



RAPPORTI TECNICI INGV

The GNSS data and metadata
management system through the
EPOS-GNSS framework and the IRN
Node (INGV)



ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

488

Direttore Responsabile

Daniela VERSACE

Editor in Chief

Milena MORETTI (editorinchief.collane-editoriali@ingv.it)

Editorial Board

Laura ALFONSI (laura.alfonsi@ingv.it)
Christian BIGNAMI (christian.bignami@ingv.it)
Simona CARANNANTE (simona.carannante@ingv.it)
Viviana CASTELLI (viviana.castelli@ingv.it)
Luca COCCHI (luca.cocchi@ingv.it)
Luigi CUCCI (luigi.cucci@ingv.it)
Lorenzo CUGLIARI (lorenzo.cugliari@ingv.it)
Alessia DI CAPRIO (alessia.dicaprio@ingv.it)
Roberto DI MARTINO (roberto.dimartino@ingv.it)
Domenico DI MAURO (domenico.dimauro@ingv.it)
Domenico DORONZO (domenico.doronzo@ingv.it)
Filippo GRECO (filippo.greco@ingv.it)
Alessandro IAROCCI (alessandro.iarocci@ingv.it)
Mario MATTIA (mario.mattia@ingv.it)
Daniele MELINI (daniele.melini@ingv.it)
Anna NARDI (anna.nardi@ingv.it)
Lucia NARDONE (lucia.nardone@ingv.it)
Marco OLIVIERI (marco.olivieri@ingv.it)
Nicola PAGLIUCA (nicola.pagliuca@ingv.it)
Pierangelo ROMANO (pierangelo.romano@ingv.it)
Maurizio SOLDANI (maurizio.soldani@ingv.it)
Sara STOPPONI (sara.stopponi@ingv.it)
Umberto TAMMARO (umberto.tammaro@ingv.it)
Andrea TERTULLIANI (andrea.tertulliani@ingv.it)
Stefano URBINI (stefano.urbini@ingv.it)

Ufficio Editoriale

Francesca DI STEFANO Coordinatore - Segreteria di Redazione

Produzione e grafica-redazionale

Barbara ANGIONI
Massimiliano CASCONI
Francesca DI STEFANO
Patrizia PANTANI

REGISTRAZIONE AL TRIBUNALE DI ROMA N.174 | 2014, 23 LUGLIO

© 2014 INGV Istituto Nazionale di Geofisica e Vulcanologia | Rappresentante legale: Carlo DOGLIONI

Sede: Via di Vigna Murata, 605 | Roma



ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

RAPPORTI TECNICI INGV

The GNSS data and metadata management system through the EPOS-GNSS framework and the IRN Node (INGV)

Patrizia Pizzulo^{*1}, Daniele Randazzo², Juliette Legrand⁴, Jean-Luc Menut³, Sergio Bruni⁵, Antonio Avallone¹

¹INGV | Istituto Nazionale di Geofisica e Vulcanologia, Sezione Irpinia

²INGV | Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna

³CNRS | Université Côte d'Azur, Observatoire de la Côte d'Azur (OCA)

⁴ROB | Royal Observatory of Belgium

⁵INGV | Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Nazionale Terremoti

**Corresponding author*

Accepted October 28, 2024 | Accettato 28 ottobre 2024

How to cite | Come citare Pizzulo P., Randazzo D., Legrand J., Menut J.-L., Bruni S., Avallone A., (2024). The GNSS data and metadata management system through the EPOS-GNSS framework and the IRN Node (INGV). Rapp. Tec. INGV, 488: 1-24, <https://doi.org/10.13127/rpt/488>

Cover RING GNSS stations (courtesy of R. Moschillo, INGV - Sezione Irpinia) | In copertina Stazioni GNSS RING (per gentile concessione di R. Moschillo, INGV - Sezione Irpinia)

488

INDEX

Abstract	7
Introduction	7
1. The EPOS-GNSS GLASS framework	8
2. Software components of the EPOS-GNSS GLASS framework	9
3. The INGV contribution on the GLASS node	11
3.1 Bash script to synchronise the RINEX files	12
3.2 RINEX files management	12
3.3 RINEX file Indexing and Quality Check	15
4. The INGV contribution in GNSS station metadata synchronisation	16
5. Conclusions	17
Bibliography	17
Sitography	17
APPENDIX A	19

Abstract

This document describes the services for managing the Global Navigation Satellite System (GNSS) data that have been developed at the Istituto Nazionale di Geofisica e Vulcanologia (INGV) within the context of the European-funded research infrastructure EPOS (European Plate Observing System) program. These services, optimised for solid Earth research applications, seek to harmonise and standardise the collection and quality control of GNSS data and metadata.

The EPOS Thematic Core Services (TCS) GNSS developed an architecture based on Geodetic Linkage Advance System Software (GLASS) nodes which hosts the metadata information and allows access to the entire database from the web and a datacenter which hosts the data to distribute.

INGV contributes to the EPOS TCS GNSS initiative by managing the GLASS Italian Ring Node (IRN) [1], to provide data and metadata from the Rete Integrata Nazionale GNSS (RING) network [2]. The RING network currently consists of approximately 250 permanent stations all over Italy and abroad (Greece and Malta) installed for the study of active tectonic deformation at regional (Eurasia-Africa plate boundary) and local (seismogenic faults) scales.

In support of this infrastructure, INGV is developing and managing different software: *Flask Web Service Server* (FWSS) which is an interface with the local database to store the metadata; *bancadati_sync* which is the script to synchronise the RINEX data repository from the source to the local IRN node; *gnss_indexer* which automates the process to populate the database; *ring_synchronization* to upload the sitelog metadata of the GNSS stations, and other software tools to let the system working.

All software was developed in Python language (version 3.10) while the PostgreSQL database is used to store the information.

Keywords EPOS; GNSS; RINEX

Introduction

The Global Navigation Satellite System (GNSS) refers to a system consisting of a network of satellites orbiting the Earth that provide positioning, navigation, and timing information to GNSS receivers, which use them to determine their geographic position with high precision. The GNSS allows the determination of three-dimensional positions worldwide, 24 hours a day, with a precision of a few millimetres, thanks to simultaneous measurements by several GNSS satellites. This system is a primary source of data in the field of geodesy, the science that studies the shape and size of the Earth, its gravitational field, and geodynamic phenomena. The accuracy of geodetic solutions (i.e. position time series) mainly depends on the quality of the acquired GNSS data and, thus, on remote stations' reliability (i.e. stable geodetic monumentation, presence of bedrock, high-quality instrumentation).

The number of permanent GNSS stations has significantly increased worldwide in recent years and continues to grow. Nowadays, in Europe, many private and public organisations collect, manage, and distribute their own GNSS data with minimal or no coordination efforts. Data from approximately 5000 GNSS stations in the Mediterranean region are available and of interest for Solid Earth sciences [e.g., Bruni et al., 2022].

The national GNSS network of stations managed by INGV is called RING and provides data and metadata which are stored on a centralised repository and a database server called "Bancadati" which is the Italian acquisition centre.

The need for a public archive of GNSS data has been addressed by the European Plate Observing

System (EPOS [3]), where a specific working group, the Thematic Core Service GNSS Data and Products (TCS GNSS) has been working on this issue since 2015 to disseminate data made available by EPOS partners (some EU countries) participating in the initiative and proposing standards in terms of data and metadata formats as well as standard tools for their open distribution. Nowadays about 3.000 permanent GNSS stations contribute to EPOS.

The TCS products also help to study the long-term and transient deformation. Indeed, the integration of GNSS products with those provided by other EPOS TCS facilitates multidisciplinary research either at regional or pan-European level.

At the national level, in Italy, EPOS is organised through a Joint Research Unit (JRU) EPOS Italy, a partnership between Public Research Institutions, Universities, and Consortia, coordinated by INGV, with the main objective of supporting the Italian scientific community in integrating their data and services into EPOS, promoting the development of additional national infrastructures. EPOS Italy promotes the implementation of the Open Science paradigm, and its partners provide EPOS with data and services from their research infrastructures in the fields of seismology, volcanology, geodesy, geology, petrology, and geochemistry.

The system outlined in this technical report was created as part of the EPOS TCS GNSS which developed an architecture based on a federation of local (i.e. national) Geodetic Linkage Advance System Software (GLASS) nodes, where each node hosts the information of its infrastructure and is managed by the central European node (Data Gateway - DGW) which collects the information from the various nodes and allows access to the entire database from the web, providing also a Product Portal which distributes the products.

INGV contributes by providing data and metadata to the IRN node and developing software to manage the local node and in support of the European infrastructure.

At the same time INGV has developed a system that manages an Italian repository, called Mediterranean GNSS Archive (MGA), which regularly collects GNSS RINEX files from a large number of Continuously Operating Reference Station (CORS) located across a wide region of mainly the European and African plates. During the development of the MGA project, many ideas and tools were inspired by the EPOS system; the difference between the MGA and EPOS archives is that the latter only collects data provided by certified Agencies that have formally agreed to provide standardised data and metadata. The contributing Agencies are also responsible for updating station metadata and submitting it to the EPOS database. The guidelines followed in MGA lead instead to archive data provided by heterogeneous networks (commercial, governmental, scientific or application oriented), many of which do not provide station sitelog, nor adopt standardised formats, but regularly provide GNSS data [Bruni et al., 2020] [Bruni et al., 2022].

1. The EPOS-GNSS GLASS framework

The TCS EPOS-GNSS group [4] has developed a framework based on a distributed architecture to manage and disseminate the quality-controlled GNSS Data and Products from the infrastructures participating in Europe.

As shown in Figure 1, the EPOS-GNSS platform allows users to access the following portals to retrieve information:

- M³G portal [5]: it is managed by Royal Observatory of Belgium (ROB) and contains validated GNSS metadata, that means information about the characteristics of the GNSS stations (instrumentation, monumentation, approximate position, etc.) and the organisations that manage them;
- Data Gateway [6]: managed by Observatoire de la Cote d'Azur, France (CNRS-OCA), the portal contains GNSS data provided by individual research infrastructures, allows selection

of stations of interest, including on a map, to obtain corresponding metadata sets such as station code, site name, longitude, latitude, altitude, data availability, instrumentation, etc. Currently, EPOS-GNSS distributes daily RINEX data with a 30-second sampling frequency;

- Product Gateway [7]: managed by University Beira Interior, Portugal (UBI) the Product Portal provides derived GNSS products such as daily solutions, position time series, and velocities.

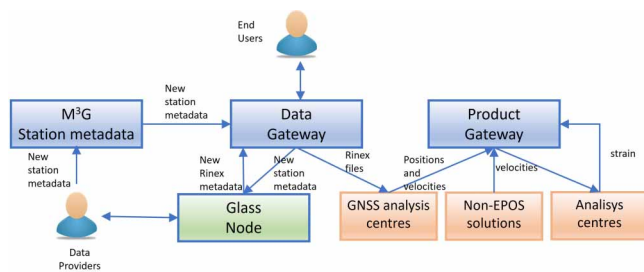


Figure 1 Data Flow within the EPOS-GNSS framework (modified from [Fernandes et al., 2022]).

The system is distributed on different European national research infrastructures, each one based on an independent EPOS-GNSS GLASS node (called “Glass Node”), which hosts a database with metadata and is connected to its Data Repositories. Each data node provides the GNSS RINEX data and metadata, the Quality Check validation referred to the station metadata, and synchronises with the Data Gateway (DGW) which also coordinates the information flow as shown in Figure 2.

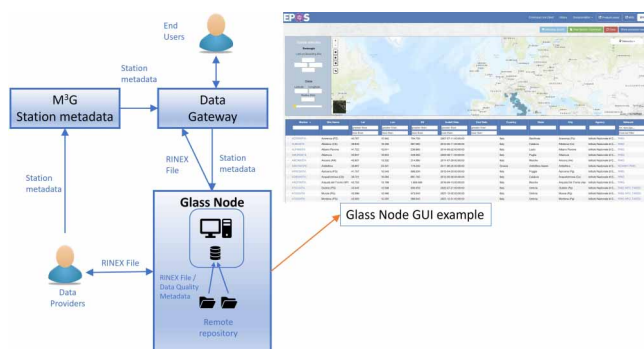


Figure 2 Interface between Glass Node and Data Gateway.

2. Software components of the EPOS-GNSS GLASS framework

To disseminate quality-controlled GNSS data and products, a distributed software architecture has been developed in support of the EPOS-GNSS framework (Figure 3), thanks to the joint effort of the following European groups: UBI (Portugal), CNRS-OCA (France), INGV (Italy), ROB (Belgium) with previous participation of IMO (Island). The architecture includes:

- A common database schema;
- Two web services to interact with the database: one for inserting data called Flask Web Service Server (FWSS - in Python) and one to query the metadata called EPOS GLASS Framework (in Java);
- A portal with a user-friendly graphical interface, to select and download data of interest;
- Quality control tools;
- Metadata submission and validation;
- Tools to ensure synchronisation and consistency of databases;
- Tools to ensure data and metadata redundancy and uniqueness.

The EPOS-GNSS GLASS follows a three-tier standard in which functional logic, data access, data storage, and user interfaces are located. The components are installed in a distributed environment where multiple nodes manage separate repositories (accessible remotely) and are coordinated by the Data Gateway.

The software package is specifically designed to collect and store various types of metadata:

- GNSS Data Center information, the Data Repository path and Node information;
- GNSS sitelog station metadata;
- GNSS RINEX File metadata;
- RINEX File quality information.

Station metadata can only be input through the Metadata Management and distribution system for Multiple GNSS Networks (M³G portal), which allows station managers to upload, validate, and distribute GNSS station metadata such as sitelog in IGS format, local network information, DOI, etc. The station metadata automatically flows into the Data Gateway database from where it is then synchronised to the nodes.

RINEX File metadata and data quality information are input into the local Glass Node database through the Python Flask Web Services Server (FWSS) software and transferred to the DGW through a synchronisation software (Epos Sync System).

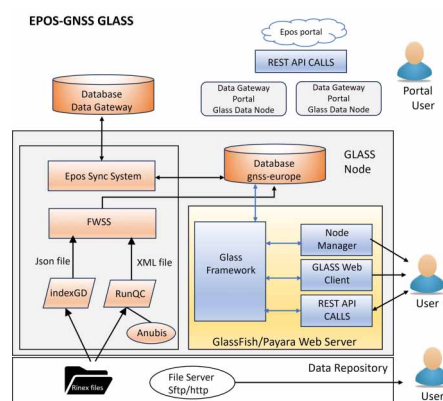
The extraction of type metadata from the RINEX file is carried out by a Python script *indexGD* which returns a JSON file that is passed to the FWSS for inserting the metadata into the Database. The Quality Check (QC) metadata is extracted by a Perl script *RunQC* that uses *G-Nut/Anubis* to analyse the quality of the RINEX file and return an XML file containing data quality information. The FWSS reads the XML provided by Anubis and loads into the database the metadata related to the Quality Check.

Anubis requires broadcast navigation files that are downloaded daily from the Geodetic Observatory Pecný (GOP) historical archive or the ROB backup server [8]. The orbit files contain detailed information about the orbital position of each satellite within the GNSS constellation. This information includes orbital parameters, the satellite's position at a given time, and other orbit-related data. Using precise information about the satellite's orbital position allows receivers to more accurately calculate the user's position on Earth by calculating the travel time of GNSS signals from the satellites to the receiver.

Quality control of the data is essential to determine which data can be made available to users. Data with critical inconsistencies between the RINEX file header and the information uploaded to M³G will not be available in the Data Gateway, but only in local nodes. Data with minor inconsistencies will be available on the Data Gateway with a specific warning status.

All metadata related to the GNSS universe are stored in a PostgreSQL database, which is free, high-performing, extensible, and scalable, and can be queried through the EPOS GLASS Framework web service.

Figure 3 The main components of the EPOS-GNSS GLASS system (modified from [Fernandes et al., 2022]).



3. The INGV contribution on the GLASS node

INGV participates in the EPOS project by developing a suite of software to manage data and metadata, starting with retrieving the RINEX file, the metadata storage into the database, the quality check and the dissemination of validated metadata to the DGW. All software was developed in Python language, version 3.10.

Within EPOS program, the IRN makes available the data from the RING network, composed of approximately 250 permanent GNSS stations that provide daily observations of the main GNSS constellations (GPS from 2000, and more recently from GLONASS, GALILEO, BEIDOU), amounting up to about 1 million of RINEX files. The data is collected from the INGV Irpinia acquisition center, and then stored and analysed in the IRN node. The structure shown in Figure 4 consists of:

- Synchronisation of data between RING data repository (“Bancadati”) and the IRN (*Bancadati_sync*, Paragraph 3.1);
- RINEX and Quality Check metadata extraction from RINEX files (*gnss_indexer*, Paragraph 3.2);
- IRN database population (*FWSS*, par. 3.3).

The RINEX files are stored in the Data Repository in a tree structure ordered by Year and DOY and are processed by the *gnss_indexer* software which lies on third-party software to extract the metadata and populates the database through the *FWSS* software.

The Glass Node, hosted on a different server, interfaces the PostgreSQL database which is populated by the *FWSS* software (managed by INGV) and stores the metadata.

The information in the database must be consistent to ensure the quality of the metadata and, to allow the matching between sitelogs and RINEX header metadata, the database is synchronised with the Data Gateway in both directions (as shown in Figure 2): RINEX File metadata is transferred from the local Node (IRN) to the DGW while the Station metadata is transferred from the DGW to the local Node.

To avoid data loss due to any software or hardware disasters, a bash script performs a daily backup of the IRN node database, retaining only the last two weeks, and sends a copy of the backup to an external server.

The INGV Data Repository represents one of the crucial and most important nodes from a scientific perspective because it actively contributes to the study of deformation in tectonic and volcanic areas, among the most active in Europe. Furthermore, it alone provides 15% of the files that can be downloaded from the EPOS Data Gateway.

Hereafter, we are going to describe each software.

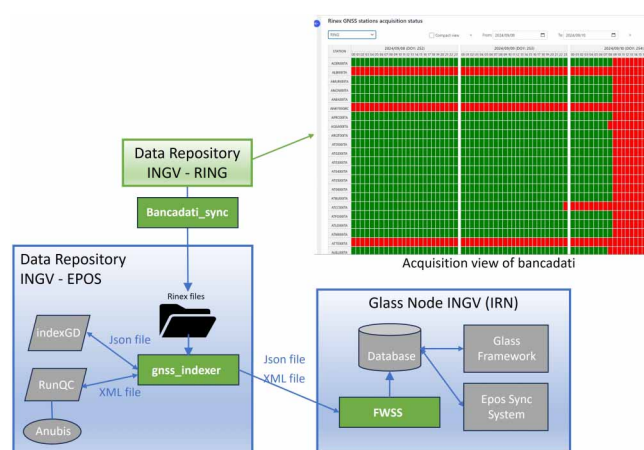


Figure 4 Architecture of the INGV Glass Node (in green the developed software).

3.1 Bash script to synchronise the RINEX files

The first step in processing RINEX files is to retrieve them from a source repository. The *bancadati_sync* is a linux bash script (Figure 5) based on the *rsync* command to synchronise files between the INGV Irpinia server and the IRN data repository, creating an incremental image copy of missing and/or newer files. The list of new files to be processed is generated by the *rsync* command through an output file, these files are then copied into a working folder (RING_STAGING) and indexed by the *gnss_indexer* software.

```
#!/bin/bash
my_date=`date +%Y-%m-%d-%H-%M-%S`
my_year=$(date +%Y)

#=====
# set rsync params
#=====
source_folder="bancadati****:/OUTGOING/RINEX30/RING/${my_year}/"

dest_folder="/mnt/data/GNSS-ARCHIVE/RING/30/${my_year}/"
export dest_folder

out_files="/home/ubuntu/sync_ingv/list/${my_date}_ftpl_files.list"
export out_files

#=====
# run rsync
#=====
rsync -av --update $source_folder $dest_folder >> $out_files

#=====
# copy all new files downloaded into STAGING folder with Python3
#=====
echo '
import os, shutil
import re

# Access environment variable
dest_folder= os.environ.get("dest_folder")
out_files= os.environ.get("out_files")

staging=dest_folder.replace("RING","RING_STAGING")

sub=re.compile("^d{3}(\w{4}d{4})\.d{2}d\.Z$")
with open(out_files, "r") as file:
    for line in file:
        if sub.match(line):
            src = f"{dest_folder}{line}".strip()
            dest=f"{staging}{line}".strip()
            os.makedirs(os.path.dirname(dest), exist_ok=True)
            shutil.copy(src, dest)

' | python3 -

#=====
# call gnss_indexer.py
#=====

cd /home/ubuntu/gnss_indexer
python3 gnss_indexer.py -m -r /mnt/data/GNSS-ARCHIVE/RING_STAGING/30
-b /mnt/data/GNSS-ARCHIVE/RING_TO_BE_INDEXED/30 -w /mnt/data/GNSS-ARCHIVE/RING_WRONG/30
-n RING -d DEBUG

echo "end ..."
```

Figure 5 The *bancadati_sync* code.

3.2 RINEX files management

Once the RINEX files are in the working folder (RING_STAGING), the *gnss_indexer* software automates the process by extracting and inserting its metadata into the database using third-party software. In particular, the *indexGD* software is called to extract information relating to station identification and provides the information in JSON format as shown in the following example:

```
{
  'station_marker': 'ACCE',
  'file_name': 'ACCE00ITA_R_20220880000_01D_30S_MO.crx.gz',
  'file_type': 'RINEX3',
  'sampling_window': '24h',
```

```

'sampling_frequency': '30s',
'relative_path': '2022/088',
'data_center': 'RING',
'reference_date': '2022-03-29 00:00:00',
'published_date': '2024-01-26 18:32:37',
'revision_date': '2022-04-08 11:38:21',
'creation_date': '2022-04-08 11:38:25',
'file_size': 1570446,
'md5_checksum': '57fc211ec60fb72468614f718903fdcd',
'md5_uncompressed': '92c5d28bdf23f6674e5dea02b3cb99ea'
}

```

The *gnss_indexer* sends the metadata through the call to an endpoint of the API web service exposed by the FWSS server (described in a later paragraph), which takes care of inserting it into the database. The API service returns a *status_code* which depends on the outcome of the processing (see APPENDIX A). If the insertion process is successful, the *gnss_indexer* proceeds with the Quality Check by processing and estimating the quality of the geodetic measurement. To process the Quality Check, the *gnss_indexer* calls the *RunQC* code to extract the information from the RINEX file relating to the satellite acquisition, and then calls the *G-Nut/Anubis* software [9] to analyse the information, giving back the quality parameters in XML format, as shown in Figure 6.

```

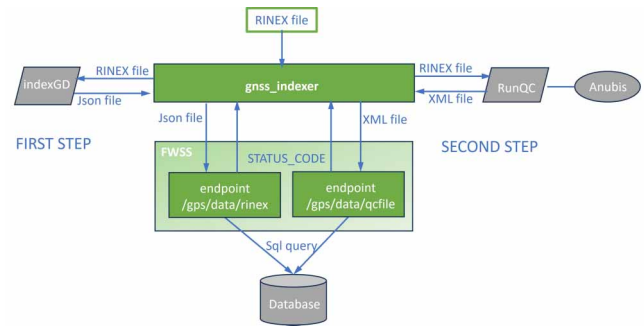
<?xml version="1.0"?>
<QC_GNSS gml:id="GRIS_20240126T000000_20240126T235930" xmlns="https://software.pecny.cz/anubis"
xmlns:gml="http://www.opengis.net/gml/3.2" xmlns:xsi="https://www.w3.org/2001/XMLSchema/instance"
schemaVersion="1.1">
  <meta>
    <created>2024-01-31T11:23:19</created>
    <program>G-Nut/Anubis Free [3.7]</program>
    <settings>
      <time_beg>2024-01-26T00:00:00</time_beg>
      <time_end>2024-01-27T00:00:00</time_end>
      <data_int>30.000</data_int>
      <elev_min>10.00</elev_min>
      <system type="GPS" />
      <system type="GL0" />
    </settings>
  </meta>
  <navi>
    <system type="GPS" nsat="31">
      <have>435</have>
    </system>
    <system type="GL0" nsat="24">
      <have>1152</have>
    </system>
  </navi>
  <data>
    <time_beg>2024-01-26T00:00:00</time_beg>
    <time_end>2024-01-26T23:59:30</time_end>
    <data_lag>4.4748380</data_lag>
    <data_int>30.00</data_int>
    <numb_epo>2880</numb_epo>
    <numb_gap>0</numb_gap>
    <total>
      <expt>56858</expt>
      <have>46507</have>
      <expt_usr>44155</expt_usr>
      <have_usr>42340</have_usr>
      <elev_min>4.95</elev_min>
      <cyc_slps>145</cyc_slps>
      <clk_jmps>0</clk_jmps>
    </total>
    <excluded>
      <xbeg>0</xbeg>
      <xend>0</xend>
      <xint>0</xint>
      <xsys>0</xsys>
    </excluded>
    <system type="GPS" nsat="30" xele="0">
      <epo>
        <expt>2880</expt>
        <have>2880</have>
        <dual>2880</dual>
      </epo>
      <amb>
        <nepo>0</nepo>
        <nsat>317</nsat>
        <nsig>24</nsig>
      </amb>
    </system>
  </data>
</QC_GNSS>

```

Figure 6 Example of XML file extracted from Anubis output.

Also these parameters are forwarded to the *FWSS* web server through another endpoint of the API web service, which takes care of inserting them into the database table and connecting them with the other information which identifies the same RINEX file. Since the Quality Check process requires time and computing resources, to optimise the operations this code is run only if required at the end of the first indexing phase, based on the returned *status_code* of the first step. Figure 7 illustrates in detail the data flow managed by the software.

Figure 7 Data flow managed by *gnss_indexer*.



The *gnss_indexer* software needs some mandatory parameters to run. Follows a table with a brief description of the usage of this python code and options to elaborate the RINEX files from the repository:

```

$ python3 gnss_indexer.py -m
-r /source/folder/GNSS-ARCHIVE/RING_STAGING/30
-b /folder/to/GNSS-ARCHIVE/RING_TO_BE_INDEXED/30
-w /folder/to/GNSS-ARCHIVE/RING_WRONG/30
-n RING
-d DEBUG

```

Parameter	Description
-m or --remove_files	Remove files from root_path after elaboration
-b or --to_be_indexed_path	Path where the RINEX files are copied when small recoverable issue occurs after the indexGD processing
-w or --wrong_path	Path where the RINEX files are copied when not recognized by indexGD (empty or unreadable)
-r or --root_path	Path to source files
-n --net	Specific network
-d --debug	Debug level [default: WARN]

Table 1 Parameter description for *gnss-indexer* software.

There are other non mandatory options which allow to manage the elaboration i.e. to select the *year* and *day* of the period to be processed, to select the *marker* of the stations or to indicate a file with the list of stations to process.

3.3 RINEX file Indexing and Quality Check

The Flask Web Service Server FWSS is a python software based on the Flask framework, it's a part of the GLASS software package used to insert GNSS station metadata, RINEX file metadata and quality indicators in the database. This web server provides REST-API services to interact with the database and it's distributed to all nodes in order to harmonise the installations and facilitate the management. In particular, the FWSS web server exposes the API service at the url endpoint */gps/data/rinex* to insert into the database the metadata relating to the RINEX file. The request is made by calling the URL with the POST method, through which the metadata is provided in JSON format. After an initial check to verify that the metadata is reliable, the software verifies the presence of the same RINEX file within the database through the following parameters:

- `md5_checksum`: is the unique identifier of the RINEX file which includes data and metadata;
- `file_name`: include the marker, the country code and the measurement's date;
- `file_type`: default values are RINEX2, RINEX3, RINEX4;
- `data_center`: is the organisation which manage the acquisition;
- `station_id`: identifier of the GNSS station into the database;
- `reference_date`: is the measurement's date;
- `creation_date`: is the first date the record was inserted into the database;
- `revision_date`: is the date of insert/update of the RINEX file (same or later than the `creation_date`).

A first query checks the presence of the RINEX file inside the database through the unique identifier *md5_checksum* and alternatively through its equivalent parameters (*file_name*, *file_type_id* and *data_center_id*). If there is no matching record in the database, the API service proceeds with inserting the new RINEX file otherwise, if the RINEX file already exists, based on the comparison between the JSON and database parameters, one of the following actions may be required:

- UPDATE: when the RINEX in the database has different *file_name* or *reference_date* or *creation_date* compared with the parameters obtained from JSON (no Quality Check is required);
- RESET / UPDATE: when the RINEX in the database has a different *station_id* or *file_type_id* or *reference_date* compared with the JSON, then a new inserting is requested and also the Quality Check elaboration.

The procedure is shown in the following Flow Chart in Figure 8.

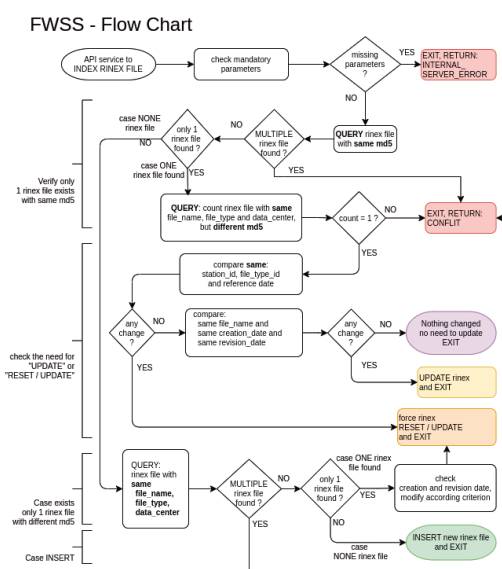


Figure 8 RINEX file indexing procedure.

The sequence for inserting a RINEX file depends on several factors, the table in APPENDIX A summarises the possible conditions that may be encountered and the actions to be taken. At the end of this first step, the FWSS web server returns a *status_code* to the calling software *gnss_indexer*, which decides whether to proceed with the evaluation of the quality check (i.e. the quality of the geodetic data).

Another API service exposed by the FWSS web server is at the url endpoint */gps/data/qcfile* used to insert information about the Quality Check analysis. This service is called with the POST method to provide the parameters in the XML format (after having been processed by *RunQC/Anubis*), relating to the constellations acquired during the GNSS measurement. Based on the Quality Check outcome, a *status flag* on the database is set, which can take on positive values from 0 to 3 if the RINEX file is accepted, has passed the QC tests and is available to be disseminated, otherwise the flag takes on negative values (*status* = -3) to indicate that the RINEX metadata has not passed the quality checks, in which case the RINEX file data is not made available.

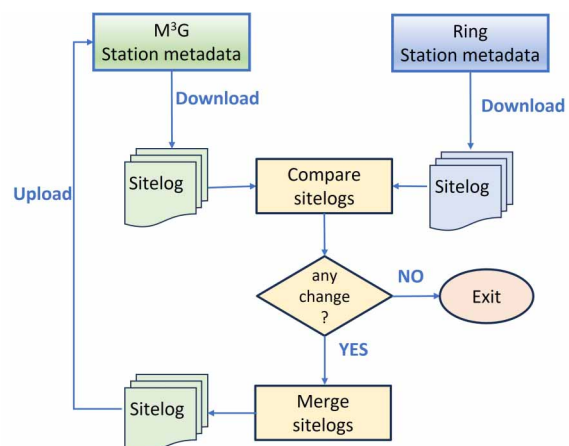
4. The INGV contribution in GNSS station metadata synchronisation

The Station metadata (or *sitelog*) is a relevant information to operate the Quality Check, which is the process to validate the RINEX data and metadata before the dissemination. It provides detailed information about the GNSS station, which is basically composed by a GPS antenna and a receiver to acquire satellites, but this equipment may change during the acquisition epochs and, to reach high measurement accuracy, it's important the coherence between what is declared on the RINEX header (day by day) and what is declared on the *sitelog* file which resumes the history of all the equipments installed on the GNSS station. Hence the need to have the *sitelog* information available on the M³G portal, in order to allow the matching. During the Quality Check the metadata which shows critical inconsistencies between the RINEX file header and the related information in the *sitelog* files will not be made available in the Data Gateway, but only in the local node.

To harmonise the *sitelog* upload from each Node of the EPOS project, the M³G portal provides API services [10] which also allow the GNSS station manager to upload the *sitelog* files.

Because any *sitelog* may be updated more times by the operators at the acquisition centre of the GNSS network, to avoid discrepancies with the M³G portal, the Python software named *ring_synchronization* has been developed: on a daily base this software downloads the *sitelogs* for all the stations of the Italian RING Network, compares them with the corresponding information from the M³G portal and, in case of differences, it generates a new *sitelog* to upload to the M³G portal through the exposed API service (Figure 9), keeping synchronised the information and improving the Quality Check.

Figure 9 Update station metadata system.



The operation of creating sitelogs deserves an explanation: in the acquisition center of the RING Network these files are generated by a software through which operators can update the metadata of the individual GNSS stations. The *ring_synchronization* software integrates this system and interacts with the web interface to generate updated sitelogs on demand, by logging in the “Bancadati” portal and clicking a sequence of buttons which allow to download the sitelogs. The developed software automates this process using the *Selenium framework* which allows to simulate user interactions with the web application by using the *Gecko web driver*, which is specifically designed to run on Firefox. In this way the sitelogs are always updated with the latest changes made by the Ring Network operators.

5. Conclusions

The codes developed for the EPOS project contribute to distributing data and metadata of geodetic measurements with high quality standards in compliance with the requirements dictated at European level, and constitute a reliable basis for scientific research. The software of the EPOS project is freely downloadable from the gitlab platform [11].

The Quality Check made it possible to highlight numerous inconsistencies between the metadata in the sitelogs and the related RINEX files, generated by the normal activity of the operators during the updating of the parameters on the geodetic stations. Such updates must always be propagated to all levels to keep the geodetic measurements reliable.

Many problems have been overcome thanks to the software developed which allows timely updating of information on the reference servers, and has allowed us to maximise the number of files distributed, limiting the case of critical errors.

The system is being updated, currently only RINEX data with sampling frequency of 30 seconds are processed. In the future it is expected to include high frequency (1Hz) RINEX data in the same database, this involves modifications to the database which is now well tested and lends itself to being scaled while maintaining the same structure.

Bibliography

- Bruni S., Randazzo D., Cavaliere A., Avallone A., Serpelloni E., Devoti R., D’Agostino N., Pietrantonio G., Riguzzi F., Cecere G., D’Ambrosio C., Falco L., Martelli L., (2020). *WP3 Sviluppo di procedure automatiche per l’integrazione nelle sale operative di misure, modelli predittivi e prodotti di ricerca dell’INGV - Attività sui dati, metadati e prodotti per le stazioni GNSS in area Mediterranea*. Misc. INGV, 57, pp. 122–126. <https://doi.org/10.13127/misc/57/22>
- Bruni S., Randazzo D., Devoti R., Serpelloni E., Cavaliere A., Martelli L., Pietrantonio G., D’Agostino N., (2022). *The INGV Mediterranean GNSS Archive (MGA)*. Rapp. Tec. INGV, 456, 1-122. <https://doi.org/10.13127/rpt/456>
- Fernandes R., Bruyninx C., Crocker P., Menut J.-L., Socquet A., Vergnolle M., Avallone A., Bos M., Bruni S., Cardoso R., et al., (2022). *A new European service to share GNSS Data and Products*. Annals of Geophysics, 65, (3), p. DM317. doi: 10.4401/ag-8776. <https://dx.doi.org/10.4401/ag-8776>

Sitography

- [1] <http://glass.ingv.it:8080/glasswebui/#/site>
 [2] <http://webring.gm.ingv.it>
 [3] www.epos-eu.org

- [4] <https://gnss-epos.eu/>
- [5] <http://gnss-metadata.eu>
- [6] <http://gnssdata-epos.oca.eu>
- [7] <http://gnssproducts.epos.ubi.pt>
- [8] <https://doi.org/10.24414/c4ba-kf16>
- [9] <https://gnutsoftware.com/software/anubis/>
- [10] <https://gnss-metadata.eu/site/api-docs>
- [11] <https://gitlab.com/gpseurope>

APPENDIX A

Table of possible events, action and *status_codes* when a RINEX file is indexed.

Status code and Meaning (Standard)	EVENT	Status	Action
201 CREATED	RINEX file does not exist on the database and is inserted	RINEX file has been indexed with id ...	status set to "0" execute => RunQC
200 OK	events which require reset/update condition: rinex found with same: <i>file_name</i> , <i>file_type</i> and <i>data_center</i> , but	<i>station_id</i> is different on <i>rinex_file</i> table <i>reference_date</i> is different on <i>rinex_file</i> table	status set to "0" execute => RunQC
202 ACCEPTED	RINEX file found with same MD5, status is not changed,	<i>new file_name</i> from body <i>new data_center</i> from body <i>new relative_path</i> from body	update => RunQC is not necessary
304 NOT_MODIFIED	RINEX file found same MD5, no update	RINEX file was found by id: ... Nothing seems changed. File has not been updated!	status is not changed => RunQC is not necessary
409 CONFLICT	conflict occurs	Found multiple records with the same marker: ... <i>md5_checksum</i> ... result duplicated inside the database <i>md5_checksum</i> ... <i>file_name</i> + <i>file_type</i> + <i>data_center</i> must be unique. You are trying to set it to <i>file_name</i> - <i>sampling_window</i> - <i>sampling_frequency</i> , but these values already exist for file with another <i>md5_checksum</i> <i>md5_checksum</i> ... <i>DataCenterStructure</i> multiple entry More than one RINEX file with name: ... + type: ... + <i>sampling_window</i> : ... and data center: ... exist on DB ...	File not inserted Action necessary from DGW

500 INTERNAL_ SERVER_ ERROR	internal error occurs	Station with marker ... was not found. Station must be registered in the database before files can be indexed.	File not inserted Action necessary from DGW
		Data center with acronym: ... not in database. Please contact the DGW, the <i>data_center</i> information should be inserted by the DGW.	
		<i>file_type</i> ... not in database. Please contact the DGW, the <i>file_type</i> information should be inserted by the DGW.	
		Missing body mandatory keys: [list of missing parameters]	

QUADERNI di GEOFISICA

ISSN 1590-2595

<https://istituto.ingv.it/it/le-collane-editoriali-ingv/quaderni-di-geofisica.html/>

I QUADERNI DI GEOFISICA (QUAD. GEOFIS.) accolgono lavori, sia in italiano che in inglese, che diano particolare risalto alla pubblicazione di dati, misure, osservazioni e loro elaborazioni anche preliminari che necessitano di rapida diffusione nella comunità scientifica nazionale ed internazionale. Per questo scopo la pubblicazione on-line è particolarmente utile e fornisce accesso immediato a tutti i possibili utenti. Un Editorial Board multidisciplinare ed un accurato processo di peer-review garantiscono i requisiti di qualità per la pubblicazione dei contributi. I QUADERNI DI GEOFISICA sono presenti in "Emerging Sources Citation Index" di Clarivate Analytics, e in "Open Access Journals" di Scopus.

QUADERNI DI GEOFISICA (QUAD. GEOFIS.) welcome contributions, in Italian and/or in English, with special emphasis on preliminary elaborations of data, measures, and observations that need rapid and widespread diffusion in the scientific community. The on-line publication is particularly useful for this purpose, and a multidisciplinary Editorial Board with an accurate peer-review process provides the quality standard for the publication of the manuscripts. QUADERNI DI GEOFISICA are present in "Emerging Sources Citation Index" of Clarivate Analytics, and in "Open Access Journals" of Scopus.

RAPPORTI TECNICI INGV

ISSN 2039-7941

<https://istituto.ingv.it/it/le-collane-editoriali-ingv/rapporti-tecnici-ingv.html/>

I RAPPORTI TECNICI INGV (RAPP. TEC. INGV) pubblicano contributi, sia in italiano che in inglese, di tipo tecnologico come manuali, software, applicazioni ed innovazioni di strumentazioni, tecniche di raccolta dati di rilevante interesse tecnico-scientifico. I RAPPORTI TECNICI INGV sono pubblicati esclusivamente on-line per garantire agli autori rapidità di diffusione e agli utenti accesso immediato ai dati pubblicati. Un Editorial Board multidisciplinare ed un accurato processo di peer-review garantiscono i requisiti di qualità per la pubblicazione dei contributi.

RAPPORTI TECNICI INGV (RAPP. TEC. INGV) publish technological contributions (in Italian and/or in English) such as manuals, software, applications and implementations of instruments, and techniques of data collection. RAPPORTI TECNICI INGV are published online to guarantee celerity of diffusion and a prompt access to published data. A multidisciplinary Editorial Board and an accurate peer-review process provide the quality standard for the publication of the contributions.

MISCELLANEA INGV

ISSN 2039-6651

https://istituto.ingv.it/it/le-collane-editoriali-ingv/miscellanea-ingv.html

MISCELLANEA INGV (MISC. INGV) favorisce la pubblicazione di contributi scientifici riguardanti le attività svolte dall'INGV. In particolare, MISCELLANEA INGV raccoglie reports di progetti scientifici, proceedings di convegni, manuali, monografie di rilevante interesse, raccolte di articoli, ecc. La pubblicazione è esclusivamente on-line, completamente gratuita e garantisce tempi rapidi e grande diffusione sul web. L'Editorial Board INGV, grazie al suo carattere multidisciplinare, assicura i requisiti di qualità per la pubblicazione dei contributi sottomessi.

MISCELLANEA INGV (MISC. INGV) favours the publication of scientific contributions regarding the main activities carried out at INGV. In particular, MISCELLANEA INGV gathers reports of scientific projects, proceedings of meetings, manuals, relevant monographs, collections of articles etc. The journal is published online to guarantee celerity of diffusion on the internet. A multidisciplinary Editorial Board and an accurate peer-review process provide the quality standard for the publication of the contributions.

Coordinamento editoriale

Francesca DI STEFANO
Istituto Nazionale di Geofisica e Vulcanologia

Progetto grafico

Barbara ANGIONI
Istituto Nazionale di Geofisica e Vulcanologia

Impaginazione

Barbara ANGIONI
Patrizia PANTANI
Massimiliano CASCONI
Istituto Nazionale di Geofisica e Vulcanologia

©2024

Istituto Nazionale di Geofisica e Vulcanologia
Via di Vigna Murata, 605
00143 Roma
tel. +39 06518601

www.ingv.it



Creative Commons Attribution 4.0 International (CC BY 4.0)



ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

