Bias and Precision of Different Sampling Methods for GPS Positions

Michel Arnaud and Albert Flori

Abstract
Based on GPS readings taken four mornings running, this article uses statistical methods, taking account of observation dependence, to study bias and variance of errors. It is shown that with or without filtering based on criteria such as the number of satellites or PDOP, mean longitude and latitude values are unbiased. In the experiment carried out, filtering was used to improve the precision of longitude readings, but had a lesser effect on latitude. However, filtering can cause long delays before obtaining a reading. Replicating readings can increase their precision, provided certain operating conditions are respected.

Introduction
Researchers working in the agricultural and environmental fields are making increasing use of the well known Global Positioning System (GPS) technique (Brown, 1992; Hurn, 1989; Wells, 1986) to position themselves in spatial terms. Unfortunately, like any measuring device, the precision of the readings obtained by GPS varies.

Two techniques are generally used to reduce the errors of various kinds (Hurn, 1989; Puterski *et al.*, 1990; Wilkie, 1989) that falsify measurements when only one device is used (GPS in absolute mode): either readings are eliminated based on a criterion linked *a priori* to the extent of the errors, or measurements, are replicated. When two receivers are available, differential correction can also be carried out (August *et al.*, 1994), which substantially improves measurement quality.

This article uses statistical methods taking account of observation dependence to study the errors affecting GPS measurements. Bias and the variance of errors are studied and conclusions drawn with a view to improving precision both by filtering and by replicating measurements.

Methodology
The data supplied by a Garmin 75 GPS receiver with an aerial fixed to the perfectly obstacle-free roof of the Maison de la Télédétection (MTD) in Montpellier (France) were recorded on microcomputer via a Garmin connection cable. The coordinates of the aerial were known to the nearest metre: 43° 38.707' North, 3° 52.586' East (V. Freycon *et al.*, 1996). A simple program in Qbasic transferred the data received to a file. Continuous recording sessions were carried out on four different dates in February 1996, during the morning and lasting around four and a half hours. The data were recorded as soon as they were received by the microcomputer, i.e., roughly every two seconds.

The data supplied by the Garmin GPS receiver that were used were the geographical coordinates in the WGS 84 reference system (latitude and longitude in degrees, minutes, and thousandths of a minute), the number of satellites used (NSAT), and the geometric (3D) and horizontal (2D), dilutions of precision (PDOP and HDOP).

Results

Bias and Precision of the Non-Filtered Series
Generally, the position indicated by the GPS at a given time does not coincide with the exact position of the apparatus. In fact, the difference between the coordinate readings and the true values can be seen as the result of two errors:

- a systematic error due to the GPS system itself, which remains the same irrespective of the measurement date and the point measured; and
- a random error that differs with each measurement, due to atmospheric conditions, among other things. Moreover, as atmospheric conditions change slowly, the random errors corresponding to two measurements at close intervals generally have similar values: they are said to be correlated (Figure 1).

When studying the precision of a measurement method, two questions have to be answered:

- Is there a systematic error?
- What is the extent of the random errors?

More formally, supposing that each position read P_i is equal to the true known value v under the effect of a systematic error s and a random error e_i ($P_i = v + s + e_i$), the problem is, on the one hand, estimating systematic error s and testing to see whether it can be assumed to be nil and, on the other hand, in estimating the variance of the e_i.

However, the conventional formulae for calculating means and variances cannot be used in this case due to the self-correlation of the random errors. Our estimates were therefore made by modeling the errors by a first-order autoregressive process. This means that the correlation between two errors e_i and e_j concerning the observations on dates t_i and t_j solely depends on the time lapse $|t_i - t_j|$. It is equal to $\rho^{|t_i - t_j|}$. The MIXED procedure of the SAS software (SAS Institute Inc., 1996) was used for computerized calculation.

Moreover, estimating and testing a single bias over the observation period as a whole (four days) risked hiding possible temporary deviations. To increase the power of the test, we therefore chose to estimate the value of the systematic error hour by hour.

The following results were obtained: for the four mornings covered by the experiment, the estimated systematic errors for latitude and longitude were never over 1/100th of a minute. For each hourly time lapse, there was not a single estimate that differed significantly from 0 at the 5 percent

Photogrammetric Engineering & Remote Sensing,
Vol. 64, No. 6, June 1998, pp. 597–600.

0099-1112/98/6406–597$3.00/0

CIRAD, Av du Val de Montferrand, BP 5035, 34032 Montpellier, Cedex 1, France (arnaud@cirad.fr).

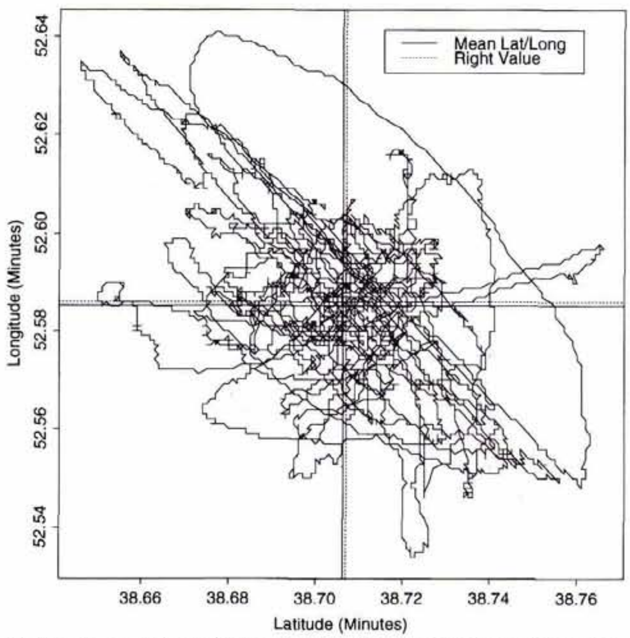

Figure 1. Positions indicated by the GPS during a morning.

threshold. We can therefore conclude that GPS is not biased and that the position indicated is correct on average.

The standard deviation of the random errors was estimated at 0.022 minutes for latitude and 0.019 minutes for longitude. Roughly speaking, this means that 95 percent of the position readings are within a 120-m radius of the true position.[1]

Effect of Filtering: Bias and Precision

Certain GPS manufacturers recommend keeping only those observations for which NSAT ≥ 4 and PDOP ≤ 2. In our experiment, we filtered our observations, keeping only those for which NSAT ≥ 4 and PDOP ≤ 2 or 2.2. Do these conditions lead to bias and how do they affect measurement precision?

First, filtering eliminates a large number of observations. For instance, in the case of filtering according to NSAT ≥ 4 and PDOP ≤ 2, around two thirds of the observations are eliminated, leaving an average of 2,000 out of 7,000 observations for each morning (Figure 2). There can be very long gaps before the conditions laid down are respected (around 50 minutes, Figure 2). Observers may have to wait some considerable time before being able to record the latitude and longitude values.

Irrespective of the filtering criteria, using the procedure described in the previous section, filtering does not introduce any bias in statistical terms.

As regards precision, filtering has a different effect on longitude and latitude: For longitude, the least stringent filter (NSAT ≥ 4 and PDOP ≤ 2.2) substantially reduces the estimated value of the standard deviation, from 0.019 minutes to 0.013. The most stringent filter (NSAT ≥ 4 and PDOP ≤ 2) improves precision even further: the estimated standard deviation is 0.009 minutes, roughly halving the standard deviation for the non-filtered series. For latitude, the gain in precision is less marked: from 0.022 minutes to 0.018 for the least stringent and 0.019 with the most stringent filter.

In short, the effect of filtering on precision is not clear-cut. In our experiment, it did not have the same effect on latitude and longitude. The two coordinates are no longer determined with the same degree of precision. Moreover, filtering eliminates a large number of observations and the observer therefore risks having to wait for some time.

Replication: OK, But Not Just Any Old Way

Although the observations were correlated, we were able to obtain an estimate of the extent of the random errors affecting each of the two coordinates. Unfortunately, the two coordinates themselves (latitude and longitude) do not vary independently (Figure 1). As a result, the two standard deviations are not sufficient to calculate the mean distance between the positions indicated by the GPS and the true position of a given point. The problem is even more complicated when considering the mean of several GPS readings rather than a single reading, which is why we used simulation techniques to determine the precision of the values resulting from calculating the mean of several readings taken at regular intervals.

More precisely, what needs to be done is to evaluate the distance separating the position obtained by calculating the mean of a certain number of GPS readings and the actual position of the point where the apparatus is installed.

To do this, we proceeded as follows:

- (i) For the four days as a whole, we eliminated all the invalid observations (where the number of satellites NSAT was less than 4) and the observations that were unlikely to be precise (where the HDOP was over 2[2]);
- (ii) We set a number of replicates N to be varied from 2 to 6

[1]At a latitude of 43°, one minute is equivalent to 1,852 m and one minute of longitude to 1,852*cos(43°)=1,315 m.

[2]Filtering was carried out by the HDOP procedure, as PDOP-based selection would have eliminated too many data.

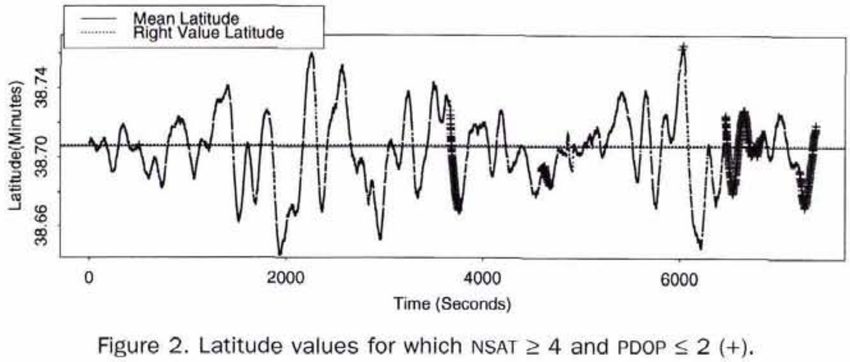

Figure 2. Latitude values for which NSAT ≥ 4 and PDOP ≤ 2 (+).

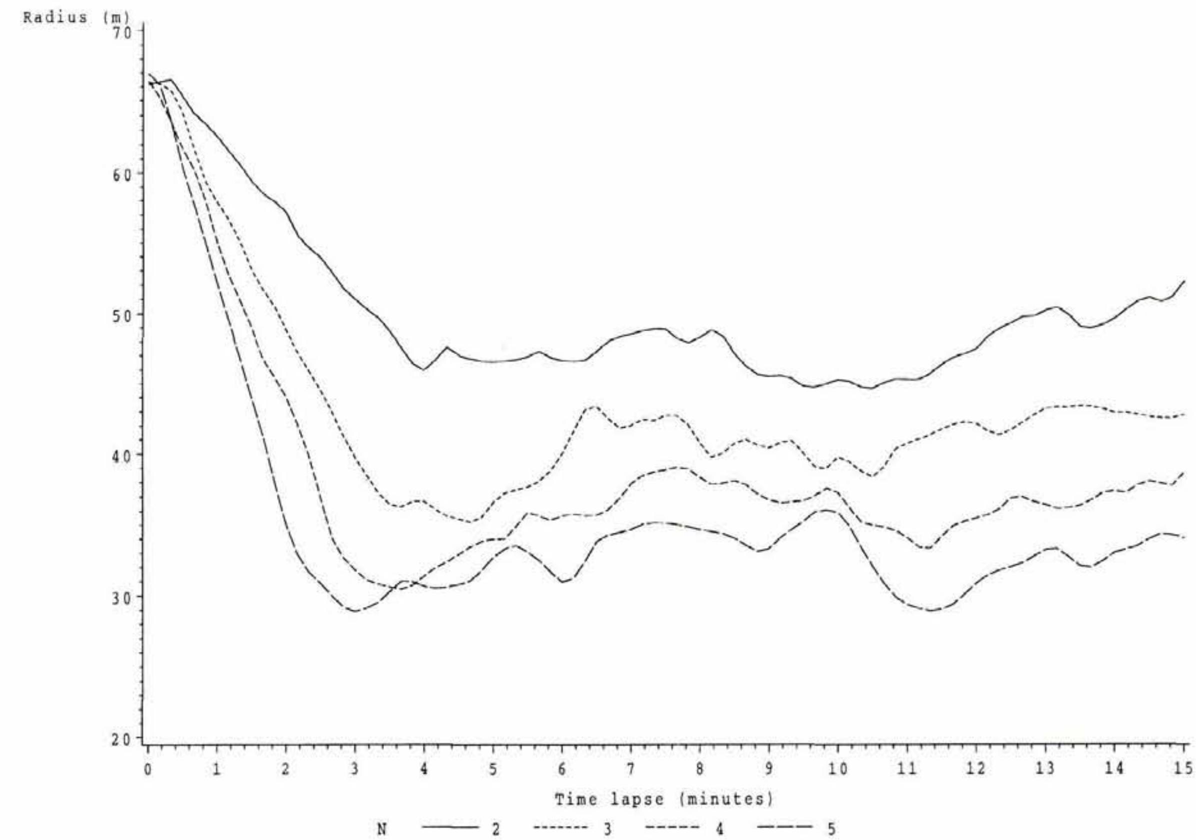

Figure 3. Confidence Radius (R) of the mean point ($R = f$(time lapse)).

and a time lapsed between two replicates of 20 to 600 seconds;

(iii) From the observations not eliminated under stage (i), we chose a random series of N observations made d seconds apart. We then plotted the mean point for this sample on a Latitude by Longitude graph; and

(iv) We repeated stage (iii) 1,000 times to determine the radius R of the circle containing 95 percent of the mean points represented (i.e., 950 points).

The radii R obtained for the different N and d values are shown in Figure 3. For a fixed N value, when d increases, the radius R decreases until it reaches a plateau for time lapses of more than 3 to 4 minutes. The level of the plateau obviously depends on the number of observations N, and is around $65/\sqrt{N}$.

This means that, to benefit fully from the improved precision obtained by increasing the number of replicates, a certain time lapse has to be allowed between two successive readings to ensure that they are independent. In our experiment, at a certain level of SA and with a certain type of equipment, the time lapse was found to be around 4 minutes.

Conclusion

Filtering and replicating observations ensure substantial increases in precision, but also mean spending more time taking measurements.

For filtering, restricting the observations used to periods when NSAT is over 4 and PDOP under 2 reduces the standard deviation of longitude values from 25 to 12 m in our experiment. However, this means sacrificing four-fifths of the observations and running the risk of waiting over an hour before being able to take a measurement.

For replicates, the time-precision link is even more direct: if it is to provide information, each additional observation has to be made after a certain time lapse. We obtained 3 to 4 minutes. Conley (1992) and Kremer (1990) found similar times. It would be worth determining whether the results are the same at other times and with other types of receiver.

However, users will have to strike a balance between the time taken to measure each point and the degree of uncertainty he deems acceptable.

For applications where such a compromise would be unsatisfactory, differential GPS, which calls for a much greater investment, would be a highly effective technical solution for increasing the quality of position measurements.

Acknowledgments

This work would not have been possible without G. Dewispeleare and G. Forgiarini (CIRAD), who lent us the GPS; the Krystal Marine company, who enabled us to produce the Qbasic program; M. Harmel from the Institut Géographique National, who gave us the coordinates of points near Montpellier belonging to the French geodetic network; and the CIRAD GIS think tank, who gave us advice and helped to improve this article.

References

August, P., J. Michaud, C. Labash, and C. Smith, 1994. GPS for environmental applications: Accuracy and precision of locational data, *Photogrammetric Engineering & Remote Sensing*, 60(1):41–45.

Brown, A., 1992. The GPS coordinate system explained, *GIS World*, 5(2):70–71.

Conley, R., 1992. GPS performance: What is normal?, *Proceedings, ION GPS-92*, pp. 885–893.

Conley, R., M. Fryt, and S. Scott, 1995. GPS performance characteristics and trends, *Proceedings, ION GPS-95*, pp. 1363–1371.

Freycon, V., *et al.*, 1996. Les GPS, principes de fonctionnement et conseils d'utilisation, Rapport interne Groupe SIG CIRAD, pp. 1–25.

Henriksen, J., G. Lachapelle, J. Raquet, and J. Stephen, 1996. Analysis of stand-alone GPS positioning using post-mission, *Proceedings, ION GPS-96*, pp. 251–259.

Hurn, J., 1989. *GPS: A Guide to the Next Utility*, Trimble Navigation Ltd., Sunnyvale, California, 76 p.

Kremer, J.T., R.M. Kalafus, P.V.W. Loomis, and J.C. Reynolds, 1990. The effect of selective availability on differential GPS corrections, *Navigation*, 37(1)39–52.

Puterski, R., J.A. Carter, M.J. Hewitt III, H.F. Stone, L.T. Fisher, and E.T. Slonecker, 1990. *Global Positioning Systems Technology and Its Application in Environmental Programs*, GIS Technical Memorandum 3, Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Las Vegas, Nevada.

SAS Institute Inc., 1996. *SAS/STAT Software, Changes and Enhancements through Release 6.11*, SAS Institute Inc., Cary, North Carolina, 1104 p.

Wells, D., N. Beck, D. Delikaraoglou, A. Kleusberg, E. Krakiwsky, G. Lachapalle, R. Langley, M. Nakiboglu, K. Schwarz, J. Tranquilla, and P. Vanicek, 1986. *Guide to GPS Positioning*, Canadian GPS Associates, Fredericton, New Brunswick, Canada.

Wilkie, D.S., 1989. Performance of backpack GPS in a tropical rain forest, *Photogrammetric Engineering & Remote Sensing*, 55(12): 1747–1749.

(Received 12 March 1997; accepted 23 October 1997; revised 25 November 1997)