A New Stochastic Modelling Procedure for Precise Static GPS Positioning

Chalermchon Satirapod, Jinling Wang and Chris Rizos
School of Geomatic Engineering
The University of New South Wales
Sydney Australia

ABSTRACT

Carrier phase measurements are used for all high accuracy static GPS positioning applications. It is well known that there are two important aspects to the optimal processing of GPS measurements: the definition of the functional model, and the definition of the associated stochastic model. In most cases the functional model is sufficiently well known, while the definition of the stochastic model still remains a challenging research topic. In this paper, a new stochastic modelling procedure is developed and employed in the processing of GPS data for static positioning applications. This procedure takes into account the heteroscedastic, space- and time-correlated error structure of the GPS measurements. To demonstrate its performance, both simulated and real data sets for short to medium length baselines have been analysed. The results indicate that the accuracy of GPS positioning results can be improved to the millimetre level.

INTRODUCTION

GPS has become an important tool for high accuracy relative positioning since it was introduced in the early 1980s. In order to ensure high accuracy, both the functional model and the stochastic model must be correctly defined. Since GPS measurements are contaminated by many errors, it is impossible to model all systematic errors in the functional model without some understanding, or prior knowledge, of the physical phenomena which underpin these errors. Consequently, the residuals obtained from a least-squares solution would normally represent both unmodelled systematic errors and noise. In principle it is possible to further improve the accuracy and reliability of GPS results through an enhancement of either (or both) the functional and stochastic models. This paper deals only with the enhancement of the stochastic model for the static GPS positioning case.

Previous studies have shown that GPS measurements have a heteroscedastic, space- and time-correlated error structure (e.g., Wang 1998; Wang et al., 1998a). The challenge is to find a way to realistically incorporate such information into the stochastic model. Several stochastic modelling procedures have recently been proposed to accommodate the heteroscedastic behaviour of GPS observations. A review of these procedures can be found in, e.g., Satirapod & Wang (2000). However, these procedures either do not take into account temporal correlations, or assume that all one-way GPS measurements have the same temporal correlation coefficients. It has been shown that different satellites have different temporal correlation coefficients (Wang, 1998), hence it is not appropriate to make such assumptions. An iterative stochastic assessment procedure, which takes into account all of the error features, has been proposed by Wang et al. (2001).

The basic idea behind the iterative stochastic modelling procedure is that the double-differenced (DD) carrier phase measurements are transformed into a set of new measurements using estimated temporal correlation coefficients. The transformed measurements are free of temporal correlations and thus have a block diagonal variance-covariance matrix (Wang, 1998). Consequently, the immense memory usage and computational load for the inversion of a fully populated variance-covariance matrix can be avoided, and the variance-covariance matrix for the transformed measurements can be estimated using a rigorous statistical method such as the MINQUE method (Rao, 1971). An iterative process is performed until sufficient accuracy is achieved. This procedure is suitable for short observation periods as it assumes that the temporal correlation coefficients and the variance of GPS measurements are constant for the whole observation period. Initial experiments based on this procedure have demonstrated encouraging results in the case of short observation periods and for short baselines.

However, when the iterative stochastic modelling procedure is applied to long observation period data sets, several shortcomings of the procedure needed to be addressed. For example, the assumption that the temporal correlation coefficients and the variance of GPS measurements are constant for the whole observation period is not realistic. Furthermore, in practice, an observation period of several hours may be expected for some geodetic applications. Thus the memory usage and computational load can become unbearable when the standard MINQUE technique is applied. Hence, it is necessary to develop a new stochastic modelling procedure that addresses these shortcomings.

SEGMENTED STOCHASTIC MODELLING PROCEDURE

To process long observation session data sets, a three-step procedure has been proposed to realistically estimate the stochastic models for the GPS measurements. The first step is to divide the whole session into short segments. Secondly, the inter-temporal correlation coefficients (for different satellite pairs) should be set to zero. In the third step, the MINQUE procedure for the estimation of variance-covariance components is replaced by an alternative method.

Step1: Data segmentation

Given that GPS measurements are contaminated by errors whose characteristics change slowly with time, it is appropriate to divide the whole measurement session into short segments, in which each segment has the same number of satellites and all the measurements for the same satellite pairs have an invariant stochastic model. This is illustrated in Figure 1. During a long observation period the satellite geometry changes considerably, hence the use of fixed 'window' widths in order to segment the measurements is not advisable.

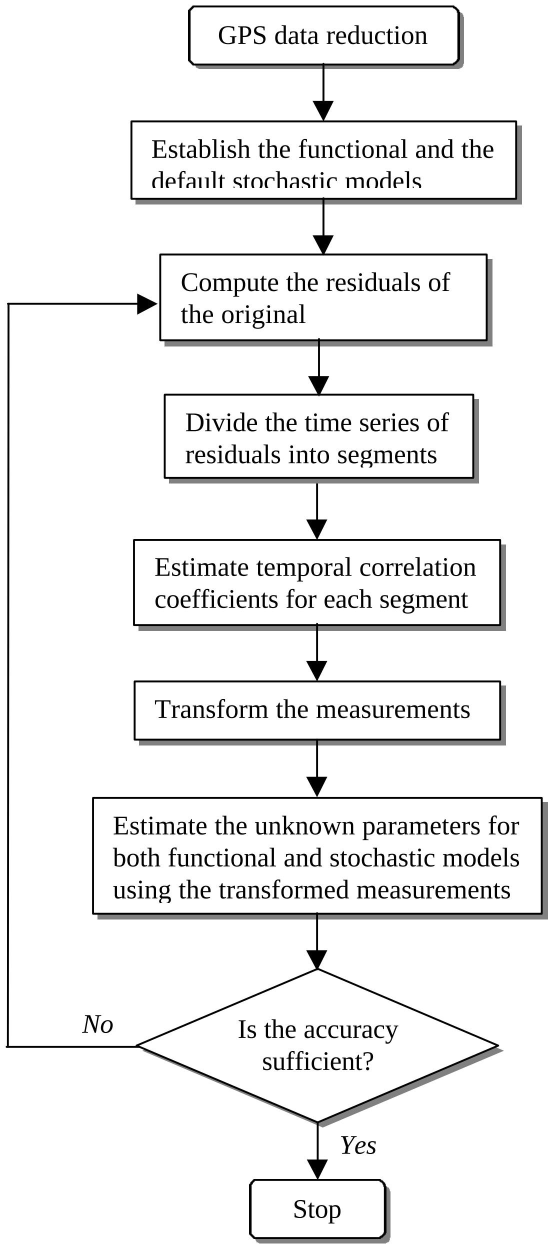

Figure 1. Flowchart of the segmented stochastic modelling procedure.

A method to circumvent this problem is proposed, in which a default window width is first selected. Then indices of when the satellite geometry has changed during the entire session are determined, and the number of observations between consecutive indices checked. If the number of observations between any pair of consecutive indices is larger than the default window width, the measurements are divided into short segments until the number of observations in the last segment is smaller than or equal to the default window width. In this case, the observations from the last segment will be combined with the ones from the previous segment. However, if the number of observations between the consecutive indices is not sufficient to form a new segment, the stochastic model estimated from the previous segment is applied to these observations.

Step2: Estimation of temporal correlation coefficients

Assuming that the inter-temporal correlations are zero, the error specification is:

$$\begin{bmatrix} e_1(t) \\ e_2(t) \\ . \\ . \\ e_n(t) \end{bmatrix} = \begin{bmatrix} r_{11} & 0 & . & . & 0 \\ 0 & r_{22} & . & . & 0 \\ . & . & & & . \\ . & . & & & . \\ 0 & 0 & . & . & r_{nn} \end{bmatrix} \cdot \begin{bmatrix} e_1(t-1) \\ e_2(t-1) \\ . \\ . \\ e_n(t-1) \end{bmatrix} + \begin{bmatrix} u_1(t) \\ u_2(t) \\ . \\ . \\ u_n(t) \end{bmatrix}$$
(1)

where r_{ii} is the temporal correlation within the i^{th} satellite pair; n is the number of satellite pairs forming the DD measurements; and e is the vector of original residuals.

According to the Durbin-Watson statistic (Durbin & Watson, 1950), the temporal correlation coefficient can be calculated by:

$$r = 1 - d/2 \tag{2}$$

with

$$d = \frac{\sum_{k=2}^{m}(e_k - e_{k-1})^2}{\sum_{k=1}^{m} e_k^2} \tag{3}$$

where m is the number of observation epochs in a segment.

The estimated temporal correlation coefficient is then used to transform the original measurements. Details of this transformation step can be found in Wang et al. (2001).

Step3: Estimation of variance-covariance components

Method 1--The standard MINQUE method
In the following Gauss-Markov model with n measurements and t unknowns:

$$l = Ax + v \tag{4}$$

$$C = P^{-1} = \sum_{i=1}^{k} q_i T_i \tag{5}$$

l and v are $n\times1$ vectors of the transformed measurements and residuals respectively; A is the $n\times t$ design matrix; x is the $t\times1$ vector of the unknown parameters; $q_1, q_2, \ldots, q_k$ are the variance-covariance components of the measurements to be estimated; k is the number of variance-covariance components; and $T_1, T_2, \ldots, T_k$ are the accompanying matrices (see, e.g., Wang et al., 1998a). P is the weight matrix of the observations. According to Rao (1971, 1979), the variance-covariance components can be estimated as:

$$\hat{q} = (\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_k)^T = S^{-1} q \tag{6}$$

where the matrix $S = \{s_{ij}\}$ with

$$s_{ij} = Tr\{RT_iRT_j\} \quad (7)$$

and the vector $q = \{q_i\}$ with

$$q_i = v^TRT_iRv \quad (8)$$

and

$$R = PQ_vP \quad (9)$$

with $Q_v = [P^{-1} - A(A^TPA)^{-1}A^T]$ being the adjusted residuals cofactor matrix.

Since the estimated variance-covariance components depend on matrix C, which includes the variance-covariance components themselves, an iterative process must be performed. It should be noted that when the MINQUE method is used in GPS data processing, the storage of the matrix R may require a huge memory. In addition, the computation relating to this matrix operation is extensive.

Method 2--The simplified MINQUE method
A simplification of the MINQUE procedure can be obtained by replacing the matrix R in Equation (9) by a diagonal matrix R^*. The matrix R^* has a block-diagonal structure and can be defined as:

$$R^* = diag(R_{kk}) , (k = 1,2,...,m) \quad (10)$$

where m is the number of the observation epoch in a segment.

Therefore, Equations (7) and (8) can be simplified to:

$$s_{ij} = Tr\{RT_iRT_j\} = \sum_{k=1}^{m} Tr(R_{kk}T_{ik}R_{kk}T_{jk}) \quad (11)$$

$$q_i = v^TRT_iRv = \sum_{k=1}^{m} Tr(v_k^T R_{kk}T_{ik}R_{kk}v_k) \quad (12)$$

This simplified procedure has been shown to produce results that are close in quality to that of the rigorous procedure. In addition, the computational time and the memory requirements of the simplified procedure are much less than in the case of the rigorous procedure.

Method 3--The proposed method
The estimation of variance-covariance components can be performed using the classical definition of the variance-covariance matrix:

$$C = E[(v-m)(v-m)^T] = E[vv^T] - mm^T \quad (13)$$

where m is the mean value. The residuals obtained from the transformed measurements are random and have zero mean (m=0). Hence, the variance-covariance matrix can be obtained by averaging the residuals within the same segment:

$$C = \frac{1}{m}\sum_{k=1}^{m} v_k v_k^T \tag{14}$$

An iterative estimation procedure is required as the estimation of the residuals is dependent on the variance-covariance matrix. Based on Equation (14) the implementation of the proposed method is relatively simple and straightforward. Its performance has been evaluated using three data sets. For ambiguity discrimination, the difference between the best and second best ambiguity combination is crucial. The statistics F-ratio (e.g., Euler & Landau, 1992) and the W-ratio (Wang et al., 1998b) are chosen for comparison (see Table 1).

Table 1. Comparison of F-ratio and W-ratio statistics.

Baseline	Method	F-ratio	W-ratio
14m	1	12.76	26.65
	2	12.77	26.67
	3	12.80	26.72
1km	1	9.20	29.92
	2	9.16	29.87
	3	9.17	29.86
2km	1	4.54	16.12
	2	4.59	16.25
	3	4.57	16.19

Remark: The observation period used here is 15 minutes and the sampling rate is 15 seconds.

From Table 1, there is no significant difference in the F-ratio and W-ratio statistics obtained from the three methods. In terms of the estimated baseline components, the results obtained from the three methods are also essentially identical. A comparison of the computational time and memory usage is presented in Table 2.

Table 2. Comparison of computational time and memory usage.

Baseline	Method	Computational time (s)	Memory usage (Kbytes)
14m	1	263.4	450.0
(5 sats)	2	7.6	7.5
	3	1.2	*None*
1km	1	1519.0	1012.5
(7 sats)	2	33.2	16.9
	3	1.9	*None*
2km	1	700.6	703.1
(6 sats)	2	18.6	11.7
	3	1.5	*None*

Remark: All solutions computed using Matlab GPS baseline software on a PentiumII-366MHz processor.

Interestingly, if the matrix R^* in the simplified MINQUE method is assumed to be constant, the variance-covariance components obtained from the proposed method are quite similar to those obtained from the simplified method. The assumption that the matrix R^* is constant is reasonable for short observation periods. This is supported by the study described in Han & Rizos (1995), which established that the maximum difference in the matrix A is about 8% for a 15 minute observation span.

TEST DATA

Simulations

There are two main advantages to using simulated data: (a) to evaluate the performance of the proposed algorithm (since it is very difficult to derive highly accurate GPS station coordinates in practice), and (b) to study the impact of incorporating different systematic errors. The data simulation involves two parts.

Simulating raw GPS observations
A simulation of the raw GPS observations was performed using the Bernese software (version 4.0). Different carrier phase observation noises were assigned to different satellites varying from 1mm to 3mm. Two data sets were simulated, a 0.5hr session for a 9km and a 5hr session for a 79km baseline.

The systematic error components
A wavelet-based technique (e.g., Chui C.K. 1992; Wickerhauser M. 1994) was applied to the GPS DD residuals in order to extract the systematic error component. The GPS DD ambiguity-fixed residuals obtained from other real data sets were decomposed into the low-frequency bias and the high-frequency noise components. The extracted systematic error component was then added to the simulated GPS observations. Two different systematic error patterns, denoted E1 and E2, were extracted and are plotted in Figures 2 and 3.

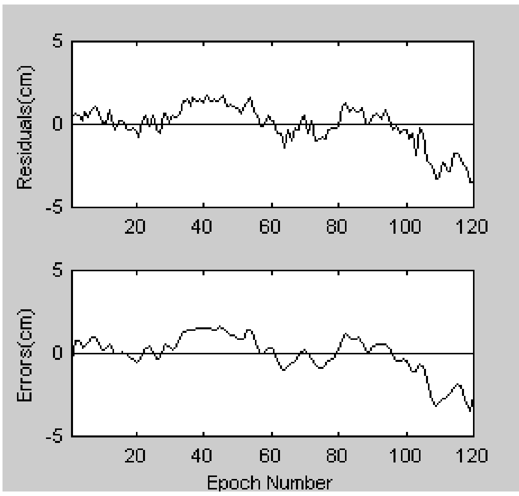

Figure 2. Signal extraction using wavelets. Top: Original DD residuals. Bottom: E1 error pattern.

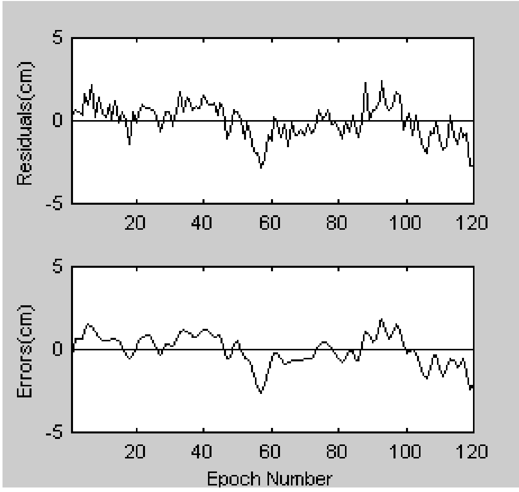

Figure 3. Signal extraction using wavelets. Top: Original DD residuals. Bottom: E2 error pattern.

Real data sets

Two data sets were downloaded from http://sopac.ucsd.edu/, collected by receivers of the Southern California Integrated GPS Network (SCIGN). For test purposes, 3hr and 5hr observation sessions were considered for 23km and 75km baseline data sets respectively.

RESULTS FROM SIMULATED DATA TESTS

In this section, the effectiveness of the proposed stochastic modelling procedure is demonstrated using both short and medium-length baseline data sets. The impact of systematic errors on the GPS positioning results was analysed for two different cases. The first case involved varying the number of satellites but adding the same error pattern to the same satellite pair. In the second case, the two error patterns were added to the satellite pairs alternately (see Tables 4 and 6). In this case, the approach consisted of the following steps: i) adding the error pattern E1 to different satellite pairs and obtaining a solution, ii) adding the error pattern E2 to different satellite pairs and obtaining a solution.

Short baseline

Matlab-coded GPS baseline processing software developed at The University of New South Wales was used to process the data sets. Only single-frequency data was used. To obtain more reliable baseline results for comparison purposes, the data set with observation noise was first processed using the MINQUE procedure and the results were used as reference values. Then, in both cases, the data set with intentionally added systematic errors was processed using the following two procedures:

A. The standard procedure with the simplified stochastic model (which includes only mathematical correlation), and assuming that temporal correlations are absent.
B. The segmented stochastic modelling procedure with a 20-epoch window width.

Case I – Varying the number of satellites

The number of satellites varied from 8 to 5 satellites, and the E1 error pattern was added to satellite pair PRN23-15 for every satellite geometry. The DD residuals for satellite pair PRN

23-15 are shown in Figure 4. The red line denotes the post-fit residuals obtained using method A, while the black line shows the residuals obtained from method B. It can be clearly seen that the systematic errors of the transformed measurements are much smaller than those of the original measurements. Similar results from real data sets, for the short baseline case, were reported in Wang et al. (2001).

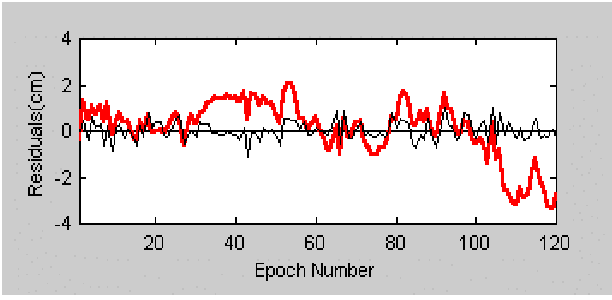

Figure 4. DD residuals obtained from the 9km baseline for satellite pair PRN 23-15.

The estimated baseline components obtained from both methods were compared with the reference values (Table 3). It can be seen that in all cases the proposed algorithm produced more reliable results than the standard procedure.

Table 3. Differences between estimated and reference coordinate values (simulated short baseline data -- Case I).

Satellites used	Method	Difference in each component		
		dN (mm)	dE(mm)	dH(mm)
8 sats	A	0.4	0.2	2.1
	B	0.0	0.0	0.0
7 sats	A	0.0	0.1	1.7
	B	0.0	0.0	0.0
6 sats	A	1.2	0.8	3.4
	B	0.0	0.0	0.0
5 sats	A	3.2	4.0	10.7
	B	0.0	0.0	0.0

Remark: E1 error pattern added to PRN23-15.

In addition, the impact of systematic errors on the positioning results tends to increase with a decrease in the number of satellites used in method A. In the case of 5 satellites, the difference in the height component is as large as 10.7mm.

Case II – Varying the error patterns and satellite pairs

The E1 and E2 error patterns were added to different satellite pairs for the geometry consisting of 5 satellites only. The DD residuals showed similar trends to those obtained from Case I. The differences in each coordinate component are shown in Table 4. It is evident that different error patterns and different satellites have a different influence on the positioning results from the use of method A, but not method B.

Table 4. Differences between estimated and reference coordinate values (simulated short baseline data -- Case II).

Error pattern /sat pair	Method	Difference in each component		
		dN (mm)	dE(mm)	dH(mm)
E1/23-15	A	3.2	4.7	10.7
	B	0.0	0.0	0.0
E1/23-03	A	3.4	4.0	10.9
	B	0.0	0.0	0.0
E1/23-31	A	0.3	0.1	0.3
	B	0.0	0.0	0.0
E2/23-15	A	0.4	2.0	5.4
	B	0.0	0.0	0.0
E2/23-03	A	1.7	2.2	5.7
	B	0.0	0.0	0.0
E2/23-31	A	0.2	0.0	0.2
	B	0.0	0.0	0.0

Medium length baseline

It is a common practice to process dual-frequency data in cases of medium length baselines. The data processing consisted of three steps. The first and second steps were carried out in the Bernese software. The DD ambiguities were solved for in the first program run. These ambiguities were then introduced as fixed values in the next program run. In this step, some information were output for further use with the Matlab-coded GPS processing software. In the first and second program runs standard parameters such as the Saastamoinen tropospheric model and the IGS precise orbit were used. In the third program run, the output information was processed using the proposed procedure. Since the dual-frequency data was used, the initial coordinates applied to simulate the GPS observations were used as true values for comparison purposes. Similar to the short baseline case, in both cases the data set containing systematic errors was processed using methods A and B.

Case I – Varying the number of satellites

The number of satellites varied from 14 to 11 satellites and the E1 error pattern was added to the satellite pair PRN25-1 for every satellite geometry. The DD residuals for the satellite pair PRN25-1 are plotted in Figure 5. The red line denotes the post-fit residuals obtained from method A, while the black line shows the residuals obtained from method B. It is clear that the post-fit residuals obtained from method B are more random than those obtained from method A.

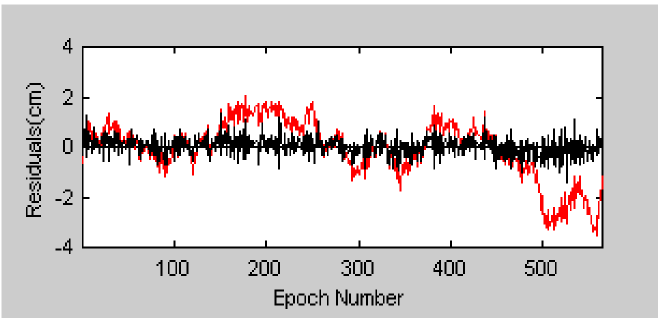

Figure 5. DD residuals obtained from the 79km baseline for satellite pair PRN25-1.

The estimated baseline components obtained from both methods were subsequently compared with the true values (Table 5). The results are similar to the short baseline case, that is, it can be concluded that the error modelling using method B is an improvement over method A.

Table 5. Differences between estimated and reference coordinate values (simulated medium baseline data -- Case I).

Satellites used	Method	Difference in each component		
		dN (mm)	dE(mm)	dH(mm)
14 sats	A	0.1	0.2	1.9
	B	0.0	0.0	0.0
13 sats	A	0.1	0.3	2.4
	B	0.1	0.0	0.1
12 sats	A	0.2	0.2	2.5
	B	0.1	0.0	0.1
11 sats	A	0.1	0.1	2.6
	B	0.1	0.0	0.2

Remark: E1 error pattern added to PRN25-1.

Case II – Varying the error patterns and satellite pairs

Once again the E1 and E2 error patterns were added to different satellite pairs for the geometry consisting of 11 satellites only. The DD residuals showed similar trends (for the sake of brevity, the residuals are not shown here). Table 6 shows the differences between the estimated baseline components and the true values. The results confirm that different error patterns and different satellites have a different influence on the positioning results obtained using method A, but not method B.

Table 6. Differences between estimated and reference coordinate values (simulated medium baseline data -- Case II).

Error pattern /sat pair	Method	Difference in each component		
		dN (mm)	dE(mm)	dH(mm)
E1/25-01	A	0.1	0.1	2.6
	B	0.1	0.0	0.2
E1/25-14	A	2.9	0.1	0.7
	B	0.0	0.0	0.0
E1/25-21	A	0.7	0.1	2.7
	B	0.1	0.0	0.3
E2/25-01	A	0.1	0.0	1.1
	B	0.1	0.1	0.2
E2/25-14	A	1.3	0.1	1.2
	B	0.1	0.0	0.1
E2/25-21	A	0.5	0.0	1.5
	B	0.1	0.0	0.2

RESULTS FROM REAL DATA SETS

The data processing strategies used here are the same as in the medium-length baseline case, except that systematic errors have not been added to the observations. The DD residuals obtained from all satellite pairs show similar trends. Selected residuals obtained from the 23km and 75km baselines are plotted in Figures 6 and 7 respectively. Again, the red line denotes the post-fit residuals obtained from using method A, while the black line shows the residuals obtained from using method B. Clearly method B generates random residuals for both data sets.

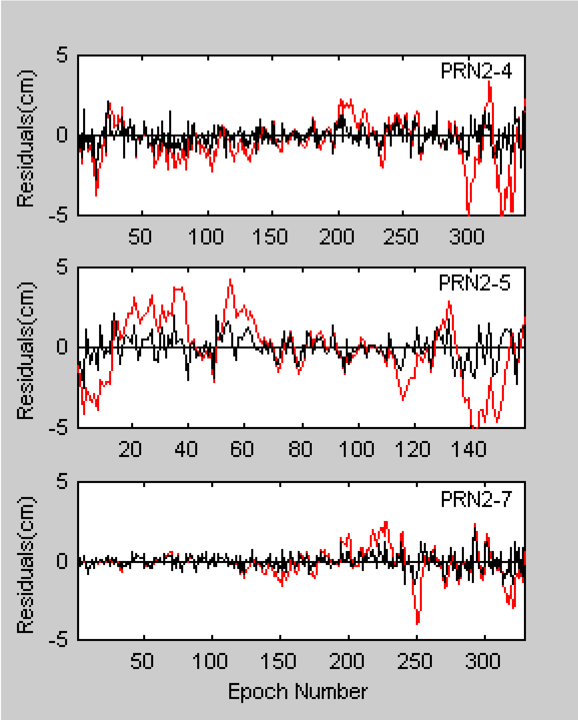

Figure 6. Selected DD residuals obtained from a 23km baseline for several satellite pairs.

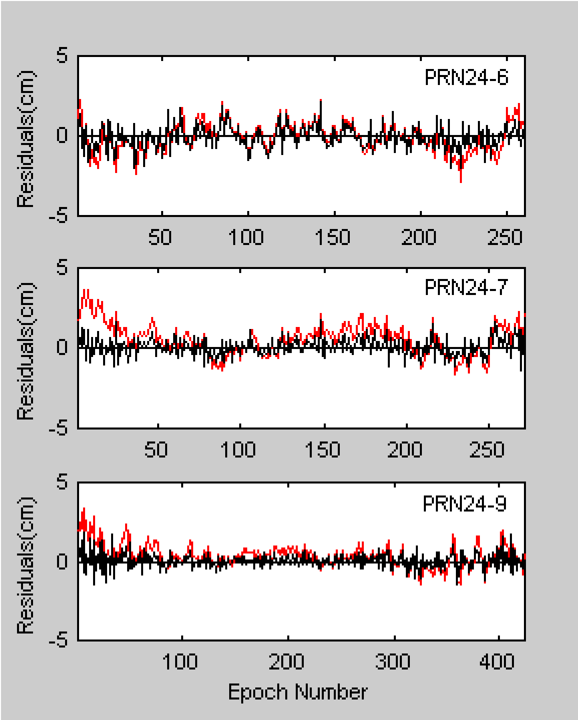

Figure 7. Selected DD residuals obtained from a 75km baseline for several satellite pairs.

CONCLUDING REMARKS

Based on a new framework of error analysis of GPS measurements, a new stochastic modelling procedure which takes into account the temporal correlations in the GPS measurements, has been introduced to effectively deal with long observation periods that are typically processed for high accuracy static positioning applications. It has been shown that the GPS measurements have a heteroscedastic, space- and time-correlated error structure, and that any mis-specification in the stochastic model may have a significant influence on the positioning results. The impact of temporal correlations was analysed using simulated and real data sets. The results indicate that there are significant effects on the positioning results when the temporal correlations are not taken into account in the stochastic model. By applying the proposed segmented stochastic modelling procedure, the residuals are more random and the accuracy of the estimated baseline components is improved to the millimetre level.

ACKNOWLEDGMENTS

The first author is sponsored in his Ph.D. studies by a scholarship from the Chulalongkorn University, Thailand. Much gratitude to Clement Ogaja for the fruitful discussions he has with the authors. The authors' warmest thanks to Pierre Fridez, of the Astronomical Institute of the University of Berne, who assisted with understanding the source code of the Bernese software. The Southern California Integrated GPS Network is also acknowledged for providing some of the data used in this study.

REFERENCES

Chui C.K. (1992) *An Introduction to Wavelets*, Academic Press, Inc., Boston, 264pp.

Durbin J. & Watson G.S. (1950) Testing for serial correlation in least squares regression I, *Biometrika*, 37, 409-428.

Euler H.J. & Landau H. (1992) Fast GPS ambiguity resolution on-the-fly for real-time applications, *6th International Symposium on Satellite Positioning*, Columbus, Ohio, 17-20 March, 650-659.

Han S. & Rizos C. (1995) Standarization of the variance-covariance matrix for GPS rapid static positioning, *Geomatics Research Australasia*, 62, 37-54.

Rao C.R. (1971) Estimation of variance and covariance components - MINQUE, *Journal of Multivariate Analysis*, 1, 257-275.

Rao C.R. (1979) MINQUE theory and its relation to ML and MML estimation of variance components, *Sankhya*, 41, Series B, 138-153.

Satirapod C. & Wang J. (2000) Comparing the quality indicators of GPS carrier phase observations, *Geomatics Research Australasia*, 73, 75-92.

Wang J. (1998) Stochastic assessment of the GPS measurements for precise positioning, *11th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation*, Nashville, Tennessee, 15-18 September, 81-89.

Wang J., Stewart M.P. & Tsakiri M. (1998a) Stochastic modelling for static GPS baseline data processing, *Journal of Surveying Engineering*, 124(4), 171-181.

Wang J., Stewart M.P. & Tsakiri M. (1998b) A discrimination test procedure for ambiguity resolution on-the-fly, *Journal of Geodesy*, 72, 644-653.

Wang J., Satirapod C. & Rizos C. (2001) Stochastic assessment of GPS carrier phase measurements for precise static relative positioning, *Submitted to Journal of Geodesy* (http://www.gmat.unsw.edu.au/snap/publications/wang_etal2001a.pdf).

Wickerhauser M. (1994) *Adapted Wavelet Analysis from Theory to Software*, AK Peters Ltd., Wellesley, Massachusetts, 486pp.

BIOGRAPHY

Chalermchon Satirapod is currently a Ph.D. student at the School of Geomatic Engineering, The University of New South Wales (UNSW), supported by a scholarship from the Chulalongkorn University. He graduated with a Bachelor of Engineering (Surveying) and Master of Engineering (Surveying) from Chulalongkorn University, Thailand, in 1994 and 1997 respectively.

Jinling Wang holds a Ph.D. from the Curtin University, Australia. He is currently an Australian Research Council Postdoctoral Fellow in the School of Geomatic Engineering, UNSW. He is a member of the Editorial Advisory Board of the journal GPS Solutions.

Chris Rizos, B.Surv. (UNSW) Ph.D. (UNSW) is professor and leader of the Satellite Navigation and Positioning (SNAP) Group at The University of New South Wales. He is Secretary of Section 1, "Positioning", of the International Association of Geodesy, and a Fellow of the Australian Institute of Navigation.