Sabine Bachmann et al.: IVS Combination Center at BKG - Robust Outlier Detection and Weighting Strategies, IVS 2012 General Meeting Proceedings, p.266–270
http://ivscc.gsfc.nasa.gov/publications/gm2012/bachmann2.pdf

IVS Combination Center at BKG - Robust Outlier Detection and Weighting Strategies

Sabine Bachmann, Michael Lösler

BKG, Federal Agency for Cartography and Geodesy, Germany

Contact author: Sabine Bachmann, e-mail: `sabine.bachmann@bkg.bund.de`

Abstract

Outlier detection plays an important role within the IVS combination. Even if the original data is the same for all contributing Analysis Centers (AC), the analyzed data shows differences due to analysis software characteristics. The treatment of outliers is thus a fine line between keeping data heterogeneity and elimination of real outliers. Robust outlier detection based on the Least Median Square (LMS) is used within the IVS combination. This method allows reliable outlier detection with a small number of input parameters. A similar problem arises for the weighting of the individual solutions within the combination process. The variance component estimation (VCE) is used to control the weighting factor for each AC. The Operator-Software-Impact (OSI) method takes into account that the analyzed data is strongly influenced by the software and the responsible operator. It allows to make the VCE more sensitive to the diverse input data. This method has already been set up within GNSS data analysis as well as the analysis of troposphere data. The benefit of an OSI realization within the VLBI combination and its potential in weighting factor determination has not been investigated before.

1. Introduction

Outliers distort the results of a least squares adjustment, if the estimator is not robust. In particular, leverage points involve useless results. Therefore, alternative methods, which are highly resistant to leverage points, have to be introduced for outlier detection. The Least Median of Square method (LMS) leads to a robust estimator and is implemented within the IVS combination. A robust outlier detection provides reliable data (random errors excepted) which is used for a variance component estimation. Precedent studies showed that the “formal error” of the combined solution is underestimated with respect to the individual solutions [2]. Applying the Operator-Software-Impact method (OSI) to the VLBI combination is a first attempt to numerically confirm the empirically determined scaling factor of $\sqrt{2}$.

2. Least Median of Square

[7] evaluates different robust estimators in the field of VLBI data analysis and concludes that the so-called BIBER estimator by [12] is an efficient and reliable estimator. During the preprocessing of the combination, the comparable data is limited to the number of contributing ACs. Hence, only a consistent check of earth orientation parameters (EOP) and station coordinates can be carried out. An employment of the BIBER estimator promises no prospect of improvement, because this estimator copes with a number of outliers of about 10-40% [8], [10]. If four solutions are available, this limit is reached if one of them is an outlier. LMS is a robust estimator with a theoretical breakdown point of 50% [9]. This means the LMS provides reliable results even if up to 50% of contaminated data exist. The objective function is given by

$$\min \operatorname*{med}_i \mathbf{v}_i^2 \tag{1}$$

where $\mathbf{v}$ is the vector of residuals. Outliers or leverage points ∇_i can be identified by comparing the robust standardized residuals to a given threshold k

$$\nabla_i = \begin{cases} \text{false, if } |v_i| \leq k\sigma_{LMS} \\ \text{true, otherwise} \end{cases} \tag{2}$$

with

$$\sigma_{LMS} = 1.4826 \left(1 + \frac{5}{n-u}\right) \sqrt{\min \operatorname*{med}_i \mathbf{v}_i^2} \tag{3}$$

where n and u depict the number of observations and unknowns, respectively [9]. If outliers are detected and the parameters are assuredly wrong, the associated AC will be excluded from the combination.

3. Weighting Strategies

The ACs are producing independently analyzed data, although the input data is similar. This means that no accuracy relations between the observations (i.e., the individual solutions of the ACs) in terms of co-factor matrices are given. Individual (group) weighting factors are introduced via variance component estimation [5]. Two approaches of VCE are applicable: with respect to the observation, and with pseudo observations. For the group definition every AC can be defined as one group, or, taking into account similarities between analysis software used by the ACs, groups are defined by uniting ACs which use the same software package. The latter approach better reflects the effective influence of the analyzed data.

3.1. Variance Component Estimation with Respect to the Observations (VCEO)

[5] proposes a method to estimate the variance components of observation sets which are stochastically independent. The observation vector $\mathbf{l}_c$ is given by

$$\mathbf{l}_c = (\mathbf{l}_1 \, \mathbf{l}_2 \, \cdots \, \mathbf{l}_n)^T \tag{4}$$

where each $\mathbf{l}_i$ describes a grouped observation subset of $\mathbf{l}_c$. If the $\mathbf{l}_i$ are uncorrelated, the stochastic model $\mathbf{C}_c = \mathbf{P}_c^{-1}$ has a block diagonal structure.

$$\mathbf{C}_c = \text{blockdiag} \, \left(\sigma_1^2 \mathbf{Q}_1 \, \sigma_2^2 \mathbf{Q}_2 \, \cdots \, \sigma_n^2 \mathbf{Q}_n\right) \tag{5}$$

Together with the design matrix $\mathbf{A}_c$, which contains the partial derivative with respect to the unknown parameters $\mathbf{x}_c$, the solution of the normal equation can be written as

$$\mathbf{x}_c = \left(\mathbf{A}_c^T \mathbf{P}_c \mathbf{A}_c\right)^{-1} \mathbf{A}_c^T \mathbf{P}_c \mathbf{l}_c = \mathbf{N}_c^{-1} \mathbf{n}_c \tag{6}$$

and the variance components for each observation set can be derived by

$$\sigma_i^2 = \frac{\mathbf{v}_i^T \mathbf{P}_i \mathbf{v}_i}{r_i} = \frac{\left(\mathbf{A}_i \mathbf{x}_c - \mathbf{l}_i\right)^T \mathbf{P}_i \left(\mathbf{A}_i \mathbf{x}_c - \mathbf{l}_i\right)}{n_i - \text{trace}\left(\mathbf{N}_c^{-1} \mathbf{A}_i^T \mathbf{P}_i \mathbf{A}_i\right)} = \frac{\mathbf{x}_c^T \mathbf{N}_i \mathbf{x}_c - 2\mathbf{x}_c^T \mathbf{n}_i + \mathbf{l}_i^T \mathbf{P}_i \mathbf{l}_i}{n_i - \text{trace}\left(\mathbf{N}_c^{-1} \mathbf{N}_i\right)} \tag{7}$$

where $\mathbf{v}$ is the vector of residuals, r the part of redundancy, and n the number of observations. Förstner's method can be applied, if the system of normal equations $\mathbf{n}_i = \mathbf{N}_i \mathbf{x}_i$, the weighted sum of reduced observations $\mathbf{l}_i^T \mathbf{P}_i \mathbf{l}_i$, and the number of observations n_i are given.

3.2. Variance Component Estimation with Pseudo Observations (VCEP)

Another way to estimate variance components is to treat each single solution $\mathbf{x}_i = \mathbf{l}_i$ as the realization of a stochastic process, whereas $\mathbf{N}_i^{-1}$ is considered as covariance matrix of $\mathbf{l}_i$. Thus, the combination can be interpreted as weighted mean with a simple functional model $\mathbf{A}_c$

$$\mathbf{A}_c = \mathbf{J}_{n,1} \otimes \mathbf{I}_n \tag{8}$$

where $\mathbf{J}_{n,1}$ is a $n \times 1$ matrix of ones, $\mathbf{I}_n$ refers to the $n \times n$ identity matrix, and n is the number of observations. Again, the stochastic model (9) has block diagonal structure and is made up of $\mathbf{N}_i$. Therefore, Förstner's method can be applied by solving (6) and (7).

$$\mathbf{C}_c = \text{blockdiag}\ \left(\sigma_1^2 \mathbf{N}_1^{-1}\ \sigma_2^2 \mathbf{N}_2^{-1}\ \cdots\ \sigma_n^2 \mathbf{N}_n^{-1}\right) \tag{9}$$

Both approaches are implemented within the IVS combination. Practically, the VCEO approach has a lower computing time, but the VCEP approach better reflects the real proportion of the influences of the data input.

3.3. First Investigation with Operator-Software-Impact (OSI)

Förstner's method is suggested by [11] for intra-technique combination. The assumption is that due to the different treatment of the observations by the ACs, the resulting normal equations $\mathbf{N}_i$ are independent, although each AC analyzes the same initial observations. [2] introduced VCEO into the combination process at the level of normal equations. Within the VCEO as well as within the VCEP, independent normal equations are assumed, which cannot be ensured. [7] suggests an advanced analysis method called Operator-Software-Impact (OSI) method. The basic strategy is to split up the vector of observations $\mathbf{l}_i$ in the following way

$$\mathbf{l}_i = \mathbf{l}_{init} + \delta \mathbf{l}_i \tag{10}$$

$\mathbf{l}_{init}$ describes the initial observations and $\delta \mathbf{l}_i$ denotes the additional modification of $\mathbf{l}_{init}$ due to individual reduction and correction steps at each AC. $\mathbf{l}_{init}$ and $\delta \mathbf{l}_i$ are uncorrelated pair by pair.

$$\mathbf{C}_{osi} = \sigma_0^2 \begin{pmatrix} \mathbf{Q}_{l_{init} l_{init}} + \mathbf{Q}_{\delta l_1 \delta l_1} & \mathbf{Q}_{l_{init} l_{init}} & \mathbf{Q}_{l_{init} l_{init}} & \mathbf{Q}_{l_{init} l_{init}} \\ \mathbf{Q}_{l_{init} l_{init}} & \mathbf{Q}_{l_{init} l_{init}} + \mathbf{Q}_{\delta l_2 \delta l_2} & \mathbf{Q}_{l_{init} l_{init}} & \mathbf{Q}_{l_{init} l_{init}} \\ \mathbf{Q}_{l_{init} l_{init}} & \mathbf{Q}_{l_{init} l_{init}} & \ddots & \vdots \\ \mathbf{Q}_{l_{init} l_{init}} & \mathbf{Q}_{l_{init} l_{init}} & \cdots & \mathbf{Q}_{l_{init} l_{init}} + \mathbf{Q}_{\delta l_n \delta l_n} \end{pmatrix} \tag{11}$$

While $\mathbf{l}_{init}$ is used by each AC, $\delta \mathbf{l}_i$ considers the influence of the individual AC as random noise with zero mean. As seen above, the OSI models the different analysis strategies as stochastic effect.

In Equation (11) $\mathbf{Q}_{l_{init}l_{init}}$ can be produced by the law of error propagation. In general, the matrix $\mathbf{Q}_{\delta l_1 \delta l_1}$ is unknown. Therefore, [7] propose two possible mathematical approaches to complete the stochastic model $\mathbf{C}_{osi}$ with the OSI-parameters $\alpha_i^2 > 0$.

$$\mathbf{Q}_{\delta l_i \delta l_i} = \alpha_i^2 \mathbf{I} \qquad \text{or} \qquad \mathbf{Q}_{\delta l_i \delta l_i} = \alpha_i^2 \mathbf{Q}_{l_{init}l_{init}} \tag{12}$$

To quantify the individual OSI-parameters α_i^2, a variance component estimation of Helmert type can be applied [1], [3].

$$\mathbf{C}_{osi} = \sigma_{l_{init}}^2 \begin{pmatrix} \mathbf{Q}_{l_{init}l_{init}} & \cdots & \mathbf{Q}_{l_{init}l_{init}} \\ \vdots & \ddots & \vdots \\ \mathbf{Q}_{l_{init}l_{init}} & \cdots & \mathbf{Q}_{l_{init}l_{init}} \end{pmatrix} + \sigma_{\delta l_1}^2 \begin{pmatrix} \mathbf{Q}_{\delta l_1 \delta l_1} & 0 \\ 0 & 0 \end{pmatrix} + \cdots + \sigma_{\delta l_n}^2 \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{\delta l_n \delta l_n} \end{pmatrix} \tag{13}$$

According to [4] and [6], the individual OSI-parameters are given through

$$\left(\alpha_1^2 \, \alpha_2^2 \, \cdots \, \alpha_n^2\right)^T = \sigma_{l_{init}}^{-2} \, \left(\sigma_{\delta l_1}^2 \, \sigma_{\delta l_2}^2 \, \cdots \, \sigma_{\delta l_n}^2\right)^T \tag{14}$$

Finally, the α_i^2 has to be introduced to the conventional accumulated combination to correct the estimated covariance matrix $\mathbf{N}_c^{-1}$. The relationship f_{osi} between the traditional accumulated combination approach and the OSI combination approach is given by [4]

$$f_{osi} = \frac{f_1}{f_2}, f_1 = 1 + \frac{1}{\sum\limits_{i=1}^{n} \frac{1}{\alpha_i^2}}, f_2 = \frac{1}{\sum\limits_{i=1}^{n} \frac{1}{1+\alpha_i^2}} \tag{15}$$

and in conclusion

$$\mathbf{N}_{osi} = \frac{1}{f_{osi}} \mathbf{N}_c \tag{16}$$

Figure 1 shows the scaling factor f_{osi} estimated with the OSI method.

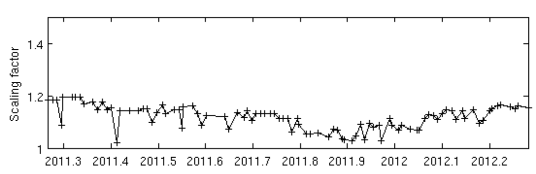

Figure 1. Scaling factor for one year of data for combined solution estimated with OSI.

4. Summary and Outlook

Introducing the LMS into the outlier test procedure leads to a robust and reliable outlier detection within the VLBI combination even if the number of observations is small. Regardless of the independent normal equations, the use of VCEO increases the degree of freedom disproportionately high, because the number of initial observations is multiplied by the number of ACs and treated as independent observations. The VCEP method does not use information about the number of initial observations during the variance component estimation. Unfortunately, the convergence of VCEP is noticeably inferior and sometimes fails. The aim of the OSI method was to proof the empirically determined scaling factor ($\sqrt{2}$) for the combined solution with respect to the individual solution as described in [2]. A first investigation of the OSI method for the VLBI combination showed a slightly lower influences factor of 1.1 to 1.2. Improving weighting strategies and investigating the individual influences of the contributing ACs will be continued at BKG.

References

[1] Bähr, H., Altamimi, Z., Heck, B.: Variance Component Estimation for Combination of Terrestrial Reference Frames. Universitätsverlag Karlsruhe, 2007/6, ISBN: 978-3-86644-206-1, 2007.

[2] Böckmann, S., Artz, T., Nothnagel, A., Tesmer, V.: International VLBI Service for Geodesy and Astrometry: Earth orientation parameter combination methodology and quality of the combined products, J. of Geophysical Research, Vol. 115, doi:10.1029/2009JB006465, 2010.

[3] Crocetto, N., Gatti, M., Russo, P.: Simplified formulae for the BIQUE estimation of variance components in disjunctive observation groups. J Geod, Vol. 74, pp 447-457, doi:10.1007/s001900000109, 2000.

[4] Fang, X.: Statistische Analyse von Ausgleichungsergebnissen auf der Basis teilweise identischer Messwerte. Diploma Thesis, Geodätisches Institut, Leibniz Universität Hannover, 2007.

[5] Förstner, W.: Ein Verfahren zur Schätzung von Varianz- und Kovarianzkomponenten. Allgemeine Vermessungs-Nachrichten, 11-12, pp 446-453, 1979.

[6] Heinkelmann, R., Böhm, J., Bolotin, S., Engelhardt, G., Haas, R., Lanotte, R., MacMillan, D., Negusini, M., Skurikhina, E., Titov, O., Schuh, H.: VLBI-derived troposphere parameters during CONT08. J Geod, Vol. 85, doi:10.1007/s00190-011-0459-x, 2011.

[7] Kutterer, H., Heinkelmann, R., Tesmer, V.: Robust outlier detection in VLBI data analysis. In: Schwegmann, W., Thorandt, V. (eds) Proceedings of the 16th EVGA Working Meeting. Verlag des Bundesamts für Kartographie und Geodäsie, Leipzig Frankfurt/Main, pp 247-256, 2003.

[8] Neitzel, F.: Identifizierung konsistenter Datengruppen am Beispiel der Kongruenzuntersuchung geodätischer Netze. Müchen, Bayer. Akademie d. Wissenschaften, Deutsche Geodätische Kommission (DGK), Reihe C, Heft-Nr. 565. 2003.

[9] Rousseeuw, P.J., Leroy, A.M.: Robust Regression and Outlier Detection. John Wiley & Sons, New Jersey, ISBN: 978-0471488552, 2003.

[10] Salvini, D.: Konzeption und Entwicklung neuer interaktiver multimedialer Lern- und Arbeitsmethoden für die geodätische Ausgleichungsrechnung. ETH Zürich, Institut für Geodäsie und Photogrammetrie, Nr. 17968, 2008.

[11] Vennebusch, M., Böckmann, S., Nothnagel, A.: The contribution of Very Long Baseline Interferometry to ITRF2005. J Geo, Vol 81, pp 553-564, doi:10.1007/s00190-006-0117-x, 2006.

[12] Wicki, F.: Robuste Schätzverfahren für die Parameterschätzung in geodätischen Netzen. Institute of Geodesy and Photogrammetry, ETH Zürich, Switzerland, 1999.