Lab 1: Measuring the Real World

Ryan Lake

Introduction

Modeling real world systems can be fraught with difficulties, but it is nonetheless an essential step to understanding complex real world situations. It is perhaps because of this that the study of the design of models and visualizations has recently become a large focus among the scientific and other scholarly communities. Of course, modeling cannot be accomplished without data, and if one is attempting to learn how best to build models, simple data sets, such as measuring the area of a parcel of land, would be preferred.

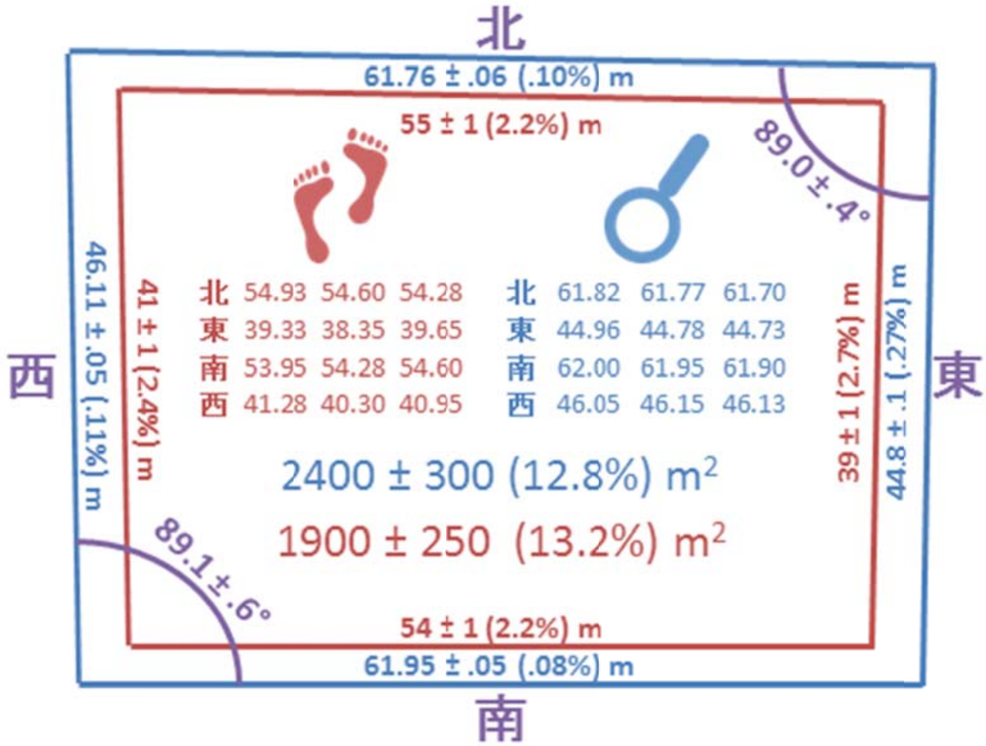
For this write-up, the area of a predetermined "rectangular" parcel of land was approximated using four different measuring tools. The collected data was then used to construct two models with the intent of not only representing the real-world land parcel, but also to allow for easy comparisons between the four different measuring methods and to highlight the relative and absolute errors of each. In constructing the graphics, "data density" was given precedence to allow as much data as possible to be represented in as small an amount of space as possible, while still maintaining relatively high legibility. Also, as previously mentioned, the models were designed so that similar types of measurements could be easily compared on the same "surface" (which in this case, is within the same graphic). Finally, the models were constructed in such a way that anyone who views them independently of any other explanation should be able to understand what they represent, even if not immediately obvious.

Data Collection

The parcel of land in question most resembled a rectangle, but in an attempt to minimize any inherent assumptions, it was not treated as such for these calculations. The parcel can instead be thought of as two triangles, divided through two opposite vertices and the areas of these triangles can be added together to find the total area. Because of this, the angles of two opposite vertices (the remaining two vertices that were NOT used to divide the area into triangles) were measured with a protractor (three separate times). The area of each triangle can then be determined with the following equation:

$$\text{Area} = \tfrac{1}{2}\,\text{Leg}_1 \times \text{Leg}_2 \times \sin\theta$$

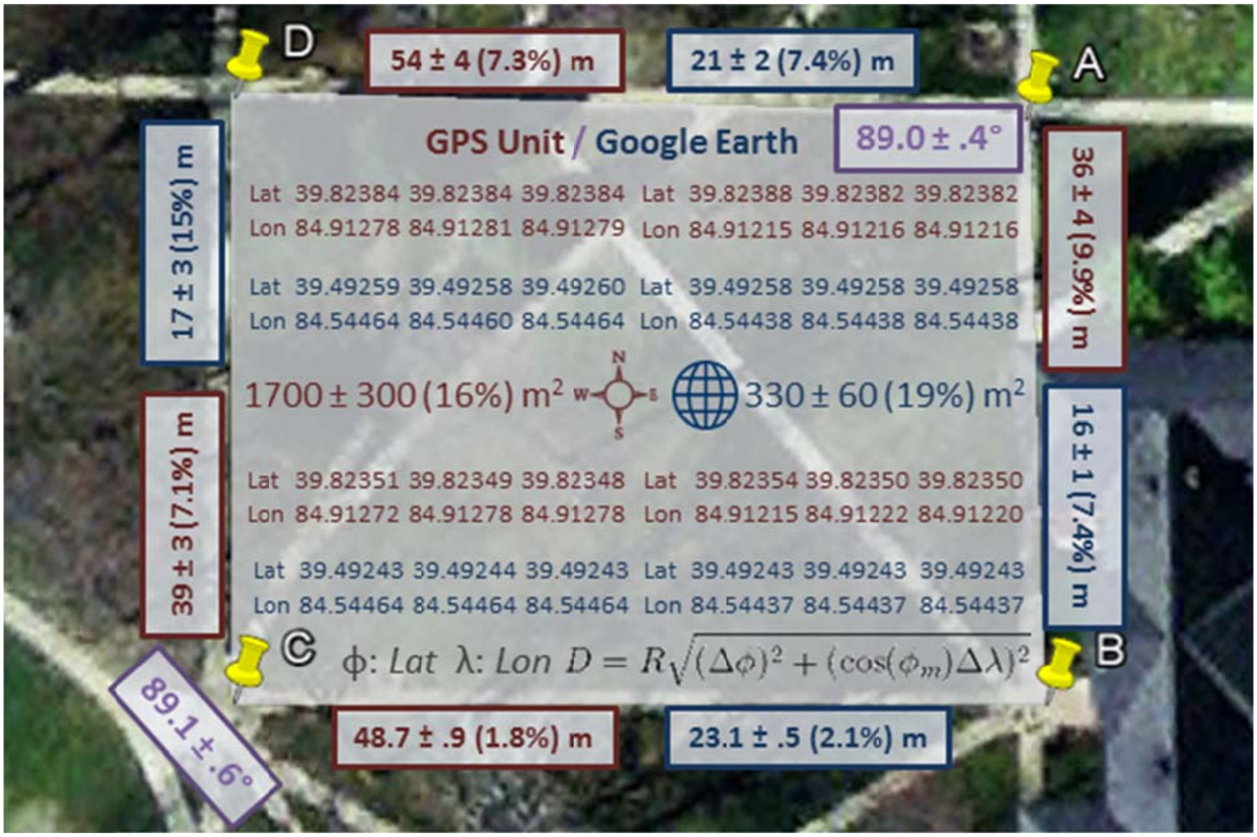
The same value for the two angles, θ, were used throughout the experiment, but in order to get an understanding for real world measurements, the lengths of each side were determined using four different tools. Two of these directly measured lengths: "pacing" the distances and a measuring wheel. The other two methods measured the "position" of each of the vertices and determined the distance between these points indirectly: a handheld GPS device and Google Earth. Because of the differing natures of these measuring devices, two different models were created to allow for a more accurate representation of the two techniques. For each of the four tools, measurements of each side/vertex were taken three separate times to allow for better error analysis.



"Pacing" the distances involved counting the number of steps it took to walk along each side of the parcel of land. This method therefore had a high likelihood of human error as each "step" will not be the same size and also because the smallest unit of measurement was "half-a-step". To convert the number of steps into a more orthodox unit of measure, a tape measure was placed on the ground, and ten steps were taken to ascertain the number of meters in one step. This process was repeated five times and the results averaged to obtain a more accurate conversion ratio.

A measuring wheel was also used to directly measure the length of each side of the parcel of land. This had the advantage of more precise measurements as every foot and inch measured should be considered to be "exactly" the same length. However, it was much more difficult keeping the wheel traveling in a straight line and it was prone to bumps and jolts which may have affected the recorded measurements.

As can be seen in the above model, the measuring wheel ultimately had a much higher precision (lower relative error) than the "pacing" method. It should also be noted that each measurement taken by the measuring wheel was significantly larger than that determined by "pacing". The calculated areas are ultimately within the accepted error of the experiment, but such a systematic difference leads one to believe that there was in fact some sort of systematic error in one of the methods of measurement. Judging by the data, there is more likely to be a systematic error in the "pacing" method because of both the uncertainty in the base measurement and also because there is an additional level of uncertainty in the process of converting "paces" into meters.



To calculate lengths using a GPS device, the global positioning measurement at each of the four corners was taken, again three separate times for each corner to better account for error. From the averages of these readings, the distance between each of the points could be determined. Because this experiment assumes that this area of land is flat and the curvature of the earth is negligible for an area of this size, the difference between the longitude coordinates must be corrected (by multiplying by cosine of the mean of the latitude coordinates). The exact equation is detailed in the above graphic. The radius of the earth, R, was assumed as 6,371,009 m. This method was rather straightforward and perhaps suffered the least from human error.

In a similar manner to the GPS, Google Earth was also used to determine the global positioning measurement of the four vertices. This was the only method that did not involve measuring something in the field directly – instead it relied upon the calculations of Google to estimate each position, and as such there was not much that anyone who didn’t work for Google could do to minimize error. As can be seen in the above graphic, the GPS and Google Earth methods had about the same measurement precisions, but they ended up with drastically different distance values. At first glance this may be a tad bewildering as the values taken by the two different methods don’t seem to be that far off, but it must be remembered that when considered on a global scale, even small discrepancies can account for large overall errors. The calculated area for Google Earth is therefore absurdly small, but the GPS device calculated area is within one standard deviation of the “pacing” method and within two standard deviations of the measuring wheel method, which means that it is still a viable method of calculating distance, although it may be better suited for larger areas.

Error Analysis

In addition to the errors listed previously for each of the different measuring devices, it would be prudent to acknowledge general errors present throughout the entire experiment, especially those due to assumptions. As mentioned in relation to the calculation of distance from GPS coordinates, the parcel of land was assumed to be flat. Not even considering any irregularities which are almost certainly present, the earth is, in fact, a sphere and so no area of land is in truth “flat”. However, it is hoped that any discrepancies due to either of these causes is negligible compared to the precision with which the lengths were able to be determined. Similarly, since the area in question was predetermined to be bounded by manmade sidewalks, it was assumed that the four sides would be straight. Upon closer examination, this was obviously not the case (although still a good approximation) but seeing as it was difficult enough to measure a relatively “straight” line with the tools used, any attempt to take into account any curvature of the bounding edges wouldn’t be worth the trouble.

Conclusion

As this experiment demonstrated, even seemingly menial tasks, like determining the area of a small parcel of land, can be difficult when any sort of accuracy or precision is desired. Similarly, this experiment also demonstrates the wide variety of tools and systems of measure that can be utilized during an experiment, and ultimately the data obtained can never be better than the inherent accuracy and precision of the tool used. Finally, graphical models are an invaluable tool to help understand data sets, from something as simple as determining what method is beast in determining the area of a small parcel of land to understanding how and when Napolean’s invasion of Russia in 1812 ultimately failed.