DETECTION OF WIND-INDUCED RESPONSE BY WAVELET TRANSFORMED GPS SOLUTIONS

Clement Ogaja[1], Jinling Wang[2] and Chris Rizos[3]

ABSTRACT: GPS solutions offer direct dynamic monitoring of long-span and tall civil structures. Timely and correct interpretation of the solutions is essential to improved monitoring and reduced number of false alarms. In this article, a new approach has been developed which employs the principal component analysis (PCA) of wavelet transformed GPS solutions to detect out-of-control signals. The methodology consists of pre-filtering the original GPS solutions via a FIR median hybrid (FMH) filter, and applying the PCA to the Haar wavelet transform (HWT) of the FMH-filtered results. It has been tested off-line using observation data from a monitoring experiment involving GPS and multi-sensors. Observation data were collected for about seven days in February 2000, including two strong wind periods. The PCA control chart of the wavelet transformed solutions was found to respond better to the wind periods compared to that of the "raw" solutions. The proposed method can be applied to monitor wind/earthquake or thermally-induced responses in the long-term deformation monitoring of large engineering structures by GPS.

[1]PhD Candidate, School of Surveying and Spatial Information Systems, University of NSW, Sydney NSW 2052, Australia. E-mail: c.ogaja@student.unsw.edu.au

[2] Postdoc. Fellow, School of Surveying and Spatial Information Systems, University of NSW, Sydney NSW 2052, Australia. E-mail: j.wang@unsw.edu.au

[3] Professor, School of Surveying and Spatial Information Systems, University of NSW, Sydney NSW 2052, Australia. E-mail: c.rizos@unsw.edu.au

keyword Terms: Global Positioning System (GPS), Principal Component Analysis (PCA), Wavelet Transform, Deformation Monitoring, FIR Median Hybrid (FMH) filter.

INTRODUCTION

Previous studies have indicated that the use of GPS technology for the long term monitoring of long-span and tall civil structures is more beneficial for continuous measurements, compared to the use of other sensors such as accelerometers which cannot readily offer real-time solutions (Ashkenazi *et al*, 1997; Celebi *et al*, 1998; DeLoach, 1989; Duth & Hyzak, 1997; Guo & Ge, 1997, Lovse *et al*, 1995). Some of the recent developments in this area of research (e.g. Ogaja et al, 2001; Roberts et al, 2000) have also indicated that such schemes tend to become more heavily instrumented with multiple GPS units and additional sensors. This, in addition to the more frequent recording by real-time kinematic (RTK) GPS, leads to the challenge of having to deal with large amounts of data, that impose heavy computational and wireless communication loads. Data reduction techniques, such as the principal component analysis (PCA) and partial least squares (PLS), therefore become necessary tools for handling the large amounts of data generated. These techniques can be easily implemented in projects where the GPS processing software runs on a centrally located control station PC, receiving real-time streams of data, and computing precise positions for all GPS stations located on the target structure. However, since GPS solutions are normally contaminated by both random and systematic errors due to multipath and other unknown sources, a proper pre-treatment of the data is necessary for their correct interpretation. Pre-filtering combined with the wavelet transformation (Wang, 1995) of the solutions can be used to better locate and identify the out-of-control signals captured during field observations.

This article looks at the results from an experiment involving GPS and auxilliary sensors. A comparative approach is used to show that the PCA of GPS solutions only is less responsive to the maximum wind conditions compared to the PCA of wavelet transformed FHM-filtered solutions. The first case considers the PCA of all the measured variables and gives the contributions to the out-of-control signals detected by setting the control limits. In the second case, an approach is developed in which the original time series of the GPS solutions are pre-filtered using FIR median hybrid (FMH) filter, and transformed into the wavelet domain before applying the PCA method. The results obtained in this way seem to be quite reasonable compared to the results from the PCA of the time-domain GPS solutions.

EXPERIMENT DESCRIPTION

The data used in this study were observed during a building monitoring experiment in Singapore. The experiment was performed over seven days (17-24 February 2000) on the rooftop of the Republic Plaza Building (RPB), the tallest building in Singapore. An RTK-GPS system was used to provide the GPS solutions in parallel with two rooftop UVW anemometers and B1 accelerometers, collecting wind velocity and acceleration response data respectively. The GPS system consisted of two Trimble 4700 receivers, one installed as a 'rover' station on the RPB rooftop, as shown in Fig. 1a, and the other as a 'base' station on the rooftop of the Finger-Pier Buidling (Fig. 1b), a low-rise building about 700m from the RPB. The GPS solutions were output every second (1Hz). Figure 2 shows the resampled historical data of the GPS observations (GPSN & GPSE), wind velocity (UE & UW), accelerometer response (δE & δN) and differential temperature (T) recorded from the experiment.

The experiment covered two periods of maximum wind conditions reaching up to more than 10 m/s, during the 19th and 22nd of February 2000.

THE PCA ALGORITHM

Theory

PCA is suitable for monitoring processes where p correlated variables are being measured simultaneously. In such circumstances it is assumed that the variables follow a p-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}' = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_p)$ and covariance matrix $\boldsymbol{\Sigma}$, where $\boldsymbol{\mu}_i$ is the mean for the i^{th} variable and $\boldsymbol{\Sigma}$ is a $p \times p$ matrix consisting of the variances and covariances of the p variables. The multivariate normality is assumed at the "start up" stage solely for the purpose of deriving the control limits. After the control limits are established, the data are assumed to be reasonably normally distributed.

In the "start up" stage, a sample of m subgroups of past data are available to estimate the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. If the i^{th} epoch observation of the p variables from the reference samples is represented by the vector $\mathbf{X}_i$, and their estimated mean vector containing the means of each variable is $\overline{\mathbf{X}}_i$, the estimated covariance matrix becomes:

$$\mathbf{S}_m = \frac{1}{m-1}\sum_{i=1}^{m}(\mathbf{X}_i - \overline{\mathbf{X}}_m)(\mathbf{X}_i - \overline{\mathbf{X}}_m)'.$$

To construct a multivariate control chart based on Hotelling's T^2 statistic, for observation $\mathbf{X}_i$ one uses the charting statistic:

$$Q_i = (\mathbf{X}_i - \overline{\mathbf{X}}_m)' \mathbf{S}_m^{-1}(\mathbf{X}_i - \overline{\mathbf{X}}_m). \qquad (1)$$

where the distribution of Q_i is approximated with a chi-square or an F distribution to obtain the control chart limits (see Appendix I).

Once the control limits are established at the "start up" stage, monitoring of future samples is realised by the charting statistic:

$$T_f^2 = (\mathbf{X}_f - \overline{\mathbf{X}}_m)' \mathbf{S}_m^{-1}(\mathbf{X}_f - \overline{\mathbf{X}}_m) \qquad (2)$$

where $\mathbf{X}_f$ denotes the p-dimensional vector of future observations on the p variables, $\overline{\mathbf{X}}_m$ is the p-dimensional mean vector of the m observations in the "clean" reference samples, $\mathbf{S}_m$ is the $p \times p$ covariance matrix associated with these observations.

Case 1 (Response evaluation by all measurements)

The results from the experiment were first evaluated by applying PCA to all the seven variables measured (GPSE, GPSN, UE, UW, δE, δN, T). This gives a linear combination of the variables that describe the major trends captured by the individual measurements. The control limits of the Q-statistic are computed for this purpose, as shown in Appendix I. Figure 3 is the multivariate control chart of the Q-statistic with the control limits set so that the number of samples outside the control limits is 1% of the entire samples when the process is in-control, representing the 99% confidence limit. The control chart shows two periods, during the 19th and 22nd, with peaks 1 and 2 lying outside the control limits, signalling that significant events occurred during these periods. It follows from the definition of Q_i in (1) that points outside the multivariate control limits are a result of one or more of the seven signals being out-of-control. Proper interpretation requires consideration of the individual contributions to these conditions.

Contribution plots can be obtained which relates to the contribution of each of the seven variables to the two 'out-of-control ' peaks. Figures 4a and 4b are examples of the contribution plots for peak 1 and peak 2 respectively, differences in the variables ' magnitude giving an indication of which variables have led to the peaks. In both figures, the variable numbers 1 to 7 on the x-axis represent the measured components GPSE, GPSN, UE, UW, δE, δN and T respectively, from which it is evident that the wind vectors have contributed most to the two peaks, followed by the GPS vectors. One can therefore evaluate the wind-induced response from the measured GPS signals.

Case 2 (Response evaluation by GPS solutions only)

This section considers evaluating the out-of-control signals by using the GPS solutions only. In this case the PCA is applied to the two variables: the GPSE vector and the GPSN vector. The proposed three-step procedure is:

I. Applying a FIR Median Hybrid (or FMH) filter (Heinonen & Neuvo, 1987) iteratively to the GPS solutions to reduce the level of noise (if any). Given a sequence $\{x_1, ... ,x_N\}$ of the GPS solutions, the $(K+1)^{st}$ FMH filter output is based on the median of the three numbers: $\hat{x}_{K+1}$ the average of K past samples, $\bar{x}_{K+1}$ the average of K future samples, and the $(K+1)^{st}$ data point, x_{K+1}:

$$x_{FMH}(k+1) = median\{\hat{x}_{K+1,} \bar{x}_{K+1}, x_{K+1}\} \quad (3)$$

where K is the window half-width. This particular filter preserves steady-state conditions as illustrated by the example in Fig. 5.

II. Computing the wavelet coefficients of the results from step I using the Haar wavelet transform (HWT). If ψ is a Haar wavelet (Olivier & Vetterli, 1991; Wang, 1995) such that $\psi_s = s^{-\frac{1}{2}}\psi(x/s)$, the HWT of a function f is defined as:

$$Tf(s,x) = \int \psi_s(x-u) f(u) du \quad (4)$$

in which the transform $Tf(s,x)$ is a function of the scale, or frequency, s and the spatial position, or time, x. In the case of the FMH-filtered GPS solutions, a discretized version of the continuous wavelet transform in (4), written as a linear transform involving $n \times n$ orthogonal matrix W, is adopted.

III. Selecting the scale and applying PCA to the wavelet coefficients obtained from step II.

Figure 6 shows the results after applying the HWT to FMH-filtered GPS vectors, each containing 1112 samples. In both panels, the top plots indicate the original GPS vectors, the middle plots are the results after applying the FMH filter using 20 passes and with a half window size of 9, and the bottom plot indicates the HWT of the FMH filtered data. Note that the steady state conditions are preserved and additional filtering was not required as it did not improve the signals. Indications of the two (or more) out-of-control events are, however, not clearly evident from these plots. In comparison, applying PCA to the wavelet transformed data revealed three significant out-of-control peaks including the two wind conditions of 19th and 22nd February 2000. The plots of the *Q*-statistic both of the original unfiltered GPS signals and the wavelet transform of the FMH-filtered GPS signals are shown in Figure 7, the horizontal lines representing the 99% confidence limits.

CONCLUDING REMARKS

A discussion is given on the suitability of applying multivariate techniques to wavelet transformed GPS solutions in order to monitor wind-induced response. PCA is applied to the Haar wavelet transform of the GPS solutions, rather than on their time history. The advantage of this is that the variability defined by the wavelet coefficients are more closely associated with the two maximum wind conditions than when the PCA is applied to the history of the data.

The study demonstrates that pre-filtering of the data with a Finite impulse response Median Hybrid filter (FMH), and developing the calibration model from a set of coefficients obtained from a Haar wavelet transform (HWT) of the FMH-filtered data, produces a monitoring model with better (sharper) classification features.

By employing contribution plots, the validity of using GPS data to evaluate the response signals is established and the monitoring model is developed, not by employing PCA on the filtered data, or on the reconstructed data after applying the wavelet transform, but on the resulting wavelet coefficients themselves. In this way, the monitoring is carried out in the wavelet domain.

ACKNOWLEDGEMENTS

This work was completed during the principal author's study at the University of New South Wales, supported by the International Postgraduate Research Scholarship funded by The University of New South Wales. The second author is the holder of an Australian Research Council Postdoctoral Fellowship.

APPENDIX I. COMPUTING THE PCA CONTROL LIMITS

If one assumes that the estimates of $\overline{\mathbf{X}}_m$ and $\mathbf{S}_m$ are the true population values $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively, then Seber (1984) has shown that the statistic Q_i is distributed as a chi-square variate with p degrees of freedom. In that case, the lower control limit is:

$$LCL = \chi^2(1-\alpha/2;\, p) \tag{5}$$

and the upper control limit is:

$$UCL = \chi^2(\alpha/2;\, p) \tag{6}$$

where $\chi^2(\alpha;\ p)$ is the $1-\alpha$ percentile of the chi-distribution with p degrees of freedom.

If one assumes that the i^{th} observation $\mathbf{X}_i$ is independent of both $\overline{\mathbf{X}}_m$ and $\mathbf{S}_m$, then the statistic Q_i (times a constant) follows an F distribution with p and $m-p$ degrees of freedom. In that case the lower control limit is:

$$LCL = \frac{p(m-1)(m+1)}{m(m-p)} F(1-\alpha/2;\, p, m-p) \tag{7}$$

and the upper control limit is:

$$UCL = \frac{p(m-1)(m+1)}{m(m-p)} F(\alpha/2;\, p, m-p) \tag{8}$$

where $F(\alpha;\ p, m-p)$ is the $1-\alpha$ percentile of the F distribution with p and $m-p$ degrees of freedom.

Since neither of the above assumptions holds true in the "start up" stage, derivation of the exact distribution of the charting statistic Q_i is desired. Gnanadesikan & Kettenring (1972) have shown that Q_i (times a constant) has a beta distribution from which it is possible to construct the needed control limits. The lower control limit based on this distribution is given by:

$$LCL = \frac{(m-1)^2}{m} B(1-\alpha/2;\ p/2, (m-p-1)/2) \qquad (9)$$

and the upper control limit is given by:

$$UCL = \frac{(m-1)^2}{m} B(\alpha/2;\ p/2, (m-p-1)/2) \qquad (10)$$

where $B(\alpha;\ p/2, (m-p-1)/2)$ is the $1-\alpha$ percentile of the B distribution with parameters $p/2$ and $(m-p-1)/2$. The relationship between random variables with the beta and F distribution:

$$\frac{(p/(m-p-1))F(\alpha;\ p, m-p-1)}{1+(p/(m-p-1))F(\alpha;\ p, m-p-1)} = B(\alpha/2;\ p/2, (m-p-1)/2) \qquad (11)$$

can then be utilized to set the following control limits:

$$LCL = \frac{(m-1)^2}{m} \times \frac{(p/(m-p-1))F(1-\alpha/2;\ p, m-p-1)}{1+(p/(m-p-1))F(1-\alpha/2;\ p, m-p-1)} \qquad (12)$$

and

$$UCL = \frac{(m-1)^2}{m} \times \frac{(p/(m-p-1))F(\alpha/2;\ p, m-p-1)}{1+(p/(m-p-1))F(\alpha/2;\ p, m-p-1)} \qquad (13)$$

in terms of the percentiles from the F distribution.

In most cases the LCL may be set to zero. The reason for this is that any shift in the mean will always lead to an increase in the statistic Q_i, and thus the LCL may be ignored.

APPENDIX II. REFERENCES

Ashkenazi, V., A.H. Dodson, T. Moore & G.W. Roberts (1997). Monitoring the movements of bridges by GPS. *10th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation GPS ION'97*, Kansas City, Missouri, 16-19 September, 1165-1172.

Celebi, M., W. Prescott, R. Stein, K. Hudnut, J. Behr & S. Wilson (1998). Structural monitoring using GPS. *11th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation GPS ION'98*, Nashville, Tennessee, 15-18 September, 929-935.

DeLoach, S.R. (1989). Continuous deformation monitoring with GPS. *Journal of Surveying Engineering*, 115(1), 93-110

Dunia, R., S.J. Qin, T.F. Edgar & T.J. McAvoy (1996). Identification of faulty sensors using principal component analysis. *AIChE J.*, 42, 2797-2812.

Duth, K. & M. Hyzak (1997). Structural monitoring with GPS. http://www.tfhrc.gov/ pubrds /spring97/gps.htm.

Gnanadesikan, R. & J.R. Kettenring (1972). Robust estimates, residuals and outlier detection with multiresponse data. *Biometrics* 28, 81-124.

Guo, J. & S. Ge (1997). Research of displacement and frequency of tall building under wind loading using GPS. *10th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation GPS ION'97*, Kansas City, Missouri, 16-19 September, 1385-1388.

Heinonen, P. & Y. Neuvo (1987). Fir-median hybrid filters. *IEEE Trans. Acoustics, Speech & Signal Processes*, 35, 832-838.

Jackson, J.E. (1980). Principal components and factor analysis: Part I – principal components. *Journal of Quality Technology*, 12, 201-213.

Lovse, J., W.F. Teskey, G. Lachapelle & E.M. Cannon (1995). Dynamic deformation monitoring of tall structures using GPS technology. *Journal of Surveying Engineering*, 121(1), 35-40.

Ogaja, C., C. Rizos, J. Wang, & J. Brownjohn (2001). High precision dynamic GPS system for on-line structural monitoring. *5th Int. Symp. on Satellite Navigation Technology & Applications*, Canberra, Australia, 24-27 July, paper 35, CD-ROM Proc.

Olivier, R. & M. Vetterli (1991). Wavelets and signal processing, *IEEE Sig. Processing*, October, 14-38.

Roberts, G.W., X. Meng & A.H. Dodson (2000). Structural dynamic and deflection monitoring using integrated GPS and triaxial accelerometers. *13th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation GPS ION 2000*, Salt Lake City, Utah, 19-22 September, 59-68.

Seber, G.A.F. (1984). Multivariate Observations. John Wiley & Sons, New York.

Wang, Y. (1995). Jump and sharp cusp detection by wavelets. *Biometrika*, 82, 2, 385-397.

APPENDIX III. NOTATION

The following symbols are used in this paper:

K	=	window half-width of FIR median hybrid filter;
m	=	number of data sample subgroups;
p	=	measured variables;
Q_{i}	=	PCA charting statistic;
$\mathbf{S}_{\boldsymbol{m}}$	=	estimated $p \times p$ covariance matrix associated with m observations;
T_f^2	=	Hotelling's charting statistic for monitoring future samples;
$Tf\,(s, x)$	=	Haar Wavelet Transform (HWT);
$\mathbf{X}_f$	=	vector of future observations on p variables;
$\overline{\mathbf{X}}_i$	=	estimated mean vector at i^{th} epoch;
$\overline{\mathbf{X}}_m$	=	the p-dimensional mean vector of the m observations;
$\hat{x}_{K+1}$	=	the average of K past samples;
$\bar{x}_{K+1}$	=	the average of K future samples;
x_{K+1}	=	the $(K+1)^{\mathrm{st}}$ data point;
α	=	statistical level of significance;
$\boldsymbol{\Sigma}$	=	$p \times p$ covariance matrix;
$\boldsymbol{\mu}'$	=	true mean vector;
ψ	=	Haar Wavelet;
χ^2	=	chi-square;

Figure captions

Fig. 1a: GPS rover on the high-rise RPB rooftop.

Fig. 1b: GPS base station(s) on the low-rise Finger-Pier Building rooftop.

Fig. 2: Simultaneous GPS and Auxilliary Sensor data. From Top to bottom: **Panel 1 & 2:** GPS signals; **Panel 3 & 4:** Wind velocity vectors; **Panel 5 & 6:** Acceleration response vectors; **Panel 7:** Differential temperature.

Fig. 3: Multivariate monitoring by simultaneous GPS and Auxilliary Sensor data. Two out-of-control signals detected

Fig. 4a: Contribution to peak 1. of Fig. 3

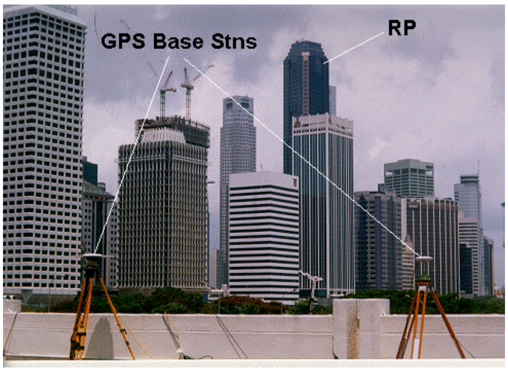
Fig. 4b: Contribution to peak 2. of Fig. 3

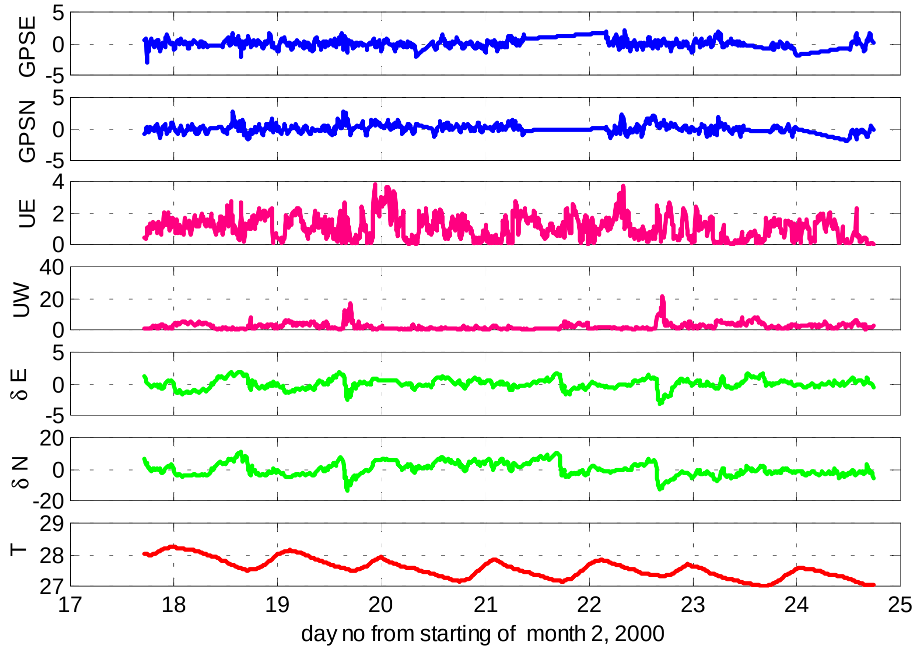
Fig. 5: FMH/Median hybrid filter example. **Top left panel:** original data signal; **top right panel:** signal with Gaussian random noise, **bottom left panel:** filtered signal after 50 consecutive passes and a window half-width of 9; **bottom right panel:** a superposition of the filtered and the uncorrupted signal.

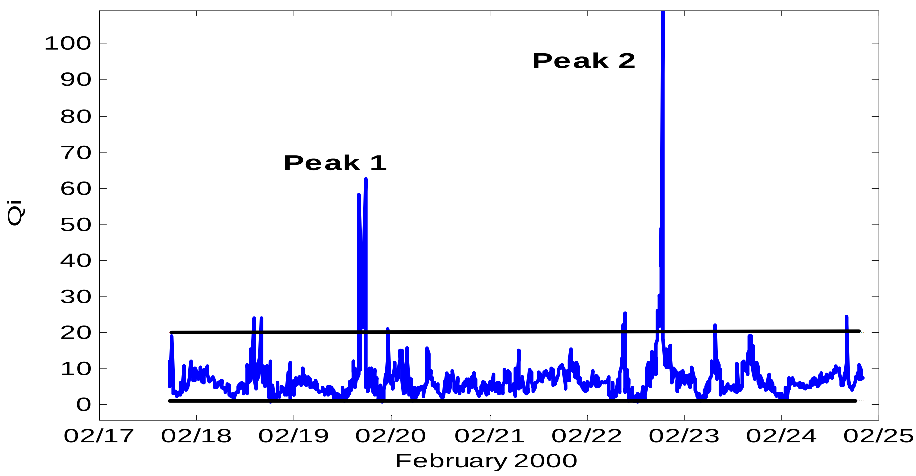
Fig. 6: Wavelet Transform of FMH filtered GPS data signals. **Top panel:** HWT of FMH filtered GPSE data signals; **bottom panel:** HWT of FMH filtered GPSN data signals.

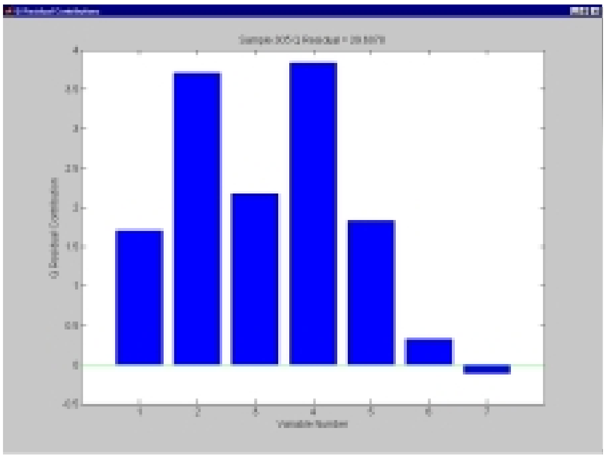
Fig. 7: Monitoring by *Q*-statistic. **Top panel:** *Q*-statistic of original GPSE and GPSN signals; **bottom panel:** *Q*-statistic of wavelet transformed FMH filtered GPSE and GPSN signals.

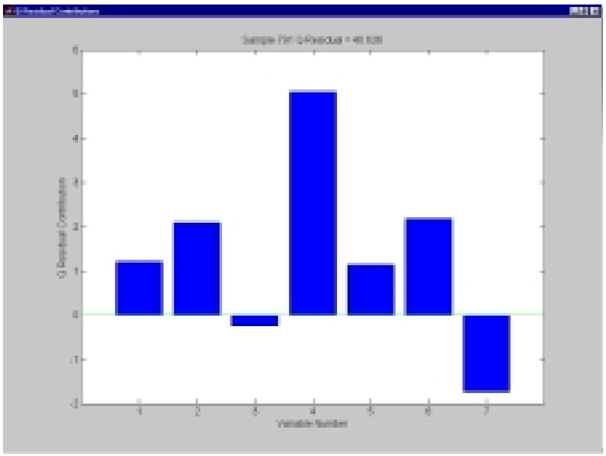

GPS Antenna -- Rover Stn.

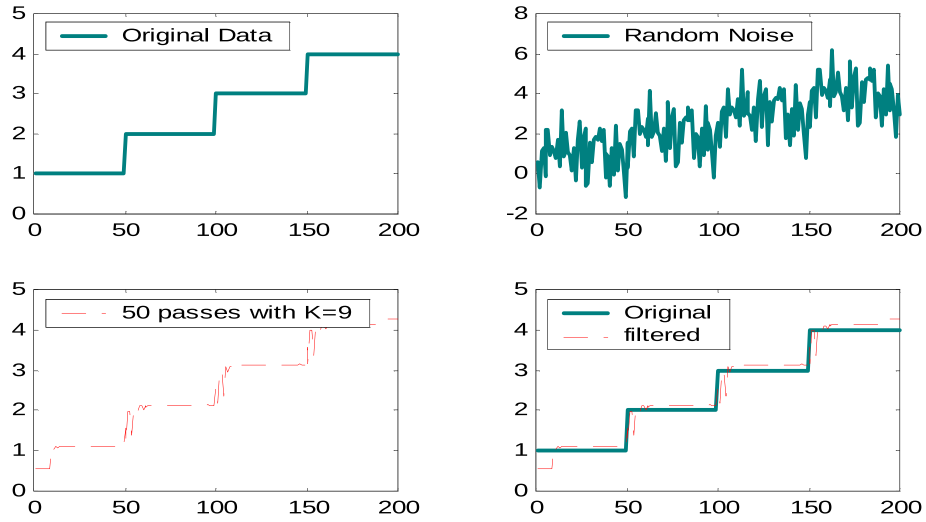

GPS Base Stns
RP

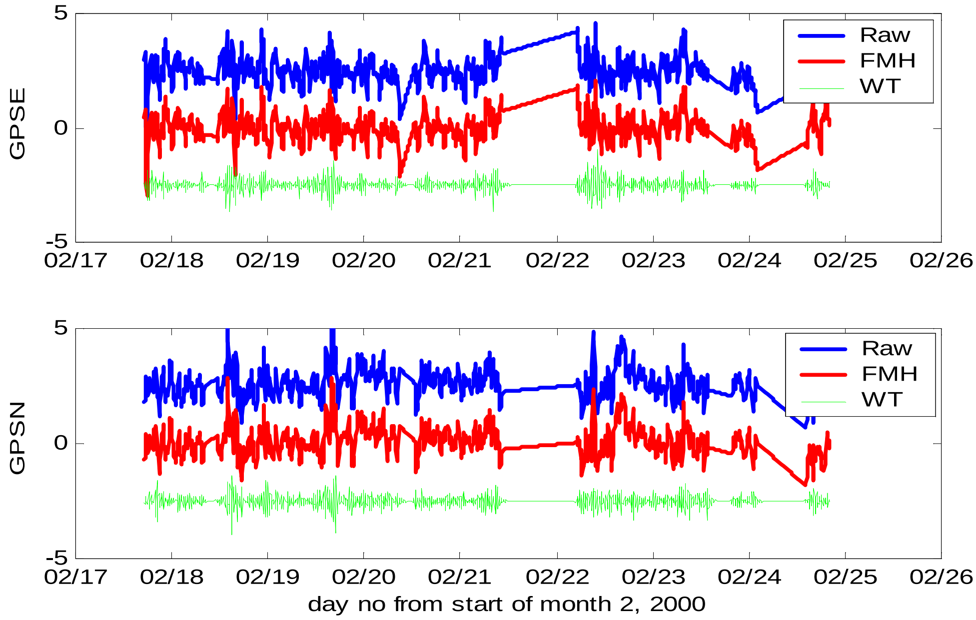

GPSE
GPSN
UE
UW
δ E
δ N
T
day no from starting of month 2, 2000

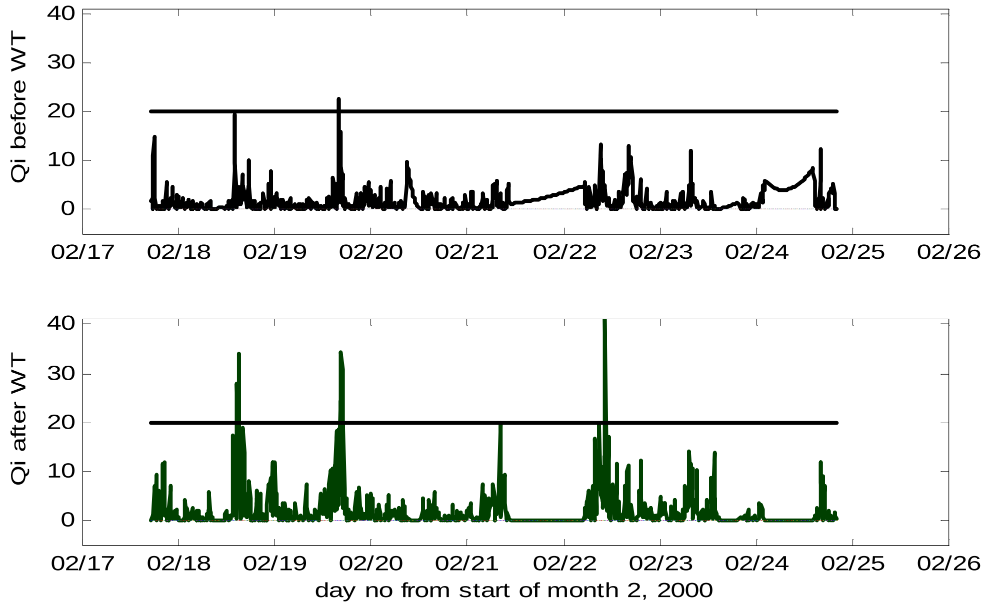

Peak 1
Peak 2
Qi
February 2000


Q Residual Contribution
Variable Number


Q Residual Contribution
Variable Number


Original Data
Random Noise
50 passes with K=9
Original
filtered


GPSE
GPSN
Raw
FMH
WT
02/17
02/18
02/19
02/20
02/21
02/22
02/23
02/24
02/25
02/26
day no from start of month 2, 2000


Qi before WT
40
30
20
10
0
02/17
02/18
02/19
02/20
02/21
02/22
02/23
02/24
02/25
02/26
Qi after WT
40
30
20
10
0
02/17
02/18
02/19
02/20
02/21
02/22
02/23
02/24
02/25
02/26
day no from start of month 2, 2000