EUMETSAT

ROM SAF

RADIO OCCULTATION METEOROLOGY

ROM SAF CDOP-3

Visiting Scientist Report 38:

Impacts on RO climatologies from under-sampling of atmospheric diurnal cycles

Stephen Leroy

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/VS/38
Version: 1.0
Date: 20 December 2020

DOCUMENT AUTHOR TABLE

	Author(s)	*Function*	*Date*
Prepared by:	Stephen Leroy	ROM SAF Visiting Scientist	20/12 2020
Reviewed by (internal):	Hans Gleisner	ROM SAF Scientist	30/11 2020
Approved by:	Sean Healy	ROM SAF Science Coordinator	18/12 2020
Approved by:	Kent B. Lauritsen	ROM SAF Project Manager	20/12 2020

DOCUMENT CHANGE RECORD

Version	*Date*	*By*	*Description*
Draft 0.1	28/10/2020	SL	First draft
Draft 0.2	4/11/2020	SL	Minor debugging, edits
Draft 0.3	10/11/2020	SL	Simulation of BnA bias
Draft 0.4	30/11/2020	SL	Addition of figures 6 and 7
Draft 0.5	20/12/2020	SL	Edit discussion of figure 2
1.0	20/12/2020	SL	Delivery

DOCUMENT DISTRIBUTION LIST

The VS report is made available at the ROM SAF website.

VS AUTHOR AND DURATION

VS Author

This VS study was carried out by Stephen Leroy, Atmospheric and Environmental Research Inc., USA; email: sleroy@aer.com

VS Duration

The VS study was performed during July 2019 – December 2020 with a one-week visit to DMI, Denmark in October 2019.

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 radio occultation (RO) data from the Metop and Metop-SG 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

Bayesian interpolation is applied to COSMIC RO data in order to analyse and map the atmospheric diurnal cycle, biases in ERA-Interim relevant to the formation of climatologies, and to produce the highest quality analyses of the diurnal cycle in the atmosphere to date. Bayesian interpolation is also applied to Metop RO data but without an analysis of the diurnal cycle in order to investigate the climatological biases incurred by under-sampling of the diurnal cycle by Metop RO. Only the migrating atmospheric features are studied herein, meaning those that are phase-locked to the Sun, even though the analysis by Bayesian interpolation permits analysis of all other diurnal features as well. Monthly maps of RO dry temperature and refractivity are produced from ~2007 through 2016 for COSMIC and Metop RO data on geopotential height levels ranging from 10 to 50 km in intervals of 0.5 km.

Two migrating diurnal features are found clearly in the analysis of COSMIC RO data: the wavenumber 1 atmospheric tide, and a large 24-hr oscillation in the mid-latitude upper stratosphere. The atmospheric tide produced by the forecasts of ERA-Interim is in near perfect agreement with that analysed from COSMIC RO data. In addition to the atmospheric tides, an oscillation spanning 30º to 60º latitude in each hemisphere and 35 to 50 km height is found. We designate it the *stratospheric diurnal oscillation* (SDO). It has been found in other data sets but is rarely studied. Its amplitude is 2 K and its maximum occurs at 20:00 solar time. The SDO is also produced in ERA-Interim forecasts, but its amplitude is too large and its maximum leads that found in RO data by approximately 2 hours.

Metop RO temperatures are found to be greater than COSMIC RO temperatures at the stratopause by 1 K, falling off with depth in the atmosphere as inverse pressure, but this error does not exist in the mid-latitude stratosphere. For level 3 climatologies formed from Metop RO data by sampling-error-removal, a bias is incurred because models—in this case, ERA-Interim—simulate the semidiurnal component of the SDO imperfectly. The result is that sampling-error-removal climatologies incur a bias of −0.2 K associated with the SDO in the mid-latitude upper stratosphere.

1. Introduction

1.1 Purpose of Document

This document contains the results from the ROM SAF Visiting Scientist activity on analysis of the atmospheric diurnal cycle as it is manifested in GPS radio occultation data processed by the ROM SAF using Bayesian interpolation with the objective to better understand what role the diurnal cycle might play in biasing atmospheric climatologies of radio occultation data.

The document is organized as follows: chapter 2 presents Bayesian interpolation and how it can be applied to analyse the diurnal cycle in GPS radio occultation data; chapter 3 describes the data used in the analysis; chapter 4 presents results; and chapter 5 is the summary and conclusions of this work.

1.2 Background

Forming monthly gridded data sets based on quantities retrieved from radio occultation data (RO) obtained using low-Earth orbiting (LEO) receivers and the transmitters of the Global Navigation Satellite Systems (GNSS) is a non-trivial task. RO data is globally, non-uniformly, and seemingly randomly distributed in longitude and latitude; thus, it seems a clear first approach to forming monthly gridded products based on RO-retrieved quantities would be binning-and-averaging. In binning-and-averaging (BnA), cells in longitude and latitude that span the entire globe are defined, all RO retrieved quantities that were retrieved from RO events that occurred in each cell at a pre-defined vertical level are averaged together and taken as the monthly average for each cell. This approach was found to be lacking almost immediately, inasmuch as the non-uniformity of the data is severe and the naturally occurring fluctuations associated with weather variability are greatly under-sampled. Moreover, some aspects of the non-uniform sampling can be systematic in nature, biasing the climatology on monthly or even infinite time-scales. Recent work has categorized the errors three-fold: spatial under-sampling of synoptic variability, temporal under-sampling of synoptic variability, and under-sampling of the diurnal cycle (Shen et al. 2020). Any approach to forming gridded climatologies from such distributions ought to take such characteristics into consideration. Two approaches that do have been introduced: sampling error removal and Bayesian interpolation.

In sampling error removal, the fluctuations of synoptic variability and the diurnal cycle in the atmosphere are removed from a BnA climatology by interpolating a weather model steered by data assimilation at the times and positions of the RO data and comparing the BnA climatology of the interpolated model to the monthly gridded field produced by the same model. The difference between the model BnA climatology and the monthly model gridded field is called the sampling error, and this sampling error is then subtracted from RO BnA climatology (Foelsche et al. 2008, 2011). The inherent assumption in this approach is that the weather model represents the fluctuations of synoptic variability and the diurnal cycle much more accurately than it represents the (mappable) monthly average of an upper air variable such as temperature or microwave refractivity. Define a BnA climatology of RO-retrieved variable y at locations r_i to be $\mathcal{B}\big(y(r_i)\big)$; the BnA climatology of the same variable interpolated from a weather model at the same locations is $\mathcal{B}\big(y_M(r_i)\big)$.

No objective requirements are placed on the resolution of these BnA climatologies in longitude and latitude. The resolution is left to the discretion of the researcher who computes and publishes the climatology. The gridded products of the same variable as published by the weather model is defined as $\mathcal{M}(y_M)$. The published climatology $\mathcal{M}(y)$ of RO data is then computed by

$$\mathcal{M}(y) = \mathcal{B}(y(r_i)) - \left[\mathcal{B}\left(y_M(r_i)\right) - \mathcal{M}(y_M)\right]. \quad (1)$$

The quantity within square brackets is the *sampling error*, is usually computed first and then used as a correction to the BnA climatology of the RO-retrieved variable. A complication arises if the same RO data set being mapped has also been assimilated into the weather model In such a case, using the analysis fields of the weather model would incur a great error in that the quantity $y(r_i) = y_M(r_i)$ for all RO data and the final RO climatology $\mathcal{M}(y) = \mathcal{M}(y_M)$. This would be incorrect, though, since there is no guarantee that the model variable y would be the same as that measured by RO when beyond some critical distance from the nearest RO data point, thus making $\mathcal{M}(y) = \mathcal{M}(y_M)$ incorrect. This problem is circumvented by interpolating the forecasts produced by the weather model rather than the analysis produced by the weather model, precisely because then $y(r_i)$ and $y_M(r_i)$ would not be the same. The sampling-error-removal approach has been applied to the computing and publication of RO climatologies many times (Lackner et al. 2011; Steiner et al. 2011, 2020; Gleisner et al. 2020).

In Bayesian interpolation, a set of basis functions are fit to RO-retrieved data within a pre-specified time window without overfitting the data. Bayesian interpolation is intended to fit irregularly gridded data with unknown error characteristics a priori (MacKay 1992). If the data has only one independent coordinate, then the basis function can be sinusoids, Legendre polynomials, Hermite polynomials, etc., whatever is best suited to the data. RO data has two independent coordinates, longitude and latitude on a sphere, and thus the natural set of basis functions to be fit are spherical harmonics because they are orthonormal on a sphere and their curvatures are easily computed (Leroy 1997; Leroy et al. 2012). It has the ability to find the effective horizontal resolution of RO data within a specified time interval and provides uncertainty estimates that are sensitive to the local density of RO data. It has been used less often than the sampling-error-removal technique, however (Ao and Hajj 2013; Leroy et al. 2018; Vergados et al. 2020; Ao et al. 2020).

The two techniques each have their advantages and disadvantages. The sampling-error-removal approach performs remarkably well at reducing the noise associated with synoptic variability in the atmosphere, especially in those regions of the atmosphere where a weather model captures that variability well, which is in the troposphere and lower stratosphere. On the other hand, it is unable to dictate the horizontal spatial resolution of the output climatology, and, as a consequence, produces small scale anomalies that derive from the weather model and not from the RO data themselves. Bayesian interpolation does yield an inference of the effective spatial resolution of the RO data, enables a simple analytic computation of horizontal gradients for the computation of geostrophic winds (Verkhoglyadova et al. 2014), and is wholly independent of outside sources of data (such as weather forecasts). On the other hand, it does not reduce the noise associated with synoptic variability, making time-series of the climatologies appear much noisier than those produced by sampling-error-removal.

Because of the broad array of GNSS radio occultation missions and satellites in orbit about the Earth, special attention must be paid in constructing climatologies of their data because of their associated orbits and how they sample the diurnal cycle. Meteorological satellite commonly are inserted into high inclination orbits, usually sun-synchronous. Over the course of a month, such the RO soundings such satellites obtain (if they have RO instruments on board) occur at a very limited range in local, or solar time (ST). If the orbit is sun-synchronous, then that limited range in solar time is the same throughout the duration of the satellite's lifetime. If there is a systematic atmospheric diurnal and/or semidiurnal cycle in the variable being retrieved, then subsampling that cycle will lead to a bias in the climatological field being built. The Metop RO data fall into this category. If the orbit has a high inclination but is not sun-synchronous, then the atmospheric diurnal and semidiurnal cycles are aliased to multi-month timescales because the solar times sampled by that satellite's RO data might be limited in limited in range, but that range drifts as the satellite undergoes its Earth oblateness-induced regression of nodes. The CHAMP RO data fall into this category. The COSMIC mission consisted of six satellites in high inclination orbits intended to be separated by 30° in ascending node, thus obtaining RO data that span all solar times in sampling throughout the duration of the mission. As consequence, COSMIC RO data should lead to easily constructed, unbiased climatologies based on their RO-retrieved variables. The distribution in ascending node of the COSMIC constellation does leave small gaps in solar time coverage at high latitudes in both hemispheres, however, but inasmuch as there is little to no systematic diurnal cycle in these latitudes, the gaps are inconsequential.

Care must be taken in the handling of RO data in the presence of a systematic atmospheric diurnal cycle in RO-retrieved variables. A simple BnA approach would certainly fail if coverage of the diurnal cycle in solar time is non-uniform. The sampling-error-removal approach works well as long as the weather model used to remove sampling error simulates the atmospheric diurnal cycle correctly. In one piece of work, a *residual* atmospheric cycle was found using the sampling-error-removal approach, computing the diurnal cycle mis-representation of the weather model using COSMIC RO data (Pirscher et al. 2010). Bayesian interpolation would also fail if the basis functions were time-independent and if coverage of the diurnal cycle in solar time is non-uniform. Bayesian interpolation could be made to succeed, however, if its basis functions were extended to cover the diurnal cycle as sinusoids. If this were done, then solar time-dependent maps of an RO-retrieved field could be produced by simply expanding the basis functions as desired after determining the most probable coefficients of a fit. A recent study extended Bayesian interpolation into the solar time domain and finding the corresponding regulariser (Leroy et al. 2020).

In this report, we apply Bayesian interpolation that analyzes the diurnal cycle to COSMIC data and evaluate the implications of its results for RO climatologies based on Metop RO data. For most of its mission, COSMIC produced RO data that covered all solar times if somewhat non-uniformly. A form of Bayesian interpolation with basis functions that extend into diurnal time can be applied to COSMIC data to analyze the atmospheric diurnal cycle on a monthly basis and produce climatologies of RO-derived variables that are not biased because of non-uniform sampling in diurnal time. Bayesian interpolation can be applied to Metop RO data as well, but there is no point in analyzing the atmospheric diurnal cycle in those data because RO data covers only a small range of diurnal times, as

explained above. The analysis of the diurnal cycle in COSMIC data can then be used to investigate how undersampling of the diurnal cycle by Metop leads to biases in its RO-derived climatologies.

2. Bayesian Interpolation and the Diurnal Cycle

Bayesian interpolation was introduced in the context of the neural sciences (MacKay 1992), and it has been described elsewhere as a technique for mapping RO data (Leroy 1997; Leroy et al. 2012). Herein we briefly present the equations of Bayesian interpolation using the same formulation and variables as used in the previous papers, with the exception of the data being represented by "$\mathbf{y}$" rather than by "$\mathbf{t}$".

In Bayesian interpolation, the basis functions are ψ_μ, they are expanded using coefficients $\mathbf{w}$ (with terms w_μ) by $\sum_\mu w_\mu \psi_\mu$, and are evaluated at the locations r_i of the RO data to form the matrix ϕ (with elements $\phi_{i\mu}$). The *first level of inference* in Bayesian interpolation is easily interpreted as a problem in minimizing χ^2:

$$\chi^2 \equiv \beta(\mathbf{y} - \phi\mathbf{w})^2 + \alpha\mathbf{w}^T\mathbf{C}\mathbf{w} \tag{2}$$

where $\mathbf{y}$ are the RO-retrieved quantities obtained on a specific vertical level. The first term on the right reduces the data misfit and the second term punishes the amplitude of the expansion coefficients, thereby imposing a smoothness condition. The latter term is known as a "penalty" or "regularising" term in most of the literature on the statistical fitting of data. The coefficients α and β are the weights to be applied to the data misfit term and the regularising term, which are unknown in this problem. The dimension of $\mathbf{y}$ is N and the dimension of $\mathbf{w}$ is k. The "most likely" values of the coefficients are determined by the first level of inference:

$$\mathbf{w}_{ML} = \mathbf{A}^{-1}\beta\phi^T\mathbf{y} \tag{3}$$

where $\mathbf{A} = \beta\mathbf{B} + \alpha\mathbf{C}$ and $\mathbf{B} = \phi^T\phi$. With a specified regulariser and specified values of the weights α and β, the most likely values for the expansion coefficients $\mathbf{w}_{ML}$ can be computed. The *second level of inference* in Bayesian interpolation finds the "most probable" values of the weights α_{MP} and β_{MP} from a probability distribution in those values as determined from the evidence function of corresponding to the first level of inference. The equation for the evidence function is not reproduced here (see MacKay 1992); it suffices to implement a peculiar property of the dimensions and weights where the evidence function is at its maximum. The most probable values of the weights α_{MP} and β_{MP} are found from

$$\alpha_{MP}\mathbf{w}_{ML}^T\mathbf{C}\mathbf{w}_{ML} = \gamma \tag{4a}$$

$$\beta_{MP}|\mathbf{y} - \phi\mathbf{w}_{ML}|^2 = N - \gamma \tag{4b}$$

in which γ is the effective number of fitted coefficients, computed by

$$\gamma = k - \alpha_{MP}\text{Trace}\,\mathbf{A}^{-1}\mathbf{C}\,. \tag{5}$$

Table 1. The parameters used in defining the on-diagonal terms of the regularizing matrix.

Condition	a	b	c
$l = m = n = 0$	0	0	0.3
$l > 0, m = n = 0$	2	0	0.3
$l, m > 0, n = 0$	2	0	1
$l = m = 0, n > 0$	0	2	1
otherwise	2	2	1

These equations are solved iteratively, starting with nominal values for the weights, computing most likely coefficients using Eq. 3, calculating the number of effective coefficients using Eq. 5, and finally computing the next values of the weights in the iteration using Eqs. 4a and 4b. The most probable values of the expansion coefficients at the evidence maximum for the weights are $\mathbf{w}_{MP}$. This iteration always converges. Obtaining 1% accuracy for the weights generally requires no more than ~6 iterations.

The basis functions in space in time are indexed by μ and dependent on longitude λ, latitude φ, and diurnal time τ as

$$\psi_\mu(\lambda, \varphi, \tau) = P_{lm}(\sin\varphi) \times \begin{cases} \cos(m\lambda)\cos(n\tau) & \text{for } \nu = 1 \\ \cos(m\lambda)\sin(n\tau) & \text{for } \nu = 2 \\ \sin(m\lambda)\cos(n\tau) & \text{for } \nu = 3 \\ \sin(m\lambda)\sin(n\tau) & \text{for } \nu = 4 \end{cases} \tag{6}$$

in which the diurnal time spans $[0,2\pi]$ to span a 24-hour period. Every integer value of the basis function index μ corresponds to a set of indices (l, m, n, ν). The $P_{lm}(\cdots)$ are associated Legendre polynomials of degree l and order m; n is the harmonic of the diurnal cycle. The Legendre polynomial degree counts from $l = 0$ to some user-specified maximum degree, and the order spans the interval $0 \le m \le l$ for every Legendre polynomial degree. The diurnal cycle harmonic counts from $n = 0$ to a user-specified maximum diurnal cycle harmonic. The index ν simply indicates whether the basis function is a sine or cosine term in the longitude or diurnal time. Notice that for $m = 0$, the $\nu = 3{,}4$ terms do not exist; and that for $n = 0$, the $n = 2{,}4$ terms do not exist. Otherwise, ν can take any value 1 through 4.

The diurnal time τ can take one of two definitions: either it is the solar time τ_d, measuring time from local midnight, or it is the synoptic time τ_s, measuring time since 00:00 UTC. Performing Bayesian interpolation with one definition of diurnal time does not translate easily into Bayesian interpolation using the other definition: there is no simple transformation relating the basis function expansion coefficients from one definition to the other because the basis function for (l, m, n, ν) is not orthogonal to the basis function for (l, m', n, ν) if $m \neq m'$. When expanding the basis functions, it is simple to account for solar time vs. synoptic time using the simple relationship $\tau_d = \tau_s + \lambda$, however.

The formula for the total number of basis functions is

$$k = (l_{\max} + 1)^2 (2n_{\max} + 1) \tag{7}$$

Bayesian interpolation requires that the number of data points N to which it is applied must be at least as great as the number of basis functions k. If fewer data points are available, then either the number of basis functions must be reduced so that $k \leq N$ or Bayesian interpolation should not be applied at all.

A definition for the regulariser was found that maximizes the evidence for it (Leroy et al. 2020). Because the basis function defined in Eq. 6 are orthonormal, the off-diagonal elements of the regularizing matrix $\mathbf{C}$ are 0. The diagonal elements are just C_μ and they are defined by

$$C_\mu = cl^a(l+1)^a n^b \tag{8}$$

with parameters a, b, c depending on the values of (l, m, n, v) that correspond to basis function index μ. In this report, the values are chosen according to Table 1. These values are chosen so as to under-penalize large global mean, time-mean values and meridional gradients for time-mean terms. This formulation for the regulariser is used in this report.

3. Analysis

We have applied the Bayesian interpolation as described in Section 2 to COSMIC RO data and to Metop RO data: we have accounted for the diurnal cycle up to the second harmonic ($n_{\max} = 2$) for COSMIC RO data but not for the Metop RO data ($n_{\max} = 0$). The COSMIC RO data commences in May, 2006, and ends in December, 2016; the Metop RO data commences in October, 2007, and ends in December, 2016. Following previous research, we have set the maximum spherical harmonic degree to 14 (Leroy et al. 2012). This means that the number of fitted coefficients for each map of COSMIC RO data is 1125 and for each map of Metop RO data is 225. In the months from November, 2007 through February, 2008, Metop collected fewer than 225 RO soundings in each month, and so Bayesian interpolation could not be applied to these months of Metop RO data. We have mapped both microwave refractivity and dry temperature (Danzer et al. 2014) as retrieved from RO data. We map refractivity and dry temperature on surfaces of constant geopotential height, spanning 10.0 km to 50.0 km at intervals of 0.5 km.

In addition to mapping the RO-retrieved variables refractivity and dry temperature, we also map the difference of these RO-retrieved quantities from the weather model that would otherwise be used in the sampling-error-removal approach to creating climatologies of RO data. Notice that the sampling-error-removal equation given as Eq. 1 can be rewritten as

$$\mathcal{M}(y) = \mathcal{M}(y_M) - \left[\mathcal{B}\left(y_M(r_i)\right) - \mathcal{B}(y(r_i))\right]. \tag{9}$$

The bracketed term on the right is the bias of the weather model with respect to the RO data, and that bias is formed by BnA. Once the bias map is computed, it is removed from the gridded climatology $\mathcal{M}(y_M)$ that can be determined from the weather model alone. The bias term is especially interesting in the context of the diurnal cycle in the sampling-error-removal approach. Recall that the sampling-error-removal approach is conditioned on the weather model producing the variability that is to be removed accurately. If the weather model does not simulate the diurnal cycle correctly, then the error in the model's diurnal cycle will lead to biases in the sampling-error-removal climatologies that are computed,

Table 2. Catalogue of RO data to which Bayesian interpolation was applied. When a variable is mentioned "less model", the quantity that was mapped is that variable as retrieved from RO data but with the same quantity as forecast by ERA-Interim interpolated in time and space to the geolocation of the RO sounding.

Mission	Variable	Dates	$l_{\max}$	$n_{\max}$
COSMIC	Dry temperature	5/2006 – 12/2016	14	2
COSMIC	Dry temperature less model	5/2006 – 12/2016	14	2
COSMIC	Refractivity	5/2006 – 12/2016	14	2
COSMIC	Refractivity less model	5/2006 – 12/2016	14	2
METOP	Dry temperature	10/2007 – 12/2016	14	0
METOP	Dry temperature less model	10/2007 – 12/2016	14	0
METOP	Refractivity	10/2007 – 12/2016	14	0
METOP	Refractivity less model	10/2007 – 12/2016	14	0

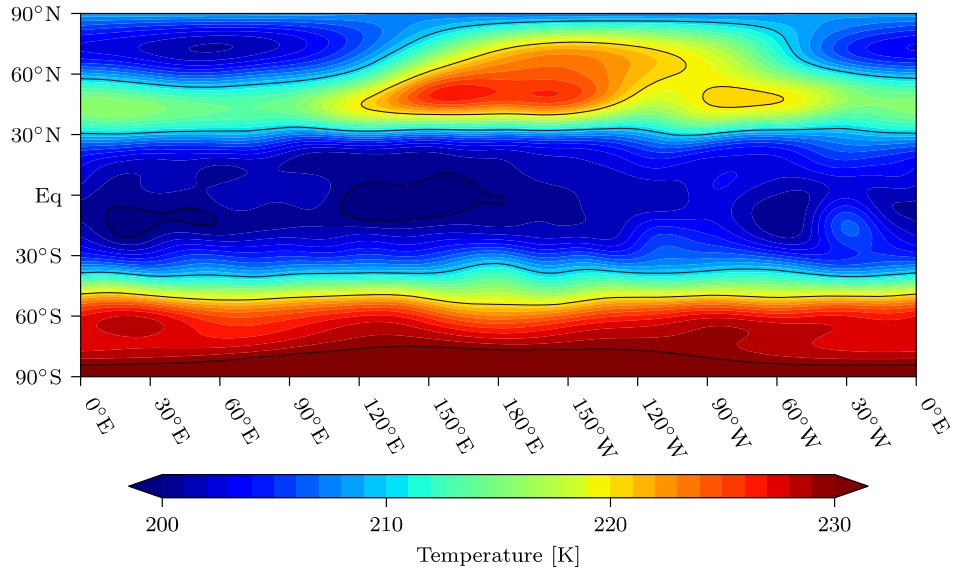

Figure 1. Map of dry temperature determined by Bayesian interpolation applied to COSMIC RO data. The map is a spherical harmonic expansion in longitude and latitude at 00:00 UTC for the month of January, 2010, at geopotential height 15.0 km.

but only if the RO data does not cover the diurnal cycle uniformly. For this reason, it will be interesting to use Bayesian interpolation to examine the model bias diurnal cycle, defined by the term $\left[\mathcal{B}\left(\mathrm{y_M(r_i)}\right) - \mathcal{B}(\mathrm{y(r_i)})\right]$.

While Bayesian interpolation is not strictly a linear operation, the model bias can be most accurately computed by a single application of Bayesian interpolation:

$$D = \mathcal{B}(y_M(r_i) - y(r_i))\,. \tag{10}$$

Mapping the model bias D requires interpolating the weather model to the times and locations r_i of the i'th RO data point. Both the refractivity and the temperature forecasts of the weather model are bilinearly interpolated in longitude and latitude and linearly interpolated in geopotential height at the locations and times of the RO data. In addition, they are linearly interpolated in time, following the implementation of Gleisner et al. (2020) using the forecast fields of ERA-Interim. Bayesian interpolation is applied to the difference between the interpolated weather model values and RO data value. The weather model used in this report is ERA-Interim (Dee et al. 2011; Poli et al. 2010).

Note that our analysis will compare "dry temperature" as retrieved by RO to actual (thermodynamic) temperature as forecast by the weather model. Through most of the stratosphere the difference between dry temperature and thermodynamic temperature should not exceed 0.01 K. In the lower stratosphere, however, the difference can approach 0.1 K in the Tropics at 10 km height, but that difference falls off rapidly with height. The difference will be very noticeable in the Bayesian interpolation mapping of the model bias D.

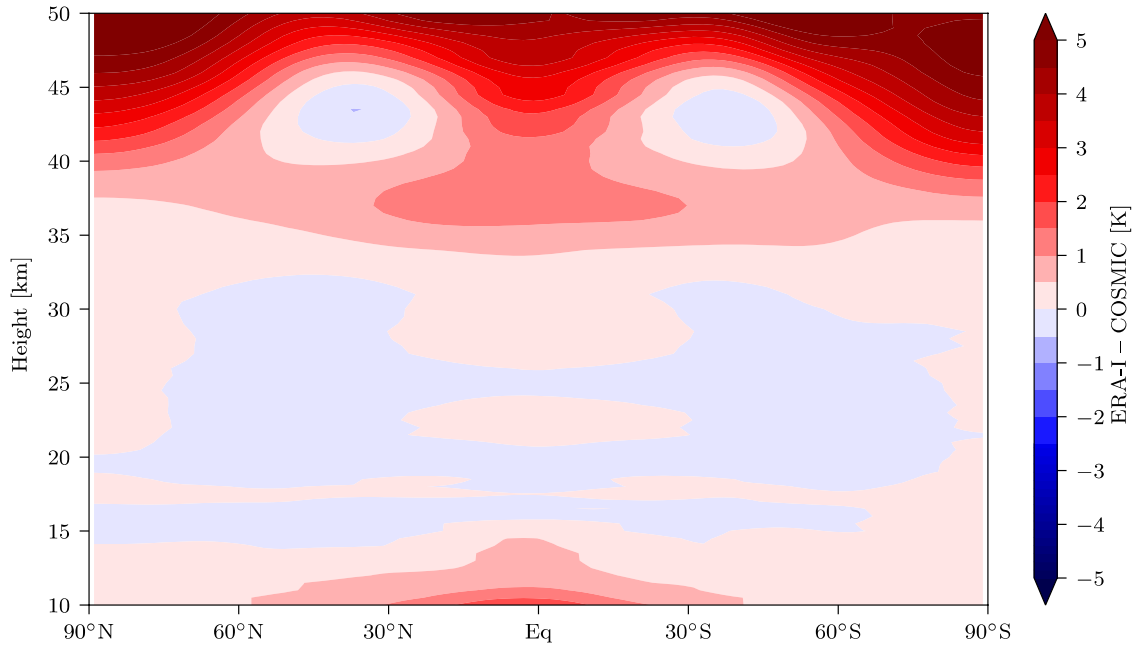

Figure 2. The temperature bias of ERA-Interim in comparison to COSMIC RO. The zonal average, diurnal average, annual average of the difference between ERA-Interim temperature and dry temperature as computed from COSMIC RO data for COSMIC data from May, 2006 through December, 2016.

Also note that we applied Bayesian interpolation to all RO data (and differences of RO data from the weather model) over the course of each month. While it may be more efficient in terms of storage and compute time than segmenting the Bayesian interpolation by 3- or 4-day periods, it can incur biases especially when RO sampling over the course of the month is non-uniform by day during seasons of rapid temporal change, such as Spring or Fall. Other applications of Bayesian interpolation to RO data have segmented data by 3- or 4-day periods before averaging those maps over the month (Ao and Hajj 2013; Ao et al. 2020).

Fig. 1 shows a monthly mean, diurnal mean map of RO dry temperature at 15 km as determined by COSMIC data for the month of January, 2010. It is simple to compute such a map given the output of Bayesian interpolation: one only needs to keep the time-mean basis functions ($n = 0$) in the expansion $\sum_\mu w_\mu \psi_\mu$.

Fig. 2 shows the bias D of ERA-Interim with respect to COSMIC RO data. To construct this field, the bias field D was computed for every month from May, 2006 through December 2016, the diurnal mean of the bias computed by retaining only $n = 0$ basis functions, and the zonal average computed by retaining only $m = 0$ basis functions for all heights. Next, the seasonal cycle was computed by averaging the model bias by month-of-year. Finally, the model bias was computed by averaging together all months of the seasonal cycle.

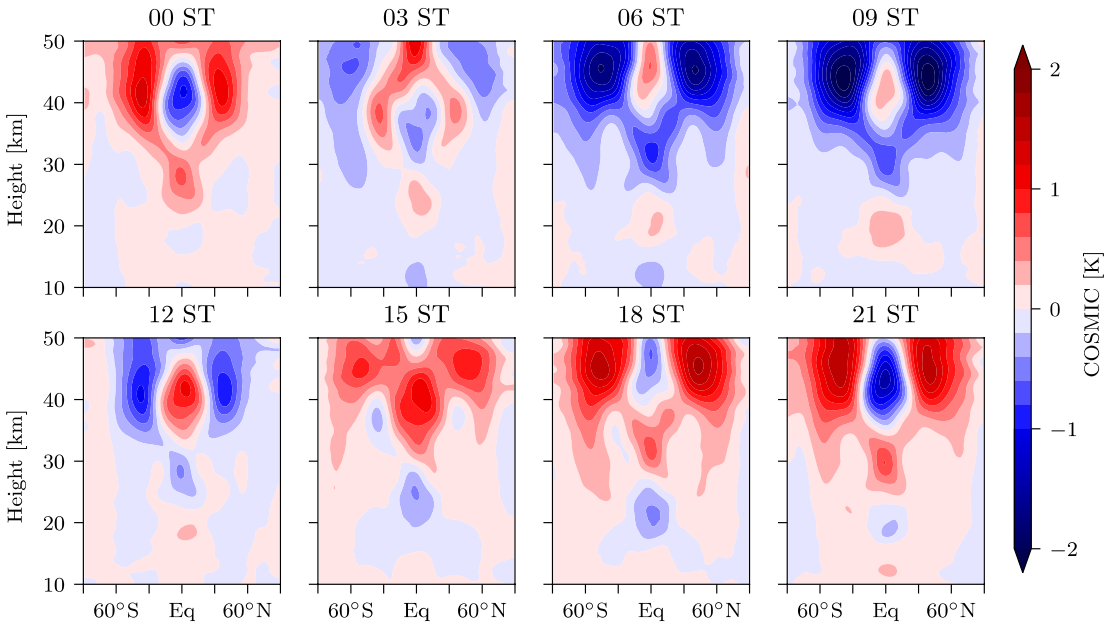

Figure 3. The zonal average, seasonal average diurnal cycle in dry temperature as computed from COSMIC RO data. The average of the seasonal cycle of the zonal mean variation associated with migrating thermal features are plotted for eight 3-hour instances of solar time (ST). "Migrating" means that these features are phase-locked to the solar angle and are thus only a function of solar time.

Several features are readily apparent. First, ERA-Interim and COSMIC agree at the level of ~0.1 K throughout the lower to middle stratosphere. Temperature in the stratosphere in ERA-Interim is anchored strongly by COSMIC RO data (Poli et al. 2010), so the agreement should not be surprising. Second, ERA-Interim temperature exceeds COSMIC RO temperature above ~40 km, and the difference increases exponentially with height, reaching ~5 K at 50 km. The degree to which this difference can be attributed to ERA-Interim or to COSMIC RO is uncertain. There are noteworthy holes in the bias in the upper stratosphere in the mid-latitude, between 30 and 50 latitude in both hemispheres. The cause of this future is likely some combination of two phenomena: it is a consequence of ionospheric residual in the RO retrieval process, and these are the latitude bands where the ionosphere is most active; and there is a possibility that ERA-Interim mis-models atmospheric tides, which are known to be difficult to model at these latitudes and heights. Third, ERA-Interim is biased warm with respect to COSMIC RO at the top of the troposphere. This is explained as the difference between "dry temperature" as computed in RO retrieval and thermodynamic temperature as represented by ERA-Interim. Dry temperature is always less than thermodynamic temperature, hence the "warm bias" of ERA-Interim in the tropical upper troposphere.

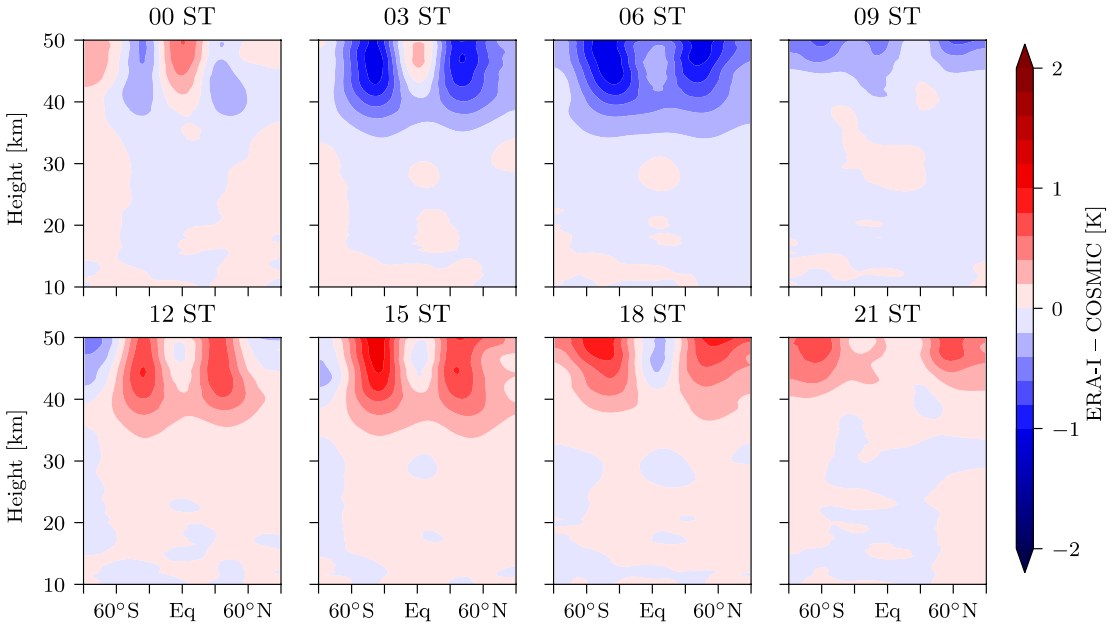

Figure 4. The bias of the diurnal atmospheric tide in temperature of ERA-Interim. These plots are the same as in Fig. 3 but for the difference between ERA-Interim and COSMIC RO. Positive values are interpreted as ERA-Interim being warmer than the migrating thermal feature in COSMIC RO dry temperature at the specified solar time.

Fig. 3 shows the annual average diurnal cycle associated with migrating tides and other thermal features as analysed in COSMIC RO-retrieved dry temperature. Bayesian interpolation was applied using a maximum diurnal harmonic $n_{\max} = 2$ and performing the mapping using $\tau = \tau_d$ as the definition of diurnal time. For each month of COSMIC RO data, the variations associated with the diurnal cycle are computed by retaining only the $n \neq 0$ basis functions; the zonal average of these fluctuations is computed by retaining only the $m = 0$ basis functions. (Computing the zonal average migrating thermal features this way is correct, only because $\tau = \tau_d$ was chosen as the diurnal time coordinate.) The variations of the diurnal cycle are expanded at three-hour intervals over the course of a day, and those variations are plotted in the figure after computing a seasonal cycle and annual average of the seasonal cycle. The simplest interpretation of this figure is that it tells a high-altitude airplane pilot what the diurnal temperature anomaly is by simply accounting for her latitude, altitude, and local time as read from a time-zone-appropriate wristwatch.

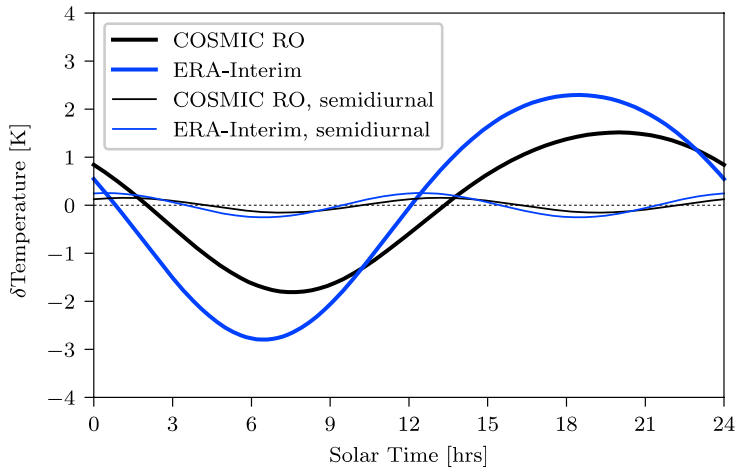

Figure 5. The midlatitude upper stratosphere diurnal oscillation (SDO) timeseries as realized in COSMIC RO data and in ERA-Interim. It is computed from Bayesian interpolation of COSMIC RO data and computing weather model bias by Bayesian interpolation. The bias map is subtracted from the COSMIC map to produce the ERA-Interim map. Each map is produced using all COSMIC RO data as in figures 2 – 4; the annual averages over the seasonal cycle of the diurnal cycles are shown. The thin lines are an expansion of the semidiurnal cycle terms only. The curves are extracted at 45ºN latitude, 48 km height.

Fig. 3 shows two prominent phenomena in the atmospheric diurnal cycle. The first is a downward propagating thermal anomaly in the deep tropical stratosphere. The phenomenon is very definitely a wave with a downward phase speed, beginning at the very top of the plotting domain (50 km) and propagating downward to the very bottom of the plotting domain (10 km). Its periodicity is 24 hours, and its vertical wavelength is ~ 23 km through most of the stratosphere. It is most prominent 6 ST and 18 ST, where two full wavelengths are easily seen. At 6 ST, positive temperature anomalies are seen at ~45 km and ~ 20 km height and negative anomalies at ~ 30 km and ~ 11 km height. The opposite is true at 18 ST, consistent with 24-hr periodicity. Other diurnal times illustrate downward phase propagation. This wave is consistent with an atmospheric tide of zonal wavenumber 1 (Xie et al. 2010). Zonally symmetric features at fixed solar times will only be those phase-locked to the Sun and with zonal wavenumbers equal to the diurnal harmonic: for atmospheric tides they are the “migrating” tides. Atmospheric tides, like all internal gravity waves, exhibit downward vertical phase speeds when the vertical component of the group velocity is upward, transporting momentum and energy from the troposphere into the stratosphere and beyond. This first feature is almost certainly the migrating atmospheric tide with zonal wavenumber 1. The second prominent feature is a diurnal warm-cold pulsing spanning the mid-latitudes (30º to 60º latitude) in both hemispheres from 35 km to 50 km height. It also is a once-per-day feature, with temperature anomalies a maximum at 21 ST and minimum at 9 ST and an overall amplitude of ~ 2 K. This feature is apparent in other data sets, such as obtained by TIMED/SABER (e.g., Sakazaki et al. 2012), but it is unexplained except as a sampling error aliased to diurnal time scales. We refer to this as the *stratosphere diurnal oscillation* (SDO).

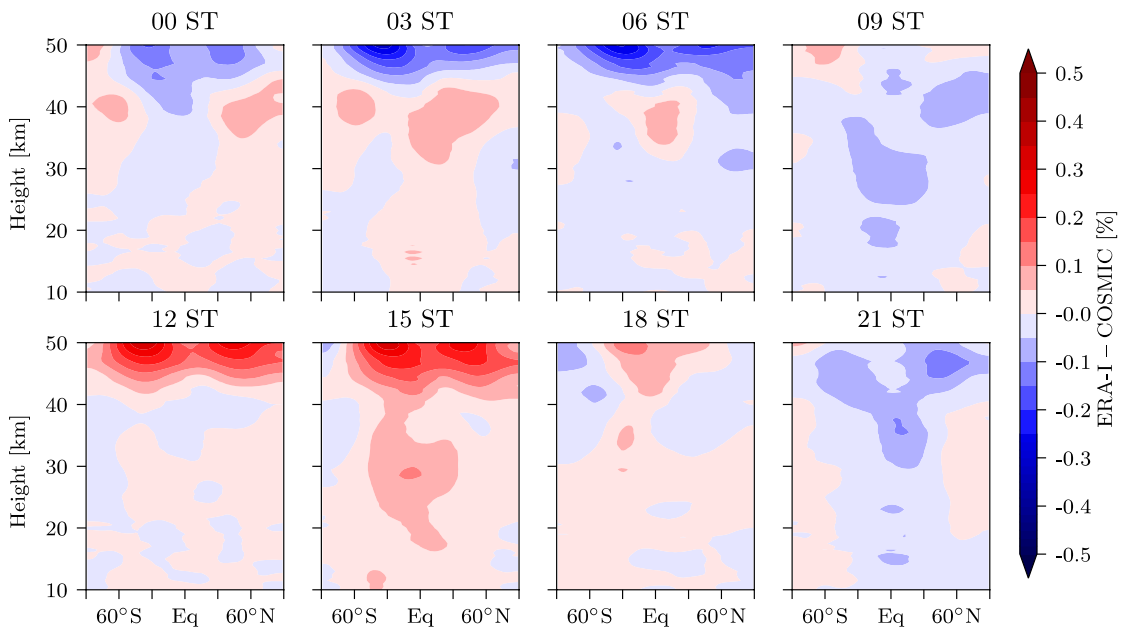

Figure 6. Diurnal cycle of the model bias in refractivity. The annual average, zonal average diurnal cycle of COSMIC RO is subtracted from the annual average, zonal average diurnal cycle of ERA-Interim at 8 solar times, the differences are normalized by the diurnal mean, zonal average of COSMIC RO refractivity, and the percentage differences are contour plotted.

Fig. 4 shows the bias of the diurnal cycle in temperature as contained in ERA-Interim in comparison to COSMIC RO dry temperature. Bayesian maps were computed for the difference between interpolated ERA-Interim temperatures and COSMIC RO temperatures as a bias D; see Eq. 10. The results shown in Fig. 4 show the same as Fig. 3 but for the model bias. It is best interpreted as the error in ERA-Interim's analysis of the migrating diurnal tides. Noteworthy is the absence of the downward phase propagating thermal feature near the equator. This means that ERA-Interim is nearly perfectly producing the migrating thermal tides in the stratosphere with respect to COSMIC RO data. On the other hand, the large migrating tidal feature in the mid-latitude upper stratosphere is simulated with much less fidelity by ERA-Interim. The bias map of the diurnal cycle in solar time shows that the model leads the data RO data because the bias in the oscillation is at its maximum approximately 6 hours before the maximum in the midlatitude stratopause oscillation in Fig. 3. Fig. 5 shows the midlatitude stratopause oscillation at 45ºN and 48 km height through 24 hours of solar time.

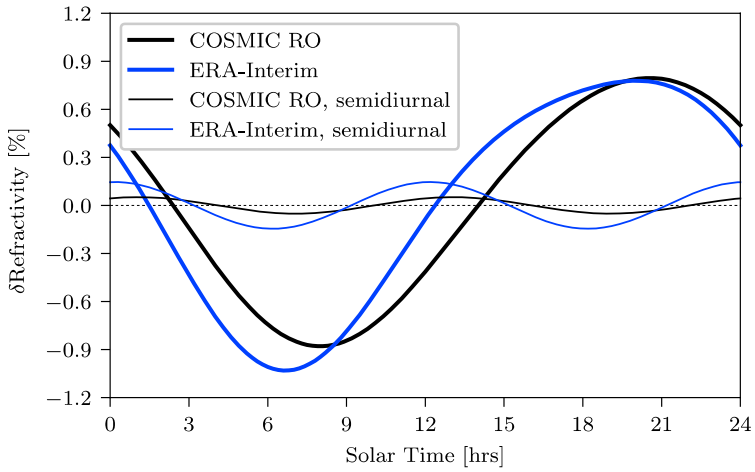

Figure 7. Same as Fig. 5 but for refractivity given as a percentage after normalization by the diurnal mean in COSMIC RO data.

Fig. 6 shows the model error in the diurnal cycle of refractivity corresponding to the model error in the diurnal cycle of dry temperature shown in Fig. 4. In GNSS RO, dry temperature anomalies occur approximately a half scale-height below the refractivity anomalies that give rise to them because pressure is obtained from refractivity by a downward integral and temperature by dividing refractivity into the pressure. The greatest error in ERA-Interim's manifestation of the diurnal cycle in refractivity occurs at the top of the analysis domain and almost certainly extends well above it. Fig. 7 shows the diurnal cycles of refractivity of ERA-Interim and COSMIC RO at 45ºN and 50 km height, following the same analysis as for temperature shown in Fig. 5.

Fig. 8 shows the refractivity bias in ERA-Interim with respect to COSMIC RO refractivity. Throughout most of the stratosphere, ERA-Interim refractivity is biased high with respect to COSMIC RO refractivity by approximately 1%. Previous studies have shown the bias to be much closer to non-existent (Gleisner 2018). The bias grows much larger in the upper stratosphere (> 2%) and disappears below the tropopause. The ripples that are readily apparent are vertical interpolation noise: we have interpolated ERA-Interim refractivity linearly in height even though refractivity is roughly exponential in height in the stratosphere. Fractional biases in refractivity as shown in Fig. 8 are related to biases in temperature as vertical gradients. Wherever there is a vertical gradient in the fractional bias in refractivity, a positive anomaly in the temperature bias appears, as shown in Fig. 2. Thus, the strong vertical gradient around 40 km height in Fig. 8 shows up as the large temperature biases centered at 40 km height as seen in Fig. 2. The ripples do not translate to temperature biases because the linear interpolation of temperature in height is more appropriate to the linear temperature profiles the atmosphere produces than it is for interpolating refractivity.

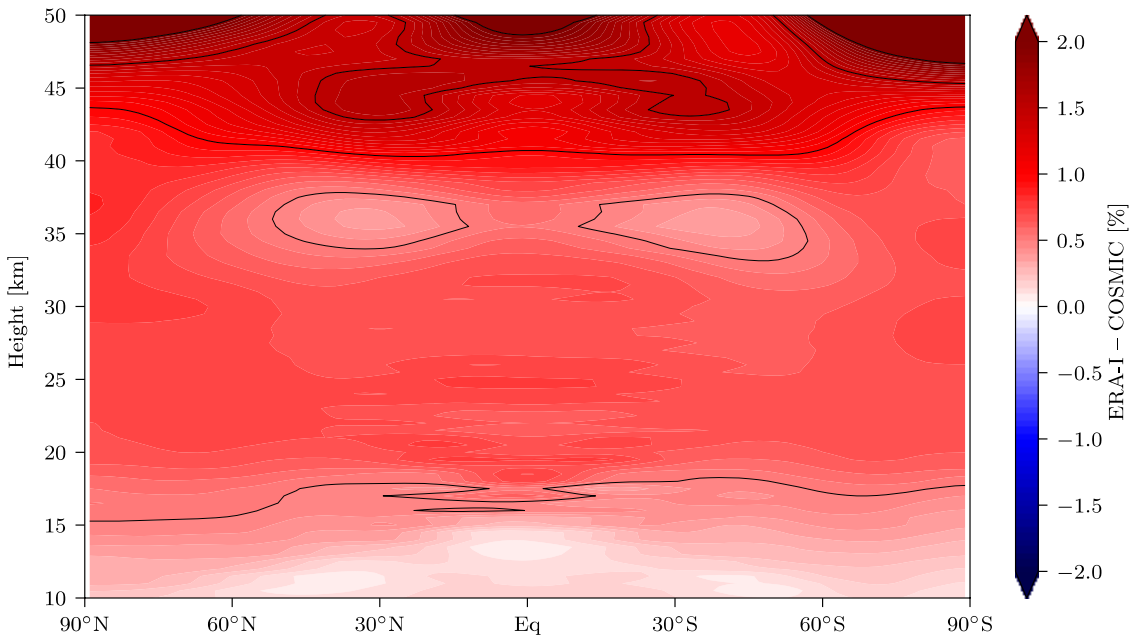

Figure 8. The bias in microwave refractivity as contained in ERA-Interim with respect to COSMIC RO data. This is the same as Fig. 2 but for refractivity rather than temperature. The units are fractional, in units of percentage, normalized by the COSMIC RO refractivity average. The contour lines correspond to the major ticks of the colourbar.

Fig. 9 shows the zonal average of the migrating diurnal cycle of refractivity as analysed in COSMIC RO data. It was computed just as for Fig. 3 for dry temperature. Fractional refractivity fluctuations are just the departures of refractivity from diurnal means and normalized by diurnal means of refractivity. Just as in Fig. 3, the downward propagating feature over the equator is the migrating atmospheric tide, dominated by zonal wavenumber 1. Fluctuations in refractivity associated with the SDO are also apparent, centered about the stratopause (~ 50 km height).

Even though, GNSS RO is considered strongest in the upper troposphere and lower stratosphere, typically 8 to 35 km height, there is evidence in this analysis that strong atmospheric signals as high as the stratopause can be mined from RO data at least in monthly climatologies. RO is processed as a three-step integral process: first convolving observed calibrated excess phase by Fourier integral operators to obtain atmospheric bending angles, second integrating observed bending angles from the top of the atmosphere downward to obtain refractivity, and finally integrating refractivity from the top of the atmosphere downward to obtain (dry) pressure. RO signals become weaker the higher one probes in the atmosphere because there is much less optical refraction where there are far fewer air molecules. Consequently, intelligent decisions must be made in the initiation of the second and third integrals at the “top” of the atmosphere in order not to create biases in monthly climatologies of RO-derived quantities.

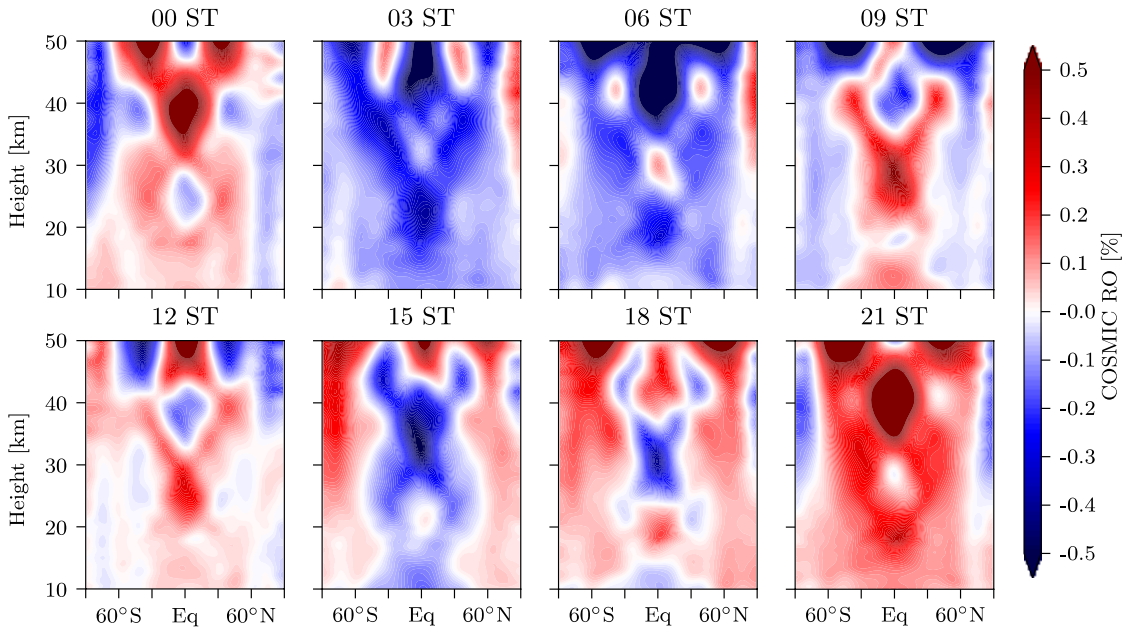

Figure 9. The zonal average, seasonal average fluctuations in microwave refractivity associated with migrating features in the stratosphere as analysed in COSMIC RO data. This figure is the same as Fig. 3 but for microwave refractivity. Fractional refractivity fluctuations are contoured in units of percentage after normalizing by the diurnal mean in refractivity as determined from COSMIC RO data.

The retrieval scheme used to generate the RO data incorporated BAROCLIM (Scherllin-Pirscher et al. 2015), a climatology of bending angles produced from COSMIC RO bending angles in the upper stratosphere and mesosphere. In the algorithm, the BAROCLIM database of bending angles was searched for the best match for each occultation, and that bending angle profile was used to initiate the second vertical integral above the stratopause. Logically, for an individual RO retrieval profile, one might assume that the refractivity and temperature near the stratopause are those that result from the bending angles of BAROCLIM, but there is evidence that atmospheric signals not embedded in BAROCLIM are being measured. The migrating diurnal tides are diurnal fluctuations that would be averaged out in the formation of BAROCLIM, which collect measured bending angles regardless of the solar time of the RO soundings. The fact that the migrating atmospheric tides are seen to propagate downward from the top of the plotting domain of Figs. 3 and 9 is an indication that information in retrieved refractivity and temperature is being realized on a single sounding basis as high as the stratopause. Note that a lot of averaging is involved in the production of the plots in Figs. 3 and 9, however.

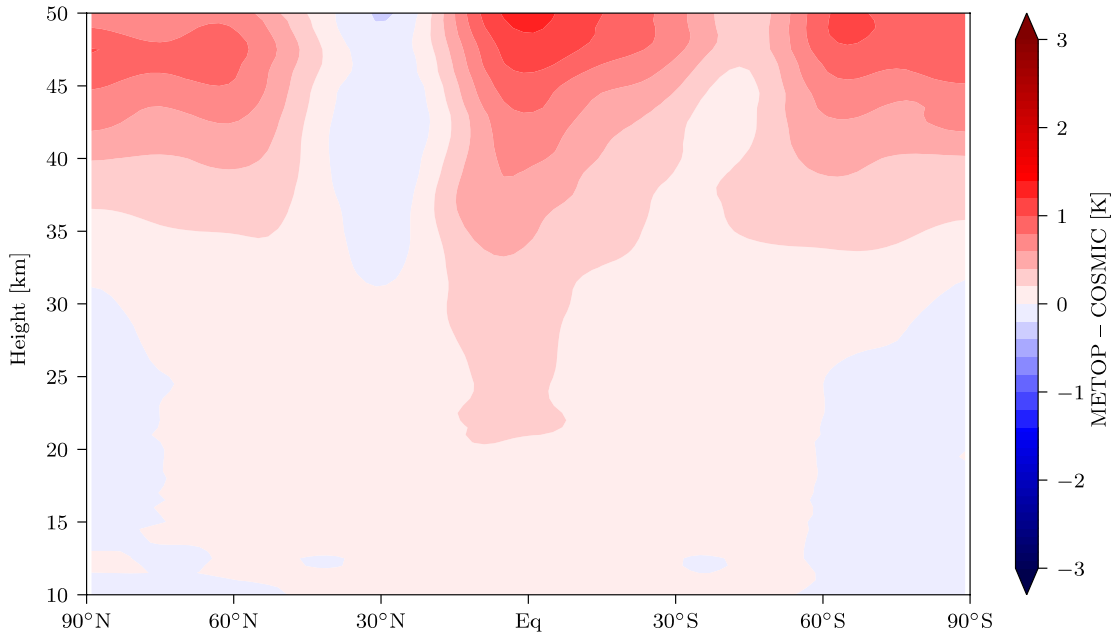

Figure 10. Difference in dry temperature climatologies produced from Metop and COSMIC RO data. Only the months of overlap of COSMIC and Metop RO data are mapped by Bayesian interpolation. The difference was formed from the diurnal mean, zonal mean of the monthly climatologies of COSMIC RO data and the zonal mean of the monthly climatologies of Metop RO data. No attempt was made to analyse the diurnal cycle in the Metop RO data.

Fig. 10 shows the climatological difference in dry temperature as retrieved from COSMIC and from Metop RO data. Bayesian interpolation was applied to Metop RO data in addition to COSMIC RO data, except no attempt was made to analyse the diurnal cycle in Metop RO data ($n_{\max} = 0$) because Metop RO data does not span the diurnal cycle. As a consequence, the differences between Metop and COSMIC climatologies of RO data will have contributions from instrument biases, structural retrieval biases, and sampling biases due to sub-sampling of the diurnal cycle by Metop RO. After the monthly climatologies of dry temperature and refractivity were constructed from Metop RO data, differences with COSMIC RO were computed for the months of overlap, a seasonal cycle of the differences was obtained, and then the mean of that seasonal cycle over all 12 months of the year computed. Overall, temperature retrieved from Metop RO data is greater than from COSMIC RO data by as much as 1 K at the stratopause, decaying exponentially with depth into the stratosphere.

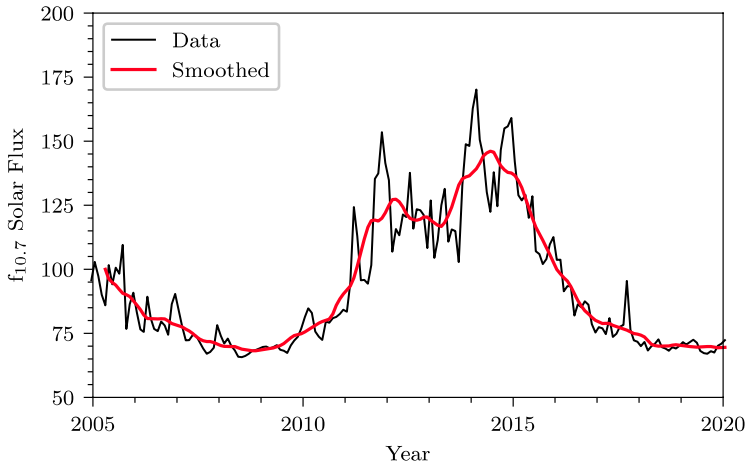

Figure 11. The radio flux of the Sun at 10.7 cm. Both a monthly timeseries and a smoothed version of that timeseries are shown. The units are "solar flux units". The data are obtained from swpc.noaa.gov.

One question remains regarding the SDO, whether it is a real atmospheric phenomenon or a consequence of ionospheric residual in the RO retrieval process. Because microwave refractivity has a substantial contribution from free electrons in the ionosphere, RO measurements track two GNSS signals at different carrier frequencies simultaneously. Because microwaves are dispersive in the ionosphere but not in the neutral atmosphere, the bending angles retrieved for each tracked signal can be linearly combined in order to remove the influence of the free electrons in the ionosphere on retrieved refractivity, leaving only the neutral atmosphere to contribute to refractivity. This linear combination can never completely remove the influence of the ionosphere on retrievals of refractivity because the two tracked signals do not follow the same ray paths through the ionosphere-atmosphere system. The residual influence of the ionosphere after the ionosphere-removing linear combination is referred to as the "ionosphere residual" (Kursinski et al. 1997). Its impact at solar maximum on temperature can be as large as 1% in daytime, or 4 K in dry temperature at 50 km height. The error falls off as inverse pressure in the atmosphere. The error should be largest where ionospheric electron concentrations are the largest, which should be during the day rather than at night, at the maximum in the 11-yr solar cycle, in the midlatitude of both hemispheres. It would have approximately the same signature as does the SDO as seen in Figs. 3 and 7.

Fig. 11 shows the 10.7 cm solar radio flux from 2005 to the present as published by the Space Weather Prediction Center of the U.S. National Oceanic and Atmospheric Administration. Solar cycle 24 commenced in ~2019, achieved solar maximum in ~2013. In year 2020, we are experiencing solar minimum and the commencement of solar cycle 25. The beginning of the COSMIC RO data saw a minimum in the solar cycle, and the later years saw a maximum.

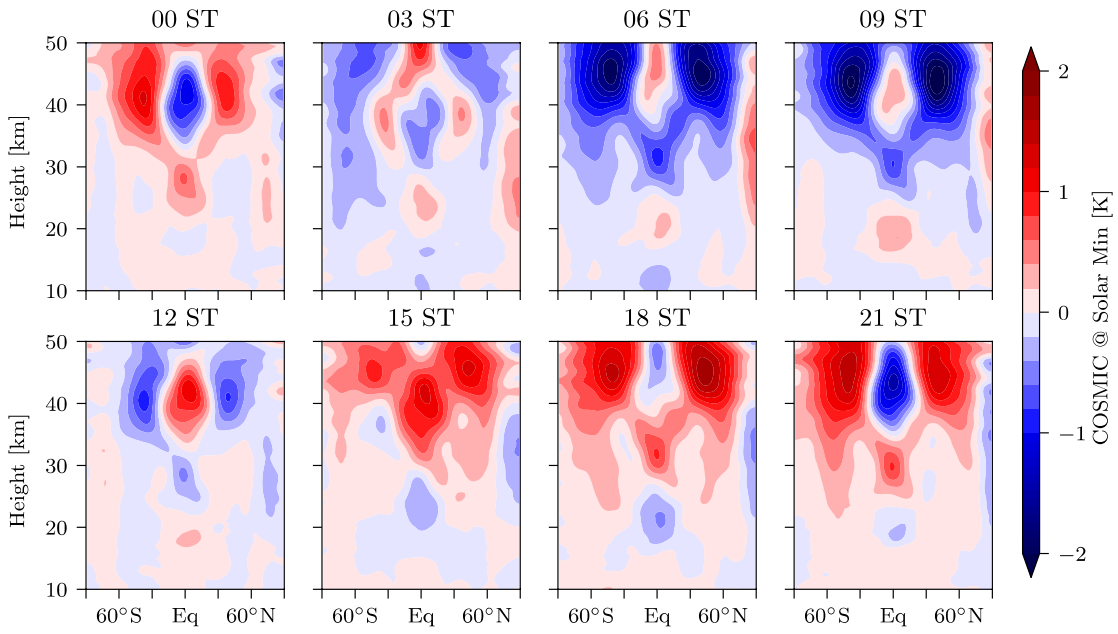

Figure 12. The zonal mean, seasonal mean migrating diurnal cycle during solar minimum. This figure is the same as Fig. 3 but with the analysis limited to COSMIC RO data obtained during solar minimum, years 2006 through 2010.

In order to compare the diurnal cycle in the COSMIC RO data to solar maximum, we restrict the analysis that produced Fig. 3 to the years 2006 through 2010, a solar minimum. The result is Fig. 12. Nearly every feature apparent in Fig. 3 is also apparent in Fig. 12, even with the same amplitude and phase. Both the downward propagating diurnal tide and the SDO are present with the same amplitude as in Fig. 3. This is convincing evidence that the SDO is an atmospheric phenomenon and not a consequence of ionospheric residual in the RO retrieval process.

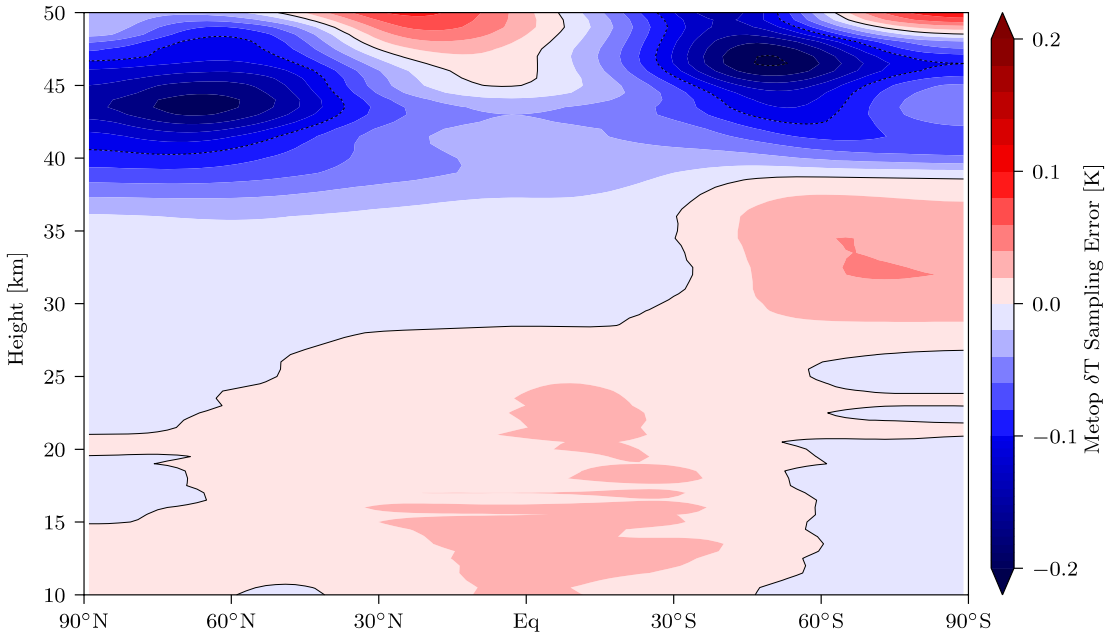

Figure 13. The predicted error in climatologies of dry temperature formed from Metop RO data by the sampling-error-removal technique. It is the opposite of the mean of the average seasonal cycle of the zonally symmetric terms of the ERA-Interim less COSMIC RO model bias D, retaining only the semidiurnal cycle terms and evaluated at 21:30 ST.

Lastly, the analysis of the bias of the weather model performed above, as shown in Figs. 4 and 5, can be manipulated to diagnose the bias in a climatology of Metop RO-retrieved variables formed by sampling-error-removal. The Metop satellites fly in sun-synchronous orbits with a 21:30 ST equator-crossing time at ascending node. Recall that sampling-error-removal works by diagnosing fluctuations in the atmosphere away from a climate average in the forecasts of a weather model and subtracting those fluctuations from the RO data before BnA. This includes the fluctuations associated with the diurnal cycle in the atmosphere, a serious concern with Metop RO because it is severely restricted in its coverage of the diurnal cycle. Our previous analysis shows that ERA-Interim predicts an SDO but with an amplitude that is too large and with a phase that is too advanced. It is the difference in the SDO diurnal cycle that can lead to errors in climatologies of Metop RO data when assembled using the sampling-error-removal technique. In particular, because Metop RO obtains occultation soundings twice per day, it is the twice-per-day, semidiurnal cycle in the bias between ERA-Interim and COSMIC RO representations of the SDO that will be errors in sampling-error-removal climatologies of Metop RO data.

Fig. 13 shows the error that is expected in sampling-error-removal climatologies due to under-sampling of the diurnal cycle and mis-modeling of the SDO by ERA-Interim. In order to construct it, we retain the zonal-average terms in the model bias D after application of Bayesian interpolation; we evaluate the semidiurnal cycle terms only at 21:30 ST for each month over the COSMIC era; we average all like months over the COSMIC era to form a seasonal cycle; and we average together all months of that seasonal

cycle. Because sampling-error-removal works by subtracting out the variations in the weather model, the opposite of the model bias D is computed. The result for dry temperature is shown in Fig. 13. The predicted error is large only in the upper stratosphere. The dominant influence is the mis-modeling of the SDO by ERA-Interim: errors as large as -0.2 K in Metop RO climatologies assembled by sampling-error-removal are expected in the mid-latitude upper stratosphere, above 35 km height. In the lower stratosphere, the expected error rarely exceeds 0.02 K, which is likely statistically insignificant.

Because RO is a limb sounding measurement, the RO soundings do not occur at precisely 21:30 ST or 9:30 ST on the descending branch of the Metop orbits but within ~2 hours of those solar times. The effect of the spread of the Metop RO soundings in solar times on the estimated sampling error bias for Metop RO data will be to reduce the size of the sampling error. The sampling error bias derives from the existence of error in the semidiurnal cycle in the upper air in ERA-Interim, the weather model used to remove sampling error. Like any harmonic of the diurnal cycle, the semidiurnal cycle is a sinusoid, and all sinusoids are reduced in amplitude yet remain sinusoids when smoothed by a boxcar smoothing function.

4. Conclusions

A brief overview of Bayesian interpolation, how it is computed, and how it can be expanded to analyse the diurnal cycle in RO data is presented. While Bayesian interpolation in one dimension was originally presented in the field of neural computation (MacKay 1992), it can be easily extended to mapping RO data on a sphere by using spherical harmonics as basis functions (Leroy 1997; Leroy et al. 2012). Furthermore, it can be extended to an analysis and interpolation of the diurnal cycle in the atmosphere by extending the spherical harmonic spatial functions in diurnal time with simple sinusoids in diurnal time. The development of the optimum regularizer for such interpolation was performed very recently (Leroy et al. 2020).

Bayesian interpolation with an analysis of the diurnal cycle out to the second harmonic was applied to COSMIC radio occultation (RO) data from May, 2006 through December, 2016. Bayesian interpolation was also applied to Metop RO data from October, 2007 through December, 2016 but without an analysis of the diurnal cycle inasmuch as Metop RO data does not span all solar times. The Metop satellites obtained insufficient numbers of RO data for November, 2007 through February, 2008 for Bayesian interpolation, so those months are not included in the analysis presented in this report. In addition to maps of the RO data produced by Bayesian interpolation, Bayesian interpolation was also applied to produce maps of the difference of ERA-Interim forecast quantities with respect to COSMIC RO data for the same months. This model bias leads to some interesting insights regarding a companion approach to the production of level 3 climatologies of RO data, that of sampling-error-removal.

ERA-Interim produces an upper stratosphere that is too warm in comparison to climatologies produced from COSMIC RO data by as large as 5 K at the stratopause. The difference falls off with depth and is no longer noticeable at 35 km height. ERA-Interim also produces microwave refractivity that is ~1% greater than in COSMIC RO data throughout much of the stratosphere, but this disagrees with other previous work of the ROM SAF showing much better agreement.

The analysis contained in this report shows a strong signal associated with the zonal wavenumber 1 migrating atmospheric tide. It is consistent with previous studies of the atmospheric tides. In addition, the analysis shows a regular 24-hr oscillation in temperature with amplitude of ~2 K in the mid-latitude upper stratosphere. We call it the SDO, the *stratospheric diurnal oscillation*. It spans the region 30º to 60º latitude in both hemispheres and between 35 and 50 km height. It migrates with the sun and experiences its maximum temperature fluctuation at ~20:00 solar time. Other publications have found the same phenomenon, but the satellite missions that were analysed were in slowly precessing orbits and it was never clear to what degree this oscillation was an artefact of temporal aliasing. Because COSMIC RO data spans all solar times, it is clear that the SDO is a real atmospheric phenomenon. Moreover, ERA-Interim does produce an SDO, but the amplitude of the SDO is too large and leads the actual SDO by ~2 hours in solar time.

Because a source of error in RO retrieval—ionospheric residual—is thought to exhibit a signature in space and diurnal time similar to the SDO, it is not clear whether or not the

SDO is an atmospheric phenomenon or an artefact of errors in RO retrieval. Because the amplitude of ionospheric residual should theoretically scale with the solar cycle, the signature of the SDO should be greatly reduced if the analysis contained herein were restricted to a few years about solar minimum, roughly 2006 through 2010. Such an analysis shows an SDO with exactly the same amplitude as exists when the analysis spans all phases of the solar cycle. It is clear that the SDO is an actual atmospheric phenomenon.

A comparison of a climatology of RO-retrieved dry temperature from Metop and COSMIC shows that Metop RO is systematically warmer than COSMIC RO in the upper stratosphere. The bias is 1 K at 50 km height and attenuates with depth as inverse pressure. The bias is canceled out in the region of the SDO, most likely a consequence of a bias induced by under-sampling of the diurnal cycle by Metop RO data.

The implication of these studies for the construction of climatologies of RO data by sampling-error-removal is that it should work well through most of the atmosphere but with a minor exception in the upper stratosphere. The weather model we used for sampling error removal is ERA-Interim, and this work shows that its simulation of the SDO is inaccurate. As a consequence, one can expect climatologies formed from Metop RO data using the sampling-error-removal technique to incur an error. This error is predicted by expanding the Bayesian interpolation maps of the model bias D at the sounding time of Metop, 21:30 ST. The most prominent error is isolated to the mid-latitude upper stratosphere, above 35 km height, and it can be as large as -0.2 K.

4.1 Acknowledgments

Acknowledgments go to Dr. Hans Gleisner for overseeing my appointment as a EUMETSAT Visiting Scientist. Likewise, acknowledgments go to Dr. Kent Lauritsen, head of the Radio Occultation Meteorology Satellite Application Facility (ROM SAF) and EUMETSAT for supporting this position. I wish to thank the internet technology staff at the Danish Meteorological Institute for providing computing resources and a virtual private network connection to those resources to enable this research.

As the VS38 position at the ROM SAF was a “stay-at-home” visiting scientist position, my home institution provided additional support for the completion of this work. The support came through a grant from the NASA Terra-Aqua-Suomi NPP Sounder Science Team program with Prof. Larrabee Strow (University of Maryland, Baltimore County) as Principal Investigator.

5. References

Ao, C.O., O.P. Verkhoglyadova, M.I. Oyola, and S.S. Leroy, 2020: A multi-mission GNSS-RO climate data record for the upper troposphere and lower stratosphere. In Preparation.

Ao, C.O. and A.J. Hajj, 2013: Monitoring the width of the tropical belt with GPS radio occultation measurements. *Geophys. Res. Lett.,* **40**, 6236–6241, doi:10.1002/2013GL058203.

Danzer, J., U. Foelsche, B. Scherllin-Pirscher, and M. Schwärz, 2014: Influence of changes in humidity on dry temperature in GPS RO climatologies. *Atmos. Meas. Tech.*, **7**, 2883–2896, doi:10.5194/amt-7-2883-2014.

Dee, D.P., and Co-authors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteor. Soc.,* **137**, 553–597, doi:10.1002/qj.828.

Foelsche, U., M. Borsche, A. Steiner, A. Gobiet, B. Pirscher, G. Kirchengast, J. Wickert, and T. Schmidt, 2008: Observing upper troposphere – lower stratosphere climate with radio occultation data from the CHAMP satellite. *Climate Dyn.,* **31**, 49–65, doi:10.1007/s00382-007-0337-7.

Foelsche, U., B. Scherllin-Pirscher, F. Ladstädter, A. Steiner, and G. Kirchengast, 2011: Refractivity and temperature climate records from multiple radio occultation satellites consistent with in 0.05%. *Atmos. Meas. Tech.,* **4**, 2007–2018, doi:10.5194/amt-4-2007-2011.

Gleisner, H., K.B. Lauritsen, J.K. Nielsen, and S. Syndergaard, 2020: Evaluation of the 15-year ROM SAF monthly mean GPS radio occultation climate data record. *Atmos. Meas. Tech.,* **13**, 3081–3098, doi:10.5194/amt-13-3081-2020.

Gleisner, H., 2018: Validation report: Reprocessed level 3 gridded data. Technical report of the Radio Occultation Meteorology Satellite Application Facility (ROM SAF), Danish Meteorological Institute, Copenhagen, Denmark, 109pp.

Lackner, B., A. Steiner, G. Kirchengast, and G. Hegerl, 2011: Atmospheric climate change detection by radio occultation data using a fingerprinting method. *J. Climate,* **24**, 5275–5291, doi:10.1175/2011JCLI3966.1.

Leroy, S.S., C.O. Ao, and O.P. Verkhoglyadova, 2018: Temperature trends and anomalies in modern satellite data: Infrared sounding and GPS radio occultation. *J. Geophys. Res.,* **123**, doi:10.1029/2018JD028990.

Leroy, S.S., C.O. Ao, and O.P. Verkhoglyadova, 2012: Mapping GPS radio occultation data by Bayesian interpolation. *J. Atmos. Ocean. Tech.,* **29**, 1062–1074, doi:10.1175/JTECH-D-11-00179.1.

MacKay, D., 1992: Bayesian interpolation. *Neur. Comp.,* **4**(3), 415–447, doi:10.1162/neco.1992.4.3.415.

Pirscher, B., U. Foelsche, M. Borsche, G. Kirchengast, and Y.-H. Kuo, 2010: Analysis of migrating diurnal tides detected in FORMOSAT-3/COSMIC temperature data. *J. Geophys. Res.,* **115**, D14108, doi:10.1029/2009JD013008.

Poli, P., S.B. Healy, and D.P. Dee, 2010: Assimilation of Global Positioning system radio occultation data in the ECMWF ERA-Interim reanalysis. *Q. J. R. Meteor. Soc.,* **136**, 1972–1990, doi:10.1002.qj.722.

Sakazaki, T., M. Fujiwara, X. Zhang, M.E. Hagan, and J.M. Forbes, 2012: Diurnal tides from the troposphere to the lower mesosphere as deduced from TIMED/SABER satellite data and six global reanalysis data sets. *J. Geophys. Res.,* **117**, D13108, doi:10.1029/2011JD017117.

Scherllin-Pirscher, B., S. Syndergaard, U. Foelsche, and K.B. Lauritsen, 2015: Generation of a bending angle radio occultation climatology (BAROCLIM) and its use in radio occultation retrievals. *Atmos. Meas. Tech.,* **8**, 109–124, doi:10.5194/amt-8-109-2015.

Shen, Z., K. Zhang, Q. He, M. Wan, L. Li, and S. Wu, 2020: Quest over the sampling error of COSMIC radio occultation climatologies. *J. Atmos. Ocean. Tech.,* In Review.

Steiner, A.K., F. Ladstadter, W.J. Randel, A.C. Maycock, Q. Fu, C. Claud, H. Gleisner, L. Haimberger, S.-P. Ho, P. Keckhut, T. Leblanc, C. Mears, L.M. Polvani, B.D. Santer, T. Schmidt, V. Sofieva, R. Wing, and C.-Z. Zou, 2020: Observed temperature changes in the troposphere and stratosphere from 1979 to 2018. *J. Climate,* **33**, 8165–8194, doi:10.1175/JCLI-D-19-0998.1.

Steiner, A., B. Lackner, F. Ladstädter, B. Scherllin-Pirscher, U. Foelsche, and G. Kirchengast, 2011: GPS radio occultation for climate monitoring and change detection *Radio Sci.,* **46**(RS0D24), doi:10.1029/2010RS004614.

Verkhoglyadova, O.P., S.S. Leroy, and C.O. Ao, 2014: Estimation of winds from GPS radio occultation. *J. Atmos. Ocean. Tech.,* **31**, 2451–2461, doi:10.1175/JTECH-D-14-00061.1.

Xie, F., D.L. Wu, C.O. Ao, and A.J. Mannucci, 2010: Atmospheric diurnal variations observed with GPS radio occultation soundings. *Atmos. Chem. Phys.,* **10**, 6889 – 6899, doi:10.5194/acp-10-6889-2010.

6. List of Acronyms

BnA	Binning and Averaging
COSMIC	Constellation Observing System for Meteorology, Ionosphere, and Climate
ECMWF	European Centre for Medium-range Weather Forecasts
ERA-Interim	ECMWF Interim Reanalysis
ERA-I	ibid.
EUMETSAT	EUropean organisation for the exploitation of METeorological SATellites
GNSS	Global Navigation Satellite System
GPS	Global Positioning System (USA)
Metop	Meteorological Operational Satellite
NWP	Numerical Weather Prediction
RO	Radio Occultation
ROM SAF	Radio Occultation Meteorology (ROM) Satellite Application Facility (SAF) (EUMETSAT)
SDO	Stratospheric Diurnal Oscillation