Variance components analysis of residual errors remaining in the GPS double-difference phase observations is presented in this paper. These errors arise due to unmodeled ionospheric, tropospheric and multipath effects and limit the accuracy of n (northwards), e (eastwards) and u (upwards) coordinate estimates. In addition, there are unavoidable pure random errors. An integral approach to variance components estimation of the residual errors is presented herein. The mathematical basis of the approach lies in the two-way nested classification, where one uses a linear model with random effects. It turned out that, for a baseline of 40km in length at mid-latitude region and during a year of the lowest sunspot activity in the 11-year cycle, daily standard deviations estimates of combined tropospheric and ionospheric effects are with intervals of ~1–11 mm, ~1–7 mm and ~4–51 mm for n, e and u coordinate, respectively. Poistemultipath effects, the intervals are ~4–12 mm, ~3–9 mm and ~8–30 mm, while with pure random error those are ~5–9 mm, ~4–7 mm and ~9–20 mm. It is important to say that the highest values arise in summer and the lowest in winter period.
1 INTRODUCTION

The theoretical basis regarding mathematical models of GPS observations which are used in processing of those observations can be found in almost every publication considering GPS. Of importance for analysis in this paper is a residual part in the double-difference phase observations arising due to several unmodeled effects limiting accuracy of positioning. The residuals exist because of lack of knowledge of a representational mathematical model for their component effects with irregular behaviour in time. Namely, the time-variable residuals of errors stem from unmodeled errors due to signal multipath, as well as tropospheric and ionospheric refraction, are analysed herein (the characteristics of these three effects are shown in subsection 1.2). The fourth effect arises due to antenna phase center offset (APCO) and antenna phase center variations (APCV). Namely, the electrical antenna phase center varies with satellite elevation and azimuth, intensity of the satellite signal, and is also frequency dependent. So, each incoming signal has its own electrical antenna phase center. The APCO defines the difference between the geometrical point on the antenna denoted as antenna reference point (ARP) and the mean position of the electrical antenna phase center (MAPC). On the other side, the APCV arise due to differences between individual electrical antenna phase centers and the MAPC. Unmodeled errors due to antenna phase centre offset and variations which, otherwise, have the greatest impact on accuracy of the coordinate in vertical plane, will not be considered because in this paper one uses a GPS baseline of 40km in length with receivers and antennas of the same type at both its ends (see subsection 1.3), and when using antennas of the same type at short- and medium-distance baselines (<1000km), there is no need to introduce corrections, because in double-difference phase observation model the effects of antenna phase centre offsets and variations of both antennas will cancel out (Kouba, 2009; El-Hattab, 2013). Of course, the cancellation is possible only if the receivers do not loose lock during the GPS observation session, in which case we have absence of cycle slips.

Above mentioned residual effects, together with the unavoidable pure random errors (pure error), limit the accuracy of GPS positioning, i.e. the accuracy of three component relative coordinate estimates of the measured GPS baseline vector in a chosen reference system. For geodetic purposes, one often uses positions in north-south (coordinate $n$) and east-west direction (coordinate $e$) in the plane of the local coordinate system containing a point on the Earth surface where the reference GPS receiver is located. The plane is orthogonal to the ellipsoid (WGS84) normal in that point. The third coordinate ($u$) represents orthogonal distance (upwards) between the point with the second GPS receiver (on the other side of the baseline) to the local horizon (the plane defined by $n$ and $e$ coordinate axes) of the reference receiver position (Leick, 2004).

Research in this paper aims to prove the possibility of using the two-way nested classification (sometimes also referred to as a hierarchical classification) linear model with random effects in variance components estimation of residuals arising due to the influence of multipath and combined influence of ionospheric and tropospheric refraction with simultaneous separation of the pure error variance estimate.

1.1 First results of the application of two-way nested classification in GPS positioning

In his doctoral dissertation, which is in its final stage, under the supervision of prof. D.Blagojević, the author of this paper presents the results of variance components researching regarding above mentioned
dominant residual effects and pure random effects in GPS double-difference phase observables, their mutual correlations and behaviour. The research has been conducted by using a common linear model with the two-way nested classification providing an integral consideration of all of these effects and until now there have been no papers about it. It is important to note that the idea for such a research came from prof. G. Perović. Guided by that idea, using mathematical formulas for two-way nested classification analysis with random effects (Searle, 1971) and testing hypotheses about influence of those effects (Hald, 1957), the author of this paper applied relational theory on relative GPS coordinates in his doctoral dissertation, whereby he greatly expanded analysis, and the results of a small part of the research, by which applicability of the mentioned idea is provided, have been cited in later published monograph (Perović, 2015). In this way, the results of integral variance components estimation for multipath, combined tropospheric and ionospheric refraction and pure random effects for a chosen GPS baseline have been published for the first time. In the monograph, actually, the PERG2FH method incorporating those results has also been presented for the first time. However, there is an essential difference between the source presentation and that in the monograph. The difference relates to the fact that the author of the dissertation, as mentioned before, analysed multipath effect among residuals, while the results for variance components he obtained have been presented by Perović (2015) as if they relate to quasi-stationary atmosphere blocks, whereby multipath effect has been rejected as if it does not exist.

In the dissertation, static GPS observations registered at each 30s over a period of four consecutive years (2008-2011) at two permanent stations of the (Montenegrin) MontePos network and eight EUREF permanent stations have been analysed. These stations were chosen so that they formed five GPS baselines with five different lengths, from 5.6km to 282km.

The behaviour of variance components estimates were separately analysed for daytime (when there was the strongest impact of ionosphere) and night (when the impact of ionosphere was reduced to a minimum), and the other relevant results were presented therein. Namely, the author has also dealt with a detailed research of the impact of outliers on the variance components estimates of residual and pure random components of the GPS station position error. Besides, the removal of the outliers has also been discussed as a part of this issue. The whole part of the research and most of the results obtained are planned to be presented in a particular paper dealing only with this issue. However, only a small part of the research is presented herein.

1.2 Main characteristics of the effects whose residuals are discussed in this paper

Multipath

Multipath is one of the most important error sources limiting GPS positioning accuracy and it is the most dominant when it comes to the baselines with smaller length. The phenomenon is conditioned by the receiver station environment characteristics. Due to periodic changes in the geometry of the positions of satellites, the influence of multipath changes with the same period, except in the case of the extraordinary, short-term circumstances. In order to discuss the two multipath components with completely different physical characteristics, the author presents the site dependent error. This error is composed of the errors coming from three effects, if the antenna setting instability error is omitted under the assumption of the high stability. So, in a simplified form it is written as $\delta S = \delta PCV + \delta MP_{\text{near-field}} + \delta MP_{\text{far-field}}$ (Wübbena,
Schmitz and Boettcher, 2006), where \( \delta_{PCV} \) is the error due to receiver antenna phase centre variations that depends on the satellite elevation angles and azimuths, while \( \delta_{MP_{near-field}} \) and \( \delta_{MP_{far-field}} \) are, respectively, errors due to antenna near-field and far field effects. Near-field effects are mainly caused by multipath interferences induced by reflectors located in the close vicinity of the antenna (e.g. surfaces of pillars or special adaptations where the antennas are mounted on (Elósegui et al, 1995)). These effects differ for various antenna types and setups and have a long-periodic (the periods of oscillation can reach several hours) behaviour with no zero mean characteristic and therefore cause a systematic error, especially in the coordinate height (Wübbena, Schmitz and Boettcher, 2006). These effects cancel out in double-difference phase observations when using antennas of the same type on both ends of baseline with similar environment characteristics of the two end receiver stations, but on such a distance that Earth’s curvature has no effect on the, conditionally speaking, identical visibility of satellites on these stations at each epoch. On the other side, far-field effects are systematic effects arising due to multipath interferences in the presence of objects located further away from the antenna (glass surface of the surrounding buildings, trees, vehicles, other large reflective surfaces, etc.) with zero mean characteristics and have a short-periodic behaviour (Wübbena, Schmitz and Boettcher, 2006) with a typical period between 15min and 30min (Seeber, 2003). Far-field effects can be averaged out by sufficient length of observation data. Residuals of these effects in double-difference phase observations have an irregular behaviour in time limiting the accuracy of GPS precise positioning.

The maximum value of multipath error in measured distance equals one quarter of the signal wavelength, therefore about 4.8cm for the observations on L1 frequency and 6.1cm for those on L2 frequency, while when using the L0 (Ionosphere-Free), LN (Narrow-Lane), LW (Wide-Lane) and LI (Geometry-Free) linear combinations of phase measurements, this error reaches, respectively, the values of 21.6, 5.4, 43.3 and 10.9cm (Wildt, 2006).

About multipath and its impact on GPS measurements one can also read in some new publications, such as Leick, Rapoport and Tatarnikov (2015) and Miller, Zhang and Spanias (2015). Various authors researched specially the influence of multipath on the GPS position estimate. For example, Fan and Ding (2006) discussed the impact of multipath on the positional accuracy at a baseline vector of 3m in length.

**Ionospheric refraction**

This effect causes a systematic error, and it is reflected in the delay of the signal emitted by satellite, i.e. refraction of the same, which depends on the TEC (Total Electron Content) values observed along the signal path as well as on the frequency of the signal. It is well known that the basic periods of TEC values change are diurnal, 27-day, annual and 11-year. However, in the time series of registered TEC values one can observe a semidiurnal, tri-diurnal, semiannual and tri-annual period (Asgari and Amiri-Simkooei, 2011). These facts are of great importance for the stochastic consideration of ionospheric influence on the results of GPS precise positioning, especially on the estimates of coordinates.

The largest part, i.e. the first order ionospheric effect, is eliminated by using linear L0 (Ionosphere-Free) combination of GPS measurements on both frequencies, L1 and L2. This part is about 99% of the total ionospheric effect and, at low elevation angles of satellites as well as the presence of ionospheric maximum, causes the error in measured distance with the magnitude that reaches a value of about 150m (Steigenberger,
However, it remains an unmodeled part of ionospheric effects which is of the greatest importance for GPS precise positioning. It consists of the second and third order ionospheric effect. The relational residual effect remaining in double-difference phase observations has an irregular behaviour in time and limits the accuracy of GPS precise positioning. It is shown that ionospheric signal delay influenced by residuals of the second order may cause the error in measured distance of about 4cm for satellite elevation of 10º, while the error caused by residuals of the third order reaches a value of about 1-4mm (Steigenberger, 2009). Residual ionospheric effects can be significantly reduced by prolonging the session. An incorrect estimate of TEC value has the effect of shortening the length of baseline vector. If minimal satellite elevation of 10º is adopted, at mid-latitudes, the error of TEC value estimate of 10 TECU (1 TECU = 10¹⁶ electrons/m²) leads to shortening the length of baseline vector of 0.7ppm (Santerre, 1989).

So, for instance, for a baselines of 10km in length the shortening is 7mm, while for those of 100km we have a value of 7cm. Thus, the size of this error depends on solar activity and sudden ionospheric disturbances, as well as on latitude of the place where the receiver is located. The ionosphere is most active in a band extending up to approximately 20º on either side of the geomagnetic equator. In this region, small-scale ionospheric disturbances, i.e. scintillations, mainly occur. The same phenomenon is also present at the high-latitude regions close to the poles (auroral regions). Scintillations are defined as rapid, short-term variations in the amplitude and phase of radio signals travelling through the ionosphere.

For an insight into current research related to aforementioned ionosphere residual effects, a reader is referred to Fritsche et al. (2005) and Hoque and Jakowski (2008).

**Tropospheric refraction**

Apart from ionospheric effects, tropospheric refraction is also dominant and causes a systematic error that may be divided into relative and absolute components. Relative component is caused by relative troposheric error that arises at one endpoint with respect to another endpoint of baseline. This component is manifested through the error of station height, \( \Delta h = \Delta T_r / \sin \alpha_{\min} \) (following Beutler et al., 1988), where \( \alpha_{\min} \) is the minimal elevation angle at which the satellite can be seen from the station, and \( \Delta T_r \) is the relative tropospheric error. So, for example, for \( \Delta T_r = 1 \text{mm} \) and \( \alpha_{\min} = 10^\circ \), we have station height estimate error of 5.8mm.

Absolute component arises due to error under the influence of tropospheric refraction, whereby an identical meteorological conditions are assumed at both ends of the baseline, and is reflected in scale error of the baseline length, \( \Delta l / l = \Delta T_a / R_E \sin \alpha_{\min} \) (following Beutler et al., 1988), wherein \( l \) and \( \Delta l \) are, respectively, the baseline length and the error in baseline length, due to absolute tropospheric error \( \Delta T_a \), while \( R_E \approx 6371 \text{km} \) is the Earth’s mean radius. Then, based on this formula, if e.g. \( \Delta T_a = 50 \text{mm} \) and \( \alpha_{\min} = 10^\circ \), we have a scale error of 45ppb (parts per billion). Thus, for \( l = 100 \text{km} \) the error in baseline length of 4.5mm is obtained.

Tropospheric refraction effects are particularly expressed at long baselines, because on their ends exist different atmospheric conditions. Besides, these effects are of great importance at large altitude differences between the endpoints of the baseline. After the application of a model for atmospheric correction (e.g. Saastamoinen model), an unmodeled residual part remains in double-difference phase observations.
and, in addition to ionospheric residuals, makes it difficult to resolve phase ambiguities and thus limits the accuracy of GPS precise positioning. The residual is also irregular in time and, as it is the case with ionospheric residual, can be reduced by prolonging the session. Tropospheric refraction usually does not change rapidly with time and is, thus, considerably reduced on the triple-difference level.

When it comes to researching models which are related to the corrections for tropospheric effects, it is important to mention earlier works such as Hopfield (1969) and Saastamoinen (1972), and, of the more recent works, e.g. Satirapod and Chalermwattanachai (2005) and Wielgosz et al. (2011) may be cited. Some of the publications where the authors have researched the behaviour of tropospheric residuals are Musa (2007) and Ibrahim and El-Rabbany (2007).

1.3 Characteristics of GPS baseline chosen for analysis

For the variance components analysis of aforementioned residuals in this paper, a daily GPS static measurements registered at each 30s during the period of the year 2008 at two of nine GPS permanent stations in MontePos (Montenegrin) network are used. One of the stations is located in Podgorica and the other is in Bar (Figure 1). At both stations, a high antenna setting stability is established, and the length of the baseline between them is 40km. So, it is a medium-distance baseline. Besides, an altitude difference of about 38m exists between the endpoints of the baseline. Both stations are equipped with Leica GRX1200 receivers and Leica AT504 Choke-Ring/Radome antennas which have the same orientation. This type of receiver provides tracking performances with strong signals and low noise, even for satellites on low elevations and in difficult conditions, with an optimal reduction of multipath and resistance to interference. In addition, the mentioned antenna type meets the requirements of Dorne-Margolin T (D/M_T) L1/L2 microstrip antenna, JPL design (http://www.nekretnine.co.me/me/Djelatnosti5.asp).
In post-processing of GPS measurements, MontePos station in Bar was used as the fixed baseline station, so all of the effects for analysis have been contained in the resulting fixed solutions for relative coordinates $n$, $e$ and $u$ which are used as input data for the variance components calculations. The year 2008 was chosen because the same is related to minimal sunspot activities within the respective well-known 11-year period, when there were no increased ionospheric influences which would make it difficult to resolve phase ambiguities in the post-processing.

Trimble Total Control v2.7 software was used herein. Before the post-processing, the following settings in the software were previously done: Processing Interval – 30s; Preference – Prefer P Code; Elevation Cutoff – 10°; Orbit Type – Precise; Frequency – Lc (for baselines longer than 5km); Antenna Model – US NGS ant_info.003; Tropospheric Model – Saastamoinen; Meteorological Model – MSIS; Processing Mode – OTF (as an auxiliary mode for phase ambiguities resolving for each epoch). In addition, in the software Processing Options the Filter for solutions was set to always choose the Best Solution among the solutions obtained from the available linear combinations of the measurements, L0 (Ionosphere-Free), LN (Narrow-Lane) and LW (Wide-Lane).

2 METHODS

In this paper, in addition to the theory based on the two-factor hierarchical classification with random effects which was presented by Searle (1971) and Hald (1957), the idea of PERG2FH method is used (Perović, 2015), but with some modifications (explanation in footnote 1). Due to the limited scope of this paper, the author briefly presents a linear model which the method is based on, but all of the used formulas have been included. For details, the reader is referred to the monograph (Perović, 2015), where the mathematical apparatus presented by Searle (1971) and Hald (1957) was skillfully and concisely given. The author of the monograph also gave a certain scientific contribution for the case of unbalanced data that occurs in the two-factor hierarchical classification issue. In sense of what has been mentioned before, research in this paper is based on linear two-factor model of hierarchical classification with random effects, with a different number of measurements per groups (Searle, 1971):

$$Y_{ijk} = \mu + \alpha_i + \beta_{ij} + \varepsilon_{ijk}; \ i = 1,2,\ldots,a, \ j = 1,2,\ldots,b_i, \ k = 1,2,\ldots,n_{ij},$$

with:

$$N_i = \sum_{j=1}^{b_i} n_{ij} = n_i, \ B = \sum_{i=1}^{a} b_i = b, \ N = \sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij}; \ n_{ij} \geq 2, \ b_i \geq 2, \ a \geq 2.$$  \hspace{1cm} (2)

In accordance to discussion in this paper, we have the following notations in (1): $Y_{ijk}$ – the fixed solution for the relative coordinate $n$, $e$ or $u$ (measurement), $\mu$ – a general mean, $\alpha_i$ – the combined random effect of tropospheric and ionospheric refraction, $\beta_{ij}$ – the random effect of signal multipath (nested within $\alpha_i$) and $\varepsilon_{ijk}$ – the pure random error, or, shortened “pure” error (nested within $\beta_{ij}$).

In this paper, because of the large number of measurements, it may be considered that $\mu$ equals “true value” of the measurement $Y_{ijk}$ denoted by $A_i$, which leads to the fact that the constant systematic error (fixed parameter denoted by $\delta$) in the model (1) can be written as $\delta = \mu - A_i = 0$. Thus, instead of the

\footnote{Perović (2015) has connected this effect to quasi-stationary atmospheric blocks}
model (1), the author uses the following linear model of true errors:

\[ \Delta_{ijk} = \alpha_i + \beta_{ij} + \epsilon_{ijk}, \quad i = 1, 2, \ldots, a, \quad j = 1, 2, \ldots, b_i, \quad k = 1, 2, \ldots, n_j. \]  

This linear model is followed by corresponding stochastic model:

\[
\begin{align*}
\alpha_i &\sim N(0, \sigma^2_{\alpha}), \\
\beta_{ij} &\sim N(0, \sigma^2_{\beta}), \\
\epsilon_{ijk} &\sim N(0, \sigma^2_{\epsilon}), \\
\text{cov}(\alpha_i, \beta_{ij}) &= \text{cov}(\alpha_i, \epsilon_{ijk}) = \text{cov}(\beta_{ij}, \epsilon_{ijk}) = 0, \\
\text{cov}(\epsilon_{ijk}, \epsilon_{ipqr}) &= 0 \quad (i \neq p \lor j \neq q \lor k \neq r),
\end{align*}
\]

and, in addition:

\[
\Delta_{ijk} \sim N(0, \sigma^2_{\Delta}); \quad i = 1, 2, \ldots, a, \quad i = 1, 2, \ldots, a, \quad k = 1, 2, \ldots, n_j.
\]

With respect to (3), (4) and (6), we have:

\[
\sigma^2_{\Delta_{ijk}} = \sigma^2_{\alpha} + \sigma^2_{\beta} + \sigma^2_{\epsilon}; \quad i = 1, 2, \ldots, a, \quad j = 1, 2, \ldots, b_i, \quad k = 1, 2, \ldots, n_j,
\]

where \(\sigma^2_{\alpha}, \sigma^2_{\beta}\) and \(\sigma^2_{\epsilon}\) are variance components which the variance \(\sigma^2_{\Delta}\) consists of, and they are determined from the true errors given in (3).

The author divided the set of all true errors for the contemplated year by days and within each established subset, particularly for each of the coordinates \((n, e\text{ and } u)\), he has carried out the following procedure:

1. Checking of behavior of the true errors using graphical presentations and removing of blunders with the magnitudes of several decimeters which can be easily detected;
2. In order to estimate variance component \(\sigma^2_{\epsilon}\), previously, all of the subsets of true errors established by days have to be divided into groups, so that all of those groups correspond to time interval within which we can consider that the factor \(\beta\) causes almost constant effect (following the condition in PERG2FH method (Perović, 2015));
3. For estimation of variance component \(\sigma^2_{\beta}\) it is necessary to divide the set of all true errors into the groups, so that all of those groups correspond to time interval within which we can consider that, now, the factor \(\alpha\) causes almost constant effect (following the condition in PERG2FH method (Perović, 2015));
4. Calculation of values given in (2);
5. Calculation of the following independent sums of squares (following Searle (1971), whereby instead of \(Y\) we put \(\Delta\) everywhere):

\[
\begin{align*}
\text{SSA} &= \sum_{i=1}^{a} N_i (\bar{\Delta}_i - \bar{\Delta})^2, \\
\text{SSB} &= \sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij} (\bar{\Delta}_{ij} - \bar{\Delta}_i)^2, \\
\text{SSE} &= \sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} (\Delta_{ijk} - \bar{\Delta}_{ij})^2, \\
\end{align*}
\]

using:

\[
\begin{align*}
\bar{\Delta}_{ij} &= \frac{\Delta_{ij}}{n_{ij}}, \\
\Delta_{ij} &= \sum_{k=1}^{n_{ij}} \Delta_{ijk}, \\
\bar{\Delta}_i &= \frac{\sum_{j=1}^{b_i} \Delta_{ij}}{N_i}, \\
\Delta_{ij} &= \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} \Delta_{ijk},
\end{align*}
\]

\[
\bar{\Delta} = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} \Delta_{ijk}}{N} \approx 0, \quad \Delta_{ij} = \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} \Delta_{ijk}.
\]

6. Calculation of mean squares (Searle, 1971):
\[ m_3^2 = MSA = SSA / (a - 1), \quad \text{with} \quad f_3 = a - 1 \ \text{d.f.} \quad (10) \]
\[ m_2^2 = MSB = SSB / (B - a), \quad \text{with} \quad f_2 = B - a \ \text{d.f.} \quad (11) \]
\[ m_1^2 = MSE = SSE / (N - B), \quad \text{with} \quad f_1 = N - B \ \text{d.f.} \quad (12) \]

7. **Hypothesis testing of influence of the nested factor \( \beta \):**

\[ H_{0,\beta} : \sigma_\beta^2 = 0, \ \text{against} \ H_{a,\beta} : \sigma_\beta^2 > 0. \quad (13) \]

*Test statistic is (Hald, 1957; cv: Perović, 2015):*

\[ F_\beta = \frac{m_2^2}{m_1^2}, \quad \text{where} \quad F_\beta \mid H_{0,\beta} \sim F(f_2, f_1). \quad (14) \]

When \( F_\beta > F_{1-\alpha, B-a,B-a} \), then we accept \( H_{a,\beta} \) and estimate variance component \( \sigma_\beta^2 \). Here we are only interested in such case. For the procedure in the case of rejecting \( H_{a,\beta} \), see Perović (2015);

8. **Hypothesis testing of influence of the first factor \( \alpha \):**

\[ H_{0,\alpha} : \sigma_\alpha^2 = 0, \ \text{against} \ H_{a,\alpha} : \sigma_\alpha^2 > 0. \quad (15) \]

*Test statistic for balance data is (Hald, 1957; cv: Perović, 2015):*

\[ F_\alpha = \frac{m_3^2}{m_2^2}, \quad (16) \]

for which, under the assumption that there are not any two groups of the second order such that the number of data in the first differs more than 50% from number of data in the second group, one can consider that \( F_\alpha \mid H_{0,\alpha} \sim F(f_3, f_2) \) (Perović, 2015), so, in the case when \( F_\alpha > F_{1-\alpha, a-1,a-a} \), we accept \( H_{a,\alpha} \) and estimate variance component \( \sigma_\alpha^2 \). Here we are also only interested in such case and for the procedure in the case of rejecting \( H_{a,\alpha} \), the reader is referred to Perović (2015);

9. In the case of accepting \( H_{a,\beta} \) and \( H_{a,\alpha} \), the complete model (3) exists, so we can calculate variance component estimates \( m_\varepsilon^2, m_\beta^2 \) and \( m_\alpha^2 \); and by using the criterion “three-sigma” in an iterative process, all of outliers among the true errors were rejected;

10. Before the calculation of \( m_\varepsilon^2, m_\beta^2 \) and \( m_\alpha^2 \), and by using the method of “three-sigma” in an iterative process, all of outliers among the true errors were rejected;

11. Calculation of the values (following Searle (1971), whereby instead of \( Y \) we use \( \Delta \) everywhere):

\[
\begin{align*}
  k_1 &= \sum_{i=1}^a n_{i1}^2 / N, \\
  k_3 &= \sum_{i=1}^a \sum_{j=1}^b n_{ij}^2 / N, \\
  k_{12} &= \sum_{i=1}^a \left( \sum_{j=1}^b n_{ij}^2 / n_{i} \right), \\
  T_A &= \sum_{i=1}^a \sum_{j=1}^b \Delta_{ij}^2 / n_{ij}, \\
  T_{AB} &= \sum_{i=1}^a \sum_{j=1}^b \Delta_{ij}^2 / n_{ij}, \\
  T_0 &= \sum_{i=1}^a \sum_{j=1}^a \sum_{k=1}^b \Delta_{ijk}^2 \quad \text{and} \quad T_\mu = \Delta_{..}^2 / N. \quad (18)
\end{align*}
\]

12. By using (17) and (18) variance components estimates \( m_\varepsilon^2, m_\beta^2 \) and \( m_\alpha^2 \) are calculated (Searle, 1971):

\[
\begin{align*}
  m_\varepsilon^2 &= (T_{0} - T_{AB}) / (N - B), \quad \text{with} \quad f_\varepsilon = N - B \ \text{d.f.} \quad (19) \\
  m_\beta^2 &= [T_{AB} - T_A - (B - a) m_\varepsilon^2] / (N - k_{12}), \quad \text{with} \quad f_\beta \ \text{d.f.} \quad (20) \\
  m_\alpha^2 &= [T_A - T_\mu - (k_{12} - k_{1}) m_\varepsilon^2 - (a - 1) m_\beta^2] / (N - k_1), \quad \text{with} \quad f_\alpha \ \text{d.f.}; \quad (21)
\end{align*}
\]
Degrees of freedom estimates are (Perović, 2015):

\[
\hat{f}_b = \frac{(m_b^2 - m_{1b}^2)^2}{(m_b^2)^2 + \frac{(m_{1b}^2)^2}{N - B}} \quad \text{and} \quad \hat{f}_a = \frac{(m_a^2 - E_2 m_a^2 + E_1 m_a^2)^2}{(m_a^2)^2 + \frac{(m_a^2)^2}{a-1} + \frac{(m_a^2)^2}{B - a} + \frac{(m_a^2)^2}{N - B}},
\]

(22)

which, in the case of a large number of data when variances are used instead of their estimates, give the following degrees of freedom:

\[
\hat{f}_b = \frac{(\sigma_b^2 - \sigma_{1b}^2)^2}{(\sigma_b^2)^2 + \frac{(\sigma_{1b}^2)^2}{N - B}} \quad \text{and} \quad \hat{f}_a = \frac{(m_a^2 - E_2 m_a^2 + E_1 m_a^2)^2}{(m_a^2)^2 + \frac{(m_a^2)^2}{a-1} + \frac{(m_a^2)^2}{B - a} + \frac{(m_a^2)^2}{N - B}},
\]

(23)

where notations \( E_1 = \frac{\pi' - \pi}{\pi} \) and \( E_2 = \frac{\pi'^2}{\pi} \) have been introduced with:

\[
\pi = \frac{1}{B - a}(N - \Sigma_{i=1}^a (\Sigma_{j=1}^{n_i} - n_i) / n_i) \quad \text{and} \quad \pi' = \frac{1}{a-1} \Sigma_{i=1}^a (\Sigma_{j=1}^{n_i} - (\Sigma_{j=1}^{n_i} / n_i) - (\Sigma_{j=1}^{n_i} / n_i) / N);
\]

(24)

13. Calculation of values (Searle, 1971):

\[
k_4 = \Sigma_{i=1}^a (\Sigma_{j=1}^{n_i} - n_i)^2, \quad k_5 = \Sigma_{i=1}^a (\Sigma_{j=1}^{n_i} / n_i), \quad k_6 = \Sigma_{i=1}^a (\Sigma_{j=1}^{n_i} / n_i)^2 / n_i,
\]

(25)

\[
k_7 = \Sigma_{i=1}^a (\Sigma_{j=1}^{n_i} - n_i)^2 / n_i, \quad k_8 = \Sigma_{i=1}^a (\Sigma_{j=1}^{n_i} / n_i), \quad k_9 = \Sigma_{i=1}^a n_i,
\]

(25')

and afterwards, by using them and (17), also calculation of values:

\[
\lambda_1 = (N - k_{12})^2[k_4(N + k_i) - 2k_9 / N],
\]

(26)

\[
\lambda_2 = k_4 k_7 + k_5 k_5^2 + k_7 k_3 k_4^2 - 2k_9 k_4 k_5 k_6 / N + 2(N - k_{12}) k_5 k_4 / N, \quad \lambda_3 = [(N - k_{12})^2(N - 1)(a - 1) - (N - k_{12})^2(a - 1)(B - a) - (k_{12} - k_3)^2(N - 1)(B - a) / (N - B),
\]

(27)

\[
\lambda_4 = (N - k_{12})^2[k_5(N + k_i) - 2k_9 / N],
\]

(28)

\[
\lambda_5 = (N - k_{12})^2(N - k_i),
\]

(29)

\[
\lambda_6 = (N - k_{12})(N - k_i)(k_{12} - k_i);
\]

(30)

14. By using (17), (25), (25') and (26)-(31), variances and covariances of estimates \( m^2_b \), \( m^2_a \) and \( m^2_\alpha \) are calculated (Searle, 1971):

\[
m^2_b = 2m^4_b / (N - B),
\]

(32)

\[
m^2_a = [2(k_7 + Nk_3 - 2k_5) m^4_b + 4(N - k_{12}) m^2_b m^2_b + 2(a - a) m^2_b / (N - a) / (N - B)] / (N - k_{12})^2, \quad m^2_\alpha = 2(\lambda_1 m^4_a + \lambda_2 m^4_b + \lambda_3 m^4_a + 2\lambda_4 m^2_a m^2_b + 2\lambda_5 m^2_a m^2_b + 2\lambda_6 m^2_b m^2_b) / [(N - k_{12})^2(N - k_{12})^2],
\]

(33)

\[
\text{cov}(m^2_b, m^2_\alpha) = [(k_{12} - k_3)(B - a) / (N - k_{12}) - (a - 1)] m^2_b / (N - k_4), \quad \text{cov}(m^2_b, m^2_b) = -(B - a) m^2_b / (N - k_{12}),
\]

(34)

(35)

(36)
\[
\text{cov}(m_{\beta}^2, m_{\alpha}^2) = \frac{2[k_3 - k_2 + (k_1 - k_4) / N]m_{\beta}^4 + 2(a - 1)(B - a)m_{\alpha}^4}{(N - B) - (N - k_1)(k_1 - k_2)m_{\beta}^2} / [(N - k_1)(N - k_1)], \tag{37}
\]

and then also correlation coefficients:

\[
\text{corr}(m_{\epsilon}^2, m_{\alpha}^2) = \text{cov}(m_{\epsilon}^2, m_{\alpha}^2) / (m_{\epsilon}^2 m_{\alpha}^2), \tag{38}
\]

\[
\text{corr}(m_{\epsilon}^2, m_{\beta}^2) = \text{cov}(m_{\epsilon}^2, m_{\beta}^2) / (m_{\epsilon}^2 m_{\beta}^2), \tag{39}
\]

\[
\text{corr}(m_{\beta}^2, m_{\alpha}^2) = \text{cov}(m_{\beta}^2, m_{\alpha}^2) / (m_{\beta}^2 m_{\alpha}^2). \tag{40}
\]

Besides the aforementioned procedure, Bartlett’s test (Bartlett, 1937) is used to test if \(k\) variance components obtained on the basis of the daily subsets of true errors are (statistically) equal, particularly for each of the factors (\(\alpha\), \(\beta\) and \(\epsilon\)), and for each of the coordinates (\(n\), \(e\) and \(u\)), and afterwards to test equality of variance components within each month in the year (then \(k = 12\)). The test implies the checking of hypothesis:

\[
H_0: \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2, \text{ against its alternative } H_1: \sigma_i^2 = \sigma_j^2, \text{ for at least one pair } (i, j). \tag{41}
\]

For that purpose, first the following variance is calculated:

\[
m = \frac{1}{f_m} \sum_{j=1}^{k} f_j m_j^2, \text{ so } f_m = \sum_{j=1}^{k} f_j \text{ st.sl.} \tag{42}
\]

where \(f_j (j = 1, 2, ..., k)\) are, respectively, degrees of freedom for variances \(m_j (j = 1, 2, ..., k)\), and after that the auxiliary value is also calculated:

\[
C = 1 + \frac{1}{3(k - 1)} \left[ \sum_{j=1}^{k} \left( f_j / f_m \right)^2 \right] - \frac{1}{f_m} \tag{43}
\]

Test statistic is:

\[
\chi^2 = \frac{1}{C} \left[ f_m \ln m^2 - \sum_{j=1}^{k} (f_j \ln m_j^2) \right], \tag{44}
\]

and it was proved that \(\chi^2 \mid_{H_0} \sim \chi^2_{k-1}\), whereby the condition \(f_j \geq 4\) must be fulfilled for each \(j\).

If \(\chi^2 > \chi^2_{1-\alpha; k-1}\), then \(H_0\) is rejected, and we should remove a suspicious variance and test the remaining variances to equality. The procedure, while it makes sense, is repeated until it comes to the case that \(H_0\) is accepted. Then, we will have the equality of \(k' (k' < k)\) variances and the variance from (42) is adopted as a definitive.

3 RESULTS

After applying the procedure described in Chapter 2, the author obtained very meaningful results which are presented below. However, it should be noted that near-field multipath component is not included in the analysis because the considered baseline has the antennas of the same type and orientation on both its ends, which are 40km away from one another, so it can be assumed that in each epoch for each visible
On the basis of what has been said in procedure steps 2 and 3 from Chapter 2, the time interval of 3min was chosen for estimation of variance component $\sigma^2_\varepsilon$, and for estimation of variance component $\sigma^2_\beta$ four time intervals of 1.5h, 2h, 3h and 4h, where it might be assumed that factor $\alpha$ causes a constant effect, were chosen. So, four variants of time intervals combining have been established. Those are the following:

- Variant A – (3min, 1.5h);
- Variant B – (3min, 2h);
- Variant C – (3min, 3h);
- and Variant D – (3min, 4h).

It is important to say that applying all variants, for each of three coordinates ($n$, $e$ and $u$), the tests given in procedure steps 7 and 8 from Chapter 2 show that influence of factor $\beta$ is present within each daily subset of data. However, in the case of factor $\alpha$, after applying variants A, B, C and D, it turned out that there is no influence of this factor, respectively, within 2, 6, 14 and 15 mentioned subsets of data for coordinate $n$, within 1, 5, 9 and 11 of them for coordinate $e$, while for coordinate $u$ we have such outcome within 1 subset and 2 subsets applying variant C and D, respectively. In these cases, which led to incomplete linear model, variance components estimates were not calculated. In tables 1-3, the extreme values of the related standard deviations estimates are presented.

**Table 1:** Extreme values of standard deviations estimates for coordinate $n$ (in mm)

<table>
<thead>
<tr>
<th>Standard deviations estimates</th>
<th>Variant A</th>
<th>Variant B</th>
<th>Variant C</th>
<th>Variant D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\varepsilon n}$</td>
<td>4.6</td>
<td>8.9</td>
<td>4.6</td>
<td>9.2</td>
</tr>
<tr>
<td>$m_{\beta n}$</td>
<td>4.2</td>
<td>11.4</td>
<td>4.2</td>
<td>11.9</td>
</tr>
<tr>
<td>$m_{\alpha n}$</td>
<td>0.9</td>
<td>10.7</td>
<td>0.9</td>
<td>11.0</td>
</tr>
</tbody>
</table>

**Table 2:** Extreme values of standard deviations estimates for coordinate $e$ (in mm)

<table>
<thead>
<tr>
<th>Standard deviations estimates</th>
<th>Variant A</th>
<th>Variant B</th>
<th>Variant C</th>
<th>Variant D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\varepsilon e}$</td>
<td>3.5</td>
<td>6.7</td>
<td>3.5</td>
<td>6.7</td>
</tr>
<tr>
<td>$m_{\beta e}$</td>
<td>3.0</td>
<td>9.1</td>
<td>3.1</td>
<td>9.4</td>
</tr>
<tr>
<td>$m_{\alpha e}$</td>
<td>0.8</td>
<td>6.8</td>
<td>0.8</td>
<td>6.9</td>
</tr>
</tbody>
</table>

**Table 3:** Extreme values of standard deviations estimates for coordinate $u$ (in mm)

<table>
<thead>
<tr>
<th>Standard deviations estimates</th>
<th>Variant A</th>
<th>Variant B</th>
<th>Variant C</th>
<th>Variant D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\varepsilon u}$</td>
<td>9.0</td>
<td>20.6</td>
<td>9.0</td>
<td>20.4</td>
</tr>
<tr>
<td>$m_{\beta u}$</td>
<td>8.1</td>
<td>28.9</td>
<td>8.2</td>
<td>29.8</td>
</tr>
<tr>
<td>$m_{\alpha u}$</td>
<td>4.1</td>
<td>51.2</td>
<td>4.1</td>
<td>51.1</td>
</tr>
</tbody>
</table>

Values in tables 1 and 2 show that between applied variants there are no significant changes in estimates $m_\varepsilon$, $m_\beta$ and $m_\alpha$ for coordinates $n$ and $e$, unlike those in table 3, where in the case of coordinate $u$ an increasing change in minimum and maximum values for $m_\beta$ is present, and that is followed by a decrease-
ing change in minimum and maximum values for \( m_\alpha \). This leads to the conclusion that by expanding time intervals in variants C and D we have partial mixing, i.e. ”overflowing” of the effect of factor \( \alpha \) in estimate \( m_\beta \). On the other side, if comparing variants A and B, there are no significant changes in \( m_\beta \) and \( m_\alpha \) values for any of three coordinates (\( n \), \( e \) and \( u \)), while \( m_\epsilon \) values are, as expected, practically equal between variants.

Based on the foregoing, and considering that, according to the previous research, the maximum period for far-field multipath component is 30min, as a reasonable choice among applied variants, the variant B is adopted. Therefore, only the results of the analysis obtained by applying this variant are presented hereinafter.

So, in Figure 2, graphical presentations of standard deviations estimates obtained by applying variant B, are presented.

![Figure 2: Behaviour of standard deviations estimates (in mm) (blue - \( m_\epsilon \), green - \( m_\beta \), red - \( m_\alpha \))](image_url)

In Table 4, extreme absolute values of correlation coefficients of obtained variance components estimates are presented. They show very low correlations, therefore, practically, there is no correlation. Such a case testifies to the fact that adequate variance components estimation method was used.

<table>
<thead>
<tr>
<th>Absolute values of correlation coefficients</th>
<th>Coordinate ( n )</th>
<th>Coordinate ( e )</th>
<th>Coordinate ( u )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>\text{corr}(m_\epsilon^2, m_\beta^2)</td>
<td>)</td>
<td>2.6 11.8</td>
</tr>
<tr>
<td>(</td>
<td>\text{corr}(m_\epsilon^2, m_\alpha^2)</td>
<td>)</td>
<td>0.0 0.0</td>
</tr>
<tr>
<td>(</td>
<td>\text{corr}(m_\beta^2, m_\alpha^2)</td>
<td>)</td>
<td>0.3 7.6</td>
</tr>
</tbody>
</table>

Applying Bartlett’s test, the equality of all variance components estimates was tested and it has turned out that for any factor (\( \epsilon \), \( \beta \) and \( \alpha \)) and any coordinate (\( n \), \( e \) and \( u \)), there is no equality. On the other side, the same test was applied within each month for all the factors and coordinates. Because of the limited scope of this paper, all results of testing are not presented herein. However, it is of an impor-
tance to note that, for the adopted significance level of 0.05, and after applying an iterative procedure of Bartlett’s test, it turned out that for coordinates \( n \), \( e \) and \( u \) we have equality of the variance components estimates within each month, but only in the case of the factor \( \alpha \), with the mean percentage of removed variances of 11%, 8% and 35%, respectively. The unified variance components estimates for each month are presented in Table 5. Because of large number of data they were obtained from, we can say they represent variance components.

Table 5: Unified variance components estimates (in mm\(^2\)) and their degrees of freedom estimates for factor \( \alpha \)

<table>
<thead>
<tr>
<th>Month</th>
<th>Coordinate</th>
<th>Coordinate</th>
<th>Coordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( m_i )</td>
<td>( \hat{f}_{m,a} )</td>
<td>( m_i )</td>
</tr>
<tr>
<td>January</td>
<td>5.0531</td>
<td>196.5</td>
<td>3.1007</td>
</tr>
<tr>
<td>February</td>
<td>7.3022</td>
<td>172.2</td>
<td>4.7019</td>
</tr>
<tr>
<td>March</td>
<td>10.2702</td>
<td>174.6</td>
<td>5.8493</td>
</tr>
<tr>
<td>April</td>
<td>9.3085</td>
<td>201.6</td>
<td>7.0534</td>
</tr>
<tr>
<td>May</td>
<td>8.1065</td>
<td>156.4</td>
<td>5.9720</td>
</tr>
<tr>
<td>June</td>
<td>42.7210</td>
<td>219.4</td>
<td>15.2606</td>
</tr>
<tr>
<td>July</td>
<td>28.9726</td>
<td>214.0</td>
<td>14.0820</td>
</tr>
<tr>
<td>August</td>
<td>34.0538</td>
<td>255.1</td>
<td>13.8782</td>
</tr>
<tr>
<td>September</td>
<td>11.9485</td>
<td>181.3</td>
<td>8.8691</td>
</tr>
<tr>
<td>October</td>
<td>9.3549</td>
<td>206.7</td>
<td>7.8849</td>
</tr>
<tr>
<td>November</td>
<td>7.1575</td>
<td>173.2</td>
<td>5.2894</td>
</tr>
<tr>
<td>December</td>
<td>8.0308</td>
<td>186.1</td>
<td>5.5280</td>
</tr>
</tbody>
</table>

Values in Table 5 show that the monthly standard deviations estimates of residuals arising due to combined tropospheric and ionospheric effects, for GPS baseline Podgorica-Bar (of 40km in length), and during the year 2008 with a minimal sunspot activity, are, for the coordinates \( n \), \( e \) and \( u \), in the intervals \(-2.2-6.5\)mm, \(-1.8-3.9\)mm and \(-9.4-20.2\)mm, respectively. By checking of equality of 12 unified variances in the same table, it turned out that for any coordinate \((n, e \text{ and } u)\) there is no equality.

On the other side, for factors \( \varepsilon \) and \( \beta \), and for the same significance level, we have no equality of variance components estimates, so they were not unified within months.

4 DISCUSSION

The research in this paper has shown that by using the two-way nested classification linear model one can integrally estimate the variance components of residual and pure random errors arising in GPS precise positioning, which is of a great importance, because a big step for the GPS data analysis has been made.

On the basis of the results obtained through the analysis, it is obvious that the highest variations of multipath and combined tropospheric and ionospheric residuals are present in summer period of the year and the lowest ones in winter period, which is in accordance with the year sunspot activities.

Each GPS site has its environmental characteristics and thus, especially, a specific periods of multipath errors. A frequencies of those periods are not the issue of this paper. However, the estimation of those
frequencies might be the issue for a further research. In the basis of such a research, the application of the nonlinear least-squares method would be relevant in searching for the related periodicity of GPS residual data, whose behaviour can be presented by trigonometric functions involving a multiple periodicity (two or more frequencies).

In addition, a further research might be also based on the application of a particular linear combination of GPS observations in the calculations of fixed solutions, where the impact of that on the magnitudes of variance components of residual errors could be analysed. The problem should also be considered depending on the GPS baseline length and/or its azimuth.

**Literature and references:**


