Wavelet transform and its application to detect the blunders in an observation vector

Amir Reza MORADI, Master of Science in Geodesy, Science &Technology University, Iran, Mosha 1379@gmail.com
Farzad ALEMI, Bachelor of Surveying engineering, Science &Technology university, Iran, farzad.alemi@gmail.com

Abstract: *Wavelet transform as a multi-resolution analysis tool can be used to detect the contingent blunders or gross errors in a set of observations which plays an important role in geomatic analysis that is now done by the use of statistical methods.*
In this paper the one dimensional wavelet transform based on Haar wavelets and scaling functions has been used to separate the fine and coarse parts of the observation vectors and after that for distinguishing the existed blunders in them by interpreting the fine part.
In a numerical example the efficiency of our desired wavelet based method has been compared with statistical approaches.

Keywords*: wavelet transform; blunder detection; decomposition*

1. Introduction

Wavelet theory has been used in applied mathematics and engineering for well under thirty years (Zuofa Li 1996), notwithstanding, by virtue of its marvelous properties, it is now very engrossing tool for applied science specialists such as surveyors and geodesists.
Mathematical transformations are exerted to signals to obtain further information, but it is not accessible in the raw signal. (In the following paper I assume a space-domain signal as a raw signal, and a transformed one beside any of the exerted transformations as a processed signal.)
Most of the signals in practice are functions of time or space in their raw format; this representation is not always the best representation of the signal and it would be more efficient and obvious study the correspondent phenomena at the frequency domain, owning to the fact, their frequency contents correlate to fine and coarse information can be readily viewed at that domain.

Regardless of the fact that, the Fourier transform is probably the most popular, but it also has its deficiency, which can only tells us how much of each frequency exists in the signal, and not tell us when or where these frequency components exist. In contrast, the Wavelet transform provides the time-frequency representation that can answer to both questions; as well as this property, it is a multi resolution analysis tool which considers the data structure and is used as a tool to link different resolution levels.

As we know, the physical observations always contain different type of errors which can be put into 3 main groups: blunders, systematic and random errors. These three types are inherently different and thus the corresponding behavior, to estimate the subsequent unknown parameter in a geomatics analysis procedure, will not be the same.

The convenient scheme to distinguish the type of the error, existing in a set of measurements made to an observable, is usually based on statistical methods and the concept of confidence intervals; in this paper, the one dimensional discrete wavelet transform has been used to determine if some measurements were affected by grass errors or blunders.
Before we proceed, it is necessary to review some foundations of the wavelet theory.

2. An Introduction to wavelet transforms

According to our knowledge a wave is the representation of an oscillating function in the time or space domain, such as a sinusoidal wave. Similarly a wavelet can be defined as a small wave with restricted energy in time or space.

Fourier and wavelet transforms are both wave based analysis methods, In Fourier transform a signal is expanded through sinusoidal waves, but the expansion tool in wavelet transform are wavelets.

If a signal or function f (t) can be expanded to a linear combination of some real-valued functions ($f(t)=\sum_i a_i\psi_i(t)$), then analysis, representation and processing of that signal will be efficiently possible, and this is the base of spectral analysis.

If the set of expansion coefficients is unique, for a given function, the expansion functions will be known as a base for the space contains *f(t)*. This base will be orthonormal if:

$$<\psi_k(t),\psi_l(t)>=\int\psi_k(t)\psi_l(t)dt=\delta_{k,l} \qquad (1)$$

In such an orthonormal system, the expansion coefficients are computed through the following inner product:

$$a_k=<f(t),\psi_k(t)>=\int f(t)\psi_k(t)dt \qquad (2)$$

In a Fourier series, exponential functions are the orthogonal expansion functions; on the other hand, these functions, in a Tailor series, are non-orthogonal polynomials t^k.
Now, in comparison with two above series, the wavelet expansion of a function f(t), will be:

$$f(t)=\sum_k\sum_j a_{j,k}\psi_{j,k}(t) \qquad (3)$$

In the above series, the base functions, which can be orthogonal or not, are the wavelets and the expansion coefficients, $a_{j,k}$ s, are known as the discrete wavelet transform of $f(t)$ and named as an inverse discrete wavelet transform, the above summation reconstructs the original signal.

Being a mapping, DWP transforms a function to a set of two dimensional discrete coefficients and the reason for which the analysis of the transformed signal form both time and frequency points of views, will be possible, similar to a set of notes in a musical score which their position shows both time and frequency related to them.

One of the characteristics of the wavelet transform is the capability of designing the base functions (wavelets) in the transformation, according to the desired application and its properties; nevertheless, all of the wavelet systems have the same following specifications:

- All of the wavelets can be constructed by translating and scaling of a main producer function known as mother wavelet:

$$\psi_{j,u}(t)=2^{j/2}\psi(2^jt-k) \qquad (4)$$

In the above formula, $\psi(t)$ is the mother wavelet and indexes j and k are respectively the translation and scale parameters. Any changes in k, results in the position of the base function (wavelet) in the time-domain and as the same any alterations in j; alter the wavelet’s width and the resolution of the extracted information.

A wavelet system has time and frequency localization simultaneously; hence, it differs between the analysis of stationary and non-stationary signals in the wavelet transform

2.1 scaling and wavelet functions

As we know in Taylor expansion of a function, the first term of the series reflects a coarse perspective of the function and the other terms, containing the differentiations of the function, shows the details of it such that the further differentiations corresponds to more details of the function expanded; similarly, this property can be obtained in the wavelet analysis through a special function known as scaling function, besides the wavelets.

The story starts with a base producer scaling function $\varphi(t)$; by translating this function a set of bases $\{\phi_k(t)\}$ will be obtained:

$$\varphi_\kappa(t) = \varphi(t-\kappa) \tag{5}$$

These functions span a subset of $L^2(R)$:

$$V_0 = \overline{span\{\phi_k(t)\}}$$
$$\forall f(t) \in V_0,\ f(t) = \sum_\kappa a_\kappa \varphi_k(t) \tag{6}$$

In a more general case $\varphi(t)$ can be dilated or stretched using a scaling parameter j, so we can reach the set $\{\varphi_{j,\kappa}(t) = 2^{j/2}\varphi(2^j t - \kappa)\}$ spans V_j

$$V_j = \overline{span\{\phi_{k,j}(t)\}},$$

$$\forall f(t) \in V_j,\ f(t) = \sum_\kappa a_k \phi_{j,k}(t) \tag{7}$$

(Note that the coefficient $2^{j/2}$ in the above base function definition is due to gain functions with a constant norm.)

Providing that we construct the spaces V_j in a such manner, the scale parameter; j, will reflect the concept of resolution and the following relation exists:

$$V_0 \subset V_1 \subset \ldots \subset L^2(R) \tag{8}$$

In conclusion, a recursive equation, as the base of multi- resolution analysis reveals the relation between the scaling functions in V spaces:

$$\phi(t) = \sqrt{2}\sum h(n)\phi(2t-n) \tag{9}$$

In which the coefficients *h*(*n*), are known as scaling coefficients.

Now, on condition we consider the orthogonal complement of V_j in V_{j+1} as W_j, and the wavelet function $\psi_{j,k}$, mentioned before, will span this space:

$$V_{j+1} = V_j \oplus W_j \\ , \\ V_0 \oplus W_0 \oplus W_1 \oplus ... = L^2(R) \tag{10}$$

Similar to the relation between the scaling's functions, the following expression exits for the wavelets by introducing a new set of coefficients, $g(n)$,s are known as wavelet coefficients:

$$\psi(t) = \sqrt{2}\sum g(n)\phi(2t-n) \tag{11}$$

By the way, the expansion of any given function belongs to $L^2(R)$, base on both scaling and wavelet functions is possible:

$$f(t) = \sum_{k=-\infty}^{+\infty} c_k \phi_k(t) + \sum_{j=0}^{+\infty}\sum_{k=-\infty}^{+\infty} d_{j,k}\psi_{j,k}(t) \tag{12}$$

In the last expression, the first part consists of the coarse information of the function and the details are reflected at the second part; such that the bigger *j*, s the more obtained details.

With an orthonormal base, the expansion coefficients are computed by using the following inner products:

$$c_k = \langle f(t), \phi_k(t) \rangle, \\ d_{j,k} = \langle f(t), \psi_{j,k}(t) \rangle \tag{13}$$

The above coefficients are the discrete wavelet transform of *f*, and will being the last summation the inverse wavelet transforms reconstructs the original signal corresponding to the function *f*(*t*).

2.2 Wavelet transforms and filters

Considering the recursive relations, mentioned before, it can be shown that:

$$cj(k) = \sum_m h(m-2k)c_{j+1}(m) \tag{14}$$

$$dj(k) = \sum_m g(m-2k)c_{j+1}(m) \tag{15}$$

And reversely:

$$c_{j+1}(k) = \sum_m cj(m)h(k-2m) + \sum_m dj(m)g(k-2m) \tag{16}$$

Now we can look at the wavelet transforms and its inverse as a filter which convolves a given sequence of a signal with some filtering coefficients and produces a new sequence in a different resolution; such that, if there exists a sequence of a function samples, these quantities will be corresponding to c_{j+1}, s at the highest resolution and the signal sequence at the lower resolution can be derived by a low- pass filtering with impulse h (n).

Meanwhile, the added details, wavelet coefficients, will be computed by using a high –pass filter with impulse g (n). This process is referred as the decomposition of the signal. Reversing this procedure, known as synthesis or reconstruction of the signal can be obtained by using the inverse discrete wavelet transform in equation (16).

3. Blunders detections by using discrete wavelet transform

Being the observations, the outliers or blunders appear to be inconsistent with the reminder of the collected data (Iglewicz, 1993). As we know, the statistical procedure is the convenient scheme to determine if there exist some blunders in a set of repeated observations made to a given physical quantity. The criterion will be a defined confidence interval at a specific confidence level, (1-α); such that any observation, existing out of this range will probably be a blunder and must be discarded in the future analysis procedure based on the measured quantities.

One of the most interesting peculiarities of the wavelet transform is the possibility of separating the high and low resolution parts in a given signal; therefore, we can obviously extract the details of a signal, corresponding to the different frequencies existing in it. Conceded that this is the characteristic of some similar mathematical transformations such as Fourier transform, but in the wavelet transform, the frequency contents of a signal contain the corresponding times or spaces of occurrence. Now if a physical quantity is measured ‘n’ times, the vector containing these quantities can be considered as a time series with a relatively constant domain in which are small fluctuations for the sake of the existing errors; accordingly, by separating the fine and coarse parts of the vector using the decomposition procedure in the wavelet transforms, we hope to encounter with an obvious difference in the fine part of the vector of observations at the elements containing gross errors.

It has been mentioned that one of the useful properties of the wavelet transform is the possibility of designing of the base wavelet functions related to the desired application; it can be shown that in a wavelet system with "N" scaling or wavelet coefficients, after applying the existence and orthogonallity conditions of the base functions, will be "N/2-1" remaining degrees of freedom in designing or choosing the base functions.

It is needless to say, the resulting details of the observation vector will be varied in each selected or designed wavelet system; here, we have used the simplest system with N=2, known as 'Haar' wavelet system without any remaining degrees of freedom; the main reasons for this selection, couple with the simplicity of the computations, is the special property of the ‘Haar’ system in which scaling and wavelet coefficients are:

$$h(n) = \{h(0) = 1/\sqrt{2}, h(1) = 1/\sqrt{2}\},$$
$$g(n) = \{g(0) = -1/\sqrt{2}, g(1) = 1/\sqrt{2}\}$$

So, the computed images of the fine and coarse vectors at the argument (time) space, choose a special form such that every two adjacent elements of the coarse vector image will be the same and equal to the mean of two corresponding elements at the raw observation vector; furthermore, the detail vector is the difference of the first observation and the coarse vector computed by the wavelet transform; therefore the following expressions will be valid in such a system:

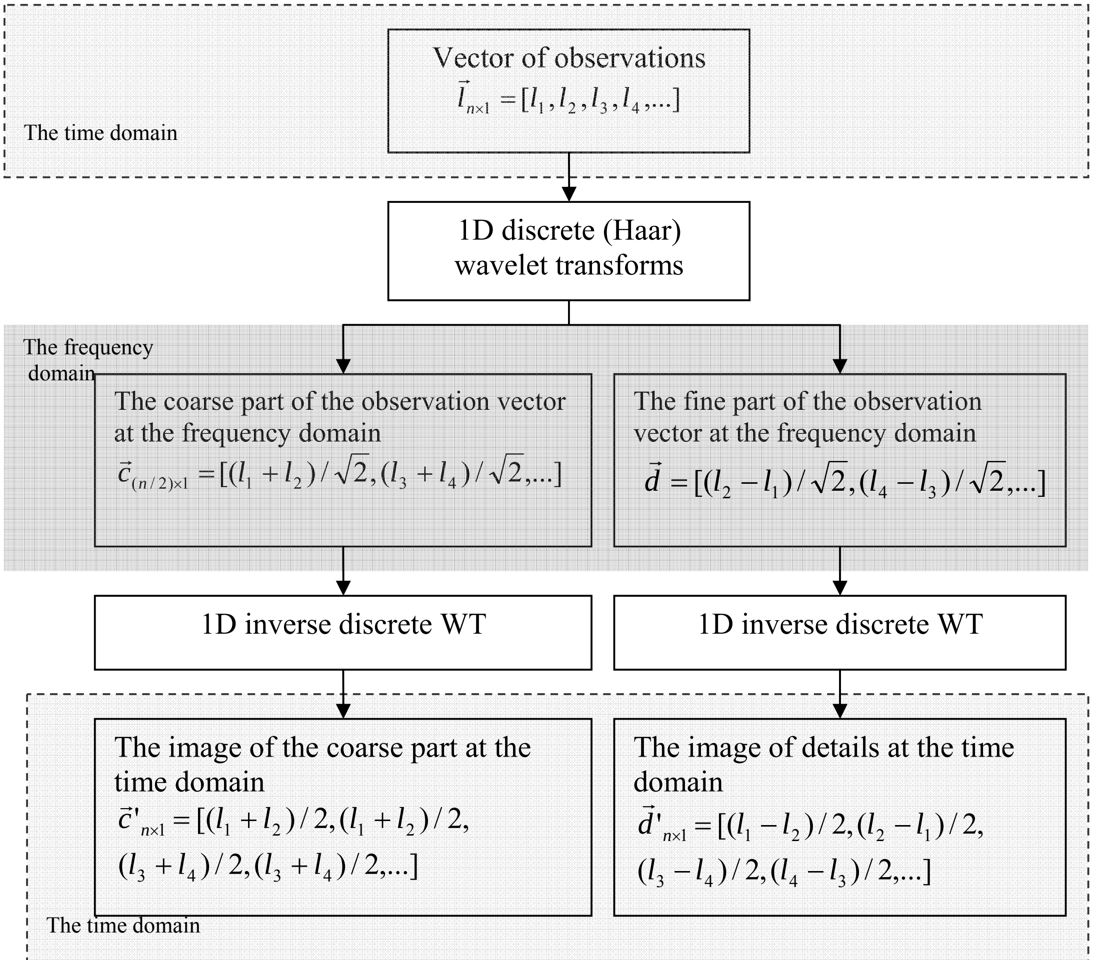

Fig 1. Diagrammatic procedure to compute the fine and coarse parts of observation vector using Haar system

As can be seen, the results of these transformations to the frequency and then back to the time domain are the images of the coarse and fine parts of the observation vector; our criteria to distinguish the probable blunders existing in the vector of observation will be based on the quantity of the elements of the image of the fine part; such that in the event there exists no blunder, the elements of the mentioned vector are very close to zero; because at the 'Haar' system used, each pair of the adjacent elements in the fine vector will be computed by subtracting the two corresponding elements of the original vector of observation. (see figure1) On the other hand, the existence of some blunders will cause an obvious deviation from zero in two corresponding elements of the fine vector.

Now, we are supposed to discuss the maximum permitted deviations from zero in the elements of the image of the fine vector.

There are some exceptions in the above scheme that I would rather explain during the following examples.

3.1 Numerical example

Illustrating the wavelet based method and related exceptions, three numerical examples will be introduced in this section.

The arc-second portion of *50* direction readings from *1"* instrument are listed below as the vector of observation may contain some blunders.

$\vec{l}_1$ = [41.9 49.5 42.6 45.5 46.3 45.5 47.2 43.4 44.6 43.3 47.4 45.5 46.1 42.6 44.7…

43.1 42.5 44.3 44.2 46.1 45.9 46.1 46.3 43.6 45.0 45.6 49.5 41.8 42.0 52.0 46.0 …

44.7 47.5 45.5 44.3 46.2 43.2 43.4 42.8 43.2 43.0 42.2 47.1 46.8 45.7 44.3 44.7 …

47.6 44.1 45.6],

At first, the two conventional statistical approaches, based on "z- score method" and the confidence interval constructed by the normal distribution have been used to determine the probable blunders, respectively. The results of these two methods, by considering a 95 percent confidence interval would be the same:

The 2nd, 27th and 30th elements of the observation vector will be distinguished as blunders.

Now, let's use the wavelet transformation to see the result of our desired scheme and compare it with two above statistical methods.

The images of the coarse and fine parts of the $\vec{l}_1$, based on the approach, shown on figure 1, will be:

$\vec{c}'_1 =$[45.70 45.70 44.05 44.05 45.90 45.90 45.30 45.30 43.95 43.95 46.45 46.45…

44.35 44.35 43.90 43.90 43.40 43.40 45.15 45.15 46.00 46.00 44.95 44.95…

45.30 45.30 45.65 45.65 47.00 47.00 45.35 45.35 46.50 46.50 45.25 45.25…

43.30 43.30 43.00 43.00 42.60 42.60 46.95 46.95 45.00 45.00 46.15 46.15…

44.85 44.85]

And

$\vec{d}'_1 =$[-3.80 3.80 -1.45 1.45 0.40 -0.40 1.90 -1.90 0.65 -0.65 0.95 -0.95 1.75…

-1.75 0.80 -0.80 -0.90 0.90 -0.95 0.95 -0.10 0.10 1.35 -1.35 -0.30 0.30 3.85…

-3.85 -5.00 5.00 0.65 -0.65 1.00 -1.00 -0.95 0.95 -0.10 0.10 -0.20 0.20 0.40…

-0.40 0.15 -0.15 0.70 -0.70 -1.45 1.45 -0.75 0.75];

Let's take a closer look at the last gained vector, as a prototype, the image of the fine part at the time domain, in figure 2. We can see the two largest picks with same quantities and opposite signs, corresponding to adjacent 29th and 30th elements of the image of the details, then the larger elements will bet the pairs (2nd, 3rd) and (27th , 28th).

We expect that the probable blunder would be investigated among tree of these six components, because in the detail part of an observation vector without the blunders all elements should be near to zero; so the larger elements' magnitude of the fine vector, the less confident observation.

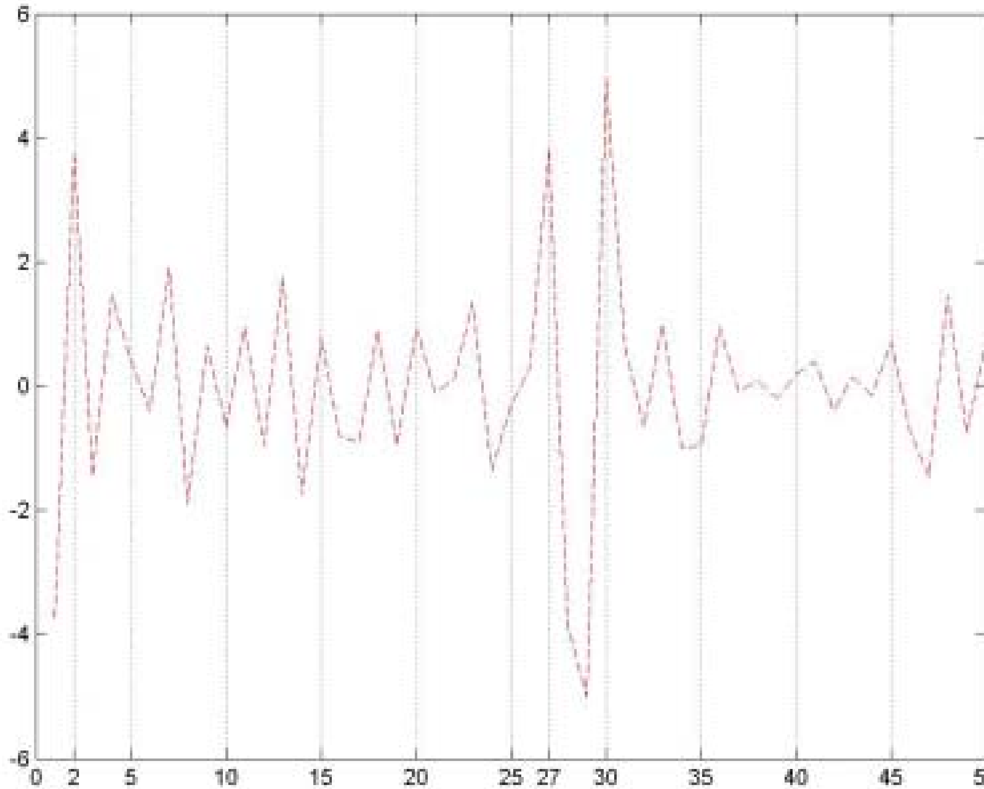

Fig 2. The image of fine part of observation vector

Now, we are encountering a problem to distinguish the actual blunder between two adjacent elements corresponding to two symmetric large picks in the above shown detail vector. To solve this problem, two adjacent doubtful observations is omitted, one at a time and by the use of the scheme presented in figure 1, the image of detail vector at the time domain is computed once more. The blunder is the one, which its corresponding large pick will not be removed as a result of its omission; on the contrary, if we abandon the precise observation the above shown large pick will be removed too; therefore, it can be stated now that the most probable blunder among the above sample observations will be 30^{th} and then 2^{nd} and 27^{th} ones, just the observation had been detected by the classic statistic methods.

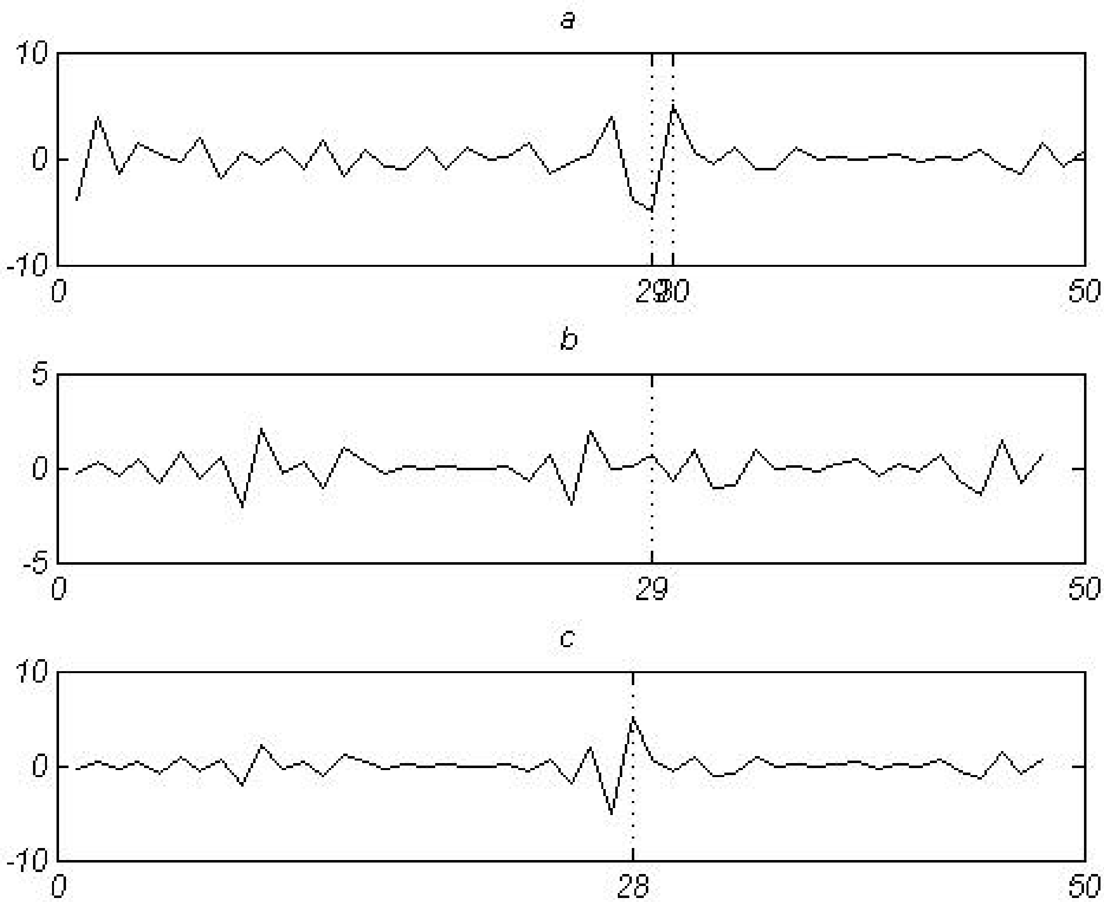

Fig 3a-c. The image of the fine part of the observation vector. a all the observations used and the largest picks can be seen at 29th and 30th elements. b the 30th element has been omitted and the corresponding pick was removed. c the 29th element has been omitted but the largest pick has not been removed

4. Conclusions and proposals

In succinct, to sum up the main points in a few words, we deduce this following statement:

- The wavelet transform can be used to analyze a given observation vector to investigate the probable gross errors or blunders as an alternative method instead of conventional statistical methods. The main advantage of the wavelet based method is its identicalness which has been the weakness of statistical approaches.
- Haar wavelet system used makes an ambiguity, in which the amplitude of the corresponding detail vector element, before or next to the blunder will be as large as the amplitude of the blunder’s fine part. This problem was solved by omitting the two adjacent elements, one by one and comparing their related details.

The above ambiguity is inherently related to the base functions used at the wavelet system. Further researches may suggest another famous wavelet system or design a
suitable one, by which the blunders would be clearly distinguished between the fine vector elements.

5. References

1. Barnett, V. and Lewis, T., 1984. Outliers in Statistical Data. John Wiley & Sons, New York
2. Keller, W., 2000. Lecture notes on wavelets. University of Stuttgart, Geodetic Institute, Department of Geodesy and Geoinformatics
3. Mallat, SG., 1989. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Patt, Anal Mach Intell II 7, 674–693
4. Zuofa, Li., 1996. Multiresolution approximation of the gravity field. Journal of Geodesy 70, 731-739
5. Barnett, V. and Lewis, T.: 1984, Outliers in Statistical Data, John Wiley & Sons, New York