Validation Report: Reprocessed Level 3 gridded data

Version 1.2

25 November 2018

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)

Ref: SAF/ROM/DMI/REP/GRD/001

DOCUMENT AUTHOR TABLE

	Author(s)	*Function*	*Date*
Prepared by:	Hans Gleisner	ROM SAF Scientist, Climate Coordinator	25/11 2018
Reviewed by (internal):	Sean Healy	ROM SAF Science Coordinator	29/5 2018
Approved by:	Kent B. Lauritsen	ROM SAF Project Manager	20/12 2018

DOCUMENT CHANGE RECORD

Version	*Date*	*By*	*Description*
0.1	25/5 2018	HGL	Version for internal review, in preparation for DRR-RE1.
1.0	8/6 2018	HGL	Version for DRR-RE1 and ORRs review. Included comments from internal review.
1.1	31/8 2018	HGL	Revised version after DRR-RE1 and ORR3/5 reviews: RID 292,294,610: Major changes implemented. RIDs 288,428,440,446,449,454,458,459,471,567,573, 576,592,598,613,618: Minor changes implemented. RIDs 287,289,291,293,425,450,460,561-564,568-572, 574,575,577-584, 587-591,593-597,599-600,602,603, 605-609,611,612,614-617: Editorial changes implemented. Sections 2.1 and 2.2: New sections, tables added.
1.2	25/11 2018	HGL	Updated version based on ICDR concept discussions at ROM SAF SG22: Page 4: Update of the ROM SAF introduction. Section 1.4: New definitions of product types. Appendix A: Now describing the ICDR and their SeSp requirements. Appendix B: New appendix describing offline data and their SeSp requirements.

ROM SAF

The Radio Occultation Meteorology Satellite Application Facility (ROM SAF) is a decentralised processing centre 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......5

1. INTRODUCTION 6

1.1 PURPOSE OF THE DOCUMENT 6
1.2 APPLICABLE AND REFERENCE DOCUMENTS 6
1.2.1 Applicable documents 6
1.2.2 Reference documents 7
1.3 ACRONYMS AND ABBREVIATIONS 8
1.4 DEFINITIONS 9
1.5 OVERVIEW OF THIS DOCUMENT 10

2. BACKGROUND 11

2.1 SATELLITE MISSIONS AND TIME COVERAGE......11
2.2 LEVEL 1A INPUT DATA......11
2.3 LEVEL 1B AND LEVEL 2 RO PROFILE DATA......12
2.4 ERA-INTERIM REFERENCE DATA......12
2.5 LEVEL 3 PROCESSING 13
2.6 QUALITY CONTROL OF PROFILES 14

3. LEVEL 3 GRIDDED MONTHLY MEAN RO DATA 16

3.1 ROM SAF CLIMATE DATA RECORDS 16
3.2 DATA NUMBERS......17
3.3 QUALITY CONTROL REJECTION RATES......18
3.4 EVOLUTION OF DATA QUALITY WITH TIME 19
3.5 SAMPLING DISTRIBUTIONS 21

4. VALIDATION 22

4.1 VALIDATION STRATEGY......22
4.2 COMPARISON WITH ERA-INTERIM REANALYSIS DATA 23
4.2.1 Figures: RO and ERA-Interim latitude-height plots......24
4.2.2 Figures: O-B time-height plots......27
4.2.3 Figures: O-B time series 37
4.2.4 Figures: Spatial distribution of O-B differences 46
4.2.5 Discussion and conclusions......50
4.3 COMPARISON BETWEEN RO MISSIONS 53
4.3.1 Figures: Mission difference time series......55
4.3.2 Figures: Metop setting/rising difference time series 64
4.3.3 Figures: Spatial distribution of mission differences 67
4.3.4 Figures: Anomaly time series......70
4.3.5 Figures: Anomaly difference time series......73
4.3.6 Discussion and conclusions......82
4.4 VALIDATION OF THE SAMPLING ERROR CORRECTION METHOD 84

5. COMPARISON WITH PRODUCT REQUIREMENTS...... 86

5.1 FIGURES: O-B TIME-HEIGHT PLOTS IN TERMS OF PRD ACCURACIES......87
5.2 FIGURES: COMPLIANCE WITH PRD REQUIREMENTS......97
5.3 DISCUSSION AND CONCLUSIONS......102

6. SERVICE SPECIFICATIONS......103

7. MAIN CONCLUSIONS 105

ANNEX A. INTERIM CLIMATE DATA RECORDS...... 105
ANNEX B. OFFLINE DATA PRODUCTS...... 108

Executive Summary

The ROM SAF Climate Data Record version 1.0 (CDR v1.0) is based on measurements by the CHAMP, GRACE, COSMIC, and Metop Radio Occultation (RO) missions. The Level 3 data consist of monthly means on a global latitude-altitude grid (5 degrees in latitude by 200 meters in altitude), averaged from a large number of near-vertical profiles of relevant geophysical variables, including bending angle, refractivity, temperature, and humidity. The profile data (Level 1B and Level 2) have been generated by the ROM SAF from excess phase and amplitude data (Level 1A) provided by UCAR (CHAMP, GRACE, and COSMIC) and EUMETSAT (Metop).

The ROM SAF CDR v1.0 consists of five broad data products, each containing 8 geophysical variables. There are four single-mission data products with limited time coverage, and a combined multi-mission data product covering the whole time period from September 2001 to December 2016. The validation has a focus on the time evolution of the data and data quality, and on understanding the degree of homogeneity of the data. A central issue is whether data from different missions can be combined to form long time series of multi-mission RO data.

The present validation includes a study of the differences between observations and reanalysis (ERA-Interim) for the different RO missions. It also includes a study of the consistency between monthly mean data retrieved from different missions, during time periods with overlaps between the RO missions. Further, the sampling error correction method used in the generation of the Level 3 data is briefly examined. The main conclusion is that there is a good overall agreement between the RO observational data and the ERA-Interim reanalysis. Some of the main differences can be explained by known effects. Another conclusion is that in a core region from 8 to 35 kilometer – the exact altitude range depending on latitude and geophysical variable – there is a high degree of consistency between the RO missions, as demonstrated both by the RO–reanalysis comparisons and by the differences between RO missions during overlap periods. There are, however, some remaining persistent biases, particularly between Metop and COSMIC, that will be addressed in a follow-on validation study, with the goal to quantify the impacts on long-term trend estimates, and if possible to suggest remedies.

The present validation shows that the ROM SAF Level 3 gridded data products meet the formal requirements as stated in the ROM SAF Product Requirements Document (PRD). It is demonstrated that the Level 3 data are consistent with the reference data (ERA-Interim reanalysis), using the methods and accuracy specifications described in the PRD. On the basis of these investigations, a set of Service Specifications are suggested for the Level 3 gridded data products. These specifications describe the commitments by the ROM SAF related to the services and products provided to the users.

1. Introduction

1.1 Purpose of the document

This document describes the validation of the ROM SAF Level 3 Climate Data Records (CDRs), which have been generated in a dedicated reprocessing activity. The data are generated by the ROM SAF processing system using Level 1B and Level 2 profile data as input, together with ancillary information from ECMWF reanalysis data. The product requirements baseline is defined in the ROM SAF Product Requirements Document (PRD) [AD.3], and the methods and algorithms used in the generation of the Level 3 data products are described in the Algorithms Theoretical Baseline Document (ATBD) [RD.6].

An extensive range of plots with a direct bearing on the validation of the climate data can also be found on the ROM SAF web site (http://www.romsaf.org). Those plots should be studied in conjunction with the present report.

Product ID	**Data product**	**Product acronym**	**Mission**	**Time coverage**	**Disse-minat.**	**Prod. ver.**
CDRs						
GRM-28-L3-R1	Gridded monthly mean CDR	R[*]GMUL	MULTI	200109-201612	web	1.0
GRM-29-L3-R1	Gridded monthly mean CDR	R[*]GMET	Metop	200612-201612	web	1.0
GRM-30-L3-R1	Gridded monthly mean CDR	R[*]GCO1	COSMIC	200607-201612	web	1.0
GRM-32-L3-R1	Gridded monthly mean CDR	R[*]GCHA	CHAMP	200109-200809	web	1.0
GRM-33-L3-R1	Gridded monthly mean CDR	R[*]GGRA	GRACE	200703-201612	web	1.0

Table 1. *Gridded Level 3 monthly mean ROM SAF CDR products. Each entry in the table covers a series of geophysical variables. The asterisk* [*] *in the product acronym is a place holder for one of the following letters: 'B' for bending angle, 'R' for refractivity, 'D' for dry temperature, 'Y' for dry pressure, 'Z' for dry geopotential height, 'T' for temperature, 'H' for humidity, and 'C' for tropopause height.*

1.2 Applicable and reference documents

1.2.1 Applicable documents

The following list contains documents with a direct bearing on the contents of this document:

[AD.1] CDOP-3 Proposal: Proposal for the Third Continuous Development and Operations Phase (CDOP-3); Ref: SAF/ROM/DMI/MGT/CDOP3/001 Version 1.2 of 31 March 2016, Ref: EUM/C/85/16/DOC/15, approved by the EUMETSAT Council at its 85th meeting on 28-29 June 2016

[AD.2] CDOP-3 Cooperation Agreement: Agreement between EUMETSAT and DMI on the Third Continuous Development and Operations Phase (CDOP-3) of the Radio Occultation Meteorology Satellite Applications Facility (ROM SAF), Ref. EUM/C/85/16/DOC/19, approved by the EUMETSAT Council and signed at its 86th meeting on 7 December 2016

[AD.3] ROM SAF Product Requirements Document, SAF/ROM/DMI/MGT/PRD/001.

1.2.2 Reference documents

The following documents provide supplementary or background information, and could be helpful in conjunction with this document:

[RD.1] ROM SAF ATBD: Level 1B bending angles, SAF/ROM/DMI/ALG/BA/001.

[RD.2] ROM SAF ATBD: Level 2A refractivity profiles, SAF/ROM/DMI/ALG/REF/001.

[RD.3] ROM SAF ATBD: Level 2A dry temperature profiles, SAF/ROM/DMI/ALG/TDRY/001.

[RD.4] ROM SAF ATBD: Level 2B and 2C 1D-Var products, SAF/ROM/DMI/ALG/1DVAR/002.

[RD.5] ROM SAF ATBD: Level 2C tropopause height, SAF/ROM/DMI/ALG/TPH/001.

[RD.6] ROM SAF ATBD: Level 3 gridded data, SAF/ROM/DMI/ALG/GRD/001.

[RD.7] The ROPP Pre-processor Module User Guide, SAF/ROM/METO/UG/ROPP/004.

[RD.8] The ROPP 1D-Var Module User Guide, SAF/ROM/METO/UG/ROPP/007.

[RD.9] The ROPP Applications Module User Guide, SAF/ROM/METO/UG/ROPP/005.

[RD.10] Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy, Observing Earth’s atmosphere with radio occultation measurements using the Global Positioning System, *J. Geophys. Res.*, 102, 23429-23465, 1997.

[RD.11] Dee, D.P., et al., The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quart. J. Roy. Meteorol. Soc.*, 137, 553–597, 2011.

[RD.12] Foelsche, U., et al., Refractivity and temperature climate records from multiple radio occultation satellites consistent with 0.05%, *Atmos. Meas. Tech.*, 4, 2007-2018, 2011.

[RD.13] Steiner, A. K., Lackner, B. C., Ladstädter, F., Scherllin-Pirscher, B., Foelsche, U., and Kirchengast, G., GPS radio occultation for climate monitoring and change detection, *Radio Sci.*, 46, RSOD24, 2011.

[RD.14] Gleisner, H., Latitudinal binning and area-weighted averaging of irregularly distributed radio occultation data, GRAS SAF Report 10, 2010.

[RD.15] Ringer, M.A., and Healy, S.B., Monitoring twenty-first century climate using GPS radio occultation bending angles, *Geophys. Res. Lett.*, **35**, L05708, 2008.

[RD.16] Angerer, B., Ladstädter, F., Scherllin-Pirscher, B., Schwärz, M., Steiner, A.K., Foelsche, U., and Kirchengast, G., Quality aspects of the Wegener Center multi-satellite GPS radio occultation record OPSv5.6, *Atmos. Meas. Tech.*, 10, 4845-4863, 2017.

[RD.17] Pirscher, B., Foelsche, U., Lackner, B.C., and Kirchengast, G., Local time influence in single-satellite radio occultation climatologies from Sun-synchronous and non-Sun-synchronous satellites, *J. Geophys. Res.*, 112, D11119, 2007.

[RD.18] Ho, S.-P., et al., Estimating the uncertainty of using GPS radio occultation data for climate monitoring: Inter-comparison of CHAMP refractivity climate records 2002-2006 from different data centers, *J. Geophys. Res.*, 114, D23107, 2009.

[RD.19] Steiner, A. K., et al., Quantification of structural uncertainty in climate data records from GPS radio occultation, *Atmos. Chem. Phys.*, 13, 1469-1484, 2013.

1.3 Acronyms and abbreviations

ATBD	Algorithm Theoretical Baseline Document
CDAAC	Cosmic Data Analysis and Archive Center
CDOP	Continuous Development and Operations Phase (EUMETSAT)
CDR	Climate Data Record
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate
DMI	Danish Meteorological Institute; ROM SAF Leading Entity
ECMWF	European Centre for Medium-range Weather Forecasts
EPS	EUMETSAT Polar Satellite System
EUMETSAT	EUropean organisation for the exploitation of METeorological SATellites
GNSS	Global Navigation Satellite System
GPAC	GNSS Processing and Archiving Center
GPS	Global Positioning System (US)
GRAS	GNSS Receiver for Atmospheric Sounding (EPS/Metop)
GRIB	GRIdded Binary (WMO)
ICDR	Interim Climate Data Record
IEEC	Institut d'Estudis Espacials de Catalunya
L1	GPS carrier frequency, 1575.42 MHz
L2	GPS carrier frequency, 1227.6 MHz
LC	L Corrected (through linear combination of L1 and L2)
LEO	Low Earth Orbit
Met Office	United Kingdom Meteorological Office
Metop	Meteorological Operational Polar satellite (EUMETSAT)
MSL	Mean Sea Level
netCDF	Network Common Data Format
NRT	Near Real Time
NWP	Numerical Weather Prediction
PRD	Product Requirements Document (ROM SAF)
RO	Radio Occultation
ROM SAF	Radio Occultation Meteorology SAF (former GRAS SAF)
ROPP	Radio Occultation Processing Package (ROM SAF)
SAF	Satellite Application Facility (EUMETSAT)
WMO	World Meteorological Organization

1.4 Definitions

RO data products from the GRAS instrument onboard Metop and RO data from other missions are grouped in *data levels* (level 0, 1, 2, or 3) and *product types* (NRT, offline, CDR, or ICDR). The data levels and product types are defined below[1]. The lists of variables should not be considered as the complete contents of a given data level, and not all data may be contained in a given data level.

Data levels:

Level 0: Raw sounding, tracking and ancillary data, and other GNSS data before clock correction and reconstruction;

Level 1A: Reconstructed full resolution excess phases, total phases, pseudo ranges, SNRs, orbit information, I, Q values, NCO (carrier) phases, navigation bits, and quality information;

Level 1B: Bending angles and impact parameters, tangent point location, and quality information;

Level 2: Refractivity, geopotential height, "dry" temperature profiles (Level 2A), pressure, temperature, specific humidity profiles (Level 2B), surface pressure, tropopause height, planetary boundary layer height (Level 2C), ECMWF model level coefficients (Level 2D), quality information;

Level 3: Gridded or resampled data, that are processed from Level 1 or 2 data, and that are provided as, e.g., daily, monthly, or seasonal means on a spatiotemporal grid, including metadata, uncertainties and quality information.

Product types:

NRT product: Data product delivered less than: (i) 3 hours after measurement (SAF Level 2 for EPS); (ii) 80 min after measurement (SAF Level 2 for EPS-SG Global Mission); (iii) 40 min after measurement (SAF Level 2 for EPS-SG Regional Mission);

Offline product: Data product delivered from less than 5 days to up to 6 months after measurement, depending on the requirements. The evolution of this type of product is driven by new scientific developments and subsequent product upgrades;

CDR: Climate Data Record generated from a dedicated reprocessing activity using a fixed set of processing software[2]. The data record covers an extended time period of several years (with a fixed end point) and constitutes a homogeneous data record appropriate for climate usage;

ICDR: An Interim Climate Data Record (ICDR) regularly extends in time a (Fundamental or Thematic) CDR using a system having optimum consistency with and lower latency than the system used to generate the CDR[3].

---

[1] Note that the level definitions differ partly from the WMO definitions: http://www.wmo.int/pages/prog/sat/dataandproducts_en.php

[2] (i) GCOS 2016 Implementation Plan; (ii) http://climatemonitoring.info/home/terminology/

[3] http://climatemonitoring.info/home/terminology/ (the ICDR definition was endorsed at the 9th session of the joint CEOS/CGMS Working Group Climate Meeting on 29 March 2018)

1.5 Overview of this document

This document is organized as follows:

Chapter 1: Contains the introduction.

Chapter 2: Contains an overview of the Level 1B and Level 2 profile data used as input to the Level 3 processing. It also contains a brief overview of the Level 3 processing, and a list of the quality screening tests.

Chapter 3: Contains a description of the validation data, with a focus on characteristics that may have an impact on the accuracy and stability.

Chapter 4: Contains the main validation results.

Chapter 5: Contains a comparison with the Product Requirements.

Chapter 6: Contains a suggestion for updated Service Specifications.

Chapter 7: Contains the main conclusions of the validation.

Annex A: Contains a list of ROM SAF ICDR products and Service Specifications.

Annex B: Contains a list of ROM SAF offline products and Service Specifications.

2. Background

2.1 Satellite missions and time coverage

The missions, their respective satellites, and the covered periods for each mission are given in Table 2. It should be noted that for COSMIC not all RO instruments in the constellation have been operating in the full time period, especially towards the end of the period, and for GRACE, both RO instruments are never in operation at the same time. Metop-B was only launched in September 2012, and RO operation started in October 2012. The very early data from CHAMP (until September 2001) has not been included.

Mission	Metop	COSMIC	CHAMP	GRACE
Satellites	Metop-A Metop-B	FM1, FM2, FM3, FM4, FM5, FM6	CHAMP	GRACE-A GRACE-B
Periods	Dec'06 – Dec'16	Jul'06 – Dec'16	Sep'01 – Sep'08	Mar'07 – Dec'16

***Table 2**. Missions, satellites, and time periods in the ROM SAF reprocessing #1, that are included in the generation of Level 3 gridded monthly mean data.*

2.2 Level 1A input data

The input Level 1A data for Metop (GRM-29-R1) are provided by the EUMETSAT Secretariat as part of their reprocessed Level 1A and 1B data, version 1.4. For COSMIC (GRM-30-R1), CHAMP (GRM-32-R1), and GRACE (GRM-33-R1), the input Level 1A data are provided by UCAR/CDAAC, more specifically, CDAAC reprocessed or post-processed data. For the CDAAC data, not all missions and time periods share the same version number. Table 3 gives an overview of the different providers and the version numbers for the data used as input to the ROM SAF reprocessing #1.

Mission	Metop	COSMIC	CHAMP	GRACE
Data provider	EUMETSAT	UCAR/CDAAC	UCAR/CDAAC	UCAR/CDAAC
Input versions	1.4	2013.3520 2014.2860 2016.1120	2014.0140	2010.2640 2014.2760

***Table 3**. Input data providers and versions.*

2.3 Level 1B and Level 2 RO profile data

The starting point for the ROM SAF Level 3 processing is a large number of near-vertical profiles, one for each occultation:

- Bending angle, $\alpha(a)$
- Refractivity, $N(H)$
- Dry pressure, $p_{\text{dry}}(H)$
- Dry temperature, $T_{\text{dry}}(H)$
- Dry geopotential height, $Z(H_{p\text{dry}})$
- 1D-Var temperature, $T(H)$
- 1D-Var specific humidity, $q(H)$

and, additionally, tropopause heights, which are derived from the dry temperature profiles:

- Tropopause height, H_{TP}

Here, H is the mean-sea level (MSL) altitude, a is the impact parameter, H_p is the pressure height (a logarithmic measure of pressure), and $H_{p\text{dry}}$ is the dry-pressure height [RD.6]. The "dry" variables are retrieved from the refractivity simply by ignoring the presence of water vapour. This is a valid assumption in the upper troposphere and in the stratosphere, where the "dry" variables are accurate approximations for the corresponding physical quantities. The retrievals of the geophysical profile data are described in the associated ATBDs [RD.1-4], while the retrieval of tropopause height has its own ATBD [RD.5].

The bending angle, refractivity, and dry profiles are provided on relatively dense vertical grids reaching up to well above the region where the RO measurements provide useful information on the neutral atmosphere. The temperature and humidity profiles are given on a standard set of vertical levels ranging from the surface up to around 80 km, near the top of the atmospheric model used as *a priori* in the retrieval. Each occultation has an associated reference location and time, which is used in binning the data.

The Level 1B and Level 2 profile data, on which the present validation is based, were retrieved with the GPAC-2.3.0 system, which includes the ROPP-8.1 software package (an internal release with adaptions made by the DMI) and the 1DV-3.3 software.

2.4 ERA-Interim reference data

ERA-Interim reanalysis data from ECMWF [RD.11] are used both for sampling-error correction of the Level 3 gridded monthly mean data, and as *a priori* in the 1D-Var retrieval of Level 2b profile data. The reanalysis fields are obtained from ECMWF as GRIB files holding data on a 1.0˚x1.0˚ latitude-longitude grid, as well as on a coarser 2.5˚x2.5˚ grid. The lower resolution is used for the sampling-error correction, as it is roughly comparable to the horizontal resolution of RO measurements.
Each occultation has reference latitude, longitude, and time associated with it. To retrieve reanalysis profiles co-located with the observations, we interpolate (bi-linearly) from the model grid to the reference location, followed by linear interpolation between adjacent

time steps. For the 1D-Var retrievals, we use short-term forecasts rather than the analysis fields. In that case, the interpolation in time is made between two model forecast lead times. Further details on this are found in [RD.4].

For each observed profile we thus obtain a corresponding co-located model profile, which is mapped to refractivity, bending angle, and 'dry' variables using forward-model routines from the ROPP-8.1 software package [RD.8].

The use of ERA-Interim data as reference for the validation has some limitations [RD.11]. First, there is a risk for circularities: RO data are assimilated by the ERA-Interim reanalysis system and ERA-Interim data are used in the generation of the some of the geophysical variables retrieved from the RO measurements. This is briefly discussed in Section 4.2. Second, the ERA-Interim data records exhibit bias shifts due to changes in the global observing system, particularly in the early years when fewer data were available for assimilation. This may obscure the validation results.

2.5 Level 3 processing

The Level 3 gridded data are generated from the Level 1B and Level 2 profile data through rather straight-forward *binning and averaging* [RD.6]. A set of equal-angle latitudinal bins are defined and all valid observations that fall within a latitude bin and calendar month undergo a weighted averaging to form a zonal mean for that latitude and month. The purpose of the weighting is to more closely approximate an area-weighted average.

The sampling errors are estimated by sub-sampling an atmospheric model (currently, the ERA-Interim reanalysis) at the observed times and locations. Based on these estimates, we do a sampling-error correction, or adjustment, by subtracting the estimated sampling errors from the observed means [RD.12,13]. The errors remaining after the sampling-error correction are referred to as residual sampling errors.

The uncertainty of the monthly mean is estimated as a combination of the per-profile measurement uncertainties and the uncertainties due to the residual errors remaining after the sampling-error correction [RD.6]. In principle, there is also a structural uncertainty due to algorithmic choices and underlying processing assumptions, but these are not explicitly quantified by the ROM SAF Level 3 algorithms. However, the ROM SAF has participated in activities with the explicit purpose to quantify structural uncertainties by comparing independent processing of the same input data. Results from these studies have been published in the scientific literature [RD.18,19].

In summary, the RO Level 3 gridded data products are generated by the following steps:

1) quality control and flagging of profiles that are identified as non-nominal ('bad')
2) vertical interpolation of profiles onto a regular Level 3 height grid
3) weighted averaging into monthly latitude bins
4) estimation of sampling errors in the monthly means
5) estimation of uncertainties (measurement and sampling) in the monthly means

6) estimation of *a priori* information in the monthly means
7) formatting of the Level 3 gridded data and meta-data into netCDF files

The generation of zonally gridded monthly mean data is followed by further averaging into seasonal and annual means, and into regional, hemispheric, and global means.

The Level 3 gridded data, which are validated in the present document, were retrieved with the GPAC-2.3.0 system, which includes the ROPP-8.1 software package (an internal release with adaptions made by the DMI) and the ROMCLIM-1.2 software.

2.6 Quality control of profiles

The purpose of the quality control is to identify profiles that are likely to provide an invalid representation of the atmosphere. Before processing the atmospheric profiles into gridded monthly-mean data, all profiles are checked against a set of criteria indicating non-nominal conditions (listed in Table 4). Some of these criteria are seldom met – they are only a basic sanity check to ensure that corrupt data do not affect the climate data (*QC-0*). Other tests are designed to identify occultations with degraded bending-angles (*QC-2*), that could be regarded as outliers (*QC-3*), or that have problems with the 1D-Var processing (*QC-4*). The actual quality control limits are effectively a compromise between the need to remove “bad” profiles and the wish to keep “good” profiles.

The first step (*QC-0*) in the quality screening procedure is a basic check to ensure that the bending angle (refractivity) profile reaches above 60 km and below 20 km impact altitude (MSL altitude). Bending angles must fall within the range -1 to 100 mrad, and refractivities must fall within the range 0 to 500 N-units. The independent variables (impact altitudes and MSL altitudes) are required to vary monotonously.

In the next step (*QC-2*), the noise properties of the L2 signal and the degree of fit of the raw LC bending angle to the background bending angle is checked. The *L2 quality score* quantifies the degradation of the L2 signal through the RMS difference of the L1 and L2 impact parameter series obtained from a radio-holographic analysis [RD.7]. The two *SO scaling factors* quantify the degree of fit to a background bending angle profile. This QC step also includes a requirement that the background bending-angle data should only play a minor role below 40 km altitude, which is indicated by the *LC weighting factor*.

The next QC step (*QC-3*) removes data identified as outliers. This is done by comparing the observed bending angles, refractivities, and dry temperatures to ECMWF reanalysis data within specified height intervals.

When an occultation does not pass a test, the whole profile is discarded. No attempt is made to identify “good” data points within a profile containing “bad” data points.

If an occultation does not pass one or several of the above tests, the bending angle, refractivity, and dry variables are marked as non-nominal. Otherwise, they are regarded as nominal, and the refractivity profiles are passed on to the 1D-Var processing. This is followed by another QC step (*QC-4*) which checks the quality of the 1D-Var solution. If the occultation passes all tests up to, and including, *QC-3*, but fails in *QC-4*, the bending angle, refractivity, and dry-variable profiles are used, while the wet profiles obtained from the 1D-Var solution are discarded.

The fraction of data rejected in each QC step varies between the RO missions, and for some RO missions it also varies with time (see Figure 2).

QC-0: basic sanity check
Identification of occultations with too small vertical extension, too few useful data points, the presence of invalid data points, or height variables that form a non-monotonous series.
- $\alpha(H_a)$ must reach below 20 km and above 60 km - $\alpha(H_a)$ values must fall within valid range: [-1,100] mrad - H_a values must form a monotonous series - $N(H)$ must reach below 20 km and above 60 km - $N(H)$ values must fall within valid range: [0,500] N-units - H must form a monotonous series
QC-1: (not used)
QC-2: bending angle quality
Checking of a) the quality of the bending angles, as quantified by the noise on the L2 impact parameter series, b) the fit of the raw LC bending angle to a background bending angle profile, and c) that the background bending-angle data only play a minor role below 40 km altitude.
- L2 quality score must be less than 30.0 - SO scaling factor 1 must fall in the interval [0.92,1.08] - SO scaling factor 2 must fall in the interval [0.60,1.40] - LC weighting factor must be larger than 0.90 below 40 km altitude
QC-3: identification of outliers
Identification of outliers by comparing with ECMWF reanalysis data mapped to refractivity, bending angle, and dry temperature.
- α must deviate from reanalysis by less than 90% between 10-40 km - N must deviate from reanalysis by less than 10% between 5-35 km - N must deviate from reanalysis by less than 20% below 5 km - T_{DRY} must deviate from reanalysis by less than 20 K between 30-40 km
QC-4: quality of 1D-Var solution
Identification of occultations that have problems converging at an acceptable 1D-Var solution.
- the 1D-var algorithm must converge within 25 iterations - the penalty function $2J/N_{obs}$ must be smaller than 5.0 at convergence

Table 4. Summary of the ROM SAF quality control of the Level 1 and 2 data used as input to the Level 3 processing. When an occultation does not pass a test, the whole profile is discarded. No attempt is made to identify “good” data points within a profile containing “bad” data points. QC-0 to QC-3 affect all variables, while QC-4 only affects the 1D-Var variables. QC-1 is currently not used operationally (only used experimentally for screening based on noise in excess phase time series).

3. Level 3 gridded monthly mean RO data

3.1 ROM SAF climate data records

The ROM SAF Climate Data Record (CDR) consists of profile data (Level 1B and Level 2) and gridded data (Level 3). The Level 3 data incorporates gridded monthly means and associated quantities (standard deviations, data numbers, sampling error estimates, etc.) of:

- Bending angle, $\alpha(a)$
- Refractivity, $N(H)$
- Dry pressure, $p_{\mathrm{dry}}(H)$
- Dry temperature, $T_{\mathrm{dry}}(H)$
- Dry geopotential height, $Z_{\mathrm{dry}} = Z(H_{p\mathrm{dry}})$
- 1D-Var temperature, $T(H)$
- 1D-Var specific humidity, $q(H)$
- Tropopause height, H_{TP}

As described in Section 2.1, H is the MSL altitude, a is the impact parameter, and $H_{p\mathrm{dry}}$ is the dry-pressure height [RD.6]. The monthly means, and the associated variables, are defined on zonal grids, 200 meters in height by 5 degrees in latitude [RD.6].

There are 5 separate Level 3 reprocessed data products: based on Metop (GRM-29-L3), COSMIC (GRM-30-L3), CHAMP (GRM-32-L3), GRACE (GRM-33-L3), and a multi-mission data set including data from all four missions (GRM-28-L3). These are listed in Table 1. The CHAMP, GRACE, and COSMIC data are based on input Level 1A (excess phases and amplitudes) from UCAR, while the Metop data are based on input Level 1A data from EUMETSAT.

The applicability of long time series of RO data depends crucially on the homogeneity of the underlying data. Inhomogeneity could, e.g., be due to degradation with time of RO instruments leading to changes in bending-angle noise, firmware or software changes with time, or subtle differences between RO instruments and/or missions. It is one of the goals of the present validation to identify such inhomogeneities.

Another source of inhomogeneity is the different sampling characteristics of the missions. A correction procedure is applied to the gridded monthly-mean data to reduce the impacts of the sampling on the RO data time series [RD.6]. The impacts of these correction methods are investigated as a part of the present validation.

3.2 Data numbers

The raw data numbers available as input to the ROM SAF processing system are shown in Figure 1. Since 2006, the COSMIC satellites have constituted the backbone of the global RO observing system – a role recently overtaken by the Metop satellites which provide more data than COSMIC since late 2015. The COSMIC mission is also important as it provides global sampling at all local times, unlike the Metop satellites which are in Sun-synchronous orbits. An important role is played by CHAMP, which provides extension of the RO time series back to 2001.

During 2007 to 2009, the number of daily occultations peaked at well above 3000, and the launch of the second Metop satellite, in combination with an update of the operational mode of the COSMIC mission, led to similar daily data numbers in 2013. The total number of occultations up to December 2016 is around 12 million. The variation of data numbers with time is primarily a consequence of the launch or loss of RO satellites, but in some cases also due to problems with the individual satellites (e.g., power problems onboard COSMIC satellites after 2009).

Data numbers are important for the accuracy of the gridded monthly-mean data, but it is not the only factor; the data quality and the quality control procedures also play a role (see Sections 3.3 and 3.4), as well as factors related to sampling (see Section 3.5).

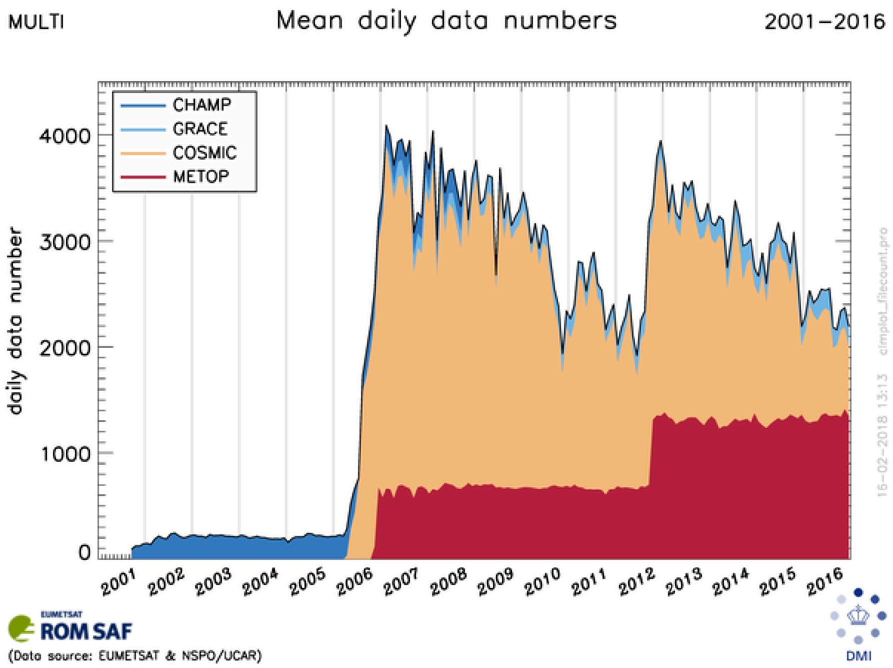


Figure 1. *Number of occultations available for climate data record generation. Since 2015, Metop provides the backbone of the global RO observing system. CHAMP is important, as it provides an extension of the time series back to 2001, while COSMIC dramatically increased the total number of occultations available during 2006 to 2016. Unlike Metop, COSMIC samples the full diurnal cycle, which is important for the generation of climate data records.*

3.3 Quality control rejection rates

As described in Section 2.4, the Quality Control (QC) procedures are applied to the input profile data, and the data that pass the QC are used as input to the Level 3 processing [*RD.6*]. The QC thus reduces the data numbers available for the Level 3 processing.

The percentages of data remaining after the sequence of QC tests are shown in Figure 2. On average, around 10-20% of the occultations are rejected, although with large differences between the RO missions. These differences reflect the intrinsic quality of the measurements, but there are also differences between the missions in the pre-processing of data up to Level 1A which may be reflected in the QC rejection rates. Metop and GRACE show the highest throughput of data. COSMIC and CHAMP (after February 2002) have roughly similar overall rejection rates, but for somewhat different reasons. For COSMIC, the QC-2 set of tests (dominated by the L2 score test) is an important factor, while for CHAMP a larger fraction of data are lost already in the initial processing up to bending angles. CHAMP also exhibits very high rejection rates up to around February-March 2002.
The semi-annual variation in the QC rejection rates are also seen in the number of bending-angle profiles with high noise levels (Figure 3b).

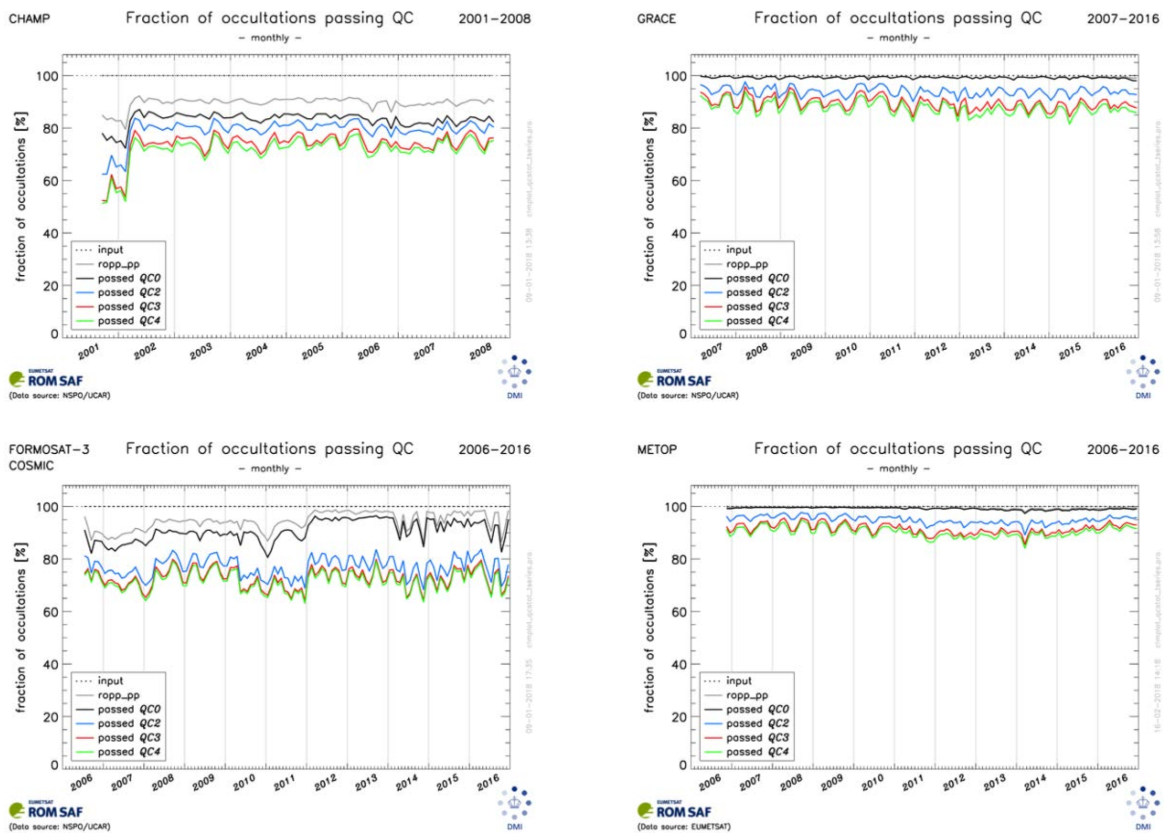


Figure 2. *Fraction of occultations available for the Level 3 gridded data generation, after the consecutive QC steps described in Section 2.4. QC-0 is a fundamental sanity check of bending-angle and refractivity profiles, QC-2 is a check based on the L2 and SO quality scores, QC-3 consists of systematic removal of outliers through comparison with ECMWF, and QC-4 is a check on the 1D-Var solution.*

3.4 Evolution of data quality with time

The evolution of data quality with time, for the different RO missions, is an important aspect of the multi-mission RO data record. For each bending angle profile, we define the high-altitude *noise floor* as the minimum standard deviation of the bending angle within a 7.5 kilometer altitude interval (roughly the bending-angle scale height) located between 60 and 80 kilometer. This definition is an attempt to identify the instrumental noise on top of a variable background of ionospheric noise. In practice, the noise floor includes contributions from both noise sources.

Figure 3a shows the noise floor distributions for the four RO missions. They are roughly consistent with the known differences between RO instruments. 80% of the bending angle profiles have noise floors smaller than about 2.9 μrad for CHAMP, 1.9 μrad for GRACE, 1.4 μrad for COSMIC, and 0.8 μrad for Metop. The COSMIC data exhibit a long tail of bending angle profiles with large noise floors, much more pronounced than for, e.g. the Metop data.

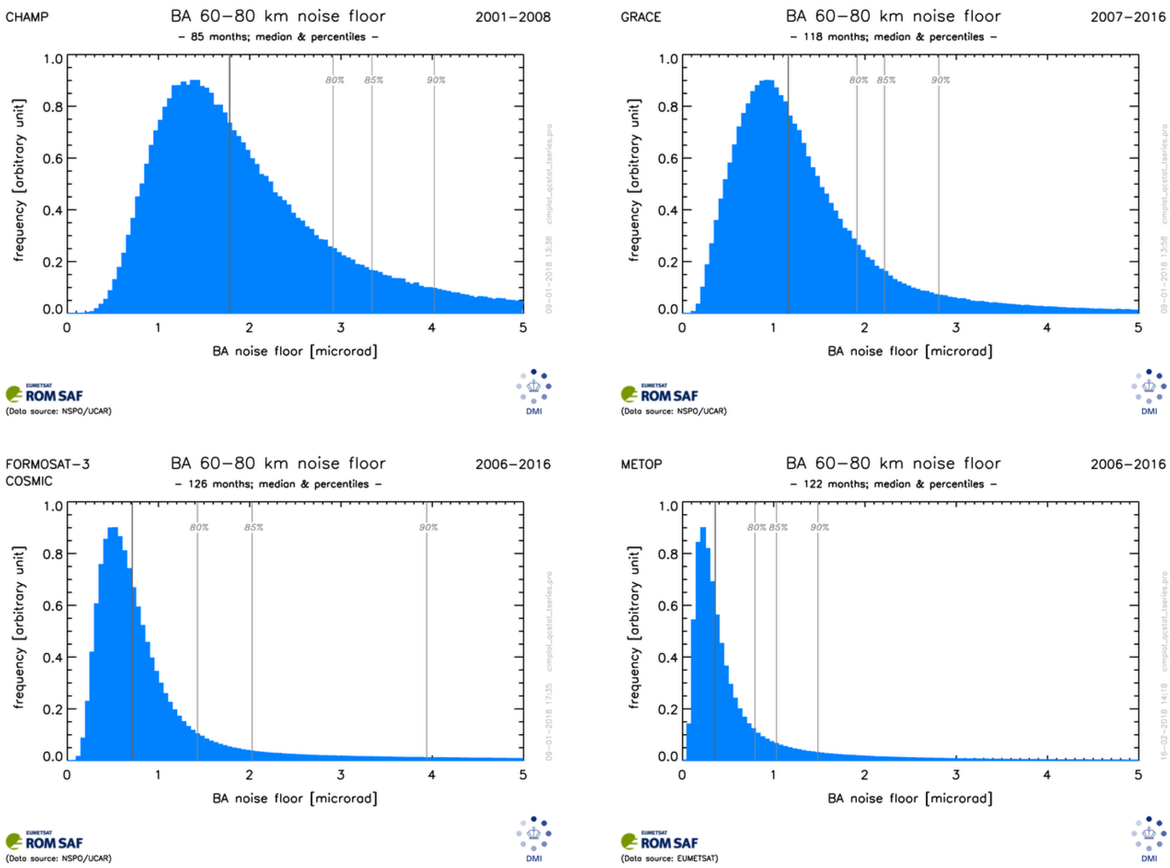


Figure 3a. *The bending angle noise floor for the four RO missions. The medians and percentiles of the distributions, indicated in the figures, are plotted as function of time in Figure 3b.*

For all missions, there is some degree of variation of data quality with time (Figure 3b). CHAMP has a significantly larger bending-angle noise during the first months of the data record. Following an instrument software update in March 2002, the bending-angle noise settles at an almost constant level. The bending-angle noise in the GRACE data is relatively constant, except for a sudden decrease in early 2014. The cause of this change is not known, and need to be further investigated.

For COSMIC, there is a substantial, although highly irregular, increase with time of a long tail of profiles with high bending-angle noise. Investigations show the long tail to be caused mainly by rising occultations, and the variation over time, including the upward trend shown in Figure 3b can be traced to individual COSMIC satellites with an apparently degraded performance [RD.16]. It should be noted that the rising occultations are rejected by the QC procedures to a much higher degree than the setting occultations.

Metop shows an interesting pattern which may be related to the solar cycle. This is a topic for future investigations.

All four missions exhibit – to a larger or smaller extent – a semi-annual pattern in the overall bending-angle noise. In most cases, the fraction of high-noise occultations seems to be larger around December/January and June/July.

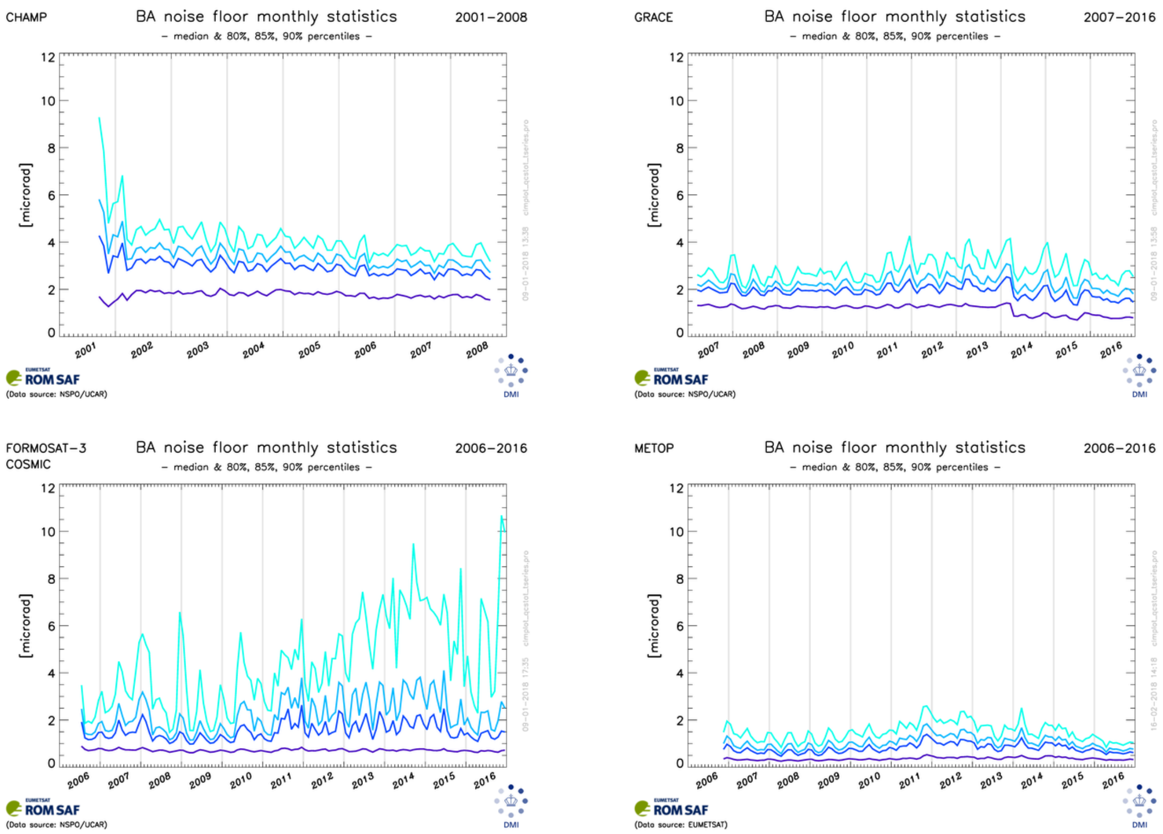


Figure 3b. *The bending angle noise floor as function of time for the four RO missions. The noise floor is here quantified by the medians (lowest, dark blue curves) and the 80%, 85%, and 90% (uppermost, light blue curves) percentiles of the monthly noise floor distributions (see Figure 3a).*

3.5 Sampling distributions

All four RO missions have a relatively uniform distribution of data numbers in longitude, whereas the latitude distribution is non-uniform [RD.14]. Figure 4 shows the longitudinal (upper panels) and latitudinal (middle panels) distributions of data numbers per unit area for CHAMP, COSMIC, and Metop. The non-uniform latitudinal distribution is a consequence of the GNSS and LEO satellite orbits in combination with the limb-sounding observing geometry of the RO instrument. The distributions are relatively similar for the different missions, with the main differences due to the inclination of the orbits.

The distribution in local time is only uniform for the COSMIC mission (Figure 4, lower panels). The reason is that the COSMIC mission includes several satellites in different orbits, and that these orbits drift in local time. The Sun-synchronous nature of the Metop orbits, makes the local-time distribution very narrow. CHAMP, which is drifting in local time, has a highly non-uniform local-time distribution on a monthly time scale, but a more uniform distribution on seasonal, or longer, time scales.

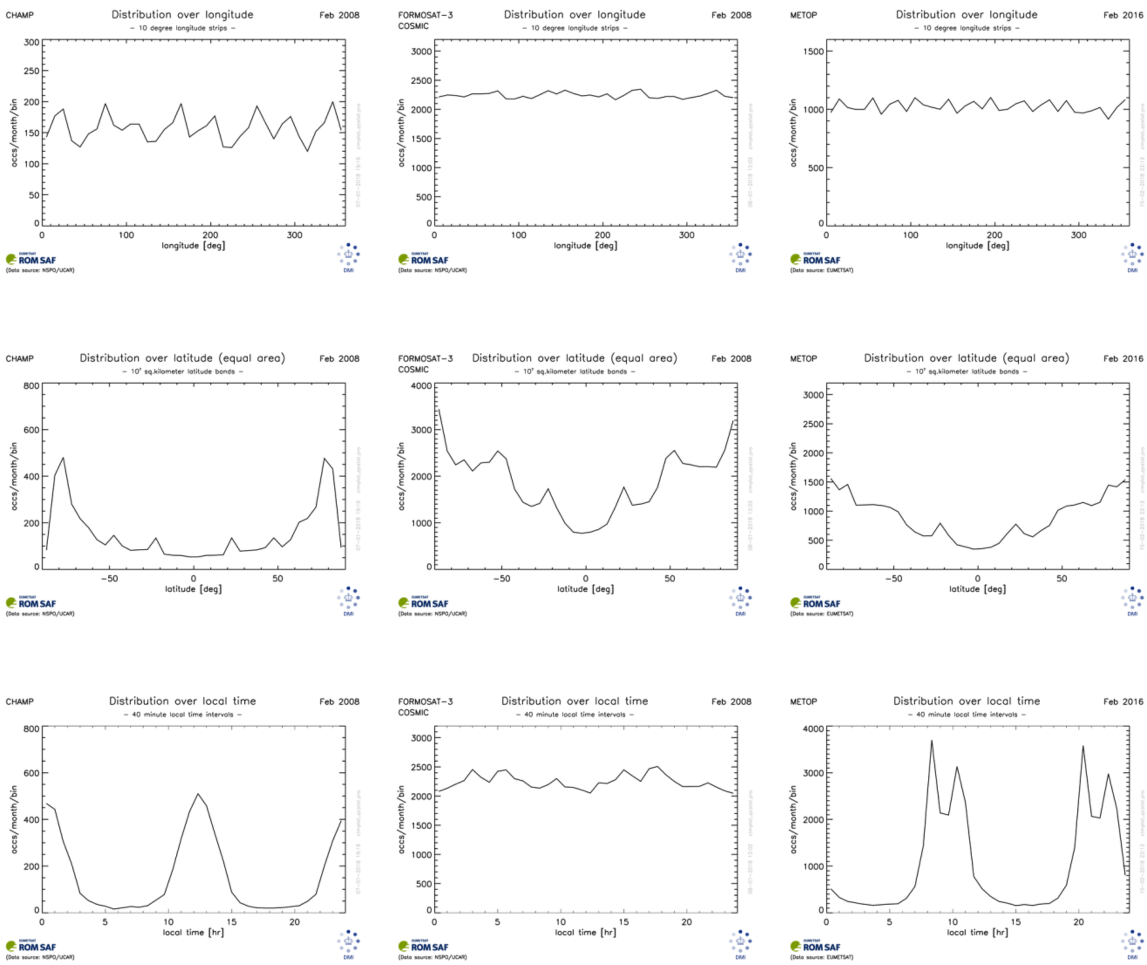


Figure 4. *Distribution of occultations over longitude (upper panels), latitude (middle panels), and local time (lower panels) for the CHAMP, COSMIC, and Metop missions. On a monthly time scale the distributions over longitude are relatively uniform for all mission, but the distribution in local time is only near uniform for the COSMIC mission which includes several satellites drifting in local time.*

4. Validation

4.1 Validation strategy

The purpose of the validation is to demonstrate that the ROM SAF Level 3 reprocessed gridded data products, and the associated time series, have the expected quality and characteristics, and that the Level 3 data products meet the formal requirements as stated in the ROM SAF Product Requirements Document (PRD) [*AD.3*]. The latter is done by demonstrating that the ROM SAF Level 3 gridded monthly mean data are consistent with the reference data (here, ERA-Interim), using the methods and accuracy specifications described in the PRD.

The Level 3 data set to be validated covers more than 15 years. It contains 5 broad data products and 8 geophysical variables. Some selection of data and properties to be included in the validation report is obviously required. The time-series aspects of the Level 3 data are particularly important. This means that understanding the time evolution of the data and data quality, and understanding the degree of homogeneity of the data record, is crucial, as it determines the usefulness of the RO data for, e.g, climate trends studies. A central issue is whether data from different missions can be combined to form long time series of multi-mission RO data.

The primary means of validation of the Level 3 gridded monthly-mean data records is a comparison with the corresponding data from ERA-Interim reanalyses (Section 4.2), where it must be acknowledged that ERA-Interim data is not the truth. In some cases, we are able to spot problems with the reanalysis data from differences with respect to the observational RO data. The RO–ERA-Interim comparisons can also be used to check the consistency of data from different RO missions, and they are further used to demonstrate the formal compliance with the ROM SAF PRD requirements (Section 5).

In addition to the comparison with ERA-Interim, the differences between RO missions during mission overlap periods are investigated (Section 4.3). The overlap periods are also used to validate the methods for sampling-error correction (Section 4.4).

The present validation thus incorporates the following parts:

- Comparison of observed Level 3 RO data with the corresponding ERA-Interim reanalysis data (Section 4.2);
- Comparison of Level 3 RO data retrieved from different RO missions during mission overlap periods (Section 4.3);
- Validation of the methods for sampling-error correction based on the mission overlap periods (Section 4.4);
- Use of the RO – ERA-Interim comparison for a formal check of compliance with the ROM SAF PRD requirements (Section 5);

4.2 Comparison with ERA-Interim reanalysis data

In this section, the Level 3 gridded monthly-mean RO data are compared to the corresponding means generated from co-located ERA-Interim reanalysis short-term forecasts.

There are a few issues to consider when comparing observed RO data with ERA-Interim data. First, RO data are assimilated by the ERA-Interim reanalysis system [RD.11]. Second, the ERA-Interim short-term forecasts have been used as *a priori* in the 1D-Var processing of the Level 2 profile data used as input to the Level 3 processing [RD.4].

These two factors complicate the choice of reference for the comparison. We have chosen to use ERA-Interim short-term forecast data as reference, the forecast time range varying from 3 to 12 hours. In this way, we avoid comparison of two data sets (observational and ERA-Interim) containing the same occultations, which would be the consequence of using ERA-Interim analysis data as reference. However, the ERA-Interim forecast data have a significant influence on the comparisons of the temperature and humidity, particularly at altitudes and/or latitudes where the background information dominates the 1D-Var solution.

In Section 4.2.1 we show examples of ROM SAF Level 3 gridded monthly-mean data and the corresponding ERA-Interim data (Figs. 6a-b), together with differences between the two data sets (Fig. 7). The examples are taken from Metop, April 2014, and are presented as monthly latitude-height plots.

In Section 4.2.2 we average the observation minus background (O-B) differences in 5 latitude bands, to get time series of the O-B differences. Data are shown for Metop (Figs. 8a-e) and COSMIC (Figs. 9a-e). They include 5 geophysical variables, and the O-B differences are presented as time-height plots covering the length of the time series.

In Section 4.2.3 we present time series of O-B differences for the full length of the CHAMP, COSMIC, and Metop time series. The O-B differences have been vertically and latitudinally averaged. Results are presented for bending angles (Figs. 10a-c), refractivity (Figs. 10d-f), and dry temperature (Fig. 10g-i).

In Section 4.2.4 we show the spatial pattern of differences between RO data and ERA-Interim, obtained by temporal averaging over the length of the time series. We do this for the 8 geophysical variables and the 4 missions, and the results are shown in Figs. 11a-d.

In Section 4.2.5 the results presented in Sections 4.2.1-4.2.4 (Figs. 6-11) are discussed.

4.2.1 Figures: RO and ERA-Interim latitude-height plots

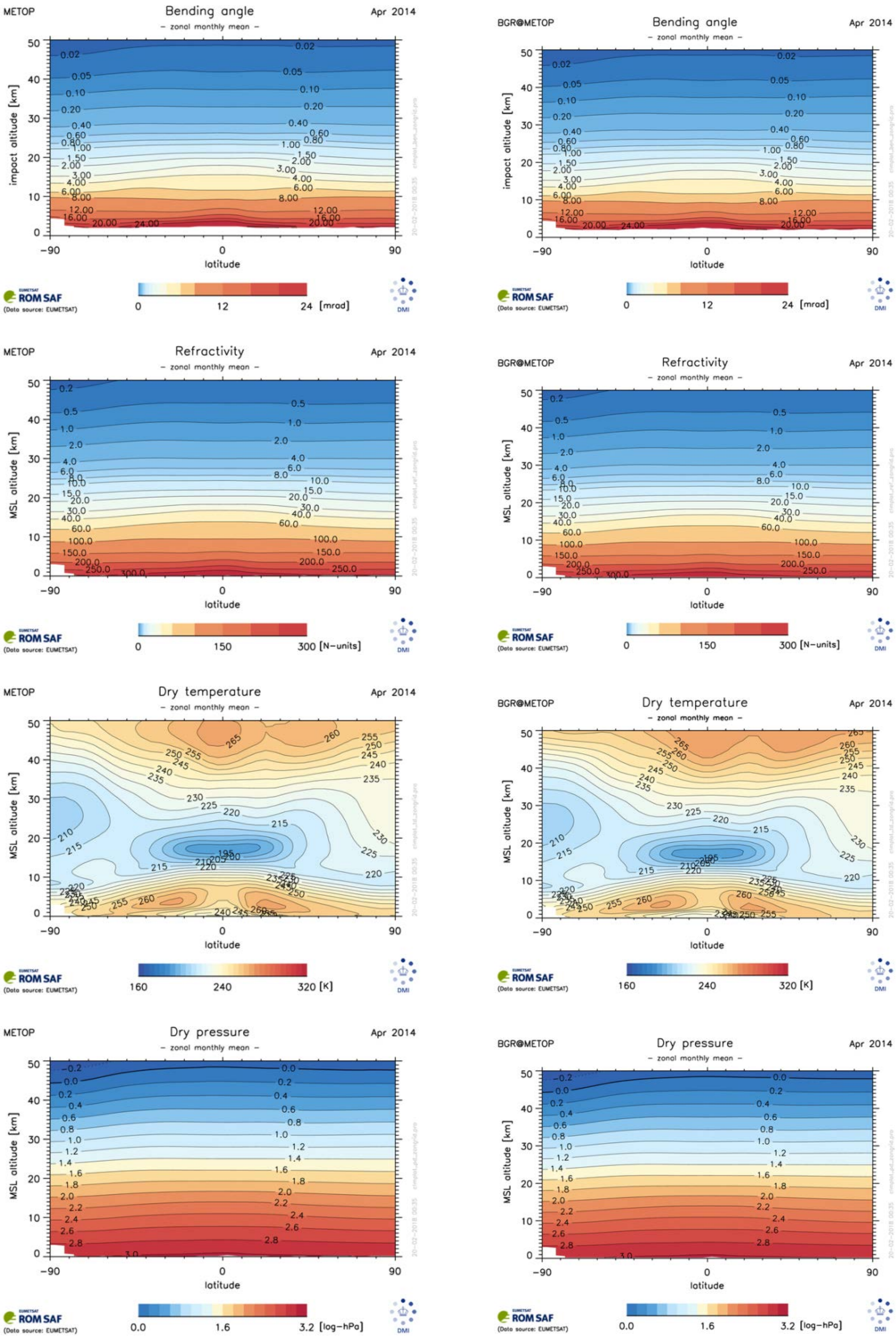


Figure 6a: *Examples of ROM SAF Level 3 gridded monthly-mean data (left panels) and the corresponding ERA-Interim data (right panels), for bending angle, refractivity, dry temperature, and dry pressure based on Metop data, April 2014.*

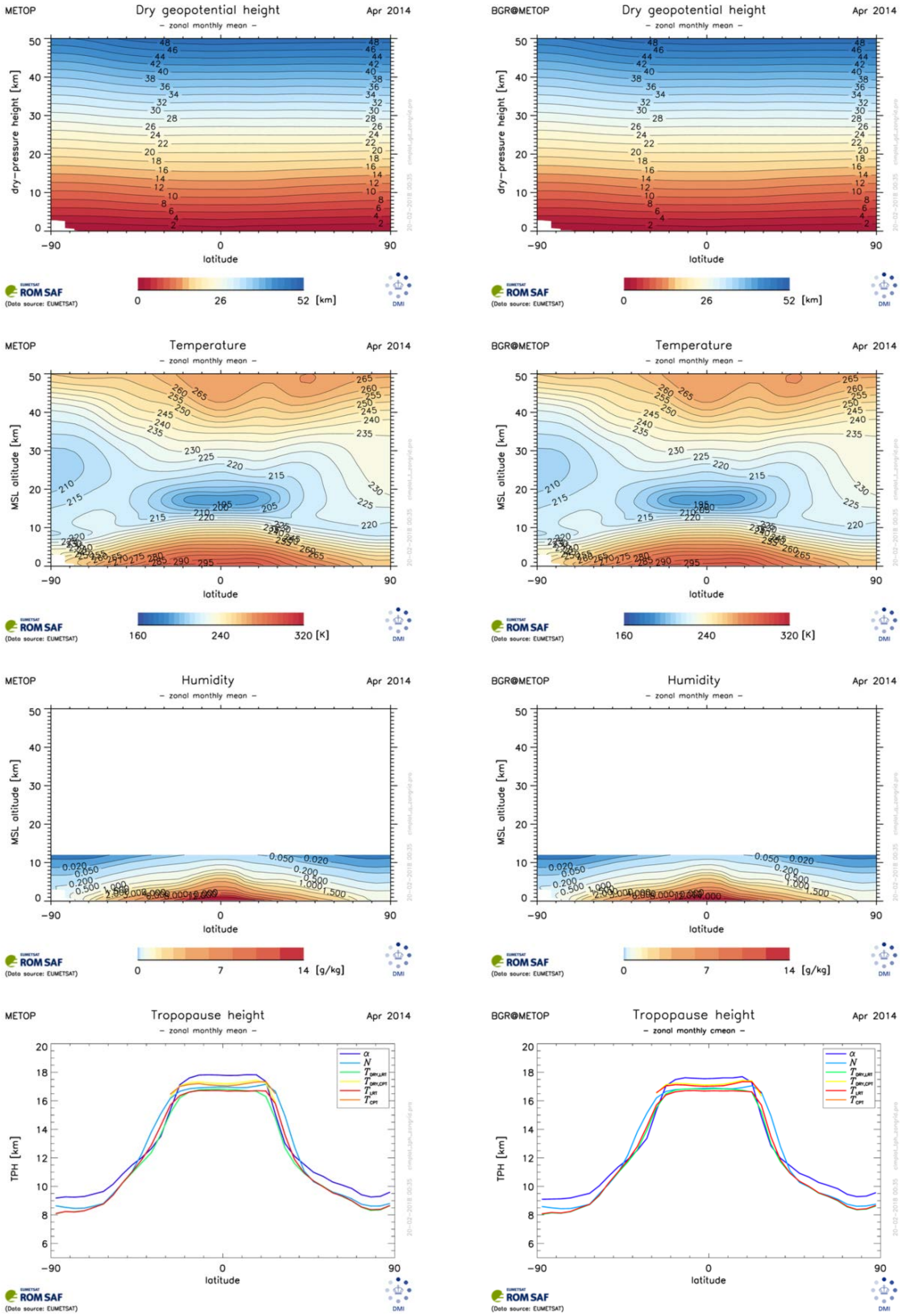


Figure 6b: *Examples of ROM SAF Level 3 gridded monthly-mean data (left panels) and the corresponding ERA-Interim data (right panels), for dry geopotential height, temperature, humidity, and tropopause height based on Metop data, April 2014.*

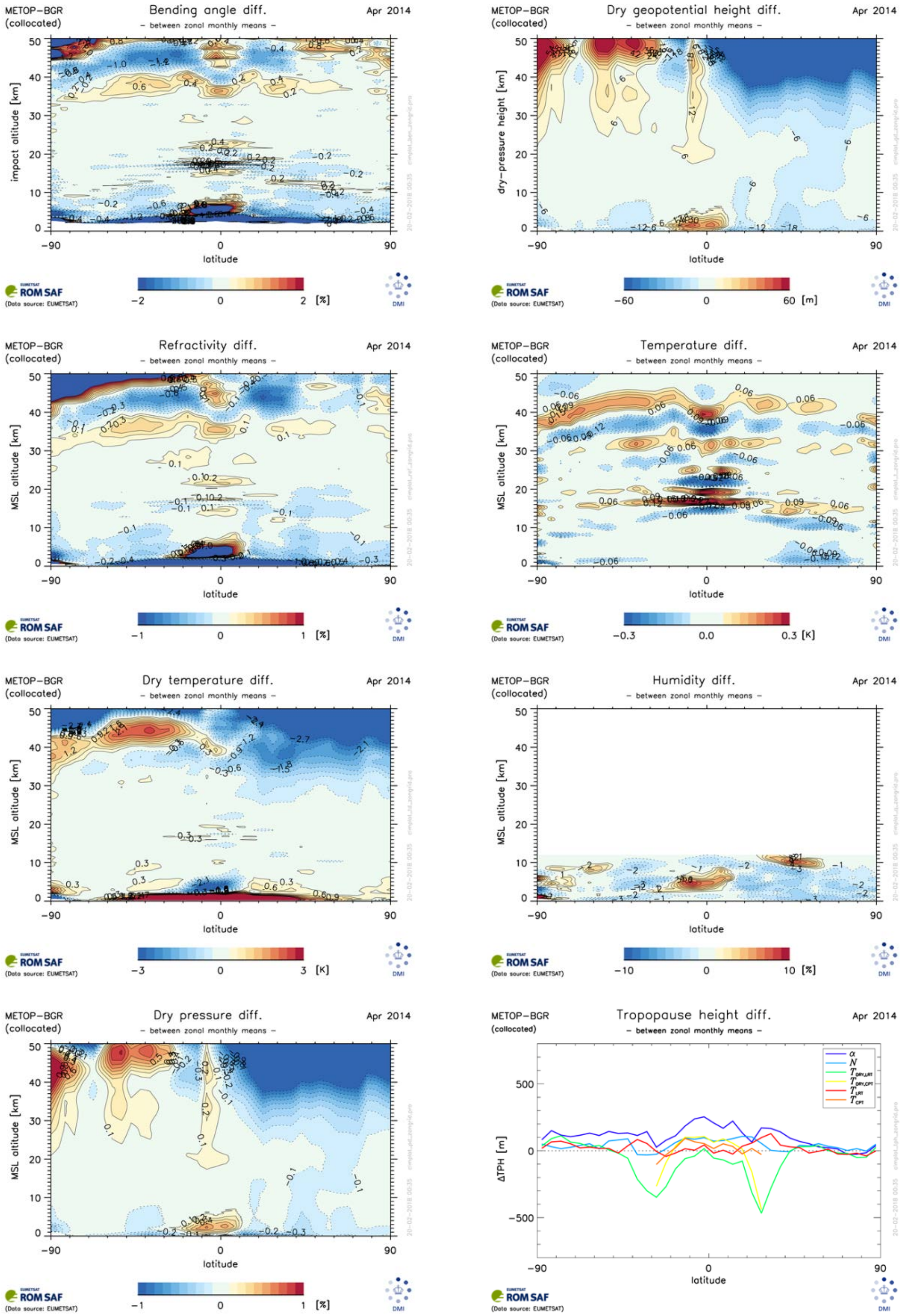


__Figure 7:__ Differences between observed monthly mean RO data and ERA-Interim for Metop data from April 2014. Computed as left panels minus right panels in Figure 6.

Metop – ERAI
Bending angle

4.2.2 Figures: O-B time-height plots

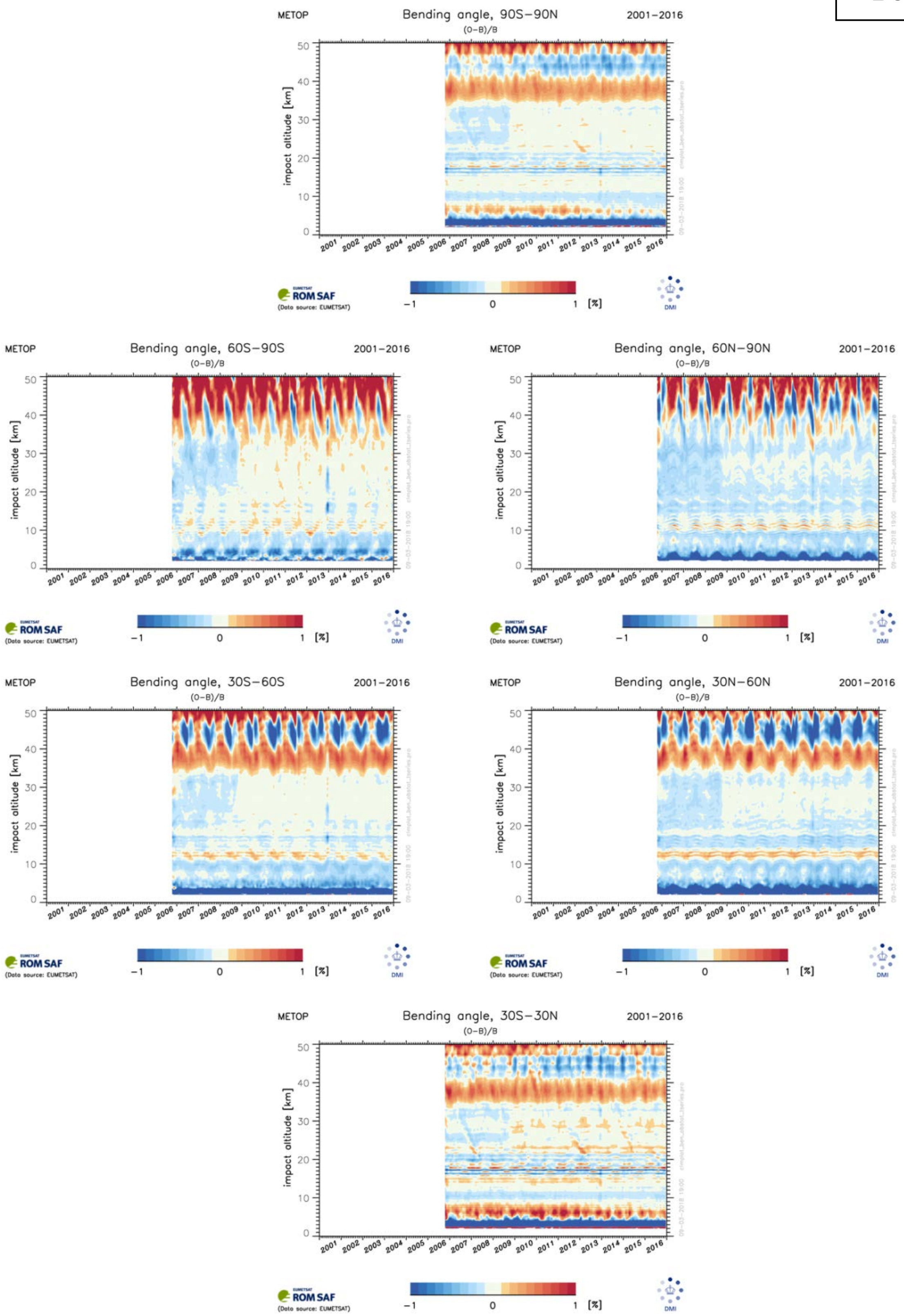


Figure 8a. *Observation minus co-located ERA-Interim bending angle for monthly mean Metop data, aggregated into 5 broad latitude bands plus global data.*

Metop – ERAI
Refractivity

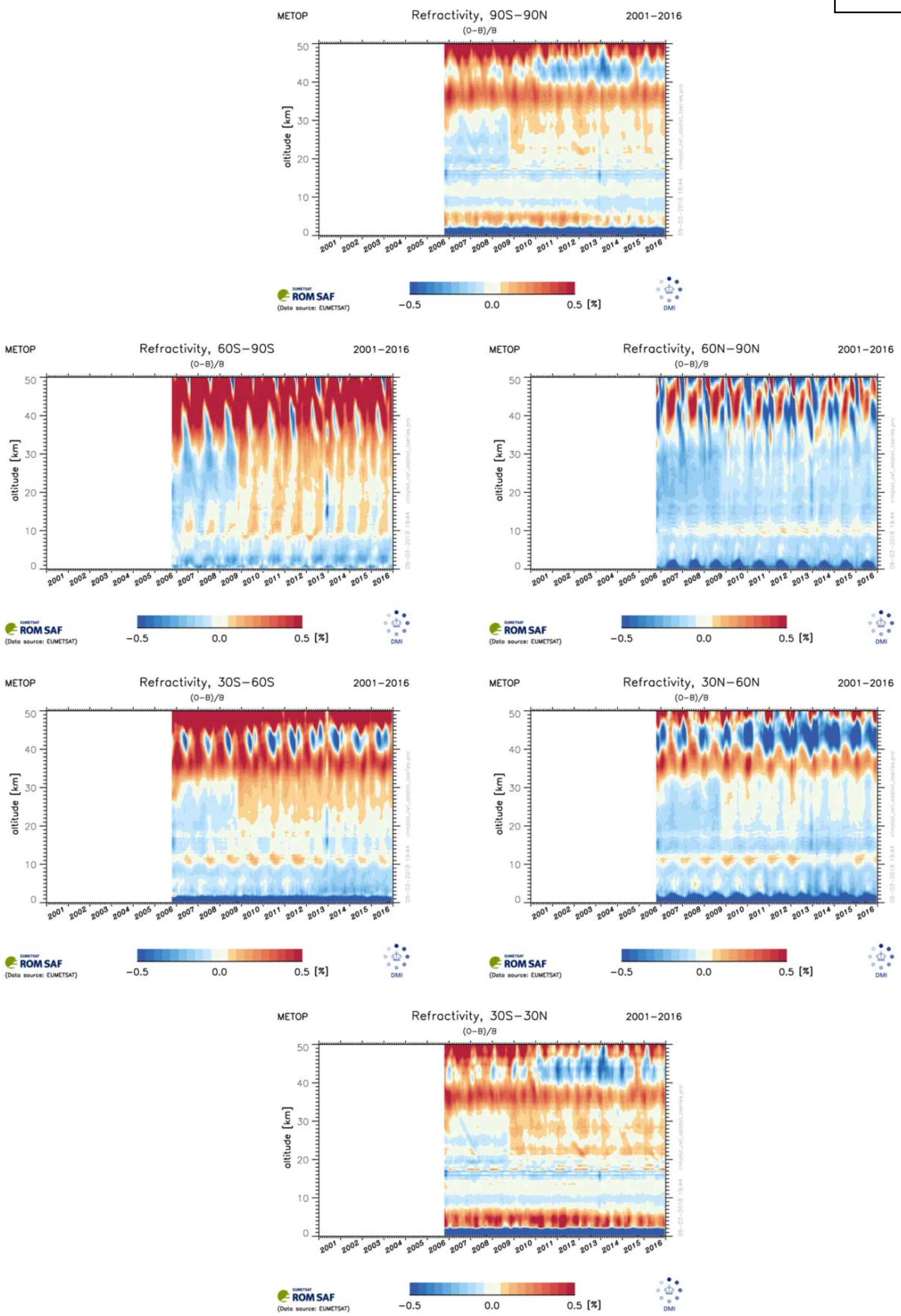


Figure 8b. *Observation minus co-located ERA-Interim refractivity for monthly mean Metop data, aggregated into 5 broad latitude bands plus global data.*

Metop – ERAI
Dry temperature

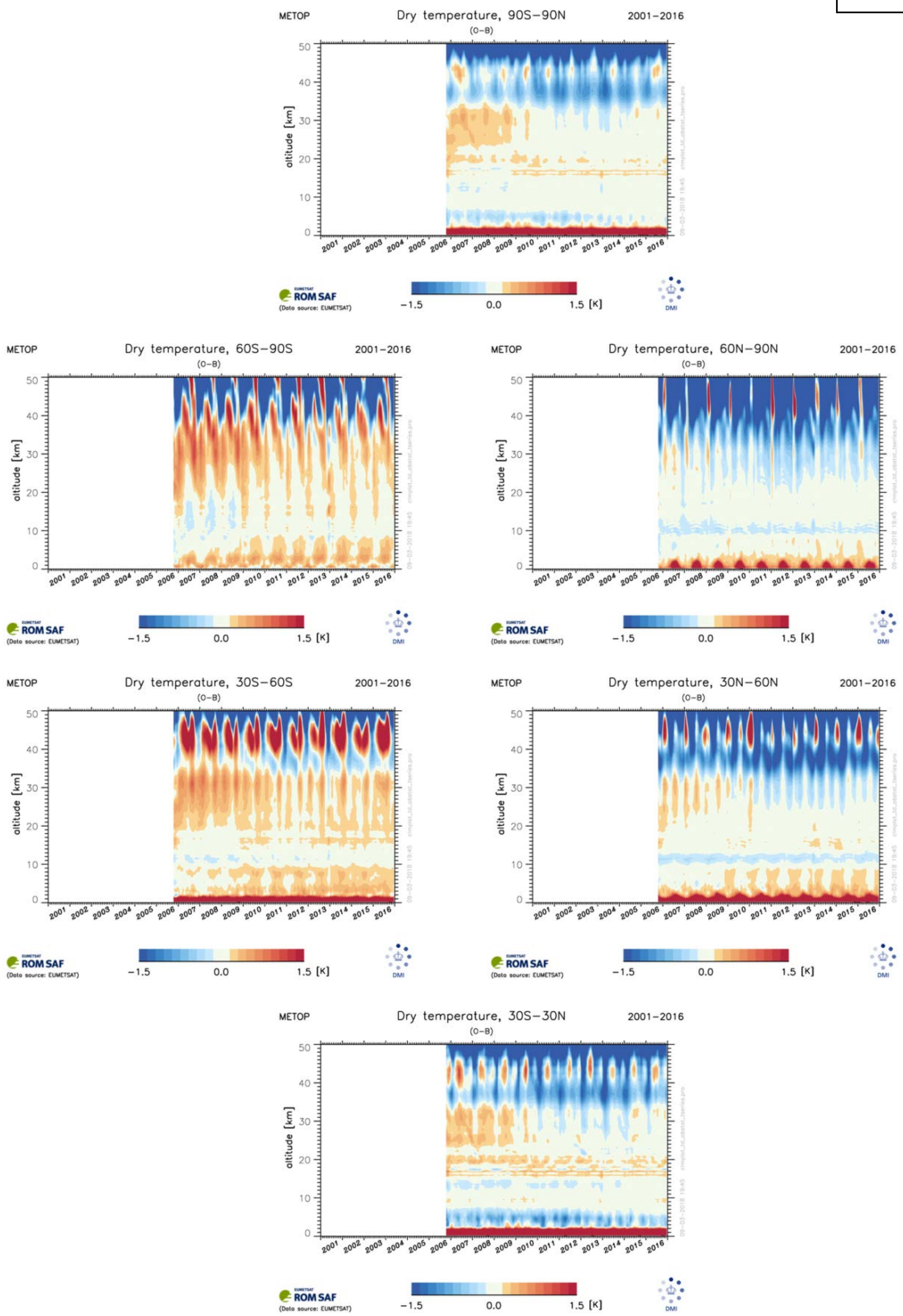


Figure 8c. *Observation minus co-located ERA-Interim dry temperature for monthly mean Metop data, aggregated into 5 broad latitude bands plus global data.*

Metop – ERAI Temperature

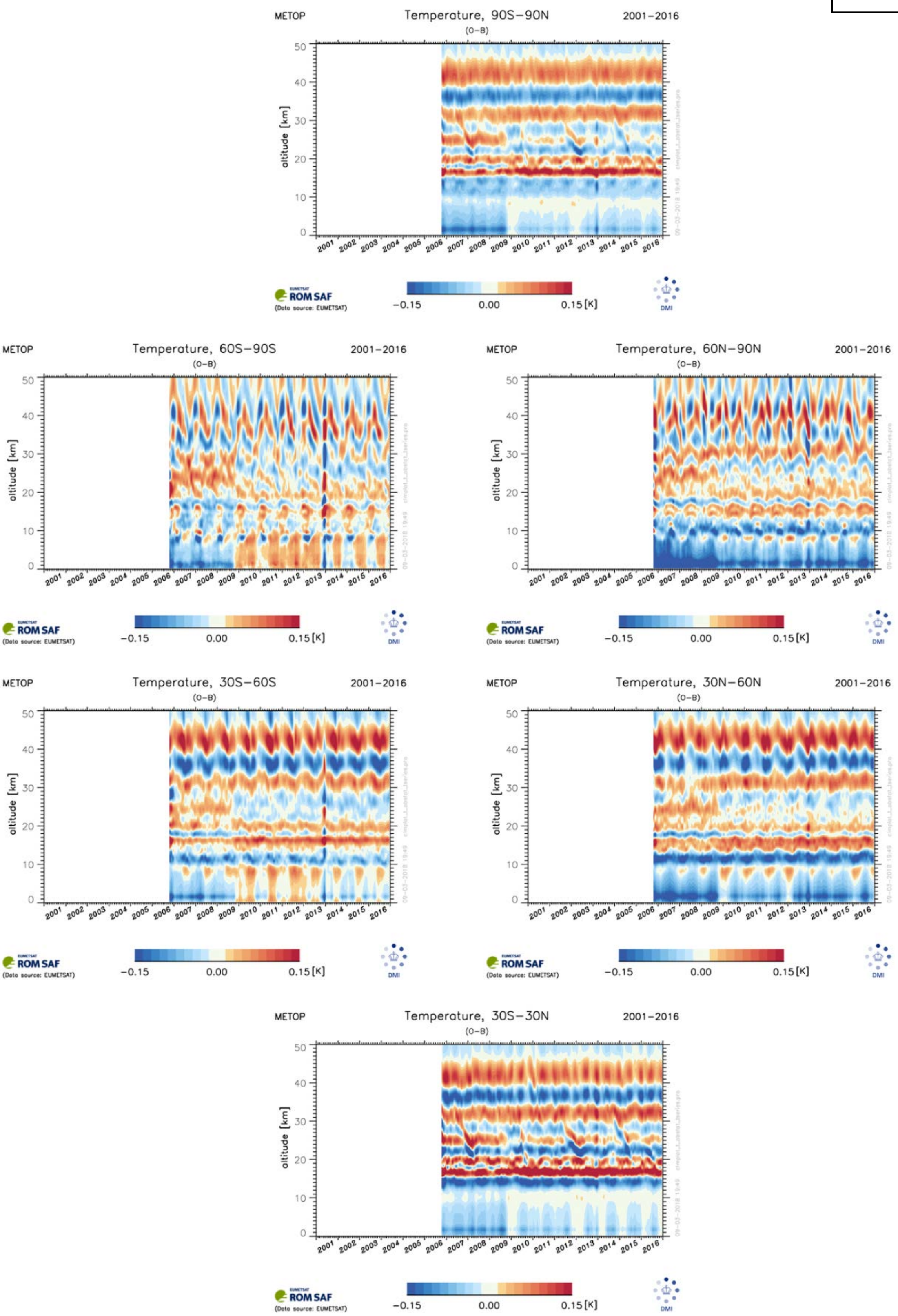


Figure 8d. *Observation minus co-located ERA-Interim temperature for monthly mean Metop data, aggregated into 5 broad latitude bands plus global data.*

Metop – ERAI Humidity

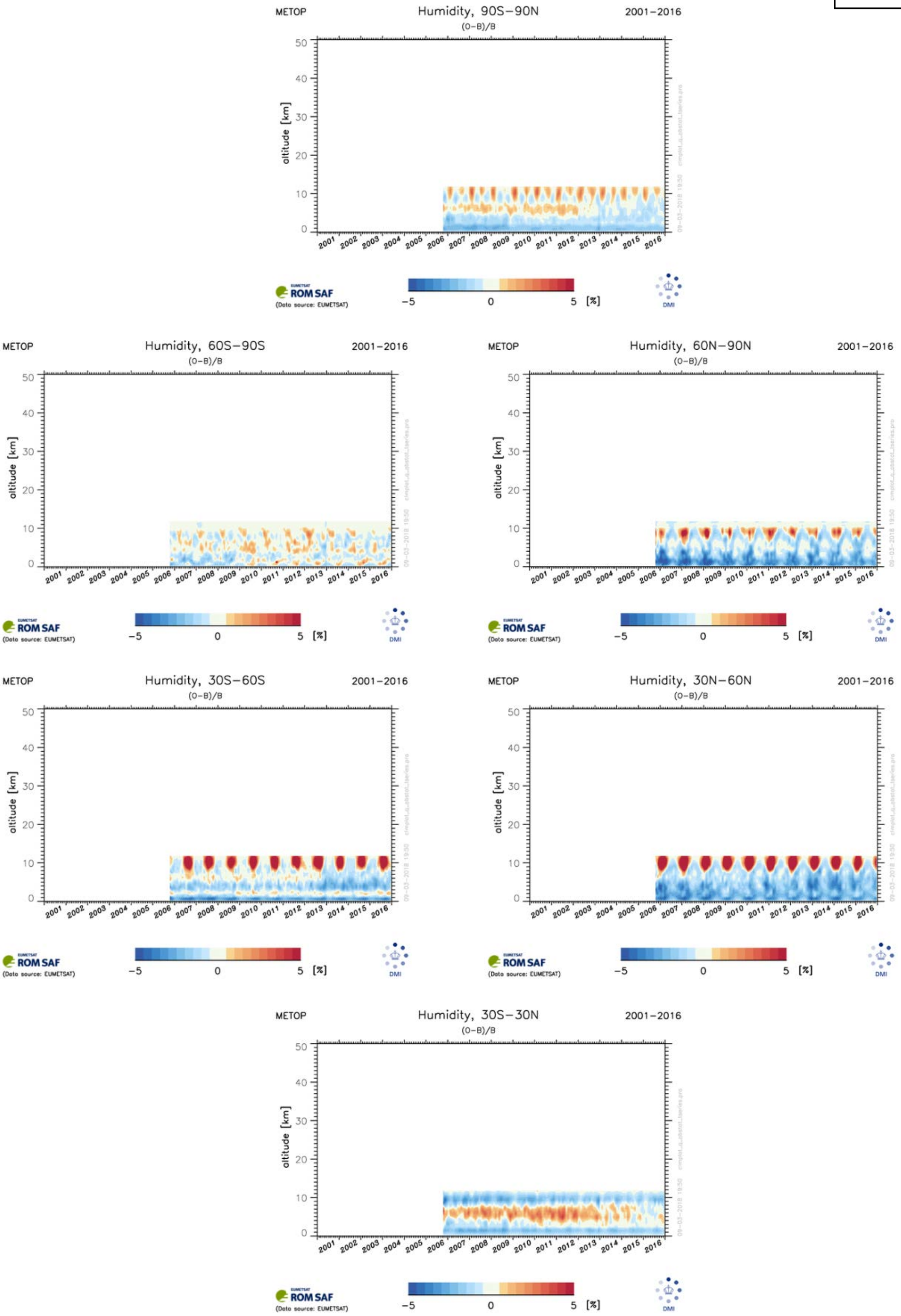


Figure 8e. *Observation minus co-located ERA-Interim humidity for monthly mean Metop data, aggregated into 5 broad latitude bands plus global data.*

COSMIC – ERAI
Bending angle

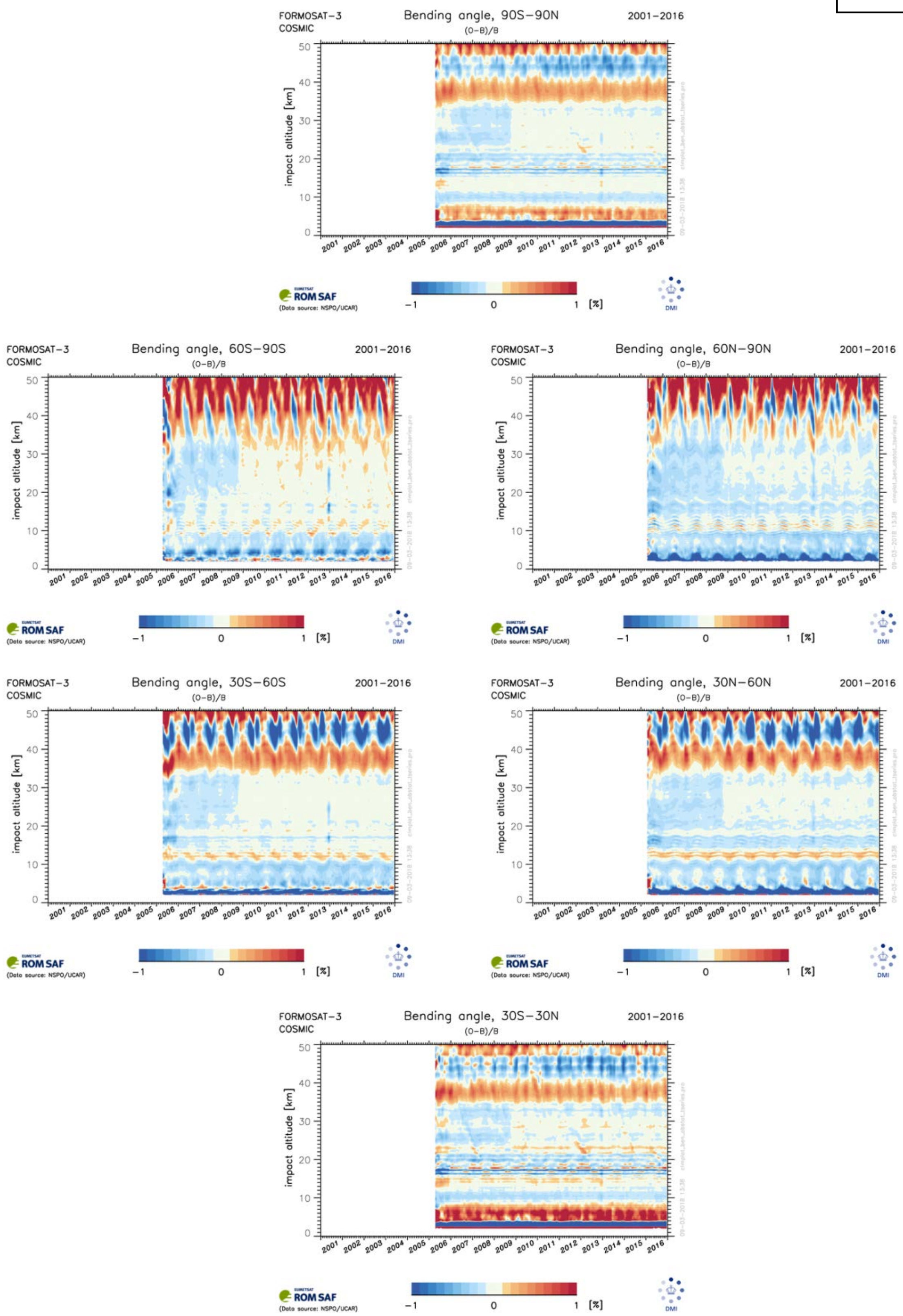


Figure 9a. *Observation minus co-located ERA-Interim bending angles for monthly mean COSMIC data, aggregated into 5 broad latitude bands plus global data.*

COSMIC – ERAI
Refractivity

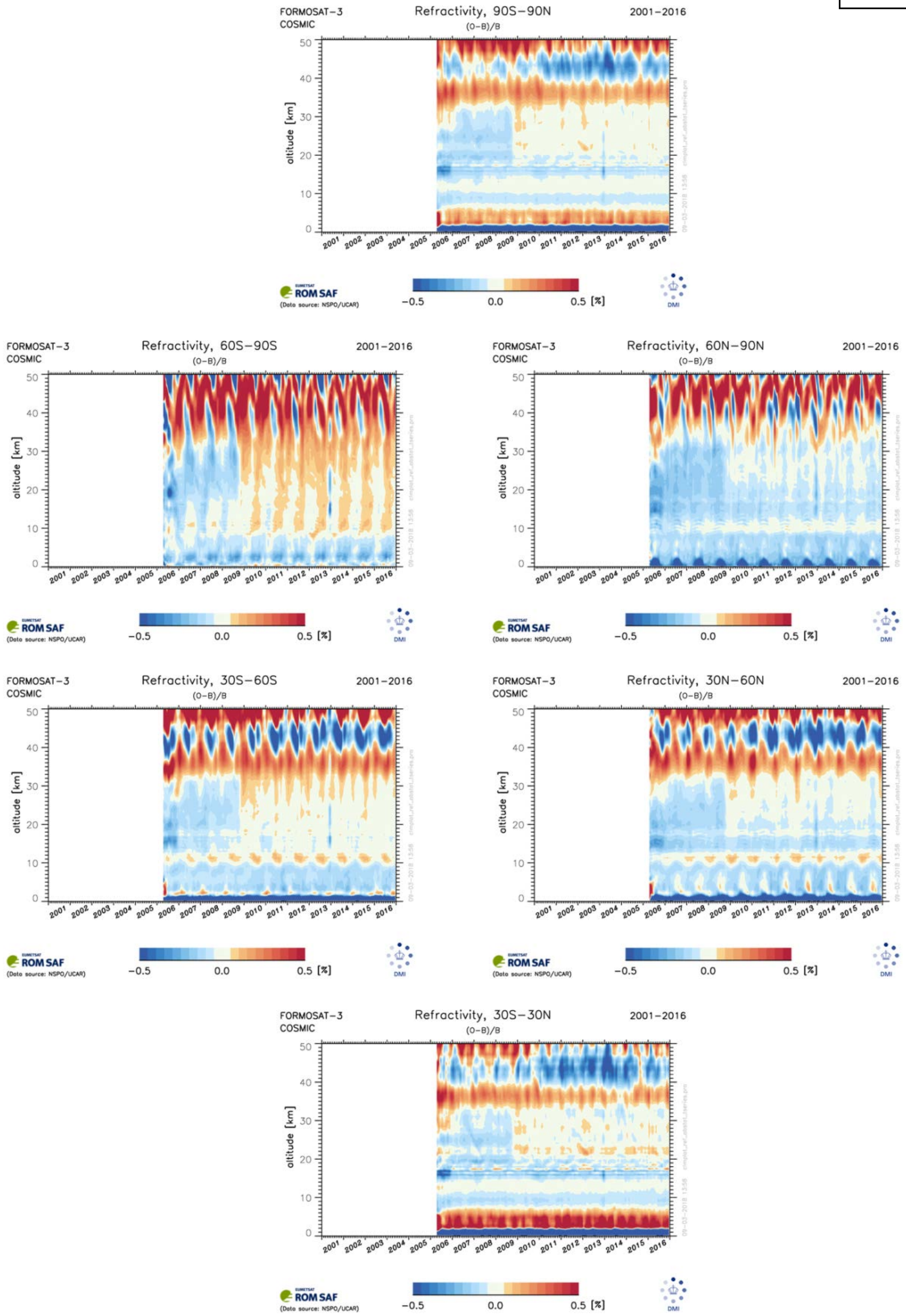


Figure 9b. *Observation minus co-located ERA-Interim refractivity for monthly mean COSMIC data, aggregated into 5 broad latitude bands plus global data.*

COSMIC – ERAI
Dry temperature

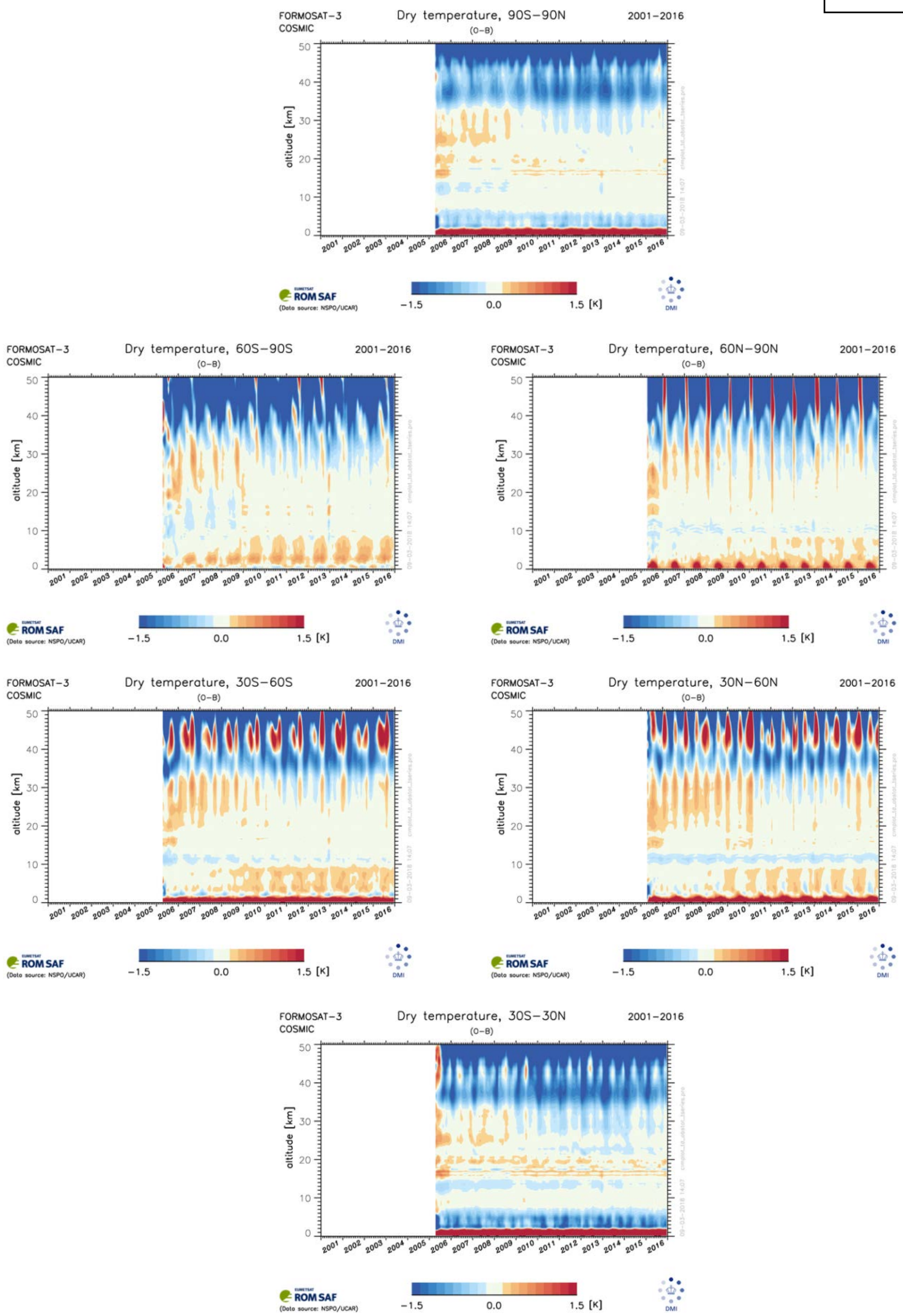


Figure 9c. *Observation minus co-located ERA-Interim dry temperature for monthly mean COSMIC data, aggregated into 5 broad latitude bands plus global data.*

COSMIC – ERAI
Temperature

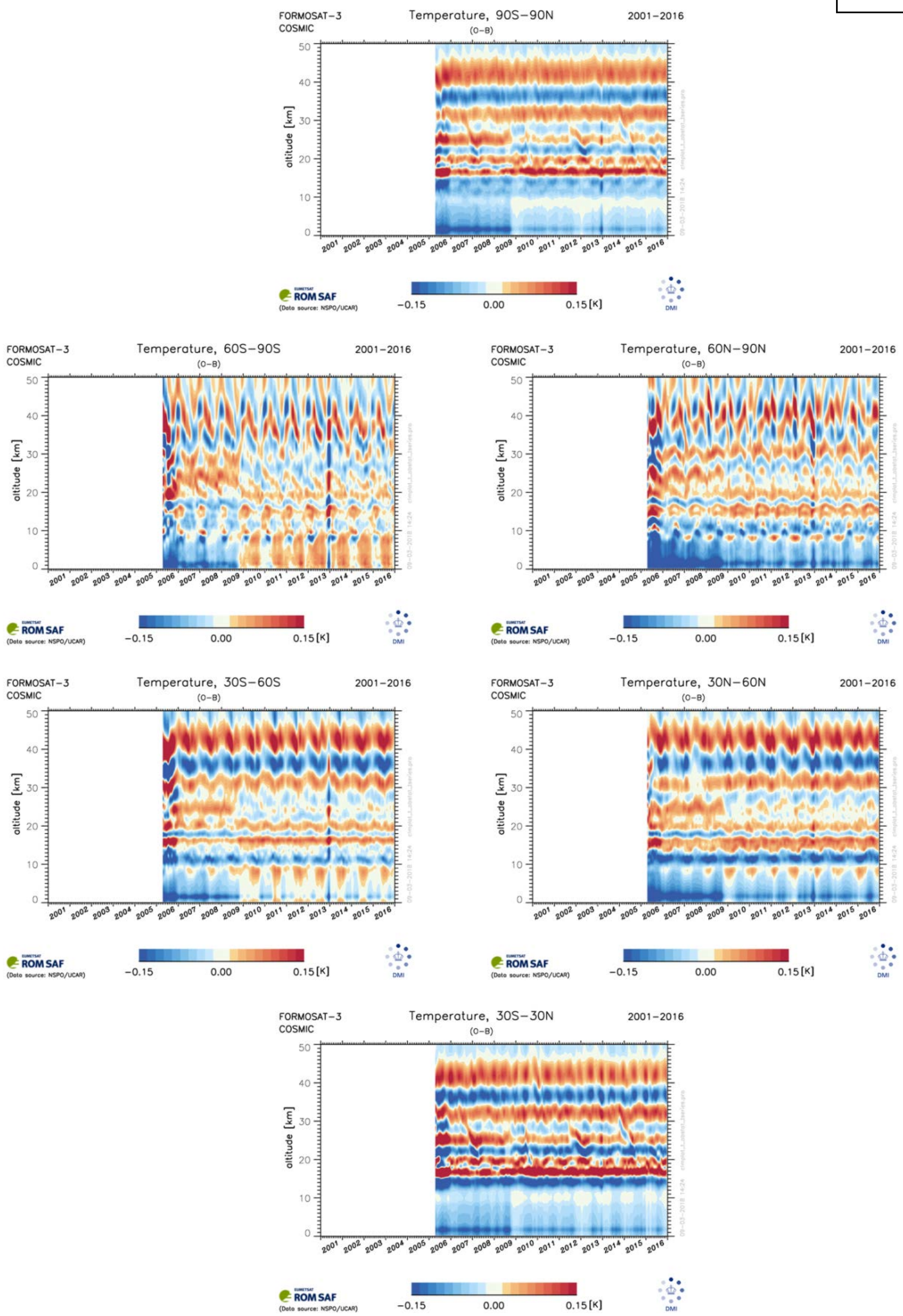


Figure 9d. *Observation minus co-located ERA-Interim temperature for monthly mean COSMIC data, aggregated into 5 broad latitude bands plus global data.*

COSMIC – ERAI
Humidity

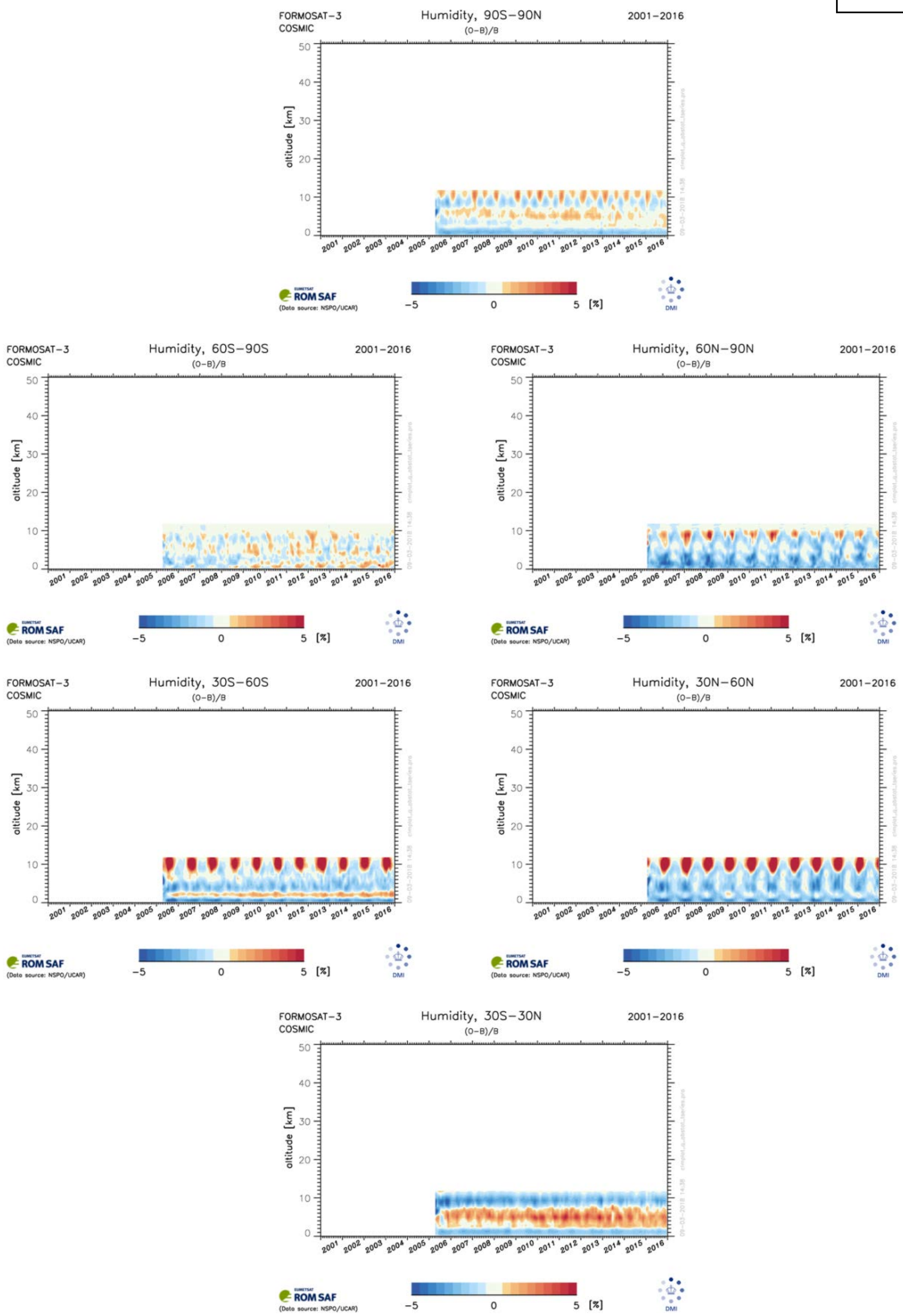


Figure 9e. *Observation minus co-located ERA-Interim humidity for monthly mean COSMIC data, aggregated into 5 broad latitude bands plus global data.*

RO – ERAI
Bending angle

4.2.3 Figures: O-B time series

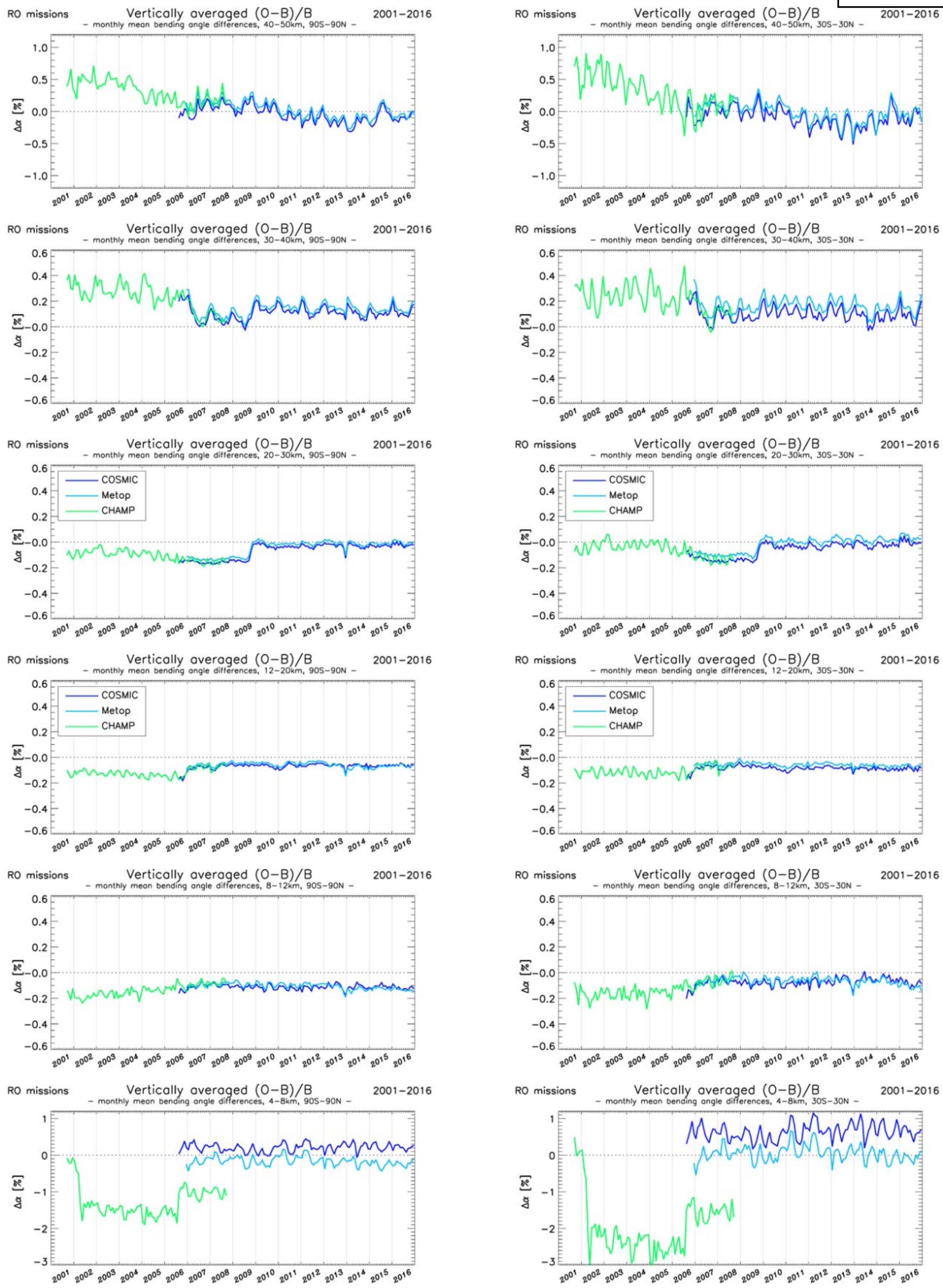


Figure 10a. *Observation minus co-located ERA-Interim bending angle for monthly mean CHAMP (green lines), COSMIC (dark blue lines), and Metop (light blue lines) data. Global averages in the panels to the left, low-latitude averages to the right. Note that the y-axis has been expanded for the lowest and highest layers.*

RO – ERAI
Bending angle

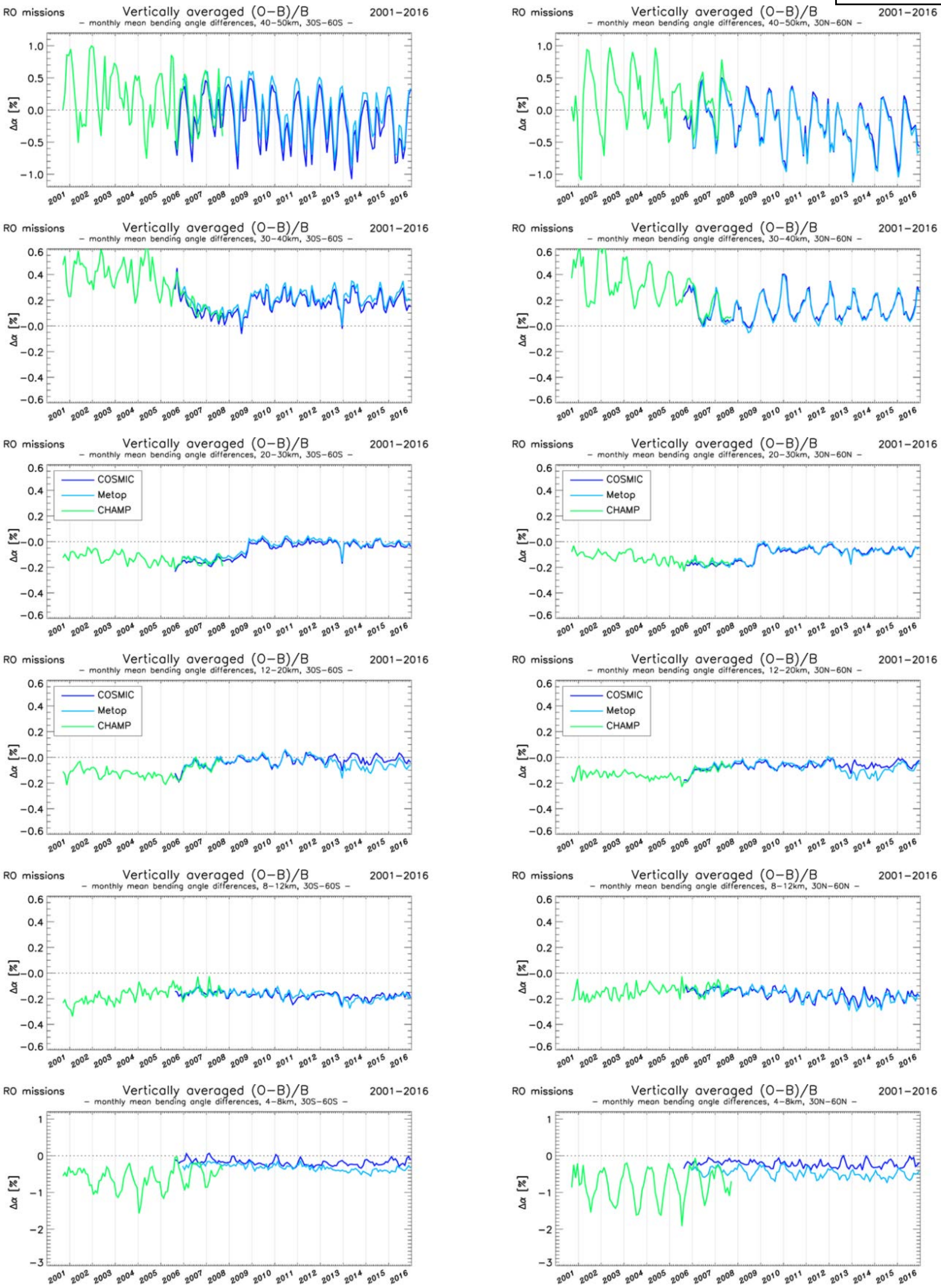


Figure 10b. *Observation minus co-located ERA-Interim bending angle for monthly mean CHAMP (green lines), COSMIC (dark blue lines), and Metop (light blue lines) data. Southern (northern) mid-latitude averages in the panels to the left (right). Note that the y-axis has been expanded for the lowest and highest layers.*

RO – ERAI
Bending angle

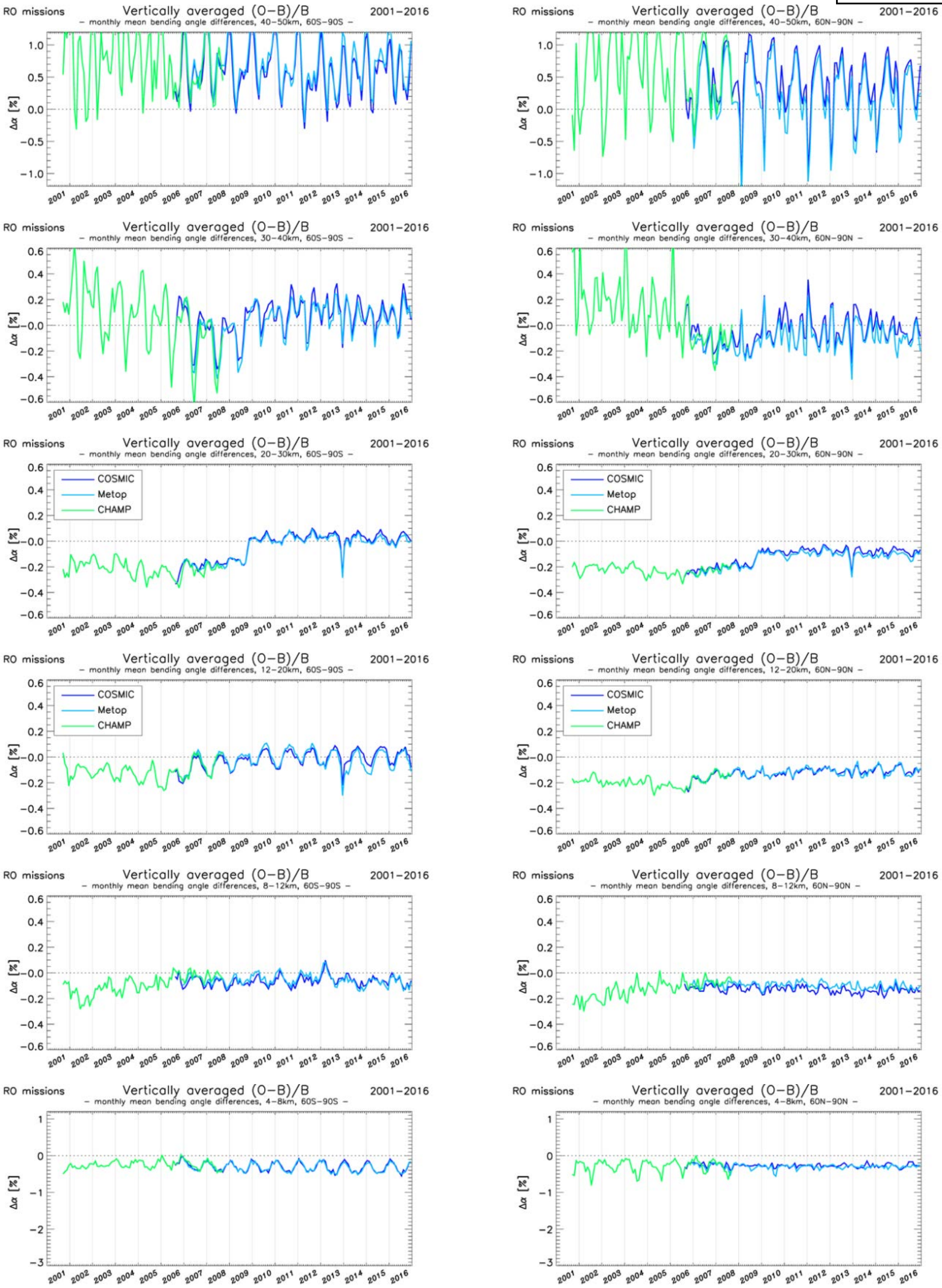


Figure 10c. *Observation minus co-located ERA-Interim bending angle for monthly mean CHAMP (green lines), COSMIC (dark blue lines), and Metop (light blue lines) data. Southern (northern) high-latitude averages in the left (right) hand panels. Note that the y-axis has been expanded for the lowest and highest layers.*

RO – ERAI Refractivity

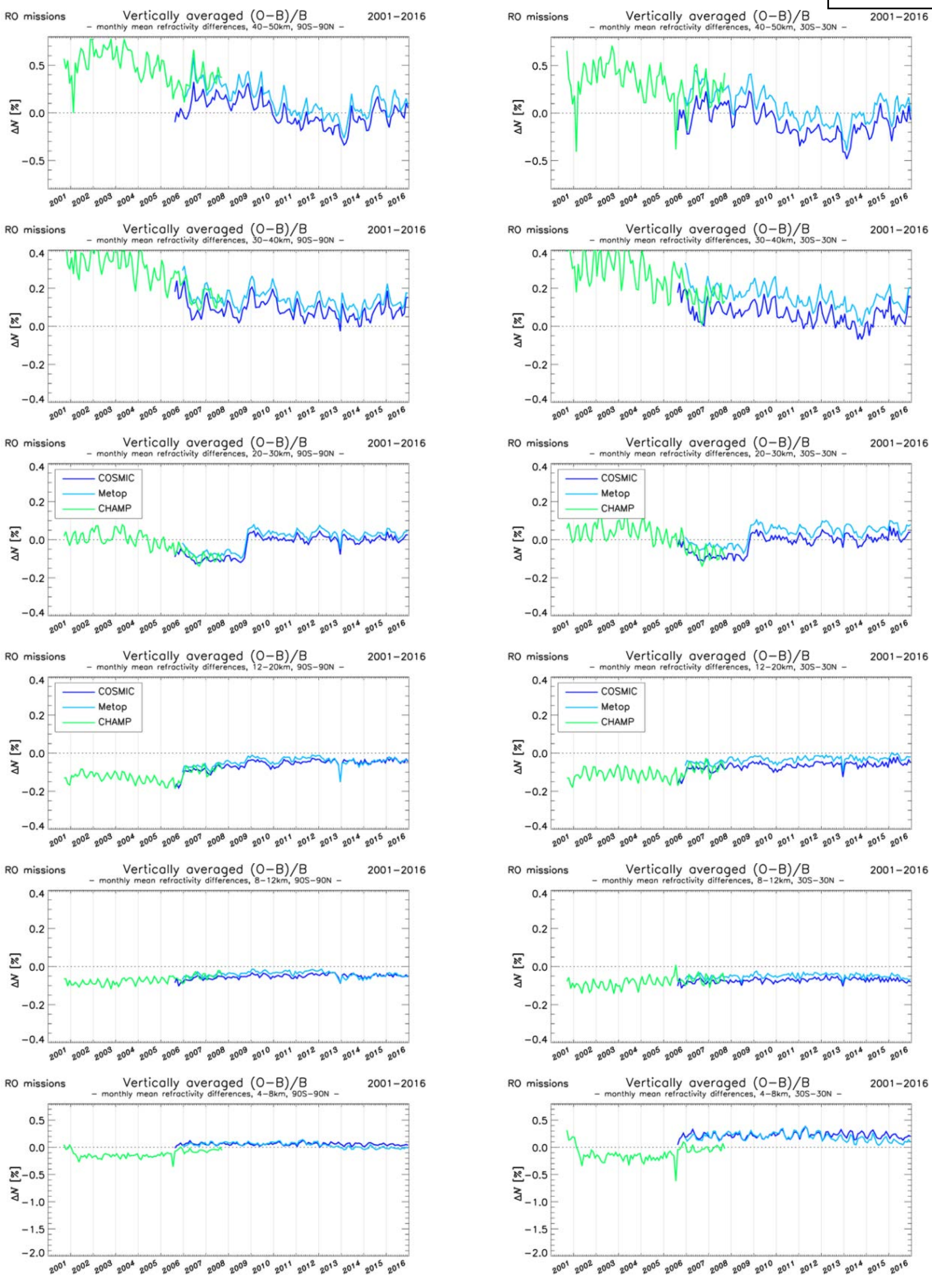


Figure 10d. *Observation minus co-located ERA-Interim refractivity for monthly mean CHAMP (green lines), COSMIC (dark blue lines), and Metop (light blue lines) data. Global averages in the panels to the left, low-latitude averages to the right. Note that the y-axis has been expanded for the lowest and the highest layers.*

RO – ERAI
Refractivity

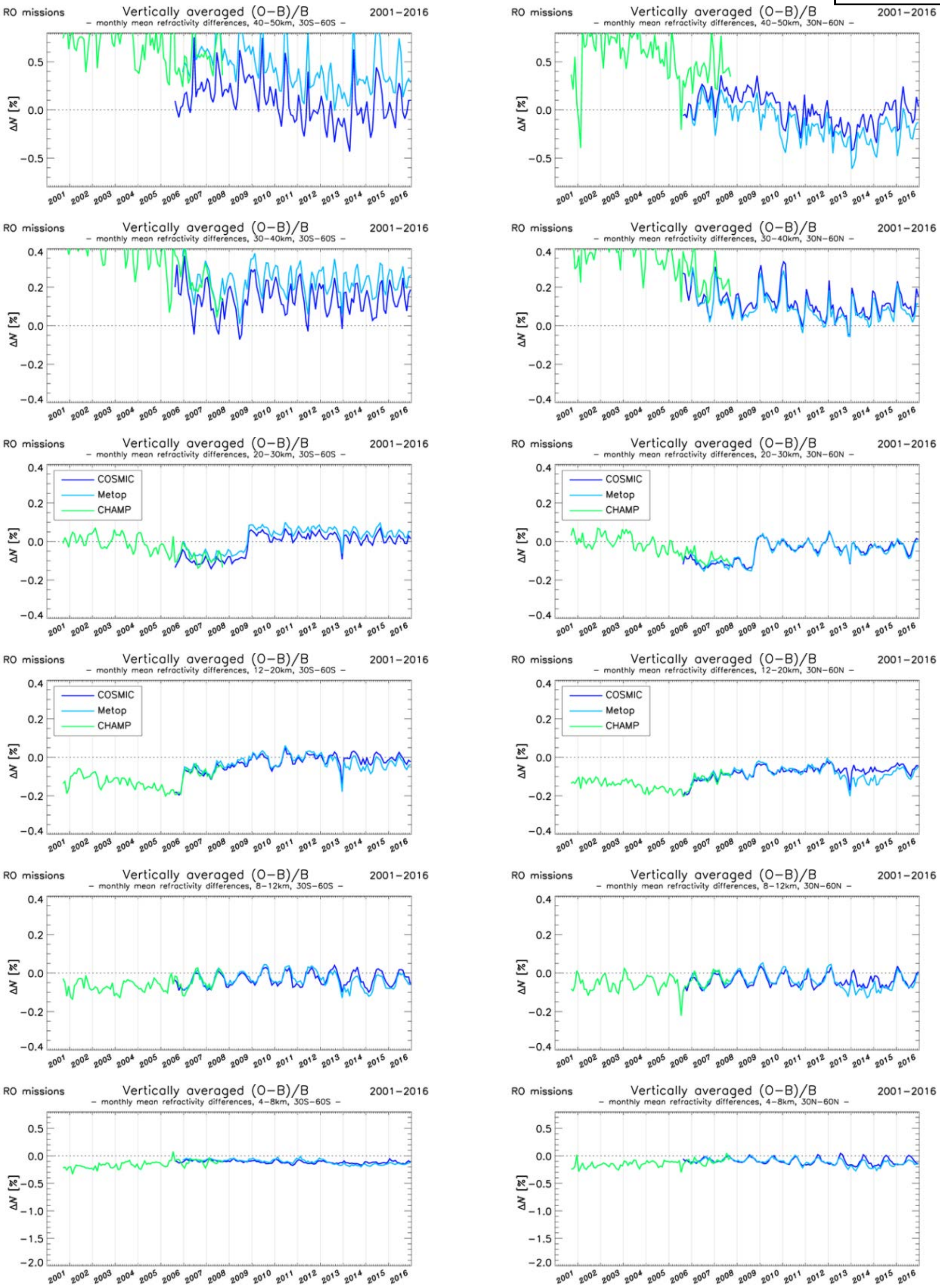


Figure 10e. *Observation minus co-located ERA-Interim refractivity for monthly mean CHAMP (green lines), COSMIC (dark blue lines), and Metop (light blue lines) data. Southern (northern) mid-latitude averages in the panels to the left (right). Note that the y-axis has been expanded for the lowest and the highest layers.*

RO – ERAI Refractivity

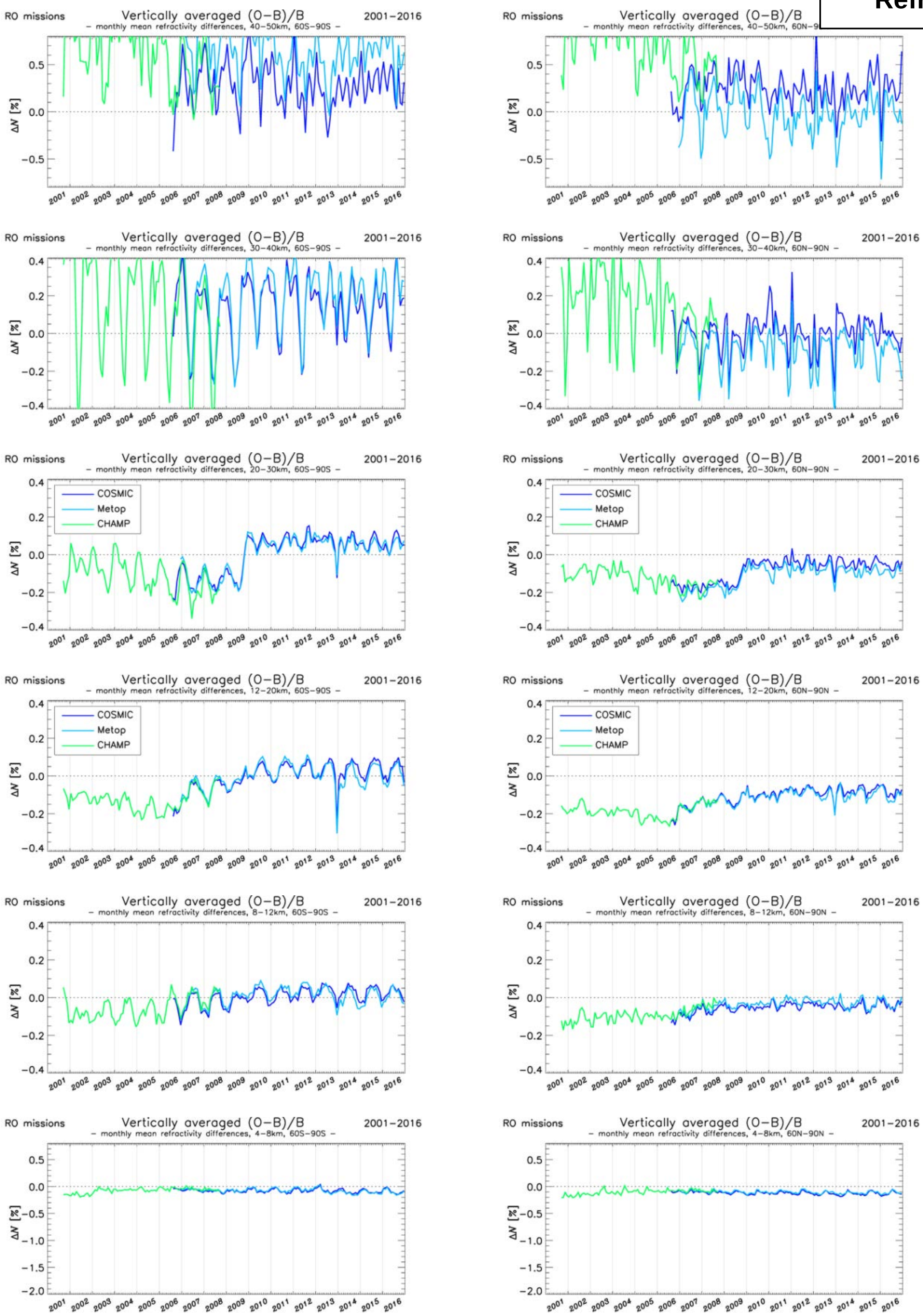


Figure 10f. *Observation minus co-located ERA-Interim refractivity for monthly mean CHAMP, COSMIC, and Metop data. Southern (northern) high-latitude averages in the left (right) hand panels. Note that the y-axis has been expanded for the lowest and highest layers.*

RO – ERAI
Dry temperature

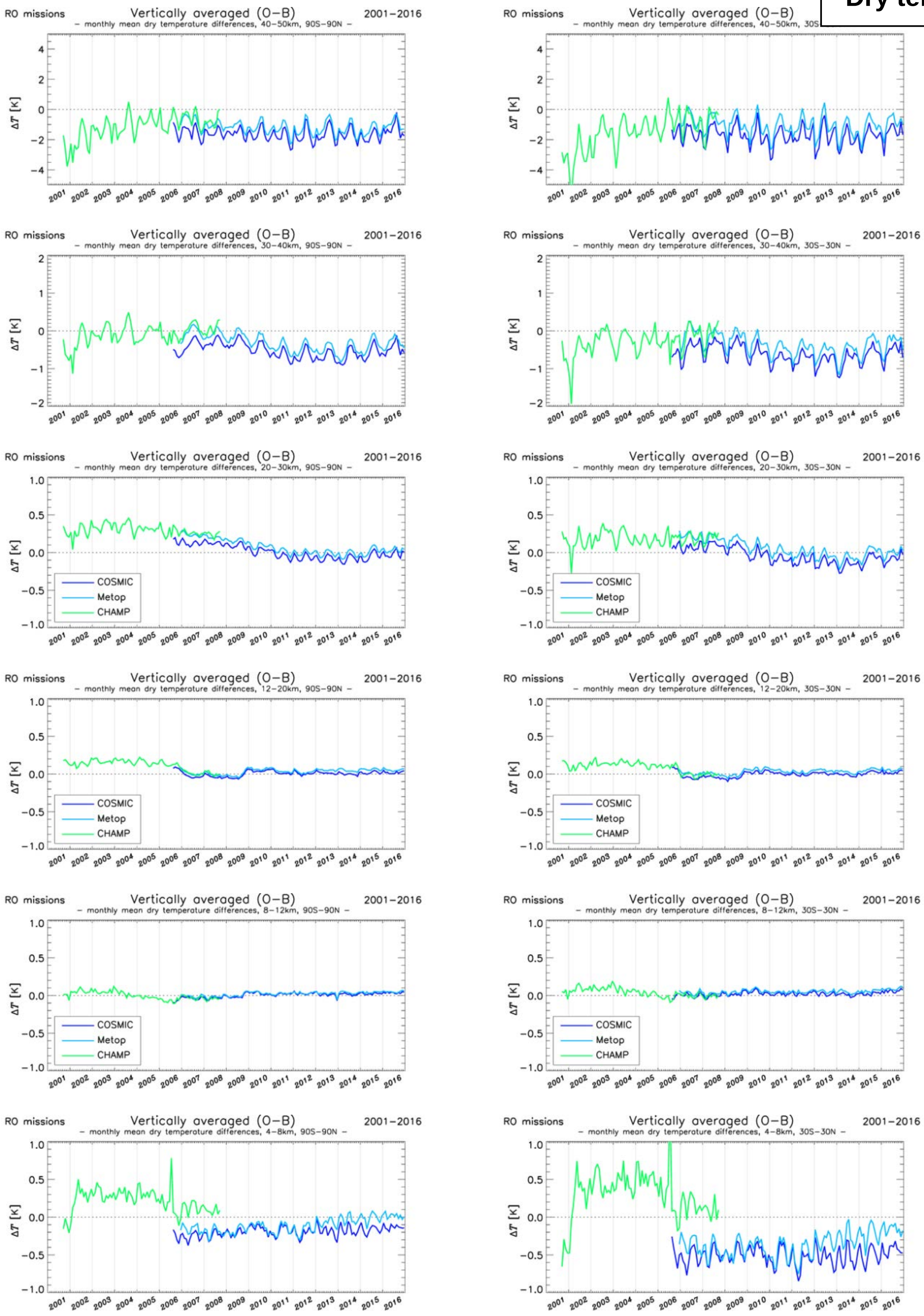


Figure 10g. *Observation minus co-located ERA-Interim dry temperature for monthly mean CHAMP, COSMIC, and Metop data. Global averages in the panels to the left, low-latitude averages to the right. Note that the y-axis has been expanded for the lowest and the highest layers.*

RO – ERAI
Dry temperature

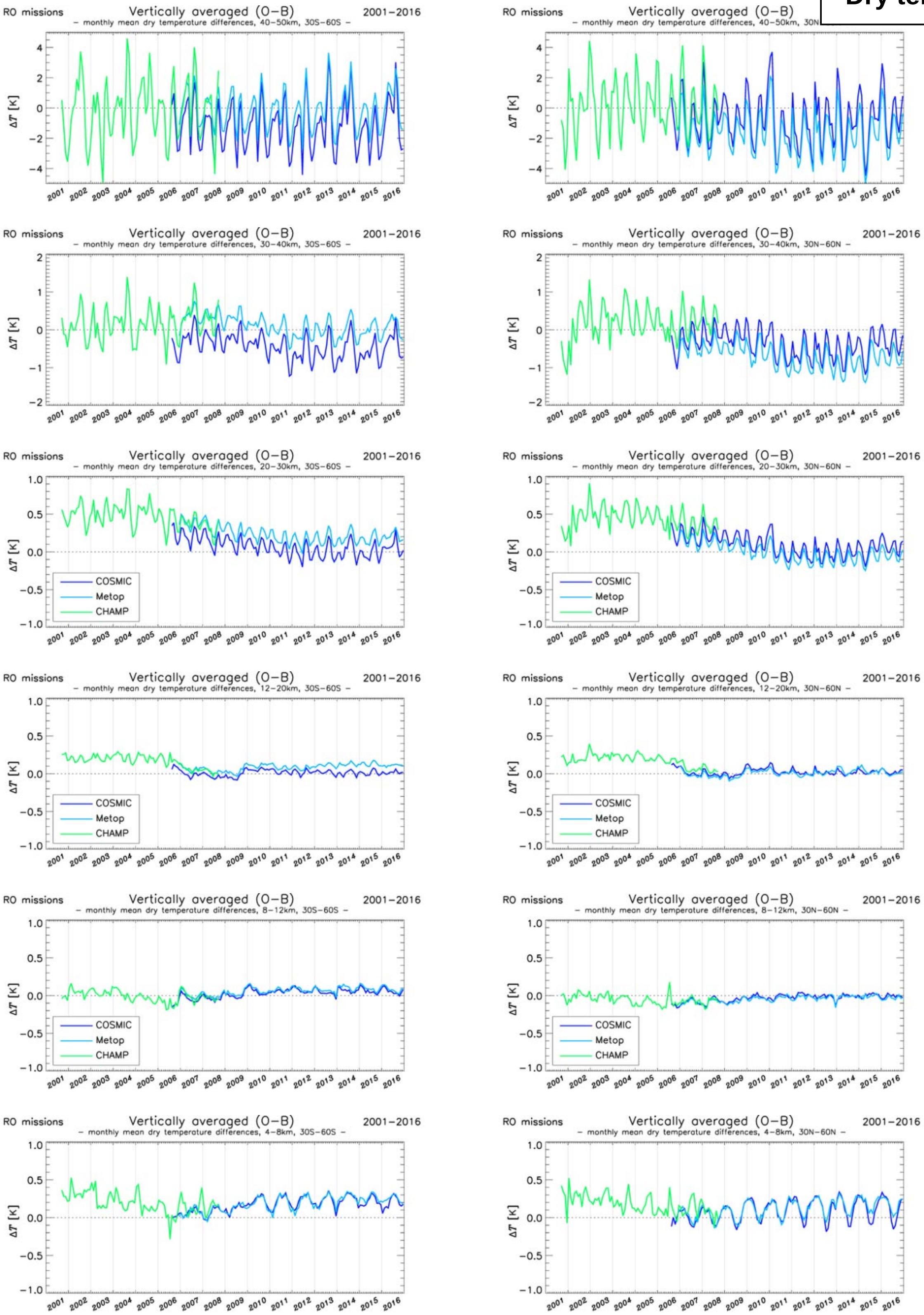


Figure 10h. *Observation minus co-located ERA-Interim dry temperature for monthly mean CHAMP, COSMIC, and Metop data. Southern (northern) mid-latitude averages in the panels to the left (right). Note that the y-axis has been expanded for the lowest and the highest layers.*

RO – ERAI
Dry temperature

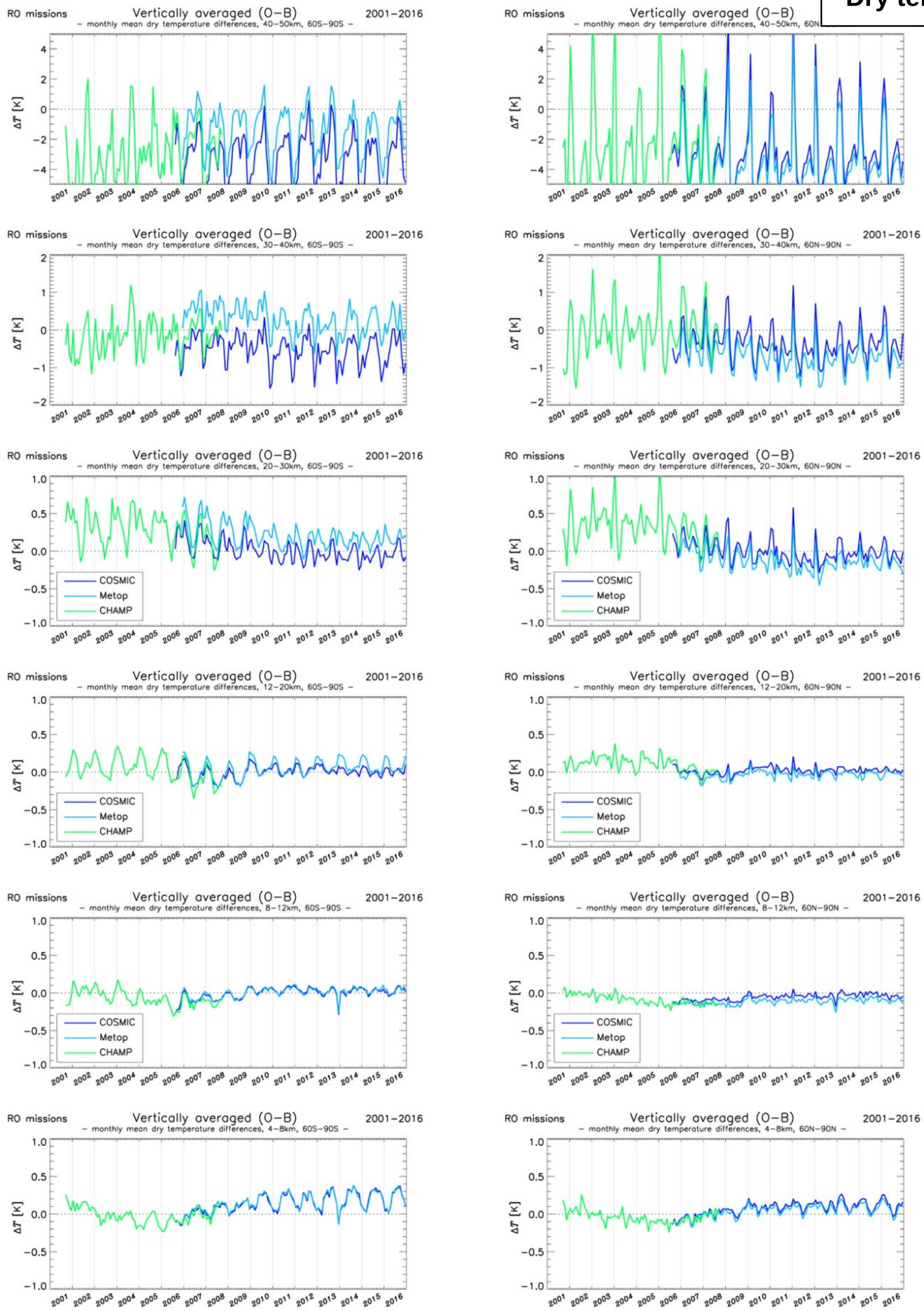


Figure 10i. *Observation minus co-located ERA-Interim dry temperature for monthly mean CHAMP, COSMIC, and Metop data. Southern (northern) high-latitude averages in the left (right) hand panels. Note that the y-axis has been expanded for the lowest and highest layers.*

Metop – ERAI
all variables

4.2.4 Figures: Spatial distribution of O-B differences

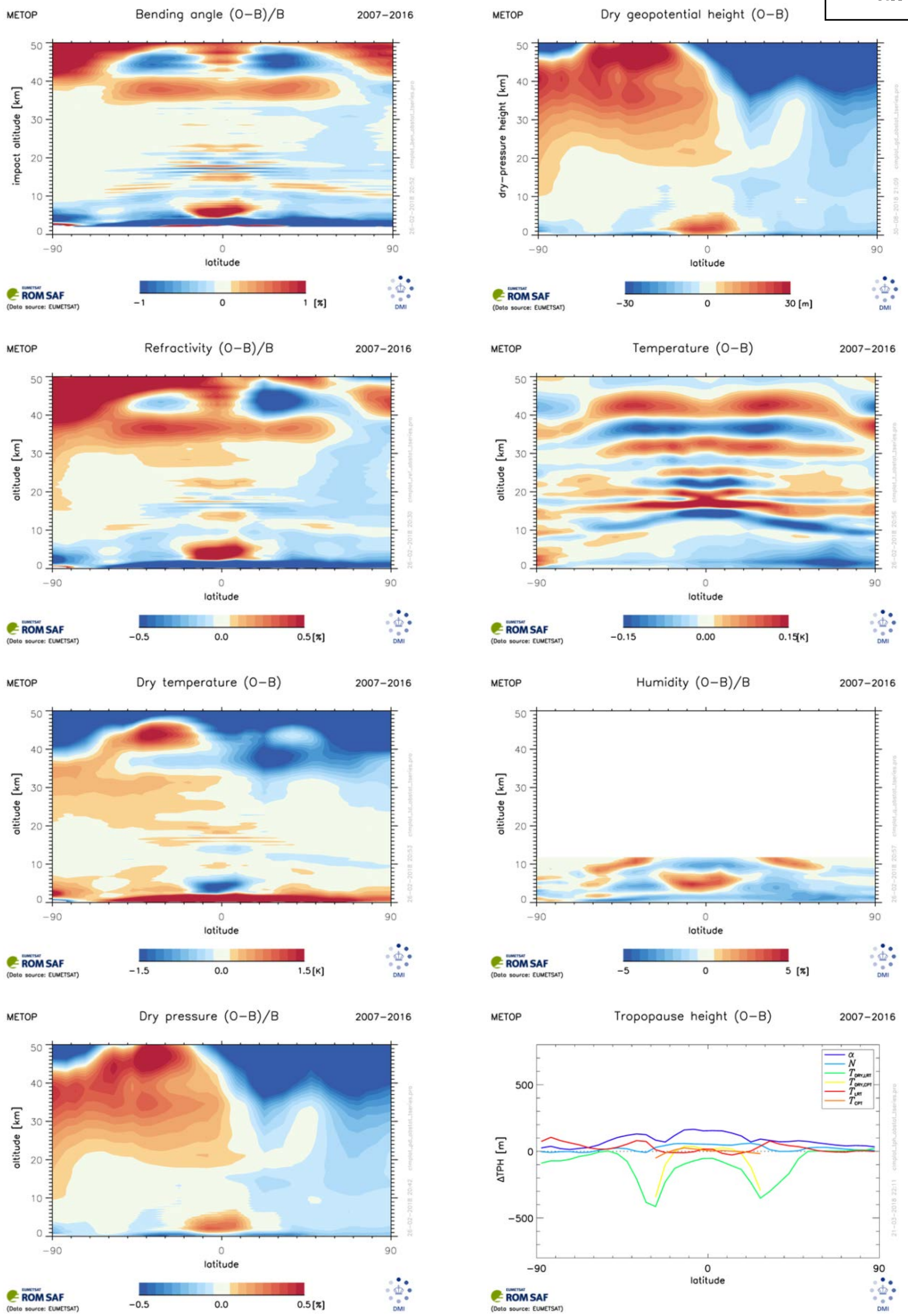


Figure 11a. *Long-term averages of observation minus co-located ERA-Interim differences for monthly mean Metop data, computed over the time period Jan 2007 to Dec 2016.*

COSMIC – ERAI all variables

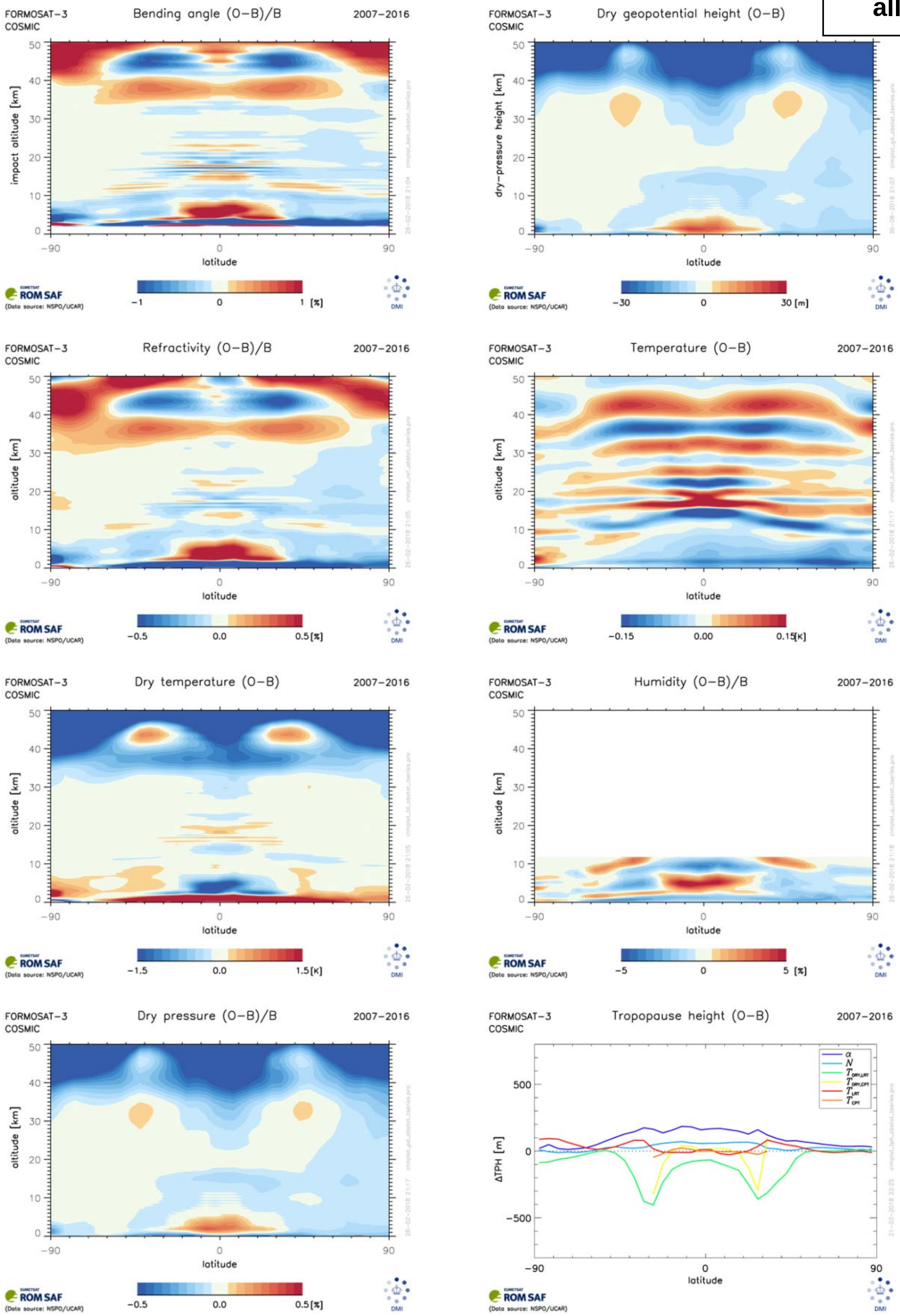


Figure 11b. *Long-term averages of observation minus co-located ERA-Interim differences for monthly mean COSMIC data, computed over the time period Jan 2007 to Dec 2016.*

CHAMP – ERAI all variables

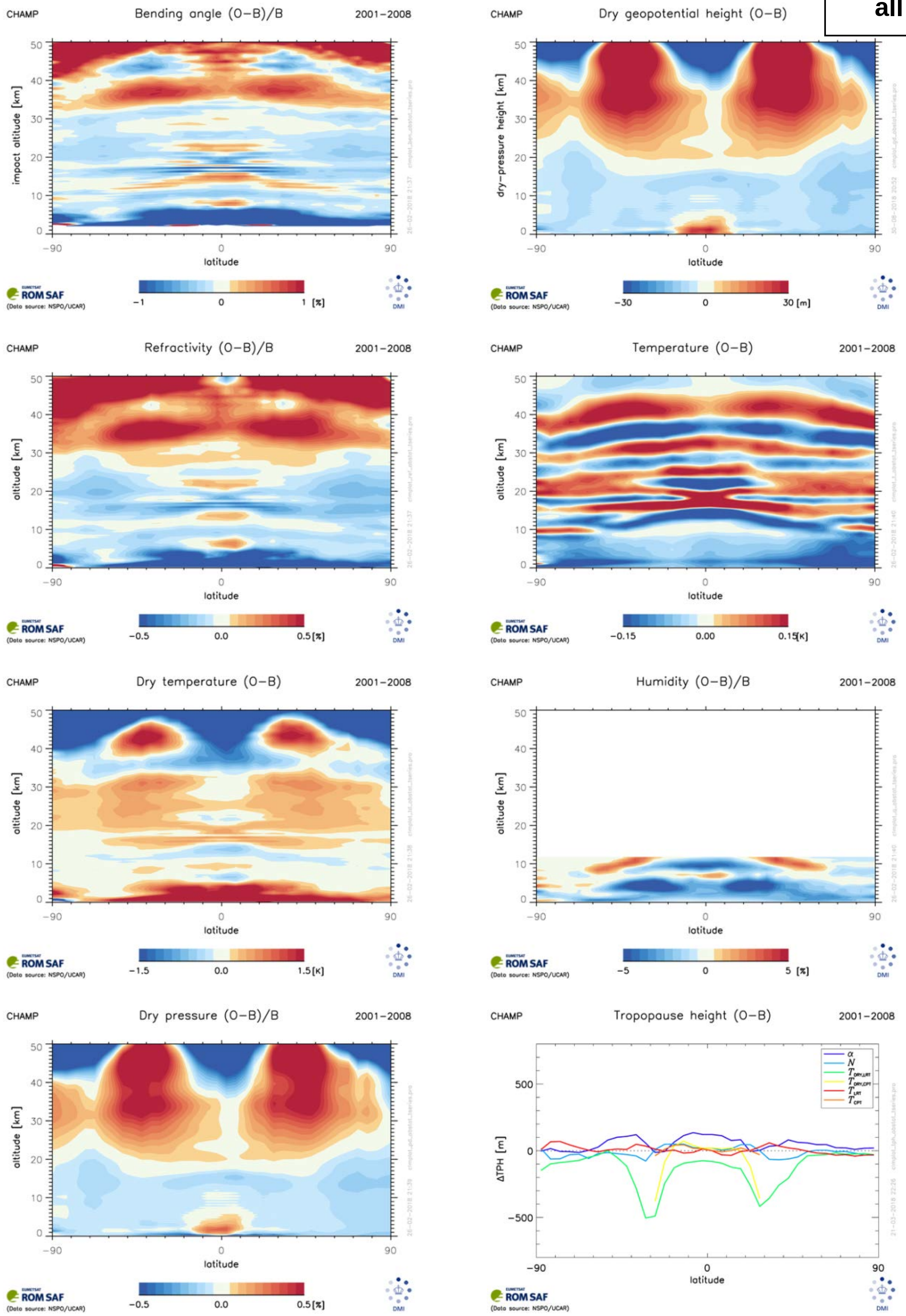


Figure 11c. *Long-term averages of observation minus co-located ERA-Interim differences for monthly mean CHAMP data, computed over the time period Oct 2001 to Sep 2008.*

GRACE – ERAI
all variables

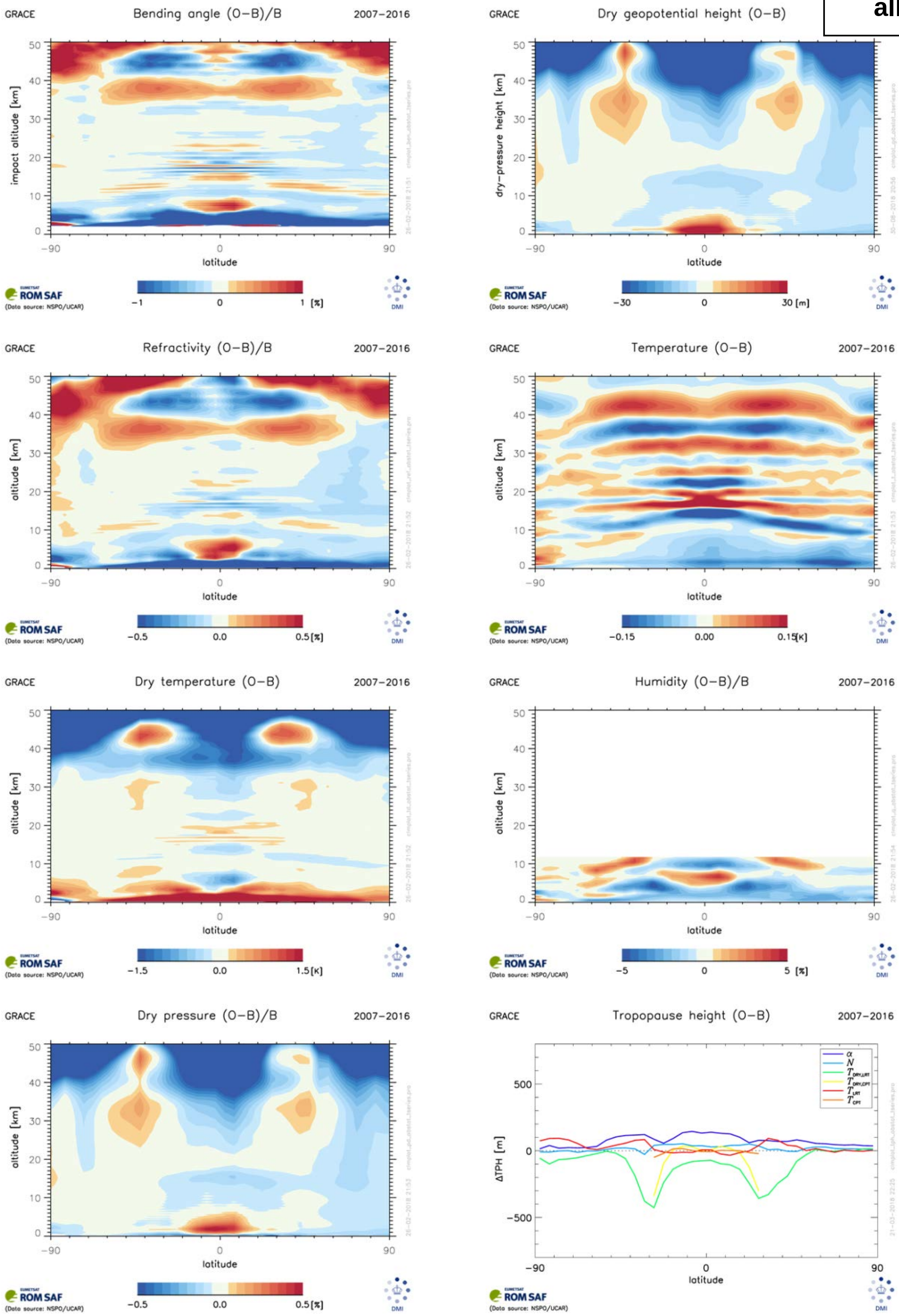


__Figure 11d.__ Long-term averages of observation minus co-located ERA-Interim differences for monthly mean GRACE data, computed over the time period Mar 2007 to Dec 2016.

4.2.5 Discussion and conclusions

Comparison of the left and right panels in Figure 6, Section 4.2.1, (Metop and co-located ERA-Interim from April 2014) shows the overall consistency between observation and reanalysis, while Figure 7 shows the differences between the observations and ERA-Interim. The O-B (observation minus background) differences in Figure 7 exhibit a structured pattern that is largely persistent over time, but with a certain degree of seasonal variations. The patterns in bending angle and refractivity are also relatively consistent between RO missions, while the differences for the “dry” variables are somewhat larger, particularly at high altitudes (see Figure 11). The horizontal bias structures in bending angle and refractivity at low- and mid-latitudes, as well as the characteristic bias structures in 1D-Var temperature and humidity, are partly related to biases in ERA-Interim and do not properly reflect biases in RO measurements.

The dry-temperature lapse rate tropopause shows negative biases of a few hundred meters around 30 degrees latitude. These may be related to the presence of multiple tropopauses, where the RO observations primarily find the lowest tropopause, while ERA-Interim is unable to resolve them. The bottom right panels of Figures 11a-d show that this pattern is largely consistent between the missions. It should be noted that the tropopause heights based on other definitions (bending angle tropopause, refractivity tropopause) do not show the same bias structure.

The time-height plots in Section 4.2.2 (Metop-ERAI and COSMIC-ERAI for bending angle, refractivity, dry temperature, temperature, and humidity) reveal a range of bias features. To understand the relevance of these biases it is important to point out that RO was assimilated in ERA-Interim from December 2006. The most pronounced systematic O-B differences in the figures are:

- All missions exhibit a bias change around November 2009. This appears to be caused by a bias in ERA-Interim due to a known bias in the operational COSMIC data that were assimilated in the reanalysis model.
- All missions exhibit a bias around November 2013. This is caused by a known error in ERA-Interim (by mistake, RO data were not assimilated in ERA-Interim).
- Seasonally varying biases at mid- and high latitudes above about 30-35 km.
- In the low-latitude bending angle, refractivity, and dry temperature plots, there are weak bias structures in the 20-30 km altitude range that have the same slope and are coincident in time with QBO structures as seen in the same variables. The most prominent is the one in 2007-2008.
- The humidity exhibits a persistent bias structure at the lowest latitudes, while the biases at mid-latitudes show a pronounced seasonal variation.

The time series plots in Section 4.2.3 provide some additional insights on the temporal characteristics:

- The bias change in November 2009 is evident in the 20 – 30 km layer, as is also the bias feature in November 2013.

- There is generally an excellent agreement between the RO missions above 8 kilometers and below 35–40 km. At high latitudes the RO missions agree all the way down to the surface. Even minor, small-scale O-B difference features are similarly described by the different RO missions.

Even though there is an overall consistency between the different RO missions, we can identify some discrepancies. These may be signs of underlying differences between the measurements, but they may also be due to differences in the spatial and temporal sampling characteristics of the RO missions. For example, misrepresentations of the diurnal cycle by ERA-Interim could lead to differences between the O-B time series of COSMIC and Metop [RD.17].

- Below 8 km, at low- and mid-latitudes, there are differences between the RO missions – in particular between CHAMP and the other missions, but also between COSMIC and Metop (Figures 10a,d,g).
- Below 8 km, the CHAMP O-B data record exhibits abrupt bias changes (first half of 2002 and mid-2006) that appear to be due to CHAMP data itself (Figure 10a). This is confirmed by the CHAMP anomaly plots in Section 4.3.4 (Figure 17a), which does not depend on ERA-Interim.
- There are differences between the RO missions above 35 – 40 km, which are large enough to affect the generation of a multi-mission data record.
- At low latitudes, above 12 km there is a persistent bending-angle and refractivity bias between the O-B time series for COSMIC and Metop (Figures 10a,d,g).
- At mid-latitudes, at altitudes 12 to 20 km, there is an increased difference between the COSMIC and Metop data after first half of 2013 (Figure 10b). This coincides in time with a Metop firmware update.

The figures in Section 4.2.4 show the spatial distribution of O-B differences for all four missions. We note the following:

- For the 1D-Var temperature and humidity, the spatial patterns for the 4 RO missions are very similar. The one mission that deviates the most is CHAMP, which exhibit temperature biases that are structurally similar, but with larger magnitudes, and in addition have low-latitude humidity biases considerably more negative than the other missions. It should be noted that CHAMP data are available for a different time period compared to other missions. The time period from September 2001 to mid-2006 is only covered by CHAMP.
- CHAMP refractivites, dry temperatures, and dry pressures have more positive biases than COSMIC and GRACE.
- COSMIC and GRACE bias structures are very similar for all geophysical variables.
- The bias structures for COSMIC, CHAMP, and GRACE are latitudinally symmetric, whereas Metop exhibit an O-B pattern that appears to have a north-south asymmetry.

We conclude that in general there is a very good agreement between the RO observational data and the ERA-Interim reanalyses. The different RO missions exhibit very similar O-B spatial and temporal characteristics – they are often nearly identical. At least some of the O-B features are due to known deficiencies in ERA-Interim.

Another conclusion is that there is a high degree of inter-mission consistency above 8 km and below about 35 – 40 km. Outside of this altitude region, there are biases between the RO missions that may be due to differences in the underlying measurement data, or due to differences in the sampling characteristics. In particular the CHAMP data below 8 km, at low- and mid-latitudes, are difficult to reconcile with data from the other RO missions. Below 8 km, the CHAMP data records themselves also exhibit abrupt bias changes (see discussion in Section 4.3.5).

4.3 Comparison between RO missions

Overlap in time between the RO missions allows investigation of the consistency between monthly mean data retrieved from different missions. The purpose of this investigation is to identify systematic differences between RO missions that cannot be explained by statistical measurement errors or by sampling effects. Systematic differences between RO missions, or between subsets of RO data, may constitute a problem, particularly if they are time varying. In the present report, we present results from the following overlap periods:

- CHAMP – COSMIC: September 2006 to September 2008
- GRACE – COSMIC: March 2007 to December 2016
- Metop – COSMIC: January 2007 to December 2016

For each pair of RO missions, for each geophysical variable, and for each latitude-altitude-month bin, we compute the following double difference:

$$\Delta\bar{X}_{A-B} = \left(\bar{X}_A^{obs} - \bar{X}_A^{bgr}\right) - \left(\bar{X}_B^{obs} - \bar{X}_B^{bgr}\right) \tag{3}$$

or, alternatively, the fractional double difference

$$\Delta\bar{X}_{A-B} = 100 \cdot \left(\left(\bar{X}_A^{obs} - \bar{X}_A^{bgr}\right) - \left(\bar{X}_B^{obs} - \bar{X}_B^{bgr}\right)\right) / \bar{X}_B^{bgr} \tag{4}$$

where $\bar{X}_A^{obs}$ and $\bar{X}_B^{obs}$ are the observed monthly means for missions A and B, respectively, and $\bar{X}_A^{bgr}$ and $\bar{X}_B^{bgr}$ are the corresponding means for the background (ERA-Interim) data. The background means are computed from background data co-located with the observed data. The purpose of this type of differencing is to reduce the impact of sampling effects on the inter-mission differences. It is equivalent to comparing sampling-error corrected monthly means (assuming that we use the background model for the sampling-error correction). Section 4.4 gives an idea of the efficiency of this method in reducing the sampling effects in the gridded data.

We also present anomaly time series for the CHAMP, COSMIC, and Metop missions, and anomaly difference time series for the RO missions with respect to an all-mission mean. The latter provides an overview of the consistency between RO missions, and the variation of that consistency with time.

In Section 4.3.1, we do area-weighted averaging of the differences $\Delta\bar{X}_{A-B}$ in 5 latitude bands, and plot them as time-height plots. Data are shown for three pairs of missions: CHAMP-COSMIC, GRACE-COSMIC, and Metop-COSMIC, and for bending angle, refractivity, and dry temperature.

In Section 4.3.2, we show time-height plots for the Metop-COSMIC differences, but with Metop separated in setting only and rising only. We also show Metop rising-setting differences.

In Section 4.3.3, we show the spatial pattern of differences between the RO missions, obtained by temporal averaging over the overlap periods.

In Section 4.3.4 we show plots of sampling-error corrected anomaly time series for CHAMP, COSMIC, and Metop averaged into height layers and latitude bands. The methods used for the computation of anomalies are described in [RD.6]. A common reference is used for all RO missions, which is necessary to reveal overall biases between the RO missions.

Finally, in Section 4.3.5 we show plots of anomaly difference time series for CHAMP, COSMIC, and Metop, with respect to an all-mission mean.

CHAMP – COSMIC
Bending angle

4.3.1 Figures: Mission difference time series

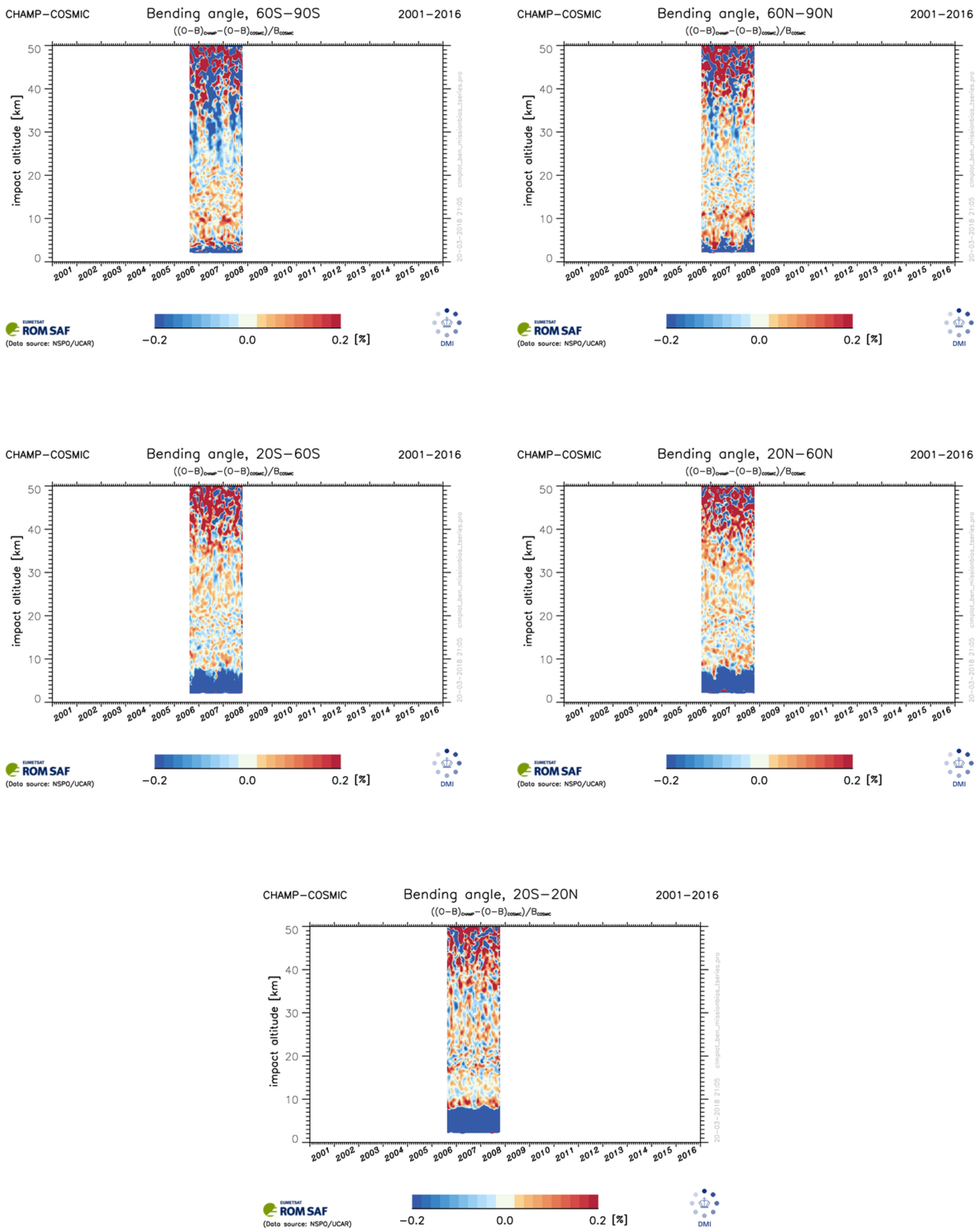


Figure 12a. *Differences between CHAMP and COSMIC monthly mean bending angles, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes. We find a systematic negative CHAMP-COSMIC bias at low and mid-latitudes below about 8 km.*

GRACE – COSMIC
Bending angle

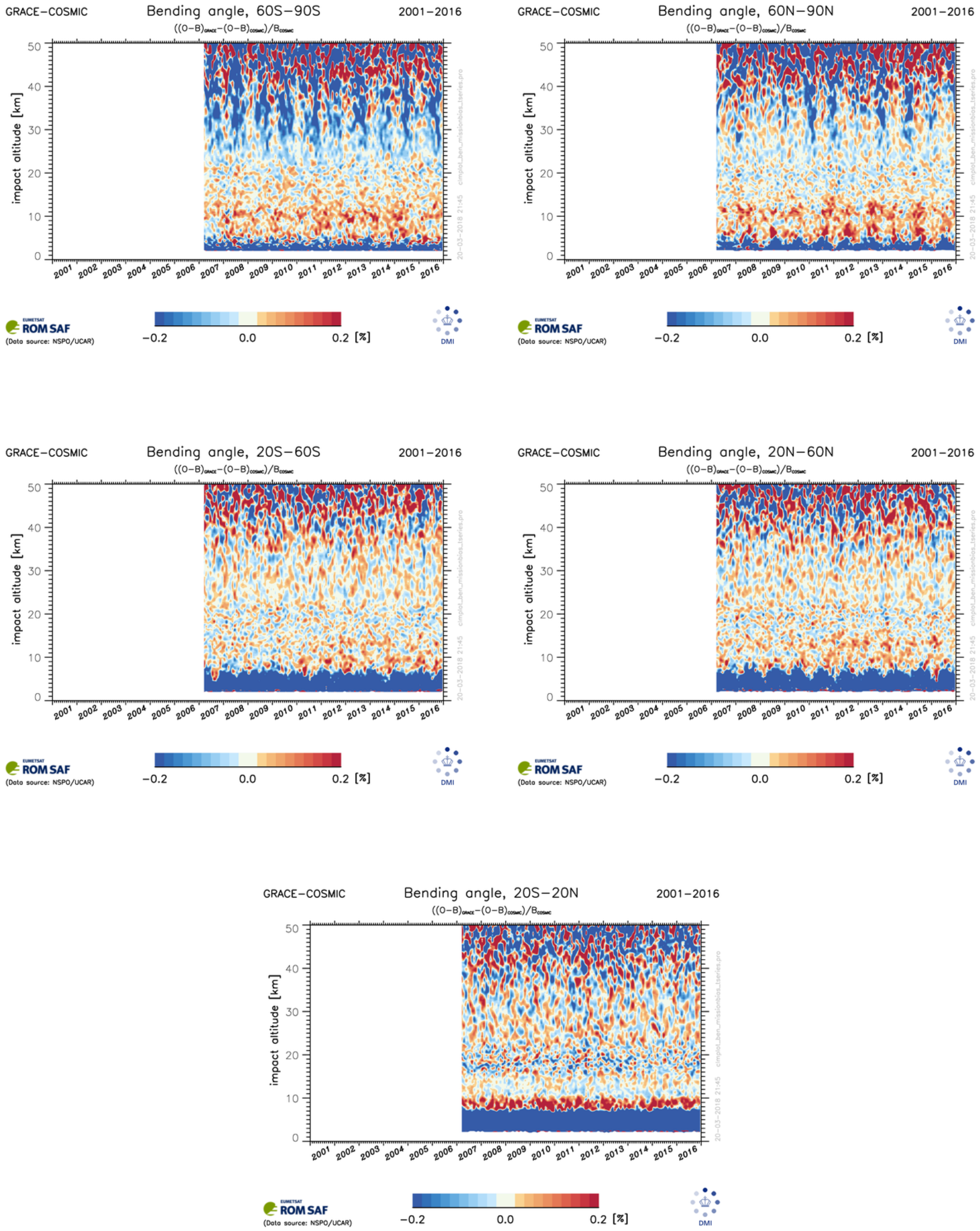


Figure 12b. *Differences between GRACE and COSMIC monthly mean bending angles, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes. We find a systematic negative GRACE-COSMIC bias at low and mid-latitudes below about 8 km, with tendencies to a seasonal cycle.*

Metop – COSMIC
Bending angle

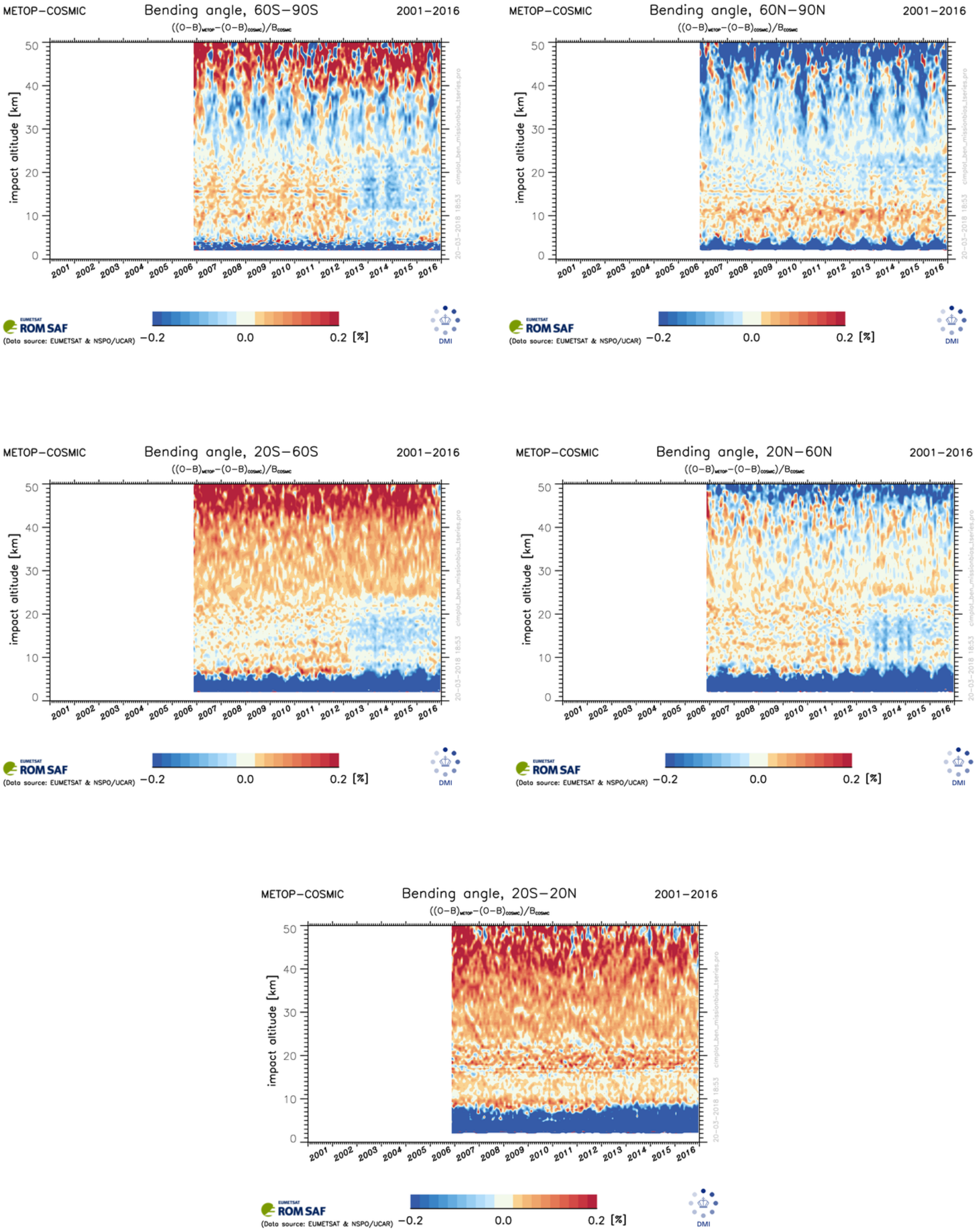


Figure 12c. *Differences between Metop and COSMIC monthly mean bending angles, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes. There are several features worth noting: a) seasonal cycle in the differences at high latitudes between 25 and 40 km, and (b) a systematic bias from 2013 below about 25 km, and c) a systematic negative Metop-COSMIC bias at low and mid-latitudes below about 8 km, with tendencies to a seasonal cycle.*

CHAMP – COSMIC Refractivity

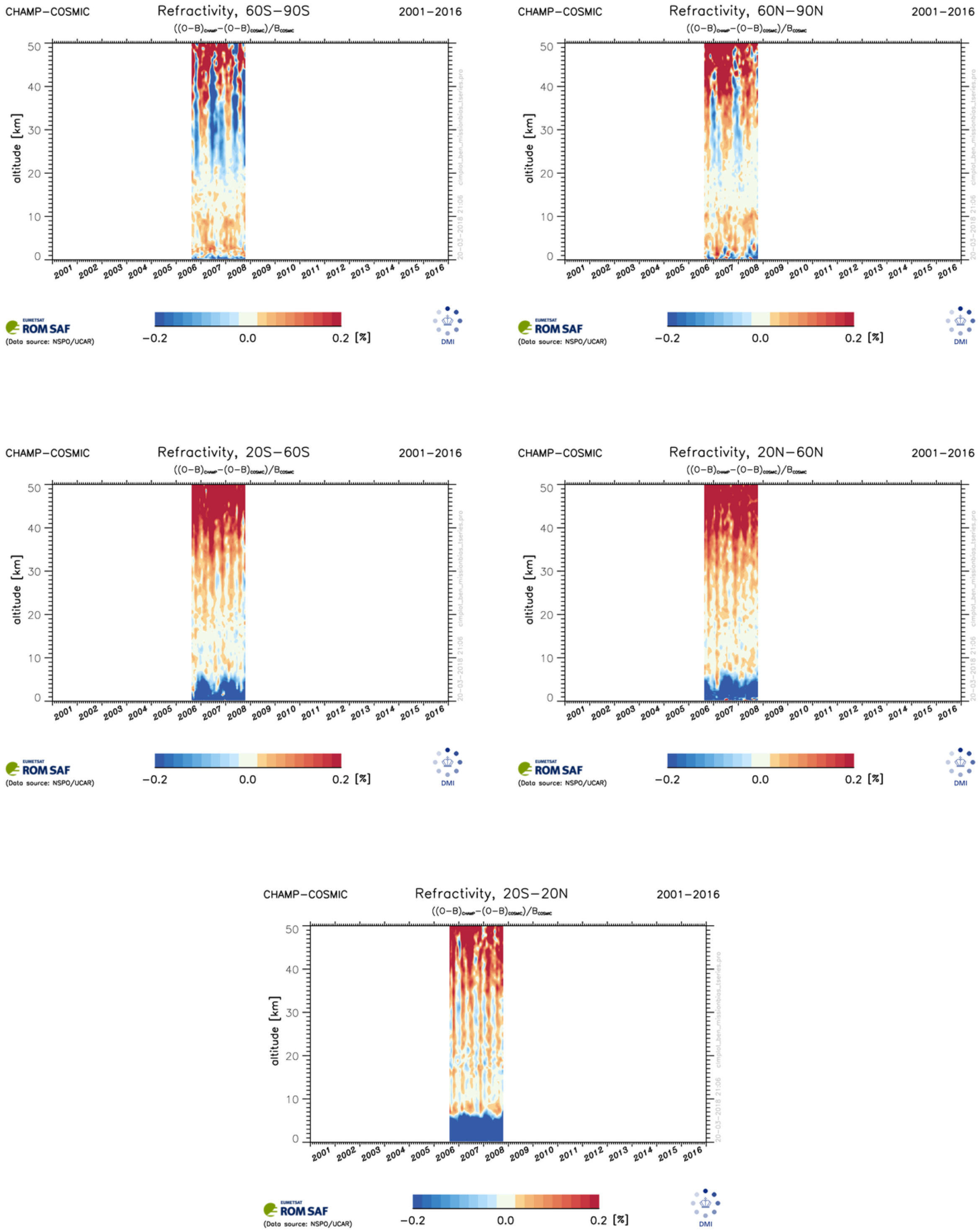


Figure 13a. *Differences between CHAMP and COSMIC monthly mean refractivity, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes. In addition to the features noted in the bending angles, we find an oscillating pattern at low and mid-latitudes possibly related to sampling effects due to the orbital drift of CHAMP.*

GRACE – COSMIC
Refractivity

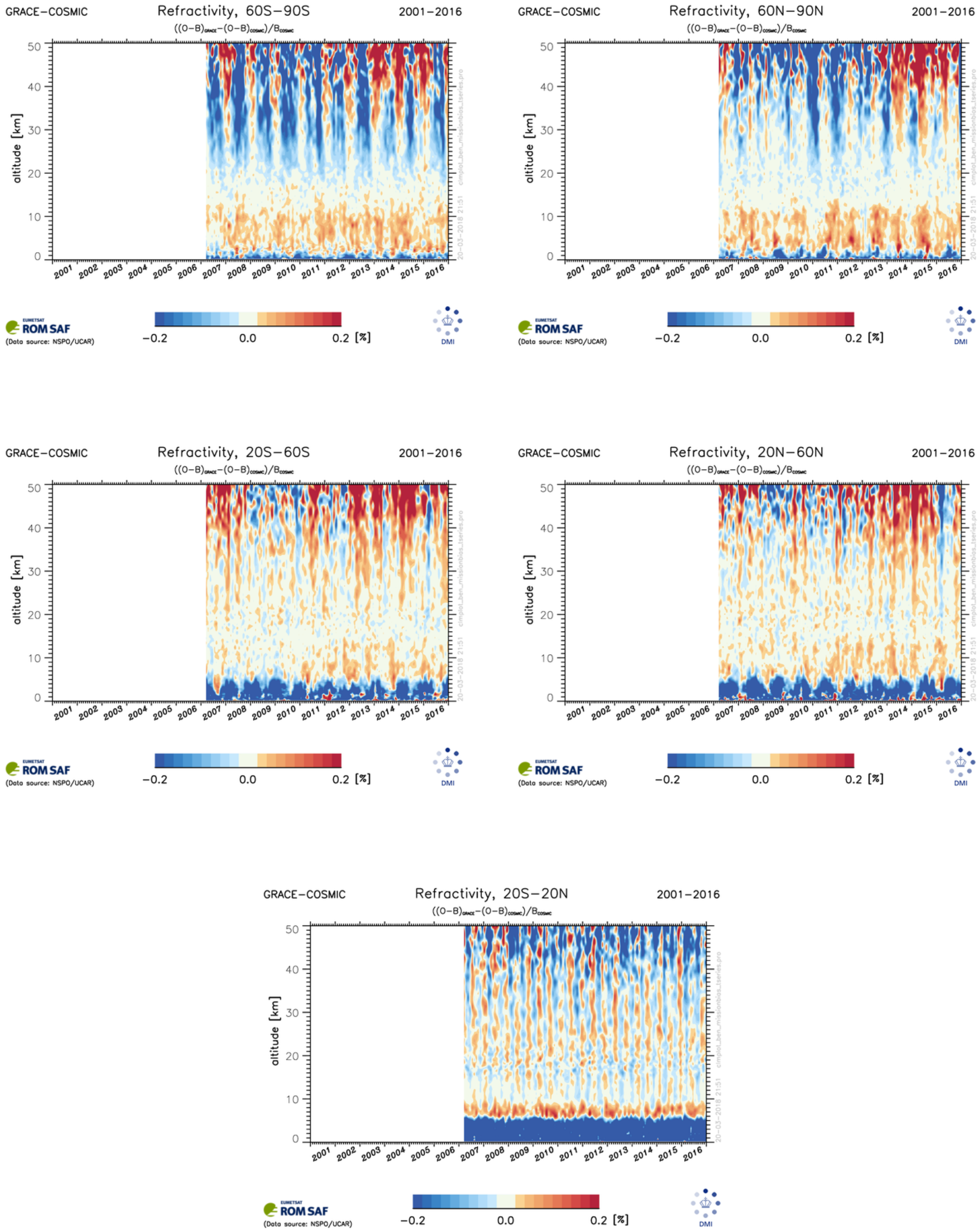


Figure 13b. *Differences between GRACE and COSMIC monthly mean refractivity, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes.*

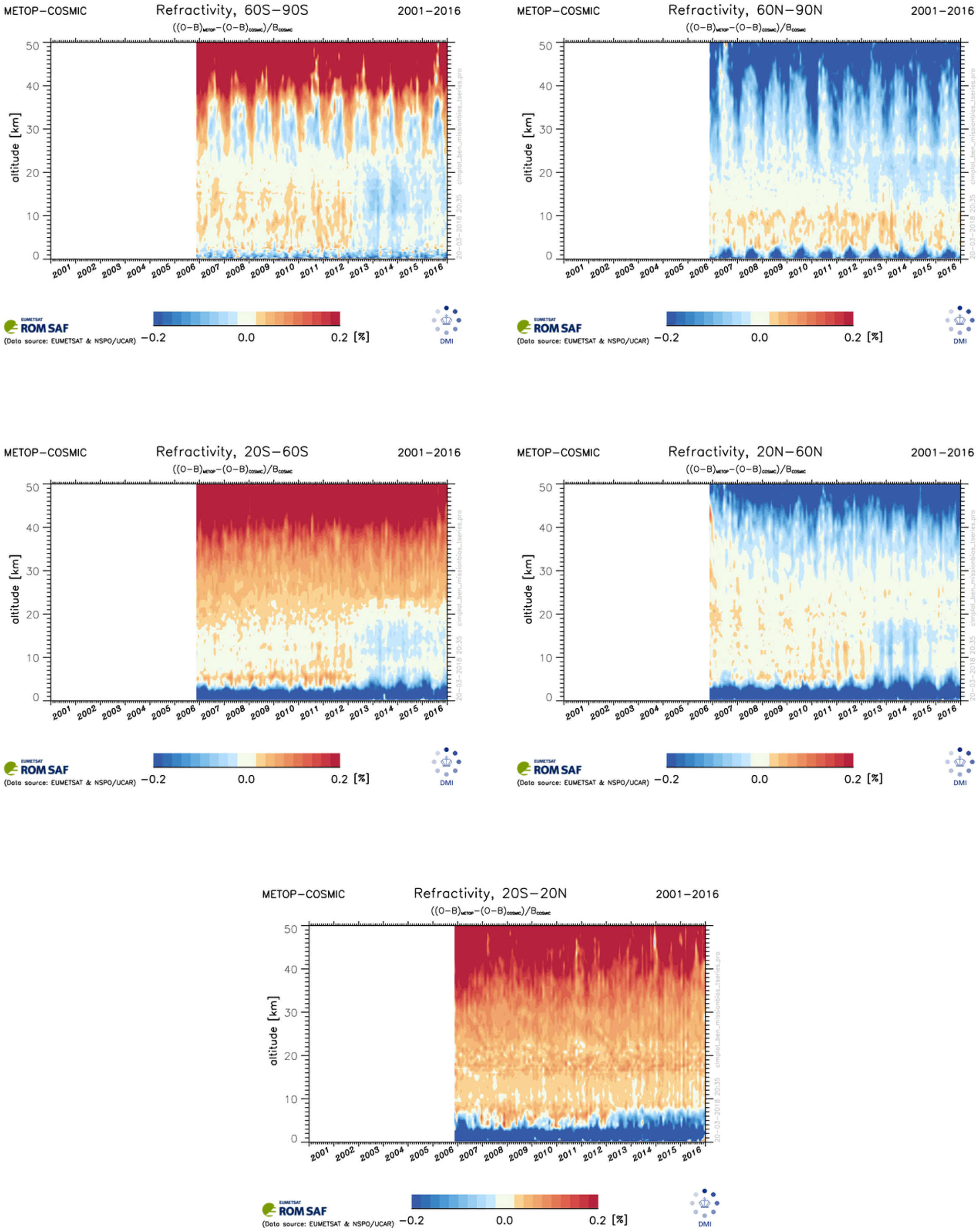


Figure 13c. *Differences between Metop and COSMIC monthly mean refractivity, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes.*

CHAMP – COSMIC
Dry temperature

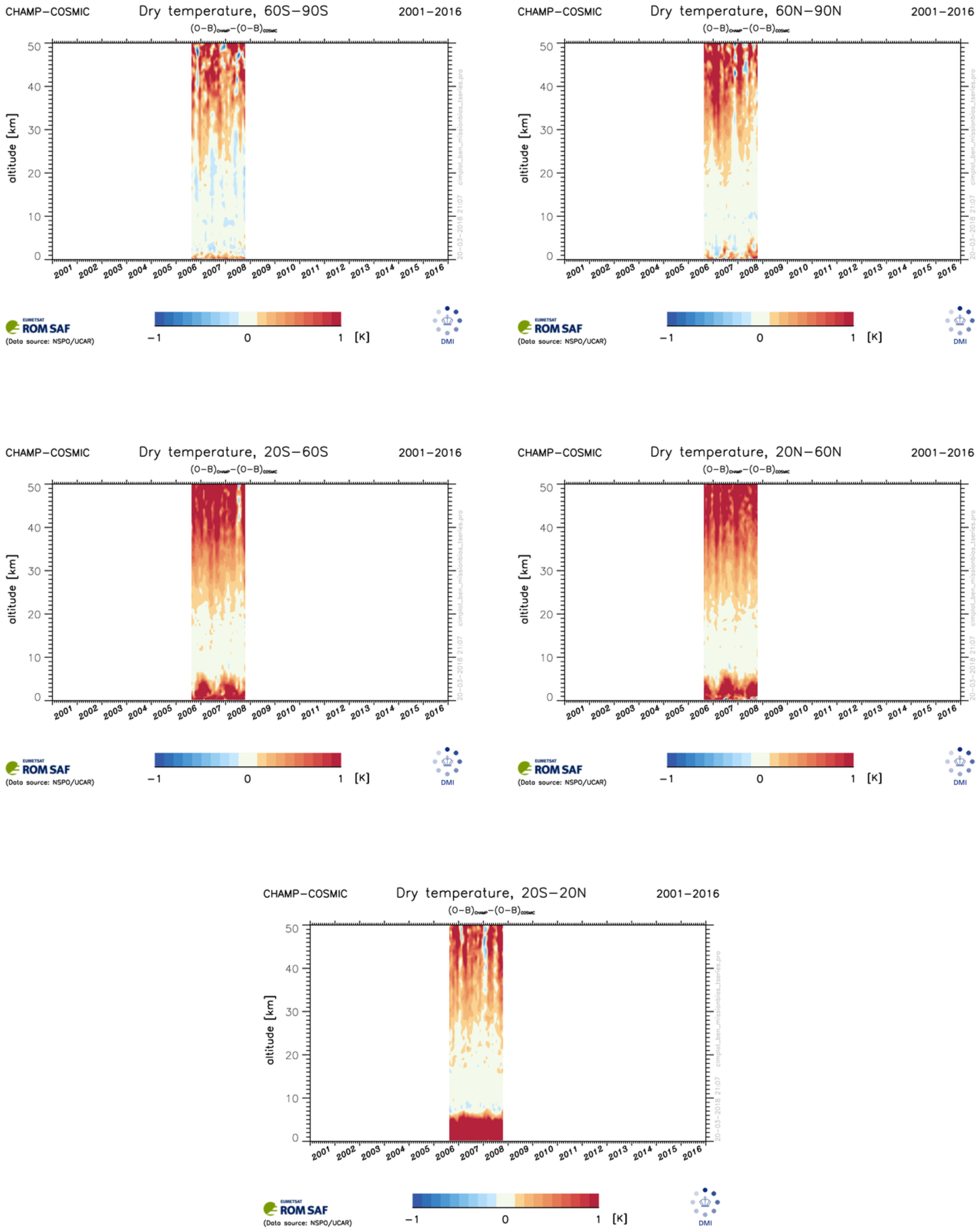


Figure 14a. *Differences between CHAMP and COSMIC monthly mean dry temperatures, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes.*

GRACE – COSMIC

Dry temperature

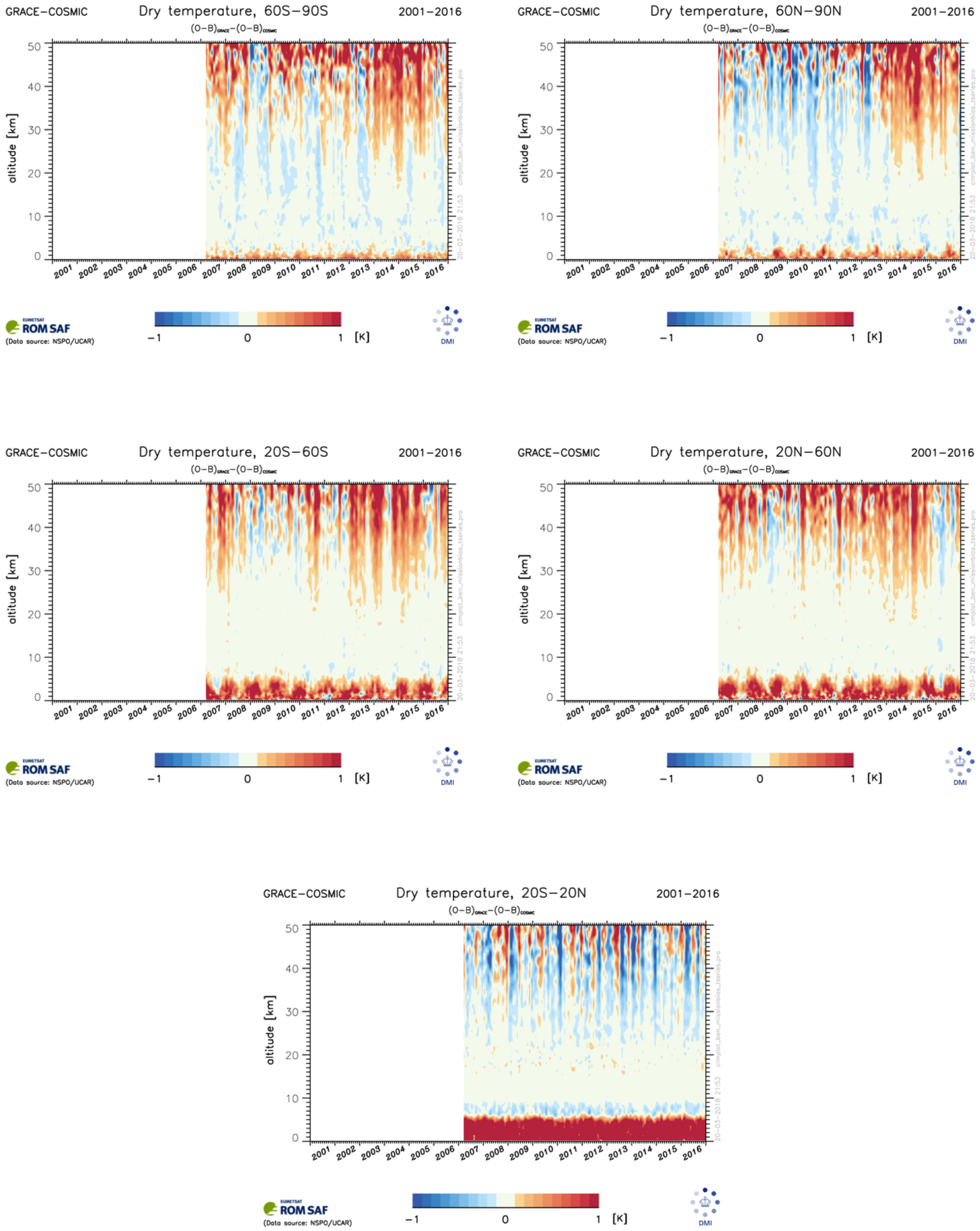


__Figure 14b.__ Differences between Metop and COSMIC monthly mean dry temperatures, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes.

Metop – COSMIC
Dry temperature

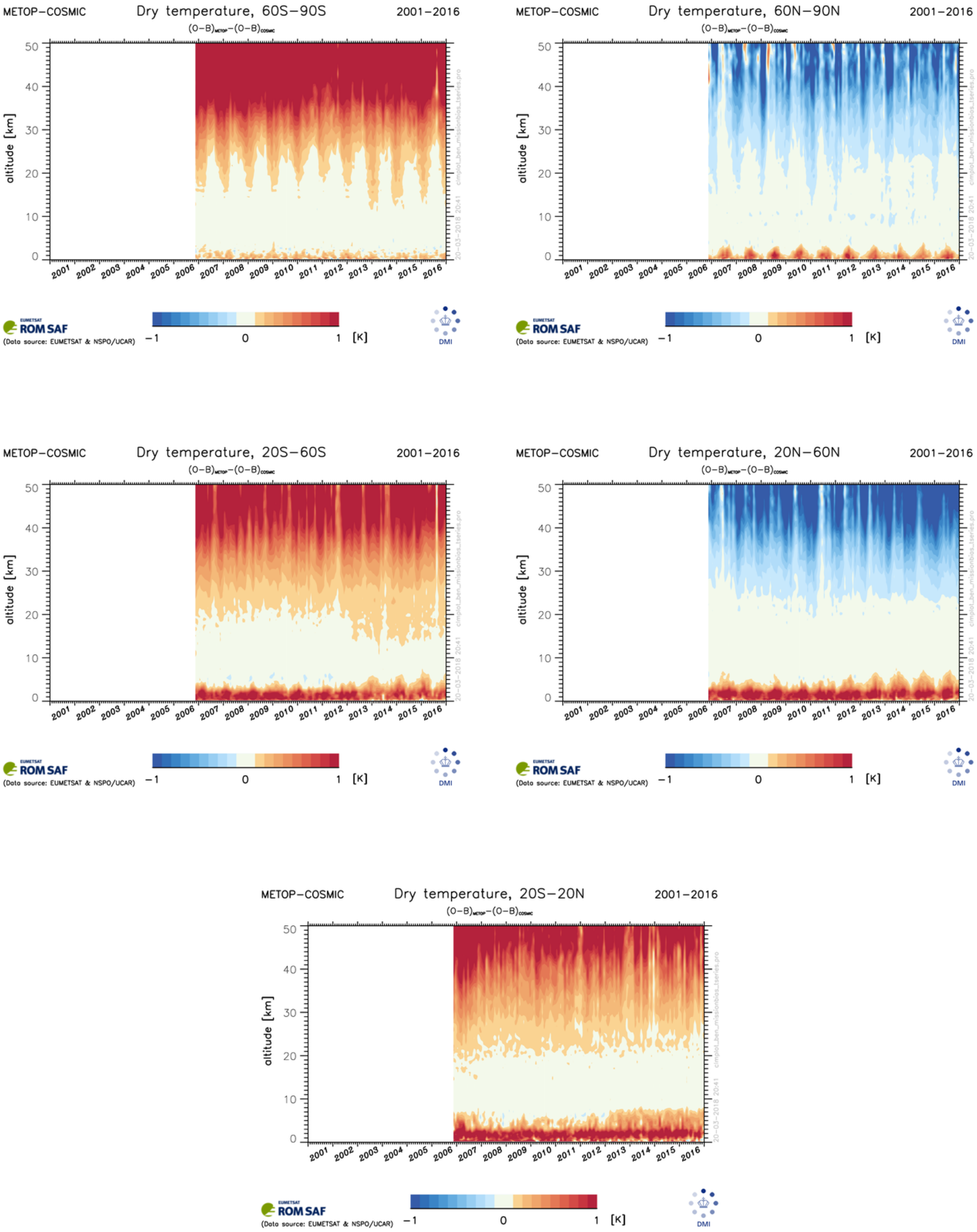


Figure 14c. *Differences between Metop and COSMIC monthly mean dry temperatures, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes.*

Metop – COSMIC setting – all

4.3.2 Figures: Metop setting/rising difference time series

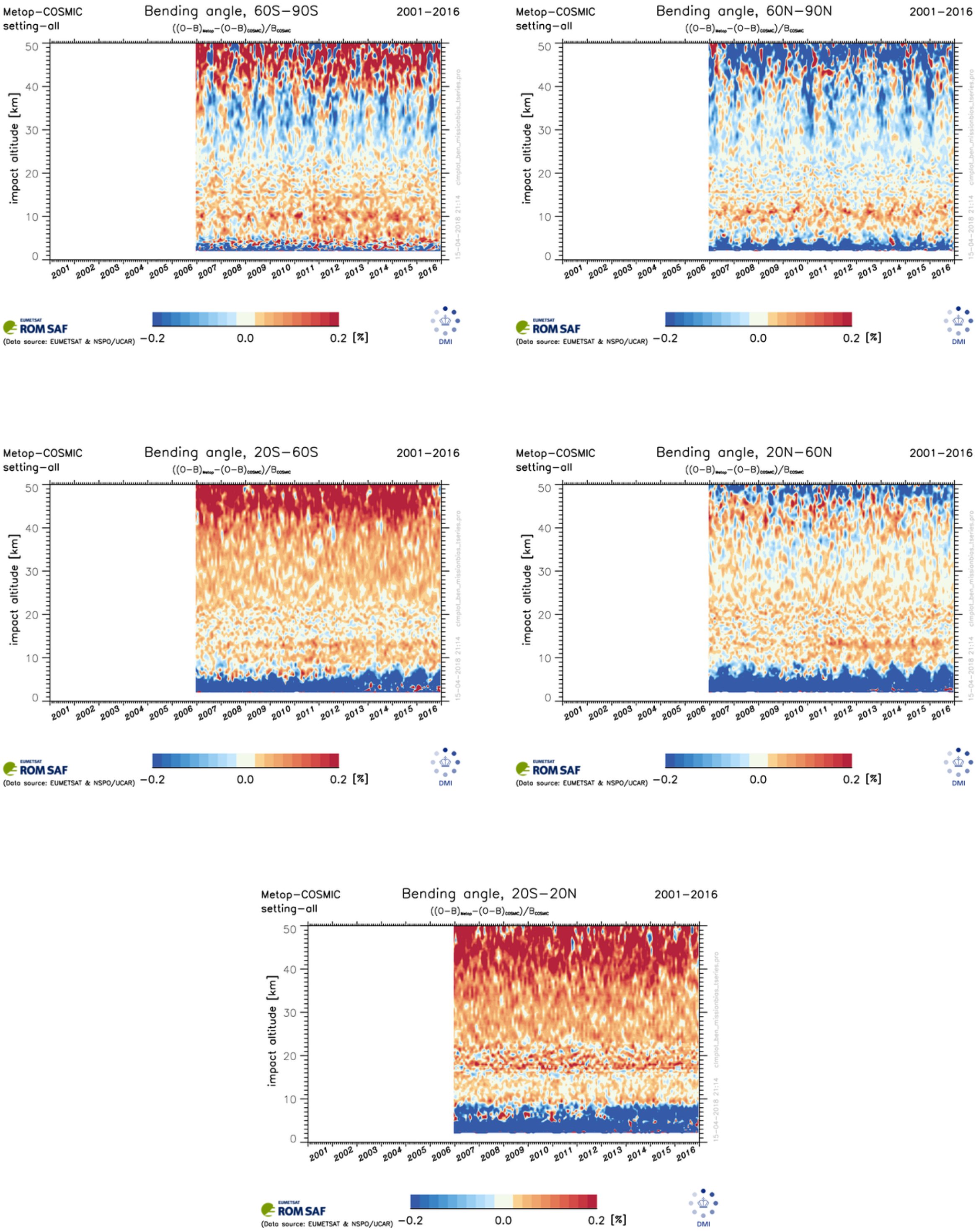


Figure 15a. *Differences between Metop (setting only) and COSMIC (all) monthly mean bending angle, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes. The Metop data here only include setting occultations, while the COSMIC data include both rising and setting.*

Metop – COSMIC
rising – all

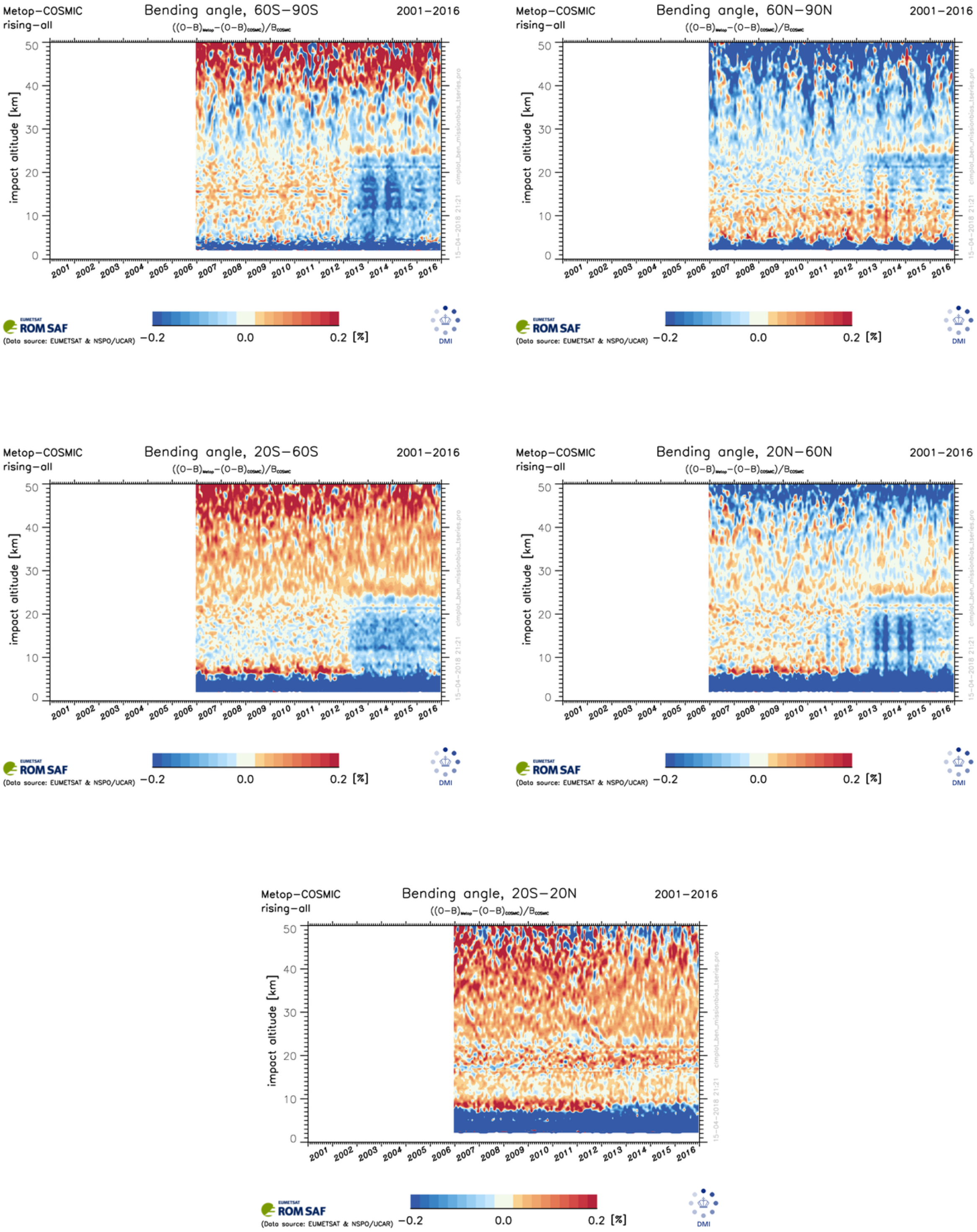


Figure 15b. *Differences between Metop (rising only) and COSMIC (all) monthly mean bending angle, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes. The Metop data here only include rising occultations, while the COSMIC data include both rising and setting.*

Metop – Metop rising – setting

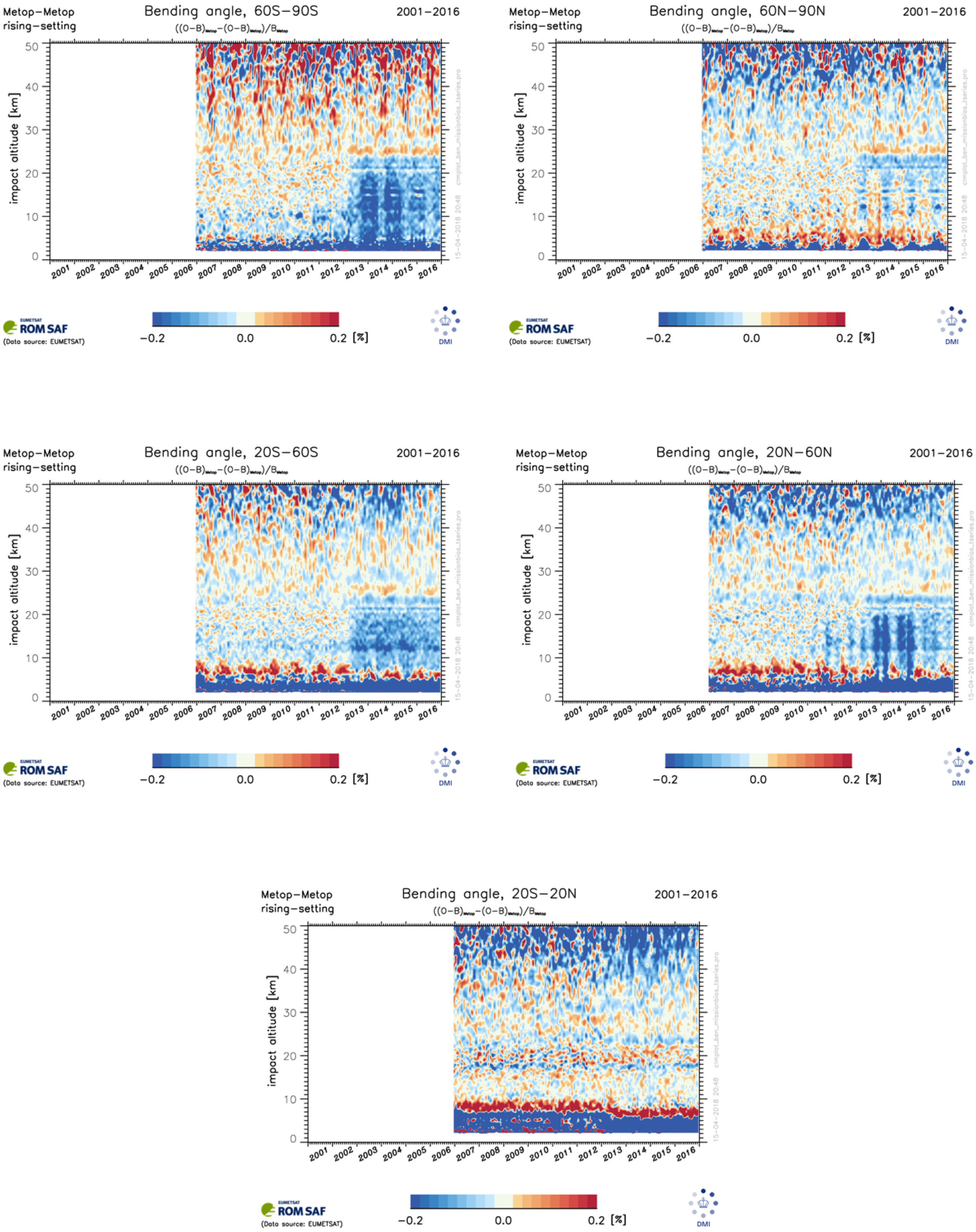
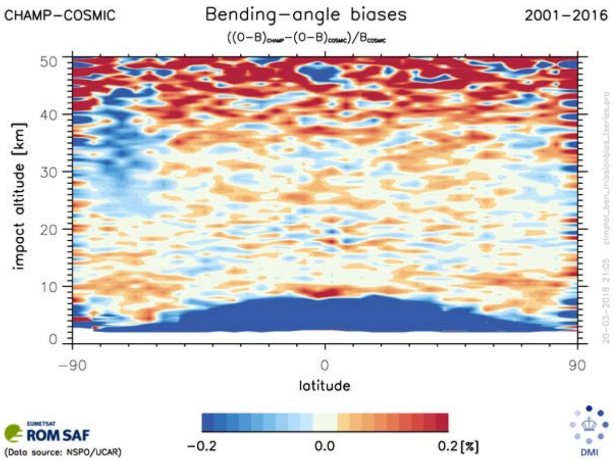
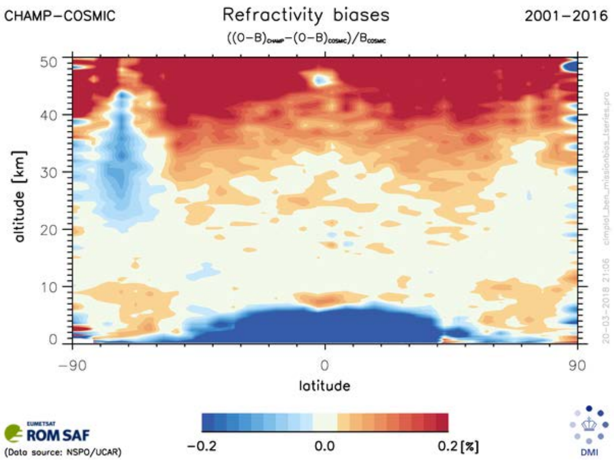


Figure 15c. *Differences between Metop (rising only) and Metop (setting only) monthly mean bending angles, after sampling-error correction, at high latitudes, mid-latitudes, and low latitudes.*

CHAMP-COSMIC
Ben, Ref, Tdry

4.3.3 Figures: Spatial distribution of mission differences





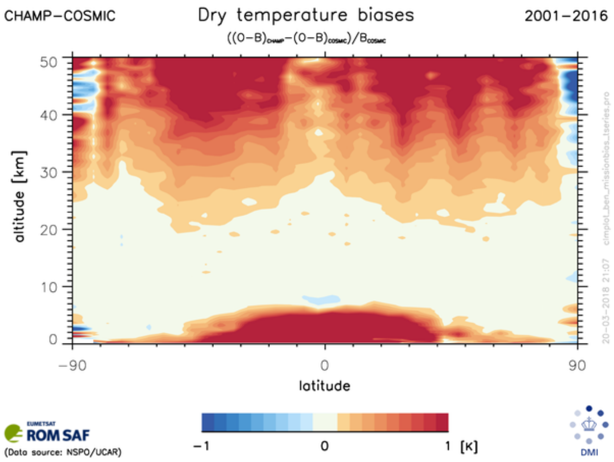


Figure 16a. *Average differences between CHAMP and COSMIC during the mission overlap period Sep 2006 to Sep 2008, for bending angle (upper panel), refractivity (middle panel), and dry temperature (lower panel).*

GRACE-COSMIC
Ben, Ref, Tdry

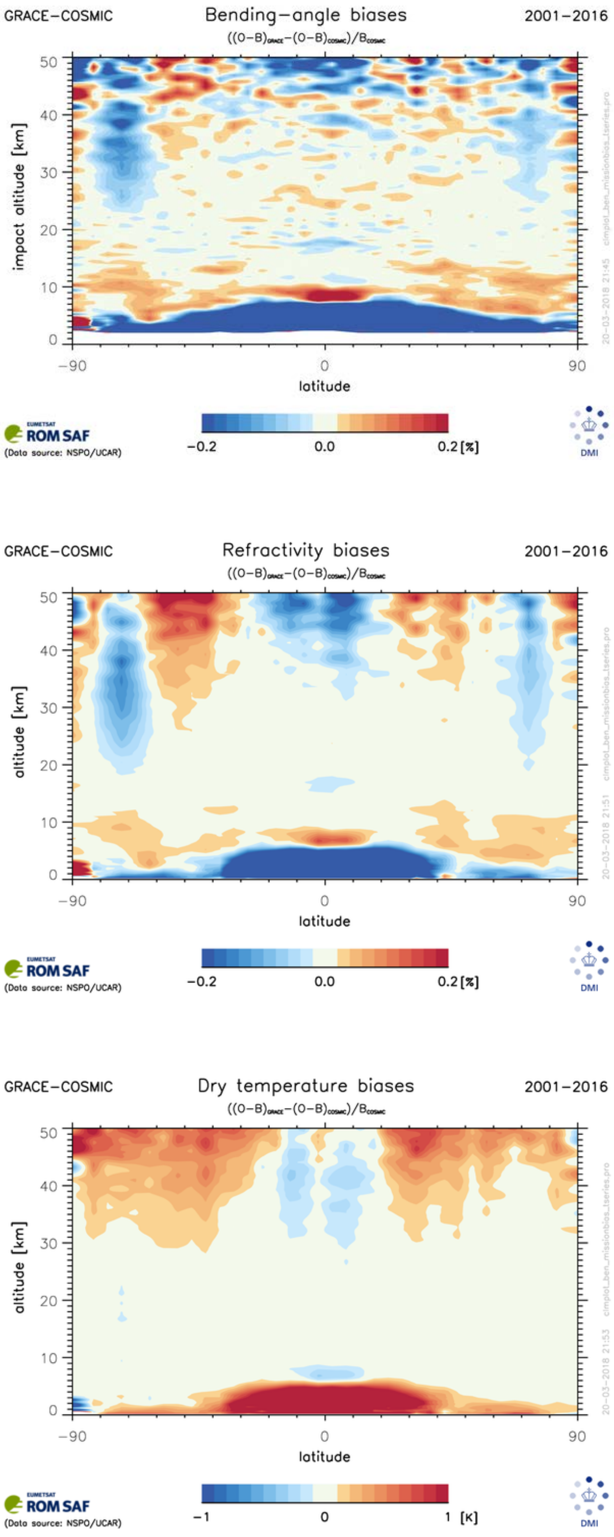


Figure 16b. *Average differences between GRACE and COSMIC during the mission overlap period Jan 2007 to Dec 2016, for bending angle (upper panel), refractivity (middle panel), and dry temperature (lower panel).*

Metop-COSMIC
Ben, Ref, Tdry

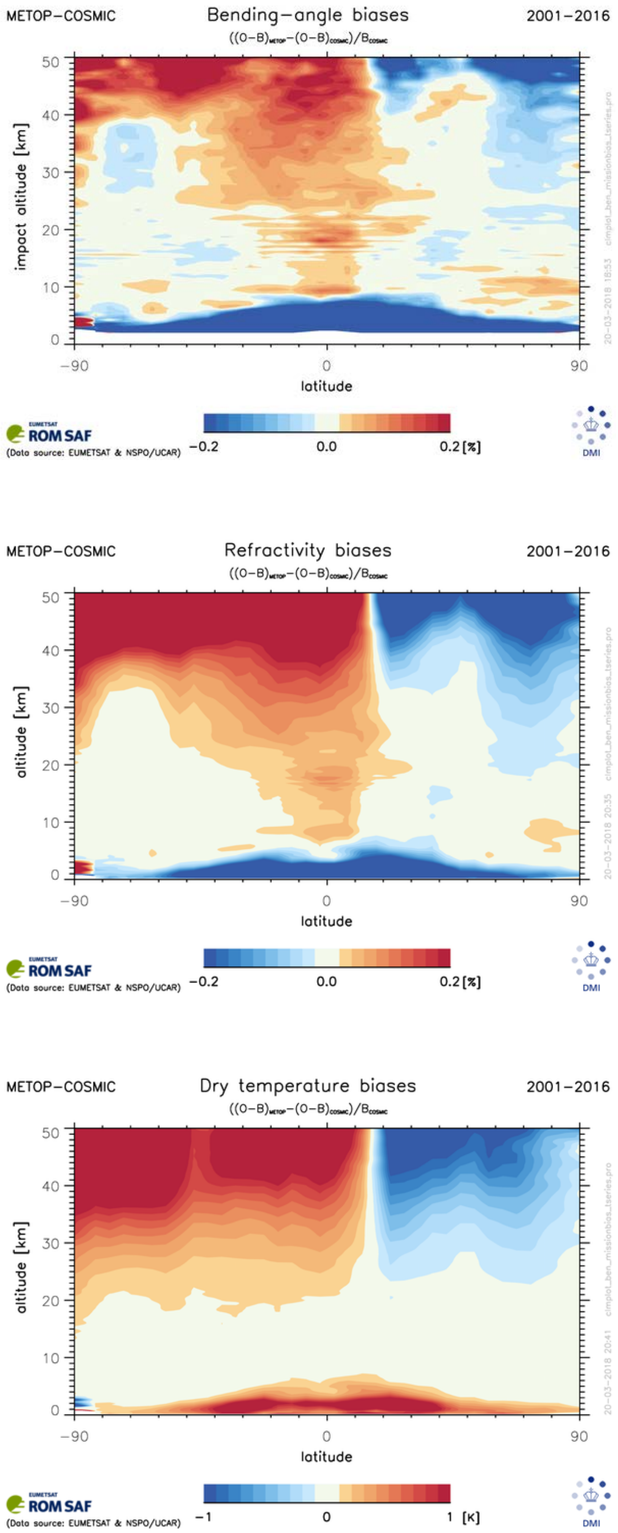


Figure 16c. *Average differences between Metop and COSMIC during the mission overlap period Jan 2007 to Dec 2016, for bending angle (upper panel), refractivity (middle panel), and dry temperature (lower panel).*

RO missions
Bending angle

4.3.4 Figures: Anomaly time series

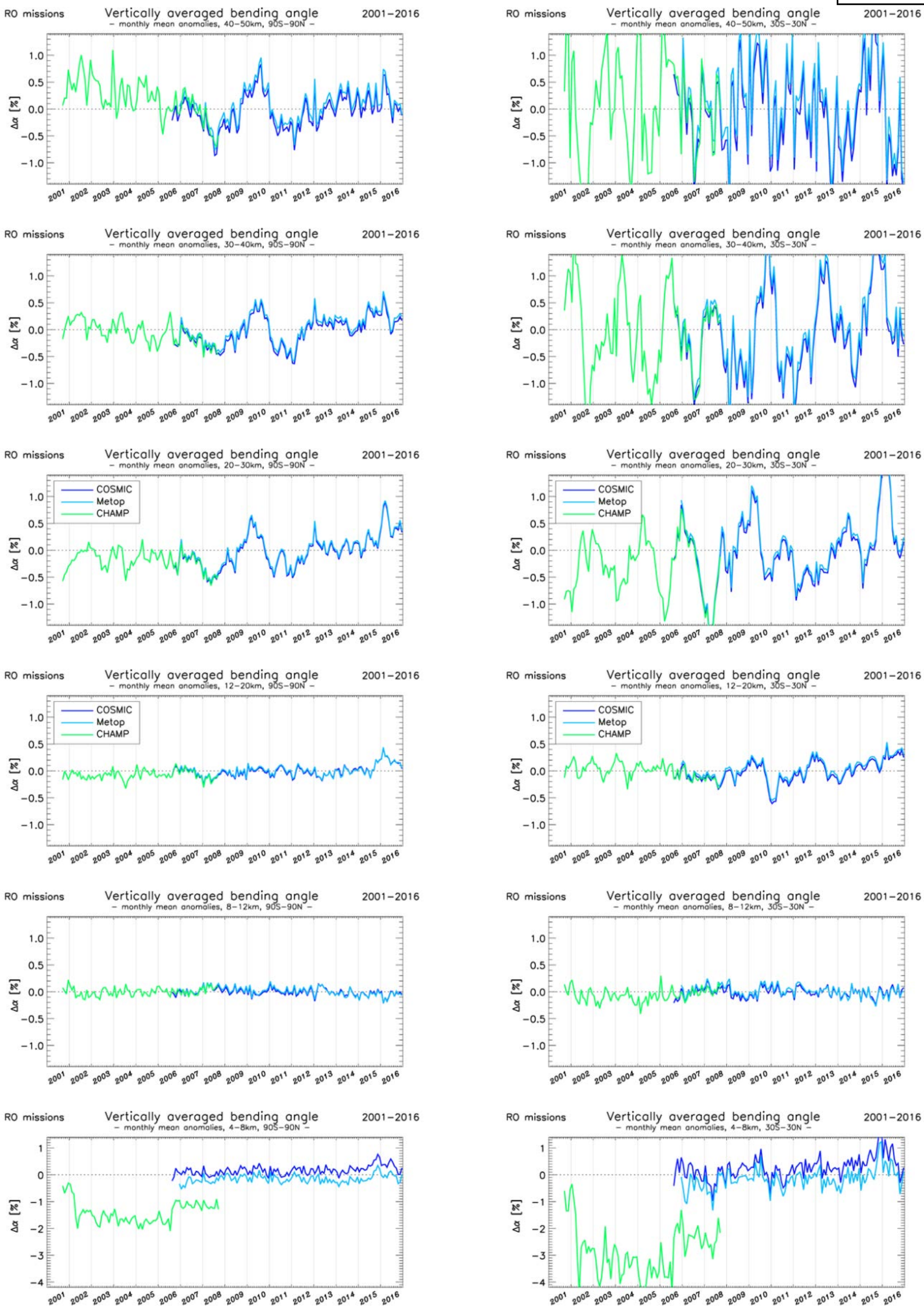


Figure 17a. *Bending angle anomaly time series (sampling-error corrected) for the CHAMP, COSMIC, and Metop missions. Global data to the left and low-latitude data to the right. Note that the y-axis has been expanded for the lowest layer.*

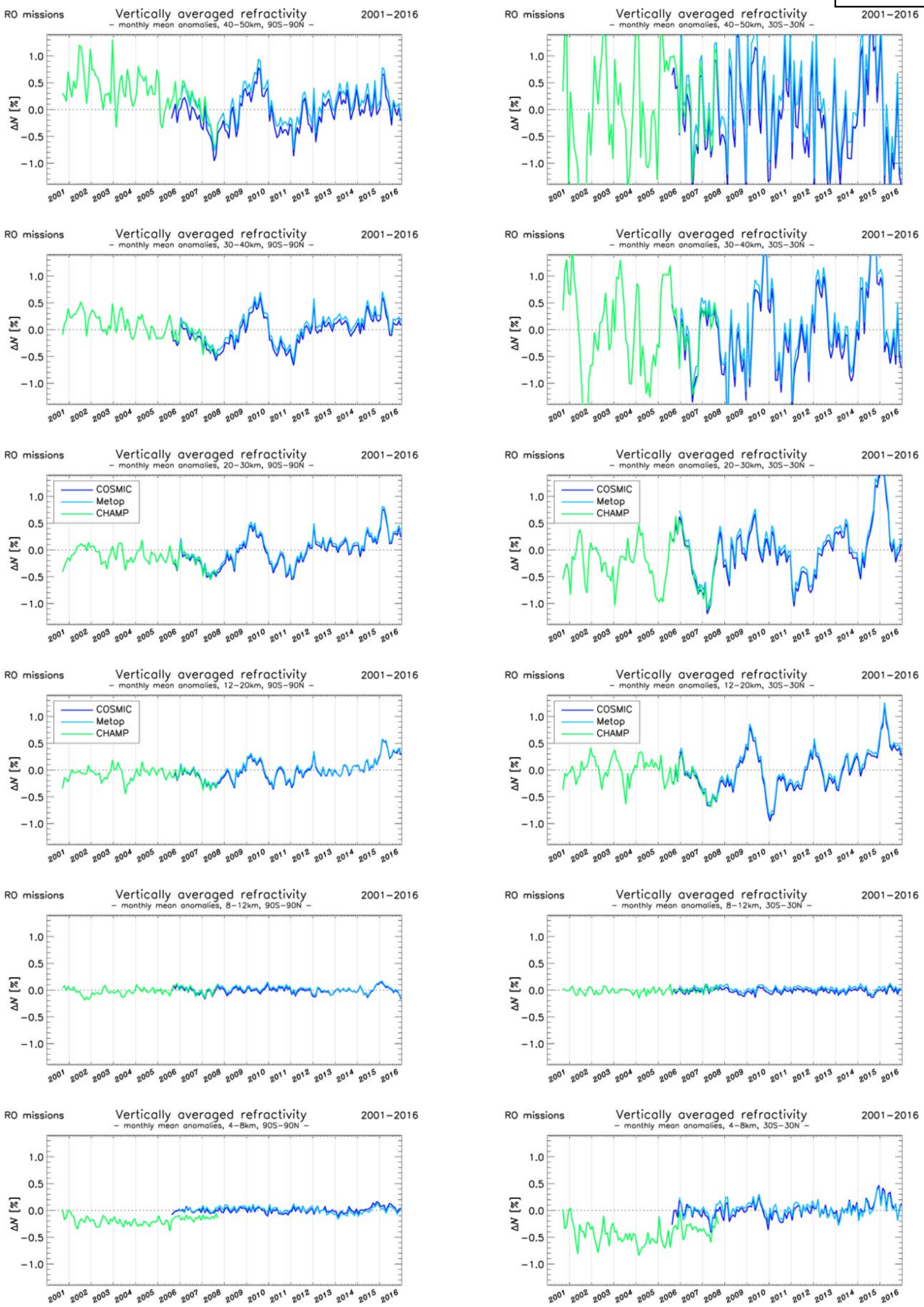


Figure 17b. *Refractivity anomaly time series (sampling-error corrected) for the CHAMP, COSMIC, and Metop missions. Global data to the left and low-latitude data to the right.*

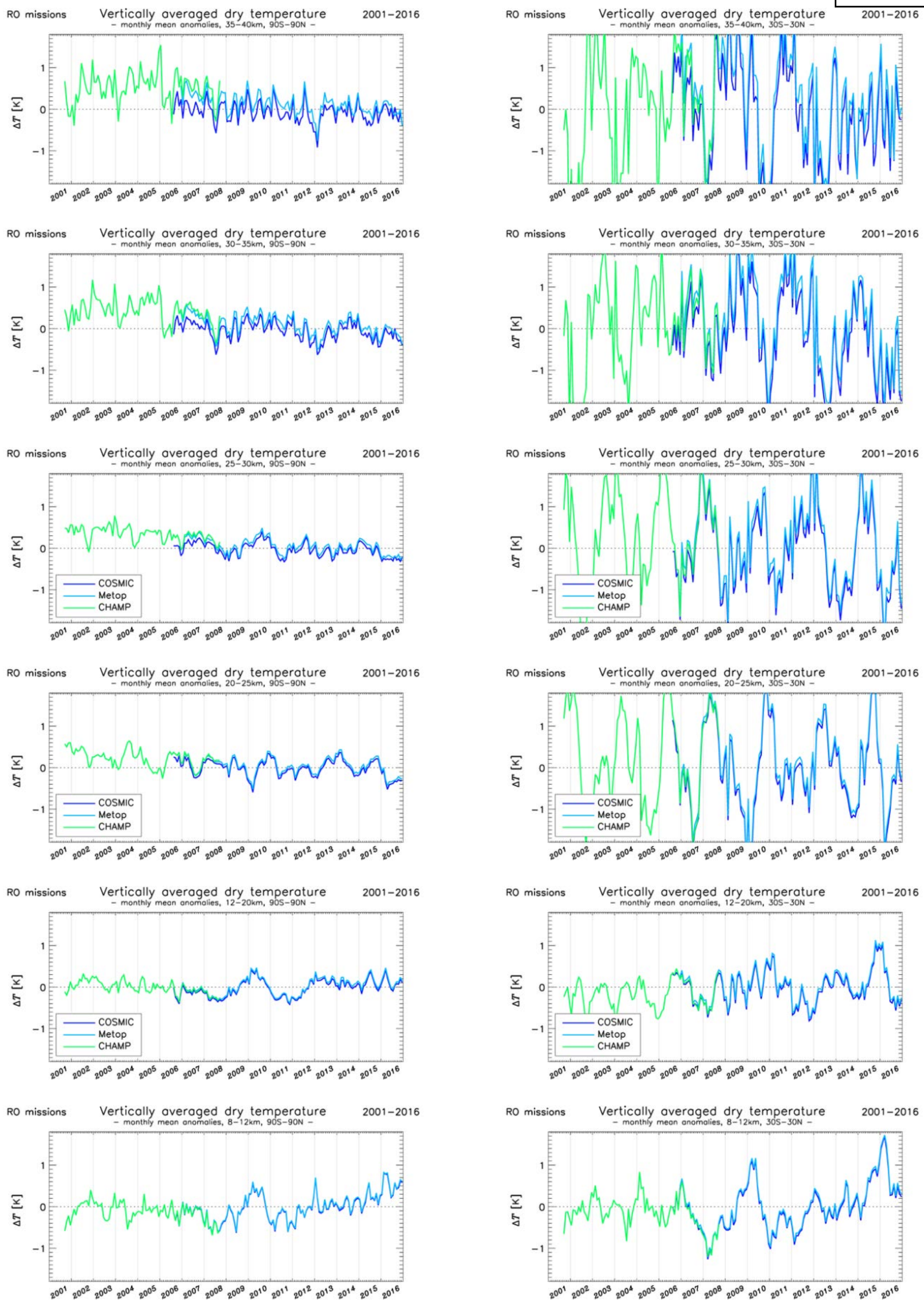


Figure 17c. *Dry temperature anomaly time series (sampling-error corrected) for the CHAMP, COSMIC, and Metop missions. Global data to the left and low-latitude data to the right.*

RO missions
Bending angle

4.3.5 Figures: Anomaly difference time series

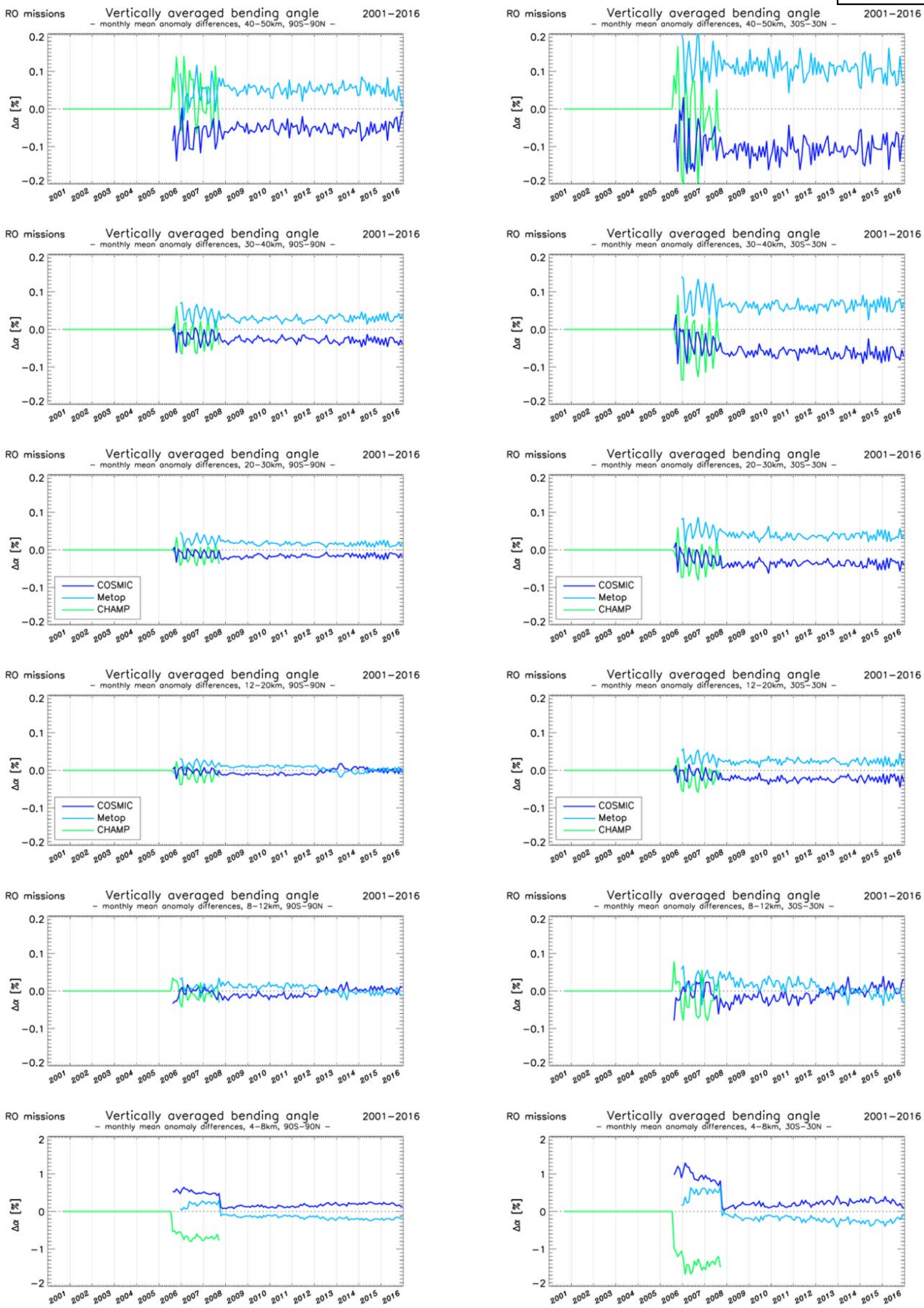


Figure 18a. *Bending angle* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Global data to the left and low-latitude data to the right.*

RO missions
Bending angle

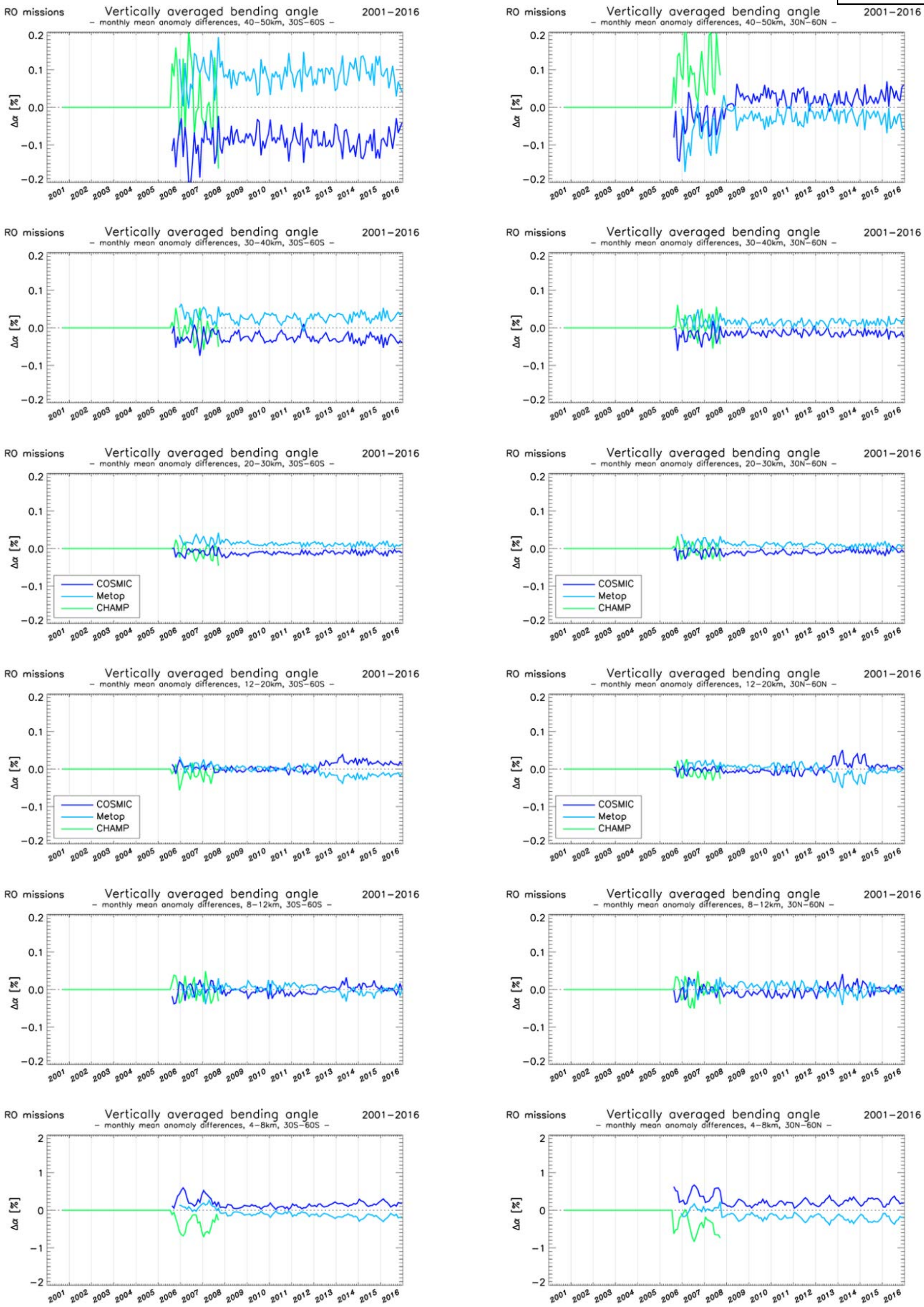


Figure 18b. *Bending angle* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Southern mid-latitudes to the right, northern mid-latitudes to the left.*

RO missions
Bending angle

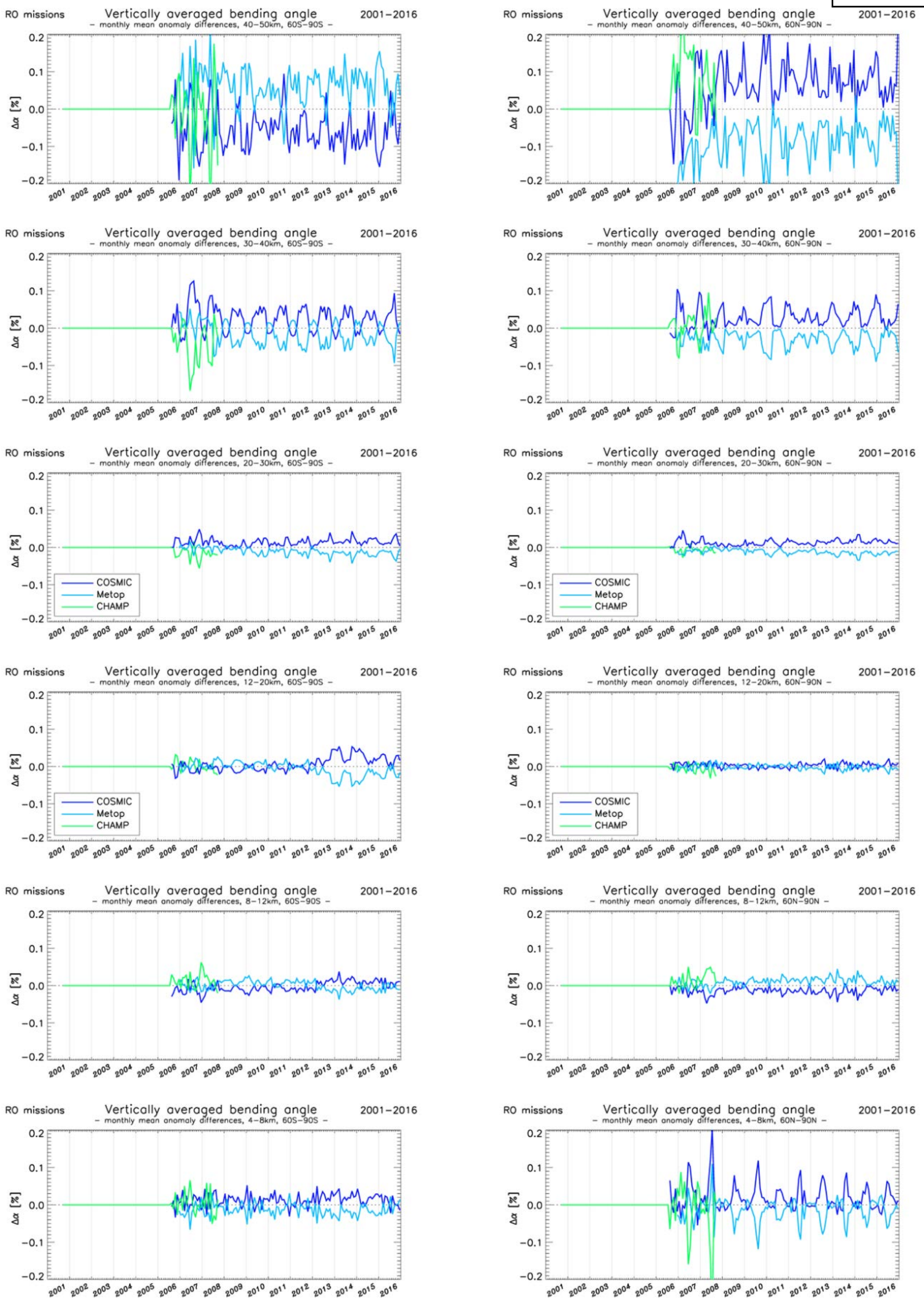


Figure 18c. *Bending angle* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Southern high latitudes to the right, northern high latitudes to the left.*

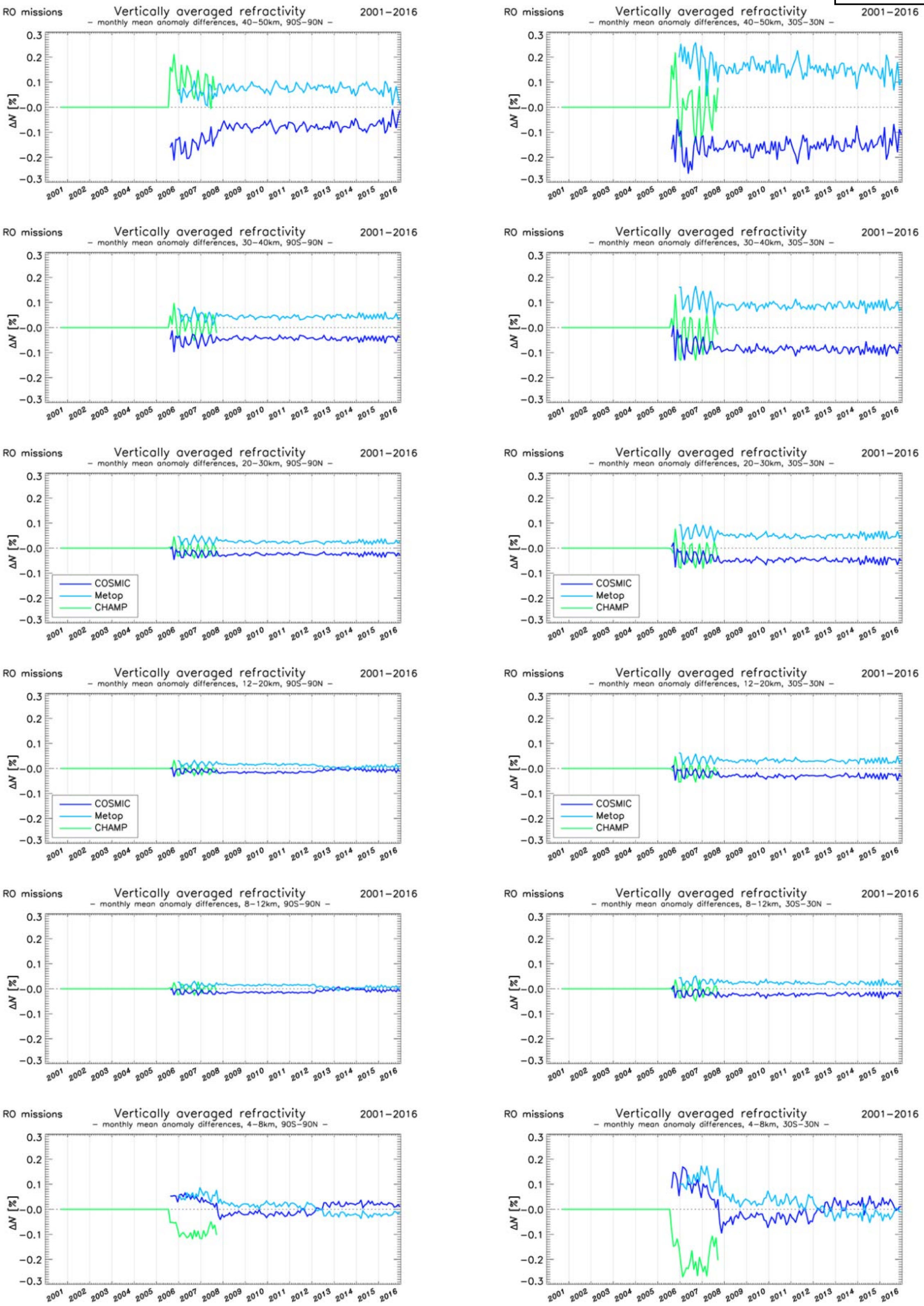


Figure 18d. *Refractivity* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Global data to the left and low-latitude data to the right.*

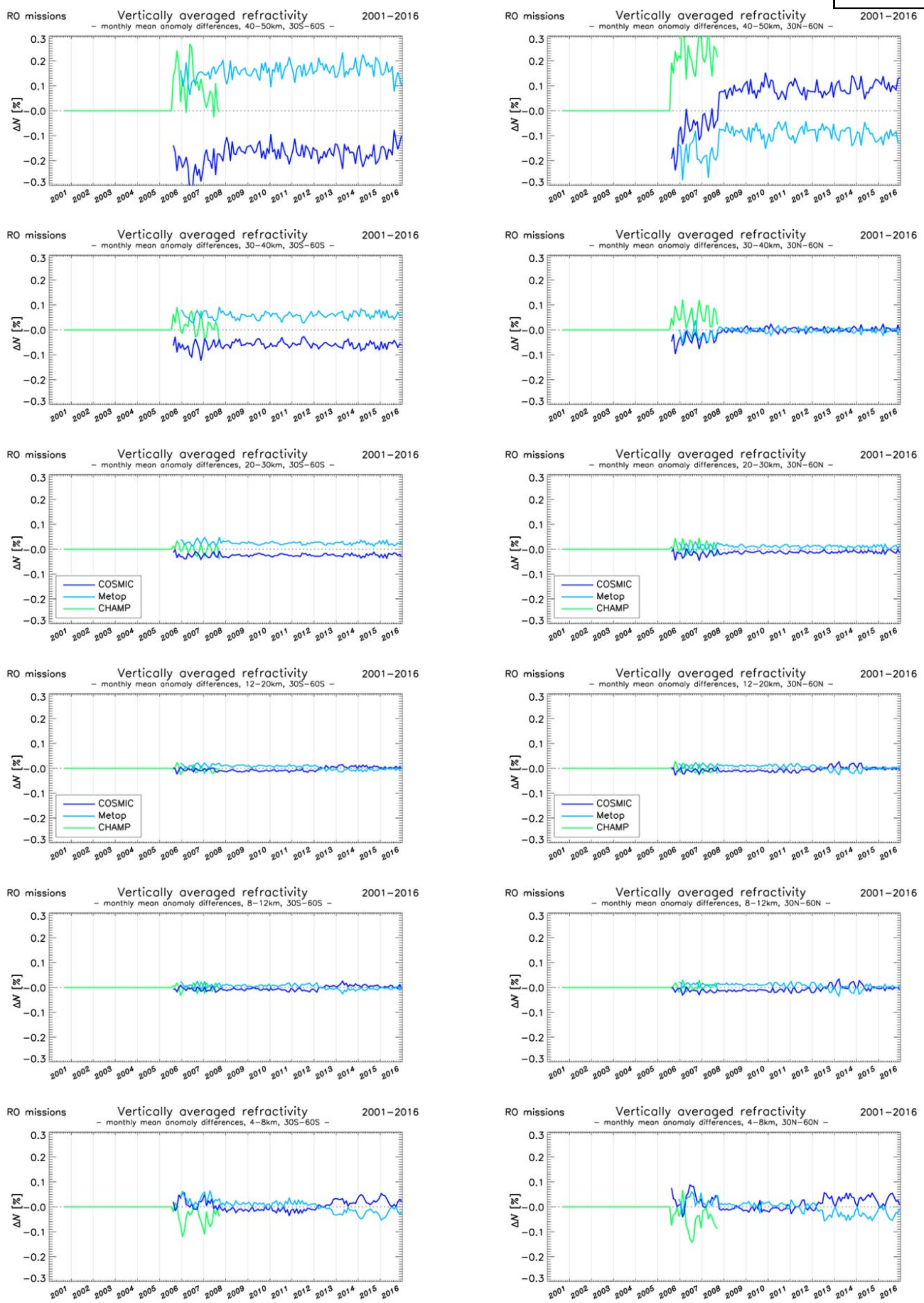


Figure 18e. *Refractivity* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Southern mid-latitudes to the right, northern mid-latitudes to the left.*

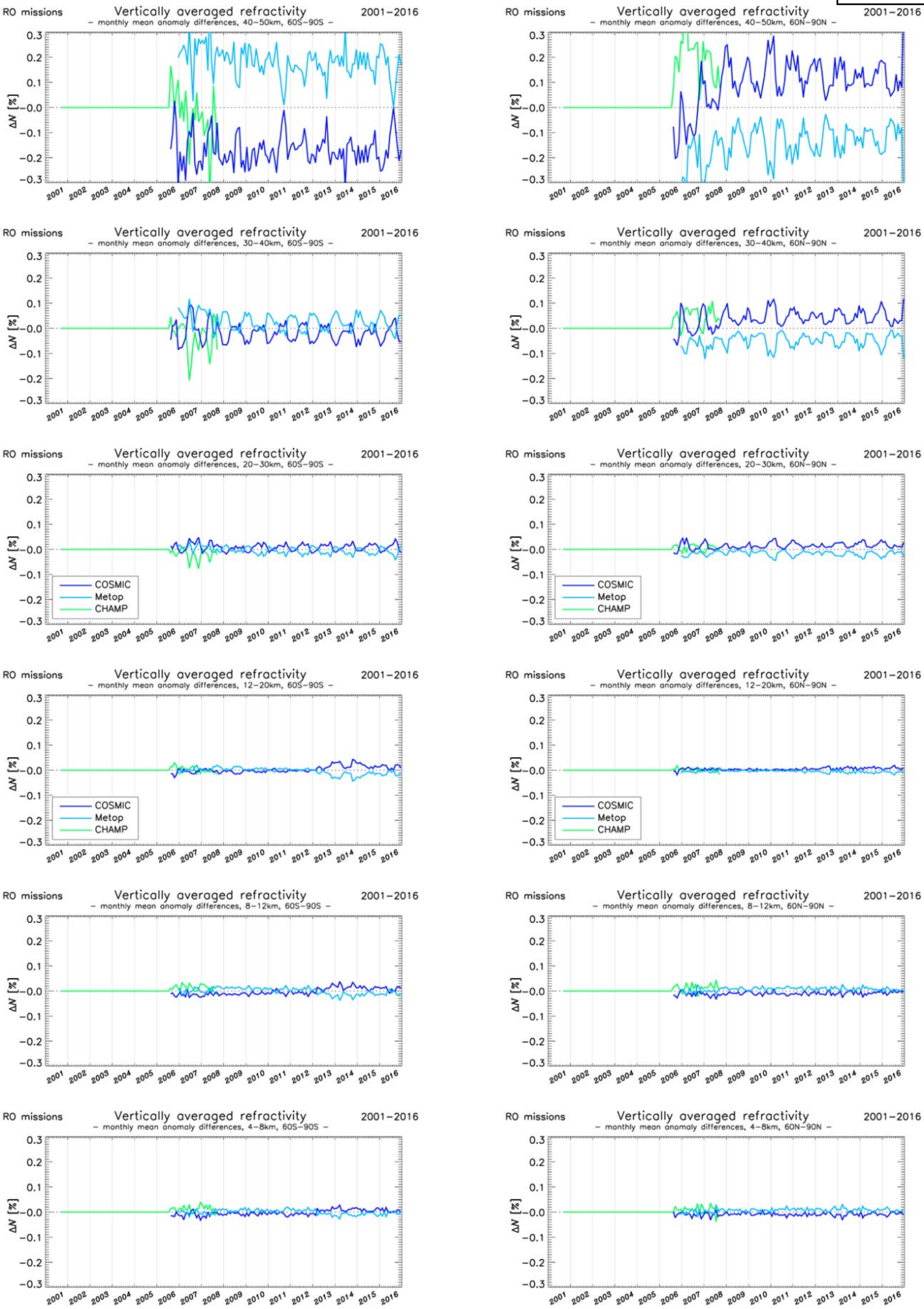


***Figure 18f.** Refractivity **anomaly difference** time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Southern high latitudes to the right, northern high latitudes to the left.*

RO missions

Dry temperature

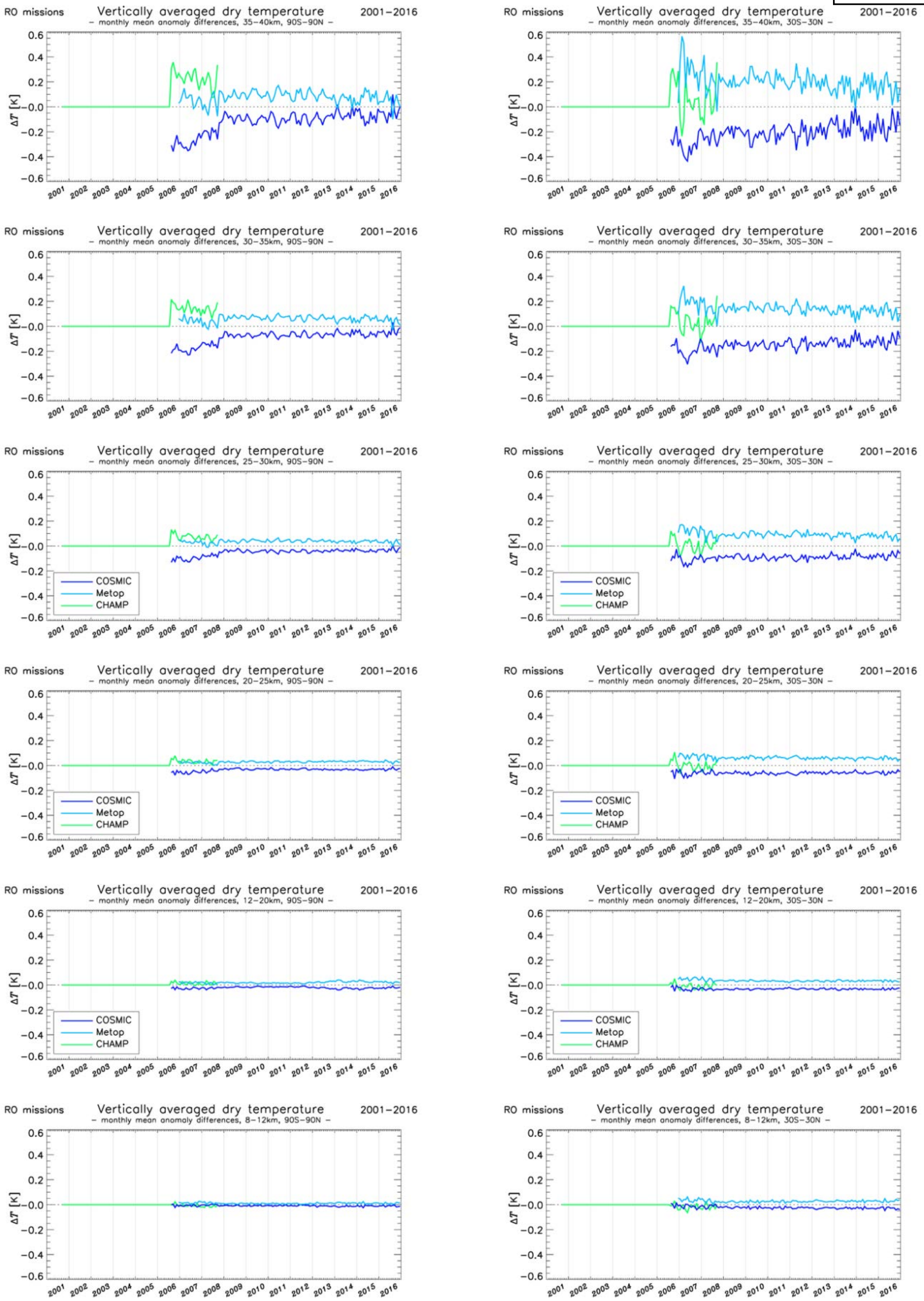


Figure 18g. *Dry temperature* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Global data to the left and low-latitude data to the right.*

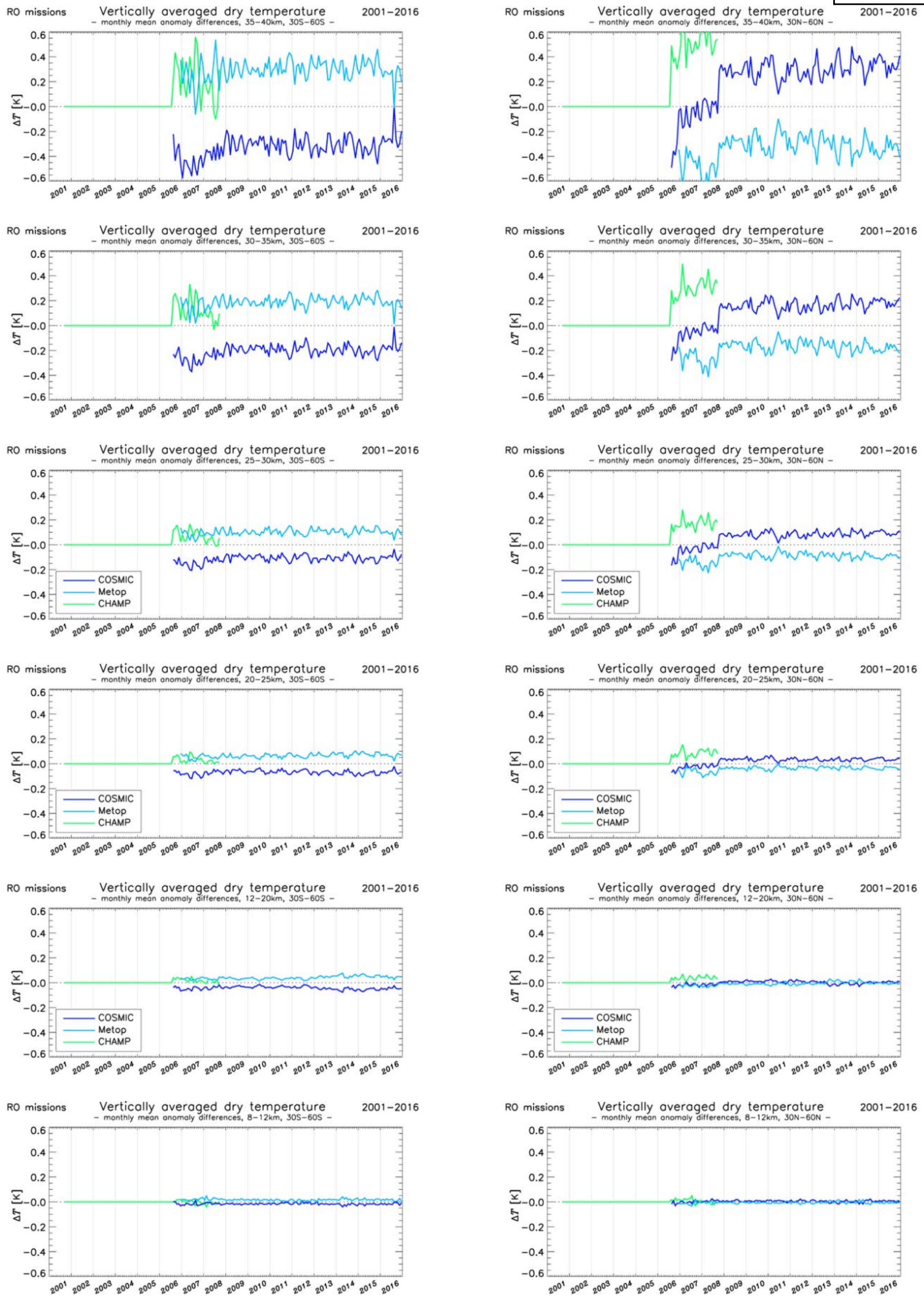


Figure 18h. *Dry temperature* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Southern mid-latitudes to the right, northern mid-latitudes to the left.*

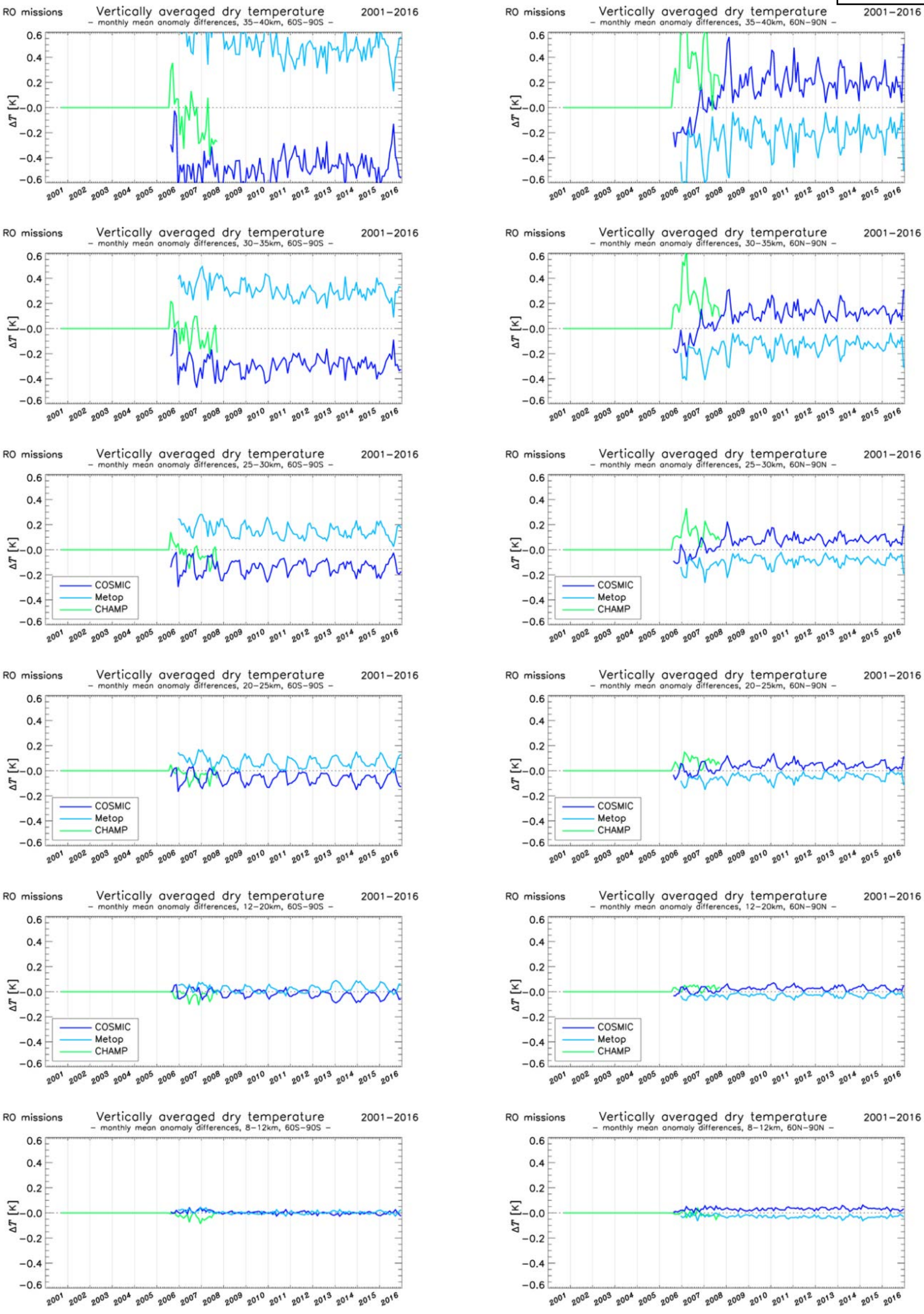


Figure 18i. *Dry temperature* ***anomaly difference*** *time series for CHAMP, COSMIC, and Metop. The differences are computed with respect to the mean of the three missions. Southern high latitudes to the right, northern high latitudes to the left.*

4.3.6 Discussion and conclusions

The inter-mission differences between sampling-error corrected monthly means are partly due to statistical measurement errors and residual sampling errors, and partly due to systematic differences caused by the input data and/or the processing. There may also be systematic sampling effects that are not accounted for by the sampling-error correction procedure.

The bending angle plots (Figs. 12a-c) in Section 4.3.1 are dominated by the quasi-random measurement errors and residual sampling errors. The systematic differences that stand out in the figures are:

- Systematic biases at low- and mid-latitudes below about 8 kilometer. This is related to the steep refractivity gradient caused by water vapour, which explains the latitudinal dependency.
- Seasonally changing biases at high latitudes between about 25 and 40 kilometer.
- The Metop-COSMIC differences at high- and mid-latitudes between 8 and 25 kilometer exhibit an abrupt change in first half of 2013 (Figure 12c). This is due to the Metop-A and -B firmware updates in 2013.
- The Metop-COSMIC differences above about 30-40 kilometer exhibit a pronounced north-south asymmetry (Figure 12c). This can be traced to differences between the input Level 1A data obtained from EUMETSAT and UCAR.

The refractivity plots (Figs. 13a-c) and the dry temperature plots (Figs. 14a-c) confirm the above findings. The Metop-COSMIC north-south asymmetry starts at lower altitudes, compared to the bending-angle plots, but are otherwise similar. In addition, the CHAMP-COSMIC refractivity plots at low latitudes show another type of systematic bias:

- The CHAMP-COSMIC refractivity differences at low- and mid-latitudes show an oscillating pattern with a period of about 4 months (Figure 13a). The oscillations are also found in the bending-angle (Figure 12a) and dry-temperature (Figure 14a) plots, although a bit weaker. This phenomenon is believed to be a sampling effect related to the orbital drift of the CHAMP satellite.

The spatial distributions of average biases between the missions, shown in Section 4.3.3, confirm the above findings. There is a tendency to systematic latitudinal variations in inter-mission biases. This tendency is most pronounced in the Metop-COSMIC differences, which exhibit a north-south asymmetry above about 20-40 kilometer, while the GRACE-COSMIC differences are more symmetric. We note that this Metop-COSMIC asymmetry is somewhat similar to the Metop-ERAI asymmetry found in Figure 11a, although of a different magnitude. The asymmetry is also seen in the Metop-COSMIC time series plots in Section 4.2.3 (Fig. 10c, upper panels). It should, however, be noted that even if the north-south asymmetry is consistent, it is still small, corresponding to less than 0.5 K at an altitude of 40 kilometer.

The time series plots in Figs. 17 and 18 demonstrate the consistency of the sampling-error corrected anomaly time series based on CHAMP, COSMIC, and Metop data. Anomaly time series of this type are used to study climate trends, and any biases between RO missions may have consequences for the ability to quantify the trends.

Figs. 17 and 18 show:

- There is a high degree of consistency between the monthly-mean bending angle, refractivity, and dry temperature retrieved from different RO missions in the height interval 8 to 30–40 km. This confirms that within certain limits, RO data can be used without inter-calibration between satellites and missions. In combination with the results in Section 4.2.3, it provides a strong indication that it is possible to use CHAMP to extend the RO time series back to 2001 or 2002. If 'consistency' is defined as the maximum difference between an RO mission and a cross-mission mean, then its variation with altitude is roughly:

Bending angle:	0.05 % (8-30 km),	0.10 % (30-40 km),	0.20 % (40-50 km)
Refractivity:	0.05 % (8-30 km),	0.10 % (30-40 km),	0.30 % (40-50 km)
Dry temperature:	0.10 K (8-25 km),	0.25 K (25-30 km),	0.35 K (30-35 km)

- The above differences are small, but they may still have an impact on trends computed from time series combining data from several RO missions. These limitations will be investigated in a limited follow-on validation study.
- Below 8 km, at low- and mid-latitudes, there are biases between the missions. The COSMIC-Metop biases in bending angle are nearly constant over the 10 year overlap period, whereas in refractivity there is a time variation in the COSMIC-Metop biases.
- Below 8 km, at low- and mid-latitudes, there are several abrupt biases changes in the CHAMP data: in the first half of 2002 there is a 1.0-2.0 % negative bending-angle bias change, and in mid-2006 there is another 0.5-1.0 % positive bias change.
- The inter-mission biases, and the large bias shifts for CHAMP data, below 8 km constitute serious problems for the generation of longer multi-mission time series of data, particularly at low- and mid-latitudes.

We conclude that there are systematic differences between the RO missions that should be addressed (seasonal variations at high altitude, a north-south asymmetry in Metop-COSMIC differences), but that most of these are of a magnitude small enough to still allow for the construction of multi-mission time series of data. The exceptions to this are the differences above 40 km, or below 8 km at low- and mid-latitudes, where in particular the differences between CHAMP and the other missions can be large enough to potentially cause problems for the construction of a multi-mission dataset. Below 8 km, there are also problems with time-varying biases in the CHAMP data itself, with implications for longer time series of data.

4.4 Validation of the sampling error correction method

The sampling error is estimated by sub-sampling in the ERA-Interim reanalysis field at the same times and locations as the observations, and computing the difference between the sub-sampled mean for a bin and the corresponding mean constructed from the full reanalysis field [RD.6]. The sampling error of the observed mean is thus assumed to be the same as the sampling error in the model, and the observed mean is "corrected" simply by subtracting the estimated sampling error.

We now use the mission overlap periods, defined in Section 4.3, to investigate how the mission differences are affected by the sampling-error correction. For each pair of RO missions, we compute the differences

$$\Delta\bar{X}_{A-B} = \bar{X}_A^{obs} - \bar{X}_B^{obs} \tag{5}$$

or, alternatively, the fractional differences

$$\Delta\bar{X}_{A-B} = 100 \cdot \left(\bar{X}_A^{obs} - \bar{X}_B^{obs}\right)/\bar{X}_B^{obs} \tag{6}$$

where $\bar{X}_A^{obs}$ and $\bar{X}_B^{obs}$ are the observed monthly means for missions A and B, respectively. We do this both for the original, non-adjusted, means, and for the sampling-error corrected means.

In the absence of sampling errors (and any other systematic errors), the monthly means derived from two different RO missions would only differ by the statistical measurement errors. The mission differences would be more or less quasi-random. The same would be true if the sampling-error estimate was a perfect description of the true sampling error.

Figure 19 gives an example of the results. The effect of the sampling-error correction is to substantially reduce the differences between the missions. The upper and middle panels in Figure 19 show that systematic features in the difference plots, such as the pronounced vertical stripes, have either disappeared or have been substantially decreased in magnitude. The differences between the missions have effectively become more quasi-random. However, systematic biases not due to sampling effects, such as those present in the lowest few kilometres, are not affected by the sampling error correction.

The lower panels of Figure 19 show the overall reduction of the differences between the missions, in the 10-30 kilometer altitude region. The width of the difference distribution is substantially reduced. The narrow purple histograms are now caused by the statistical measurement errors, systematic errors, and around 20-30 % of the sampling errors that we expect to remain, presumably as random errors.

We conclude that the sampling-error correction method is quite effective in reducing the sampling errors, decreasing systematic inter-mission biases and making the differences more quasi-random. This conclusion is based on many plots like Figure 19, and on similar results from other RO missions and geophysical variables.

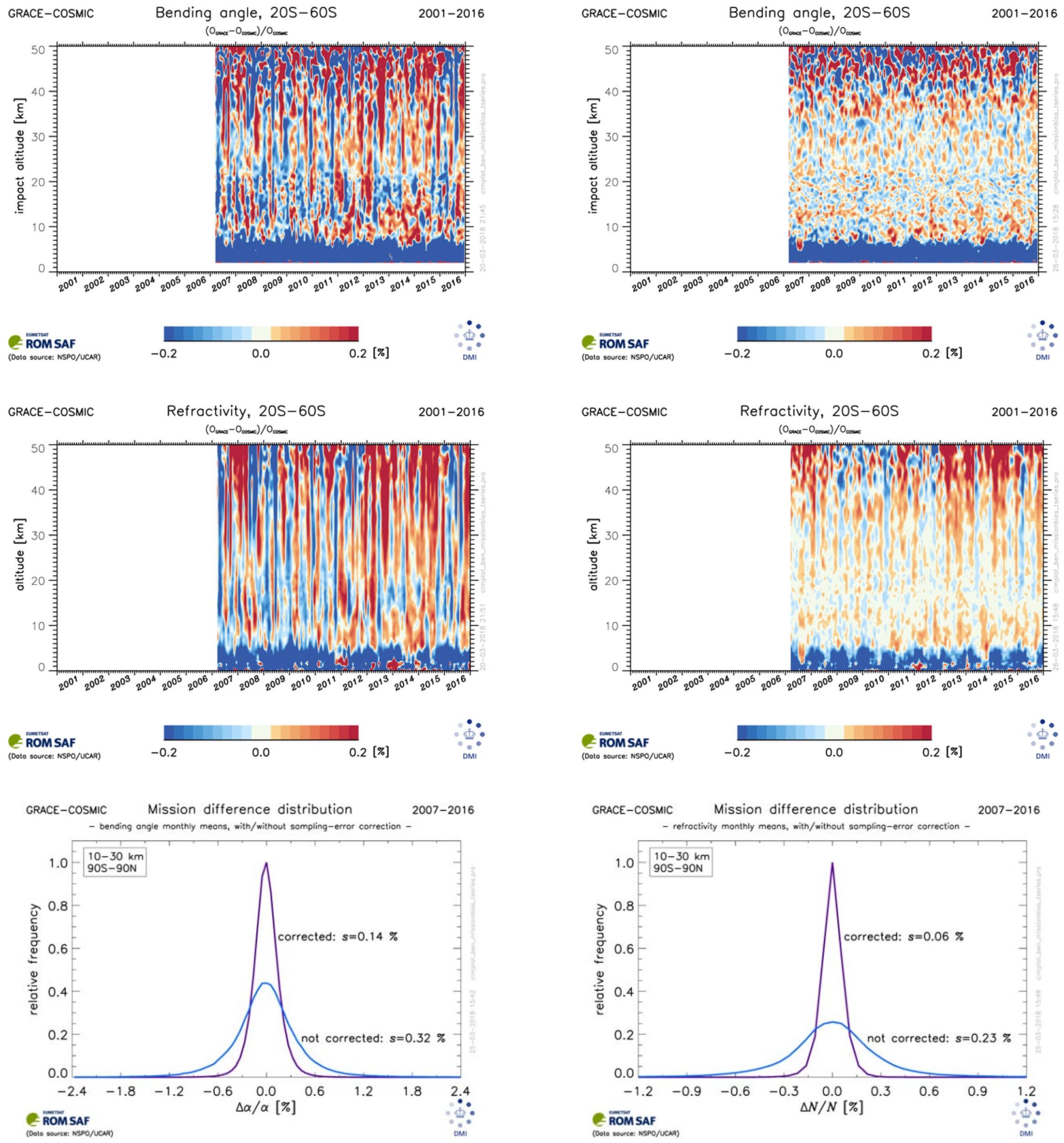


Figure 19: *Example of the impact of sampling error correction. Upper panels: GRACE-COSMIC bending angle difference data with and without sampling error correction. Middle panels: same but for refractivity. Lower panels: distribution of GRACE-COSMIC monthly mean differences, with and without sampling-error correction. The variable s in the figures is a robust measure of the spread: the normalized median absolute deviation (numerically identical to the standard deviation for a gaussian distribution).*

5. Comparison with product requirements

The formal requirements for the ROM SAF data products are stated in the PRD [*AD.2*]. There are three sets of accuracy requirements (*threshold*, *target*, and *optimal*), defined separately for the RO missions. The requirements are defined as functions of height (except for the tropopause height requirements). In the present report, the PRD accuracy requirements are colour coded: orange for *threshold*, yellow for *target*, and green for *optimal*. Data that do not reach the threshold are coded with *red* colour.

For the present comparisons with the PRD accuracy requirements, we aggregate the monthly mean data in two different ways:

i. Averaging of monthly mean O-B within 6 latitude bands (Section 5.1). These plots are similar to the plots presented in Section 4.2.2, but are here colour coded according to the PRD accuracies.

ii. Definition of 'typical' monthly |O-B| data within 9 broad latitude-height regions: low/mid/high latitudes and low/mid/high altitudes (Section 5.2). This also forms the basis for the definition of Service Specifications.

We define 9 broad latitude-height regions as follows: tropics (30˚S–30˚N), mid-latitudes (30˚N–60˚N and 30˚S–60˚S), and polar (60˚N–90˚N and 60˚S–90˚S), and low (0-8 km), middle (8-20 km), and high (20-50 km) altitudes.

Each of these 9 regions includes several hundred monthly values (though not independent, the data are more or less strongly correlated). For each observed monthly-mean value, we compute the absolute deviation from ERA-Interim, |O-B|, and determine whether it is smaller than the threshold, target or optimal accuracies. We then determine the PRD compliance for the latitude-height region by requiring that at least 60% of the monthly-mean data values within that region reach the corresponding accuracy.

Hence, the *formal compliance with the PRD requirements is based on the 60% percentiles of the absolute deviation from ERA-Interim, within 9 broad latitude-height regions*. This quantity is also used in the definition of the Service Specifications [AD.4],

In Section 5.1, we present plots of O-B within 6 latitude bands for the Metop and COSMIC missions and for bending angle, refractivity, dry temperature, temperature, and humidity, colour coded according to the accuracies stated in the PRD document.

In Section 5.2, we present plots of the formal compliance with the PRD requirements (based on the 60% percentile of |O-B| within 9 latitude-height regions) for Metop, COSMIC, CHAMP, GRACE, and the multi-mission data record, for all eight geophysical variables.

In Section 5.3, the results shown in the plots are discussed.

5.1 Figures: O-B time-height plots in terms of PRD accuracies

Metop – ERAI
Bending angle

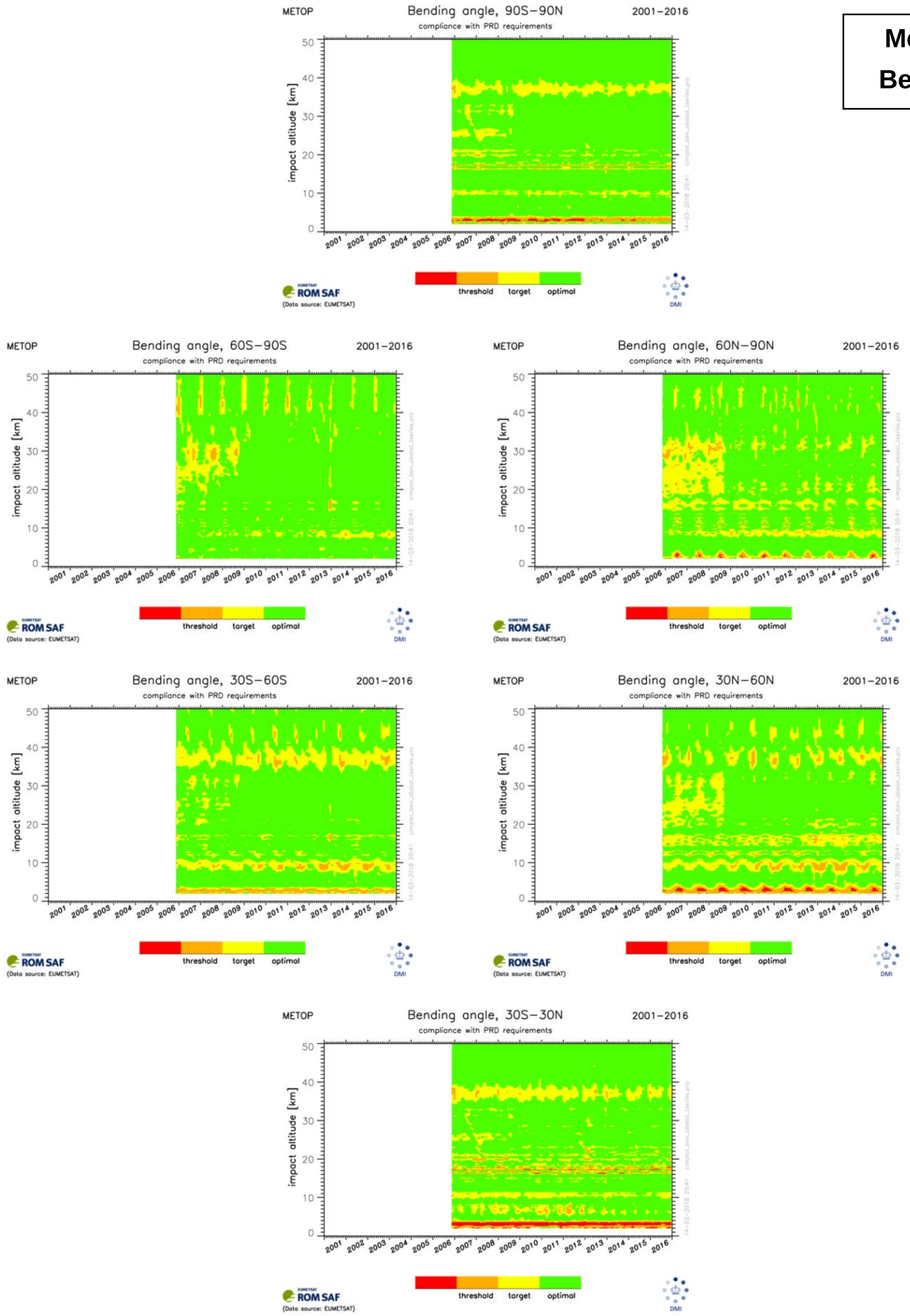


Figure 20a. *Metop bending angle O-B data. Similar to the plots in Figure 8a, but colour coded according to the PRD accuracies.*

Metop – ERAI
Refractivity

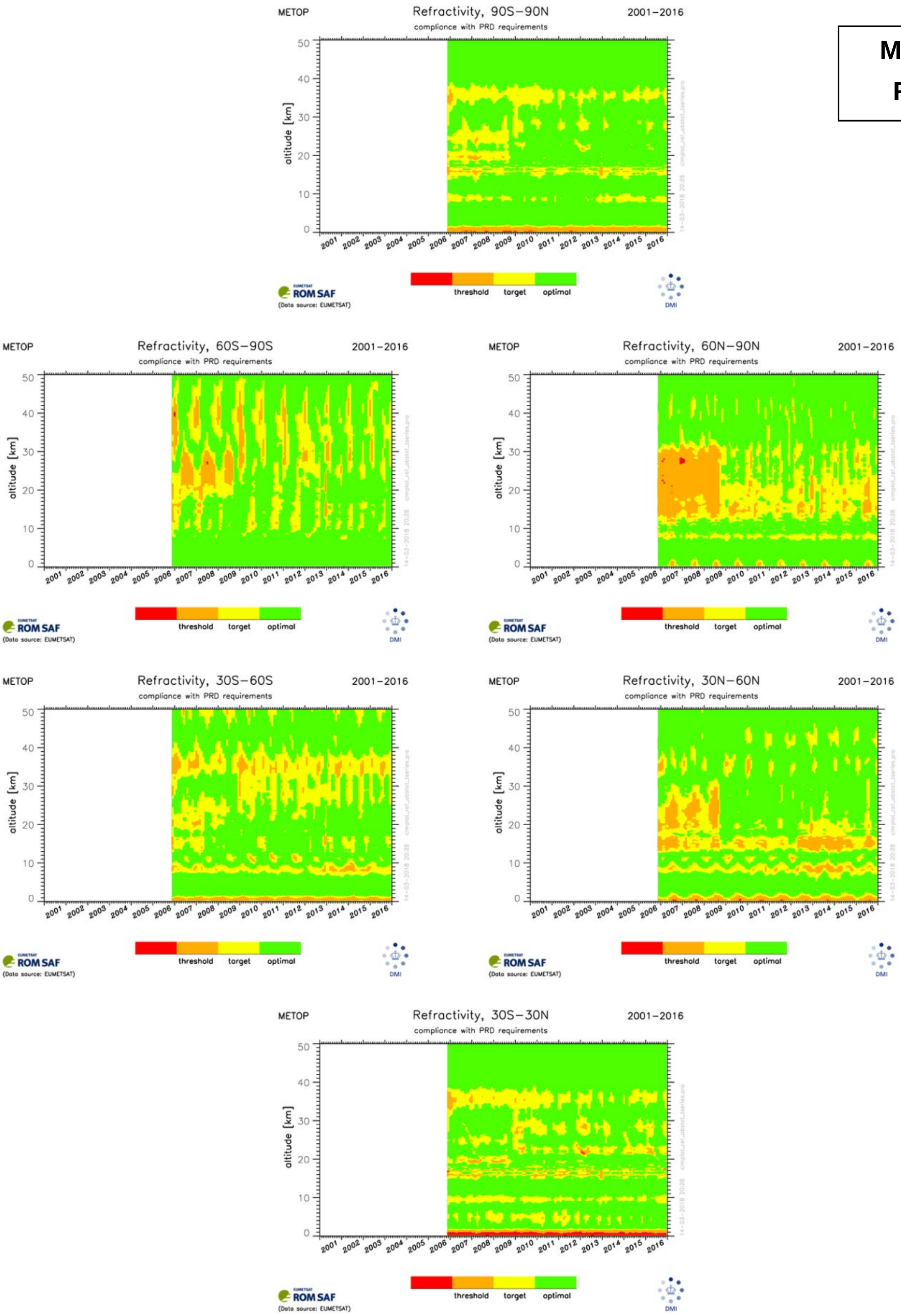


Figure 20b. *Metop refractivity O-B data. Similar to the plots in Figure 8b, but colour coded according to the PRD accuracies.*

Metop – ERAI
Dry temperature

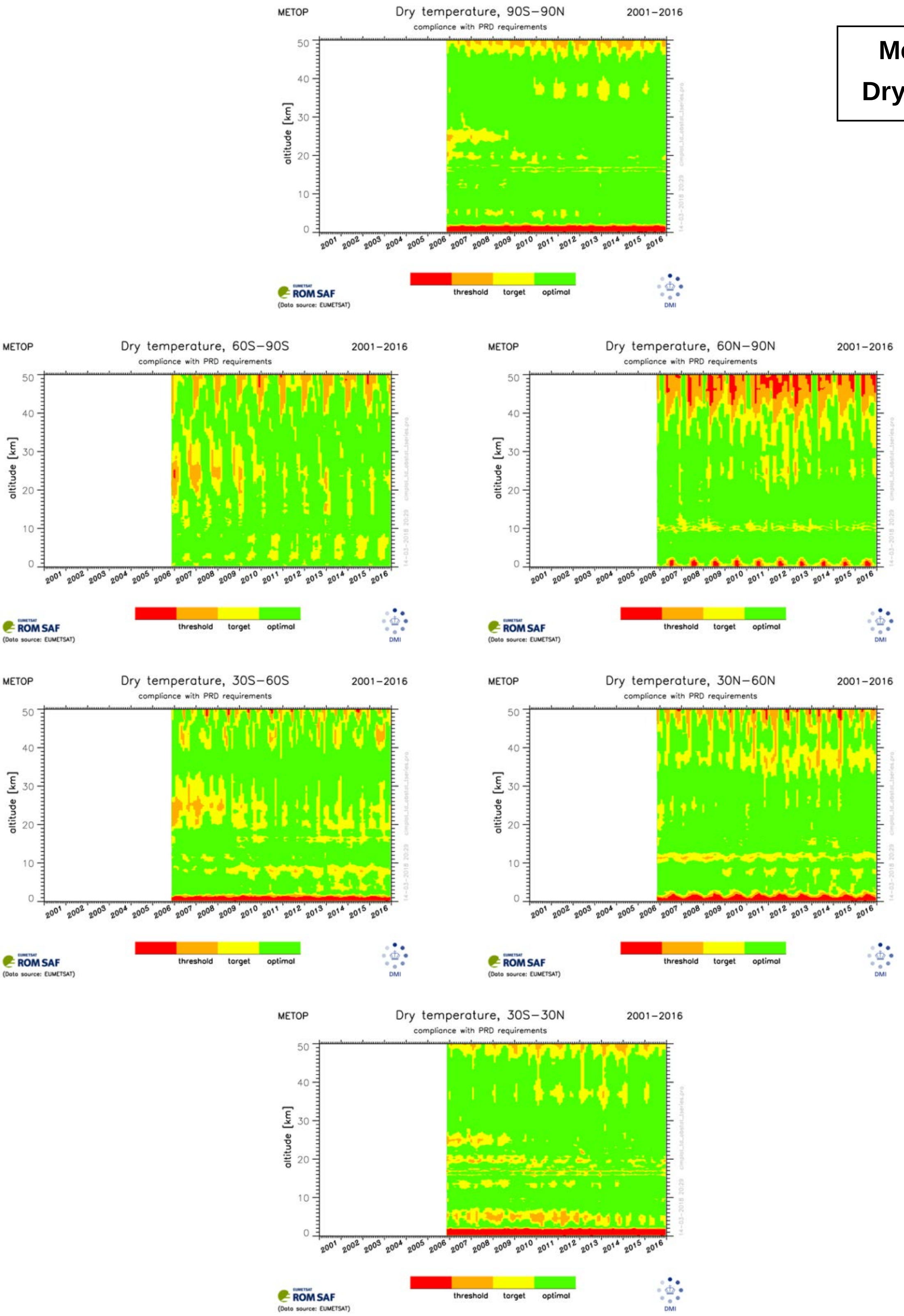


Figure 20c. *Metop dry temperature O-B data. Similar to the plots in Figure 8c, but colour coded according to the PRD accuracies.*

Metop – ERAI Temperature

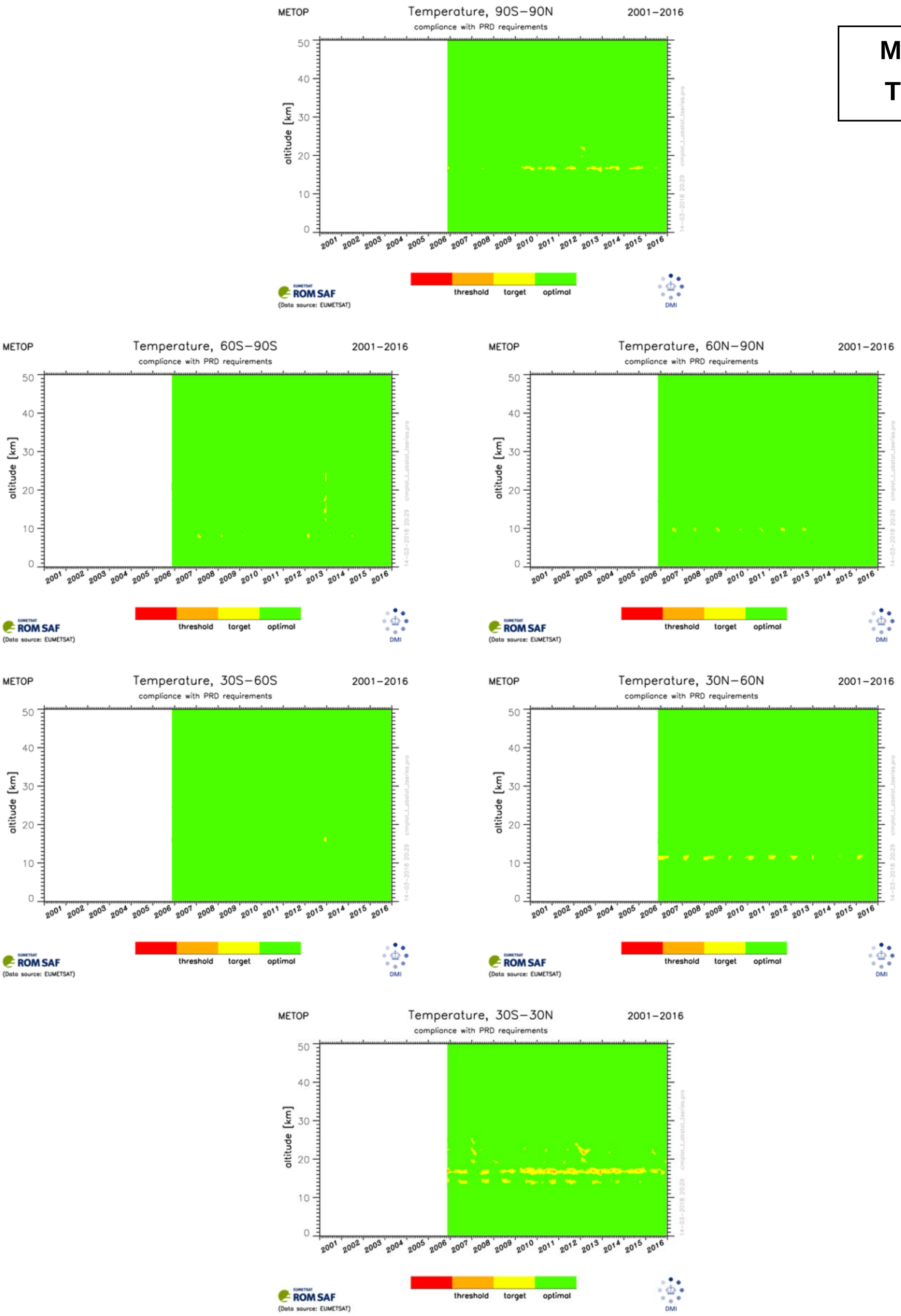


Figure 20d. *Metop temperature O-B data. Similar to the plots in Figure 8d, but colour coded according to the PRD accuracies.*

Metop – ERAI
Humidity

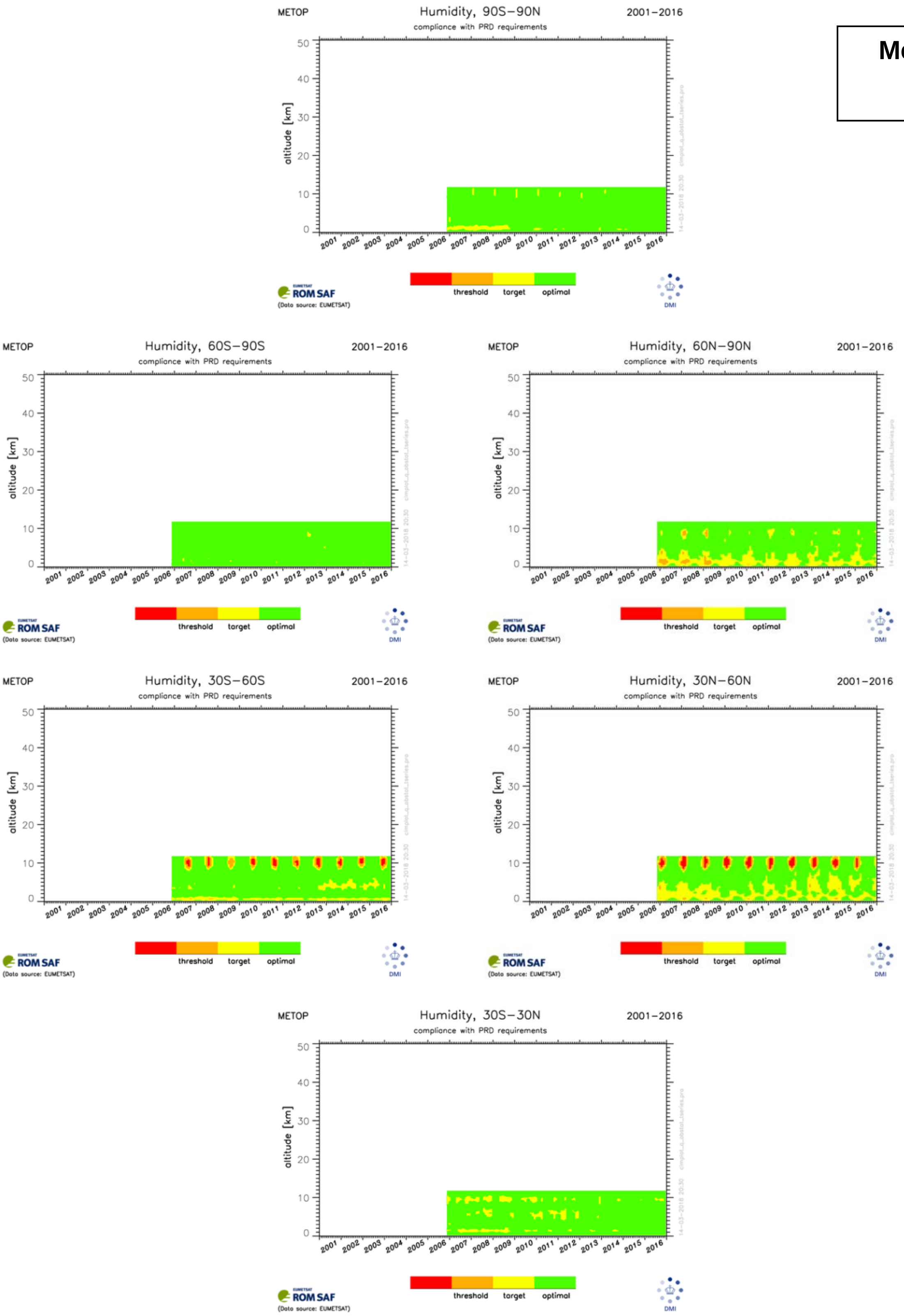


Figure 20e. *Metop humidity O-B data. Similar to the plots in Figure 8e, but colour coded according to the PRD accuracies.*

COSMIC – ERAI
Bending angle

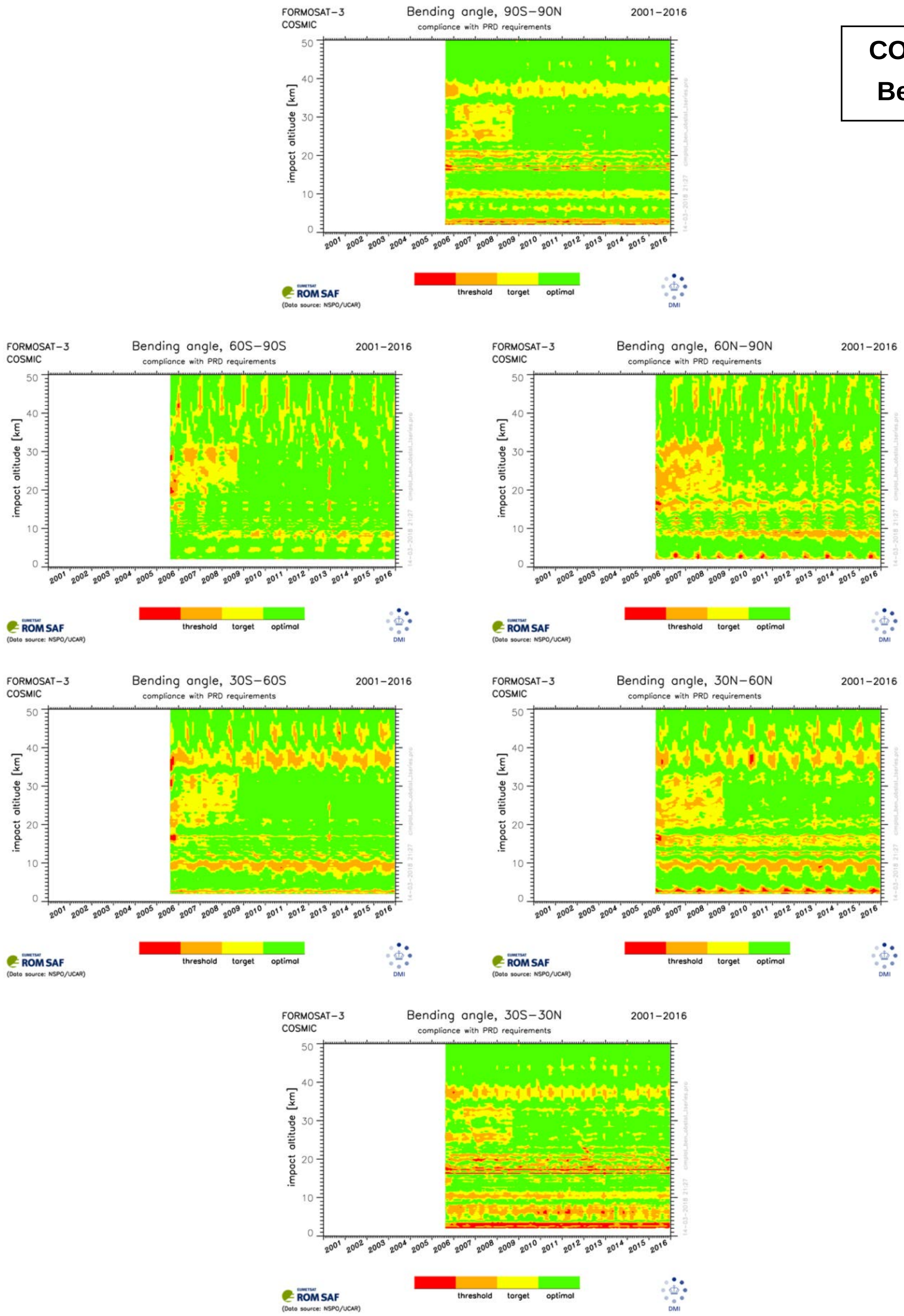


Figure 21a. *COSMIC bending angle O-B data. Similar to the plots in Figure 9a, but colour coded according to the PRD accuracies.*

COSMIC – ERAI
Refractivity

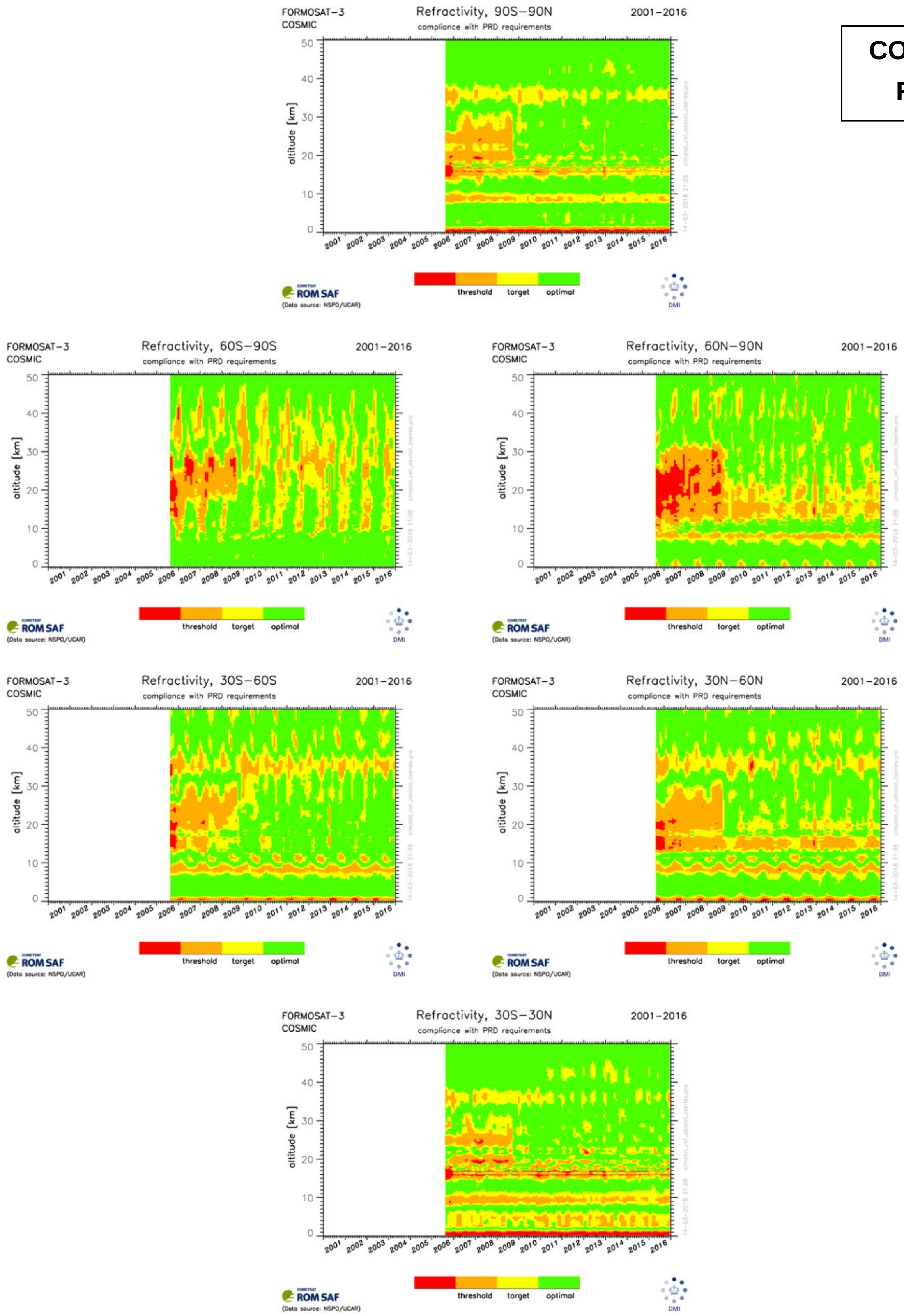


Figure 21b. *COSMIC refractivity O-B data. Similar to the plots in Figure 9b, but colour coded according to the PRD accuracies.*

COSMIC – ERAI
Dry temperature

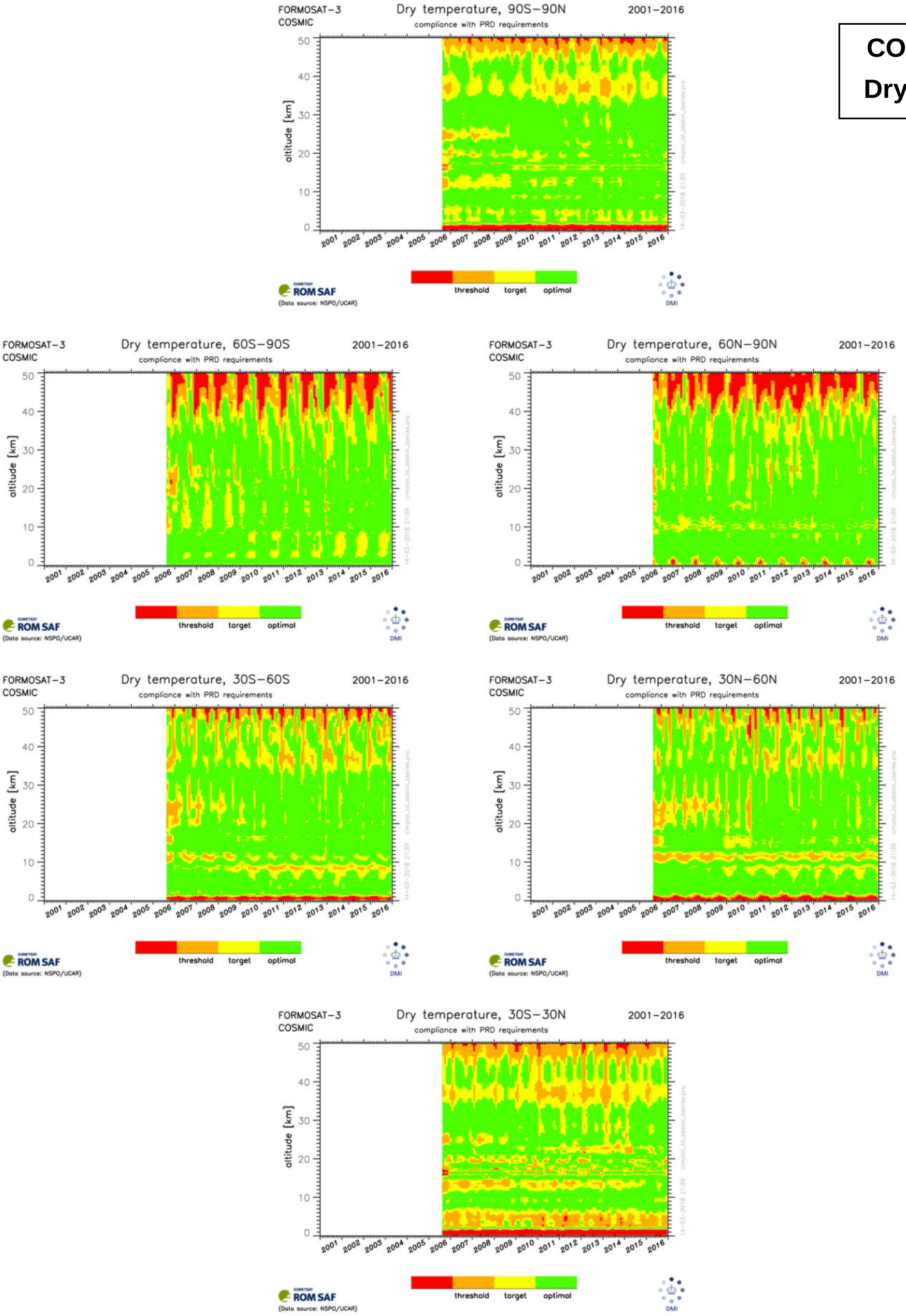


Figure 21c. *COSMIC dry temperature O-B data. Similar to the plots in Figure 9c, but colour coded according to the PRD accuracies.*

COSMIC – ERAI
Temperature

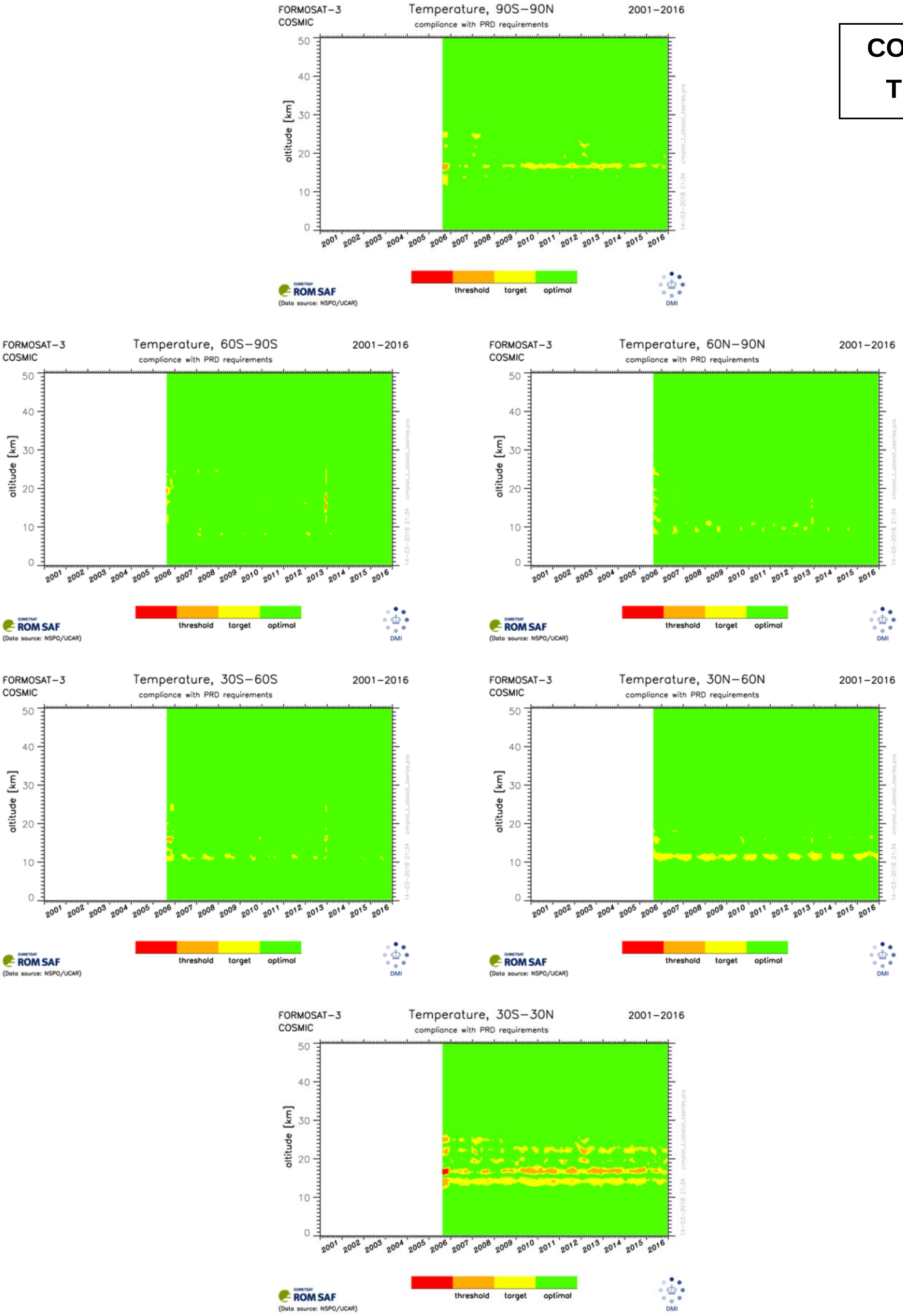


Figure 21d. *COSMIC temperature O-B data. Similar to the plots in Figure 9d, but colour coded according to the PRD accuracies.*

COSMIC – ERAI
Humidity

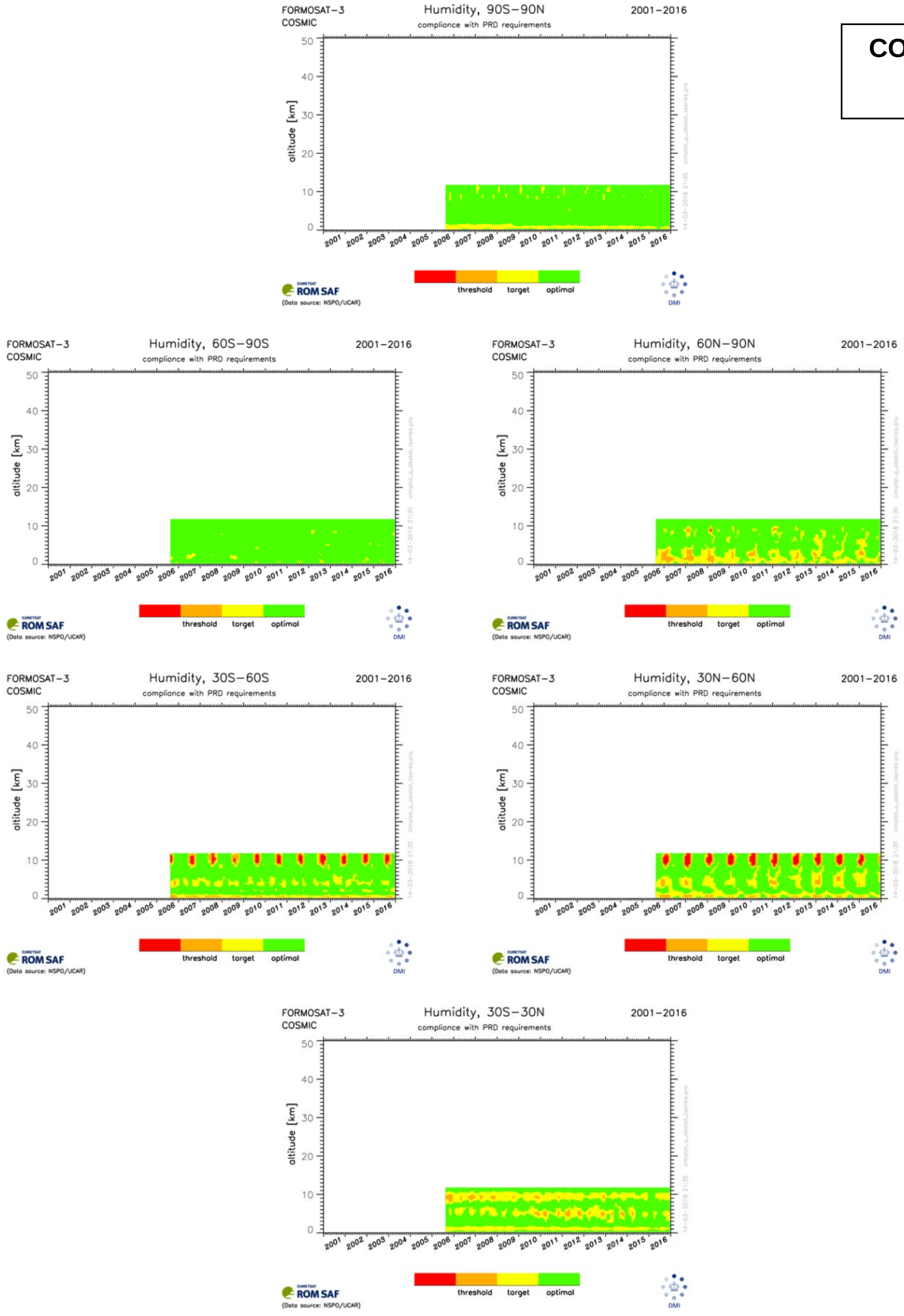


Figure 21e. *COSMIC humidity O-B data. Similar to the plots in Figure 9e, but colour coded according to the PRD accuracies.*

5.2 Figures: Compliance with PRD requirements

Metop

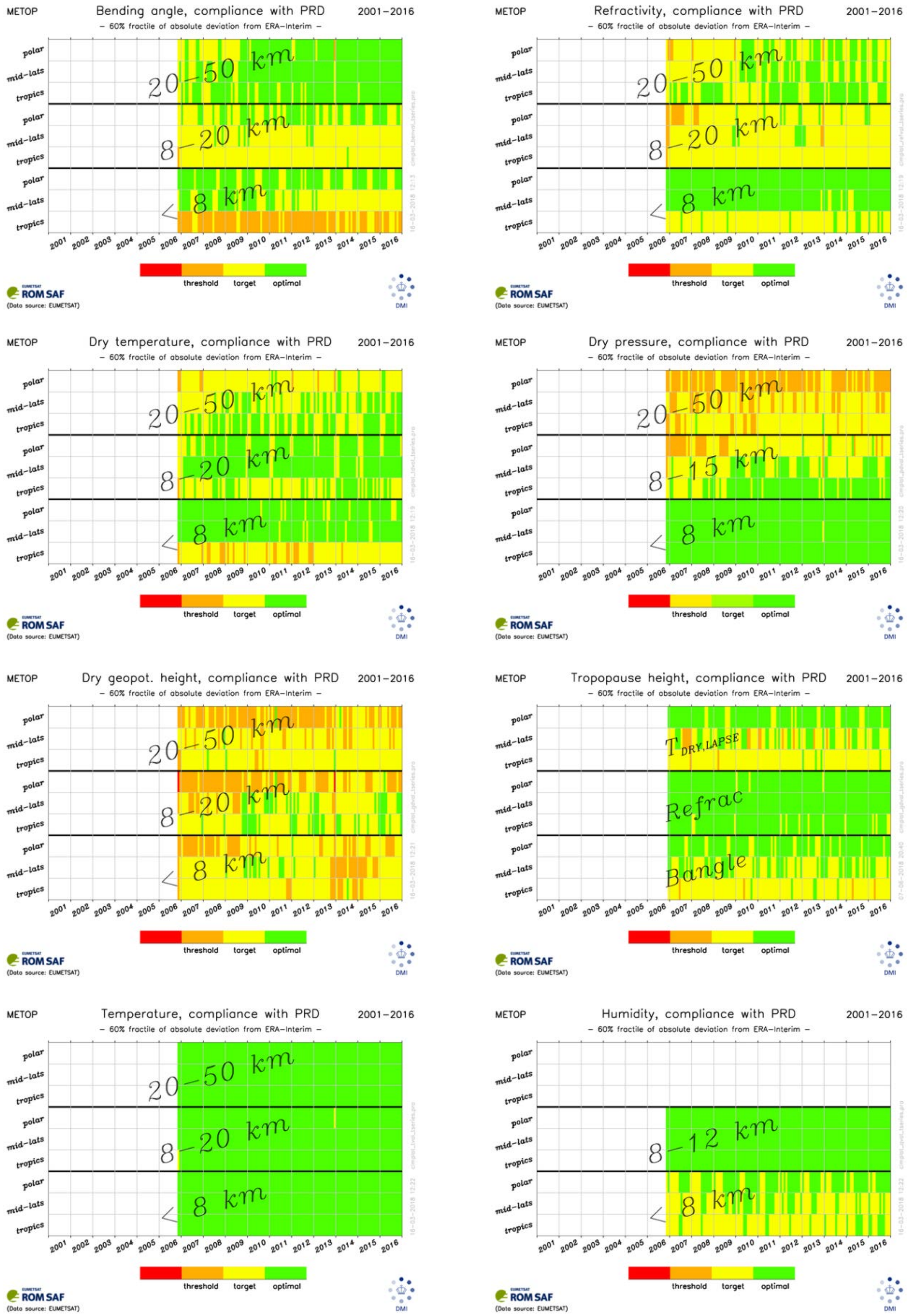


Figure 22a. *Compliance with the PRD requirements for monthly mean Metop data, based on the 60% percentile of the |O-B| distribution within 9 broad latitude-height regions. Red colour indicates non-compliance (see Section 5.3 for a discussion).*

COSMIC

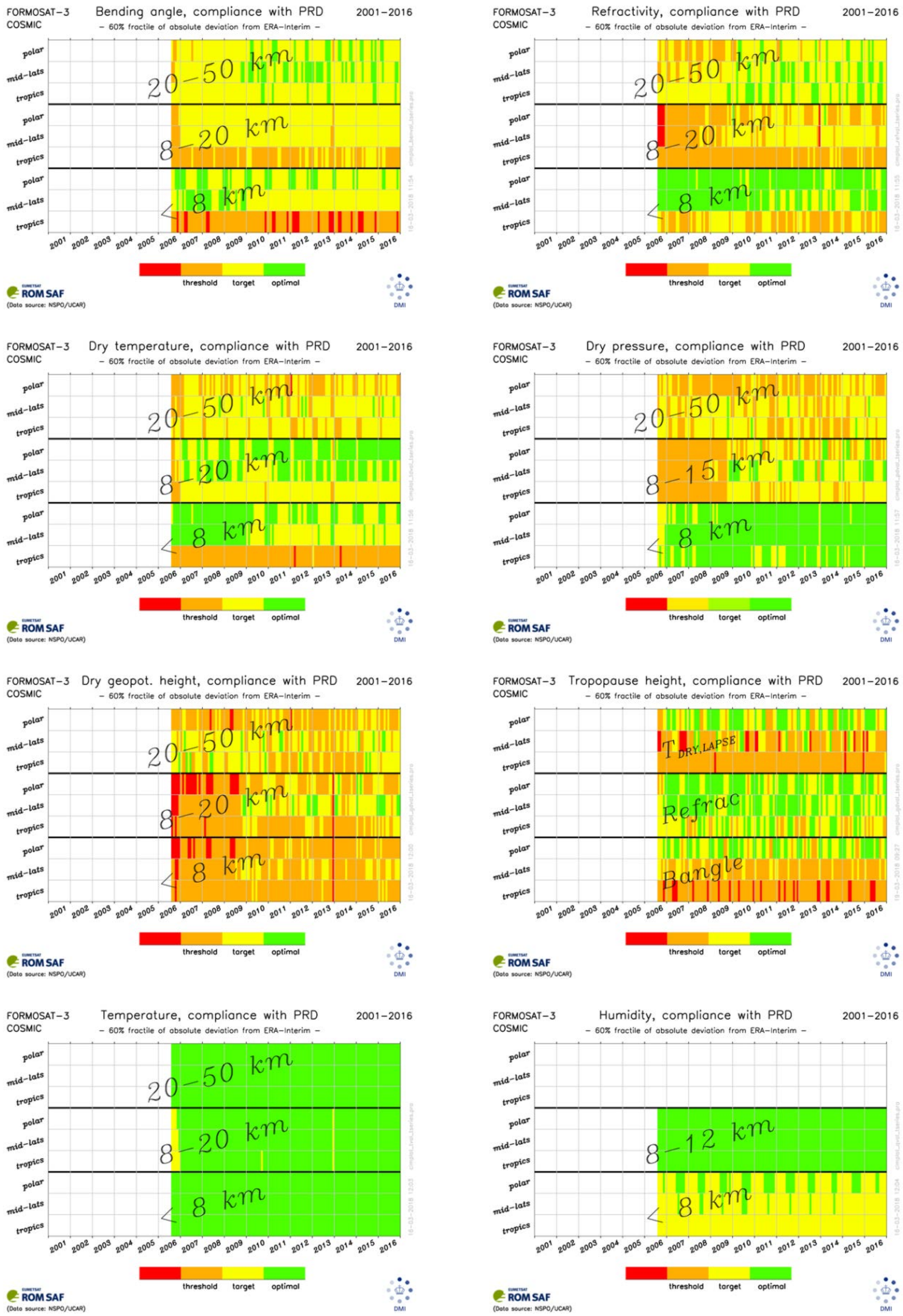


Figure 22b. *Compliance with the PRD requirements for monthly mean COSMIC data, based on the 60% percentile of the |O-B| distribution within 9 broad latitude-height regions. Red colour indicates non-compliance (see Section 5.3 for a discussion).*

CHAMP

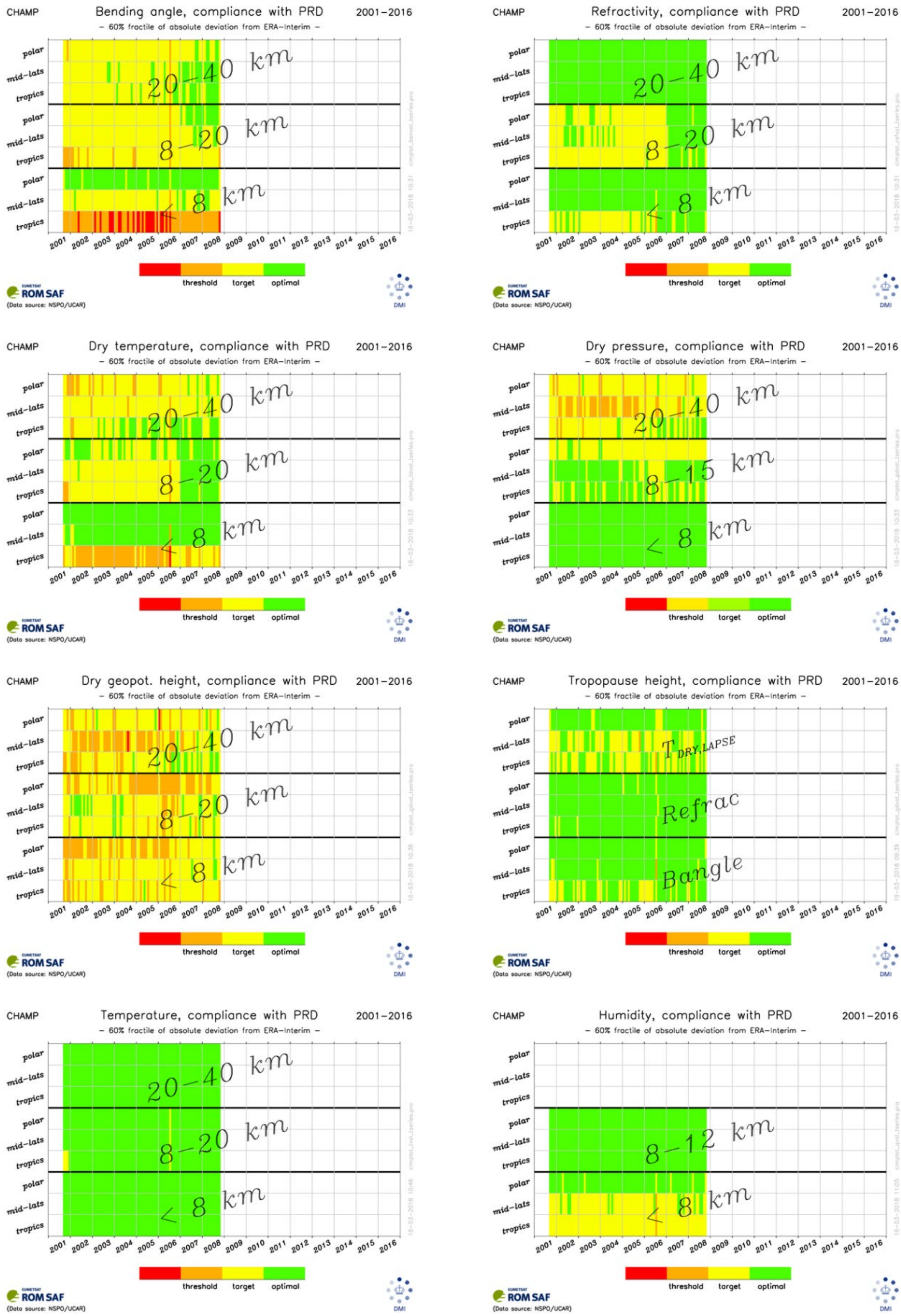


Figure 22c. *Compliance with the PRD requirements for monthly mean CHAMP data, based on the 60% percentile of the |O-B| distribution within 9 broad latitude-height regions. Red colour indicates non-compliance (see Section 5.3 for a discussion).*

GRACE

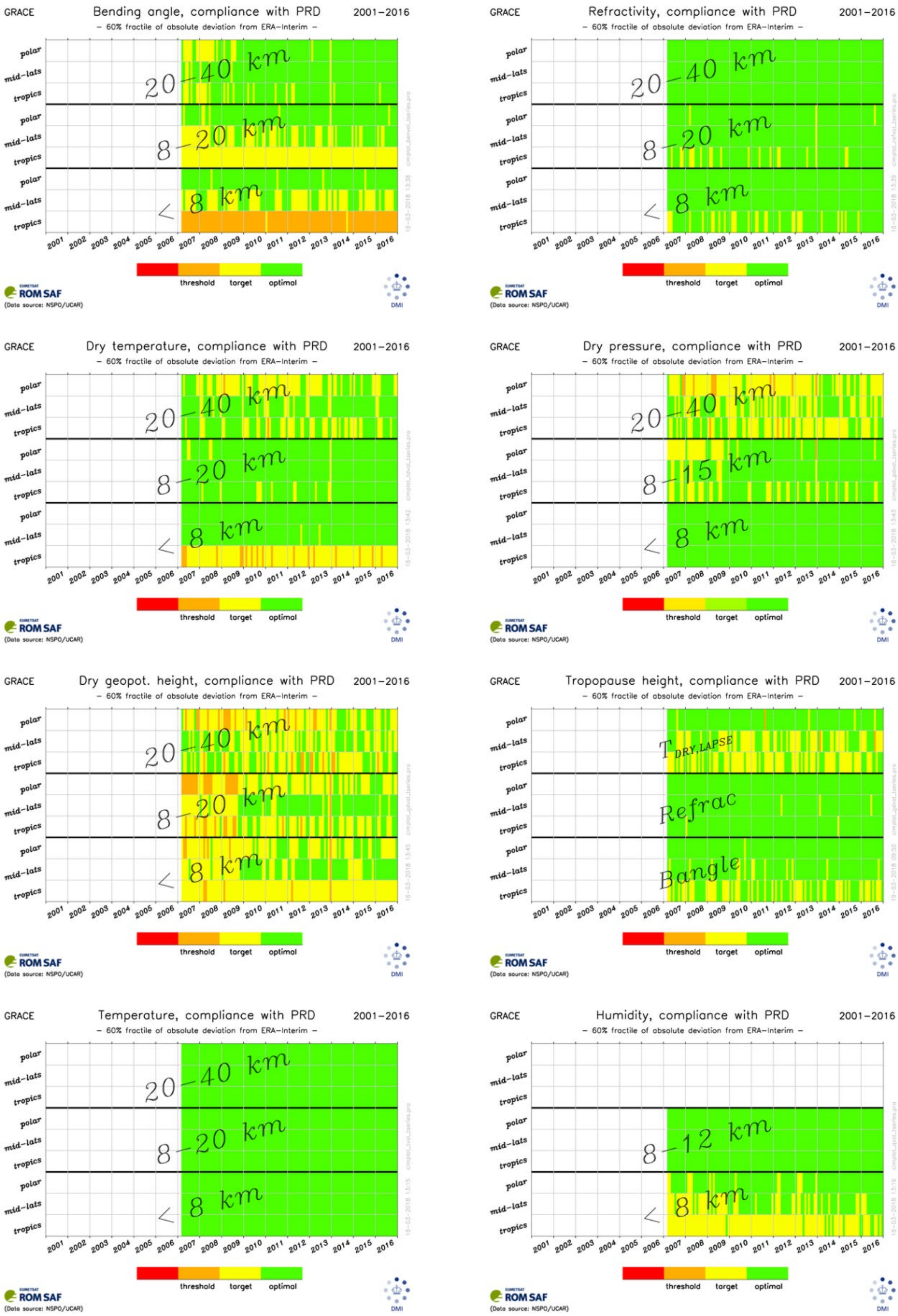


Figure 22d. *Compliance with the PRD requirements for monthly mean GRACE data, based on the 60% percentile of the |O-B| distribution within 9 broad latitude-height regions. Red colour indicates non-compliance (see Section 5.3 for a discussion).*

MULTI

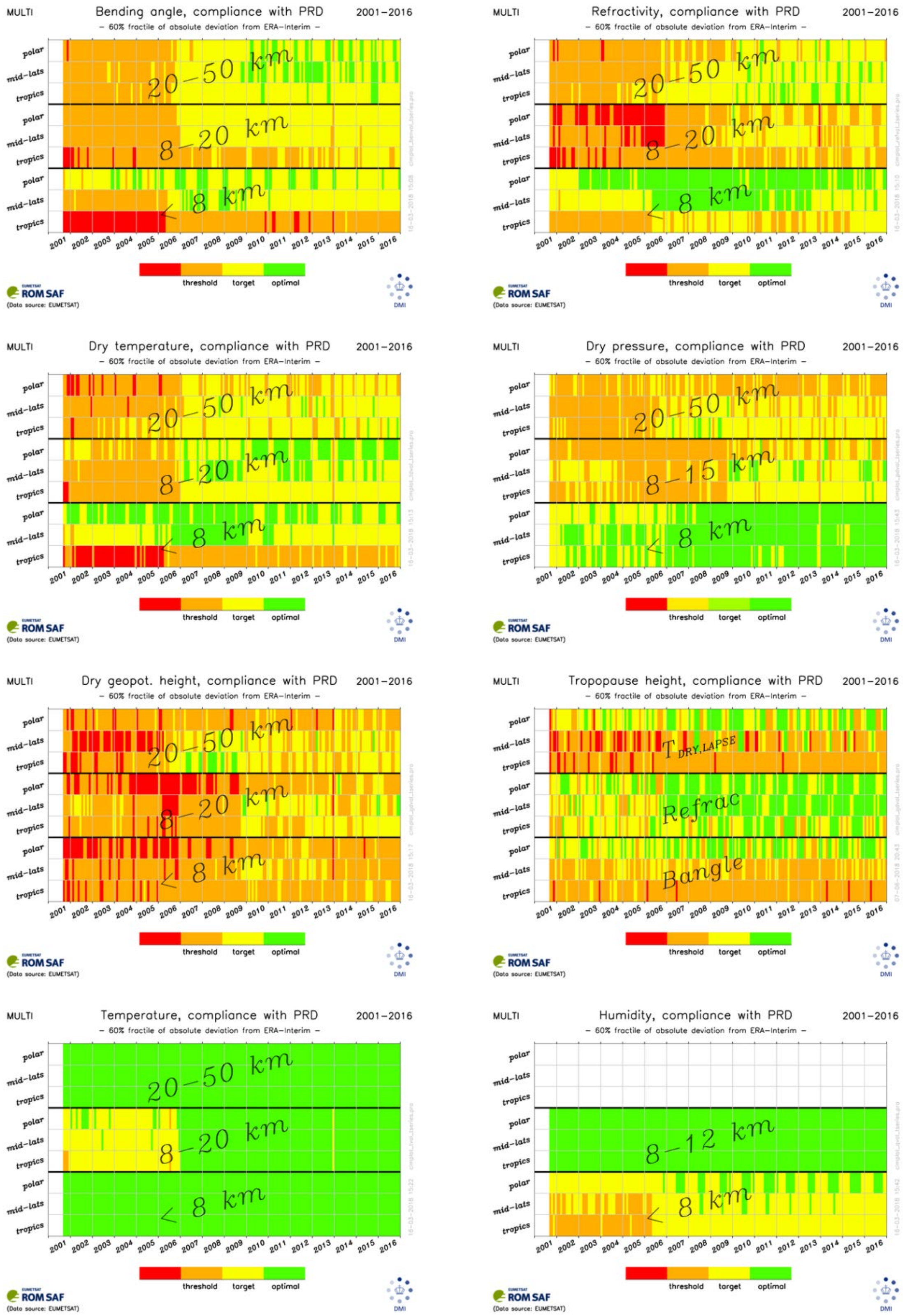


Figure 22e. *Compliance with the PRD requirements for monthly mean multi-mission data, based on the 60% percentile of the |O-B| distribution within 9 broad latitude-height regions. Red colour indicates non-compliance (see Section 5.3 for a discussion).*

5.3 Discussion and conclusions

The Product Requirements [RD.3] express the commitment of the ROM SAF team for the development of data products. As described in the beginning of Section 5, the compliance of the data products with the PRD accuracies are checked from the statistics of monthly-mean data within 9 broad latitude-height regions.

Section 5.1 shows the RO deviations from ERA-Interim in terms of the PRD accuracies. Comparison with the corresponding figures in Section 4.2.2 provides a context for the accuracy requirements. We find that the stepwise change in 2009, due to ERA-Interim (discussed in Section 4.2.5), has a rather large magnitude relative to the PRD accuracy requirements. Also other features, e.g., the persistent bias between 35 and 40 kilometer and seasonal oscillations in O-B, are significant.

One practical conclusion is that at least some of the deviations from the optimal accuracies relative to the PRD requirements are most likely caused by biases in ERA-Interim, and are not due to the observed RO data. The degree of compliance with the PRD requirements partly depends on the choice of reference data set.

Section 5.2 shows the formal compliance with the PRD accuracy requirements, for all RO missions and for all eight geophysical variables. The compliance is checked for nine latitude-height regions (three latitude regions for the tropopause height), summarized in the plots presented in Figs. 22a-e – one plot for each mission and geophysical variable.

The Metop data are all compliant with the PRD requirements (no red cells in Figure 22a). The lowest degree of compliance is found for dry pressure and temperature at high altitudes and high latitudes, and at low altitudes in the tropics. These are regions where biases are to be expected. A certain degree of seasonality in the compliance is also observed, particularly at high altitudes, but also for humidity in the troposphere. COSMIC exhibits a similar pattern, although with a lower degree of compliance. For COSMIC, we find non-compliant data (red cells in Figure 22b). It should be noted that the non-compliance in November 2013 is caused by an error in ERA-Interim, which provides an example of the dependence on the choice of reference data.

CHAMP and GRACE exhibit a higher degree of compliance than COSMIC. The requirements on these missions are also set lower than the requirements on Metop and COSMIC, such that they are easier to fulfill. The requirements on the MULTI data set, on the other hand, is roughly similar to the requirements on Metop and COSMIC. Hence, the early period of the MULTI data, which consist solely of CHAMP data, contains many non-compliances (red cells in Figure 22e), which is to be expected.

We conclude that the ROM SAF Level 3 gridded monthly-mean data products, addressed by this validation study, are formally compliant with the PRD accuracy requirements. The main non-compliances are found for the COSMIC data (2006-2008 period and at low altitudes) and for the early part of the MULTI data set, due to the reliance on CHAMP data.

6. Service specifications

The Service Specifications describe the commitments by the ROM SAF related to the services and products provided to the users. These commitments include a set of operational accuracy targets that should be met by the Level 3 gridded data products, and which should be regularly monitored and documented as a part of normal operations. Hence, they are relevant for the operational generation of ICDRs and offline data, as well as for the validation of the CDRs generated as a reprocessing activity. Even though these targets should be consistent with the PRD [*AD.3*], they are not necessarily identical to the PRD requirements.

The accuracies proposed to be included in the service specifications for the Level 3 gridded data products are listed in Table 5, Table A2, and Table A4. The methods used for comparing RO data with the service specifications are identical to the methods defined in the PRD. The service specification accuracies agree closely to the *target accuracies* in the PRD requirements, with some exceptions.

We monitor the compliance to the service specification by comparing a *certain percentile of the absolute deviation from the reference data* with the accuracies stated in the Service Specifications Document. Hence, the procedure is similar to the validation against the PRD described in Section 5. The outcome of the regular monitoring against the service specifications is provided on the ROM SAF web page.

Table 5. *Proposed ROM SAF Level 3 Service Specifications for the CDRs (reprocessed data products). The accuracies are stated separately in three height layers; below 8 km, 8-25 km, and 25–50 km (for humidity only up to 12 km, and for the CHAMP and GRACE missions only up to 40 km). Where both absolute and relative numbers are given, the requirement is given by the greater of these two.*

GRM-28-L3 (MULTI) GRM-29-L3 (Metop) GRM-30-L3 (COSMIC)	**GMR-33-L3 (GRACE)**	**GRM-32-L3 (CHAMP)**
Bending angle		
25–50 km: 0.2 % or 0.4 μrad[1] 8–25 km: 0.2 % 0– 8 km: 2.0 – 0.2 %	25 – 40 km: 0.3 % or 0.6 μrad[1] 8 – 25 km: 0.3 % 0 – 8 km: 3.0– 0.3 %	25 – 40 km: 0.4 % or 0.8 μrad[1] 8 – 25 km: 0.4 % 0 – 8 km: 4.0 – 0.4 %
Refractivity		
25 – 50 km: 0.08 % or 0.004 N-units[1] 8 – 25 km: 0.08 % 0 – 8 km: 0.8 – 0.08 %	25 – 40 km: 0.12 % or 0.006 N-units[1] 8 – 25 km: 0.12 % 0 – 8 km: 1.2 – 0.12 %	25 – 40 km: 0.20 % or 0.008 N-units[1] 8 – 25 km: 0.20 % 0 – 8 km: 2.0 – 0.20 %
Dry temperature		
25 – 50 km: 0.2 – 2.0 K 8 – 25 km: 0.2 K 0 – 8 km: 1.0 – 0.2 K	25 – 40 km: 0.3 – 3.0 K 8 – 25 km: 0.3 K 0 – 8 km: 1.5 – 0.3 K	25 – 40 km: 0.40 – 4.0 K 8 – 25 km: 0.4 K 0 – 8 km: 2.0 – 0.4 K
Dry pressure		
25 – 50 km: 0.08 – 0.40 % 8 – 25 km: 0.08 % 0 – 8 km 0.40 – 0.08 %	25 – 40 km: 0.12 – 0.60 % 8 – 25 km: 0.12 % 0 – 8 km: 0.60 – 0.12 %	25 – 40 km: 0.20 – 1.00 % 8 – 25 km: 0.20 % 0 – 8 km: 1.00 – 0.20 %
Dry geopotential height		
25 – 50 km: 4 – 40 m 8 – 25 km: 4 m 0 - 8 km: 4 m	25 – 40 km: 6 – 60 m 8 – 25 km: 6 m 0 – 8 km: 6 m	25 – 40 km: 8 – 80 m 8 – 25 km: 8 m 0 – 8 km: 8 m
Temperature		
25 – 50 km: 0.2 – 2.0 K 8 – 25 km: 0.2 K 0 – 8 km: 1.0 – 0.2 K	25 – 40 km: 0.3 – 3.0 K 8 – 25 km: 0.3 K 0 – 8 km: 1.0 – 0.3 K	25 – 40 km: 0.4 – 4.0 K 8 – 25 km: 0.4 K 0 – 8 km: 2.0 – 0.4 K
Specific humidity		
8 – 12 km: 3.0 % 0 – 8 km: 3.0 %	8 – 12 km: 4.0 % 0 – 8 km: 4.0 %	8 – 12 km: 6.0 % 0 – 8 km: 6.0 %
Tropopause Height		
100.0 m	200.0 m	300.0 m
[1] Whichever is greater. [2] An accuracy interval means a linearly changing quantity between the two values over the given vertical coordinate.		
Methods for validation		
Nine broad latitude-height regions (tropics, mid-latitudes, high latitudes and low, middle, high altitudes) are defined. The absolute values of the differences between the monthly-mean RO data and the ERA-Interim reanalysis data are computed on the Level 3 grid. Each value is compared to the service specification valid for that altitude. The compliance with the Service Specifications are determined, within each region and for each calendar month, by requiring that 60% of the absolute differences are smaller than the corresponding specification.		

7. Main conclusions

We can now summarize of the findings described and discussed in Section 4:

- There is a good overall agreement between the RO observational data and the ERA-Interim reanalysis data. The main observation minus background (O-B) differences found are: a bias change in late 2009, a bias feature in late 2013 lasting about a month, seasonally varying biases at mid- and high latitudes above 30-35 kilometer, positive bending-angle and refractivity biases at low and mid-latitudes between 35-40 kilometer and at low latitudes around 5 kilometers, negative bending-angle and refractivity biases at low- and mid-latitudes in the lowest few kilometers, weak bias features coincident with QBO-related structures at low latitudes, a persistent humidity bias structure at low latitudes and seasonally varying humidity biases at mid-latitudes.
- There is generally a good agreement between the RO missions above 8 kilometer and below about 35–40 kilometer, where the different RO missions exhibit very similar O-B spatial and temporal characteristics.
- Below 8 kilometer, at low- and mid-latitudes, there are systematic differences between the O-B characteristics of the RO missions: in the 4-8 kilometer altitude range, there is an up to 2% bias in bending angle between CHAMP and the other missions, and up to 0.5% bias in bending angle between Metop and COSMIC.
- The differences between RO missions during mission overlap periods show: systematic biases at mid- and low latitudes below about 8 kilometers, seasonally varying biases at high latitudes between about 25 and 40 kilometer, differences with respect to Metop in first half of 2013 (due to Metop-A and –B firmware updates), Metop-COSMIC north-south asymmetry above about 30-40 kilometer.
- The hemispheric asymmetry between Metop and the other RO missions, clearly visible at high altitudes, seems to be related to the use of Level 1A input data from different centers. It is not an issue with the Metop data itself, or the processing from Level 1A data.
- There is a high degree of consistency between the RO missions in the height interval 8 to 35-40 kilometer (depending on latitude and geophysical variable), as seen in the sampling-error corrected anomaly time series. If ‘consistency’ is defined as the maximum difference between an RO mission and a cross-mission mean, then its variation with altitude is roughly:

Bending angle:	0.05 % (8-30 km),	0.10 % (30-40 km),	0.20 % (40-50 km)
Refractivity:	0.05 % (8-30 km),	0.10 % (30-40 km),	0.30 % (40-50 km)
Dry temperature:	0.10 K (8-25 km),	0.25 K (25-30 km),	0.35 K (30-35 km)

These findings lead up to the ***main conclusion***s:

1. There is a high degree of consistency between the RO data and the corresponding ERA-Interim reanalysis short-term forecast data, from about 5 to 35 kilometers altitude, somewhat depending on RO mission, latitude, and geophysical variable.

Some of the main bias features can be identified as known ERA-Interim bias features.

2. There is a high degree of consistency between the RO missions, particularly between 8 and 35 kilometers altitude, the exact altitude range depending on latitude and geophysical variable. However, there are still small persistent biases between the missions, which may need to be taken into account depending on the application (see 3 and 4 below).

3. The main biases between the RO missions are a) those between CHAMP and the other missions below 8 kilometers, b) near-constant biases between COSMIC and Metop increasing with height and exhibiting an hemispheric asymmetry at high altitudes, and c) an additional near-constant bias between COSMIC and Metop between 10 and 20 kilometers occurring after mid-2013 (due to Metop firmware updates).

4. The inter-mission differences are small, and do not appear to drift in time. For many applications, RO data are homogeneous enough to construct longer time series of RO data between about 8 and 35 kilometers, the exact altitude limits depending on latitude and geophysical variable. However, there are certain applications (e.g. longer-term trend estimations) which may be affected by the inter-mission differences. These effects will be further investigated and quantified in a follow-on ROM SAF validation study.

Annex A. Interim Climate Data Records

As described in the definitions section (Section 1.4), the CDRs have been generated in a dedicated reprocessing activity using the same algorithms throughout the length of the data records, while ICDRs are generated on a regular basis with the same algorithms as the CDRs, but using currently available input data. The main rationale for the ICDRs is that they extend the CDRs until data from a new reprocessing become available. There is a strong focus on the consistency between the ICDRs and the CDRs, and the ICDRs are expected to have very similar characteristics as the corresponding CDRs. The validation described in Sections 4 and 5 of the present document also applies to the corresponding ICDRs.

Table A1. *Gridded Level 3 monthly mean ROM SAF ICDR products. Each entry in the table covers a series of geophysical variables. The asterisk* [*] *in the product acronym is a place holder for one of the following letters: 'B' for bending angle, 'R' for refractivity, 'D' for dry temperature, 'Y' for dry pressure, 'Z' for dry geopotential height, 'T' for temperature, 'H' for humidity, and 'C' for tropopause height.*

Product ID	Data product	Product acronym	Mission	Time coverage	Disse-minat.	Prod. ver.
ICDRs						
GRM-29-L3-I1	Gridded monthly mean ICDR	I[*]GMET	Metop	201701-	web	1.0

Table A2. *Proposed ROM SAF Level 3 Service Specifications for the ICDRs. The requirement on the ICDR is that it should be fully consistent with the CDR. As a consequence, identical Service Specifications are applied.*

GRM-29-L3-I1 (Metop)
Accuracy targets, all geophysical variables
Same as for GRM-29-L3-R1
Methods for validation
Same as for GRM-29-L3-R1

Annex B. Offline Data Products

As described in the definitions section (Section 1.4), the CDRs have been generated in a dedicated reprocessing activity using the same algorithms throughout the length of the data records, while ICDRs are generated on a regular basis with the same algorithms as the CDRs but using currently available input data. The offline data products, on the other hand, are generated based on algorithms that may have evolved somewhat since the last reprocessing, in order to reflect the latest scientific development. The validation described in Sections 4 and 5 of the present document also applies to the corresponding offline data products.

Table A3. *Gridded Level 3 monthly mean ROM SAF offline products. Each entry in the table covers a series of geophysical variables. The asterisk* [*] *in the product acronym is a place holder for one of the following letters: 'B' for bending angle, 'R' for refractivity, 'D' for dry temperature, 'Y' for dry pressure, 'Z' for dry geopotential height, 'T' for temperature, 'H' for humidity, and 'C' for tropopause height.*

Product ID	Data product	Product acronym	Mission	Time coverage	Disse-minat.	Prod. ver.
Offline data						
GRM-93–99, 191	Gridded monthly mean offline data	O[*]GMEA	Metop-A	201701-	HTTP	1.0
GRM-53–59, 192	Gridded monthly mean offline data	O[*]GMEB	Metop-B	201701-	HTTP	1.0
GRM-83–89, 194	Gridded monthly mean offline data	O[*]GMET	Metop	201701-	HTTP	1.0

Table A4. *Proposed ROM SAF Level 3 Service Specifications for the offline data products. The accuracies are stated separately in three height layers; below 8 km, 8-25 km, and 25–50 km (for humidity only up to 12 km). Where both absolute and relative numbers are given, the requirement is given by the greater of these two.*

GRM-93–99, GRM-191 (Metop A) **GRM-53–59, GRM-192 (Metop B)** **GRM-83–89, GRM-194 (Metop)**
Bending angle
25 – 50 km: 0.2 % or 0.4 μrad[1] 8 – 25 km: 0.2 % 0 – 8 km: 2.0 – 0.2 %
Refractivity
25 – 50 km: 0.08 % or 0.004 N-units[1] 8 – 25 km: 0.08 % 0 – 8 km: 0.8 – 0.08 %
Dry temperature
25 – 50 km: 0.2 – 2.0 K 8 – 25 km: 0.2 K 0 – 8 km: 1.0 – 0.2 K
Dry pressure
25 – 50 km: 0.08 – 0.40 % 8 – 25 km: 0.08 % 0 – 8 km 0.40 – 0.08 %
Dry geopotential height
25 – 50 km: 4 – 40 m 8 – 25 km: 4 m 0 - 8 km: 4 m
Temperature
25 – 50 km: 0.2 – 2.0 K 8 – 25 km: 0.2 K 0 – 8 km: 1.0 – 0.2 K
Specific humidity
8 – 12 km: 3.0 % 0 – 8 km: 3.0 %
Tropopause Height
100.0 m
[1] Whichever is greater. [2] An accuracy interval means a linearly changing quantity between the two values over the given height interval.
Methods for validation
Nine broad latitude-height regions (tropics, mid-latitudes, high latitudes and low, middle, high altitudes) are defined. The absolute values of the differences between the monthly-mean RO data and the ERA-Interim reanalysis data are computed on the Level 3 grid. Each value is compared to the service specification valid for that altitude. The compliance with the Service Specifications are determined, within each region and for each calendar month, by requiring that 60% of the absolute differences are smaller than the corresponding specification.