The elimination of vertical refraction in trigonometric measurements of height differences

Štefan Sokol [1] and ***Marek Bajtala*** [1]

Refraction effect is caused by an inhomogeneous propagation environment, where the measurement is realized. Refraction can be regarded as major source of systematic errors and affects on the reliability and accuracy of surveying measurements. In this article a PCA (Principal Component Analysis) is used for the elimination of multicollinearity between variables and for the reduction to a lower dimension. Subsequently, a regression model with new variables is applied to correct an influence of vertical refraction.

Key words: *Trigonometric heighting, refraction, multicollinearity, PCA (Principal Component Analysis), multiple regression model*

Introduction

The environment in which geodetic measurements are realized is usually inhomogeneous. A Light beam spreading in such environment is attenuated and its path gradually curves. A diffraction of light rays in an inhomogeneous environment is referred to as the refraction of line of sight. The resulting refraction angle can be decomposed into two components according to the spatial curvature of path of the light beam:

- horizontal (lateral) refraction, which affects the measurement of horizontal directions and lengths,
- vertical refraction, which affects the measured zenithal distances or elevation angles.

With the development of instrumentation, also the trigonometric method for the determination of height differences finds its application in solving practical problems. In the past, an accurate determination of length as an intermediate value for the calculation of height difference was limited by available instrumentation. Nowadays, this problem is solved by using total stations, which allow simultaneous angle and distance measurements [2], [13]. However, the influence of environment in the form of vertical refraction remains as a critical factor in the use of this method [5], [16].

This paper is dedicated to repeated trigonometric measurement of height differences and the use of regression model for the reduction of influence of vertical refraction on determined height differences. The results of experimental measurements from the locality of Petržalka served as a basis for the analysis and adjustment.

The experimental measurement

The experimental measurement was realized on July 26, 2007 on points of the setting-out network of the Apollo bridge in Petržalka. A sight line between points VS7 – VS13 with a length of approximately 182 m was the main object of observation. A zenithal distance, horizontal distance and meteorological parameters were measured at the survey station and at the targeted point at two-hour intervals during the 14 hour measurement.

Points are monumented by lined boreholes that are filled with concrete and they are adjusted as pillars on the surface. Pillars are equipped with a steel plate with a hole for forced centering of measuring instruments and targets using a mounting screw. Considering a permanent deep marking of control stations, we can assume that there are no changes in the stability of observed points during the measurement [8], [9].

The Zeiss Trimble 3602 DR was used for measurement of zenithal distances and horizontal distances, while one unit of measurement of zenithal distance consisted of 10 measurements with double targeting in the first and second position of the telescope, corresponding to the one group of 10 determined height differences. The resulting file of height differences consists of 8 groups due to the time difference between individual groups (2 hours – we can assume a change of meteorological conditions between groups). Meteorological parameters, temperature, relative humidity and atmospheric pressure were recorded by using the GFTB 100 device (thermometer, hygrometer, and barometer) or by using the GFTH 200 digital hygrometer/thermometer and GPB 1300 digital barometer.

[1] *prof. Ing. Štefan Sokol, PhD., Ing. Marek Bajtala, PhD.,* Department of Surveying, Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Radlinského 11, 813 68 Bratislava, stefan.sokol@stuba.sk , marek.bajtala@stuba.sk

The processing of measurement results

The results of experiment were processed in a several successive steps. It is necessary to determine whether changes in measured height differences and explanatory variables (meteorological parameters) affect measured height differences in groups so that we can proceed to the multiple regression analysis between these components [3], [4]. The Fisher's test [15], which demonstrated the influence of changes in meteorological parameters on measured height differences at the selected reciprocal level of significance α = 0,05, was used for this purpose.

The independence, or non-correlation of explanatory variables is another precondition for the applicability of multiple regression model. Otherwise, the used model shows a so-called multicollinearity, i.e. variables are approximately linearly dependent [10], [11]. The presence of multicollinearity has several consequences [7]:

- the acceptance of hypothesis of the zero value of regression coefficients, although it actually does not apply,
- regression coefficients can get wrong signs,
- the unacceptability of regression function for predictions and others.

The presence of multicollinearity in our sample is examined using pairwise correlation coefficients (Tab. 1), while its presence is demonstrated if the following applies for any of the correlation coefficients [12]:

$$\left|r_{i,j}\right| > 0{,}80\,. \tag{1}$$

At first, we proceed to the centering and reduction of data according to the following equation, considering various units of variables [1]:

$$X_{i,j} \rightarrow Y_{i,j} = \left(X_{i,j} - \overline{X}_i\right) / \sqrt{D\left(X_i\right)}\,, \tag{2}$$

where:

$i = 1,...,p$, p is the number of variables,

$j = 1,...,n$, n is the number of observations of variable,

$\overline{X}_i$ – arithmetic average of the *i*-th variable,

$D\left(X_i\right)$ – dispersion of the *i*-th variable.

Tab. 1. The selection correlation matrix of centered and reduced explanatory variables.

	temperature$_{stat.}$	**humidity$_{stat.}$**	**pressure$_{stat.}$**	**temperature$_{target}$**	**humidity$_{target}$**	**pressure$_{target}$**
temperature$_{stat.}$	1					
humidity$_{stat}$	-0.99	1				
pressure$_{stat}$	-0.65	0.67	1			
temperature$_{target}$	0.95	-0.97	-0.64	1		
humidity$_{target}$	-0.93	0.93	0.57	-0.97	1	
pressure$_{target}$	-0.27	0.28	0.83	-0.22	0.15	1

We can assume that there is a linear dependence between variables on the basis of correlation coefficients listed in Tab. 1. Although there is no generally valid method to remove multicollinearity, the method of principal components is one of possibilities, which is also used to reduce a dimension, i.e. the number of new variables, so that there is a minimal loss of information [6].

Method of principal components

It is a method for the creation of new, non-correlated variables that are linear combination of original variables. The maximum number of newly created uncorrelated variables is equal to the number of original variables. The solution of principal components is based on the matrix of original variables and their covariance matrix. The covariance matrix is basically the correlation matrix of original variables, if we work with centered and reduced values of variables. A new random variable is obtained according to [14]:

$$Z_i = a_i^T \cdot Y \tag{3}$$

where:

Z_i represents the *i*-th principal component of the Y matrix,

a_i – *i*-th eigenvector of the Σ_Y covariance matrix.

If eigenvalues of the matrix Σ_Y are denoted as $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$ and $a_1,...,a_p$ as their corresponding eigenvectors, than the Z_1 is the first principal component of the matrix Y that represents the largest part of the total variability of data and each other component represents the largest part of the remaining data variability. In our case, the values of eigenvalues are graphically displayed in Fig. 1 and the percentage in the total variability of data is presented in Tab. 2.

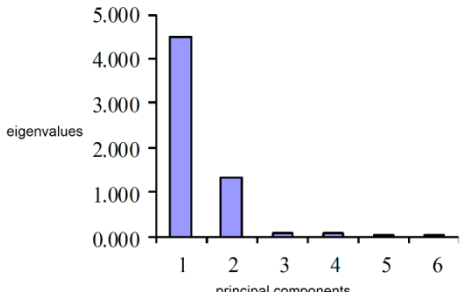

Fig. 1 Graphical representation of eigenvalues

Tab. 2 Eigenvalues

	Eigenvalue	Percentage
1	4.500	75.013
2	1.316	21.937
3	0.093	1.550
4	0.064	1.067
5	0.023	0.383
6	0.003	0.050

As can be seen in Tab. 2, the first two principal components depleted more than 96% of the total variability of data and the share of remaining components is basically negligible. Therefore, it follows that the original six variables (Tab. 3) are replaced by two new and uncorrelated variables Z_1, Z_2 (Tab. 4). The relation between original variables and new values of variables can be assessed by using correlation coefficients. As follows from Tab. 5, the temperature and humidity at survey station and targeted point affect the distribution around the first principal component, while the pressure at the targeted point has the maximum impact on the second component.

Tab. 3. The matrix of original centered and reduced explanatory variables.

temperature$_{stat.}$	-1.964	-1.089	0.233	0.311	0.583	0.992	0.622	0.311
humidity$_{stat}$	1.881	1.164	-0.215	-0.093	-0.841	-0.920	-0.598	-0.379
pressure$_{stat}$	0.880	1.062	0.971	0.562	-0.165	-0.801	-1.164	-1.346
temperature$_{target}$	-1.610	-1.363	-0.106	0.368	1.131	0.884	0.574	0.121
humidity$_{target}$	1.568	1.515	-0.032	-0.835	-0.916	-0.550	-0.625	-0.125
pressure$_{target}$	-0.126	0.883	0.883	0.883	0.883	-1.135	-1.135	-1.135

Tab. 4. The matrix of new variables.

Z_1	3.479	2.895	0.377	-0.323	-1.444	-2.052	-1.765	-1.167
Z_2	-0.989	0.191	1.222	1.318	1.284	-0.657	-0.987	-1.382

Tab. 5. The matrix of new variables.

	temperature$_{stat.}$	**humidity$_{stat.}$**	**pressure$_{stat.}$**	**temperature$_{target}$**	**humidity$_{target}$**	**pressure$_{target}$**
Z_1	-0.97	0.97	0.80	-0.96	0.93	0.44
Z_2	0.18	-0.17	0.57	0.23	-0.30	0.88

Regression model

The regression function of model was determined in the following form, considering the results of the method of principal components, by using which two new and uncorrelated explanatory variables were created:

$$\eta = \beta_0 + \beta_1 \cdot Z_1 + \beta_2 \cdot Z_2 , \tag{4}$$

The estimates b_0, b_1, b_2 of regression parameters $\beta_0, \beta_1, \beta_2$ are determined by the least squares method and subsequently point estimates of the regression function can be calculated according to (Tab. 6):

$$h'_i = b_0 + b_1 \cdot Z_{1i} + b_2 \cdot Z_{2i} \tag{5}$$

where Z_{1i}, Z_{2i} represent values of the *i*-th observation of Z_1, Z_2 .

Tab. 6 The comparison of height differences.

h_{mer} [m]	h' [m]	h_{PN} [m]	$h_{PN} - h_{mer}$ [mm]	$h_{PN} - h'$ [mm]
2.9837	2.9831	2.9785	-5.2	-4.6
2.9807	2.9818		-2.2	-3.3
2.9795	2.9788		-1.0	-0.3
2.9778	2.9781		0.7	0.4
2.9777	2.9772		0.8	1.3
2.977	2.9781		1.5	0.4
2.9785	2.9786		0.0	-0.1
2.9801	2.9794		-1.6	-0.9

Conclusion

The two major components, reproducing around 96 % of the variability of original variables, were extracted using the method of principal components and subsequently the regression function of model was determined by using them. The statistical significance of model's coefficients was assessed using the t-test. The null hypothesis was rejected for all coefficients and the alternative hypothesis, confirming the statistical significance of coefficients, was accepted on the reciprocal level of significance $\alpha = 0{,}05$.

Measured height differences, height differences determined by regression model and their differences from the reference value of height difference, which was obtained by the method of precise leveling, are presented in Tab. 6. As may be seen from the comparison of differences, the constructed regression model contributes to the partial elimination of vertical refraction, by imposing the size of changes of meteorological parameters affecting individual group measurements.

However, it should be noted that the given regression function is not generally applicable also for other cases. This follows from the fact that the principal components are defined by the variability of original variables (meteorological parameters).

***Acknowledgment**: This paper was created by the realization of the grant project No. 1/0133/14 of the Scientific Grant Agency of the Slovak Republic Detection of areal and discrete displacements of unstable areas on the basis of low-cost photogrammetry and satellite technology.*

References

[1] Anděl, Ji. : Mathematic statistic. *Praha: SNTL, ALFA. 1985. 352 p.*.

[2] Baran, M., Pukanská, K., Bartoš, K.: Setting up a geodetic network for the purposes of geodetic documentation of the current state of historical monument Zborov castle. *Geodesy, Cartography and Geographic information systems 2012. 7. scientific and technical international conference: proceedings of articles: 24. - 25.október 2012, Tatranská Lomnica, Slovakia. TU in Košice, FBERG, 2012, p. 1-7. ISBN 978-80-553-1173-9.*

[3] Gašincová S., Gašinec J., Staňková H., Černota P.: Comparison of LSM and alternative estimation methods of the processing results of geodetic measurements. *SDMG 2011: Proceedings of the 18th conference: Praha, 5.-7. 9. 2011. Ostrava, VŠB-TU, 2011, 40-50, ISBN 978-80-248-2489-5.*

[4] Gašincová,S., Gašinec, J., Weiss,G., Labant, S.: Application of robust estimation methods for the analysis of outlier measurements. *GeoScience Engineering. 57, 2011, 3, p. 14-29. ISSN 1802-5420. Available at: http://gse.vsb.cz/2011/LVII-2011-3-14-29.pdf*

[5] Gašinec, J., Gašincová, S.: Estimation Vertical Components of Refraction from Terrestrial and Satellite Geodetic Measuring. *Acta Montanistica Slovaca. 14, 2009, 14, special issue 1, 47-53. ISSN 1335-1788. Available at: http://actamont.tuke.sk/pdf/2009/s1/8gasinec.pdf*

[6] Hajduová, Z., Weiss, E., Mixtaj, L.: Application of statistical methods at copper wire manufacturing. *Metalurgija. 48, 2009, 1, 55-58, ISSN 0543-5846.*

[7] Hančlová, J., Nevima, J.: Regression analysis as a tool of financial analysis. *Available at: http:// moodle.vsb.cz*

[8] Kukučka, P., Weiss, G., Labant, S., Weiss, R.: Comparison of deformation analyses performed on the basis of GNSS observations using single-frequency and double-frequency measurement receivers. *Acta Montanistica Slovaca. 17, 2012, 3, 195-203, ISSN 1335-1788 Available at: http://actamont.tuke.sk/pdf /2012/n3/9kukucka.pdf.*

[9] Labant, S., Weiss, G., Zuzik, J., Baran, M.: Graphical interpretation deformation analysis of stability area using of strain analysis. *Acta Montanistica Slovaca, 19, 2014, 1, 31-40. ISSN 1335-1788. Available at: http://actamont.tuke.sk/pdf/2014/n1/5labant.pdf*

[10] Labant, S., Weiss, G.: Adjustment of geodetic network and deformation analysis. *1st edition, Ostrava, VŠB - Technical University of Ostrava, 2012, 137 p. ISBN 978-80-248-3220-3.*

[11] Labant, S., Weiss, G.: Analysis of a simulation of missing satellite observations in the deformation network. *Acta Montanistica Slovaca. 17, 2012, 3, 158-166, ISSN 1335-1788. Available at: http://actamont.tuke.sk/pdf/2012/n3/3labant.pdf*

[12] Pavelka, F., Klímek, P. : Applied Statistics. *Brno University of Technology. 2000, 129 p. ISBN 80-214-1545-2. Available at: http://www.vscht.cz/ktk/www_324/lab/texty/statistika/as.pdf*

[13] Rákay, Š. jr., Rákay, Š., Weiss, R., Zuzik, J.: Adjusting wheel alignment on machinery equipment. *Acta Montanistica Slovaca. 17, 2012, 4, 348-353, ISSN 1335-1788. Available at: http://actamont.tuke.sk/pdf/2012/n4/17rakay.pdf.*

[14] Rao, R, C.: Linear methods of statistical inference and its applications. *Praha: ACADEMIA. 1978. 668 p.*

[15] Sokol, Š. Effect of atmospheric conditions on trigonometric measurement of heights. *Bratislava: STU. 2002, 80 p., ISBN 80-227-1695-2.*

[16] Šíma, J., Seidlová, A.: Geodetic work for project of chairlift. *XXII: Slovak – Polish – Russian Seminar „Theoretical Foundation of Civil Engineering". Žilina, 9.9. – 13.9.2013. Moskva 2013, p. 521-528. ISBN 978-5-93093-986-6.*