Hansjörg Kutterer: Reliability Measures for Geodetic VLBI Products, IVS 2004 General Meeting Proceedings, p.301–305
http://ivscc.gsfc.nasa.gov/publications/gm2004/kutterer

Reliability Measures for Geodetic VLBI Products

Hansjörg Kutterer

Deutsches Geodätisches Forschungsinstitut (DGFI)
e-mail: kutterer@dgfi.badw.de

Abstract

The reliability of geodetic VLBI products depends essentially on the checkability of the observation data and the reference frame points. In statistics, reliability has two meanings. The first one deals with the detectability of incomplete or incorrect parts of the model by means of statistical hypothesis tests (internal reliability). The second one studies the impact of non-detectable model errors on the estimated parameters (external reliability). In this paper the theoretical background of reliability is shortly reviewed. The focus lies on its application to the estimation procedure with respect to possible errors in the reference frame. The potential influence of non-detectable errors in the station and source coordinates on VLBI products is presented and discussed based on the data from the CONT02 campaign.

1. Theoretical Background

Precision and reliability are quality issues of their own right. Whereas precision measures such as formal errors or error ellipses, respectively, can already be computed from minimum observation configurations, reliability measures are only meaningful in the case of sufficient partial redundancy. In other words, each observation must be checkable by at least some other observations. Such a check is based on a statistical hypothesis test.

Two different forms of reliability can be distinguished. On the one hand, internal reliability describes the detectability of errors in the data and is quantified by means of, e.g., the marginally detectable blunder in a hypothesis test with a certain Type I error probability α. The associated probability for successful blunder detection is the test power $1-\beta$ with β the Type II error probability. On the other hand, external reliability means the maximum influence of a non-detected blunder on the parameters of interest. For details see [1] or [2].

In VLBI data three types of data are of interest: the original delay observations, the parameter soft constraints in terms of pseudo-observations, and hard constraints for datum definition in terms of condition equations. In this paper only the third type is considered as the focus lies on possible outliers in the (terrestrial or celestial) reference frame positions. The terrestrial datum components are defined by means of a NNT (No Net Translation) and a NNR (No Net Rotation) condition with respect to the ITRF2000. The celestial datum components are defined by means of a NNR condition with respect to the ICRF.

The basic test idea is that the datum definition should be identical (in a statistical meaning) for each arbitrarily chosen set of datum points. Hence, the estimated parameters based on a datum definition using all stations and sources should be approximately the same as those derived from a datum definition when, e.g., one arbitrary datum point is discarded. In the following this is briefly described in mathematical terms.

The datum defining condition equation for the parameter updates $\mathbf{\Delta x}$ reads as

$$\mathbf{G}\ \mathbf{\Delta x} = \mathbf{0}. \quad (1)$$

The constraint matrix $\mathbf{G}$ is represented in transposed form point by point as

$$\mathbf{G}^T = \begin{bmatrix} \mathbf{G}_1^T & \mathbf{G}_2^T & \dots & \mathbf{G}_n^T \end{bmatrix}. \tag{2}$$

with

$$\mathbf{G}_i = \begin{bmatrix} \mathbf{G}_{terr,NNT,i} & \mathbf{G}_{terr,NNR,i} & \mathbf{G}_{cel,NNR,i} \end{bmatrix}. \tag{3}$$

In case of terrestrial datum points (stations) it is

$$\mathbf{G}_{terr,NNT,i} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{G}_{terr,NNR,i} = \begin{bmatrix} 0 & -z_i & y_i \\ z_i & 0 & -x_i \\ -y_i & x_i & 0 \end{bmatrix}, \quad \mathbf{G}_{cel,NNR,i} = \mathbf{0} \tag{4}$$

and in case of celestial datum points (radio sources)

$$\mathbf{G}_{terr,NNT,i} = \mathbf{0}, \quad \mathbf{G}_{terr,NNR,i} = \mathbf{0}, \quad \mathbf{G}_{cel,NNR,i} = \begin{bmatrix} \cos(RA) \cdot \sin(DE) & -\sin(RA) \\ \sin(RA) \cdot \sin(DE) & \cos(RA) \\ \cos(DE) & 0 \end{bmatrix}. \tag{5}$$

The remaining entries are all equal to 0. The definition of two different geodetic datums based on the same set of reference points can be described by

$$\mathbf{G}_1 = \mathbf{P}_1\mathbf{G}, \quad \mathbf{G}_2 = \mathbf{P}_2\mathbf{G} \tag{6}$$

where $\mathbf{P}_1$ and $\mathbf{P}_2$ denote two selection matrices, i.e., diagonal matrices with entries 1 on the main diagonal for datum point coordinates and 0 for all other parameters. The parameter differences and their cofactor matrix, respectively, are then given by

$$\mathbf{w} = (\mathbf{S}_2 - \mathbf{S}_1)\Delta\mathbf{x}, \quad \mathbf{Q}_{\mathbf{ww}} = (\mathbf{S}_2 - \mathbf{S}_1)\ \mathbf{Q}_{\mathbf{xx}}\ (\mathbf{S}_2 - \mathbf{S}_1)^T. \tag{7}$$

with $\mathbf{Q}_{\mathbf{xx}}$ the cofactor matrix of $\boldsymbol{\Delta}\mathbf{x}$ and the matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ for the projection of $\boldsymbol{\Delta}\mathbf{x}$ into the two different datums

$$\mathbf{S}_1 = \mathbf{I} - \mathbf{G}\left(\mathbf{G}^T\mathbf{P}_1\mathbf{G}\right)^{-1}\mathbf{G}^T\mathbf{P}_1, \quad \mathbf{S}_2 = \mathbf{I} - \mathbf{G}\left(\mathbf{G}^T\mathbf{P}_2\mathbf{G}\right)^{-1}\mathbf{G}^T\mathbf{P}_2. \tag{8}$$

A proper test statistic for the significance of the parameter difference is given by

$$T = \frac{1}{\sigma^2}\mathbf{w}^T\mathbf{Q}_{\mathbf{ww}}^{-}\mathbf{w} \sim \begin{cases} \chi^2_{3,\lambda}, & \text{for terrestrial stations} \\ \chi^2_{2,\lambda}, & \text{for radio sources} \end{cases} \tag{9}$$

with the formal variance factor σ^2, the superscript "-" denoting a reflexive generalized inverse and with the non-centrality parameter $\lambda = \lambda(\alpha, \beta)$ which is characterized by

$$\lambda \quad \begin{cases} = 0 & |H_0 : E(\mathbf{w}) = \mathbf{0} \\ > 0 & |H_a : E(\mathbf{w}) \neq \mathbf{0} \end{cases} \tag{10}$$

Here H_0 denotes the null hypothesis ("identical datum by $\mathbf{G}_1$ and $\mathbf{G}_2$") and H_a the alternative hypothesis ("non-identical datum").

The results presented in the following sections were derived from 14 days (October 17-30, 2002) of the CONT02 campaign. For the computations a preliminary extension of the software package OCCAM 6.0 (see [3]) was used permitting the estimation of source coordinates. The source and station coordinates were estimated from the normal equations system which was accumulated over the complete campaign. All other parameters were removed in advance in a rigorous way.

2. Influence of Datum Point Selection

2.1. Station Effects

In a first scenario the terrestrial stations are tested according to Eq. (9). Two datum definitions are compared. The first datum is defined by the coordinates of all stations. The second datum is defined by taking all points but one as datum points. Hence, the greater the value of the test statistics, the less compatible is the estimated position of this point and its reference frame position. If all values are small or at least on the same level, the null hypothesis seems to be quite likely. A clear hint regarding a lacking compatibility is given if there is a particular test value significantly greater than the others.

The results are shown in figure 1. If all eight stations which participated in the CONT02 campaign are selected as datum points, the test statistic values presented in the back row are obtained (dark grey). There is a clear indication of ONSALA60 and a minor indication of GILCREEK. In a next step ONSALA60 is discarded in advance so that it does not influence the reference datum. The corresponding values are given in the front row (light gray). Note that GILCREEK is not indicated any longer. The situation looks now quite homogeneous. Obviously the ITRF 2000 position and the estimated position of ONSALA60 during the CONT02 campaign do not agree well. Taking some reference frame noise level into account the positions of the other stations are rather consistent.

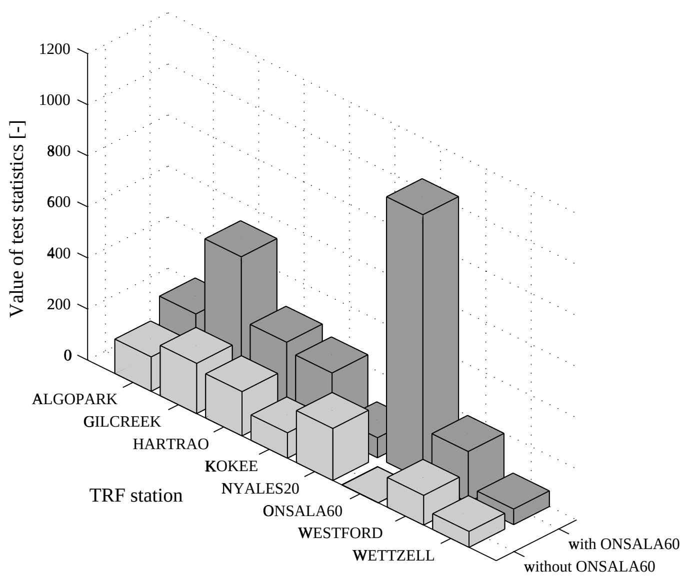

Figure 1. Test statistic values if the respective station is discarded from the datum definition. If ONSALA60 is discarded before testing, the values of the other stations are reduced and homogenized significantly.

The most important results concerning possible discrepancies between the estimated and the reference positions are the following. The coordinate uncertainty induced by the datum point selection is about 2 mm, the biases detectable by the test according to Eq. (9) (considering $\lambda = 10.90$ derived from $\alpha = 0.05$ and $\beta = 0.2$) are less than 3.6 mm on average and less than 10.7 mm at maximum. Hence, the datum point selection for the CONT02 is rather uncritical for the estimated station positions.

2.2. Radio Source Effects

In a second scenario the radio source positions are tested as it was outlined in Section 2.1. Again, the initial datum was defined using all sources. Afterwards, one particular source was discarded from the definition of the datum. This was done in turn for all sources. The values of the test statistics are presented in figure 2. It can clearly be seen that some of the sources have significantly larger values than the others. Again, this points to discrepancies between the estimated and the reference positions.

Reliability ellipses (see figure 3) can be calculated based on the test statistics according to Eq. (9) in order to assess the detectable errors in a reference point position. These ellipses represent the locus of an erroneous position which can be detected with $\beta = 0.2$ when testing with $\alpha = 0.05$ according to Eq. (9). The smaller the ellipses, the higher is the reliability. Significant semimajor axes indicate directions with low reliability. Hence, the sources with mean declination are rather reliable whereas the others reveal some deficiencies. Note that the size and orientation of the ellipses is mainly caused by the session configuration and scheduling. There is no clear assignment of the reliability measures to the ICRF classes ("defining", "candidate", "others").

The coordinate uncertainty of the radio sources induced by the datum point selection is about 10 μas, the biases detectable by the test according to Eq. (9) (considering $\lambda = 9.63$ derived from $\alpha = 0.05$ and $\beta = 0.2$) are less than 0.24 μas on average and less than 3 mas at maximum. Hence, the datum point selection for the CONT02 is mostly uncritical for the estimated source positions. Note that there isn't any impact on the estimated station positions and vice versa.

2.3. EOP Effects

In order to assess the influence of possibly erroneous datum point positions on the EOP, the reference datum was defined using all stations and all sources. An alternative datum was defined by discarding the station ONSALA60 and the radio sources 0106+013 and 4C39.25 which were indicated by the tests. The effect of the two different datums on the EOP turned out to be mainly a constant offset. For the pole coordinates -14 μas (x-component) and +45 μas (y-component) were found. The offset in Universal Time (ΔUT1) yielded +13 ms. In case of nutation the offset in obliquity was stable (+4 μas) whereas the values of the offset in longitude varied about an average value of -10 μas. Although these effects are rather small, they indicate the possible amount of biases in the parameters due to typical inconsistencies in the reference frames.

3. Conclusions

Proper reliability measures for VLBI products can be derived in a rigorous way using statistical test theory as background. The reliability analysis of the CONT02 campaign is well-suited for the illustration of the procedure but the presented results must not be overrated. They are certainly capable to give some background for the interpretation of the CONT02 results. For a thorough assessment of station and source coordinates it is rather recommended to assess a large number of different VLBI sessions in the outlined way.

At present, there is one open question concerning methodology which will soon be solved: the presented reliability measures are relative as an absolute accuracy level was not yet introduced. For this purpose the test statistic according to Eq. (9) needs to be extended in order to handle empirical variances in a statistically proper way. This will be solved soon.

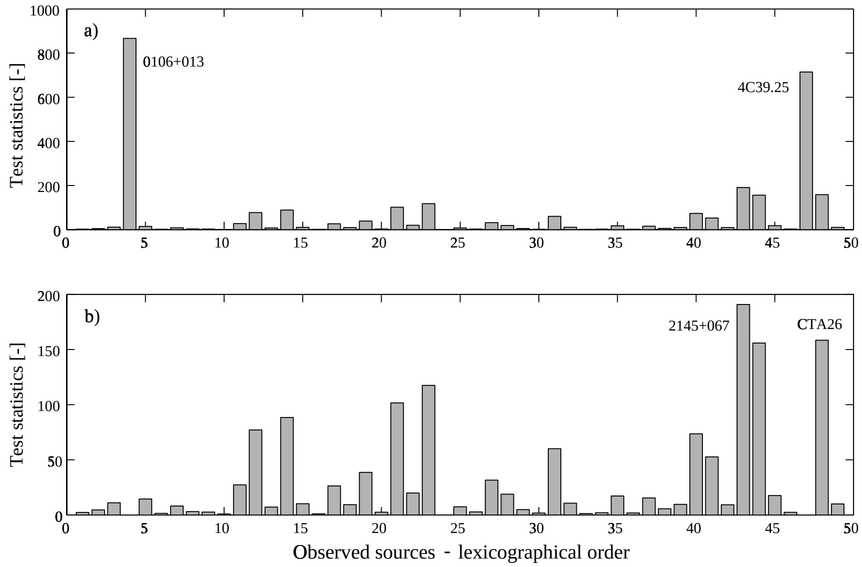

Figure 2. Test statistic values if the respective radio source is discarded from the datum definition. Note the different ranges on the y-axis. Some sources with significant test values are named. a) Datum defined using all radio sources. b) Datum defined as in a), but without the sources 0106+013 and 4C39.25.

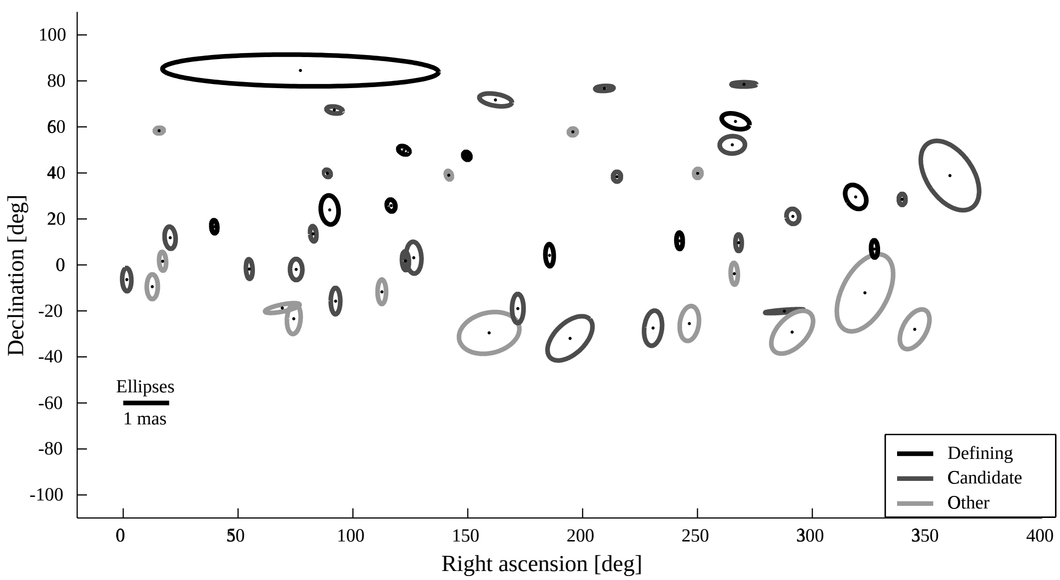

Figure 3. Reliability ellipses for the estimated radio source positions (Legend: ICRF classification).

References

[1] Baarda, W., A testing procedure for use in geodetic networks, Publications on Geodesy, Vol. 2, No. 5, Netherlands Geodetic Commission, Delft, 1968.

[2] Koch, K.R., Parameter Estimation and Hypothesis Testing in Linear Models, Springer, Berlin, 1999.

[3] Titov, O., J. Boehm, V. Tesmer, OCCAM v.6.0 software for VLBI data analysis, In: IVS 2004 General Meeting Proceedings, edited by N. R. Vandenberg and K. D. Baver, NASA/CP-2004-212255.