

ESTIMATION OF SUB-DIURNAL NOISE LEVEL IN GPS TIME SERIES

Janusz BOGUSZ ¹⁾ and Bernard KONTNY ^{2)*}¹⁾ Centre of Applied Geomatics, Military University of Technology²⁾ Institute of Geodesy and Geoinformatics, Wrocław University of Environmental and Life Sciences, Poland

*Corresponding author's e-mail: bernard.kontny@igig.up.wroc.pl

(Received January 2011, accepted April 2011)

ABSTRACT

This paper contains the studies of the noise level in GPS (Global Positioning System) time series. As the data for this research the authors used changes of the geodetic coordinates of the ASG-EUPOS and associated (Slovak and Czech) sites. The method of precise GPS observations processing in short-time intervals was worked out in the Centre of Applied Geomatics, Military University of Technology. The authors focus on the diurnal and sub-diurnal frequency bands (tidal) as the tidal effects influence the sites' positions at the most and the model used in the standard processing software does not contain geodetic coefficients. Thus the residual values of the geodetic coordinates time series should hold some information in the shape of coloured noise. The paper comprise short description of the Earth tides phenomenon, the concept of the tidal parameters determination, character of the white and coloured noise and the assessment of the noise estimation from GPS data. As the processed network contains 130 sites the spatial distribution of the noise's parameters is also investigated.

KEYWORDS: GPS, short-time solutions, noise, ASG-EUPOS**INTRODUCTION**

The GNSS (Global Navigation Satellite System) residual time series after filtration (trend, periodical components removal) should in theory have the character of white noise. To prove this thesis the analyses of residuals received from two independent models (deterministic tidal HW95 and stochastic ARIMA) were performed. For comparison the white noise of variance equal to the variance from tidal model was distinguished. The white noise characterizes the complete independence of observation (the lack of autocorrelation of signal), so our null hypothesis states that the autocorrelation coefficients (AC) are equal to zero, alternative hypothesis founds, that at least one AC differs from zero. White noise is a random signal (or process) with a flat power spectral density. White noise has zero value of autocorrelation function (ACF) for each delay. Coloured noise – noise that may contain a wide audible spektrum but shows a greater intensity in a narrow band of frequencies. It does not vary completely randomly, there is a co-variance between the sample values at different time-indexes.

STOCHASTIC MODEL

ARIMA is the autoregressive integrated moving average model. It is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). They are applied in some

cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to remove the non-stationarity. The model is generally referred to as an ARIMA(p, d, q) model where p , d , and q are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively (Box and Jenkins, 1970).

Given a time series of data X_t where t is an integer index and the X_t are real numbers, then an ARIMA(p, d, q) model is given by (Box and Jenkins, 1970):

$$\left(1 - \sum_{i=1}^p \phi_i \cdot L^i\right) (1-L)^d X_t = \left(1 - \sum_{i=1}^q \theta_i \cdot L^i\right) z_t \quad (1)$$

where:

 L – lag (or backshift) operator; ϕ_i – parameters of the autoregressive part of the model; θ_i – parameters of the moving average part; z_t – error terms proportional to the white noise ($z_t, MVN(0, \sigma^2)$)

For detection of the deterministic components masked in a random background the autocorrelation function (ACF) may be used, because autocorrelation functions of deterministic data persist over all time displacements, while autocorrelation functions of stochastic processes tend to zero for large time displacement. The ACF at lag k could be described as:

Table 1 Frequency intervals for estimating noise amplitudes (Wenzel, 1977).

from [cpd]	to [cpd]	description
0.007	0.193	longperiodic waves
0.800	1.193	diurnal tidal waves
1.733	2.127	semidiurnal waves
2.800	3.193	terdiurnal waves
3.800	4.193	quaterdiurnal waves

$$\rho_k = \frac{\sum z_t \cdot z_{t-k}}{\sum z_t^2} \sim MVN \quad (2)$$

and it has the nature of multivariate normal distribution (*MVN*).

The autocorrelations are uncorrelated with variances:

$$Var(\rho_k) = \frac{n-k}{n(n+2)} \approx \frac{1}{n} \quad (3)$$

so that the statistics:

$$n \sum_{k=1}^m \rho_k^2 \sim \chi^2(m) \quad (4)$$

should have χ^2 distribution with m -degrees of freedom (dimension of the moving average window).

If we replace the ACF with the ACF estimator we obtain test function BP (Box-Pierce) which means that the test statistic has Chi-squared distribution (Box and Pierce, 1970):

$$BP = n \sum_{k=1}^m \hat{\rho}_k^2 = n \hat{\rho}' \hat{\rho} \sim \chi^2(m-p) \quad (5)$$

And this function should also has χ^2 distribution with $m-p$ degrees of freedom (p is the degree of autoregression model).

Box and Pierce have also extended BP test statistics for any ARIMA(p,d,q) model (Box and Pierce, 1970):

$$BP = n \sum_{k=1}^m \hat{\rho}_k^2 \sim \chi^2(m-p-q) \quad (6)$$

with $m-p-q$ degrees of freedom, but it requires n large relative to m .

Better test statistic was made by Ljung and Box (function LB) as simple modification yielding substantially improved approximation (Ljung and Box, 1978):

$$LB = n(n+2) \sum_{k=1}^m (n-k)^{-1} \hat{\rho}_k^2 \sim \chi^2(m-p-q) \quad (7)$$

TIDAL DATA ADJUSTMENT

We can write equations of observation corresponding to each reading l_i in the form:

$$l_i + \hat{v}_i = f(\hat{x}, \hat{y}, \dots) \quad (8)$$

where:

\hat{x}, \hat{y} – estimated unknown tidal parameters;

\hat{v}_i – estimated observation residuals.

The solution is obtained from the normal equations giving the most probable estimators of the unknowns and the corresponding errors under the assumption that they are normally distributed and uncorrelated. For that matter, the geodetic records are disturbed not only by accidental errors, but also by a non stationary noise function $g(u, v, \dots)$. The equations of observation become:

$$l_i + \hat{v}_i = f(\hat{x}, \hat{y}, \dots) + g(u, v, \dots) \quad (9)$$

Elimination of the noise external to the tidal band:

- energy at frequencies higher than the Nyquist frequency. Elimination by using suitable sampling rate;
- energy at frequencies higher than the main tidal bands. This part of the spectrum is frequently referred to as “noise” but can incorporate harmonic constituents due to non linearity effects. Elimination by a low pass filtering;
- energy at frequencies lower than the diurnal band. This part of the spectrum is conventionally called „drift”, although it contains the long period tides. Elimination by a high pass filtering;
- energy between the tidal bands which is the most difficult to eliminate in the filtering procedure. Part of it will remain in the filtered data.

The error estimation for the adjusted tidal parameters made by Eterna 3.4 package (Wenzel, 1996) is done in two steps:

- by least squares adjustment procedure, which generally gives too optimistic estimates due to the negligence of error correlations;
- by Fourier amplitude spectrum of the residuals of the adjustment using average spectral amplitudes over individual tidal bands.

The error estimation procedure uses average noise amplitudes from the computed residual spectrum in intervals presented in Table 1.

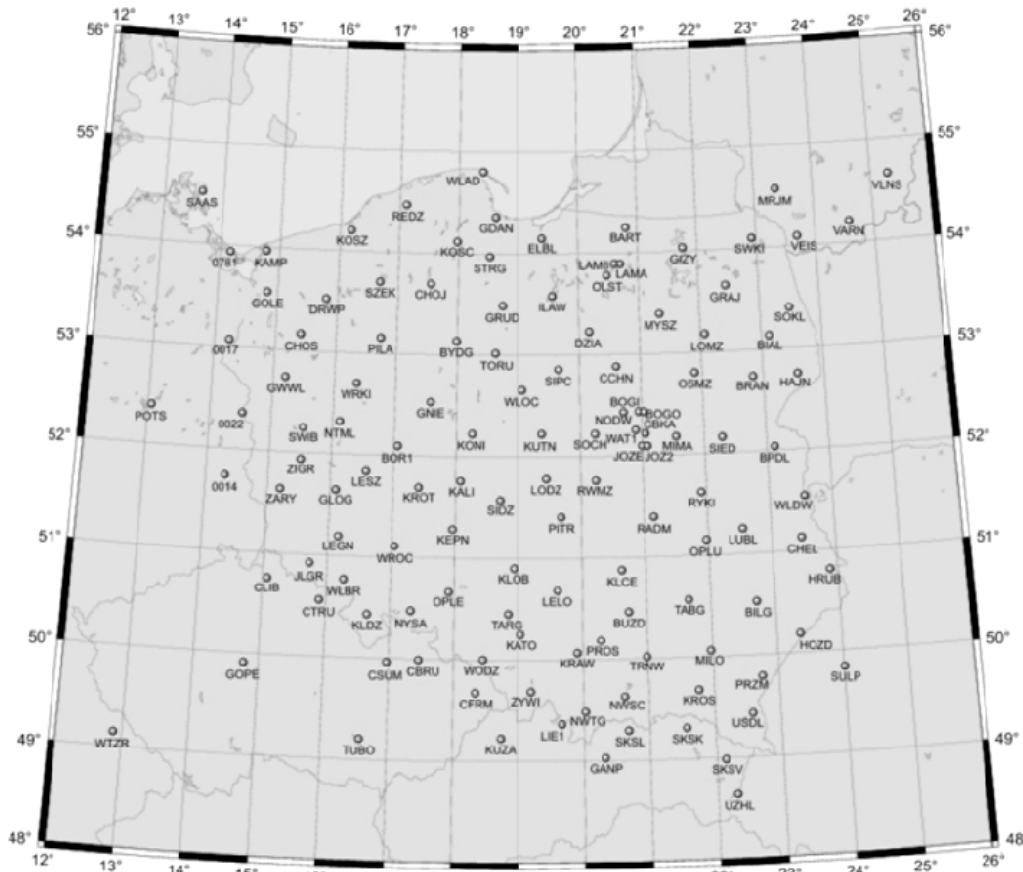


Fig. 1 Adjusted GPS network.

The standard deviations computed by the least squares adjustment are scaled up by the relation:

$$scale = \frac{noise(f)}{w_n} \quad (10)$$

where

w_n – the estimated Fourier amplitude of noise in the frequency band 0.01...6.0 cpd.

The equation (8) could be rewritten for each epoch t_i under the form:

$$l_i + v_i = \sum_j \sum_k d_j \cdot A_{j,k}^T \cdot \cos(\alpha_{i,j,k}^T + \varphi_j) \quad (11)$$

where:

- i – epoch;
- j – tidal group;
- k – wave inside tidal group;
- $d_j = \frac{A_j}{A_j^T}$ – the ratio (amplification) of the observed to the theoretical amplitude of the group j ;
- $A_{j,k}^T$ – the theoretical amplitude of a tidal constituent k inside of the group j ;

$\alpha_{i,j,k}^T$ – the argument at instant t_i of a tidal constituent (j,k)

φ_j – the phase difference supposed constant inside group j .

And the determined values during the least squares adjustment are d_j and φ_j . Moreover value of d_j depends on the assumed deterministic model of the solid Earth tides.

DETERMINISTIC MODEL

The HW95 tidal potential catalogue has 12935 tidal waves containing 19300 adjusted coefficients. It has been computed using the DE200 numerical ephemeris of Jet Propulsion Laboratories, Pasadena, of the solar system bodies between 1850 and 2150. The computation has been carried out by an iterative procedure combining spectral analysis and least squares adjustment. The HW95 tidal potential catalogue has been validated using different gravity tide benchmark series. The error of gravity tides computed at mid latitude stations from HW95 is 0.0015 nm/s² RMS and 0.010 nm/s² at maximum between years 1850 and 2150 (Hartmann and Wenzel, 1995).

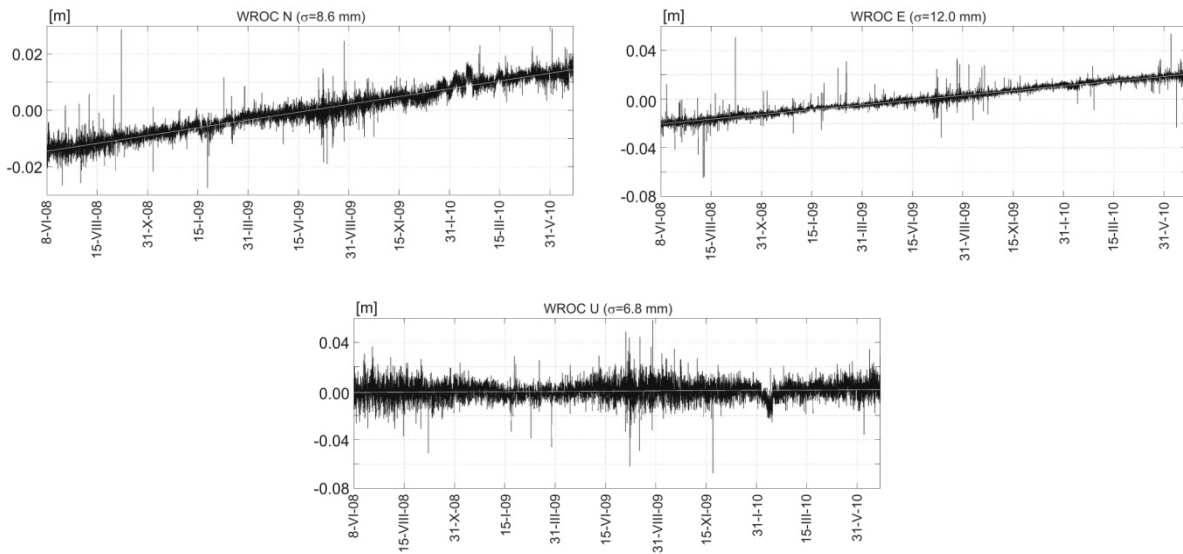


Fig. 2 North, East and Up components [m] of WROC site.

GPS DATA

In order to determine the stations' positions the data from ASG-EUPOS (Polish Active Geodetic Network) was used. ASG-EUPOS multifunctional precise satellite positioning system established by the Head Office of Geodesy and Cartography in 2008 (Bosy et al., 2008). The adjusted network consisted of 130 stations (Fig. 1), the period covered observations collected from 8.06.2008 to 18.06.2010.

Data processing strategy as well as the applied models are widely described in the paper by Araszkiewicz et al. (2010).

For practical calculations the WROC station, which belong to the Institute of Geodesy and Geoinformatics, Wrocław University of Environmental and Life Sciences, was taken. The data collected at this site is presented in Figure 2.

RESULTS AND DISCUSSION

For LB test for autocorrelation of residuals ($n=17\ 640$ samples) $m=24$, $\alpha=0.05$ and residual time series of 735 days were taken. The residuals from ETERNA (Fig. 3) gave the value $LB=18\ 960$ and $\chi^2=36.42$ which reject the null hypothesis, as well as the residuals from ARIMA(1,1,1) (Fig. 4) ($LB=1\ 392$, $\chi^2=36.42$).

Only the residuals from simulated white noise ($LB=18.95$, $\chi^2=36.42$) (Fig. 5) do not reject the null hypothesis. Therefore, residuals of both models (ETERNA and ARIMA) can not be accepted as a white noise.

Figure 6 shows a comparison chart of autocorrelation coefficients for the 24-hour window. For the simulated white noise autocorrelation coefficients are not significantly different from zero

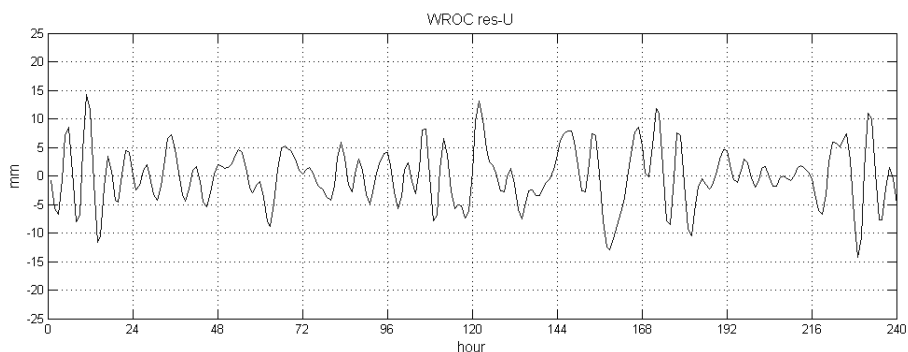


Fig. 3 Residuals from ETERNA (example of 10 days period).

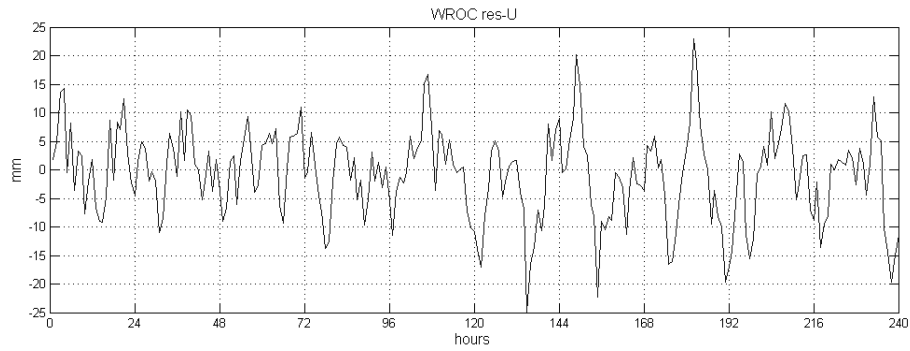


Fig. 4 Residuals from ARIMA (example of 10 days period).

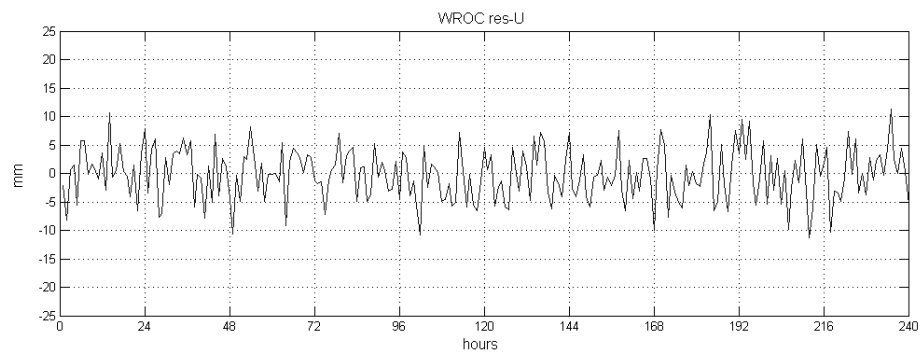


Fig. 5 Residuals from simulated white noise (example of 10 days period).

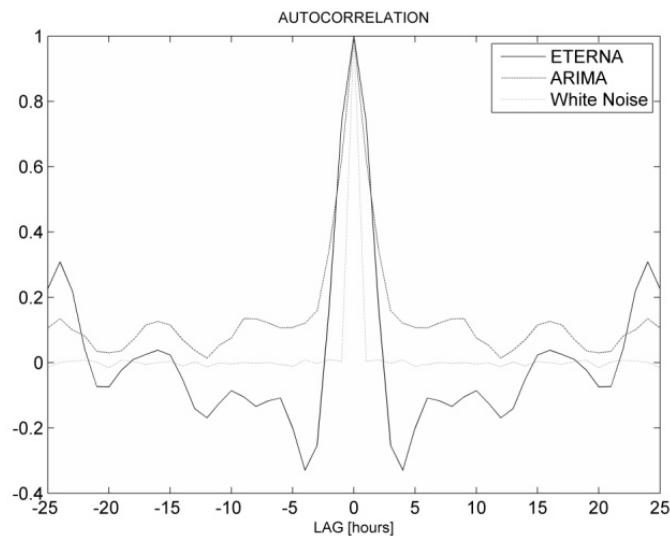


Fig. 6 Comparison chart of autocorrelation coefficients.

which means that the power of white noise is spread evenly across the frequency range, which cannot be said about clearly visible maxima and minima for local tidal model.

The ARIMA model noise power spectrum is close to $1/f$, which characterizes the coloured noise, in

particular the so-called flicker noise (Fig. 7). In the model ETERNA signal is initially filtered and does not contain a power spectrum of high and low frequencies. The nature of the spectrum partly corresponds to the coloured noise (frequency range 0.125-0.25 1/h – periods from 4 to 8 h), partly to the

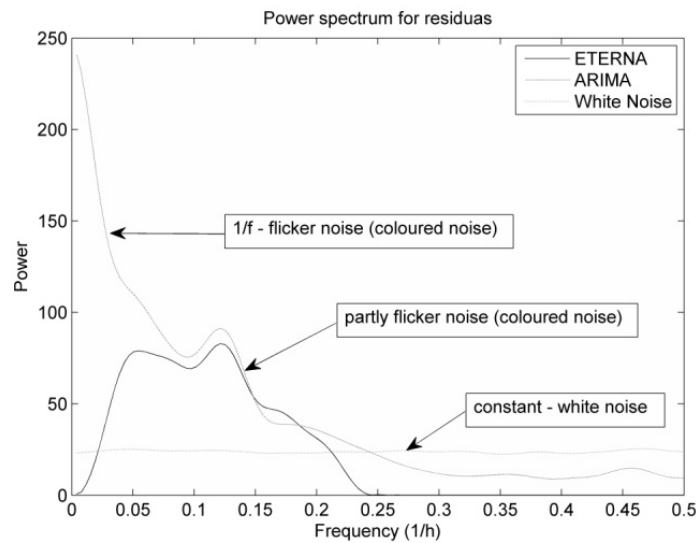


Fig. 7 Smoothed power spectrum and noise model identification.

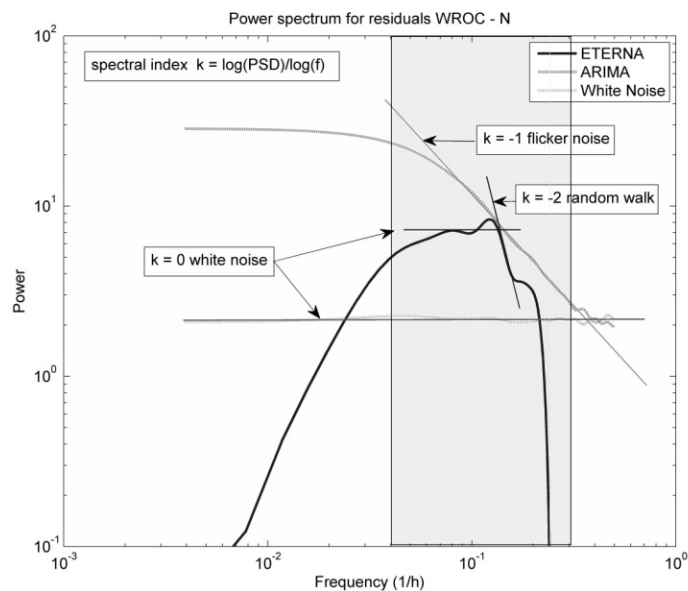


Fig. 8 Spectral index for the N component (WROC site).

white noise (frequencies 0.04-0.15 1/h – periods from 8 to 24 h). For comparison, the white noise power spectrum is represented by the horizontal line.

As shown by S. Williams, the nature of the noise is characterized by spectral index, calculated as the ratio of the slope of the power spectrum drawn on a logarithmic scale (Williams et al., 2004). Index of around -1 indicates the coloured noise (flicker), corresponds to -2 – random walk, of course $k=0$ is the white noise. Figure 8 shows the spectral index for the N component, Figure 9 for E and Figure 10 for U for WROC station. The grey boxes indicate the range of diurnal and sub-diurnal frequencies.

For comparison, the spectral index for the U component of some other ASG stations (BOR1, BOGI, LAMA, JOZE, KRAW, KATO) is shown on the Figure 11. The coloured noise nature of residuals for sub-diurnal frequencies is visible for all stations.

Figure 12 shows the distribution of noise of U component of ASG stations taken for analysis and the spatial distribution of noise in selected frequencies (diurnal and sub-diurnal) calculated for the U component using ETERNA software. The spatial distribution of the horizontal components of N and E are similar, but several times lower noise level. The distribution indicates the existence of several stations

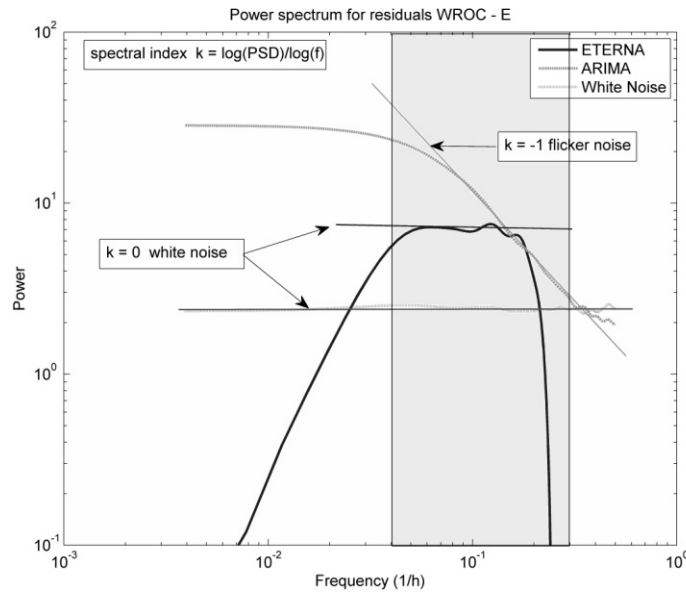


Fig. 9 Spectral index for the E component (WROC site).

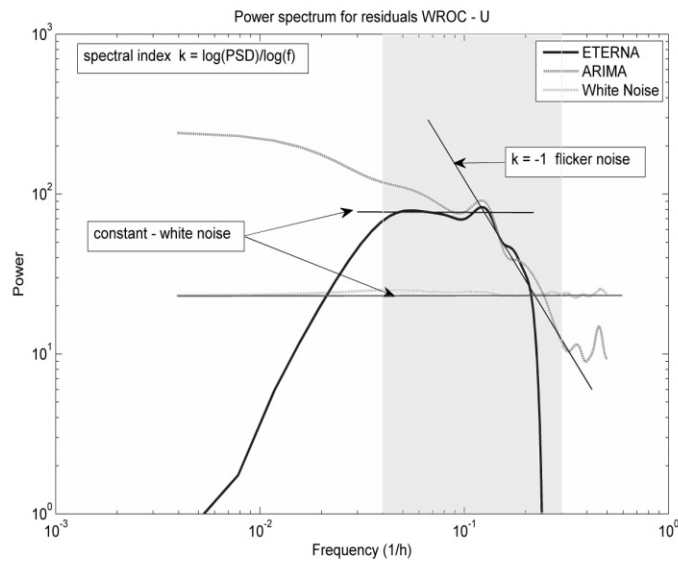


Fig. 10 Spectral index for the U component (WROC site).

as the outliers in noise level (single aggregations) and the generally higher level of noise for the Eastern and Southern (Carpathian) than the Western stations.

SUMMARY

The noise of GPS time series in diurnal and sub-diurnal frequencies cannot be considered as a white noise. It is rather coloured noise, similar to flicker noise. Following (Amiri-Simkooei, 2009) we can assume that $z(t)$ is a linear combination of independent, identically distributed unit-variance

random variables $\alpha(t)$, and a sequence of temporally correlated random variables $\beta(t)$, such that:

$$z(t_i) = a \cdot \alpha(t_i) + b_k \cdot \beta(t_i) \tag{12}$$

Scalar factors a and b_k are the amplitude of white and coloured noise of spectral index k respectively. The covariance matrix for the measurements x can then be expressed as:

$$C_x = a^2 I + b_k^2 J_k \tag{13}$$

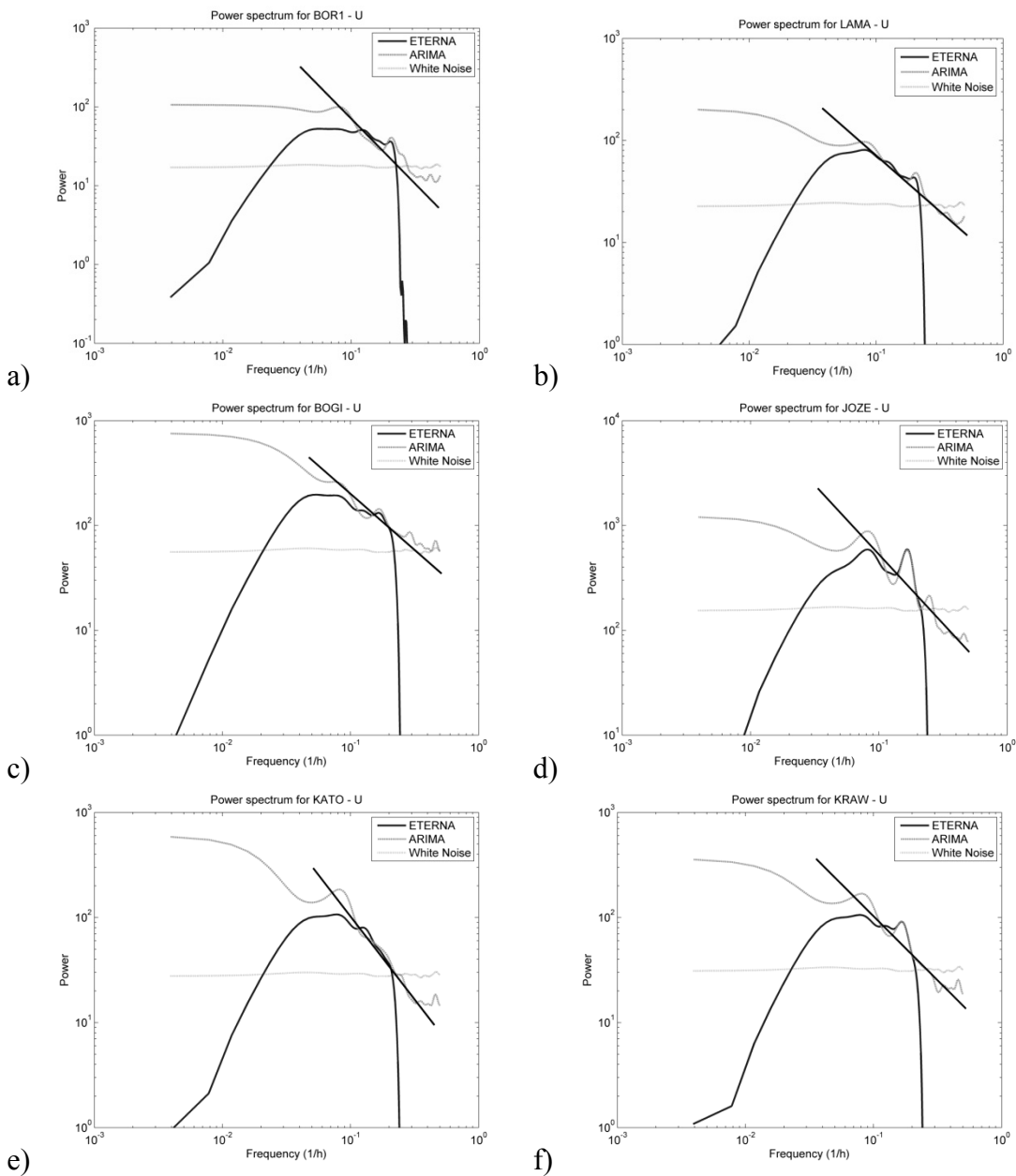


Fig. 11 Spectral index for the U component of selected ASG stations:
 a) BOR1, b) LAMA, c) BOGI, d) JOZE, e) KATO, f) KRAW

where I is the identity matrix and J_k is the covariance matrix for the appropriate coloured noise.

Average level of noise of diurnal and sub-diurnal frequency bands of GPS time series of ASG-EUPOS sites does not exceed the level of 1 mm. Generally sites on the Eastern part of the network are relatively more influenced by high frequency noise than sites on the Western part.

ACKNOWLEDGMENTS

Dr. Mariusz Figurski (CGS WAT) for his commitment to the GPS data processing is gratefully acknowledged.

REFERENCES

Amiri-Simkooei, A.R.: 2009, Noise in multivariate GPS position time-series, *Journal of Geodesy*, 83,175–187, doi: 10.1007/s00190-008-0251-8.

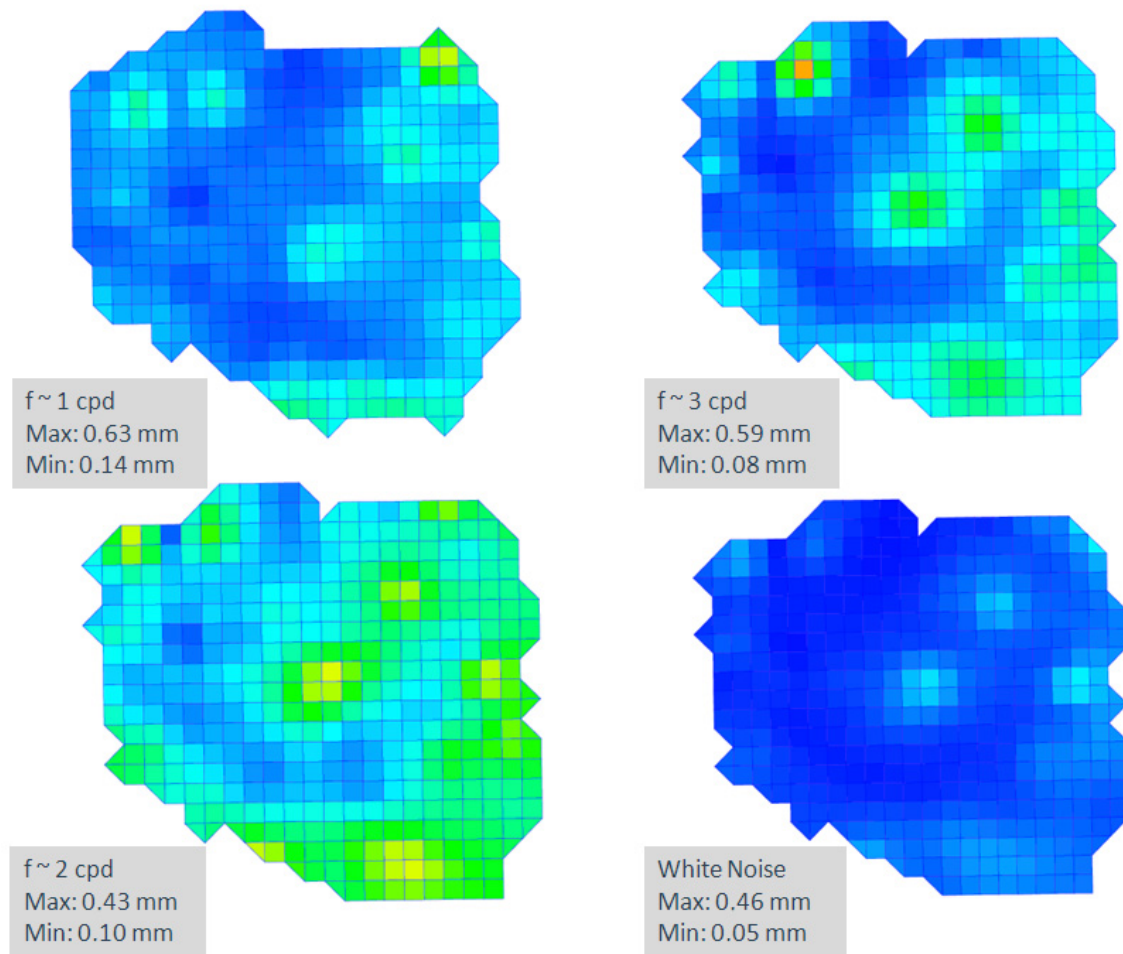


Fig. 12 Distribution of the noise level of ASG sites.

- Araszkiewicz, A., Bogusz, J., Figurski, M. and Szafranek, K.: 2010, Application of short-time GNSS solutions to geodynamical studies – preliminary results, *Acta Geodyn. Geomater.*, 7, No. 3 (159), 295–302.
- Bosy, J., Oruba, A., Graszka, W., Leończyk, M. and Ryczywolski, M.: 2008, ASG-EUPOS densification of EUREF permanent Network on the territory of Poland, *Reports on Geodesy*, 2 No. 85, Warszawa 2008, 105–112
- Box, G.E.P. and Jenkins G.: 1970, *Time series analysis: Forecasting and control*, San Francisco: Holden-Day.
- Box, G.E.P. and Pierce, D.A.: 1970, Distribution of the autocorrelations in autoregressive moving average time series models, *Journal of American Statistical Association*, 65, 1509–1526.
- Hartmann, T. and Wenzel, H.-G.: 1995, The harmonic development of the Earth tide generating potential due to the direct effect of the planets, *Geophysical Research Letters*, 22, No. 24, 3553–3556.
- Ljung, G.M. and Box, G.E.P.: 1978, On a measure of a lack of fit in time series models, *Biometrika*, 65, 297–303. doi:10.1093/biomet/65.2.297.
- Wenzel, H.-G.: 1977, Estimation of accuracy for the Earth tide analysis results, *Bulletin d'Informations Mares Terrestres (BIM)*, Bruxelles, 76, 4427–4445.
- Wenzel, H.-G.: 1996, The nanogal software: Earth tide processing package ETERNA 3.30, *Bulletin d'Information des Marées Terrestres (BIM)*, Bruxelles, No. 124, 9425–9439.
- Williams, S.D.P., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R.M., Prawirodirdjo, L., Miller, M. and Johnson, D.J.: 2004, Error analysis of continuous GPS position time series, *Journal of Geophysical Research*, 109, B03412, doi:10.1029/2003JB002741.