EUMETSAT
ROM SAF
RADIO OCCULTATION METEOROLOGY

ROM SAF CDOP-3

Visiting Scientist Report 37:

Radio Occultation Measurements as the Primary Anchor in a Hierarchy of Anchor Observations for Numerical Weather Predictions

Jordis S. Tradowsky

ROM SAF Consortium
Danish Meteorological Institute (DMI)
European Centre for Medium-Range Weather Forecasts (ECMWF)
Institut d'Estudis Espacials de Catalunya (IEEC)
Met Office (UKMO)

DOCUMENT AUTHOR TABLE

	Author(s)	Function	Date
Prepared by:	Jordis S. Tradowsky	ROM SAF Visiting/Associate scientist	30/04/19
Reviewed by (Internal):	Neill Bowler	Project Supervisor	25/04/19
Reviewed by (Internal):	John Eyre	Project Mentor	08/05/19
Approved by:	Sean Healy	ROM SAF Science Coordinator	26/06/19
Approved by:	Kent Lauritsen	ROM SAF Project Manager	05/07/19

DOCUMENT CHANGE RECORD

Version	Date	By	Description
Version 0.1	23/04/19	JST	The draft report to be reviewed.
Version 0.2	30/04/19	JST	Draft including Neill's comments.
Version 0.3	12/06/19	JST	Version including comments from John Eyre and Tom Gardiner, as well as final comments from Neill Bowler.
Version 1	05/07/19	JST	Including Sean Healy's and Kent Lauritsen's comments. Adjusting some text based on Neill Bowler's suggestions.

VS Authors

This visiting/associate scientist report was produced by Jordis Tradowsky (jordis@bodekersceintific.com)

VS Duration

The VS study was performed during four months within the period of 1 December 2018 to 30 April 2019 with a 7 weeks visit to Met Office, UK, during January - March 2019 and with a one week visit to DMI in March 2019.

ROM SAF

The Radio Occultation Meteorology Satellite Application Facility (ROM SAF) is a decentralised processing center under EUMETSAT which is responsible for operational processing of GRAS radio occultation (RO) data from the Metop satellites and radio occultation data from other missions. The ROM SAF delivers bending angle, refractivity, temperature, pressure, humidity, and other geophysical variables in near real-time for NWP users, as well as reprocessed Climate Data Records (CDRs) and Interim Climate Data Records (ICDRs) for users requiring a higher degree of homogeneity of the RO data sets. The CDRs and ICDRs are further processed into globally gridded monthly-mean data for use in climate monitoring and climate science applications.

The ROM SAF also maintains the Radio Occultation Processing Package (ROPP) which contains software modules that aid users wishing to process, quality-control and assimilate radio occultation data from any radio occultation mission into NWP and other models.

The ROM SAF Leading Entity is the Danish Meteorological Institute (DMI), with Cooperating Entities: i) European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, United Kingdom, ii) Institut D'Estudis Espacials de Catalunya (IEEC) in Barcelona, Spain, and iii) Met Office in Exeter, United Kingdom. To get access to our products or to read more about the ROM SAF please go to: `http://www.romsaf.org`.

Intellectual Property Rights

List of Contents

Executive Summary

For the purpose of numerical weather prediction (NWP), both radio occultation (RO) and radiosonde (RS) measurements are used to anchor the temperature. Thus, it is important to know if these observation types are consistent and to eliminate any existing biases between them prior to their assimilation into NWP systems. Following the development of a new method to compare RO and RS departures (with respect to model background) using a tangent linear dry temperature retrieval (Tradowsky [25]), and the use of this method to estimate RS temperature biases (Tradowsky [24]), this report will focus on the comparison of RO with specifically high-quality RS types and reference-quality reprocessed RS data. Besides, this report will further investigate structural uncertainties in the tangent linear RO retrieval.

Three work packages are addressed in this report, i.e.:

1. Comparison of RO data with a reference-quality RS data product developed by the Global Climate Observing System Reference Upper-Air Network (GRUAN, `https://www.gruan.org/`).
   As both, GRUAN and RO data, are commonly used as reference in the comparison with other observation types or models, good agreement between GRUAN and RO is expected. If discrepancies between the observation types is found, this indicates either a remaining bias in the GRUAN data, issues with the RO retrieval, or weaknesses in the method.

2. Comparison of RO data with high-quality RS measurements from the operational upper-air network. Globally around 800 upper-air sites measure temperature, humidity, wind speed, wind direction direction, and, depending on instrument type, pressure, using balloon-borne RS that are launched from the surface and typically reach altitudes up to 35 km. These measurements are mainly intended for assimilation into NWP models, however, they can also be used for process studies and climate change analysis. The data are usually corrected using a vendor dependent correction software and are then distributed through the Global Telecommunication System, which distributes meteorological measurements. The measurements from RS types which are known to provide especially good measurements are compared with RO measurements here to analyse if it would be advisable to assimilate them into NWP models without applying a bias correction beforehand.

3. Analysis of structural uncertainties in the tangent linear RO retrieval. As described in Tradowsky [25], Tradowsky et al. [26], the technique used here for the calculation of the RO dry temperature departures uses bending departures, which are set to zero above an impact height of 35 km before they are propagated into dry temperature departures with a tangent linear retrieval. Like every assumption that is made in a retrieval, reasonable variations in the choice of this upper cut-off impact height lead to a spread in the departure. This spread is also referred to as structural uncertainty here. This is analogous to structural uncertainty related to choosing a climatology/weighting to be used in the calculation of statistically optimized bending angles.
   Here, these structural uncertainties will be analysed on a global grid to get a better understanding of which areas are more sensitive to the choice of upper cut-off height.

The comparison of RO and GRUAN data has been performed for six upper-air sites, namely

Ny Ålesund, Sodankylä, Lindenberg, Cabauw, Barrow, Southern Great Plains, and Teneriffe, which have a sufficient number of profiles and associated background data available for the years 2014 to 2016. Instead of using near real-time RO data, the reprocessed ROM SAF Climate Data Record for the GRAS satellite is used. Overall, the two observation types compare well, however, there are some unexplained differences at two upper-air sites in high Northern latitudes, as well as at nighttime at the Southern Great Plains location. A warm bias in the GRUAN data product for the RS92 that has been found in other studies has been confirmed here.

The comparison of RO with high-quality sonde types has provided mixed results which complicates drawing a clear conclusion from the results. Overall, the globally distributed RS41 measurements show the best agreement with operational RO measurements. The RS92 sonde, which has been used in Europe, Canada, South America and the Pacific shows a warm bias at high levels which is strongest during daytime and the Lockhead Martin RS (mainly US sites) shows a slight warm bias at high levels during high solar elevation angles (and to a much lesser degree also at other solar elevation angles) and otherwise tends to be slightly too cold. The Modem M10 RS type shows very mixed results depending on the upper-air site, which would lead to a risk when assimilating the sonde type without a bias correction. This sonde types is launched in France but also at sites in a variety of locations. Based on the results, a bias correction prior to the assimilation is advisable for most sonde types and solar elevation angles and such a correction can for example be determined using RO measurements as an unbiased reference.

The spread of dry temperature departures for different upper cut-off impact heights has been calculated on a 10° latitude by 30° longitude grid globally. This structural uncertainty is largest close to the equator and smallest at high Northern latitudes, with a mean value over all grid points of 0.69 K at 10 hPa. While the structural uncertainty decreases towards high Southern latitudes, it increases again above Antarctica.

1 Introduction

Within numerical weather predictions (NWP) vast amounts of measurements are integrated into the model through a process called data assimilation. Among those, radiosonde (RS) profiles have been assimilated for decades and, since 2006, radio occultation (RO) data have been assimilated, demonstrating a positive impact on weather forecasts (see e.g. Healy [11], Poli et al. [19], Rennie [20]). Most observations have biases, which require correction for their integration in NWP models. These corrections can be applied either ahead of the assimilation into the model, or within a variational bias correction scheme (Derber and Wu [6], Dee [5], Auligné et al. [1]). Variational bias correction ingests the measurements and applies a correction which is constrained by so-called 'anchor measurements', that are not corrected as part of the variational bias correction. RO measurements are an important anchor observation as they do not require any correction prior to or within the assimilation. The RO bending angles, which are nowadays assimilated into most NWP models, are calculated from highly accurate measurements of the time delay which a radio signal experiences when passing through the atmosphere (see e.g Kursinski et al. [17, 16] for a description of the method). However, there are also other anchor observations, such as those from balloon-borne RSs, which may require a correction prior to their assimilation. One method to calculate the bias corrections for RSs using RO measurements as a reference and NWP model background as transfer medium is described in Tradowsky [25, 24], Tradowsky et al. [26]. However, there are also other methods to calculate RS biases, for example by analysing direct co-locations between RO and RS (see Sun et al. [23, 22]).

This ROM SAF visiting scientist report will investigate if RO data from the ROM SAF Climate Data Record (CDR) are in agreement with reference-quality RS data produced by the Global Climate Observing System Reference Upper-Air Network (GRUAN). GRUAN data products are reference-quality in that all known biases have been corrected and an estimate of the uncertainty is provided with every datum. The comparison between these two reference-quality datasets can reveal any unknown biases in the GRUAN data set, or potential problems in the RO retrieval. If the two entirely independent datasets agree to within their uncertainties, it shows the consistency of both datasets. It is also possible to robustly point out model biases using GRUAN and RO observations, as it will be shown in Chapter 4.

This project furthermore aims to develop a hierarchy of anchor measurements using RO measurements as the primary anchor (level 1 anchor). Through comparison with RO, chosen high-quality RS types are tested as to whether they are in agreement with RO. If agreement is found, the given RS type may be assimilated into NWP models without applying a correction prior to assimilation. RS types for which agreement with RO is found, build the second level in the hierarchy. If the RO and RS data are not in agreement, a bias correction should be applied prior to the assimilation of the RS profiles. By applying such a correction prior to assimilation, these measurements can still anchor observation that are corrected as part of the variational bias correction. Eyre [8] showed that the influence of model biases on the background and analysis biases increases with a decreasing proportion of anchor measurements. Thus, it is essential to leave measurements that can be corrected prior to their assimilation out of the variational bias correction scheme. RO can serve as an unbiased reference to correct these measurements prior to their assimilation. As RS typically show the largest biases during daytime - caused by solar radiation affecting the sensor - an assimilation without correction may be possible for nighttime measurements while daytime measurements may

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT
ROM SAF

require a bias correction.

While the raw RO observations are highly accurate, assumptions that are required in the retrieval lead to uncertainties in the dry temperature (Tdry) which can be calculated from RO bending angles assuming that the humidity in the atmosphere is negligible. Within this study, the humidity is assumed to be negligible when the difference between Tdry and the physical temperature taking into account the humidity (from the UK Met Office NWP system) reaches values below 0.09 K. Thus, Tdry is a valid estimation of the physical temperature where the atmosphere is extremely dry, which is mainly the case above the tropopause, but, depending on the individual profile, can also include levels within the higher troposphere. To study how changes in the assumptions affect the retrieval, it will be investigated how variations in one parameter - the cut-off impact height used within the tangent linear RO retrieval - affect the retrieved Tdry. The resulting spread in Tdry gives an indication of the uncertainty caused by reasonable variations of the cut-off impact height.

This visiting scientist project builds on former projects which developed a tangent linear RO retrieval and develop a bias correction scheme to be tested within the UK Met Office global NWP system. The ROM SAF Visiting Scientist Report Report 26 (Tradowsky [25]) described a new technique to calculate RS temperature biases in the stratosphere, using RO bending angles as the reference. Instead of using direct co-locations, the Met Office Unified Model is used as a transfer medium. The difference between the background departures of RO and RS estimates the RS temperature bias, under the assumptions that the RO observations are unbiased and that the model background cancels out when the differencing is performed. For this purpose RO bending angle departures are propagated into Tdry departures using a tangent linear Tdry retrieval, which, rather than using statistical optimization, calculates the Tdry departures from a subset of the bending angle departures. This is achieved by setting the bending angle departures above an impact height of 35 km to zero. This method allows better control of the flow of information than using statistical optimization, and it is shown, that the applied cut-off gives results similar to smoothing the bending angles with a climatological value as it is done in the conventional (non-linear) Tdry retrieval chain (Tradowsky et al. [26]). There are ongoing discussions within the ROM SAF about the effect of applying a cut-off to the bending angle departures. Chapter 6 of this report aims to shed some light on the size of the effect of this cut-off.

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

ROM SAF

2 Data and model

Within this study RS data, RO data, and model fields from the UK Met Office Unified Model are used and these datasets are described in the following sections.

2.1 Radio Occultation data

This study mainly focuses on analysing RO data from EUMETSAT's GNSS Receiver for Atmospheric Soundings (GRAS) satellites - GRAS-A and GRAS-B. However, for the comparison of RO with high-quality RS data in Chapter 5, additionally the UCAR near real-time data of the US-Taiwanese Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) - FORMOSAT-3 mission (here referred to as COSMIC) are used. For GRAS measurements, two different datasets are used. In Chapter 5 the near real-time GRAS data provided by the ROM SAF (originating centre DMI, ID 94) are used. These RO data and their associated model background fields are extracted from the UK Met Office Observation Processing System (OPS). ROM SAF's RO data, which are also operationally assimilated at the UK Met Office, are based on near real-time data processed by EUMETSAT, however, applying additional quality control.

For the comparison between RO and GRUAN, GRAS data which have been reprocessed by the ROM SAF are used. This Climate Data Record (CDR v1.0, Product ID: GRM-29-R1, `doi:10.15770/EUM_SAF_GRM_0002`) is processed with the same software for all years, which is not given for the near real-time data. Therefore, using the CDR, it is possible to analyse data from several years, in this case from 2014-2016 which is the time frame for which CDR, GRUAN data, as well as NWP background fields were available. To use the CDR data instead of near real-time data as they are available in the monitoring files extracted from the OPS, CDR profiles need to be mapped to operational GRAS profiles in the monitoring files. For this matching a maximum time difference of 2 minutes and a maximum combined difference in longitude and latitude of 0.5° is applied. The criteria were chosen based on careful evaluation of the available profiles, i.e. using criteria that find a match, but excluding double matches and wrong matching of profiles. The slight mismatch between the longitude and latitude given for near real-time versus reprocessed occultations might be caused by different software versions in the ROM SAF processing.

The ROM SAF CDR v1.0 also contains COSMIC CDR data (derived from low-level reprocessed phases and amplitudes provided by UCAR), which generally could also be used in the RO versus GRUAN comparison. Therefore, also the ROM SAF COSMIC CDR data were mapped to the operational data. However, to obtain regular matching, larger differences in longitude and latitude were required. A combined latitude/longitude difference threshold of 2° has been applied. When analysing the resulting data set it occurred, however, that the Tdry statistics show a weird shape which is very different to the statistics calculated from GRAS only. This should not be the case, as it has been shown in Tradowsky [25] that the statistics are only minimally affected by the combination of both satellites. This indicated that there has been an issue with the mapping of CDR data to the operational UCAR COSMIC data. To avoid this issue the RO - GRUAN comparison in Chapter 4 uses GRAS CDR data only.

2.2 Radiosonde data

The operational RS data as they have been submitted by upper-air sites to the Global Telecommunications System (GTS) are used, as well as a reference-quality RS data product from a subset of GRUAN sites. For the comparison of RO with high-quality RS data in Chapter 5, RS measurements from the operational network comprising approximately 800 globally-distributed upper-air sites are used. These sites mainly operate for the purpose of NWP and the RS profiles are usually corrected by the vendor software at the upper-air sites - a process which commonly uses black-box software. Therefore, it is assumed that we do not know what corrections have been applied at operational upper-air sites. The vendor corrections will also change over the years which implies that any hierarchy of anchor observations developed may change over time. More detail about the operational upper-air network and changes that it is currently undergoing is given in Section 2.2.1

GRUAN processed measurements, in contrast, have a well documented correction of all known biases applied Dirksen et al. [7] that does not change within one version of the dataset, and include a traceable estimate of the uncertainty on every value. These uncertainty estimates can be propagated through the calculations as is described in Section 2.2.2. Agreement or consistency between GRUAN and RO data will be investigated in Chapter 4.

2.2.1 The operational radiosonde network

Currently the global RS network is transitioning from alphanumeric TEMP code to high-resolution BUFR code (see Ingleby and Edwards [15]). Currently around 31% of sites are not providing data in BUFR format, while the others provide one of the three options for BUFR data:

- High-resolution BUFR data. Native BUFR profiles which are based on the raw measurements from the RS. These profiles have many hundred vertical levels and are the most valuable option of data for NWP.

- Low-resolution BUFR data. Native BUFR profiles which are based on raw measurements from RS. However, low-resolution BUFR data include significantly fewer vertical levels than the high-resolution BUFR profiles. There may, therefore, be an issue with deciding if a RS profile is available in low-resolution BUFR or in TEMP format.

- Non-native BUFR data. TEMP RS profiles transferred into BUFR format. These profiles may be worse than the TEMP reports and weather forecast centres may prefer to use the TEMP profiles instead.

Figure 2.1 gives an overview of the current status of the TEMP to BUFR transfer. Some areas, such as South-East Asia are mostly providing both TEMP profiles and reformatted (non-native) BUFR profiles (compare 2.1(a) and (b)).

The UK Met Office started assimilation of BUFR data one country at a time after checking that the quality is sufficient. Currently, however, for some sites the BUFR data are rejected and TEMP profiles are assimilated instead, as it has been the case for the German upper-air site in Lindenberg.

Choosing the correct RS profiles to be included in the analysis has been a complication in

this report as some sites have a variety of profiles in the dataset extracted from the Met Office Observation Processing System (OPS). When, for example, looking a site 71836, there are instances where four different profiles are available for the same launch time. Of course, only one of them should be used here to ensure that each RS profile is included only once into the statistics. Typically the profiles available for the same launch time include:

- A low-resolution TEMP profile

- A high-resolution BUFR profile

- A BUFR profile which only reaches 100hPa

However, even more profiles, or profiles that do not agree with each other, as it has occurred in one instance at site 71836, may be present in the dataset. It is unclear why this is the case. Unfortunately, a variety of other complications with the dataset occurred, some of which may be taken care of by the new RS monitoring system under development at the UK Met Office. Below, some issues that have been seen are summarised. These issues have been discussed with the team at the UK Met Office, however, as the team is currently changing, it was not possible to unravel or solve some of the underlying issues during the course of this research project.

- For the upper-air site in Lindenberg, it seems to be the case that TEMP data are assimilated instead of available high resolution BUFR data.

- For some sites a full profile as well as a profile up to $100\,\mathrm{hPa}$ has been assimilated (Lawrence Morgan, personal communication). This should not be the case as it increases the weight of the observation.

- For the upper-air site 06260 (which is the GRUAN site Cabauw), only profiles up to $100\,\mathrm{hPa}$ are available in the dataset.

- For the upper-air site 02836 (which is the GRUAN site Sodankylä) usually three profiles are available, i.e. a TEMP, a BUFR, and a BUFR to $100\,\mathrm{hPa}$. However, for some cases instead of the BUFR profile another low resolution profile is included in the dataset. While reaching the same pressure level as the TEMP profile, this has less levels than the TEMP profile. The reason for this is not known.

For this report, the profile which reaches the lowest pressure level (highest altitude) will be used. If there are two profiles that reach the same lowest pressure, the one with the most valid levels will be used. If, after these test, there are still several profiles that are equally likely to be the correct one, none of them is used to avoid including incorrect data.

All these issues will require a careful investigation by the new RS monitoring team at the UK Met Office. It is likely that the value of RS is currently not fully exploited, which becomes especially obvious as some of the sites for which issues seem to exist are part of the GRUAN network which has highly trained staff and is very likely to provide high-quality RS observations.

The handling of the profiles will in general be the same independent of whether they are given in BUFR or TEMP, however, if the RS profile is available in low resolution, statistics will only be available for a subset of the pressure levels given below.

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT
ROM SAF

Analysed pressure levels: 1000, 950, 925, 900, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 200, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 48, 46, 44, 42, 40, 38, 36, 34, 32, 30, 28, 26, 24, 22, 20, 18, 16, 14, 12, 10, 9, 8, 7, 6, 5

For comparison, the standard pressure level in the TEMP format are given below. If for a chosen site, some of the RS profiles have high and others low resolution, the sample size will differ between the pressure levels.

Analysed pressure levels TEMP and BUFR with less than 400 levels: 1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10.

A possible interesting future activity could be a denial study, i.e. to run the analysis for a high resolution data set, then sub-sample to a lower resolution and analyse differences in the results.
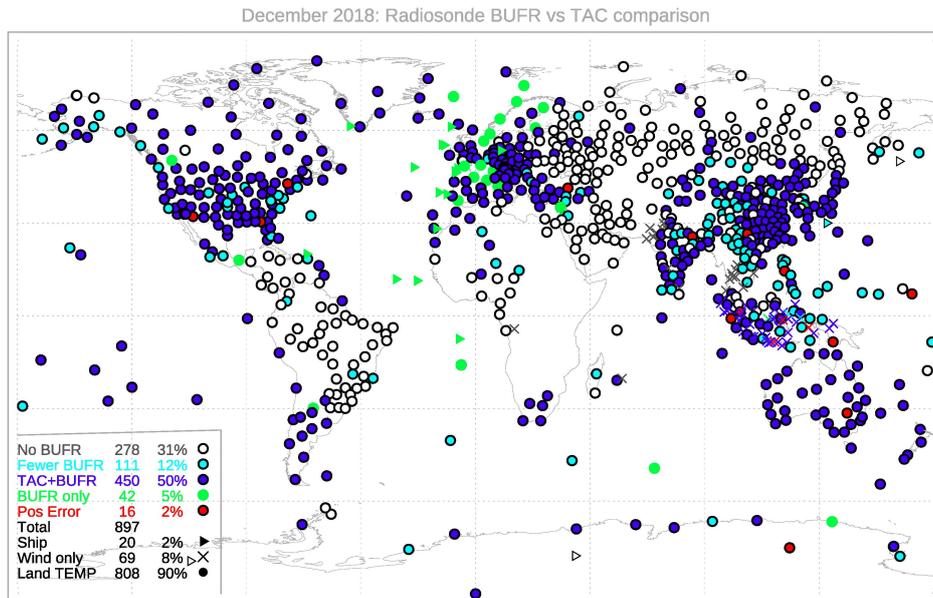
(a) BUFR vs TEMP (TAC) availability
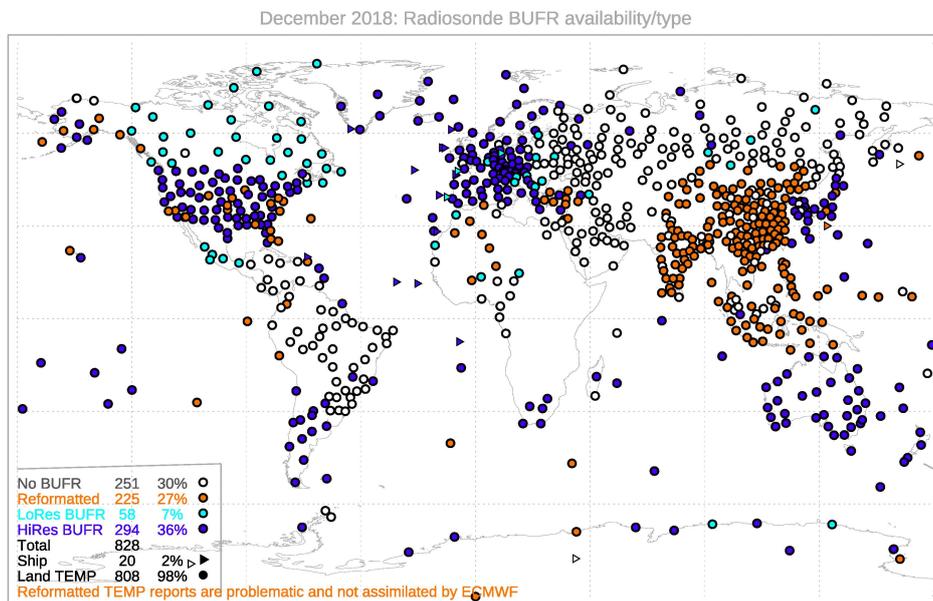
(b) Types of BUFR submitted to GTS

Figure 2.1: *Availability of RS data through the GTS as of December 2018. Figures from Bruce Ingleby,* `https://confluence.ecmwf.int/display/TCBUF/Data+availability`*, last accessed: 16 January 2019*

2.2.2 The Global Climate Observing System Reference Upper-Air Network (GRUAN)

The GCOS Reference Upper-Air Network (GRUAN; www.gruan.org) was established to fill a long-recognized need of the international climate research community for reference-quality measurements of upper-air essential climate variables. GRUAN was established following a series of meetings held jointly by GCOS and the United States National Oceanic and Atmospheric Administration (NOAA) between 2005 and 2007. As detailed in GCOS-112 [9], the purpose of GRUAN is to: 1. Provide long-term, high-quality climate records; 2. Constrain and calibrate data from more spatially-comprehensive global observing systems (including satellites and current RS networks); and 3. Fully characterize the properties of the atmospheric column.

As of late 2018, GRUAN comprises 28 sites (see Fig.2.2), of which 17 have a GRUAN data product for the Vaisala RS92 RS and 10 have been certified. All 28 sites measure vertical profiles of essential climate variables such as temperature, pressure, and water vapour. GRUAN data products are reference-quality measurements in that all systematic biases have been accounted for and measurement uncertainties are provided as part of the GRUAN data products and these are traceable to internationally recognized measurement standards (Immler et al. [13]). Currently, only a GRUAN data product for the Vaisala RS92 RS is available, but data products for other RS, ozonesondes, frost point hygrometer, lidar, microwave radiometer and the Global Navigation Satellite System (GNSS) precipitable water vapour are under development.
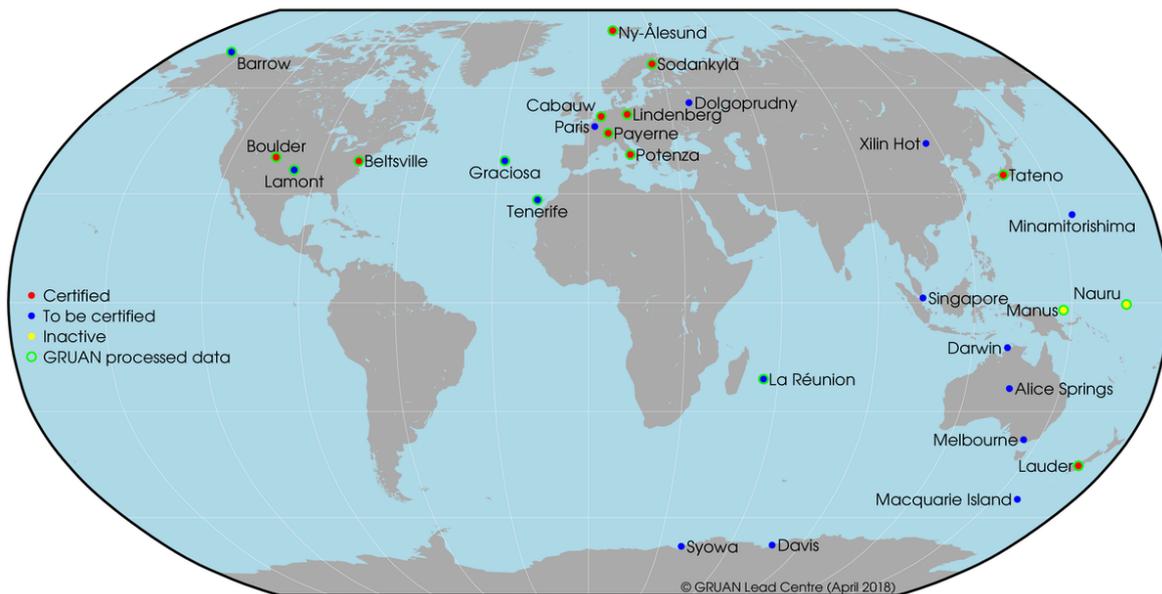
Figure 2.2: *Upper-air sites that contribute to GRUAN. Courtesy GRUAN Lead centre, figure from* `http://www.dwd.de/EN/research/international_programme/gruan/sites.html`*, last accessed: 19 December 2018*

Within this project the Vaisala RS92 GRUAN data product version 2 will be analysed for the years 2014 to 2016, to be coincident with the years for which reprocessed ROM SAF GRAS data and coincident model backgrounds are available.

The GRUAN minus model departures are calculated on up to 80 pressure levels from 1000 hPa

to 5 hPa and can subsequently be compared with GRAS RO departures from model backgrounds. Some of the GRUAN sites also serve as operational upper-air sites, which means that the vendor supplied RS product for the same launch is provided to the Global Telecommunication System. This would allow the measurements made at a GRUAN site (not the GRUAN data product which is not available in near-real time) to be assimilated into NWP models. Currently, effort within the GRUAN community is ongoing to receive a WMO station ID for all GRUAN sites and to enable all sites of the network to submit data to the Global Telecommunication System. Table 2.1 gives an overview of all GRUAN sites that provide a Vaisala RS92 product. Where available the WMO station IDs is included in the table.

For the GRUAN sites Ny Ålesund (NYA), Sodankylä (SOD), Lindenberg (LIN), Cabauw (CAB), Barrow (BAR), Southern Great Plains (SGP) and Tenerife (TEN) sufficient profiles for the calculation of departure statistics are available, and thus, the data from these stations are analysed here.

Station Name	WMO ID	Lat	Lon	Alt [m]	Sample size
Barrow (BAR)	70027	71.32°	-156.62°	8	**1362**
Beltsville (BEL)		39.05°	-76.88°	53	
Boulder (BOU)	72471	39.95°	-105.20°	1743	
Cabauw/De Bilt (CAB)	06260	52.10 °	5.18°	1	**732**
Graciosa (GRA)	0507	39.09°	-28.03°	30	
Lauder (LAU)	93817	-45.05°	169.68°	370	
Lindenberg (LIN)	10393	52.21°	14.12°	98	**3898**
Manus (MAN)		-2.06°	147.43°	6	
Nauru (NAU)		-0.52°	166.92°	7	
Ny Ålesund (NYA)	01004	78.92°	11.92°	5	**1117**
Payerne (PAY)	06610	46.81°	6.95°	491	
Potenza (POT)		40.60°	15.72°	720	
La Reunion (REU)		55.38°	-21.08°	2200	
Southern Great Plains (SGP)	74646	36.61°	-97.49°	320	**2653**
Sodankylä (SOD)	02836	67.37°	26.63°	179	**1235**
Tateno (TAT)	47646	36.06°	140.13°	25	
Teneriffe (TEN)	60018	28.32°	-16.38°	115	**613**

Table 2.1: GRUAN upper-air sites that have a RS92 GRUAN data product version 2 available. The official WMO station ID is given where available, even if the station ID is currently still missing in the WMO OSCAR Surface system (`https://oscar.wmo.int/surface/index.html`). The station name is written bold if the sample size (number profiles available from 2014 to 2016) is sufficient to calculate the departure statistics. The abbreviation of the station name as used within GRUAN is given in brackets. The SGP site is also known under the name Lamont. Longitude and Latitude values are taken from the GRUAN data product.
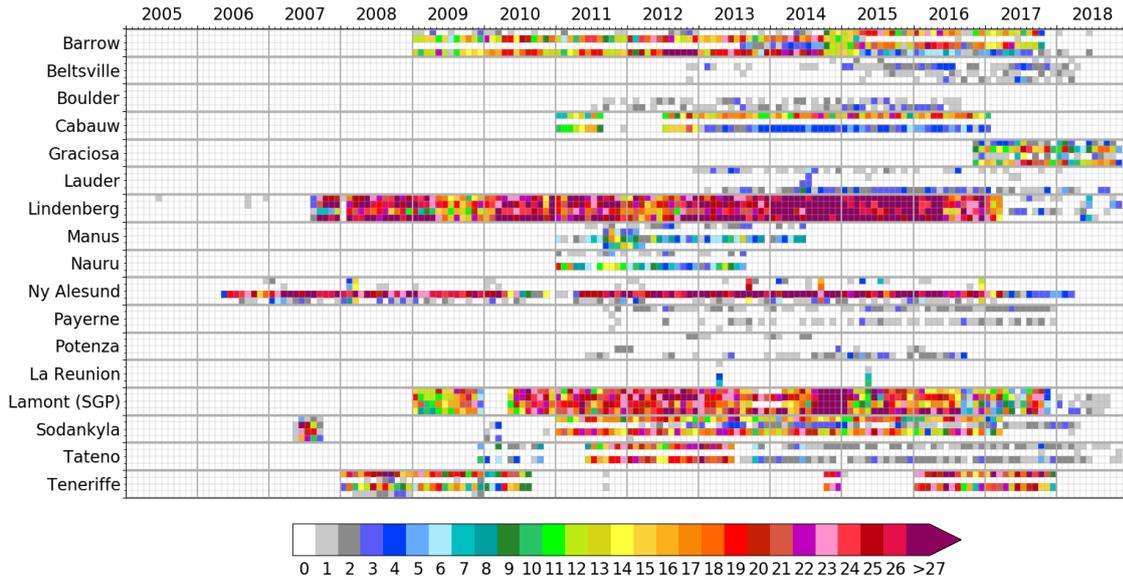
Figure 2.3: *GRUAN data availability at the beginning of December 2018. The colour scale gives the number of RS launches per month (horizontal pixels) and per 6-hour launch time interval (vertical pixels).*

2.3 Model

The Met Office NWP systems are upgraded approximately twice per year. A summary of the main changes to the NWP system between 2014 and 2018 are given in Table 2.2 Most of these changes will have little impact on the suitability of using the model as a transfer medium. One notable upgrade is the change of dynamical core and model resolution on 15 July 2014. The update of the dynamical core is described in detail in Walters et al. [29]. It is also worth noting that the introduction of variational bias correction (VarBC) on 15 March 2016 made a substantial change to the behaviour of the NWP system.

Operational suite version	Date of change	Summary of updates
OS41	25 Sep 2018	- Update to Global Land 8.1 configuration. - Assimilation of cloud-affected microwave observations from ATOVS. - Increased background frequency for observation processing. - Upgrade to UM10.9.
OS40	13 Feb 2018	- En-4DEnVar based covariances with rebalanced hybrid weighting and waveband localisation. - Assimilation of GMI observations. - Observation thinning changes to IASI and ATOVS. - Use NEMOVAR derived Ostia sst and sea-ice data. - Upgrade to UM10.8.
OS39	11 Jul 2017	- Increase horizontal resolution to N1280 (10 km). - Improved snow analysis reconfiguration. - Upgrade to UM10.6.
OS38	8 Nov 2016	- Use ATOVS, and AIRS and CrIS surface channels over land. - Use of geostationary (SEVIRI, MVIRI and AHI) WV channels over low cloud. - Introduction of FY-3B microwave humidity data. -Upgrade to UM10.4.
OS37	15 Mar 2016	- Variational bias control (VarBC) for satellite radiances. - CVT covariances and swapped vertical and horizontal transforms. - New ASCAT soil wetness climatology. - Assimilation of aircraft RH observations. - Upgrade to UM10.2.
OS36	25 Aug 2015	- Migration to Cray XC40 Supercomputer. - Upgrade to UM 10.1.
OS35	3 Feb 2015	- Package of satellite changes. - Tropical cyclone central pressure assimilation.
OS34	15 Jul 2014	- Major UM upgrade including implementation of Global Atmosphere 6.1 and Global Land 6.1. - ENDGame dynamical core. - Various physics upgrades. - Increase in model resolution 25 km to 17 km. - Increase in assimilation resolution from 60 km to 40 km.
OS33	4 Feb 2014	-Migration of operational suite to Rose. -Upgrade to UM8.5.

Table 2.2: Main upgrades to the Met Office global NWP system between 2014 and 2018.

3 Method used to compare radio occultation and radiosonde statistics

The method used here to compare RS with RO data was initially developed to calculate RS temperature bias corrections on a station-by-station basis, as described in Tradowsky [25, 24], Tradowsky et al. [26]. A double-differencing approach, using the Met Office global NWP system as a transfer medium is applied to the mean RS temperature departures and the mean RO Tdry departures. The departures are calculated with respect to the model background fields. For every RS profile at a given upper-air site, the departure from the model background (short-range forecast) is calculated on up to 91 pressure levels ranging from 1000 to 5 hPa. The number of pressure levels on which the profile was calculated depends on the vertical resolution of the profile which was submitted to the Global Telecommunication System for the purpose of assimilation and on the altitude that the sonde reached. Then, the mean RS temperature departure (observation minus background, i.e. O-B) is calculated from all profiles at a given site. For RO data, the The RO bending angle departures from model background are calculated using all GRAS (and COSMIC) RO profiles measured within a circle with a radius of 500 km[1] around the upper-air site. This means, that the method does not rely on direct co-locations in time, as the model background is minimizing the effect of imperfect co-locations as described below and in Tradowsky et al. [26]. However, the method is susceptible to seasonal biases in models for upper-air sites that have a preferential sampling of profiles in a specific season. RS are typically launched at specific times of the day (typically a subset of 00 UTC, 6 UTC, 12 UTC, and 18 UTC). These fixed times can lead to preferential sampling during one season, especially in high latitudes, i.e. the dusk/dawn launches could all be made within two months of the year when the standard launch time coincides with dusk or dawn. As the RO departures are calculated from the RO profiles measured in all seasons, a potentially existing seasonal model bias could influence the results. This has been further investigated in Tradowsky [24].

The RO bending angle is used rather than a retrieval of the temperature as it is a comparably raw variable, which does not include *a priori* information. This is also one of the reasons why most NWP centres nowadays assimilate RO bending angles.

A tangent linear retrieval is used to propagate the bending angle departures to Tdry departures within this project. This retrieval has been developed within a former ROM SAF visiting scientist project, and has been described in detail in Tradowsky [25, 24], Burrows and Healy [2], Tradowsky et al. [26], including a discussion of the upper cut-off impact height above which bending angle departures are set to zero within the retrieval, and the uncertainty associated with choosing this cut-off. Tdry provides a valid estimate of the physical temperature only at levels of the atmosphere where water vapour is negligible and, thus, the Tdry departures are calculated for the stratosphere and dry areas of the upper troposphere only. The lowest dry level is calculated from the model humidity individually for every RO profile allowing the use of RO measurements as low in the atmosphere as is reasonable for a given profile. Thus, at high latitude sites, the comparison between RO and RS can be performed down to lower altitudes than at low latitude sites based on the varying altitude of the tropopause (about 12 km at high latitudes and about 18 km in the Tropics).

The difference between the two observation types is calculated on each pressure level as a

---

[1] For a discussion of the choice for this vicinity radius see Tradowsky [25].

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

ROM SAF
EUMETSAT

double-difference expression, i.e.:

$$\overline{O_{RO} - B_{RO}} - \overline{O_{RS} - B_{RS}} \qquad (3.1)$$

where O is the observation and B is the background, forward-modelled into observation space. The assumption in our analysis is that that B_{RO} and B_{RS} are equally representative of the true values at the RO and RS locations respectively, i.e the central assumption is, that the NWP forecast bias does not vary between the RO and RS locations. This is a more robust assumption compared to direct co-locations between measurements, which are made assuming that the atmosphere does not vary over the separation distance. A similar double differencing approach is used by Haimberger et al. [10] to homogenize RS temperature records. The method further assumed that the observation bias is constant over the separation distance, i.e. the RS would show the same bias at a different location within a circle of 500 km.

Readers interested in the details of the RO retrieval are encouraged to read Tradowsky [24], Burrows and Healy [2]. A sensitivity study investigating for example the sensitivity to the vicinity radius around the upper-air site for which the RO departures are calculated can be found in Tradowsky [25], and an investigation of the uncertainty caused by setting high-level bending angle departures to zero is described in Tradowsky [24].

Over the course of this ROM SAF visiting scientist project, there has been some discussion between ROM SAF members as to whether or not using a tangent liner retrieval including an upper cut-off impact height is the best possible way to compare two observation types, and, if this happens to be the best-possible way, which cut-off impact height would be ideal. While the tangent linear retrieval has a variety of benefits for the comparison of observation types, it is not part of this project to decide what would be the ideal method. However, one important benefit of the tangent linear retrieval is the absence of any need for statistical optimisation of bending angle profiles at high altitudes. Such statistical optimisation is essential in the conventional RO retrieval and is typically applied above altitudes of 35-40 km. Depending on the climatology used in the statistical optimisation, the resulting dry temperature profiles differ. Analysing the the mean differences between Tdry profiles from individual centres with the inter-centre mean value where the retrievals are calculated by different processing centres based on the same bending angles gives an estimate of the structural uncertainty, see Ho et al. [12], Steiner et al. [21]. For global RO profiles, this uncertainty based on the choices made in the retrieval, such as the climatology and smoothing algorithm, was found to be around 0.4 K at 10 hPa, see Ho et al. [12]. The spread of dry temperatures departures calculated from using different upper cut-off impact heights between 35 and 55 km is in the same order for global COSMIC FM6[2] profiles, but differs significantly from site to site, as has been shown in Tradowsky et al. [26].

---

[2] FM stands for Flight Module; COSMIC has six FM's, i.e. individual satellites

4 Analysis of consistency between RO and reference-quality radiosonde measurements from GRUAN

RO data, as well as GRUAN processed RSs are commonly used as reference observations of the upper-air (see e.g. Ladstädter et al. [18], Calbet et al. [3], Carminati et al. [4]). The comparison of these two observation types, therefore, can reveal problems in the RS bias correction or in the RO retrieval. If the uncertainties of both observation types are estimated correctly, and all biases are corrected, then the two observation types should agree within their uncertainties. The double differencing method applied here, which uses model background fields of the UK Met Office NWP field as transfer medium, is assumed to get rid of any differences caused by imperfect co-location, see Tradowsky et al. [26]. While the assumption of a not-varying background error required here might not be perfect, it is a less severe assumption compared with the assumption of a not-varying atmospheric field which is required for direct co-locations [see e.g. 23, 22]. Therefore, statistics can be calculated from a smaller sample size than for direct co-locations as the standard deviation is significantly smaller as has been shown in Tradowsky [25].

In the following sections RO profiles from the GRAS satellites are compared with GRUAN profiles by selecting all GRAS occultations within 500 km of the respective GRUAN site for the years 2014-2016. Rather than using the near-real time data, the ROM SAF CDR is used. While it would increase the sample size to use COSMIC data from the CDR as well, this led to some issues and, as it is currently not clear why the departure statistics look so different when including COSMIC data, they are excluded from the comparison. Based on results from Tradowsky [25], it is expected that the departure statistics would look very similar. One reason for the large differences in RO departure statistics when using only GRAS in comparison to GRAS and COSMIC, may be a difficulty that occurred when mapping the reprocessed COSMIC CDR data to the operational data for which a background profile is available in the data set of NWP background fields that has been used here. As the model background field is required for the analysis, reprocessed CDR data are mapped to their near real-time counterparts. This mapping is done based on time, longitude, and latitude of the RO profile.For COSMIC, the co-location criteria required significant relaxing in comparison to GRAS in order to get regular matches. This leads to the possibility of wrong matching, i.e. the wrong RO profile could be chosen so that background and measurement are not agreeing in time and location. To avoid the risk of the results being affected by incorrectly matched occultations for which the background would be at a different location then the observation, COSMIC data were not used.

Propagation of GRUAN uncertainties   Uncertainties in the GRUAN data product have a variety of sources, such as the uncertainty in the calibration, the uncertainty in the radiation correction of the warm daytime bias which is typical for RS instruments, or the statistical uncertainty estimated from the standard deviation. These uncertainties are estimated individually for every value, i.e. every temperature profile has an associated, individual uncertainty profile. This per-value estimate of the uncertainty is used here and is interpolated to the required pressure level, which is discussed in detail in the ROM SAF Visiting Scientist Report 31 Tradowsky [24]. Some of the GRUAN uncertainty components, such as the standard deviation, can be decreased through the averaging of many profiles, but how should uncertainties

in the corrections decrease by averaging, even if the correction relies on the same principle each day? As such an decrease of uncertainties through averaging only seems realistic for those parts of the uncertainty that are completely uncorrelated in time, the total uncertainty given in the GRUAN profiles will be divided into two parts - one that is assumed uncorrelated in time, and one that is assumed to be correlated in time and for which averaging will not diminish the uncertainty by the square root of the number of profiles.

Within the GRUAN data product three components of the GRUAN uncertainty budget are provided (see Dirksen et al. [7]), i.e. the total uncertainty (variable 'u_temp'), the random uncertainty of temperature (variable 'u_std_temp'), and the absolute uncertainty of the temperature sensor calibration (variable 'u_cor_temp'). These uncertainties are given for 1σ (k=1).

To separate the correlated and uncorrelated uncertainty components, the random uncertainty based on the standard deviation is subtracted from the total uncertainty, and the rest of the uncertainty is assumed to be correlated in time. Within the GRUAN community work is underway do develop a more detailed breakdown of the uncertainty correlations and this could be very useful in future research. Following the naming convention in Dirksen et al. [7], this can be expressed as:

$$u_{uncorr,i} = \frac{u_{std\_temp,i}}{\sqrt{N'}} \tag{4.1}$$

with $N' = 11$ being the effective sample size of the kernel of the smoothing filter applied in the GRUAN processing (see Dirksen et al. [7]), and

$$u_{corr,i} = \sqrt{u_{temp,i}^2 - u_{uncorr,i}^2} \tag{4.2}$$

It is possible that there are other components of the uncertainty which may be assumed uncorrelated in time, however, these can't be separated based on what is given in the data files.

The correlated/uncorrelated uncertainty components are then propagated separately following the suggestion by Immler et al. [13] for correlated and uncorrelated components of the uncertainty. The uncorrelated uncertainty of the mean temperature is calculated as:

$$\overline{u_{uncorr}} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} u_{uncorr,i}^2} \tag{4.3}$$

where the uncertainty decreases with $1/\sqrt{N}$. For the component of the uncertainty which is assumed to be correlated in time, the uncertainty on the mean is calculated as:

$$\overline{u_{corr}} = \frac{1}{N} \sum_{i=1}^{N} u_{corr,i} \tag{4.4}$$

The total uncertainty on the mean temperature is calculate from the correlated and uncorrelated component as:

$$\sigma_{GRUAN} = \sqrt{\overline{u_{corr}}^2 + \overline{u_{uncorr}}^2} \tag{4.5}$$

The variable σ_{GRUAN} gives an upper bound for the uncertainty on the averaged GRUAN profiles as potentially some of the components that are now treated as being correlated in time, may be uncorrelated. If, during the comparison of RO and GRUAN statistics the two observation types show a much closer agreement than being just inside the uncertainty bounds, this may indicate an overestimation of uncertainties.

The uncertainty in GRUAN temperatures is largest for daytime measurements as RS temperatures are typically biased warm during daytime due to solar radiation warming the sensor. While a correction for the radiation bias has been applied to the GRUAN data, this correction introduces an uncertainty which is part of the overall uncertainty budget. Former studies (e.g. Ladstädter et al. [18]) have indicated a remaining warm bias in daytime GRUAN data, which may be caused by an underestimation of the radiation bias in the GRUAN bias correction scheme. In the following sections, we will analyse the difference between the GRAS and GRUAN data for the GRUAN sites Lindenberg, Southern Great Plains, Barrow, Ny Ålesund, Sodankylä, Tenerife, and Cabauw. These are the upper-sir sites that have a reasonable amount of data in the years 2014-2016. Some GRUAN sites, such as Lauder, New Zealand launch only one RS per week, which does not lead to a sufficient sample size for a meaningful comparison.

In the context of GRUAN, consistency between two independent measurements m_1 and m_2, with uncertainties u_1, u_2 is achieved if

$$|m_1 - m_2| < k \cdot \sqrt{u_1^2 + u_2^2} \qquad (4.6)$$

is true for k=1 (Immler et al. [13]). k is the coverage factor "which determines an interval about the mean value as a multiple of standard uncertainty" (Immler et al. [13]). If Eq.4.6 is true for $k = 2$ the data are in statistical agreement.

4.1 Lindenberg

The GRUAN station in Lindenberg, Germany, is operated by *Deutscher Wetterdienst*, the German weather service and is situated at the GRUAN Lead-Centre. RSs are launched four times daily and the Vaisala RS product is submitted to the GTS regularly. As of 2018, Lindenberg is launching the Vaisala RS41 instrument, but in the years 2014-2016 which are analysed in this section, the Vaisala RS92 has been operational.

Figure 4.1 shows the departure statistics with respect to model background fields for GRUAN RS data (pink) and GRAS RO data (blue), which have been extracted from the ROM SAF CDR archive. The difference between the statistics is shown in green including uncertainty bars for this double difference which are calculated as: $\sqrt{(\sigma_{RO})^2 + (\sigma_{GRUAN})^2}$. The GRUAN uncertainties at the k=1 level include the full uncertainties as described above, while the RO uncertainties only include the sampling uncertainty based on the standard deviation of departure profiles. Data for the years 2014-2016 are averaged to produce the statistics shown in these plots. The statistics are calculated separately for four solar elevation angle (SEA) ranges, i.e. high, low, dusk, night given in Table 4.1.

The RO Tdry departure is only meaningful in the dry part of the atmosphere. Therefore, the humidity in the atmosphere is analysed for every individual RO profile and the data are only used as far down in the atmosphere as is reasonable. This leads to a decreasing RO sample size at lower levels. While those profiles that are still available in dry conditions are valid, they may be sampled in preferential conditions and not be representative for the whole sample. Therefore, the lower cyan coloured horizontal line gives the pressure level above which at least 80% of the RO profiles are available. The upper cyan coloured horizontal line shows the last pressure level at which at least 80% of the GRUAN profiles are available. The amount of GRUAN profiles decreases with altitude due to balloon burst. The pressure range for which a comparison between GRUAN and RO is possible differs from site to site based on the humidity in the atmosphere, as well as on the balloon size used for RS ascent.

At the upper-air site in Lindenberg most balloons reach high altitudes and a meaningful comparison is possible to even above 10 hPa for high SEAs (no upper cyan coloured line). Consistency between RO and GRUAN is found at most levels for Lindenberg, however, there is a tendency for the GRUAN temperatures to be warmer at low pressure levels for high and low solar elevations. As the uncertainties are also largest for high solar elevation angles, the deviation between the observation types in mainly still within the uncertainty range of the averaged GRUAN profiles.

Name	SEA range [degree]
High	SEA >= 22.5°
Low	7.5° <= SEA < 22.5°
Dusk	−7.5° <= SEA < 7.5°
Night	SEA < −7.5°

Table 4.1: *Solar elevation angle (SEA) ranges.*
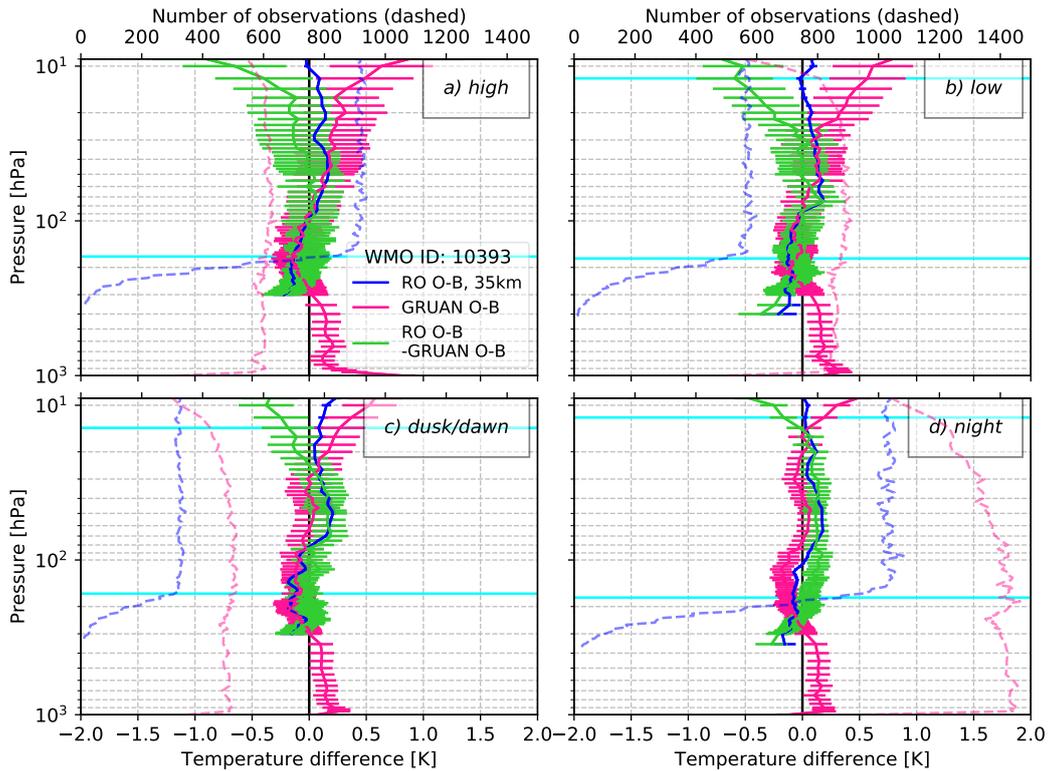
Figure 4.1: *Departure statistics for GRUAN (pink), GRAS RO (blue), and the difference between the two observation types (green) for different solar elevation angle ranges in a) to d). The sample size is given in dashed lines, and the cyan coloured horizontal lines give the highest/lowest pressure level for which 80% of the RO data or 80% of the RS data are available. The GRUAN uncertainty is calculated as described in Section 4, the RO uncertainty bars are standard errors (sampling uncertainty), taking into account the sample size and the standard deviation of the departures. Uncertainty bars of the difference is calculated as* $\sqrt{(\sigma_{RO})^2 + (\sigma_{GRUAN})^2}$.

4.2 Southern Great Plains

The GRUAN site Southern Great Plains in Lamont, Kansas, is a super-site operated by the US *Department of Energy* as part of the *Atmospheric Radiation Measurement* program. RSs are typically launched at least twice daily, however, they do not reach very high altitudes due to the comparably small balloons that are used for most launches. Therefore, the comparison between GRUAN and RO is possible only up to the altitude of the 30 hPa pressure level, see Fig. 4.2. For all SEA ranges, there is a tendency for the double difference between the observation types to be positive, which is equivalent to the GRUAN data giving a colder temperature than RO, however at many levels, the difference is still within the uncertainties. The effect is the least distinct at dusk and the most pronounced at high SEAs and at nighttime. For high SEAs the sign of the double difference reverts at the highest levels, however, an insufficient amount of profiles ia available at these levels and therefore this result is not robust.
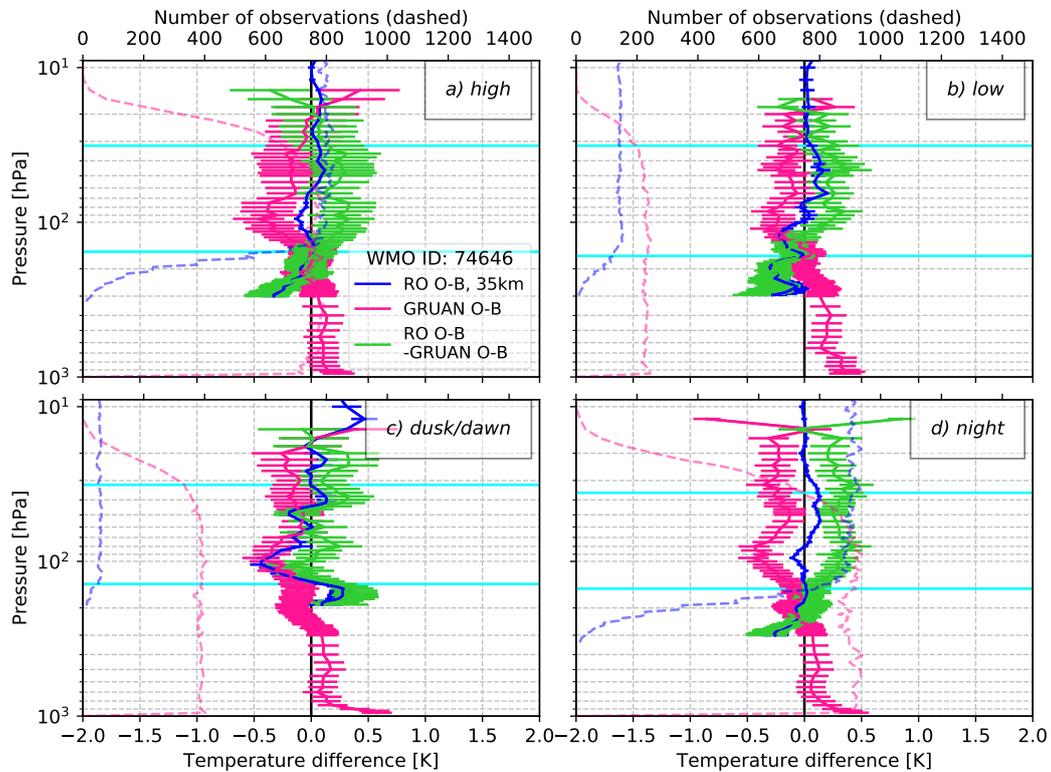
Figure 4.2: *As Fig. 4.1, but for the GRUAN site Southern Great Plains.*

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT
ROM SAF

4.3 Barrow

The GRUAN site in Barrow, Alaska, is operated by the US *Department of Energy* as part of the *Atmospheric Radiation Measurement* program. As for the Southern Great Plains site, the balloons do typically not reach high altitudes and the range of pressure levels to be analysed reaches from about 150 hPa to 40-50 hPa. In this range, observations are in agreement within the uncertainties. For high and low SEA's there is a tendency for the departure statistics to deviate at high altitudes as the GRUAN observation minus background becomes more positive. This is the effect expected from a radiation bias. However, as the sample size for RSs is diminished at these altitudes, a quantitative analysis is not possible.
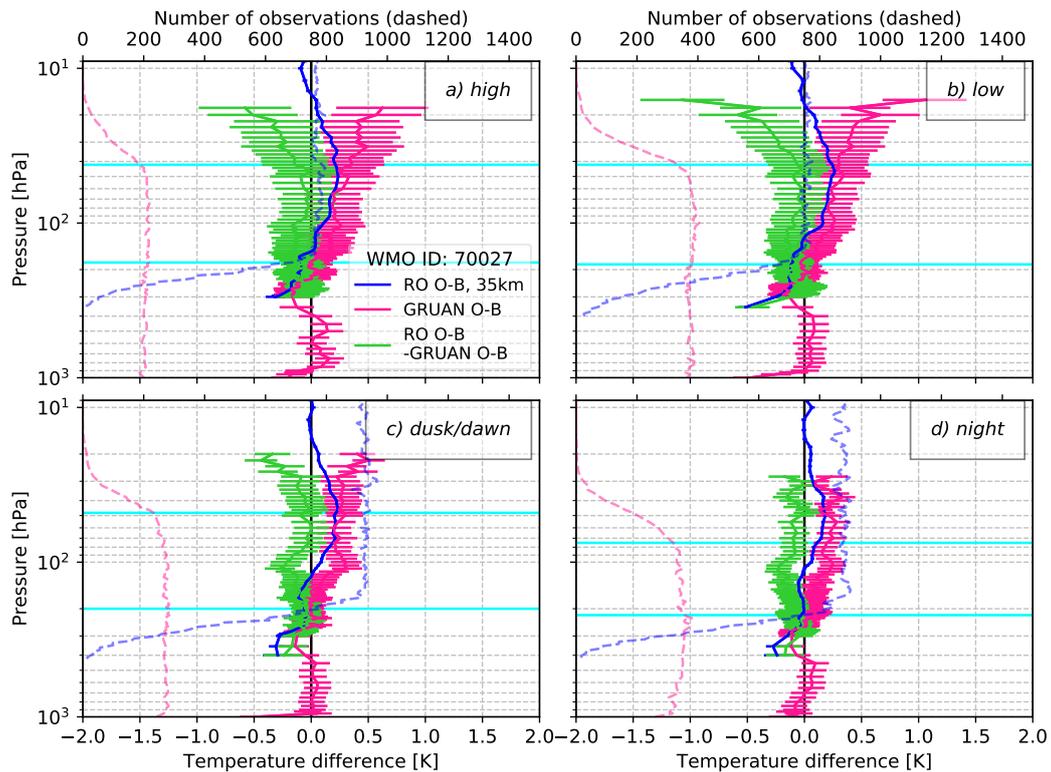
Figure 4.3: As Fig. 4.1, but for the GRUAN site Barrow.

4.4 Ny Ålesund

The upper-air site in Ny Ålesund, Svalbart, Norway, at a latitude of 78.92° is the northern-most site of GRUAN. The station is operated by *Alfred-Wegener-Institut* based in Potsdam, Germany.

RSs in Ny Ålesund typically reach at least 10 hPa and therefore a comparison with RO is possible to nearly 10 hPa, depending on SEA range analysed. Due to the high latitude of this site, the tropopause is comparably low and a comparison is possible at altitudes above the 200 hPa level. For sites at lower latitudes, comparisons are typically only possible at altitudes above the 150 hPa level.

At Ny Ålesund GRUAN and RO data are consistent between 200 hPa and about 20-30 hPa. At altitudes above these pressure levels, the observation types do not agree for any of the SEAs, but always exhibit a negative double difference (green line). While such a negative difference between RO and GRUAN profiles could be caused by a remaining warm radiation bias in the GRUAN data during daytime, it is unclear why the negative difference is visible also at nighttime.

Interestingly, the nighttime statistics show a model bias which is visible in the departure statistics of both observation types. Between 150-100 hPa both, RO and RS show a negative departure, and between 100-40 hPa, both observation types show a positive departure from the model background. The observation types agree to within their uncertainties with each other (they are consistent). This is an example of how two completely independent observation types can be used to confidently exhibit issues with other data types such as model fields.
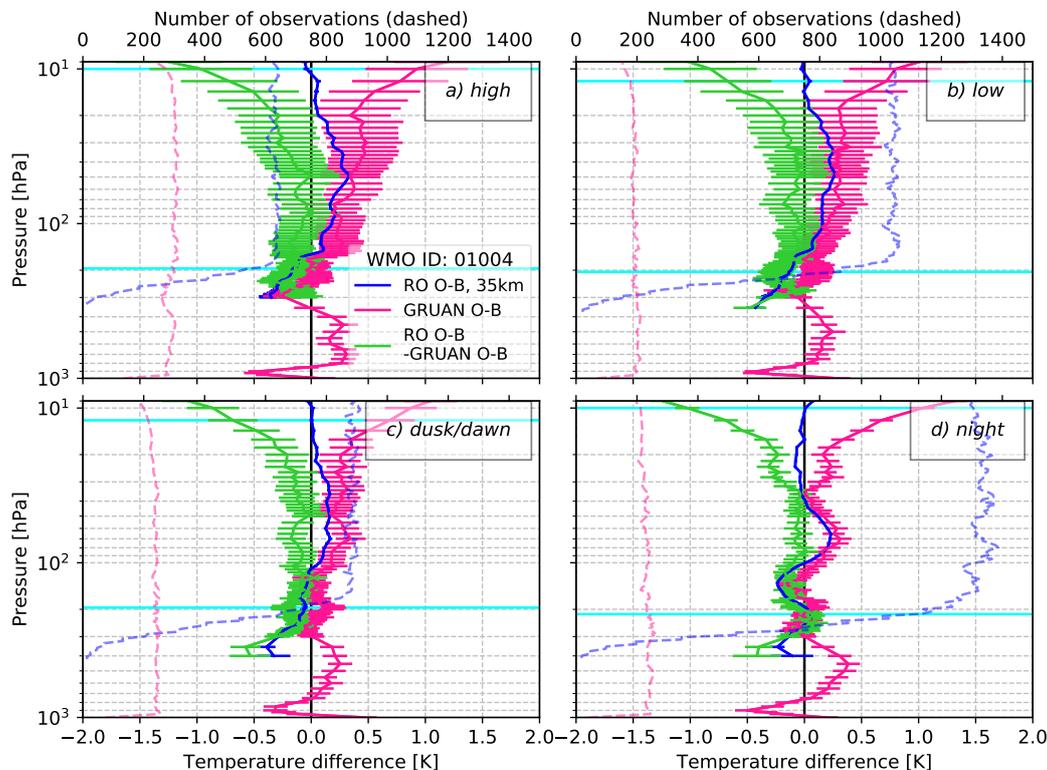
Figure 4.4: *As Fig. 4.1, but for the GRUAN site Ny Ålesund.*

4.5 Sodankylä

The GRUAN site in Sodankylä, Finland, is operated by the *Finnish Meteorological Institute*. For the GRUAN site Sodankylä, a comparison with RO is possible from about 150 hPa to 40-30 hPa depending on the SEA. At higher altitudes, less than 80% of the GRUAN profiles are available due to balloon burst. The model biases that have been seen at Ny Ålesund are also visible at Sodankylä, where the RO and GRUAN departures show an even better agreement with each other at nighttime up to about 30 hPa. Similarly to Ny Ålesund a negative double difference is visible at all SEA's at high altitudes. As less than 80% of the GRUAN data are available at these high altitudes, the results may, however, be affected by sampling biases.

This behaviour of a negative double difference for all SEAs may be a feature re-occurring at Arctic sites, as it is also indicated at Barrow. This feature would deserve further analysis. In Tradowsky [24] GRUAN and operational RO data have been analysed for summer and winter separately to analyse if a seasonal model bias might affect the results. However, no clear conclusion was found.
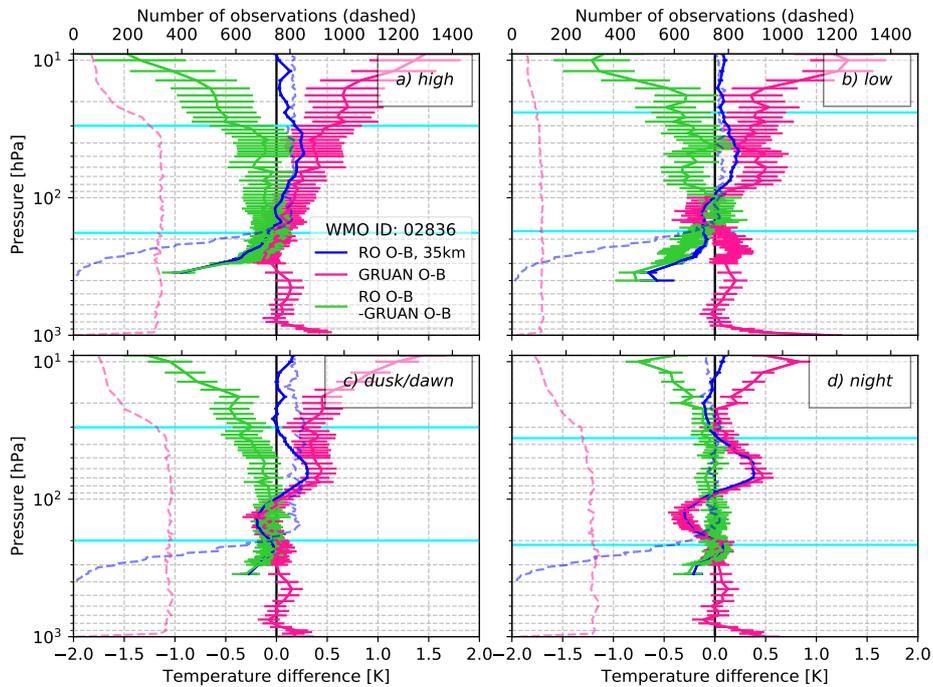
Figure 4.5: *As Fig. 4.1, but for the GRUAN site Sodankylä.*

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT ROM SAF

4.6 Tenerife

The upper-air site in Tenerife has recently joined the GRUAN network and is maintained by the *Spanish Weather Service, AEMET*. A GRUAN data product is available for some years from 2008 onwards. Data are available for all of 2016 and parts of 2014, however there are only very few GRUAN RS profiles for Tenerife in 2015.

Statistics could only be calculated for high SEAs and nighttime launches which can be compared between 150 hPa and 20 hPa. Given the large uncertainties in the GRUAN daytime data, the observation types are found to be consistent at daytime. At nighttime, the uncertainties are smaller and the observation types are consistent only at some pressure levels. Interestingly, RO and RS generally show the same pattern, however, with a stronger development in the GRUAN profile, i.e. where the GRUAN departure is negative, the RO departures are also negative, but with a smaller absolute value. This indicates model biases which are picked up by both observation types. RO profiles have a horizontal resolution of approximately 300 km while RS profile provide point measurement. Thus, a local effect leading to the negative departures in both observation types, would explain the stronger extend of departures in RS than in RO measurements.
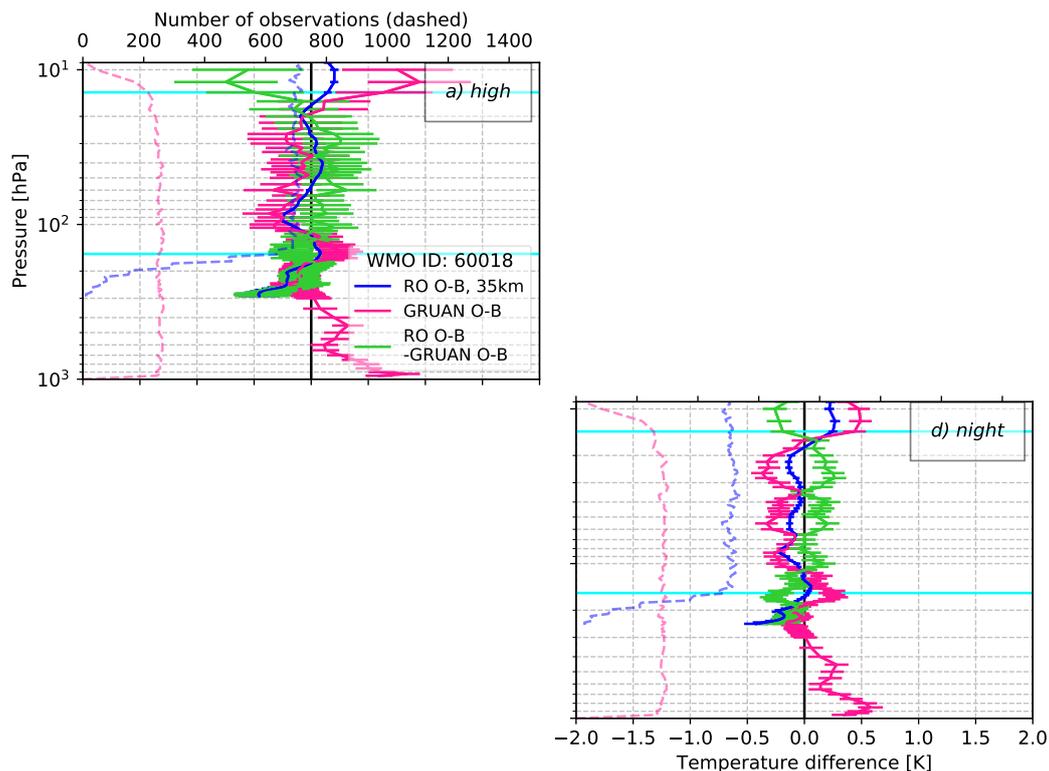
Figure 4.6: *As Fig. 4.1, but for the GRUAN site Tenerife.*

4.7 Cabauw

The measurement station Cabauw/De Bilt in the Netherlands is operated by the Dutch Weather Service *Koninklijk Nederlands Meteorologisch Instituut* (KNMI).

For those pressure levels at which enough profiles are available for a comparison, the observation types are mostly consistent. For high SEAs, the double difference at low pressures is negative and the RO and RS departures are not consistent with the RS measurements being too warm in comparison to the RO data. This points again to a remaining warm bias at high altitude in daytime GRUAN measurements.
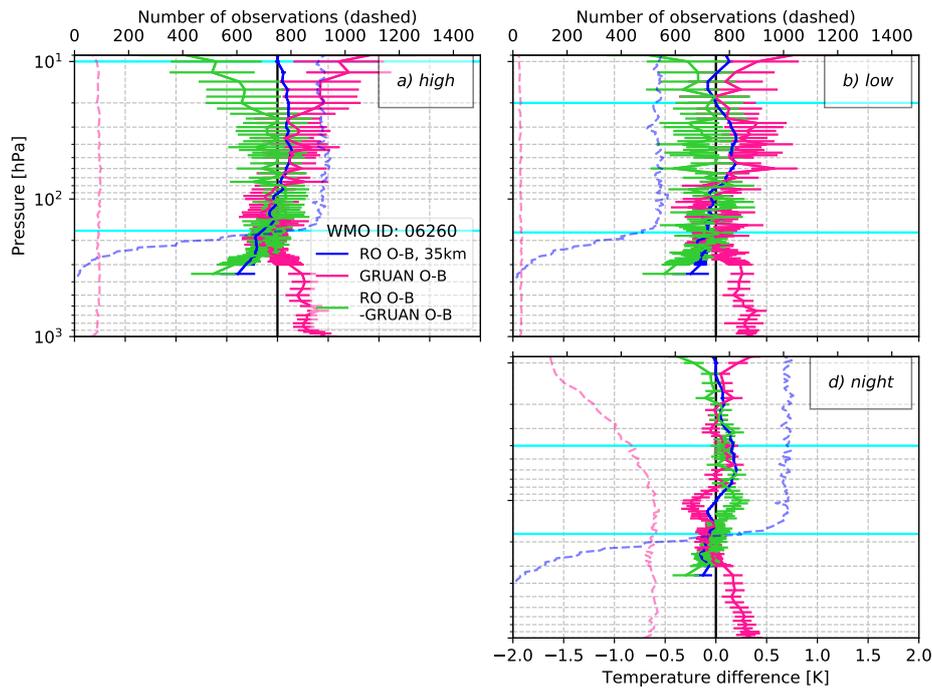
Figure 4.7: *As Fig. 4.1, but for the GRUAN site Cabauw.*

5 Analysis of consistency between RO and operational radiosonde measurements from high-quality sonde type

In this chapter, observation minus background departures from GRAS and COSMIC are compared with statistics for high-quality RS types. The comparison is performed for the year 2018 and the operational GRAS and COSMIC data as they were assimilated into the UM NWP model are used. Based on the results described in Ingleby [14] (and personal conversation with Bruce Ingleby), the following RS types have been analysed:

RS name	Manufacturer	RS typesWMO-No. 306 [30]
RS92	Vaisala (Finland)	79, 80, 81, 13, 14, 113, 114
RS41	Vaisala (Finland)	23, 24, 25, 42, 41, 123, 124,125, 141, 142
Modem M10	Modem (France)	77, 177
LMS-6	Lockhead Martin (USA)	82, 182

Table 5.1: High-quality RS types analysed in this report. Some RSs have several RS types given in WMO-No. 306 [30]. This accounts for different ground-stations and separates auto-launcher systems from manually-launched instruments.

While the RS produced by Meisei are known to be of good quality as well, the sample size was evaluated to be too small to produce meaningful results. This may not mean that there are generally too few sites that launch this type, but as described in Section 2.2.1 issues with the input data might have decreased the sample size as well.

In the following sections, the consistency between the different RS types and the RO measurements will be analysed based on one year of operational data. As the statistics may differ for different latitude bands, a separation into 30° latitude bands is applied based on the location of the upper-air site which is launching the sonde. The reference coordinate of the RO will be within 500 km of the location of the upper-air site. Statistics are only compared if at least 80 RS profiles and at least 80 RO profiles were available within a given SEA range. Furthermore, at least 80% of the analysed profiles must provide a value, i.e. no double difference is calculated above the altitude where 20% of the balloons have burst or below the altitude where more than 20% of the RO profiles are measured under too moist conditions to assume Tdry to provide a valid estimate of the temperature. These conditions reduce the amount of upper-air sites available for comparison and for some RS types comparisons may not be possible at all SEA ranges.

5.1 Comparing RO and Vaisala's RS41 radiosonde

During the year 2018 there have been 90 upper-air sites[1] which have been exclusively launching Vaisala's RS41 instrument. These sites are distributed globally with the highest density

---

[1] 04220, 04270, 04320, 04339, 04417, 06260, 08302, 08383, 10035, 10113, 10184, 10393, 10410, 10548, 10739, 10868, 12843, 14240, 14430, 15420, 16045, 16080, 16113, 16144, 16320, 16429, 16546, 33041, 40373, 40394, 40437, 41112, 41217, 41316, 45004, 47155, 48845, 48900, 60018, 61901, 63985, 67083, 67197, 70027, 74626, 74646, 78762, 78807, 84628, 85442, 85469, 85586, 85799, 85934, 88889, 89664, 93112, 93844, 93997, 94302, 94312, 94332, 94510, 94653, 94711, 94776, 94821, 95527, 96315

in Europe as can be seen in Fig. 5.1.
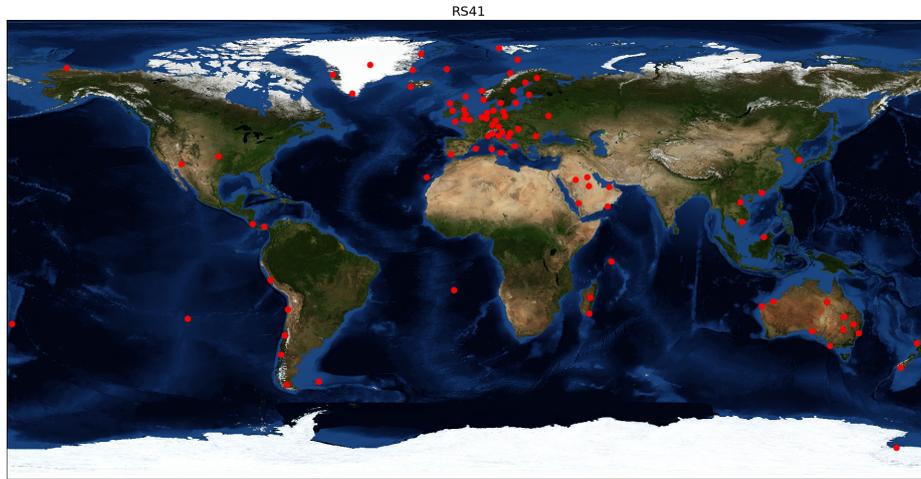
Figure 5.1: *Upper-air sites included in the analysis of Vaisala RS41 statistics.*

The observation minus background departures for these sites are analysed here to see how the RS departures agree with the RO departures (both with respect to the model background). Figure 5.2 shows the double difference of RO minus RS departures from model background fields for different SEA ranges. Each thin line shows the double difference for one upper-air site and the different colours indicate different latitude bands. If ten or more profiles are available for a SEA range and a latitude range, a thick line in the respective colour gives the mean value of the profiles. This does, however, not mean that ten or more profiles are available at each of the pressure levels.

For high SEAs the mean of the statistics is calculated for high Northern latitudes and mid Northern latitudes. Good agreement between the observation types is found between 100 and 20 hPa, with negative double differences at altitudes above 20 hPa and slightly positive double differences at altitudes below the 100 hPa level. A negative difference means that the temperature measured by the RS is warmer than what has been implied by RO, which is likely caused by the warm radiation bias that radiosondes experience during daytime. At low SEAs fewer profiles are available, but the mean profile for high and mid Northern latitudes look similar to the statistics for high SEA, which is also true to high Northern latitudes at dusk. It must be noted, however, that the number of profiles used in the averaging is small at the highest levels. Surprisingly, a negative double difference is also visible at nighttime for the highest levels, although to a smaller extent and based on a low number of profiles.

For the RS41, the combined uncertainty in sounding is given as 0.3 K below 16 km and as 0.4 K above 16 km (see Vaisala [27]). Thus, assuming that the uncertainties are correlated, if RO and the Vaisala RS41 agree to within ($0.4\,\mathrm{K} + \sigma_{RO}$) in the stratosphere, the data can be assumed to be consistent following Immler et al. [13]. As the RS uncertainties won't be fully correlated, this however will be an overestimate of the uncertainty. Overall, consistency is found for the averages over latitude bands, however, there is a tendency for the RS to measure

too warm temperatures at the highest pressure levels. For NWP also biases that are smaller than 0.4 K are critical and, thus, bias corrections may still be applied to all these profiles, or to a subset of the analysed pressure levels or latitude bands.

To get a better understanding of what causes the outliers in the above plots, statistics per SEA range for some of the sites that are peculiar in the above plots will be discussed.
At low SEA, one double difference profile for mid Southern latitudes clearly diverts from all other profiles. This profile is from upper-air site 85799 located in Chile. The statistics for this site are shown in Fig. 5.3. Only low and dusk/dawn SEAs have a sufficient amount of profiles to be included into the analysis per RS type. While the double difference profiles agree well at other pressure levels, large differences are found between 40 and 70 hPa. The reason for these differences are currently unclear. The upper air site at the airport of Puerto Montt is located next to a large lagoon and West of the Andes, making local effects possible which may not be picked up by RO given the larger horizontal scale of the measurements.

For high SEAs, some peculiar profiles occur in low Southern latitudes and two of these cases are shown in Fig. 5.4 and Fig. 5.5. Both of these sites only have statistics for high SEAs and have nearly 200 and around 300 profiles, respectively. Site 61901 is a British site located on the Island of St Helena, West of the African continent. It provides high resolution BUFR data. While there have not been launches every day of the year, but 324 in total, they are distributed over the whole year, which makes a seasonal sampling bias based on the sampling times of RS unlikely. All of the sondes are launched between 11 and 12 UTC. It is unclear at what exact UTC times the RO data have been collected, but if there is a model bias that depends on the time of the day and is not equally sampled in the RS and RO dataset, this could explain the discrepancy between the observation types. While COSMIC data will be randomly distributed over the day, the GRAS overpass times are constant from day to day.
The upper-air site 63985, see Fig. 5.5, is located in the Seychelles and low resolution TEMP profiles are used for this site. While the RO and RS departures from model background generally follow each other, the extent of the negative departure at 70 hPa is stronger in the RS statistics, leading to a positive double difference of about 0.75 K. It is possible that this is an effect of the different horizontal scales on which the different observation types measure, i.e. RS provide point measurements, while RO measure on a horizontal scale of approximately 300 km. Local effects might therefore be extenuated in the RO measurement, given that the RO measurement is sampling a wider area around the site. The RS in contrast samples the air along the flight path and can pick up local effects. While this could explain why some features are stronger in the RS departures than in the RO departures, it is unclear if this is the reason.

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT ROM SAF

(a) high
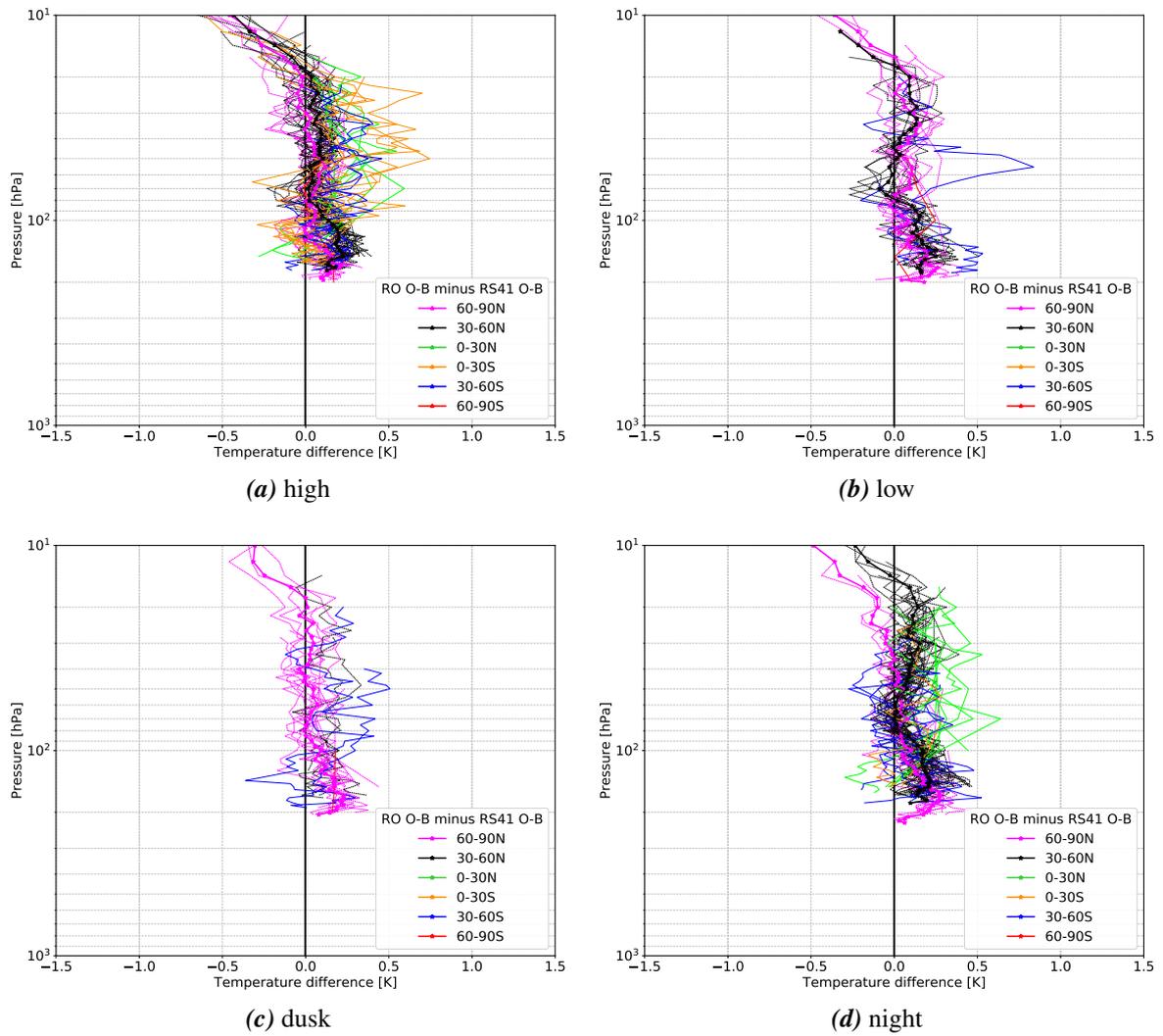
(b) low

(c) dusk

(d) night

Figure 5.2: *Double difference between Vaisala RS41 measurement minus model and GRAS/COSMIC minus model background for different solar elevation angle ranges, (a) high SEA, (b) low SEA, (c) dusk SEA, (d) nighttime SEA.*
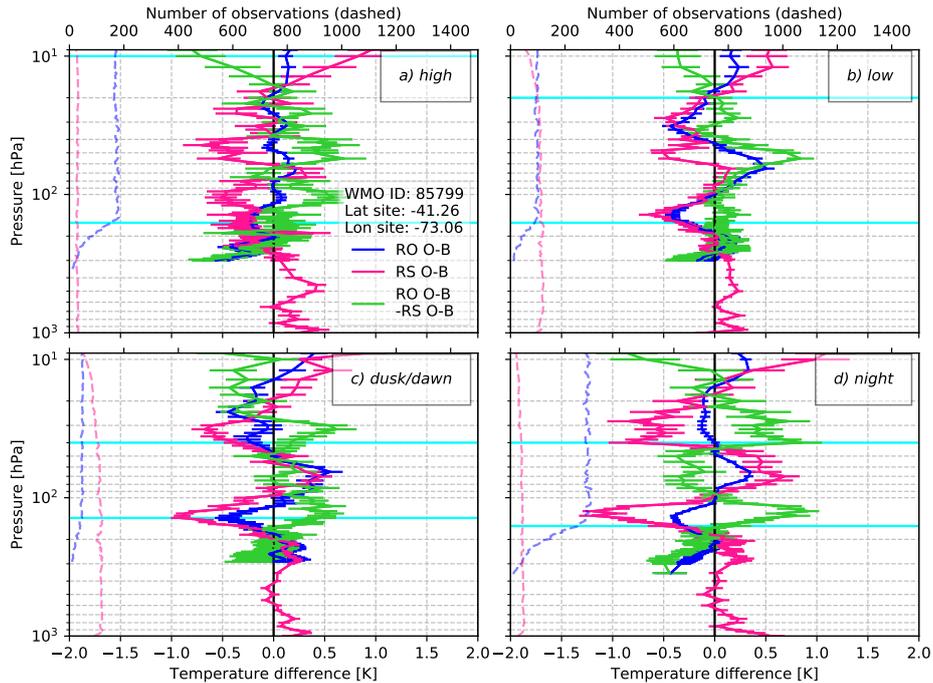
Figure 5.3: *RO departures (RO O-Bs, blue), RS departures (RS O-Bs, pink) and estimated bias (RO O-B minus RS O-B, green) for the upper air site 85799 in Chile. The error bars represent the standard errors. The horizontal cyan lines indicate the highest/lowest pressure level where at least 80% of the RO/RS profiles are included in the statistics.*
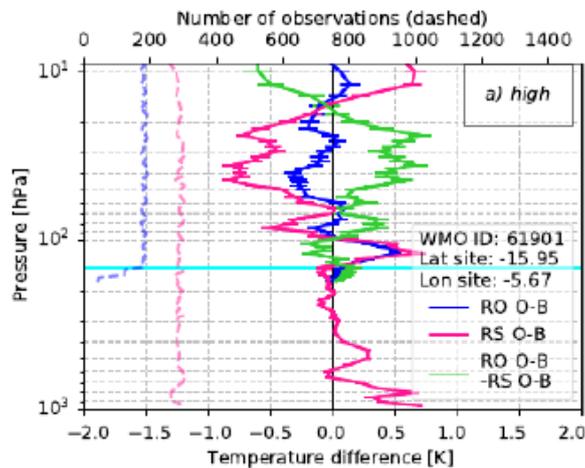
Figure 5.4: *RO departures (RO O-Bs, blue), RS departures (RS O-Bs, pink) and estimated bias (RO O-B minus RS O-B, green) at the upper-air site 61901 in St Helena. The error bars represent the standard errors. The horizontal cyan lines indicate the highest/lowest pressure level where at least 80% of the RO/RS profiles are included in the statistics.*

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT
ROM SAF
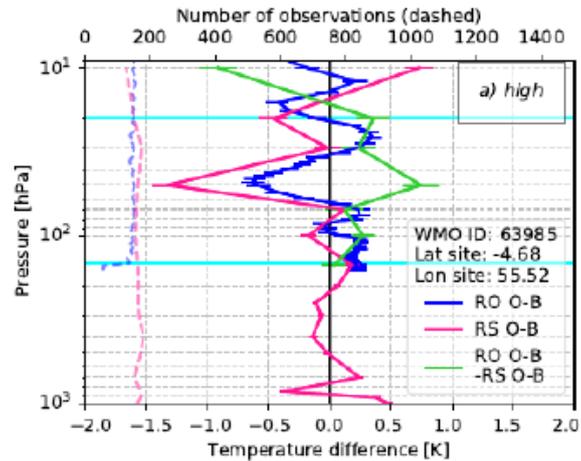
Figure 5.5: *RO departures (RO O-Bs, blue), RS departures (RS O-Bs, pink) and estimated bias (RO O-B minus RS O-B, green) at the upper-air site 63985 in the Seychelles. The error bars represent the standard errors. The horizontal cyan lines indicate the highest/lowest pressure level where at least 80% of the RO/RS profiles are included in the statistics.*

5.2 Comparing RO and Vaisala's RS92 radiosonde

During the year 2018 there have been 23 upper-air sites[2] which have been exclusively launching Vaisala's RS92 instrument, which is the precursor of Vaisala's latest model - the RS41. The sites can be seen in Fig. 5.6 and are located in Europe, Australia, North America and South America.
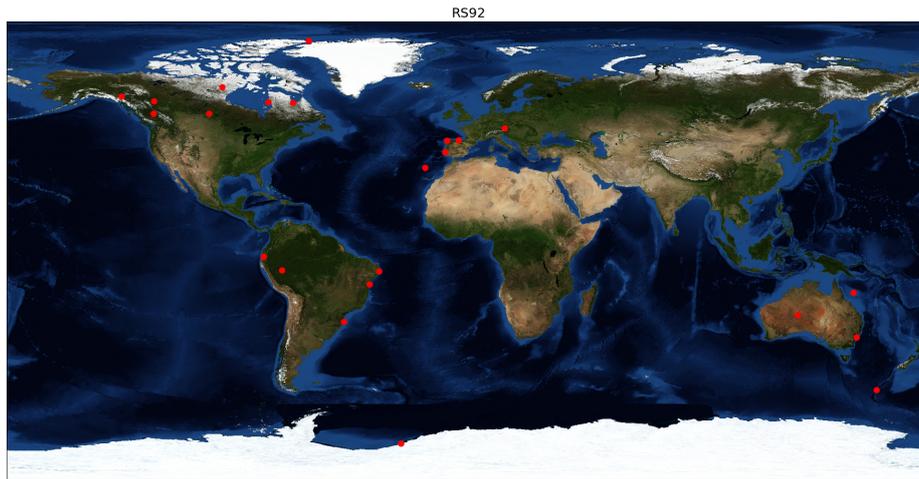
Figure 5.6: *Upper-air sites included in the analysis of Vaisala RS92 statistics.*

As for the RS41, the double difference between RS and RO are analysed here to evaluate if the two observation types are consistent, see Fig. 5.7. The total uncertainty in sounding is given as 0.5 K for the Vaisala RS92 (Vaisala [28]). As only 23 upper-air sites launched only RS92 sonde in 2018, and most of them do not cover all SEA ranges, the sites that produce statistics per SEA range are limited. Below the 30 hPa level, consistency within the RS92 uncertainty range is found, however, as mentioned above also smaller differences can matter for NWP. At altitudes above the 30 hPa level, a negative double difference occurs for all SEA ranges, which is in contrast to the expectation of such a warm RS bias only at daytime. While the sample size is small, this is still a peculiar feature deserving further investigation. Looking at the statistics from individual sites reveals that the upper-air sites that show these negative double differences at nighttime for high altitudes are located in Canada. One such example is shown in Fig. 5.8, where statistics are available for all SEA down to a pressure level of 10 hPa. At high and low SEA, the two observation types start to deviate at 30 hPa and at night a significant deviation is only visible at 10 hPa. At lower altitudes the profiles agree reasonably well. While the negative double differences during the day are likely caused by a warm radiation bias in the radiosondes, it is unclear what leads to the negative double difference at nighttime. A similar feature has been found for GRUAN sites located in the Arctic, however, the reason for this behaviour is currently not understood.

---

[2] 08001, 08023, 08522, 08579, 11010, 71082, 71867, 71906, 71907, 71908, 71926, 71945, 71964, 82705, 82900, 83229, 83899, 84203, 89022, 94299, 94461, 94767, 94998,

(a) high
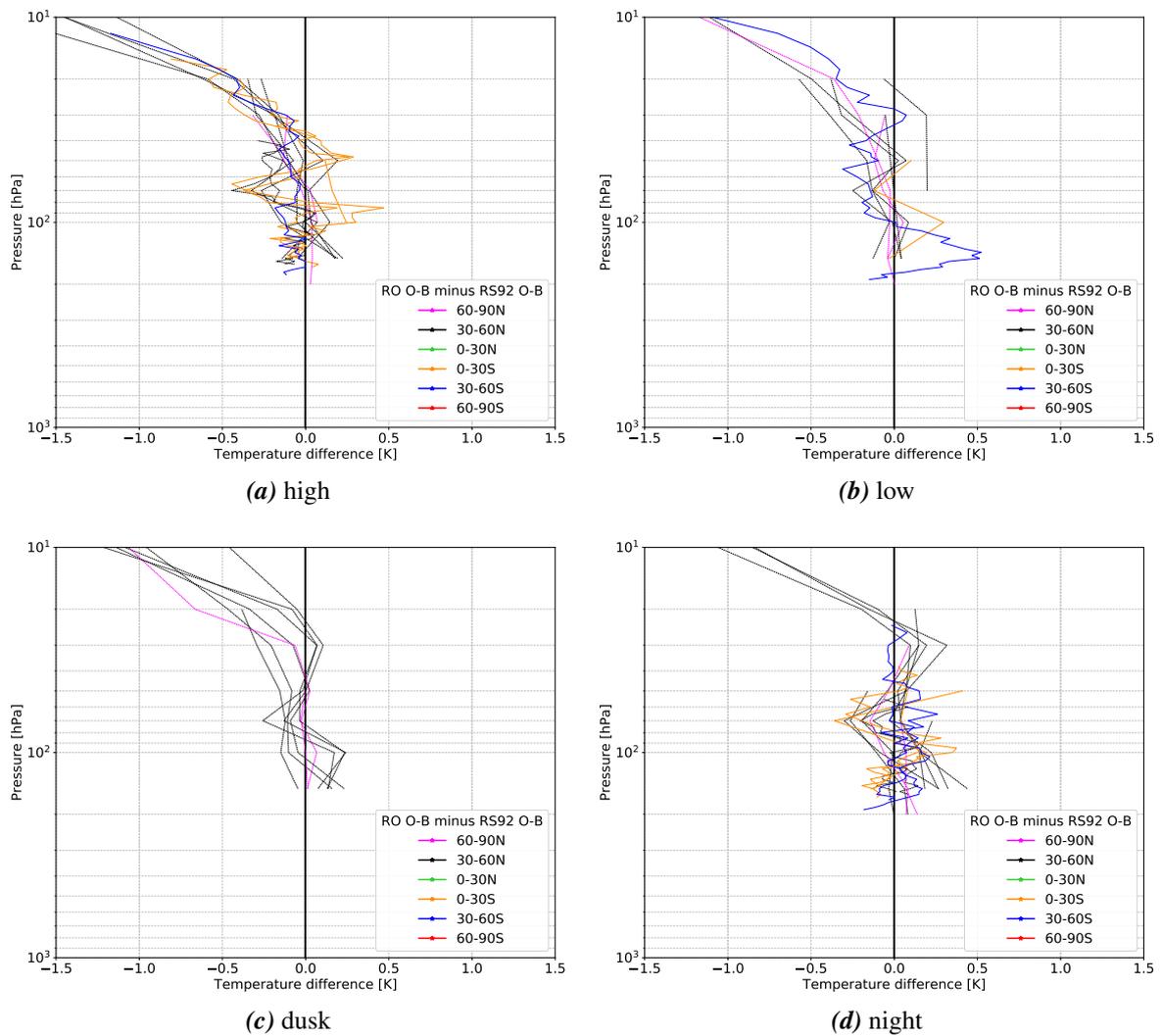
(b) low

(c) dusk

(d) night

Figure 5.7: *Double difference between Vaisala RS92 measurement minus model and GRAS/COSMIC minus model background for different solar elevation angle ranges, (a) high SEA, (b) low SEA, (c) dusk SEA, (d) nighttime SEA.*
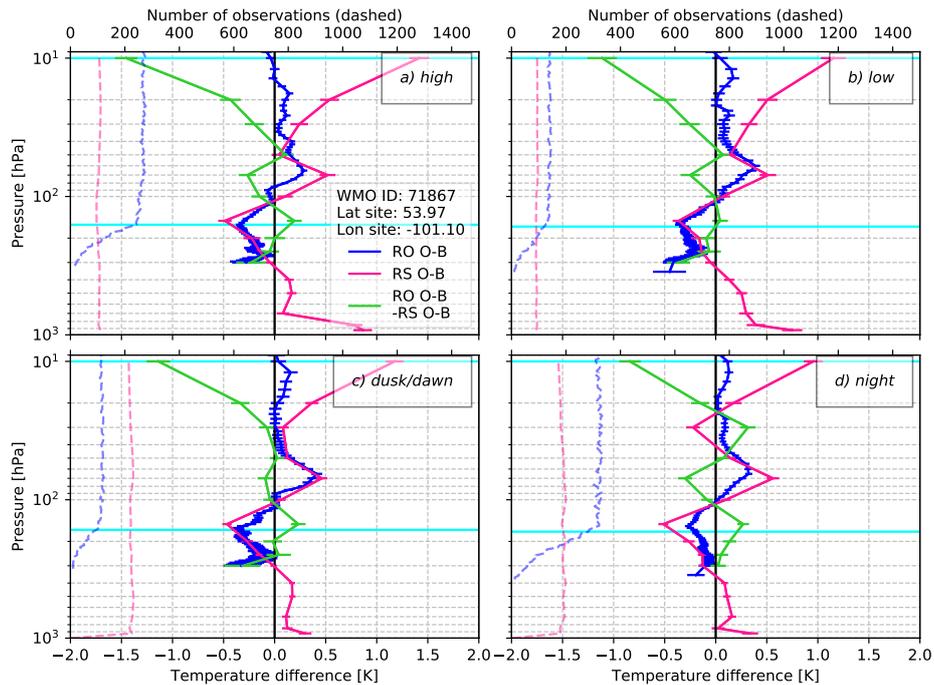
Figure 5.8: *RO departures (RO O-Bs, blue), RS departures (RS O-Bs, pink) and estimated bias (RO O-B minus RS O-B, green). The error bars represent the standard errors. The horizontal cyan lines indicate the highest/lowest pressure level where at least 80% of the RO/RS profiles are included in the statistics.*

5.3 Comparing RO and the Lockhead Martin LMS-6 radiosonde

During the year 2018 there have been 46 upper-air sites[3] which have been exclusively launching the Lockhead Martin LMS-6. These sites are mainly located in mainland United States and on Pacific Island belonging to, or having a close relationship with The United States, such as Hawaii, the Marshall Islands and American Samoa as can be seen in Fig. 5.9.
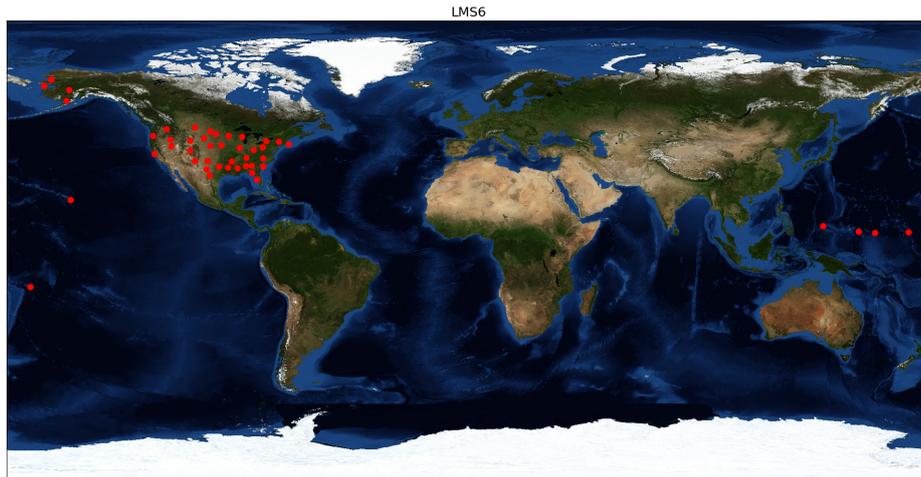
Figure 5.9: *Upper-air sites included in the analysis of Lockhead Martin LMS-6 statistics.*

The double differences for the LMS-6 are shown in Fig. 5.10. The LMS-6 sonde measurements are used in TEMP format, i.e. they are only available on standard pressure levels, see Section 2.2.1. Based on the locations of upper-air sites using this sonde, sufficient sites to calculate an average are only available at mid Northern latitudes. For high SEA, the agreement between RO and RS is good for pressures down to the 30 hPa level, with larger differences at 20 and 10 hPa, where a warm RS bias is the likely cause of differences. At low and dusk/dawn SEAs, not many upper-air sites provide data, however, there is a tendency for the double difference to be positive below 20 hPa and negative above. The statistics for nighttime look similar and also show a slightly negative double difference at 10 hPa. The similarity between nighttime and dusk/dawn observations indicates that a combination of these SEA ranges could be possible and would improve the sample sizes.

---

[3] WMO station IDs 70133, 70200, 70231, 70326, 72208, 72210, 72214, 72215, 72230, 72235, 72248, 72249, 72261, 72265, 72317, 72327, 72340, 72363, 72365, 72426, 72476, 72493, 72518, 72520, 72528, 72558, 72562, 72582, 72645, 72649, 72659, 72662, 72672, 72681, 72694, 72764, 72768, 72786, 74494, 74560, 91285, 91334, 91348, 91376, 91413, 91765

(a) high
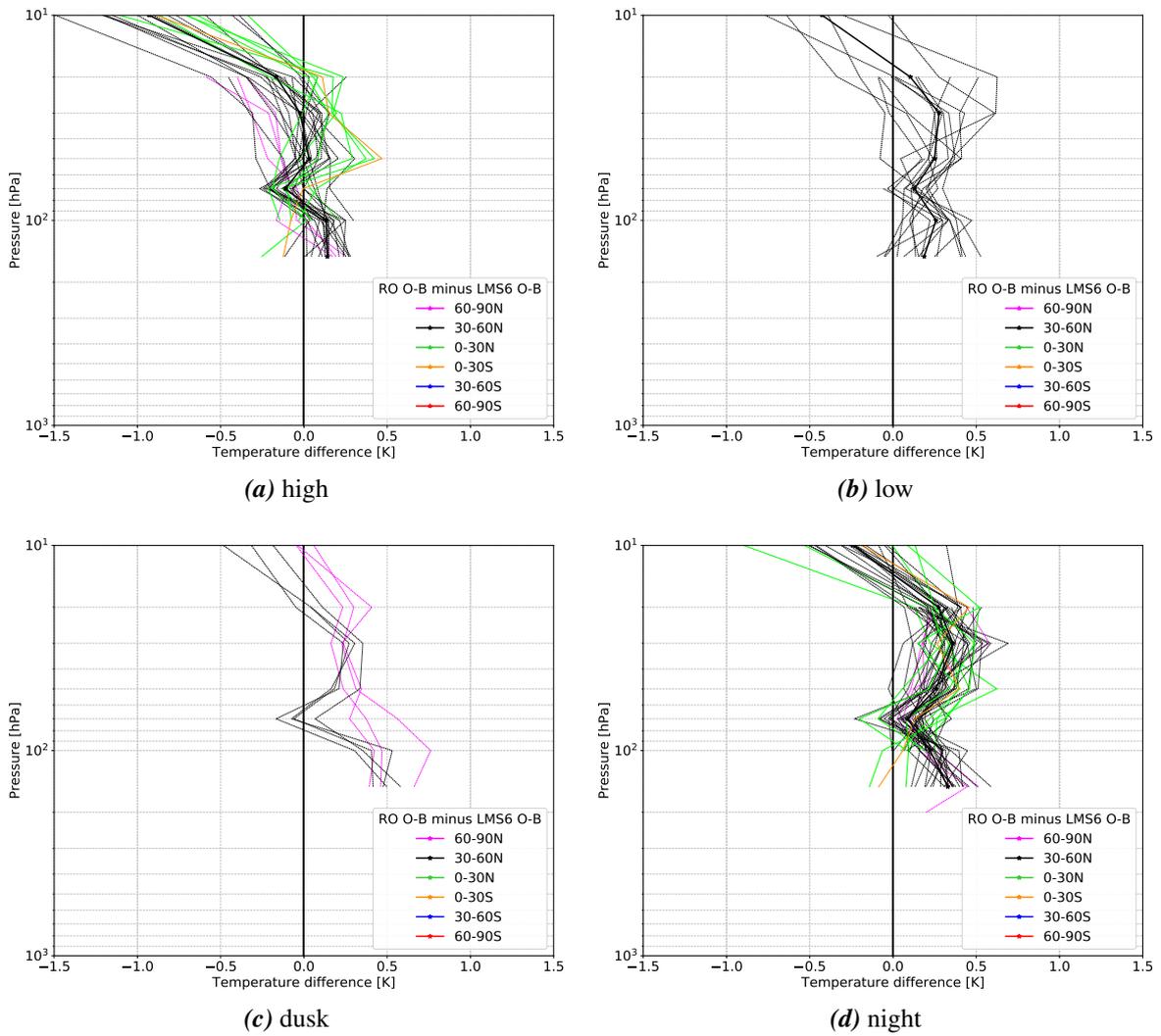
(b) low

(c) dusk

(d) night

Figure 5.10: *Double difference between Lockhead Martin LMS6 measurement minus model and GRAS/COSMIC minus model background for different solar elevation angle ranges, (a) high SEA, (b) low SEA, (c) dusk SEA, (d) nighttime SEA.*

5.4 Comparing RO and the Modem radiosondes

During the year 2018 there have been 29 upper-air[4] sites which have been exclusively been launching the M10 RS produced by Modem. This sonde type is used in France as well as in other areas around the globe as can be seen in Fig. 5.11.
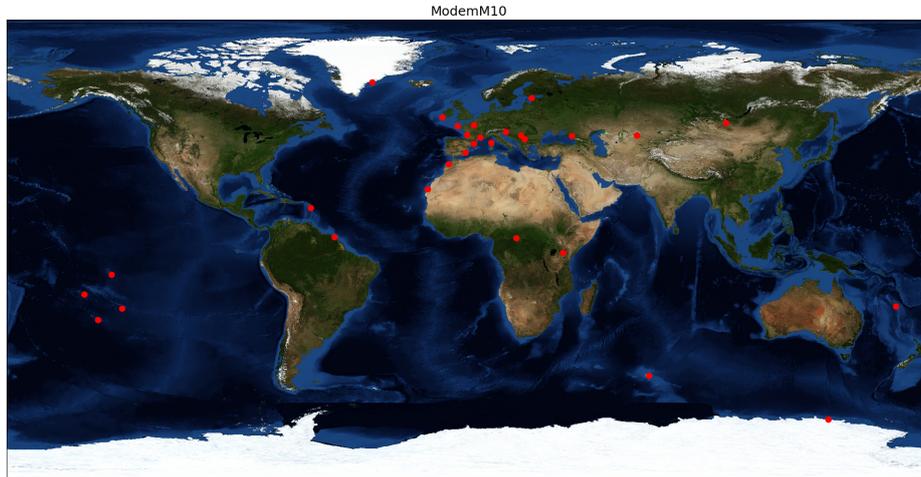
Figure 5.11: Upper-air sites included in the analysis of Modem M10 statistics.

Based on the very limited amount of data available at low and dusk SEAs, statistics are only analysed for high SEA and nighttime measurements, see Fig. 5.12. The profiles for mid Northern latitudes overall show a negative double difference at most pressure levels, with increasing magnitude at higher altitudes. For the other latitude bands less than ten profiles are available and no averaging has been performed. However, there is tendency for the double differences to be negative, i.e. a warmer RS temperature than the temperature that has been implied by RO measurements.

At nighttime, the agreement between the Modem M10 sonde and RO for Northern mid-latitudes is better, however a tendency to be slightly negative is visible.

One obvious outlier appears at high SEA, i.e. the line ending at 100 hPa at a double difference of approximately -1.2 K. This outlier is caused by upper-air site 38064 located in Kazakhstan. The statistics for this site are shown in Fig 5.13. The statistics for this site are very suspicious, and although many pressure levels do not have a sufficient amount of data, larger differences than expected are visible at some pressure levels. As there are more examples of sites that launch Modem M10 sondes which show suspicious statistics, this sonde type can not be recommended to be assimilated without bias correction. Further investigation would be required to analyse why these large departures of Modem M10 from model background, which are not shown in the co-located RO departures occur. Likely this will happen as part of RS monitoring system.

---

[4] 07761, 08190, 08430, 13275, 13388, 14015, 26038, 37789, 38064, 44231, 60096, 60155, 61998, 63741, 64650, 78897, 81405, 89642, 91592, 91925, 91938, 91948, 91958
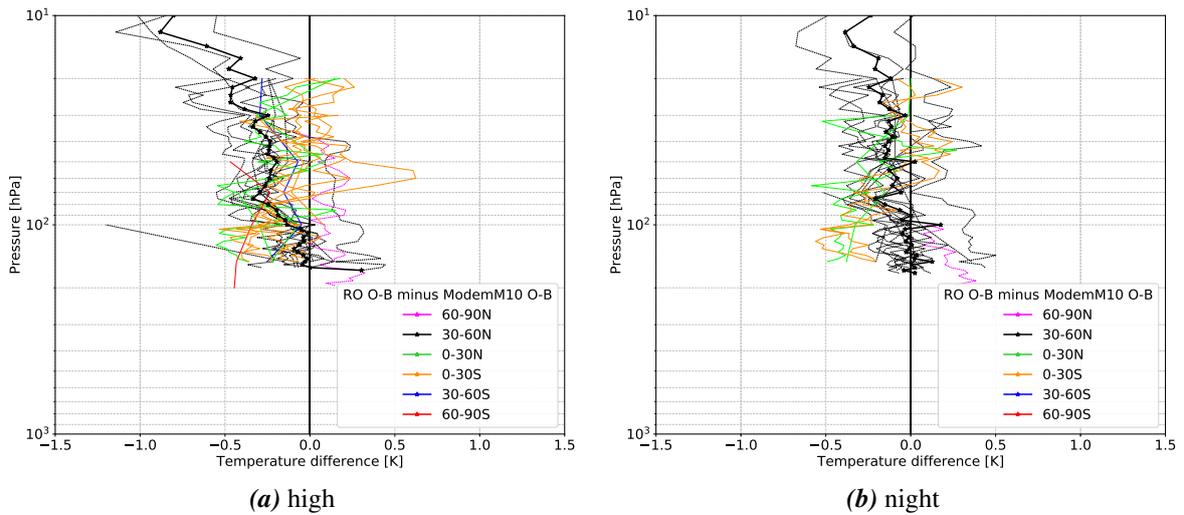
(a) high   *(b)* night

Figure 5.12: *Double difference between Modem M10 measurement minus model and GRAS/COSMIC minus model background for different solar elevation angle ranges, (a) high SEA, (b) nighttime SEA.*
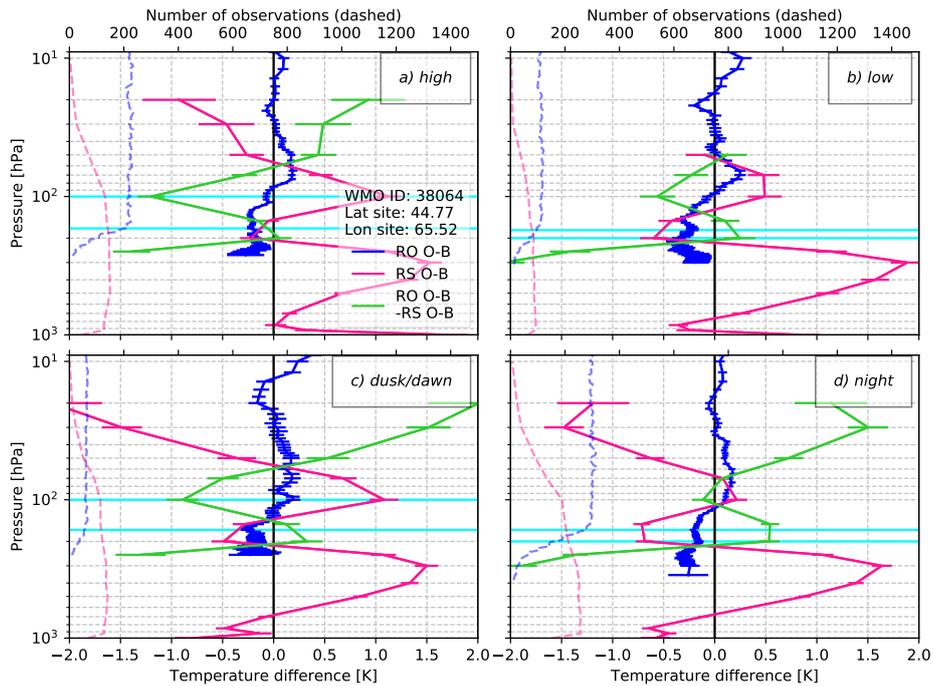
Figure 5.13: *RO departures (RO O-Bs, blue), RS departures (RS O-Bs, pink) and estimated bias (RO O-B minus RS O-B, green). The error bars represent the standard errors. The horizontal cyan lines indicate the highest/lowest pressure level where at least 80% of the RO/RS profiles are included in the statistics.*

6 Analysis of the structural uncertainty in a tangent linear RO retrieval

Structural uncertainty arises from the retrieval and assumptions made within a retrieval. In the conventional RO retrieval a source of structural uncertainty is the choice for a smoothing algorithm to smooth noisy high level bending angle data and the choice of the climatology used for smoothing. The differences arising from smoothing and climatology choices made by different processing centres can be used to estimate the structural uncertainty in the conventional RO retrieval as is has been done by e.g. Ho et al. [12], Steiner et al. [21]. The tangent linear retrieval uses neither a climatology nor a smoothing algorithm. Instead, bending angle departures from a specific altitude range are propagated into Tdry departures. Outside of the chosen altitude range, the departures are set to zero. One source of structural uncertainty in the tangent linear retrieval is thus the choice of altitude above which the departures are set to zero, which is called cut-off here. Tradowsky et al. [26] analysed the structural uncertainty for global COSMIC FM6 departures, and Tradowsky [24] additionally showed the spread of Tdry departures for some GRUAN sites.

Figure 6.1 shows the Tdry departures for all SEAs at the upper-air site Lindenberg (10393) for different upper cut-off height and no cut-off. While the departures are similar at the upper troposphere and lower stratosphere, they diverge at higher altitudes (lower pressure) levels. The departures tend to become more positive at high cut-off impact heights. Burrows and Healy [2] show that this is caused by the model background fields as they diverge in the opposite direction when using model fields from ECMWF. Excluding the influence of model background from the RO and RS comparison motivates the cut-off used within the tangent linear retrieval.

While the statistics shown in the preceding sections were calculated using a cut-off impact height of 35 km in the vicinity of individual sites, in this chapter the Tdry departures are calculated on a global 10° latitude by 30° longitude grid for cut-off impact heights of 35 km, 40 km, 45 km, 50 km, 55 km. The spread (range of the calculated departures) of the Tdry departures at 10 hPa is then plotted in Fig. 6.2 (points) and Fig. 6.3 (contour plot based on statistics for the points shown in Fig. 6.2). The spread between the departures is largest in the Tropics from −20° to 10° latitude, where it reaches values up to nearly 1.2 K. Strong vertical gradients, caused by gravity waves in the Tropics may explain this behaviour. The spread is smallest at high Northern latitudes and generally decreases with increasing Northern latitude. In the Southern hemisphere, however, the spread of the departures first decreases when getting closer to the pole, but then it increases again at −70° and −80°, i.e. above and in close proximity to the Antarctic continent.

The calculation in the spread of departure statistics between 35 and 55 km impact height improves the understanding of the method. However, for the purpose of comparing RO and RS departures the calculated spread will be an overestimation of the uncertainty, as a cutoff impact height of 55 km would invalidate the central assumption for the double-differencing technique; that is, the model bias is constant over the separation distance (see also Burrows and Healy [2]). Using a double-difference minimises the effects of co-location errors without introducing model biases into the calculation. However, the interaction of the Tdry calculation with the cut-off is not fully understood and may introduce model errors in this range into the calculation, leading to an overestimation of the uncertainty. Thus, the results from this chapter are not included into the plots showing RO and RS comparisons, but should rather

be understood as an exercise to improve the understanding of the tangent linear RO retrieval.
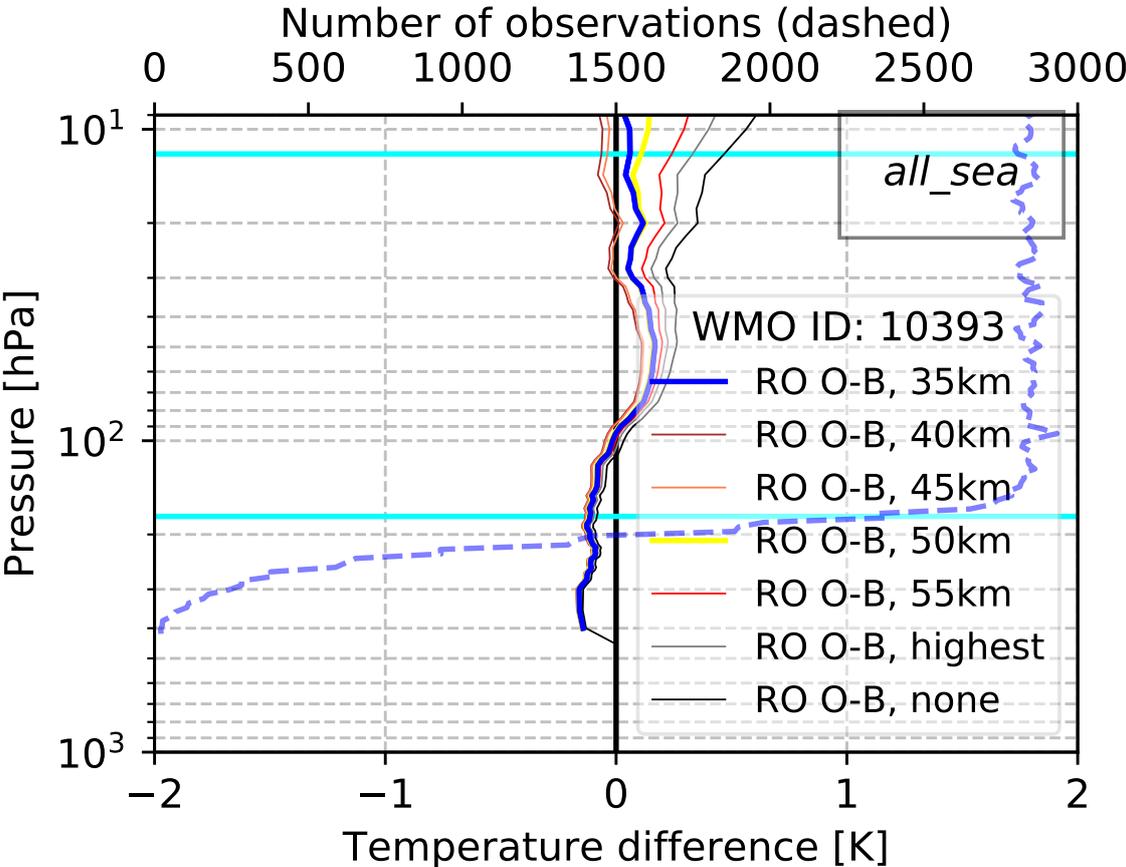
Figure 6.1: *Tdry departure statistics for cut-off impact heights of 35 km, 40 km, 45 km, 50 km, 55 km, cut-off for highest value, and without cut-off, respectively. Statistics are shown for the upper-air site Lindenberg in Germany.*

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT
ROM SAF
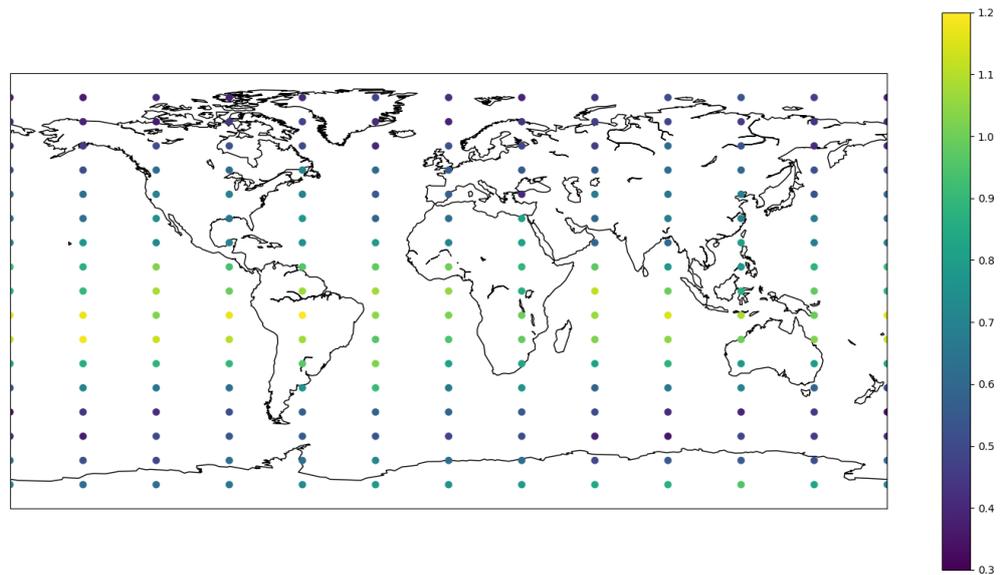
Figure 6.2: *Spread of the departure statistics for cut-off impact heights between 35 km and 55 km at 10 hPa. Colour scale in Kelvin.*
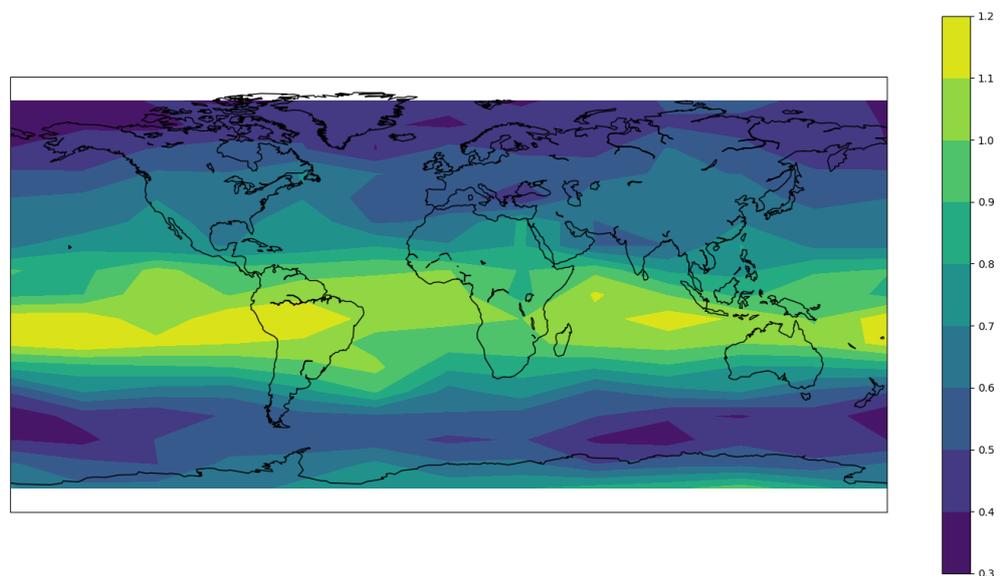
Figure 6.3: *Spread of the departure statistics for cut-off impact heights between 35 km and 55 km at 10 hPa. Colour scale in Kelvin.*

7 Conclusions

This ROM SAF visiting scientist report analyses the agreement between RO and high-quality RS. This should serve the purpose of developing a hierarchy of anchor observations for NWP, and test the hypotheses that GRAS RO data and reference-quality RS data from GRUAN are in agreement. Furthermore, a better understanding of uncertainties in the RO retrieval should be gained through this report.

The comparison between GRUAN and RO departures, where the departures are calculated with respect to UK Met Office NWP background fields, has been performed for the years 2014-2016 using the ROM SAF Climate Data Product for GRAS which became available during the course of this project. Generally, consistency between RO and GRUAN is found at many pressure levels, exhibiting the value of these entirely independent data sets. The comparison of departure statistics for RO and RS can robustly reveal model biases, i.e. a bias is detected if both observation types are consistent with each other, but not with the model. Differences between RO and GRUAN are mainly found at high altitudes where the the negative double difference indicates that the GRUAN temperature is warmer than the RO retrieval. This is likely caused by a residual warm bias in the GRUAN profile caused by solar radiation. However, it is surprising that this negative double difference is also visible at nighttime for a high latitude site. A second high latitude site shows this effect, however, the small sample size decreases the reliability. It is not clear what causes this negative double difference for nighttime measurements, which indicated GRUAN temperatures are warmer than those implied by RO. Similar results have, however, been found for some operational high latitude sites using the RS92 sonde which is the same sonde types of the GRUAN data product, however without GRUAN processing applied.

The comparison between the ROM SAF CDR data and the GRUAN data exhibits consistency between these entirely independently derived datasets at many pressure levels and their value in determining model biases. While GRUAN data are not available in near-real time and can therefore not serve as an anchor for NWP models, they can be used in validation studies. The availability of full uncertainty information in the GRUAN RS data product greatly assists the consistency assessment in such studies. Further details of the split between correlated and uncorrelated uncertainties would be of use in future work and is discussed within the GRUAN community. Most of the GRUAN sites also submit their measurements to the GTS and these operational data products (which are different from the GRUAN processed profiles) are often assimilated into NWP.

Future work could concentrate on investigating specifically the statistics for high latitude sites from GRUAN, i.e. Ny Ålesund, Sodankylä and Barrow, as well as the data from operational upper-air sites at high latitudes. Such a study could include separation into different seasons, analysing the GRUAN data product for the Vaisala RS41 and the updated GRUAN Vaisala RS92 once they become available, as well as comparison with the Tdry data product produced by the ROM SAF as part of the CDR, which uses a conventional RO retrieval rather than the tangent linear retrieval applied here. Ideally, an independent third data source would be used in the comparison to understand what is causing the differences between the RO and RS statistics. Finding a suitable datasets for such a comparison could be part of future work and a good starting point could be by checking which other measurements are performed at the well-equipped GRUAN sites in the high latitudes. Measurements from e.g. a lidar or microwave radiometer might be of use in a future study. Another possibility would be the comparison to another space-based instrument such as ESA's Infrared Atmospheric Sound-

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT
ROM SAF

ing Interferometer (IASI). Such a comparison could be done in Tdry space or in radiance space, which is the native measurement of IASI. A tool to propagate GRUAN profiles into radiance space has been developed by the UK Met Office (see Carminati et al. [4]) and could be adapted for such a task, i.e. it could also be used to propagate RO retrievals into radiances.

High-quality operational RS data have been compared with RO data to evaluate if measurements from some RS types can be assimilated into NWP models without applying any bias correction. The results obtained from the comparison are mixed and complicated by some issues with the RS dataset that are described in Section 2.2.1.Therefore, further analysis after the new RS monitoring system is in place at the UK Met Office is suggested prior to making any final decision on which sondes can used without bias correction.

From the results obtained here, all RS types require a bias correction at high altitudes, as there seems to be a remaining warm bias in the RS data, even after the vendor correction has been applied. It is likely that the vendor corrections are conservative to avoid the overcorrection of the radiation bias. There seem to be a tendency for such a negative double difference to occur at all SEA ranges, although a warm RS bias would theoretically only occur at daytime. From all the analysed RS types, Vaisala's RS41 has the most available profiles and the smallest double differences. It could therefore be considered to use the RS41 as the secondary anchor without a bias correction at lower altitudes and using a bias correction at altitudes above the 20 hPa level.

Where a correction for RS needs to be applied, this correction could, for example, be calculated with respect to RO measurements as described in Tradowsky [25, 24]. As RO Tdrys are used in the investigation, comparisons can only be performed in the dry part of the atmosphere. While this is a limitation of the method, RS biases are typically larger in the stratosphere than in the troposphere.

In addition to the work presented here, the report and the software produced as part of the report will be used in support of a forecast impact study to test the bias corrections developed within Tradowsky [24]. It may be advisable to start this forecast impact study after some issues with the assimilation of RS data have been taken care off. The UK Met Office is at the moment setting up a new team which will develop a comprehensive RS monitoring system. This will likely solve the issues that have complicated this report and increase the value of RS in the assimilation system.

In Chapter 6, the structural uncertainty was discussed. One source of information about the structural uncertainty in the retrieval applied here, is the spread of departures for different cut-off impact heights from 35 to 55 km[1]. This spread is analysed at the 10 hPa pressure level for a global grid, showing that the spread is largest in the Tropics, reaching nearly 1.2 K and smallest in the high Northern latitudes. The mean spread over all grid points at 10 hPa is 0.69 K. While the spread is decreasing in the Southern hemisphere with increasing latitude, the spread increases again above Antarctica.

While the calculated spread gives a good indication of the sensitivity of the method to the cut-off impact height, this spread is not a good indication for the uncertainty in the RO and RS comparison. As explained in Tradowsky et al. [26], using a double-difference minimises the effects of co-location errors without introducing model biases into the calculation. However, the interaction of the Tdry calculation with the cut-off is not fully understood and may introduce model errors in this range into the calculation, leading to this estimate of the un-

---

[1] For the calculation of statistics in all other chapters, a cut-off impact height of 35 km is used.

certainty being an overestimation when estimating RS biases.

This report uses a method which has been developed in two former ROM SAF visiting scientist projects to compare RO data with measurements from balloon-borne instruments. Model background fields, i.e. short range forecasts, from the UK Met Office NWP model are used as a transfer medium to eliminate effects caused by imperfect co-location. Using this method RS from the operational network as well as from GRUAN are compared with RO. Furthermore the structural uncertainty in the tangent linear RO retrieval is analysed for a better understanding of the method. Future work will include a forecast impact study that will use results and software from this code. The RS monitoring system under development at the Met Office might also benefit from this report. Once this monitoring system is in place, it will likely be possible to strengthen suggestions for sonde types and pressure levels that may not require a bias correction.

Future work could then concentrate on determining if there are other observation types that could be used as additional anchors, potentially including a bias correction prior to their assimilation. Aircraft-based measurements could be a candidate for such an investigation. While their bias correction would be challenging, they provide an increasing amount of data and, therefore, the potential benefit of correcting them prior to the assimilation rather than including them into the variational bias correction is large.

7.1 Acknowledgement

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

ROM SAF
EUMETSAT

Bibliography

[1] Auligné, T., McNally, A. P., and Dee, D. P. (2007). Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, 133:631–642.

[2] Burrows, C. P. and Healy, S. B. (2016). Sensitivity of radio occultation-based dry temperature retrievals to upper-level information and its relevance to radiosonde bias corrections. Forecasting Research Technical Report 615, Met Office. Available at `http://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/library/frtr_615_2016_2p.pdf`.

[3] Calbet, X., Peinado-Galan, N., Rípodas, P., Trent, T., Dirksen, R., and Sommer, M. (2017). Consistency between gruan sondes, lblrtm and iasi. *Atmos. Meas. Tech.*, 10:2323–2335.

[4] Carminati, F., Migliorini, S., Ingleby, B., Bell, W., Lawrence, H., Newman, S., Hocking, J., and Smith, A. (2019). Using reference radiosondes to characterise nwp model uncertainty for improved satellite calibration and validation. *Atmospheric Measurement Techniques*, 12(1):83–106.

[5] Dee, D. P. (2004). Variational bias correction of radiance data in the ECMWF system. ECMWF conference paper, European Centre for Medium-Range Weather Forecasts. Available at `http://www.ecmwf.int/sites/default/files/elibrary/2004/8930-variational-bias-correction-radiance-data-ecmwf-system.pdf`.

[6] Derber, J. C. and Wu, W.-S. (1998). The Use of TOVS Cloud-Cleared Radiances in the NCEP SSI Analysis System. *Mon. Wea. Rev.*, 126:2287–2299.

[7] Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H. (2014). Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde. *Atmos. Meas. Tech.*, 7:4463–4490.

[8] Eyre, J. R. (2016). Observation bias correction schemes in data assimilation systems: a theoretical study of some of their properties. *Quarterly Journal of the Royal Meteorological Society*, 142(699):2284–2291.

[9] GCOS-112 (2007). GCOS REFERENCE UPPER-AIR NETWORK (GRUAN): Justification, requirements, siting and instrumentation options. Technical Report GCOS-112 (WMO/TD No. 1379), World Meteorological Organization. Available at `https://www.wmo.int/pages/prog/gcos/Publications/gcos-112.pdf`.

[10] Haimberger, L., Tavolato, C., and Sperka, S. (2012). Homogenization of the Global Radiosonde Temperature Dataset through Combined Comparison with Reanalysis Background Series and Neighboring Stations. *J. Climate*, 25:8108–8131.

[11] Healy, S. B. (2008). Forecast impact experiment with a constellation of GPS radio occultation receivers. *Atmos. Sci. Let.*, 9:111–118.

[12] Ho, S.-P., Hunt, D., Steiner, A. K., Mannucci, A. J., Kirchengast, G., Gleisner, H., Heise, S., von Engeln, A., Marquardt, C., Sokolovskiy, S., Schreiner, W., Scherrlin-Pirscher, B., Ao, C., Wickert, J., Syndergaard, S., Lauritsen, K. B., Leroy, S. S., Kursinski,

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT ROM SAF

E. R., Kuo, Y.-H., Foelsche, U., Schmidt, T., and Gorbunov, M. (2012). Reproducibility of GPS radio occultation data for climate monitoring: Profile-to-profile inter-comparison of CHAMP climate records 2002 to 2008 from six data centers. *J. Geophys. Res.*, 117.

[13] Immler, F. J., Dykema, J., Gardiner, T., Whiteman, D. N., Thorne, P. W., and Vömel, H. (2010). Reference Quality Upper-Air Measurements: guidance for developing GRUAN data products. *Atmos. Meas. Tech.*, 3(5):1217–1231.

[14] Ingleby, B. (2017). An assessment of different radiosonde types 2015/2016. Technical Memorandum 807, European Centre for Medium-Range Weather Forecasts.

[15] Ingleby, B. and Edwards, D. (2015). Changes to radiosonde reports and their processing for numerical weather prediction. *Atmos. Sci. Let.*, 16:44–49.

[16] Kursinski, E. R., Hajj, G. A., Leroy, S. S., and Herman, B. (2000). The GPS Radio Occultation Technique. *Terrestrial, Atmospheric and Oceanic Sciences*, 11(1):53–114.

[17] Kursinski, E. R., Hajj, G. A., Schofield, J. T., Linfield, R. P., and Hardy, K. R. (1997). Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, 102:23429–23465.

[18] Ladstädter, F., Steiner, A. K., Schwärz, M., and Kirchengast, G. (2015). Climate intercomparison of GPS radio occultation, RS90/92 radiosondes and GRUAN from 2002-2013. *Atmos. Meas. Tech.*, 8(3):1819–1834.

[19] Poli, P., Healy, S. B., and Dee, D. P. (2010). Assimilation of Global Positioning System radio occultation data in the ECMWF ERA-Interim reanalysis. *Q. J. R. Meteorol. Soc.*, 136:1972–1990.

[20] Rennie, M. P. (2010). The impact of GPS radio occultation assimilation at the Met Office. *Quart. J. Roy. Meteor. Soc.*, 136:116–131.

[21] Steiner, A. K., Hunt, D., Ho, S.-P., Kirchengast, G., Mannucci, A. J., Scherllin-Pirscher, B., Gleisner, H., von Engeln, A., Schmidt, T., Ao, C., Leroy, S. S., Kursinski, E. R., Foelsche, U., Gorbunov, M., Heise, S., Kuo, Y.-H., Lauritsen, K. B., Marquardt, C., Rocken, C., Schreiner, W., Sokolovskiy, S., Syndergaard, S., and Wickert, J. (2013). Quantification of structural uncertainty in climate data records from gps radio occultation. *Atmospheric Chemistry and Physics*, 13(3):1469–1484.

[22] Sun, B., Reale, A., Schroeder, S., Seidel, D. J., and Ballish, B. (2013). Toward improved corrections for radiation-induced biases in radiosonde temperature observations. *J. Geophys. Res.*, 118:4231–4243.

[23] Sun, B., Reale, A., Seidel, D. J., and Hunt, D. C. (2010). Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics. *J. Geophys. Res.*, 115.

[24] Tradowsky, J. (2016). Radiosonde Temperature Bias Corrections using Radio Occultation Bending Angles as Reference. ROM SAF Visiting Scientist report 31, Radio Occultation Meteorology Satellite Application Facility.

[25] Tradowsky, J. S. (2015). Characterisation of radiosonde temperature biases and errors using radio occultation measurements. ROM SAF Visiting Scientist report 26, Radio Occultation Meteorology Satellite Application Facility. Available at `http://www.romsaf.org/Publications/reports/romsaf_vs26_rep_v12.pdf`.

[26] Tradowsky, J. S., Burrows, C. P., Healy, S. B., and Eyre, J. R. (2017). A new method to correct radiosonde temperature biases using radio occultation data. *Journal of Applied Meteorology and Climatology*, 56(6):1643–1661. doi:10.1175/JAMC-D-16-0136.1.

[27] Vaisala (2013a). Vaisala Radiosonde RS41 Measurement Performance. Ref. B211356EN-A.

[28] Vaisala (2013b). Vaisala Radiosonde RS92-SGP. Available at: https://www.vaisala.com/sites/default/files/documents/RS92SGP-Datasheet-B210358EN-F-LOW.pdf. Ref. B210358EN-F.

[29] Walters, D., Wood, N., Vosper, S., and Milton, S. (2014). ENDGame: A new dynamical core for seamless atmospheric prediction. Met Office Report, Met Office. Available at `http://www.metoffice.gov.uk/media/pdf/s/h/ENDGameGOVSci_v2.0.pdf`.

[30] WMO-No. 306 (2011). Manual on Codes, International Codes, Volume I.2. Technical Report WMO-No. 306, World Meteorological Organization. 2015 edition, available at `http://library.wmo.int/pmb_ged/wmo_306-I2_en.pdf`.

Ref: SAF/ROM/DMI/REP/VS/37
Version: 1.0
Date: 05/07/2019

ROM SAF CDOP-3
Visiting Scientist Report 37

EUMETSAT
ROM SAF

A Acronyms and abbreviations

COSMIC	Constellation Observing System for Meteorology, Ionosphere, and Climate
DMI	Danish Meteorological Institute (ROM SAF Leading Entity)
ECMWF	The European Centre for Medium-Range Weather Forecasts
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GCOS	Global Climate Observing System
GNSS	Global Navigation Satellite System
GRAS	GNSS Receiver for Atmospheric Sounding (onboard Metop)
GRUAN	GCOS Reference Upper-Air Network
GTS	Global Telecommunication System (used to submit measurements to be used in NWP)
IEEC	Institut d'Estudis Espacials de Catalunya
Met Office	United Kingdom Meteorological Office
Metop	Meteorological Operational Polar satellite (EUMETSAT)
NCEP	National Centers for Environmental Prediction
O-B, OB	Observation minus Background
RO	Radio Occultation
ROM SAF	Radio Occultation Meteorology SAF (former GRAS SAF)
RS	Radiosonde
SAF	Satellite Application Facility (EUMETSAT)
SEA	Solar Elevation Angle
Tdry	Dry temperature