Detection of measurement errors in the time series of the coordinates of surface movement monitoring points with the use of satellite navigation systems

Zbigniew SIEJKA and Violetta SOKOŁA-SZEWIOŁA, Poland

Key words: GNSS monitoring, measurement errors, post mining, seismic activity

SUMMARY

Nowadays, we have a wide range of modern, internally very complex navigation satellite systems (GNSS), with which we can conduct precise geodetic observations. It should be noted, however, that these systems have some significant limitations that constitute the main barrier to their widespread use. As shown by previous studies in this field, time series of coordinate changes obtained on the basis of constantly operating GNSS stations are burdened with many effects, which leads to large discrepancies in observations, also for stations operating in relatively small areas. Changes in the coordinates of such stations are due to several main reasons: large-scale effects such as tectonic effects, strictly local effects such as the position of the observation antenna or anthropogenic factors, and the effects of systematic errors of the GNSS system. Therefore, the time series of coordinates, in addition to the useful measurement signal, includes systematic effects ("color" noise) and random measurement noise ("white" noise). Appropriate modeling of these effects allows for precise and reliable determination of changes in the coordinates of observation points. The paper deals with the problem of detecting outliers and random errors (“white” noise) in time series of coordinates generated at GNSS stations. The results of the continuous monitoring of the movements of the earth's surface carried out under the European research project * are presented. One of the aims of the project is to develop methods and plans for long-term monitoring of post-mining areas in order to reduce seismic hazards during and after the closure of coal mines. Daily solutions of high frequency (1 Hz), implemented in real time, in a continuous mode, were analyzed. The observations were carried out simultaneously at three checkpoints, which are monitoring stations of the automatic measurement and processing system of GNSS observations. On their basis, time series of coordinate changes were determined, for which systematic components, periodic components (fluctuations) and random components (noise) were calculated. Additionally, the mutual spatial-temporal correlation in the analyzed time series was examined. The results of the analyzes presented in the paper were used to develop the results of long-term measurements, which are the subject of the above mentioned project.

Detection of measurement errors in the time series of coordinates of the surface movement monitoring points with the use of satellite navigation systems

Zbigniew SIEJKA and Violetta SOKOŁA-SZEWIOŁA, Poland

1. INTRODUCTION

A time series is a series of numbers that define the dynamics of changes in the analyzed phenomenon, registered in successive moments of time (Zagdański and Suchwałko, 2016). In the time series analyzed in the study, it was assumed that they consist of a signal and uncorrelated random noise. The data used were real daily data of coordinate changes coming from three stations monitoring land surface movements in three dimensions (N, E, H), in the form of time series recorded with a 1-second interval. The stations are located in the area of the Upper Silesian Coal Basin "in the mining area of the closed" Kazimierz-Juliusz "mine, which is currently being flooded. They are part of the measuring module of the automatic surface movement monitoring system, developed for the implementation of a research project with the acronym PostMinQuake *, concerning the process of closing mines in seismically active areas. This module consists of the above-mentioned three monitoring stations and five reference stations, which constitute reference points for monitoring stations. They are located outside the impact area of mining exploitation in order to ensure the stability of the reference system. The sketch of the GNSS control and monitoring network at the Kazimierz-Juliusz facility and the location of the monitoring stations in the field is shown in Figure 1.

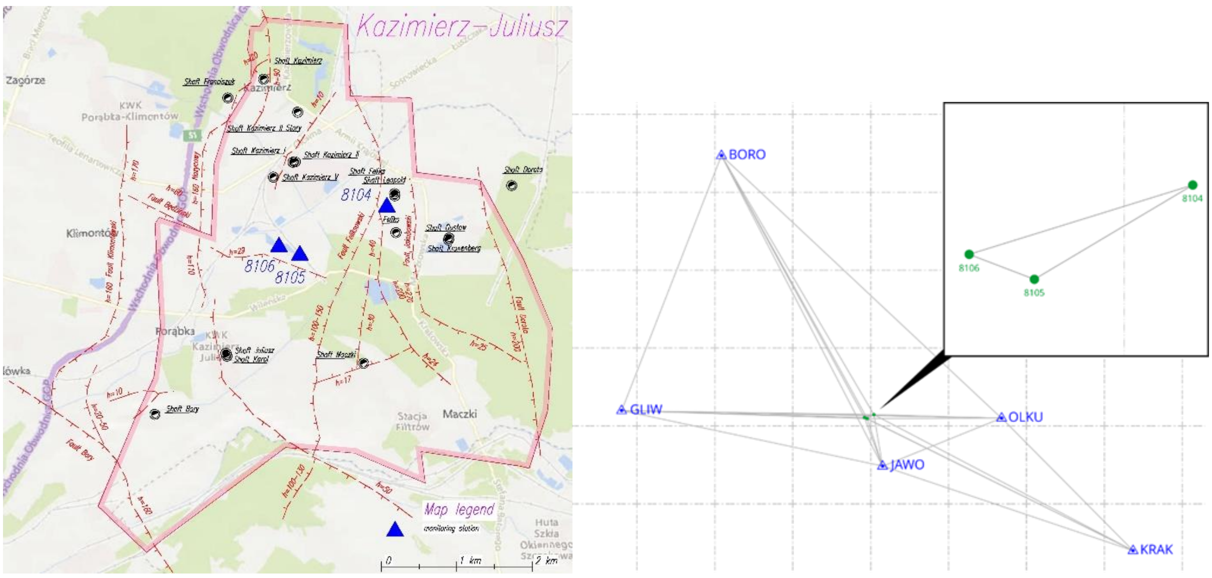

Fig. 1. Sketch of the GNSS control and monitoring network at the Kazimierz-Juliusz facility. and the location of the monitoring stations in the field

Time series of changes in the coordinates of monitoring stations were analyzed in terms of the occurrence of data discontinuities in time and the so-called jumps in the form of outliers. Discontinuities in time series are caused by three types of causes: technical, human activity and environmental causes (Kowalczyk 2015, Rapiński and Kowalczyk 2016). The most common causes include: interruptions in station operation, change of antennas, changes in stabilization, software and receiver update, changes in the troposphere and ionosphere, changes in reference systems and systems, station subsidence, tectonic movements, movements caused by human activity.

Discontinuities in time series can be defined as:

- discontinuity in time - no data for some epochs,
- discontinuity in the data of recorded coordinate changes (jumps) - in successive epochs in relation to previous epochs.

This subject was taken up by many authors from various research centers in the country and abroad, including (Kowalczyk and Rapiński 2017), who presented the methodology of aligning the network of vertical movements from data from permanent GNSS stations. The problem of detecting the detection of jumps in time series, based on the example of the study of long-term mountain uplift, was investigated by (Goudarzi M.A. et all. 2012). On the other hand, research was conducted on the use of MATLAB software for noise reduction in GNSS time series (He X., et all 2020). Detection of missing data in the time series of GNSS monitoring (Liu H. and Li L., 2022).

2. RESEARCH MATERIAL - AND RESEARCH METHODOLOGY

Since the commencement of large-scale operation of satellite measurement systems operating in a permanent manner, it has become possible to create ever longer time series of vertical and horizontal movements of the earth's crust on the basis of data from GNSS stations (Kontny and Bogusz 2012).

In this paper, high-frequency GNSS observations observed with a time resolution of 1 Hz were used as research material. Nine series of observations made on 3 control points covering the time period of 24h were used. At each point, 86,400 observations were recorded, which were summarized into time series showing the differences in 3D coordinates for individual components: north, east and altitude (ΔdN, ΔdE, ΔdH). The analyzed data covers the period from February 20, 2022 from 00:19:01 to February 21, 2022 to 00:19:00. For the original time series prepared in this way, first of all, the discontinuity of the measurement epochs of coordinate changes in the time series was identified as the lack of data in a set time interval.

For this purpose, two methods were used: visual (Fig. 2, 3, 4) and numerical. On their basis, no data discontinuities in the time series were found. On the other hand, the numerical method confirmed that each of the time series contained the complete number of observations (86,400) for the analyzed period of time.

In the second stage of the study, a visual analysis of the original time series in the charts was performed (Fig. 2, 3, 4). On this basis, the occurrence of certain cycles of periodic fluctuations in the analyzed measurements was found. In addition, basic descriptive statistics and parameters of the normal distribution were also determined, in the form of values: minimum (MIN), maximum (MAX), mean (AV), range (R), median (ME) and standard deviation (SD). Table 1 summarizes the empirical values of these parameters. It has been shown that the daily solutions from individual stations for certain time periods are characterized by high differentiation and the occurrence of outliers in time series.

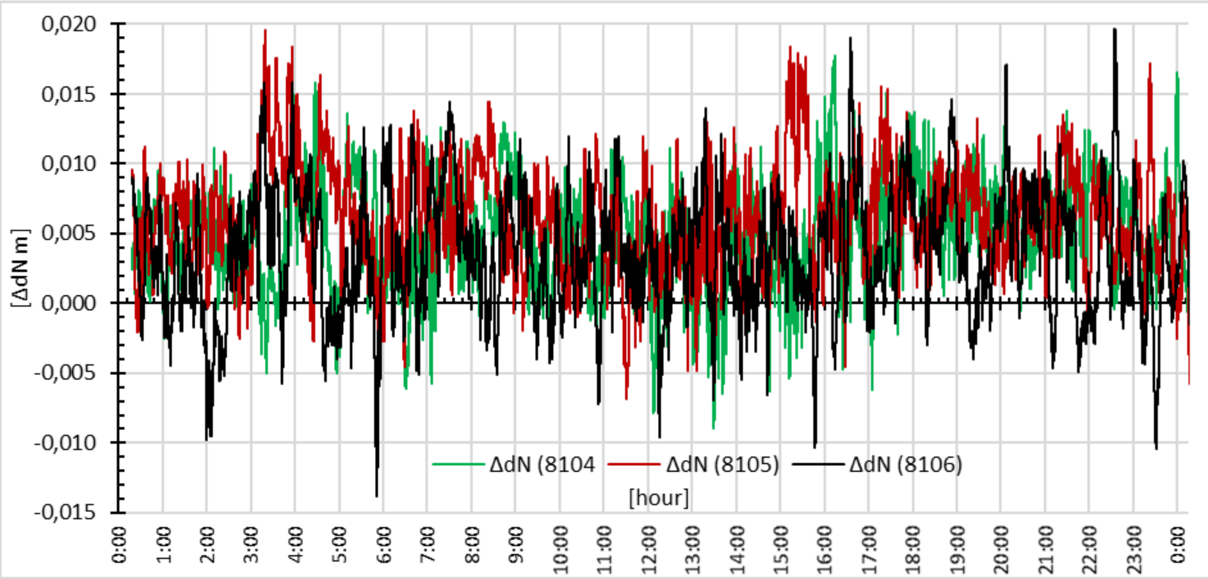

Fig. 2. Original time series (ΔdN) of changes in the northern component for points 8104, 8105, 8106.

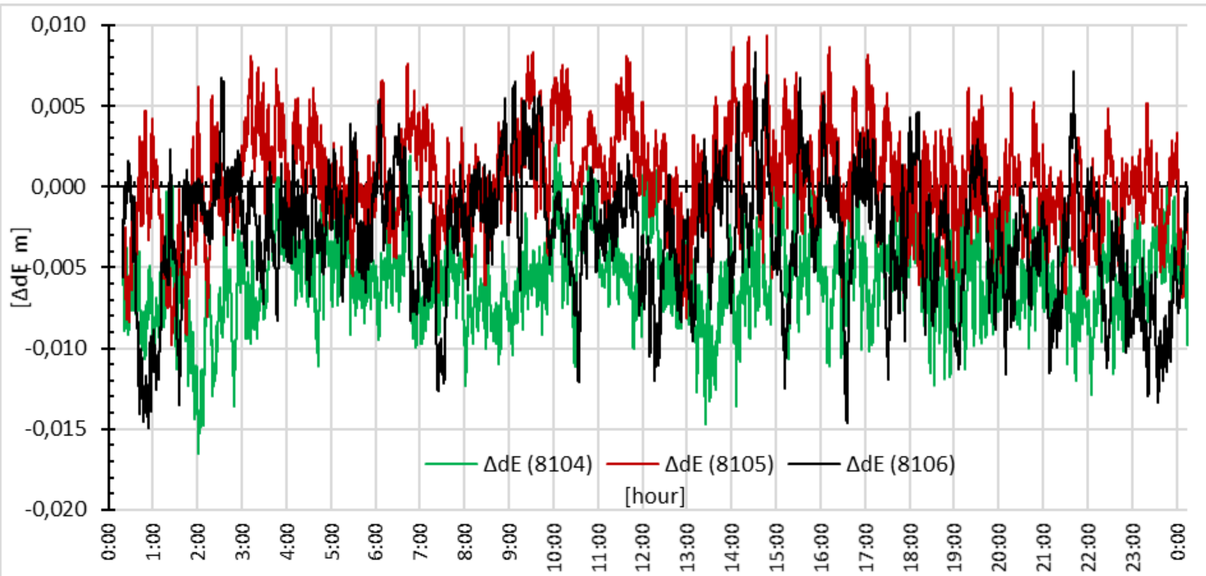

Fig. 3. Original time series (ΔdE) of changes in the eastern component for points 8104, 8105, 8106.

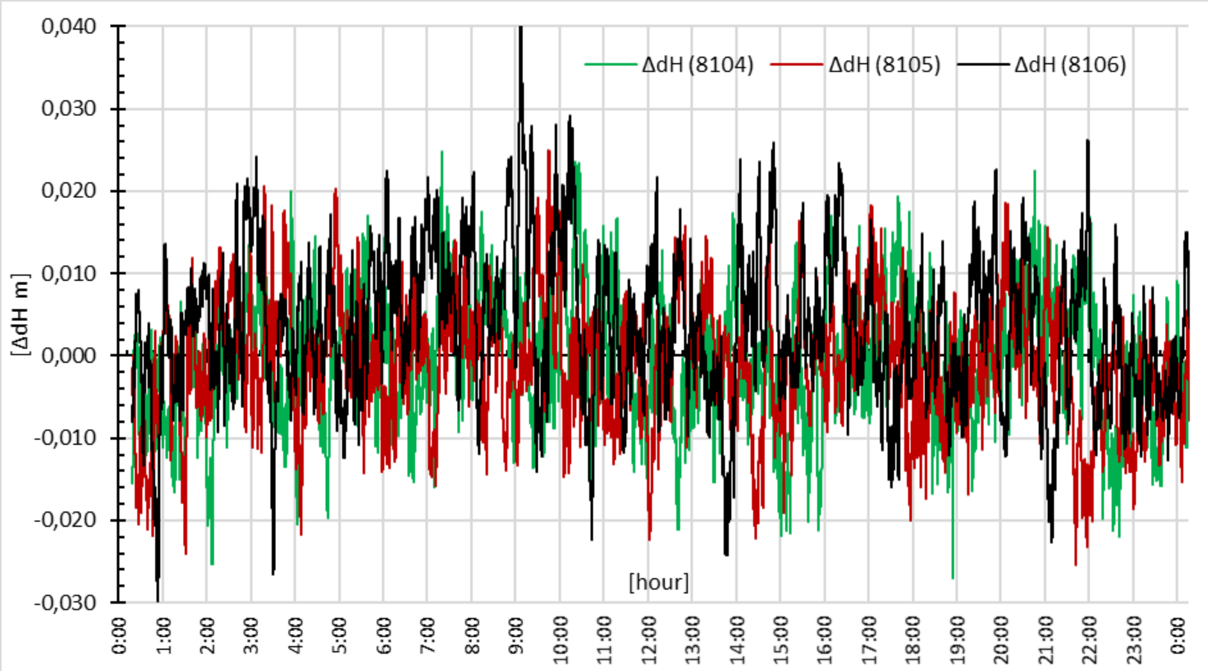

Fig. 4. Original time series (ΔdH) of changes in the eastern component for points 8104, 8105, 8106.

Table 1. Summary of the basic descriptive statistics of the original time series of changes in the coordinate components.

statistics	ΔdN [m]			ΔdE [m]			ΔdH [m]		
	8104	8105	8106	8104	8105	8106	8104	8105	8106
MIN	-0.0090	-0.0068	-0.0138	-0.0165	-0.0098	-0.0149	-0.0270	-0.0253	-0.0297
MAX	0.0178	0.0195	0.0197	0.0026	0.0094	0.0083	0.0248	0.0250	0.0411
R	0.0268	0.0264	0.0335	0.0191	0.0191	0.0232	0.0517	0.0503	0.0708
AV	0.0048	0.0061	0.0036	-0.0056	0.0005	-0.0030	-0.0002	-0.0009	0.0034
ME	0.0049	0.0061	0.0035	-0.0056	0.0004	-0.0025	-0.0001	-0.0007	0.0031
SD	0.0037	0.0038	0.0045	0.0027	0.0031	0.0039	0.0077	0.0075	0.0091

The next stage of the time series analysis included the filtration of the obtained observational data, the aim of which was, among others, the reduction of certain outliers in time, the occurrence of which was found visually in the diagrams, and empirically confirmed in Table 1, using the descriptive statistics provided. In order to eliminate outliers, the 3 sigm rule was used, according to which, in the case of a normal or close to normal distribution, only a small number of all observations exceed the range $(\bar{x} - 3\sigma, \bar{x} + 3\sigma)$.In the analyzed time series, it was shown that the number of outliers defined according to this principle ranged from 0.02% to 0.79% (Table 2). In line with this rule, outliers were taken as observations that did not fall within the limit of 3 standard deviations from the mean. These observations were eliminated from the analyzed time series and temporarily replaced with a special no data code, and then the no data codes were replaced with the simple arithmetic mean based on n = 10 observations preceding no data.

In the next stage of the study, a moving average was used as a method of time series equalization. In the paper, time series equalization was used with the use of a simple moving average, calculated from an odd number of adjacent words in the series. This approach allows the obtained equalization result to be assigned to the value of t located in the middle of the time interval considered in the calculations. If the number of original terms of the series used to calculate the mean is $2q$ + 1 (where q is any natural number), then the mean is calculated from the formula:

$$\overline{y_t} = \frac{1}{2q+1} \sum_{r=-q}^{q} y_{t+r}$$

where:

$t = q, q+1, q+2, \ldots, n-q$

$\overline{y_t}$ – smoothed series value for time t

Table 2. Summary of outliers for individual time series.

Number of observations	ΔdN [m]			ΔdE [m]			ΔdH [m]		
	8104	**8105**	**8106**	**8104**	**8105**	**8106**	**8104**	**8105**	**8106**
All	86400	86400	86400	86400	86400	86400	86400	86400	86400
removed	519	438	610	488	151	18	228	140	686
[%] removed	**0.60 [%]**	**0.51 [%]**	**0.71 [%]**	**0.56 [%]**	**0.17 [%]**	**0.02 [%]**	**0.26 [%]**	**0.16 [%]**	**0.79 [%]**

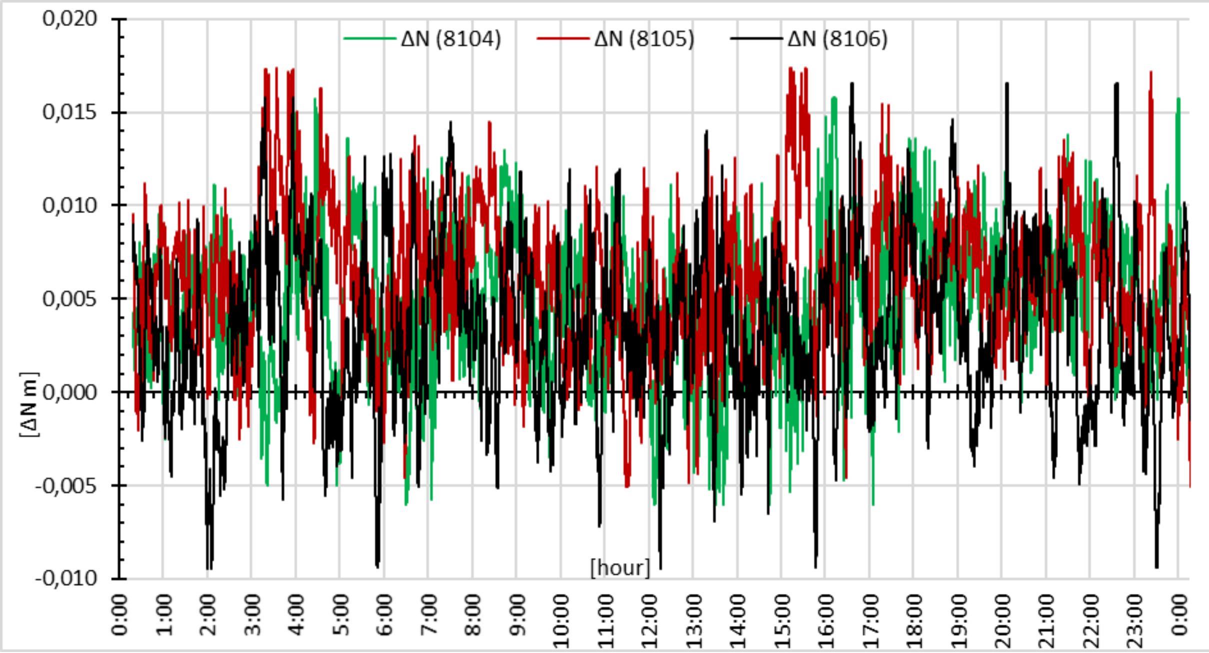

Fig. 5. Transformed time series (ΔN) of changes in the northern component for points 8104, 8105, 8106

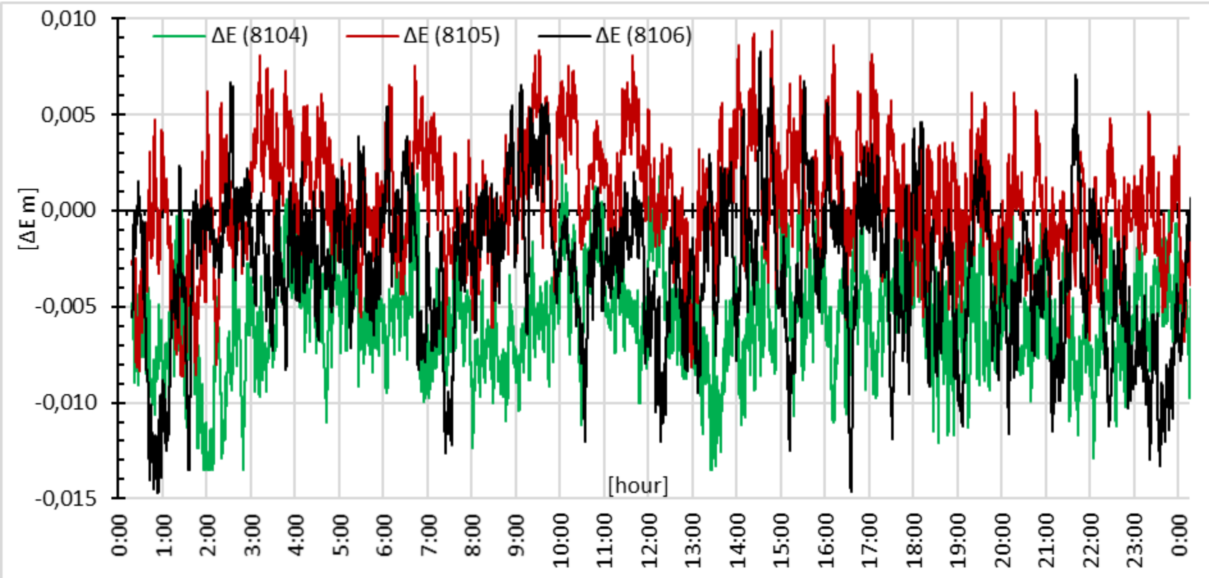

Fig. 6. Transformed time series (ΔE) of changes in the eastern component for points 8104, 8105, 8106.

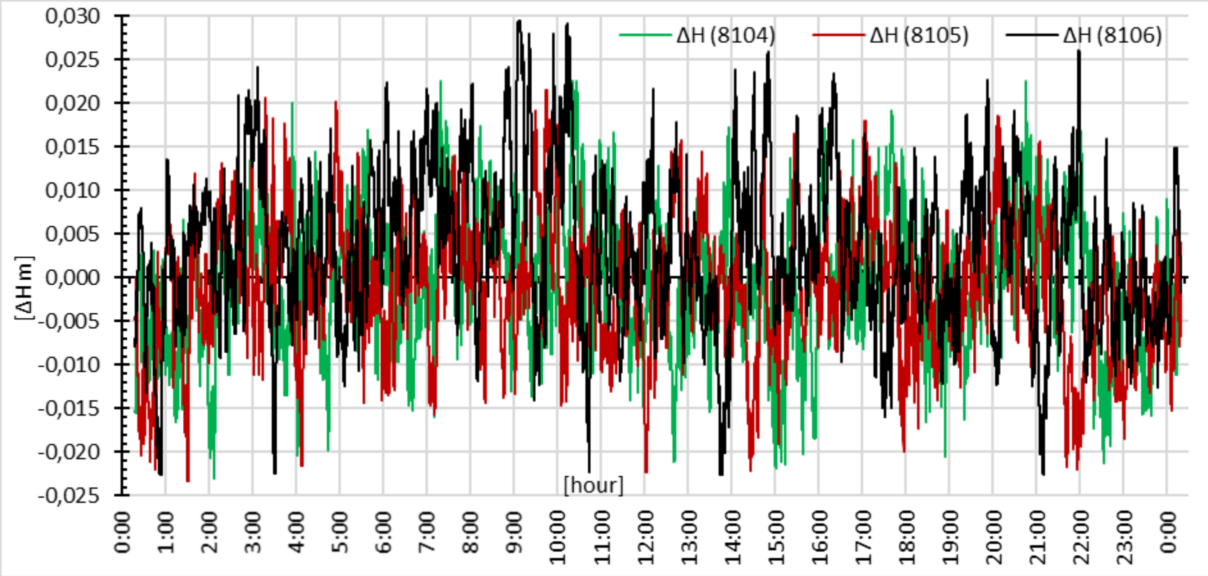

Fig. 7. Transformed time series (ΔH) of changes in the height component for points 8104, 8105, 8106.

Table 3. Summary of basic descriptive statistics of aligned time series, changes in coordinate components

statistics	ΔdN [m]			ΔdE [m]			ΔdH [m]		
	8104	**8105**	**8106**	**8104**	**8105**	**8106**	**8104**	**8105**	**8106**
MIN	-0.0061	-0.0051	-0.0094	-0.0135	-0.0086	-0.0147	-0.0230	-0.0234	-0.0226
MAX	0.0158	0.0173	0.0166	0.0024	0.0094	0.0083	0.0226	0.0215	0.0295
R	**0.0218**	**0.0224**	**0.0260**	**0.0159**	**0.0180**	**0.0230**	**0.0456**	**0.0449**	**0.0522**
AV	**0.0048**	**0.0061**	**0.0036**	**-0.0056**	**0.0005**	**-0.0030**	**-0.0002**	**-0.0009**	**0.0034**
ME	**0.0049**	**0.0061**	**0.0035**	**-0.0056**	**0.0004**	**-0.0025**	**-0.0001**	**-0.0007**	**0.0031**
SD	**0.0037**	**0.0038**	**0.0044**	**0.0027**	**0.0031**	**0.0039**	**0.0077**	**0.0075**	**0.0090**

3. DISCUSSION AND CONCLUSIONS

The analysis of time series assumes that the original observations contain a factor describing the studied phenomenon and a measurement error that significantly hinders the identification of its model. The condition for the correct identification of the studied phenomenon is the correct analysis of errors in the sequence and its equalization. The applied approach to the time series equalization allows for the elimination of random fluctuations as well as periodic fluctuations from the series. Adjustment based on a moving average from more words smoother the time series. On the other hand, the alignment of a smaller number of words better reflects the changes taking place in relation to the values of the original variables, but the random fluctuations of the variable have a greater impact on the smoothed series.

The applied methodology allowed for an improvement of about 22% in the precision of determination of observation time series by removing outliers and leveling the series. At the same time, the empirical standard deviations (SD) for individual components of the time series, determined in Tables 1 and 2, characterizing the average accuracy of the determined position, remained at the same level, which means that the equalization did not significantly change the characteristics of the original observation series, despite the fact that in set 2 outliers were removed.
This approach, used in the process of developing the results of observations carried out as part of the research project in question, allows the smoothing of the time series without interfering with its internal nature and structure.

The presented results are the subject of the European research project - PostMinQuake: "The project leading to this application has received funding from the Research Fund for Coal and Steel under grant agreement No 899192". "Scientific work published as part of an international project co-financed by the program of the Minister and Science of Higher Education entitled "PMW" PostMinQuake financed by the Coal and Steel fund co-financed in the Silesian University of Technology, by the Ministry of Science and Higher Education 5124 / FBWiS / 2020/2 for the implementation of a co-financed international project No. W 62 / FBWiS / 2020

REFERENCES

Goudarzi M. A., Cocard M., Santerre R.; Woldai T., 2012, GPS interactive time series analysis software. Springer-Verlag Berlin Heidelberg.

He X., Yu K, Montillet J.P., Xiong Ch., Lu T., Zhou S., Ma X., Cui H., Ming F.,. 2020. GNSS-TS-NRS: An Open-Source MATLAB-Based GNSS Time Series Noise Reduction Software. Remote Sensing 2022, 12, 3532; doi:10.3390/rs12213532

Kontny, B., Bogusz, J., 2012.,Models of vertical movements of the earth crust surface in the area of Poland derived from levelling and GNSS data. Acta Geodynamica et Geomaterialia, 9(3), 167.

Kowalczyk K., 2015, The creation of a model of relative vertical crustal movement In the Polish territory on the basis of the data from Active Geodetic Network EUPOS (ASG EUPOS). Acta Geodyn. Geomater., Vol. 12, No. 3 (179), pp. 215–225, DOI: 10.13168/AGG.2015.00220039

Krasuski K., Wierzbicki D., 2020, Monitoring Aircraft Position Using EGNOS Data for the SBAS APV Approach to the Landing Procedure.

Liu H. and Li L., 2022, Missing Data Imputation in GNSS Monitoring Time Series Using Temporal and Spatial Hankel Matrix Factorization.

Rapiński J., Kowalczyk K., 2016, Detection of discontinuities in the height component of GNSS time series. Acta Geodyn. Geomater., Vol. 13, No. 3(183), pp. 315-320. DOI: 10.13168/AGG.2016.0013

Rapiński J., Kowalczyk K., 2016, Detection of discontinuities in the height component of GNSS time series. Acta Geodyn. Geomater., Vol. 13, No. 3(183), pp. 315-320. DOI: 10.13168/AGG.2016.0013

Tucikešić S., Blagojević D., 2019, Modelling of the time-series of GNSS coordinates and their interaction with average magnitude earthquakes. Geodetski Vestnik 63 (01): 525-540, DOI: 10.15292/geodetski-vestnik.2019.04.525-540

Zagdański, A., i Suchwałko, A., 2016, Analiza i prognozowanie szeregów czasowych. Praktyczne wprowadzenie na postawie środowiska. Wydawnictwo Naukowe PWN SA

BIOGRAPHICAL NOTES

Zbigniew Siejka, PhD, DSc, Eng - research worker at the Department of Geodesy, University of Agriculture and Silesian University of Technology in Gliwice. His research interests include accuracy and reliability of positioning using GNSS (Global Navigation Satellite System) in geodesy and related disciplines. He also has experience in displacement monitoring using GNSS. He is active in the professional field as a consultant and expert on the use of satellite positioning systems in surveying practice. He is an author and co-author of numerous national and international publications on the use of modern measurement techniques in Geodesy.

Violetta Sokoła-Szewioła, PhD, DSc, Eng, Prof. SUT- Violetta - research worker at the Silesian University of Technology in Gliwice. Head of the Team of Cartography and Spatial Information Systems. Member of the Association of Polish Surveyors (SGP), President of the SGP Branch in Katowice, board member of the Polish ISM committee. In her scientific research she deals with issues related to the observation of land surface movements using modern measurement techniques, in particular in mining and post-mining areas. Author of over 100 publications of national and international range. Author or co-author of several patent applications.

CONTACTS

PhD, DSc, Eng, Zbigniew Siejka
University of Agriculture in Krakow
Al. A. Mickiewicza 21,
31-120 Krakow,
POLAND
Tel: +48 12 662 45 39
zbigniew.siejka@urk.edu.pl

Prof. SUT, PhD, DSc, Eng, Violetta Sokoła-Szewioła
Silesian University of Technology
Akademicka 2,
Gliwice,
POLAND,
Tel: +48 32 347 29 90
violetta.sokola-szewiola@polsl.pl