

The First Release of ν Solve

Sergei Bolotin, Karen D. Baver, John M. Gipson, David Gordon, Daniel S. MacMillan

NVI, Inc./NASA Goddard Space Flight Center

Contact author: Sergei Bolotin, e-mail: `sergei.bolotin@nasa.gov`

Abstract

In this paper we describe the first release of ν Solve. ν Solve is software that will replace the current interactive part of the SOLVE system. We review the capabilities of the software, its current status, and our plans for future development.

1. Introduction

The deployment of the VLBI2010 network of stations will result in a significant increase in observations. The current VLBI data analysis software was not designed for this amount of data. VLBI2010 will also introduce new types of observables.

To face this challenge, development of the next generation VLBI data analysis software was initiated at the NASA Goddard Space Flight Center. First, the IVS Working Group on VLBI data structures (IVS WG4) was established in 2007. The results of the group's work made it possible to begin designing the new VLBI data analysis software. In 2010 the general design of the software and an overview of its architecture were presented at the IVS General Meeting at Hobart (Tasmania). Subsequently, we began writing the software and in 2011 at the EVGA Meeting in Bonn we demonstrated a prototype of the new software system.

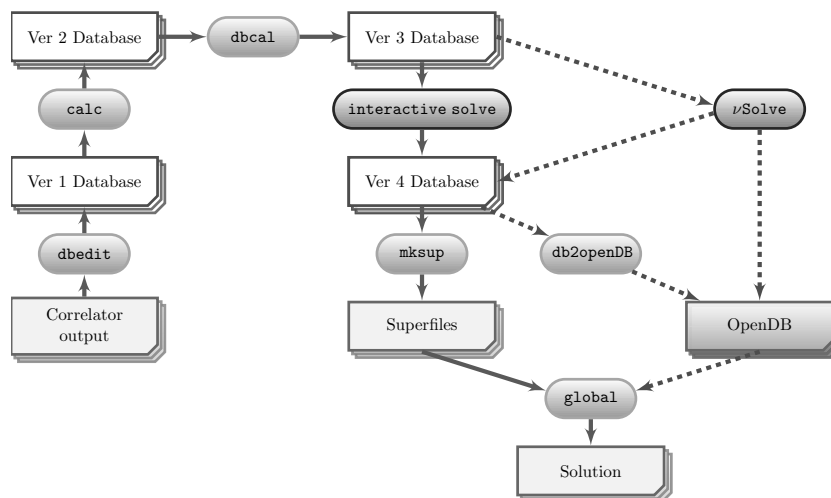


Figure 1. Geodetic VLBI data flow.

In Figure 1 a simplified data flow diagram of geodetic VLBI observations is shown. Currently (solid arrows on the diagram), the correlator output passes through a program called `dbedit` and

is converted into a *database version 1* file. Then, the program `calc` adds theoretical values into the database and the program `dbcal` adds cable calibration data and meteorological parameters (if available). The resulting *database version 3* files (for X- and S-bands) are processed with an interactive tool, called `interactive solve`, and the results are stored in a *database version 4* file. Many things are done during the last transformation: ionospheric calibrations are computed, ambiguities are resolved, and outliers are marked for exclusion from solutions. In general, *database version 4* files are assumed to be free of all anomalies and can be safely processed in automatic modes. The program `mksup` creates a binary *superfile*, which contains a subset of the data in the database necessary for analysis. Finally, the non-interactive part of the SOLVE system, `global`, generates final solutions using superfiles. Since the current CALC/SOLVE system is involved in data preparation process, it will be replaced with the new software piece by piece.

In the whole chain of data flow, the most time-consuming part is `interactive solve`. This part imposes some restrictions on input data (e.g., the number of observed stations has to be no greater than 32) and is a source of analyst errors. In the future (dashed lines on the diagram), this part will be replaced first by ν Solve. In parallel, improvements to `global` are under way as well as the replacement of the Mark III databases by OpenDB files.

The openDB format described elsewhere in this volume in [5] has many advantages over the custom Mark III database format. The openDB format stores most VLBI data in netCDF files where each netCDF file contains a few items of related VLBI data, for example meteorological parameters at a site, or editing criteria. The netCDF files are organized by a special ASCII “wrapper” file which provides the location of files requested for a solution. This gives great flexibility. For instance, one can easily see the effect of alternative editing by modifying the wrapper. In addition, netCDF files are platform and operating system (OS) independent and there are netCDF interfaces to many programming languages. This allows the data to be accessed by anyone in the VLBI community. Furthermore the netCDF user community has developed many freely available tools for accessing and manipulating data in netCDF files.

In this article we will discuss features and abilities of the ν Solve software, as a replacement for the interactive part of the SOLVE system.

2. New VLBI Data Analysis Software

2.1. Software Development Environment

The new VLBI data analysis software is being developed under the GNU/Linux operating system. However, there is no explicit demand for GNU/Linux OS and ν Solve is designed to operate on any POSIX compatible OS. We chose C++ as the programming language due to its power, flexibility, and portability. The well known and widely used Qt library provides sets of widgets for the graphical user interface and containers of data and auxiliary tools.

Currently, the new VLBI data analysis software consists of two parts: a *Space Geodesy Library*, a library where data structures and algorithms are implemented (about 90% of total source code); and an executable ν Solve – a driver that calls library functions and organizes work with an end-user (about 10% of total source code). However, in this article we will refer to all the source codes as ν Solve software.

2.2. Modular Structure

ν Solve uses a modular structure. This allows us to keep our system flexible and stable. A module is a logical part of source code that is loosely tied to other parts of the software. Several modules are represented by external libraries (e.g., GUI module of Qt library). The modular structure of the software was discussed in detail in [4].

2.3. Functionality

There are two modes to work with VLBI observations: ν Solve can operate with data through the CALC/SOLVE catalog subsystem or work in a standalone mode.

The capabilities of ν Solve are evolving in accordance with the adopted model of the software production process (see [3]). All necessary key functional points are implemented in the software. It is able to read VLBI observations from Mark III database files, process a single multiband VLBI session and save results, display various data on the plots, estimate various parameters, detect and process clock breaks, resolve group ambiguities, perform ionospheric correction, calibrate weights of observations, and eliminate outliers. We will next discuss the most important functional points in detail.

Plotting subsystem. A simple plotting subsystem has been developed for ν Solve. It allows a user to display various values that are either stored in the files, evaluated “on the fly”, or estimated from VLBI observations analysis. Also, it permits a user to edit observations by eliminating outliers, resolving ambiguities, and adjusting clock break parameters. In Figure 2, the residuals for the session \$10DEC06XA are shown by the plotter. A 1-second clock jump is clearly visible for baselines to the NYALES20 station.

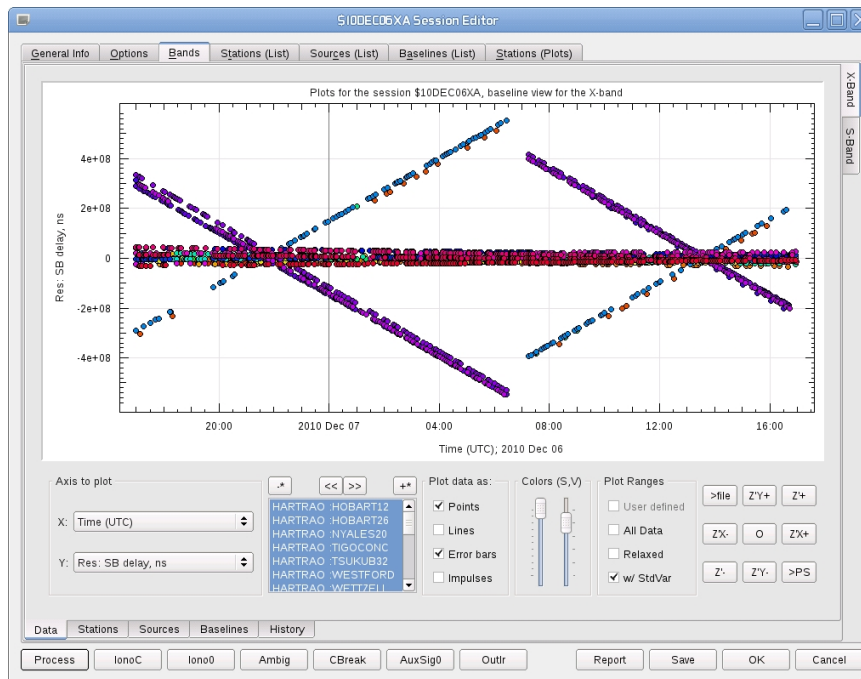


Figure 2. Plotting subsystem: a screenshot of a plot with residuals.

Estimator. In ν Solve we employ a square root information filter (SRIF, see [1] for details) for parameter estimation. SRIF is a powerful and flexible algorithm that allows us to implement types of parameters that are realized in the current CALC/SOLVE system (local parameters, parameters modeled as piecewise linear functions) as well as to introduce stochastic parameters. Also, it is less restrictive for segmented parameters, e.g., piecewise parameters can have different lengths of segment intervals or have overlapping segments.

Clock breaks. To compensate for a clock break, ν Solve adds a step-wise linear function to the clocks of a station. There are session-wide and band-dependent clock break models. A simple algorithm was developed to detect clock breaks in automatic mode and evaluate parameters of a clock break model. Alternatively, the user can add or adjust the parameters of clock break model manually.

Ambiguity resolution. We applied the same ideas for ambiguity resolution as used in SOLVE. However, the realization is different. ν Solve does not make any assumptions about ambiguity spacing and is able to resolve ambiguities automatically for a session with mixed (due to lost channels at a band) ambiguity spacing. Manual editing of ambiguities is also available.

Reweighting. To normalize the χ^2 distribution of residuals, we use the same approach implemented in SOLVE. Additional standard deviations are computed either for a whole session or for each baseline.

Outliers. ν Solve treats outliers as observations where the absolute value of normalized residuals is greater than the user specified threshold. The normalization is made either for a whole session network or for each baseline. There are two modes in outlier processing: 1) elimination where outliers are tagged for exclusion from data analysis; and 2) restoration, where previously excluded observations are marked for inclusion back into analysis. Outlier processing is performed in conjunction with the reweighting process.

2.4. Processing a Session

In this section we describe a typical sequence of operations that are necessary to process a VLBI session with the ν Solve software. First, the user reads the observations. Depending on the selected mode of working, ν Solve either reads files specified by the user from the hard drive or consults with CALC/SOLVE to figure out where the databases are expected to be and reads them. Then the user selects clock parameters to be estimated and obtains a single band delay solution. This solution is used in the first stage of data processing to check clock breaks and resolve group ambiguities.

The user inspects residuals obtained from a single band solution, eliminates outliers, and checks for clock breaks. Real clock breaks have the same value of jump in clock behavior for both bands. In contrast, pseudo clock breaks have different magnitudes in each band. These are discussed further below.

After clock breaks are fixed, the user switches from single band delay to group delay and resolves group ambiguities for each band. Resolving group ambiguities is an iterative process; after each iteration the user has to obtain a new solution. Usually, two or three iterations are sufficient. However, sometimes it is necessary to turn on estimation of wet zenith delays and station positions to correctly resolve ambiguities.

At the same stage of processing, small pseudo clock breaks should be checked. There are some effects that manifest themselves as clock breaks (e.g., manual phase calibration which was

applied at a correlator in several segments). To correct these anomalies, band-dependent clock break models are introduced by ν Solve. After resolving ambiguities, fixing small clock breaks and obtaining the ionospheric corrections, the user can evaluate ionosphere corrections for each band.

With resolved ambiguities, fixed clock breaks and obtained ionosphere corrections the user can turn on estimation of wet zenith delays, station positions, angles of nutation, rate of Earth rotation, etc. Obtaining a solution with all these conventional parameters allows the user to detect and mark for exclusion large (for example, $10 - 20\sigma$ and greater) outliers.

The next step is estimating clocks, wet zenith delays, and tropospheric gradients as piecewise linear functions. Again, large outliers should be excluded.

The last two operations, reweighting and outlier processing, are done in combination. First, the user “reweights” the solution – that is adjusts the sigmas of the observations until $\chi^2 = 1$. Then, outliers are eliminated. If some observations are removed from data analysis, the reweighting should be repeated and outlier processing should be performed too.

At the end of session processing, the user saves results as a new version of the database. If the session was accessed through the CALC/SOLVE catalog subsystem, the catalog is updated to reflect that a new version of this session is available. Also, the user can generate an NGS file with the observations to export data for external software.

3. Public Release

At this time, ν Solve has all the necessary capabilities to analyze a VLBI session. Now we are making tests comparing the results and intermediate values produced by the interactive part of current SOLVE and the new software.

ν Solve has the following advantages over interactive SOLVE: 1) better user interface; 2) ability to plot more kinds of data; 3) ability to automatically resolve baseline dependent ambiguities, and 4) automatic clock breaks processing. In general, it significantly decreases an analysts’ time to process a VLBI session.

A first public release of ν Solve is expected with the forthcoming release of the CALC/SOLVE system. After public release we expect user feedback to improve the software.

References

- [1] Biermann, G.J., Factorization Methods for Discrete Sequential Estimation, V128, Mathematics in Science and Engineering Series, Academic Press, 1977.
- [2] G. Booch, Object-Oriented Analysis and Design with Applications. Addison-Wesley, 2nd ed., ISBN 0-8053-5340-2, 1994.
- [3] S. Bolotin, J.M. Gipson and D. MacMillan, Development of a New VLBI Data Analysis Software. In: IVS 2010 General Meeting Proceedings, NASA CP 2010-215864, NASA GSFC, Maryland, D. Behrend, and K.D. Baver (eds.), 197–201. 2010.
- [4] S. Bolotin, J.M. Gipson, D. Gordon and D. MacMillan, Current Status of Development of New VLBI Data Analysis Software In: Proc. of the 20th Meeting of the EVGA, Bonn, Germany, W. Alef, S. Bernhart and A. Nothnagel (eds.), 86–88. 2011.
- [5] J.M. Gipson, The Report of IVS-WG4, In: IVS 2012 General Meeting Proceedings, this volume.