Detection of Outliers in GPS Measurements by Using Functional-Data Analysis

C. Ordoñez¹; J. Martínez²; J. R. Rodríguez-Pérez³; and A. Reyes⁴

Abstract: The identification of outliers in global positioning system (GPS) observations—to compare equipment, positioning methods, or working conditions—has traditionally been performed using univariate or multivariate statistics. However, these methods have certain drawbacks when processing data collected by GPS receivers. Such data can be more suitably handled as observations at discrete points of a smooth stochastic process and, consequently, other statistical approaches to the analysis of functional data may prove more suitable. We analyzed the applicability of the concept of functional depth to the identification of outliers in GPS observations. The proposed method was applied to 12 series of GPS receiver data collected in an open space and in similar signal reception conditions. The results obtained adapted better to the expected results, given the signal-reception conditions, than those obtained by the classical statistical approaches used by other writers to compare GPS observations. DOI: 10.1061/(ASCE)SU.1943-5428.0000056. © 2011 American Society of Civil Engineers.

CE Database subject headings: Global positioning systems; Measurement; Data analysis; Surveys.

Author keywords: GPS measurements; GPS accuracy; Functional data analysis; Functional depth; Outliers.

Introduction

Global positioning system (GPS) devices are used in many applications that require accurate point measurements, such as geodesy, surveying, mapping, transportation, navigation, and the geosciences. GPS positional accuracy depends on many factors (Van Sickel 1996; Guochang 2003). Errors in the satellite clock, satellite ephemerides, receiver clock, and atmospheric delays degrade accuracy, as does satellite geometry.

The level of accuracy required in GPS measurements varies greatly from application to application. A number of studies have been carried out to assess differences in accuracy between receivers (Serr et al. 2006), positioning methods (Naesset and Jonmeister 2002), and working environments (Hasegawa and Yoshimura 2003; Tiberius and Kensellar 2003; Rodríguez-Pérez et al. 2007). The process generally consists of calculating parameters characterizing measurement precision, which are then compared using statistical tests that determine whether or not the observations come from the same distribution (Rodríguez-Pérez et al. 2007). From the point of view of statistics, an outlier is defined as an observation that lies outside the overall pattern of a distribution. One of the typical criteria for identifying observations considered to be outliers is the 3$\sigma$ rejection criterion, i.e., outliers are observations that deviate from the mean in an amount equal to or greater than three times the standard deviation. Nickitopoulou et al. (2006) conducted a large number of experiments to simulate harmonic movements using a rotating GPS-receiver antenna; the accuracy of the observations was determined by comparing the recorded coordinates with the real ones and two types of deviations from the theoretical curves were identified according to their scale. Finally, these writers detected outliers from these deviations using the 3$\sigma$ criterion, finding 1.5% of outliers in midlatitudes when at least six satellites were visible.

Pismoulis and Stiros (2008) used equipment composed of a GPS receptor and a robotic total station to determine the oscillation parameters for major structures. They also used the 3$\sigma$ rejection criterion to discard gross errors. Some of these errors were related to data gaps, but others were according to their interpretation.

Wdowinski et al. (1997) considered two types of outliers, each due to a specific cause: (1) observations with error bars larger than three times the root-mean square (RMS) and (2) observations that deviate from the mean of the series more than the number of times the RMS scatters.

Nonetheless, this kind of analysis is not the most appropriate for detecting outliers when only a dense set of data collected over time is available, such as the data collected by GPS receivers, which can be considered as observations at discrete points of a curve. The main reasons univariate or multivariate statistical methods are not suitable for working with functional data are that (1) the time correlation structure is ignored by these methods, (2) the infinite dimensionality of the functions means that methods for multivariate samples are greatly affected by the curse of dimensionality, (3) many of these methods are implicitly restricted to Gaussian or elliptical populations, and (4) outliers may not be detected for each time instant in a curve, yet the curve itself may be an outlier (Febreo et al. 2008).

To overcome these drawbacks, a number of writers have proposed methods for detecting outliers in sets of functional data (Fraiman and Muniz 2001; López-Pintado and Romo 2007; Cuevas et al. 2006; Febreo et al. 2008). We analyze, as an alternative to traditional methods, the usefulness of one of these functional-depth methods in order to identify outliers in GPS observations.
Functional depth aims to measure the centrality of a given curve within a group of curves.

The article is organized as follows: the section on materials and methods describes how to build functions from a set of discrete points using a smoothing process, the next section reviews the concept of functional depth and its usefulness for identifying functional outliers, the section on the case study describes the experimental data used in the study, and the results section describes the results obtained for the functional analysis and compares these results with those obtained from applying statistical tests for comparing non-Gaussian distributions, namely, the Kruskal-Wallis and Mann-Whitney tests (Hollander and Wolf 1999). The article concludes with the conclusions drawn from an analysis of the results.

Materials and Methods

Smoothing

Functional data are observations of a random continuous process observed at discrete points (Ramsay and Silverman 2005). Given a set of observations \( x(t_j) \) in a set of \( n_p \) points, \( t_j \in \mathbb{R} \), in which \( t_j \) is each instant of time, all the observations can be considered as discrete observations of the function \( x(t) \in \mathbb{C}F \), in which \( F \) is functional space. To estimate the function \( x(t) \), it is considered that \( F = \text{span}(\phi_1, \ldots, \phi_{n_p}) \), in which \( (\phi_k)_{k=1, \ldots, n_p} \) is a set of basis functions. In view of this expansion, the solution to this problem is given by (Ramsay and Silverman 2005)

\[
C = (\Phi'\Phi + \lambda R)^{-1}\Phi'z
\]

By using a Fourier development as the basis functions, the matrix \( R \) in Eq. (4) is the identity matrix.

Smoothing then basically consists of fitting a curve to the point observations. The adjustment is on the basis of minimizing the error estimate obtained by calculating the difference between the observed and adjusted errors. Smoothing also avoids excessive curvature in the fitted curves by penalizing second derivatives.

Functional Data Concept

Depth measurement was introduced originally in multivariate analysis to measure the centrality of a point with respect to a cloud of points. Depth provides a way of ordering points in a Euclidean space from the center to the periphery in such a way that the points closest to the center will have greater depth. The notion of depth has recently been extended to functional data (Fraiman and Muniz 2001; Cuevas et al. 2006). Functional depth measures the centrality of a curve \( x_i \) within a set of curves \( x_1, \ldots, x_n \).

The most popular depth measurements are

- Fraiman-Muniz depth (FMD): Let \( F_{nm}[x(t)] \) be the cumulative empirical-distribution function (Fraiman and Muniz 2001) for the values of the curves \( [x_i(t)]_{i=1}^n \) in an instant of time \( t \in (a, b) \) given by

\[
F_{nm}[x(t)] = \frac{1}{n} \sum_{i=1}^n I[x_i(t) \leq x(t)]
\]

in which \( I(\cdot) \) is the indicator function.

The Fraiman-Muniz depth for a curve \( x_i \) with respect to the set \( x_1, \ldots, x_n \) is given by

\[
\text{FMD}_{n}[x_i(t)] = \int_{a}^{b} D_{n}[x_i(t)] dt
\]

in which \( D_{n}[x_i(t)] \) is the point depth of \( x_i(t) \) for \( \forall t \in (a, b) \), given by

\[
D_{n}[x_i(t)] = 1 - \left( 1 - F_{nm}[x_i(t)] \right)
\]

- H-modal depth (HMD): The functional mode (which is based on the mode concept) is defined as the curve most densely surrounded by the other curves in the sample. H-modal depth (Cuevas et al. 2006) is expressed as

\[
\text{HMD}_{n}(x_i, h) = \frac{1}{n} \sum_{i=1}^{n} K(\|x_i - x_k\|/h)
\]

in which \( K: \mathbb{R}^+ \rightarrow \mathbb{R}^+ \) is kernel function; \( \| \cdot \| \) is norm in functional space; and \( h \) is bandwidth parameter (see Cuevas and Fraiman 1997 for more information on the choice of the bandwidth). One of the most widely used norms for a functional space is \( L^2 \), given as

\[
\|x_i(t) - x_j(t)\|_2 = \left( \int_a^b (x_i(t) - x_j(t))^2 dt \right)^{1/2}
\]

The infinite norm \( L^\infty \) is sometimes used.

\[
\|x_i(t) - x_j(t)\|_{L^\infty} = \sup_{t \in (a,b)} |x_i(t) - x_j(t)|
\]
Different kernel functions $K(\cdot)$ can also be defined, among them the truncated Gaussian kernel (Cuevas et al. 2006)

$$K(t) = \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \quad t > 0$$

**Functional Outliers**

A functional-sample set may have elements that, although they do not incorporate error in themselves, may feature patterns different from the others. The depth measurement described previously is used to identify outliers in functional samples. This enables sets of observations over time fitted to curves to be compared, rather than just the mean values in the measurement time interval.

Depth and outlier are inverse concepts; hence, an outlier for a functional sample will have considerably less depth. The curves with the least depth are sought in order to identify functional outliers.

H-modal depth and a truncated Gaussian kernel generate the outlier selection criterion, and the bandwidth parameter chosen was the 15th percentile of the empirical distribution of $\frac{1}{\sigma_j}(x_j - \mu_j)^2$, $j = 1, ..., n$ (Febrero et al. 2008). A cutoff $C$ is selected in such a way that the percentage of correct observations poorly identified as outliers (type I error) is $\alpha$ (usually 5% or 1%), given by

$$\Pr\{\text{HMD}_n[x_i(t)] \leq C\} = \alpha, \quad i = 1, ..., n$$

Unfortunately, the distribution of the chosen functional depth is not known and so the value for $C$ had to be estimated. Of the different approaches to estimating this value (Febrero et al. 2008), we chose for the purpose of this research a method based on bootstrapping (Peng and Qi 2008) the curves of the original set with a probability proportional to depth.

The bootstrapping approach can be summarized as follows:

1. A new sample is extracted from the original sample by means of sampling with replacement; in other words, each extracted element is replaced after extraction and so may be selected again;
2. Based on this new sample, the populational parameter of interest is estimated on the basis of the construction of a statistic;
3. The two steps above are repeated until a large number of estimates are obtained; and
4. Finally, the empirical distribution of the statistic is determined.

Fig. 1 shows a flowchart depicting the process used to identify outliers. First, point observations are transformed into curves by a smoothing process. Then functional depth for each curve is calculated. Functional depth of each curve measures the distance to the central curve with the highest depth. In the next step, the probability distribution of the functional depth is estimated through a process

![Flowchart showing the methodology used to identify outliers](image-url)
of sampling (bootstrap). Finally, from that probability distribution, the value of the quantile \( C \) corresponding to a specific significance level \((\alpha)\) is calculated. Curves with a depth less than or equal to \( C \) are classified as outliers.

### Case Study

#### Data Collection

The sample used for this research consisted of error measurements for a GPS receiver in a fixed location with no obstacles to the reception of GPS signals.

The GPS experimental data were collected for 4 days over periods of 5-6 h between August 20 and 23, 2007, using one double-frequency GPS receiver (HiperPlus, Topcon Positioning Systems, Inc., Livermore, CA) observing the GPS pseudorange and carrier phase. The antenna heights ranged from 1.45 and 1.60 m, and the logging rate was 1 s. Collection of observations lasted for at least 1.5 h and the process was repeated three times a day. GPS data were revised to ensure continuity and were cut to obtain 12 data sets of 1 h (three data sets per day).

Then the sample \( [(x_{i1}, x_{i2}, \ldots, x_{i600})]^{T}_{i} \) consists of 12 observations, in which \( x_{ij} \) represents measurement in instant \( i \) (in seconds) over the period of an hour at point \( j \).

The observation point was located at a latitude of 42°41’08.79872” N, a longitude of 6°38’03.210587” W (WGS84) and at an ellipsoidal height of 933.829 m. These coordinates were obtained by static positioning at an ellipsoidal height of 933.829 m. These coordinates were obtained by static positioning and postprocessing correction were carried out using the base station called PONF, which was the nearest reference station in the Regional GNSS Network (Regional GNSS Network). The geographic coordinates of the observation point were projected and this position was set up as the true position for calculating horizontal and vertical accuracies. The Universal Transverse Mercator (UTM) coordinates were \( X_{true} = 693,814.623 \) and \( Y_{true} = 4,728,635.531 \) (Datum ETRS89; zone 29N). The \( Z_{true} \) is the ellipsoidal height.

#### Data Characterization

Several GPS parameters were selected to try to justify possible outliers. Because accuracy and precision of GPS position depend on the number of satellites, this parameter was considered as the mean number of satellites \((N)\) in which the signal was received. Mean elevation angle \((EA)\) and position dilution of precision \((PDOP)\) are two useful indicators of satellite geometry for three-dimensional (3D) positioning. The GPS L1 and L2 carrier frequencies are susceptible to interference from other transmitted signals or to disturbances caused by obstacles. The channel signal-to-noise ratio (ratio of received signal power to the noise power accompanying the GPS signal) relating to the delay-lock loops for coarse acquisition (CA) code (SNCA), PL1 code (SNL1), and PL2 code (SNL2) were considered for this research. Table 1 shows the basic statistics for the studied parameters. The mean number of satellites was not converted to an integer so the observations corresponding to the minimum and maximum values could be better distinguished.

#### Error Calculation

The error in each time instant was estimated by comparing the coordinates observed in each second with the coordinates considered as true:

\[
\begin{align*}
\sigma_{XY,acc} &= \sqrt{(X_i - X_{true})^2 + (Y_i - Y_{true})^2} \\
\sigma_{Z,acc} &= |Z_i - Z_{true}|
\end{align*}
\]  

in which \( \sigma_{XY,acc} \) and \( \sigma_{Z,acc} \) = horizontal (planimetric) and vertical (alitmetric) accuracy, respectively; \( X_i, Y_i, \) and \( Z_i \) = measured positions in the instant \( i \); and \( X_{true}, Y_{true}, \) and \( Z_{true} \) = true positions along the easting, northing, and ellipsoidal height directions, respectively.

#### Results

The first step in identifying outliers was to fit the error values for each point to a curve using a smoothing process. An important aspect of this process was the selection of the number of basis functions to include in the expansion; this was done using the coefficient of determination (\( RSQ \)) value obtained in the process. The \( RSQ \) is given by

\[
RSQ = \frac{\sigma_{xy,acc}^2}{\sigma_{xy,est}^2 + \sigma_{xy,real}^2}
\]  

in which \( \sigma_{xy,est}^2 \) and \( \sigma_{xy,real}^2 \) = square covariance of estimate and real variables (horizontal and vertical errors in our case); and \( \sigma_{xy,est}^2 \) and \( \sigma_{xy,real}^2 \) = their variances, respectively.

![Fig. 2. RSQ in the smoothing obtained for the different sizes of basis-function sets considered](image-url)
Of interest was the equilibrium between the number of basis functions and the error in the fit, resulting in small errors for a number of basis functions that did not represent an excessive computational cost. Fig. 2, shows that in the least favorable case (the observation set requiring most basis functions to achieve a satisfactory fit), the \( R^2 \) asymptotically approached 1 from 3,000 basis functions. Consequently, this was the number of basis functions used for the smoothing.

A sample \( \{x_i\}_{i=1}^{12} \) was obtained, in which \( x_i \) represents a function. Fig. 3 shows the functional sample obtained after smoothing for both horizontal and vertical errors. In each of the graphs are 12 curves, one for each data set, showing changes over time of the error. The functional complexity of the sample is evident, with a large number of local minimums and maximums and very pronounced slopes. At first sight, there are no noticeable differences between the error curves, and this could lead to the conclusion that none of the curves differ considerably from the rest.

For this set of curves, the H-modal depth measurement was taken by using the norm \( L^2 \). Outliers were then identified as described previously. The picture to the right in Fig. 4 shows the only outlier identified for the horizontal error, corresponding to the curve number 3 for \( \alpha = 0.001 \). No outlier was identified for \( Z \) error.

This result differs from that obtained by using the Kruskal-Wallis test (the fact that the errors do not follow a normal distribution was first checked, for which reason a nonparametric test was used) for all the positions except number 3. This test rejected the null hypothesis at a 99% significance level that the 11 error observations (all except number 3) came from the same population. Recall that the Kruskal-Wallis test compares the medians of the groups, so it is possible to have a tiny \( p \) value—clear evidence that the population medians are different—even if the distributions overlap considerably. Moreover, when the values for each pair of curves are compared using a Mann-Whitney test, the null hypothesis is rejected. That means that according to this test, none of the 12 curves comes from the same population. In Fig. 5, two of the horizontal error curves (8 and 9) represented in Fig. 3 are shown. It can be seen that they are quite similar. Also, the parameters in Table 1 for these two curves are very similar.

We also implemented the classic outlier detection study from a vectorial perspective, i.e., assessing the Z-score value to see whether the difference for each observation and the overall mean surpassed three standard deviations:

\[
Z_i = \frac{V_i - V_r}{\sigma}
\]

in which \( V_i \) = mean value of the observations by receptor \( i \); \( V_r \) = reference value (the global mean); and \( \sigma \) = standard deviation of the sample.

According to this criterion, a set of data is regarded as an outlier for \( \alpha = 0.01 \) if the corresponding Z-score is greater than 3. Thus, according to this approach, there are no outliers in the set of 12 curves, including curve number 3, identified as outlier by the functional method proposed.

So, by using functional depth to compare groups of sample data, we obtain totally different results from those for classical statistical hypotheses. This is because the classical test compares...
central values summarizing the whole statistical distribution, whereas functional depth allows curves to be compared.

The explanation for the outlier identified via functional analysis was given by an analysis of the mean values for the GPS parameters recorded for each observation that had a bearing on the precision of the measurements given in Table 1. As can be observed, position 3 recorded the smallest mean values for the variables $N$, $EA$, $SNL1$, and $SNL2$, one of the lowest for the variable $SNCA$, and the highest for $PDOP$. Position 3 overall, in fact, had the variable values that most negatively affected the precision of the measurements. This may have given rise to larger errors in this set of observations than in the other sets (it had the greatest mean-error value) and, consequently, explains why it was identified as an outlier when compared with the other observations by the depth measurement.

Conclusions

We analyzed the application potential of functional-data analysis to the identification of outliers in errors in a set of observations made with a GPS receiver. The functional focus considers measurement errors as observations for a smooth random process observed at discrete time points. Outliers are then identified by comparing curves rather than central values as in traditional statistical tests for comparing distributions. This represents a considerable advantage because all the information in the measured interval of time is used rather than a summary value. Furthermore, because the method is nonparametric, the functional analysis does not assume a normal distribution for the errors.

Application of the method to a set of 12 GPS observations made at the same location but on different days enabled us to determine that the observation set for one of the days could be considered an outlier with respect to the other 11 observation sets.

Analysis of the set of variables that had a bearing on error, furthermore, enabled an explanation of why the errors recorded on that particular day could be considered as outliers with respect to the errors obtained on the other days.

References


Technical Papers

113 Deformation Assessment Considering an A Priori Functional Model in a Bayesian Framework
Barbara Betti, Noemi Emanuela Cazzaniga, and Vincenza Tornatore

120 An Improved Weighted Total Least Squares Method with Applications in Linear Fitting and Coordinate Transformation
Xiaohua Tong, Yanmin Jin, and Lingyun Li

129 MINQUE of Variance-Covariance Components in Linear Gauss-Markov Models
Peng Junhuan, Shi Yun, Li Shuhui, and Yang Honglei

140 Machine Learning Techniques Applied to the Assessment of GPS Accuracy under the Forest Canopy
Celestino Ordóñez, José R. Rodríguez-Pérez, Juan J. Moreira, J. M. Matias, and Enoc Sanz-Ablanedo

150 Detection of Outliers in GPS Measurements by Using Functional-Data Analysis
C. Ordóñez, J. Martínez, J. R. Rodríguez-Pérez, and A. Reyes

Case Studies

156 Evaluation of GPS/Galileo RTK Network Configuration: Case Study in Greece
Maria Tsakiri

Technical Notes

167 Second Order Design of Geodetic Networks by the Simulated Annealing Method
Sergio Baselga

Book Reviews

John V. Hamilton

Reviewers

175 Reviewers