Mapping GPS Positional Errors with Spatial Linear Mixed Models

A. F. Militino, M. D. Ugarte, J. Iribas, and E. Lizarraga-Garcia

Received: date / Accepted: date

Abstract Nowadays, GPS receivers are very reliable because of their good accuracy and precision; however, uncertainty is also inherent in geospatial data. Quality of GPS measurements can be influenced by atmospheric disturbances, multipathing, synchronization of clocks, satellite geometry, geographical features of the observed region, low broadcasting coverage, inadequate transmitting formats, or human or instrumental unknown errors. Assuming that the scenario and technical conditions that can influence the quality of GPS measurements are optimal, that functional and stochastic models that process the signals to a geodetic measurement are correct, and that all the GPS observables are taken in the same conditions, it is still possible to estimate the positional errors as the difference between the real coordinates and those measured by the GPS. In this paper, three spatial linear mixed models, one for each axis, are used for modelling real-time kinematic (RTK) GPS accuracy and precision, of a multiple-reference-station network in dual-frequency with carrier phase measurements. We interpret “accuracy” as unbiased and “precision” as a variance, but not for a particular signal or a set of signals when they are processed to become an estimated coordinate. Along the paper, the proposed models provide an estimate of the “accuracy” in terms of bias defined as the difference between real coordinates and measured coordinates after being processed. Moreover, our proposed models also provide an estimate of the “precision” through the standard errors of the estimated differences. This is done using ten different transmitting formats. Mapping and quantifying these differences can be interesting for users and GPS professionals. The performance of these models is illustrated by mapping positional error estimates within the whole region of Navarre, Spain. Sampled data have been taken in 54 out of the 211 geodetic vertex points of this region. Maps show interesting error patterns depending on transmitting formats, the different axes, and the geographical characteristics of the region. Higher differences are found in regions with bad broadcasting coverage, due to the presence of mountains and high degree of humidity.

Keywords Statistical estimation · Kriging · Geophysical measurements · Statistics · Spatial correlation

1 Introduction

The Global Positioning System (GPS) determines positions by measuring distances to Earth-orbiting satellites. Hofmann-Wellenhoff and Lichtenegger (2001) provide a comprehensive study on how GPS works. However, the measurement of a particular location is not exact. For example, Teunissen (1998) shows closed-form expressions for the minimal detectable biases of single and dual-frequency pseudo-range and carrier-phase data with three different single-baseline models. There are still some relevant factors that can reduce the quality of the measurements, mainly due to: a) Atmospheric...
disturbances, b) Timing of clocks, c) Multipathing, d) Satellite ephemerides and e) Satellite geometry. a) Atmospheric disturbances consist mainly of the slowdown that the signal experiences while passing through the ionosphere, which delays the signal’s arrival at the receiver, thereby affecting the distance calculation. Lejeune et al. (2012) develop software to compute the positioning error due to the ionosphere for all baselines of the Belgian GPS network, the so-called Active Geodetic Network (AGN). Hernández-Pajares et al. (2011) provide a summary of the ionosphere as related to space geodetic techniques. The amount of water vapor in the atmosphere can also affect the time delay. Lenderink and van Meijgaard (2001) study structures in predicted Integrated Water Vapor model for three cyclones (Kerstin, Liane, Monica) in 1995. Two versions of the model with differences in the physics package were compared: standard ECHAM4, and ECHAM4 with revised cloud and turbulence physics. The delay due to water vapor, the wet delay, is difficult to model using ground surface data. Thus, it is often estimated from the GPS data. In order to obtain the most accurate results from the GPS processing, a model of the horizontal distribution of the wet delay may be necessary. Emardson and Jarlemark (1999) evaluate these models through simulations. Banville and Langley (2013) provide a geometry-based approach with rigorous handling of the ionosphere. Bock et al. (2010) show that the use of modified GPS antennas can improve the propagation delays in the neutral atmosphere (troposphere). b) An additional source of uncertainty is timing because the clocks in the artificial satellites are very near perfect (Han et al. 2001), but the clocks in the receivers are not as good (Huang and Zhang 2012, Kenneth et al. 2008). At this regard, Liu et al. (2004) model the between observation-type delays for the purpose of precise positioning, under practical circumstances. c) Multipathing is the reflection of the satellite signal off a building or some other large reflective object before reaching the receiver, which again delays its arrival as indicated by Mekik and Can (2010). Under the assumption that the surrounding environment remains unchanged, Phan et al. (2012) show that multipath contamination of GPS measurements can be formulated as a function of the sidereal repeatable geometry of the satellite with respect to the fixed receiver. If the direct signal is also acquired, the software of most modern receivers can reject the indirect signals. However, multipathing can confuse the calculations for position. Iwase et al. (2013) detect the multipath errors in the satellite signals and exclude these signals to improve the positioning accuracy of GNSS, in particular in an urban canyon environment. d) Satellite ephemerides can also reduce the quality of measurements because the gravitational pull of the Sun and Moon on a satellite can distort its orbit. Desmars et al. (2009) introduce a new method to estimate the accuracy of predicted positions at any time, in particular outside the observation period. e) The relative position of the satellites, known as the satellite geometry, can also affect the accuracy. Cai et al. (2012) study the position accuracy depending on the satellite geometry, and satellite elevation. According to the existing correlation between day-to-day residual estimates, Wu and Hsieh (2010) prove that a multipath mitigating algorithm improves the accuracy of the GPS height determination by at least 40%. Teunissen (1997) consider the time-averaged GPS single-baseline model. This author studies its relation with the geometry-free model and the geometry-based model in a qualitative sense. The least-squares estimators of the model are also derived and their properties discussed. Finally, Luo et al. (2011) illustrate the discrepancies between the normal distribution assumption and reality, based on a large and representative data set of GPS phase measurements covering a range of factors, including multipath impact, baseline length, and atmospheric conditions.

Estimation of GPS positional reliability has been extensively studied in the literature from different points of view. For example, Hasewaga and Yoshimura (2007) estimate the GPS positional accuracy under different forest conditions. Rodriguez-Solano et al. (2012) study the impact of Earth radiation pressure on the positional accuracy. Cai and Gao (2012) combine GPS and GLONASS observations to improve accuracy. Odolina (2012) takes into account the temporal correlation for improving accuracy in real-time kinematic (RTK) network data. However, the study of the GPS positional accuracy after processing data is scarce from the stochastic perspective. Cressie and Kornak (2003) review a method for adjusting spatial inference in the presence of data location error, particularly for data that have a continuous spatial index. Specifically, their paper derives a kriging methodology in the presence

Fig. 1 Lakora, one of the 211 vertex points in Navarre; and map of Navarre
of location error. The authors study two location error models known as the coordinate-positioning model and the feature-positioning model. In Cressie’s work, the aim is to assess the effect on spatial prediction without making inferences about true feature locations. In Barber et al. (2006), map positional errors are defined as differences between locations represented in a spatial database and the corresponding unobservable true locations. The authors focus on GIS positional errors, explaining how estimates of positional error can be obtained, and providing estimates of the true location. Three bivariate variables are used: a feature location on the map, an associated GPS location, and a true location for that feature. The authors cast the map positional error problem as one registering GPS locations against the given map followed by inferring the true location based upon their knowledge of GPS accuracy. Wang et al. (2002a) study the dependence of the GPS positioning precision on the station location. They use a weighting function to account for an averaged number of satellites at different latitudes and demonstrate that GPS positioning precision reduces when the latitude increases. They also quantify the positioning precision in the north-south direction, showing that it is worse than in the east-west direction at all latitudes. The same conclusion is given by Meng et al. (2004) and Bosser et al. (2007), who study the impact of satellite geometry. Lack of temporal dependence specification between carrier phase measurements and the effect on precise positioning is studied by Wang et al. (2002b).

The step of processing raw data to transform them into geodetic measurements requires the specification of a mathematical model consisting of a functional model and a stochastic model (Tiberius, 2000). This process involves computing unknown parameters from a set of measurements or observations, preferably incorporating some measures that express the quality of the estimated parameter values (Tiberius, 1999). Least squares estimation constitutes the most popular technique to obtain the desired estimates. Then, it is clear that stochastic models are a compulsory requirement for obtaining final GPS measurements (Wang et al. 1998, Satirapod et al. 2002, Wang et al. 2005, Amiri-Simkooei et al. 2009). However, this kind of models is not the objective of our paper. Assuming that all these stochastic models are already specified and correctly estimated, we consider that the differences between real and final measured coordinates in GPS are originated by unknown reasons: if they were known, they could be corrected. That is why three spatial linear mixed models are proposed to estimate the unknown sources of uncertainty. Our goal is not to improve estimations by modifying any parameter of these stochastic models. On the contrary, we are going to assume that all these stochastic models are correct and perfectly estimated. Nonetheless, there may exist some misspecification or any other kind of uncertainty that we cannot change due to atmospheric disturbances, satellite geometry, satellite elevation, multipathing or any other already cited. In that case, we only assume that we can estimate the positional errors as the differences between the GPS measurements and the real ones. A map of these differences in X, Y, and Z axes will provide an additional and useful information to the user.

Then, this work addresses the estimation of GPS positional errors. The observations have been taken with a multiple-reference-station network in a real-time-kinematic system of dual-frequency using carrier-phase measurements. The application is done providing prediction of positional errors and their standard errors in the whole region of Navarre, Spain.

The paper is organized as follows. Section 2 explains the application and the terminology. Section 3 reviews the theory of the spatial linear mixed-effects model and defines the specific models for the three different axes X, Y, and Z. It also explains how these models allow for the estimation of positional errors. Section 4 provides the results and, finally, the Summary and Conclusions are presented.

### 2 Application

The application illustrates the performance of GPS techniques based on the positioning in real time services provided by Navarre active geodetic network infrastructure, called RGAN (see RGAN 2009). This network consists of 14 beacon receiver stations strategically located along the region. The static network consists of 211 vertex points whose locations are fixed and spread in the 10,000 km² of the region. The RGAN network is based on Leica Geosystems technology and uses Spider as the maintenance’s software from the same company. The equipment used in this work consists of a Leica

---

### Table 1 Abbreviations of network and the closest formats

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N_1)</td>
<td>NET_i-max_CMR</td>
</tr>
<tr>
<td>(N_2)</td>
<td>NET_i-max_CMR+</td>
</tr>
<tr>
<td>(N_3)</td>
<td>NET_i-max_Leica</td>
</tr>
<tr>
<td>(N_4)</td>
<td>NET_i-max_RTCM_2.3</td>
</tr>
<tr>
<td>(N_5)</td>
<td>NET_i-max_RTCM_3.1</td>
</tr>
<tr>
<td>(N_6)</td>
<td>NET_max_RTCM_3.1</td>
</tr>
<tr>
<td>C_1</td>
<td>CLOSEST_CMR</td>
</tr>
<tr>
<td>C_2</td>
<td>CLOSEST_Leica</td>
</tr>
<tr>
<td>C_3</td>
<td>CLOSEST_RTCM_2.3</td>
</tr>
<tr>
<td>C_4</td>
<td>CLOSEST_RTCM_3.1</td>
</tr>
</tbody>
</table>

---
ATX 1230 and a GNSS Smart Rover made up of a Leica 1230 Sensor, a GNSS antenna and a RX 1250 XC Controller. As Brown et al. (2006) point out, combining the Leica GPS Spider RTK network software with the Leica System 1200 GPS receivers clearly improve the measurements, even when using network correction data at a sampling rate of only 5s. The geodetic data are given with the European Terrestrial Reference System 1989 (ETRS89) (see Bruyninx et al. 2001), which is the reference system used by RGAN and validated by the Spanish National Geographic Institute (IGN). A preliminary work (Lizarraga-Garcia 2009) accomplished a simulation study where 54 sampled vertex points could be chosen to minimize the mean squared prediction error for the whole province. Following that study, the Department of Civil Engineering of the Navarre Government continued taking field measurements during 2011 and 2012. The sampled vertex points are regularly spread along the whole territory with a mean distance of 40km between them. They correspond to round and blue dots, while square and red dots are the beacon stations in Figures 2, 3, and 4. The sampled vertex points are also located at different heights, varying between 207m and 1844m with a mean height of 840m. Moreover, the corresponding ground-truth data is also known for these sampled vertex points. Ground-truth data have been obtained in the same conditions as the sampled measurements. That is, measurements for elaborating ground-truth data are taken at the same time and with the same rover as the sampled measurements in order to avoid influence of different satellite geometry, atmospheric variability, broadcast connections or device handling. However, the way of processing the information is different: while the sampled measurements are a limited number of real-time measurements using the real-time kinematics technique, RTK, the ground-truth data are the result of static postprocessing measurements during at least one hour. The difference between sample and ground-truth coordinates, in X, Y, and Z will provide the response variables $DiffX$, $DiffY$, and $DiffZ$ of the three stochastic positioning error models.

In this paper, real time kinematic (RTK) technique with phase, dual-frequency and a multiple-reference-station network is used. Thus, GPS positioning is corrected by adjusting measurements taken by two or more receivers. Positioning measurements of sampled vertex points have been taken using 10 different transmitting formats and they are based on both the closest reference station (“CLOSEST” type) and the network of closer stations (“NET” type). The names of the formats used and their abbreviations are summarized in Table 1.

Radio technical position commission for maritime services (RTCM) is a widely used protocol for communication between the reference stations and the mobile receivers at the location being determined. The use of RTCM standards allows for interoperability between equipment from different manufacturers. Edition 3.1 includes an interoperable definition for RTK Network operation, which supports centimeter-level accuracy positioning service over large regions. Leica SpiderNet was used to calculate single sites; however, MAX and iMAX corrections are used in RTCM 2.3 and 3.1.

The RTCM “Version 3” message format is optimized for RTK operation, which requires more data than earlier differential correction techniques, such as in “Version 2” (see Brown et al., 2006). Among the different formats, MAX format is very competitive because it selects automatically the optimum number of sites for the cell to generate master- auxiliary corrections for each rover. The auxiliaries are chosen from the surrounding stations to provide the best possible set of corrections for the rover’s position. For older rovers, Leica GPS Spider provides an individualized version of the master- auxiliary corrections, known as iMAX, that may be transmitted using older versions of RTCM. It has the advantage of using a lower bandwidth single site RTCM 2.3 or 3.0 format that can also be interpreted by older receivers that do not support the new network messages. The use of the closest reference station provides very accurate precision in each of the response variables and even better precision than some iMAX formats. MAX and iMAX show very similar values and a better performance than other RTK network formats. “NET_i-max_CM” and “NET_i-max_CM+” are Trimble formats and “NET_i-max_Leica” is a very specific format used by Leica.

3 Spatial Linear Mixed-Effects Model

Traditionally, kriging has been largely used in spatial modelling, as a minimum squared prediction method of spatial variables where spatially correlated observations come from a single realization of a spatial stochastic process. In this work, repeated measurements in each location have been taken in several transmitting RTK GPS formats, yet the number of transmitting formats for each location can differ. In order to analyze positional errors, spatial linear mixed-effects models are proposed as an alternative to kriging. The models contain both fixed and random effects. Fixed effects allow for modeling positional errors, and random effects allow for modelling their variance-covariance structure. The inclusion of random effects reduces the number of parameters in the specification of the variances and gives more flexible models (see McCullogh and Searle,
In particular, the same random effects associated with the repeated measurements in each location express a kind of correlation within this group of data. Moreover, spatially dependent data may share the same covariance matrix of the error term as is done in kriging. In this case, spatial linear mixed models can incorporate the spatial dependence into the covariance matrix of the error term in a similar way as kriging does, but they can also incorporate the correlation between samples taken under the same transmitting format through a common random effect. Thus, spatial linear mixed models are an alternative for modelling spatial linear models with repeated measurements (Militino et al. 2008). The general spatial linear mixed model is written as

\[ Y = X\beta + Zu + \epsilon, \]

(1)

where \( X \) is the design matrix \((n \times (p + 1))\) of the unknown fixed effects, \( \beta = (\beta_0, \beta_1, \ldots, \beta_p)^t \) is the \((p + 1)\) vector of fixed parameters and \( u = (u_1, \ldots, u_q)^t \) is the \( q \)-vector of random effects with an associated \((n \times q)\) design matrix \( Z \). This matrix consists of a collection of zeros and ones that indicate the association between random effects and observations. The random variables \( u \) and \( \epsilon \) are assumed to have mean equal to \( 0 \) and covariance matrix

\[ \text{cov}[u, \epsilon^t] = \sigma^2 \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix}. \]

Hence, the mean and variance of \( Y \) are given by \( E[Y] = X\beta \) and \( V = ZGZ^t + R \) respectively.

Historically, one of the first solutions to the estimation process of fixed and random effects was proposed by Henderson (1953). In principle, normality for the known fixed effects, \( \hat{\beta} \) and \( \hat{\theta} \), respectively, where \( \hat{\theta} \) is the vector of variance components and \( \hat{\theta} \) is its estimator. The estimates for \( \beta \) and \( \hat{u} \) are then called empirical BLUE (EBLUE) and empirical BLUP (EBLUP), respectively.

The vector of variance components \( \theta \) may be estimated by using the fitting of constants method or likelihood-based methods, such as the maximum likelihood (ML) or the restricted maximum likelihood (REML) method. In this paper, the REML method is used, and the restricted log likelihood function of Model 4 is given by

\[
\begin{align*}
l_R(\theta | Y) &= -\frac{1}{2} \log |V_\theta| - \frac{1}{2} (Y - X\hat{\beta}_\theta)^t V_\theta^{-1} (Y - X\hat{\beta}_\theta) \\
&\quad - \frac{1}{2} \log |X^t V_\theta^{-1} X| - \frac{n - p}{2} \log (2\pi). \tag{4}
\end{align*}
\]

Here, \( \hat{\beta}_\theta \) is the estimate of \( \beta \) given by expression \( 2 \) and both \( \hat{\beta}_\theta \) and \( V_\theta \) depend on the vector of parameters \( \theta \). The function given in expression \( 4 \) does not involve \( \beta \), and therefore the resulting variance parameter estimators are unbiased. For this and other advantages, REML is usually recommended.

The asymptotic covariance matrix of the parameter estimators can be obtained using the standard theory of maximum likelihood, that is, calculating the inverse of the observed information matrix evaluated at the optimum. Empirically, the approximate covariance matrix of \( \beta \) and \( \hat{u} \) is calculated as the covariance of the prediction error, because neither of the predictors \( \hat{u} \) nor \( \beta \) account for the uncertainty in the prediction of the other. This covariance matrix is given by

\[
D = E \left[ \begin{pmatrix} \hat{\beta} - \beta \\ \hat{u} - u \end{pmatrix} \begin{pmatrix} \hat{\beta} - \beta \\ \hat{u} - u \end{pmatrix}^t \right] = \begin{pmatrix} D_{11} & D_{21} \\ D_{21} & D_{22} \end{pmatrix},
\]

where

\[
D_{11} = \text{var}(\hat{\beta} - \beta) = \text{var}(\hat{\beta}) = (X^t V^{-1} X)^{-1} = A,
\]

\[
D_{21} = \text{cov}(\hat{\beta} - \beta, (\hat{u} - u)^t) = ZGZ^t X D_{11},
\]

\[
D_{22} = \text{var}(\hat{u} - u) = (Z^t R^{-1} Z + G^{-1})^{-1} - D_{21} X^t V^{-1} X Z G.
\]

In practice, it is estimated by substituting \( G \) and \( V \) by \( G_\theta \) and \( V_\theta \), respectively.

Let \( x_{i,j}, y_{i,j} \), and \( z_{i,j} \) be the average of the three measurements taken in each sampled vertex point, where \( i = 1, \ldots, n_i \) indicates the observation within each format and \( n_j \) is the total number of observations of the \( j \)-th format \((j = 1, \ldots, 10)\). Coordinates \( x_{i,j} \) and \( y_{i,j} \) correspond to the UTM coordinates, and \( z_{i,j} \) to the orthometric elevation. Let \( x_{i,true}, y_{i,true} \), and \( z_{i,true} \) be the corresponding ground-truth coordinates. Three types of measurement errors are considered here: differences in the \( X \) axis calculated as \( \text{Diff}X_{i,j} = x_{i,true} - x_{i,j} \); differences in the \( Y \) axis, \( \text{Diff}Y_{i,j} = y_{i,true} - y_{i,j} \); and finally differences in the \( Z \) axis (orthometric elevation) given by \( \text{Diff}Z_{i,j} = z_{i,true} - z_{i,j} \).
Besides the measured coordinates, two additional covariates are used as fixed effects in every model: qalt and qposic corresponding to the standard deviations of the observable height and position, respectively.

The three different response variables are fitted with spatial linear mixed models, using libraries “nlme” for parameter estimation (Pinheiro and Bates, 2000) and “AlCemodavg” (Mazerolle, 2006) for standard errors of the predictions. Both libraries can be obtained from the free statistical package R (R Development Core Team 2012). See also Ugarte et al. (2008) for linear statistical modelling in R. The stochastic models are given by

\[
\text{Diff} \ X_{i,j} = \beta_{0,x} + \beta_{1,x} x_{i,j} + \beta_{2,x} y_{i,j} + \beta_{3,x} z_{i,j} + \epsilon_{i,j,x},
\]

\[
\text{Diff} \ Y_{i,j} = \beta_{0,y} + \beta_{1,y} x_{i,j} + \beta_{2,y} y_{i,j} + \beta_{3,y} z_{i,j} + \epsilon_{i,j,y},
\]

\[
\text{Diff} \ Z_{i,j} = \beta_{0,z} + \beta_{1,z} x_{i,j} + \beta_{2,z} y_{i,j} + \beta_{3,z} z_{i,j} + \epsilon_{i,j,z}.
\]

Let us denote by \( w \) the directions \( x \), \( y \), or \( z \). Then, \( u_{j,w} \) is the \( j \)th random effect common to the observations taken at the same transmitting format \( j \), such that the overall vector \( \mathbf{u}_w = (u_{1,u}, \ldots, u_{10,u}) \sim N(0, \sigma_w^2 \mathbf{G}_w) \); and \( \epsilon_{i,j,w} \), where \( i = 1, \ldots, n_j \), is the \( i \)th \( j \)th random error, such that \( \epsilon_{i,1,w}, \ldots, \epsilon_{54,10,w} \sim N(0, \sigma_w^2 \mathbf{R}_w) \). The heteroscedasticity present in all the three models has been corrected expressing model variance as a power function of the covariates, in this case \( qalt_{i,j,w} + \text{qposic}_{i,j,w} \gamma \). Heteroscedasticity has been tested by the likelihood ratio test in all the models. Estimation of covariance parameters has been obtained by restricted maximum likelihood, and the exponential covariance function has been used. At lag \( h \) it corresponds to

\[
C(h) = c_0 + \sigma_w^2 \exp(-h / r_w),
\]

where \( c_0 \) is the nugget effect and \( r_w \) is the range.

### 4 Results

Measurement units are important when working with statistical models as kriging, because the kriging procedure could be sensitive to these units. However, from a statistical point of view, the projected coordinates are preferable to angular coordinates such as longitude and latitude because of two reasons. Firstly, singularities in matrix inversion processes are avoided and, secondly, they provide a constant distance relationship anywhere on the map. On the other hand, UTM coordinates are easy to interpret for users that are non-experts in Geodesy. Conversion between coordinates can be done directly through mathematical expressions. Anyway, similar results could be drawn when using alternative coordinates.

Estimates of the covariance range \( (r_w) \), variance power \( (\delta_w) \), and residual standard error \( (\sigma_w) \) are given in Table 2. Covariance ranges of \( \text{Diff} \ X, \text{Diff} \ Y, \) and \( \text{Diff} \ Z \) vary between 6.15 and 1.7mm. As expected, the spatial dependence shows that similar positional errors are only found at very short distances, behavior particularly pronounced in the \( X \) direction. In the \( Y \) and \( Z \) direction spatial dependence is smaller. The lowest variance power \( (\delta_w) \) of the heteroscedastic variance corresponds to the \( X \) direction and the biggest one to the \( Z \) direction. Thus, larger heteroscedasticity is expected in the \( Z \) or \( Y \) direction with regard to \( X \). Residual standard errors (0.27, 0.15, and 0.23) are roughly similar, showing that the covariates explain accurately the variability of the response variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range (r)</th>
<th>var power (δ)</th>
<th>residual s.e. (σ_w)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiffX</td>
<td>6.15</td>
<td>0.99</td>
<td>0.27</td>
</tr>
<tr>
<td>DiffY</td>
<td>1.70</td>
<td>1.32</td>
<td>0.15</td>
</tr>
<tr>
<td>DiffZ</td>
<td>2.59</td>
<td>1.34</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 3 Summary of the predictions and their standard errors in mm by network and the closest formats for DiffX
Fig. 2 Prediction map of the positional errors (left) and their standard errors (right) in $X$ for NET$_j$-max_Leica format

Fig. 3 Prediction map of the positional errors (left) and their standard errors (right) in $Y$ for NET$_j$-max_Leica format

Fig. 4 Prediction map of the positional errors (left) and their standard errors (right) in $Z$ for CLOSEST_Leica format
Table 4 Summary of the predictions and their standard errors in mm by network and the closest formats for \( \text{DiffY} \)

<table>
<thead>
<tr>
<th>format</th>
<th>min</th>
<th>max</th>
<th>range</th>
<th>mean</th>
<th>prediction s.e. min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_1 )</td>
<td>-2.89</td>
<td>15.56</td>
<td>18.45</td>
<td>1.10</td>
<td>0.41</td>
<td>4.35</td>
</tr>
<tr>
<td>( N_2 )</td>
<td>-3.87</td>
<td>6.49</td>
<td>10.36</td>
<td>0.89</td>
<td>0.44</td>
<td>3.17</td>
</tr>
<tr>
<td>( N_3 )</td>
<td>-5.53</td>
<td>2.78</td>
<td>8.31</td>
<td>-0.57</td>
<td>0.43</td>
<td>2.06</td>
</tr>
<tr>
<td>( N_4 )</td>
<td>-3.41</td>
<td>4.76</td>
<td>8.16</td>
<td>0.01</td>
<td>0.38</td>
<td>2.29</td>
</tr>
<tr>
<td>( N_5 )</td>
<td>-2.53</td>
<td>6.69</td>
<td>9.23</td>
<td>0.41</td>
<td>0.40</td>
<td>2.79</td>
</tr>
<tr>
<td>( N_6 )</td>
<td>-2.56</td>
<td>6.48</td>
<td>9.05</td>
<td>0.09</td>
<td>0.40</td>
<td>2.33</td>
</tr>
</tbody>
</table>

\( C_1 \) | -2.29 | 8.08  | 10.37 | 2.15  | 0.43                | 2.87|
\( C_2 \) | -4.20 | 3.89  | 8.09  | -0.08 | 0.38                | 2.08|
\( C_3 \) | -2.30 | 9.51  | 11.81 | 0.89  | 0.46                | 3.45|
\( C_4 \) | -2.56 | 6.48  | 9.04  | 0.84  | 0.44                | 2.81|

Table 5 Summary of the predictions and their standard errors in mm by network and the closest formats for \( \text{DiffZ} \)

<table>
<thead>
<tr>
<th>format</th>
<th>min</th>
<th>max</th>
<th>range</th>
<th>mean</th>
<th>prediction s.e. min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_1 )</td>
<td>-38.34</td>
<td>-11.77</td>
<td>26.57</td>
<td>-23.87</td>
<td>3.69</td>
<td>8.50</td>
</tr>
<tr>
<td>( N_2 )</td>
<td>-26.16</td>
<td>-6.35</td>
<td>19.81</td>
<td>-16.69</td>
<td>3.70</td>
<td>6.48</td>
</tr>
<tr>
<td>( N_3 )</td>
<td>-6.48</td>
<td>12.48</td>
<td>18.95</td>
<td>5.48</td>
<td>3.71</td>
<td>5.01</td>
</tr>
<tr>
<td>( N_4 )</td>
<td>-24.29</td>
<td>-5.95</td>
<td>18.34</td>
<td>-14.04</td>
<td>3.69</td>
<td>5.30</td>
</tr>
<tr>
<td>( N_5 )</td>
<td>-9.78</td>
<td>13.85</td>
<td>23.63</td>
<td>4.90</td>
<td>3.69</td>
<td>5.95</td>
</tr>
<tr>
<td>( N_6 )</td>
<td>-10.11</td>
<td>16.50</td>
<td>26.61</td>
<td>6.98</td>
<td>3.70</td>
<td>5.35</td>
</tr>
</tbody>
</table>

| \( C_1 \) | -13.19 | 12.52  | 25.71 | 1.42   | 3.70                | 6.12|
| \( C_2 \) | -5.98  | 12.51  | 18.50 | 6.19   | 3.69                | 5.02|
| \( C_3 \) | -9.11  | 18.50  | 27.70 | 6.13   | 3.71                | 6.85|
| \( C_4 \) | -11.83 | 9.64   | 21.48 | 2.91   | 3.71                | 6.01|

 Whereas the UTM coordinates \( X \) and \( Y \) and the orthometric height \( Z \) are known, \( \text{qalt} \) and \( \text{qposic} \) covariates are not. As a result, \( \text{qalt} \) and \( \text{qposic} \) sampled values of the closest sampled location are used for every location of the prediction grid. That is why the maps of the predictions have a polygonal shape that is defined by the closest locations to the sampled one (see Figures 2 and 3). Tables 3-4 show the transmitting format, the minimum, maximum, range, and mean of the prediction measurement errors, and their minimum and maximum standard errors, for the whole Navarre region. A summary of these numerical results is plotted in Figures 2 and 3 (for network solutions) and 4 (for the closest reference station), where blue bars indicate the range of the predicted positional errors and red bars depict the range of their standard error. The prediction range and the standard errors of \( \text{DiffX} \) and \( \text{DiffY} \) are quite similar, yet in \( \text{DiffZ} \) standard errors are a bit bigger. Looking jointly at the range and mean of the positional error prediction, and taking into account their standard errors, it can be concluded that “NET-i-MAX Leica” (\( N_3 \)) is the most accurate format in the three axes. In the \( X \) direction, network formats are more accurate than the closest formats, but in the \( Y \) or \( Z \) direction the closest formats are very competitive, specifically “CLOSEST Leica” (\( C_2 \)) format. In \( X \) and \( Z \) a positive bias always exists because prediction means are always positive, but in the \( Y \) direction some formats have a negative bias. Also, it is observed that \( N_3 \) format can be even more accurate in \( Y \) than in \( X \), although \( Z \) is always less accurate than planar coordinates. In \( N_1, N_5, C_1, \) and \( C_3 \) transmitting formats, the West-East or \( X \) direction has less range error than \( Y \), which agrees with the result obtained by Wang et al. (2002a), Meng et al. (2004), and Bosser et al. (2007). However, this behavior does not happen in the rest of the formats. Moreover, \( X \) error mean is always bigger than \( Y \) error mean, yet range error in \( X \) is not always bigger than in \( Y \). Overall, network formats are not necessarily more precise than the closest formats in all the axes. In the \( X \) direction, the prediction error range for all formats is between -4mm and 9mm, in \( Y \) between -4 and 16mm, and in \( Z \) between -38mm and 19mm. In the best format (\( N_3 \)), \( \text{DiffX} \) can vary between -4mm and 6mm, and \( \text{DiffY} \) between -4mm and 3mm, and in the \( C_2 \) format \( \text{DiffZ} \) varies between -6mm and 12.5mm. In this work, we illustrate several maps where the \( X \), \( Y \), and \( Z \) axes are expressed in UTM coordinates that were scaled to Km and moved to a new origin defined by the minimum of \( X \) and \( Y \). With this new coordinates origin, we could easily interpret distances to check spatial dependence and compare small geographical regions with similar errors. The positional errors are mapped in a grid of 1km \( \times \) 1km of the whole region of Navarre. Figures 2 and 3 show predicted positional errors of \( \text{DiffX} \) and \( \text{DiffY} \) on the left, and their standard errors on the right, for the whole Navarre in “NET-i-MAX Leica” format. Figure 4 shows the predictions of \( \text{DiffZ} \) and its standard errors in “CLOSEST Leica” format. The geographical pattern of these maps is similar in all directions. North and North-East directions present larger positional and standard errors. They correspond to the Pyrenees region, where humidity and mountains are frequent and, perhaps, broadcast connections are also weak. Finally, some spots in the center West direction with larger positional errors than on the rest are also seen. Perhaps a bad broadcast connection is also present here.

Summary and Conclusions

The primary aim of this work is to analyze the quality of the GPS measurements taken with a multiple-reference
station network on RTK system with dual-frequency and carrier-phase measurements. We assuming that all uncontrollable sources of uncertainty are random and they can be explained using spatial linear mixed-effects models. These models provide a good methodology for jointly predict the positional errors in different formats when spatial dependence is assumed. The scenario for illustrating the measurements is based on a high quality device and a geodetic infrastructure of 14 beacon reference stations over a region of 10,000 km². RTK has been the differential process used for correcting position measurements in real time. Also, 10 different formats have been used to transmit data, allowing for a connection to the closest or to a network of closer reference stations. The variables \( \text{DiffX, DiffY, and DiffZ} \), calculated as differences among ground-truth coordinates and measured GPS coordinates, have shown different accuracy and precision performances between them and within the formats, yet all of them provide centimeter-accuracy positions, less than 1cm for \( X \) or \( Y \), and less than 1.5cm for \( Z \). Higher errors for \( Z \) have also been reported in the literature (Meng et al. 2004).

The main contribution of our paper is the methodology used for providing positional error maps when using a GPS receiver. Thus, this methodology could be easily extended to other regions, although our study is limited to Spain. We have focused on a multiple-reference-station RTK network in dual-frequency with carrier-phase measurements because a high performance is attained. However, other alternatives could also be used. In addition, the statistical methodology proposed in this paper is very useful to check the reliability of a beacon station network in a certain region or to ascertain the optimal location of a new beacon station.

Consequently, although GPS commercial devices present a high level of accuracy and precision, this can be increased using appropriate transmitting formats. In particular, network formats are not always better than the closest formats. Here, the closest formats are very competitive when measuring on the \( Y \) or \( Z \) axes, and network formats are better for the \( X \) coordinate.

The transmitting formats are not a critical issue, but it can influence the quality of the measurements. Thus, we consider different models depending on the formats, and deduce that there are slight differences between them. This issue allows us to check if the location-based error distributions are in fact independent or not of these transmitting formats.
In addition, precision in \( Y \) is not always less than in \( X \) for all the transmitting formats, contradicting the theory. In specific formats, the precision of \( Y \) can be even better than the prediction of the \( X \) coordinate. On average, \( Y \) direction is more precise than \( X \) direction (because its means are lower than \( X \) means), but it shows greater variability (because \( Y \) ranges are greater than \( X \) ranges). Unfortunately, \( Z \) is always less accurate than planar coordinates and here the closest format is very competitive with regard to network formats.

Final findings show that positional errors are not completely independent of the locations. Valley or mountain regions with high degrees of humidity present both greater positional and standard errors, possibly due to a weak broadcasting coverage.

To summarize, it has been shown that positional error maps derived from spatial linear mixed models can help the user to evaluate the quality of positional measurements.

Acknowledgements  The authors would like to thank the comments made by the associate editor and the three referees. They have contributed to improve this manuscript.

The authors would also like to thank the Department of Civil Engineering of the Navarre Government for providing the data, and more specifically to Miguel Ángel Jiménez de Cisneros. This research has been supported by the Spanish Ministry of Science and Education (project MTM2011-22664 jointly sponsored with Feder grants).


Lizarraga-Garcia E (2009) Estimation of the measurement error model when applying GNSS techniques which are based on the positioning in real times services provided by RGAN. Technical Report: Final Degree Project in Industrial Engineering, ETSIIT, Public University of Navarre, Spain.


the time-averaged satellite geometry. J Geod 1997 71(5):290-301.


