GCOS
Reference
Upper-
Air
Network

GRUAN Technical Note 13

User Guide for the RS41 GRUAN Data Product Version 1 (RS41-GDP.1)

Michael Sommer, Christoph von Rohden and Tzvetan Simeonov

Publisher

GRUAN Lead Centre

Number & Version

GRUAN-TN-13
1.0 (2022-11-21)

Document info

Title:	User Guide for the RS41 GRUAN Data Product Version 1 (RS41-GDP.1)
Topic:	Radiosonde
Authors:	Michael Sommer, Christoph von Rohden and Tzvetan Simeonov
Publisher:	GRUAN Lead Centre, DWD
Document type:	Technical Note
Document number:	GRUAN-TN-13
Page count:	60
Version:	1.0 (2022-11-21)

Abstract

This document describes the format of the data files in the GRUAN Data Product (GDP) of the RS41 radiosonde. GDPs are stored in NetCDF files. The data fields and attributes within the RS41 radiosonde NetCDF files are comprehensively explained in this document. The GDP includes raw data, processed data, uncertainty estimates, and selected meta data.

The measurement practice of GRUAN-compliant radiosoundings as well as the algorithms for correction and processing of the raw data including the underlying motivations, assumptions, and uncertainty considerations are extensively described in the main RS41 Technical Document *Sommer et al.* (2022a, in review).

In this version of the document, the release version 1 of the RS41 GRUAN Data Product (RS41-GDP.1), which is in operational use since 2021-11-09 (release date), is described (*Sommer et al.*, 2022b).

Editor Remarks

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by GRUAN.

Contacting GRUAN Lead Centre

Please contact the GRUAN Lead Centre (gruan.lc@dwd.de) if you have any questions or comments in relation to this document, or in relation to GRUAN's use of Vaisala RS41 radiosonde data.

Revision history

Version	Author / Editor	Description
1.0 (2022-11-21)	M. Sommer, C. von Rohden, T. Simeonov	First published version as GRUAN Technical Note 13 (GRUAN-TN-13)
0.9 (2022-10-17)	Michael Sommer	Prepared for TT-internal review
0.4 (2022-06-29)	Christoph von Rohden, Michael Sommer	LC-internal review
0.3.3 (2022-05-03)	Michael Sommer	Prepared for LC-internal review
0.2 (2020-06-30)	Michael Sommer	Add available content from GRUAN-TD-4 (RS92 document)
0.1 (2020-06-26)	Michael Sommer	Development of general document structure; add file content as appendix
0.0 (2020-06-25)	LC	Template with a base structure, no contents

Table of contents

1 Introduction

The GCOS Reference Upper-Air Network (GRUAN) data product, derived from Vaisala RS41 radiosonde measurements, was developed to meet the criteria for reference measurements. These criteria stipulate the collection of meta-data, the use of well-documented correction algorithms, and estimates of the measurement uncertainty (*Immler et al.*, 2010). An important and novel aspect of the GRUAN processing is that the uncertainty estimates are vertically resolved.

The RS41 radiosonde is used to measure atmospheric profiles of temperature (T), relative humidity (U), pressure (p), as well as wind speed and wind direction from the surface to the lower stratosphere with a time resolution of one second. The GRUAN data processing is based on the raw data of the measurements, and the algorithms include corrections for systematic and random effects, resulting in a manufacturer-independent data product which should be free of biases (*von Rohden et al.*, 2022; *Sommer et al.*, 2022a). The uncertainties, which are estimated for each variable as essential part of the processing, are calculated following the recommendations outlined in the "Guide for expressing uncertainty in measurement", *JCGM* (2008). The combined total uncertainties generally consist of statistical components (e.g. from sensor calibration) and uncertainties of the corrections for systematic effects (bias).

The results of the RS41 processing are stored in GRUAN data product files which comply with the Network Common Data Format (NetCDF). The files contain the processed data together with the raw calibrated measurements, relevant supplementary variables, and comprehensive meta-data. This document provides guidance on the names, locations, and access of the various RS41 GDP variables in the NetCDF product files and summarises information about their contents and also their origin with regard to the processing.

2 GRUAN processing of RS41 data

2.1 Processing concept and inclusion of uncertainties

The GRUAN processing of the data files produced in a RS41 sounding is a two-stage process, where the first step involves collection, validation and pre-processing of the sensor measurement data and accompanying meta-data, which are subsequently stored in a database. The actual processing (i.e. the application of the GRUAN correction algorithms to the measured raw data) is performed during the second step provided that the relevant (meta) data files are available.

The RsLaunchClient, a software tool that is written and provided by the Lead Centre, facilitates the collection of (meta)data and their subsequent submission to the GRUAN database. The RsLaunchClient is particularly useful for collecting meta-data from soundings with a complex configuration, e.g. rigs consisting of multiple sondes or non-radiosonde instruments (see GRUAN-TD-3: *Sommer*, 2014).

Stations using autolaunchers or launch setups that are not changed during longer periods of routine operation can alternatively use the GruanToolRsLaunch (gtRsl), which is also written and provided by the Lead Centre. This tool is useful for an automated data transfer to the GRUAN database (see GRUAN-TN-6: *Sommer*, 2020).

Upon submission using one of the tools, the meta-data are stored in the GRUAN Meta-Data Base (GMDB), while the radiosonde raw data are stored in a file archive at the GRUAN Lead Centre (LC). The GRUAN data processing is started as soon as all meta-data and raw data for a radiosonde launch are available. In the first step of the processing the original Vaisala MW41 sounding files (*.mwx) are converted into the NetCDF format, preserving all information contained in the MWX file. The raw PTU data are available from the tables *RawPtu* and *GSupport*, and the GPS position data are taken from the tables *GpsRawMeasurements* and *GpsResults*.

An extensive description of the correction algorithms of the GRUAN data processing is given in *Sommer et al.* (2022a, in review), and in particular for temperature in daytime soundings in *von Rohden et al.* (2022). A brief summary of these algorithms is given below. The GRUAN processing is applied to calibrated sensor data (raw data) after application of corrections from manufacturer-prescribed ground checks.

An important component is the estimation of uncertainties associated with the processing steps for each variable. The individual uncertainties contribute to the combined total uncertainty of a variable, which is given as standard uncertainty with a coverage factor $k = 2$ alongside with the data columns for the actual variable.

In addition, uncertainties are also evaluated with regard to their correlation properties, i.e. a particular type of correlation is assigned to each uncertainty component. Three types are defined: ‘uncorrelated’ (‘ucor’), ‘spatially correlated’ (‘scor’), and ‘correlated in time’ (‘tcor’). These types stand for independent (random) uncertainties, correlated uncertainties within a single sounding profile (but no correlation with the uncertainties of other soundings), and correlated uncertainties amongst separate profiles measured at different times and/or locations (sites), respectively. In this concept, temporal correlation in general includes the spatial (vertical) correlation within individual profiles.

The above classification does not follow from strict mathematical-statistical considerations or comprehensive covariance analyses because of complexity. Instead a simplified approach is chosen, where the 'scor' and 'tcor' uncertainties assume a full correlation, according to a correlation coefficient of $r = 1$, and uncorrelated ('ucor') assume $r = 0$. The correlation categories for each variable are specified 'by hand', based on evidence or experience.

This classification scheme allows for the combination of the complete set of correlated and uncorrelated uncertainties into a consistent total uncertainty value, equivalent to the original combined uncertainty:

$$u = \sqrt{u_{\text{ucor}}^2 + u_{\text{scor}}^2 + u_{\text{tcor}}^2}. \tag{2.1}$$

In *Sommer et al.* (2022a) more motivation and information on uncertainties with regard to their correlation type is given. Special flow chart diagrams are presented for temperature, relative humidity, and pressure from both dedicated sensor and derived from GPS, illustrating the uncertainty progression and the implementation in the processing.

The combined uncertainty (to be identified by the appendage *_uc* in the variable name), and at least one of the uncertainties with respect to the correlation type (*_uc_ucor*, *_uc_scor*, and *_uc_tcor*) are provided in the GRUAN data files for all variables (see also Section 3.2.1). Not all of the three correlated components are always present, or one or more may be present but filled with NaN. In such cases they are not existent or not evaluated for the actual variable. For example, the uncertainty component *_uc_scor* is associated only to daytime measurements of temperature.

> **Note:** Uncertainty estimates for variables and attributes are consistently specified in the product files at the $k = 2$ confidence level.

2.2 Pre-checks of raw data

During a sounding, which can take up to several hours, the RS41 transmits data to the ground station every second. The raw data set of a sounding encompasses the whole period between radiosonde initialisation until the end of data recording with the ground receiver or the complete loss of the signal. Individual data packets can also be lost intermittently, which may produce gaps in the time series (or the profile) of variable size during a sounding. Analyses of numerous soundings indicate that the appearance or pattern of data transmission errors (frequency, extent, distribution) is variable. The processing software is designed to recognise and tackle such irregularities, however problems still may arise in case they are of unexpected type.

As an important step to detect problems which potentially may hinder the subsequent processing, a pre-check of the raw data is routinely performed. All measured raw data are analysed and undergo a pre-processing:

- *outlier detection* – Outliers are detected and removed.

- *gap detection* – Gaps (missing data) are located and the length of the gaps is determined.
- *gap interpolation* – Existing gaps up to a size according to 10 seconds are filled with linearly interpolated values. Removed outliers are treated as gaps.

The raw data and the results of the pre-check analyses are available as variables with appendage *_raw* for the respective variable name and *_raw_flags* for associated flags. The list below contains the concerned variables:

- *temp_raw* – Raw temperature,
- *ihum_raw* – Raw humidity at temperature of humidity sensor,
- *thum_raw* – Raw temperature of humidity sensor,
- *rh_raw* – Raw humidity, given in relation to the air temperature,
- *press_sens_raw* – Raw sensor pressure before adjustment after ground check,
- *posx_raw* – Raw geographical Cartesian coordinate X,
- *posy_raw* – Raw geographical Cartesian coordinate Y,
- *posz_raw* – Raw geographical Cartesian coordinate Z,
- *lat_raw* – Raw latitude (WGS84),
- *lon_raw* – Raw longitude (WGS84),
- *alt_raw* – Raw altitude (WGS84),
- *wzon_raw* – Raw zonal wind component,
- *wmeri_raw* – Raw meridional wind component,
- *hodp_raw* – Raw Horizontal Dilution Of Precision (HDOP) (*no raw data flags available*),
- *vdop_raw* – Raw Vertical Dilution Of Precision (VDOP) (*no raw data flags available*).

All *_raw_flags* variables contain a bit-wise combined flag in which several flags related to the raw data variable are encoded:

- *1* – not a number (NaN), missing data,
- *2* – value classified as outlier and removed (set to NaN),
- *4* – value detected to be out of the valid range,
- *8* – value identified to be exactly zero, and
- *16* – value which is linearly interpolated by the gap interpolation step.

2.3 Ground checks

Before launching a RS41 radiosonde, various ground checks can be performed. The global attributes "g.GroundCheck." contain details and results of these checks (see Appendix A.1).

If performed and documented, the following ground checks can be analysed within the processing:

- *RI41* – Manufacturer's standard check for radiosonde initialisation and functionality test. Only the pressure sensor readings are directly compared to a reference sensor.
- *SHC* – Independent ground check recommended by GRUAN, which is carried out indoors using a Standard Humidity Chamber (SHC). The main purpose is a humidity check

at 100 % relative humidity. The temperature can simultaneously be compared to a reference thermometer. In addition, an independent comparison of the pressure sensor with a reference sensor can be performed during that period.

- *SHELTER* – Optional outdoor ground check, which can e.g. be carried out in a weather hut at the launch site for comparison of humidity, temperature, and pressure.

For each of the performed ground checks, a set of analysis results is provided, which refers to all examined variables (e.g. humidity, temperature, pressure). The results include the measured radiosonde and reference sensor values, the pre-calculated differences, and uncertainty estimates. In case a sounding is carried out without prior performance of the GRUAN SHC ground check, a statistically evidenced uncertainty estimate based on existing SHC check results is assigned to account for the lack of information due to the missing check.

2.4 Position, altitude, and geopotential height

Altitude and horizontal position variables are based on GPS measurements of the RS41. The position is given by geographical longitude (*lon*) and latitude (*lat*). In the RS41-GDP.1 data product, altitudes based on three different reference systems are provided to the user: WGS84 altitude (*alt_wgs84*), altitude above mean sea level (*alt_amsl*), and geopotential height (*alt_gph*).

The geopotential height is selected as the standard altitude variable *alt*, since it is used in the standardised data output of the WMO (WIGOS), which is particularly relevant for weather forecasting. The trajectory of the radiosonde measurement, i.e. the coordinate triples to which all other measured variables are assigned, are *lat*, *lon*, *alt*.

List of relevant main variables:

- *alt* – Default altitude. *alt_gph* is chosen as default.
- *alt_wgs84* – Altitude WGS84 [m],
- *alt_amsl* – Altitude above mean sea level [m],
- *alt_gph* – Geopotential height [m],
- *lat* – Latitude [°N],
- *lon* – Longitude [°E].

2.5 Pressure

The barometric pressure is derived from the GPS altitude (*press_gnss*) based on the barometric altitude formula, and taking into account the actual temperature and humidity profile.

For the RS41-SGP model variant, which is equipped with a physical pressure sensor, the GDP offers the sensor pressure (*press_sens*) separately to the GPS-based calculated pressure. In that case, the sensor readings are compared to a reference sensor at the ground before launch. The determined offset is applied as a relative correction (factor) to restore the original calibration (*press_sens_corr_cal*). As final step in the processing, smoothing is applied (*press_sens_corr_sm*), which removes sensor noise. Both the raw data (*press_sens_raw*) as well as the full correction (*press_sens_corr*) and its individual components can be accessed separately (see Section 3.2.2).

The pressure derived from GPS is used as the standard pressure variable (*press*). In case of the RS41-SGP the two pressure variables are not combined or merged in any way.

List of pressure-related main variables:

- *press* – Default pressure. *press_gnss* is chosen as default.
- *press_sens* – Barometric air pressure measured using silicon sensor [hPa]
- *press_gnss* – Barometric air pressure derived from GNSS-altitude [hPa].

In addition, the combined uncertainty (*_uc*) and correlated uncertainty components (*_uc_ucor*, *_uc_scor*, and *_uc_tcor* according to availability) are provided (see also Section 3.2.1).

2.6 Temperature

For the temperature measurement, the most significant correction is related to the solar short-wave induced sensor warming during day-time ascents (radiation error, *temp_corr_rad*). The radiation correction step in the processing includes information from a number of additional sources, e.g. sun elevation, model estimates of short-wave actinic fluxes, and ventilation derived from sonde movements. In a final step, a smoothing procedure is applied to remove typical patterns that arise from the rotation or pendulum movement of the radiosonde (*temp_corr_sm*). The effective length of the smoothing kernel and therefore the effective vertical resolution (*temp_res*) is adjusted to day and night soundings to avoid over-correction, and it can be variable within a profile, e.g. under twilight conditions.

Both the raw data (*temp_raw*) as well as the full amount of correction (*temp_corr*) including the individual components of the correction amount can be accessed separately (see Section 3.2.2).

List of temperature-related main variables:

- *temp* – Temperature [K].

In addition to the combined uncertainty (*_uc*) and the correlation-related uncertainty components (*_uc_ucor*, *_uc_scor*, *_uc_tcor*, see Section 3.2.1), several (sub)components are provided:

- *temp_uc_rad* – Combined uncertainty related to temperature radiation correction. It is composed of its own correlation-related uncertainties, namely *_uc_ucor_rad*, associated with the calculation of ventilation and pressure (the latter of which is negligibly small), *_uc_scor_rad*, the origin of which is the solar radiation model (RTM), and *_uc_tcor_rad*, a temporally correlated uncertainty associated with the laboratory experiments for the radiation sensitivity of the temperature measurement.
- *temp_uc_tcor_cal* – calibration-related uncertainty after ground preparation (temporally correlated)
- *temp_uc_ucor_sm* – Component related to smoothing (uncorrelated).

The uncertainty (sub)component *_uc_scor_rad* (and therefore *_uc_scor*) should be treated as correlated when comparing two or more sonde units of the same type attached to the same rig and using the same product version for evaluation.

A high amount of correlation (i.e. 'full' correlation expressed in *_uc_scor*) can still be assumed for separate flights with the same radiosonde model when launched at the same site close in time, because the illumination conditions are very similar.

For a summary of information on the uncertainty components and their linking see the graphical representation in Fig. 6.1 in *Sommer et al.* (2022a).

2.7 Humidity

For relative humidity, the most important variables provided in the product file are the corrected relative humidity *rh* (the final product) and the associated uncertainty *rh_uc*. In the processing, the final humidity is derived from the original humidity measured by the continuously heated sensor at its 'internal' 5 K-elevated temperature. That is, the measured values from the sensor are 'converted' to the actual relative humidity of the air according to the temperature difference between air and sensor. The internal raw data are given for the sake of completeness in the columns *ihum_raw* and *thum_raw*, respectively. The raw humidity in accordance to the measured ambient air temperature is saved in *rh_raw*. Note that the measured air temperature is also 'raw', i.e. uncorrected for e.g. the solar radiative effect. To explain the humidity variables stored in the GDP files, the raw humidity *rh_raw* can be considered as the initial humidity variable to which several corrections are subsequently applied.

The overall humidity correction is given in *rh_corr* (= *rh* − *rh_raw*). It is composed of three (additive) components, which are related to the humidity sensor time-lag, the radiation error of the air temperature measurement, and data smoothing, respectively. To offer the users more flexibility for customised data evaluations, these components are saved separately in *rh_corr_tl*, *rh_corr_trc*, and *rh_corr_sm*. They are briefly explained in the following.

The response of the sensor to changes in relative humidity increases considerably with decreasing absolute temperature. This means that the sensor inherently smoothes the measured profile and introduces a temporal shift compared to the real course of the humidity. The effect, which is called time-lag, is corrected by reconstructing the expected 'true' profile from the measured data. The profile of correction amounts is saved in *rh_corr_tl*. The correction procedure includes a smoothing of the reconstructed profile to dampen noise components that are amplified by the time-lag correction algorithm. The effective time resolution of the corrected profile is generally lower than that of the raw profile, and also variable with altitude because of the use of a smoothing kernel with a kernel length that is related to temperature and therefore to the time-lag. The resolution is saved in (*rh_res*).

In daytime, the air temperature measurement using the radiosonde's temperature sensor may be biased by solar radiative warming. This effect can be transferred in a respective humidity bias that forms another part of the overall correction of humidity. It is saved in *rh_corr_trc*.

The third correction component comes from a slight smoothing which is again applied to the fully corrected humidity profile in a final step. It is saved in *rh_corr_sm*.

The following uncertainties are provided in the product file:

- *rh_uc* – Combined overall uncertainty of the corrected relative humidity, consisting of the two correlation-related combined components *rh_ucor* and *rh_tcor* (next two bullets).

- *rh_ucor* – Combined uncorrelated uncertainty, consisting of
  - *rh_uc_ucor_sm* – Smoothing-related uncertainty
  - *rh_uc_ucor_tair* – Uncertainty related to the air temperature
  - *rh_uc_ucor_tlc* – Time-lag related uncertainty
- *rh_tcor* – Combined temporally correlated uncertainty, consisting of
  - *rh_uc_tcor_cal* – Uncertainty related to the calibration after ground preparation
  - *rh_uc_tcor_tint* – Uncertainty of the 'internal' temperature of the humidity sensor
  - *rh_uc_tcor_tlc* – Time-lag related uncertainty

  The uncertainty of the time-lag correction is provided separately:
- *rh_uc_tlc* – Combined uncertainty related to time-lag correction, composed of the two correlation-related uncertainty components *rh_uc_ucor_tlc* and *rh_uc_tcor_tlc* (listed at the end of each of the two previous bullets).

Combined uncertainties are to be understood as the square root of the sum of squared components. For classification of the humidity uncertainty components see also Fig. 6.2 in *Sommer et al.* (2022a).

Relative humidity is considered the most common way of expressing the atmospheric water vapour content, at least up to the tropopause. However, depending on the application, various other humidity measures are in use. Therefore, several of these measures of humidity are provided in the product file:

- *rh* – Relative humidity [%RH],
- *dp* – Dew point temperature [K],
- *wvmr_mass* – Water vapour mass mixing ratio [ppm],
- *wvmr_vol* – Water vapour volume mixing ratio [ppmv],
- *wvpp* – Water vapour partial pressure [hPa],
- *wvsp* – Water vapour saturation pressure [hPa].

The combined uncertainty (*_uc*) and several of the correlation-related components (*_uc_ucor*, *_uc_scor*, *_uc_tcor*) are provided for all of these variables (see also Section 3.2.1).

2.8 Wind

Wind speed and wind direction are derived from the radiosonde's measurements of geographical position changes using the GPS. To account for the pendulum movements of the radiosonde relative to the position of the balloon, the data are smoothed accordingly before evaluating wind parameters.

For convenience, unsmoothed wind speed and direction (*wspeed_usm*, *wdir_usm*) and raw data of the wind speed components (*wzon_raw*, *wmeri_raw*) as well as the correction amounts related to pendulum smoothing (*wspeed_corr_sm* and *wdir_corr_sm*) are provided (see Section 3.2.2).

In addition to the combined uncertainty (*_uc*) and one of the correlation-related uncertainty components (*_uc_ucor*, Section 3.2.1), two source-related uncertainty components are provided for both wind variables:

- *_ucor_sm* – Smoothing-related part of uncorrelated uncertainty
- *_ucor_gnss* – GNSS related part of uncorrelated uncertainty

List of wind-related main variables:

- *wdir* – Wind direction (measured clockwise from North) [°]
- *wspeed* – Wind speed [m s^{-1}]
- *wzon* – Zonal wind component [m s^{-1}]
- *wmeri* – Meridional wind component [m s^{-1}].

2.9 Supplementary variables

In support of a deeper understanding of the results of individual processing steps, a number of supplementary variables are provided in addition to the main variables. Some of them may be useful in conjunction with the radiation correction of temperature, e.g. sun elevation, day and night conditions, or results of solar radiation flux simulations:

- *sza* – Solar zenith angle [°]
- *sea* – Solar elevation angle [°]
- *saa* – Solar azimuth angle [°]
- *hca* – Horizon correction angle [°]
- *dorn* – Day or night flag (Flag which decodes day, twilight, or night: 0 – unknown, 1 – day, 2 – night, 4 – civil twilight, 8 – nautical twilight, 16 – astronomical twilight, 32 – partly sun) [1]
- *swdir* – Simulated solar direct flux [W m^{-2}]
- *swdif* – Simulated solar diffuse flux [W m^{-2}].

The ventilation speed is important for the daytime radiation correction of the temperature. It is composed of the vertical ascent speed and the horizontal pendulum speed. Intermediate variables for the analysis of the pendulum motion are provided with:

- *vspeed* – Vertical speed of radiosonde (ascent/descent speed) [m s^{-1}]
- *lat_sm* – Pendulum-smoothed trajectory latitude [°N]
- *lon_sm* – Pendulum-smoothed trajectory longitude [°E]
- *latmf* – Factor to convert latitude from degree to meter [m]
- *lonmf* – Factor to convert longitude from degree to meter [m]
- *penper* – Pendulum period [s]
- *penspeed* – Pendulum speed [m s^{-1}]
- *penrad* – Pendulum radius [m]
- *vent* – Ventilation at sensor boom [m s^{-1}].

Furthermore, some more specific variables are provided that can be used in appropriate contexts. List of supplementary main variables:

- *grav* – Acceleration of gravity [m s^{-2}]

- *geoid* – Difference between measured geoid and ellipsoid WGS84 [m]
- *tau_rh* – Timelag (63.2 %-time) of humidity sensor [s]
- *ciwv* – Cumulated integrated water vapour [kg m^{-2}]
- *icesat* – Relative humidity (over water) at ice saturation [%RH]
- *band* – Measurement band flag (Flag which decodes measurement band: 1 – flight, 2 – ascent, 4 – descent, 8 – troposphere, 16 – stratosphere, 32 – utls, 64 – prelaunch, 128 – postflight) [–].

The combined uncertainty (*_uc*) and correlation-related components of uncertainty (*_uc_ucor*, *_uc_scor*, *_uc_tcor*) are provided for some of these supplementary variables (see also Section 3.2.1).

3 Basic structure and use of NetCDF files

GRUAN Data Product (GDP) files are provided in NetCDF format. NetCDF is a self-describing binary format developed by UCAR/Unidata. Software libraries for handling NetCDF files are available for a number of programming languages (e.g. FORTRAN, C, JAVA, Python) and can be downloaded at https://www.unidata.ucar.edu/software/netcdf, or are directly integrated in data analysis tools or programming software (e.g. IDL, Matlab).

The GRUAN processing uses the capabilities of NetCDF to include meta-data. The following common terms are explained here to help understanding the data product: global attributes, variables and variable attributes.

3.1 Global attributes

Global attributes are intended to store information (meta-data) relevant to the entire content of the file. Two groups of meta-data attributes are distinguished:

- General information using the Climate and Forecast (CF) meta-data standard CF-1.7 (*Eaton et al.*, 2017, https://cfconventions.org/)
- GRUAN-specific details, including
  - product and file
  - site, measurement system and measurement setup
  - measurement (event) and surface observations
  - used equipment (instruments): ground system, main sonde, telemetry sonde
  - performed ground checks: RI41, SHC and SHELTER.

Global attributes describe the data product in general, the time and location of the measurement, meteorological conditions at the ground, the software version of the ground station equipment, etc.

Global attributes are saved in the following format:

> *g.<Category>.<AttributeName>*: value (always stored as string)

See Appendix A.1 for a description of all global attributes.

3.2 Variables

Various variables are included in each GDP NetCDF file. The 1D-arrays of the variables are always of the same length. In that sense they can be considered analogous to columns of a single table. Each file contains one 1D-array with the time variable in s. The arrays of all variables relate to this time array.

The files furthermore contain columns with vertically resolved uncertainty estimates related to the data columns for most of the variables. Variables for which no corresponding uncertainty estimate is given should be considered as ancillary data. A list of all variables (i.e. of all columns in the table) is provided in Appendix A.2.

The variables can be grouped according to the following scheme:

- time, altitude, position
- pressure, temperature, relative humidity
- other humidity measures, wind
- supplementary variables
- raw data, flags related to raw data variables.

3.2.1 Uncertainty components

Note: Uncertainty estimates in the file are consistently specified at the 95 % confidence level (coverage factor $k = 2$) for all variables and attributes.

For each main variable, combined overall uncertainties (*_uc*) are given. These overall uncertainties are combinations of individual components which represent evaluated or estimated uncertainties assigned to (physical) magnitudes or variables, e.g. calibration, corrections, smoothing, other sensor, etc. They are – after conversion to the dimension of the main variable according to their individual sensitivities – combined by common error propagation:

$$u = \sqrt{u_{\mathrm{p1}}^2 + u_{\mathrm{p2}}^2 + u_{\mathrm{p3}}^2 + \dots} \qquad (3.1)$$

However, each of these components is assigned one of the three correlation properties: uncorrelated (correlation coefficient $r = 0$), fully spatially correlated ($r = 1$), or fully temporally correlated ($r = 0$), designated with *_uc_ucor*, *_uc_scor*, or *_uc_tcor*, respectively (see list below). Consistently, after combining uncertainty components of the same correlation type, the combination of these (un)correlated components using error propagation again gives the overall uncertainty of the main variable:

$$u = \sqrt{u_{\mathrm{ucor}}^2 + u_{\mathrm{scor}}^2 + u_{\mathrm{tcor}}^2}. \qquad (3.2)$$

In case not all of the three correlation-related components are reported for a variable, Eq. (3.2) is reduced accordingly.

- *_uc_ucor* – uncorrelated uncertainty, i.e. not correlated or correlated at most over a very short distance or time (see resolution of variable)

- *_uc_scor* – spatially correlated uncertainty, correlated within the actual profile. In case of simultaneous flights of more than one RS41 radiosonde on the same rig, this component is also correlated between the profiles of the different sondes.
- *_uc_tcor* – temporally correlated uncertainty for profiles sequentially measured at a site over time (weeks, months, possibly even years) or simultaneously (i.e. multi-GDP flights), also including the (vertical or spatial) correlation within the actual profile.

3.2.2 Corrections and their components

For some parameters, applied total corrections are given in separate columns (*_corr*). If raw data (*_raw*) are available, the correlation between product and raw data can be traced using the correction:

$$x = x_{\text{raw}} + x_{\text{corr}}. \tag{3.3}$$

For some parameters, the total correction is again composed of individual components which are additive:

$$c = c_{\text{p1}} + c_{\text{p2}} + \dots \tag{3.4}$$

3.3 Attributes assigned to variables

The NetCDF files provide attributes associated with each variable, e.g. physical units. Each column is attached a standard (or short) name, a long name, a description (comment), and a physical unit as attributes. Additional attributes are given for internal use. See appendix A.3 for a brief description of the complete set of attributes.

3.4 Usage of files

The GDP file contents are designed to serve as a basis for various potential applications. While the product data is always given at high resolution (1 s or $\sim$5 m s^{-1}, respectively), some applications require adapted (coarser) resolutions. In connection with an appropriate conversion or compression of the column data by the user (averaging or gridding), it should be noted that when combining the measurement uncertainties, their correlation properties should be taken into account. This is explained below in Section 3.4.1.

The GDP files contain further information which may be less relevant for certain applications. Therefore, it may be advantageous for users to optimise the GDP files before use, i.e. to reduce their contents and therefore size. A number of tools for that is presented in Section 3.4.2.

3.4.1 Use of uncertainties in spatial and temporal gridding

For applications that require product data resolutions other than that provided in the files, an adjustment (in time or space) may be appropriate. This can be achieved by a spatial (vertical) or a (long-term) temporal gridding. Spatial gridding (or binning) always refers to a particular measured profile (sounding). Gridding is usually done by applying averaging windows running along the profiles or time series, where the use of different kernel shapes enables appropriate types of weighting.

For example, a regular grid can be defined (e.g. 100 m) for a set of profiles in order to determine monthly means of the vertically averaged data in the grid layers in a second step. Another example is the comparison of radiosonde data with a much lower resolution instrument, e.g. satellites, which can be made after such a gridding using appropriate weighting functions.

Temporal gridding means averaging over a (sufficient) number of soundings from the same location over time, e.g. to calculate a monthly mean. Strictly speaking, the columns of the GDP-files for each radiosounding represent flight trajectories, i.e. not vertical profiles but measurement points distributed in time and three-dimensional space. In that sense they are also (short) time series. Therefore, spatial gridding in a first step enables assignment of averaged times to the gridded values in a next step, as a prerequisite for temporal gridding.

The actual gridding of the particular variables is usually straight forward. In contrast, the determination of the uncertainties of the gridded values need more attention, since the correlation characteristics have to be taken into account.

A rough outline on how uncertainties might be considered within such a gridding procedure is given in the following.

3.4.1.1 Spatial gridding (vertical altitude bins)

For spatial gridding the measured data along the sounding trajectory are considered an instantaneous ‘vertical’ profile representing the atmospheric state for the region above the sounding station which can be condensed into a vertical profile with lower resolution. The spatial mean of a variable ν_s in the actual grid layer in case of a simple ‘boxcar’ averaging (uniform kernel with constant weighting) can be calculated as:

$$\nu_\mathrm{s} = \overline{\nu} = \frac{1}{N}\sum_{j=1}^{N}\nu_j. \tag{3.5}$$

The individual uncorrelated uncertainty components of the data points in the actual layer which form the mean ν_s contribute to the uncertainty of that mean. The uncertainty is again uncorrelated (_uc_ucor):

$$u_\mathrm{ucor}(\nu_\mathrm{s}) = u_\mathrm{ucor}(\overline{\nu}) = \frac{1}{N}\sqrt{\sum_{j=1}^{N}u^2_\mathrm{ucor}(\nu_j)}. \tag{3.6}$$

In addition, the variability of the measurement points ν_j in the actual grid layer contributes to the uncorrelated uncertainty component of the layer mean. The uncertainty $u_\mathrm{var}(\nu_\mathrm{s})$ of the mean due to this variability is estimated in a first approach based on the standard deviation of the mean:

$$u_{\text{var}}(\nu_{\text{s}}) = u_{\text{var}}(\overline{\nu}) = \sqrt{\frac{1}{N(N-1)}\sum_{j=1}^{N}(\nu_j - \overline{\nu})^2}. \tag{3.7}$$

Both components combine to the overall uncorrelated uncertainty:

$$u_{\text{uc}}(\nu_{\text{s}}) = u_{\text{uc}}(\overline{\nu}) = \sqrt{u^2_{\text{ucor}}(\nu_{\text{s}}) + u^2_{\text{var}}(\nu_{\text{s}})}. \tag{3.8}$$

The first component of the *correlated* uncertainty for the actual grid layer $u_{\text{sc}}(\nu_{\text{s}})$ is equivalent to the mean of the spatially correlated uncertainties (*_uc_scor*) (more exactly it is the linear sum of the product of the correlated uncertainties of each data point with the uniform weight $1/N$):

$$u_{\text{sc}}(\nu_{\text{s}}) = u_{\text{sc}}(\overline{\nu}) = \frac{1}{N}\sum_{j=1}^{N} u_{\text{scor}}(\nu_j). \tag{3.9}$$

The second component of the correlated uncertainty for the actual grid layer $u_{\text{tc}}(\nu_{\text{s}})$ is the mean of the temporally correlated uncertainties (*_uc_tcor*):

$$u_{\text{tc}}(\nu_{\text{s}}) = u_{\text{tc}}(\overline{\nu}) = \frac{1}{N}\sum_{j=1}^{N} u_{\text{tcor}}(\nu_j). \tag{3.10}$$

Both formulas are special cases under the assumption of a full degree ($r = 1$) of correlation between adjacent data points.

The combined uncertainty for each grid layer is then the sum of the squares of the two components:

$$u(\nu_{\text{s}}) = u(\overline{\nu}) = \sqrt{u^2_{\text{uc}}(\nu_{\text{s}}) + u^2_{\text{sc}}(\nu_{\text{s}}) + u^2_{\text{tc}}(\nu_{\text{s}})}. \tag{3.11}$$

3.4.1.2 Temporal gridding (time series)

If a temporal gridding is to be carried out (e.g. a monthly mean), it should follow the spatial gridding as second step. That is, a temporal gridding should be applied to vertical profiles which were first converted to the same altitude grid.

The temporal mean ν_{t} of a variable within an altitude layer and over a specific time interval is:

$$\nu_{\text{t}} = \overline{\nu_{\text{s}}} = \frac{1}{N}\sum_{k=1}^{N} \nu_{\text{s},k}. \tag{3.12}$$

The uncorrelated standard uncertainty of the profile data for the average in the actual time interval, $u_{\text{uc}}(\nu_{\text{t}})$, is given by the uncertainty of the mean. In case of a simple unweighted mean this is

$$u_{\text{uc}}(\nu_{\text{t}}) = u_{\text{uc}}(\overline{\nu_{\text{s}}}) = \frac{1}{N}\sqrt{\sum_{k=1}^{N} u^2_{\text{uc}}(\nu_{\text{s},k})}. \tag{3.13}$$

Another uncorrelated uncertainty component of the mean, $u_{\text{var}}(\nu_{\text{t}})$, comes from the variability of the ν_k in the actual averaging time interval. Just as with spatial averaging, this component can be estimated based on the standard deviation:

$$u_{\text{var}}(\nu_{\text{t}}) = u_{\text{var}}(\overline{\nu_{\text{s}}}) = \sqrt{\frac{1}{N(N-1)}\sum_{k=1}^{N}(\nu_{\text{s}} - \overline{\nu_{\text{s},k}})^2}. \tag{3.14}$$

The spatially correlated uncertainty of the spatially gridded values (Eq. 3.9) is considered uncorrelated between profiles at different locations or at the same location at different times. It contributes therefore as a third component to the uncorrelated uncertainty for the averaging time interval:

$$u_{\text{sc}}(\nu_{\text{t}}) = u_{\text{sc}}(\overline{\nu_{\text{s}}}) = \frac{1}{N}\sqrt{\sum_{k=1}^{N} u_{\text{sc}}^2(\nu_{\text{s},k})}. \tag{3.15}$$

The overall uncorrelated uncertainty is the combination of the three components:

$$u_{\text{unc}}(\nu_{\text{t}}) = u_{\text{unc}}(\overline{\nu_{\text{s}}}) = \sqrt{u_{\text{uc}}^2(\nu_{\text{t}}) + u_{\text{var}}^2(\nu_{\text{t}}) + u_{\text{sc}}(\nu_{\text{t}})}. \tag{3.16}$$

The correlated uncertainty for the actual time range $u_{\text{cor}}(\nu_{\text{t}})$ is given by

$$u_{\text{cor}}(\nu_{\text{t}}) = u_{\text{cor}}(\overline{\nu_{\text{s}}}) = \frac{1}{N}\sum_{k=1}^{N} u_{\text{tc}}(\nu_{\text{s},k}), \tag{3.17}$$

which again is equivalent to the mean of the temporally correlated uncertainties.

The combined total uncertainty for each time bin follows from the sum of squares of the constituting components:

$$u(\nu_{\text{t}}) = u(\overline{\nu_{\text{s}}}) = \sqrt{u_{\text{unc}}^2(\nu_{\text{t}}) + u_{\text{cor}}^2(\nu_{\text{t}})}. \tag{3.18}$$

3.4.2 Optimising GDP files

The GDP files are formatted as standard NetCDF files, which enables the use of common NetCDF tools for manipulation. There are various tools freely available for working with NetCDF files. For the command line, for example, the well-known tools ‘ncdump’, ‘nccopy’ and ‘ncgen’ from Unidata are useful (see e.g. https://www.unidata.ucar.edu/software/netcdf/docs/netcdf_working_with_netcdf_files.html).

Tool ‘ncdump’ To obtain an overview of the contents of a NetCDF file, the tool ‘ncdump’ can be used. It helps to extract the defined groups, variables and attributes as a structured text file. If desired, the entire content (the actual data) can also be extracted.

```
# Usage of ncdump
#
# print header (definitions and attributes) to a text file (no data)
> ncdump -h input.nc > output.dump
#
```

```
# print full content to a text file
> ncdump input.nc > output.dump
#
# include also special ‘virtual’ attributes
> ncdump -s input.nc > output.dump
```

The created ‘dump’ files are so-called CDL formatted text files, which could be used to generate new NetCDF files using the tool ‘ncgen’.

Tool ‘nccopy’ This tool can be used to create a copy of an already existing NetCDF file. The copy can be manipulated, e.g. for filtering variables, changing compression, ...

It is useful e.g. to reduce the size of large NetCDF files. The following example shows how a list of selected variables can be copied into a new NetCDF file.

```
# Usage of nccopy
#
# copy only some required variable to a new (much smaller) nc file
> nccopy -V time,alt,lon,lat,press,temp,rh,wdir,wspeed input.nc output.nc
```

Tool ‘ncgen’ This tool can be used to generate a NetCDF file from a CDL text file.

For example, the complete CDL text files generated by ‘ncdump’ can be used here as input files.

```
# usage of ncgen
#
# create a nc file based on given CDL input file
> ncgen -4 -b -o output.nc input.cdl
```

3.4.3 Reading GDP files

The GDP files for the RS41 radiosonde measurements may reach a size of 5 MiB. Due to the use of internal compression, the actual extracted data table requires considerably more memory (up to a factor of 100). This should be taken into account when processing the data. With the standard settings of common tools to read NetCDF, the files are often completely loaded into the memory (RAM), which can lead to memory problems when working simultaneously with a larger number of files. For this reason it is recommended to take advantage of the read-in routines and memory management capabilities of the programming or scripting language used.

In general not all contents (variables, auxiliary variables, uncertainties, corrections, raw data, flags, etc.) of the GDP files are needed at once in practice. Therefore, functionalities and capabilities to load only the variables of interest for the actual application should be extracted, also in favour of processing time.

4 Distribution and naming of files

4.1 Distribution

GRUAN data files will be made available at end of 2022 on a FTP server at the National Centers for Environmental Information (NCEI, ftp://ftp.ncei.noaa.gov/pub/data/gruan). In the directory structure, the GRUAN data files are organised according to product type, version, site, and year:

- processing/ – Processed GRUAN data
- level<level number>/ – Data level, here level2 (see below)
- <product code>/ - Product code, here RS41-GDP (RS41 GRUAN Data Product)
- version-<version number>/ – Product version, here e.g. version-001
- <site code>/ – Site code, e.g. SNG for Singapore (see https://www.gruan.org/network/sites)
- <year>/ – Year of measurement, e.g. 2019.

Note: The access path to RS41-GDP.1 at NCEI will be activated soon (end of 2022).

Access to the entire RS41-GDP.1 archive is alternatively provided via the GRUAN website after login at https://www.gruan.org/data/file-archive/rs41-gdp1-at-lc.

Note: Please consider the GRUAN data policy for use of the data, especially for publication in scientific articles (see https://www.gruan.org/data/data-policy).

4.2 File naming

The file name convention for the product files is as follows:

Convention : SSS-MM-NN_L_<product name>_VVV_<date>_<id>.nc

Example : SNG-RS-01_2_RS41-GDP_001_20191018T120000_1-000-001.nc

- *SSS* – The three letter abbreviation for the station name, e.g. SNG for Singapore (see GRUAN website at https://www.gruan.org/network/sites)
- *MM* – A two letter code for the measurement system, e.g. RS for Radiosonde
- *NN* – The number of the measurement system, e.g. 01 (first radiosonde system at Singapore)
- *L* – The data level coded as follows:

  - *0* – Original raw data (files)
  - *1* – Pre-processed raw data
  - *2* – Processed data from one measuring system
  - *3* – Best composite of several measuring systems

- *<product name>* – The name of the GRUAN Data Product (arbitrary length), e.g. RS41-GDP
- *VVV* – Version of data product, e.g. 001
- *<date>* – Scheduled start time [UTC] of the flight in the format YYYYMMDDThhmmss
- *<id>* – Identification of the measuring event in the format B-PPP-VVV where:

  - *B* – is the number of the balloon (one or a specific number in case of simultaneous launches of more than one balloon), e.g. 1
  - *PPP* – [internal use only] the number of reported item which is the related telemetry software/hardware, e.g. 000
  - *VVV* – the version of the data, e.g. 001

5 Known issues

Note: Bugs, errors and other issues noticed after the release of the version of the GRUAN Data Product described in this document are listed here and, if possible, advice is given how to deal with them.

Some bugs and feature requests have been identified during the use of data files of RS41 GDP (version 1):

- Occasional gaps (NaN values) in columns of temporally correlated uncertainty of altitude parameters (see Section 5.1)

Bugs and errors will be fixed, and the above listed feature requests will be included/improved in an updated version of the GRUAN data product.

Note: Please contact the GRUAN Lead Centre (gruan.lc@dwd.de) in case of any bugs or anomalies identified in the NetCDF files of this GRUAN Data Product. Suggestions or requests for changes regarding existing content or the addition of further contents to be included in an updated version of the GDP are welcome.

5.1 Occasionally undefined temporal uncertainty "tcor" for altitude parameters

The evaluation of the "tcor" uncertainty component of any altitude parameter (amsl, gph, wgs84) includes information given with the VDOP parameter (stored in the file as "vdop"). Occasionally, certain points in the profile with valid numbers for the "lat", "lon" and "alt" variables have NaN values for "vdop". This is unexpected because a DOP analysis is mandatory for the output of valid GPS position data. As a result, also the values for the "tcor" uncertainty component are set NaN. The actual GDP version does not check for this, a fix will be implemented in the next update.

The following workaround is proposed: On occurrence of missing valid VDOP values, i.e. NaNs set also for the "tcor" uncertainty component, the latter should be replaced with the maximum "tcor" uncertainty of the data set which means a conservative estimate.

The impact of this fix on the quality of the data set is limited. Essentially, the replacement of the NaNs with a conservatively estimated uncertainty would assign somewhat lower weights to the respective data included in potential analyses or applications. However, NaNs and thus data gaps are avoided in favour of successful processing.

Appendix

A Details of GDP file contents

A.1 Global attributes (metadata)

Note: Uncertainty estimates for variables and attributes in the product files are consistently specified at the 2-σ confidence level (coverage factor $k = 2$).

Table A.1: List of global attributes included in NetCDF file

Name	Example	Description
Attributes of CF convention		
Conventions	CF-1.7	List of applied conventions
title	RS41 GRUAN Data Product Version 1 (RS41-GDP.1)	Title of data product
institution	GRUAN Lead Centre (DWD, MOL-RAO)	Institution of producer of file / data product
source	Vaisala Radiosonde RS41	Source of measurement (instrument)
history	2022-01-17T10:54:09.765Z Gruan radiosonde data processing based on radiosonde raw data file (MWX) using GRUAN Data Processing System v1.5.2 (2021-09-02)	History of file changes (starts with creation)
references	See the GRUAN documentation on RS41.	List of relevant references
comment	This is a RELEASE version of a data product file.	Specific comment to the version of data product
Group g.Product		
Id	752292	ID in GMDB table 'ProductItem'
FullKey	RS41-GDP.1	Full unique key of data product version

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
Key	RS41-GDP	Key of data product
Name	RS41 GRUAN Data Product Version 1	Name of data product
Version	1	Version of data product
Title	RS41 GRUAN Data Product Version 1 (RS41-GDP.1)	Title of data product
Level	2	Data level of this data product (see Section 4.2)
LevelDescription	Processed data	Short description of the data level
Revision	1	Revision of this product item
History	2022-01-17T10:54:09.765Z Gruan radiosonde data processing based on radiosonde raw data file (MWX) using GRUAN Data Processing System v1.5.2 (2021-09-02)	History of file changes (starts with creation)
References	See the GRUAN documentation on RS41.	A reference related to the data product
Producer	GRUAN Lead Centre (DWD, MOL-RAO)	Producer (organisation)
ProducerTool	GRUAN Data Processing System v1.5.2 (2021-09-02)	Name and version of producer tool
ProducerContact	data@gruan.org	Email address to contact the producer
TimeResolution	1.0 s	Time resolution of data
Status	COMPLETED	Status of processing
StatusDescription	Processing is completed without errors.	Description of processing status
ProcessingSteps	{...}	List of all processing steps incl. used modules and versions
PrelaunchIncluded	no	Is pre-launch data included? (always no) [yes|no]
AscentIncluded	yes	Is ascent data included? (always yes) [yes|no]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
DescentIncluded	no	Is descent data included? (always no) [yes|no]
PostflightIncluded	no	Is post-flight data included? (always no) [yes|no]
Group g.File		
Type	GNC-DATA	Specific file type/format
TypeVersion	1.0	Version of file type/format
TypeName	GRUAN NetCDF Radiosonde Data File	Name of file type/format
Timestamp	2022-01-17T10:54:09.765Z	Timestamp of file creation [ISO 8601]
Comment	This is a RELEASE version of a data product file.	A comment to this file
Group g.Site		
Id	1	ID in GMDB table 'Site'
Key	LIN	Unique key of GRUAN site (defined in GMDB table 'Site')
Name	Lindenberg	Unique name of GRUAN site (defined in GMDB table 'Site')
Institution	Lindenberg Meteorological Observatory (MOL), German Weather Service (DWD)	Organisation operating the GRUAN site
Country	Germany	Country where the GRUAN site is located ('nil' in case of no info)
District	Brandenburg	First order district where the GRUAN site is located ('nil' in case of no info)
TimeZone	UTC+01:00	Time zone of GRUAN site location ('nil' in case of no info)
Longitude	14.12 °E	Longitude of GRUAN site [°E]
Latitude	52.21 °N	Latitude of GRUAN site [°N]
Altitude	98 m	Altitude of GRUAN site [m]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
IsCooperation	no	Is the GRUAN site in cooperation with other organisations? [yes|no]
IsDistributed	no	Does the GRUAN site consist of facilities spread over several locations? [yes|no]
IsMobile	no	Is the site mobile (e.g. a ship)? [yes|no]
Climate	moderate mid-latitude climate at the transition between marine and continental influences	Short description of the climate at site ('nil' in case of no info)
EnvironmentType	rural	General type of environment around the site ('nil' in case of no info)
LandUse	grassland / cropland 60 %, pine forest 30 %, open water 5 %, settlements 5 %	Land use around the site ('nil' in case of no info)
Population	nil	Population around the site ('nil' in case of no info)
SurfaceRoughness	nil	Roughness of surface around the site ('nil' in case of no info)
Topography	small hills	Topography around the site

Group g.MeasurementSystem

Id	126	ID in GMDB table 'MeasuringSystem'
Key	NYA-RS-01	Unique key of measurement system (defined in GMDB table 'MeasuringSystem')
Name	Ny-Aalesund Radiosonde Launch Site	Name of measurement system
Type	Radiosonde	Type of measurement system (defined in GMDB table 'MeasuringSystemType')
Longitude	11.92271 °E	Longitude of measurement system [°E]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
Latitude	78.92301 °N	Latitude of measurement system [°N]
Altitude	15.67 m	Altitude of measurement system [m]
OperatedBy	Research Unit Potsdam (AWI-POTSDAM), Alfred Wegener Institute for Polar and Marine Research (AWI), Helmholtz Association (HELMHOLTZ)	Organisation operating the measurement system
WmoCode	01004	WMO code related to the measurement system (or site)
WmoName	NY-ALESUND II	WMO name related to the measurement system (or site)
WmoRegion	IV - Europe	Number and name of WMO region
WigosSid	0-20001-0-10393	WIGOS site ID
Group g.MeasurementSetup		
Id	217	ID in specific GMDB table ‘Measuring’
Key	ROUTINE2	Key of measurement setup (defined in GMDB table ‘Measuring’)
Name	Ny-Alesund Routine Sounding (RS41 with MW41)	Name of measurement setup
Type	Radiosounding	Type of measurement setup (defined in GMDB table ‘MeasuringType’)
Group g.SurfaceObs		
Pressure	1013.7065 hPa	Pressure at launch site [hPa]
Temperature	268.6 K	Temperature at launch site [K]
RelativeHumidity	60.0 %	Relative humidity at launch site [%RH]
WindSpeed	1.0 m/s	Wind speed at launch site [$\mathrm{m\,s^{-1}}$]
WindDirection	158.0 °	Wind direction at launch site [°]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
SynopClouds	782//	WMO synop coded clouds
CloudsText	St fra Cu med Sc str op	Textual coded clouds
Comment	after drizzle	Free comment to weather conditions at launch site
Group g.Measurement		
Id	151081	ID in GMDB table 'MeasuringEvent'
InternalKey	LG2022_10_Mmrs	Site internal key set for this measurement
LastChange	2022-01-24T15:08:50.600Z	Timestamp of last change in GMDB [ISO 8601]
StandardTime	2022-01-24T12:00:00.000Z	Scheduled or standard date & time of measurement [ISO 8601]
StartTime	2022-01-24T11:03:37.171Z	Exact start time (launch time) of measurement; zero point of the *time* variable [ISO 8601]
StartTimeManufacturer	2022-01-24T11:03:37.187Z	Original start time (launch time) detected or set by manufacturer [ISO 8601]
StartTimeDiff	1.250 s	Difference between used and original start time (launch time) [second]
StartTimeSource	pressure	Source/variable for start time estimate. Options are: altitude, pressure, manual, pre-defined
BalloonNumber	1	Number of balloon at 'StandardTime' of measurement (default is 1)
Version	1	Stored version of actual measurement (default is 1)

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
Equipment	{DC-MW41/DPS, DC-RI41-B/CheckTool, POT100-SPFE1/CheckTool, LAUNCH-SHELTER/WeatherStation, TA600/Balloon, 0-NONE/Parachute, UW-V30-4/Unwinder, SOLO/Rig, RS41-SGP/Sonde}	Ground equipment and rig composition (full list; keys defined in GMDB tables 'Instrument' and 'InstrumentType')
Balloon	TA600	Key of instrument model of balloon (defined in GMDB table 'Instrument')
FillingWeight	1000 g	Balloon lift (mass which is lifted by filled balloon)
Parachute	TP-103	Key of instrument model of parachute (defined in GMDB table 'Instrument')
Unwinder	UW-V55-4	Key of instrument model of unwinder (defined in GMDB table 'Instrument')
FullStringLength	55.0 m	Overall string length between balloon and main sonde [m]
Rig	SOLO	Key of instrument model of rig (defined in GMDB table 'Instrument')
Payload	200.0 g	Total weight of payload (excluding the balloon) [g]
GrossWeight	850.0 g	Total weight of balloon chain [g]
TimeOfDay	twilight	Term to describe the time of day of flight (daytime, nighttime, twilight)
DaytimeCount	2097	Count of daytime data points [–]
NighttimeCount	4846	Count of nighttime data points (all twilight phases are part of night) [–]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
TwilightCount	1626	Count of relevant twilight data points (only civil twilight is included) [–]
CivilTwilightCount	1626	Count of civil twilight data points [–]
NauticalTwilight-Count	1847	Count of nautical twilight data points [–]
Astronomical-TwilightCount	1373	Count of astronomical twilight data points [–]
PartlyVisibleSun-Count	91	Count of data points with the sun partly covered by the horizon (is handled as part of daytime) [–]
PositionRangeLat-North	1.38746 °N	Northernmost degree of latitude of the flight trajectory [°N]
PositionRangeLat-South	1.32548 °N	Southernmost degree of latitude of the flight trajectory [°N]
PositionRangeLon-East	103.94989 °E	Easternmost degree of longitude of the flight trajectory [°E]
PositionRangeLon-West	103.57969 °E	Westernmost degree of longitude of the flight trajectory °E]
LocGpsAltitudeRef	24.82 m	Altitude (MSL) of local GPS/GNSS antenna at ground [m]
LocGpsAltitude-Mean	56.90 m	Mean altitude (WGS84) measured by local GPS/GNSS antenna at ground during full measurement cycle (incl. pre-flight and flight) [m]
LocGpsAltitude-Geoid	35.98 m	Geoid altitude (MSL) at position of measurement system [m]
LocGpsAltitude-Offset	32.08 m	Difference between measured WGS84 altitude and reference altitude [m]
LocGpsAltitude-Residual	-3.90 m	Difference between offset and geoid [m]
BurstpointAltitude	36022.8 m	Altitude (MSL) of burst point [m]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
BurstpointPressure	5.8 hPa	Pressure of burst point [hPa]
BurstpointGeopot-Height	35871.4 gpm	Geopotential height of burst point [m]
BurstpointTime	7355.0 s	Time after launch of burst point. It is equal to length of measurement. [s]
BurstpointDatetime	2017-05-25T10:51:50.250Z	Timestamp of burst point [ISO 8601]
BurstpointSource	altitude	Source of burst point. Following options are possible: altitude, pressure, manual, pre-defined
PrecipitableWater-Column	4.68 kg/m²	Precipitable water column (measured by main sonde) [$\mathrm{kg\,m^{-2}}$]
PrecipitableWater-ColumnUc	0.225 kg/m²	Uncertainty of precipitable water column (measured by main sonde) [$\mathrm{kg\,m^{-2}}$]
TropopauseDetected	yes	Was the algorithm able to detect the tropopause? [yes|no]
TropopauseGeopot-Height	7912.2 gpm	Geopotential height of detected tropopause [m]
TropopauseAltitude	7939.0 m	Altitude (MSL) of detected tropopause [m]
Tropopause-Temperature	219.7 K	Temperature at detected tropopause [K]
TropopausePressure	337.1 hPa	Pressure at detected tropopause [hPa]
TropopausePot-Temperature	299.8 K	Potential temperature at detected tropopause [K]
TropopauseTime	1717.0 s	Time after launch of detected tropopause [s]
Tropopause-Datetime	2017-05-25T09:17:52.250Z	Timestamp of detected tropopause [ISO 8601]
Group g.MainSonde		
SerialNumber	M4140267	Serial number of main sonde

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
Model	RS41-SGP	Key of instrument model of main sonde (defined in GMDB table ‘Instrument’)
Name	Vaisala Radiosonde RS41-SGP	Full name of main sonde
ModelFamily	RS41	Family of instrument model of main sonde
Type	RS	Short key of instrument model of main sonde
Manufacturer	Vaisala (VAISALA)	Name and key of manufacturer of main sonde
Group g.TelemetrySonde		
SerialNumber	M4140267	Serial number of telemetry sonde
Model	RS41-SGP	Key of instrument model of telemetry sonde (defined in GMDB table ‘Instrument’)
Name	Vaisala Radiosonde RS41-SGP	Full name of telemetry sonde
ModelFamily	RS41	Family of instrument model of telemetry sonde
Type	RS	Short key of instrument type of telemetry sonde (defined in GMDB table ‘InstrumentType’)
Manufacturer	Vaisala (VAISALA)	Name and key of manufacturer of telemetry sonde
XData	yes	XData sensors (instruments) are linked at this telemetry sonde? [yes|no]
XDataSensors	OIF411 CFH	List of linked XData sensors/instruments (‘nil’ in case of no sensors)
Group g.GroundSystem		
Id	518084	ID in GMDB table ‘MeasuringEventAssembly’

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
Model	DC-MW41	Instrument model of ground system (defined in GMDB table 'Instrument')
Name	Vaisala DigiCORA Sounding System MW41	Full name of ground system
ModelFamily	DIGI-CORA	Family of instrument model of ground system
Type	PROCESSING	Short key of instrument type of ground system (defined in GMDB table 'InstrumentType')
InternalKey	MW41-SYSTEMS	Internal key of ground system (defined in GMDB table 'MeasuringSystemAssembly')
SerialNumber	L29403	Serial number of a main part of the ground system (e.g. SPS)
Manufacturer	Vaisala (VAISALA)	Name and key of manufacturer of ground system
Software	MW41 v2.6.0	Software version used with ground system
Group g.GroundCheck – Related information to all performed ground checks		
List	{RI41, SHC, SHELTER}	A list of all performed (and available) ground checks
Sub-group g.GroundCheck.RI41 – Related information to the mandatory ground check using the initialisation tool RI41 or RI41-B		
CheckDate	2017-05-25T08:17:05.816Z	Timestamp of start of check period (may alternatively describe the end of check period) [ISO 8601]
CheckLength	NaN s	Length (duration) of check period in seconds [s]
CheckType	GC-RI41	Unique key of check type (always GC-RI41; defined in GMDB table 'InstrumentCheck')
RefPressure	1013.58 hPa	Reference pressure [hPa]
SensorPressure	1013.0041 hPa	Sensor pressure [hPa]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	**Example**	**Description**
OffsetPressure	-0.58 hPa	Detected offset of RS-sensor pressure related to the reference (Used to adjust the calibration of pressure sensor) [hPa]
PressureQc	good	QC pressure flag [no|bad|fair|good]
Altitude	17.43 m	Altitude (above MSL) of ground check tool [m]
OffsetAltitudeRef-Pressure	25.00 m	Altitude offset of reference pressure [m]
Sub-group g.GroundCheck.SHC – Related information to the recommended ground check using a Standard Humidity Chamber (SHC)		
CheckDate	2017-05-25T08:26:06.250Z	Timestamp of start of check period [ISO 8601]
CheckLength	338.0 s	Length (duration) of check period in seconds [s]
CheckType	GC-POT100	Unique key of check type (always GC-PTU100 or GC-SHC; defined in GMDB table 'InstrumentCheck')
Analysed	yes	Could the check period be found? [yes|no]
AutoDetected	yes	Could the check period be automatically detected? [yes|no]
RefRelative-Humidity	100.00 %	Mean reference relative humidity [%RH]
RefRelative-HumidityStdDev	NaN %	Standard deviation of reference relative humidity during check period [%RH]
RefRelative-HumidityUc	0.031 % (k=2)	Uncertainty of mean of reference relative humidity [%RH]
RefRelative-HumidityUcFinal	0.501 % (k=2)	Uncertainty of mean of reference relative humidity including calibration and environment [%RH]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
RefRelative-HumidityCount	-1	The count of reference relative humidity values which was available in check period. [–]
RefRelative-HumidityLinSlope	NaN %	Slope of a relative humidity linear trend during check period [%RH]
RefTemperature	297.03 K	Mean reference temperature [K]
RefTemperature-StdDev	0.031 K	Standard deviation of reference temperature values during check period [K]
RefTemperatureUc	0.002 K (k=2)	Uncertainty of mean of reference temperature [K]
RefTemperature-UcFinal	0.050 K (k=2)	Uncertainty of mean of reference temperature including calibration and environment [K]
RefTemperature-Count	175	The count of reference temperature values which was available in check period. [–]
RefTemperature-LinSlope	-0.00015 K	Slope of a temperature linear trend during check period [K]
RefPressure	1013.53 hPa	Mean reference pressure [hPa]
RefPressureStdDev	0.002 hPa	Standard deviation of reference pressure during check period [hPa]
RefPressureUc	0.000 hPa (k=2)	Uncertainty of mean of reference pressure [hPa]
RefPressureUFinal	0.100 hPa (k=2)	Uncertainty of mean of reference pressure including calibration and environment [hPa]
RefPressureCount	339	The count of reference pressure values which was available in check period. [–]
RefPressureLin-Slope	-0.00007 hPa	Slope of a pressure linear trend during check period [hPa]
SensorRelative-Humidity	100.79 %	Mean sensor relative humidity [%RH]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
SensorRelative-HumidityStdDev	0.072 %	Standard deviation of sensor relative humidity during check period [%RH]
SensorRelative-HumidityUc	0.004 % (k=2)	Uncertainty of mean of sensor relative humidity [%RH]
SensorRelative-HumidityUcFinal	1.500 % (k=2)	Uncertainty of mean of sensor relative humidity including calibration and environment [%RH]
SensorRelative-HumidityCount	339	The count of sensor relative humidity values which was available in check period.
SensorRelative-HumidityLinSlope	-0.00011 %	Slope of a relative humidity linear trend during check period [%RH]
SensorTemperature	297.15 K	Mean sensor temperature [K]
SensorTemperature-StdDev	0.011 K	Standard deviation of sensor temperature during check period [K]
SensorTemperature-Uc	0.001 K (k=2)	Uncertainty of mean of sensor temperature [K]
SensorTemperature-UcFinal	0.100 K (k=2)	Uncertainty of mean of sensor temperature including calibration and environment [K]
SensorTemperature-Count	339	The count of sensor temperature values which was available in check period. [–]
SensorTemperature-LinSlope	-0.00002 K	Slope of a temperature linear trend during check period [K]
SensorPressure	1013.53 hPa	Mean sensor pressure [hPa]
SensorPressure-StdDev	0.056 hPa	Standard deviation of sensor pressure during check period [hPa]
SensorPressureUc	0.003 hPa (k=2)	Uncertainty of mean of sensor pressure [hPa]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
SensorPressure-UcFinal	0.200 hPa (k=2)	Uncertainty of mean of sensor pressure including calibration and environment [hPa]
SensorPressure-Count	339	The count of sensor pressure values which was available in check period. [–]
SensorPressure-LinSlope	-0.00010 hPa	Slope of a pressure linear trend during check period [hPa]
OffsetPressure	0.00 hPa	The detected offset of RS-sensor pressure related to the reference [hPa]
OffsetPressureUc	0.224 hPa (k=2)	Uncertainty of detected offset of RS-sensor pressure related to the reference [hPa]
OffsetTemperature	0.12 K	The detected offset of RS-sensor temperature related to the reference
OffsetTemperature-Uc	0.112 K (k=2)	Uncertainty of detected offset of RS-sensor temperature related to the reference [K]
OffsetRelative-Humidity	0.79 %	The detected offset of RS-sensor relative humidity related to the reference [%RH]
OffsetRelative-HumidityUc	1.581 % (k=2)	Uncertainty of detected offset of RS-sensor relative humidity related to the reference [%RH]
RelativeHumidity-SensorsDiff	NaN %	[not used in case of RS41] [%RH]
PressureQc	good	QC pressure flag [no|bad|fair|good]
TemperatureQc	bad	QC temperature flag [no|bad|fair|good]
RelativeHumidityQc	good	QC relative humidity flag [no|bad|fair|good]
RelativeHumidity-SensorsQc	no	[not used in case of RS41] [no|bad|fair|good]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
Altitude	17.43 m	Altitude (above MSL) of ground check tool [m]
OffsetAltitudeRef-Pressure	0.70 m	Altitude offset of reference pressure [m]
Sub-group g.GroundCheck.SHELTER – Related information to an additional ground check using a shelter or weather hut outdoors (list of attributes are equal to ground check with a SHC)		
CheckDate	2017-05-25T08:34:31.250Z	Timestamp of start of check period [ISO 8601]
CheckLength	288.0 s	Length (duration) of check period in seconds [s]
CheckType	GC-SHELTER	Unique key of check type (always GC-SHELTER; defined in GMDB table 'InstrumentCheck')
Analysed	yes	Could the check period be found? [yes|no]
AutoDetected	no	Could the check period be automatically detected? [yes|no]
RefRelative-Humidity	59.94 %	Mean reference relative humidity [%RH]
RefRelative-HumidityStdDev	1.042 %	Standard deviation of reference relative humidity during check period [%RH]
RefRelative-HumidityUc	0.061 % (k=2)	Uncertainty of mean of reference relative humidity [%RH]
RefRelative-HumidityUcFinal	1.002 % (k=2)	Uncertainty of mean of reference relative humidity including calibration and environment [%RH]
RefRelative-HumidityCount	289	The count of reference relative humidity values which was available in check period. [–]
RefRelative-HumidityLinSlope	-0.00216 %	Slope of a relative humidity linear trend during check period [%RH]
RefTemperature	268.45 K	Mean reference temperature [K]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
RefTemperature-StdDev	0.037 K	Standard deviation of reference temperature values during check period [K]
RefTemperatureUc	0.002 K (k=2)	Uncertainty of mean of reference temperature [K]
RefTemperature-UcFinal	0.150 K (k=2)	Uncertainty of mean of reference temperature including calibration and environment [K]
RefTemperature-Count	284	The count of reference temperature values which was available in check period.
RefTemperature-LinSlope	0.00029 K	Slope of a temperature linear trend during check period [K]
RefPressure	1013.84 hPa	Mean reference pressure [hPa]
RefPressureStdDev	0.007 hPa	Standard deviation of reference pressure during check period [hPa]
RefPressureUc	0.000 hPa (k=2)	Uncertainty of mean of reference pressure [hPa]
RefPressureUcFinal	0.100 hPa (k=2)	Uncertainty of mean of reference pressure including calibration and environment [hPa]
RefPressureCount	289	The count of reference pressure values which was available in check period. [–]
RefPressureLin-Slope	-0.00021 hPa	Slope of a pressure linear trend during check period [hPa]
SensorRelative-Humidity	64.92 %	Mean sensor relative humidity [%RH]
SensorRelative-HumidityStdDev	2.331 %	Standard deviation of sensor relative humidity during check period [%RH]
SensorRelative-HumidityUc	0.137 % (k=2)	Uncertainty of mean of sensor relative humidity [%RH]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	Example	Description
SensorRelative-HumidityUcFinal	1.506 % (k=2)	Uncertainty of mean of sensor relative humidity including calibration and environment [%RH]
SensorRelative-HumidityCount	289	The count of sensor relative humidity values which was available in check period. [–]
SensorRelative-HumidityLinSlope	0.00343 %	Slope of a relative humidity linear trend during check period [%RH]
SensorTemperature	268.40 K	Mean sensor temperature [K]
SensorTemperature-StdDev	0.066 K	Standard deviation of sensor temperature during check period [K]
SensorTemperature-Uc	0.004 K (k=2)	Uncertainty of mean of sensor temperature [K]
SensorTemperature-UcFinal	0.100 K (k=2)	Uncertainty of mean of sensor temperature including calibration and environment [K]
SensorTemperature-Count	289	The count of sensor temperature values which was available in check period.
SensorTemperature-LinSlope	0.00034 K	Slope of a temperature linear trend during check period [K]
SensorPressure	1013.41 hPa	Mean sensor pressure [hPa]
SensorPressure-StdDev	0.050 hPa	Standard deviation of sensor pressure during check period [hPa]
SensorPressureUc	0.003 hPa (k=2)	Uncertainty of mean of sensor pressure [hPa]
SensorPressure-UcFinal	0.200 hPa (k=2)	Uncertainty of mean of sensor temperature including calibration and environment [hPa]
SensorPressure-Count	289	The count of sensor pressure values which was available in check period. [–]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Name	**Example**	**Description**
SensorPressure-LinSlope	0.00059 hPa	Slope of a pressure linear trend during check period [hPa]
OffsetPressure	-0.43 hPa	The detected offset of RS-sensor pressure related to the reference [hPa]
OffsetPressureUc	0.224 hPa (k=2)	Uncertainty of detected offset of RS-sensor pressure related to the reference [hPa]
OffsetTemperature	-0.06 K	The detected offset of RS-sensor temperature related to the reference [K]
OffsetTemperature-Uc	0.180 K (k=2)	Uncertainty of detected offset of RS-sensor temperature related to the reference [K]
OffsetRelative-Humidity	4.98 %	The detected offset of RS-sensor relative humidity related to the reference [%RH]
OffsetRelative-HumidityUc	1.809 % (k=2)	Uncertainty of detected offset of RS-sensor relative humidity related to the reference [%RH]
PressureQc	bad	QC pressure flag [no|bad|fair|good]
TemperatureQc	good	QC temperature flag [no|bad|fair|good]
RelativeHumidityQc	bad	QC relative humidity flag [no|bad|fair|good]
RelativeHumidity-SensorsQc	no	[not used in case of RS41] [no|bad|fair|good]
Altitude	17.44 m	Altitude (above MSL) of ground check tool [m]
OffsetAltitudeRef-Pressure	-1.72 m	Altitude offset of reference pressure [m]

A.2 Variables (data)

Note: Uncertainty estimates for variables and attributes in the product files are consistently specified at the 2-σ confidence level (coverage factor $k = 2$).

Table A.2: List of all variables included in NetCDF file

Name	Type	Unit	Description
Time			
time	float (F8.1)	s	Time in seconds since launch time
Altitude			
alt	float (F8.1)	m	Default altitude; Geopotential height is chosen as default altitude; equal to alt_gph
alt_uc	float (F8.2)	m	Uncertainty of default altitude; Geopotential height is chosen as default altitude; equal to alt_gph_uc; (k=2)
alt_amsl	float (F8.1)	m	Altitude above mean sea level
alt_amsl_uc	float (F11.2)	m	Uncertainty of altitude above mean sea level; (k=2)
alt_amsl_uc_ucor	float (F17.2)	m	Uncorrelated part of uncertainty of altitude above mean sea level; (k=2)
alt_amsl_uc_tcor	float (F17.2)	m	Temporally correlated part of uncertainty of altitude above mean sea level; (k=2)
alt_wgs84	float (F9.1)	m	Altitude WGS84
alt_wgs84_uc	float (F12.2)	m	Uncertainty of altitude WGS84; (k=2)
alt_wgs84_uc_ucor	float (F17.2)	m	Uncorrelated part of uncertainty of altitude WGS84; (k=2)
alt_wgs84_uc_tcor	float (F17.2)	m	Temporally correlated part of uncertainty of altitude WGS84; (k=2)
alt_gph	float (F8.1)	m	Geopotential height
alt_gph_uc	float (F10.2)	m	Uncertainty of geopotential height; (k=2)
alt_gph_uc_ucor	float (F15.2)	m	Uncorrelated part of uncertainty of geopotential height; (k=2)
alt_gph_uc_tcor	float (F15.2)	m	Temporally correlated part of uncertainty of geopotential height; (k=2)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
Position			
lon	double (F10.5)	°E	Longitude
lon_uc	double (F11.6)	°E	Uncertainty of longitude; (k=2)
lon_uc_ucor	double (F11.6)	°E	Uncorrelated part of uncertainty of longitude; (k=2)
lon_uc_tcor	double (F11.6)	°E	Temporally correlated part of uncertainty of longitude; (k=2)
lat	double (F10.5)	°N	Latitude
lat_uc	double (F11.6)	°N	Uncertainty of latitude; (k=2)
lat_uc_ucor	double (F11.6)	°N	Uncorrelated part of uncertainty of latitude; (k=2)
lat_uc_tcor	double (F11.6)	°N	Temporally correlated part of uncertainty of latitude; (k=2)
Pressure			
press	float (F8.3)	hPa	Default pressure; GNSS pressure is chosen as default pressure; equal to press_gnss
press_uc	float (F8.4)	hPa	Uncertainty of default pressure; GNSS pressure is chosen as default pressure; equal to press_gnss_uc; (k=2)
press_sens	float (F10.3)	hPa	Sensor pressure; Barometric air pressure using silicon sensor
press_sens_uc	float (F13.4)	hPa	Total uncertainty of barometric air pressure using silicon sensor; (k=2)
press_sens_uc_ucor	float (F17.4)	hPa	Uncorrelated part of uncertainty of barometric air pressure using silicon sensor; (k=2)
press_sens_uc_tcor	float (F17.4)	hPa	Temporally correlated part of uncertainty of barometric air pressure using silicon sensor; (k=2)
press_sens_uc_tcor_cal	float (F22.4)	hPa	Calibration related part of temporally correlated uncertainty of barometric air pressure using silicon sensor; (k=2)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
press_sens_uc_ucor_sm	float (F21.4)	hPa	Smoothing related part of uncorrelated uncertainty of barometric air pressure using silicon sensor; (k=2)
press_sens_corr	float (F15.4)	hPa	Correction of barometric air pressure using silicon sensor
press_sens_corr_cal	float (F19.4)	hPa	Part of correction of barometric air pressure using silicon sensor which is related to calibration
press_sens_corr_sm	float (F18.4)	hPa	Part of correction of barometric air pressure using silicon sensor which is related to smoothing
press_gnss	float (F10.3)	hPa	GNSS pressure; Barometric air pressure derived from GNSS-altitude
press_gnss_uc	float (F10.4)	hPa	Uncertainty of barometric air pressure derived from GNSS-altitude; (k=2)
press_gnss_uc_ucor	float (F18.4)	hPa	Uncorrelated part of uncertainty of barometric air pressure derived from GNSS-altitude; (k=2)
press_gnss_uc_tcor	float (F18.4)	hPa	Temporally correlated part of uncertainty of barometric air pressure derived from GNSS-altitude; (k=2)
Temperature			
temp	float (F6.2)	K	Temperature
temp_uc	float (F7.3)	K	Uncertainty of temperature; (k=2)
temp_uc_ucor	float (F12.3)	K	Uncorrelated part of uncertainty of temperature; (k=2)
temp_uc_scor	float (F12.3)	K	Spatially correlated part of uncertainty of temperature; (k=2)
temp_uc_tcor	float (F12.3)	K	Temporally correlated part of uncertainty of temperature; (k=2)
temp_uc_rad	float (F11.3)	K	Radiation correction related part of uncertainty of temperature; (k=2)
temp_uc_ucor_rad	float (F16.3)	K	Radiation correction related part of uncorrelated uncertainty of temperature (sources are pressure and ventilation); (k=2)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
temp_uc_scor_rad	float (F16.3)	K	Radiation correction related part of spatially correlated uncertainty of temperature (source is RTM); (k=2)
temp_uc_tcor_rad	float (F16.3)	K	Radiation correction related part of temporally correlated uncertainty of temperature (source is experiment); (k=2)
temp_uc_tcor_cal	float (F16.3)	K	Calibration after ground preparation related part of temporally correlated uncertainty of temperature; (k=2)
temp_uc_ucor_sm	float (F15.3)	K	Smoothing related part of uncorrelated uncertainty of temperature; (k=2)
temp_corr	float (F9.4)	K	Correction of temperature
temp_corr_rad	float (F13.4)	K	Part of correction of temperature which is related to radiation
temp_corr_sm	float (F12.4)	K	Part of correction of temperature which is related to smoothing
temp_res	float (F8.1)	s	Time resolution of temperature
Relative humidity			
rh	float (F6.2)	%RH	Relative humidity
rh_uc	float (F7.3)	%RH	Uncertainty of relative humidity; (k=2)
rh_uc_ucor	float (F10.3)	%RH	Uncorrelated part of uncertainty of relative humidity; (k=2)
rh_uc_tcor	float (F10.3)	%RH	Temporally correlated part of uncertainty of relative humidity; (k=2)
rh_uc_tlc	float (F9.3)	%RH	Time-lag correction related part of uncertainty of relative humidity; (k=2)
rh_uc_ucor_tlc	float (F14.3)	%RH	Time-lag correction related part of uncorrelated uncertainty of relative humidity
rh_uc_tcor_tlc	float (F14.3)	%RH	Time-lag correction related part of temporally correlated uncertainty of relative humidity; (k=2)
rh_uc_tcor_cal	float (F14.3)	%RH	Calibration after ground preparation related part of temporally correlated uncertainty of relative humidity; (k=2)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
rh_uc_ucor_sm	float (F13.3)	%RH	Smoothing related part of uncorrelated uncertainty of relative humidity; (k=2)
rh_uc_tcor_tint	float (F15.3)	%RH	Internal T sensor related part of temporally correlated uncertainty of relative humidity; (k=2)
rh_uc_ucor_tair	float (F15.3)	%RH	Air T sensor related part of uncorrelated uncertainty of relative humidity; (k=2)
rh_corr	float (F7.3)	%RH	Correction of relative humidity
rh_corr_tl	float (F10.3)	%RH	Part of correction of relative humidity which is related to the sensor time-lag
rh_corr_trc	float (F11.3)	%RH	Part of correction of relative humidity which is related to the radiation correction of temperature sensor
rh_corr_sm	float (F10.3)	%RH	Part of correction of relative humidity which is related to smoothing
rh_res	float (F6.1)	s	Time resolution of relative humidity
Other humidity measures			
dp	float (F6.2)	K	Dew point temperature
dp_uc	float (F7.3)	K	Uncertainty of dew point temperature; (k=2)
dp_uc_ucor	float (F10.3)	K	Uncorrelated part of uncertainty of dew point temperature; (k=2)
dp_uc_scor	float (F10.3)	K	Spatially correlated part of uncertainty of dew point temperature; (k=2)
dp_uc_tcor	float (F10.3)	K	Temporally correlated part of uncertainty of dew point temperature; (k=2)
wvmr_mass	float (F9.2)	ppm	Water vapour mass mixing ratio
wvmr_mass_uc	float (F12.3)	ppm	Uncertainty of water vapour mass mixing ratio; (k=2)
wvmr_mass_uc_ucor	float (F17.3)	ppm	Uncorrelated part of uncertainty of water vapour mass mixing ratio; (k=2)
wvmr_mass_uc_scor	float (F17.3)	ppm	Spatially correlated part of uncertainty of water vapour mass mixing ratio; (k=2)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
wvmr_mass_uc_tcor	float (F17.3)	ppm	Temporally correlated part of uncertainty of water vapour mass mixing ratio; (k=2)
wvmr_vol	float (F8.2)	ppmv	Water vapour volume mixing ratio
wvmr_vol_uc	float (F11.3)	ppmv	Uncertainty of water vapour volume mixing ratio; (k=2)
wvmr_vol_uc_ucor	float (F17.3)	ppmv	Uncorrelated part of uncertainty of water vapour volume mixing ratio; (k=2)
wvmr_vol_uc_scor	float (F17.3)	ppmv	Spatially correlated part of uncertainty of water vapour volume mixing ratio; (k=2)
wvmr_vol_uc_tcor	float (F17.3)	ppmv	Temporally correlated part of uncertainty of water vapour volume mixing ratio; (k=2)
wvpp	float (F6.2)	hPa	Water vapour partial pressure
wvpp_uc	float (F7.3)	hPa	Uncertainty of water vapour partial pressure; (k=2)
wvpp_uc_ucor	float (F12.3)	hPa	Uncorrelated part of uncertainty of water vapour partial pressure; (k=2)
wvpp_uc_scor	float (F12.3)	hPa	Spatially correlated part of uncertainty of water vapour partial pressure; (k=2)
wvpp_uc_tcor	float (F12.3)	hPa	Temporally correlated part of uncertainty of water vapour partial pressure; (k=2)
wvsp	float (F6.2)	hPa	Water vapour saturation pressure
wvsp_uc	float (F7.3)	hPa	Uncertainty of water vapour saturation pressure; (k=2)
wvsp_uc_ucor	float (F12.3)	hPa	Uncorrelated part of uncertainty of water vapour saturation pressure; (k=2)
wvsp_uc_scor	float (F12.3)	hPa	Spatially correlated part of uncertainty of water vapour saturation pressure; (k=2)
wvsp_uc_tcor	float (F12.3)	hPa	Temporally correlated part of uncertainty of water vapour saturation pressure; (k=2)
Wind			
wdir	float (F5.1)	°	Wind direction; Wind from direction with 0°:north, 90°:east, 180°:south, 270°:west

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
wdir_uc	float (F7.2)	°	Uncertainty of wind direction; k=1 limit is 180°; (k=2)
wdir_uc_ucor	float (F12.2)	°	Uncorrelated part of uncertainty of wind direction; k=1 limit is 180°; (k=2)
wdir_uc_ucor_sm	float (F15.2)	°	Smoothing related part of uncorrelated uncertainty of wind direction; k=1 limit is 180°; (k=2)
wdir_uc_ucor_gnss	float (F17.2)	°	GNSS related part of uncorrelated uncertainty of wind direction; k=1 limit is 180°; (k=2)
wdir_corr_sm	float (F12.2)	°	Smoothing related part of correction of wind direction
wdir_usm	float (F8.1)	°	Unsmoothed wind direction
wspeed	float (F6.2)	$m\,s^{-1}$	Wind speed
wspeed_uc	float (F9.3)	$m\,s^{-1}$	Uncertainty of wind speed; (k=2)
wspeed_uc_ucor	float (F14.3)	$m\,s^{-1}$	Uncorrelated part of uncertainty of wind speed; (k=2)
wspeed_uc_ucor_sm	float (F17.3)	$m\,s^{-1}$	Smoothing related part of uncorrelated uncertainty of wind speed; (k=2)
wspeed_uc_ucor_gnss	float (F19.3)	$m\,s^{-1}$	GNSS related part of uncorrelated uncertainty of wind speed; (k=2)
wspeed_corr_sm	float (F17.3)	$m\,s^{-1}$	Smoothing related part of correction of wind speed
wspeed_usm	float (F10.2)	$m\,s^{-1}$	Unsmoothed wind speed
wzon	float (F7.2)	$m\,s^{-1}$	Zonal wind component
wzon_uc	float (F7.3)	$m\,s^{-1}$	Uncertainty of zonal wind component; (k=2)
wzon_uc_ucor	float (F12.3)	$m\,s^{-1}$	Uncorrelated part of uncertainty of zonal wind component; (k=2)
wzon_uc_ucor_sm	float (F15.3)	$m\,s^{-1}$	Smoothing related part of uncorrelated uncertainty of zonal wind component; (k=2)
wzon_uc_ucor_gnss	float (F17.3)	$m\,s^{-1}$	GNSS related part of uncorrelated uncertainty of zonal wind component; (k=2)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
wzon_corr_sm	float (F12.3)	$m\,s^{-1}$	Smoothing related part of correction of zonal wind component
wmeri	float (F7.2)	$m\,s^{-1}$	Meridional wind component
wmeri_uc	float (F8.3)	$m\,s^{-1}$	Uncertainty of meridional wind component; (k=2)
wmeri_uc_ucor	float (F13.3)	$m\,s^{-1}$	Uncorrelated part of uncertainty of meridional wind component; (k=2)
wmeri_uc_ucor_sm	float (F16.3)	$m\,s^{-1}$	Smoothing related part of uncorrelated uncertainty of meridional wind component; (k=2)
wmeri_uc_ucor_gnss	float (F18.3)	$m\,s^{-1}$	GNSS related part of uncorrelated uncertainty of meridional wind component; (k=2)
wmeri_corr_sm	float (F13.3)	$m\,s^{-1}$	Smoothing related part of correction of meridional wind component

Supplementary variables

vent	float (F6.2)	$m\,s^{-1}$	Ventilation speed
vent_uc	float (F7.3)	$m\,s^{-1}$	Uncertainty of ventilation speed; (k=2)
vent_uc_ucor	float (F12.3)	$m\,s^{-1}$	Uncorrelated part of uncertainty of ventilation speed; (k=2)
vspeed	float (F6.2)	$m\,s^{-1}$	Vertical (ascent/descent) speed
vspeed_uc	float (F9.3)	$m\,s^{-1}$	Uncertainty of vertical speed; (k=2)
vspeed_uc_ucor	float (F14.3)	$m\,s^{-1}$	Uncorrelated part of uncertainty of vertical speed; (k=2)
sza	float (F7.3)	°	Solar zenith angle
sea	float (F7.3)	°	Solar elevation angle
saa	float (F8.3)	°	Solar azimuth angle
hca	float (F8.3)	°	Horizon correction angle
grav	float (F9.5)	$m\,s^{-2}$	Acceleration of gravity
geoid	float (F9.2)	m	Difference between geoid and ellipsoid WGS84

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
dorn	integer (I5)	1	Day or night. Flag which decodes day, twilight, or night: 0 – unknown, 1 – day, 2 – night, 4 – civil twilight, 8 – nautical twilight, 16 – astronomical twilight, 32 – partly sun
tau_rh	float (F6.1)	s	Time-lag of humidity sensor (the inverse time-lag constant)
tau_rh_uc	float (F9.2)	s	Uncertainty of time-lag of humidity sensor
tau_rh_uc_tcor	float (F13.2)	s	Temporally correlated uncertainty of time-lag of humidity sensor
ciwv	float (F6.2)	$kg\,m^{-2}$	Cumulated integrated water vapour
ciwv_uc	float (F7.3)	$kg\,m^{-2}$	Uncertainty of cumulated integrated water vapour; (k=2)
icesat	float (F6.1)	%RH	Humidity of ice saturation
lat_sm	float (F10.5)	°N	Strong smoothed trajectory latitude
lon_sm	float (F10.5)	°E	Strong smoothed trajectory longitude
latmf	double (F10.2)	m	Factor to convert latitude from degree to meter
lonmf	double (F10.2)	m	Factor to convert longitude from degree to meter
penper	float (F6.1)	s	Pendulum period
penspeed	float (8.2)	$m\,s^{-1}$	Pendulum speed
penspeed_uc	float (F11.3)	$m\,s^{-1}$	Uncertainty of pendulum speed; (k=2)
penspeed_uc_ucor	float (F16.3)	$m\,s^{-1}$	Uncorrelated part of uncertainty of pendulum speed; (k=2)
penrad	float (F6.1)	m	Pendulum radius
penrad_uc	float (F9.2)	m	Uncertainty of pendulum radius; (k=2)
penrad_uc_ucor	float (F14.2)	m	Uncorrelated part of uncertainty of pendulum radius; (k=2)
swdir	float (F7.2)	$W\,m^{-2}$	Solar direct flux
swdir_uc	float (F10.3)	$W\,m^{-2}$	Uncertainty of solar direct flux; (k=2)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
swdir_uc_scor	float (F13.3)	$W\,m^{-2}$	Spatially correlated part of uncertainty of solar direct flux; (k=2)
swdif	float (F7.2)	$W\,m^{-2}$	Solar diffuse flux
swdif_uc	float (F10.3)	$W\,m^{-2}$	Uncertainty of solar diffuse flux; (k=2)
swdif_uc_scor	float (F13.3)	$W\,m^{-2}$	Spatially correlated part of uncertainty of solar diffuse flux; (k=2)
band	integer (I6)	1	Measurement band flag. Flag which decodes measurement band: 1 – flight, 2 – ascent, 4 – descent, 8 – troposphere, 16 – stratosphere, 32 – utls, 64 – prelaunch, 128 – postflight
Raw data			
temp_raw	float (F8.2)	K	Raw temperature
ihum_raw	float (F8.2)	%RH	Raw humidity at internal temperature of sensor
thum_raw	float (F8.2)	K	Raw temperature of humidity sensor
rh_raw	float (F8.2)	%RH	Raw humidity pre-calculated to air temperature
press_sens_raw	float (F14.2)	hPa	Raw sensor pressure before adjustment using ground check
posx_raw	double (F13.3)	m	Raw geographical Cartesian coordinate X
posy_raw	double (F13.3)	m	Raw geographical Cartesian coordinate Y
posz_raw	double (F13.3)	m	Raw geographical Cartesian coordinate Z
lat_raw	double (F10.5)	°N	Raw latitude (WGS84)
lon_raw	double (F10.5)	°E	Raw longitude (WGS84)
alt_raw	float (F8.1)	m	Raw altitude (WGS84)
wzon_raw	float (F8.2)	$m\,s^{-1}$	Raw zonal wind component
wmeri_raw	float (F9.2)	$m\,s^{-1}$	Raw meridional wind component
hdop_raw	float (F8.3)	1	Raw Horizontal Dilution Of Precision (HDOP)
vdop_raw	float (F8.3)	1	Raw Vertical Dilution Of Precision (VDOP)

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Type	Unit	Description
Flags related to raw data variables			
			Flag which decodes several flags of raw data: 1 – nan, 2 – outlier, 4 – out_of_range, 8 – zero, 16 – interpolated
temp_raw_flags	integer (I14)	1	Flags of raw temperature
ihum_raw_flags	integer (I14)	1	Flags of raw internal humidity
thum_raw_flags	integer (I14)	1	Flags of raw internal humidity temperature
rh_raw_flags	integer (I12)	1	Flags of pre-calculated raw relative humidity
press_sens_raw_flags	integer (I20)	1	Flags of raw sensor pressure
posx_raw_flags	integer (I14)	1	Flags of raw X coordinate
posy_raw_flags	integer (I14)	1	Flags of raw Y coordinate
posz_raw_flags	integer (I14)	1	Flags of raw Z coordinate
lat_raw_flags	integer (I13)	1	Flags of raw latitude
lon_raw_flags	integer (I13)	1	Flags of raw longitude
alt_raw_flags	integer (I13)	1	Flags of raw altitude
wzon_raw_flags	integer (I14)	1	Flags of raw zonal wind component
wmeri_raw_flags	integer (I15)	1	Flags of raw meridional wind component

A.3 Attributes of variables (metadata)

Table A.3: List of all attributes of variables included in NetCDF file

Name	Example	Description
CF attributes		
ancillary_variables	"alt_wgs84_uc alt_raw"	A list of all directly related ancillary variables
axis	"T"	Mark the variable as a specific axis
calendar	"gregorian"	Define related calendar to the time axis
comment	"Barometric air pressure using silicon sensor"	An additional comment to describe the variable in more detail
coordinates	"lon lat alt"	Coordinates to define the position of measurement
flag_masks	1S, 2S, 4S, 8S, 16S (short)	List of defined bit flags (can be combined)
flag_meanings	"nan outlier out_of_range zero interpolated"	List of meanings of defined bit flags
long_name	"Time"	The long name of the variable
standard_name	"time"	A variable name which is defined in CF standard
units	"seconds since 2020-02-05T10:54:14.049Z"	The unit
valid_min	-400.0f (float)	Valid minimum value
valid_max	45000.0f (float)	Valid maximum value
_ChunkSizes	1048576U (uint)	[file internal attribute]
General GRUAN specific attributes		
g_column_type	"original data"	[not used yet]
g_content_type	"TIME"	Column can be marked with a content type (GRUAN Java library)
g_coverage_factor	2.0f	Uncertainty estimates for variables and attributes in the product files are consistently specified at the 2-σ confidence level (coverage factor $k = 2$).

This table is continued on the next page.

Table A.3 – *Continued from previous page*

Name	**Example**	**Description**
g_format_format	"F8.1"	An optimal format code (Fortran like) for this variable
g_format_nan	"NaN"	A specifier for NaN in case of converting to text
g_format_type	"FLT"	A data type code (GRUAN IDL library)
g_format_width	"8"	Optimal column width in a formatted text table
g_resolution	"1.0 s (time)"	The real resolution as effective sample size
g_source_desc	"GpsResults.Wgs84Altitude"	Source of data (e.g. the table and column in case of raw data in MWX file)

Acronyms

CF	Climate and Forecast
FTP	File Transfer Protocol
GCOS	Global Climate Observing System
GDP	GRUAN Data Product
GMDB	GRUAN Meta-Data Base
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GRUAN	GCOS Reference Upper-Air Network
HDOP	Horizontal Dilution Of Precision
IDL	Interactive Data Language
ISO	International Organization for Standardization
LC	Lead Centre
MSL	Mean Sea Level
MWX	Zipped MW41 sounding archive file
NaN	Not a Number
NetCDF	Network Common Data Format
NCEI	National Centers for Environmental Information
PTU	Pressure Temperature hUmidity
QC	Quality Control
RS	RadioSonde
RTM	Radiative Transfer Model
SHC	Standard Humidity Chamber
UCAR	University Corporation for Atmospheric Research
UTC	Coordinated Universal Time
VDOP	Vertical Dilution Of Precision
WGS84	World Geodetic System 1984
WIGOS	WMO Integrated Global Observing System
WMO	World Meteorological Organization

Bibliography

Eaton, B., *et al.*, NetCDF Climate and Forecast (CF) Metadata Conventions. Version 1.7, Technical report, cfconventions.org, 2017, URL https://cfconventions.org/Data/cf-conventions/cf-conventions-1.7/cf-conventions.html.

Immler, F. J., *et al.*, Reference quality upper-air measurements: Guidance for developing GRUAN data products, *Atmospheric Measurement Techniques*, **3**(5), 1217–1231, 2010, URL https://doi.org/10.5194/amt-3-1217-2010.

JCGM, *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (JCGM, 2008), 1st edition, URL https://www.bipm.org/en/publications/guides/gum.html, [Last accessed: May 04, 2020].

von Rohden, C., M. Sommer, T. Naebert, V. Motuz, and R. Dirksen, Laboratory characterisation of the radiation temperature error of radiosondes and its application to the GRUAN data processing for the Vaisala RS41, *Atmospheric Measurement Techniques*, **15**, 383–405, doi: amt-15-383-2022, 2022, URL https://doi.org/10.5194/amt-15-383-2022.

Sommer, M., User Guide for the GRUAN RsLaunchClient, GRUAN Technical Document 3 (GRUAN-TD-3), GRUAN Lead Centre, 2014, URL https://www.gruan.org/documentation/gruan/td/gruan-td-3, v0.7 (2014-02-05).

Sommer, M., Brief Description of GruanToolRsLaunch (gtRsl), GRUAN Technical Note 6 (GRUAN-TN-6), GRUAN Lead Centre, 2020, URL https://www.gruan.org/documentation/gruan/tn/gruan-tn-6, v1.0 (2020-08-24).

Sommer, M., *et al.*, GRUAN characterisation and data processing of the Vaisala RS41 radiosonde, 2022a, DRAFT, v0.9.1 (2022-04-12), in review.

Sommer, M., *et al.*, RS41 GRUAN Data Product Version 1 (RS41-GDP.1), GDP, GRUAN Lead Centre (DWD), 2022b, URL https://doi.org/10.5676/GRUAN/RS41-GDP.1.