

# The Stratospheric Diurnal Cycle in COSMIC GPS Radio Occultation Data: Scientific Applications

Stephen Leroy<sup>1,1,1</sup> and Hans Gleisner<sup>2,2,2</sup>

<sup>1</sup>Atmospheric and Environmental Research

<sup>2</sup>Danish Meteorological Institute

November 30, 2022

## Abstract

The diurnal cycle throughout the stratosphere is analyzed by applying Bayesian interpolation to COSMIC GPS radio occultation (RO) data and three scientific applications of the analysis are introduced. COSMIC RO data is the only data set that completely samples the diurnal cycle in the stratosphere continuously for a decade, providing a unique opportunity to explore several scientific topics related to the diurnal cycle in the stratosphere. First, the migrating thermal tides are analyzed with unprecedented accuracy and precision, with an uncertainty in the analysis of the vertically propagating tides ranging from 0.1 in the lower stratosphere to 0.6 K in the upper stratosphere for an individual month of RO data and the uncertainty in a ten-year climatological diurnal cycle a factor of 10 less. This enables potential observational studies of the forcing of the tides in the troposphere and the stratosphere. Moreover, the mid-latitude trapped tide is found to be smaller than what is produced by an atmospheric model and lags the model in phase, a likely consequence of a faulty parameterization of eddy diffusivity in the upper stratosphere. Second, a clear signal of solar cycle influence on the diurnal cycle is evident in this analysis, but whether the cause is the systematic bias of ionospheric residual associated with RO retrieval or it is an actual atmospheric phenomenon is less clear. The pattern suggests ionospheric residual is dominant, but modeling studies will be necessary to address the question. Third, RO satellites and missions that obtain inadequate coverage of the diurnal cycle will be biased by under-sampling it, whether or not sub-sampling weather forecasts is used to removal sampling error. The analysis of the diurnal cycle in COSMIC RO data can be used to diagnose the systematic sampling error incurred by incomplete coverage of the diurnal cycle, which is of the order of 0.2 K for a Metop-based RO climatology.

# The Stratospheric Diurnal Cycle in COSMIC GPS Radio Occultation Data: Scientific Applications

S. S. Leroy<sup>1\*</sup> and H. Gleisner<sup>2</sup>

<sup>1</sup>Atmospheric and Environmental Research

<sup>2</sup>Danish Meteorological Institute

## Key Points:

- The propagating stratospheric tides in ERA-Interim are proven accurate but the symmetric trapped mode is biased in amplitude and phase.
- The ionospheric residual systematic error inherent to GNSS radio occultation is analyzed.
- Radio occultation climatologies based on single satellites are biased by the diurnal cycle in the upper stratosphere.

---

\*Atmospheric and Environmental Research, 131 Hartwell Ave., Lexington, Massachusetts, USA

Corresponding author: Stephen S. Leroy, [sleroy@aer.com](mailto:sleroy@aer.com)

## Abstract

The diurnal cycle throughout the stratosphere is analyzed by applying Bayesian interpolation to COSMIC GPS radio occultation (RO) data and three scientific applications of the analysis are introduced. First, the migrating thermal tides are analyzed with unprecedented accuracy and precision, with an uncertainty in the analysis of the vertically propagating tides ranging from 0.1 in the lower stratosphere to 0.6 K in the upper stratosphere for an individual month of RO data and with an uncertainty in a ten-year climatological diurnal cycle a factor of 10 less. Moreover, the mid-latitude trapped tide is found to be smaller than what is produced by an atmospheric model and lags the model in phase, a likely consequence of a faulty parameterization of eddy diffusivity in the upper stratosphere. Second, a clear signal of solar cycle influence on the diurnal cycle is evident in this analysis, but whether the cause is the systematic bias of ionospheric residual associated with RO retrieval or it is an actual atmospheric phenomenon is less clear. Third, RO satellites and missions that obtain inadequate coverage of the diurnal cycle will be biased by under-sampling it, whether or not sub-sampling weather forecasts is used to removal sampling error. The analysis of the diurnal cycle in COSMIC RO data can be used to diagnose the systematic sampling error incurred by incomplete coverage of the diurnal cycle, which is of the order of 0.2 K for a Metop-based RO climatology.

## Plain Language Summary

The most familiar tides are those that track the Sun and Moon as sea level floods and ebbs, something that boating enthusiasts must always consider. A different kind of tide exists in the atmosphere, and these tides not only propagate horizontally as do ocean tides, but also vertically. Atmospheric tides have been extremely difficult to observe, generally because satellites have hardly ever flown in the right orbits to observe them, until the radio occultation mission COSMIC came to pass. If COSMIC data are analyzed using a statistical technique known for mining data for all of the information they have to offer, it becomes possible for the first time to analyze the atmospheric tides in the stratosphere, thus yielding a treasure trove of interesting science. For one, we should be able to confirm theories of how the atmospheric tides are forced. Second, we can diagnose a difficult physical problem with atmospheric models, namely how turbulence mixes air in the upper atmosphere. And third, we can also figure out what biases exist in other satellite-based data sets on temperature and pressure due to the orbits those other satellites fly in.

## 1 Introduction

Radio occultation (RO) using the transmitters of the Global Position System (GPS) (Kursinski et al., 1997, 2000) and more recently the transmitters of all Global Navigation Satellite Systems (GNSS) has provided unique constraints on observations of the Earth's troposphere and stratosphere by virtue of its global coverage, accuracy to 0.1 K, vertical resolution approaching 100 m (Foelsche et al., 2011), having geopotential height as the independent coordinate of retrieval (Leroy, 1997; Scherllin-Pirscher et al., 2017), direct sensitivity to water vapor in the lower troposphere (Kursinski et al., 1995; Kursinski & Gebhardt, 2014), and stationary (but non-homogeneous) coverage of the diurnal cycle (Pirscher et al., 2010). This paper will expound on science applications that arise when a novel analysis technique, Bayesian interpolation, is applied to the coverage of the diurnal cycle by GPS radio occultation. Those applications are (1) a detailed analysis of the thermal tides, (2) a diagnosis of systematic bias incurred in RO retrieval that is otherwise difficult to detect, and (3) an examination of sampling biases that can be incurred in climatologies produced from RO data. Below are extended introductions to these applications.

First, the diurnal cycle has been studied by atmospheric scientists for over a century, beginning with its manifestation in the surface air, then in the troposphere and lower stratosphere, and then up through the mesosphere (Chapman & Lindzen, 1970). Almost all of the observed diurnal variability in the atmosphere is due to the thermal tides, a form of global-scale internal gravity wave modified by the Earth’s rotation and forced by the absorption of solar radiation by ozone in the upper stratosphere and by latent heating and absorption of solar radiation by water vapor in the troposphere. At the surface the tides are manifested as fluctuations in surface pressure as measured by meteorological stations; in the troposphere and stratosphere they can be measured *in situ* by radiosonde; and in the stratosphere and mesosphere by remote sensing. Each entails its own limitation. Surface pressure measurements are extremely limited in global coverage; radiosonde measurements require more than the usual twice daily or four-times daily sampling and are also limited in global coverage; and measurements from space require non-sun-synchronous sampling, which is only obtained by satellites in rapidly precessing orbits. The latter was obtained with measurements by the SABER instrument on the TIMED satellite (Zhang et al., 2006; Sakazaki et al., 2012) and radio occultation measurements from the CHAMP satellite (Pirscher et al., 2010). Yet SABER analysis of the tides was limited to the uppermost stratosphere and the mesosphere, leaving most of the stratosphere un-sampled. Because both the TIMED and CHAMP satellites sampled diurnal time by rapid regression of nodes, the analyses of the tides from these data are subject to bias by aliasing of the seasonal cycle and planetary-scale waves to diurnal time-scales and are limited in their ability to quantify long time-scale variability of the tides.

RO data obtained by the COSMIC mission (Anthes et al., 2008) provides a unique opportunity to study the thermal tides in the stratosphere because it was constituted of six satellites spaced by  $30^\circ$  in right ascension of ascending node for a period of over 10 years. The data are global in distribution and span the diurnal cycle, albeit unevenly, at all latitudes for most of the lifetime of the COSMIC mission. RO soundings of stratospheric temperature and pressure as functions of geopotential height are precise and accurate. If a physical optics RO retrieval of bending angle is performed, the vertical resolution is approximately 100 m; if a geometric optics retrieval is performed—as is most often the case in the stratosphere—the vertical resolution is closer to 1.5 km, easily sufficient to resolve the thermal tides. In theory, little to no bias is incurred in retrievals of temperature and pressure in day-time vs. night-time soundings except ionospheric residual; see below. The data are ideally suited to studies of the global structure of atmospheric tides in the stratosphere and their inter-annual variability. As a result, no thorough, continuous observational study of the tides in the global stratosphere was possible until COSMIC. More recently, the COSMIC-2 mission obtained dense diurnal coverage, thus enabling the study of tides and other equatorial waves (Randel et al., 2021). COSMIC-2 RO only covers the Tropics, however.

Secondly, systematic bias relevant to climatologies of RO and long-term trend analysis is incurred in day-time vs. night-time RO sounding because of the ionosphere’s influence on RO sounding. The microwave signals of the GNSS satellites are refracted in the ionosphere because free electrons contribute significantly to the microwave index of refraction. In order to eliminate the influence of the ionosphere on retrievals of neutral atmospheric variables, the ionosphere’s contribution is eliminated by a linear combination of GNSS signals at two different carrier frequencies, which itself is possible because the ionosphere’s contribution to microwave refractivity is dispersive. The dispersion leads to a complication, however, because the two signals do not follow the same path through the ionosphere and atmosphere, meaning any linear combination will be imperfect, leaving a residual second order error after eliminating more significant first order effects (Vorob’ev & Krasil’nikova, 1994). The effect was first identified by Kursinski et al. (1997). A method has been proposed to reduce this second order *ionospheric residual* in RO (Healy & Culverwell, 2015), but it has proven exceedingly difficult to isolate ionospheric residual in data because of naturally occurring atmospheric variability on the same time-scales and

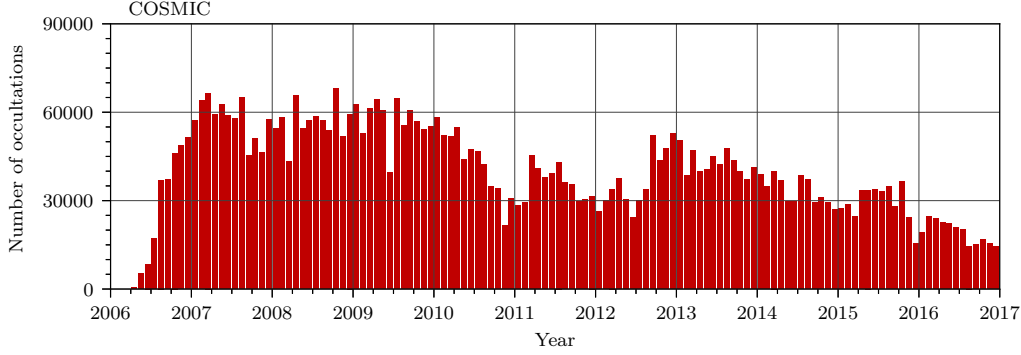
because other sources of error in RO retrieval have similar envelopes in their vertical structure.

RO data obtained by COSMIC are well suited to detecting the signature of ionospheric residual, because of COSMIC's complete coverage of the diurnal cycle over a complete solar cycle. The ionospheric residual is a greater systematic error in RO retrieval when the ionosphere, especially the F-layer, is more active, as it is during the day rather than at night, at solar maximum rather than at solar minimum. Analysis of the diurnal cycle in COSMIC RO data should show inter-annual variability in the amplitude and phase of the diurnal cycle that correlates with solar activity. A study like this is only possible with a data set that is longer than the solar cycle and completely covers solar time over the period of the data set. COSMIC RO data satisfy these conditions.

Thirdly, much of the RO data that have been obtained in the last 25 years have not covered all solar times, making the climatologies that are constructed from them subject to aliasing and biasing effects. RO data collected by a sun-synchronous orbiter sample only two solar times, aliasing the semi-diurnal cycle in the atmosphere to a systematic sampling bias. Such orbiters include the three Metop satellites, and NASA satellites including TerraSAR-X and TanDEM-X, and FengYun-3C. Similarly, climatologies constructed from RO data collected from rapidly precessing orbiters are also biased by the atmospheric semi-diurnal cycle, but because the regression of nodes of the satellites' orbits proceeded on seasonal time-scales, the incurred bias appears as intra-annual to inter-annual variability of the atmosphere. Such orbiters include CHAMP, C/NOFS, and GRACE. A method to remove biases due to uneven sampling, including sampling in solar time, has been proposed for RO data that works by sub-sampling the forecasts of a numerical weather prediction system at RO data geolocations, estimating a bias by subtracting a climatology formed by the interpolated forecasts from the monthly average climatology published by the numerical weather prediction system, and subtracting that bias from the RO climatology constructed from the RO data themselves (Foelsche et al., 2008, 2011). Such approaches rely upon the atmospheric forecasts not creating systematic biases of their own, however.

RO data obtained by COSMIC provide an opportunity to investigate the bias incurred by RO sampling patterns that do not fully cover the diurnal cycle. Herein we do so for the Metop-A and Metop-B RO sounders, both in sun-synchronous orbits. An analysis of the diurnal cycle obtained from COSMIC data can be expanded at the solar times of Metop RO soundings to obtain the bias that would be incurred in a climatology formed from binning and averaging Metop RO data. Similarly, an analysis of the diurnal cycle of the difference of COSMIC data and an atmospheric analysis at the solar times of Metop RO soundings would give the bias that would be incurred in climatologies of RO data obtained by Metop even with sampling error correction applied. The consequences are expected to be significant for periods of RO when the diurnal cycle was not fully covered or when RO satellites covered the diurnal cycle only by regression of nodes, such as the era of CHAMP (Wickert et al., 2002).

The studies mentioned above are enabled by applying Bayesian interpolation to map the RO data obtained by COSMIC and Metop, and this paper focuses on the scientific applications made possible by this type of analysis. The applications introduced above, except an empirical analysis of ionospheric residual, have been undertaken elsewhere. Relevant work will be addressed in the body of this paper. In no case, however, has Bayesian interpolation been used in the analyses, and those works have suffered from a lack of error analysis and a lack of detection sensitivities as a consequence. This paper is organized as follows. The second section contains a description of the data and how they are analyzed. The third section presents the scientific applications introduced above are addressed by Bayesian interpolation applied to RO data. The fourth section presents a discussion.



**Figure 1.** Timeseries of the numbers of monthly COSMIC RO profiles considered in this work.

## 2 Data and method

The data sets and the way in which they are analyzed are presented. First comes an introduction to the RO data used. Second comes a brief introduction to a weather analysis and its forecasts that are used in this work. Finally, a brief overview of Bayesian interpolation is given.

### 2.1 Radio occultation data

The first Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC Anthes et al., 2008) obtained radio occultation (RO) data using the transmitters of the Global Positioning System (GPS) from April, 2006, through April, 2020. The six COSMIC satellites were deployed from the same launch vehicle, and mutual regression of nodes was used to spread the satellites by  $30^\circ$  in ascending node in order to obtain approximately uniform coverage of the diurnal cycle. The constellation achieved its final configuration in November, 2007, but with one of the satellites (F3) never achieving final orbit (Fong et al., 2008).

Microwave refractivity and dry temperature are retrieved by standard retrieval techniques (e.g. Hajj et al., 2002; Gorbunov et al., 2011). The excess phase data is obtained from the COSMIC DAAC (<https://cdaac-www.cosmic.ucar.edu/>). In the first step of retrieval, canonical transform of the second type is used to unwrap the effects of diffraction and atmospheric multi-path below 25 km (Gorbunov & Lauritsen, 2004; Gorbunov et al., 2004) and geometric optics above 25 km to recover bending angle as a function of impact parameter. In the second step of retrieval, the Abel integral inversion is used to retrieve a refractivity profile from a bending angle profile. The integral is initialized between 40 and 120 km altitude with the assistance of an upper atmosphere bending angle climatology (BAROCLIM; Scherllin-Pirscher et al., 2015). In the third step of retrieval, the hydrostatic integral is initialized at 120 km under the assumption of an exponential profile of refractivity above that. Profiles of microwave refractivity, dry pressure, and dry temperature as functions of geopotential height (Leroy, 1997; Scherllin-Pirscher et al., 2017) result.

Implementation of the BAROCLIM climatology to aid in initializing the Abel inversion integral was formulated so as to minimize any bias in retrieval incurred by the initialization. BAROCLIM is a climatology of bending angle data between 40 and 120 km impact height derived solely from actual COSMIC RO data bending angle data. It is computed by averaging bending angle data in  $10^\circ \times 10^\circ$  longitude-latitude bins by

month. Each month-bin contains approximately 11,000 soundings that are averaged, thus yielding an ensemble of smooth bending angle profiles, one per month-bin. In a retrieval, the BAROCLIM ensemble of smoothed bending angle profiles is searched for the best match to the observed bending angle profile, where best match is determined *after* a two-parameter tuning. The two-parameter tuning itself is a simple shift and slope adjustment in the logarithm of the bending angle. The search-tune-match algorithm is specifically designed so that no bias is introduced into the Abel inversion while preserving climatological realism above 40 km. While the ensemble of BAROCLIM smooth bending angle profiles may contain less variance than an ensemble of single soundings, the tuning should theoretically restore the single-sounding variance, including the variance associated with the diurnal cycle. Any diurnal cycle that exists in the measurements should persist through BAROCLIM initialization of the Abel inversion without increase or decrease.

This is the same data set used in Gleisner et al. (2020) and described in algorithm theoretical baseline documents (<http://www.romsaf.org>) except that no attempt is made to account for ionospheric residual by “kappa-correction”. Figure 1 shows the timeline of the number of COSMIC RO data considered monthly. At its peak performance, COSMIC generated  $\approx 60,000$  soundings monthly tapering off to approximately 20,000 monthly by 2016.

## 2.2 ERA-Interim

In order to evaluate the performance of sampling error correction as a technique for constructing climatologies of RO data, the forecasts of a numerical weather prediction system are compared to the RO data. Here, the forecasts of ERA-Interim (Dee et al., 2011) are interpolated to the time and location of COSMIC and Metop-A,B RO data. The forecasts of ERA-Interim are interpolated rather than the analyses because ERA-Interim has assimilated COSMIC and Metop-A,B RO data: because the analyses nearly perfectly replicate the RO soundings themselves at the times and locations of the soundings, an RO climatology constructed using the sampling error correction technique would simply recreate the climatology of ERA-Interim. ERA-Interim generated forecasts at 3-hr intervals from the analyses at 00 and 12 UTC hours. Consequently, in order to cover all UTC times, the 3-hr, 6-hr, 9-hr, and 12-hr forecasts were interpolated. Microwave refractivity is computed from the model forecasts of temperature, pressure, and water vapor using the Smith-Weintraub formula. The sampling error correction technique itself is described elsewhere (Foelsche et al., 2008, 2011; Gleisner et al., 2020).

## 2.3 Bayesian interpolation

Bayesian interpolation fits data with *a priori* unknown error characteristic that are non-uniformly distributed without overfitting the data. It does so by fitting a superposition of basis functions to the data and penalizing the fit. In a second level of inference, the weighting between the mis-fit and the penalty is determined optimally. Originally, a thorough exposition of the method was given for fitting data with one independent coordinate (MacKay, 1992). It was adapted to mapping RO data by considering spherical harmonics as the basis functions (Leroy, 1997) and the penalty, or regularizer, tuned rigorously and its performance evaluated later (Leroy et al., 2012). Because RO climatologies depend strongly on how the RO soundings sample the diurnal cycle, the technique was expanded by multiplying the spherical harmonic basis function by sinusoids in diurnal time, by optimizing the regularizer, and evaluating the performance (Leroy et al., 2021). The approach has been presented in depth in all of those previous works; Leroy et al. (2021) can be considered a companion work.

The RO retrievals of temperature and refractivity are interpolated onto surfaces of constant geopotential height ranging from 10 to 50 km at intervals of 0.5 km. The dry



temperature (Danzer et al., 2014) of the RO soundings and ERA-Interim output are interpolated linearly in geopotential height as are the logarithms of microwave refractivity for both the RO profiles and ERA-Interim output. ERA-Interim is interpolated bilinearly in longitude and latitude to the position of each RO sounding. ERA-Interim is interpolated linearly in time to the time of each RO sounding. Bayesian interpolation is then applied up to spherical harmonic degree 14 and to the fourth harmonic of the diurnal cycle. Even though Leroy et al. (2021) recommends that RO data be collected into 3- to 5-day bins before applying Bayesian interpolation, in part to account for temporal synoptic variability as a bias in a climatology, Bayesian interpolation is applied to monthly collections of RO data. While doing so would create biases in a climatology because of uneven daily sampling, it is not expected to impact climatologies of the diurnal cycle, which is the goal of this work.

Not only are the RO refractivity and dry temperature mapped by Bayesian interpolation, so are the differences between RO data and atmospheric model forecasts of the same fields. The difference between data and model forecasts bears on monthly average assemblages of RO data, also referred to as *climatologies*. Traditionally, most atmospheric science uses atmospheric reanalyses, and research on phenomena that occur on seasonal to multi-decadal time-scales involve monthly average *gridded* reanalysis products. RO climatologies are desired because RO data are understood to be far better calibrated than reanalyses and hence more useful for climate trend studies. RO climatologies are noisy, however, because RO data are sparse and greatly under-sample atmospheric variations associated with synoptic variability. It has become common to attempt to remove the noise in RO climatologies by estimating sub-sampling biases and subtracting those from RO data-only *bin-and-average* climatologies. The biases are estimated by a bin-and-average calculation applied to reanalysis fields interpolated in time and location to the times and locations of the RO soundings and then subtracting from that calculation the gridded monthly average fields of the reanalysis. Finally, the published RO climatology is the bin-and-average RO data with the sub-sampling bias subtracted. This is the *sampling error removal* algorithm.

Mathematically, the *sampling error removal* algorithm can be written as

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

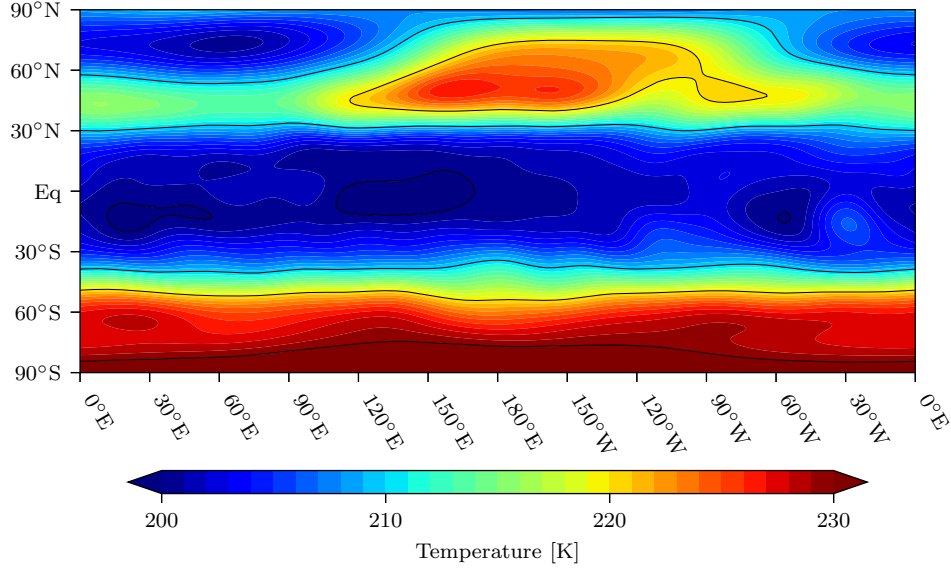
in which  $y$  is the variable to be mapped,  $y(r_i)$  is the value of that variable as retrieved by RO at location  $r_i$ , and  $y_M(r_i)$  is the model-interpolated value of that variable at locations  $r_i$ .  $\mathcal{M}(y_M)$  is a gridded monthly average published by the model,  $\mathcal{M}(y)$  is the published RO climatology,  $\mathcal{B}(y(r_i))$  is the bin-and-average of RO data, and  $\mathcal{B}(y_M(r_i))$  is the bin-and-average of the model output fields interpolated to the times and locations  $r_i$  of the RO data. The term in square brackets is the under-sampling bias of RO data. This can be rearranged to obtain

$$\mathcal{M}(y) = \mathcal{M}(y_M) - \mathcal{B}(y_M(r_i) - y(r_i)) \quad (2)$$

because of linearity of the bin-and-average operation. Equation 2 provides a new interpretation of the *sampling error removal* algorithm: the published RO climatology is just the gridded reanalysis monthly average  $\mathcal{M}(y_M)$  with a model bias  $\mathcal{B}(y_M(r_i) - y(r_i))$  removed.

The model bias—the second term on the right of equation 2—is mapped by Bayesian interpolation as well as maps of RO data alone. As is apparent in equation 2, the difference maps are expected to yield insights into the performance of the sampling error removal technique: if systematic errors exist in the model, which is ERA-Interim in this case, then the sampling error removal technique can also introduce unwanted biases into the published climatology  $\mathcal{M}(y)$ . The same regularizer as is used in mapping RO data by themselves is used in mapping the differences between RO data and model-interpolated data.





**Figure 2.** Diurnal average, monthly average dry temperature at 15 km geopotential height as produced by Bayesian interpolation applied to COSMIC RO data for January, 2010.

The analysis by Bayesian interpolation allows for research of migrating and non-migrating tides alike, but this work concerns only the migrating tidal features. Bayesian interpolation makes such an analysis simple. Recall from Leroy et al. (2021) that the basis functions are defined as

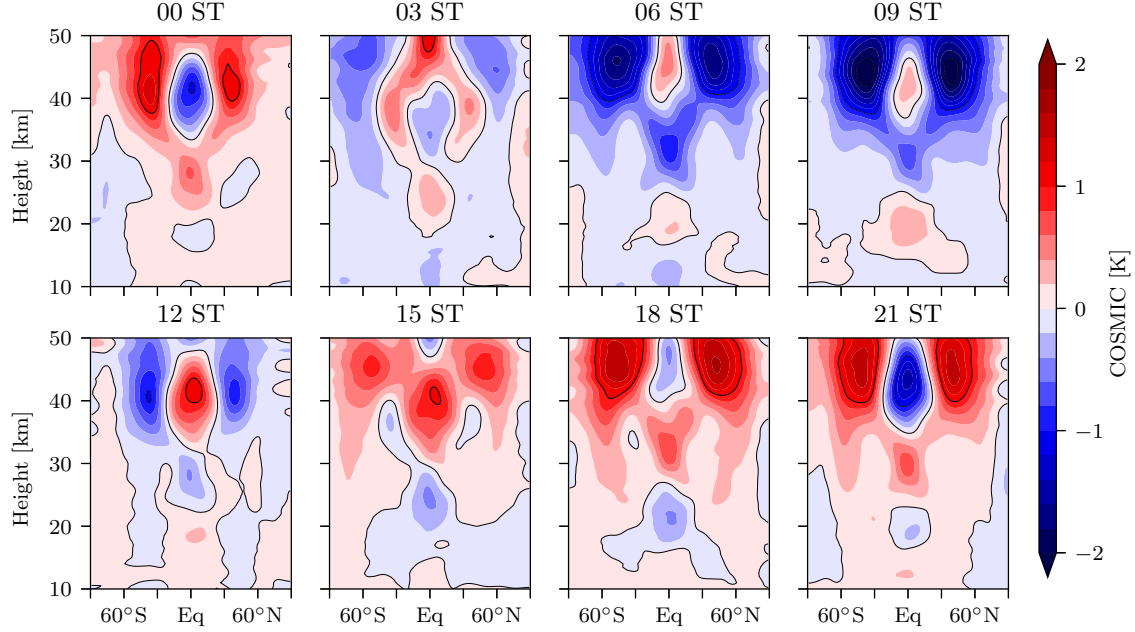
$$\psi_{\mu,\nu}(\lambda, \varphi, \tau) \equiv P_{lm}(\sin \varphi) \times \begin{cases} \cos(m\lambda) \cos(n\tau) & \text{for } \nu = 1 \\ \cos(m\lambda) \sin(n\tau) & \text{for } \nu = 2 \\ \sin(m\lambda) \cos(n\tau) & \text{for } \nu = 3 \\ \sin(m\lambda) \sin(n\tau) & \text{for } \nu = 4 \end{cases}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

with  $P_{lm}(x)$  the associated Legendre polynomial of degree  $l$  and order  $m$ ,  $\lambda, \varphi, \tau$  are longitude, latitude, time, all in radians, and  $\mu$  a proxy index for unique combinations of degree  $l$ , order  $m$ , and diurnal cycle harmonic  $n$ . For the analyses presented below, one only needs to perform the Bayesian interpolation with time  $\tau$  defined as diurnal time and retain only the  $m = 0$  terms in the expansion of the basis functions. Sampling error analysis then requires only the consideration of the rows and columns of the error covariance matrix of the expansion coefficients corresponding to the retained terms.

### 3 Analysis

#### 3.1 Example map

An example of Bayesian interpolation applied to COSMIC data is shown in figure 2. It is an analysis of the diurnal average dry temperature at 15 km geopotential height for the month of January, 2010. This level lies in the stratospheric middle-world, with 15 km lying in the upper troposphere in the Tropics and in the lower stratosphere at higher latitudes. The enormous meridional gradient in temperature can be considered a demarcation of the extent of the Tropics. The warm summer (southern) stratosphere is prominent as is a winter (northern) midlatitude stationary eddy with zonal wavenumber 1, the latter of which is anticyclonic over the Pacific with an amplitude of  $\approx 20$  K. To form such a map, only the terms of the basis function expansion using diurnal cycle harmonic  $n = 0$  are retained. Maps such as this are formed for any month between September



**Figure 3.** Dry temperature anomaly associated with the migrating thermal tides, harmonics 1 through 4, seasonally averaged and averaged over the COSMIC RO record, September 2006 through December 2016. The amplitude of the anomaly is contoured as functions of latitude and geopotential height a 3-hr intervals in solar time (ST), beginning with 0 hrs ST and progressing up to 21 hrs ST.

2006 and December 2016 for all geopotential heights between 10 and 50 km at 0.5 km intervals. The same is done for the *difference* between COSMIC RO data and the forecasts of ERA-Interim; see figure 8 and §3.5.

### 3.2 Migrating thermal tides and uncertainty

Bayesian interpolation extended to account for the diurnal cycle lends itself ideally to studies of atmospheric thermal tides. Figure 3 is an animation of the migrating tides averaged over the entire extent of the COSMIC RO record and for all seasons. It shows the temperature anomaly associated with the migrating tides for diurnal harmonics 1 through 4 at 3-hr intervals in solar time. These plots can be interpreted as combinations of two distinct phenomena: the vertically propagating migrating tides near the equator, and a vertically trapped oscillation that occurs in the upper stratosphere in mid-latitudes. The vertically propagating migrating tides show downward vertical propagation beginning at the stratopause (50 km height) and ending at the tropical tropopause (17 km). The positive temperature anomaly first appears at the stratopause at  $\sim 03$  hrs solar time. After 24 hrs it descends to 25 km height, and after 33 hrs it reaches the tropical tropopause. The negative temperature anomalies lag by  $\approx 12$  hrs, as the tides are dominated by the  $n = 1$  diurnal harmonic. The amplitude of the migrating tides varies with altitude. In addition to the vertically propagating mode within  $10^\circ$  latitude of the equator, a pulsing mode appears in the upper stratosphere in mid-latitudes, approximately between 35 and 50 km height and between  $20^\circ$  and  $60^\circ$  latitude in both hemispheres. It is sinusoidal in solar time without any signature of vertical propagation. It

reaches peak amplitude at  $\approx 18$  hrs solar time and is also dominated by the  $n = 1$  diurnal harmonic.

The vertically propagating and vertically trapped modes are known from other studies. Until the COSMIC project, they were seen in remote sensing data from satellites in non-sun-synchronous orbits, such as LIMS (Hitchman & Leovy, 1985), UARS (Wu et al., 2008), CRISTA (Ward et al., 1999), and TIMED/SABER (Mukhtarov et al., 2009); see Sakazaki et al. (2012) for discussion and other references. The study of thermal tides using these satellites' data examined the tides' structure in the mesosphere and lower thermosphere primarily, mostly because the tides' amplitudes were large enough to be easily detected. The precision and sensitivity of RO in the stratosphere enabled their study in the stratosphere, first with the CHAMP mission (Zeng et al., 2008; Xie et al., 2010) and later using COSMIC RO data (Pirscher et al., 2010). Before COSMIC, all of these studies were done with data from rapidly precessing orbits, begetting the concern that the precision could alias seasonal and other long term variability to diurnal timescales.

It is customary in studies of migrating thermal tides to contour wave amplitude and phase separately, as is done here in figure 4 for the first harmonic of the diurnal cycle. With the output expansion coefficients of Bayesian interpolation, the  $n$ 'th harmonic migrating thermal tidal temperature fluctuation at latitude  $\varphi$ , height  $h$  and for month  $i$  can be written as

$$T'_n(\tau; \varphi, h, i) = C_n(\varphi, h, i) \cos n\tau + S_n(\varphi, h, i) \sin n\tau \quad (4)$$

with  $\tau$  being solar time measured in radians. The amplitude  $A_n$  and phase  $\tau_n$  of the  $n$ 'th migrating tidal harmonic are computed by

$$A_n(\varphi, h, i) = [C_n^2(\varphi, h, i) + S_n^2(\varphi, h, i)]^{1/2} \quad (5a)$$

$$\tau_n(\varphi, h, i) = (1/n) \arctan[S_n(\varphi, h, i), C_n(\varphi, h, i)] \quad (5b)$$

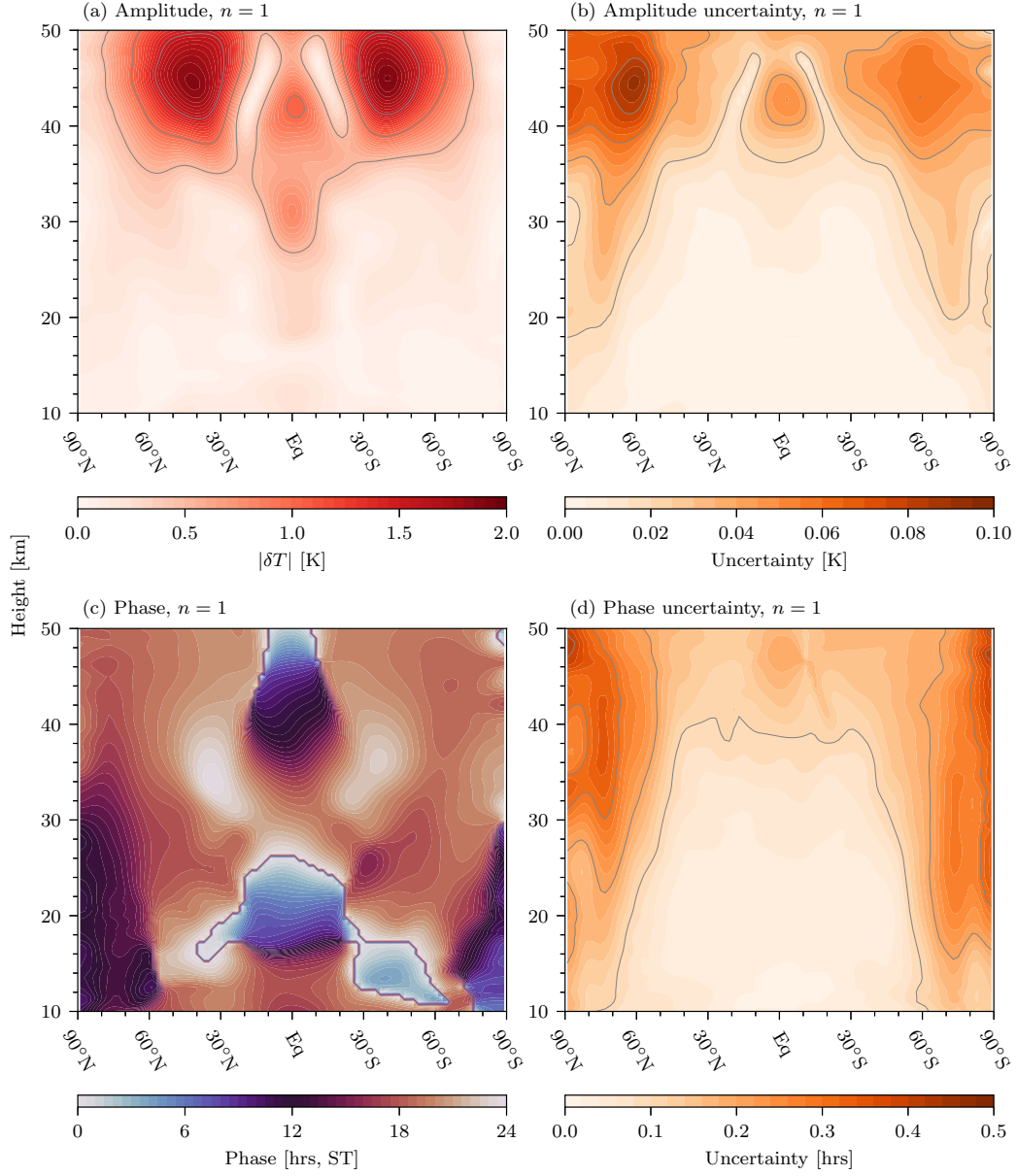
in which  $\arctan(\dots, \dots)$  is a four-quadrant arctangent. The error analysis for the inference of the amplitude and phase of the migrating tides derives from RO sampling error and from inter-annual variability in the amplitude and phase of the migrating tides and is complicated by the strong seasonality in the tides, typified by meridional motion in the peak of the tides according to the sub-solar latitude. In this error analysis, the monthly time-series of cosine coefficients  $C_n(\varphi, h, i)$  and sine coefficients  $S_n(\varphi, h, i)$  are de-seasonalized, yielding  $\tilde{C}_n(\varphi, h, i)$  and  $\tilde{S}_n(\varphi, h, i)$  for month  $i$  in the time-series. Anomalies in the time-series of these coefficients are caused by inter-annual variability in the tides and sampling error due to under-sampling of the atmosphere by COSMIC RO data. The de-seasonalized time-series of the cosine and sine terms of the tidal expansion can then be used to estimate the uncertainty in each of those terms along with their inter-correlation  $\rho$ :

$$\sigma_c = \left[ \frac{1}{M-12} \sum_{i=1}^M \tilde{C}_n^2(\varphi, h, i) \right]^{1/2} \quad (6a)$$

$$\sigma_s = \left[ \frac{1}{M-12} \sum_{i=1}^M \tilde{S}_n^2(\varphi, h, i) \right]^{1/2} \quad (6b)$$

$$\rho = \frac{1}{(M-12) \sigma_c \sigma_s} \sum_{i=1}^M \tilde{C}_n(\varphi, h, i) \tilde{S}_n(\varphi, h, i). \quad (6c)$$

COSMIC RO data for November, 2007, through December, 2016, are used in this analysis because the COSMIC constellation achieved its final formation only in November, 2007; hence,  $M = 110$  months. The reason 12 degrees of freedom are removed in the computation of root-mean-squares is because the monthly seasonal cycle (12 months) is computed and subtracted from the time-series of cosine and sine coefficients in the construction of the anomaly time-series  $\tilde{C}_n$  and  $\tilde{S}_n$ . If the annual average climatology of the



**Figure 4.** Amplitude and phase of the seasonal average migrating thermal tide, harmonic  $n = 1$ . All months of COSMIC RO data are used in contouring (a) the amplitude of the diurnal cycle in temperature (K), (b) the uncertainty in the amplitude, (c) the phase of the same harmonic (hours solar time), and (d) the uncertainty in the phase. The seasonal cycle is removed before estimating the uncertainties, following equations 6 and 7.

cosine and sine terms is just the average of the mean-monthly climatologies, resulting in  $C_n(\varphi, h, i)$  and  $S_n(\varphi, h, i)$ , and, assuming stationarity of the anomaly time-series  $\tilde{C}_n(\varphi, h, i)$  and  $\tilde{S}_n(\varphi, h, i)$ , the uncertainties in the amplitude  $\sigma_{\text{amplitude}}$  and in the phase  $\sigma_{\text{phase}}$  become

$$\sigma_{\text{amplitude}}^2 = (C_n^2(\varphi, h) \sigma_c^2 + S_n^2(\varphi, h) \sigma_s^2 + 2\rho C_n(\varphi, h) S_n(\varphi, h) \sigma_c \sigma_s) / (M A_n^2(\varphi, h)) \quad (7a)$$

$$\sigma_{\text{phase}}^2 = (S_n^2(\varphi, h) \sigma_c^2 + C_n^2(\varphi, h) \sigma_s^2 - 2\rho C_n(\varphi, h) S_n(\varphi, h) \sigma_c \sigma_s) / (M n^2 A_n^4(\varphi, h)). \quad (7b)$$

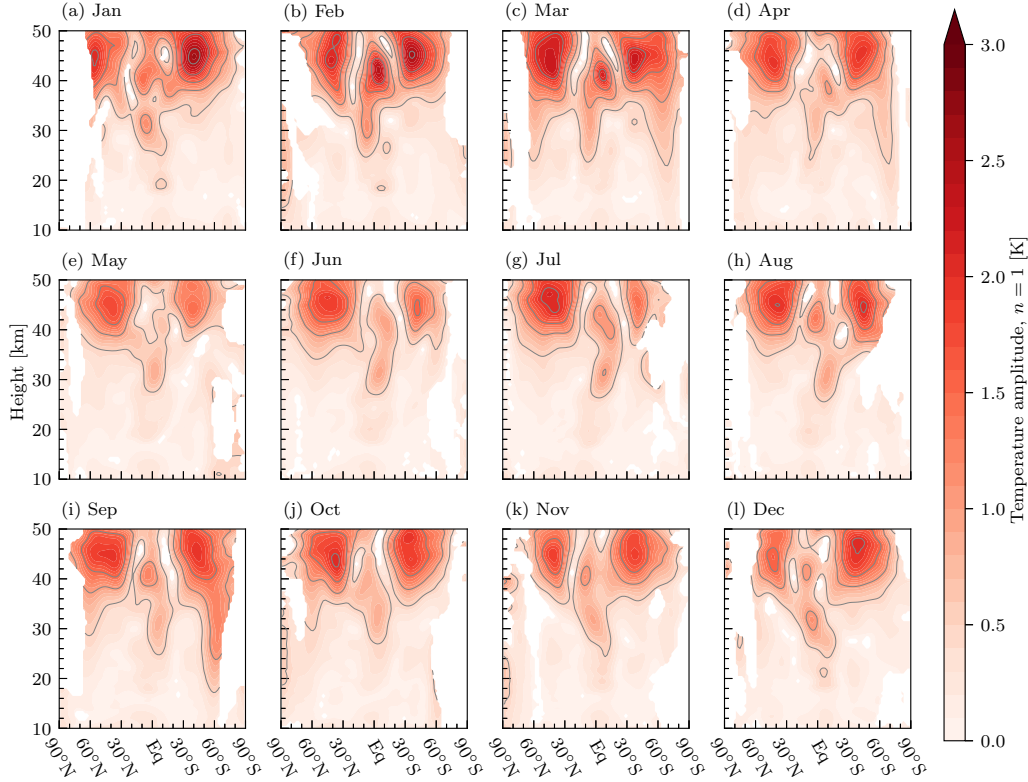
The coefficients  $C_n(\varphi, h)$ ,  $S_n(\varphi, h)$ , and  $A_n(\varphi, h)$  are averages of the de-seasonalized coefficients. An explanation and derivation of these equations are given in Appendix A. An extra factor of  $M$  is included in these equations because  $\sigma_{\text{amplitude}}$  and  $\sigma_{\text{phase}}$  are standard errors rather than standard deviations such as  $\sigma_c$  and  $\sigma_s$ .

In figure 4, downward phase propagation over the equator is evident. The amplitude (plot a) is  $\approx 1$  K, which occurs at 42 km height, with a secondary maximum of  $\approx 0.7$  K at 30 km height. Multiple maxima in amplitude are likely caused by the background zonal wind shear; see Hagan et al. (1999), for example. Downward phase propagation is evident as phase increasing with depth over the equator (plot c). Most prominent, though, is the tidal feature in the mid-latitude upper stratosphere, at heights above 37 km and between  $20^\circ$  and  $70^\circ$  latitude in each hemisphere. The feature has an amplitude of 2 K and shows no sign of vertical propagation and only very weak propagation poleward in each hemisphere. Figure 4 shows the uncertainty associated with the amplitude as plot (b). Over the equator, the uncertainty in the amplitude reaches 0.05 K at 44 km height and decreases to less than 0.02 K beneath 38 km. The signal-to-noise ratio of detection of the  $n = 1$  migrating thermal tide over the COSMIC-1 record is in excess of 20. Uncertainty in the amplitude is maximized in the mid-latitude upper stratosphere, especially in the northern hemisphere, reaching 0.08 K at  $65^\circ\text{N}$  and 44 km height. The maxima in both hemispheres overlie the poleward edge of the mid-latitude non-propagating tidal feature. Detection of this non-propagating feature has a signal-to-noise ratio in excess of 10. Lastly, figure 4 also shows the uncertainty in the phase of the  $n = 1$  migrating thermal tides. Plot (d) shows that the uncertainty in the phase over the equator is maximized at 48 km height is 12 minutes. Beneath 40 km, the uncertainty in the phase is less than 6 minutes.

Figure 5 shows the annual cycle of the  $n = 1$  migrating thermal tide as sampled by COSMIC-1 RO data. Only the amplitudes are shown for each month, but they are masked where the noise exceeds one-third the amplitude. The uncertainties are computed using the same formalism of equations 6 and 7. Because ten years of COSMIC RO data are used, however, the equations of 6 are exercised independently for each month with  $M = 10$  and a denominator (inside the square root) of  $M - 1$ , there being  $M - 1$  independent degrees of freedom in the calculations of the standard deviations and correlations. While this is not the first time the seasonality of the stratospheric tides has been analyzed using RO data, it is the first time it has been done with an error analysis, showing the power of Bayesian interpolation.

Figure 6 shows the amplitude and phase, along with their uncertainties, for the  $n = 2$  harmonic of the diurnal cycle. The tidal amplitude is maximized in the upper stratosphere in latitudes that span the Tropics, reaching 0.6 K at 45 km height. As opposed to the  $n = 1$  harmonic, the phase plot indicates upward phase propagation. Figures such as 4, 5, and 6 are possible up to  $n = n_{\text{max}}$ . In this analysis,  $n_{\text{max}} = 4$ .

The climatology of the diurnal cycle constructed here is monthly, giving expansion coefficients for every month from late 2006 through the end of 2016. When considering the diurnal cycle in just one month, the only uncertainties come from the leakage of synoptic (weather) variability and measurement error into the inference of the coefficients. In Bayesian interpolation, this sampling error is computed directly under the assumption that it is spatially and temporally uniform. Instructions on computing this uncer-



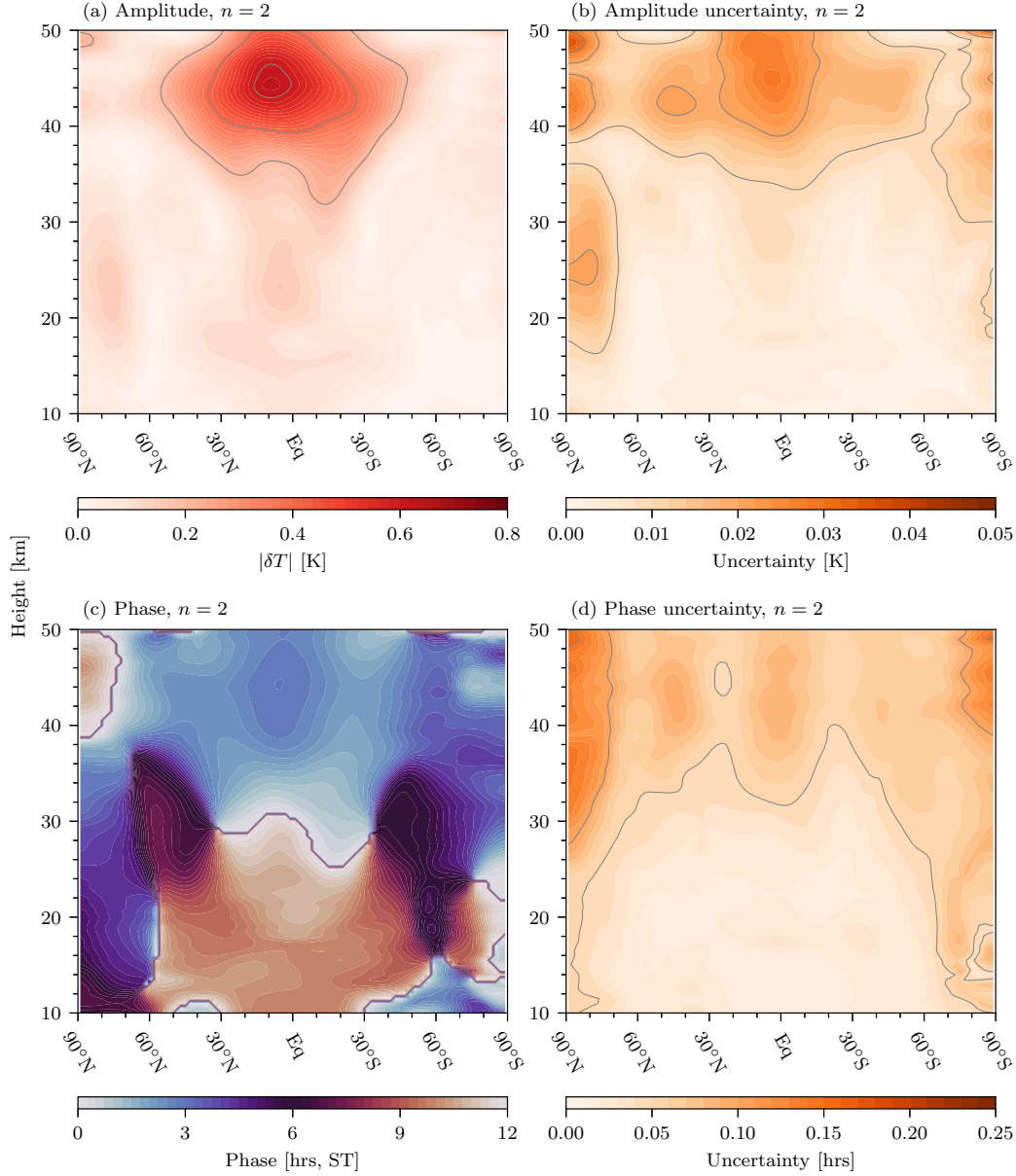
**Figure 5.** The annual cycle of the  $n = 1$  migrating thermal tide, amplitudes only. Each plot shows the amplitude of the  $n = 1$  harmonic of the diurnal cycle, zonally averaged, for a month-of-year. The contour maps are masked where the signal-to-noise ratio of detection falls beneath 3.

tainty from the output of Bayesian interpolation can be found in Leroy et al. (2021), equation 6. Figure 7 shows the impact of sampling error on the determination of the amplitude of four harmonics of the diurnal cycle. It is the average of the sampling error over all of the months from January, 2007, through December, 2016, having neglected the last few months of 2016 because the COSMIC RO did not yet fully sample the diurnal cycle. For the vertically propagating equatorial tides, the uncertainty in amplitude due to sampling error decreases with harmonic  $n$  and increases with height. For the diurnal harmonic ( $n = 1$ ), the uncertainty in amplitude ranges from 0.01 K in the lower stratosphere to 0.06 K at the stratopause over the equator. The uncertainty decreases by approximately 30% for the semi-diurnal harmonic ( $n = 2$ ). The uncertainties always decrease from the Tropics to the mid-latitudes but see strong maxima over both poles, where wintertime synoptic variability is greatly enhanced; see figure 5 of Leroy et al. (2021), for example. The reason the analyses of figures 4 through 6 are much more precise than for figure 7 is that they involved averaging over many months whereas figure 7 shows uncertainties for just one month.

### 3.3 Mid-latitude oscillation

The vertically propagating migrating thermal tides over the equator have been documented elsewhere, but the trapped mid-latitude upper stratospheric structure for the  $n = 1$  harmonic seen in figure 4 has not been thoroughly investigated. For this reason, the annual average analyzed time-series of dry temperature in the peak of the feature



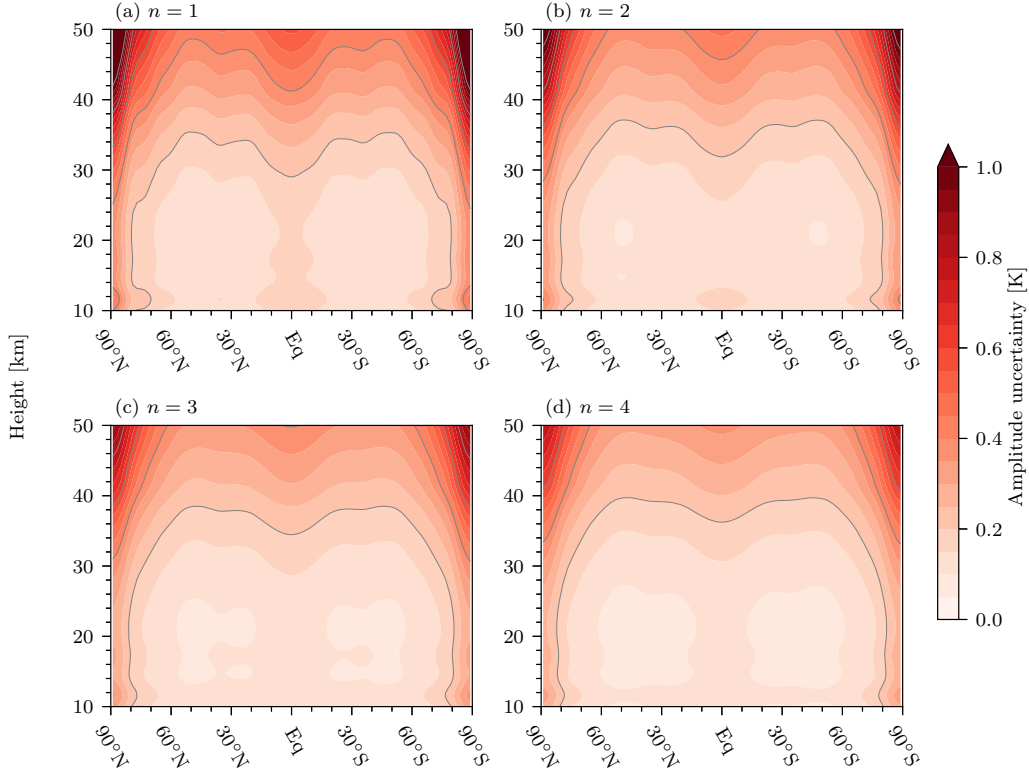


**Figure 6.** Same as figure 4 but for the  $n=2$  harmonic.

are shown in figure 8. This figure shows an expansion of the Bayesian interpolation at 45°N latitude and 48 km height for both the analyzed RO data and ERA-Interim. For each year, the average expansion coefficients over the 12 months of that year were computed before the expansion is performed. The diurnal cycle of ERA-Interim is computed by expanding the Bayesian interpolation fit of the RO data and adding the fit of model less data, the latter being the last term on the right of equation 2. All harmonics of the Bayesian interpolation analysis are included in the expansion.

The analysis of the mid-latitude trapped feature in figure 8 reveals two new findings. First, RO data and the model disagree on the amplitude and phase of the trapped, upper stratospheric mid-latitude feature. RO data shows the maximum amplitude to be





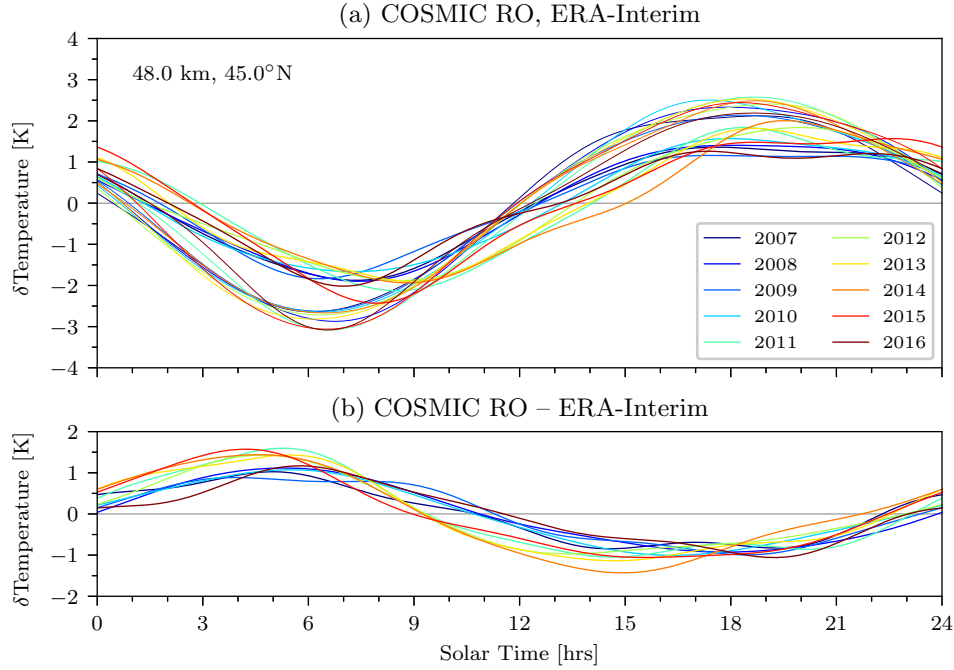
**Figure 7.** The uncertainty in the inference of diurnal cycle amplitude for four harmonics of the diurnal cycle for an individual month of COSMIC-1 RO data. The sampling error for all months from January, 2007, through December, 2016, is computed for diurnal harmonics 1 through 4 (plots a–d) then averaged over all 120 months.

$\approx 2$  K while ERA-Interim shows the amplitude to be  $\approx 3$  K. The time-series of the feature in RO data also lags the time-series in ERA-Interim by  $\approx 2$  hrs. Second, the RO data show more inter-annual variability in this feature than does ERA-Interim. There are notable departures from a mean diurnal cycle in the analysis of RO data especially for years 2012 through 2015. The departures are particularly noticeable in figure 8(b). The years 2012 through 2015 are years of maximum activity in the solar cycle. This is evident in figure 9, which shows the f10.7 cm flux that is commonly used as an indicator of solar activity.

The difference in the diurnal cycle between data and model in solar maximum years is very similar to the same difference in solar minimum years as seen in figure 8(b): the model bias is robust. The change in the data-model difference from solar minimum to solar maximum is much smaller than the mean data-model bias but is robust nevertheless. We conclude both that the upper stratospheric trapped feature is poorly analyzed in ERA-Interim and that a further bias is incurred by the solar cycle. The former is a much larger effect than the latter.

### 3.4 Ionospheric residual?

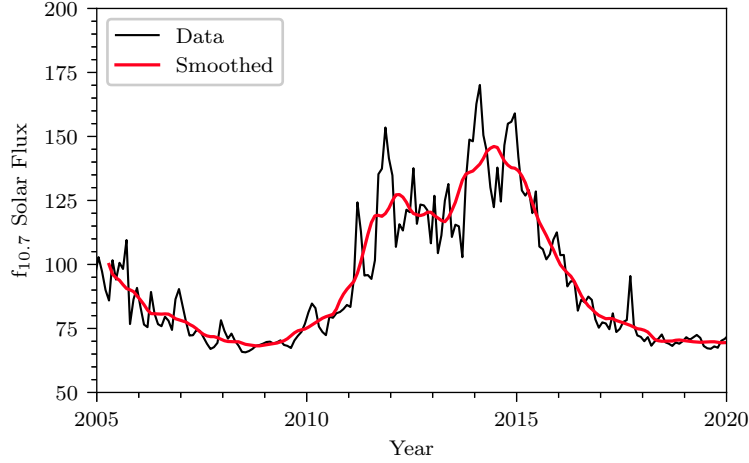
Two explanations are possible for the correlation between solar activity and the diurnal time-series of the mid-latitude upper stratospheric trapped feature. The first is that it is an actual feature of the atmosphere, that changing forcing by solar ultraviolet-



**Figure 8.** Diurnal time-series of zonal average dry temperature at 45°N, 48 km height for every year from 2007 through 2016 as analyzed in COSMIC RO data and as realized in ERA-Interim. Plot (a) shows expansions of diurnal harmonics  $n = 1$  through  $n = 4$  of the Bayesian interpolation analysis for both RO data (thicker lines) and ERA-Interim (thinner lines). Lines are color-coded by year according to the legend. Plot (b) shows the difference of COSMIC RO dry temperature less ERA-Interim temperature at the same latitude and height as in plot (a) following the same color-coding by year as in plot (a).

let changes the stratospheric thermal tides. The second is that solar activity induces a bias in RO temperature sounding. The first is not without merit, as solar ultraviolet forcing of the stratosphere and the tropospheric climate through downward influence has been hypothesized and discussed for over two decades (Lean, 2010). The second, however, is possible as well, as residual ionospheric error in RO remote sensing is a well known yet elusive phenomenon.

It is exceedingly difficult to disentangle the different contributors to error in RO retrieval in the stratosphere using data alone. Estimating the contributors theoretically is possible and has been done elsewhere (Danzon et al., 2020), but never empirically. In order to explore the impact of the 11-yr solar cycle on the stratospheric diurnal cycle, a triple-difference analysis is formed, one that reduces sampling error, isolates the solar cycle and analyzes the change in the diurnal cycle. Figure 10 shows this triple difference. Specifically, the calculations that yield the figure are (1) the difference of daytime (13:00 ST) less nighttime (4:00 ST) chosen to maximize the day-night contrast, (2) the difference of the diurnal cycle as seen in RO data and in ERA-Interim, and (3) the average diurnal cycle for solar maximum years (2012–2015) less that for solar minimum years (2007–2009). The first difference is an index for the quantity of interest, which is the diurnal cycle. The second difference serves to remove noise due to sampling error, following Foelsche et al. (2008, 2011), and thus yield a less noisy analysis. The third difference serves to attribute results of this analysis to the 11-yr solar cycle. The best interpretation of this analysis and the resulting figure is that it shows the influence of the

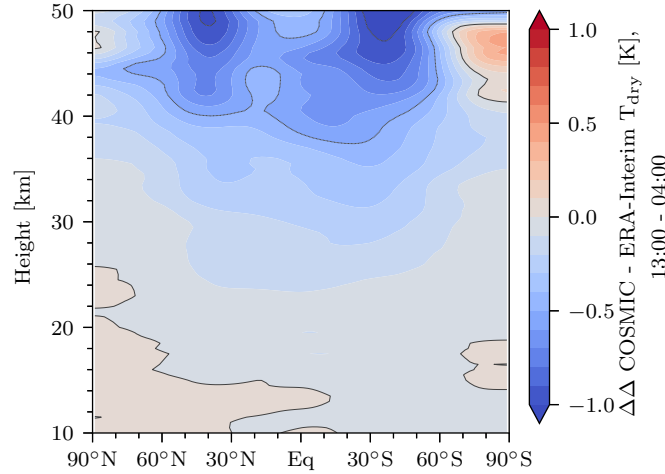


**Figure 9.** Monthly average solar activity, as indicated by the f10.7 cm flux, from 2005 through 2019. The black line is for monthly averages, and the red line is the same but smoothed using a 13-month weighted smoothing filter. The units of flux are  $10^4$  Jansky, or  $10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$ . The data are obtained from the NOAA Space Weather Prediction Center.

solar cycle on the diurnal cycle in dry temperature in the stratosphere as it is mis-analyzed by ERA-Interim. It could represent either mis-modeling of the atmosphere’s response to forcing by the solar cycle as seen in the diurnal cycle or ionospheric residual associated with RO.

The triple-difference of figure 10 shows a pattern that is symmetric about the equator, is strongest in the upper stratosphere, and tapers off with depth and at high latitudes. The dominant extrema occur in midlatitudes in the upper stratosphere in both hemispheres, specifically at the stratopause (50 km) and  $40^\circ$  latitude in both hemispheres, and the amplitude of the extrema is a change of approximately  $-1 \text{ K}$ . The response in the Deep Tropics, again at the stratopause, is less but still approximately  $-0.3 \text{ K}$ . Further analysis (not shown) shows the pattern to be robust. The pattern remains largely unchanged in all combinations of individual solar maximum years (2012, 2013, 2014, 2015) less solar minimum years (2007, 2008, 2009), which forms 12 independent combinations. Even the positive response over in the upper stratosphere over the Antarctic is robust in these combinations.

Figure 10 cannot be easily explained theoretically. A method for removing ionospheric residual from RO climatologies has been proposed that is based on the intensity of solar activity, making it possible to generate RO climatologies with and without the correction applied. The difference between those would be a theoretical estimate for the influence of ionospheric residual in temperature. Figure 7 of Danzer et al. (2020) shows the impact of ionospheric residual correction in the diurnal average, zonal average temperature response for one month at low solar activity and for another month at high solar activity. The extremum impact of ionospheric residual is approximate  $-3.5 \text{ K}$  at high solar activity and  $-0.5 \text{ K}$  at low solar activity. The theoretical signature is symmetric about the equator and peaked at the stratopause over the equator. The same study does not present the diurnal signature of ionospheric residual, however. While the theoretical response is similar in form to that of figure 10, its magnitude is substantially larger than that in figure 10 and it does not show extrema at the mid-latitude stratopause.



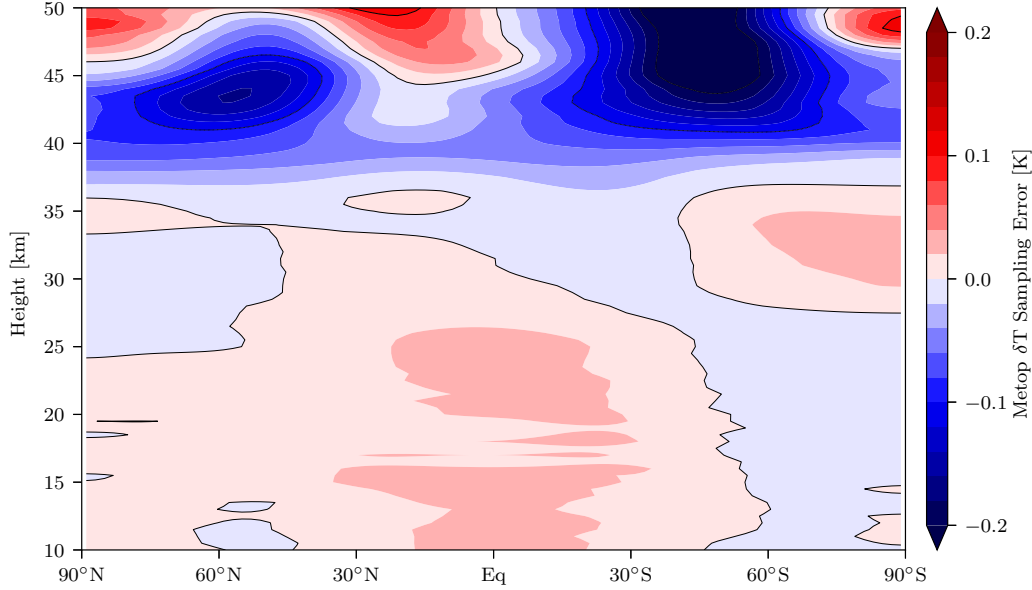
**Figure 10.** Impact of the solar cycle on the diurnal cycle. The change in the diurnal cycle in temperature is analyzed as a triple difference: daytime (13:00 ST) less nighttime (04:00), RO data less ERA-Interim, and solar maximum (years 2012–2015) less solar minimum (2007–2011). Light grey lines are shown for contour levels corresponding to the major ticks of the color bar.

### 3.5 Sampling error

Figure 8 shows clear evidence that ERA-Interim mis-models the stratospheric diurnal cycle, especially the trapped mid-latitude stratopause feature, and figure 10 hints at mis-modeling elsewhere near the stratopause. The mis-modeling is systematic in the diurnal cycle; thus, the model can be expected to induce errors in sampling error-adjusted climatologies of RO data if it, or another model that includes such a bias, is used to correct for sampling errors. Recall that, because RO data still greatly under-sample synoptic variability in the atmosphere, one frequently used method to remove the sampling error that results from under-sampling in order to yield less noisy climatologies is to sub-sample model forecasts to the geolocations of the RO soundings and compare monthly average model fields to climatologies of the sub-sampled model constructed in the same way as climatologies of RO data. The computed sampling bias is then subtracted from the RO climatology. While intended to reduce the random error normally associated with under-sampling of synoptic variability, the technique of *sampling error removal* can also introduce biases in climatologies due to biases in model realizations of the diurnal cycle, which is possible only when the distribution of RO data under-samples the diurnal cycle.

The Metop-A, Metop-B, and Metop-C satellites fly in sun-synchronous orbits, with the ascending nodes at 20:46, 21:30, and 21:30 solar time, respectively. Each of the Metop satellites carries a GPS RO instrument on board, and the high latitude orbits of Metop can deliver RO distributions that cover the globe. The distribution of RO soundings, however, severely under-samples the diurnal cycle, however, and a mis-modeled diurnal cycle will yield sampling-error-corrected climatologies that are biased. The resulting bias is computed as an expansion of the Bayesian interpolation fit to the COSMIC–ERA-Interim difference—the last term on the right of equation 2—at 9:30 and 21:30 solar time. The result is figure 11.

The biases in Metop climatologies of RO data after *sampling error removal* are generally small in the upper troposphere and the lower and middle stratosphere but become more pronounced in the upper stratosphere, as seen in figure 11. The incurred bias is



**Figure 11.** Bias in climatology of Metop RO data incurred by sampling error removal. Bayesian interpolation maps of ERA-Interim forecasts of temperature less RO retrievals of dry temperature are expanded at 9:30 and 21:30 solar time and zonally averaged. The result is an estimate of the bias in climatologies of Metop RO data incurred because of incomplete coverage of the diurnal cycle and systematic ERA-Interim mis-modeling of the diurnal cycle.

less than 0.04 K beneath 37 km height. Above that level, however, the bias reaches an amplitude of approximately  $-0.5$  K in the mid-latitude upper stratosphere in the Southern Hemisphere and approximately  $-0.2$  K in the Northern Hemisphere. The largest biases correspond to the trapped mid-latitude migrating tide seen in the mid-latitude upper stratosphere seen in figure 4a. Mis-modeling of this feature is as large as 1.5 K as seen in figure 8, much larger than the incurred bias in figure 11. This is a fortuitous consequence of Metop sampling at solar times that correspond to nulls in the mis-modeling of the diurnal cycle; see figure 8b. If the Metop satellites crossed the equatorial plane at 3:00 solar time instead of 9:30, the incurred temperature bias would instead have been approximately 1.0 K.

#### 4 Discussion

The theory and observation of migrating thermal tides in the upper atmosphere have been written about extensively in the past, but only recently has ionospheric residual bias in RO begun to be explored, and little work has been done to examine sampling biases incurred in RO climatologies when the diurnal cycle is under-sampled. Each of these three topics is addressed above by analyzing the diurnal cycle as observed by COSMIC RO using Bayesian interpolation and is discussed in the context of previous work below.

For thorough descriptions of previous research on thermal tides in the upper atmosphere, stratosphere through thermosphere and ionosphere, see a recent review (Pancheva & Mukhtarov, 2011) or some of the most recent papers on the topic (Mukhtarov et al., 2009; Sakazaki et al., 2012, 2018). As discussed in the Introduction, most observational work on atmospheric tides uses the observations of the SABER instrument on the TIMED

satellite (Russell et al., 1999), but the analyses are focussed on the mesosphere and thermosphere with little analysis in the stratosphere and without coverage poleward of  $50^\circ$  latitude in both hemispheres. Prior to TIMED/SABER, observational work on tides was based on data of the Microwave Limb Sounder on the UARS satellite. Both satellites sampled the diurnal cycle, which is necessary for diagnosing tides, by undergoing rapid regression of nodes.

The work above showed how complete coverage of the diurnal cycle obtained by the COSMIC RO mission can be analyzed for migrating tides throughout the stratosphere. Certainly the diurnal and semi-diurnal tides have been analyzed in satellite data previously. What is particularly noteworthy here is the accuracy obtained when analyzing the COSMIC RO data by Bayesian interpolation as extended into the diurnal cycle. Bayesian interpolation is used in the previous section to establish a climatology of the migrating tides that is accurate and precise to less than 0.1 K in amplitude for each harmonic of the diurnal cycle and each month of the annual cycle. Only over the polar regions can the amplitude not be inferred with a signal-to-noise ratio greater than 3. The uncertainty analysis considers both the random error resulting from under-sampling of synoptic variability and inter-annual variability of the diurnal cycle. The accuracy and precision extends to estimates of the phase of migrating tides. The diurnal harmonic migrating tide (figure 4) shows predominantly downward phase propagation while the semi-diurnal harmonic (figure 6) shows upward phase propagation in the tropical stratosphere. Because upward phase propagation corresponds to downward group velocity for tides and other forms of atmospheric gravity waves, the analysis of the previous section reveals that the dominant forcing of the diurnal migrating tide is tropospheric and the that of the semi-diurnal tide is upper stratospheric. The forcing mechanism in the upper stratosphere is radiative heating by the absorption of solar ultra-violet by ozone, and the forcing in the troposphere is latent heating by deep convection. For explanations of the forcing, see, for example, Hagan and Forbes (2002, 2003). The detail and precision of the phase plots of figures 4 and 6 enables such discernment.

The mid-latitude non-propagating feature in the upper stratosphere has been diagnosed previously in TIMED/SABER data (Mukhtarov et al., 2009; Sakazaki et al., 2012, 2018) but with some question. The previous work identified it as a predominantly diurnal feature with maximum at 16:00 solar time but cast some doubt on the diagnosis because the TIMED/SABER data coverage was obtained by nodal regression. This analysis, based on complete continuous coverage of the diurnal cycle, diagnoses it also as a predominantly diurnal feature, but its maximum occurs later in the day, at approximately 18:30 solar time. The feature corresponds to a tidal mode long-known as a solution to the tidal equations, namely the  $\Theta(-1, 2)$  symmetric Hough mode (Chapman & Lindzen, 1970; Mukhtarov et al., 2009). It does not propagate meridionally or vertically and is thus a trapped mode, one most likely forced by solar ultra-violet absorption by ozone because of its proximity to that forcing. Propagating modes are dissipated by propagation, but a trapped mode such as this can be dissipated only by eddy dissipation or long-wave radiative relaxation. Atmospheric models in the form of reanalyses do reproduce this mode (Sakazaki et al., 2018), but this analysis shows that at least one of them, ERA-Interim, over-estimates its amplitude and produces a maximum in phase too early in the day. Both the increased amplitude and the phase lead of this mode produced by the model can be explained by mis-modeling of eddy diffusivity in the upper stratosphere. An increase in the amount of eddy diffusivity would both decrease the amplitude of the mid-latitude trapped mode and lag the maximum to later in the day. If so, such analyses could prove a useful diagnostic for the physics of diffusion in the upper atmosphere.

The same analysis also shows that the mid-latitude trapped feature shows systematic differences in solar maximum vs. solar minimum years. While it is possible that the large changes in solar ultra-violet associated with solar activity could induce a temperature change in the upper stratosphere associated with the diurnal cycle, it is more likely

that the change in the diurnal cycle seen in figure 10 is associated with ionospheric residual error in RO retrieval. It comes about because of the divergence of the paths of the two GNSS signals tracked during an RO sounding, a consequence of the dispersive nature of the ionosphere. The greater the electron density in the ionospheric F-layer, the more the GPS L1 and L2 signals diverge, and the greater the error in the retrieval of temperature in RO sounding. Moreover, an increase in ultra-violet solar radiation can amplify the tidal response but it cannot change the phase, which is clearly seen in figure 8. The BAROCLIM climatology of bending angles used in initializing the Abel inversion for refractivity is as subject to ionospheric residual as each RO sounding. Because it is searched and tuned to individual RO soundings, it cannot introduce any ionospheric residual error in RO that is not already present in each RO sounding. A more thorough study is needed, however, to disentangle a real atmospheric temperature change in the diurnal cycle vs. a signature of ionospheric residual. This could be accomplished using an RO data set that has been corrected to eliminate ionospheric residual, following approaches recently described (Healy & Culverwell, 2015; Fan et al., 2017; Danzer et al., 2020).

Finally, few remote sensing techniques possess the absolute accuracy that is needed for climate monitoring, but RO is one of them. RO possesses this property because its fundamental observable is traceable to the international definition of the second, which is presently known to approximately 1 part in  $10^{15}$ , as opposed to a radiance standard, which is known much more poorly. In the stratosphere, however, three issues must be considered when using RO for climate monitoring: (1) the possibility of local multi-path, (2) ionospheric residual, and (3) non-uniform sampling. (Local multi-path occurs when the GNSS receiver tracks a GNSS signal reflected from an element on board the satellite in addition to the same GNSS signal that propagates directly from the transmitter to the receiving antenna.) Factors (1) and (2) were described in Kursinski et al. (1997); factor (1) can be mitigated by systematic inter-comparison between missions, and factor (2) as discussed in the previous paragraph. The third factor, though, may not be easily corrected if the RO data systematically under-sample the diurnal cycle. The method most commonly used to correct for sampling error is the *sampling error removal* technique, wherein the sampling error corresponding to an RO sampling pattern is computed by sub-sampling atmospheric forecasts to the times and locations of RO soundings. This technique was originally intended to address the random noise associated with synoptic variability and so make time-series of temperature appear smoother and a better approximation of reality, but it can induce a bias if it is applied to RO data that systematically under-sample the atmosphere where the forecast model contains systematic biases. This study showed the bias that would result for a temperature climatology that would be assembled from Metop RO data when sampling-error-corrected using the forecasts of ERA-Interim. As above, ERA-Interim mis-models the diurnal cycle in the upper stratosphere, and the climatology that results would be biased negative in that region. The bias would have been worse if the Metop orbits did not coincide with nulls in the *difference* between the RO-inferred diurnal cycle and the model diurnal cycle.

It is crucial that RO obtain a sampling pattern that completely covers the diurnal cycle even if non-uniformly. *Sampling error removal* cannot completely eliminate the bias incurred by under-sampling of the diurnal cycle. As of the year 2021, the COSMIC-2 RO mission is obtaining complete coverage of the diurnal cycle in low latitudes. While the Metop RO instruments obtain global coverage, they do not obtain complete coverage of the diurnal cycle, leaving gores in the coverage of the diurnal cycle at high latitudes. It is a problem best fixed with a targeted deployment of a sun-synchronous RO satellite near a nodal crossing time of 3:30/15:30 solar time. Absent such an RO satellite, the best way to address the gap in coverage is careful use of a climatology of the diurnal cycle established using COSMIC RO data.



## Acknowledgments

The RO data were provided by the Radio Occultation Meteorology Satellite Application Facility (ROM SAF; <https://www.romsaf.org>), which is a decentralized RO processing center under EUMETSAT. We specifically used data from the ROM SAF climate data record v1.0 (doi:10.15770/EUM\_SAF\_GRM.0002 and doi:10.15770/EUM\_SAF\_GRM.0003). The analyses of the migrating thermal tides are publicly available through Zenodo, doi:10.5281/zenodo.5259994. This research was funded by EUMETSAT's Visiting Scientist program, by the Terra, Aqua, Suomi-NPP Science Team (NASA grant 80NSSC18K0946), and by Atmospheric and Environmental Research, Inc.

## References

- Anthes, R., Bernhardt, P., Chen, Y., Cucurull, L., Dymond, K., Ector, D., ... Zeng, Z. (2008). The COSMIC/FORMOSAT-3 mission: Early results. *Bull. Amer. Meteor. Soc.*, 89(3), 313–334. doi: 10.1175/BAMS-89-3-313
- Chapman, S., & Lindzen, R. (1970). *Atmospheric tides*. Dordrecht, Netherlands: D. Reidel.
- Danzer, J., Foelsche, U., Scherllin-Pirscher, B., & Schwärz, M. (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
- Danzer, J., Schwaerz, M., Kirchengast, G., & Healy, S. (2020). Sensitivity analysis and impact of the kappa-correction of residual ionospheric biases on radio occultation climatologies. *Earth and Space Sci.*, 7, e2019EA000942. doi: 10.1029/2019EA000942
- Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., ... Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteor. Soc.*, 137, 553–597.
- Fan, Z., Sheng, Z., Guo, P., & Shi, H. (2017). An idealized method of simulating residual ionospheric errors in radio occultation. *Nature Scientific Reports*, 7, 16632. doi: 10.1038/s41598-017-16972-4
- Foelsche, U., Borsche, M., Steiner, A., Gobiet, A., Pirscher, B., Kirchengast, G., ... Schmidt, T. (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., Scherllin-Pirscher, B., Ladstädter, F., Steiner, A., & Kirchengast, G. (2011). Refractivity and temperature climate records from multiple radio occultation satellites consistent within 0.05%. *Atmos. Meas. Tech.*, 4, 2007–2018. doi: 10.5194/amt-4-2007-2011
- Fong, C., Yang, S., Chu, C., Huang, C., Yeh, J., Lin, C., ... Chi, S. (2008). FORMOSAT-3/COSMIC constellation spacecraft system performance: After one year in orbit. *IEEE Trans. Geosci. Rem. Sensing*, 46(11), 3380–3394. doi: 10.1109/TGRS.2008.2005203
- Gleisner, H., Lauritsen, K., Nielsen, J., & Syndergaard, S. (2020). Evaluation of the 15-year ROM SAF monthly mean GPS radio occultation climate data record. *Atmos. Meas. Tech.*, 13, 301–3098. doi: 10.5194/amt-13-3081-2020
- Gorbunov, M., Benzon, H., Jensen, A., Lohmann, M., & Nielsen, A. (2004). Comparative analysis of radio occultation processing approaches based on Fourier integral operators. *Radio Sci.*, 39(6), doi:10.1029/2003RS002916.
- Gorbunov, M., & Lauritsen, K. (2004). Analysis of wave fields by Fourier integral operators and their application for radio occultations. *Radio Sci.*, 39(RS4010), doi:10.1029/2003RS002971.
- Gorbunov, M., Shmakov, A., Leroy, S., & Lauritsen, K. (2011). COSMIC radio occultation processing: Cross-center comparison and validation. *J. Atmos. Ocean. Tech.*, 28, 737–751.
- Hagan, M., Burrage, M., Forbes, J., Hackney, J., Randel, W., & Zhang, X. (1999).

- GSWM-98: Results for migrating solar tides. *J. Geophys. Res.*, *104*(A4), 6813–6827. doi: 10.1029/1998JA900125
- Hagan, M., & Forbes, J. (2002). Migrating and nonmigrating diurnal tides in the middle and upper atmosphere excited by tropospheric latent heat release. *J. Geophys. Res.*, *107*(D24). doi: 10.1029/2001JD001236
- Hagan, M., & Forbes, J. (2003). Migrating and nonmigrating semidiurnal tides in the upper atmosphere excited by tropospheric latent heat release. *J. Geophys. Res.*, *108*(A2). doi: 10.1029/2002JA009466
- Hajj, G., Kursinski, E., Romans, L., Bertiger, W., & Leroy, S. (2002). A technical description of atmospheric sounding by GPS occultation. *J. Atmos. Solar Terr. Phys.*, *64*(4), 451–469. doi: 10.1016/S1364-6826(01)00114-6
- Healy, S., & Culverwell, I. (2015). A modification to the standard ionospheric correction method used in GPS radio occultation. *Atmos. Meas. Tech.*, *8*, 3385–3393.
- Hitchman, M., & Leovy, C. (1985). Diurnal tide in the equatorial middle atmosphere as seen in LIMS temperatures. *J. Atmos. Sci.*, *42*, 557–561. doi: 10.1175/1520-0469(1985)042<0557:DTITEM>2.0.CO;2
- Kursinski, E., & Gebhardt, T. (2014). A method to deconvolve errors in GPS RO-derived water vapor histograms. *J. Atmos. Ocean. Tech.*, *31*, 2606–2628. doi: 10.1175/JTECH-D-13-00233.1
- Kursinski, E., Hajj, G., Hardy, K., Romans, L., & Schofield, J. (1995). Observing tropospheric water-vapor by radio occultation using the Global Positioning System. *Geophys. Res. Lett.*, *22*(17), 2365–2368.
- Kursinski, E., Hajj, G., Leroy, S., & Herman, B. (2000). The GPS radio occultation technique. *Terr. Atmos. Ocean. Sci.*, *11*(1), 53–114.
- Kursinski, E., Hajj, G., Schofield, J., Linfield, R., & Hardy, K. (1997). Observing Earth’s atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, *102*(D19), 23429–23465. doi: 10.1029/97JD01569
- Lean, J. (2010). Cycles and trends in solar irradiance and climate. *WIREs Climate Change*, *1*, 111–122. doi: 10.1002/wcc.018
- Leroy, S. (1997). Measurement of geopotential heights by GPS radio occultation. *J. Geophys. Res.*, *102*(D6), 6971–6986. doi: 10.1029/96JD03083
- Leroy, S., Ao, C., & Verkhoglyadova, O. (2012). Mapping GPS radio occultation data by Bayesian interpolation. *J. Atmos. Ocean. Tech.*, *29*, 1062–1074. doi: 10.1175/JTECH-D-11-00179.1
- Leroy, S., Ao, C., Verkhoglyadova, O., & Oyola, M. (2021). Analyzing the diurnal cycle by Bayesian interpolation on a sphere for mapping GNSS radio occultation data. *J. Atmos. Ocean. Tech.*, *38*, 951–961. doi: 10.1175/JTECH-D-20-0031.1
- MacKay, D. (1992). Bayesian interpolation. *Neur. Comp.*, *4*(3), 415–447. doi: 10.1162/neco.1992.4.3.415
- Mukhtarov, P., Pancheva, D., & Andonov, B. (2009). Global structure and seasonal interannual variability of the migrating diurnal tide seen in the SABER/TIMED temperatures between 20 and 120 km. *J. Geophys. Res.*, *114*, A02309. doi: 10.1029/2008JA013759
- Pancheva, D., & Mukhtarov, P. (2011). Atmospheric tides and planetary waves: Recent progress based on SABER/TIMED temperature measurements (2002–2007). In M. Abdou & D. Pancheva (Eds.), *Aeronomy of the Earth’s Atmosphere and Ionosphere* (pp. 19–56). New York: Springer. doi: 10.1007/978-94-007-0326-1\_2
- Pirscher, B., Foelsche, U., Borsche, M., Kirchengast, G., & Kuo, Y. (2010). Analysis of migrating diurnal tides detected in FORMOSAT-3/COSMIC temperature data. *J. Geophys. Res.*, *115*(D14108). doi: 10.1029/2009JD013008
- Randel, W., Wu, F., & Podglajen, A. (2021). Equatorial waves, diurnal tides

- and small-scale thermal variability in the tropical lower stratosphere from COSMIC-2 radio occultation. *J. Geophys. Res.*, *126*, e2020JD033969. doi: 10.1029/2020JD033969
- Russell, J., Mlynczak, M., Gordley, L., Tansock, J., & Esplin, R. (1999). An overview of the saber experiment and preliminary calibration results. In A. Larar (Ed.), *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III* (Vol. 3756, pp. 277–288).
- Sakazaki, T., Fujiwara, M., & Shiotani, M. (2018). Representation of solar tides in the stratosphere and lower mesosphere in state-of-the-art reanalyses and in satellite observations. *Atmos. Chem. Phys.*, *18*, 1437–1456. doi: 10.5194/acp-18-1437-2018
- Sakazaki, T., Fujiwara, M., Zhang, X., Hagan, M., & Forbes, J. (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., Steiner, A., Kirchengast, G., Schwärz, M., & Leroy, S. (2017). The power of vertical geolocation of atmospheric profiles from GNSS radio occultation. *J. Geophys. Res.*, *122*, 1595–1616. doi: 10.1002/2016JD025902
- Scherllin-Pirscher, B., Syndergaard, S., Foelsche, U., & Lauritsen, K. (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
- Vorob'ev, V., & Krasil'nikova, T. (1994). Estimation of the accuracy of the atmospheric refractive index recovery from Doppler shift measurements at frequencies used in the NAVSTAR system. *Phys. Atmos. Ocean.*, *29*, 602–609.
- Ward, W., Oberheide, J., Riese, M., Preusse, P., & Offerman, D. (1999). Tidal signatures in temperature data from CRISTA 1 mission. *J. Geophys. Res.*, *104*(D13), 16391–16403. doi: 10.1029/1998JD100109
- Wickert, J., Beyerle, G., Hajj, G., Schwieger, V., & Reigber, C. (2002). GPS radio occultation with CHAMP: Atmospheric profiling utilizing the space-based single difference technique. *Geophys. Res. Lett.*, *29*(8), doi:10.1029/2001GL013982.
- Wu, Q., Ortland, D., Killeen, T., Roble, R., Hagan, M., Liu, H., ... Niciejewski, R. (2008). Global distribution and interannual variations of mesospheric and lower thermospheric neutral wind diurnal tide: 1. Migrating tide. *J. Geophys. Res.*, *113*, A05308. doi: 10.1029/2007JA012542
- Xie, F., Wu, D., & Mannucci, A. (2010). Atmospheric diurnal variations observed with GPS radio occultation soundings. *Atmos. Chem. Phys.*, *10*, 6889–6899. doi: 10.5194/acp-10-6889-2010
- Zeng, Z., Randel, W., Sokolovskiy, S., Deser, C., Kuo, Y., Hagan, M., ... Ward, W. (2008). Detection of migrating diurnal tide in the tropical upper troposphere and lower stratosphere using the Challenging Minisatellite Payload radio occultation data. *J. Geophys. Res.*, *113*(D03102). doi: 10.1029/2007JD008725
- Zhang, X., Forbes, J., Hagan, M., Russell III, J., Palo, S., Mertens, C., & Mlynczak, M. (2006). Monthly tidal temperatures 20–120 km from TIMED/SABER. *J. Geophys. Res.*, *111*, A10S08. doi: 10.1029/2005JA011504

## Appendix: Uncertainty in amplitude and phase of a tide

In Bayesian interpolation, uncertainty in the posterior coefficients of expansion is returned as a fully populated covariance matrix. Any particular tidal mode is characterized by just two terms in the expansion of the basis functions: a cosine term with amplitude  $C_n$  and a sine term with amplitude  $S_n$ . The covariance matrix corresponding to these terms is

$$\Sigma = \begin{bmatrix} \sigma_c^2 & \rho\sigma_c\sigma_s \\ \rho\sigma_c\sigma_s & \sigma_s^2 \end{bmatrix} \quad (8)$$

in which  $\sigma_c$  and  $\sigma_s$  are the uncertainties in the cosine and sine terms and  $\rho$  the correlation of those uncertainties. The expansion in diurnal time  $\tau_d$  is

$$T_n(\tau_d) = C_n \cos(n\tau_d) + S_n \sin(n\tau_d). \quad (9)$$

Recall that diurnal time ranges from 0 to  $2\pi$  over the course of 24 hours. The diurnal cycle can be written also as

$$T_n(\tau) = A_n \cos(n(\tau - \phi_n)) \quad (10)$$

with  $A_n$  the amplitude of the mode and  $\phi_n$  the phase. These can be derived from  $C_n$  and  $S_n$  according to

$$A_n^2 = C_n^2 + S_n^2 \quad (11a)$$

$$\phi_n = \arctan(S_n, C_n)/n \quad (11b)$$

with  $\arctan(\dots, \dots)$  a four-quadrant arctangent defined such that  $\tan \phi_n = S_n/C_n$ . Ordinary error propagation assuming Gaussian uncertainties is performed following the standard formulas:

$$\sigma_{A_n}^2 = \left(\frac{\partial A_n}{\partial C_n}\right)_{S_n}^2 \sigma_c^2 + \left(\frac{\partial A_n}{\partial S_n}\right)_{C_n}^2 \sigma_s^2 + 2 \left(\frac{\partial A_n}{\partial C_n}\right)_{S_n} \left(\frac{\partial A_n}{\partial S_n}\right)_{C_n} \rho\sigma_c\sigma_s \quad (12a)$$

$$\sigma_{\phi_n}^2 = \left(\frac{\partial \phi_n}{\partial C_n}\right)_{S_n}^2 \sigma_c^2 + \left(\frac{\partial \phi_n}{\partial S_n}\right)_{C_n}^2 \sigma_s^2 + 2 \left(\frac{\partial \phi_n}{\partial C_n}\right)_{S_n} \left(\frac{\partial \phi_n}{\partial S_n}\right)_{C_n} \rho\sigma_c\sigma_s. \quad (12b)$$

The partial derivatives are computed from equations 11 and inserted into equations 12 to obtain the following uncertainties in  $A_n$  and  $\phi_n$ :

$$\sigma_{A_n}^2 = (C_n^2\sigma_c^2 + S_n^2\sigma_s^2 + 2\rho C_n S_n \sigma_c \sigma_s)/A_n^2 \quad (13a)$$

$$\sigma_{\phi_n}^2 = (C_n^2\sigma_s^2 + S_n^2\sigma_c^2 - 2\rho C_n S_n \sigma_c \sigma_s)/(n^2 A_n^4). \quad (13b)$$

With determinations of  $\sigma_{C_n}$ ,  $\sigma_{S_n}$ , and  $\rho$ , whether by Bayesian interpolation or by computing inter-annual variability, these equations can be applied to estimate the uncertainty in amplitude and phase of a tide.