

An Empirical Approach for the Estimation of GPS Covariance Matrix of Observations

Rodrigo Leandro, University of New Brunswick
Marcelo Santos, University of New Brunswick
Karen Cove, University of New Brunswick

BIOGRAPHY

Rodrigo Leandro is Ph.D. candidate at the Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada under supervision of Dr. Marcelo Santos. He holds a M. Sc. Eng. Degree in Civil Engineering from the University of São Paulo, in São Paulo, Brazil. He has been involved in research in the fields of Geodesy and satellite navigation.

Marcelo Santos is Associate Professor at the Department of Geodesy and Geomatics Engineering at UNB. He holds a M.Sc. in Geophysics from the National Observatory, in Rio de Janeiro, and a Ph.D. in Geodesy from UNB. He has been involved in research in the fields of space and physical geodesy, GNSS and navigation

Karen Cove received her B.Sc. in Geomatics Engineering in 2002 and will be graduating in October 2005 with a M.Sc. from the Dept. of Geodesy and Geomatics Engineering at the University of New Brunswick, Canada. She is currently employed at CARIS as a Project Engineer. Her research interests include hydrography, GPS positioning, and marine boundaries.

ABSTRACT

The covariance matrix of observations plays an important role in GPS data processing. For example, the weighting of the observations are based on the covariance matrix of the observations. The covariance matrix of the observations also plays a role in the estimation of the covariance matrix of the parameters. In this paper we present an empirical stochastic approach to create covariance matrices for GPS. Our approach aims at a more realistic and complete information about the stochastic behaviour of GPS observations. As a result, improvement in quality and quality control of estimated coordinates using GPS measurements is obtained.

In the present case study, baselines are processed using pseudorange. The observations are considered as a time series after double differencing. The advantages and

disadvantages of this empirical stochastic approach are explored, comparing it with usual methods for observations covariance matrix estimation. An analysis is made in terms of obtained coordinates where baselines between points with known coordinates were processed.

Two baselines with length of approximately 2km and 20km were processed using three different weighting schemes for observations. In general, when the empirical approach was used a good improvement in the bias was obtained. The mean biases were at least 44% smaller in the first case and at least 21 % in the second one. The model certainly brings improvement to the solution, with better results achieved for height determination, with improvements up to 72% in mean bias for the shorter baseline.

INTRODUCTION

The covariance matrix of observations plays a fundamental role in GPS data adjustment. Usually the method of least squares is used to compute the coordinates of a given receiver. In the adjustment the covariance matrix of observations drives how each one of the observations will contribute for the update of the parameters. This matrix includes not just variances, but also a relation between all observations, the covariances. An usual approach to estimate the covariance matrix of observations is to set a weight for all observations and then, in case of baseline, propagate it using the double difference operator. It is not assumed any correlation between observations made in two different epochs, the so called autocorrelations. The weights can be set based on different types of information, such as elevation angle of the satellite or signal-to-noise ratio. Sometimes even the identity matrix (equally weighted observations) is used.

In this work, an approach to build the covariance matrix of observations for GPS data processing is presented. The main objective of such technique is to populate the covariance matrix with realistic information, estimated by means of a stochastic analysis of the raw data. Trying to

have as much as possible information inside the covariance matrix is a way to carry into the adjustment model a realistic picture of the quality and behavior of the observations. Eventually, the procured parameters will be adjusted using this matrix. The ultimate goal of this work is to improve the quality of GPS data processing.

In order to analyze the data assuming a stochastic behavior a few requirements need to be satisfied, such as stationarity. Because of that, the GPS data need to undergo some modifications before it can be used. In our Empirical Stochastic model (which will be called herein with ESto model) all analysis is made based on raw data, without any external information, before the adjustment.

In next section a background explanation about stochastic processes and their analysis will be made. Next the treatment given to GPS data in order to make its analysis as a stochastic process possible will be shown. A case study was carried out and its results are shown, as well as conclusions and recommendations for future work.

STOCHASTIC PROCESSES

A time series is a series of observations made through time. An important feature of a series is that, usually, the observations made at subsequent epochs are dependent on each other. The analysis of a time series is based on this dependency. In order to carry out the analysis, generally stochastic models are used.

The utilization of mathematical models to describe physical phenomena is very common. Sometimes a model based on physical laws can be capable of determining the value of a given variable for a specified time. This is possible when the variable (in our case the GPS observable) is totally deterministic.

A stochastic process is a statistical phenomena which occurs through time according to probability laws, [Box et. al., 1994]. A time series can be considered as a realization of a stochastic process.

There is a particular process, called a stationary process, which has a particular statistic equilibrium state. A stochastic process is called widely stationary when its properties remain unaffected when the time origin is changed. This means that the joint probability function of a process with n observations $z_{t_1}, z_{t_2}, \dots, z_{t_n}$, observed at time instants t_1, t_2, \dots, t_n is the same as a process with n observations $z_{t_1+k}, z_{t_2+k}, \dots, z_{t_n+k}$, observed at time instants $t_1+k, t_2+k, \dots, t_n+k$ for every integer k .

If a process is stationary, we can say that:

$$\rho(z_t) = \rho(z_{t+k}) \quad \forall k \geq 0, t \geq 0, \quad (1)$$

where $\rho(z_t)$ is the probability density function, and because of that, we can conclude that the mean of a stationary process is constant:

$$\mu_t = E(z_t) = \int_{-\infty}^{\infty} z \cdot \rho(z) dz, \quad (2)$$

and therefore:

$$\mu_t = \mu_{t+k}, \quad (3)$$

for every integer k . The mean of a process can be estimated with the equation:

$$\bar{z} = \frac{1}{n} \cdot \sum_{t=1}^n z_t. \quad (4)$$

And the variance of a process can be defined as:

$$s^2 = E[(z_t - \bar{z})^2] = \int_{-\infty}^{\infty} (z - \bar{z})^2 \rho(z) dz. \quad (5)$$

The variance of a process can be estimated with the following equation:

$$s^2 = \frac{1}{n} \cdot \sum_{t=1}^n (z_t - \bar{z})^2. \quad (6)$$

Based on this equations and the assumption behind them, the variances for GPS observables are computed in the ESto model.

The stationarity of a process implies that the joint probability function $\rho(z_{t_1}, z_{t_2})$ is the same for any instants t_1 and t_2 , and for the same time interval between them. This means that we can estimate the joint probability function of a process for different time intervals k . The covariance between the values z_t and z_{t+k} , separated by a time interval k , which is constant for every t for a stationary process is called the auto-covariance function, and can be defined by:

$$\gamma_k = \text{cov}(z_t, z_{t+k}) = E[(z_t - \bar{z})(z_{t+k} - \bar{z})]. \quad (7)$$

It is possible to estimate the auto-correlation function of a process for a given lag k with the following equation:

$$\rho_k = \frac{E[(z_t - \bar{z})(z_{t+k} - \bar{z})]}{\sqrt{E[(z_t - \bar{z})^2] \cdot E[(z_{t+k} - \bar{z})^2]}} = \frac{\sum_{t=1}^{n-k} [(z_t - \bar{z})(z_{t+k} - \bar{z})]}{s^2}. \quad (8)$$

ESTO MODEL

In ESto model, the previously described approaches are used to estimate the correlations between different observations (made at different time and/or from different satellite pairs), as well as the variances for each one of the observed satellites. The input of the model is the raw data, and the output is the stochastic parameters.

As stated in the previous section, a stochastic process needs to have certain characteristics in order to be used for estimation of variances and covariances. Since the objective of this work is to analyze GPS data as a stochastic process, it is necessary to modify the original time series of observations to satisfy the assumptions made in stochastic analysis. The main assumption that will be discussed here is the stationarity of the process. It is clear that GPS observations are not stationary, because they are related to the satellite orbits and always vary with time. Figure 1 shows an example of double differenced pseudoranges through time.

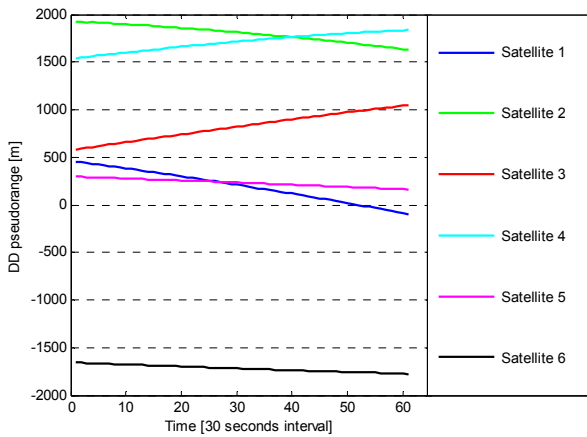


Figure 1. Double difference pseudoranges time series.

Clearly these observations could not be used directly in a stochastic analysis, and therefore need to be modified. In order to get a stationary series from the measurements, they are reduced by using the approximate geometric distances between receiver and satellite antennas, according to:

$$\Delta \nabla P_{red} = \Delta \nabla P_{obs} - \Delta \nabla \rho, \quad (9)$$

where $\Delta \nabla P_{obs}$ is the double differenced observed pseudorange, $\Delta \nabla \rho$ is the double differenced geometric distance between receiver and satellite antennas and $\Delta \nabla P_{red}$ is the double differenced reduced pseudorange. Figure 2 shows the double differenced C/A pseudoranges before and after the reduction made in the ESto model, for each one of the satellites shown above.

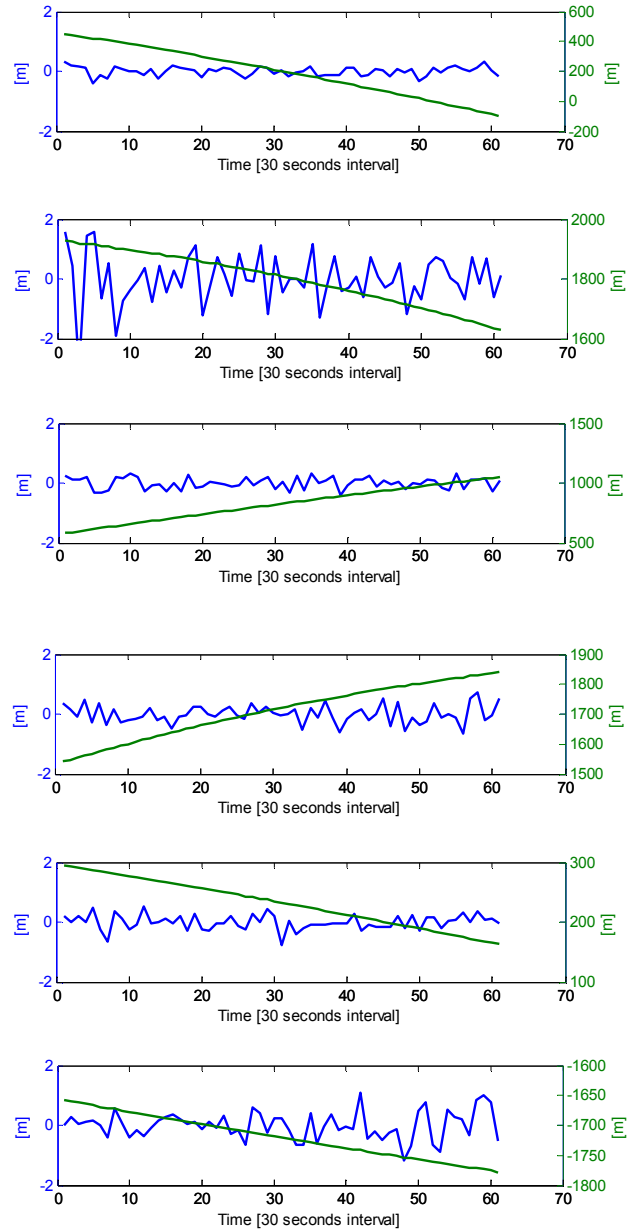


Figure 2. Reduction of DD pseudoranges, in the ESto model.

In the plots above, the green lines represent the original series, and the blue lines represent the reduced series. After this step the reduced series are used in the stochastic analysis. This part of the ESto model is based on the approaches mentioned in the previous section.

As an illustration of the stochastic part of the model, Figure 3 shows the auto-correlation function estimated with ESto model and also computed based on the adjustment residuals, for satellites with different elevation angles (65, 31 and 17 degrees). The observables are C/A pseudoranges, observed for a short baseline (5 km) during

half hour. The sampling interval is 30 seconds, thus one epoch lag represents 30 seconds in time.

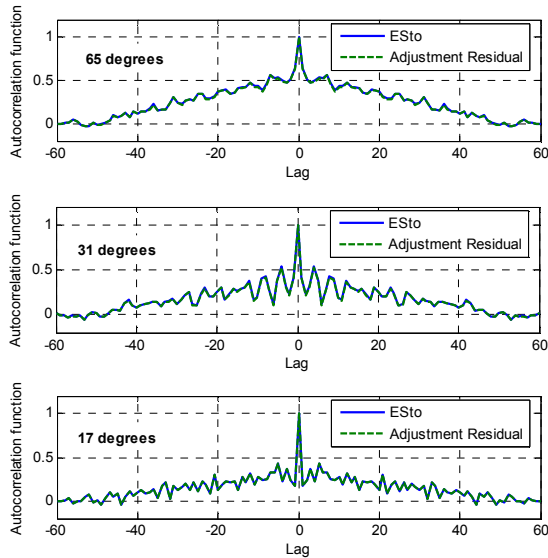


Figure 3. Auto-correlation function for satellites with different elevation angles.

It can be seen that ESto provides a good approximation of the residuals auto-correlation and therefore for the observations. It can be also noticed how the correlations get smaller as the elevation angle gets lower.

Another issue is the variance which is assigned to each satellite. In ESto model, as said previously, it is also computed by analyzing the raw data. Figure 4 shows the pseudorange standard deviations estimated by ESto for different elevation angles.

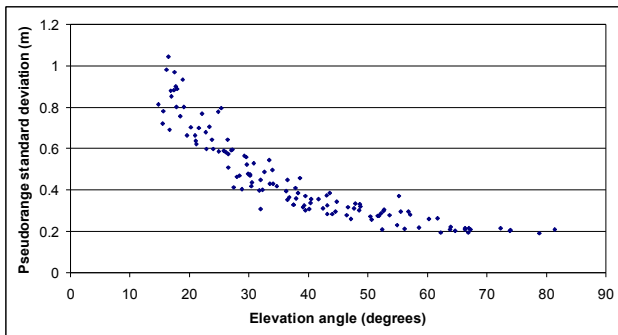


Figure 4. Estimated standard deviation for different elevation angles.

The standard deviation varies with the elevation angle in an exponential like way, between around 0.2 m and 1 m. In Figure 5, it is shown the agreement between the estimated standard deviations and the residuals standard deviation after the adjustment.

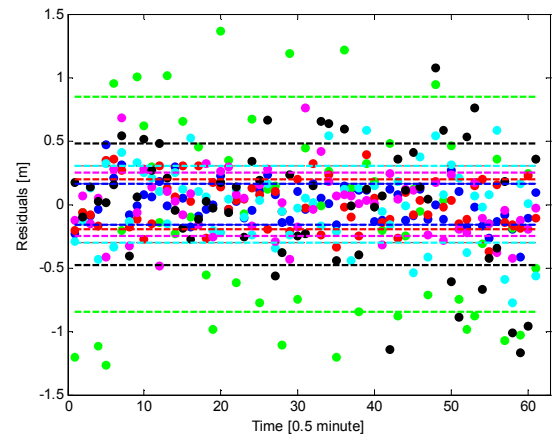


Figure 5. Standard deviations estimated with ESto and adjustment residuals.

In Figure 5, the dots represent the residuals of each of the observations, during a half hour observation session with 30 seconds of sampling interval (7 satellites observed). The dashed lines represent the one sigma standard deviations estimated with ESto. For each satellite, the color of the lines and dots is the same. Table 1 shows the computed values for both cases.

Table 1

Std Deviation	Residuals [m]	ESto [m]
1 (blue) – 65°	0.1588	0.1589
2 (green) – 17°	0.8452	0.8453
3 (red) – 52°	0.1931	0.1931
4 (cyan) – 44°	0.3043	0.3043
5 (magenta) – 45°	0.2516	0.2516
6 (black) – 31°	0.4764	0.4763

ESto model has been developed in order to allow an estimation of a fully populated observation covariance matrices for GPS. In terms of validation, it has been compared with other techniques to build such matrices. The first one, called here Formal DD, is the simplest of the ones explored in this work, is a matrix generated by propagating an identity matrix with the double difference operator, thus all satellites have the same weight. The second approach uses the elevation angle as a parameter to estimate the standard deviation of the observations (here this approach is being called Elevation based). After the standard deviations are estimated, they are propagated using the double difference operator as well. Figure 6 shows a representation of covariance matrices estimated using these three schemes.

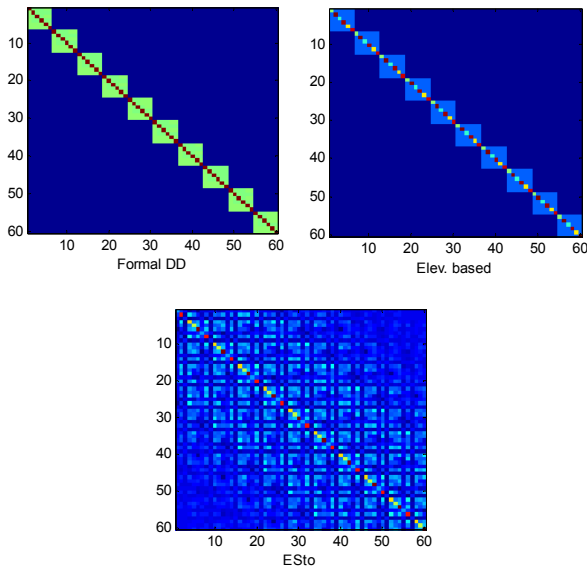


Figure 6. Representation of covariance matrices estimated using different schemes.

In the figure above, different colors represent different values within the matrices. Each one of them has its own scale, and they were generated from the same data set. In the case of Formal DD, it can be noticed that the elements on the diagonal have always the same color, what means the same variance. The green squares around the diagonal are the correlation between different satellite pairs within the same epoch. This correlation exists with this pattern due to the use of a common reference satellite. In the second case (Elev. based), the main difference from the previous is in the different weights for different satellites (this can be noticed by the different colors in the diagonal). The blue squares are due to the same as the green ones in the previous. The dark blue means zero in both of them. When the ESto model is used, it can be noticed that a large amount of information is placed into the covariance matrix. These values were derived from raw data, and represent the correlation of each observation with respect to each other. The next step to validate the ESto model should be an analysis with real data processing, explored in next session.

DATA PROCESSING

In order to investigate the efficiency of this empirical stochastic approach with respect to the conventional techniques, it was used to process two short baselines using C/A pseudoranges. The choice of such data set is justified by the elimination of several errors in the measurements, such as clock errors and atmospheric refraction, which would need a special treatment in the stochastic analysis. It was used 24 hours of data, which were processed in half hour batch adjustments along the day. It was done for two baselines, UNB1-FRDN and GAGS-NRC1, with approximately 5 km and 20 km,

respectively. It were used the C/A pseudoranges as observables. Figure 7 shows the results of the adjustment for the UNB1-FRDN baseline.

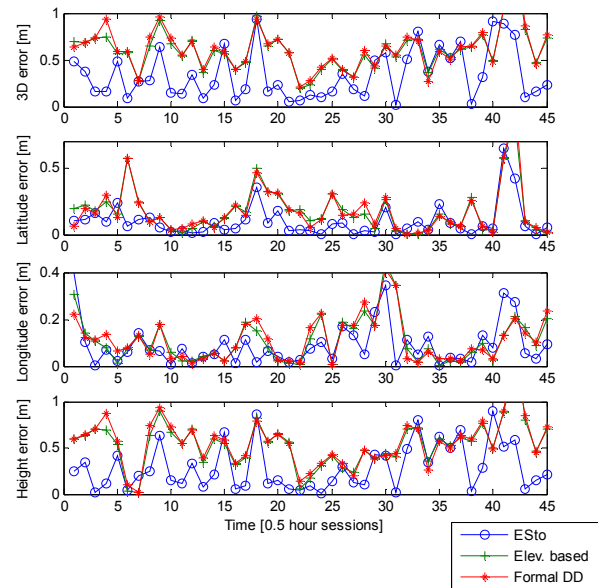


Figure 7. Results for the adjustment of the baseline UNB1-FRDN.

Although all results are fairly good, when ESto was used, the mean bias was quite lower than other solutions. However the standard deviation didn't show a good improvement. The use of ESto shows some errors quite larger than their mean values, which can be affecting standard deviation values. The use of ESto model provided better solutions in 90% of the situations for this baseline. Table 2 shows the results summary for this case.

Table 2

		Formal DD	Elev. based	ESto
Mean error (m)	Latitude	0.146	0.137	0.051
	Longitude	-0.081	-0.076	-0.024
	Height	0.543	0.523	0.153
	3D	0.632	0.614	0.341
RMS (m)	Latitude	0.257	0.248	0.155
	Longitude	0.146	0.146	0.128
	Height	0.623	0.599	0.380
	3D	0.689	0.664	0.430

Tables 3 and 4 show the absolute and relative improvements achieved by the use of ESto with respect to the other models.

Table 3

		w.r.t. Formal DD	w.r.t. Elev. based
Mean error (m)	Latitude	0.10	0.09
	Longitude	0.06	0.05
	Height	0.39	0.37
	3D	0.29	0.27
RMS (m)	Latitude	0.10	0.09
	Longitude	0.02	0.02
	Height	0.24	0.22
	3D	0.26	0.23

Table 4

		w.r.t. Formal DD	w.r.t. Elev. based
Mean error	Latitude	66%	64%
	Longitude	70%	68%
	Height	72%	71%
	3D	46%	44%
RMS	Latitude	40%	38%
	Longitude	12%	12%
	Height	39%	37%
	3D	38%	35%

ESto model provided improvements ranging from 5cm to 37cm in bias and 2cm to 26cm in RMS. The greatest improvements were achieved for height determination, and consequently also for three-dimensional positioning. In terms of relative results, the improvements are very good, ranging from 44% to 72% for mean bias and 12% to 40% for RMS. These results show that the use of the empirical approach can definitely bring advantages in terms of quality of the estimated coordinates.

Further analysis was carried out with the 20km baseline, in order to validate the model. Figure 8 shows the results for the GAGS-NRC1 baseline.

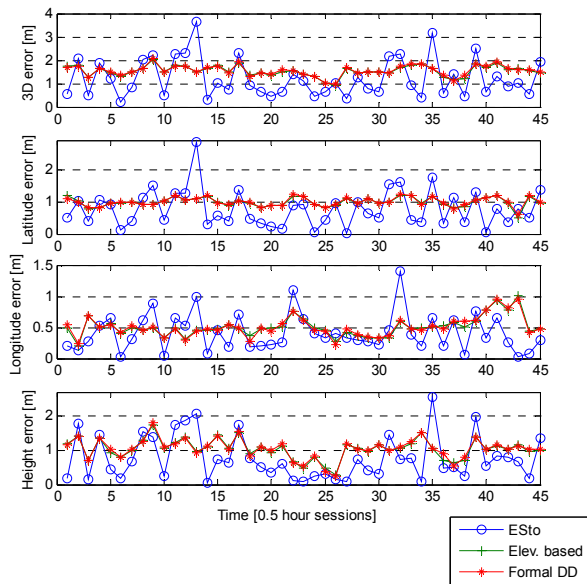


Figure 8. Results for the adjustment of the baseline GAGS-NRC1.

In this case, when ESto was used, the mean bias was lower than the other solutions. The standard deviation is higher for ESto solutions. As in the previous baseline, the use of ESto shows some solutions quite larger than the mean values of the errors, which can be affecting standard deviation values. The use of ESto model provided better solutions in only 70% of the situations in this case. Table 5 shows the results summary for this case.

Table 5

		Formal DD	Elev. based	ESto
Mean error (m)	Latitude	-0.993	-0.986	-0.762
	Longitude	0.508	0.506	0.395
	Height	1.050	1.036	0.696
	3D	1.558	1.543	1.223
RMS (m)	Latitude	1.003	0.996	0.948
	Longitude	0.533	0.529	0.508
	Height	1.090	1.076	1.000
	3D	1.574	1.560	1.473

Tables 6 and 7 show the absolute and relative improvements achieved by the use of ESto with respect to the other models.

Table 6

		w.r.t. Formal DD	w.r.t. Elev. based
Mean error (m)	Latitude	0.23	0.22
	Longitude	0.11	0.11
	Height	0.35	0.34
	3D	0.34	0.32
RMS (m)	Latitude	0.06	0.05
	Longitude	0.03	0.02
	Height	0.09	0.08
	3D	0.10	0.09

Table 7

		w.r.t. Formal DD	w.r.t. Elev. based
Mean error (m)	Latitude	23%	23%
	Longitude	22%	22%
	Height	34%	33%
	3D	22%	21%
RMS (m)	Latitude	5%	5%
	Longitude	5%	4%
	Height	8%	7%
	3D	6%	6%

For the GAGS-NRC1 baseline ESto model provided improvements ranging from 11cm to 34cm in bias and 2cm to 10cm in RMS. The greatest improvements were achieved for height determination, and consequently also for three-dimensional positioning, as in the previous solution. In terms of relative results, the improvements are very good, ranging from 21% to 34% for mean bias and 4% to 8% for RMS. These results show that for this case the empirical approach did not show great improvements, although the results are generally better than using other weighting schemes.

CONCLUSIONS AND FUTURE WORK

It was shown that the empirical model is capable of providing a good stochastic modeling of observables without any external source of information, before any adjustment.

The implementation of the model is computationally efficient, and can be easily implemented for GPS data processing. Although its use makes the process slower, it is not a great computational cost when compared with the gain in terms of the adjustment results. The greater time is due the computation of auto-correlation and cross-correlation functions that involve several combinations in order to have almost fully populated covariance matrices.

Two baselines were processed using three different weighting schemes for observations. In general, when ESto was used a good improvement in the bias was obtained. The mean biases were at least 44% smaller in the first case and at least 21 % in the second one. The RMS values had improvements of at least 12% for the shorter baseline and at least 4% for the longer one.

The greater improvements were achieved for height determination, with improvements up to 72% in mean bias for the shorter baseline.

In terms of standard deviation, the solutions obtained with ESto showed small or no improvements. There are still some outliers in the solution when ESto is used, what means that there is room for improvement in this sense. An attenuation of those outliers would help to make the standard deviations lower.

The results for the shorter baseline (2km) seemed to provide greater improvements what implies that maybe even for relatively short baselines (20km), some additional treatment in the stochastic analysis can be necessary.

The model certainly brings improvement to the solution, and future investigation is needed to investigate the source of solution outliers and improvements in the stochastic analysis for longer baselines.

Further research to apply ESto to point positioning and carrier phase measurements is also a desirable future step in the development and validation of the model.

REFERENCES

Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. (1994). Time Series Analysis – Forecasting and Control. Prentice-Hall International, London.

Collins, J. P. and R. B. Langley (1999). Possible weighting schemes for GPS carrier phase observations in

the presence of multipath. Final contract report for the U.S. Army Corps of engineers Topographic Engineering Center, No DAAH04-96-C-0086 / TCN 98151, March, 33pp.

El-Rabbany, A. (1994). The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning. Ph.D. Thesis - University of New Brunswick, Department of Geodesy and Geomatics Engineering, Fredericton, NB.

Hofmann-Wellenhof, B., H. Lichtenegger and J. Collins (2001). Global Positioning System: Theory and Practice. Springer-Verlag Wien New York.

Kim, D and R. Langley (2001). Estimation of the Stochastic Model for Long-Baseline Kinematic GPS Applications. ION National Technical Meeting, January 22-24, 2001, Long Beach-CA.

Langley, R. B. (1997). GPS observation noise. GPS World, Vol. 8, No. 6, April, pp 40-45.

Vanicek, P. and E. Krakiwsky (1982). Geodesy: The Concepts. North-Holland Publishing Company, Amsterdam – New York – Oxford.