Looking into potential issues with GPS devices

David Jansen, Jake Gordon, Niki Learn & Beth Archie

2022-12-15

After the team started to use some new GPS devices (Garmin GPSMAP 64st) in 2019 we noticed that there were some issues with the accuracy. In this document we look at this and visualize some of these issues.

The new Garmin GPSMAP 64st is the first GPS in the field that can use a combination of GPS and Galileo. We found out that in May-Aug of 2019 Galileo was having accuracy issues. The Galileo satellites might have been given the wrong locations (the accuracies would still look good). The system was also not fully operational yet and more satellites were added in 2019 and 2020. Now that Galileo is fulling functional with adequate satellite availability, it should actually have a better accuracy than the GPS system. The issue of the summer 2019 data is not fixable so data should be used with some care.

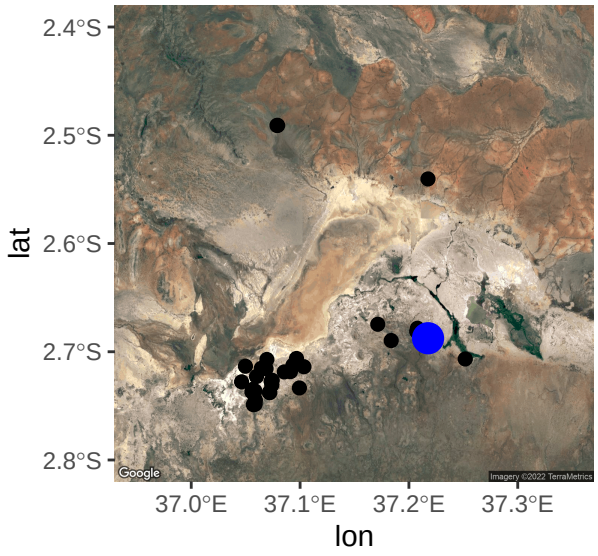
Looking at departing from camp

We first looked at the departure from camp. The team takes a GPS point at every departure from the field. As camp is a known fixed location, it is the easiest point to check for accuracy.

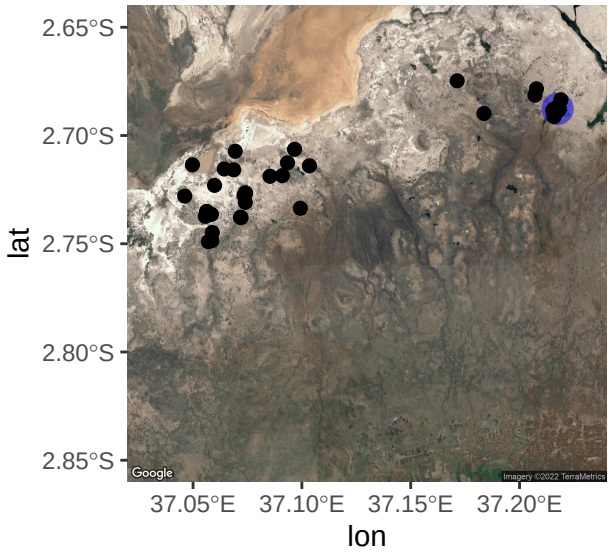
In the maps that follow the blue dot represents camp. All distances are in meters. The location of camp is set as the median of the lon and lat of the locations of departures from camp (x = 301770; y = 9703049; lon = 37.217; lat -2.685).

Visualization on a map

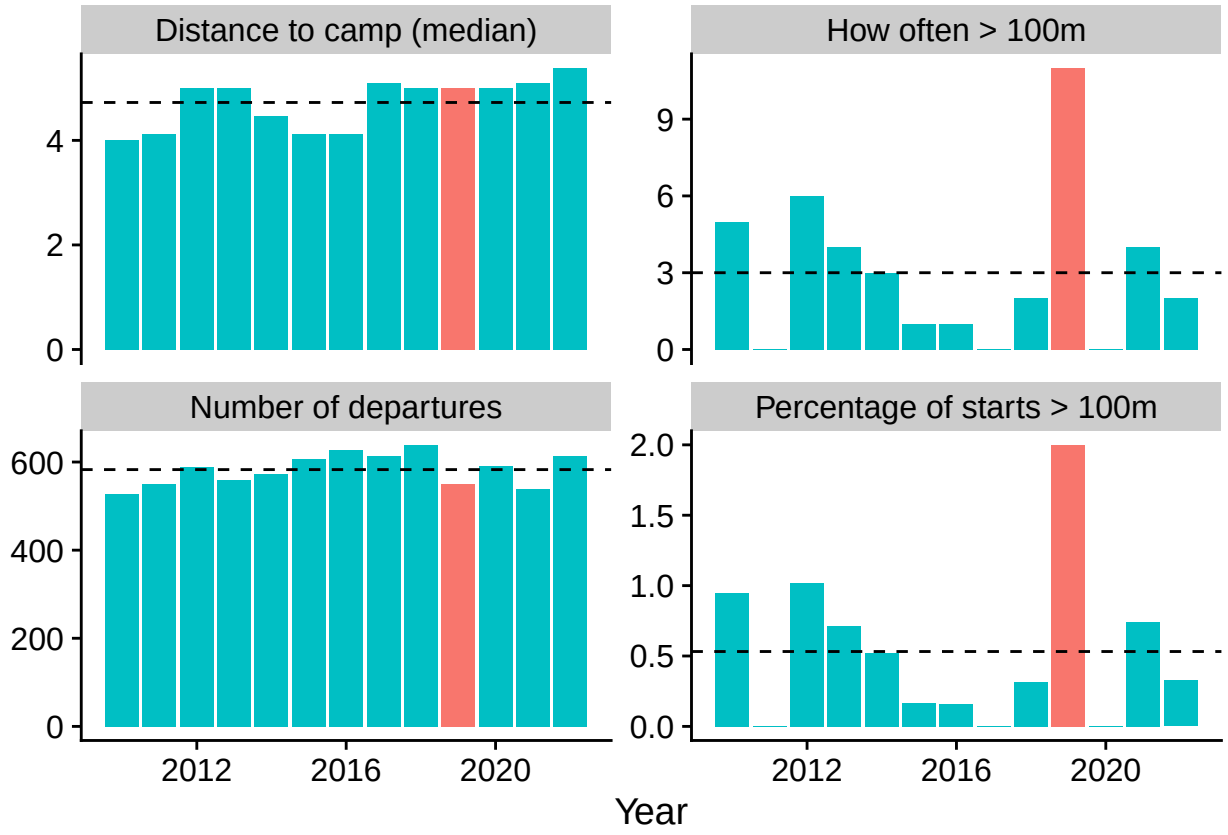
There are a few points that are very far way from camp.



If we ignore those for a moment and look at points that fall into the area that includes camp and baboon groups' home ranges, this is the map. Note that many points in camp overlap, so we can't see all the points.



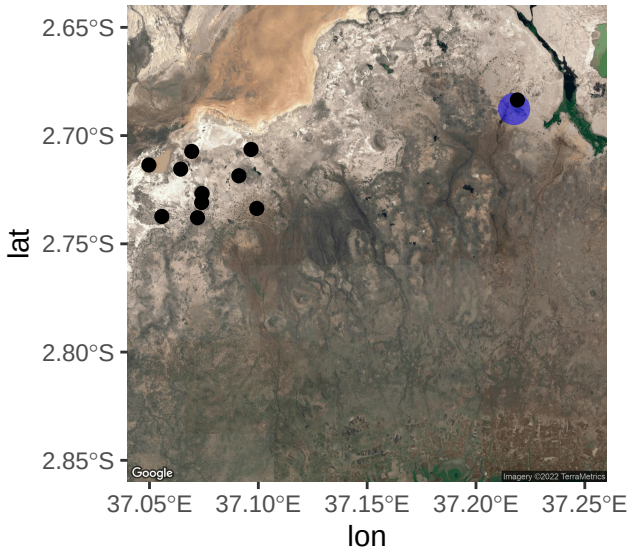
Next we look at the distance to camp and how many of these departures were more than 100m and 2019 is highlighted. The dashed line is the mean of the whole data set.



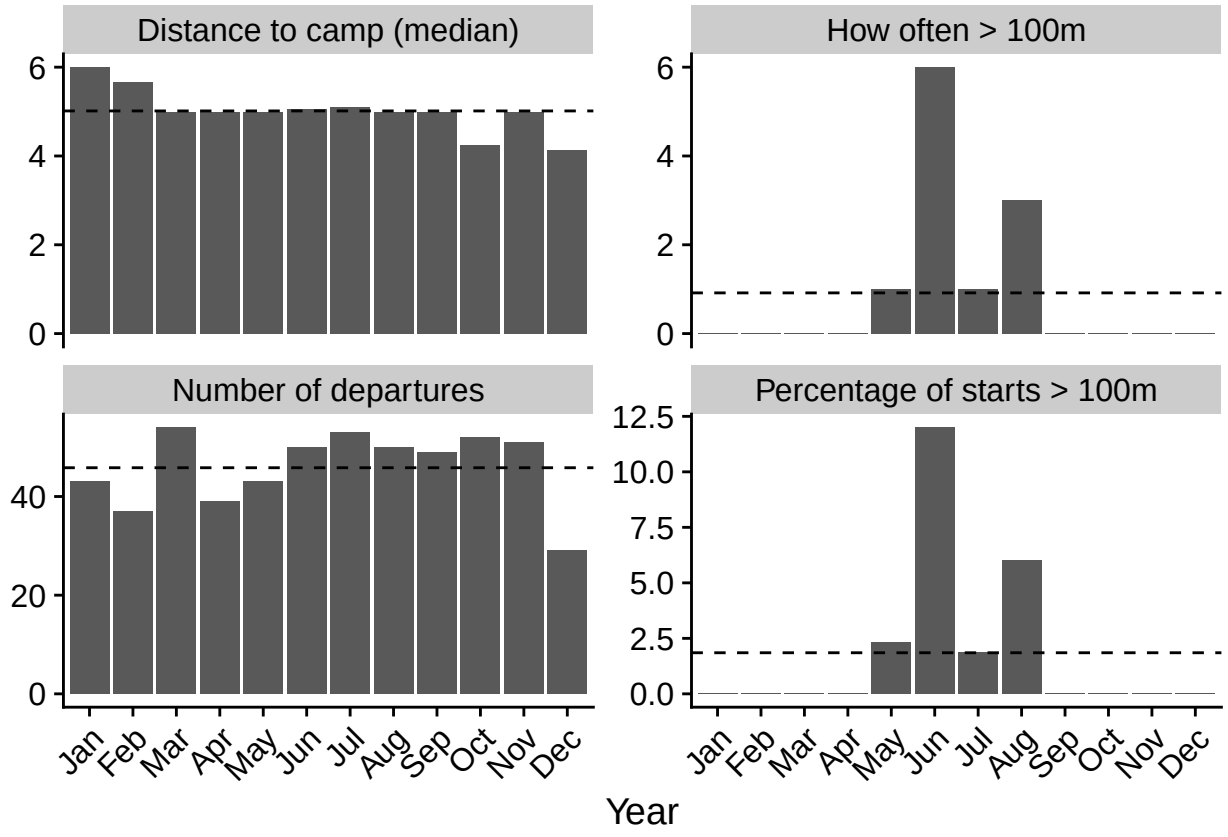
2019 stands out due to the higher number of GPS points that are more than a 100m away from camp. Importantly to note is that these large discrepancies still only account for approx 2 percent of the data.

Focus on 2019

As the problem seemed to appear mainly in 2019 we looked at it in more detail.

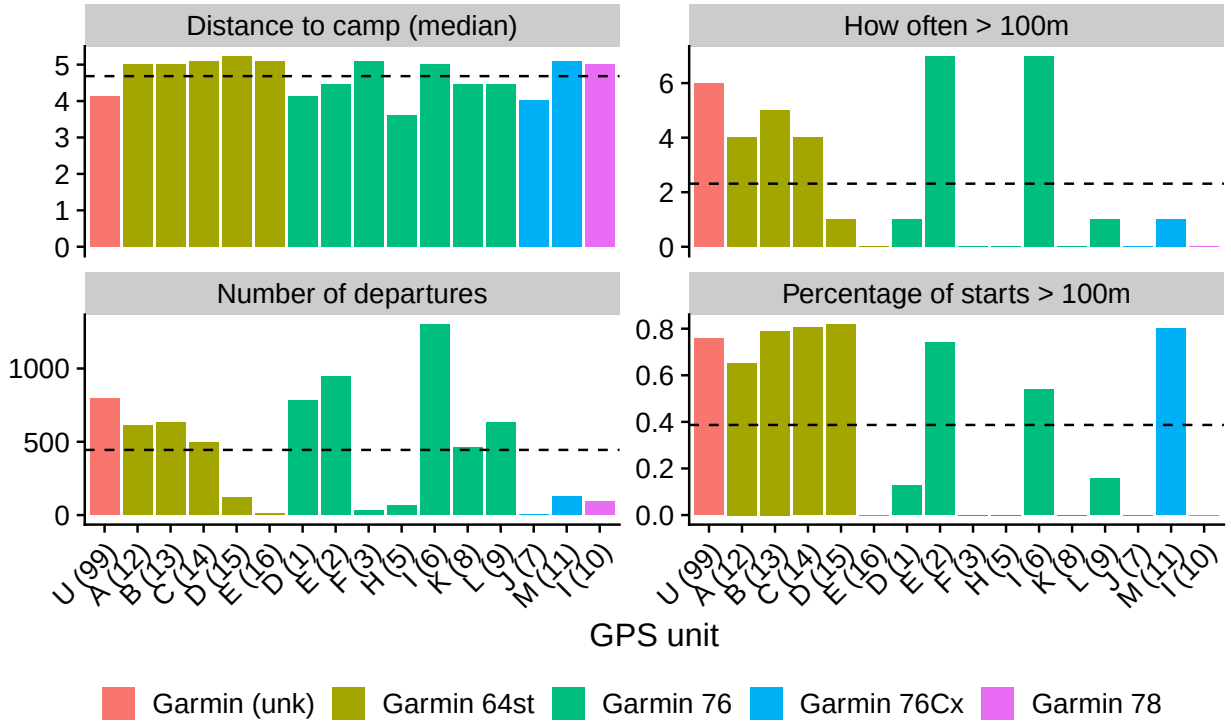


All the departures that were > 100 m from camp occurred between May and August. The problem is especially large in June 2019, where over 12% of the GPS points taken at departure are more than 100m away from the camp reference coordinates.



Overview per GPS unit

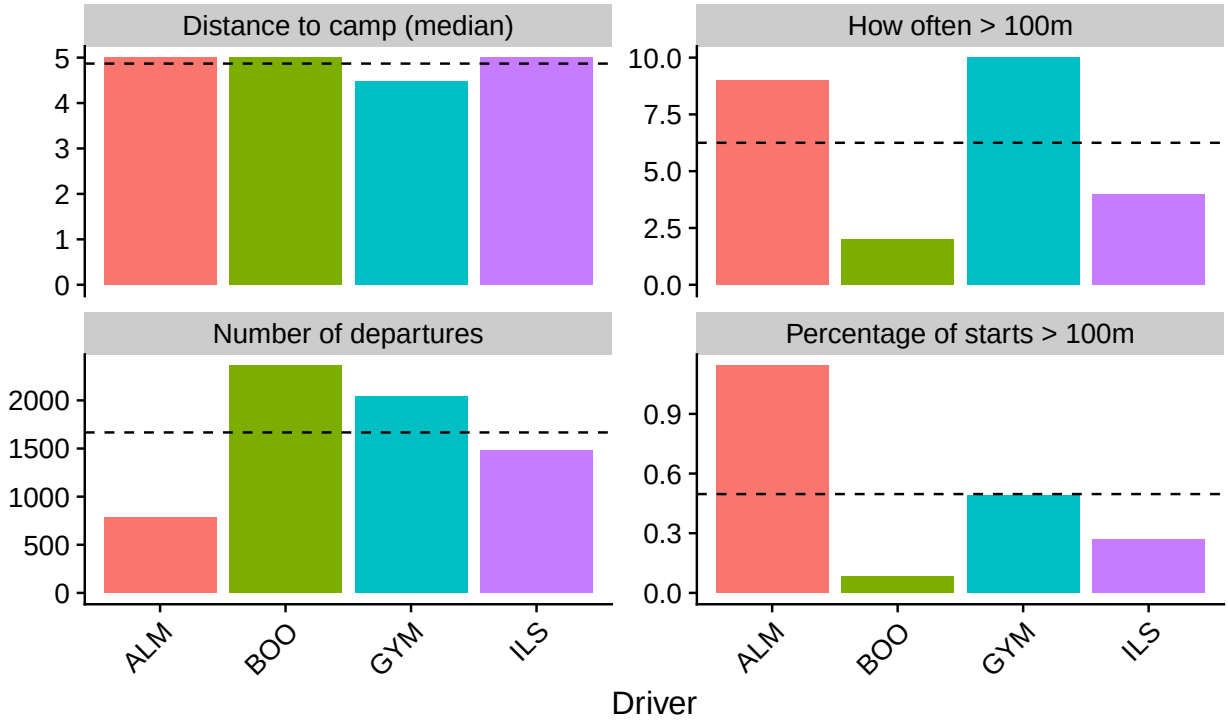
Leaving 2019 and returning to all of the data, we investigated whether the issues were unique to specific models of our GPS units.



While some specific units appear to have higher error rates than others, the only model that appears to be consistently wrong is the Garmin 64st, which was the only model in use in May-Aug 2019.

Overview per driver

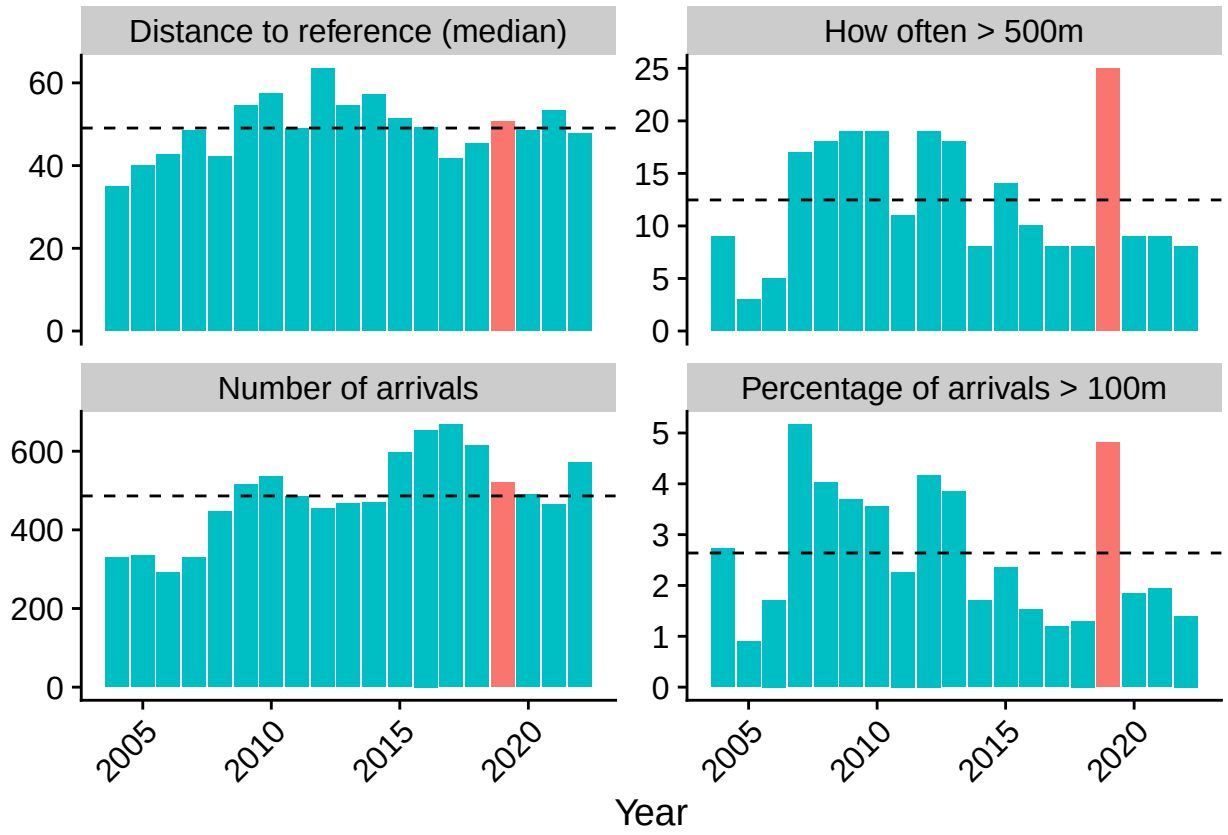
We also investigated whether the identity of the driver,the person who usually collects all GPS points, is related to this issue.



The issue seems to occur more often for some drivers than it does for others, but importantly: all drivers are affected. This does not seem to be a driver-specific issue.

Waterholes and groves

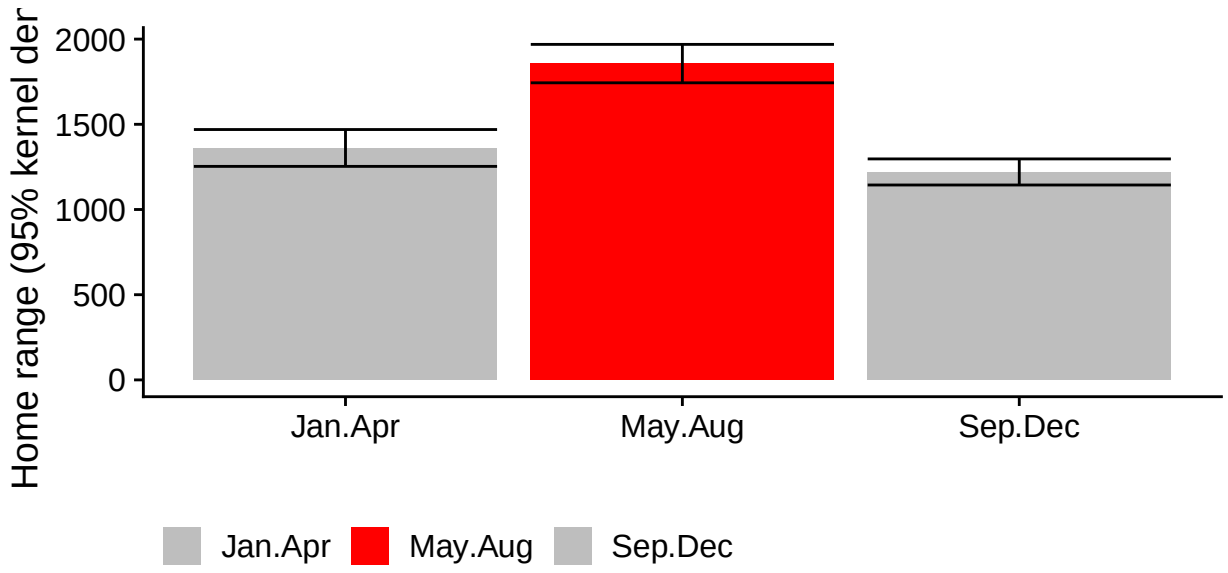
In this section we looked if the GPS disparities where also an issue for the GPS points at groves and waterholes (including the initial arrivals in the field with a know location). Specifically we checked how far the first GPS point was from the reference oordinates for the location the team says they are. Here we used 500m as a cut off, since it is a known issue that the team does not always take the GPS close to the actual waterhole or grove.



Home ranges

For the GPS points that are are during the focals there is no reference points we can use to calculate the error. Therefore we used a diffent approach for these points. Using all GPS points taken during visits to the baboon groups, we calculated home range sizes per group. Here we show the results of the 95% kernel estimation. Kernel density estimators, which map a utilization distribution, are one of the most popular methods for measuring home ranges. The 95% indicates that 95% of the points collected are included. This is to prevent outliers from artificially increasing home range size. If GPS has a lower accuracy the points are likely less clumped and this would lead to a larger estimated home range. As a reminder May-Aug 2019 were the months that were suspicious.

As you can see the problem months in 2019 jump out. It seems very unlikely that the home ranges increased in size just for those few months. The temporal increase in home range is more likely due to more outliers due to bad GPS points and those outliers don't fully get filtered out with the 5% of points are excluded in the kernel density estimation.



Conclusion

As with all of our data, the GPS data are imperfect. Each year's data has some points whose coordinates are implausible in some way that we are unlikely to ever be able to explain. However, May, June, July, and August in 2019 are different; these implausible coordinates occur at a higher rate. We now know can be explained by satellite issues and which do not seem to be explained by any other factor. We leave it to the users to decide how (or if) they want to use the data from these months.