system is knowing the sources and effects of these errors and how to best minimize and control them. Through collecting and analyzing GPS data in different states, we were able to create connections and draw conclusions about why these errors may be larger in certain situations and when certain situations are more fit to use this type of technology.