An Experiment for Outdoor GPS Localization Enhancement using Kalman Filter with Multiantenna Consumer-Grade Sensors

Phudinan Singkhamfu¹, Parinya Suwansrikham²
College of Arts, Media and Technology
Chiang Mai University, Chiang Mai Thailand

Abstract—Consumer-Grade global positioning system (GPS) is widely used in many domains. The obvious issue of this consumer-grade device is low accuracy and reading fluctuation results. In terms of using an application that requires a more precise location, the output could be difficult. In this study, the authors deploy various methods to reduce the global positioning system data fluctuation and present field test results. Two main types of the device worked together to collect data from global positioning systems, such as Microcontroller for algorithm processing and presenting data and global positioning system receivers for receiving data from a satellite. We combine three global positioning system modules to received signals in a single device and test calculated data compared with the Kalman filtering methods in many cases, including moving and static devices. Implementing the Standard Kalman Filter to multiple global positioning system Modules has improved the constancy of cheap global positioning system equipment. The experiment algorithm is presented significant improvement to overcome the retrieved data fluctuation problem. This study's contribution will enable creating a cheap global positioning system locator device for various applications that require more accuracy than the standard consumer-grade receiver.

Keywords—Global positioning systems accuracy; Kalman; multi global positioning systems; global positioning systems pointer; global positioning systems enhance; filtering algorithm

I. INTRODUCTION

It is widely known that Global Positioning System or GPS [1], which was invented during the 1960s–1970s, has been broadly used in several sectors such as service, academics, economics, and development. It can safely be said that GPS is a fundamental technology commonly found in our daily lives.

Even though the positioning system of GPS is relatively new and has been further developed into numerous inventions in the last five decades, it does not particularly mean that GPS is the most accurate system, especially when compared to GNSS (Global Navigation Satellite Systems), which is a more expensive specialized navigation system [2,3].

Although, a consumer-grade GPS is less accurate, and current computer technology can improve its precision with algorithm commands. Kalman Filter is an algorithm used to estimate possible variables and lower the discrepancy of GPS. In consequence, it is making the inexpensive GPS locator for many projects that limited fund is complicated, for example, the guidance device in entre level drone, personal location device, and forest fire locator for the rescue team.

However, there can still be an unsatisfying discrepancy if Kalman Filter is solely applied to just one device [4]. On the other hand, if several GPS devices are integrated with Kalman Filter to determine a more reliable statistical means, the results can be more efficient compared to using only one GPS device [5,6]. The prototype also has the limitation of hardware durability due to using a prototype grade sensor and Universal printed circuit board (PCB).

This experiment aspires to present a new concept derives from combining two calculation techniques using different algorithms but sharing the same objectives. This innovation can elevate the efficiency of the system using only one of the calculation techniques. It is expected that this innovation is an alternative to better technological development.

II. BACKGROUND

For technological development, consumer-grade smart devices typically contained parts or sensors that could easily be found in the market due to cheap costs and accessibility while still generating acceptable precision. For example, a Quadcopter drone could solely control the Hover Control System by itself using the Microcontroller and Inertial Measurement Unit (IMU), which could be found in general markets [7].

Lower prices and convenient accessibility came with lower efficiency compared with other more expensive specialized devices. Moreover, there have been many times that the instability of the devices results in inaccuracy. One of the most encountered problems was the instability of GPS in navigating and positioning. The accuracy of 95% of the reviewed literature was approximately 10 – 15 meters from the designated location, both Latitude and Longitude [8]. This was since several environmental factors were affecting the accuracy of the results of consumer-grade GPS devices; for example, there was a Doppler Shift phenomenon where the increased speed of GPS devices generated very low discrepancy [9], and the weather during a clear sky generated 0 – 2 meter discrepancy, while during a closed canopy condition, the discrepancy could be up to 9 meters [10].
It is common that the minor discrepancy the device generated, the more reliable it was. Nevertheless, the acceptable discrepancy of GPS devices has never been determined. Instead, it differed according to the objectives and the application of each device. For example, ground vehicles’ navigation system needed few-meter accuracy, whereas the land surveying drone needs centimeter accuracy because the pictures should be in high resolution. When the pixels were smaller, less discrepancy was essential. Therefore, for Waypoint Tracking and Stitch Image, which needed more precision, Real-Time Kinematic (RTK) devices with more accuracy were used instead [11].

After many problems with instability GPS signals were reported, many inventors have developed algorithms to be used with such devices to increase their efficiency for the most satisfying results [12]. Upon further literature reviews, there have been many published articles on hardware development via algorithms. Among those, there were three interesting experiments that were in accordance with the mentioned principle. The first experiment was the estimation and improvement of GPS coordinates in UAV to be more stable via an algorithm called Kalman Filter (a command set estimating possible data via variance variables (noise covariance) of sensors [13]) with GPS and Barometer based on Position-Velocity-Acceleration model [14]. From this experiment, it was found that the PVA method using Kalman Filter generated relevant results for further studies related to Extended Kalman Filter (the further research of the owner of this experiment).

Kalman Filter was also further developed for more specific purposes; for example, Extended Kalman Filter (EKF) was further developed from Standard Kalman Filter to estimate variant data as Non-Linear [15,16]. The second experiment presented here was the integration of EFK with small UAV, UWB (Ultra-Wild Band), cheap MUI devices, and indoor vision-based sensors [17]. The result of this experiment showed that the EKF application generated approximately a 10-centimeter discrepancy from actual positions. This result proved that hardware GPS positions could be improved by optimizing results via a software algorithm and further improved efficiency.

In addition to using Kalman Filter to stabilize GPS devices, the other interesting method was averaging outcomes from more than one GPS device for more statistically accurate results and for the inaccuracy distribution to be nearer to Normal Distribution compared with using only one GPS device [18]. The third experiment was using Extended Kalman Filter with several Low-cost GPS receivers to increase the efficiency of GPS devices on UAVs by installing one u-box GPS receiver on each arm of the quadrotor, and one more on the center point, totaling five receivers. This experiment showed that the outcomes were more reliable than using only one GPS device [19].

III. METHODOLOGY

A. Structure

For this experiment, two types of equipment were used to collect data from GPS: a Microcontroller for processing and presenting data and GPS receivers for receiving data from a satellite with Arduino Mega 2560 R3 and GY-GPS6MV2 [20]. GPS modules were divided into three serials and received signals via Rx Tx pin, as shown in Fig. 1.

Data were collected by running commands via Arduino IDE and then logged into a Serial Monitor. Each round of data collection lasted approximately 2 – 3 seconds. There were 3 sets of data collected. Each set composed of Latitude, Longitude, and True Altitude (the height above mean sea level) collected from each GPS receiver.

After 30 seconds of GPS sensor calibration, the GPS receivers then started to collect data from the satellite. Each loop started after a 2-second delay since the average time measured before the experiment was 2 seconds.

There were two main scenarios for data collection; one was when the sensors were completely still (no movement), and the other was when the sensors were moving for at least 30 seconds. All data were later used for average calculation and Kalman Filter implementation in order to decrease discrepancy, as shown in Table I. The results from all six methods were analyzed to determine the most efficient method for data stability, while all three GPS receivers were entirely still with no sensor movement.

<table>
<thead>
<tr>
<th>N o</th>
<th>Method</th>
<th>No Movement</th>
<th>Moving</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Implementing Kalman Filter</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>Measuring for an average</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>Implementing Kalman Filter and then</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>measuring for an average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Implementing Kalman Filter and then</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>measuring for an average for every 2 to N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>term</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Measuring for an average and then</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>implementing Kalman Filter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Measuring for an average and then</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>implementing Kalman Filter for every 2 to N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>term</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Circuit Diagram Chart for GPS Signal Receiver.
B. Standard Kalman Filter Approach

Kalman Filter is a set of computer commands used to predict possible outcomes of linear equations based on estimations from Mean Square Error from historical data.

This experiment also used Kalman Filter with the same algorithm as the previous study [21], which consisted of three steps: 1) Initialization: initialized the variables used in prediction, 2) Prediction: calculated data for the possible outcomes, and 3) Update: currently collected data for the prediction of the next set of data. Prediction and Update functioned together recursively for data prediction using Kalman Gain as variables determining the future’s possible outcomes would be according to the current data. The equations for Standard Kalman Filter Data Prediction were:

\[ \hat{x}_{k-1} \] and \[ P_{k-1} \]--------------(1)

\[ \hat{x}_k^- = A\hat{x}_{k-1} + Bu_k \] \[ \hat{x}_k^- \]--------------(2)

\[ P_k^- = AP_{k-1}A^T + Q \] \[ P_k^- \]--------------(3)

\[ K_k = P_k^-H^T(HP_k^-H^T + R)^{-1} \] \[ K_k \]--------------(4)

\[ \hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \] \[ \hat{x}_k \]--------------(5)

\[ P_k = (1 - K_kH)P_k^- \] \[ P_k \]--------------(6)

The objective of these equations was to find an estimated value of the data at time \( K \), aka \( \hat{x}_k \), based on data collected from the current time \( z_k \) according to \( K_k \) (Kalman Gain). This was a crucial variable that varied directly with data from the past (1). Equation (2) and (3) were Prediction State where roughly estimated data were stored in \( \hat{x}_k^- \) (Prior estimate) and \( P_k^- \) (Prior error covariance) before being used later in Update, which was related to (4), (5), and (6) to finally yielding results in the estimate called \( \hat{x}_k \).

IV. EXPERIMENT

Upon turning on the signal receivers, the calibration took 30 seconds before data collection started. Every collected GPS data would be displayed on the Serial Monitor of Arduino IDE. The data collection lasted at least 30 minutes, timed by a time switch. Every read needed approximately 2 – 3 seconds, and after which, all collected data were stored for further calculation.

The first experiment was to collect GPS data while the sensors were completely still. This experiment's location was the open space near the reservoir with no high buildings within a 100-meter radius from the receivers' position. The experiment was conducted at around 5.30 pm, during a clear sky with no visible cloud. The total time spent was 33 minutes and 2 seconds.

The following experiment was to collect GPS data while the sensors were continually moving. The location for this experiment was in the city, surrounded by no higher than 4-story buildings. The experiment was conducted at around 5.32 pm, during a clear sky with no visible cloud. The GPS receivers were sticking out from a backpack while the backpack carrier walked for 2.76km with the average speed at 7 – 8 m/hr, referring to Nike Run Club Application. The total time spent was 41 minutes and 30 seconds.

There were two rounds of data collection, one when the sensors were completely still and the other when the sensors were continually moving. For the one when the sensors were completely still, there were 663 sets of data collected, while for the one when the sensors were.

The picture on the left of Fig. 2 showed the location of 663 sets of GPS data read from all three sensors with no movement after visualizing on Grafana Application. This data collection lasted 1,982 seconds, or 33 minutes and 2 seconds. On average, each read took 2.99 seconds.

The GPS data read from each of the moving sensors was shown in the right picture of Fig. 2. This data collection contained 839 sets of GPS data and lasted 2,326 seconds, 3 minutes, and 46 seconds. On average, each read took 2.77 seconds.

The altitude above sea level read when the sensors were completely still and when they were constantly moving were represented by red, green and blue lines, respectively. On the left of Fig. 3, the range of the sensors with no movement was at 16.5 meters, with the lowest at 316.7 meters and the highest at 343.2 meters. Meanwhile, the range of the moving sensors was at 39.8 meters, with the lowest at 267.6 meters and the highest at 307.4 meters, as shown on the right of Fig. 3.

The collected data would then be calculated for the distribution of data, namely the Maximum, Minimum, Range, Standard Deviation, Mean Deviation, and Variance, derived from each sensor.

The distribution of Latitude and Longitude were shown in Table II, and True Altitude (Altitude Above Sea Level) was shown in Table III. However, the experiment with moving sensors was not calculated for the distribution of data because the actual position was changed continuously, meaning that all data could not be used to find the current location's distribution.
Latitude, Longitude and True Altitude were calculated to improve the stability of data using six methods. From all of the six methods, there were three interesting methods when applied to all ten scenarios as presented here.

A. Implement Kalman Filter to Data at a Certain Time and then Measure the Averages

This method conducted two calculations. The first calculation was implementing Kalman Filter to data from each sensor since it was found in previous studies that Kalman Filter could lower discrepancy to a certain level. However, the results were not efficient enough to stabilize the data [11]. Therefore, the second calculation for this method aimed to elevate the data improvement by measuring the averages using (9).

\[
\text{Lat}_{\text{avg}} = \frac{\sum_{i=1}^{n} C_{\text{lat}_i}}{n} \quad (1)
\]

\[
\text{Lng}_{\text{avg}} = \frac{\sum_{i=1}^{n} C_{\text{lng}_i}}{n} \quad (8)
\]

\[
\text{Alt}_{\text{avg}} = \frac{\sum_{i=1}^{n} C_{\text{alt}_i}}{n} \quad (9)
\]

The results from (7), (8), and (9) were the GPS locations and altitudes when the sensors were completely still as shown in Fig. 4.

The purple area was the one where Kalman Filter was implemented before measuring the averages. It was noticeable that the area was narrower compared to the other three sets of unprocessed data from three sensors due to the decreased data distribution. Table IV showed the statistical data with significantly decreased deviation compared with unprocessed data in Table II and Table III.

B. Measure the Averages, and then Implement Kalman Filter

This method was similar to the first method, and the difference was only that each set of data from all three receivers was used to calculate the averages before implementing Kalman Filter.

Even though the Ranges of latitude, Longitude, and True Attitude of this method were the same as the first method, this method’s statistical variance was significantly lower. As shown in Fig. 5, the second method’s data distribution was remarkably similar to that of the first method, making it hard to distinguish via observation. From Table V, the variances of the GPS positions of both methods were slightly different, while the altitudes bore no difference at all at two decimal places.
C. Use the Data to Measure the Averages and then Implementing Kalman Filter for Every 2 to N Terms

It was found that using GPS data to calculate for averages before implementing Kalman Filter yielded better results; therefore, for this third method, every N Term was measured for averages before Kalman Filter was implemented, N being the Interval Number of data calculated for averages. For example, if N = 3, the system would read GPS data 3 times and then used these three values for calculation. The average gained from each GPS receiver were then added together and divided by the number of receivers (3 in this particular case) to find the average of Multiple Sensors, which were then implemented with Kalman Filter. This method aimed to observe the tendency of data in the case that the Interval of finding averages kept increasing while the GPS receivers bore no movement.

Table VI found that data distribution tended to keep decreasing when N (average Interval) increased. Upon checking the range of distribution, when Interval equaled 2, 3, and kept going to the total number (N), it could be seen that for every increasing N, the standard deviation decreased and tended to keep decreasing. The change of the graph's trend was noticeable in Fig. 6, which showed the comparison of data calculated with this method with average calculation at 2, 3 intervals and from 1 to 663 terms.

For the case that the equipment was constantly moving, this method calculating for averages from 1, 2, to 663 terms would not be used. This was due to the fact that, when the Interval of the averages were increased, the data would start moving towards the center of the data as shown in Fig. 7 where the path of data at Interval 1 to the total number at 663 sets for the case that the equipment was continually moving. The Purple Line and the Blue line represented calculations with both Kalman Filter and Average Measurement. It is evident that when time passed, the path was compressed towards the center of the data. Therefore, this method was not used for moving equipment.
V. RESULT ANALYSIS

This experiment demonstrated three methods used to lower the distribution of data: 1) Implementing Kalman Filter, 2) Finding Averages and 3) Implementing Kalman Filter and Finding Averages. From the distribution data shown in Table VII, the least effective method was solely implementing Kalman Filter. The Standard Deviations of GPS Coordinates (Latitude, Longitude) while the GPS receivers were completely still were decreased by 3.39% and 10.7% on average; whereas the Standard Deviations of True Altitude were decreased by 2.99% and 6.23% for the non-moving equipment and the moving equipment, respectively. Even though it is evident that Kalman Filter could help reduce the distribution of data, but its efficiency was too low to be used with projects which needed data stabilization, as previously mentioned in earlier studies regarding Kalman Filter [11].

Next up was the method where data were used to find averages. This method significantly increased the stability and decreased data distribution better than the one implementing only Kalman Filter. The Standard Deviation of Latitude and Longitude for non-moving equipment were decreased by 49% and 75.21%. The Standard Deviation of True Altitude for non-moving equipment and moving equipment was decreased by 65.45% and 40.94%, respectively. They were resulting in more stability compared to the method solely implementing Kalman Filter.

Implementing both Kalman Filter and average measuring to improve data stability could be further divided into four sub-methods: 1) They were using results after implementing Kalman Filter to find averages, 2) using Kalman Filter results to find averages of every data from 1 to N loop, 3) using data after finding averages to implement Kalman Filter, and 4) using the averages of every data from 1 to N loop to implement Kalman Filter. Based on all these sub-methods statistical data, it was found that implementing Kalman Filter and average findings could better stabilize the data compared to applying only one method. From Table VII, in the case that Kalman Filter was implemented before average measuring with non-moving equipment, it was found that the Ranges of Latitude, Longitude, and Altitude for this particular method was narrower at 0.23, 0.55, and 0.38 meters, respectively, when compared with the method with average measuring only. The tendency of lower data distribution was similar for the case with moving equipment. For the case with non-moving equipment, using more loops to find data averages yielded more stability. As time passed, every increasing Interval of average finding statistically significantly lowered the distribution of data. However, the method of finding averages before implementing Kalman Filter yielded more stable data distribution when N increased compared to implementing Kalman Filter before finding averages. However, statistics showed that when the equipment was entirely still, finding averages and then implementing Kalman Filter at any N, the variances were so close to solely implementing Kalman Filter at any N that the differences were unnoticeable with bare eyes. Similarly, with moving equipment, data measuring for averages before implementing Kalman Filter yielded slightly higher variances compared to the other method; therefore, hardly bearing any effect upon implementation. Nevertheless, the method of implementing Kalman Filter together with measuring for averages with increasing loops was incompatible with the case of moving equipment since the average of a specific position at any time required data from that particular position; otherwise, the results would be incorrect as shown in Picture 7. For example, every loop required 10 meters of a straight line. If data from the current position were combined with data from the previous position 10 meters away and calculated for an average, the result would be the 5-meter average between these two positions, which was 5 meters away from where it was supposed to be. This was the reason why calculations with average loops were unsuitable to be used with moving equipment to lower the variances of GPS positioning data.

### TABLE VII. STANDARD DEVIATION AND RANGE OF DATA FROM EACH CALCULATION METHOD

<table>
<thead>
<tr>
<th>Method</th>
<th>No Movement</th>
<th>Moving</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GPS Coordinate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S.D. (Lat, Lng)</td>
<td>Range (m)</td>
</tr>
<tr>
<td>Raw data</td>
<td>0.0000084, 0.0000159</td>
<td>6.12, 6.78</td>
</tr>
<tr>
<td>KF</td>
<td>0.0000082, 0.0000156</td>
<td>5.67, 6.23</td>
</tr>
<tr>
<td>Find average</td>
<td>0.0000043, 0.0000039</td>
<td>2.34, 2.22</td>
</tr>
<tr>
<td>KF + Average</td>
<td>0.0000042, 0.0000035</td>
<td>2.11, 1.67</td>
</tr>
<tr>
<td>KF + Average every 2 terms</td>
<td>0.0000042, 0.0000035</td>
<td>2.11, 1.67</td>
</tr>
<tr>
<td>KF + Average every 3 terms</td>
<td>0.0000041, 0.0000035</td>
<td>2.11, 1.67</td>
</tr>
<tr>
<td>KF + Average every 1…663 terms</td>
<td>0.0000012, 0.0000011</td>
<td>0.78, 0.55</td>
</tr>
<tr>
<td>Average + KF</td>
<td>0.0000041, 0.0000035</td>
<td>2.11, 1.78</td>
</tr>
<tr>
<td>Average + KF every 2 terms</td>
<td>0.0000040, 0.0000033</td>
<td>1.89, 1.56</td>
</tr>
<tr>
<td>Average + KF every 3 terms</td>
<td>0.0000038, 0.0000031</td>
<td>1.78, 1.45</td>
</tr>
<tr>
<td>Average + KF every 1…663 terms</td>
<td>0.0000011, 0.0000011</td>
<td>0.78, 0.55</td>
</tr>
</tbody>
</table>
VI. CONCLUSIONS

It can be concluded from this experiment that measuring for averages together with implementing Standard Kalman Filter to three sets of GY-GPS6MV2 Modules to improve the stability of cheap GPS equipment can indeed help reduce the variances of data both when the equipment is constantly moving and when they are completely still. The most effective method is measuring for averages before implementing Standard Kalman Filter. For the case with non-moving equipment, the increasing average loops can lower the variances, whereas, for the case with moving equipment, the increasing average loops reduce data reliability. Even though the increasing loops for average measuring help reduce data variance, it directly varies with time spent collecting data; in other words, the more loops for average measuring, the more time needed for data gathering for return output. From the result, there are limitations of moving measurement. The algorithm will slow down the reading cycle to calculate an average and filter of each reading. That would be the primary direction for future research to overcome these limitations.

In conclusion, this experiment has proved that integrating Standard Kalman Filter with average finding for multiple consumer-grade GPS equipment is another suitable alternative for projects that need to reduce variances from GPS equipment at a lower cost. This innovation can elevate data management with variance through computer commands for technological science and geoinformatics. For example, it can be used with the guidance system searching for missing persons, improving the small projects with customer-grade sensors, or being used to develop future technology and so on continuously.

REFERENCES