

# **IVS Memorandum 2008-017v01**

**07 September 2008**

**“Towards Automated Data Analysis”**

***Zinovy Malkin***

## 1. Introduction

VLBI2010 concept envisages (near)-real-time correlation processing following by rapid automated analysis aimed, in the first place, at determination of EOP. In other words, a seamless data flow from antennas backend to EOP uploading to IVS and IERS combining centers is anticipated. To achieve this goal, automated reliable procedures should be developed at all stages of the VLBI data processing, when yet not implemented. Existing procedures and data formats should be updated and unified, when necessary, to provide a final result expected from the VLBI2010 network.

It seems to be natural if the automated rapid analysis aimed at operational EOP computation will be organized at the same IVS Network Control and Processing Center, where the network control and correlator processing is performed. Further complete scientific analysis will be conducted in IVS Analysis Centers, and not discussed here.

In this memo, some problems related to organization of automated data analysis are discussed. We will start our consideration from the delays computed on correlator for each frequency band and stored in the appropriate data structure along with other information extracted from raw correlator output, schedule, log files, etc. (data files). This data storage is close to the current database v. 1 enriched with supplement info extracted from log files and needed for further data processing.

## 2. Overview of the automated data analysis

The whole set of operational determination of EOP includes the following steps.

- 1.1. Retrieve data files from correlator and/or IVS data center.
- 1.2. Retrieve or compute *in situ* other data used for analysis, such as a priori EOP, atmospheric loading, mapping function, tropospheric gradients, master file, etc.
- 1.3. Perform data analysis.
- 1.4. Perform quality check.
- 1.5. Upload result to IVS data center.

The main operational data analysis tasks are:

- 2.1. Compute and apply ionosphere correction.
- 2.2. Resolve ambiguities.
- 2.3. Interpolate meteo parameters at the epochs of the scans observed.
- 2.4. Compute EOP.

During EOP computation the analyst usually has to solve several tasks, as a rule in an interactive mode:

- 3.1. Choice of clock reference station.
- 3.2. Elimination of outliers.
- 3.3. Detection of clock breaks.
- 3.4. Cable cal handling.
- 3.5. Adjusting of parameterization.
- 3.6. Detection of abnormal stations behavior and corresponding adjustment of the estimation procedure.

Many analysis steps listed above are fully or partly automated at different IVS Analysis and Correlator Centers, others are yet under development. Fully automated computation of UT1 from Intensive sessions was implemented at the IAA IVS Analysis Center in 2001. The procedure includes all the necessary steps from downloading all data needed for processing to uploading the result to IVS data center [1,2]. This procedure started with the database v. 4 provided in the IVS data center. Advanced automated analysis procedures, including earlier steps, such as ambiguity resolution, and thus providing the whole automated data flow from correlation to UT1, were recently developed at the NICT Kashima Space Research Center [2].

However, the automated analysis of 24h sessions with computation of full set of EOP along with troposphere and parameters of interest is a more complicated task, and analysis of these sessions often requires decisions made by the analyst. Our experience shows that about 99% of intensive sessions processed in the automated mode do not require further re-visiting by an analyst, whereas only 80–85% R1/R4 sessions give a satisfactory results being processed semi-automatically (i.e. w/o manual corrections). The rest of sessions require manual intervention, mainly due to clock breaks, and, to less extent, due to other reasons, such as choice of the clock reference station, excessive station noise and so on.

### **3. Automation of specific analysis tasks**

(to be developed in the next versions)

### **4. Supplement requirements to VLBI systems and data files**

To make automated data analysis more simple and reliable, some changes in the existing procedures will be valuable. Those includes (subject to discussion and amplification):

- Standardize (use more strict formats for) data files delivered from stations: log files and meteo files in the first place.
- Standardize and extend correlator report format to provide easy automated extraction of the info needed for analysis, such as manual phase cal applied (or equivalent procedures which may be used during automated correlator processing), clock breaks detected, etc.
- Standardized messages of the station and network control software on important events which may affect data analysis.

### **5. References**

1. Malkin Z., Skurikhina E., Sokolskaya M., Krasinsky G., Vasilyev M., Gubanov V., Surkis I., Kozlova I., Rusinov Yu. IAA VLBI Analysis Center Report 2001. In: N.R.Vandenberg, K.D.Baver (Eds.), IVS 2001 Annual Report, NASA/TP-2002-210001, 2002, 220-223.
2. Malkin Z., E. Skurikhina. OCCAM/GROSS Software Used at the IAA EOP Service for processing of VLBI Observations. Transactions. IAA, 2005, v. 12, 54-67. (in Russian)
3. Koyama Y., Sekido M., Hobiger T. Takiguchi H., Kondo T. Developments of an Automated Data Processing System for Ultra Rapid dUT1 e-VLBI Sessions. 2008 IVS General Meeting Proceedings, in press.