Stochastic Modelling of GPS Phase Observations for Improved Quality Estimation

Neil Brown¹, Allison Kealy²
and
Ian Williamson³

¹PhD Candidate
Email: neb@sunrise.sli.unimelb.edu.au

²Lecturer
Email: akealy@unimelb.edu.au

³Professor and Head of Department
Email: ianpw@unimelb.edu.au

Department of Geomatics
The University of Melbourne, Australia
Tel: 03 8344 6806 / Fax: 03 9347 2916

ABSTRACT

Data quality information has been recognised as essential in assessing the fitness for use of any spatial dataset, and fundamental to enabling efficient and effective data integration through spatial data infrastructure (SDI). Missing or inaccurate data quality information can result in inappropriate use of the data with associated consequences of poor decision making, reduced utility and decreased market value.

The increasing use of the Global Positioning System (GPS) as a primary data acquisition source for spatial databases highlights the significance of this problem. At present the measures of quality for GPS derived coordinates given by commercial software packages tend to be unrealistic and are more often than not optimistic. This is because not all of the systematic and random errors present in the observations are fully modelled through the standard functional or stochastic models used. This paper presents some of the current problems in identifying the quality of GPS data as derived from commercial processing software. Common GPS processing strategies are reviewed in the context of error modelling and data quality. Finally, current research activities into strategies for maximizing GPS data quality are presented.
INTRODUCTION

Recognition of spatial data as a valuable resource for decision making and asset management is driving the development of spatial data infrastructure (SDI) at global, regional, national and local scales. In general the purpose of SDI is to enable efficient collection, management, access, delivery and utilisation of spatial data (Coleman and McLaughlin, 1998, Rajabifard et al., 2000, GI2000, 1996). SDI are initiatives that aim to share consistent data sets in an easy and secure manner with a variety of users and typically consist of an institutional framework, technical standards, fundamental data sets, access network and human resources (Rajabifard et al., 2000, Coleman & McLaughlin, 1998). The potential applications of SDI are unlimited, however they are generally most recognised for their benefits in resource maximization and decision making.

At the core of SDI is the desire to integrate spatial data sets from a range of sources into a GIS in order to manage assets or aid decision-making. The geodetic framework is one of the fundamental data sets in SDI because it is necessary to reference spatial data and facilitate the integration of data from a variety of sources and regions. Without a common reference framework data sets will be inconsistent and not match. Therefore, in order to integrate two data sets consideration must be given to the reference frames (geodetic datums) used and the transformation between them. Unmodelled systematic errors in the datum definition can result in loss of data quality if the transformation does not account for them. The quality of individual records within a dataset also affect data integration, as real-world relationships may not be represented within the data. Without understanding of the data quality such misrepresentations have the potential to adversely affect decisions made based on the data.

Spatial data quality is expressed in terms of accuracy and precision and is a function of the geodetic infrastructure, equipment, field practices and processing methodology used to gather the data. Often accuracy is difficult to measure and precision is used instead. This has the potential to give misleading estimates of the accuracy and is a major focus of this paper. The reader is referred to Hunter (1999) for information on the handling of spatial data quality in GIS. Precision, or repeatability, is a statistical term that describes the random error in a solution and is represented by a standard deviation or RMS (root-mean-square). A system in which the calculated precision estimate of the parameters is representative of the true error is said to have high fidelity. If a measurement system has high fidelity, the user is better able to ensure that measurements are within tolerance. The difficulties in determining good (high fidelity) data quality estimates shall be discussed with reference to GPS.

The accuracy requirements of data collected for use within SDI depend on the intended application of the data. Regardless of how coarse the data are there is a requirement to have a measure of the data quality. Knowledge of data quality is essential to assessing data’s fitness for purpose and increasing its potential useability within a SDI. It is recognised that for spatial data to be used correctly and wisely, metadata that describes data quality, in addition to other attributes, should accompany the spatial data (Nebert, 2000, GI2000, 1996). Depending on the particular data and its possible uses this may mean that metadata down to the individual feature, object or record level is necessary (Nebert, 2000). Data quality for positional information is generally given in terms of a standard deviation and results from the data capture process. Mis-specification of this quality parameter may result in inappropriate use and reduced utility and
market value of the data. At present, the measures of GPS quality given by common software packages are either over-optimistic, or conversely, are overly conservative and therefore have low fidelity (Keenan and Cross, 2001, Barnes et al., 1998, Wang, 1999). This lack of fidelity can potentially impact SDI development starting with the geodetic infrastructure and at every jurisdictional level and is an active area of research worldwide.

The Global Positioning System (GPS) has established itself as the primary means for spatial data capture and is increasingly being used in a wide range of applications from personal navigation to geodesy. In recent years there has been a lot of discussion regarding GPS data quality. Much of the research in this area has focused on the stochastic models used in GPS processing and their importance to ambiguity resolution (eg Han, 1997, Teunissen, 1997, Wang, 1999). This paper outlines the fundamental theory of GPS and the methodology behind current processing algorithms and models in order to illustrate the quality concerns in GPS positioning.

GPS POSITIONING

Each satellite in the GPS constellation transmits signals on two sinusoidal carrier waves with frequencies of 1575.42 MHz and 1227.60 MHz, designated L<sub>1</sub> and L<sub>2</sub>. Modulated onto L<sub>1</sub> are two pseudo-random noise (PRN) codes, C<sub>1</sub> and P<sub>1</sub> in RINEX (Gurtner, 2000) notation. A second P-code (P<sub>2</sub>) is modulated onto the L<sub>2</sub> frequency. The C<sub>1</sub> (or civilian/access or C/A) code repeats every millisecond and is unique for every satellite.

A GPS receiver generates internal copies of the PRN codes. Assuming the clocks in the satellite and receiver are both synchronised to GPS time, the signal travel time can be determined by measuring the shift between the internal and incoming versions of the code. The PRN codes are designed so that they have a low auto correlation allowing the shift to be measured precisely and unambiguously. Multiplication of the transmission time by the speed of light gives the range between the satellite and receiver. However, because the satellite and receiver clocks are not perfectly synchronised with GPS time and the influence of other error sources these ranges are called pseudoranges. It is possible to measure the carrier beat phase used to transmit the codes more precisely than the codes themselves, however the so-called integer ambiguities must be resolved for centimetre and better positioning (Hoffman-Wellenhof et al., 1994). Both phase and code measurements can be written in the form of pseudoranges.

A phase pseudorange Φ<sub>r</sub> between receiver <i>r</i> and satellite <i>s</i> (in metres) can then be expressed as (van der Marel, 1998)

\[
\Phi_r^s = \|s^r(t) - r^s(t)\| + c\delta t^s(t) - c\delta t_r(t) + I_r^s(t) + T_r^s(t) + M_r^s(t) + D_r^s(t) + \lambda N_r^s + e_r^s(t)
\]

(1)

where:

- \(s^r, r^s\) are the vectors of satellite and receiver coordinates at time of transmission (t);
- \(\delta t^s, \delta t_r\) are the satellite and receiver clock offsets from GPS time;
\( I_r^s \) is the apparent error in the range caused by the ionosphere;
\( T_r^s \) is the apparent error in the range caused by the troposphere;
\( M_r^s \) is the multipath error;
\( D_r^s \) is the antenna phase centre offset and variation;
\( e_r^s \) is the measurement noise;
\( \lambda \) is the wavelength of the signal; and
\( c \) is the speed of light.

Pseudoranges are simultaneously measured from a number of satellites. Measurements to at least four satellites are required to calculate a three-dimensional position because the receiver clock must also be estimated. This reduces to three if Doppler measurements are available (Hoffman-Wellenhof et al., 1994). Data screening is performed to remove outliers and (for phase data) detect and repair cycle slips. The least squares algorithm is used to estimate the unknown coordinates of the receiver and (optionally) other parameters of interest along with their precisions. Consequently an observation equation of the form:

\[
\Phi = f(X_r, Y_r, Z_r, \Delta \delta t_r)
\]  

(2)

can be constructed where

- \( \Phi \) pseudorange measurement
- \( X_r, Y_r, Z_r \) coordinates of the receiver
- \( \delta t_r \) receiver clock offset

This must be linearised using a Taylor’s Series expansion for use in the least squares algorithm. On linearising the equations become:

\[
\Phi - f(X'_r, Y'_r, Z'_r, \Delta \delta t'_r) = \frac{\partial f}{\partial X_r} \Delta X_r + \frac{\partial f}{\partial Y_r} \Delta Y_r + \frac{\partial f}{\partial Z_r} \Delta Z_r + \frac{\partial f}{\partial \delta t_r} \Delta \delta t \]

(3)

where \( X'_r, Y'_r, Z'_r, \Delta \delta t'_r \) are approximate values (a priori estimates) of the receiver’s position and clock offset and \( \Delta X_r, \Delta Y_r, \Delta Z_r, \Delta \delta t \) are corrections to the approximate values and are the unknowns being sought.

The \( m \) measurements and \( u \) parameters are then represented in the linear form using the indirect method as:

\[
v = A \begin{bmatrix} x \end{bmatrix} + w
\]

(4)

where \( v \) is a vector least-squares residuals, \( A \) is the design matrix resulting from the mathematical (functional) model which relates the observations to the unknown parameters being sought, \( x \) is the vector of unknown parameters (corrections to the a priori estimates) and \( w \) is a vector of computed minus observed measurements. A \( m \) by \( m \) variance-covariance matrix \( C_1 \) is
constructed based on the stochastic model which describes the statistical behaviour of the observations in context of the mathematical model. The diagonal elements of $C_1$ are the variances ($\sigma^2$) of the measurements.

The least-squares optimised estimates for the parameters is then obtained by iterating (Mikhail, 1976)

$$\hat{x} = -(A^T PA)^{-1} A^T Pw$$

(5)

where $P$ is the weight matrix, defined as the inverse of the variance-covariance matrix $C_1$.

The variance-covariance matrix of the least squares estimates is obtained via (Mikhail, 1976)

$$C_\hat{x} = \sigma_0^2 (A^T PA)^{-1} \quad \text{where} \quad \sigma_0^2 = \frac{v^T P v}{(m-u)}$$

(6)

Obviously, the mathematical and stochastic models are critical to parameter estimation. It is the task of the mathematical model to account for systematic errors in the observations whilst the stochastic model describes their random nature. Failure to correctly model the systematic biases will result in inaccurate coordinates. Inadequately modelled random errors, which are generally small, have only a minor effect on the parameters but cause low fidelity quality estimates and can make ambiguity resolution more difficult. Unfortunately many of the error sources listed above are difficult to model, especially if the GPS antenna is moving. As such several processing algorithms are commonly adopted when trying to obtain positions accurate to between the decimetre and millimetre level from GPS, all of which impose restrictions on the equipment and measurement procedure that is used. The most widely known of these methods is differential positioning (or DGPS). Note the term DGPS is normally only used for positioning based on the PRN codes. Another technique employs linear combinations of the code and/or phase measurements on one or more frequencies to eliminate or estimate some sources of error.

**Linear Combinations**

If both code and phase measurements are available on one or more frequencies linear combinations of the measurements may be formed. A number of linear combinations are widely used to estimate or eliminate errors in the GPS signals. For instance, a linear combination of code and phase measurements on a single frequency can be used to estimate the code multipath (Estey and Meertens, 1999). Other linear combinations are used to form new observation equations as part of the mathematical model. Two such linear combinations (in Bernese notation) are the $L_3$ ionosphere-free and the $L_5$ wide-lane linear combinations.

The $L_3$ combination is calculated as (Hugentobler et al., 2001)

$$L_3 = \frac{1}{t_1^2 - t_2^2} (t_1^2 L_1 - t_2^2 L_2)$$

(7)
and is very useful for ambiguity resolution because it eliminates almost all of the ionospheric delay. However, the noise on L₃ is approximately three times greater than that of L₁. Also the ambiguities lose their integer nature when L₃ is formed meaning that L₃ cannot be directly used to resolve L₁ and L₂ ambiguities. Instead L₃ is used in conjunction with another combination such as L₅.

Calculated as (Hugentobler et al., 2001)

\[
L₅ = \frac{1}{f₁ - f₂}(f₁L₁ - f₂L₂)
\]

(8)

the wide-lane linear combination is named because of its long (~86cm compared with 19cm for L₁) wavelength. Whilst the noise is greatly magnified the influence of the ionosphere is relatively small in relation to the wavelength and the ambiguities retain their integer nature. If the wide-lane ambiguities are resolved the determination of the L₁ (and hence L₂) ambiguities is greatly simplified.

**Differencing**

Whilst the previously mentioned errors are unique to a particular receiver-satellite combination, errors between another receiver and the same satellite will be similar if the receivers are close to one another (see Figure 1). The same is true for one receiver and two satellites, or between one receiver and one satellite at different times. These spatial and temporal correlations in the observations are a function of baseline length and arise from errors caused by the satellite and receiver clock, troposphere, ionosphere and orbits. The spatial correlation between the measurements is the basis for the double-differencing procedure, whereby one receiver is positioned relative to another receiver with known coordinates. A typical example is the case of a roving receiver being positioned relative to receiver from a permanent array of reference stations (such as the Australia-wide Australia Regional GPS Network (ARGN) operated by AUSLIG or Land Victoria’s GPSnet in Victoria).

The double difference for observations taken simultaneously with two receivers (i and j) from two satellites (m and n) is given by:

\[
\nabla Δϕ_{ij}^mn = (ϕ_i^n - ϕ_j^n) - (ϕ_i^m - ϕ_j^m)
\]

(9)

where \( \nabla Δ \) is the double difference operator and \( ϕ \) is the code or phase observation. This can be performed as a matrix operation:

\[
Δ\nabla \bar{ϕ} = D\bar{ϕ}
\]

(10)

where \( \bar{ϕ} \) is a vector of code or phase measurements and \( D \) is the double differencing matrix.

Differencing has the effect of eliminating the spatially correlated errors and is most effective with short baselines. Double-differencing effectively removes errors resulting from the satellite and receiver clocks and significantly reduces errors caused by the ionosphere, troposphere and
satellite orbits. The downside of this procedure is that each double difference requires four one-way measurements.

**Figure 1.** Correlations in the GPS signals (Barnes and Cross, 1998)

**LIMITATIONS IN ERROR HANDLING**

Using the differencing and linear combination algorithms outlined above most of the systematic error can be eliminated from the observations. This results in greatly simplified observation equations and, hence, a simplified mathematical model. The functional models used in GPS positioning are well defined and documented and are generally accepted (Bock et al., 1986, Hoffman-Wellenhof et al., 1994, Strang and Borre, 1997). However, some systematic errors (such as multipath) are not removed by these methods and residual error remains from other sources. Empirically derived mathematical models for the troposphere and antenna phase centre variations can help in reducing this error. The magnitude of the remaining error depends on:

- multipath,
- the processing algorithms used (eg L₃ used to eliminate the ionosphere),
- baseline length,
- antenna phase centre model and tropospheric model used,
- ephemeris type (broadcast or precise), and
- the parameters estimated explicitly by the mathematical model (eg tropospheric estimates).

In addition to the systematic error, correlations and noise exist in the data and should be accounted for in the stochastic model. Modern receivers are able to measure the carrier phase to
better than 0.01 cycles, which corresponds to millimetre precision (Hoffman-Wellenhof et al., 1994). This noise is amplified when some linear combinations are formed (e.g. $L_3$ mentioned previously).

Correlations result from a number of processes. The relatively slow variation in the atmosphere compared to the GPS observation intervals introduces a time correlation in the errors. When anti-spoofing is enabled, receivers must apply reconstruction techniques to enable measurement of the P-code and carrier phase on $L_2$. This can result in a correlation between the observables on the two frequencies (Bona, 2000). Some receivers also filter the measurements in an effort to reduce noise introducing a time correlation (Bona, 2000). Receivers have also been seen to have inter-channel biases and $L_1/L_2$ differential delays (Gao et al., 1999, Lin and Rizos, 1996).

**CURRENT STOCHASTIC MODELS**

The least squares algorithm relies on correct weighting of the measurements to provide unbiased estimates of the parameters. Residual and unmodelled systematic errors in the functional model and correlation between the GPS observables mean that determination of the variance-covariance matrix (and hence the weight matrix) is difficult (Barnes et al., 1998, Wang, 1999, Keenan and Cross, 2001). It is common practice (Hoffman-Wellenhof et al., 1994, Strang and Borre, 1997) to assume that the observations from every satellite have equal variance and are uncorrelated. I.e.

\[
E[e_i(t)] = 0 \\
E[e_i(t)e_j(t)] = \begin{cases} 0 & (i \neq j) \\ \end{cases} \\
E[e_i(t)e_i(t)] = \begin{cases} 0 & (t \neq v) \\ \end{cases} \\
V[e_i(t)] = E[e_i(t)e_i(t)] = \sigma^2
\]

(11a) \hspace{1cm} (11b) \hspace{1cm} (11c) \hspace{1cm} (11d)

The stochastic model for the one-way observations is therefore $C_1 = \sigma^2 I$. The mathematical correlations that arise from the formation of the double difference observations can be modelled using $C_\Pi = D^T C_1 D$ where $C_\Pi$ is the variance-covariance matrix of the double differenced observations and $D$ is the differencing matrix from Equation (10). Such an approach ignores differences in the noise characteristics that result from signal propagation and radio frequency (RF) interference and differences in the satellite antennas. It also ignores the correlations between the measurements that were mentioned previously.

Some software uses a slightly more sophisticated process that calculates individual weights for each satellite based on its elevation, signal to noise ratio (SNR) or the baseline length, but which also ignore correlations between the measurements (Barnes et al., 1998, Bock et al., 1986). Even high-end software uses these stochastic models, though they may also have the option to estimate a stochastic model by post-fitting the residuals from the processing (MIT and Scripps, 2000).

Whilst satellite elevation and SNR can be used as quality indicators, they do not always reflect reality because of signal diffraction and other influences (Satirapod and Wang, 2000, Brunner et al., 1999, Gerdan, 1995). Barnes et al. (1998) demonstrates the inadequacies of these approaches
and the increased precision and fidelity obtained by using the correct stochastic model which they calculate using a reverse engineering approach. In general, an incorrect stochastic model will result in unreliable results, over-optimistic measures of precision and a reduced ability to detect errors (Nicolai, 1988). Improving the stochastic model also has benefits in faster and more reliable integer ambiguity determination (Euler, 2000, Han, 1997).

DEVELOPMENTS IN STOCHASTIC MODELLING

A number of techniques are available to estimate variance-covariance information of geodetic data, including Minimum Norm Quadratic Unbiased Estimation (MINQUE) developed by Rao (1971) and the Generalized Maximum-Likelihood Estimation (GML) in Grodecki (2001). Unfortunately, these approaches are not ideal because they assume the covariance matrix is block diagonal, thereby making it difficult if not impossible to estimate temporal correlations using these methods. Wang (1998) uses the MINQUE algorithm and adapts the double-difference functional model to include time correlation. An initial solution for the parameters is made using the standard double-difference mathematical and stochastic models, from which the correlation coefficients are estimated. Results presented by Wang (1998) show a considerable improvement of this method over normal processing algorithms. However, this methodology assumes the baselines are short and is computationally intensive.

A generalised yet representative stochastic model is yet to be developed. One factor inhibiting the generation of such a model is multipath. Phase multipath on L₁ can be as high as 4.7cm, which is well above the noise level (Wanninger and May, 2000). Calculation of the multipath error requires that the position, orientation and reflectance of all reflective surfaces in the vicinity of the antenna be known. Tests by Mora-Castro et al. (1998) show that, if this information is available, multipath can be modelled quite well. Providing there is no change in the reflectance of the reflecting surface (as might be caused by rain or snow) the multipath at a fixed site will repeat each 23hr 56min when the configuration of the satellite constellation is the same. Radovanovic (2000) shows that multipath correlation between consecutive days is highest within 12 seconds of 23hr 56min. Modelling of multipath by its day to day repeatability can show improvements of 25% to 40% but require well known receiver coordinates, a fixed antenna and stable reflectors and reflectances (Radovanovic, 2000).

The authors plan to analyse large amounts of data from the Victorian GPSnet reference network (GPSnet) in order to try and establish trends in the residuals and improve understanding of receiver behaviour. If the stochastic model is correct and systematic errors (such as multipath) have been removed the residuals may be expected to have a constant (zero) mean and variance. If not a modification is required to the stochastic model, perhaps requiring the addition of one or more variables. Due to modelling restrictions the one-way observations are unsuitable for analysis at the noise level, while the double difference observations have a more complicated structure with the noise being a function of eight individual measurements. Analysis of zero-baseline data (two receivers connected to a single antenna) shows that, in the absence of residual systematic errors, the double difference measurements are uncorrelated and Gaussian (white-noise) in nature. The variances vary widely between different satellite pairs showing elevation dependence. Plotting the empirical covariance function of non-zero baselines shows that there is a strong time correlation in the double differences. The aim is to separate out each error source,
starting with multipath, from the residuals leaving the random error, which can then be modelled. From the empirical covariance functions it is possible to fit a covariance model for the systematic and random parts of the error (the signal and the noise). Least squares collocation, which includes a stochastic signal in the parameter estimation process, can then be used to provide more precise estimates of the parameters of interest. The difficulty lies in determining the source and nature of the correlations in the data. This means breaking down the resultant error into its constituent parts, and is the focus of ongoing research.

CONCLUSION

Examination of current techniques used to handle the numerous error sources in GPS has shown that most errors can be effectively removed given the right circumstances (equipment, baseline length, time etc). However, some systematic errors are still problematic, especially multipath. Additionally current stochastic modelling is not necessarily able to model the noise and correlations in the GPS measurements. Continuing research into stochastic modelling is expected to result in higher fidelity quality estimates of GPS derived positions than are produced by current commercial software packages. Improved quality estimation will have follow-on effects for data utility and use through SDI by allowing for more sound assessment of the suitability of the data and decisions made using the data.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the support of Land Victoria and the Centre for Spatial Data Infrastructures and Land Administration, The University of Melbourne in preparing this paper.

REFERENCES


Nicolai, I.R., (1990), Some practical examples of the user of standardised statistical testing of positioning for seismic and engineering surveys, *FIG XIX International Congress*, Helsinki, Finland, 17pp.


**BIOGRAPHICAL NOTES**

Neil Brown
Neil graduated from the University of Melbourne in 1999 with a Bachelor of Arts and a Bachelor of Geomatics (Honours). He is currently undertaking a Doctor of Philosophy degree in the Department of Geomatics, University of Melbourne on stochastic modelling of the GPS observables.

**Allison Kealy**

Allison Kealy is currently a lecturer in the Department of Geomatics at the University of Melbourne, specialising in the research areas of GPS, GLONASS and integrated systems. Allison received her PhD in Geodesy from the University of Newcastle upon Tyne, UK in 1996, after which she spent 2 years in industry providing technical support for GPS/GLONASS manufacturers Ashtech Ltd.

**Professor Ian Williamson**

PhD (NSW), DrHC (Olsztyn), FTSE, FIS Aust, FIE Aust, HonM FIG, HonF MSIAust, LS, CPEng

Ian Williamson is Professor of Surveying and Land Information, Head, Department of Geomatics and Director, Centre for Spatial Data Infrastructures and Land Administration at the University of Melbourne. His teaching and research is concerned with designing, building and managing land administration, cadastral, and land and geographic information systems in both developed and developing countries. He is a Licensed Land Surveyor and Chartered Professional Engineer.

He has undertaken research and consultancies worldwide including for Australian governments, AusAID, many individual country governments, the United Nations and the World Bank. He was Chairperson of Commission 7 (Cadastre and Land Management) of the International Federation of Surveyors 1994-98, and is currently Director, FIG/UN Liaison 1998-2002. He is an Honorary Member of the FIG. At the University of Melbourne he has been President of the Academic Board and Pro-Vice-Chancellor. He is currently Chairperson of the Victorian Government’s Geospatial Information Reference Group and Chairperson of Working Group 3 (Cadastre) of the United Nations sponsored Permanent Committee for GIS Infrastructure for Asia and the Pacific (2001-2004).