



**GCOS
Reference
Upper-
Air
Network**

GRUAN Technical Note 10

**The GRUAN Policy for the handling,
sharing, and distribution of GRUAN
measurement data**

Ruud Dirksen and Michael Sommer

Publisher

GRUAN Lead Centre

Number & Version

GRUAN-TN-10
1.0 (2026-02-27)

Document info

<i>Title:</i>	The GRUAN Policy for the handling, sharing, and distribution of GRUAN measurement data
<i>Topic:</i>	Documentation
<i>Authors:</i>	Ruud Dirksen and Michael Sommer
<i>Publisher:</i>	GRUAN Lead Centre, DWD
<i>Document type:</i>	Technical Note
<i>Document number:</i>	GRUAN-TN-10
<i>Page count:</i>	15
<i>Version:</i>	1.0 (2026-02-27)

Abstract

This Technical Note outlines the policy for the handling, sharing and distribution of the data that are submitted by the sites to the Lead Centre (LC).

A data-level and user classification scheme is defined, where access privileges to particular data levels are defined for various user categories. In this scheme data streams for which a certified GRUAN Data Product (GDP) exists, are provided to the global user-community with unrestricted access. Raw measurement data are only accessible to the Lead Centre, the site that provided the data, the processing centre, and the data product developers. Intermediate data, pre-release GDP versions, and data from accompanying instruments are restricted to GRUAN-internal and special users.

For measurement activities outside a site's GRUAN commitment, e.g. campaign measurements, data retention for up to 12 months can be requested by site managers to ensure first-publication rights. This data distribution strategy is consistent with the WMO unified data policy (resolution 1 of Cg-Ex. 2021).

GRUAN-processing of data from non-GRUAN sites is available upon request as a service to the research community. However, these data shall be referred to as GRUAN-processed data to distinguish these from GDP data files, which are generated exclusively for measurements performed by GRUAN sites.

Finally, the GRUAN data policy for proper acknowledgement and referencing of GRUAN data usage and the guidelines for co-authorship are presented.

For completeness, a concise summary of the relevant aspects of the WMO unified data policy is included in the appendix.

Revision history

Version	Author / Editor	Description
1.0 (2026-02-27)	Ruud Dirksen and Michael Sommer	First published version as Technical Note 10

Table of contents

1	Introduction	6
2	Classification of data levels and data users	7
2.1	Data levels	7
2.2	Data users	8
3	GRUAN strategy for data distribution	9
3.1	Normal operations	9
3.2	Data retention for campaigns	9
3.3	Data from non-GRUAN sites	10
4	Rules for GRUAN data usage	11
4.1	Data retrieval by users	11
4.2	Acknowledgement of GRUAN data usage	11
4.2.1	Co-Authorship	11
A	The WMO unified data policy	13
A.1	Introduction	13
A.2	Application to GRUAN	13
	Acronyms	14

1 Introduction

The purpose of this document is to redefine the GRUAN policy for the handling, sharing and distribution of the data that are submitted to the Lead Centre (LC). This is motivated by requests from the sites who on one side want to increase visibility and recognition for their efforts, but also want to have a mechanism in place for limiting the distribution of certain measurement data, e.g. for data that are collected beyond the scope of their GRUAN commitment, or during measurement campaigns. On the other hand, GRUAN is a World Meteorological Organisation (WMO) co-sponsored programme, and therefore aims to comply with the WMO policy for the exchange of data (discussed in more detail in Appendix A). In addition, a clear mechanism is needed to distinguish data that has been processed on request for external parties, such as for measurement campaigns, from GDP data files that were generated for regular GRUAN sites to ensure proper visibility and acknowledgement.

These issues were topic of discussion during various GRUAN Implementation and Coordination Meetings (ICMs), and this Technical Note aims to find the balance between the WMO policy for free exchange of meteorological data and sharing GRUAN knowledge and achievements with the wider community, and on the other hand acknowledging the interests and visibility of the GRUAN sites.

As part of their GRUAN commitment, all GRUAN sites perform regular radiosoundings whose data are routinely submitted to the Lead Centre. In addition, a considerable number of sites perform balloon soundings with extended payloads that include dedicated instruments for measuring other atmospheric constituents such as ozone, stratospheric water vapour, or aerosol¹. Typical instruments employed for these kinds of measurements are ozonesonde, CFH, FPH, Skydew, POPS, COBALD, and other in situ research instruments, some of which can be beyond the site's GRUAN commitment. Usually these additional instruments are not equipped with a radio transmitter and instead rely on the radiosonde to transmit their measurement data to the ground station. The data from the radiosonde and the instruments connected to it are stored in the radiosounding data file, which after completion of the sounding is submitted to the GRUAN Lead Centre to end up in the GRUAN file archive. In the normal course of operations, data from instruments for which a GRUAN Data Product (GDP) exists will be processed and disseminated to the user community without issue. As a rule, raw measurement data and data from instruments for which no GDP exists are not made publicly available, but these may be shared within the GRUAN community for research or evaluation purposes. However, if the radiosounding file contains data from instruments that are not part of a site's GRUAN commitment, a site may require that steps are taken to ensure that these non-GRUAN data are not distributed further.

In addition, it is common practice for sites that participate in a scientific measurement campaign, that the campaign data are under embargo for certain period. The reason for this is that the campaign partners and participants reserve the right to be the first to write a paper on the campaign's results before publishing the data. This form of data retention should be possible for specific campaign measurements performed at GRUAN sites, but this should not apply to the operational measurements that are performed as part of the site's normal GRUAN commitment.

¹Radiosoundings and observations of stratospheric water vapour are priority-1 parameters and therefore basic requirements for a GRUAN site. Ozone is a priority-2 parameter.

The issue of data retention currently predominantly applies to radiosoundings since the majority of the GDPs exist for radiosondes, but it will become relevant for other measurement systems in the near future, when GDPs for these become available.

To accommodate the wish for control over the data distribution with the GRUAN- and user-community², a classification system for both data and users is introduced, where access privileges to particular data levels are defined for various user categories.

Section 2 defines the various data and user levels within GRUAN, and Section 3 presents the access privileges for the user groups to the various data levels within GRUAN. In addition it presents the options for the sites to request data retention for selected measurements.

Section 4 presents the rules for referencing, acknowledgement and co-authorship when using data provided by GRUAN.

Throughout this document the terms data flow and data stream are used, that have specific meaning and refer to different entities. They are defined as follows:

- **Data flow** – Any movement or transfer of data files, be it between data archive and processing centre, from site to LC, from LC to data portal etc. A more abstract, high-level term, without signifying a specific instrument, data product or site.
- **Data stream** – This refers to all data related actions (e.g. transfer, processing, archiving, dissemination) for a specific instrument type or model, e.g. the RS41 or M10 radiosonde. The term data stream can apply to an individual measurement system, an individual site, or to the complete GRUAN network.
- **Data set** – A collection of data files, e.g. the GDP version 2 data files of iMS-100 soundings at Tateno.

2 Classification of data levels and data users

Data that is submitted to the LC undergoes various steps of (pre)processing to generate the data product files that are eventually disseminated to the user community. Some of these intermediate products, as well as the initial (raw) and the final product data, are stored in the GRUAN file archive. Using the classification given in Section 8.1 of the GRUAN Guide (*GCOS-171, 2013*), we define various data levels and user categories that are used in the data policy in Section 3.

2.1 Data levels

- **Original Raw** – Data file produced by the sounding system (e.g. DC3DB, MWX, GSFZ, ZIP, ...) in a manufacturer-defined format. This file, or collection of files, usually includes raw measurement data³, primary raw data (the 'real' output of the sensors, e.g.

²The GRUAN community refers to those who are active in GRUAN, such as the sites, TT or WG members, or those contributing to GDP development. The user community refers to all using GRUAN data, without being affiliated or having an active role in the network

³Here we follow the commonly used GRUAN definition of raw data, which means unprocessed, calibrated measurement data, with no corrections for measurement errors applied.

voltages or frequencies as recorded by the radiosonde processor), metadata, manufacturer-processed data, and data from additional instruments that were connected to the radiosonde (e.g. ozonesonde, CFH, FPH, Skydew, COBALD, POPS, CPS, etc.).

- **Converted Raw (NetCDF)** – Same as Original Raw, but converted into NetCDF format. The various files and tables from the *Original Raw* data are represented as groups in the NetCDF file.
- **Intermediate Raw (NetCDF)** – Filtered data. Same as *Converted Raw*, but with the data from certain additional instruments excluded.
- **EDT (NetCDF)** – Manufacturer-processed data, in NetCDF format. Examples of this are the Vaisala-processed RS41-EDT or RS92-EDT data products that have been converted to a NetCDF file at the GDP processing centres. Main application within GRUAN: intercomparisons with GRUAN Data Products.
- **GDP** – GRUAN Data Product (NetCDF); the GDP includes amongst others corrected measurement values, uncertainties, applied corrections, raw measurement data, and meta-data.
- **Integrated GDP** – This is a product that results from the combination of measurements from multiple instruments. Example: radiosonde GDP and data from a frostpoint hygrometer. This data type is tailored for specific applications and can for example be used for instrument comparisons.

The development of a GRUAN Data Product (GDP) is a laborious process, where the data processor goes through several iterations of testing, evaluation, adjustment and improvement. Therefore, three levels of product maturity are defined:

- **Alpha** – Development stage.
- **Beta** – Ready for analysis, validation and certification.
- **Release** – Numbered version of the final and certified data product (e.g. GDP.2 for version 2 of the GDP).

2.2 Data users

Data users are classified as follows:

1. LC and relevant specialists (Site, Processing Centre, GDP developer).
2. GRUAN-internal / GRUAN community (e.g. Working Group and Task Teams).
3. External on request. These are people with special interest in GRUAN data, e.g. related to campaigns and not part of group 2.
4. World (everyone).

In the scheme that is presented in Section 3, users from group 1 (Lead Centre and relevant specialists) have access to all data types,

The access to **Original Raw** or **Converted Raw (NetCDF)** by members of user groups 2 and 3 who are associated with a manufacturer is evaluated on an individual basis to prevent potential conflicts of interest.

3 GRUAN strategy for data distribution

3.1 Normal operations

The current situation is that only certified GRUAN data products (as of February 2026: RS92-GDP.2, RS-11G-GDP.1, RS41-GDP.1, and IMS-100-GDP.2) are publicly available and disseminated via the GRUAN website (see Section 4.1). Data streams of instruments for which no GDP exists, such as radiosondes whose GDP is still under development, additional instruments such as CFH, Skydew, FPH, ozonesonde, COBALD, etc. as well as the various forms of raw data and GDP alpha versions reside in the Lead Centre data archive, and are not distributed further. During their respective development phases, the beta versions of the GDPs for the RS41 and iMS-100 were made available for evaluation to a selected group of users.

A data access scheme for the different data levels is shown in Table 1, which defines data access rights for the various data user groups. According to this scheme, users from group 1 (Lead Centre and relevant specialists) have access to all data types, and final data products (both GRUAN and manufacturer-processed) are available to the world. As indicated in Section 2.2, these specialists include the site, the processing centre, and the data product developers, who obviously should have access to the raw data.

GRUAN internal users (group 2) will have access to all data types except the original raw data, and external users from group 3 will have access to intermediate raw and all types of data products (including alpha and beta releases). The rationale behind this is that users from the GRUAN community and the data supplier have access to otherwise restricted data for the purpose of evaluation. Finally, users from group 4 will only have access to the EDT and the final GDP data.

Table 1: Data access scheme to different data levels for various user groups.

	LC & specialists (Group 1)	GRUAN internal (Group 2)	External on request (Group 3)	World (Group 4)
Original Raw	✓			
Converted Raw (NetCDF)	✓	✓		
Intermediate Raw (NetCDF)	✓	✓	✓	
EDT (NetCDF)	✓	✓	✓	✓
GDP (NetCDF)	Alpha	✓	✓	
	Beta	✓	✓	
	Final	✓	✓	✓

3.2 Data retention for campaigns

In short: Site managers can apply for data retention for campaign data to ensure first publication rights. The retention only applies to campaign measurements, and the latency period should not exceed one year.

Deliberate data retention for certain measurements that are performed in the framework of a scientific measurement campaign is a common practice to ensure first-publication rights for the campaign participants. Data retention can be applied to campaign data collected at a GRUAN site at the site's request, although it should not be applied indiscriminately to the entire data stream from the site. Eligible for retention are additional measurements in the framework of the campaign effort, such as radiosoundings outside the regular schedule, or soundings with other radiosonde types or additional non-regular instruments attached to the radiosonde. This campaign data and the processed GDP files will not be distributed, or made available, until the data retention period has expired. Data from campaign instruments attached to the regular GRUAN radiosondes are included in the **Original Raw** and **Intermediate Raw** files that reside in the Lead Centre data archive and are not distributed further, as is shown in Table 1. If for these additional instruments GDPs exist, which will be the case in the future for CFH, Skydew or FPH, these data will be processed and the GDP files will remain in escrow until the end of the retention period. The one-year limit for the retention of campaign data is consistent with the WMO unified data policy ([WMO, 2022](#)), which calls for a maximum delay of one year for climate data.

The site manager can request data retention by submitting a well-founded and well-documented application. This application should clearly indicate the time frame and duration of the campaign, the measurements involved, and the desired data retention period.

When GDP developers use data for evaluation of the data product or its processing, it is required that they involve the data providers (the sites) in this activity.

3.3 Data from non-GRUAN sites

On several occasions data has been processed for sites that are not part of GRUAN. Examples of these are the MOSAiC expedition ([Shupe et al., 2022](#); [Maturilli et al., 2022](#)), the StratoClim campaign ([Brunamonti et al., 2018](#)), and a satellite comparison study over Eureka ([Weaver et al., 2019](#)). Although the Lead Centre gladly supports external research projects by providing GRUAN's reference-quality data, there must be a clear distinction between these campaign data and the data stream from regular GRUAN sites. The rationale for this is to ensure that GRUAN sites get proper acknowledgement for their efforts, and that some sites need visibility within their organisation to justify the resources that go into maintaining a GRUAN measurement program. Publishing campaign data from non-GRUAN sites as regular GRUAN data, that is indistinguishable from the data stream from GRUAN sites, could confuse the data users and may lead to a distorted view of GRUAN, incorrectly identifying a non-GRUAN site as part of the network, suggesting the existence of more sites, or different sites, than those that are actually part of the network. This could ultimately be disadvantageous to the GRUAN sites.

Therefore, GDP files, or **GRUAN data**, is a term exclusively reserved for measurements performed by GRUAN sites, whereas data from non-GRUAN sites that has been processed by the Lead Centre or another processing centre will be referred to as **GRUAN-processed data**.

4 Rules for GRUAN data usage

4.1 Data retrieval by users

As of February 2026, GRUAN GDP radiosonde data are available at a GRUAN server, which can be reached via the GRUAN website <https://www.gruan.org/data/data-products>. Data on this server is accessible for all users after registration and login (data user group 4)⁴.

4.2 Acknowledgement of GRUAN data usage

Whenever GRUAN data have been used in scientific work that is being published, the data origin must be acknowledged and referenced. A minimum requirement is to reference

1. GRUAN, as a reference network of GCOS.
2. The Technical Document describing the GDP in question.

If only data from one GRUAN site (or a limited number of sites) has been used,

3. Additional acknowledgement to the reference site(s) and their sponsoring institutions or organisations shall be given. See <https://www.gruan.org/network/sites> for information on the sites.

Proper referencing shall include the DOIs of the GDP datasets (Table 2). When DOIs for data from individual sites are available, these shall be used for referencing the data.

Table 2: Available GDP datasets with reference.

GDP name	DOI	Certification year
RS92-GDP.2	10.5676/GRUAN/RS92-GDP.2	2012
RS-11G-GDP.1	10.5676/GRUAN/RS-11G-GDP.1	2019
RS41-GDP.1	10.5676/GRUAN/RS41-GDP.1	2022
IMS-100-GDP.2	10.5676/GRUAN/IMS-100-GDP.2	2022

When GRUAN data are used for a scientific publication, sending an electronic copy of the paper to the Lead Centre would be appreciated.

4.2.1 Co-Authorship

In short:

- When using radiosonde GDP data, co-authorship is not required, but would be appreciated.

⁴In addition, GRUAN GDP data for RS92-GDP.2 and RS-11G-GDP.1 from before 2021 are also available at NCEI <https://www.ncei.noaa.gov/pub/data/gruan/processing/level2/>

- When using GRUAN-processed radiosonde data from non-GRUAN sites, co-authorship should be offered to (a representative of) the GDP-developer(s) and to (a representative of) the processing centre.
- When using non-radiosonde GDP data, the GDP developer and the instrument specialist from the site that provides the data should be offered co-authorship.

GRUAN was developed by an international panel of scientists including experts in radiosondes and remote sensing measurements for water vapour, temperature and pressure in the upper troposphere and lower stratosphere. Apart from operational radiosonde systems, GRUAN sites are equipped with sophisticated, state-of-the-art instrumentation and comply with strict requirements of station maintenance, exposure of instruments, calibration, quality assurance procedures and the like. To ensure that the goal of long-term high-quality climate records is reached, site scientists, who are leading experts for the instruments used at the respective GRUAN sites, often take responsibility for individual instruments operated at the GRUAN site.

Inclusion of GRUAN scientists as co-authors of papers making extensive use of GRUAN (non-radiosonde) data is justifiable and highly recommended, in particular if a site scientist or instrument specialist has responded to questions raised about data quality and/or suitability for the specific study in question, or has been directly involved in contributing to the paper in other ways. In case of radiosonde data, the GDP developers are the relevant experts, who are qualified to answer questions on the data product and who should be considered for co-authorship. The co-authorship is not a pre-condition for release of GRUAN data. However, it is highly recommended that any data user should contact the responsible site scientist or GDP developer and ask if he/she wants to become co-author, or if an acknowledgement would be sufficient. Users of GRUAN data are encouraged to establish direct contact with site scientists/GDP developers for the purpose of complete interpretation and analysis of data for publication purposes.

In case of GRUAN-processed data, data processed for sites that are not part of GRUAN, co-authorship of a member of the GDP-development team and/or the data processing centre is required.

A The WMO unified data policy

A.1 Introduction

As described in the Chapter “Development and Adoption of the WMO Unified Data Policy” of [WMO \(2022\)](#): International collaboration in weather and climate was first formalized in 1873, with the founding of the International Meteorological Organisation (IMO). In 1947 the WMO replaced the IMO, and data exchange became formalized via WMO’s World Weather Watch (WWW) and the Global Atmospheric Research Programme (GARP) in 1967. To counteract the upcoming desire for monetisation of weather data and the resulting unravelling of international exchange of meteorological data, WMO Cg-XII adopted resolution 40 to reinforce its commitment to free and unrestricted international exchange of meteorological data. Resolution 40 mainly dealt with weather data, and was later complemented with Resolution 25 (Cg-XIII) for hydrological data and Resolution 60 (Cg-17) for climate data. The improvement in weather prediction quality due to ongoing development of models, assimilation and progress in (satellite) observations and computing power renewed the interest in the commercial value of meteorological data. This, together with the recognition that the Earth’s climate must be studied and understood as an integrated and interacting system of atmosphere, ocean and land, led to the adoption of the WMO unified data policy (Resolution 1 (Cg-Ex. 2021)) that combines the before-mentioned resolutions 40, 25, and 60 and expands the data exchange beyond weather, climate and water to include atmospheric composition, oceanography, cryosphere and space weather.

The purpose of the WMO unified data policy ([WMO, 2022](#)) is to give WMO-members access to timely and reliable weather, climate, water and related environmental data and information. In essence it calls for the exchange in real or near-real time of core weather data, and for the timely exchange, preferably within one year, of core observational data for climate.

The full definitions of core data for weather and climate are given in the appropriate documents cited in Resolution 1, suffice to say that the (operational) measurement programs performed for GRUAN can be regarded core observation data.

A.2 Application to GRUAN

It is stated in Chapter 8.2 of the GRUAN Guide ([GCOS-171, 2013](#)), that GRUAN is a WMO co-sponsored programme, and therefore it is appropriate to comply with WMO policy, practice, and guidelines for the exchange of meteorological and related data and products. The Guide refers to the then-applicable WMO data policy defined in Resolution 40 (Cg-XII) and the new unified data policy from 2021 basically expands the uninhibited data exchange to observational and model data from other areas.

It is also mentioned in the Guide that the GRUAN data policy only applies to data submitted to and processed and distributed by GRUAN. Hence, data that are provided by the sites to other networks are not considered GRUAN data.

When joining GRUAN, sites are explicitly reminded of this in the confirmation letter that is sent by GCOS secretariat, and thereby implicitly agree with the exchange of data within GRUAN and beyond.

With the strategy for data distribution and retention described in Section 3, GRUAN complies with the WMO guidelines for exchange of meteorological and climate data outlined in the WMO unified data policy.

Acronyms

CFH	Cryogenic Frostpoint Hygrometer
COBALD	Compact Optical Backscatter Aerosol Detector
CPS	Cloud Particle Sensor
DC3DB	DigiCORA 3 DataBase File Format
DOI	Digital Object Identifier
EDT	Manufacturer data product
FPH	Frostpoint Hygrometer
GARP	Global Atmospheric Research Programme
GCOS	Global Climate Observing System
GDP	GRUAN Data Product
GRUAN	GCOS Reference Upper-Air Network
GSFZ	Zipped Graw Sounding File
ICM	Implementation and Coordination Meeting
IMO	International Meteorological Organisation
LC	Lead Centre
MWX	Zipped MW41 sounding archive file
NCEI	National Centers for Environmental Information
NetCDF	Network Common Data Format
POPS	Portable Optical Particle Spectrometer
TT	Task Team
WG	Working Group
WMO	World Meteorological Organisation
WWW	World Weather Watch
ZIP	Compressed archive file format

References

Brunamonti, S., *et al.*, Balloon-borne measurements of temperature, water vapor, ozone and aerosol backscatter at the southern slopes of the Himalayas during StratoClim 2016-2017, *Atmospheric Chemistry and Physics*, **18**(21), 15,937–15,957, <https://doi.org/10.5194/acp-18-15937-2018>, 2018.

GCOS-171, The GCOS Reference Upper-Air Network (GRUAN) GUIDE, Technical report, WIGOS Technical Report No. 2013-03, 2013, URL https://library.wmo.int/doc_num.php?explnum_id=7196.

Maturilli, M., *et al.*, MOSAiC radiosonde data (level 3), 2022, <https://doi.org/10.1594/PANGAEA.943870>.

- Shupe, M. D., *et al.*, Overview of the mosaic expedition: Atmosphere, *Elementa: Science of the Anthropocene*, **10**(1), 00,060, <https://doi.org/10.1525/elementa.2021.00060>, 2022, ISSN 2325-1026.
- Weaver, D., *et al.*, Comparison of ground-based and satellite measurements of water vapour vertical profiles over Ellesmere Island, Nunavut, *Atmospheric Measurement Techniques*, **12**(7), 4039–4063, <https://doi.org/10.5194/amt-12-4039-2019>, 2019.
- WMO, WMO unified data policy, 2022, URL <https://library.wmo.int/idurl/4/58009>, [Accessed 20 October 2025].