

Status of the EPN Historical Data Center

EPN CB Team

Royal Observatory of Belgium

Contact: epncb@oma.be

Background

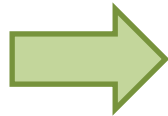
- Upgrade of our Data Center to comply to FAIR data principle
- FAIR-GNSS project
 - Partners: **Royal Observatory of Belgium & Ghent University**
 - Funding Agency: **Belgian Science Policy Office (BELSPO)**
 - Duration: **15/12/2020 – 14/03/2023**
 - Goal: **Create an open data portal for European and Belgian GNSS reference station data collections, built upon FAIR guiding principles**

FAIR-GNSS webinar material available at <https://fair-gnss.oma.be/webinar.php>

OLD database

| DataCenters | | |
|------------------|---|--|
| * id | # | |
| * LongMarkerName | c | |
| * Centre | t | |
| * RnxDir | c | |
| * DATE | d | |
| * Filename | t | |
| RnxVersion | c | |
| Size | # | |
| DateFirst | d | |
| DateLast | d | |
| * DateOfUpdate | d | |
| Flag | # | |

| ROBDataRepository | | |
|-------------------|---|--|
| * id | # | |
| * LongMarkerName | c | |
| * RnxDir | c | |
| * DATE | d | |
| * Filename | t | |
| RnxVersion | c | |
| Size | # | |
| DateFirst | d | |
| DateLast | d | |
| * DateOfUpdate | d | |
| Flag | # | |



New database

Decide what to download

| DataCenters | | |
|------------------|---|--|
| * id | # | |
| * longmarkername | c | |
| * Centre | t | |
| filedate | d | |
| * Filename | t | |
| RnxVersion | c | |
| Size | # | |
| DateFirst | d | |
| DateLast | d | |
| * DateOfUpdate | d | |
| Flag | # | |

Origin of the files

| gnssdata_from | | |
|-------------------------|---|--|
| * idfilefrom | # | |
| * idfile | # | |
| filelib | # | |
| centrefrom | t | |
| url | t | |
| filename | t | |
| filesize | # | |
| checksum_md5 | c | |
| checksum_sha1 | c | |
| publication_date_online | d | |
| * table_timestamp | d | |

| gnssdata_from_headers | | |
|--------------------------|---|--|
| * idfilefrom | # | |
| filename | t | |
| rinex_version | t | |
| rinex_type | t | |
| marker_name | t | |
| marker_number | t | |
| observer | t | |
| agency | t | |
| receiver_sn | t | |
| receiver_type | t | |
| receiver_vers | t | |
| antenna_sn | t | |
| antenna_type | t | |
| antenna_delta_h_string | t | |
| antenna_delta_n_string | t | |
| antenna_delta_e_string | t | |
| approx_position_x_string | t | |
| approx_position_y_string | t | |
| approx_position_z_string | t | |
| interval_string | t | |
| interval_obs | # | |
| time_first_obs_string | c | |
| time_system | c | |
| time_last_obs_string | c | |
| table_timestamp | d | |

Provenance of each data file

Link to the original file

Information on the files

| gnssdata | | |
|-----------------|---|--|
| * idfile | # | |
| fileversion | # | |
| longmarkername | c | |
| rinex_format | c | |
| filedate | d | |
| filename | t | |
| filesize | # | |
| checksum_md5 | c | |
| checksum_sha1 | c | |
| datefirst | d | |
| datelast | d | |
| flag | # | |
| deleted | # | |
| check_mode | # | |
| table_timestamp | d | |

| gnssdata_headers | | |
|-------------------|---|--|
| * idfile | # | |
| filename | t | |
| rinex_version | t | |
| rinex_type | t | |
| marker_name | t | |
| marker_number | t | |
| observer | t | |
| agency | t | |
| receiver_sn | t | |
| receiver_type | t | |
| receiver_vers | t | |
| antenna_sn | t | |
| antenna_type | t | |
| antenna_delta_h | # | |
| antenna_delta_n | # | |
| antenna_delta_e | # | |
| approx_position_x | # | |
| approx_position_y | # | |
| approx_position_z | # | |
| interval_header | # | |
| interval_obs | # | |
| time_first_obs | d | |
| time_system | c | |
| time_last_obs | d | |
| * table_timestamp | d | |

| gnssdata_pgm | | |
|-----------------|---|--|
| * id_pgm | # | |
| idfile | # | |
| pgm | t | |
| pgm_run_by | t | |
| pgm_date_string | t | |
| pgm_date | d | |
| from_comment | # | |

| gnssdata_headers_mandatory | | |
|----------------------------|---|--|
| * idfile | # | |
| obstypes | # | |
| sys_phase_shift | # | |
| glonass_slot_frq | # | |
| glonass_cod_phs_bis | # | |

Operations done on the files

| file_actions | | |
|-------------------|---|--|
| * id_file_actions | # | |
| idfile | # | |
| fileversion | # | |
| metadata_used | c | |
| action_name | # | |
| * table_timestamp | d | |

| file_errors | | |
|------------------|---|--|
| * id_file_errors | # | |
| id_file_actions | # | |
| error_type | # | |

| file_changes | | |
|-------------------|---|--|
| * id_file_changes | # | |
| id_file_actions | # | |
| changetype | t | |
| changefrom | t | |
| changeto | t | |

| validated_observer | | |
|--------------------|---|--|
| observer | t | |
| agency | t | |
| action | # | |
| * timestamp | # | |

Operations done on each data file

Implementation

- Software design and development (started in Summer of 2021)
- Testing on recent data and historical data (since November 2021)
- Re-population of the database with additional required information (since February 2022, restarted in June 2022 after a meeting with the Regional DC)
 - Data after 1st January 2022 + all incoming files (recent or historical)
 - ⇒ Full information: provenance,
changes applied on the files,
header content,
warnings and errors
 - Data before 1st January 2022
 - ⇒ Limited information: header content,
warnings and errors
- Cleaning of the flags, enrichment with information from Anubis quality checks, time series monitoring (Not Started Yet)

Warnings and errors handled

Error cannot be corrected → RINEX file is not published

| |
|---|
| Station stopped RINEX 2 |
| Duplicated RINEX (e.g. _R and _S) |
| RINEX 3 long and short exists! |
| RINEX 3 bad name (not 0000) |
| crz2rnx or rnx2crz failed |
| <Site log> indicates a receiver or antenna was removed (or installed) during this day |
| No site log file for <station> in M3G |
| Invalid or isolated RINEX 3 data |
| Incorrect conversion from RNX 2 to RNX 3 |
| Former EPN station |
| Data are not valid |
| Incorrect position |
| Less than 20 epochs with observations |
| Invalid RINEX 3 data header |
| Observation sampling rate is <x sec> |
| Few epochs available |
| Mandatory line with label END OF HEADER is missing. |
| Mandatory line with label END OF HEADER is missing. |
| Misformatted observation epoch |
| Mismatch between Epoch of first observation <2020 11 12 > and date extracted from the RINEX file name |
| File contains consecutive observations from two different days |
| RINEX format in header is not a RINEX 2 or 3 version |

File with new Observer → RINEX file is quarantined

Error can be corrected → RINEX file is published

| |
|---|
| Special character detected and removed |
| Mismatch between YYYY MM DD in header string TIME OF FIRST OBS and YYYY MM DD of Epoch of first observation |
| Epoch of first observation is after time in header string TIME OF LAST OBS |
| Misformatted <YYYY MM DD HH MM SS.SSS> in header string TIME OF FIRST OBS has been corrected |
| Misformatted <YYYY MM DD HH MM SS.SSS> in header string TIME OF LAST OBS has been corrected |
| Misformatted header label has been corrected |
| Removed line with duplicate header label |
| Missing Mandatory line |
| Mismatch at the decimal second level between Epoch of first observation <2020 11 12 > and date extracted from the RINEX file name |
| Epoch of first observation is before time in header string TIME OF FIRST OBS |

Conclusion on the Status of the EPN Historical Data Center

- Upgrade of our Data Center with more information:
 - provenance,
 - changes applied on the files,
 - header content,
 - warnings and errors
- Homogeneous error detection
- Remaining work:
 - cleaning of the flags
 - add external information (Anubis QC, EPN products)