GPS positional accuracy of Seagliders

Jim Bennett & Fritz Stahr
University of Washington, School of Oceanography
August, 2014

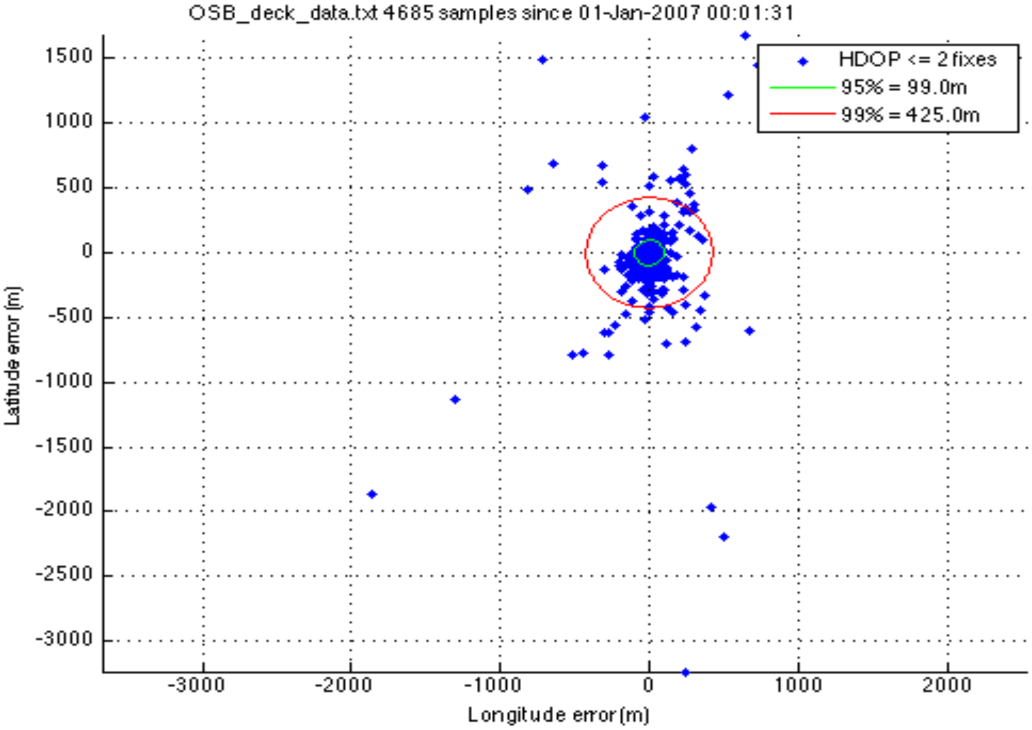
While on the surface, Seagliders acquire positional fixes by powering up a Garmin GPS unit, waiting until it reports acquiring valid fixes (status of 'A' in GPRMC sentence from the unit), then taking the first valid fix with a reported HDOP (horizontal dilution of precision) of <= 2.0. However, the glider code waits for no more than $N_GPS valid fixes or $T_GPS minutes total elapsed time since power up to record a position. The HDOP criteria was chosen because we assumed it corresponds to a positional error of ~10 meters[1]. The value of $N_GPS is typically 20, and $T_GPS is typically 5 minutes. Fixes from the unit are reported roughly once per second.

A review of various glider deployments and simulated dives suggests that the accuracy of glider positions might be poor in spite of using HDOP to indicate an accurate fix.

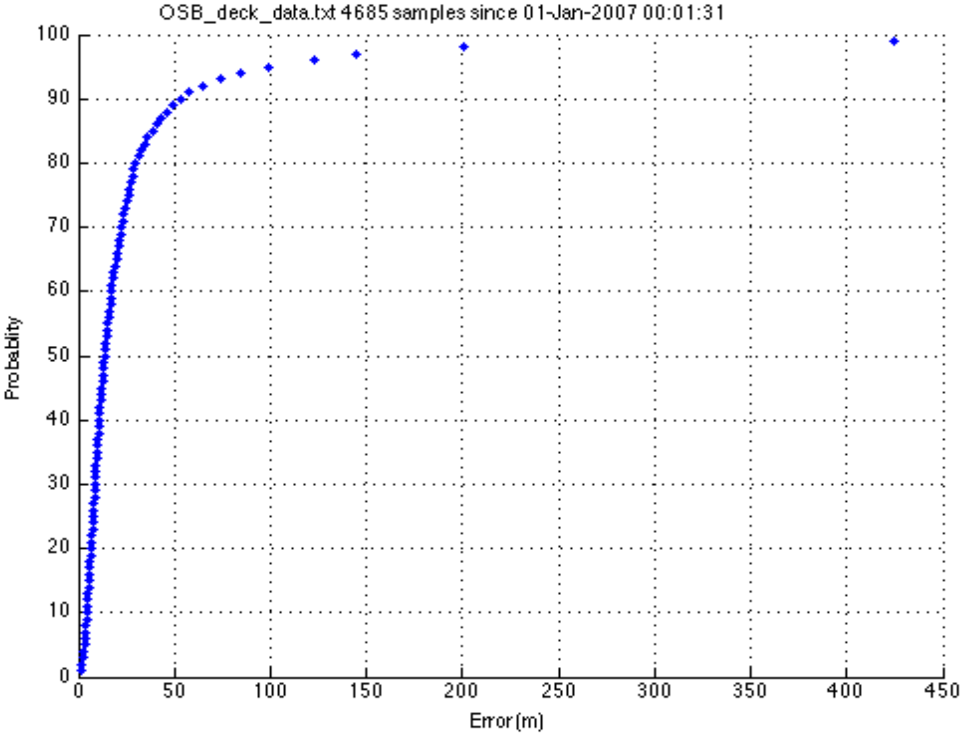
We analyzed several years of simulated dives on the 3rd floor deck of the Ocean Sciences Bldg (OSB) at the University of Washington. These dives exercised a number of different models of GPS units from Garmin (primarily 15H-W, 15xH-W) and Seaglider antennae located within 1-2 meters of one another. The plot below shows the actual error (in meters) from the mean deck latitude and longitude. As the figure below shows, the actual error is often significantly larger than the expected 10 meters based on HDOP <= 2.

---

[1] Our standard formula for computing expected error in meters from HDOP was 3.04**hdop* + 3.57 meters. In contrast to our acceptance criteria, Dana Swift reports the current Argo GPS code takes the first valid ('A') fix with 3 or more satellites; neither HDOP nor HPE is used.



The position errors are not distributed normally, instead falling off exponentially, with 95% of the points within 99 meters but 99% within 425 meters, nearly a half kilometer. These data suggest that this level of positional inaccuracy has been with the Seaglider program since adopting the HDOP cut-off heuristic around 2003.



In addition, the data does not show the expected correlation between HDOP values and actual position error that a cursory reading of the literature implies (data not shown). It appears that HDOP is intended to indicate the maximum *achievable* accuracy given the current satellite configuration, not the *actual* accuracy of any particular fix. As Sarah Webster remarked, this is all unfortunately interesting.

We next ran experiments using a fixed antenna on the OSB roof[2] with older Garmin model 15H and newer model 15xH units. We modified the glider code to ignore the HDOP cutoff, forcing the collection of 60 good fixes ($N_GPS,-60) before latching a final fix. The actual error variance of the final fixes (reviewed in more detail below) was substantially better, typically about 10 meters, indicating that fix accuracy improves with increased acquisition time.

Thus it appears that our current HDOP-based criteria prematurely latches GPS fixes yielding the large positional errors. Our deck statistics suggest that all the GPS data collected during missions to date should assume a 100 meters estimated error rather than the expected 10 meters, and it could be sometimes much worse (this estimate is refined below). Unfortunately, there is no way of knowing on a per-fix basis how inaccurate it might be. Comparing adjacent $GPS1 and $GPS2 fixes can suggest implausible surface drifts but not which of the two fixes is bad, or even whether both fixes are bad.

Inaccurate fixes have implications for depth-average current (DAC) calculations on historical missions. Dives in weak currents require a longer time between surfacing fixes to reduce the effect of the two position errors and hence resolve the current. In general, the minimum DAC we can resolve based on the average expected error, e, of the GPS units is $2e/t$, where t is the time of the dive[3]. Thus, for example, assuming a 100 meter error for GPS fixes the 45 minute dives we typically perform in Puget Sound are unable to reliably resolve currents below 7.5 cm/s (bottom-mounted ADCP measurements suggest maximum flood currents in Port Susan of ~6 cm/s). Typical Seaglider open-ocean dives, however, lasting 8 to 9 hours could resolve currents around 0.6 cm/s assuming the flight model and parameters are accurate[4].

Inaccurate fixes may also explain why recovery crews have been at the reported position of a surfaced glider and yet are unable to see it, or report an acoustic range some distance away.

Improving positional accuracy

---

[2] Observations suggest that GPS fixes via the OSB roof antenna are susceptible to intermittent, severe multi-path effects, likely due to the cycling of nearby HVAC equipment. Actual errors from fixes during these times can reach 1-2 kilometers regardless of HPE and the number of satellites. This interference is easily detected and was absent or removed before the analyses presented here. Fixes acquired on the OSB deck are not affected by this source of error.

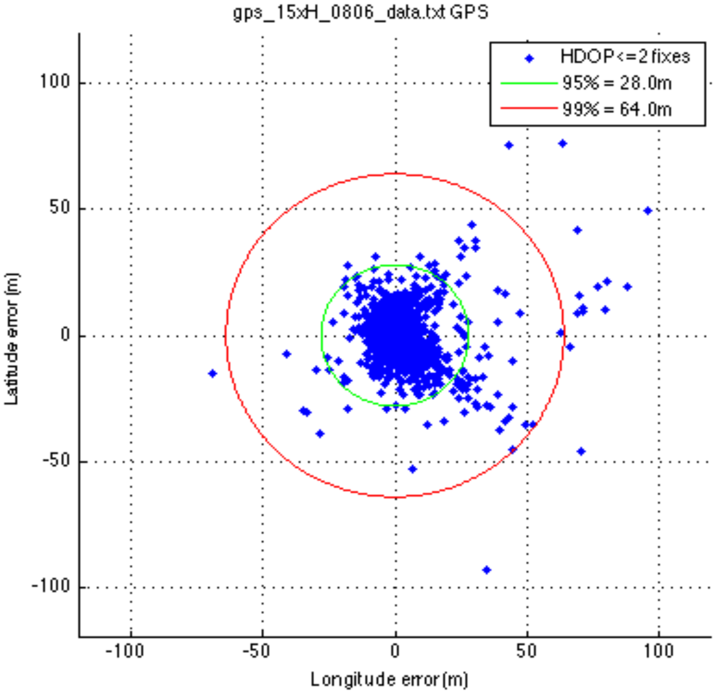
[3] Consider a dive in still water starting at a reported $GPS2 position and surfacing at a $GPS1 position some time t later. Suppose our flight model is accurate, in form and parameters, and it predicts that, starting at $GPS2, the vehicle would indeed surface at the $GPS1 position. In this case we would predict no depth-averaged current, as appropriate for the still water. However, each GPS fix has an expected error, e, in meters. In the worst case, the actual dive position could have been on the side of the error circle away from $GPS1 and the surfacing position could be on the opposite side of the $GPS1 error circle, away from the starting $GPS2 fix. The predicted depth-averaged current in this case is $2e/t$ and is the minimum resolvable current for dives of time t.

[4] The trade-off between an error magnitude and its probability allows us to ask what is the chance, with the historical GPS scheme, that we could resolve, say, a minimum depth-average current of 1 cm/s on a 45 minute dive in Port Susan, assuming an accurate flight mode: 1 cm/s = $2*e$(m)*100(cm/m)/45*60(s) or $e <= 13.5$ meters . This occurs 52% of the time per fix and since both fixes have to be on that order, the probability is $.52^2$ or 27% of the dives. We can also ask, given the conservative threshold of a 1% chance of the error being greater than 100m, what is the minimum dive duration required to resolve a 1cm/s DAC: $t = 2*100$(m)*100(cm/m)/1(cm/s) = $2*10^4$(s) or 5.5 hours.

We next investigated new criteria to improve positional accuracy and estimate the cost in time (thus energy). Currently, our glider code relies on only the standard NMEA strings that all GPS units report (regardless of manufacturer), particularly the GPRMC and GPGGA sentences. The Garmin manual documents a proprietary sentence (PGRME) that provides a horizontal, vertical and total estimated position error (HPE, VPE and EPE, all in meters). This sentence has been a part of Garmin's firmware since ~1999. We modified the glider's code to request that sentence in addition to GPRMC and GPGGA during the forced $N_GPS experiments.

Using a fixed antenna on the roof of OSB we collected data from a model 15H unit running firmware level 2.8 and a newer 15xH running firmware level 3.9[5]. Each data set consisted of ~2000 acquisition cycles, each cycle collecting up to 60 'A' fixes.

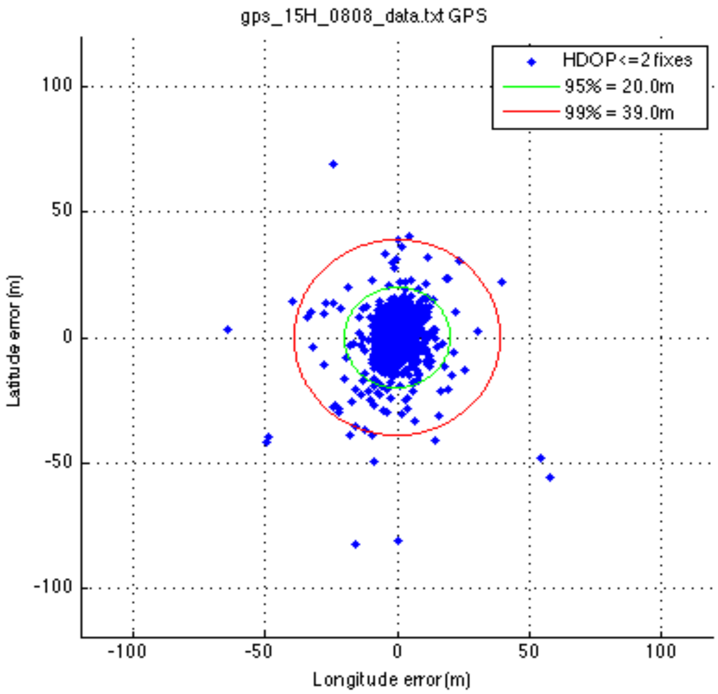
We first confirmed the poor positional accuracy of the HDOP-only scheme by plotting the error of the first 'A' fixes with HDOP <= 2.0 (note change of scale). Surprisingly, the newer 15xH performs worse under the HDOP criteria than the 15H (99% at 64 meters vs. 39 meters)[6].



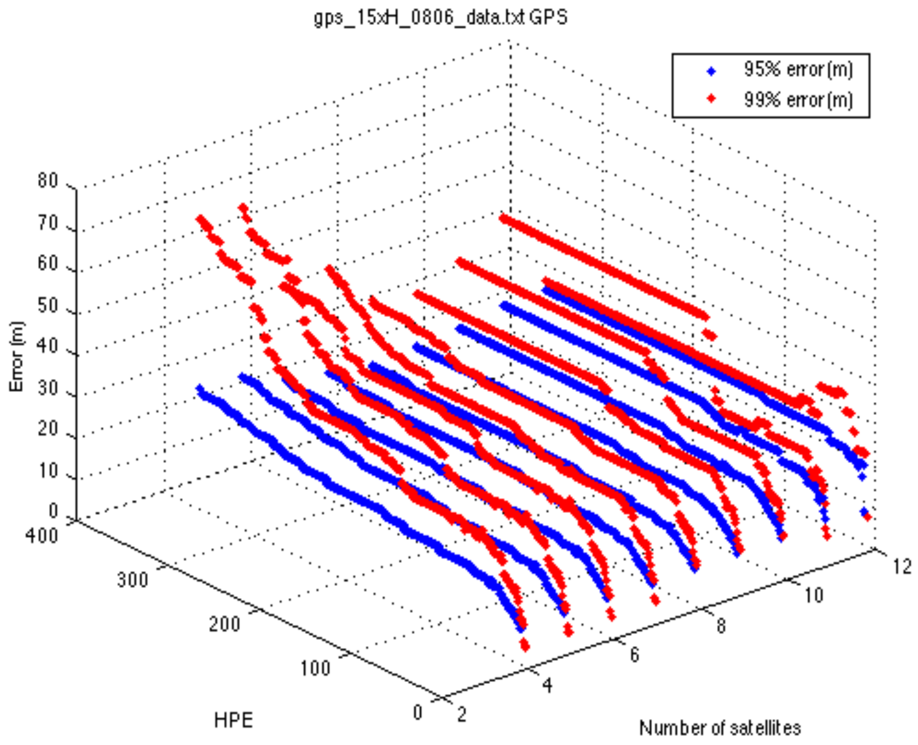
---

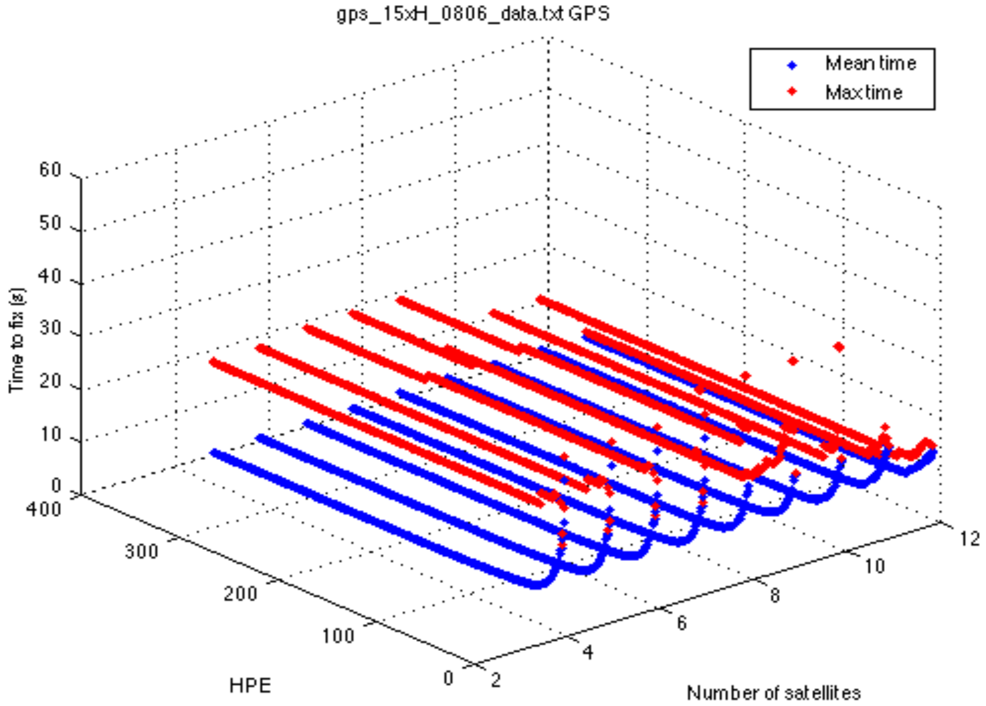
[5] The firmware version numbers from Garmin can be confusing. For the 15xH unit, the firmware version numbers run from 3.0 to 3.9; for the 15H, the version numbers run from 2.0 to 3.3, then jump to 4.0 and continue to 4.4. Thus it is possible for an older 15H unit to have an apparently newer firmware revision (4.4, say) than the 15xH (at 3.9). The firmware is not interchangeable between units.

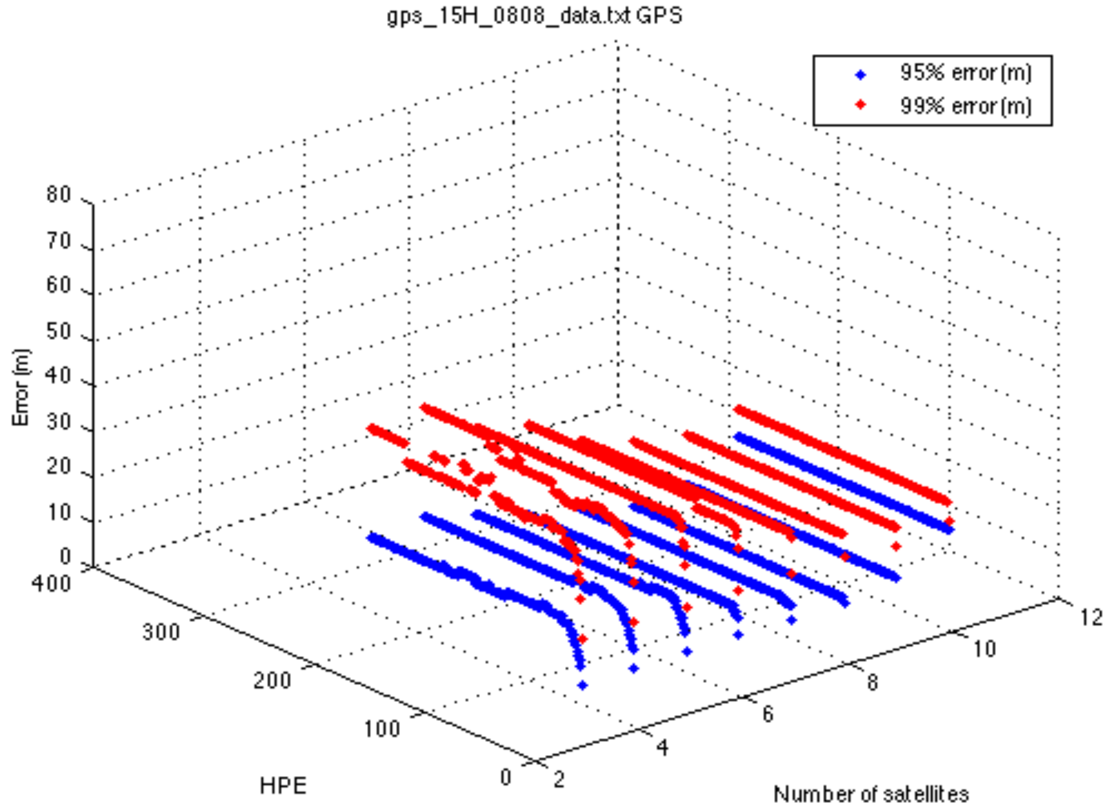
[6] We did not attempt to determine the GPS model used by the different deck dives. Indeed, some of the deck data may reflect the performance of older Garmin model 25 units.

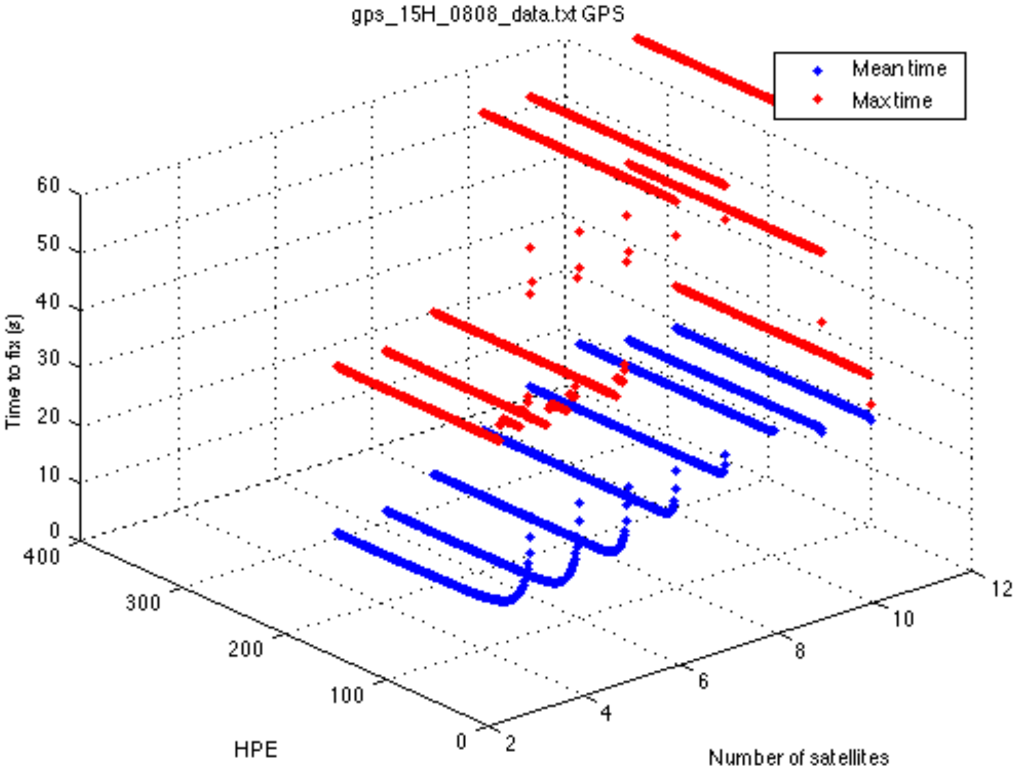


We analyzed how the actual error varied with different HPE and satellite acceptance criteria for the different units. Specifically for each cycle, we investigated the actual error and elapsed time to the first fix in a cycle, if any, with HPE less than a given error threshold employing a minimum number of satellites. Total acquisition time can exceed 60 seconds depending on the number of invalid 'V' sentences received before valid 'A' sentences.


gps_15xH_0806_data.txt GPS
95% error(m)
99% error(m)
Error (m)
0
10
20
30
40
50
60
70
80
HPE
400
300
200
100
0
Number of satellites
2
4
6
8
10
12

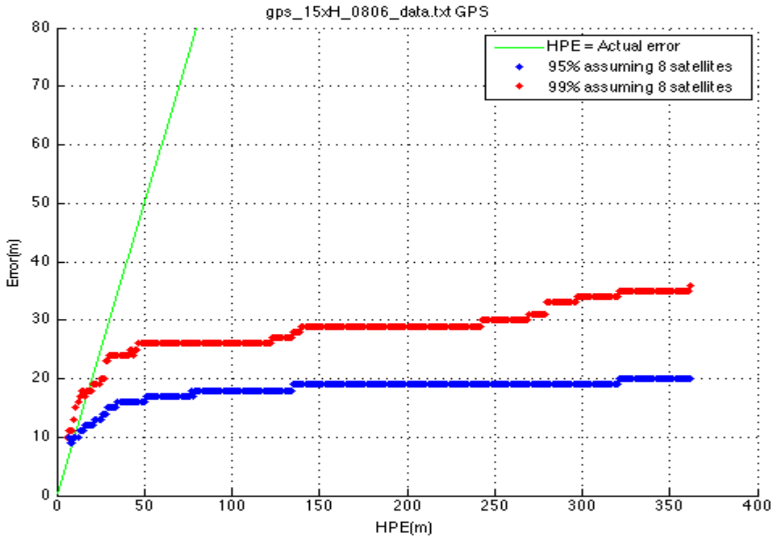

gps_15xH_0806_data.txt GPS
Mean time
Max time
Time to fix (s)
0
10
20
30
40
50
60
HPE
400
300
200
100
0
Number of satellites
2
4
6
8
10
12

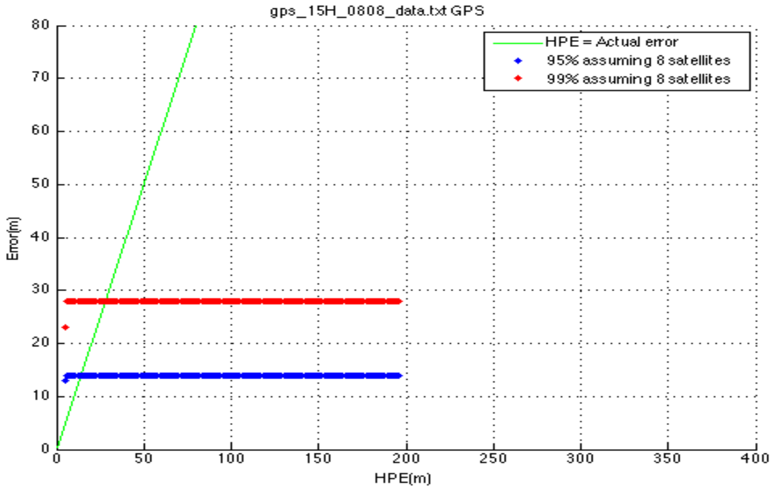

gps_15H_0808_data.txt GPS
95% error (m)
99% error (m)
Error (m)
80
70
60
50
40
30
20
10
0
400
300
200
100
0
HPE
2
4
6
8
10
12
Number of satellites


gps_15H_0808_data.txt GPS
Mean time
Max time
Time to fix (s)
60
50
40
30
20
10
0
400
300
200
100
0
HPE
2
4
6
8
10
12
Number of satellites

In all cases, the positional accuracy improves with increasing time (number of fixes per cycle), increasing satellites, and decreasing reported HPE. However, the units have very different error and timing profiles. The 15xH appears to have sacrificed accuracy for speed of initial acquisition compared to the 15H (explaining the poor HDOP-only performance noted above) but nevertheless requires less time overall to achieve a given positional error. In particular, for fixes involving a minimum of 8 satellites, the 15xH requires roughly 14 fixes to achieve a 97% chance of 10 meter accuracy compared to the 15H, which requires nearly 35 fixes to achieve a 90% chance of 10 meter accuracy.

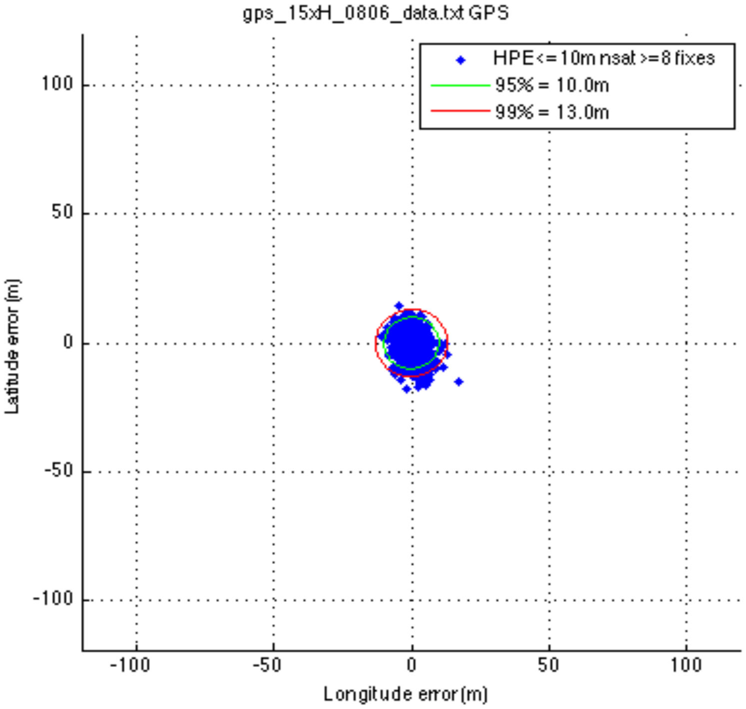
There appears to be no simple correlation between (combinations of) reported values from the unit and positional accuracy. For both units, the HPE value largely overstates the actual error, which is comforting, but both units begin to underestimate the actual error as reported HPE drops below 30 meters or so. Compare the green 1:1 line in the figures below: Points to the right of the line suggest the HPE predicted error is worse than the actual error while points to the left of the line indicate the actual error is worse than HPE suggests. Nevertheless, using HPE from the unit provides a reasonable estimate of actual accuracy below 30 meters and is certainly an improvement over using HDOP as an error predictor for GPS fixes.

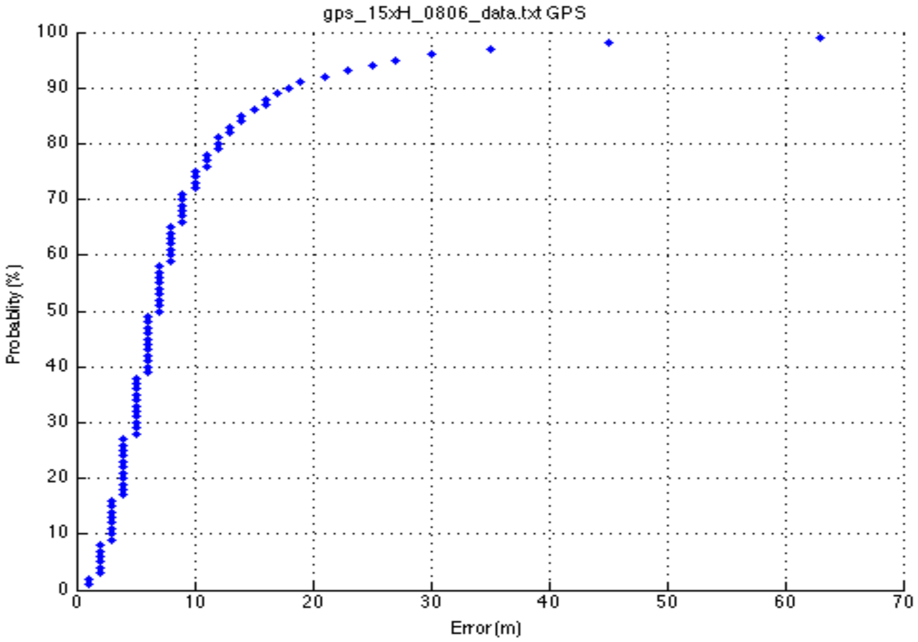


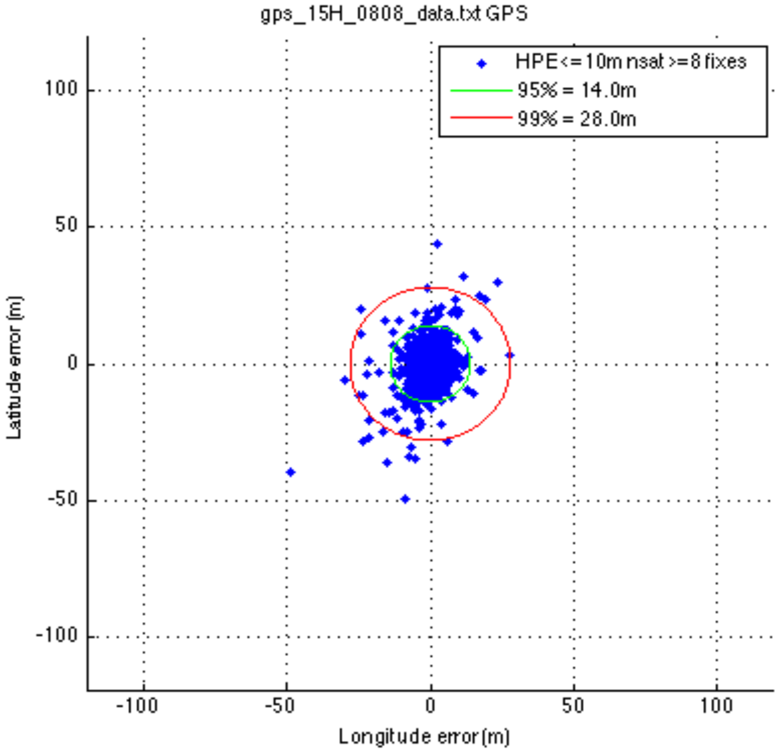


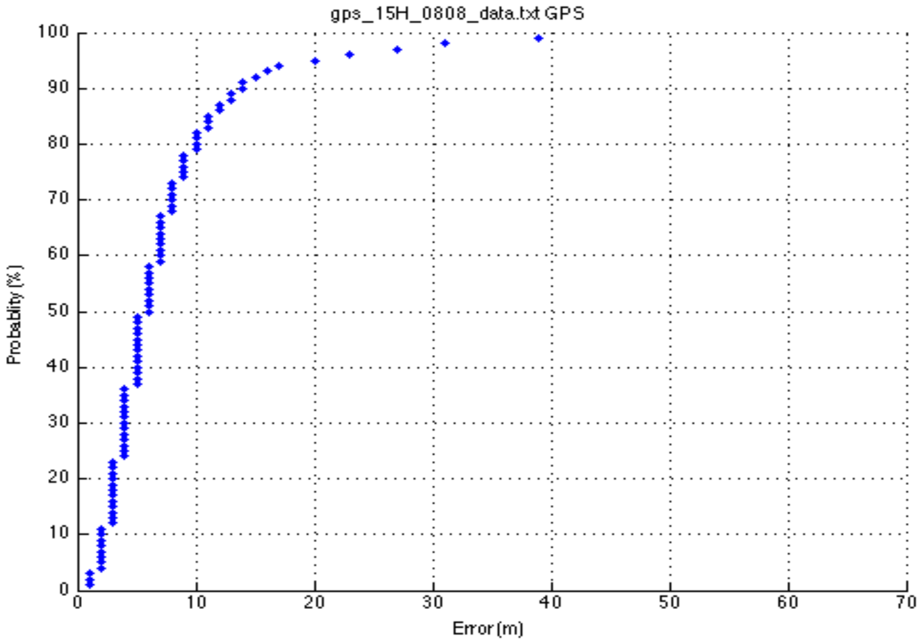
Additional tests comparing the use of HPE against known locations and simulating a drifting glider by walking a glider-in-a-bucket around campus show that the Garmin proprietary estimate remains quite good, even when moving.

To achieve 10 meter accuracy based on, say, at least 8 satellites requires around 15 valid fixes for the 15xH, 35 for the older 15H. The error probability curves for the units are significantly improved with this termination criteria (note x-axis scale changes from the second figure):






gps_15H_0808_data.txt GPS
HPE<=10m nsat>=8 fixes
95% = 14.0m
99% = 28.0m
Latitude error (m)
Longitude error (m)
100
50
0
-50
-100
-100
-50
0
50
100


gps_15H_0808_data.txt GPS
Probability (%)
Error (m)
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70

Based on these observations, the glider code will now employ a pilot-specified criteria for latching a fix in addition to HDOP <= 2, which is always used. The $N_GPS parameter value takes the form *eennff*, specifying the maximum HPE permitted (*ee,* in meters), the minimum number of satellites required (*nn*) , and the maximum number of valid fixes to acquire (*ff*). If either *ee* or *nn* are zero, the corresponding criteria is ignored; if both are zero the GPS code operates as before, terminating acquisition upon the first fix with HDOP <= 2.0 and after no more than *ff* fixes[7].

For the test results shown above, for example, $N_GPS would be 100815 for a 15xH unit or 100835 for a 15H, which requires a fix with HDOP <=2.0 and HPE <= 10 meters from at least 8 satellites, but waiting for no more than 15 or 35 valid fixes respectively. An $N_GPS of 505 would ignore HPE but require at least 5 satellites and wait for no more than 5 valid fixes. In the absence of information about which unit type is installed, the new $N_GPS default value will be 100840.

GPS fixes in the log file will now report HPE, estimated drift speed (m/s)[8], drift heading (degrees true), and the number of satellites used for the fix. If these values are not available they are reported as -1.

The basestation will estimate a depth-average current magnitude error based on the reported HPE of each fix and the time of the dive[9]. This estimated current error is a *lower bound* that assumes the flight model and the vehicle parameters are accurate; the actual current error could be worse if, for example, the vehicle parameters are not estimated correctly. If HPE is not available (i.e., historical deployments), the basestation will assume a conservative positional error of 100 meters; this value can be overridden in the `sg_calib_constants.m` file using the variable `GPS_position_error`.

---

[7] If the proprietary PGRME sentence is not available, *ee* ignored.

[8] The Garmin units apparently have a floor of 1 knot or 0.5 m/s.

[9] The basestation will also estimate a surface drift current magnitude error based on the GPS errors and the time between the $GPS1 and $GPS2 fixes.