GCOS
Reference
Upper-
Air
Network

GRUAN Technical Document 4

Brief Description of the RS92 GRUAN Data Product (RS92-GDP)

Michael Sommer, Ruud Dirksen and Christoph von Rohden

Publisher

GRUAN Lead Centre

Number & Version

GRUAN-TD-4
Rev. 2.1.1
(2021-07-14)

Document info

Title:	Brief Description of the RS92 GRUAN Data Product (RS92-GDP)
Topic:	Radiosonde
Authors:	Michael Sommer, Ruud Dirksen and Christoph von Rohden
Publisher:	GRUAN Lead Centre, DWD
Document type:	Technical Document
Document number:	GRUAN-TD-4
Page count:	28
Version:	Rev. 2.1.1 (2021-07-14)

Abstract

This document describes the format of the data files for the RS92 radiosonde GRUAN data product. GRUAN data products are stored in NetCDF files. The data fields and attributes within the RS92 radiosonde NetCDF files are comprehensively described in this document. The GRUAN data products include raw data, processed data, uncertainty estimates, and selected meta data.

The correction algorithms that are applied in the GRUAN processing and the RS92 radiosonde measurements itself are extensively described in *Dirksen et al.* (2014).

Revision 2 of this document describes the release version 2 of the RS92 GRUAN data product (RS92-GDP) which is in operational use since 1 September 2012. This GRUAN data product is published as RS92-GDP.2 (*Sommer et al.*, 2012).

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 the 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 RS92 radiosonde data.

Revision history

Version	Author / Editor	Description
2.1.1 (2021-07-14)	Michael Sommer	Small typos and formatting issues are fixed
2.1 (2021-06-16)	Michael Sommer, Christoph von Rohden, E: Susanne Körner	Technical reformatting to LaTeX and new Section 5.1 added
2.0 (2016-02-11)	Michael Sommer, Ruud Dirksen, Christoph von Rohden	Revised version with modifications relating to version 2 of RS92-GDP
1.1 (2011-12-07)	Michael Sommer	Revised version with optimisation of structure and additional content
1.0 (2011-08-30)	Franz Immler	First version of this technical document describing the RS92 GRUAN data product (RS92-GDP) release version 1

Table of contents

1 Introduction

The GCOS Reference Upper Air Network (GRUAN) data product derived from Vaisala RS92 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 the vertical resolution of the uncertainty estimates. The RS92 radiosonde measures profiles of temperature (T), humidity (U), pressure (p), as well as wind speed and direction with a time resolution of one second. The GRUAN data processing algorithms include a correction for systematic and random effects, resulting in a data product which should be free of biases (*Dirksen et al.*, 2014). The uncertainties are calculated following the recommendations outlined in the "Guide for expressing uncertainty in measurement" (*JCGM*, 2008). The total uncertainty is assessed from estimates of calibration uncertainty statistical errors, and uncertainty resulting from any bias corrections.

The GRUAN data product consists of the processed data together with the raw calibrated measurements (*RAWPTU*).

2 GRUAN processing of RS92 radiosonde measurements

The GRUAN processing of each RS92 sounding is a two-stage process, where the first step involves collection, validation and pre-processing of the sensor measurements and accompanying meta-data, which are subsequently stored in a database. The actual processing (i.e., the application of the GRUAN correction algorithms) is performed during the second step provided that all 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 that consisting of multiple sondes, or non-radiosonde instruments (*Sommer*, 2014).

Upon submission, the meta-data are stored in the GRUAN meta-data-base (GMDB) whereas the raw data are stored in a file archive at the GRUAN Lead Centre. 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 DigiCora III database files (*.dc3db) are converted into the NetCDF format, preserving all information held in the dc3db file. The raw PTU data are stored in the table *FRAWPTU*, and the GPS position data are stored in the table *GPSDCC_RESULT*.

An extensive description of the correction algorithms of the GRUAN data processing is given in *Dirksen et al.* (2014). A summary of these algorithms is given below. The GRUAN processing is applied to the calibrated sensor data (raw data) to which the ground check corrections have been applied. However, the ground check correction for the humidity sensors is reversed prior to the processing.

2.1 Temperature

Temperature measurements are corrected for heating by solar radiation and for temperature spikes due to warm air flowing off the balloon and the sensor housing.

The radiation bias depends on the actinic flux (i.e., the sum of direct and scattered solar irradiance), ambient pressure and ventilation. The dependence of the radiation bias on these three parameters is determined through laboratory studies. The actinic flux affecting the temperature sensor is estimated using a radiative transfer model which takes into account the solar elevation angle during ascent. The uncertainty of the correction of the radiative temperature bias results from the quality of the assumptions on the albedo in the respective scene (clouds versus cloud-free), sensor orientation, ventilation, and the reliability of the radiation model derived from the laboratory measurements, and is therefore correlated.

Spike removal is accomplished by smoothing/filtering the temperature profile.

The uncorrelated uncertainty is determined by the statistical contributions (i.e. the standard deviation with respect to a smoothed profile) and is determined during the spike removal.

The correlated uncertainty of the temperature consists of the uncertainty associated with the radiation correction and the calibration uncertainty of the temperature sensor. For more details regarding the temperature correction see Section 5 of *Dirksen et al.* (2014).

2.2 Humidity

Humidity data are corrected for radiation dry bias, sensor time-lag, and temperature dependent calibration error.

The radiation dry bias is caused by solar heating of the humidity sensor. The correction for this effect takes into account the higher saturation pressure within the heated humidity sensitive sensor material. The temperature offset (effect) of the humidity sensor is estimated from the heating of the temperature sensor multiplied by an enhancement factor to account for the higher sensitivity of the (uncoated) humidity sensor to radiative heating. For more details regarding the radiation dry-bias correction, see Section 6.4 of *Dirksen et al.* (2014).

Sensor time-lag becomes relevant below -40 °C where the response time exceeds approximately 20 seconds. The time-lag causes flattening and smoothing of gradients and structures in the humidity profile predominantly in the upper troposphere and tropopause region. The correction for time-lag involves the numerical inversion of a low-pass filter. The inversion method amplifies the noisy components in the measured data. Therefore it is followed by another. This smoothing is used to estimate the uncorrelated uncertainty contribution from (statistical) noise. For more details regarding the time-lag correction, see Section 6.5 of *Dirksen et al.* (2014).

The correction of temperature dependent calibration error is based on comparisons with simultaneous Cryogenic Frost-point Hygrometer (CFH) soundings, which reveal a temperature dependent dry bias for ambient temperatures between -30 and -70 °C. For more details regarding the time-lag correction, see Section 6.3 of *Dirksen et al.* (2014).

2.3 Pressure, altitude and geopotential height

The RS92 is equipped with a pressure sensor and a GPS sensor which provide two independent methods to retrieve the vertical coordinate. In the lower part of the profile, where the signal-to-noise performance of the pressure sensor is sufficient, the geopotential height is derived from the pressure measurements. In the upper part of the profile, the GPS sensor is preferred to retrieve the geopotential height. The altitude of the switch between both sensors is typically between 9 and 17 km and follows from a statistical analysis. The geopotential height, geometric altitude, and pressure scale are consistent. For more details regarding the retrieval of the vertical coordinate, see Section 8 and Appendix B of *Dirksen et al.* (2014).

2.4 Wind

The horizontal position of the radiosonde is given in the original table *GPSDCC_RESULT* of the dc3db files. These data are smoothed using a digital filter with a cut-off frequency of 0.025 Hz (i.e. a smoothing over 40 seconds is applied). A statistical uncertainty is calculated from the noise on the signal that is assumed to arise from the pendulum motion of the radiosonde and from noise created by the GPS receiver itself. The zonal (u) and meridional (v) wind speed given in the data table (see Appendix B of *Dirksen et al.*, 2014) are directly derived from the raw data without smoothing applied.

For more details on the retrieval of wind data, see Section 9 of *Dirksen et al.* (2014).

2.5 Smoothing

Data are intermediately smoothed where necessary and associated uncertainties (see below) are calculated. A final smoothing is performed using a digital filter with a cut-off frequency (f_c), which is specific for each measurement parameter. The effective time resolution defined by $1/f_c$ is provided in the attributes for each variable (attribute *g_resolution*). The resolution of the data product is again given at the original time resolution of one second.

2.6 Uncertainties, corrections and resolution

For each measured variable other than latitude and longitude, the combined uncertainties are given in a separate data field named *u_<variable tag name>*. This uncertainty generally consists of a correlated and an uncorrelated component. The uncorrelated component is given as standard deviation of the mean resulting from smoothing and therefore depends on the applied smoothing filter. The uncertainties arising from the calibration and from systematic effects are typically correlated.

The correlation length, i.e. the range within the profile where the uncertainties of the individual data points are not independent from each other, can be as large as the entire profile, as is the case for the calibration uncertainty. When the correlation length is restricted to limited sections of the profile, it is typically still considerably larger than the original vertical time resolution. The reduced effective time resolution N' of the profile after application of the smoothing filter in the final step of the processing is stored in the attribute *g_resolution*.

For the temperature T and the relative humidity U, the correlated and the uncorrelated uncertainty components are additionally available as standard uncertainties in separate data fields, individually for each data point. For other parameters (wind, altitude, and pressure) only the combined uncertainty is supplied, which in the case of wind (speed and direction) is uncorrelated (standard deviation).

For further data analyses with the GRUAN products, the user is encouraged to consider the provided uncertainties. When for example calculating averages, the combined standard (=total) uncertainty requires separate evaluation of the components from random (uncorrelated) sources and those from systematic (correlated) sources (e.g. corrections for systematic effects applied to profile sections or the whole profile), before combining them. The following formulae should be used for the uncertainty components of those averages:

Correlated uncertainties:

$$\bar{k}_i = \sum_{j=i-M}^{i+M} (c_j k_{i+j}) \tag{2.1}$$

Uncorrelated uncertainties:

$$\bar{r}_i = \sqrt{\sum_{j=i-M}^{i+M} (c_j r_{i+j})^2} \tag{2.2}$$

$\bar{k}_i$ is a mean value for the uncertainty representative for the averaging interval, whereas $\bar{r}_i$ takes into account the compensational effect when combining statistically independent data points. The combined (total) uncertainty u of the average signal is then the sum of the squared components:

$$u = \sqrt{\bar{k}^2 + \bar{r}^2} \tag{2.3}$$

The c_j in Eqs. (2.1) and (2.2) are the kernel weights for a filter with $2M+1$ elements $c_{-M} \ldots c_{+M}$, which is defined by the users according to their specific applications. It should be normalised such that $\sum c_j = 1$. This kernel is not related to that used for GRUAN processing (see Section 2.5).

In the current version 2 of the data product, the field *u_std_<parameter>* contains the standard deviation instead of the uncorrelated statistical uncertainty. Therefore the field *u_std_<parameter>* must be divided by the resolution of *<parameter>* (attribute *g_resolution*) to get the uncorrelated uncertainty. For the relative humidity, the variable *res_rh* has to be used instead of attribute *g_resolution*, because the resolution is variable due to the time-lag correction. See following examples:

- Use of attribute *g_resolution* as N' of variable *temp*, which is the scalar value of 10 s

$$u(T) = \sqrt{u^2_{\text{cor}}(T) + \frac{u^2_{\text{std}}(T)}{10}} \tag{2.4}$$

- Use of variable *res_rh* instead of attribute *g_resolution* as N'

$$u(U) = \sqrt{u^2_{\text{cor}}(U) + \frac{u^2_{\text{std}}(U)}{\text{res\_rh}}} \tag{2.5}$$

> **Note:** There are systematic inconsistencies in the uncertainty budgets of some variables, e.g. temperature (*temp*), relative humidity (*rh*), with reference to the values of the attribute *g_resolution* (N'). See Section 5.1 for dealing with this error.

3 Locating and naming of files

GRUAN data files are published on a ftp server at the National Climatic Data Center – NCDC (ftp://ftp.ncdc.noaa.gov/pub/data/gruan).

The directory structure organises GRUAN data files by product, version, site and year:

- *processing/* – Processed GRUAN data
- *level<level number>/* – Data level, here *level2* (see below)
- *<product code>/* – Product code, here *RS92-GDP* (RS92 GRUAN Data Product)
- *version-<version number>/* – Product version, here e.g. *version-002*
- *<site code>/* – Site code, e.g., *CAB* for Cabauw (see https://www.gruan.org/network/sites)
- *<year>/* – Year of measurement, e.g., *2011*

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

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

Example : CAB-RS-01_2_RS92-GDP_001_20110203T120000_1-000-001.nc

- *SSS* – The three letter abbreviation for the station name, e.g. CAB for Cabauw (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 Cabauw)
- *L* – The data level coded as follows:
  - 0 Original raw data (files)
  - 1 Preprocessed 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., RS92-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 part which is the related telemetry software/hardware, e.g., 000
  - *VVV* – the version of the data, e.g., 001

4 Handling and basic structure of NetCDF files

GRUAN data product files (GDP files) are provided in NetCDF format. NetCDF is a self-describing binary format that was developed by UNIDATA. Software libraries for handling NetCDF files are available for a variety of programming languages (e.g., FORTRAN, C, JAVA, IDL) at https://www.unidata.ucar.edu/software/netcdf. The GRUAN processing uses NetCDF capabilities to include meta-data. Some general terms are discussed here to explain the structure of the data products: global attributes, variables and variable attributes.

4.1 Global attributes

Global attributes are intended to store information (meta-data) with a global meaning for the content of the actual file. Two groups of meta-data attributes are distinguished:

- General information using the climate and forecasting meta-data standard CF-1.4 (*Eaton et al.*, 2009)
- GRUAN details, which include
  - measurement event (launch)
  - instruments used (sondes)
  - processing software used
  - status of overall quality

The global attributes describe the data product, 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.

4.2 Variables

Each variable is stored as a separate column as part of a single table that forms the entire data content in a GDP file. The 1d-data arrays for each variable have same length (number of data points), equal to that of a common column for time [s]. An overview list of the stored variables is included in each NetCDF file.

The data of the variables for the RS92 GRUAN data product are generally complemented by uncertainty estimates. Generally, uncertainty estimates are provided for the variables in the RS92 GRUAN data product in separate columns. Variables without corresponding uncertainty estimates should be considered as ancillary data. A list of all variables (table columns) is provided in Appendix A.2.

4.3 Attributes to variables

Attributes associated with each variable (e.g. units) are provided in the NetCDF files as well. The column standard name, long name, a description (comment), and the physical unit of the actual variable are given in the attributes attached to each column. Additional attributes are provided for internal use. See appendix A.3 for a brief description of the available variable attributes.

5 Known issues

A number of small errors (bugs) have been identified during the use of data files of RS92 GRUAN data product (version 2), and requests for extensions or improvements came up:

- **Bug** – Wrong wind values (*wdir* and *wspeed*) in the first 23 seconds after launch and in the last 23 seconds before the balloon burst.
- **Bug** – Sometimes the relative humidity (*rh*) is reported as 0.0 instead of a null value (NaN).
- **Bug** – Incorrect tropopause heights are occasionally (rarely) reported in the global attributes.
- **Bug** – There are systematic inconsistencies in the uncertainty budgets of some variables, e.g. temperature (*temp*), relative humidity (*rh*). See Section 5.1 for dealing with this error.
- *Feature* – Super-adiabatic lapse rates often occur after passage of wet clouds. Identification and flagging of affected data points will be implemented in the next version of the data product.
- *Feature* – Ground check data are not stored in the NetCDF files.
- *Feature* – More detailed flagging of affected data points and variables in terms of a number of evaluation criteria such as filtering, contamination, quality rating.

The errors will be fixed, and the features listed above will be included or improved in the next version 3 of this GRUAN data product.

Note: If you have found any further bugs or errors in the NetCDF files of this GRUAN data product or have any suggestions for improvement, please contact the GRUAN Lead Centre (gruan.lc@dwd.de).

5.1 Inconsistencies in uncertainty budgets

Temperature

Temperature profiles were smoothed with a running non-uniform kernel. With that smoothing, a kernel-weighted standard deviation $\sigma_{\text{std}}(T)$ was calculated (stored in *u_std_temp* in the GDP file), as well as the effective sample size $N_T^{'}$, which is stored as constant numerical value of 10.0 in the attribute *g_resolution* (see also Section 2.6). Both were used to calculate a statistical uncertainty for the smoothed data, which reflects the small-scale noise of the original measurements, and should then be taken as uncorrelated uncertainty component $u_{\text{ucor}}(T)$:

$$u_{\text{ucor}}(T) = \frac{\sigma_{\text{std}}(T)}{\sqrt{N'_T}}. \tag{5.1}$$

The stored value for N'_T is incorrect. Using it for recalculation of $u_{\text{ucor}}(T)$, and – together with the correlated uncertainty $u_{\text{cor}}(T)$ stored in the product file (*u_cor_temp*) – the total temperature uncertainty

$$u(T) = \sqrt{u_{\text{cor}}(T)^2 + u_{\text{ucor}}(T)^2}, \tag{5.2}$$

would lead to inconsistencies with the correctly calculated value for $u(T)$ stored in the product file (*u_temp*). Furthermore, the correct value of the effective sample size, which is generally $N'^{*}_T = 11.409255$, is not completely constant but changes if the running kernel includes missing data points (NaN) or approaches the ends of the actual temperature profile.

Correction instructions

It is recommended to recover point by point the uncorrelated uncertainty component with:

$$u_{\text{ucor}}(T) = \sqrt{u(T)^2 - u_{\text{cor}}(T)^2} \tag{5.3}$$

Alternatively, the true (variable) effective sample size can be recovered using

$$N''_T = \frac{\sigma_{\text{std}}(T)^2}{u(T)^2 - u_{\text{cor}}(T)^2}, \tag{5.4}$$

and from that the uncorrelated uncertainty:

$$u_{\text{ucor}}(T) = \frac{\sigma_{\text{std}}(T)}{\sqrt{N''_T}}. \tag{5.5}$$

Relative humidity

Similar to temperature, the uncertainty budget for relative humidity is inconsistent and does not fully comply to the description in Section 2.6. The components of the budget can be restored by recalculation.

The following variables of the uncertainty budget concerning the relative humidity are involved:

- *u_rh* – total (combined) uncertainty $u(U)$
- *u_cor_rh* – correlated part of uncertainty $u_{\text{cor}}(U)$
- *u_std_rh* – uncorrelated part of uncertainty as standard deviation $\sigma_{\text{std}}(U)$
- *res_rh* – effective sample size (resolution) N'_U which is variable during flight

The stored value for the effective sample size (or resolution) N'_U is incorrect. That is, using Eq. (2.5) in Section 2.6 for recalculation would lead to inconsistencies.

Correction instructions

It is recommended to recover point by point the uncorrelated uncertainty component by using:

$$u_{\text{ucor}}(U) = \sqrt{u(U)^2 - u_{\text{cor}}(U)^2}. \tag{5.6}$$

Alternatively, the correct (variable) effective sample size can be recovered with

$$N''_U = \frac{\sigma_{\text{std}}(U)^2}{u(U)^2 - u_{\text{cor}}(U)^2}, \tag{5.7}$$

and from that the uncorrelated uncertainty:

$$u_{\text{ucor}}(U) = \frac{\sigma_{\text{std}}(U)}{\sqrt{N''_U}}. \tag{5.8}$$

Overview of corrected values for effective sample size

Resolution (effective sample size) N' is given as attribute *g_resolution* of all main variables and should be corrected in general:

- *time, lat, lon, u, v* – "1.0 s (time)" $\longrightarrow N'^* = 1.0$
- *temp, FP* – "10.0 s (time)" $\longrightarrow N'^* = 11.409255$
- *press, geopot, alt* – "15.0 s (time)" $\longrightarrow N'^* = 17.114460$
- *wdir, wspeed* – "40.0 s (time)" $\longrightarrow N'^* = 45.639629$
- *asc* – "60.0 s (time)" $\longrightarrow N'^* = 68.459496$

Appendix

A Appendix

A.1 Global attributes

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

Category/Attribute	Example	Description
CF specific global attributes [see Eaton et al., 2009]		
Conventions	'CF-1.4'	Used convention → CF-1.4: NetCDF Climate and Forecast CF Metadata Convention, Version 1.4, 2009-02-27
title	'RS92 GRUAN Data Product (Version 2)'	Name of the data product
institution	'DWD - German Weather Service'	Institution, where the measurement was performed
source	'RS92-SGP'	Source of measurement data – the instrument
history	–	Sequence of processing steps
references	–	References to publications or documentations, describing the data product
comment	–	Description of data product
GRUAN specific global attributes → in the format *g.* <*Category* >. <*AttributeName* > : value (*always stored as string*) → Example: *g.Product.ID* : '15775'		
Product (category)		Several specific facts of this data product file.
ID	15775	[*internal use only*] – Identifier in the GRUAN meta-data-base (GMDB).
Code	'RS92-GDP'	Code of data product (same as in the file name).
Name	'RS92 GRUAN Data Product'	Name / title of data product

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Category/Attribute	Example	Description
Version	2	Version of data product (same as in the file name)
Level	2	Level of data file (same as in the file name) – see chapter 3
LevelDescription	–	Description of *Level*
History	–	Sequence of processing steps
References	–	References to publications or documentations,describing the data product
Producer	'GRUAN Lead Centre'	Institution, where the data file was created
OrgResolution	'1.0 s (time)'	Original resolution of measurement with unit and corresponding variable
Status	'Data_approved'	Quality status of data file. Only status *Data_approved* will be published at NCEI. Options are: • *Data_approved* – data successfully processed and passed all quality checks (GRUAN stamp) • *Data_checked* – minor issue. Data successfully processed but did not pass all quality checks. • *Discarded – Do not use this data!*
StatusDescription	–	Long description of status

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Category/Attribute	Example	Description
ProcessingCode	'TRC, cc, HRC, TL, pGPS, or'	List of codes of used processing modules, e.g.: • *TRC* – Temperature radiation correction • *cc* – Calibration correction (of humidity sensors) • *HRC* – Humidity radiation correction • *TL* – Time-lag correction • *pGPS* – Combination of pressure and GPS • *or* – Outlier removal (remove temperature spikes)
General (category)		General facts of file and GRUAN site.
FileTypeVersion	0.8	Version of file type definition. *If the version has changed, it is possible that attributes/variables have been added or removed.*
Timestamp	2011-12-01T 09:09:49	Date and time of file creation (UTC)
SiteCode	'LIN'	GRUAN station code (see https://www.gruan.org/network/sites)
SiteName	'Lindenberg'	GRUAN station name (see https://www.gruan.org/network/sites)
SiteWmoId	10393	WMO number of GRUAN site (only if existent)
SiteInstitution	'DWD - German Weather Service'	Institution, which operates the site
MeasuringSystem (category)		The measurement system at the specific GRUAN site. *This is always a specific radiosonde launch site in case of the RS92 GRUAN data product.*
ID	'LIN-RS-01'	Code of measurement system (same as in the file name).
Type	'Radiosonde'	Type of measurement system.
Longitude	'14.12 °'	Longitude of the measurement system (e.g. launch site). [degree East]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Category/Attribute	Example	Description
Latitude	'52.21 °'	Latitude of the measurement system (e.g. launch site). [degree North]
Altitude	'103.8 m'	Altitude of the measuring system (e.g. launch site). If a pressure reference sensor is used to recalibrate the pressure sensor – this altitude is stored here. [m]
SurfaceObs (category)		Surface observations from launch site at launch date.
Pressure	'1007.90 hPa'	Surface pressure at launch site [hPa]
Temperature	'0.70 °C'	Temperature at launch site [°C]
RelativeHumidity	'88.0 %'	Relative humidity at launch site [%RH]
Ascent (category)		Radiosonde ascent (launch) information
ID	4329	[*internal use only*] – ID of the event (radiosonde launch) from the GRUAN meta database (GMDB).
StandardTime	2011-12-01T06:00:00	Standard (or scheduled, synop) time for launch (same as in the file name). *'Scheduled' usually means WMO-required, e.g. 00, 06, 12, 18 UTC.*
StartTime	2011-12-01T05:07:36	Actual time of the launch (UTC)
BalloonNumber	1	Number of balloon (in case of multiple balloons launches at the same *StandardTime*; same as in the file name).
BalloonType	'TA600'	Code of balloon type (codes are defined in the GRUAN meta database – GMDB)
UnwinderType	'UW-V30'	Code of unwinder type (codes are defined in the GRUAN meta database – GMDB)
FillingWeight	'300.0 g'	Balance weight determines the balloon gas filling [g]
Payload	'300.0 g'	Weight of all components attached to the balloon [g]
GrossWeight	'900.0 g'	Weight of all components to launch (including balloon)[g]

This table is continued on the next page.

Table A.1 – *Continued from previous page*

Category/Attribute	Example	Description
IncludeDescent	‘no’	Are the descent data included in this file? [yes/no]
BurstpointAltitude	‘31782.0 m’	Altitude of burstpoint [m]
BurstpointPressure	‘8.60 hPa’	Pressure of burstpoint [hPa]
PrecipitableWater-Column	‘9.3 kg m-2’	Precipitable water column derived from the sounding [kg/m^2]
PrecipitableWater-ColumnU	‘0.3 kg m-2’	Uncertainty of precipitable water column [kg/m^2]
TropopauseHeight	‘12092.1 m’	Height/altitude of WMO tropopause [m]
Tropopause-Temperature	‘203.9 K’	Temperature at WMO tropopause [K]
TropopausePressure	‘187.8 hPa’	Pressure at WMO tropopause [hPa]
Tropopause-PotTemperature	‘328.8 K’	Potential temperature at WMO tropopause [K]
Instrument (category)		Instrument of measuring (sonde)
SerialNumber	‘F3641355’	Serial number of the instrument
Type	“RS92-SGP’	Identifier of the instrument type (codes are defined in the GRUAN meta database – GMDB).
TypeFamiliy	‘RS92’	Identifier for the instrument family (codes are defined in the GRUAN meta database – GMDB).
Manufacturer	‘Vaisala’	Name of the instrument manufacturer
Weight	‘285.0 g’	Weight of the instrument [g]
TelemetrySonde	‘RS92’	[*not relevant*] – Code of the instrument family (sonde) used for telemetry. This is only relevant for devices, which do not have own telemetry functionality (like CFH, COBALD).
SoftwareVersion	‘3.64.1’	Software version used for telemetry/processing.
Comment	–	Additional comments to the instrument

A.2 Variables (data columns)

Note: The units for the variables follow the CF-1.4 convention (see Appendix A.3, attribute *units*). And there are several special unit definitions, i.e.:

- degree – decimal degree
- degree East – decimal degree (W → E; –180 → +180)
- degree North – decimal degree (S → N; –90 → +90)
- 1 – means dimensionless (no unit)
  - If percentage [%] is required, such as for relative humidity, multiply the value in the column by 100.
  - Mixing ratio is always given in volume mixing ratio with no unit. If parts per million [ppm] is required, multiply the value in the column by 1,000,000.

Table A.2: Variables for RS92 GRUAN data product files

Variable	Unit	Description
time	s	Time after launch
press	hPa	Air pressure derived from pressure sensor measurement and GPS altitude measurement (see Section 8 of *Dirksen et al.*, 2014).
temp	K	Ambient temperature (see Section 5 of *Dirksen et al.*, 2014).
rh	1	Relative humidity over liquid water, using the *Hyland and Wexler* (1983) formula (see Section 6 of *Dirksen et al.*, 2014).
wdir	degree	Wind direction derived from smoothed GPS data. Compass reading (0°= North) (see Section 9 of *Dirksen et al.*, 2014).
wspeed	$m{\cdot}s^{-1}$	Wind speed derived from smoothed GPS data. (see Section 9 of *Dirksen et al.*, 2014).
geopot	m	Geopotential height calculated from air pressure and GPS altitude (see Section 8 of *Dirksen et al.*, 2014).
lon	degree East	Longitude, taken from GPS data.
lat	degree North	Latitude, taken from GPS data.

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Unit	Description
alt	m	Geometric altitude above sea level calculated from air pressure and GPS altitude (see Section 8 of *Dirksen et al.*, 2014).
u	$m \cdot s^{-1}$	Zonal wind speed. Unfiltered, retrieved by DigiCora from GPS data.
v	$m \cdot s^{-1}$	Meridional wind speed. Unfiltered, retrieved by DigiCora from GPS data.
FP	K	Frost-point temperature calculated from *rh* and *temp* using Hyland and Wexler's formula (*Hyland and Wexler*, 1983)
WVMR	1	Volume mixing ratio of water vapour calculated from *rh*, *temp* and *press*.
asc	$m \cdot s^{-1}$	Ascent rate of the radiosonde calculated from *alt*.
SWrad	$W \cdot m^{-2}$	Short-wave radiation (actinic flux) on the sensor, retrieved from a radiation transfer model taking into account the actual solar zenith angle (see Section 5.2.2 of *Dirksen et al.*, 2014).
u_SWrad	$W \cdot m^{-2}$	Correlated uncertainty of the actinic flux on the sensor, estimated from the difference between a cloudy and a cloudfree scenario (see Section 5.2.2 of *Dirksen et al.*, 2014).
cor_temp	K	Radiation bias correction of temperature. $T_{GRUAN} = T_{RAW}$ + cor_temp (see Section 5 of *Dirksen et al.*, 2014).
u_cor_temp	K	Correlated uncertainty component of the air temperature derived from estimated of uncertainty in calibration and radiation correction (see Section 5.6 of *Dirksen et al.*, 2014).
u_std_temp	K	Standard deviation from the mean of the air temperature, see Section 5.1. It characterises the noise in the measurement (see Section 5.6 of *Dirksen et al.*, 2014).

This table is continued on the next page.

Table A.2 – *Continued from previous page*

Name	Unit	Description
u_temp	K	Combined uncertainty of temperature composed of the correlated and uncorrelated components (see Section 5.6 of *Dirksen et al.*, 2014).
u_alt	m	Combined uncertainty of the geometric altitude, dominated by correlated uncertainty (see Section 8 of *Dirksen et al.*, 2014).
u_press	hPa	Combined uncertainty of the air pressure dominated by correlated uncertainty (see Section 8 of *Dirksen et al.*, 2014).
res_rh	s	Effective time resolution of the relative humidity after time-lag correction (1/f_c, derived from altitude-dependent kernel width used for averaging) (see Section 6.5 of *Dirksen et al.*, 2014).
u_std_rh	1	Standard deviation of the relative humidity, see Section 5.1. It characterises the noise after the time-lag correction (see Section 6.6 of *Dirksen et al.*, 2014).
cor_rh	1	Correction applied to the relative humidity. $RH_{GRUAN} = RH_{RAW}$ + cor_rh (see Section 6 of *Dirksen et al.*, 2014).
u_cor_rh	1	Correlated uncertainty component due to uncertainties in relative humidity derived from calibration, radiation correction, and time-lag correction (see Section 6.6 of *Dirksen et al.*, 2014).
u_rh	1	Combined uncertainty of relative humidity (correlated and uncorrelated components are provided in separate variables) (see Section 6.6 of *Dirksen et al.*, 2014).
u_wdir	degree	Uncorrelated uncertainty of the wind direction. It is calculated as the statistical uncertainty of the mean (see Section 9 of *Dirksen et al.*, 2014).
u_wspeed	$m \cdot s^{-1}$	Uncorrelated uncertainty of the wind speed. It is calculated as the statistical uncertainty of the mean (see Section 9 of *Dirksen et al.*, 2014).

A.3 Variable Attributes

The variable attributes include standardised attributes from CF-1.4 (*Eaton et al.*, 2009) and special attributes, defined for GRUAN purposes (starting with *g_*).

Table A.3: Variable attributes saved in the GRUAN data product

Attribute	**Example *(variable wspeed)***	**Description**
CF specific variable attributes (see *Eaton et al.*, 2009)		
standard_name	'wind_speed'	Standard name of the field. (compatible with the standard CF-1.4[1], where applicable)
units	'm·s^{-1}'	SI unit (conform to CF-1.4[1])
long_name	'Wind Speed'	Name of the variable
comment	'Wind speed'	Brief description of variable, may include some information about the processing
coordinates	'lon lat alt'	List of relevant coordinate variables (mostly 'lon lat alt')
related_columns	'u_wspeed'	List of related variables (columns) (e.g. uncertainties, corrections, resolution)
positive	–	Direction of increasing coordinate ['up' / 'down'] (relevant only for variable *alt*)
GRUAN specific variable attributes		
g_source_desc	'RS92'	Source table (in dc3db-file) from which raw data were retrieved and codes for processing steps
g_resolution	'40.0 s (time)'	$1/f_c$ of applied smoothing filter
g_format_type	'FLT'	[*internal use*] – Internal code of format type (e.g. FLT for floating-point number)
g_format_format	'F6.2'	[*internal use*] – Format code (Fortran-like) for output in ASCII files.
g_format_width	'6'	[*internal use*] – Width of formatted value
g_format_nan	'NaN'	[*internal use*] – Internal format for missing values

This table is continued on the next page.

Table A.3 – *Continued from previous page*

Attribute	**Example** *(variable wspeed)*	**Description**
g_column_type	'original data'	Type of data (of this variable): • original data • derived data product • total uncertainty • standard deviation • correlated uncertainty • uncorrelated uncertainty • resolution • correction
g_processing_flag	'uncertainty calculated, smoothed'	Description of processing steps that were performed: • *raw*: data were processed from raw data, i.e. the table FRAWPTU was used, not FLEDT or other preprocessed sources • *corrected*: data was corrected for known bias(es). • *uncertainty calculated*: the uncertainty was calculated and added as an additional variable to the table • *bias corrected*: a bias correction was applied to this field • *individually smoothed*: the field was smoothed with a digital filter with a cut-off-frequency f_c=1/resolution • *GC checks are positive*: ground check correction are within specified limits • *additional GC positive*: deviation observed in 100 % pot are within specifications • *spikes removed*: positive spikes in temperature profile have been removed

[1] CF-1.4: NetCDF Climate and Forecast CF Metadata Convention, Version 1.4, 2009-02-27, (*Eaton et al.*, 2009)

Bibliography

Dirksen, R. J., *et al.*, Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, *Atmospheric Measurement Techniques*, **7**, 4463–4490, doi:10.5194/amt-7-4463-2014, 2014, URL https://doi.org/10.5194/amt-7-4463-2014.

Eaton, B., *et al.*, NetCDF Climate and Forecast (CF) Metadata Conventions. Version 1.4, Technical report, cfconventions.org, 2009, URL http://cfconventions.org/Data/cf-conventions/cf-conventions-1.4/build/cf-conventions.html.

Hyland, R. and A. Wexler, Formulations for the thermodynamic properties of the saturated phases of H_2O from 173.15 K to 473.15 K, *ASHRAE Tra*, **89**, 500–519, 1983.

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.

Sommer, M., User Guide of GRUAN RsLaunchClient. GRUAN Technical Document 3 (GRUAN-TD-3), GRUAN Technical Document 3 (GRUAN-TD-3), GRUAN Lead Centre, 2014, URL https://www.gruan.org/documentation/gruan/td/gruan-td-3, GRUAN internal used DRAFT Version 0.7.

Sommer, M., R. J. Dirksen, and F. J. Immler, RS92 GRUAN Data Product Version 2 (RS92-GDP.2), GRUAN Data Product, GRUAN Lead Centre, 2012, URL http://dx.doi.org/10.5676/gruan/rs92-gdp.2.