Statistical methods of the progress report of the tight adjustment of realization geodetic warps

Janusz DĄBROWSKI, Poland

Keywords: matrix realization, evaluation of the credibility of matrices.

SUMMARY

High availability of various surveying programs makes that the users of these programs – surveyors - often wonder about the possibilities of the use of individual applications for specific purposes. The inspiration to write this article was to conduct a study based on a comparative analysis of the three popular surveying programs: Geonet, C-Geo and Winkalk used for tight adjustment of realization geodetic warps . The authors of unpublished studies based on source documentation of geodetic service of the town Wojnicz bypass have chosen, among the three results, a route of elimination of a program guided by the principle of rejecting the results of the most outliers. In this article, the author made an attempt to discuss and carry out a much broader analysis of the results obtained on the basis of three programs of geodetic calculations. The preliminary data for alignment and comparison were adopted from the above-mentioned report, and all calculations were the same. The author in the article does not use the names of the programs but labels them A, B and C. The introduction of new names aims at making the names of specific programs 'secret', since the purpose of writing the article was to develop a comparative methodology, discussion of the results obtained and the proposed use of the optimum methodology for the verification of possible program implemented by each surveyor.

Our results confirmed the validity of the work to use a global parameter, which in many cases could be due to one size determine the choice of an appropriate surveying program by a specific user.

Statistical methods of the progress report of the tight adjustment of realization geodetic warps

Janusz DĄBROWSKI

1.INTRODUCTION

The assessment of the accuracy of matrix realization can be done in many ways. Frequently we use the so-called local indicators, the covariance matrix for each point or the covariance matrix for selected elements of the matrix realization [Czaja, 1993]. The most probable values of the coordinates of points of the matrix realization can be obtained by aligning closely the measured data by which coordinates are estimated with the use of multiparameter of Gaussian-Markov model [Czaja 1993, Gocał 2007, Ferguson, Takane 1997]. In Polish technical guidelines for realization measuring G-3.1, the average error the least-designated side of the network was used as the main criterion for assessing the accuracy of the designation of a horizontal warp realization. However, the mean adjustment error after aligning expressed in mm/1km was considered as a criterion for the accuracy of the designation of a network realization (Collective work 1994). The guidelines give the formula for the average error side, angle, direction, coordinate in the matrix method closely adjusted according to the method of least squares. Errors of the general formula are calculated according to the average error matrix function (Czaja 1993, Gocał 2007, Przewlocki 2009) mF:

$$m_F = m_0 \cdot \sqrt{f^T (A^T P A)^{-1} f} \qquad (1)$$

where:

A - matrix of coefficients of the unknowns in equation of amendments observation,

P - matrix of weights - diagonal matrix with elements equal to the inverse of squared error of observation,

m_0- the average error after aligning individual observations,

f - single-column matrix of coefficients of the unknowns as a function of F.

In matrices aligned by means of approximate methods, side length average error is calculated from the approximate analysis of the accuracy of such errors by analyzing the components, using the right transmission errors.

In both sides polygonal traverses, the average error of the length of the side is calculated using the formula [6]:

$$M_d = m_d \cdot \sqrt{\frac{n}{n+1}} \qquad (2)$$

where:

m_d - the average side length measurement error,

n - number of points designated.

The precision of conducting angle and linear measurements result from the criteria set for the particular classes of geodetic accuracy and delineation of the details of the planned construction resulting from the implementation of the condition:

$$m_t < \frac{k \cdot dl}{r} \qquad (3)$$

where:

m_t - average error of marking out details of the building,

dl - acceptable deviation of determining details of the building,

k - rate taking into account the influence of error of building-assembly duties, being included within the limits of $0.4 < k < 1.0$

r - rate which the value depends from the required plausibility of the correctness of ranging and by the degree of the randomness of errors of ranging.

The average error of ranging details of the building depends above all on errors of ranging warp; of geometry of warp and the accuracy of performing measurements of elements of this warp.

Another way of the evaluation of the accuracy of warp is applying global indicators which are a simultaneous exponent (measuring instrument) of all measured realization network elements .

In the next chapters, results of leveling of dense binding made on the same measuring and coordinate data were subjected to the comparison with three applications C-Geo, GEONET and Winkalk with statistical methods using the Statistica 9.0 package.

2. DESCRIPTION OF RESULTS OF LEVELING REALIZATION WARP

In the article we used the results of measurements of realization warp put on for the service of the ring road of Wojnicz in 2007. After analysing the results of compensation achieved with the help of three programs C-Geo, Geonet and Winkalk we stated, that at applying each of the programs we obtained results, in terms of the accuracy, in accordance with geodetic guidelines. All theoretical assumptions quoted in the introduction were fulfilled and from a formal point of view it is possible to recommend all programs for applying. However, a deeper analysis showed, that coordinate values and errors of appointing them for each of the programs are comparable but differ from each other. In order to examine which program gives reliable results, surveyors counted differences between individual results obtained from the compensation with three programs. The ultimate choice of the most reliable results, in their opinion, were made by choosing two most coincident results (Kochmański 1952), and then the one, in their view, the most credible.
Choice of two results of calculations which are most coincident from the three possible is in a way rejecting the outliers. Such proceedings are right when we have a lot of results of measurements and the average value is the most probable value for us. The values :Xwyr, Ywyr., dx, dy, mx, my and components and ellipses of the error ah, B and ϕ were a base of the detailed comparative analysis.

3. STATISTICAL COMPARATIVE ANALYSIS

A statistical comparative analysis is above all comparing basic accuracy parameters for coordinate points after leveling for object warp obtained with the help of programs C-Geo, Geonet and Winkalk. In the next part of the article, the names of programs are "hidden" and marked with letters A, B and C, since a commercial evaluation and choice of the best program aren't the only intention of the author but a proposal to conduct discussion of achieved results after applying the Statistica package for any programs.

The data taken from the geodetic study contained the comparison of the results of calculations of coordinates with three programs and methodology of choice of optimal results by rejecting the most diverging results. In the geodetic study mainly the X values and Y were being compared after leveling for the same approximate coordinate values and basic statistics of the incorrectness of coordinate points after the adjusting .

Below an excerpt from the table is added, in which two most similar mx and my values were overshadowed for two programs from three results calculated for every point and for the same "input parameters".

Table No. 1 Putting the mx and my value together obtained with the help of the program A, B and C

Nr pts	mX_A [mm]	mX_B [mm]	mX_C [mm]	mY_A [mm]	mY_B [mm]	mY_C [mm]
1	21,2	9,3	14,6	15,6	6,9	13,5
2	16,1	8,7	6,3	13,4	10	6,7

3	20,7999	9	6,7	18,6	10,1	7,2
4	13,8	11,4	9,8	10,4	11,4	8
…	…	…	…		…	…
24	20,5	11,8	10,1	15,3	13,3	11,7
Sum of pts	4	23	21	15	14	19
Share in %	8,3%	47,9%	43,8%	31,3%	29,2%	39,5%
Number A→B	3	3	0	5	5	0
Number A→C	1	0	1	10	0	10
Number B→C	0	20	20	0	9	9

Source: own study

The table shows an excerpt from putting the mX value together and mY obtained as a result of dense compensation with the help of the programs A, B and C. Three from two values were marked (overshadowed) : mX and mY for every case. Next, a number of pairs closest appearing to the value were summed up and a percentage value of the similarity of results was determined for individual pairs of results. In case of mX, results of leveling in the B and C program are coinciding appropriately in 47.9% and 43.8%, and for the mY value the most coinciding results are in A and C programs, though the divergences among A,B and C programs for the mY value are rather small. The comparison of the mX and mY value, as well as the visual evaluation of shading the respective elements of the table coincide fully with results of recapitulating at the bottom of the table. The graphic and valuable evaluation of rests from the variance does not fully match the fundamentals of statistics, since it does not relate to square sums of the value of rests. Additionally, the disintegration of rests and their sum are also a parameter which should be analysed. The sums of rests and the sum of squares of rests were compared below in the table. Additionally, a proposal of the evaluation of described sizes was put in the bottom part of table.

Table no 2 Calculation of the sum of residuals and the sum of squares of residuals by means of three programs

Number pts	mXA [mm]	mXB [mm]	mXC [mm]	mYA [mm]	mYB [mm]	mYC [mm]
Sum R	464	446	425	274	317	237
Sum R2	10469	11208	10570	3461	4951	2636
points for R	0	1	2	2	0	1
Points for R2	4	0	2	2	0	4
Total number of points	4	1	4	4	0	5

Source: own study

Total sum of points for programme A from estimation mx

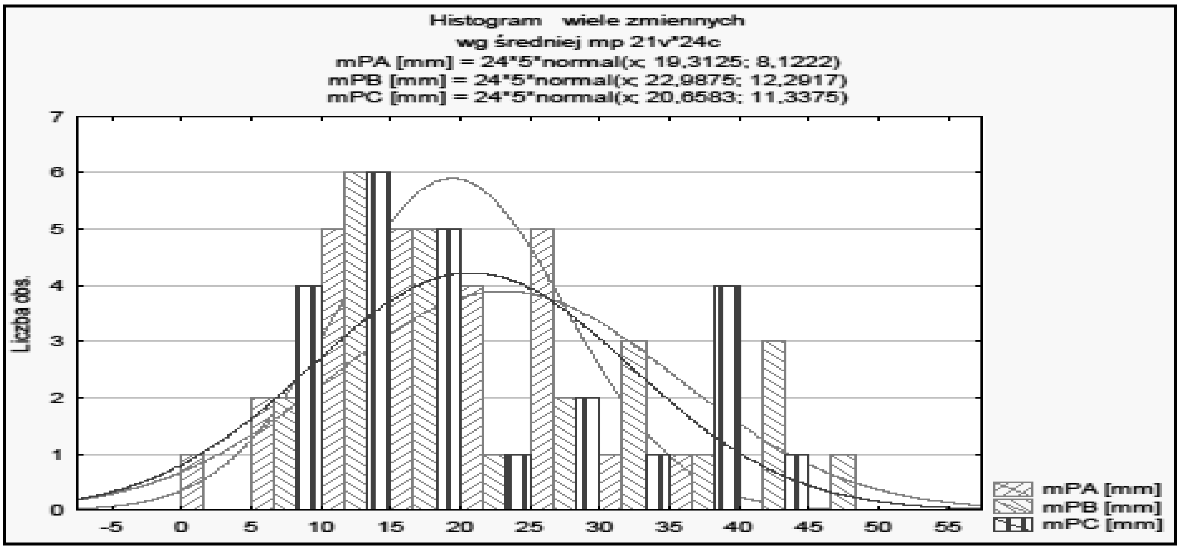
and my = 8
Total sum of points for programme B from estimation mx and my = 1
Total sum of points for programme C from estimation mx and my = 9

To make a comparative estimate the following grading scale was assumed:

The value of the sum of residuals: 0 (for ΣR→Value max.); 1 (for ΣR→Median value); 2 (for ΣR→Value min.).
The value of the sum of the squares of residuals: 0 (for ΣR^2→Value max.); 2 (for ΣR^2→Median value); 4 (for ΣR^2→Value min.).
Assumed grading score is of discretionary character and fulfills only supportive role in programme estimation. In the present work the author puts much more emphasis on the sum of the squares of the residuals than to the sum of the residuals which can be seen in assuming the grading scale in the above table. According to these criteria it can be stated that programs C and A adapt the model to the measuring results in the best way.

In all types of comparisons of statistical data very often the values of residuals are compiled in terms of the normal distribution. In the author's opinion in this case asimmetricality of the distribution show that the network doesn't cover the area in a fully uniform way or in different places of the network we deal with different numbers of observations.

Drawing no 1 Histogram of residuals for falsity of points for programs A, B i C

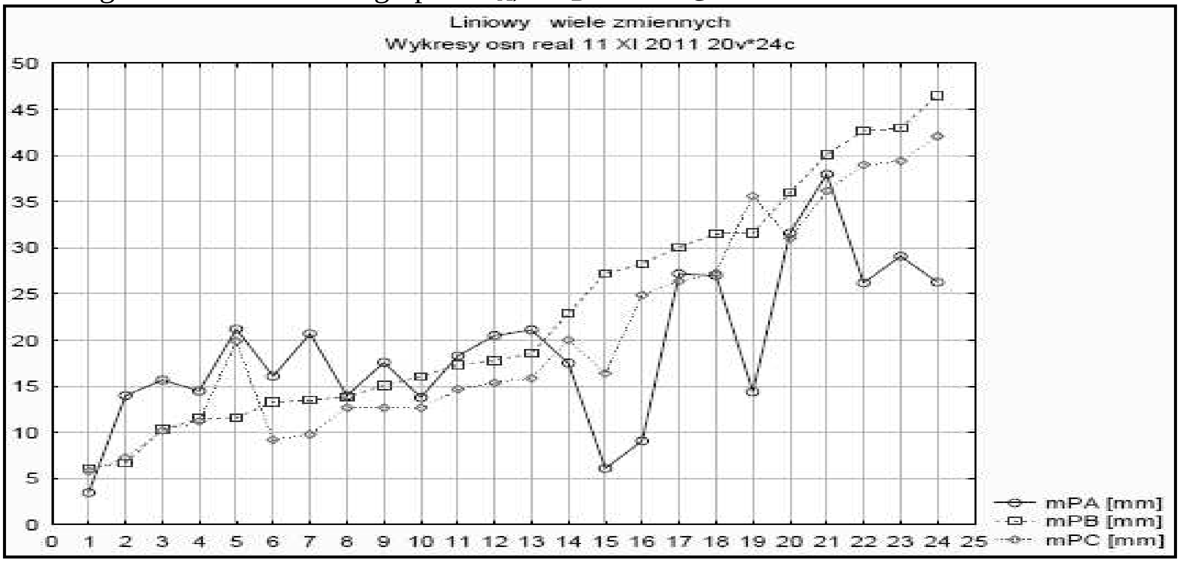


Source: own study

Histogram illustrates large concurrence and similarity for the distribution of value of the position error of points received by means of programme B and C.

Line graph of many variables show in drawing no 2 also confirms better internal concurrence of the results of calculations for programme B and C compared to programme A.

Drawing 2 Line graph mP_A, mP_B and mP_C

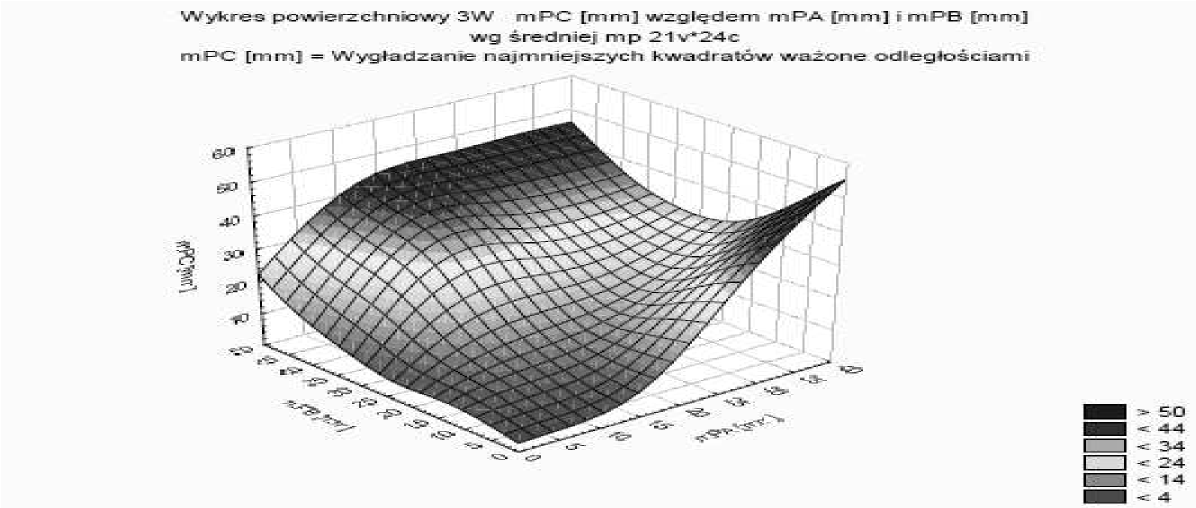


Source: own study

Values mp for points 15, 16 and 19 calculated by means of programme A are clearly lower than analogous received by means of programme B and C. Points on the list were chosen according to growing distances from the middle of local coordinate system. As for the rule the value mX, mY and mP should grow as it goes away from the beginning of the local coordinate system. The diagram below shows and confirms this rule for results achieved by means of programme A, B and C.

If measured dimensions of point location error were examined by means of spatial diagram, so the more it recedes from the beginning of the system the theoretical dispersion of the three examined variables should present uniform sloping surface looking similar to the lateral surface of a cone. Diagram no 3 shows the so-called saddle and slight departures of the diagram from its theoretical form.

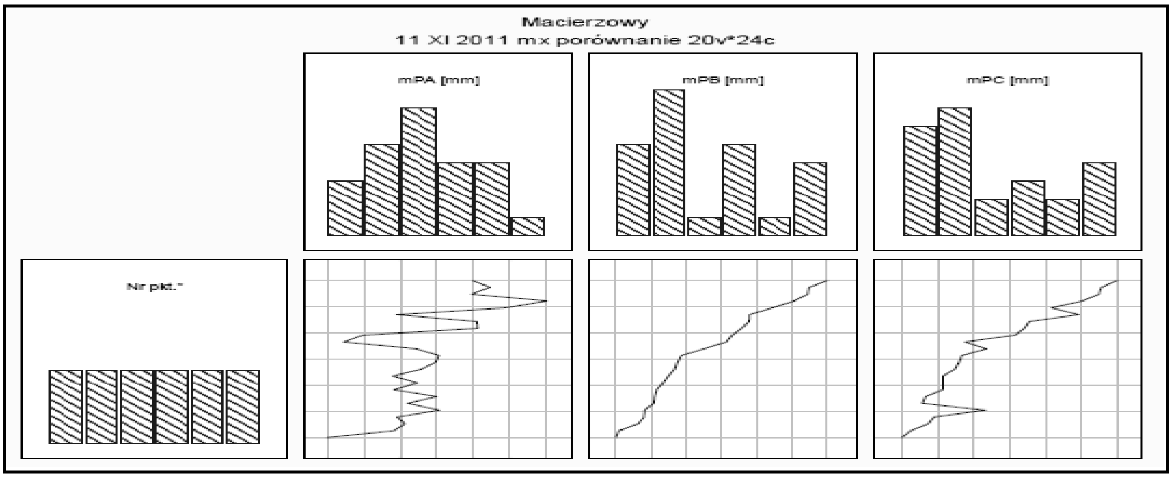
Drawing no 3 Surface diagram of point location errors for programme A, B and C



Source: own study

Very interesting to interpret is matrix diagram for values of residuals made simultaneously for three programmes.

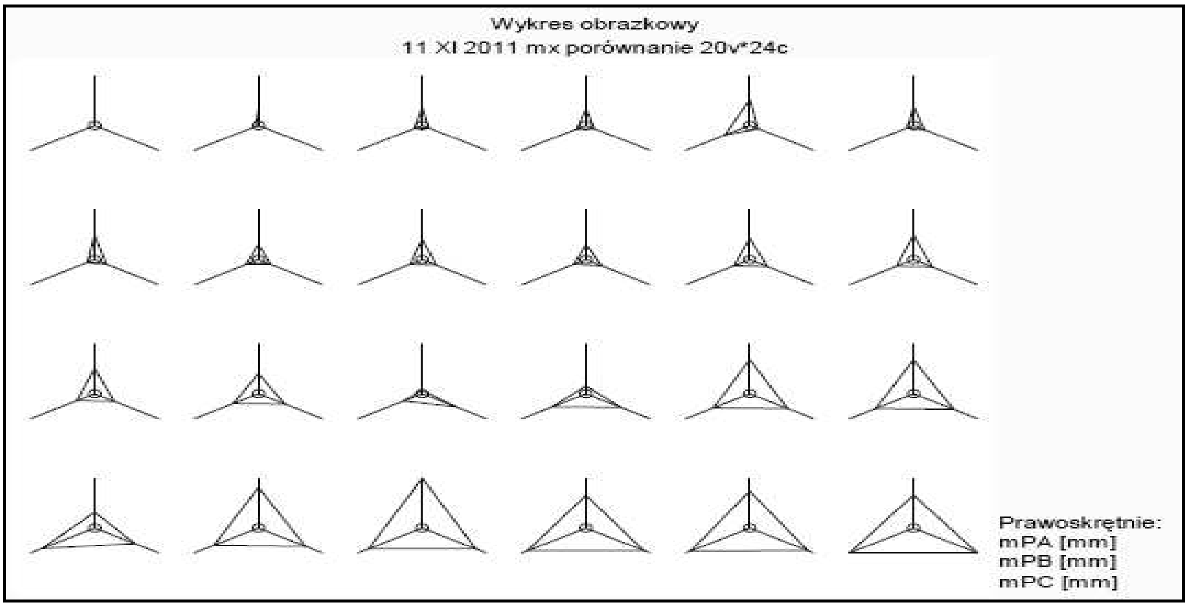
Drawing no 4 Matrix diagram of point location error in relation to point numbers



Source: own study

According to the distribution of the point location error programme A adjusts model best whereas the graph itself being a compound of numbers of points and distribution definitely „points to" programme B. Other form of depicting earlier conclusions is a pictorial chart which confirms the earlier quoted regularity of growth of the value mP as the distance from the middle of the coordinate system and symlik points grows.

Drawing no 5 Pictorial chart of residuals received from programmes A, B and C

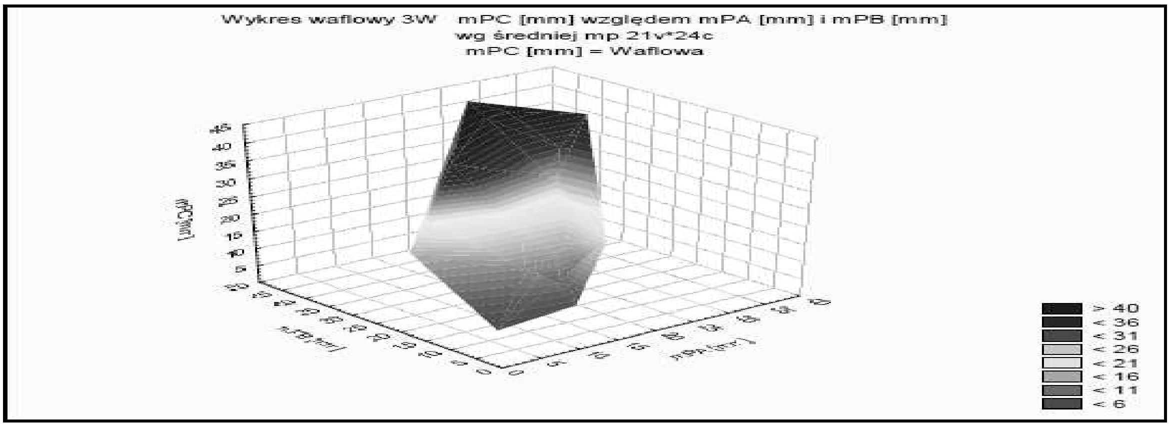


Source: own study

The length of the sides of the triangle informs about the value of residuals for particular programmes compared to a given matrix point.

Another possibility of graphical discussion of results offered by Statistica package is waffle diagram, and in particular gradient lines and contours of values of point location errors received for particular programmes.

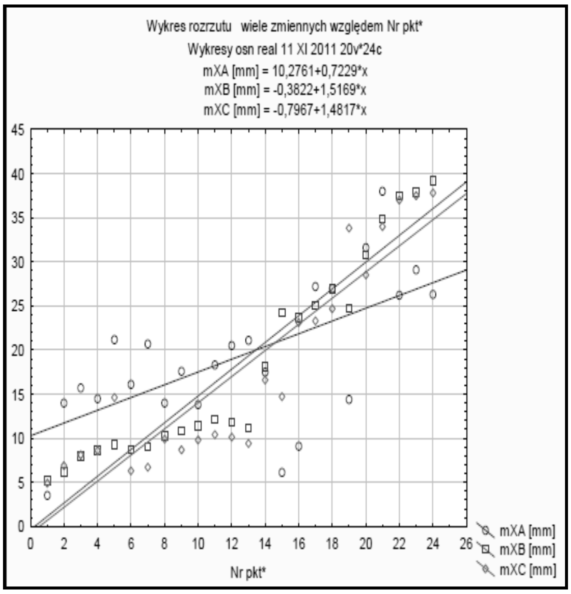
Drawing no 6 Waffle diagram for mP_A, mP_B and mP_C

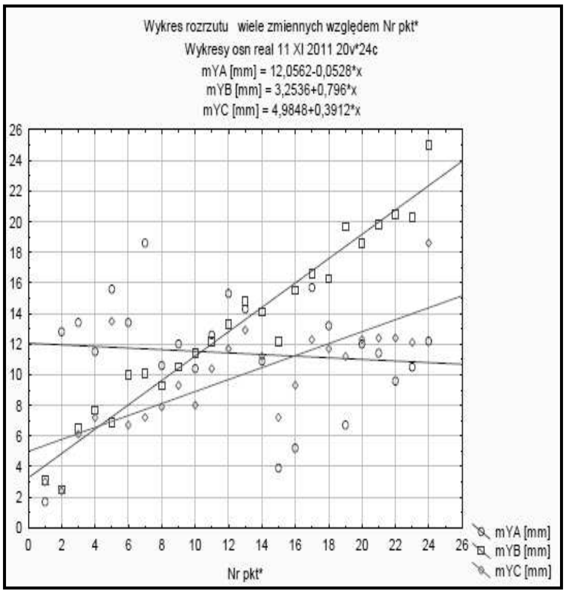


Source: own study

Waffle diagram is relatively difficult to interpret. Gradient lines show directions of maximum growth of value of point falsity. For a perfectly fitted model the gradient lines should be as a straight line joining the beginning of the system and the maximum value. Comparison of value mx and my on the graphs of dispersion with the lines of trend on drawings 7 and 8 confirms better fitting of the model along the X axis and definitely worse along axis Y. It is particularly visible for values mY received by means of programme A.

Drawing 7 and 8 Graph of dispersion mx and my with trend lines.

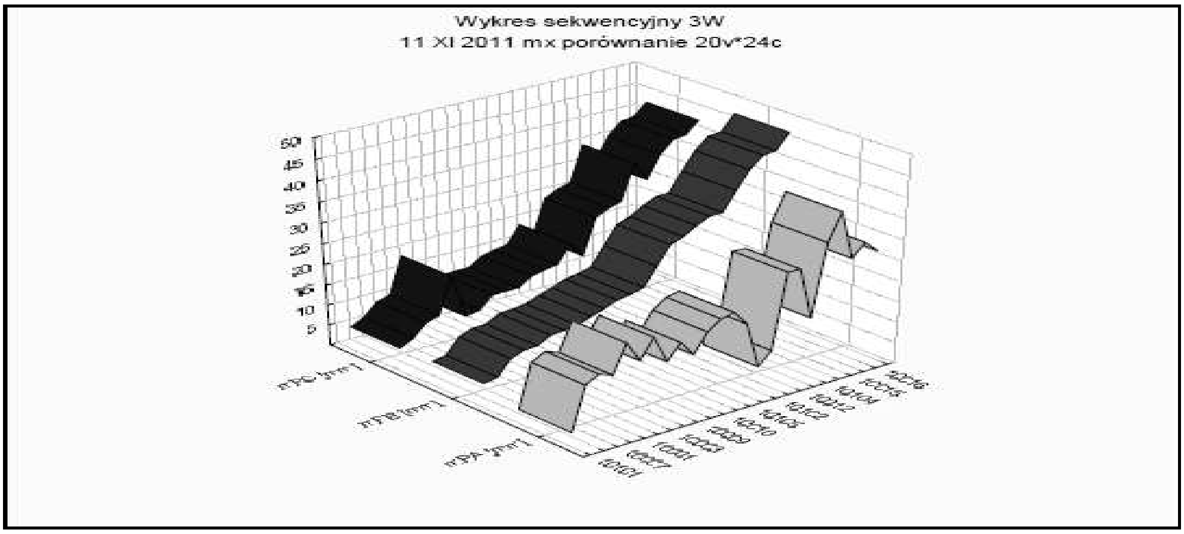




Source: own study

Parallel position of trend lines *mx* for programmes B and C suggests using by their creators similar alignment algorithms for both programmes. A little disturbing is the course (slope) of the regression line for values my received by means of programme A. Derogation from the rule of accumulation of residual value with the growing distance from the beginning of coordinate system is quite clear. Another way of graphical analysis of variability of residuals is a sequence ribbon diagram for three programmes.

Drawing 10 Sequence diagram for values *mp* received by means of programmes A, B and C



Source: own study

Ribbon diagram in a very synthetic way illustrates values of residuals of error location points calculated by using three programs. In a particular examined case after examining many aspects of graphics and value of residuals display it can be said that to choose a program it would be enough to compare only sums of squares of residuals and graphical of the ribbon diagram.

4. CONCLUSIONS

All considered and tested programs such as C-Geo, Geonet and Winkalk meet the requirements for the software to level geodetic bench marks.

Basic elements of the additional analysis of results of leveling of the geodetic software beyond standard mentioned in the report (mx, my, mp, A, B, fi, [pvv], mo, md, mα, mk) should be:

- dispersion diagrams of the elements of analysed quantities e.i. mx and my,
- graphic evaluation of the regularity of growth of the value mx and my as the distance from the beginning of the coordinate system grows for local systems or for reparse points.

In ambiguous situations of the optimal choice of program the analysis should be extended by diagrams of distribution of residuals: surface diagram, pictorial (triangular) diagram and ribbon diagram.

After appropriate preparation of data the process leveling itself takes only a few minutes for a typical realization bench mark. Therefore it is advisable, using Statistica package, to execute several charts allowing for the optimal selection of software.

It also seems advisable to develop synthetic global parameter which is a function of the basic precision parameters measured and designated elements of the control network.

In the author's opinion the choice of component indexes of the global parameter should be made with the help of factor analysis (Ferguson 1997, Dąbrowski 2009) which allows for monotonic ranking of the parameters carrying the same information (Sobczyk 1997, Zakrzewska 1993).

Theoretically, the best indicator parameter would cover all the possible combinations of measurement compared to actual measurements taken because it would carry also the information allowing to optimize the number of measuring elements . For this purpose seems to be ideal to use "neural network", but unfortunately all practical attempts to use this tool for leveling the control networks known to the author were not confirmed in practical application.

BIOGRAPHICAL NOTES

Czaja J., 1993. Wybrane zagadnienia z Geodezji Inżynieryjnej. Wydawnictwa AGH. Kraków

Dąbrowski J., 2009 Zastosowanie metod i algorytmów statystycznych do wyznaczenia parametrów globalnych… Studia i Materiały TNN. Olsztyn, str.31-46.

Ferguson G. A., Takane Y., 1997. Analiza statystyczne w psychologii i pedagogice. Wydawnictwo Naukowe PWN s. 149.

Gocał Jan 2007. Geodezja Inżynieryjno-Przemysłowa Cz. I. Uczelniane Wydawnictwa Naukowo-Dydaktyczne AGH. Kraków.

Kochmański T., 1952 Rachunek Wyrównawczy. Państwowe Wydawnictwa Szkolnictwa Zawodowego, str. 55-66. Warszawa.

Operat geodezyjny „Obwodnica Miasta Wojnicza" 2007.

Collective work Geodezja Inżynieryjna I, II i III 1994. PPWK Warszawa.

Przewłocki Stefan 2009. Geodezja Inżynieryjno-Drogowa. Wydawnictwo Naukowe PWN. Warszawa.

Sobczyk M. 1997. Statystyka, Wydawnictwa Naukowe PWN. Warszawa 1997.

Zakrzewska M. 1993. Słownik terminów analizy czynnikowej.

CONTACTS

Janusz Dąbrowski
Institute of Technical Engineering The Bronisław Markiewicz State Higher School of Technology and Economics in Jarosław
ul. Czarnieckiego 16
37-500 Jarosław

POLAND
Tel. + 48 16 632 76 93
Fax + 48 16 632 76 93
e-mail: jd@pwste.edu.pl
Web site: www.pwste.pl