Copernicus Climate Change Service

P1: 012

The GNSS Repository for the Global RINEX Data Files

Kalev Rannat (1), Hannes Keernik (1, 2), Fabio Madonna (3, 4)
(1)Tallinn University of Technology, Department of Software Science
(2) University of Tartu, Institute of Physics
(3) University of Salerno, Department of Physics
(4) National Research Council of Italy, CNR-IMAA

The concept and main objectives

The efforts spent within the C3S 311a Lot3 (“Access to observations from baseline and reference networks”) and C3S 311c Lot2 (“Historic In Situ Upper Air Database”) contracts of Cop1 have been widen in Cop2 to facilitate the access to and the homogenization of comprehensive upper-air, baseline and reference observations for a subset of GCOS relevant ECVs.

An action of the 2022 GCOS Implementation Plan (GCOS-244) asks for the development of a global repository of fundamental climate data records for the observations collected by GNSS receivers. In the frame of an activity of the Copernicus Climate Change Service named “C3S2 311 Lot2”, aiming at harmonizing the access to data from Baseline, Reference and Comprehensive Upper-Air datasets (Madonna et al., 2022; Rannat et al., 2023), such a global GNSS REPository (GREP) has been developed and is currently under testing for the first release planned for August 2024.

The initial idea of GREP is to create a unified collection of GNSS hourly observational data in RINEX with a number of stations enabling an average horizontal coverage of at least 100 km over land. “Global” in the acronym GREP means that the repository is not limited to a specific GNSS networks, country, geographical areas or continents. The GREP will be like one huge data lake where find data by latitudes/longitudes, in an area of interest, for the GNSS data reprocessing. GREP QC is performed automatically with the same settings for all RINEX-logfile pairs from any data source. It helps to out-filter practically unusable data and documents the mismatches or missing metadata from the logfiles.

GREP – what it is (what it does)

GREP aims to serve the climate research community worldwide, being a data lake for GNSS data re-processing for climate (not for International Terrestrial Reference Frame (ITRF) or operative meteorology (e.g., EUMETNET’s E-GVAP)).

GREP simplifies further GNSS reprocessing by offering a unified (harmonized) data lake. GREP performs preliminary QC and documents the inconsistencies between the observational and metadata.

GREP data harmonization and QC are based on data available for an epoch where the observations were made. All RINEX/logfile pairs are documented by their names and can be easily tracked for actuality in future. Data harmonization in GREP means unifying the data compression methods, giving 9-character site filenames and long RINEX3 filenames for RINEX.

.... and what it is not

GREP is not an alternative GNSS data sharing point for national or network DCs.

GREP data is (initially) not open-access (the access can be obtained from ECMWF). However, it contains only open-access data from a number of networks and agencies worldwide. No policy for private or restricted data yet.

GREP is not solely oriented on data from well-known sites with IGS-standard installations. The data for GREP, collected worldwide (not talking about single networks or DSs), is quite heterogeneous and uncomfortable to use “as is”. However, the data, as fundamental observations, are valuable and should be kept in unified way for further usage.

Data providers

GREP data providers are the GNSS networks/agencies sharing their data. Initially, due to the open-Access data policy, GREP started with data from the following providers: **Geoscience Australia, NOAA CORS Network, Canadian Active Control System, EUREF Permanent GNSS Network and IGS**.

Australian Government Geoscience Australia

Government of Canada

NOAA National Oceanic and Atmospheric Administration U.S. Department of Commerce

EUREF

IGS International GNSS Service

For smooth and dense coverage of the globe, it will be the highest priority to initiate connections/collaborations with local and national agencies (otherwise, we cannot cover the white areas in Africa and Asia we can notice today). GREP data providers are (and will be) referred as supporters for world climate monitoring and research actions. The data providers (having data in GREP) will be cited as valuable supporters of the reprocessing efforts. With citing and acknowledgements , the DCs (and networks) gather a reputation as supporters of worldwide climate research.

The main principles of data and metadata collection

The data is collected worldwide from open-access portals. The number of sites grows in the course of years (for example, the sites from Geoscience Australia and NCN Fig. 2). One of the data collection principles in GREP is avoiding site duplicates (i.e., not downloading RINEXes for the sites in a data portal/network, including sites form others). The process of data harmonization and QC will not change any initial data or metadata.
GREP does not choose sites by hardware or installation criteria (however, collecting only data from 2-frequency geodetic grade receivers), not initially distinguishing sites by quality.

GREP keeps track of all sites in the repository by site names and approximate coordinates (Table 1). Recording the coordinates helps to search/collect sites by geographical longitude/latitude from the repository. This option makes GREP global; the user does not need to search for data from different networks.

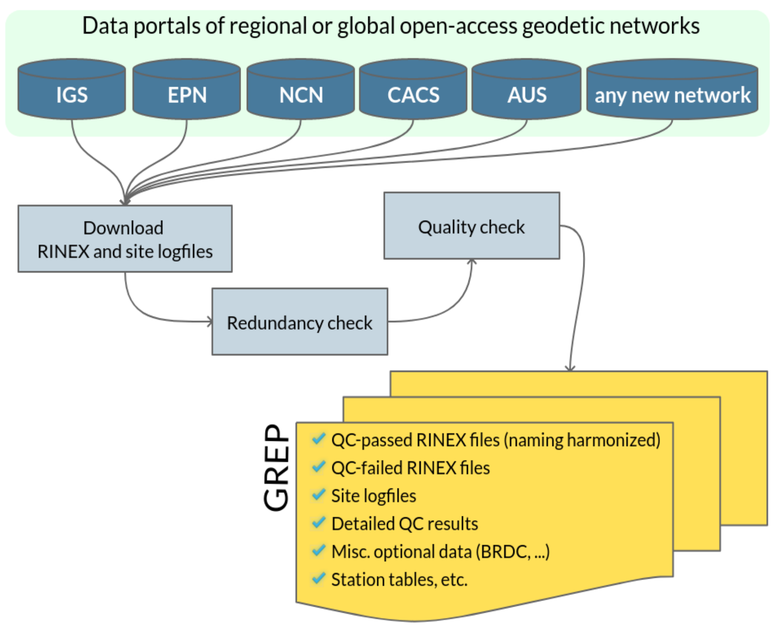

Figure 1. Data flow in GREP.

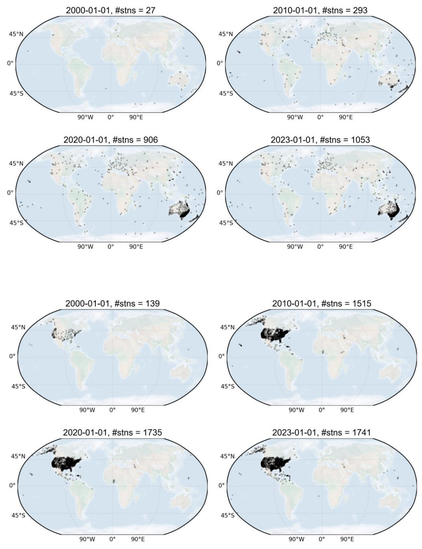

Figure 2. GNSS sites in Geoscience Australia and NCN data portal in 2000-2023.

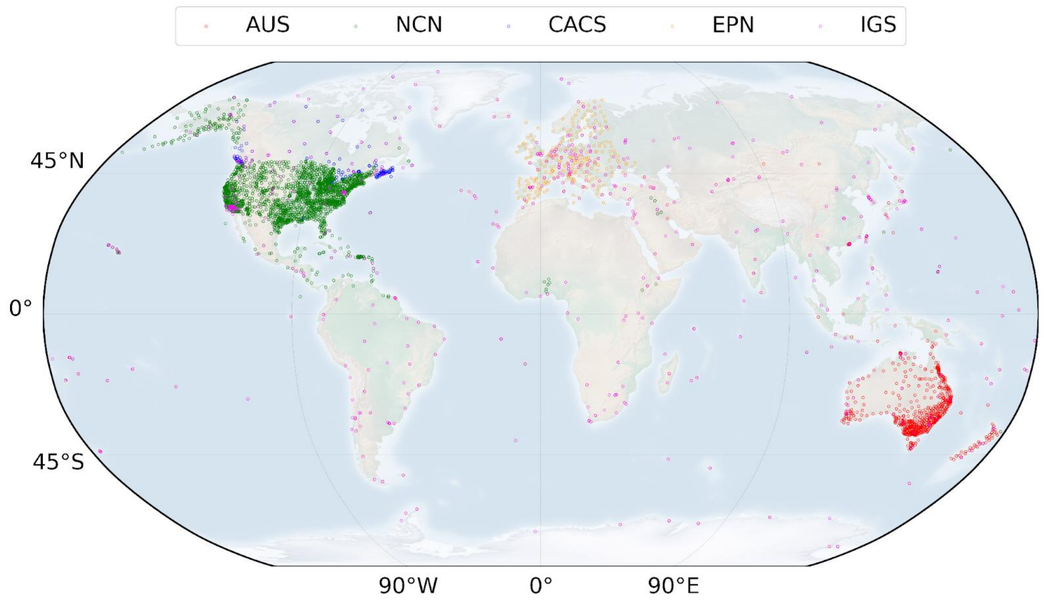

Figure 3. Global coverage of sites from different data sources - AUS (Geoscience Australia), NCN (NOAA CORS Network), CACS (Canadian Active Control System), EPN (EUREF Permanent GNSS Network) and IGS (International GNSS Service).

GREP data structure, QC, tables:

```
GREP
├── brdc
├── df_backups
├── logfiles
├── logging
├── pickles
├── RINEX
│   └── YYYY
│       └── DOY
├── RINEX_err
│   └── YYYY
│       └── DOY
└── stns
```

brdc – daily broadcasted ephemeris files

df_backups – dataframe backups
logfiles – site’s logfiles
logging – operative GREP daily logfiles
RINEX – RINEX archive by Year (YYYY) and Day of the Year (DOY)
RINEX_err - RINEX archive by Year (YYYY) and Day of the Year (DOY) for erroneous files (e.g., missing XYZ)
stns – stations table

Figure 4. GREP directory tree structure.

Table 1: An excerpt of the list of sites represented in GREP (saved as stn_table.csv and updated during each network-based data process).

site_name	country_code	lat	lon	elev	x	y	z
1LSU	USA	30.407423887	268.819740343	-5.9263	-113402.6875	-5504361.9251	3209404.4422
1NSU	USA	31.750799597	266.902397459	28.7726	-293349.4472	-5420742.9009	3336981.0869
1ULM	USA	32.529033486	267.924094479	16.2283	-194982.0023	-5379222.1684	3410046.9185
5RMK	AUS	-34.175435516	140.733920331	27.5358	-4089693.8623	3343334.3936	-3562578.421
AB02	USA	52.9706021169	191.145326542	193.3064	-3776808.7894	-744083.7365	5068728.2767
AB07	USA	55.349274891	199.523227388	89.7442	-3425750.5967	-1214685.0687	5223662.8645
AB08	USA	60.3848284	193.79914382	27.532	-3068712.8914	-753699.5299	5521814.0355
AB09	USA	65.614978427	191.937874079	162.8584	-2583615.2747	-546237.0636	5786502.4278
AB13	USA	56.307323472	201.496202438	488.0993	-3299992.0035	-1299648.7598	5283907.1548
AB14	USA	59.1081628	200.908468157	658.2124	-3066910.0363	-1171658.4616	5450696.8755
AB17	USA	63.886356932	199.30526863	118.1769	-2656728.4685	-930647.3462	5704256.3083
AB21	USA	51.864145991	183.337350848	60.1465	-3940204.4601	-229767.9973	4993530.2385
AB27	USA	67.055884674	203.095107581	508.8682	-2293838.0383	-978173.3421	5851314.3402
AB28	USA	62.093811716	207.185081196	1583.2086	-2663005.6277	-1367722.5546	5614815.0262
AB33	USA	67.251010852	209.827454178	335.3628	-2145891.4888	-1230329.5106	5859604.0517
AB35	USA	60.079068101	217.610119637	537.2333	-2526854.5901	-1946651.0663	5505342.0403
AB36	USA	65.030396865	209.25601448	817.0802	-2355803.0029	-1319637.9453	5759882.0031
AB37	USA	62.967324013	214.548116404	1136.9918	-2394433.9192	-1648611.1922	5659336.4877
AB39	USA	66.559344668	214.787366749	148.3555	-2089696.9784	-1451694.0079	5829174.6627
AB43	USA	58.198841176	223.359189827	27.2562	-2449678.8081	-2312243.1412	5397464.296
AB45	USA	68.760498376	211.128827644	507.0689	-1983815.5605	-1198075.8173	5922819.3527
AB48	USA	56.245061136	225.352998692	5.7524	-2496256.3718	-2527206.2635	5279656.7888
AB49	USA	55.580101131	226.931512007	17.4796	-2467568.0664	-2639809.3662	5238172.2091
AB50	USA	58.416774828	225.45469675	51.8871	-2348976.5636	-2386558.3455	5410237.8763
AB51	USA	56.797623668	227.086434468	76.2236	-2383749.7818	-2564002.6155	5313656.4117
ABER	GBR	57.144011321	357.91978202	54.685	3466272.6191	-125904.0828	5334663.6473
ABMF	GLP	16.262307258	298.472464363	-24.8342	2919786.1238	-5383745.6145	1774605.1264
ABPO	MDG	-19.018302364	47.22921262	1554.0592	4097217.3834	4429119.9323	-2065771.3323
ABQ1	USA	34.957347483	253.505518269	1719.3088	-1486208.7295	-5019131.6638	3634974.8782
ABQ2	USA	34.957106549	253.505507571	1720.0126	-1486214.1798	-5019146.628	3634953.3688
AC07	USA	65.961291717	198.713382708	176.8288	-2467774.3562	-835938.0287	5802352.141
AC08	USA	58.928772728	206.355297512	644.7796	-2957192.8489	-1465088.6346	5440397.768
AC12	USA	54.830957816	200.41043455	84.9618	-3450877.299	-1284085.1974	5190636.6278
AC13	USA	55.821897781	204.377594439	222.4771	-3271200.0783	-1482340.0104	5253511.9965
AC17	USA	60.663898993	207.596153114	883.1206	-2776807.4797	-1451442.4756	5537860.0904
AC19	USA	62.519211332	206.39266542	547.1188	-2643624.5169	-1311884.8257	5635935.2622
AC23	USA	60.475091439	209.122043021	81.424	-2752836.1727	-1533596.4174	5526824.0272
AC26	USA	58.214552789	205.849706081	190.6156	-3030887.8909	-1468433.6912	5398525.1478
AC27	USA	59.252504535	205.837116185	417.9353	-2942366.686	-1424747.8675	5458729.3726
AC31	USA	64.637973221	197.76087437	228.8335	-2609009.1566	-835696.3599	5740743.3122
AC37	USA	60.439684992	206.134616151	1627.2727	-2832787.1123	-1389890.9273	5526223.4831
AC38	USA	57.753685783	206.658123296	44.4792	-3048695.7027	-1530543.3273	5371187.8288
AC39	USA	58.609717153	207.605923657	157.9228	-2951219.6484	-1542248.875	5421553.365
AC40	USA	56.930350821	201.381426303	40.2511	-3248413.624	-1271824.7519	5321705.5472
AC42	USA	54.471773957	197.216346793	258.3845	-3548311.3522	-1099495.301	5167645.1477
AC45	USA	56.564454507	205.819033222	451.5596	-3171145.8327	-1534291.6133	5299707.2137
AC47	USA	60.081448586	207.376053326	922.8883	-2832480.6066	-1466716.8924	5505808.5299

```
|                    SITE DAILY RINEX AND LOGFILES CONSISTENCY CHECK                    |
=========================================================================================

   1. Comparing approximate coordinates XYZ with tolerance +/-50 m
   2. Comparing Receiver TYPE, SERIAL NUMBER and FIRMWARE VERSION
   3. Comparing Antenna TYPE, SERIAL NUMBER, RADOME
        and ARP (up, north, east). tolerance +/- 0.5 mm

   ERROR CODES:

   0  - consistent (no errors)

   21 – mismatch of X
   22 – mismatch of Y
   23 – mismatch of Z
   201 - messy XYZ values in RINEX header

   30 – messy section 3 (GNSS Receiver Information). Usually caused by faulty
     Date Removed : CCYY-MM-DDThh:mmZ written into subsections 3.x
   31 – mismatch of receiver type
   32 – mismatch of receiver serial number
   33 – mismatch of receiver’s firmware version (checking will be discontinued)

   301 - incomplete receiver specification in RINEX header
   302 - incomplete receiver specification in logfile

   40 – messy section 4 (GNSS Antenna Information). Usually caused by faulty
     Date Removed : CCYY-MM-DDThh:mmZ written into subsections 4.x
   41 – mismatch of antenna type
   42 – mismatch of antenna’s serial number
   43 – mismatch of antenna’s radome
   44 – mismatch of ARP up
   45 – mismatch of ARP north
   46 – mismatch of ARP east

   47 – non-numeric values (or emty fields) in site RINEX header ANTENNA: DELTA H/E/N fields
   48 – non-numeric values (or empty fields) in logfile section 4 (Antenna Marker → ARP) fields

   401 - incomplete antenna specification in RINEX header
   402 - incomplete antenna specification in logfile

   WARNING: The error codes just indicate the inconsistencies, the data user must check and decide
   if, what and how to correct!

Data in columns:
=========================================================================================
| daily observational RINEX | site log-file | comma-separated list of   | comma-separated list of   |
|                           |               | ACs and reanalysis where  | error codes from comparing|
|                           |               | the site has been used    | rinex and log-file        |
=========================================================================================

01NA00AUS_R_20183340000_01D_30S_MO.crx|01NA_2024-04-25_log.xml|| 0
1LSU00USA_R_20183340000_01D_30S_MO.crx|1lsu.log.txt|| 0
1NSU00USA_R_20183340000_01D_30S_MO.crx|1nsu.log.txt|| 0
1ULM00USA_R_20183340000_01D_30S_MO.crx|1ulm.log.txt|| 0
30NA00AUS_R_20183340000_01D_30S_MO.crx|30NA_2024-04-25_log.xml|| 0
31NA00AUS_R_20183340000_01D_30S_MO.crx|31NA_2024-04-25_log.xml|| 0
AB0200USA_R_20183340000_01D_30S_MO.crx|ab02.log.txt|TUG IGS-repro3| 0
AB0400USA_R_20183340000_01D_30S_MO.crx|ab04.log.txt|| 0
AB0600USA_R_20183340000_01D_30S_MO.crx|ab06.log.txt|| 0
AB0700USA_R_20183340000_01D_30S_MO.crx|ab07.log.txt|TUG IGS-repro3| 0
AB0800USA_R_20183340000_01D_30S_MO.crx|ab08.log.txt|| 0
AB0900USA_R_20183340000_01D_30S_MO.crx|ab09.log.txt|TUG IGS-repro3| 0
AB1100USA_R_20183340000_01D_30S_MO.crx|ab11.log.txt|TUG IGS-repro3| 0
AB1200USA_R_20183340000_01D_30S_MO.crx|ab12.log.txt|| 0
AB1300USA_R_20183340000_01D_30S_MO.crx|ab13.log.txt|| 0
AB1400USA_R_20183340000_01D_30S_MO.crx|ab14.log.txt||48
AB1500USA_R_20183340000_01D_30S_MO.crx|ab15.log.txt|TUG IGS-repro3| 0
AB1700USA_R_20183340000_01D_30S_MO.crx|ab17.log.txt||48
```

Figure 5. An example of a daily QC output as an ASCII file (comparing_*.txt), saved at the end of the daily processing cycle after filtering the data frame based on the day of the year.

References:

1. Global Climate Observing System (GCOS). (2022). GCOS Implementation Plan 2022.
2. Madonna, F., Tramutola, E., SY, S., Serva, F., Proto, M., Rosoldi, M., et al. (2022). The new Radiosounding HARMonization (RHARM) data set of homogenized radiosounding temperature, humidity, and wind profiles with uncertainties. Journal of Geophysical Research: Atmospheres, 127, e2021JD035220. https://doi.org/10.1029/2021JD035220.
3. Rannat, K.; Keernik, H.; Madonna, F. The Novel Copernicus Global Dataset of Atmospheric Total Water Vapour Content with Related Uncertainties from GNSS Observations. Remote Sens. 2023, 15, 5150. https://doi.org/10.3390/rs15215150.

Acknowledgements

We acknowledge NASA’s Crustal Dynamics Data Information System (CDDIS), The Royal Observatory of Belgium, NOAA/National Geodetic Survey, Natural Resources Canada and Canadian Active Control System (CACS) and The GNSS Data Centre of Australian Geosciences for open-access GNSS data.

This work was done on behalf of the European Union’s Copernicus Climate Change Service implemented by ECMWF.

European Commission

Copernicus Europe’s eyes on Earth

IMPLEMENTED BY ECMWF