Christoph Steinforth et al.: Outlier Detection in the Combination of VLBI EOP, IVS 2002 General Meeting Proceedings, p.265–271
http://ivscc.gsfc.nasa.gov/publications/gm2002/steinforth

Outlier Detection in the Combination of VLBI EOP

Christoph Steinforth, Axel Nothnagel

Geodetic Institute of the University of Bonn
Contact author: Christoph Steinforth, e-mail: steinforth@uni-bonn.de

Abstract

On March 1, 2002, a new combined IVS EOP series has been established which is linked directly to the ITRF2000 reference frame. As compared to the input series submitted previously, the ITRF2000 based input series provided by the IVS Analysis Centers agree much better. In the process of automatic combination, the detection of outliers is of great importance in order to reduce human intervention to a minimum. The procedure used for outlier elimination responds to the level of the scatter in the residuals. As a consequence the new combination seems to be more robust against outliers.

1. Introduction

For almost 18 months now, the combined IVS Earth orientation parameter (EOP) series IVS01001 has been based on a number of different realizations of terrestrial reference systems and an alignment to the IERS C04 series [2] [3]. Depending on the quality of the nutation components of the input series, weighting factors have been assigned which are being updated from time to time. Current biases as subtracted before the combination and weight factors are summarized in table 1. A weight factor smaller than one means that the input of the respective series is downweighted correspondingly in the combination process.

Table 1. Biases and weight factors used in IVS01001 (determination period: 1.1.1999 - 31.9.2000); Weight factors greater than 1 increase the weight of the input series.

	x_p [μas]	y_p [μas]	$dUT1$ [μs]	w.f. [-]
AUS	-35.9	367.3	6.9	1.22
BKG	91.2	-8.0	-17.0	0.93
GSF	-43.0	281.7	8.4	0.90
IAA	-135.5	263.9	9.3	1.00
SPU	-46.3	155.1	16.6	1.11
USN	2.3	295.7	6.5	0.90

Comparing the formal errors of the input series as reported by the IVS Analysis Centers with the scatter of the post fit residuals as represented by the weighted RMS differences helps to assess the precision of the input series. In table 2 the mean formal errors of all input series and all EOP components show, on average, a fairly good level of agreement. As expected, some components have a higher post combination scatter than anticipated from the formal errors. However, most Analysis Centers show higher formal errors than post combination scatter except for the UT1-UTC component.

The mean formal precision of the combined series IVS01001 as reported in table 2 is reduced considerably as compared to the input series. Still lacking is the proper treatment of correlations between the input series which are expected to be as high as 70% due to the use of almost identical

observations, it seems quite natural that the formal precision of the combination series is so much better.

Table 2. Mean formal errors (MFE) of input and WRMS relative to combination IVS01001 (from 1.1.1999 to 31.12.2001)

	$x_p[\mu as]$		$y_p[\mu as]$		$dUT1[\mu s]$		$d\epsilon[\mu as]$		$d\psi \sin\epsilon_0[\mu as]$	
	MFE	WRMS	MFE	WRMS	MFE	WRMS	MFE	WRMS	MFE	WRMS
AUS	114.1	122.7	94.0	118.5	5.1	6.3	63.9	61.1	65.1	57.3
BKG	109.1	88.1	90.1	86.9	5.3	7.1	65.4	61.1	66.5	67.8
GSF	109.4	96.4	88.7	86.1	4.6	6.8	80.2	61.2	82.4	67.4
IAA	92.9	78.5	78.8	72.0	4.0	4.1	70.2	58.6	71.5	65.1
SPU	90.2	61.0	73.0	59.3	3.9	4.4	63.9	50.7	65.4	58.1
USN	109.2	98.3	88.4	70.6	4.6	3.5	80.4	73.1	82.3	76.1
mean formal precision of combination										
IVS	61.2		51.4		2.7		37.8		38.6	

A more realistic assessment of the errors can be carried out by comparing the IVS combined series with EOP series of other techniques or combination products where other techniques play an important role. For this reason, the combined IVS EOP series is regularly compared with the IERS C04 and the Bulletin B series [5]. Most important is the comparison of the combined VLBI EOP series with an EOP series derived from the results of an independent technique of similar quality, for example GPS. Weighted mean differences are -11.5 μas, 23.1 μas and -3.5 μs for x pole, y pole and UT1-UTC with weighted RMS differences of 101.1 μas, 93.7 μas and 5.6 μs. Figure 1 shows the weighted differences for the period from Jan. 1, 1999 to Dec. 31, 2001 for the y pole component. The rate of 26.5 $\mu as/y$ which can be seen is due to the use of different TRF.

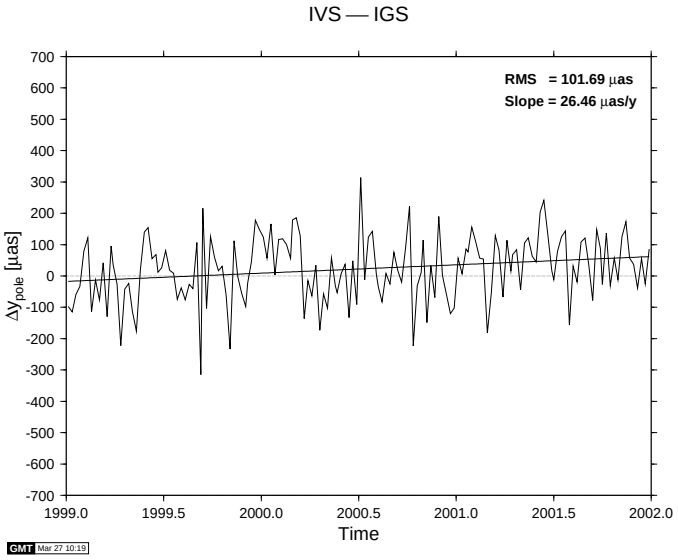

Figure 1. Differences between IVS01001 and IGS (y_p)

2. Outlier Detection

In order to increase the reliability of the combined EOP series, the input data is routinely being checked for outliers prior to the final combination. In a preliminary combination, residuals are computed for all input series relative to EOP from a provisional combination. As a test statistic the ratio of the combination residual v_k and its postfit standard deviation σ_{v_k} is then calculated for each component and each input series independently (eq. 1):

$$\tau_k = \frac{v_k}{\sigma_{v_k}} \tag{1}$$

$$|\tau_k| > \tau_{\alpha;n-u;n} \tag{2}$$

with $\tau_{\alpha;n-u;n}$ = percentage point of τ–distribution; n = number of observations; $n-u$ = degree of freedom; $1-\alpha$ = significance level. An outlier at the siginificance level of $1-\alpha$, for example 95%, is detected if the absolute value of the test statistic is greater than the respective percentage point of the central τ–distribution [4]. Since all components are always combined rigorously, i.e. with the corresponding covariances, in this example n is 12 (3 components × 4 ACs) and n - u is 9. The outliers will be reported to the respective Analysis Centers with a request for further tests or for recomputation of the data point.

As a numerical example for the outlier detection, table 3 lists the $dUT1$-residuals for Jan. 5 2000 with, at that time, only four IVS Analysis Centers. The percentage point of the τ distribution is 2.44. Looking at the values for all input series we find that the IAA data point has a value above the τ–distribution limit. Hence this data point is marked as suspicious and is not included in the final combined data point.

Table 3. Numerical example for outlier detection; epoch Jan. 5, 2000; parameter $dUT1$

AC	v_k [μs]	σ_{v_k} [μs]	$\lvert\tau_k\rvert$ [–]	outlier ?
BKG	-5.1	3.3	1.55	–
GSF	1.1	3.7	0.30	–
IAA	14.7	5.7	2.57	yes
SPU	-3.2	5.7	0.56	–

3. New Combined Series IVS02001

The use of ITRF2000 station coordinates as the basis for the IVS combined series is the most recent step towards the generation of a consistent chain from the quasi-inertial frame of radio sources to a commonly accepted conventional terrestrial reference system. While the IVS01001 series was referred to the IERS C04 series through constant biases as described in 1 the new combined series IVS02001 is consistently linked to the ITRF2000. This has been made possible through the fact that almost all IVS Analysis Centers produce EOP series which either use ITRF2000 station coordinates as fixed input parameters or constrain their solutions to ITRF2000 on a no-net-translation and no-net-rotation basis.

The way the new combined series is computed differs only slightly from the way in which the series IVS01001 has been generated. The weight factors are computed in the same fashion as it is described in [2]. Although one would expect that all ITRF2000 based input series have a zero mean relative to the combined series, this is not the case at the level of several tens of microarcseconds. In order to eliminate the effects of these small scale systematics, we first generated a preliminary reference series for the period between January 1, 1999, and December 31, 2000. After extensive

tests an arithmetic mean of four input series from Geoscience Australia (OCCAM), Bundesamt für Kartographie und Geodäsie (CALC/SOLVE), Goddard Space Flight Center (CALC/SOLVE) and Institute for Applied Astronomy (OCCAM) has been computed. These series were chosen to give a good balance between the two different analysis software packages which are used at the IVS Analysis Centers. This preliminary reference series meets the consistency requirements as best as is possible under the current circumstances.

For all input series we then computed bias and rate terms with respect to the reference series in order to minimize possible AC-specific impacts on the combination (table 4). The bias terms are not really informative here since they depend on the reference epoch. Average magnitudes are -35.0 μas to 30.0 μas at the center of the determination period.

Table 4. Rates and weight factors used in combination IVS02001 (determination period: 1.1.1999 - 31.12.2000

	x_p [μas/y]	y_p [μas/y]	$dUT1$ [μs/y]	w.f. [-]
AUS	-20.6	-11.1	0.3	1.22105
BKG	-24.2	10.2	0.0	0.94785
GSF	15.9	22.6	-0.2	1.13326
IAA	38.5	18.1	-1.0	1.08235
SPU	-45.6	-9.6	1.3	0.68767
USN	17.5	28.6	-0.2	1.15541

As compared to the IVS01001 series some of the weight factors had to be adjusted considerably. The validity of the reduction in the weight of the SPU series can easily be justified with the comparably large scatter of the nutation residuals (figure 2). It is quite understandable that a large scatter in the nutation residuals also has an impact on the other EOP components.

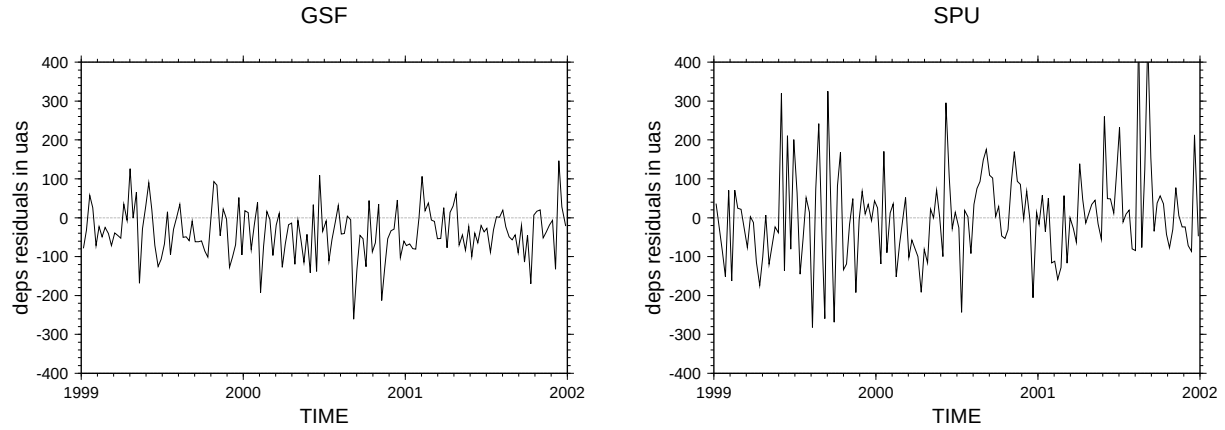

Figure 2. Comparison of $d\epsilon$-residuals relative to combination

Due to the fact that by now six IVS Analysis Centers regularly submit EOPs to the IVS Data Centers and due to the better agreement of the series (see below) outliers can be eliminated more reliably.. As a numerical example for the outlier detection in the new combined series IVS02001 table 5 summarizes the y_p-residuals for July 2nd, 1999. In this example, the corresponding percentage point is $\tau = 2.70$. There is one data point in the series of the IAA which seems to be an outlier relative to the combination.

The final combination of polar motion and UT1-UTC is calculated using the full variance and

covariance information according to

$$x_{j,combi} = \frac{\sum_{i=1}^{n} f_i p_{x_{ij}} (x_{ij} - (rate_{i,x} \cdot MJD_j - b_{i,x}))}{\sum_{i=1}^{n} f_i p_{x_{ij}}} \quad (3)$$

with x_{ij} = observation; $p_{x_{ij}}$ = input weight; f_i = weight factor; $rate_{i,x}$ = rate relative to reference series (cf. table 4); $b_{i,x}$ = y-axis intercept of straight line fit; j = epoch; n = number of Analysis Centers. For the combination of the nutation offsets no rates and biases are applied. The covariances are omitted in eq. 3 for clarity reasons.

Table 5. Numerical example for outlier detection in IVS02001; July 2, 1999; parameter y_p

AC	v_k $[\mu as]$	σ_{v_k} $[\mu as]$	$\lvert\tau_k\rvert$ $[-]$	outlier ?
AUS	-48.9	130.7	0.37	–
BKG	-175.2	210.7	0.83	–
GSF	-95.0	109.8	0.87	–
IAA	340.8	113.1	3.01	yes
SPU	-183.8	280.3	0.66	–
USN	-98.6	111.3	0.89	–

The averaged formal errors of the input series, the resulting WRMS relative to the combination and the mean internal precision of the combined series are listed in table 6. In some cases Analysis Centers used the change over from the IVS01001 to the IVS02001 solution for their transition to a different TRF realization. Station coordinates which were determined in pure VLBI solutions were replaced by fixed ITRF2000 coordinates. Other Analysis Centers map the VLBI coordinates onto ITRF2000 by using no-net-translation and no-net-rotation constraints. For this reason, the weighted RMS scatter of the post combination residuals also changed as a consequence of the new combination. In addition, the fixed ITRF2000 coordinates sometimes lead to slightly increased mean formal errors. The results of the new combination are, as usual, published on the IVS Analysis Coordinator's webpage both in graphical and in numerical representation.

Table 6. Mean formal errors (MFE) of input and WRMS relative to combination IVS02001 (from 1.1.1999 to 31.12.2001)

	$x_p[\mu as]$		$y_p[\mu as]$		$dUT1[\mu s]$		$d\epsilon[\mu as]$		$d\psi \sin\epsilon_0[\mu as]$	
	MFE	WRMS	MFE	WRMS	MFE	WRMS	MFE	WRMS	MFE	WRMS
AUS	114.1	95.0	94.0	98.2	5.1	4.4	63.9	52.6	65.1	56.3
BKG	140.4	128.6	119.3	113.2	5.7	7.1	74.4	72.7	75.8	74.6
GSF	109.4	57.9	88.7	49.6	4.6	2.6	80.2	56.5	82.4	57.8
IAA	91.6	98.9	77.3	90.6	3.9	4.3	68.9	64.5	70.3	61.4
SPU	129.8	112.5	99.3	85.6	5.4	4.5	76.5	101.0	79.4	115.6
USN	109.2	61.3	88.4	51.6	4.6	2.9	80.4	56.0	82.3	57.6
mean formal precision of combination										
IVS	43.1		38.2		2.1		35.7		36.6	

Compared to the results listed in table 2 the IVS02001 series shows a better internal precision and a better consistency, but the formal precision seems to be even more optimistic than of the old

combined series. The reason for the better consistency is the more homogenous use of identical or very closely matching station coordinates. On March 1st, 20002, the new combined series IVS02001 has replaced the old series IVS01001.

The comparison of IVS02001 with the IGS series shows a much better agreement in terms of the WRMS differences. However, the bias in the y component of more than 300 μas which is equivalent to almost 1 cm at one Earth radius is rather puzzling. Although the IGS only changed from ITRF97 to ITRF2000 station coordinates on December 2, 2001, the EOPs on the basis of ITRF97 should be fairly consistent with those after Decenber 2001 since both ITRFs are fairly consistent and since there is no obvious impact on the differences (see figure 3). The use of the ITRF97 coordinates by IGS in the first period of the comparison can, thus, not be the reason for the significant bias. Without a more detailed investigation there is no obvious explanation yet.

Table 7. Weighted mean differences and wrms IVS02001 - IGS (from 1.1.1999 to 31.12.2001)

	$x_p[\mu as]$	$y_p[\mu as]$	$dUT1[\mu s]$
WMEAN	-77.0	319.5	6.5
WRMS	95.4	85.4	5.9

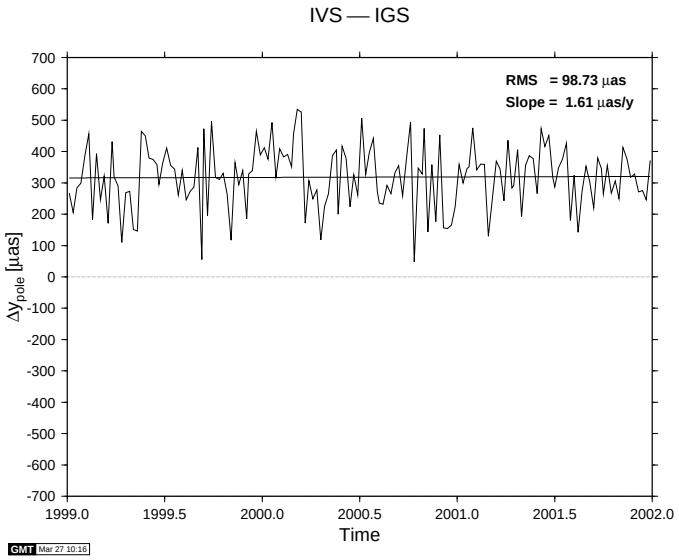

Figure 3. Differences between IVS02001 and IGS (y_p)

References

[1] Cannon, W. H.,Comments on VLBI as a Fundamental Geodetic Positioning Technique and IVS Combined VLBI Products, Memo electronically available under http://giub.geod.uni-bonn.de/vlbi/IVS-AC/divers/cannon.html

[2] Nothnagel, A., C. Steinforth, Analysis Coordinator Report 2000, In: International VLBI Service for Geodesy and Astrometry 2000 Annual Report, NASA/TP–2001-209979, N. R. Vandenberg and K. D. Baver (eds.), 55–64, 2001.

[3] Nothnagel, A., C. Steinforth, Analysis Coordinator Report 2001, In: International VLBI Service for Geodesy and Astrometry 2001 Annual Report, N. R. Vandenberg and K. D. Baver (eds.), to be published

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

[5] http://hpiers.obspm.fr/eop-pc