DETECTION OF DISCONTINUITIES IN THE HEIGHT COMPONENT OF GNSS TIME SERIES

Jacek RAPINSKI [1] * and Kamil KOWALCZYK [2]

[1] *Institute of Geodesy, University of Warmia and Mazury in Olsztyn, Oczapowskiego 2, 10-719 Olsztyn, Poland*
[2] *Department of Land Surveying and Geomatics, University of Warmia and Mazury in Olsztyn, Heweliusza 12, 10-724 Olsztyn, Poland*

Corresponding author's e-mail: jacek.rapinski@uwm.edu.pl

ABSTRACT

This paper describes the algorithm to be used in the estimation in the vertical crustal movement speed estimation. A GNSS height between two permanent stations calculated daily are used as input time series. The main goal of this algorithm is to estimate the value of the rate of change of height difference between two stations despite the discontinuities in time and in the height data ("jumps"). Presented algorithm was implemented and tested on real data. The algorithm is detecting the discontinuities and determines the value of "jumps". Article presents the algorithm, data used for testing and results of processing.

1. INTRODUCTION

Vertical crustal movements are an important research topic used in many fields. In Europe the new European Vertical Reference Frame 2007 (EVRF 2007) was introduced on the basis of the European Vertical Reference System 2007 (EVRS 2007). This system is a kinematic height reference system (Sacher et al., 2009). Leveling data from 27 European countries were unified to epoch 2000 using the "official" NKG land uplift model computed by the Working Group for height determination vertical crustal movements model NKG2005LU (Agren and Svensson, 2007). Both the system and the frame are constantly modified due to the updates in the leveling data (Liebsch et al., 2014). The update of leveling data is a long time process. The leveling campaigns are conducted at average interval of 20 years or in shorter intervals as supplementary leveling (Kowalczyk, 2008; Kowalczyk and Rapinski, 2013). This makes the development of local and global models difficult (Kowalczyk et al., 2010). A solution to this problem may be the use of GNSS data from permanent stations. This is a major topic of the EUREF committee (EUREF symposium Budapest 2013, EUREF symposium Vilnius 2014).

The availability of permanent GNSS networks gives the possibility to compute the height differences between many points in regular time intervals. The problem is the accuracy of this kind of measurements. The accuracy between two consecutive epochs is lower than the one required to compute the movement. However the data from permanent GNSS stations are nowadays available for a period of few years. This is enough to allow the computation of the vertical movement using daily height differences as time series (Figurski et al., 2014). The motivation for this article was taken mostly from a previous work on vertical crustal movements estimated on the basis of GNSS time series. This research shows that there are outliers and "steps" in the raw results of GNSS data processing (Rapinski, 2014; Kowalczyk et al., 2010; Kowalczyk and Rapinski, 2013).

The discontinuities in time series are caused by three types of reasons: technical, human activity and environmental (for example: brakes in station operation, change of antennas, change in stabilization, software and receiver updates, changes in troposphere and ionosphere, change of reference frames, station subsidence, tectonic movements, movements caused by human activity). Discontinuities in the time series can be defined as:

- discontinuity in time – lack of data for some epochs;

- discontinuity in height difference data ("jumps") – quasi constant change of height difference in subsequent epochs in relation to previous epochs.

The discontinuities can be identified manually (Völksen and Hackl, 2012) or the process can be automated.
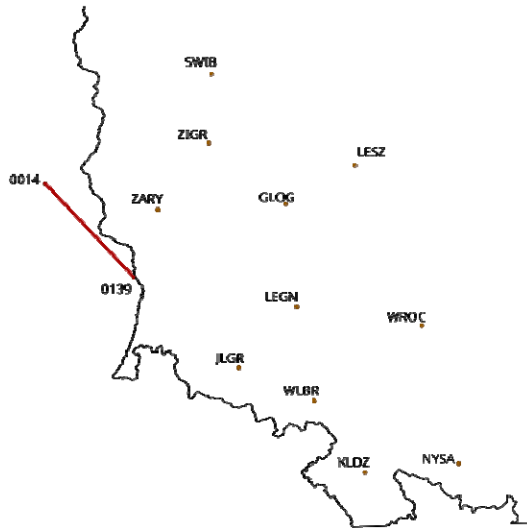
Fig. 1 Location of 0014 (Cottbus) and 0139 (Rothenburg) stations.

The main goal of our research was to calculate the parameters of the change of height differences between two GNSS stations despite the "steps" in the data. This problem was investigated by other authors in recent years which proves it is present (Argialas and Mavrantza, 2004; Roggero, 2012).

## 2. DATA USED IN THE ALGORITHM VERIFICATION

In order to verify the correctness of the proposed methodology, an example was selected from the data provided by the German Satellitenpositronierungsdienst

Table 1 Data statistics.

Number of data points	1091
Minimum	$-80.6123m$
Maximum	$-80.5774m$
Mean	$-80.59m$
Standard deviation	$0.0069m$
Variance	$4.8210^{-5}m^2$
Skewness	-0.51
Kurtosis	-0.90

(*SAPOS*) system. Height differences were computed daily from 0014 (Cottbus) - 0139 (Rothenburg) permanent GNSS stations. These two stations are located close to the southern part of Polish-German border. The location of these stations is depicted in Figure 1. The character of height differences data is described in Table 1.

Figure 2 depicts the data and a straight line fitted using least squares method. Looking at the data one can notice a significant "step" in the height differences around epoch 2011. The most possible cause of this was some modification of hardware or software on one of the stations. This step is causing an obvious error in the estimation of the vertical crustal movement, which can be derived from the plot as a tangent of the regression line. This effect is unwanted and is introducing outliers into vertical crustal movement network adjustments. In the data presented in this article, there is only one clearly visible "step", thus there can be more in case of other GNSS vectors.
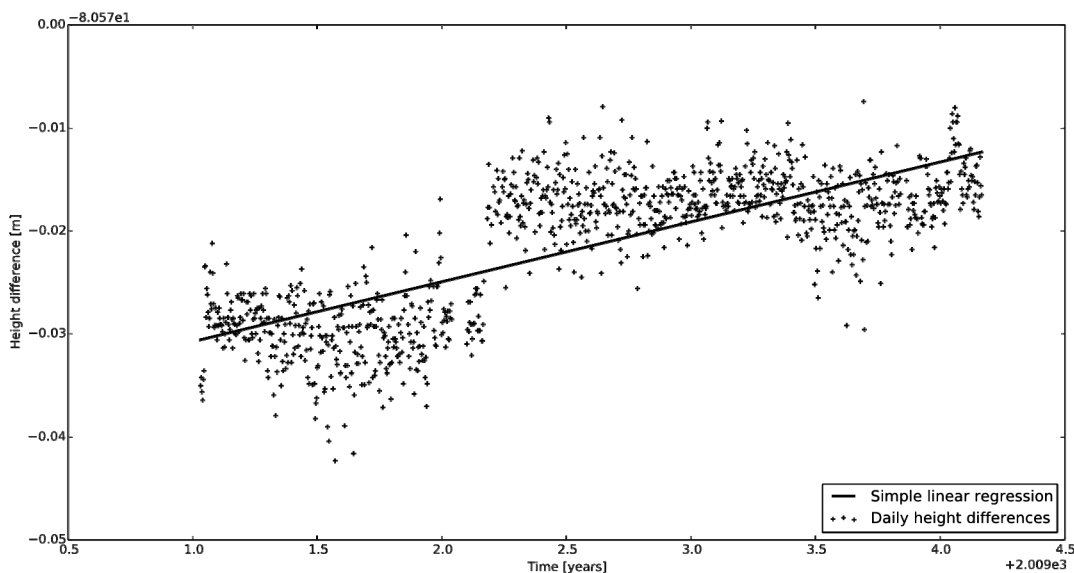
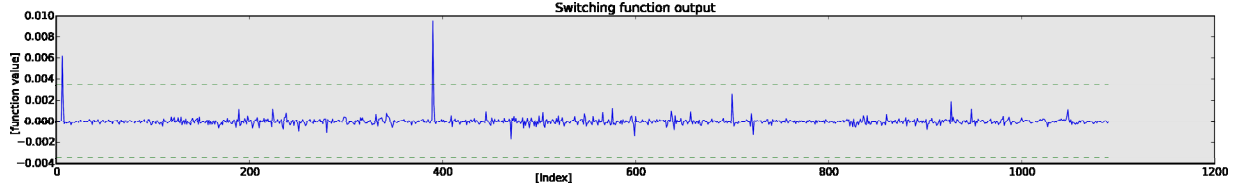**Fig. 2** Time series used in this article and a line fitted using least squares.

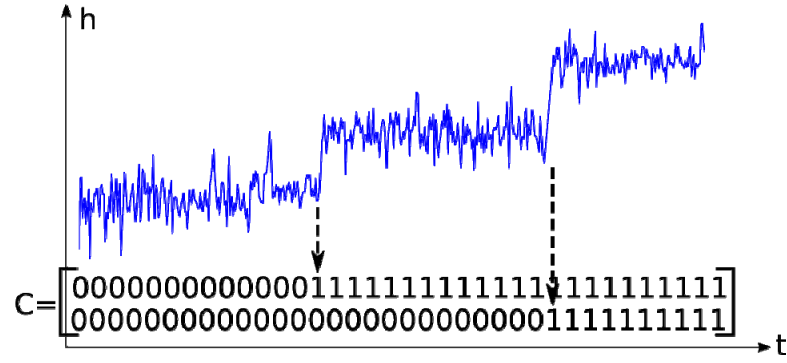**Fig. 5**   The output function of switching edge detector.

Fig. 6   The design of the matrix C.

There are a lot of approaches to resolve this issue. Most of them assume validation of the data and removing of those "steps" prior to the calculation of the movement. In our approach, we are trying to calculate the rate of movement and size of the "steps" simultaneously in one adjustment (Smith, 1998; Argialas and Mavrantza, 2004). The algorithm can be divided in two parts:

- Detection of the epochs in which "step" occurs. (Switching Edge Detection)

- Estimation of a vertical movement parameters and a magnitude of "steps". (Preparing the model and Least Squares estimation of the parameters)

3.   SWITCHING EDGE DETECTOR

To detect places in which "steps" occurr, the switching edge detector algorithm was used (Smith, 1998). Assuming a time series h_i defined at time point t_i, moving averages are constructed using the n-point window (equation 1). There are two moving averages first one from point k-n to n (denoted with - in the lower index) and from k to $k+n$ (denoted with + in the lower index). The same convention of naming is used in this section for all of the equations. Resulting moving averages for a one month window are presented in Figure 3.

$$H_{i\pm} = \frac{\sum_{k=1}^{n} h_{i\mp k}}{n} \tag{1}$$

The next step is to construct moving variances (1) on the basis of the moving averages (1). Resulting function is depicted in Figure 4.

$$S_{i\pm} = \frac{\sum_{k=1}^{n} \left(h_{i\mp k} - H_{i\pm}\right)^2}{n} \tag{2}$$

The output function of the switching edge detector is defined as:

$$H_{out}\left(t_i\right) = g_{i+}H_{i+} + g_{i-}H_{i-} \tag{3}$$

where g_{i-} and g_{i+} are the switching factors defined as:

$$g_{i+} = \frac{S_{i-}^{2r}}{S_{i+}^{2r} + S_{i-}^{2r}} \qquad g_{i-} = \frac{S_{i+}^{2r}}{S_{i+}^{2r} + S_{i-}^{2r}} \tag{4}$$

Figure 5 depicts the shape of the switching edge detection output function. Dashed line represents a certain threshold defined for the particular dataset. When output function crosses this threshold, it indicates the "step" in the data. In the presented example the threshold was set to the $\pm$ half of the standard deviation of the data set.

Switching edge detector is resulting in a C matrix containing zeroes and ones.

The matrix C has as many columns as there are epochs in the data set and as many rows as the number of detected "steps". To obtain matrix C the algorithm iterates over each epoch of the output function. If the value of output function is grater then the threshold, one row is added to the matrix C with zeroes in the columns before this epoch and ones from this epoch to the last epoch. The example of this approach is depicted in Figure 6. This example is taken from a different data set then presented in the rest of the article just for a better visualization of the idea.

In case of the example data (0014–0139) two "steps" were detected – in epoch number 6 (year 2010.04795) and 390 (year 2011.18219). Resulting matrix $\underline{C}$ has two rows: first one with zeros from epoch 0 to 5 and ones form epoch 6 to the last one. In the second row there are zeroes from epoch 0 to 389 and ones from 390 to the end.

4. MATHEMATICAL MODEL

The mathematical model used in this article is a straight line with "steps" in particular epochs:

$$h(t) = vt + h_0 + c_1 s_1 + c_2 s_2 + \cdots + c_m s_m \tag{5}$$

where:
t = epoch, $h(t)$ = height difference in epoch t, v = velocity between stations
h_0 = height difference at epoch 0
$c_1, c_2, ..., c_m$ = elements from matrix $\underline{C}$
$s_1, s_2, ..., s_m$ = magnitude of "jumps". So the equation of residuals is:

$$v_t = vt + h_0 + c_1 s_1 + c_2 s_2 + \cdots + c_m s_m - h(t) \tag{6}$$

where v_t is the residuum in epoch t. The model of observations without "steps" can be written in matrix notation as:

$$\underline{V} = \underline{A}\underline{X} + \underline{L} \tag{7}$$

The standard least squares method with the following objective function can be then used to estimate the parameters of the model (5):

$$\underline{V}^T \underline{V} = \min \tag{8}$$

If no "steps" are detected the design matrix will take the following form:

$$\underline{A_0} = \begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots \\ t_n & 1 \end{bmatrix} \tag{9}$$

In the presence of "steps" matrix $\underline{A_0}$ must be concatenated with matrix $\underline{C}^T$:

$$\underline{A} = \begin{bmatrix} \underline{A_0} | \underline{C}^T \end{bmatrix} \tag{10}$$

Thus design matrix and the matrix of unknown parameters will take the forms:

$$\underline{A} = \begin{bmatrix} t_1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots \\ t_k & 1 & 1 & 0 & \cdots & 0 \\ \vdots \\ t_n & 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \quad \underline{X} = \begin{bmatrix} v \\ h \\ s_1 \\ s_2 \\ \vdots \\ s_m \end{bmatrix} \tag{11}$$

Then the set of equations (7) is solved using least squares method resulting in estimated values of unknown parameters $h(t), v, h_0, s_1, s_2, ..., s_m$.

Table 2 Estimated parameters.

	v	h_0	s_1	s_2
	mm/y	m	Mm	mm
estimated value	0.3	−80.6049	5.0	12.4
$\sigma 2$	0.6	0.004	4.0	1.1

5. RESULTS

The results of the application of the algorithm presented in previous sections are depicted in Figure 7 and summarized in Table 2.

The value of the first "jump" was estimated with relatively low accuracy. This was caused by the location of this "step" at the beginning of the time series (small number of observations before the "jump"). The second "jump" was estimated with satisfactory accuracy. The accuracy of the vertical movement was estimated on the level of 0.6 mm. This was expected according to the calculations presented in (Kowalczyk, 2015).

Many GNSS time series are burdened with a certain periodic vertical movement component (caused mostly by seasonal changes). Proposed methodology can successfully detect "jumps" in the presence of periodic components. An example is presented in Figure 8.

Application of this algorithm to such data, requires more attention from the user. The parameters of switching edge detector (size of the moving window and threshold) must be selected with care in order to avoid detection of too many "jumps".

6. SUMMARY

The article presents the algorithm to estimate the rate of change of height differences between two GNSS stations despite the "steps" in the data. To complete this task it introduces the switching edge detection algorithm combined with least squares solution. The algorithm was successful in detecting jumps in the height difference time series. The size of the jumps is estimated together with rates of changes and initial height difference in a least squares adjustment. The performance of this approach is shown using the real life example. This method will be tested on the entire network of the GNSS stations in Poland.

There is a lot of place for further research in this topic. One of them is to introduce a more complex model of observations including acceleration or seasonal oscillations in the data or use of more sophisticated method for fitting the model into the data.

REFERENCES

Agren, R. and Svensson, J.: 2007, Postglacial land uplift model and system definition for the new Swedish height system RH 2000. Technical report, Forfattare Jonas AgrenochRunarSvensson, Gavle.

Argialas, D.P. and Mavrantza, O.D.: 2004, Comparison of edge detection and hough transform techniques in the extraction of geologic features. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 34, 790–795.

Figurski, M., Kowalczyk, K. and Bogusz, J.: 2014, On the possibility of using GNSS data to model the vertical crustal movements. 14th SGEM GeoConference on Informatics, Geoinformatics and Remote Sensing, 2, 567–574. DOI: 10.5593/SGEM2014/B22/S9.071

Kowalczyk, K.: 2008, Vertical crustal movements in Poland for instance any fragment three levellings network. 7 th International Conference on Environmental Engineering, 3, 1354–1358.

Kowalczyk, K.: 2015, The creation of a model of relative vertical crustal movements in the Polish territory on the basis of the data from active geodetic network EUPOS (ASG EUPOS). Acta Geodyn. Geomater, 12, No. 3 (179), 215–225.
DOI: 10.13168/AGG.2015.0022

Kowalczyk, K. and Rapinski, J.: 2013, Evaluation of levelling data for use in vertical crustal movements model in Poland. Acta Geodyn. Geomater., 10, 4 (172), 401–410. DOI: 10.13168/AGG.2013.0039

Kowalczyk, K., Rapinski, J. and Mroz, M.: 2010, Analysis of vertical movements modelling through various interpolation techniques. Acta Geodyn. Geomater, 7, No. 4 (160), 399–409.

Liebsch, G., Rülke, A., Sacher, M. and Ihde, J.: 2014, Definition and realization of the EVRS: How do we want to proceed ? EUREF Symposium 2014, Vilnius, Lithunia.

Rapinski, J.: 2014, Robust adjustment of vertical crustal movements in Poland, based on GNSS observations. 9th International Conference on Environmental Engineering, Vilnius, Lithuania, 1–5.

Roggero, M.: 2012, Discontinuity detection and removal from data time series. In: N. Sneeuw, P. Novak, M. Crespi and F. Sanso (Eds.), VII Hotine-Marussi Symposium on Mathematical Geodesy, Roma 2009, 137, Springer Berlin Heidelberg, 135–140.
DOI: 10.1007/978-3-642-22078-4_20

Sacher, M., Ihde,J., Liebsch, G. and Makinen, J.: 2009, EVRF2007 as realization of the European Vertical Reference System. Bollettino di Geodesia e Scienze Affini, 68, No. 1, 35–50.

Smith, D.: 1998, A quantitative method for the detection of edges in noisy time-series. Phil. Trans. R. Soc. Lond.. Series B, 353, 1969–1981.

Völksen, C. and Hackl, M.: 2012, Crustal deformation and seismicity in Southern Bavaria revealed by GNSS observations. EGU General Assembly 2012, Vienna, Austria, 8444.
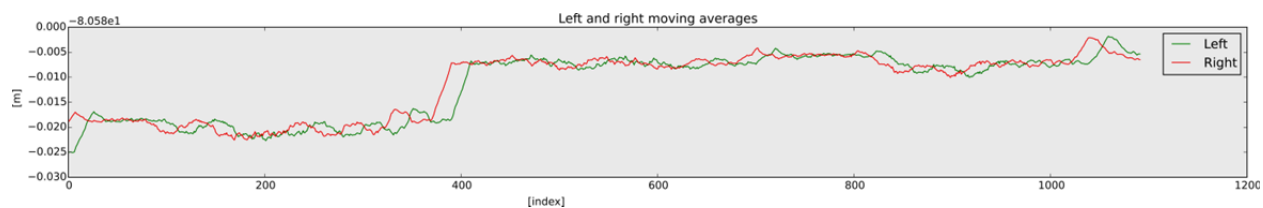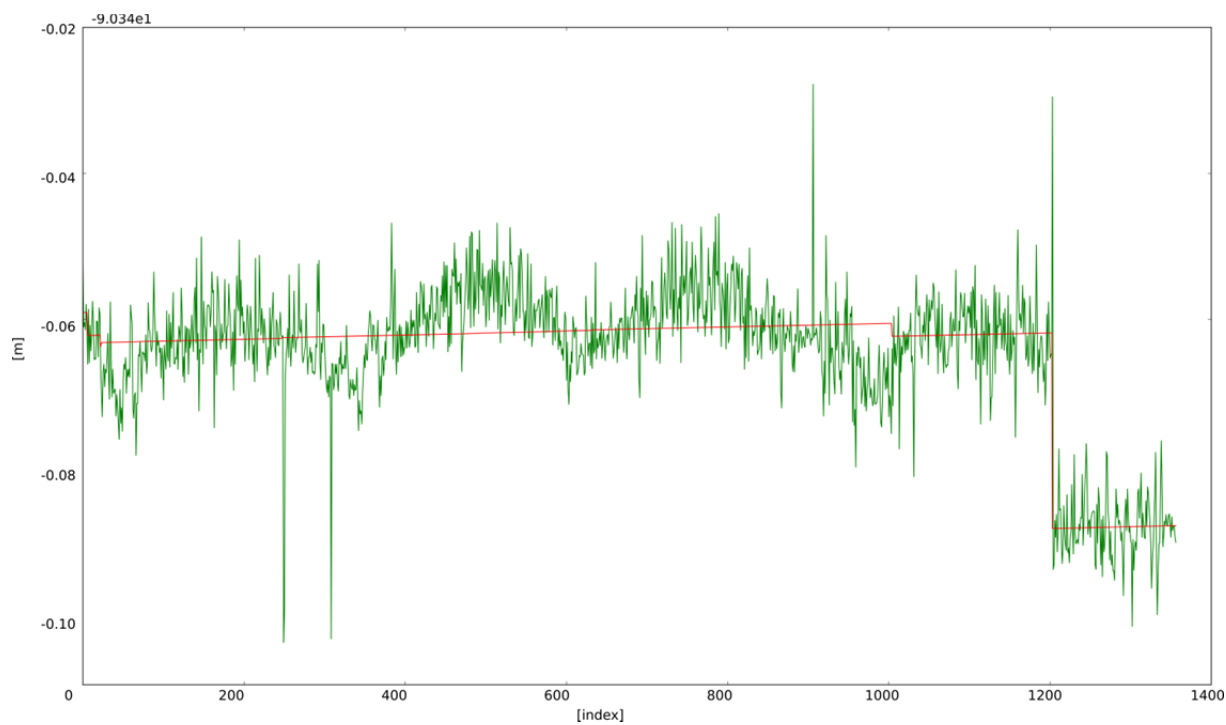
Fig. 3   Moving averages for one month window.
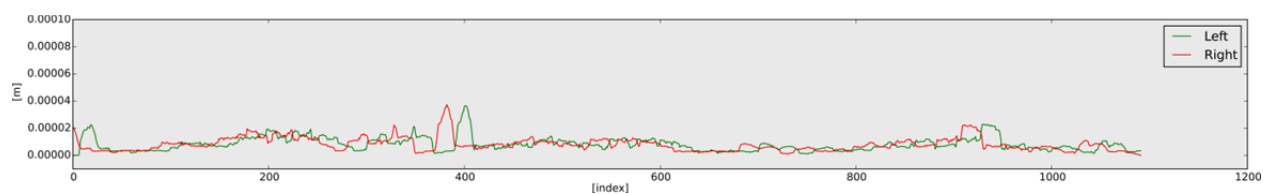
Fig. 4   Moving variances.
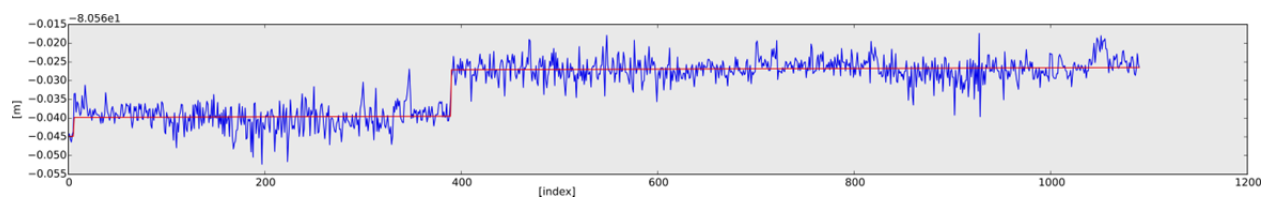
Fig. 7   The results of fitting model to data.

Fig. 8   The results of fitting linear model to data burdened with periodic component.