

ANUBIS – A TOOL FOR QUALITY CHECK OF MULTI-GNSS OBSERVATION AND NAVIGATION DATA

Pavel Vaclavovic⁽¹⁾, Jan Dousa⁽¹⁾

⁽¹⁾ Research Institute of Geodesy, Topography and Cartography, Ústecká 98, Zdiby 250 66, Czech Republic, Email: pavel.vaclavovic@pecny.cz, jan.dousa@pecny.cz

ABSTRACT

The GNSS software library G-Nut has been developed at the Research Institute of Geodesy, Topography and Cartography since 2011. Along with the PPP applications for positioning and troposphere monitoring, the third application recently built and released is called Anubis. Its main purpose is to support data editing, cutting and splicing, quality and quantity monitoring for GNSS from various formats – currently RINEX2x and RINEX3x files, but RTCM real-time streams in future. The Anubis supports observation data from all global navigation satellite systems and their augmentations (GPS, GLONASS, Galileo, BeiDou, SBAS and QZSS).. It can be used for observation quality and quantity monitoring which includes including pre-processing. Whenever supported with navigation message, it provides also elevation and azimuth dependent characteristics. The pre-processing mode is used for the reconstructing observations affected by cycle slips or receiver clock jumps. The first version of Anubis was released in the mid of 2013 under the GNU General Public Licence version 3.

1. INTRODUCTION

Geodetic observatory Pecny (GOP) operates the IGS and EUREF analysis centre for processing GNSS data for coordinate and velocity estimation, troposphere monitoring and GNSS orbit determination. Data from national, European and global sites are used for these tasks, i.e. stemming from various sources. Data are usually disseminated in the RINEX format (Receiver Independent Exchange) [1] usually without support of information about the quality and content. Any corrupted file may cause unexpected behaviour in analyses thus usually requiring special manual interventions. Data quality monitoring can provide information not only for data processing activities, but also for high-quality data collection and archiving by individual providers.

New challenges arise from emerging new multi-GNSS signals, frequencies and constellations. The RINEX 3.x format has been standardized for including all new data, but not all software tools are ready to support it. Such data needs to be properly monitored and tested what suggests developing new tools for this purpose. The

paper describes an open-source tool Anubis, which has been developed for editing and quality checking of observation and navigation multi-GNSS data in RINEX 2.x and RINEX 3.x file formats as well as for the RTCM real-time streams in future.

The Anubis application is derived from the G-Nut library [2] developed at GOP of the Research Institute of Geodesy, Topography and Cartography in Czech Republic. The library is written in C++ [3] and it is designed mainly for a command-line utilization supported by the XML configuration file. The G-Nut library is not published as whole, but applications are released under the GNU public license V3.0 from the web <http://www.pecny.cz/>.

The first section describes editing capabilities of RINEX files, a standard format for archiving GNSS data. The second section concerns the quality checking functionality of the software. This section is the most important since observation data can be corrupted by several ways such as multipath, cycle slips, receiver clock jumps, data gaps and others. The third section demonstrates some examples of the software utilizations and outputs. The final section concludes the paper.

2. DATA EDITING

There are several situations when editing of a file is necessary. For instance, to merge hourly files into the daily one or vice versa, to convert data from streams to files, to cut data with respect to a particular time span, data re-sampling and many others. It can be useful for the debugging of pre-processing algorithms, e.g. particular number of cycles can be added to carrier phase observations while trying to identify and estimated them afterwards. Advantage of such procedure is that we a priori know the number of introduced cycle slips as well as all affected epoch for comparing with our estimation.

To support editing mode data must be decoded, saved in internal structures, modified with regards to the user requirements and, finally, stored in a file in required format. Currently, all steps have been implemented in the Anubis, with exception of the last one. Although the RINEX 2.x and 3.x decoders have been already

finalized and seriously tested exploiting data from the IGS MGEX campaign [4]. The RINEX encoder is, however, scheduled for the implementation in the second software development phase. The first phase aimed for the proper multi-GNSS data decoding, monitoring and pre-processing, i.e. providing qualitative and quantitative statistics of data content. In the second phase the decoder for the RTCM binary format will be developed so data disseminated in real-time could be handled too.

3. QUANTITATIVE STATISTICS

High precision GNSS data usually consist of carrier phases and pseudo-ranges from at least two frequencies and usually more global navigation satellite systems, GPS, GLONASS, Galileo, BeiDou, and their regional augmentations, such as QZSS, WAAS/EGNOS or IRNSS. We analyzed data from the Multi-GNSS Experiment (MGEX) campaign, the project of the International GNSS Service (IGS) aimed for collecting observations from all existing GNSS constellations and available signals of all satellites. Many inconsistencies and incompleteness can be noticed already by a visual inspection since the data set is truly an experimental one. As the first step, we checked contents of data files disseminated in RINEX 3.x. The content is highly variable due to frequent receiver firmware updates – providing new observations, frequencies or satellites.

This was the main reason for implementing RINEX data content monitoring in Anubis. Achieved quantities are extracted into a summary file which is divided into several sections with an individual verbosity. One of the sections summarizes the RINEX metadata and compares them with that set up by user. Special section collects observation type statistics using information from the RINEX header and from data decoded. Another section summarizes RINEX body content, such as enumerate all available satellite systems, satellites, signals and bands. A rough position is computed based on optional navigation message and compared with approximate coordinates provided in the RINEX header. It is well known that processing of dual- or multi-frequency data is more accurate due to the possibility of forming ionosphere linear combinations. For this reason numbers of bands (frequencies) are monitored for each satellite and observation types (pseudo-range and carrier-phase). Other quantitative factors, such as frequent data gaps or small data pieces, causing potential difficulties in a data processing, are monitored too.

4. QUALITATIVE CHARACTERISTICS

Two strategies are applied for the qualitative assessment of data files. These are described in the next sections – a) observation pre-processing (clock jump and cycle slip

detection) and b) code multi-path estimation.

4.1. Observation pre-processing

The pre-processing for searching and repairing clock jumps and phase cycle slips is very important part of any software dealing with GNSS carrier phase data analysis. In high-accurate applications periods with continuous satellite tracking can be only used efficiently, because in such case a single initial ambiguity is set for each satellite. If the continuity is broken for a satellite and cycle slip cannot be estimated, specific ambiguity must be added meaning an additional estimated parameter. Anubis uses various time differenced linear combinations (LC) that are compared with predefined thresholds. Cycle slip detection algorithm is based on Melbourne-Wübbena [5] and geometry-free LC due to their useful properties. This linear combination (L_4) is defined by equation

$$\begin{aligned} L_4 &= L_1 - L_2 = \lambda_1 N_1 - \lambda_2 N_2 - I_1 + I_2 \quad (1) \\ &= \lambda_1 N_1 - \lambda_2 N_2 - \left(1 - \frac{f_1^2}{f_2^2}\right) I_1 \end{aligned}$$

where subscript 1 and 2 stands for band number, L is carrier frequency in meter unit, λ denotes wavelength, N is initial ambiguity, I is ionosphere delay and f means frequency. L_4 is independent of receiver clock errors and geometry (satellite/receiver position) and contains only ionosphere delays and initial ambiguities for both frequencies. All other frequency-independent terms are neglected. The first two terms on the right site of Eq. 1 are constant in time meaning that unexpected jump in L_4 must be caused by cycle slip. The detection is based on the following criterion

$$L_4(t_2) - L_4(t_1) > k\sigma_{L_4} + \Delta I_{max} \quad (2)$$

The maximal ionosphere delay ΔI_{max} is implicitly defined as 0.4 m/hour in Anubis and the factor k is usually set to 4. The advantage of such approach is that it is based on carrier phase data only. Otherwise, it should be noted that in case of a positive test, we do not know whether any of L_1 , L_2 or both are corrupted.

If dual-frequency carrier-phase L and pseudo-range P are available, the Melbourne-Wübbena LC (L_6) can be formed

$$\begin{aligned} L_6 = L_W - P_N &= \frac{1}{f_1 - f_2} (f_1 L_1 - f_2 L_2) \quad (3) \\ &\quad - \frac{1}{f_1 + f_2} (f_1 P_1 + f_2 P_2) \\ &= \lambda_W N_W \end{aligned}$$

where $\lambda_W = \frac{c}{f_1 - f_2}$ and $N_W = N_1 - N_2$ are called wide-lane wavelength and ambiguity, respectively.

Advantage of using L_6 combination consists in the elimination of ionosphere, troposphere, geometry (satellite and receiver positions) and satellite and receiver clocks. The wavelength of such combination is about 86 cm . On the other hand the inclusion of pseudo-range data increases the noise of the linear combination. Comparing L_6 for epoch k and $k-1$ give us information whether slip occurs or not. It should be noted that slips on L_1 or L_2 cannot be checked directly, but their difference $N_1 - N_2$ only. Due to a constant property of the right term in Eq. 3 we can easily check any cycle-slip by temporal differencing L_6 observation. The detection is based on the criteria

$$L_6(t_2) - L_6(t_1) > k\sigma_{L_6} \quad (4)$$

where the coefficient k is again set to 4 and σ_{L_6} is the sigma of L_6 observation. Since a cycle-slip on any specific frequency cannot be detected, but the L_1 - L_2 cycle slip only, any cycle-slip common to L_1 and L_2 become undetectable. An improvement of the technique thus lies in differencing of L_6 within a single epoch and in using a mean value over all previous epochs since the last occurring cycle slips.

The second purpose of the pre-processing consists of searching and correcting for receiver clock jumps. Due to a low quality of some receiver oscillators, clocks are shifted by one or few milliseconds when clock bias becomes too large. Observations at particular epoch as well as observations in all subsequent epochs are affected in the same way and must be corrected otherwise ambiguity re-initialization and new convergence interval would regularly appear. The principle of our algorithm resists in the pseudo-range compensations of clock jumps, while carrier phases for each satellite could still contains the same cycle slip [6]. Fortunately, we know that the slip is exactly a millisecond or a few milliseconds; therefore, we can repair it precisely. Anubis can be thus used for recovering the coherency between range and phase data. One section of Anubis extraction provides results from cycle slip and receiver clock jump detections. Values of cycle slips and clock jumps as well as epochs at which these occur are reported.

4.2. Code multipath estimation

The multipath affects both basic GNSS observations – pseudo-ranges and carrier phases, however, the former is much larger and highly variable between various receiver types. The multipath error has substantial contribution to the accuracy of observed pseudo-ranges, which are mainly used in a single point positioning technique (navigation, precise point positioning etc.). The knowledge of the multipath effect can be useful for proper pseudo-range weighting. Such information can

also provide specific quality assessment of the receiver. When dual-frequency data are available, the pseudo-range multipath is estimated from the linear combination eliminating the satellite-receiver geometry and all atmospheric effects. The combination, however, does not eliminate ambiguities and any differential bias. While the latter is constant over time, it is not always true for ambiguities due to a presence of cycle slips. The pre-processing (and optionally cycle slip reparation) is thus important for the multipath estimation.

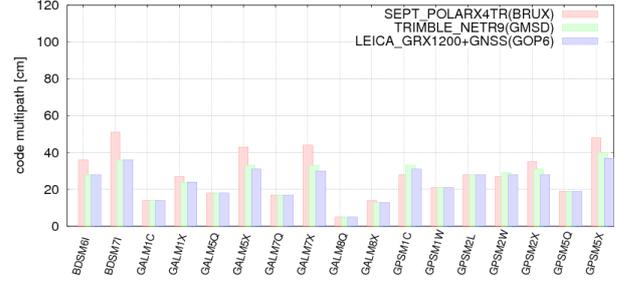


Figure 1. Average pseudo-range multipath for all available codes estimated for three different receiver types at stations BRUX, GMSD and GOP6.

We have developed a generalized formula for Anubis using linear combination P_{mp} supporting all pseudo-range multipath estimation, i.e. for any frequency, any available signal and any GNSS constellation providing dual-frequency observations:

$$P_{mp} = P_k - L_i - \alpha (L_i - L_j) \frac{1}{f_i + f_j} \quad (5)$$

with

$$\alpha = \frac{(f_i^2 + f_k^2)}{(f_i^2 - f_j^2)} \cdot \frac{f_j^2}{f_k^2} \quad (6)$$

The multipath is estimated as a standard deviation over a sequence of consecutive epochs (usually 15-30) where calculated mean represents all remaining biases. We do not require any specific pre-processing (e.g. for every involved GNSS) because a simple cycle slip detection algorithm is implemented as a part of the multipath estimation. Fig.1 shows the example of pseudo-range multipath estimation calculated for all signals of all available satellites at three M-GEX stations – BRUX, GMSD and GOP6, which represents three different receiver types – SEPT POLARX4T, TRIMBLE NETR9 and LEICA GRX1200+GNSS, respectively. The figure displays average values while Anubis can also provide time-dependent satellite-specific multipath estimate, see Fig.2 for GPS pseudo-range on L1 and L2 frequencies for station BSCN in day 001 of 2013. The low-elevation observation has typically larger multipath errors as well as pseudo-ranges on L1 have lower multipath than L2.

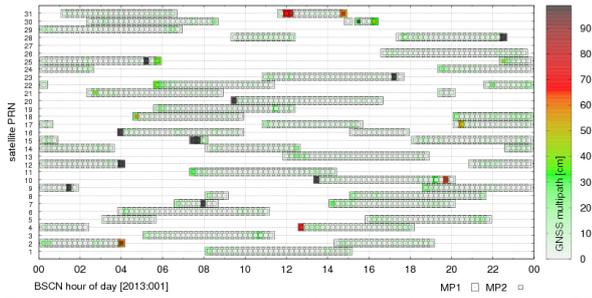


Figure 2. Time-dependent, satellite-specific multipath for GPS pseudo-ranges on L1 (squares) and L2 (circles) frequencies.

5. USER CONFIGURATION

The software can be executed from a command line with a single parameter defining configuration filename in the XML format. The format has been chosen because of its flexibility, easy extension, self-contained structure and a support in various end-user editors. The format is applied for all applications derived from the G-Nut library while different sections correspond to specific functionalities. For instance, common elements are defined for the input, output and general settings, while specific element `<qc>` is used by Anubis only. Following lines shows an example of the configuration:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<!DOCTYPE config>
<config>
<gen>
<beg> "2013-02-09 00:00:00" </beg>
<end> "2013-02-09 23:59:30" </end>
<sys> GPS GLO GAL BDS SBS QZS </sys>
<int> 30 </int>
<rec> BRUX GOPE MATE </rec>
</gen>
<inputs>
<rinexo> RINEX/mate0400.13o </rinexo>
<rinexo> RINEX/gope0400.13o </rinexo>
<rinexo> RINEX/brux0400.13o </rinexo>
<rinexn> RINEX/brux0400.13n </rinexn>
<rinexn> RINEX/brux0400.13g </rinexn>
<rinexn> RINEX/brux0400.13l </rinexn>
</inputs>
<qc
sec_sum="1"
sec_hdr="1"
sec_est="1"
sec_obs="1"
sec_gap="1"
sec_bnd="2"
sec_pre="1"
sec_ele="1"
sec_mpt="2"
int_stp="1200"
int_gap="600"
int_pcs="1800"
mpx_nep="15"
mpx_lim="3.0"
/>
<outputs verb="1" >
<xtr> LOG/IGSv3/2013/040/$(rec)_130400.xtr </xtr>
<xml> LOG/IGSv3/2013/040/$(rec)_130400.xml </xml>
</outputs>
</config>
```

The section `<gen>` defines general information, such as the beginning and the end epoch of data (`beg`, `end`), list of requested satellite systems (`sys`), sampling interval (`int`) and the list of marker names (`rec`). The section `<inputs>` defines all input files in various formats, such as observation (`rinexo`) and navigation (`rinexn`) data in RINEX format. If navigation file is defined, various extracted quantities are supported with azimuth and elevation information.

The section `<qc>` provides specific settings for Anubis – first attributes defines verbosity level for specific extractions – summary information (`sec_sum`), header meta-data and user requests (`sec_hdr`), observation statistics (`sec_obs`), data gaps and small data pieces (`sec_gap`), available band counting for all observations (`sec_bnd`), detection of cycle slips and clock jumps (`sec_pre`), azimuth and elevation information (elevation) and multipath estimation (`sec_mpt`). Last attributes concern specific setting – interval step in seconds for all time-dependent characteristics (`int_stp`), intervals in seconds for detecting gaps and small data pieces (`int_gap`, `int_pcs`) and, finally, settings for the multipath estimation – number of epochs used for multipath calculation (`mpx_nep`) and factor for sigma multiplication for cycle-slip detection (`mpx_lim`).

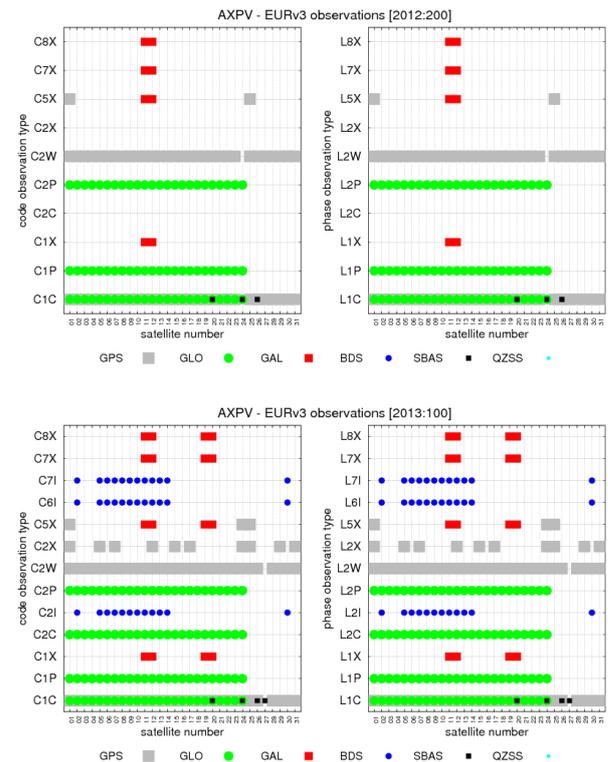


Figure 3. Observation types from the RINEX header (y-axis) compared with real data content (marks) for all satellites (x-axis) collected from AXPV station, the day 200 in 2012 (top) and the day 100 in 2013 (bottom).

The last section *<outputs>* defines requested output files (pre-defined uniquely for all processed sites). When configuring Anubis to process more RINEX files, these can be processed in parallel.

6. RESULTS

Selected plots demonstrate the current functionality of Anubis. Fig.3 shows provided RINEX header information together with real data types collected from data itself. Two plots for two different epochs are shown (day 200 in 2012, day 100 in 2013) for station AXPV from EUREF's RINEX3.0 testing repository.

The Y-axis shows all observation types reported in the RINEX header and the X-axis shows satellite PRN numbers. The plots then display the presence of specific observation types for various satellite systems shown in different colours. Individual satellites transmitting particular types can be identified as well as emerging signals after receiver firmware upgrades. In the first plot the receiver reports new signal in the header only (blank lines for the specific signals, e.g. C2X or C2C), while later real data are collected for these signals (filled specific satellites columns for these new signals).

Another example provides a count of bands (frequencies) available for phase and code observations as well as all GNSS constellation (Fig.4).

Sometimes it is useful to know how data quality evolves in time (e.g. for re-analyses efforts). By default, Anubis provides a simple text file output that can be easily searched by linux *grep* command so the time-series of specific parameters can be generated. Some example sections of the Anubis output is shown in appendix.

ACKNOWLEDGEMENT

The work was supported by the Czech Science Foundation (project no P209/12/2207).

REFERENCES

1. Gurtner, W. and Estey, L. (2009). Online at: http://igsceb.jpl.nasa.gov/igsceb/data/format/rinex30_1.pdf.
2. Vaclavovic, P., Dousa, J., Gyori, G. (2013). G-Nut software library - state of development and first results, *Acta Geodyn. Geomat.*, 10: 172.
3. Musser, D. (2001). STL tutorial and reference guide: C++ programming with the standard template library. *Addison Wesley*.
4. Weber, R. (2012). The IGS Multi-Signals Tracking Campaign MGEX – Planning, Status and Perspectives, Presented at the IGS Workshop 2012, July 23-27, Olsztyn, Poland.

5. Wübbena, G. (1985). Software development for geodetic positioning with GPS using TI 4100 code and carrier measurements. In: *Proceedings 1st international symposium on precise positioning with the global positioning system*, Rockville, Maryland, pp 403-412
6. Guo, F. and Zhang, X. (2013). Real-time clock jump compensation for precise point positioning. *GPS solution*, doi: 10.1007/s10291-012-0307-3.

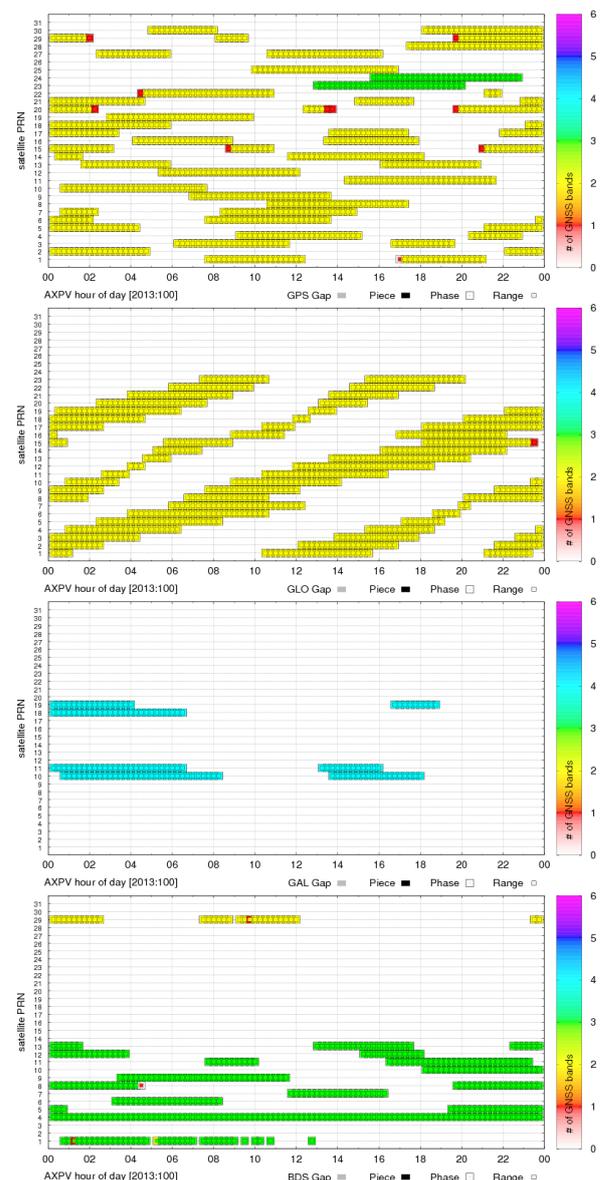


Figure 4. Plots of counted bands (frequencies) for all satellites available from GPS, GALILEO, Galileo and BeiDou systems.

