3. KSP VLBI System

3.3 Data Analysis System

3.3.1 Data Analysis Software

By

Yasuhiro KOYAMA, Kosuke HEKI, Yukio TAKAHASHI, and Masato FURUYA

(Received on)

Running Title: 3.3.1 Data Analysis Software

ABSTRACT

New softwares have been developed to automate all of the necessary procedures for analyzing observation data of the Very Long Baseline Interferometry system in the Key Stone Project. As a result, no human interaction is required for the data analysis and the results can be obtained automatically. The final results are placed in the publicly accesiible space on the Internet so that anyone can use the data immediately after experiments. The automated data analysis system allows frequent experiments and minimizes the time required to produce results. This paper describes the features and concepts of the data analysis system.

(Keywords: VLBI, geodesy, data analysis)

1. Introduction

In the data analysis of the Very Long Baseline Interferometry (VLBI) technique for making precise geodetic measurements, coordinates of all the observation sites except for one reference site are estimated from a set of observed time delays and their rates of change. The time delay and its time derivative are direct products obtained from data correlation processing and bandwidth-synthesis processing. All of the data obtained and relevant information are stored in a database file and data analyses are performed based on the database file. In the data analysis, site coordinates of a reference site are fixed to *a priori* values and the coordinates of other sites are estimated by using a least-squares estimation method. However the initial database contains bad data and delay ambiguities that should be removed first to obtain correct estimations. Before the Key Stone Project (KSP) VLBI data analysis system was developed, these procedures required extensive knowledge of the theoretical aspects of VLBI data analysis methods and took a long time to obtain final results. Frequent VLBI experiments were not realistic unless the data analysis procedures were automated. We have developed softwares for the KSP VLBI data analysis system to make all the necessary data analysis procedures fully automated. To make the softwares work reliably, various conditions had to be taken into account and a deep insight of the VLBI data analysis was required. This paper describes the details of the data analysis system.

2. Data analysis system

2.1 Hardware system and network

The KSP VLBI data analysis softwares were developed on a dedicated workstation (HP9000 model 715/100 running on HP-UX 9.05 operating system). The system is located at Koganei, where the correlator and data processing facilities are located. The workstation has two network interfaces and is configured as a firewall system between the network for the KSP VLBI system and the network for the laboratory. The network for the laboratory is connected to the Internet, so the workstation can be accessed from any hosts on the Internet and on the KSP VLBI system network; however the KSP VLBI system network is protected from unauthorized access from outside. The data analysis workstation provides Domain Name Service for the hosts on the KSP VLBI network and is configured as a network mail server to relay alarm messages from the KSP VLBI observation system.

2.2 Initial data analysis

In both tape-based and real-time VLBI modes, the output data from the correlator system at Koganei are processed by the bandwidth synthesis program **KOMB** to determine time delays and their time derivatives for all combinations of pair of the participating stations. The set of obtained time delays and their rates of change is then combined with auxiliary data, including meteorological data and delay calibration measurements, and is stored in two database files, one for X-band data and the other for S-band data. Only the X-band database is used for data analysis and the S-band database is mainly used for obtaining correlated amplitudes of observed radio sources. The databases are created using Mark-III database handler software developed at the Goddard Space Flight Center of the National Aeronautics and Space Administration. The file format of the databases is based on the Mark-III database file format, and can be handled by various software programs based on the Mark-III database handler routines. Fig. 1 illustrates the relationships between various data files and programs used to generate database files. In the figure, rectangles represent programs and rounded rectangles represent data files. The program **KOMB** processes cross correlation output files generated from the correlator and creates **KOMB** files on the data processing workstation. All programs in Fig. 1 run on the data processing workstation, except **PUTM3**, which runs on the data analysis workstation. Therefore, the data files used by **PUTM3** are placed in a common region where the files can be accessed from both the data processing workstation and the data analysis workstation. The common region is realized by mounting the disk area on the the data processing workstation from the data analysis workstation using the Network File System protocol. The program **PUTM3** merges observation data and *a priori* information based on the IERS Convention 1996 published by International Earth Rotation Service (IERS).[(1)] The one exception is the planetary ephemeris for which the DE200 model[(2)] is used instead of the DE403 model[(3)] recommended in the IERS Convention 1996.

Once the X-band database file has been created on the data analysis workstation, a set of data analysis programs processes the database file until the initial analysis results are obtained. Fig. 2 illustrates the entire process of data analysis. First, the **DBUPDATE** routine updates the database with the latest *a priori* information. The *a priori* information to be updated includes coordinates (δx and δy) of the Earth's rotation pole with respect to the conventional pole, $\mathrm{UT1} - \mathrm{UTC}$, nutation offsets from the IAU80 standard model, terrestrial coordinates of reference points of VLBI sites, and coefficients for site displacements due to ocean loading effects. The same routine is used in the re-analysis procedures when a part of the *a priori* information is revised. In the next step, the **CALC** routine calculates both theoretical delays and delay rates based on the *a priori* information as well as a Jacobian matrix with various parameters for the least-squares estimates. After the **CALC** routine, the database is ready for the least-squares analysis. **DATSET** then extracts necessary information from the database and stores it in a work file.

After these steps, the **VLBEST** routine runs three or more times and least-squares analyses are performed using different combinations of estimation parameters depending on the stage of the analysis. In the first step, only the clock offsets and their rates of change with respect to a reference site are estimated and the other parameters are fixed. In this step, the observed delays may include delay ambiguities, which are certain amounts of time multiplied by integers. The ambiguities are the results of the bandwidth synthesis processing, and the minimum step of ambiguity is defined by the inverse of the greatest common divisor of frequency spacings of the local frequencies of observation channels, which is 100 nanoseconds for the present frequency assignments. The ambiguities are resolved using the residuals of delays after the least-squares adjustment. In the second step, clock offsets at a time interval of 1 hour and site coordinates of all sites except for a reference site are estimated by the routine **VLBEST**. Observed delays having large residuals are flagged and removed from the following analysis by the routine **MRKOBS**. If the residual is greater than three times the root mean square (RMS) of the residual delays and also greater than 200 picoseconds (ps), it is regarded as a bad datum. This step is iterated until there are no more data to be removed. In the last step, zenith wet tropospheric delays at a time interval of 3 hours are added to the parameters to be estimated, and the final results are obtained. The quality of the results are evaluated by the final RMS of the delays. If the RMS residual is smaller than a certain criterion (e.g. 100 ps), the results are regarded as reliable and are made accessible over the Internet. If the RMS residual exceeds the criterion, an alarm message is sent to pre-registered e-mail addresses. In this case, the database is analyzed manually and the results are finally publicized after reliable results are obtained.

2.3 Updating data analysis results

In addition to the normal data analysis procedure, results are updated when any of the *a priori* information is improved. The most frequent case of this is when the IERS issues a bulletin, and the procedures to update the analysis results are also automated as shown in Fig. 2. The IERS is responsible for providing up-to-date series of parameters related to the Earth's rotation and issues Bulletin-A twice a week and Bulletin-B once a month. At the time of the initial data analysis, only the predicted values of δx, δy, and $\mathrm{UT1} - \mathrm{UTC}$ are available as *a priori* information. No predictions are provided for the nutation offsets, so these values should be obtained using an empirical model(4) tabulated in the IERS Convention 1996. The values of these parameters are improved in the bulletins issued after the experiment. The bulletins are received by e-mail and are processed automatically by the data analysis software. All the databases for which *a priori* information is revised are updated with the revised *a priori* information by the routine **DBUPDATE**. These databases are then processed by the routine **CALC**, and finally the parameters are reanalyzed by the routines **DATSET** and **VLBEST**.

The Terrestrial Reference Frame and Celestial Reference Frame are also refined occasionally by IERS. The latest reference frames available at present are respectively ITRF96(5) and RSC(WGRF)95R01.(6) Since the revisions of the reference frames affect all of the KSP VLBI analysis results, all databases have to be updated and analyzed. These procedures are not automated since the revision is not frequent and the format of the data file for the reference frames are not the same every time. Instead, the *a priori* data files should be prepared and the databases have to be updated manually.

2.4 Publication of the analysis results

The KSP VLBI system provides precise site positions from a set of observations performed during an experiment session that spans about 24 hours under the present operation arrangement. The obtained site positions can therefore be interpreted as the average positions of the reference points of the antennas during the 24 hours. If one of the observation sites moves with respect to the other three sites, the motion can be detected promptly. The rapidity and the precision are two valuable aspects of the KSP VLBI results. To make full use of the KSP VLBI data, the analyzed results are placed in an area accessible from any hosts on the Internet by hyper text transfer protocol (http) and file transfer protocol (ftp). One can access the latest results of the KSP VLBI data by using a World Wide Web (WWW) browser or a ftp client program. Unified resource locators of the server are `http://ksp.crl.go.jp/` for the WWW access and `ftp://ksp.crl.go.jp/dist/` for the anonymous ftp access. The Japan Meteorological Agency accesses the KSP VLBI results twice every day and is monitoring the data along with other data such as from the Global Positioning System (GPS), strain meters, and tilt meters. When a regional deformation associated with seismic activities occurs, the deformation will be detected by multiple techniques. In addition to the WWW and ftp access methods, the analysis results can be sent to pre-registered addresses by e-mail for users who require rapid results.

3. Concluding remarks

The KSP VLBI data analysis system has been designed so that it does not require human operations. It enables regular geodetic VLBI experiments to be conducted with four VLBI sites at a time interval of one or two days. The experimental results are produced quickly: in the tape-based VLBI mode, they can be obtained within two days after the last observation in the experiment, whereas in the real-time VLBI mode, they can be obtained within just a half an hour after the last observation. The results are also automatically updated when the updated Earth Orientation Parameters are delivered by e-mail from IERS. The results are made available to the public over the Internet.

To improve the results in the future, we should switch the planetary ephemeris to the DE403 model. The atmospheric delay correction can be improved by introducing an anisotropic troposphere model. Detailed studies of the tropospheric delay are being performed by using VLBI and GPS data at present.[7] The performance of the hardware system, in particular the performance associated with the limited disk space, can be improved by replacing the workstation system with one of the latest models. The operating system currently used for the workstation is reported to be unreliable beyond the year 2000 and the replacement of the workstation system is now under preparation.

The usefulness of the VLBI data obtained from the regular KSP experiments can be greatly improved if the results are generated in SINEX (Solution Independent Exchange) format. This format would enable the results to be combined with other geodetic measurements, including GPS and Satellite Laser Ranging (SLR) measurements. A detailed comparison of the VLBI results and results from GPS and SLR measurements is one of the major purposes of the KSP and is expected to contribute to the unification of the terrestrial reference frames constructed by multiple space geodetic techniques.

Acknowledgments

The authors would like to thank colleagues at the Goddard Space Flight Center of the National Aeronautics and Space Administration for permission to use the CALC and related softwares in the KSP VLBI data analysis system.

References

(1) D. D. McCarthy, "IERS Conventions (1996)," IERS Technical Note 21, issued by International Earth Rotation Service, 1996.

(2) E. M. Standish, "The Observational Basis for JPL's DE200, the Planetary Ephemerides of the Astronomical Almanac," Astron. Astrophys., vol. 233, pp. 252-271, 1990.

(3) E. M. Standish, X. X. Newhall, J. G. Williams, and W. F. Folkner, "JPL Planetary Ephemerides of the Astronomical Almanac," JPL IOM 314.10-127, issued by Jet Propulsion Laboratory, 1995.

(4) T. A. Herring, "VLBI Studies of the Nutations of the Earth," in The impact of VLBI on astrophysics and geophysics; Proceedings of the 129th IAU Symposium, Kluwer Academic Publishers, pp. 371-375, 1988.

(5) C. Boucher, Z. Altamini, and P. Sillard (eds), "Results and Analysis of the ITRF96," IERS Technical Note 24, issued by International Earth Rotation Service, 1998.

(6) C. Ma and M. Feissel (eds), "Definition and realization of the International Celestial Reference System by VLBI astrometry of extragalactic objects," IERS Technical Note 23, issued by International Earth Rotation Service, 1998.

(7) R. Ichikawa, Y. Koyama, and Y. Takahashi, "Effects of Atmospheric Valiability at KSP Network - Preliminary Results," J. Commun. Res. Lab., this issue, 1999.

Fig. 1. Relationships between processes and data files required to create database files. The rectangles represent programs and the rounded rectangles represent data files.

Fig. 2. Schematic data flow and block diagram of the automated VLBI data analysis. Both initial data analysis and the reanalysis after the latest Earth's Orientation Parameters become available are illustrated.

Fig. 1

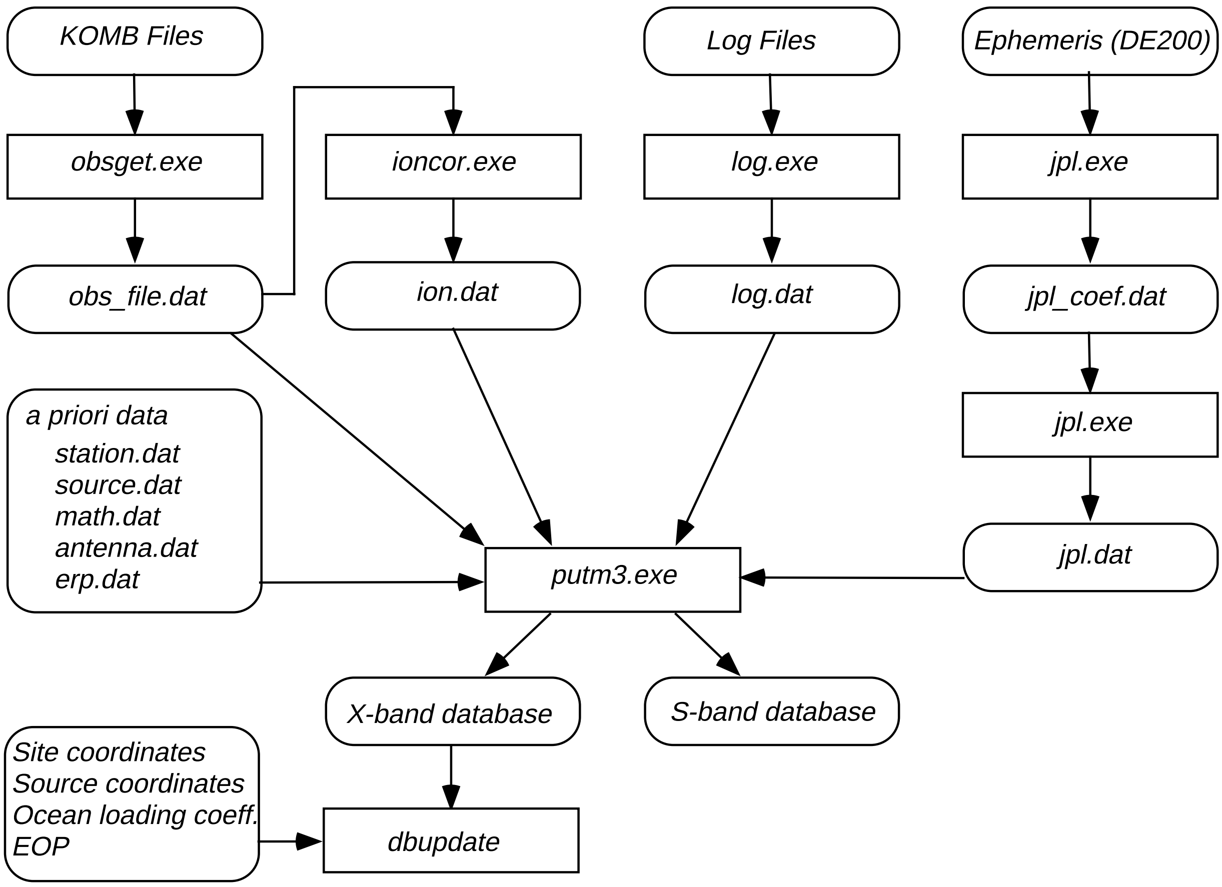
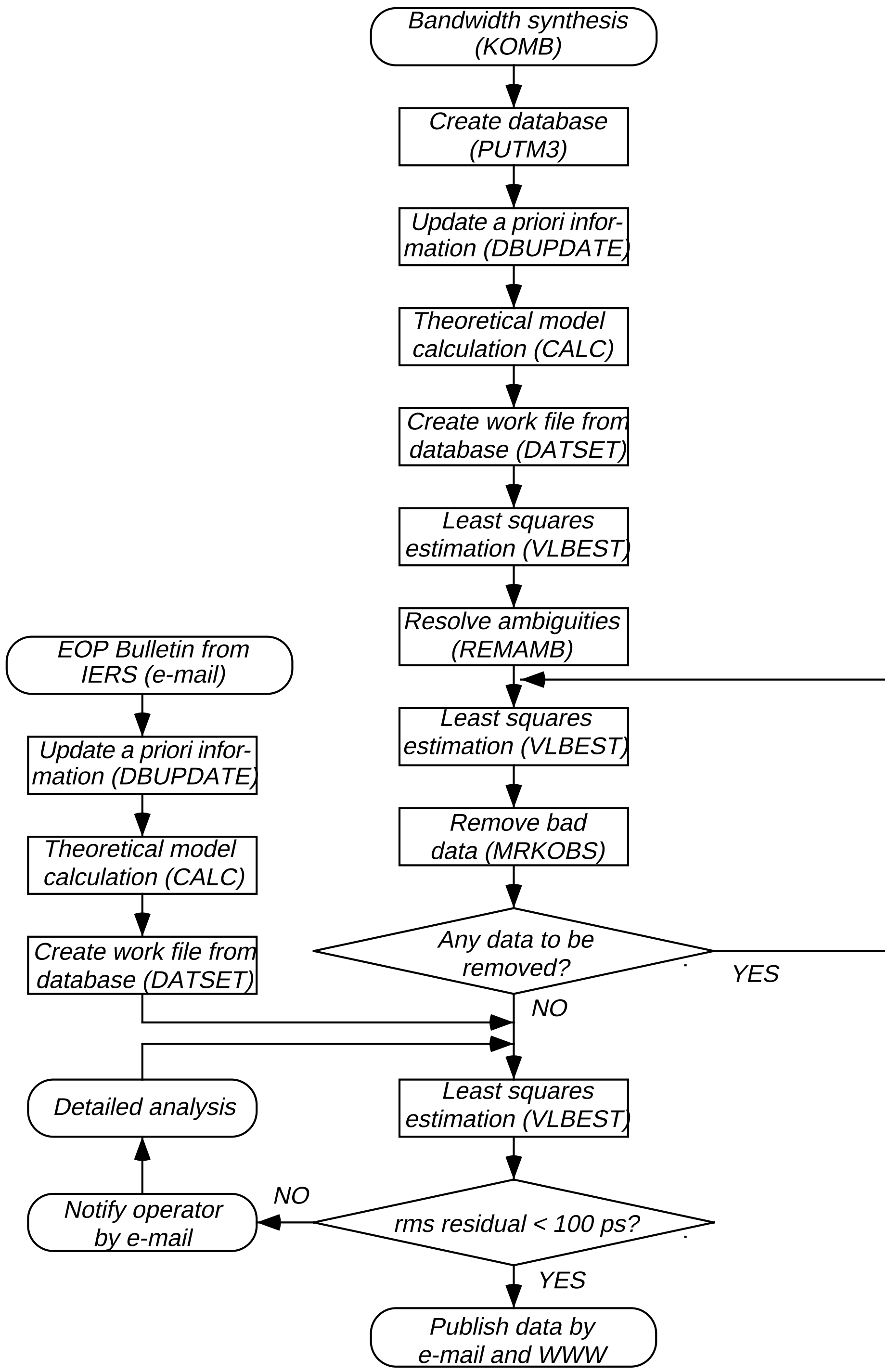

KOMB Files
Log Files
Ephemeris (DE200)
obsget.exe
ioncor.exe
log.exe
jpl.exe
obs_file.dat
ion.dat
log.dat
jpl_coef.dat
a priori data
station.dat
source.dat
math.dat
antenna.dat
erp.dat
jpl.exe
putm3.exe
jpl.dat
X-band database
S-band database
Site coordinates
Source coordinates
Ocean loading coeff.
EOP
dbupdate

Fig. 2



Bandwidth synthesis (KOMB)
Create database (PUTM3)
Update a priori information (DBUPDATE)
Theoretical model calculation (CALC)
Create work file from database (DATSET)
Least squares estimation (VLBEST)
Resolve ambiguities (REMAMB)
Least squares estimation (VLBEST)
Remove bad data (MRKOBS)
Any data to be removed?
YES
NO
EOP Bulletin from IERS (e-mail)
Update a priori information (DBUPDATE)
Theoretical model calculation (CALC)
Create work file from database (DATSET)
Detailed analysis
Least squares estimation (VLBEST)
Notify operator by e-mail
NO
rms residual < 100 ps?
YES
Publish data by e-mail and WWW