Automatic Processing of INT Sessions with nuSolve

Sergei Bolotin, Karen Baver, Mario Bérubé, John Gipson

Abstract IVS INT sessions are conducted to determine the change in Earth rotation, which is measured as a correction to Universal Time, or dUT1. This correction varies unpredictably with time. Because dUT1 is used in precise navigation, particularly GNSS, rapid turnaround is very important. The elapsed time from observation to obtained results can be shortened using automatic data processing at the analysis stage. To develop an application for this analysis, we used the script mode of the VLBI data processing software nuSolve. In this script mode, nuSolve reads commands from a script file and executes them. The script engine is implemented by the Qt library and supports the EC-MAScript (standardized Java script) programming language.

A user of the script mode has access to the same functionality of nuSolve as a user of the GUI mode. In this paper we discuss the application of the script for the automatic processing of INT sessions. We conducted a comparison of results from the script execution with manual data analysis for all INT S/X and VGOS sessions performed during the last five years. For 50% of the sessions the automatic script gives a solution that is identical to the one obtained by an operator. For 75% of the sessions, the difference of the dUT1 estimation between the two types of solutions is less than 5 microsec. We discuss the anomalies of INT sessions that cause significant differences between automatic and operator solutions.

Keywords VLBI, data analysis, automatization, software

NVI, Inc., NASA GSFC Code 61A, 8800 Greenbelt Road, Greenbelt, Maryland 20771, USA

1 Introduction

One of the practical goals of the geodetic Very Long Baseline Interferometry (VLBI) technique is the determination of the change in Earth rotation. These variations are caused by interactions of Earth, ocean, and atmosphere and cannot be predicted. Using VLBI observations it is possible to obtain this unmodeled change of Earth rotation, which is measured as a correction to Universal Time, or dUT1. A special type of VLBI observations, an INT session, is organized to estimate dUT1 only. The International VLBI Service for Geodesy and Astrometry (IVS) [Nothnagel et al., 2017] INT sessions are planned to last one hour (though a VLBA part of the IVS-IN1 network was scheduled to run two hours). Because dUT1 is used in precise navigation, particularly GNSS, rapid turnaround is very important.

At NASA GSFC, we developed software that automates operations of downloading files with VLBI observations, performes necessary calculations and calibrations, and prepares a new VLBI session ready for processing by an analyst. When the analyst has processed the session, the software obtains a solution and submits the results to the IVS Data Centers for further use.

The procedure of session analysis is performed manually by an analyst with the nuSolve software [Bolotin et al., 2014]. During the initial processing of a session the following steps should be made: resolving of group delay ambiguities; elimination of outliers; checking for presence of clock break effects; and evaluation of ionosphere corrections among others. This is an important part of the geodetic VLBI data flow.

To reduce the latency from observations to obtained results we developed an application for automatic analysis using the script mode of the VLBI data processing software nuSolve. In this script mode, nuSolve reads commands from a script file and executes them. A script is a plain ASCII file that nuSolve reads and executes line by line. The script engine is implemented by Qt library and supports the ECMAScript (standardized Java script) programming language. nuSolve exports its types, objects, and functions to make them accessible to the script engine. That makes it possible for a user to have access to the same functionality of nuSolve as available in GUI mode. In a script a user can read a session, change models and parameterization, obtain a solution, create a report on an obtained solution, store a new version of the session, and so on.

The script mode appeared in nuSolve in 2018, in version 0.6.0 of the software distribution. Since then it has been used by NASA GSFC and other VLBI Analysis Centers for data processing, investigations, and testing the software.

2 Automatization of Data Analysis

Initially, when the script mode appeared in nuSolve, the script `pia4INT.js` was created and included in the distribution. The purpose of this script was to demonstrate how routine operations performed by an analyst can be conducted in script mode. Script `autoINT.js` is derived from script `pia4INT.js` and basically performs the same operations. Additional functionality and the ability to adjust the configuration is added to this script. The new script also checks for some known problems of INT sessions, e.g., unusable data at very short baselines if a notch filter was not applied during correlation.

The script `autoINT.js` performs the following procedures:

- Read a database in vgosDb format [Gipson, 2012]. If the version is greater than 3, reset all editing.
- Check for known anomalies and set up a clock reference station.
- Set parameters for estimation: clock offsets and rates only. Get a single band delay solution for S-band.
- Check group delay ambiguities in S-band. Get a solution and check outliers.

- Check group delay ambiguities in X-band. Get a solution and check outliers.
- Evaluate ionosphere corrections.
- Set parameters for estimation: clocks, zenith delays, and baseline vector(s) or dUT1.
- Perform a reweighting/outlier processing loop.
- Obtain a final solution.
- Save a report in spool file format.
- If it was specified by the user, save a new version of the database.

The script processes one database at a time. A spool file and additional output are stored in predefined directories.

To compare results of the script for automatic analysis with a solution obtained by an analyst, a simplified script was derived from the `autoINT.js` script. This script reads a database, sets up parameters in the same way as it is done by the `autoINT.js` script, and obtains a final solution with the editing stored in the database. Then, a spool file and similar additional output are stored. A solution obtained with this script serves as a reference solution.

To compare the two versions of editings, the following values of a final solution were collated:

- Number of potentially usable and processed observations.
- WRMS of the solution.
- Estimation of dUT1: a posteriori value, an adjustment and its standard deviation.

For a given INT session, if these six values of automatic and operator solutions are the same, then both solutions are equal.

3 Data

To test the `autoINT.js` script we analyzed 2,494 INT sessions observed from January 2017 until July 2022 (S/X, VGOS, and mixed mode). That includes the IVS-IN1, IVS-IN2, IVS-IN3, VGOS-INT, VG, and SI series. 19 stations from 15 sites participated in the observations: GGAO12M, HART15M, HOBART12, ISHIOKA, KASHIM34, KOKEE, KOKEE12M, MACGO12M, MK-VLBA, NYALE13S, NYALES20, ONSA13SW, PIETOWN, SESHAN25, SVETLOE, WESTFORD, WETTZ13N, WETTZ13S, and WETTZELL. Two of the stations have tested new

equipment with bogus station names: NYALDBBC, WETTDBBC, and WETTVDIF. The observations were correlated at the seven correlators: BONN (Max Planck Institute for Radio Astronomy, Bonn, Germany), GSI (Geospatial Information Authority of Japan, Tsukuba, Japan), HAYS (MIT Haystack Observatory, Westford, Massachusetts), OSO (Onsala Space Observatory, Onsala, Sweden), UTAS (University of Tasmania, Hobart, Australia), WASH (U.S. Naval Observatory, Washington, DC), and WETZ (Geodetic Observatory Wettzell, Wettzell, Germany). According to IVS procedures, different series of INT sessions are processed by different IVS Analysis Centers and then submitted to the IVS Data Centers.

Most of the INT sessions are single baseline sessions, although some of the sessions have more than one baseline. A single baseline S/X INT session collects 20 to 40 observations. Due to unpredicted problems during the observations some data can get lost and the length of a session could decrease. For the processed set of INT sessions, the shortest one has 20 minutes of observations.

The single baseline VGOS INT sessions usually have 40 to 80 observations in an interval of 60 minutes. In two VGOS sessions, VGOS-T2049 (22FEB18VG) and VGOS-T2077 (22MAR18VG), all available VGOS stations participated in observations collecting more than a thousand observations. The INT sessions (like any other geodetic VLBI sessions) could have different problems (clock breaks, non-detection, etc.) that makes it impossible to process them at all. For example, session IN121-102 (database 21APR12XU) has a total of 20 observations, and only four of them have a quality code greater than zero (i.e., the rest of data are non-detections).

4 Analysis

To compare the automatic and human solutions we used results of session processing that were performed at the Goddard VLBI Analysis Center. Auxiliary software tools were developed to extract data from the two sets of solutions and to compare them. The comparison showed that for 1,259 out of the 2,494 sessions the automatic script gives a solution that is identical to the one obtained by an operator.
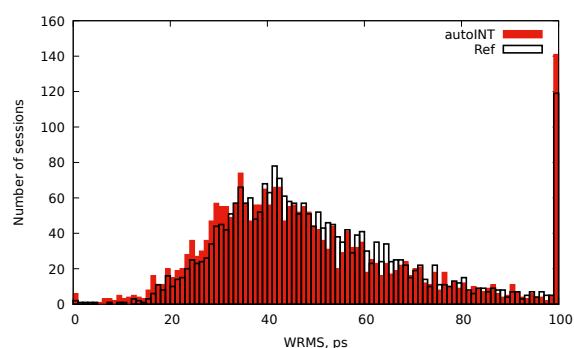

Fig. 1 Histogram of WRMS for the automatic (filled, red bars) and the reference (empty bars) solutions.

Figure 1 shows distributions of the number of sessions as a function of WRMS for the two sets of the solutions. The bars with zero WRMS display the number of sessions that have an insufficient number of usable observations for the least-squares estimation (fewer than six observations for a one-baseline session). The bars with WRMS 100 ps and more correspond to sessions that have severe problems (like clock breaks, subambiguities, problems with phase calibration, etc.). The bars of the reference solution with WRMS 0 ps and 100 ps shows the number of problematic INT sessions in the five-year period (not including sessions for which databases were not created or were not preprocessed by an operator). The corresponding bars of the automatic solutions are bigger; the difference of the numbers of sessions with the reference solution shows how many sessions the `autoINT.js` script was not able to process. The automatic script was not able to obtain a solution for a total of 26 INT sessions; these sessions have to be processed manually.
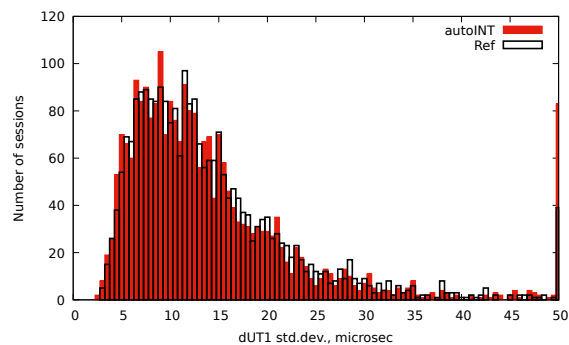

Fig. 2 Histogram of dUT1 standard deviations for the automatic (filled, red bars) and the reference (empty bars) solutions.

Distributions of the number of sessions as a function of the standard deviations of dUT1 estimation for both sets of solutions are shown in Figure 2. As one can see, most of the INT sessions have standard deviations of UT1 that are in a range of 5 to 15 microseconds. The rightmost bar represents sessions with standard deviations for UT1 of 50 microseconds or more. The difference of numbers of sessions for automatic and reference solutions for these bars is 44 sessions.

Figure 3 shows distributions of the number of sessions as a function of the difference of dUT1 estimation for the two sets of the solutions. The central bar that is representing 1,189 sessions is removed from the figure to make other bars visible. For 1,905 sessions, or 76%, the difference of the dUT1 estimation between the two types of solution is less than 5 microseconds. The range of dUT1 differences was chosen to be -50 to 50 microseconds. 89 sessions do not fall in this range, which indicates that they, perhaps, need to be processed manually too. However, checking these sessions showed that the relatively large difference in the estimated dUT1 parameter is caused by deselecting (or including) one or a few observations in the automatic solution that was/were not present (or deselected) in the reference solution. When a number of observations used in a solution is small (e.g., ten observations and six estimated parameters), adding or removing one observation can change results drastically.
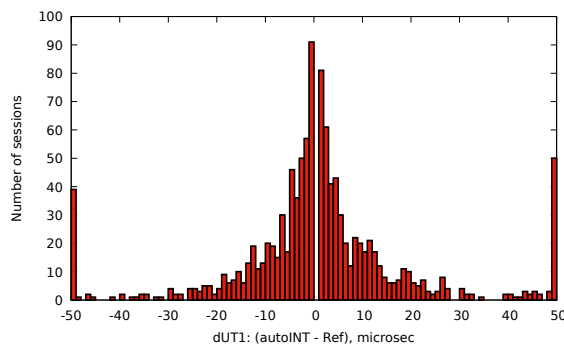
sessions, or 87%, have scaled differences of dUT1 estimation that are in the range of -1 to 1.
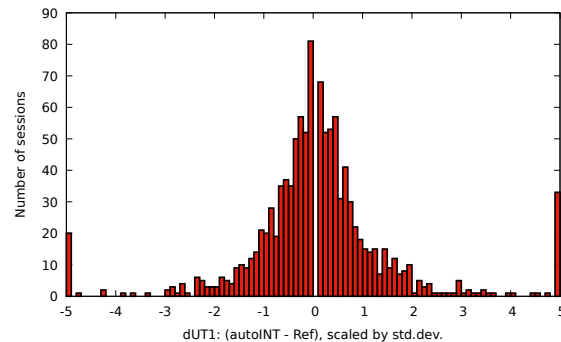

Fig. 4 Histogram of scaled by standard deviation difference of dUT1 estimation between automatic and reference solutions.

Analysis of the INT sessions that have different automatic and reference solutions showed that for most of these sessions the difference between the two solutions is caused by a very small number of analyzed observations. Having 10–20 good observations is not enough to determine outliers.

For VGOS INT sessions (which usually have more observations than the S/X sessions), a comparison of the differences between the automatic and reference solutions showed that for 138 (or 46%) of a total of 298 sessions the solutions are equal.
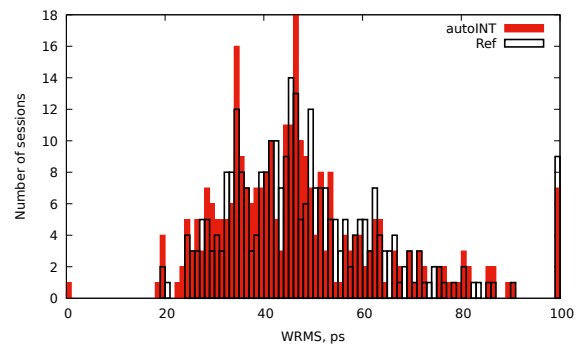

Fig. 3 Histogram of the difference of dUT1 estimation between automatic and reference solutions.

A histogram of scaled differences of dUT1 estimations (i.e., the differences divided by the standard deviations of dUT1 estimation) between the two sets of the solutions is represented in Figure 4. A total of 2,180

Fig. 5 Histograms of WRMS for the automatic (filled, red bars) and the reference (empty bars) solutions for VGOS INT sessions only.

Figure 5 shows histograms of WRMS for VGOS sessions for the automatic and the reference solutions. The script was unable to process only one of the VGOS sessions, VGOS-21349 (21DEC15VI), with nine us-

able observations. For the rest of the VGOS sessions the results of the automatic analysis and an operator solution are close.

Histograms of distributions of standard deviations of dUT1 estimations are shown in Figure 6. For VGOS INT sessions most of the sessions have standard deviations of UT1 that are in the range of 3 to 10 microseconds. For 250 (84%) of the VGOS INT sessions the difference of the dUT1 estimation between two types of solution is less than 5 microseconds. And 264 sessions, or 88%, have scaled differences of dUT1 estimation that are in the range of -1 to 1.
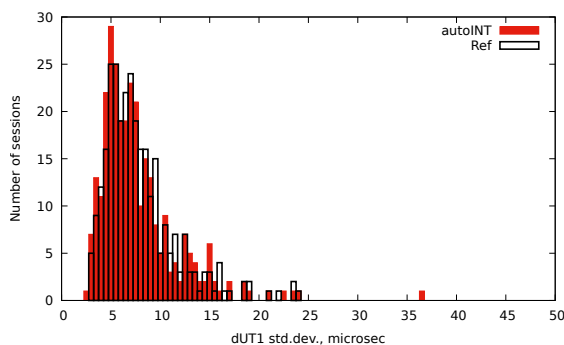

Fig. 6 Histograms of dUT1 standard deviations for the automatic (filled, red bars) and the reference (empty bars) solutions for VGOS INT sessions only.

It should be noted that work on the script `autoINT.js` is not finished. The functionality of the script can be improved; for example, there are many sessions produced by the GSI correlator where the single band delays are non-usable (e.g., IN322-059 (22FEB28XK)). Because the script starts its analysis from processing the single band delays, these sessions will fail in automatic analysis. Using the group delays and the delay rates instead of the single band delays could solve this problem.

5 Conclusions

The automatic script is capable of processing INT sessions. For 2,494 INT sessions collected over five-and-a-half years, the script has failed to process 44 sessions. The scaled differences of dUT1 estimations between the automatic and the reference solutions for 87% of INT sessions are in the range of -1 to 1.

Special treatment of outliers is required when a session with a small number of observations is processed.

The script is part of the nuSolve distribution. The work on the script is continuing. We plan to update the script and continue to include the modified versions in the future nuSolve releases.

References

[Bolotin et al., 2014] S. Bolotin, K. Baver, J.M. Gipson, D. Gordon and D. MacMillan. *The VLBI data analysis software νSolve: development progress and plans for the future*. In: D. Behrend, K.D. Baver, and K. Armstrong, editors, *IVS 2014 General Meeting Proceedings*, ISBN 978-7-03-042974-2, Science Press, Beijing, China, pages 253–257, 2014.

[Gipson, 2012] J.M. Gipson *IVS Working Group 4: VLBI Data Structures* In D. Behrend, and K.D. Baver, editors, *IVS 2012 General Meeting Proceedings*, NASA/CP-2012-217504, NASA GSFC, Maryland, pages 212–221, 2012.

[Nothnagel et al., 2017] A. Nothnagel, T. Artz, D. Behrend, Z. Malkin, Journal of Geodesy 91(7), 711 (2017). DOI 10.1007/s00190-016-0950-5.