

## Appendix A: STAR

Commercial Weather Data Pilot Round 2

# STAR CWDP Round 2 Data Processing, Analysis, and Validation Report

Shu-peng Ben Ho, Xinjia Zhou,

Xi Shao, Loknath Adhikari,

Stanislav Kireev, Bin Zhang, Erin Lynch

And the rest of STAR CWDP Science Team



## Table of Contents

### Contents

A.1 INTRODUCTION .....	4
A.1.1 Organization.....	4
A.2 STAR RO DATA PROCESSING PACKAGE .....	5
A.2.1 STAR RO Data Processing Package.....	6
A.2.2 STAR 1D-VAR Inversion Package .....	8
A.2.3 STAR ICVS Tool Development .....	10
A.2.4 STAR Multi-Sensor Validation System.....	11
A.2.4.1 Multi-Sensor Data.....	11
A.2.4.2 Validation Approach.....	14
A.3 GEOOPTICS: PROCESSING, ANALYSIS, AND VALIDATION.....	15
A.3.1 Assessment of GeoOptics L1a Data Quality .....	16
A.3.1.1 Bending angle Uncertainty due to Clock Bias .....	16
A.3.1.2 Geolocation Differences between Processing Centers .....	17
A.3.1.3 Stability of Satellite Position and Velocities .....	19
A.3.1.4 Phase Measurement Noise .....	21
A.3.1.5 Signal to Noise Ratio (SNR) Assessment.....	23
A.3.2 Assessment of GeoOptics L2 Data Quality .....	24
A.3.2.1 Bending Angle and Refractivity Comparisons .....	24
A.3.2.2 Comparison to NWP Model Fields.....	26
A.3.2.3 Data Penetration.....	31
A.3.2.4 Monitoring of Long-term Stability .....	33
A.3.2.5 Estimation of Retrieval Accuracy.....	34
A.3.2.6 Observation Uncertainty Estimates.....	37
A.3.2.7 Comparison to MW and IR Sounders.....	40
A.4 SPIRE: PROCESSING, ANALYSIS, AND VALIDATION.....	45
A.4.1 Assessment of Spire L1a Data Quality .....	45
A.4.1.1 Stability of Satellite Position and Velocities .....	45
A.4.1.2 Phase Measurement Noise .....	46
A.4.1.3 Signal to Noise Ratio (SNR) Assessment.....	48

A.4.2 Assessment of Spire L2 Data Quality ..... 49

    A.4.2.1 Bending Angle and Refractivity Comparisons ..... 49

    A.4.2.2 Comparison to NWP Model Fields..... 51

    A.4.2.3 Data Penetration..... 56

    A.4.2.4 Monitoring of Long-term Stability ..... 57

    A.4.2.5 Uncertainty of RO Water Vapor Retrievals..... 57

    A.4.2.6 Observation Uncertainty Estimates..... 61

    A.4.2.7 Comparison to MW and IR Sounders..... 64

A.5 SUMMARY, RECOMMENDATION, AND LESSONS LEARNED..... 69

    A.A.1 DATA PIPELINE TO NOAA/STAR..... 71

    A.A.2 GEOOPTICS..... 71

        A.A.2.1 Data Delivery ..... 71

        A.A.2.2 Daily Counts ..... 75

        A.A.2.3 Spatial and Temporal Characteristics ..... 76

        A.A.2.4 Lessons Learned..... 78

    A.A.3 SPIRE..... 78

        A.A.3.1 Data Delivery and Counts..... 78

        A.A.3.2 Spatial and Temporal Characteristics of Delivered Data..... 81

        A.A.3.3 Lessons Learned..... 82

    A.B.1. EXTRACTING THE OPNGNS PHASE AND RINEX FILE PHASE ..... 84

    A.B.2. ORBITAL DETERMINATION FOR LOE/GNSS POSITION AND VELOCITY ..... 84

    A.B.3 EXTRACTING THE ATTITUDE INFORMATION ..... 84

    A.B.4 EARTH COORDINATE SYSTEM CONVERSION ..... 84

    A.B.5 CYCLE SLIP DETECTION ..... 85

    A.B.6 CALCULATION OF EXCESS PHASE..... 85

    A.B.7 CALCULATION OF BENDING ANGLE..... 87

## A.1 Introduction

The National Oceanic and Atmospheric Administration (NOAA) STAR (Center for Satellite Applications and Research) is now operationally processing commercially-provided Global Navigation Satellite System (GNSS) Radio Occultation (RO) data using independently developed inversion package. This is to support the OSAAP's Commercial Weather Data Project Round 2 (CWDP R2) assessment. The STAR processed CWDP retrievals are used to understand the quality of the RO raw phase delay data and to quantify the structural uncertainty of the derived bending angle profiles and refractivity profiles when different inversion methods and initialization approaches are implemented.

GNSS RO data provided by two commercial data providers: GeoOptics and SPIRE Global are examined. The specific tasks of STAR on CWDP R2 include:

- Develop and execute strategies to compare RO data to established atmospheric data sets and models
- Augment the Integrated Cal/Val System (ICVS) to support the CWDP
- Validate ICVS CWDP-related functions using validated RO data sets (examples: COSMIC, KOMPSAT-5, MetOp)
- Test incoming commercial RO Level 1 files against technical requirements in the contract with commercial RO vendors (examples include: verify RO quantities and spatial distribution requirements are met. Validate file and data formats. Verify data quality per requirements.)
- Make monitoring/test results available via password-restricted web-based monitoring within 24 hours of data ingest
- Plot daily and cumulative distribution of RO measurements;
- Implement web-based RO files and statistics monitoring
- Implement web-based and alert-based data monitoring
- Processing CWDP 2 data and validate the results
- Develop routine reporting of data analysis results

In this report, we will summarize results for those activities that are planned (1) internally in STAR as part of the exploitation of critical observing systems data, (2) in coordination with partners such as UCAR for the routine/operational data processing, (3) in collaboration with other federal partners and Cooperative Institutes experts in radio occultation (such as JPL, and CICS respectively) on further exploiting the CWDP's RO data and maximizing their impact, (4) with users and in particular NWS, for the research to operations (R2O) transition. It is important to note also that this task will be coordinated with other RO projects being undertaken in STAR, specifically PAZ, KOMPSAT, and COSMIC-2. The CWDP project will benefit from knowledge, capabilities, lessons learned and tools gathered as part of these projects.

### A.1.1 Organization

Section A.1 provides information regarding the scope, purpose, and organization of this report. Section A.2 gives an overview of the STAR Processing Package and data evaluation methods including the STAR 1D-VAR Inversion Package, the development of tools for data monitoring with the Integrated Calibration and Validation System (ICVS), and multi-sensor validation techniques.

Section A.3 presents the analysis and validation of the level 1a and level 2 data delivered by GeoOptics during their contractual delivery period from November 2018 through September 2019.

Similarly, Section A.4 presents the analysis and validation of the level 1a and level 2 data delivered by Spire Global, Inc. during the months of June and July 2019.

Section A.5 provides a summary of conclusions and recommendations as well as lessons learned during this round of the CWDP project.

Appendix A.A contains a summary of the data delivery and distribution procedures at STAR and details about the data delivered by each vendor, including daily and total counts, geographic and temporal distributions, data types, and versions.

Appendix A.B contains further details of the STAR retrievals of level 2 data from the level 1a data provided by vendors.

The abbreviations and acronyms used in the report is listed in the end of this report.

## A.2 STAR RO data processing package

STAR has been developed as a GNSS RO processing and research center. Four major focus area of STAR RO data processing and validation are defined in Figure A-1. In particular, we have dedicated our efforts to (1) RO data processing (both L1a-L1b processing, and L1b to L2 processing), (2) developing the Integrated calibration and validation (cal/val) system (ICVS) for data monitoring, (3) performing multi-sensor validation, and (4) data assimilation. The CWDP R2 tasks performed by STAR and covered in this report are mainly based on studies from the first three focus areas.

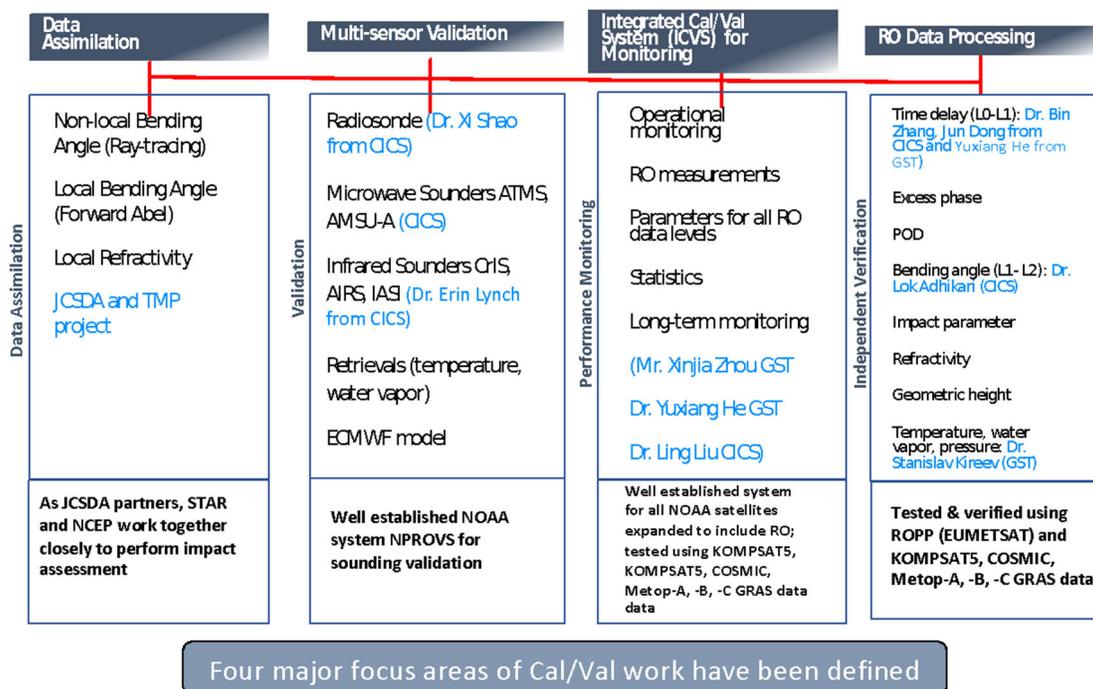


Figure A-1: Four major focus areas of STAR CWDP2 RO processing and validation system.

STAR uses (1) the Radio Occultation Processing Package (ROPP) which was developed and maintained by ROM SAF (Radio Occultation Meteorology Satellite Application Facility), and

(2) STAR developed FSI (Full Spectrum Inversion) package to process the CWDP data. The inversion methods and data filtering implementation used in the ROPP are detailed in Culverwell et al. (2015) and are not repeated in this report.

STAR has adapted the inversion modules in the ROPP package to utilize our own quality control (QC) approaches. Forward calculated bending angle profiles using GFS 6-hour forecast temperature, moisture, and pressure as inputs are used to constrain the derived L2 bending angle profiles. Only the inverted bending angle profiles that deviate by less than 5 standard deviations compared to the forward computed bending angle profiles at all vertical levels are included in this study.

The STAR FSI processing package is described in Section A.2.1 where the initial validation of CWDPR2 bending angle retrieval processed by using the STAR FSI operational inversion package is shown in the same section. The STAR 1D-var inversion package is described in Section A.2.2. As described above an Integrated calibration and validation system (ICVS) was developed to routinely monitor the quality of the derived RO data. We describe the validation and monitoring functions built in the ICVS in Section A.2.3. The STAR Multi-sensor Validation approaches are described in Section A.2.4.

### A.2.1 STAR RO Data Processing Package

To derive bending angle and refractivity profiles from RO occultation measurements, one must perform L0-L1 (from raw data to excess phase) and L1-L2 (converting excess phase to bending angle) processing. Over the past year, the STAR GNSS team has developed the capability to perform both L0-L1 and L1-L2 data processing. Figure A-2 depicts the flow chart of the STAR RO processing procedures.

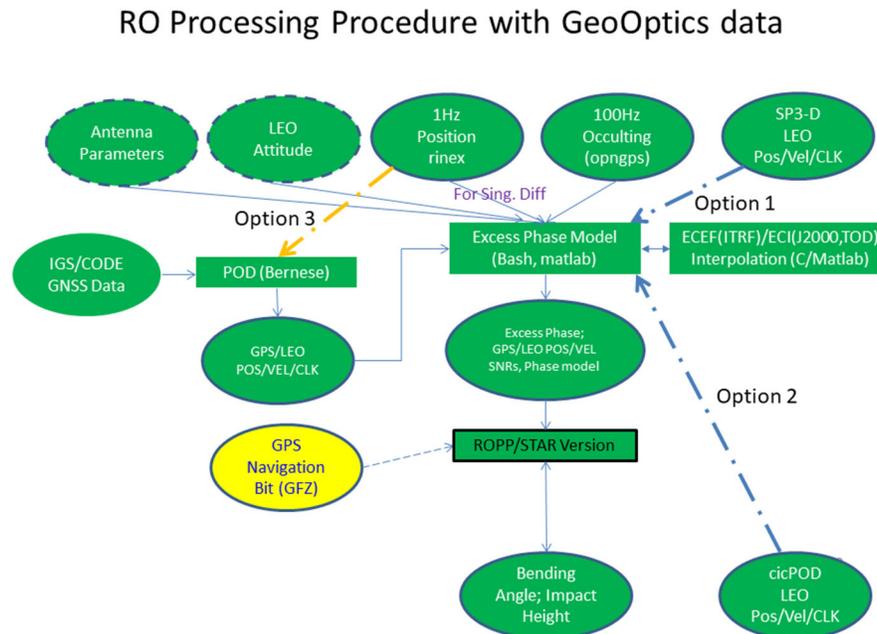


Figure A-2: The flow chart of the STAR RO processing procedures.

For L0-L1 processing, we are able to perform 1) precise orbit determination (POD) and clock synchronization to eliminate the effects of the geometric Doppler and of relative transmitter

receiver oscillator drift, 2) bending angle calculation, 3) ionospheric corrections, 4) Abel integral inversion with upper boundary conditions, and 5) quality control (QC). The general excess phase processing and refractivity inversion procedures and the sources of errors are described in Kursinski et al. (1997).

STAR processing algorithms include the following sequential processing modules:

- a. ***I/O subroutine*** to read phase, amplitude and geometry data
- b. ***Translation subroutine*** to change ECEF coordinates to local center of curvature and calculate the radius of curvature. The Flow chart to calculate the propagation of GPS and LEO orbits to circular orbits relative to a local center of curvature is described in Figure A-3.
- c. ***Projection subroutine*** to project GNSS/LEO orbit to fixed radii from local center of curvature
- d. ***FFT subroutine*** to get bending angle at each impact parameter: This is the core FSI/FFT subroutine.
- e. ***Inverse Abel subroutine*** to compute refractivity from bending angle/impact parameter.

Flow Chart

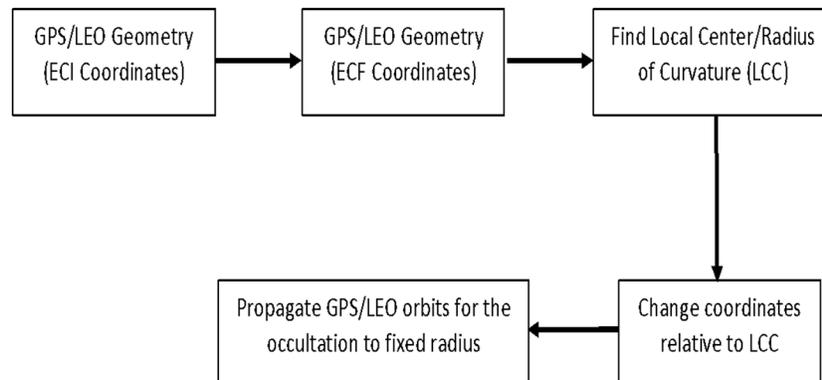
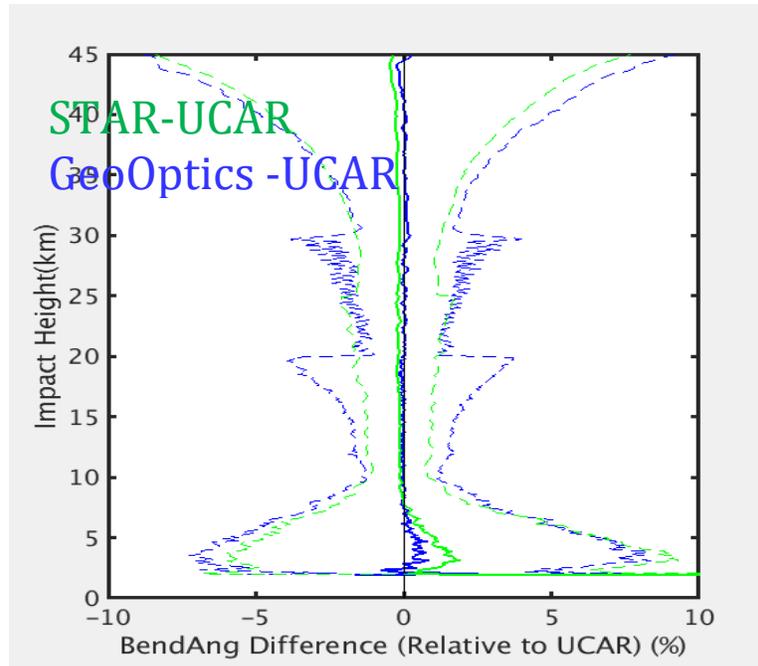


Figure A-3: Flow chart to calculate the propagation of GPS and LEO orbits to circular orbits relative to a local center of curvature.

For the CWDP data operation, we input the Phase/SNR (signal noise ratio) data purchased and processed by the commercial providers into our inversion algorithms and output the bending angle profiles. We further perform an Abel integral transform which converts atmospheric bending angles to profiles of refractivity.

Accurate bending angles processed from the raw GNSS signals received by the RO receivers are critical for numerical weather prediction (NWP) through data assimilation (DA). Using the

STAR inversion package described above, we have successfully inverted the raw phase, amplitude and geometry data provided from GeoOptics to bending angle profiles (i.e., STAR bending angle profiles).



*Figure A-4: Bending angle profiles difference between STAR and UCAR pairs (in green) and between GeoOptics and UCAR pairs (in blue). The standard deviation of the mean differences are in dashed lines.*

Figure A-4 depicts the STAR bending angle profiles compared to the corresponding profiles generated by UCAR (the green line). The bending angle difference between those provided by GeoOptics and UCAR are in blue. GeoOptics data collected for the whole month of December 2018 are used in this comparison of more than 2300 profiles. Figure A-4 shows that the STAR bending angle are very close to those of UCAR with less than 0.1% of fractional difference. The standard deviation of the mean difference (the dashed lines) for STAR-UCAR pairs are smaller than those of GeoOptics-UCAR pairs especially at the impact height between 10 km and 30 km.

Figure A-4 demonstrates the feasibility of STAR RO inversion package to process the raw RO phase, amplitude and geometry data and generate reasonable RO bending angle profiles which are of the same quality of those UCAR processed bending angle profiles.

Although not shown, we expect the STAR processed bending angle profiles will provide positive impact to NWP, as similar to the result demonstrated by JCSDA using the UCAR processed GeoOptics bending angle profiles.

## A.2.2 STAR 1D-VAR Inversion Package

The 1D-var methodology is used to retrieve temperature  $T(P)$  and water vapor partial pressure  $P_w(P)$  from observed refractivity  $N(P) = 10^6 \times (n(P) - 1)$  at atmospheric pressure level  $P$  using optimal estimation technique (Rodgers, 2000). Retrieved variables (temperature and water vapor) at each vertical level are processed independently, thus the “measurement vector”

(terminology from Rodgers, 2000) is degraded to a scalar value for the RO 1D-var algorithm, and “retrieved vector” consists of two elements, i.e. temperature and water vapor on each level. One radio occultation event provides vertical a profile of atmospheric refractivity, from which profiles of atmospheric constituents  $T(P)$  and  $P_W(P)$  are retrieved in corresponding altitude range. Refractivity  $N$  is related to the atmospheric temperature  $T$  and water vapor partial pressure  $P_W$  at the perigee of line of sight with pressure  $P$  by the following equation (Bean and Dutton, 1966):

$$N = FM(T, P_W) = 77.6 \frac{P}{T} + 3.73 \times 10^5 \frac{P_W}{T^2} \quad (0-1)$$

Here,  $P$ ,  $T$ , and  $P_W$  are in mbar, Kelvin, and mbar units respectively, and  $FM$  means RO Forward Model. For given measurement,  $Y_{OBS}$ , the optimal estimation of retrieval vector  $X = (T, P_W)$  is derived from the following iterations:

$$X_{i+1} = X_0 + (K_i^T E^{-1} K_i + B^{-1})^{-1} \times K_i^T E^{-1} \{(Y_{OBS} - Y(X_i)) + K_i(X_i - X_0)\}, \quad (0-2)$$

where  $X_0 = (T_0, P_{W_0})$  is the first guess,  $K = \left(\frac{\partial N}{\partial T}, \frac{\partial N}{\partial P_W}\right)$  is Jacobian vector,  $B$  is the a priori background covariance matrix, and  $E$  represents instrument noise and the forward model error covariance matrix. The index  $i$  corresponds to iteration number. Iterations are continued until residual between observed refractivity  $N_{OBS}$  and simulated one from retrieved state  $N_i = FM(T_i, P_{W_i})$  satisfies the condition  $\frac{|N_{OBS} - N_i|}{N_{OBS}} \leq 0.001$  (0.1 %). The retrieval is considered to be successful when two conditions are met: 1) iterations are converged, i. e.  $N$ -residual has reached the threshold value; and 2) retrieved water vapor partial pressure  $P_{W_i}$  is positive.

To take into account the seasonal and spatial variability of the Earth’s atmosphere, the background atmospheric covariance matrix  $B$  and measurement noise matrix  $E$  have been evaluated separately for five latitude zones (North Polar 90N to 60N, North Mid-Latitudes 60N to 20N, tropics 20N to 20S, South Mid-Latitudes 20S to 60S, and South Polar 60S to 90S) and by monthly averaging.

NOAA/NCEP Global Forecasting System (GFS) data have been used to evaluate the background state covariance matrix  $B(Lat, Mon)$ . Averaging has been done separately for each month and each of five latitude zones. Thus, 60 total estimates of background states have been computed using mean profiles and standard deviation for atmospheric temperature and water vapor. The diagonal elements of covariance matrix  $B$  are set equal to the square of corresponding standard deviation, while off-diagonal elements are set to zero.

The same set of GFS background states has been used to derive simulated refractivity  $N$ , using RO forward model. Then measurement noise matrix  $E(Lat, Mon)$  was defined as a matrix with off-diagonal elements equal to zero and diagonal elements proportional to the variance of  $N$ , on monthly basis and separately for each latitude zone:

$$\begin{aligned} E_{ij} &= \{\gamma_{Noise} \times \sigma(N_i)\}^2, & i = j \\ E_{ij} &= 0, & i \neq j \end{aligned} \quad (0-3)$$

Low values of  $\gamma_{Noise}$  lead to instability of the inversion, i.e. fewer successful retrievals. On other hand, high values of  $\gamma_{Noise}$  result in loss of measurement information content, and retrievals would converge to the first guess state. After trials, it was found that  $\gamma_{Noise} = 0.1$  is optimal value to tradeoff between these two factors. For this value of  $\gamma_{Noise}$ , the measurement noise is about 1% of expected observed refractivity. Finally, the first guess profiles used to initiate the retrieval iteration are taken from GFS Forecast data within 3-6 hours and 1-degree of latitude/longitude from time and location of RO observations.

STAR RO 1D-Var retrieval algorithm has been intensively validated on COSMIC and COSMIC-2 data to prove its workability and robustness before applying it for CWDP analysis. Detail validation analyses of the STAR 1D-var retrievals are not shown in this report.

### A.2.3 STAR ICVS Tool Development

NOAA/NESDIS/STAR plays an important role in the evaluation and longterm monitoring of satellite data quality. The STAR Integrated Calibration and Validation System provides routine monitoring of instrument status and state of health, calibration parameters, and data quality. As part of this pilot project, a STAR testbed center was established to routinely process the CWDP data and anICVS was also developed to routinely monitor the quality of the derived RO data.

A STAR operational inversion package was developed to process COSMIC, Metop-C, and CWDP 2 data. CWDP R2 data were obtained from both GeoOptics and SPIRE Global. An interface page of the STAR GNSS RO ICVS is shown in Figure A-5.

The references used in the RO ICVS include weather model outputs, reanalysis, data from various RO missions, and radiosonde measurements. The functions of the STAR GNSS RO ICVS include:

- Near real time and long-term instrument status, performance monitoring, and anomaly diagnosis
- Near real time and long-term level 1 data product quality monitoring
- Real time support for sensor calibration activities
- Rapid and preliminary estimates of satellite data impact in NWP applications
- Assurance of the integrity of the climate data records from all satellite instruments
- Routine comparisons of atmospheric profiles with other satellite observations and retrievals including microwave, and infrared.
- Routine comparisons of profiles with those from Radiosondes.
- Dynamic web interface with many capabilities.
- Long-term monitoring of the parameters.

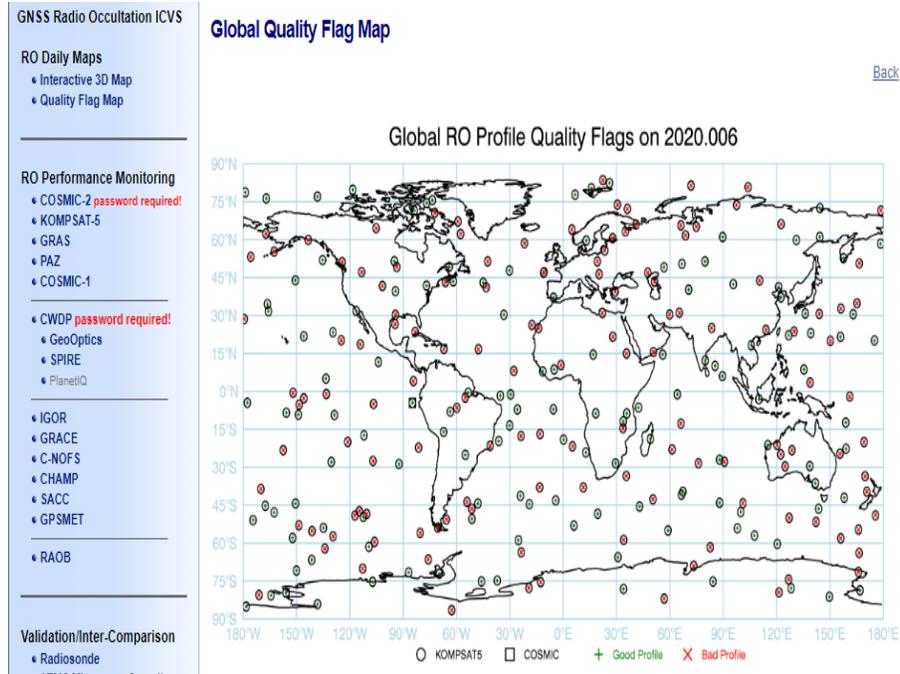


Figure A-5: An interface page of the STAR GNSS RO ICVS.

Currently, RO data from 12 publicly available missions are included, from GPSMET collected in 1995 to COSMIC2 data in 2019. RO data products from centers including UCAR, ROMSAF, ROPP and STAR (under testing) are analyzed.

## A.2.4 STAR Multi-Sensor Validation System

The main goal of the STAR Multi-sensor Validation System is to perform the evaluation of data from various RO missions with respect to background/infrared (IR)/Microwave (MW) observations through direct comparison or through converting RO data to brightness temperatures using the Community Radiative Transfer Model (CRTM). In this report, we also use RO measurements from MetOp missions as references to compare with either CWDP RO data or the IR/MW data. The comparison results for GeoOptics and SPIRE are shown in Section A-4 and Section A-4, respectively.

### A.2.4.1 Multi-Sensor Data

#### (1) Radio Occultation Data

The radio occultation data analyzed in this work and their L2 data processors are listed in Table A-1. MetOp-A/-B obtained through EUMETSAT’s Radio Occultation Meteorology Satellite Application Facility (ROM SAF) (<http://www.romsaf.org>) are used as the reference to evaluate the CWDP data. In this report, RO data from GeoOptics and SPIRE processed by vendor and UCAR, respectively, are analyzed.

Table A-1: Radio occultation data analyzed in this work

RO Instrument	L2 Data Processor
---------------	-------------------

<b>MetOp-A</b>	ROMSAF
<b>MetOp-B</b>	ROMSAF
<b>GeoOptics-085</b>	GeoOptics
<b>GeoOptics-087</b>	GeoOptics
<b>GeoOptics-085</b>	UCAR
<b>GeoOptics-087</b>	UCAR
<b>SPIRE-090</b>	SPIRE
<b>SPIRE-099</b>	SPIRE
<b>SPIRE-090</b>	UCAR
<b>SPIRE-099</b>	UCAR

## (2) Hyperspectral Infrared Sounder Data

High spectral resolution radiance data obtained from the Cross-track Infrared Sounder (CrIS) instrument are used to compare to the brightness temperatures converted from GeoOptics and SPIRE temperature and water vapor profiles.

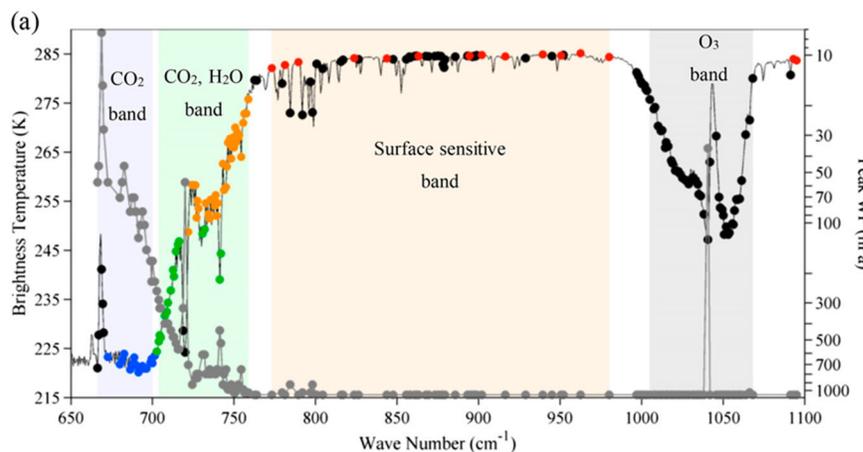


Figure A-6: Peak weighting function pressure of the CrIS LWIR band vs. spectral wave-number together with brightness temperature observed over typical scene.

CrIS is a Fourier transform spectrometer. CrIS is the first in a series of advanced operational sounders that provide more accurate, detailed atmospheric temperature and moisture observations for weather and climate applications. It is the key instrument currently flying on the Suomi NPP and NOAA-20 satellites, and represents a significant enhancement over NOAA's legacy infrared sounder—the High Resolution Infrared

Radiation Sounders (HIRS). CrIS provides soundings of the atmosphere with 2211 spectral channels, over three wavelength ranges: LWIR (9.14-15.38  $\mu\text{m}$ ), MWIR (5.71-8.26  $\mu\text{m}$ ) and SWIR (3.92-4.64  $\mu\text{m}$ ). The CrIS scans a 2200km swath width ( $\pm 50$  degrees), with 30 Earth-scene views. Each field consists of nine fields of view which are organized as 3x3 array of 14km diameter spots (nadir spatial resolution).

Figure A-6 shows the peak weighting function location of CrIS LWIR (9.14-15.38  $\mu\text{m}$ ) channel, whose corresponding pressure levels range from  $\sim 0.1$  to  $\sim 1000$  hPa. In this study, the LWIR band channels will be used for comparison.

### (3) Microwave Sounder Data

The Advanced Technology Microwave Sounder (ATMS) is a cross-track scanner with 22 channels which provides sounding observations needed to retrieve profiles of atmospheric temperature and moisture for civilian operational weather forecasting as well as continuity of these measurements for climate monitoring purposes. Like the long heritage of its predecessors, ATMS combines all the channels of the preceding AMSU-A1, AMSU-A2, and AMSU-B sensors into a single package with considerable savings in mass, power, and volume. The ATMS covers 22 channels in bands from 23 GHz through 183 GHz. ATMS is currently onboard the NPP and NOAA-20, and will be on the follow-on JPSS missions.

Figure A-7 shows the weighting functions of ATMS channels. In this work, channels 6-15 of ATMS, whose peak sounding height ranges from 5 to 50 km, are used as reference to validate the RO temperature and/or humidity profiles.

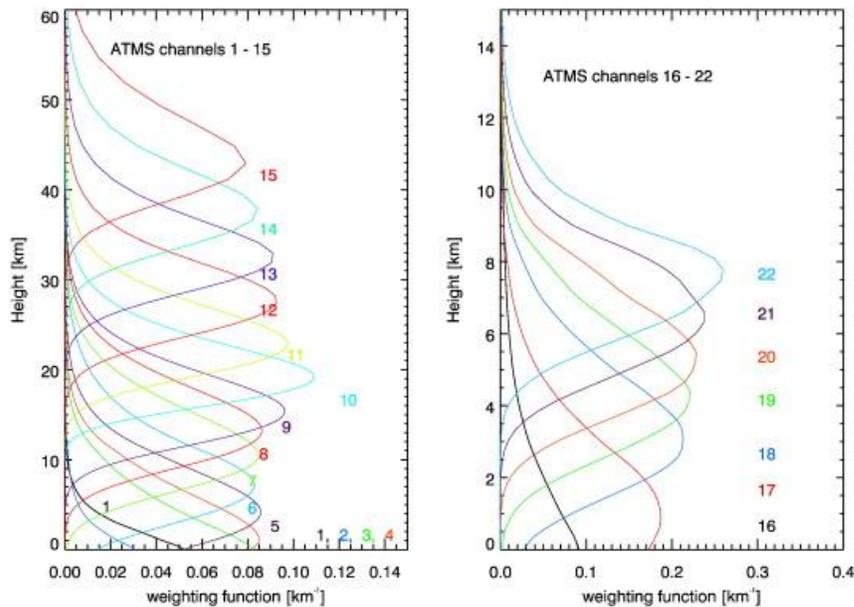


Figure A-7: Weighting function of ATMS channels.

### A.2.4.2 Validation Approach

#### (1) Radiative Transfer Model Simulations

This study makes use of the Community Radiative Transfer Model (CRTM) to convert the thermal profiles retrieved from RO observations to simulated brightness temperatures for direct comparison to satellite sounder observations. CRTM is a fast radiative transfer model for calculations of radiances for satellite infrared or microwave radiometers, and is able to output infrared/microwave radiance and brightness temperature at CrIS/ATMS's spectral resolutions.

The modeling bias analysis scheme enables evaluation of CWDP RO data bias at different pressure levels with respect to atmospheric reanalysis background data through comparison with the CRTM-modeled brightness temperature using the temperature from infrared and microwave sensors and global atmosphere reanalysis data from ECMWF.

Table A-2 summarizes the setup of the simulation model for radiative transfer modeling (RTM). In the case of cross-instrument (IR/MW vs. RO) temperature bias evaluation, the temperature and humidity profile data retrieved from RO sensor measurements are used as input to the CRTM simulation. The CRTM-simulated BT from RO, e.g. GeoOptics, SPIRE and MetOp-A/B, measurements and observed BT by IR/MW sensors are compared for cross-instrument bias evaluation.

*Table A-2: Input Variables and Parameters for RTM Simulation with CRTM.*

Category	Variable	Data source
Atmosphere	Level and layer pressure temperature	RO
	Specific humidity	ECMWF or RO
	Ozone mass mixing ratio	ECMWF
Surface	Water type	1 (sea water)
	Skin temperature	ECMWF
	Wind speed	ECMWF
	Wind direction	ECMWF
Geometry	Altitude	Satellite data
	Satellite zenith angle	Satellite data

	Satellite azimuth angle	Satellite data
	Solar zenith angle	Satellite data
	Solar azimuth angle	Satellite data
	Latitude/longitude	Satellite data
Parameters	Climatology	U.S. standard profile
	Water coverage	1 for ocean

Surface parameters including skin temperature, wind speed and wind direction are obtained from ECMWF’s ERA-Interim reanalysis model which is based on 6 hourly increments (<http://www.ecmwf.int/en/research/climate-reanalysis/era-interim>). The 0.75° spatially gridded model level product has a resolution equivalent to ~80 km.

Ozone, water vapor, and temperature (used for O-B comparison) profiles are also obtained from ECMWF’s ERA-Interim reanalysis pressure level data which are coordinated at 37 mandatory pressure levels and is available from ground to up to ~0.1 hPa.

## (2) Matchup and Screening Algorithm

In the RO vs. background/model data analysis, the matchup condition is set as restricting the co-location between RO sensor data and background/model data to be within 200 km distance difference and a 3-hour time difference. The primary datasets used for the O-B evaluations, *i.e.* RO temperature minus ECMWF temperature profile are based on June 8, 2019 measurements and no other restriction has been used. The RO vs. MW comparison are based on June 8, 2019 measurements and latitude range of RO data has been limited to be within  $\pm 60^\circ$ . For RO vs. IR measurement comparison, June 8 to June 13, 2019 satellite and ECMWF data have been used. To screen out cloud effect for IR, IR measurements with BT difference in surface channel larger than 4 K are excluded. In addition, for the comparison of RO with MW/IR observations, only those MW/IR nadir-view pixels are used to minimize the viewing geometry effect.

## A.3 GeoOptics: Processing, Analysis, and Validation

GeoOptics delivered data starting in November of 2018. Delivery of the L1a data continued through October 2019, while delivery of the L2 data ended in September 2019. Details of the data delivery, data types, and versions can be found in Appendix A.A. An analysis of the data quality at the various levels follows.

### **A.3.1 Assessment of GeoOptics L1a Data Quality**

In this section, we examine the quality of GeoOptics L0-L1 data including clock measurements, excess phase and doppler, and signal to noise ratio (SNR) which directly affect the accuracy of L2 bending angle and refractivity profiles.

#### **A.3.1.1 Bending angle Uncertainty due to Clock Bias**

The GeoOptics satellite has only one POD antenna pointing to zenith direction and one occultation antenna pointing to rear direction of the satellite movement. For comparison, COSMIC has POD antennas for high rate reference link data set recording (50HZ). This data, coupled with the high rate Occultation data, can be used in single differencing to remove receiver clock errors. However, the GeoOptics datasets do not include high link reference data (e.g. in opnGns data format). Thus, a zero differencing method was used to remove LEO clock error in excess phase calculation.

Another approach would be to use single differencing, but with 1 HZ observations from POD antenna for a selected GNSS satellite. However, to obtain 100 Hz data from 1HZ data interpolation is needed. The interpolation error could be significant depending on the clock stability. If the clock bias is large and unstable, then the interpolation error can be large. Xia et al. (2016) estimate the refractivity difference between using the non-differencing and single differencing methods can be as large as 3%. In this section, we evaluate the GeoOptics clock bias time series.

GeoOptics provides POD data with 1 HZ recording frequency in SP3-D format as well as netCDF format. (see Table A-16). These two files have consistent clock bias values. Figure A-8 shows one day clock bias and drift rate for GeoOptics satellite cic085. Figure A-8 shows that the clock is adjusted every few hours. The bias varies between 0 to -600  $\mu\text{s}$ , but the drift variation has a very small magnitude. The clock bias/error has an approximately linear drift rate about 3.5ms/day (40ns/s).

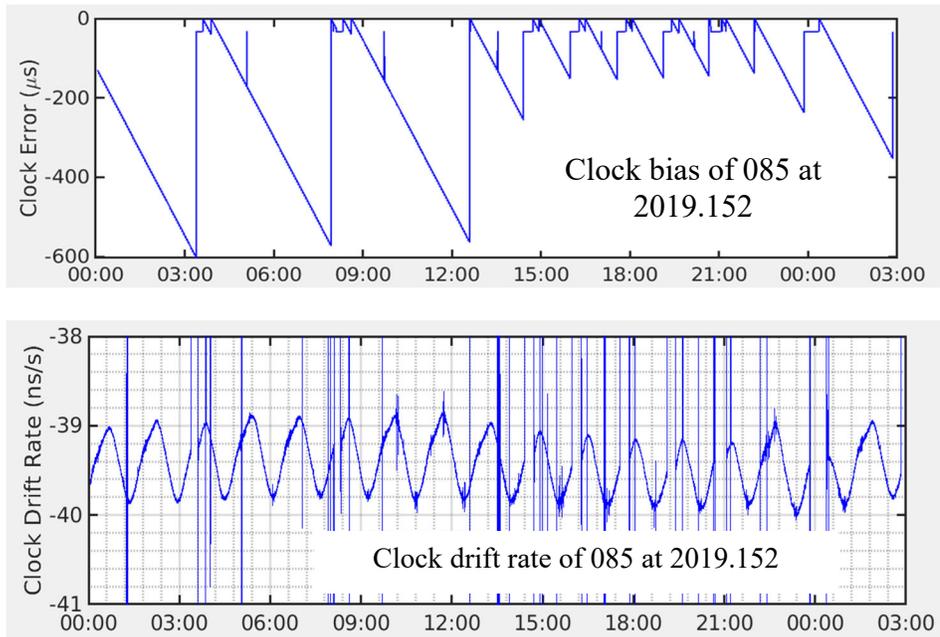


Figure A-8: GeoOptics 085 Clock bias and Drift in one day

For comparison, COSMIC (FM-1) has a very stable clock with a range of fluctuation less than  $2\mu\text{s}$  per day. The drift rate of the #85 is much larger with a value of  $40\text{ns/s}$ , or  $\sim .5\text{ms/day}$ . Moreover, the drift rate plot in Figure A-8 does not indicate a constant drift rate, but rather one that is changing with time similar to a sinusoidal function with an amplitude of  $\sim 1\text{ns/s}$ .

As illustrated in Schreiner et al. (2010), the single differencing using high rate referencing link data from POD antennas can significantly reduce bending angle error compared to zero differencing. However, the interpolation of the GeoOptics clock correction from 1HZ to 100 HZ is needed to perform zero differencing directly using POD clocks or single differencing with 1 HZ POD observations (from RINEX files). Problems can be introduced if the clock is not stable or not completely represented using a polynomial interpolation. Even if the 1HZ clock bias can be estimated well, the residual error in bending angle relies on the interpolation scheme and smoothing filter applied to the excess phase prior to the application of the bending angle calculation algorithm. Though the Clock uncertainty from the POD file is about  $0.8\text{ cm}$  ( $1\sigma$  value), the actual error in a single excess phase profile can be easily more than centimeter (with light speed multiplied).

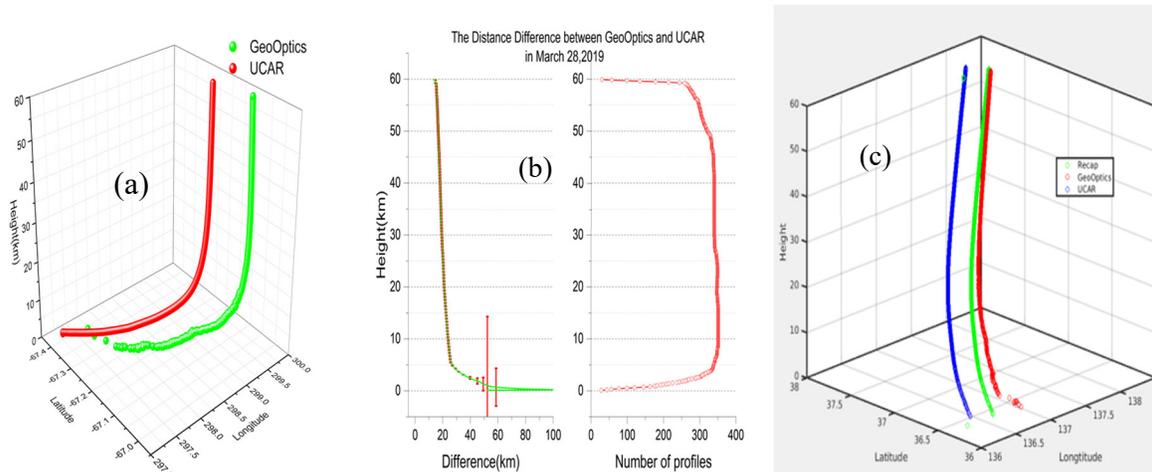
### A.3.1.2 Geolocation Differences between Processing Centers

When comparisons are made between two datasets, generally a collocation of the observations is performed in space and time using geolocation information for each observation. In the case of comparisons made between bending angle profiles produced by two different processing centers, it is sufficient to ensure the profile pairs compared come from the same occultation. These pairs can be identified using other information, e.g. LEO ID, GNSS ID, and timing of the occultation. Since the comparison is performed as a function of impact height, the geolocation information produced by each center is often not a factor.

However, different processing techniques can result in different geolocation results. These differences can arise from differences in the POD solution or in deriving the associated impact

height for the bending angle profile. In this section, we look into bending angle profile geolocation difference between the GeoOptics products and UCAR reprocessed products and try to understand the possible reasons for the differences.

Figure A-9 shows bending angle geolocation comparison between the GeoOptics and UCAR processed data. The longitude/latitude difference is quite clear in this figure and looks like a systematic bias. GeoOptics reports the tangent point latitude/longitude while UCAR uses the perigee latitude/longitude. As the bending angle approaches zero (at high impact heights), the horizontal distance between the two profiles remains about 20km.



*Figure A-9: Geolocation Difference between GeoOptics and UCAR bending angle profiles. a) 3D bending angle profile comparison. B) Mean horizontal difference between GeoOptics and UCAR products c) Considering J2000/TOE difference, the geolocation difference can be reduced significantly.*

The retrieval of bending angle from excess phase employs two methods: the geometric optics method and the wave optics methods. Both involve the position, velocity, the excess phase (and its time derivative) as well as SNR as function of time. Geometric optics considers single ray propagation from the GNSS transmitter to the LEO receiver. For wave optics approach, the atmosphere multi-path effects are taken into account and the generalized Fourier transformation (canonical transformation) can be applied to the excess phase and time domain, so that the bending angle and impact height can be derived from the transformed space.

While GNSS satellite position and velocity are given in Earth Centered Earth Fixed (ECEF) coordinates, the satellite antenna position and phase center variation are given in instrument fixed coordinate system, the attitude are given as quaternions representing rotation from the Earth Centered Inertial (ECI) reference frame to the spacecraft coordinate system, and the excess phase correction due to general relativity effects must be carried out in the ECI coordinate system. The final bending angle geolocation, longitude and latitude are also derived in the ECEF coordinate system. However, using excess phase or positioning products in different ECI coordinate systems might cause differences in geolocation.

Figure A-9b also shows the geolocation mean difference and standard deviation for all profiles over one month. The bias is persistent and the standard deviation is small. The GeoOptics POD (cicPOD) files are given in J2000 coordinate system, and UCAR usually uses the True of Epoch

(TOE) ECI coordinate system. The geolocation offset might be related to the ingestion of the POD information of GeoOptics directly into their bending angle model.

Figure A-9c shows a correction (green) after considering the difference between J2000 and TOE, the recalculated geolocation is much closer to the original GeoOptics one (red line). The relatively larger surface difference can then be explained by the bending angle difference.

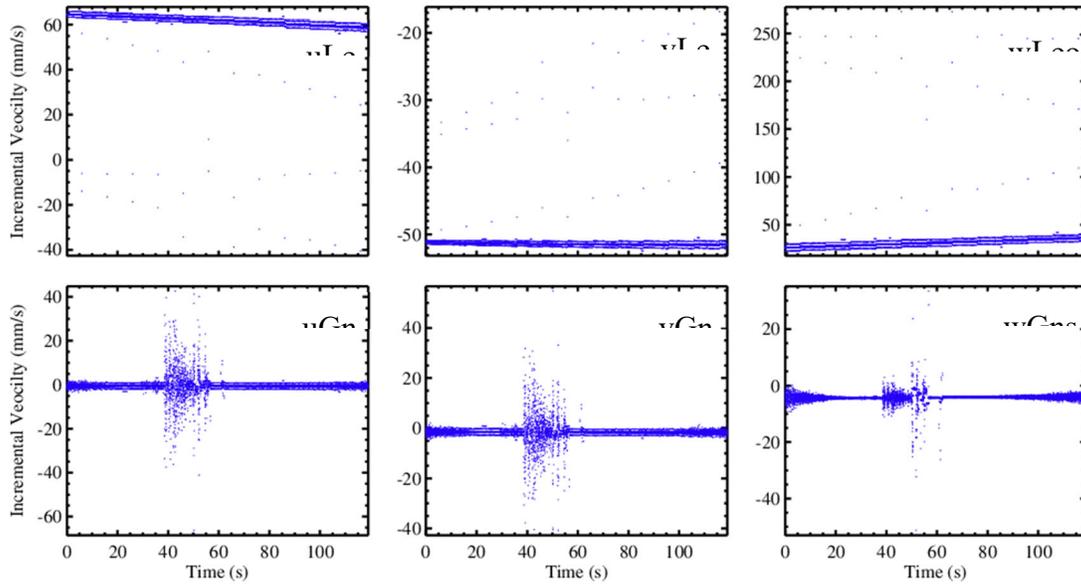
The impact of the geolocation difference might be negligible when directly comparing bending angle profiles from different products, since the bending angle profile pairs are usually selected based on their satellite numbers, time etc. instead of geolocation. However, if these profiles are put into data assimilation system, or used for O-B bias checks, the geolocation difference (about 20km) can cause systematic bias since the horizontal spatial variation of bending angle is not uniform.

### **A.3.1.3 Stability of Satellite Position and Velocities**

The geometry and phase measurements are the fundamental RO variables whose stability and precision directly impact the atmospheric bending angle and refractivity retrievals. The RO Level 1a data contain the time series of geometry, phase, and signal to noise ratio (SNR) which are used to retrieve bending angle and refractivity profiles. In this sub-section, we examine the quality of the GeoOptics geometry, phase, and SNR data.

The stability of the satellite positions are analyzed using the incremental change of the time derivative of the position vectors (velocity vectors). In the Level 1a data, the position vectors are given in Cartesian coordinates as a function of time. The difference between the velocity vectors at consecutive time steps ( $u(t_{i+1}) - u(t_i)$ ) provides information on the stability of the Cartesian coordinates of the positions. If the positions are stable, the incremental velocity vectors have small variations, whereas large variation in incremental velocity shows instability in the position vectors.

Figure A-10 shows the incremental velocity vector components for LEO in the top panel and for GNSS in the bottom panel. The increment in the LEO  $u$ -component shows that the velocity is increasing at  $\sim 60$  mm/sec during the occultation period. The variations at consecutive time steps are small (in the order of mm/second). The variations in the  $v$  – and  $w$  – components are also of similar magnitude. The GNSS incremental velocity vectors are larger, especially from 39 – 60 seconds for the  $u$  – component. It shows that the GNSS positions have larger instability than the LEO positions.



*Figure A-10: Incremental velocity vectors for (top) LEO and (bottom) GNSS for LEO cicero 085 and GNSS satellite G21 on 2018-07-01 at 00:01 Z.*

The time series of the variations in the satellite positions can be illustrated by examining the mean absolute deviation of the incremental velocity vector. Figure A-11 shows daily mean absolute deviation (MAD) of the incremental velocity vectors for GeoOptics satellites #85 and #87 for July 2019. The MAD is calculated as the average of the absolute deviation of each incremental velocity from the mean calculated over the duration of the occultation measurement. The MAD for all the profiles in for each day are then averaged to calculate the daily mean absolute deviation values. The mean absolute deviation of the LEO incremental velocity vectors are in the range of 2 – 3 mm/sec, and the GNSS incremental velocity in the range of 10 – 70 mm/sec. The large jump on day 17 and 19 are caused by a small number of profiles with very large values. For comparison, the daily averaged COSMIC2 MAD are within  $\pm 2$  mm/s for both LEO and GNSS positions.

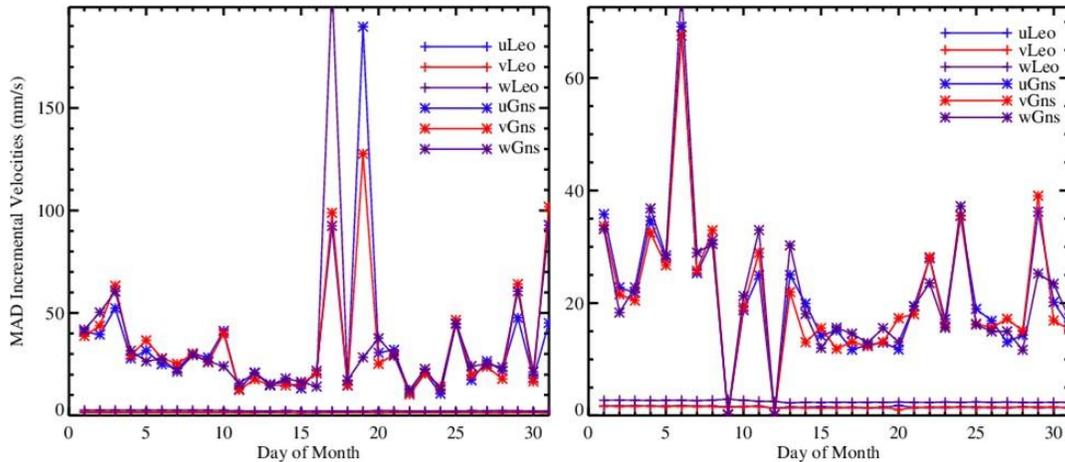


Figure A-11: Mean absolute deviation (MAD) of incremental velocity vectors for July 2019 for (left) cicero085 and (right) cicero087.

#### A.3.1.4 Phase Measurement Noise

The excess phase is the phase difference between RO signal traveling through a vacuum and passing through the earth's atmosphere, and is associated with the state of the atmosphere. The excess Doppler frequency is derived by taking the time derivative of the excess phase provided in the Level L1a data. Variations in the Doppler frequency are related to the noise in the observed excess phase. Figure A-12 shows an example of the excess phase, Doppler and SNR for GeoOptics #85 RO L1b data for GNSS satellite G30 at 00:05 Z on July 01, 2019. The L1 and L2 band data are shown in blue and red curves, respectively. The figure illustrates that excess Doppler, which is the derivative of the phase, shows the noise in the phase measurements more clearly than the excess phase. The excess Doppler is dominated by noise at impact heights below -50 km. The SNR in the bottom panel also shows that below -50 km, the SNR is at a base value.

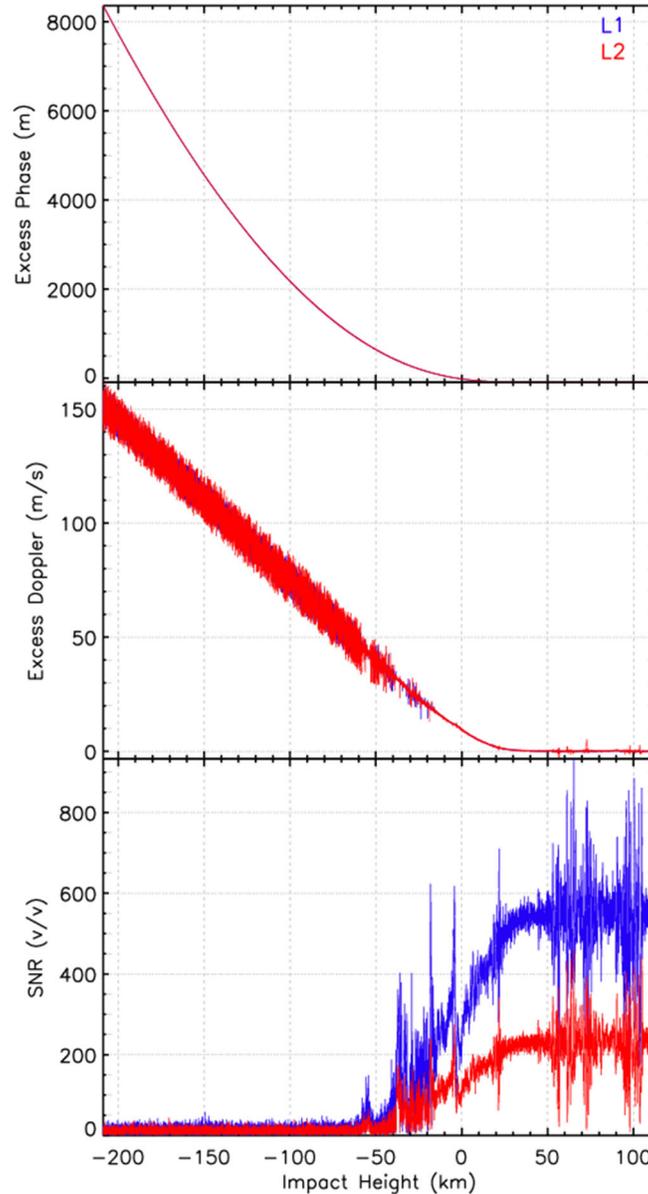


Figure A-12: Excess phase (top), Doppler (middle) and SNR for cic085 GPS satellite G30 at 00:05Z on 2019-07-01.

The variation in the Doppler frequency is analyzed by calculating the standard deviation of the Doppler frequency at intervals of 1-km in impact height. For example, the standard deviation at 5 km impact height is calculated as the standard deviation of the Doppler frequency at 4.5 km to 5.5 km impact height range. Figure A-13 shows the standard deviation of the Doppler frequency as a function of impact height for July 2019 for GeoOptics satellites cic085 and cic087. The figure also shows the mean Doppler frequency as a function of impact parameter. The Doppler frequency (thick lines) are larger closer to the surface because of both (1) decrease in SNR and increased noise and (2) rapid phase accelerations closer to the surface. The figure shows rapid increase in the standard deviation below impact parameter height of -10 km, but < 2 Hz for L1

and < 4 Hz for L2 above -10 km. The smaller standard deviation of the L1 Doppler indicates that the noise in the L1 phase is smaller than in the L2 phase. The standard deviation in the excess Doppler frequency is comparable to COSMIC2.

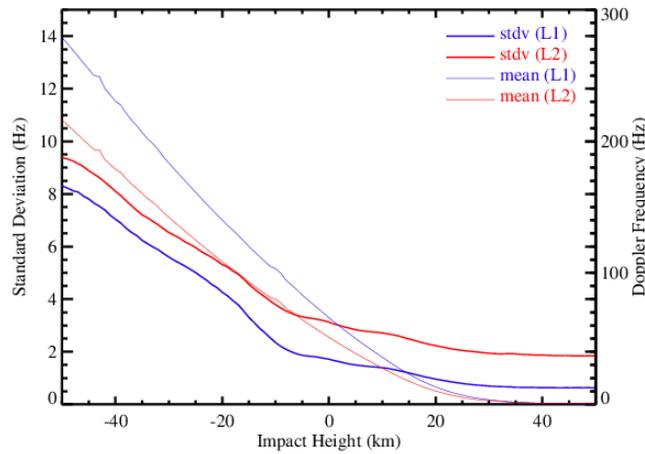


Figure A-13: Standard deviation of the GeoOptics Doppler frequency for July 2019.

### A.3.1.5 Signal to Noise Ratio (SNR) Assessment

The SNR is a measure of the signal strength at the receiver. As RO signals penetrate deeper layers of the atmosphere, the signal attenuates so the SNR is larger at higher altitudes and decreases with altitude. To find useful signal it is important to determine where the noise dominates the signal. It can be determined by finding a base SNR value. The base SNR depends on the instrument and varies for different measurements, but is generally in the range of 15 -25 v/v for both GeoOptics satellites #85 and #87. The mean SNR (calculated over straight line impact heights from 0 km to 85 km) for #85 and #87 are 600 and 525 v/v for L1 and ~225 v/v for L2 (Figure A-14 top panel). The distribution of the SNR for June 2019 is shown in the lower panel of Figure A-14. The SNR on L1 and L2 band extends from 100 v/v to 1500 v/v and 50 – 700 v/v, respectively.

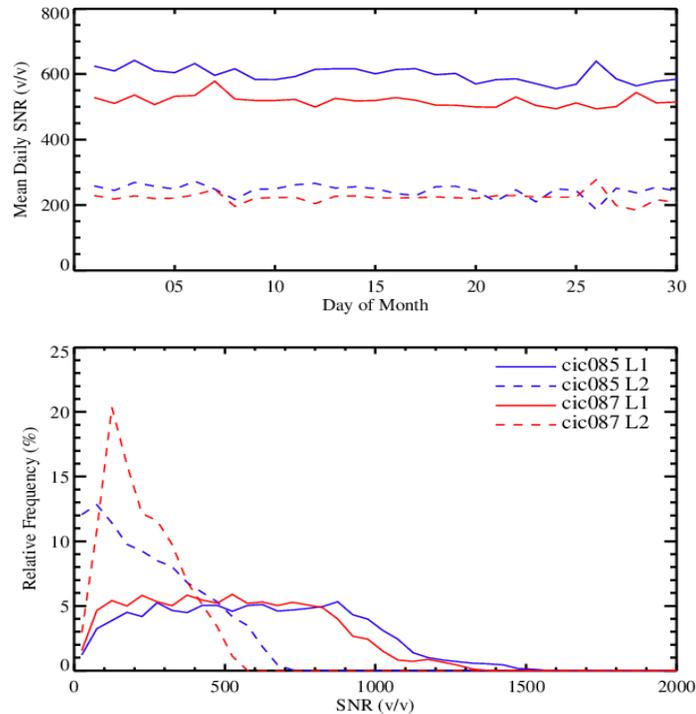


Figure A-14: Daily mean SNR (top), and SNR frequency distribution for June 2019.

### A.3.2 Assessment of GeoOptics L2 Data Quality

In this section, we present an assessment of the quality of GeoOptics L2 data including bending angle profiles, refractivity, and retrieved temperature and moisture profiles. Comparisons to other data sets, including data processed by other centers, NWP model forecasts and analyses, radiosondes, and sounding produced by other sensors, are examined.

#### A.3.2.1 Bending Angle and Refractivity Comparisons

##### A.3.2.1.1 Comparison between GeoOptics Bending Angle and STAR FSI Retrievals

STAR processing of the GeoOptics L1a data to bending and refractivity includes preprocessing of the L1 and L2 components that include (1) coordinate transformation from ECI to ECEF, (2) determining local center of curvature and establishing new center of local curvature, (3) data truncation using dynamic SNR threshold, (4) applying FSI retrievals on L1 and L2 excess phases, (5) ionospheric correction and quality control on bending angles, and (5) inversion to refractivity using Abel inversion.

An inter-comparison of the bending angles from STAR, UCAR and JPL processing (for GeoOptics) shows stable STAR bending retrievals (Figure A-15). There is a positive bias in the STAR bending angle below 3 km relative to both UCAR and JPL bending angles.

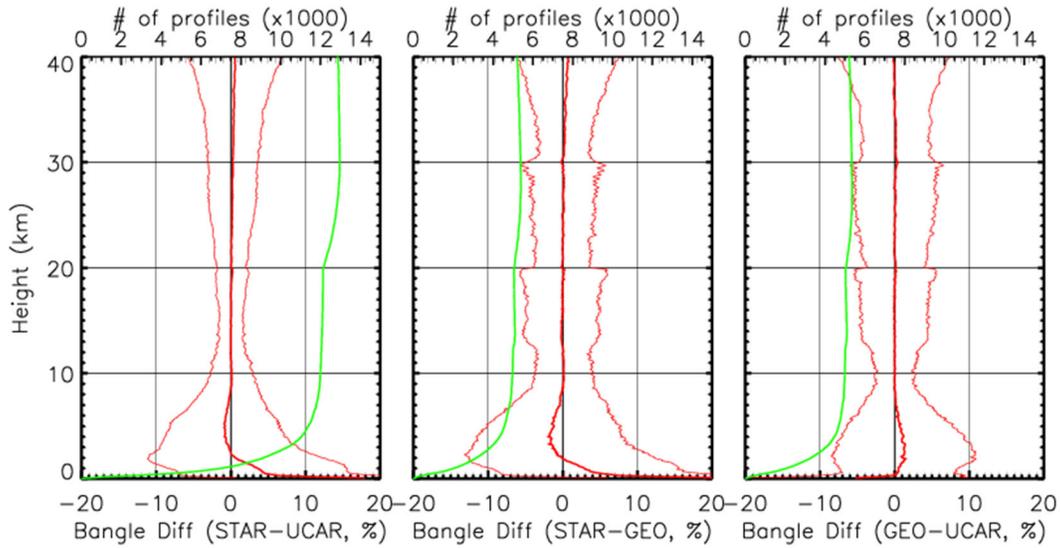


Figure A-15: Fractional bending angle differences for (left) STAR-UCAR, (center) STAR – JPL, and (right) JPL – UCAR.

### A.3.2.1.2 Comparison between GeoOptics Refractivity and STAR FSI Retrievals

The refractivity obtained from the three processing centers also show similar biases to the bending angles (Figure A-16). The mean fractional refractivity differences are close to 0 from 6 km up to 35 km.

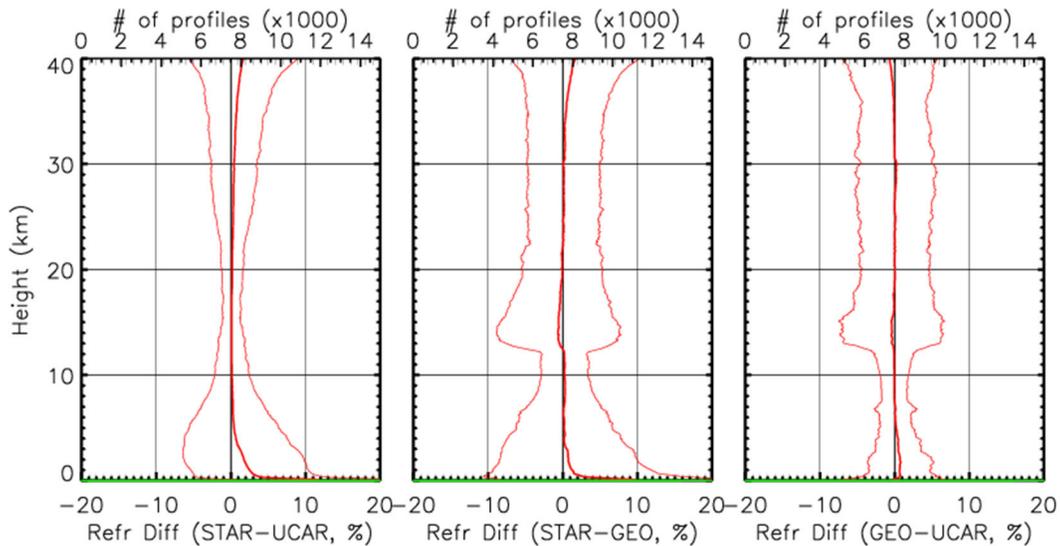


Figure A-16: Fractional refractivity differences for (left) STAR-UCAR, (center) STAR – JPL, and (right) JPL – UCAR.

### A.3.2.1.3 Comparison between GeoOptics Data and STAR ROPP Retrievals

To quantify the general quality of GeoOptics bending angel, we compare multiple months of GeoOptics bending angel with those independently processed by ROPP where GeoOptics excess phases are used as inputs. Figure A-17a and b depict the fractional bending angle difference and temperature difference between ROPP and GeoOptics. Results show that the GeoOptics bending angle profiles and temperature profiles are very close to those derived from ROPP.

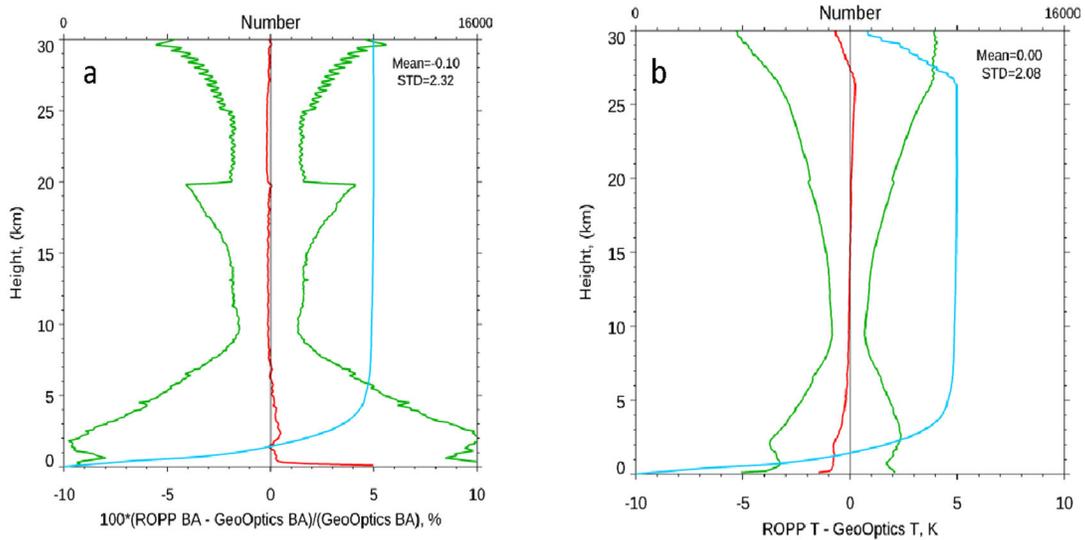


Figure A-17: (a) the fractional bending angle difference between ROPP and GeoOptics and (b) temperature difference between ROPP and GeoOptics

### A.3.2.2 Comparison to NWP Model Fields

#### A.3.2.2.1 Comparison to GFS 6-hour Forecast

Similar to Figure A-17, in Figure A-18 we compare GeoOptics N, T, and q with these obtained from GFS 6-hour forecast. Results show that GeoOptics N, T, and q are very consistent with those from GFS 6-hour forecast.

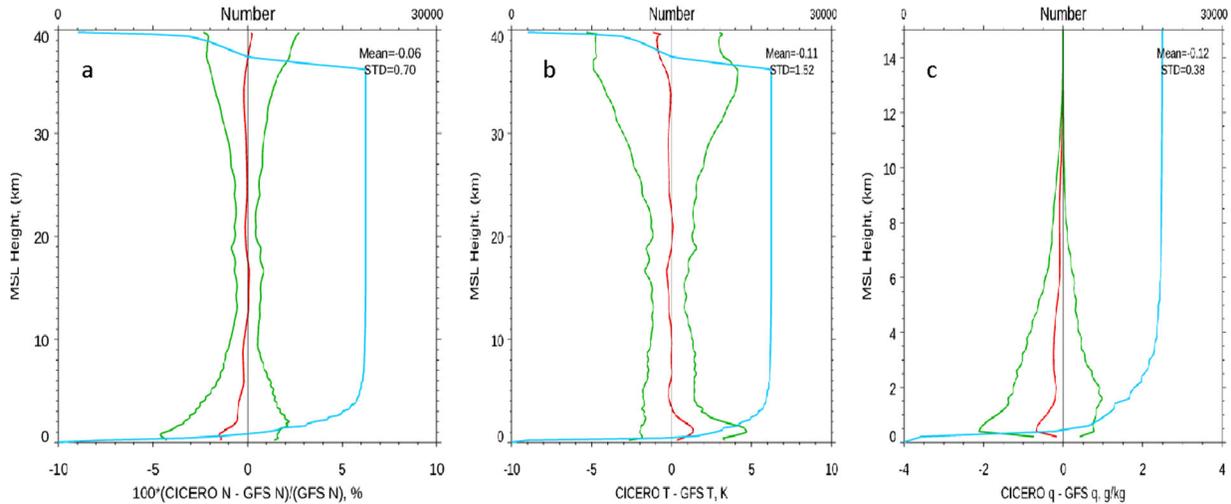


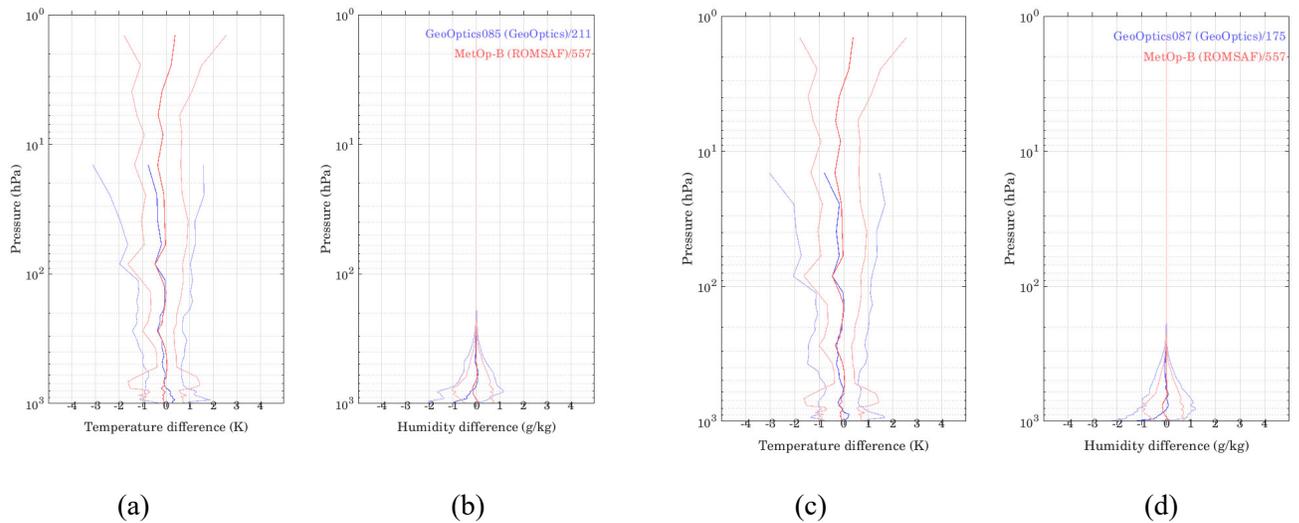
Figure A-18: (a) The fractional refractivity difference, (b) temperature difference, and (c) water vapor mixing ratio difference between GeoOptics and GFS 6-hour forecast, where the GeoOptics and RAOB pairs are within 300 km and 3 hours collected from July 1 to July 31, 2019.

#### A.3.2.2.2 Comparison to ECMWF ERA Interim Reanalysis

The evaluation of GeoOptics RO data quality is performed through O-B comparison with temperature and humidity profiles from European Centre for Medium-Range Weather Forecasts (ECMWF)-Interim global atmospheric reanalysis data, chosen for its comprehensive global and temporal coverage and integration of all available observation data. Both GeoOptics-85 and -87 RO data processed by GeoOptics and UCAR are analyzed. Quality flags in the RO data have been applied. The analysis results are organized according to analysis approach, RO sensor number, and RO L2 data processors. For O-B analysis, the data on 2019-06-08 are selected since all RO instruments of interest have data available on this day.

#### GeoOptics-85/87 (GeoOptics Version) L2 and ECMWF Comparisons

GeoOptics-85 and -87 RO L2 data processed by GeoOptics are compared with ECMWF and the results are shown in Figure A-19 which shows O-B temperature and humidity difference of GeoOptics-85/87 (GeoOptics Version), mean and standard deviation of O-B temperature and humidity bias of GeoOptics-85 and of GeoOptics-87, respectively. The overall performances in term of O-B bias and standard deviation of GeoOptics-85 and -87 RO data are consistent.



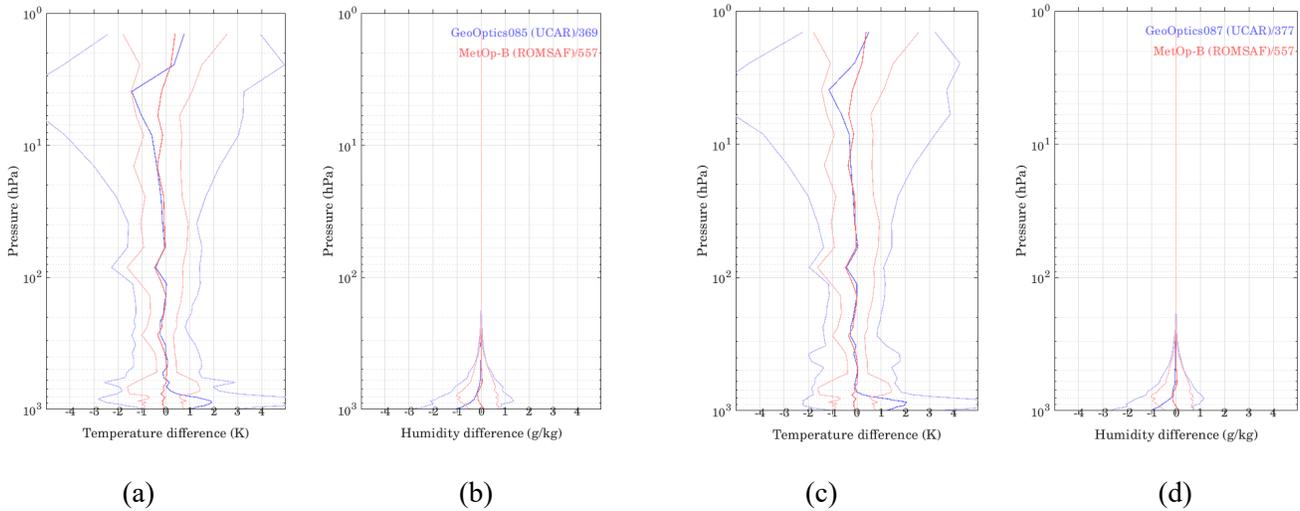
**Figure A-19: O-B bias analysis for GeoOptics-85/87 L2 RO data (GeoOptics Version): Mean and standard deviation of O-B (a) temperature and (b) humidity difference of GeoOptics-85 L2 RO data and mean and standard deviation of (c) temperature and (d) humidity difference of GeoOptics-87 L2 RO data. The mean and standard deviation of O-B bias for MetOp-B RO data are also plotted.**

One main problem to note for GeoOptics-85 and -87 RO L2 data processed by GeoOptics is that the data have a clear the cut-off below 10 hPa, *i.e.* no retrieved data below this pressure level, which prevents comparison with certain channels of sounder sensor data that measure BT below this pressure level.

In Figure A-19, the mean O-B temperature and humidity biases of GeoOptics-85/87 are consistent with MetOp-B over the pressure range between 50 and 700 hPa. For pressure levels greater than 700 hPa (especially near surface pressure level) and less than ~50 hPa, there are significant difference in O-B temperature and humidity biases between GeoOptics-85/87 and MetOp-B. The largest biases for O-B temperature and humidity reach ~0.8 K and 1 g/kg, respectively, for both GeoOptics-85 and -87 L2 data processed by GeoOptics.

### **GeoOptics-85/87 (UCAR Version) L2 Data and ECMWF Comparisons**

The GeoOptics-85/87 L2 RO data processed by UCAR are compared with ECMWF background data in Figure A-20. Figure A-20 shows the O-B (ECMWF) temperature and humidity difference of GeoOptics-85/87 (UCAR Version), mean and standard deviation of O-B temperature and humidity difference of GeoOptics-85 and of GeoOptics-87, respectively.



*Figure A-20: O-B bias analysis for GeoOptics-85/87 L2 RO data (UCAR Version): Mean and standard deviation of O-B (a) temperature and (b) humidity difference of GeoOptics-85 L2 RO data and mean and standard deviation of (c) temperature and (d) humidity difference of GeoOptics-87 L2 RO data. Mean and standard deviation of O-B bias for MetOp-B RO data are also plotted.*

The overall O-B performances of GeoOptics-85 and -87 L2 RO data processed by UCAR are consistent. The data coverage of UCAR version L2 GeoOptics RO data extends below pressure = 10 hPa which extends much more than the L2 RO data processed by GeoOptics (Figure A-19).

Results show that the mean O-B temperature and humidity biases of GeoOptics-85/87 are consistent with MetOp-B over the pressure range between 15 and 700 hPa. At pressure levels greater than 700 hPa (especially near surface pressure level), there are significant difference in O-B temperature and humidity biases between GeoOptics-85/87 and MetOp-B. Overall standard deviations of O-B bias of GeoOptics-85 and -87 are larger than MetOp-B. The largest biases for O-B temperature and humidity reach ~2 K and 1 g/kg, respectively, for both GeoOptics-85 and -87 L2 data processed by UCAR.

### Summary of GeoOptics RO Data and ECMWF Comparisons

Table A-3 summarizes the O-B (ECMWF) temperature bias and standard deviation (K) at different atmospheric pressure level regions for GeoOptics-85 and -87 RO L2 data processed by GeoOptics and UCAR. Similarly, Table A-4 summarizes the results for humidity.

*Table A-3: Summary of O-B (ECMWF) temperature bias and its standard deviation (K) at different atmospheric pressure levels for GeoOptics-85 and -87 RO L2 data processed by GeoOptics and UCAR. Derived from June 8, 2019 data for all RO data. ERA5 data from ECMWF are used as background.*

RO Instrument	L2 Data Processor	Pressure Levels (hPa)				# of Valid Profiles	Note
		600.0--1000.0	90.0-- 600.0	10.0-- 90.0	0.1-- 10.0		
MetOp-A	ROMSAF	-0.093 ± 1.025	-0.108 ± 0.621	-0.203 ± 0.988	-0.040 ± 1.382	572	As reference; Wet T.
MetOp-B	ROMSAF	-0.098 ± 1.054	-0.097 ± 0.627	-0.204 ± 1.001	-0.009 ± 1.403	557	As reference; Wet T.
GeoOptics-085	Geo-Optics	0.034 ± 0.916	-0.160 ± 1.046	-0.399 ± 1.666	N/A(too few data)	211	Wet Temperature
GeoOptics-087	Geo-Optics	-0.0784 ± 0.907	-0.138 ± 1.080	-0.381 ± 1.737	N/A (too few data)	175	Wet Temperature
GeoOptics-085	UCAR	0.785 ± 3.160	-0.076 ± 1.342	-0.250 ± 2.026	-0.899 ± 4.313	369	Wet Temperature
GeoOptics-087	UCAR	0.701 ± 2.901	-0.093 ± 1.361	-0.176 ± 1.986	-0.680 ± 4.387	377	Wet Temperature

GeoOptics L2 data processed by GeoOptics are valid up to pressure > 10 hPa. The overall mean O-B temperature bias of GeoOptics-85 and -87 (GeoOptics version) is close to MetOp-B (reference) except in the pressure region of 10-90 hPa where the GeoOptics L2 data show larger bias and standard deviation. From Figure A-20, the mean O-B bias and its standard deviation of humidity data retrieved from RO data processed by GeoOptics are both larger than those of MetOp-B in pressure regions 200-500 hPa and 800-1000 hPa.

For UCAR L2 GeoOptics data, the overall Mean O-B bias of UCAR L2 data for GeoOptics-85/87 is close to MetOp-B (reference) except larger O-B temperature bias in pressure levels 600-1000 hPa and 0.1-10 hPa in comparison with MetOp-B. The overall standard deviations of O-B temperature bias of UCAR L2 data for GeoOptics-85/87 are consistently larger than those of MetOp-B over all of the pressure levels. The mean O-B bias and its standard deviation of humidity data retrieved from RO data processed by UCAR are both larger than those of MetOp-B in pressure regions 800-1000 hPa (surface region).

Comparison of GeoOptics with UCAR version GeoOptics data shows that O-B temperature bias of UCAR version is significantly larger than GeoOptics version in pressure region 600-1000 hPa. In the pressure regions 10-90 hPa and 90-600 hPa, the O-B temperature bias of UCAR version is lower than GeoOptics version. It is also shown that the O-B humidity bias of UCAR version is larger than GeoOptics version for pressure region 500-800 hPa and 800-1000 hPa, and smaller than GeoOptics version for pressure region 200-500 hPa. It is also worth pointing out that UCAR L2 GeoOptics data produce more profiles than those processed by GeoOptics.

*Table A-4: Summary of O-B (ECMWF) humidity bias and its standard deviation (g/kg) at different atmospheric pressure levels for GeoOptics-85 and -87 RO L2 data processed by GeoOptics and UCAR. Derived from June 8, 2019 data for all RO data. ERA5 data from ECMWF ar*

RO Instrument	L2 Data Processor	Pressure Levels (hPa)			# of Valid Profiles	Note
		800.0--1000.0	500.0-- 800.0	200.0-- 500.0		
MetOp-A	ROMSAF	$-0.0492 \pm 0.6815$	$-0.0659 \pm 0.6380$	$-0.0067 \pm 0.1187$	572	As reference; Wet T.
MetOp-B	ROMSAF	$0.0007 \pm 0.6993$	$-0.0353 \pm 0.6055$	$-0.0021 \pm 0.1324$	557	As reference; Wet T.
GeoOptics-085	GeoOptics	$-0.3839 \pm 1.1920$	$0.0249 \pm 0.7750$	$-0.0185 \pm 0.2139$	211	Wet Temperature
GeoOptics-087	GeoOptics	$-0.2007 \pm 1.1772$	$0.0307 \pm 0.8200$	$-0.0252 \pm 0.2229$	175	Wet Temperature
GeoOptics-085	UCAR	$-0.5089 \pm 1.6133$	$-0.0836 \pm 0.8673$	$-0.0063 \pm 0.1507$	369	Wet Temperature
GeoOptics-087	UCAR	$-0.5421 \pm 1.4728$	$-0.0840 \pm 0.8066$	$-0.0099 \pm 0.1594$	377	Wet Temperature

### A.3.2.3 Data Penetration

Figure A-21a depicts the numbers of observation at each penetration depth for multiple RO missions while Figure A-21b shows the percentage of the observation number relative to the highest observation number at the level above 8 km altitude for various RO missions for July 2019.

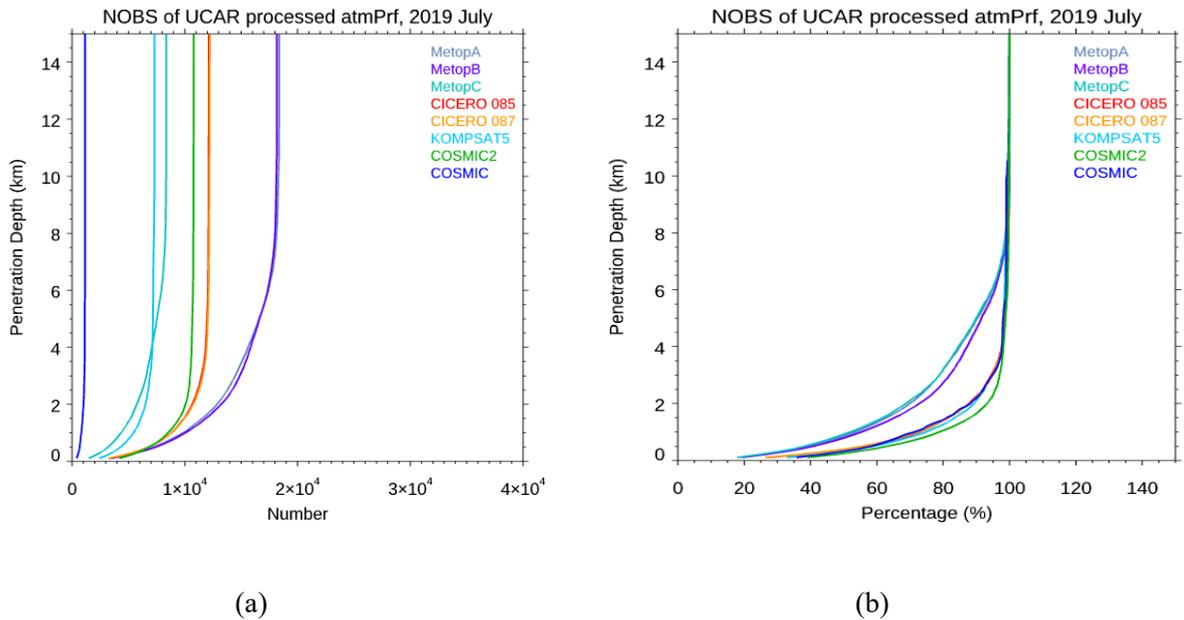


Figure A-21. (a) The numbers of observation at each penetration depth for multiple RO missions for July 2019. (b) The percentage observation number at each penetration depth relative to the highest observation number in the vertical level (usually above 8 km) for multiple RO missions for July 2019.

### COSMIC-1 vs COSMIC-2 SNRs

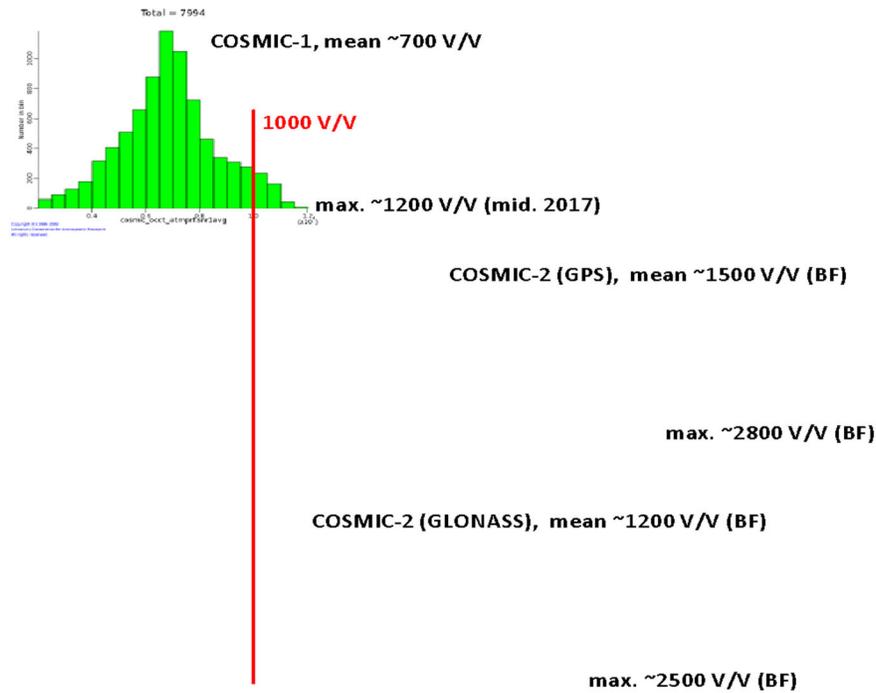


Figure A-22: The histogram of signal to noise ratio (SNR) distribution of COSMIC-1 and COSMIC-2. The middle panel is for the COSMIC-2 receiver SNR with GPS transmitter and the lower panel is for the COSMIC-2 receiver SNR but with GLONASS transmitter.

The histogram of signal to noise ratio (SNR) distribution of COSMIC-1 and COSMIC-2 is shown in Figure A-22. COSMIC-2 uses an advanced receiver known as Tri-GNSS (Global Navigation Satellite System) Receiver System (TGRS) developed by the Jet Propulsion Laboratory and a digitally beam-steered antenna. Compared to the COSMIC-1 receiver, the TGRS will have a significantly increased (SNR).

The histograms of SNR distribution of GeoOptics, KOMPSAT5, and COSMIC are shown in Figure A-23.

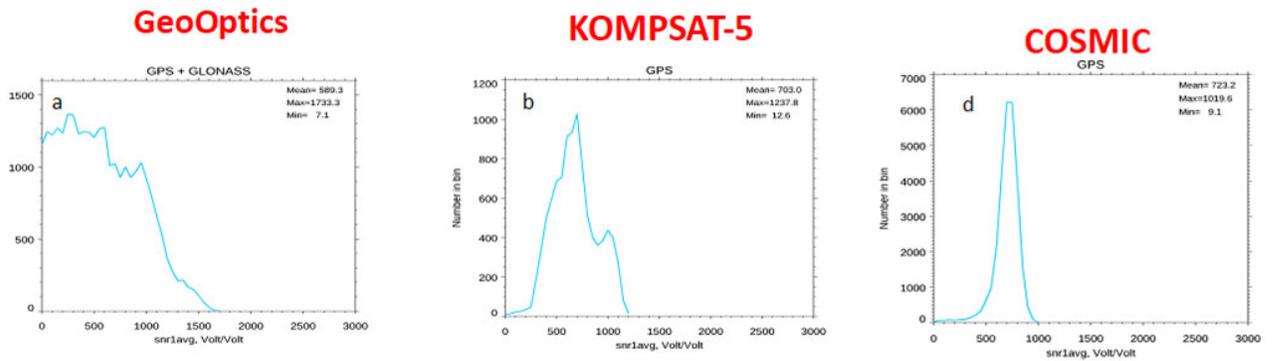


Figure A-23: The histogram of SNR distribution of (a) GeoOptics, (b) KOMPSAT5, and (c) COSMIC.

Figure A-21 shows that with a higher SNR COSMIC-2 has a deeper penetration comparing to all other RO missions. The lowest penetration height of GeoOptics is between those of COSMIC and COSMIC-2. Table A-5 depicts the lowest penetration height of 80% of the total data for different RO missions at different latitudinal zones. The lowest penetration height for both GeoOptics is only slightly higher than those of COSMIC-2 almost at all latitudinal zones where COSMIC-2 is available.

Table A-5: The lowest penetration height of 80% of the total data for different RO missions at different latitudinal zones. All the data are from July 2019.

	10N-10S	10N-30N	10S-30S	30N-45N	30S-45S	45N-60N	45S-60S	60N-90N	60S-90S
Metopa	3.2	7.2	4.0	4.9	2.1	3.2	1.2	3.0	3.6
Metopb	2.6	4.5	3.7	4.0	2.0	2.6	1.3	2.6	3.6
Metopc	2.8	4.7	4.0	4.9	1.8	3.2	1.4	3.0	3.5
Cic085	1.3	0.8	1.2	2.3	0.7	1.6	0.4	1.5	2.8
Cic087	1.4	1.6	1.3	2.1	0.7	1.6	0.4	1.6	2.7
Cosmic	1.5	1.2	1.4	1.9	0.6	1.9	0.4	1.5	2.2
Cosmic2	1.2	0.5	1.1	1.8	0.6				
kompsat5	1.7	1.4	1.4	1.8	0.6	1.2	0.3	1.1	2.8

#### A.3.2.4 Monitoring of Long-term Stability

Because the sensitivity of infrared and microwave instruments may decay in space after launch, it is critically important to have independent observations with long-term stability as climate benchmarks to which measurements from either newly-launched or aged instruments can all refer.

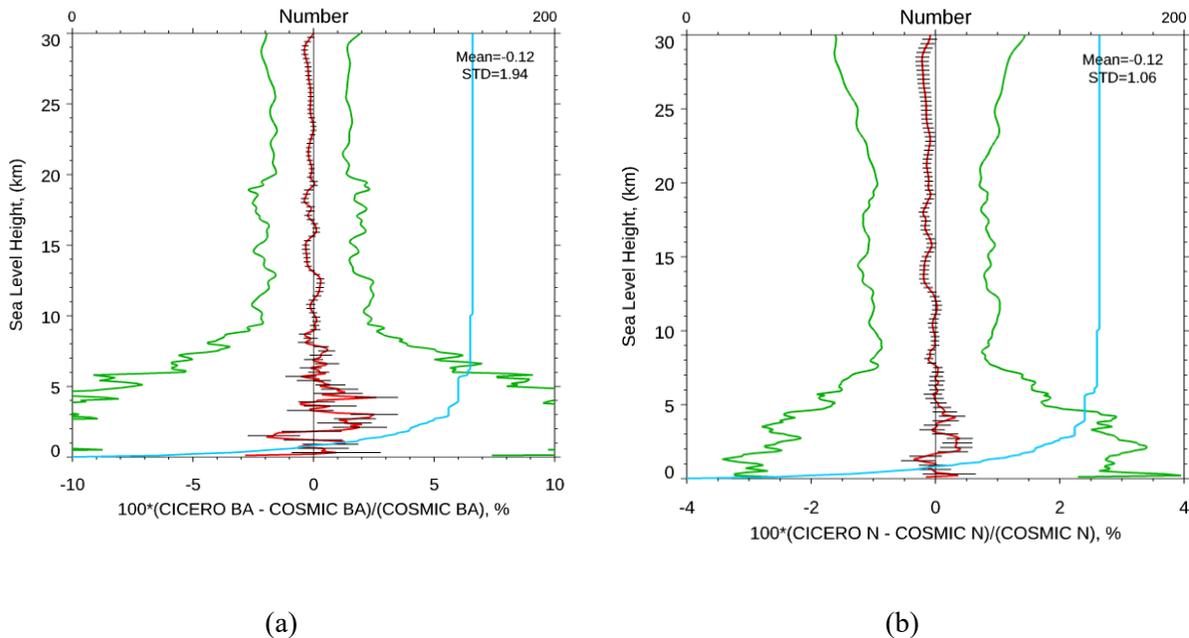


Figure A-24: The fractional (a) bending angle and (b) refractivity differences between GeoOptics and COSMIC where the COSMIC and GeoOptics pairs are within 300 km and 3 hours collected from July 1 to July 31, 2019.

Figure A-24 depicts the fractional bending angle and refractivity comparisons between GeoOptics and COSMIC, respectively. Here we used COSMIC and GeoOptics pairs collocated within 300 km and 3 hours collected from July 1 to July 31, 2019. The comparison results for GeoOptics-KOMPSAT5 pairs, GeoOptics-PAZ pairs, and GeoOptics-Metop-A/-B pairs are also very similar with those shown in Figure 0-16 and are not shown.

The fact that the mean fractional bending angle and refractivity difference in the height ranging from 10 km to 30 km is within the normalized standard error of the mean difference demonstrates long term stability of the GeoOptics RO signals.

### A.3.2.5 Estimation of Retrieval Accuracy

This section presents validation of refractivity and water vapor retrievals obtained from GeoOptics. Water vapor profiles have been derived with NOAA/STAR RO 1D-Var algorithm and compared with collocated RAOB observations. The uncertainty of retrievals is evaluated in terms of bias and standard deviation (StDv) obtained by monthly averaging.

Radiosonde profiles of temperature and moisture, taken for comparison with RO retrievals, have been obtained from UCAR (<https://rda.ucar.edu/datasets/ds351.0/>) where NCEP ADP Global Upper Air Observations Weather Data are collected. These RAOB datasets are available at 20 mandatory levels from 1000 mb to 1 mb, plus a few significant levels, and reported with time intervals from 1 hour to 12 hours daily.

We have used radiosonde type RS92 only for our comparison. Collocation criteria to compile joint RO+RAOB datasets were equal to  $\Delta t \leq 2$  hours and  $\Delta D \leq 300$  km between RO and

RAOB observations. Before comparison, both, RO and RAOB, profiles should be interpolated onto the same altitude/pressure grid. STAR RO 1D-Var retrieval algorithm is valid over the altitude range from 0 km to 39.9 km with increment 0.1 km. RAOB profiles have mean altitude resolution of about 0.5 km. At the first step, RAOB profiles have been interpolated onto RO fine grid. Then both RO retrieved and RAOB profiles have been merged in layers with thickness of 0.5 km, *i.e.* about the actual resolution of RAOB data. For merging, we have used  $\Delta P$ -weighted convolution of initial five RO levels:

$$X_{Merged} = \frac{1}{|P_6 - P_1|} \sum_{i=1}^5 \left( \frac{X_i + X_{i+1}}{2} \right) |P_i - P_{i+1}| \quad (0-1)$$

Here  $X_{Merged}$  is convolved value of temperature or moisture, which have been used for comparison.

One month of GeoOptics and RAOB pairs are collected from July 1 to 31, 2019. The comparison results for fractional refractivity difference, the temperature difference, and water vapor mixing ratio difference are shown in Figure A-25a-c, respectively. The GeoOptics 1D-var retrieved N, T, and q are very consistent with those of RAOB measurements.

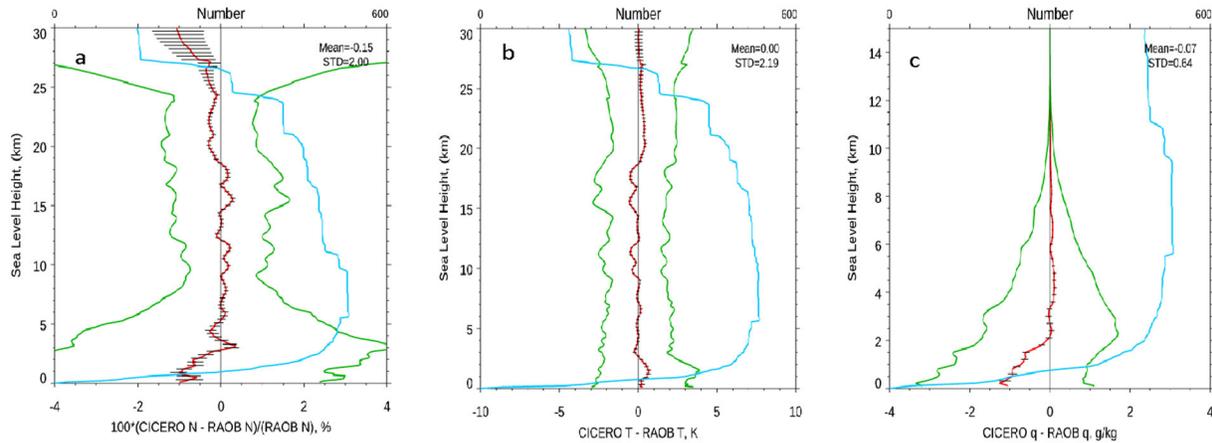


Figure A-25: (a) The fractional refractivity difference, (b) temperature difference, and (c) water vapor mixing ratio difference between GeoOptics and RAOB measurements where the GeoOptics and RAOB pairs are within 300 km and 3 hours collected from July 1 to July 31, 2019.

CWDP GeoOptics v.5 dataset includes RO measurements from two instruments, CIC085 and CIC087, collected on a daily basis over November 2018 to September 2019 (CIC085, 11 months total) and January 2019 to September 2019 (CIC087, 9 months total).

As the first step, GeoOptics bending angles and refractivity have been compared with collocated KOMPSAT5 observations. Criteria of collocation were the same as for RAOB comparison, *i.e.*  $\Delta t \leq 2$  hours and  $\Delta D \leq 300$  km, what resulted in about 80 matchup pairs on average for each month and instrument. Then bias and StDv have been computed by monthly averaging.

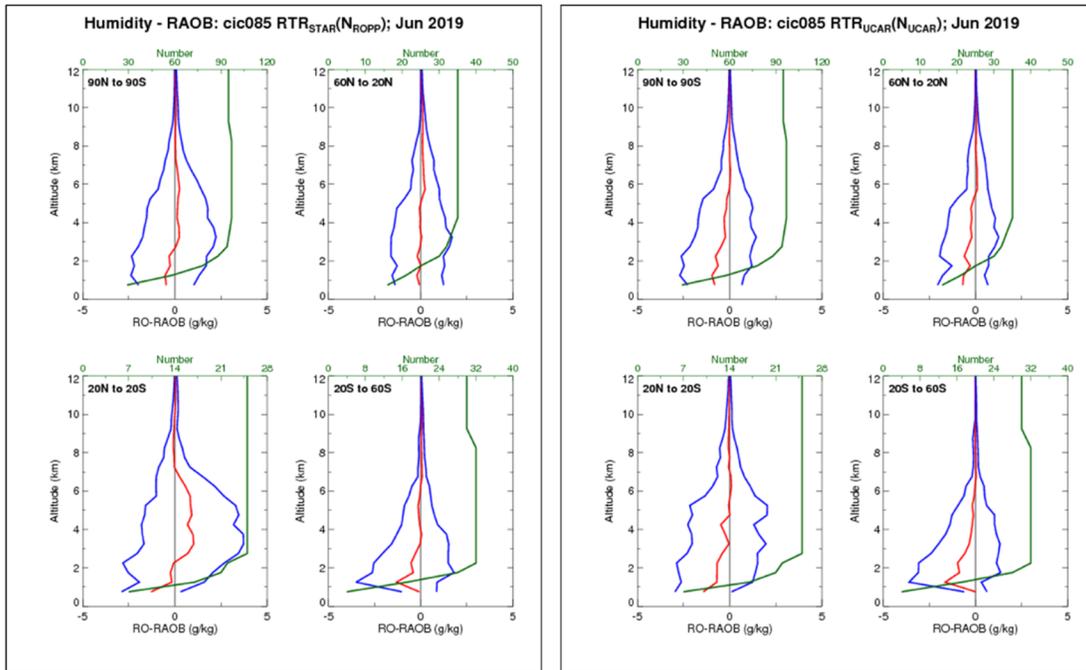


Figure A-26: CIC085-RAOB comparison for June of 2019. The four plots in the left frame depict water vapor, retrieved from ROPP N-values with STAR 1D-Var algorithm and for 4 latitude zones. The four plots in the right frame, organized the same way, present comparisons for water vapor retrieved by UCAR.

STAR RO team has produced N-values from GeoOptics bending angles with the ROPP package and then processed them with 1D-Var retrieval algorithm (referred as STAR/ROPP subscripts). Comparison statistics of STAR/ROPP water vapor with RAOB observations is presented on Figure A-26 and Figure A-27 for June 2019 and both GeoOptics instruments, respectively. In addition, comparison was also done for UCAR-retrieved water vapor taken from corresponding *wetPrf* files. Although the magnitude of GeoOptics-RAOB differences is about the same for STAR/ROPP and UCAR retrievals, one can note that STAR/ROPP humidity has less bias with RAOB in the lower troposphere up to altitude  $\sim 3$  km than UCAR one, and bigger positive bias in the middle troposphere, especially noticeable at tropics zone. StDv is about the same order for both, STAR/ROPP and UCAR, retrievals. The same feature is observed for N-residual between N-values, simulated from retrieved states and RAOB.

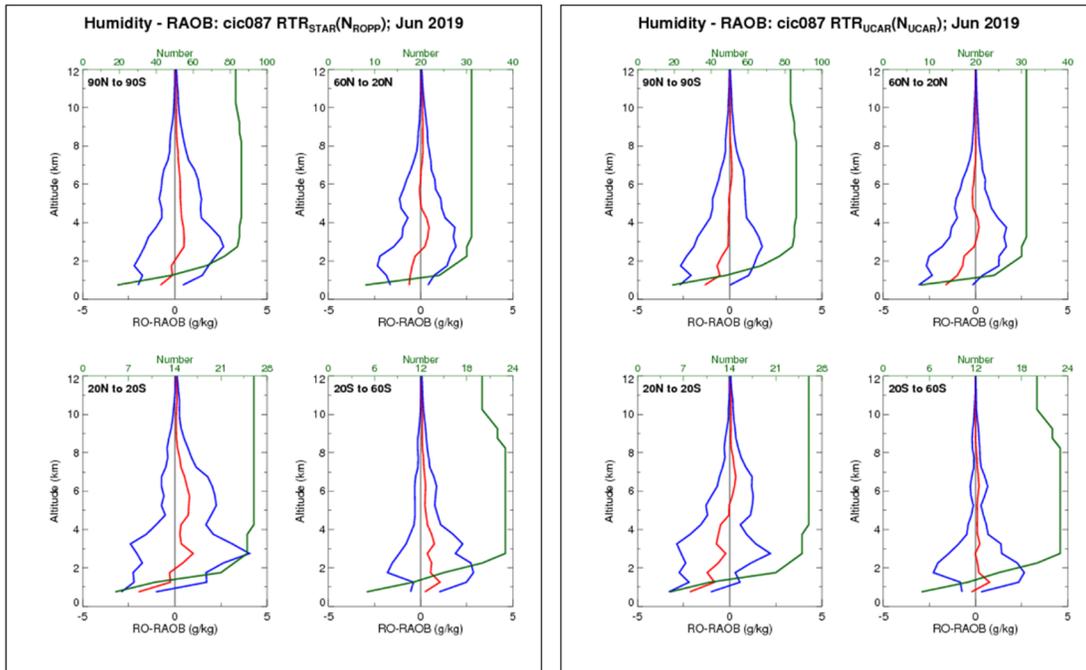


Figure A-27: The same as Figure A-26 but for CIC087-RAOB water vapor comparison

### A.3.2.6 Observation Uncertainty Estimates

Satellite data are critically important for the improvement of numerical weather prediction (NWP) through data assimilation (DA). For effective assimilation of GPS RO data, an accurate error estimate is necessary to optimally weight RO observations in the DA system when other satellite and in situ data are also used. Recently, a proxy for COSMIC Bending Angle Observation Error (BAOE) estimate based on individual observation local spectral width (LSW) has become available in the latest processing of COSMIC data.

As described in section 0, a wave optics approach can be used to retrieve bending angles in the lower troposphere where steep moisture and temperature gradients give rise to multipath effects. The canonical transformation from the space of excess phase and time to bending angle and impact height produces spectrograms like the one shown in Figure A-28a.

The spectrogram depicted is sharp, i.e. the main spectral component is well pronounced above the impact height of about 8.5 km, somewhat pronounced between 8.5 and 4 km and not pronounced below 4 km, indicating that the spread of the spectrum and thus the uncertainty of retrieved bending angle increases with decreasing height. Figure A-28b shows normalized local power spectrum at 3.75 km impact height. The spectrum does not have well pronounced main components and, thus, the bending angle determined at this impact height may have large uncertainty

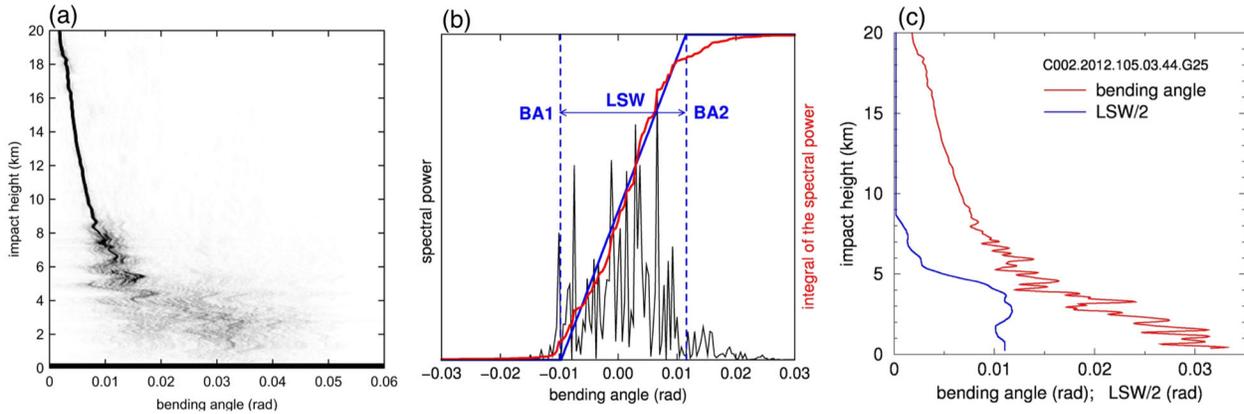


Figure A-28: (a) 2-dimensional sliding spectrogram of RO signal transformed to impact height representation for tropical COSMIC RO, (b) normed local power spectrum at 3.75 km impact height, and (c) retrieved bending angle (red) and the corresponding LSW/2 (blue) profiles.

The LSW captures the uncertainty in the estimate of the bending angle at a given impact height due to the spread of the spectrum. Further details can be found in that paper and references within. The reference bending angle profile obtained from the spectrogram is shown in Figure A-28c along with the profile of LSW/2 for each impact height.

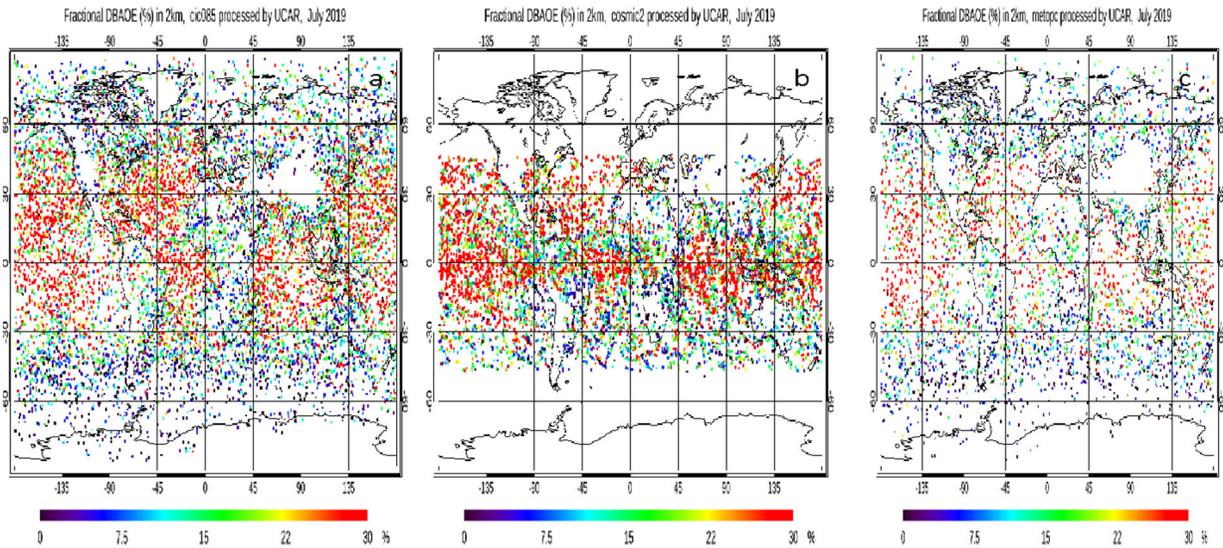
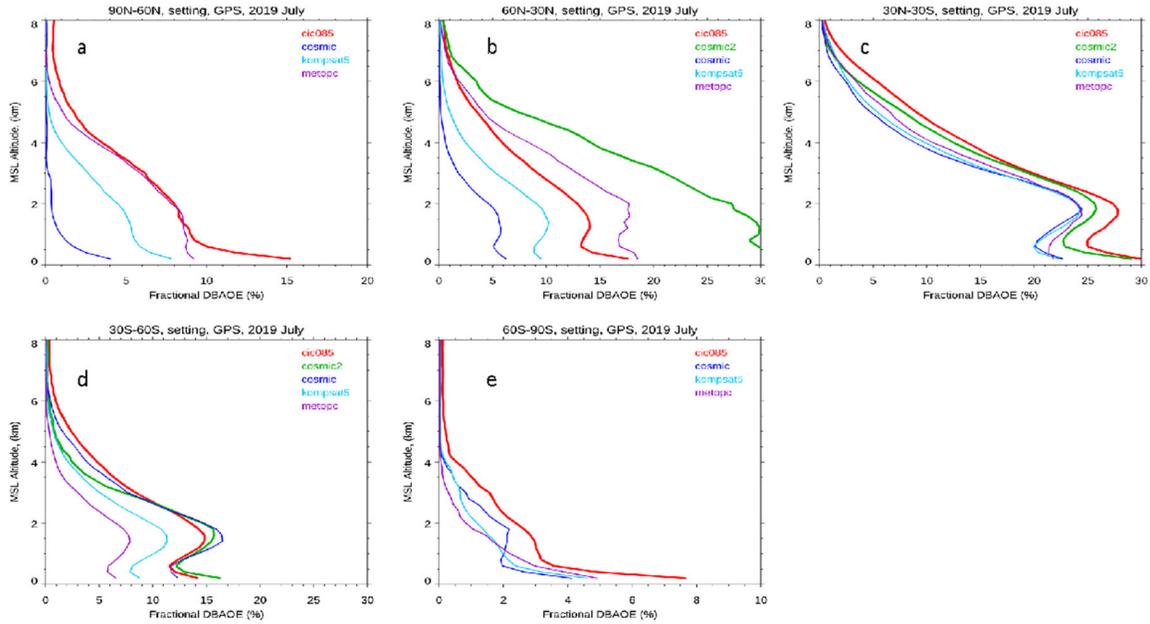


Figure A-29: Distribution of fractional DBAOE (unit: %) at 2 km MSL on July 2019 for (a) GeoOptics, (b) COSMIC-2, and (c) Metop-C.

Figure A-29a-c depicts the distribution of fractional Dynamic Bending Angle Observation Error (DBAOE) (defined as  $100\% \times \text{DBAOE}/\text{bending angle}$  at the same altitude) at 2 km MSL on July 2019 for GeoOptics, COSMIC-2, and Metop-C, respectively. It is obvious that the fractional DBAOE is larger over tropical regions over oceans than that over mid-/high-latitudes over lands. In general, the distributions for fractional DBAOE is highly correlated with the water vapor amount in the atmosphere.

Figure A-30 depicts that the fractional DBAOE for several RO missions including GeoOptics (cic085) at different latitudinal zones. Figure A-30 depicts that the fractional DBAOE for GeoOptics a little bit higher comparing with other RO missions except for the north-hemisphere mid-latitude.

**Mean Fractional DBAOE, 2019 July , setting, GPS**



*Figure A-30: Comparison of the fractional DBAOE for several RO missions over (a) North Hemisphere high-latitude, (b) North Hemisphere mid-latitude, (c) Tropical region, (d) South Hemisphere mid-latitude, and (e) South Hemisphere high-latitude.*

The zonal mean of the fractional bending angle observation error for GeoOptics, COSMIC, and KOMPSAT5 are shown in Figure A-31a-c, respectively.

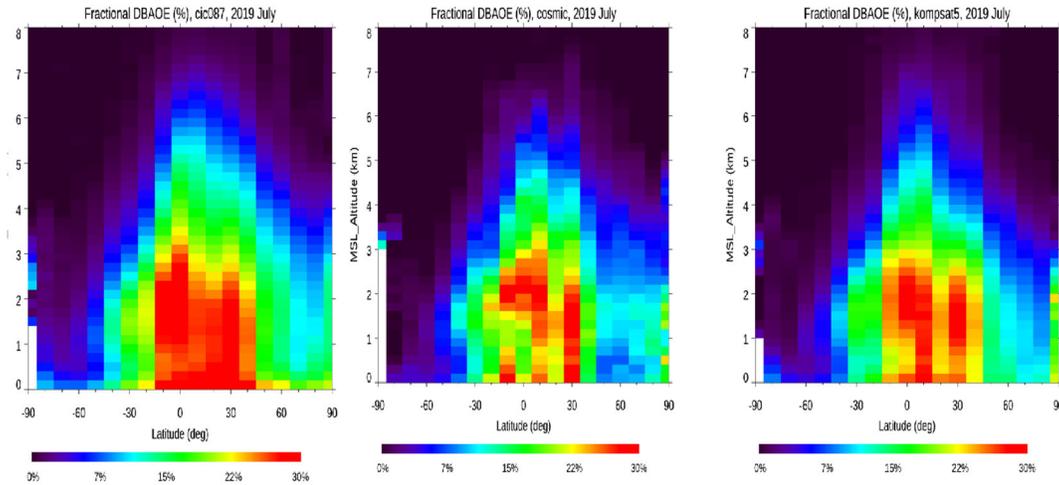


Figure A-31: The zonal mean of the fractional bending angle observation error for (a) GeoOptics, (b) COSMIC, and (c) KOMPSAT5. The contours are the total RO sounding counts (unit: x1000) in an interval of 0.5. RO soundings are sorted in latitude bins of 10o and altitude bins of 200m.

### A.3.2.7 Comparison to MW and IR Sounders

In this section, we compare RO data (i.e., Metop-A, -B, and GeoOptics) with measurements from IR and MW sounders using the Multi-Sensor Validation System described in Section A.2.4.

In this analysis, SNPP ATMS and NOAA-20 CrIS are used as the reference MW and IR sensor, respectively. To eliminate the viewing angle effect, only nadir viewing Field-of-view (FOV) data have been selected. The co-location criteria between RO and ATMS/CrIS are distance difference is within 200 km and time difference is within 3 hours.

Table A-6: Input variables and parameters for GeoOptics-85/87 RO data evaluation with CRTM.

Category	Variable	Data source
Atmosphere	Level and layer pressure temperature	RO (GeoOptics)
	Specific humidity	RO (GeoOptics)
	Ozone mass mixing ratio	ECMWF
Surface	Water type	1 (sea water)
	Skin temperature	ECMWF
	Wind speed	ECMWF
	Wind direction	ECMWF
Geometry	Altitude	Satellite data

	Satellite zenith angle	Satellite data
	Satellite azimuth angle	Satellite data
	Solar zenith angle	Satellite data
	Solar azimuth angle	Satellite data
	Latitude/longitude	Satellite data
Parameters	Climatology	U.S. standard profile
	Water coverage	1 for ocean

For GeoOptics-85/87 (GeoOptics and UCAR Version), a detailed description of input variables and parameters for simulations with CRTM is provided in Table A-6. The three-dimensional variables of temperature and water vapor are provided by GeoOptics-85/87 (GeoOptics and UCAR Version) RO data. The three-dimensional variables of ozone mixing ratio, as well as the two-dimensional variables of surface skin temperature, surface wind speed and surface wind direction, are obtained from ECMWF analyses, which are available at a 6-h interval.

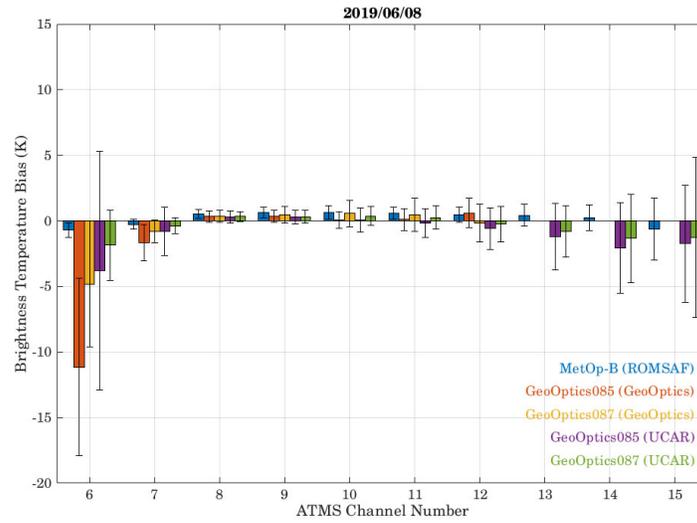
### A.3.2.7.1 RO vs. MW Comparison

Figure A-32a shows the mean and standard deviation (STD) for RO vs. MW comparison for ATMS channel 6-15 using GeoOptics-85/87 (GeoOptics and UCAR Version) RO data and Figure A-32b shows its enlarged results. Note that given the temperature and humidity cut-off at 10 hPa for GeoOptics (GeoOptics version), BT biases with GeoOptics-85 for channel 13, 14 and 15 are not shown in this figure. Table A-7 shows the summary of RO vs. MW data points.

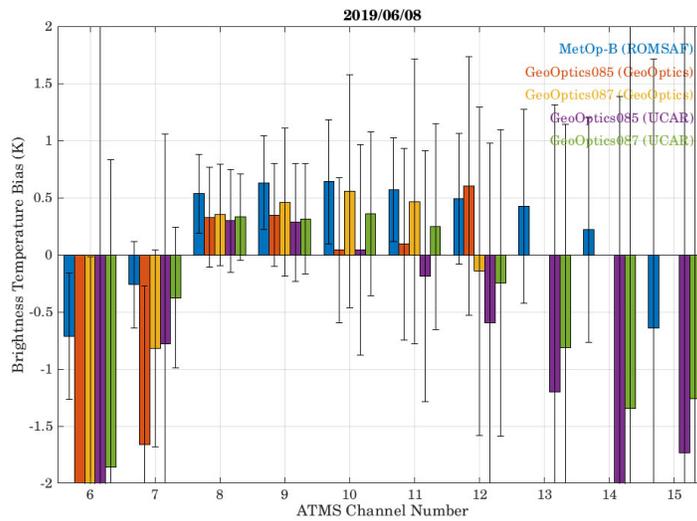
MetOp-B is used as reference and its BT differences for all channels are within  $\pm 0.7$  K and its STDs are within 2 K for all channels. For the BT comparison with GeoOptics (GeoOptics and UCAR versions), the differences with GeoOptics (GeoOptics and UCAR versions) for ATMS channel 8-12 are with  $\pm 0.7$  K (Figure A-32b), which is consistent with the O-B comparison results that the temperature and humidity biases between GeoOptics and ECMWF during 270 to 10 hPa are relatively small (see Section A.3.2.2.2).

The reason for the large bias and STD with GeoOptics-85/87 (GeoOptics version) for channel 6 and 7 is the limited amount of co-located data between 600-1000 hPa (0-15 km height) as shown in Figure A-33.

For UCAR Version, on the other hand, GeoOptics-85/87 for channel 6 also shows 2-4 K bias, which dues to the GeoOptics-85/87 temperature and humidity bias for 700-1000 hPa (Figure A-32). And the reason for the large bias and STD comparing with GeoOptics-85/87 for channel 6 and channels 14/15 are the large bias for UCAR version at 700-1000 hPa and cut-off at  $\sim 2$  hPa for GeoOptics (UCAR version) (Figure A-32). It also can be seen that the GeoOptics-87 has smaller BT difference than GeoOptics-85 for channels 6, 7, and 12~15 is because of the smaller temperature difference for 600-1000 hPa and 10-90 hPa (Table A-3). On the contrary, GeoOptics-85 has smaller BT difference for channels 8-11 compared with GeoOptics-87 is because of the smaller temperature difference of GeoOptics-85 for 90-600 hPa (Table A-3).



(a)



(b)

Figure A-32: (a) Mean and STD of BT bias for RO (GeoOptics-85/87) vs. MW (NPP ATMS) comparison for ATMS channels 6-15 and (b) its enlarged BT bias. The GeoOptics-85/87 data is based on GeoOptics and UCAR Version L2 data.

Table A-7: Summary of RO vs. MW data points. Derived from June 8, 2019 data for all GeoOptics RO data. ERA5 data from ECMWF are used as background.

RO Instrument	L2 Data Processor	# of Valid Profiles	Note
---------------	-------------------	---------------------	------

MetOp-B	ROMSAF	15	As reference; Wet T.
GeoOptics-085	Geo-Optics	10	Wet Temperature
GeoOptics-087	Geo-Optics	17	Wet Temperature
GeoOptics-085	UCAR	18	Wet Temperature
GeoOptics-087	UCAR	30	Wet Temperature

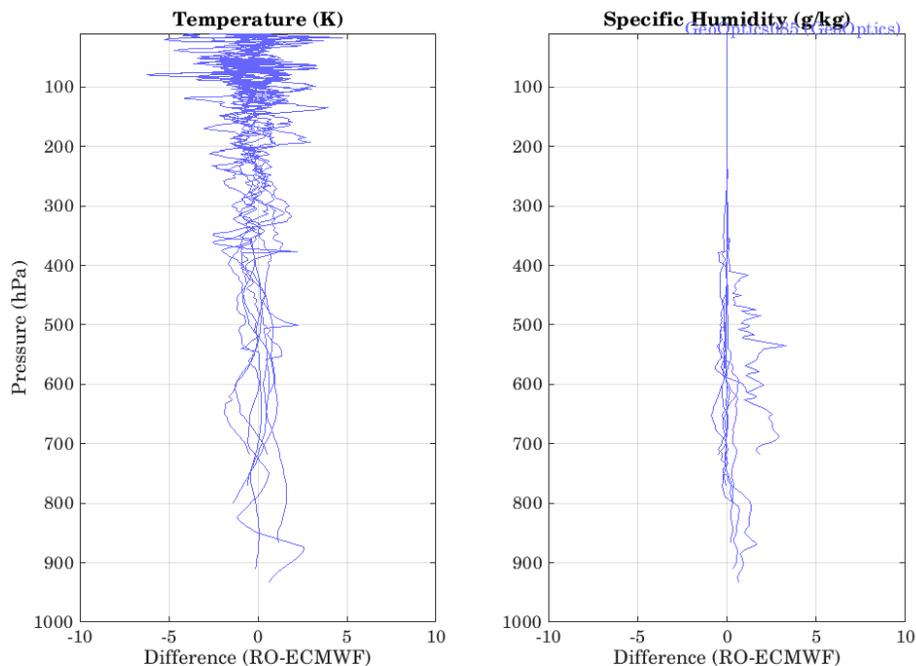


Figure A-33: GeoOptics085 (GeoOptics version) temperature and humidity profiles colocation with ATMS.

### A.3.2.7.2 RO vs. IR Comparison

Figure A-34 shows mean and STD for RO vs. IR comparison using GeoOptics-85/87 (GeoOptics and UCAR Version) RO data derived from June 8, 2019 to June 13, 2019. ERA5 data from ECMWF are used as background and MetOp-B is used as reference. Table A-8 shows the summary of RO vs. IR data points. Note that to screen out the cloud effect, the BT difference profiles whose maximum BT difference larger than 4 K for ground channels (860 to 880  $\text{cm}^{-1}$ ) have been removed.

It can be seen that generally the BT differences for MetOp-B are with  $\pm 0.5$  K except between for channels with wavenumbers between 1000 to 1080  $\text{cm}^{-1}$ , which are sensitive to  $\text{O}_3$ , and most likely the accuracy level of the  $\text{O}_3$  profile input is not high enough. The relatively large bias seen for wavenumbers between 670 and 750  $\text{cm}^{-1}$  is due to the use of the default  $\text{CO}_2$  profile.

For GeoOptics-85/87(GeoOptics and UCAR version), the BT differences for surface sensitive bands (750 to 1000 wavenumber) are stable and but show relatively large bias ( $\sim 1$  K in general) due to the data loss of GeoOptics-85/87 (GeoOptics version) at 600-1000 hPa (0-15 km height) (Figure A-34) and large temperature and humidity bias of GeoOptics-85/87 (UCAR version) at 600-1000 hPa (0-15 km height).

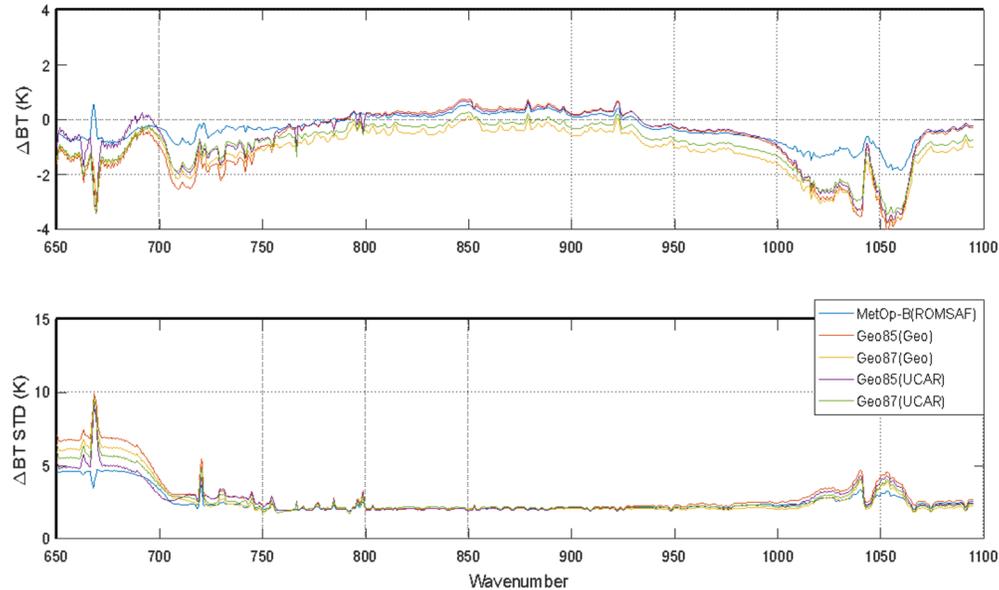


Figure A-34: Mean and STD for RO vs. IR comparison: GeoOptics-85/87 (GeoOptics and UCAR Version L2 Data)

For GeoOptics-85/87(GeoOptics and UCAR version), large bias ( $\sim 3.5$  K) and large STD at the  $670\text{ cm}^{-1}$  wavenumber given the missing data at 0~10 hPa for GeoOptics-85/87 (GeoOptics version) and missing data at 0~3 hPa for GeoOptics-85/87 (UCAR version). And similar to MetOp-B, the BT differences for channels with wavenubmers between 1000 and  $1080\text{ cm}^{-1}$  are as large as  $\sim 4$  K, probably due to inaccuracies in the  $\text{O}_3$  input profile and the large temperature difference for ground pressure levels. The relatively large bias for the 670 to 750 wavenumber region is also due to the lack of accurate  $\text{CO}_2$  profile.

Table A-8: Summary of RO vs. IR data points. Derived from June 8, 2019 to June 13, 2019 data for all GeoOptics RO data. ERA5 data from ECMWF are used as background.

RO Instrument	L2 Data Processor	# of Valid Profiles	Note
MetOp-B	ROMSAF	121	As reference; Wet T.
GeoOptics-085	Geo-Optics	80	Wet Temperature
GeoOptics-087	Geo-Optics	57	Wet Temperature
GeoOptics-085	UCAR	92	Wet Temperature
GeoOptics-087	UCAR	88	Wet Temperature

## A.4 Spire: Processing, Analysis, and Validation

Spire Global, Inc. (Spire) provided L1a and L2 data for the months of June and July 2019. Data collected by some subset of twelve satellites was provided each day to total between 500 and 600 L2 profiles during the month of June along with the corresponding lower level data. In July, the delivery was increased to 1200 profiles per day. All of the Spire processed profiles were flagged as nominal. UCAR/CDAAC also provided L2 profiles retrieved from the L1a data provided by Spire. Details of the data delivery can be found in Appendix A. An assessment of the quality of the Spire data is provided in the following section.

### A.4.1 Assessment of Spire L1a Data Quality

Spire provides L1a time series of satellite positions, phase and SNR for Galileo and QZSS in addition to GPS and GLONASS. The L1a data provided by SPIRE includes excess phase, L1 and L2 SNR, LEO positions and velocities in ECI coordinates, and GNSS positions and velocities in ECI coordinates at 50 Hz. The excess phase data contains navigation bits that need to be removed in preprocessing.

#### A.4.1.1 Stability of Satellite Position and Velocities

Analysis of the stability of the position and velocity vectors are performed for Spire L1a data using the same approach described in Section 0. The incremental velocity vectors are calculated for individual profiles, and then one value of the mean absolute deviation for the whole profile is calculated. The mean absolute deviations for all the profiles for each day are then averaged to calculate the daily mean absolute deviation.

Figure A-35 shows the incremental velocity vectors for SPIRE satellite 084 on July 1, 2019. The  $u$ - and  $v$ - components of the LEO velocities show three separate bands of values. It means that the  $u$ -component of the velocity at around 20 seconds can change by 300, 130 or -40 mm/sec. The variation in the  $w$ -component is much smaller for both the LEO and the GNSS satellites.

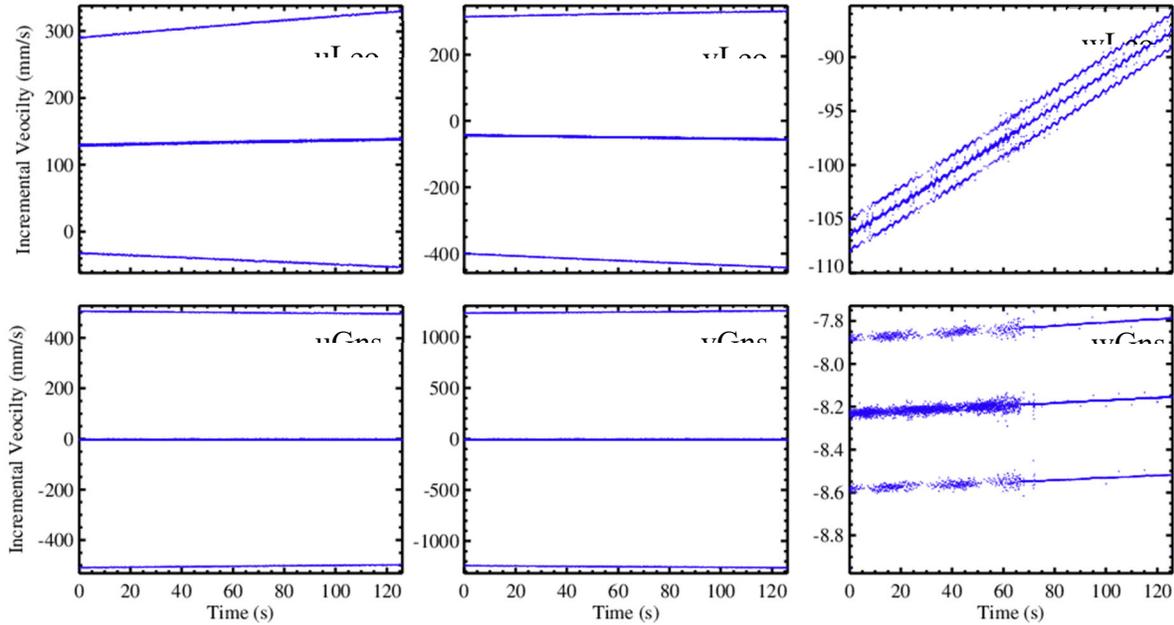


Figure A-35: Incremental velocity vectors for (top) LEO and (bottom) GNSS for SPIRE satellite 090 and GNSS satellite G05 on 2019-07-01-00:14Z.

The large incremental velocity variations for the  $u$ - and  $v$ -components are illustrated in the daily mean absolute deviation values shown in Figure A-36. The magnitude of the absolute deviation for the GNSS and LEO velocities are  $\sim 200$  and  $50$  mm/sec respectively for  $u$ - and  $v$ -components. The  $w$ -components are  $\sim 0.2$  and  $3$  mm/s for GNSS and LEO, respectively.

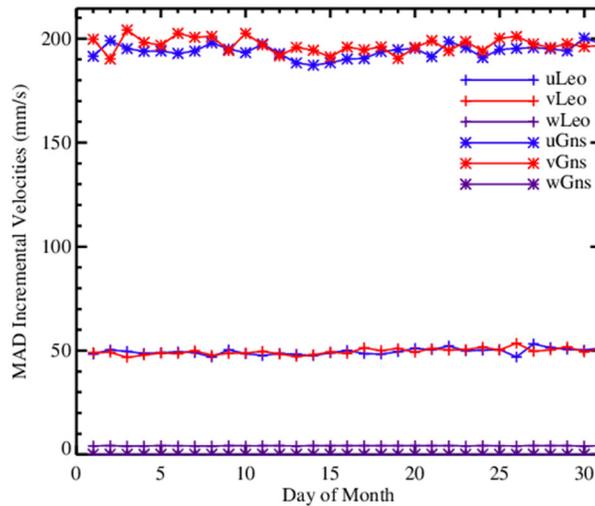


Figure A-36: Mean absolute deviation (MAD) of incremental velocity vectors for July 2019 for SPIRE satellites

#### A.4.1.2 Phase Measurement Noise

The excess phase data provided in the SPIRE L1a data contains navigation bits that causes the calculated Doppler to appear noisy even at high altitudes. Figure A-37 shows a typical example of a setting occultation. The top panel shows the L1 and L2 excess phases, the middle panel shows the excess Doppler, and the lowest panel shows the SNR. In a typical occultation data, the Doppler becomes noisier as altitude decreases, but is less noisy at higher altitudes where the SNR is large. However, in the SPIRE data, the noise in the excess Doppler remains large throughout the measurement period.

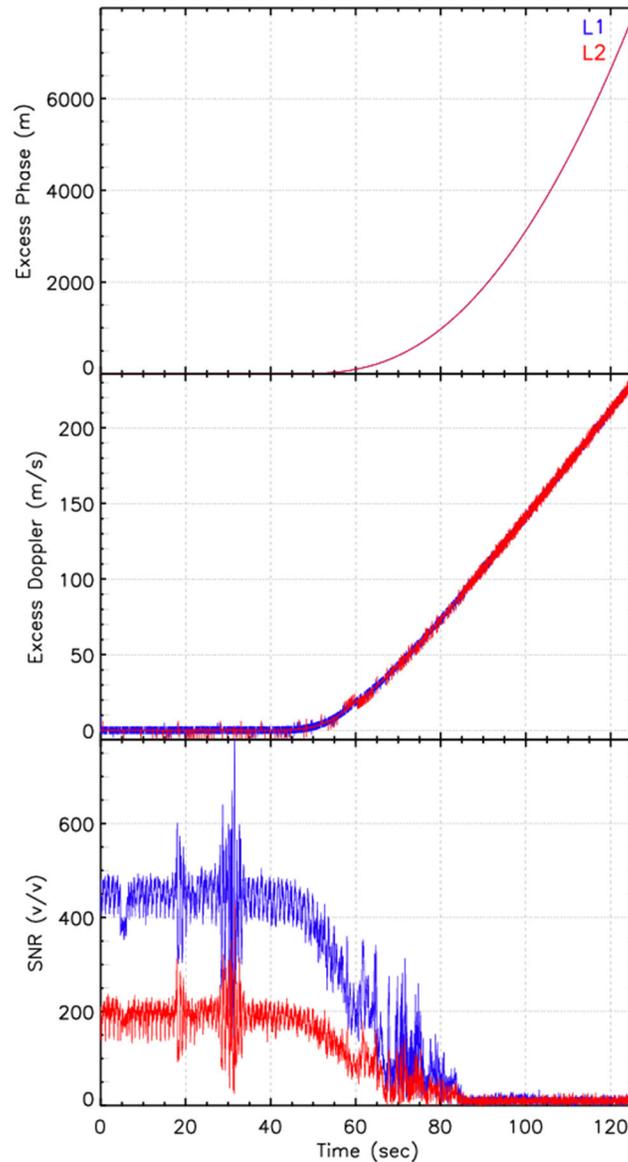


Figure A-37: Excess phase (top), Doppler (middle) and SNR for GPS satellite G05 at 00:14Z on 2019-07-01.

The standard deviation of the Doppler frequency for SPIRE data is shown in Figure A-38. For the SPIRE phase data, the standard deviation in the L1 Doppler is greater than for L2 Doppler.

The Standard deviation decreases with altitude, but it is still in the  $\sim 8$  Hz at 50 km impact height for L1.

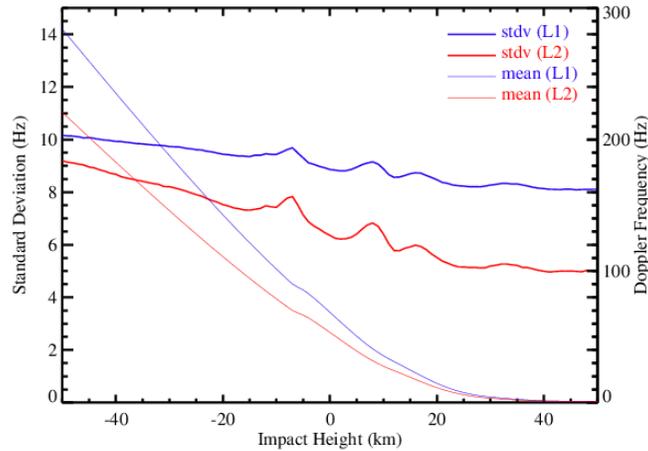


Figure A-38: Standard deviation of the SPIRE Doppler frequency for July 2019.

#### A.4.1.3 Signal to Noise Ratio (SNR) Assessment

The SNR for SPIRE L1a data varies for each occultation. However, on average, the mean SNR above the noise level for L1 and L2 bands are  $\sim 425$  and  $180$  v/v, respectively (Figure A-39). The bottom panel in Figure A-39 shows the frequency distribution of L1 and L2 band SNRs. It shows that the SNR in L1 band goes up to  $1500$  v/v with maxima near  $400$  v/v.

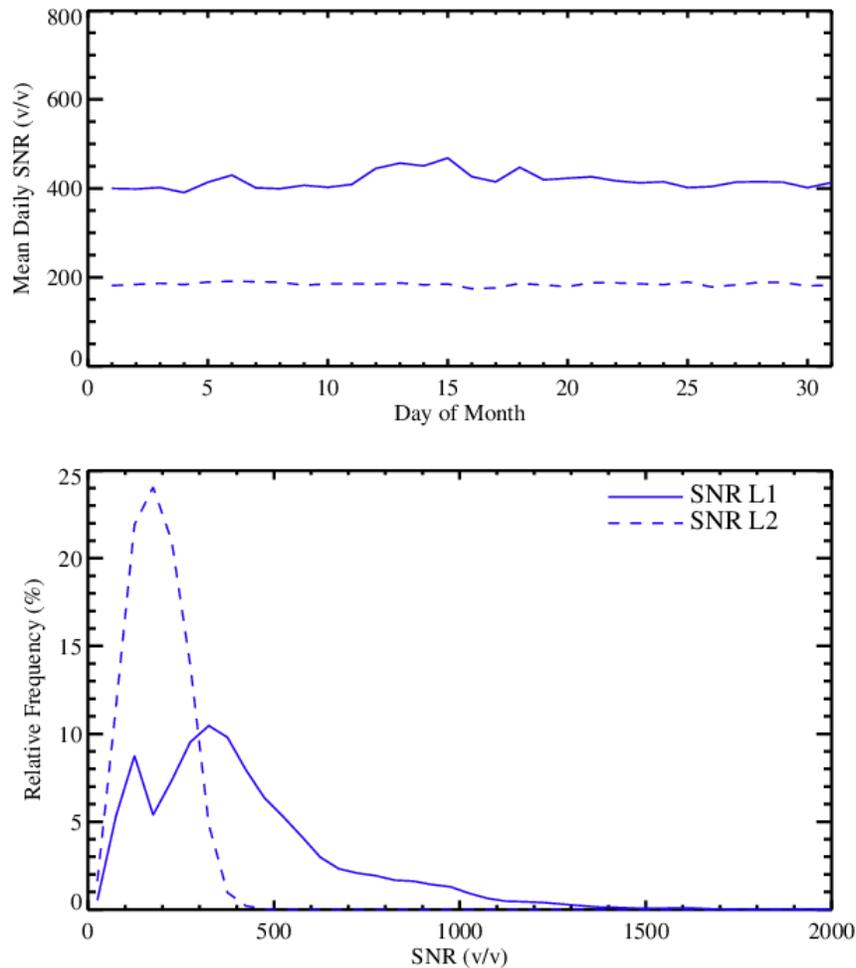


Figure A-39: Mean daily L1 and L2 SNR for all SPIRE receivers (top) and SNR frequency distribution for July 2019

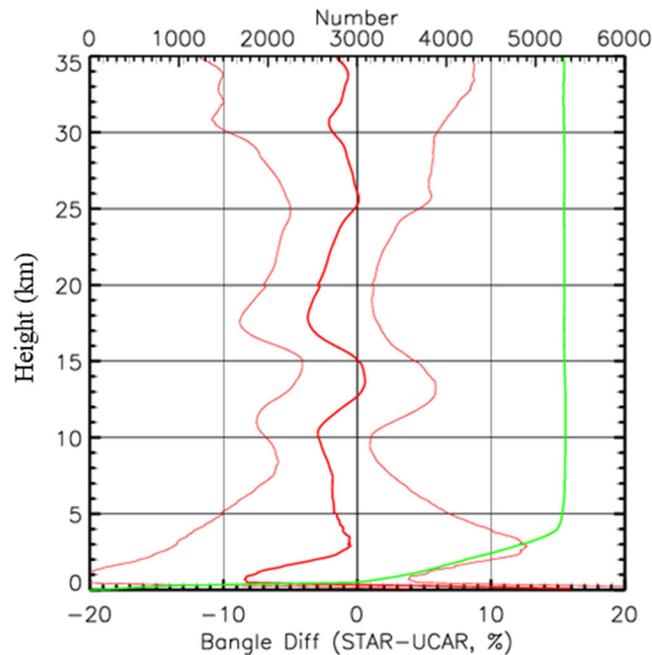
## A.4.2 Assessment of Spire L2 Data Quality

### A.4.2.1 Bending Angle and Refractivity Comparisons

#### A.4.2.1.1 Comparison between SPIRE Bending Angle and STAR FSI Bending Angle

SPIRE L1b data contains navigation bits that needs to be removed in preprocessing before bending angle can be retrieved. A navigation bit removal program was developed at STAR that used climatological refractivity profile to down-convert the signal. The navigation bit removal program uses the following components (1) use occultation information to determine (collocate) climatological refractivity profile, (2) use forward model to derive bending angle profiles as a function of impact parameter from height – refractivity model input, (3) Calculate excess (atmospheric) Doppler shift from impact parameter and bending angle profile, (4) interpolate the Doppler shift to the occultation times, and (5) Integrate the Doppler shift to calculate excess phase.

The model excess phase is then used to down-convert the measured signal with navigation bits. If the model excess phase is good, the down-converted signal frequency will be zero. Navigation bits appear as deviation of half cycle ( $\pi$ ) from the mean zero, that can be removed. The current version of the model phase calculation is not accurate enough and results in a negative bending angle bias in the STAR bending angle compared to UCAR bending angle (Figure A-40). Again, owing to the insufficient clock data provided by SPIRE, computed access phase contained larger uncertainty which results in a larger FSI bending angle biases comparing to those from FSI GeoOptics bending angle results.



*Figure A-40: Fractional bending angle difference between STAR and UCAR for July 2019 SPIRE L1b data.*

#### **A.4.2.1.2 Comparison between SPIRE Refractivity and STAR FSI Retrievals**

Figure A-41 shows fractional refractivity difference of STAR bending angle with UCAR and collocated ERA5 refractivity. The bending angle refractivity bias in STAR retrievals is observed in the STAR retrieved refractivity.

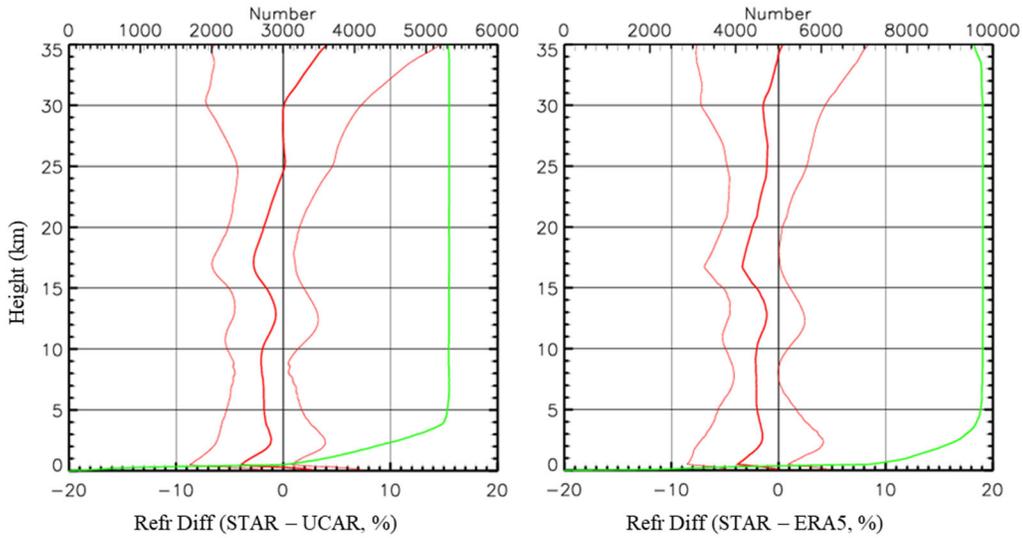


Figure A-41: Fractional refractivity difference between (left) STAR and UCAR and (right) STAR and ERA5, for July 2019 SPIRE L1b data

#### A.4.2.2 Comparison to NWP Model Fields

##### A.4.2.2.1 Comparison to GFS 6-hour Forecast

Figure A-42 depicts the (a) fractional refractivity difference and (b) water vapor mixing ratio difference between Spire and GFS 6-hour forecast. The large SPIRE N (derived from UCAR) biases are owing to the insufficient clock data provided by SPIRE.

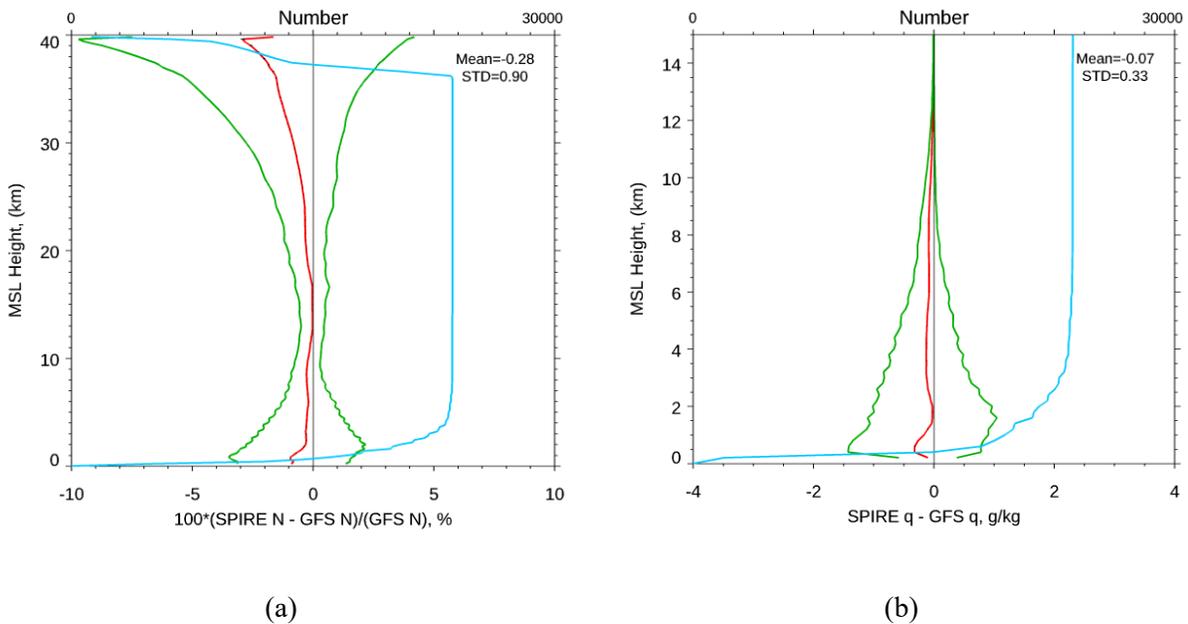


Figure A-42: The (a) fractional refractivity difference and (b) water vapor mixing ratio difference between Spire and GFS 6-hour forecast.

### A.4.2.2.2 Comparison to ECMWF ERA Interim Reanalysis

#### O-B Analysis of SPIRE-90/99 (SPIRE Version) L2 Data

There are two main issues with the SPIRE L2 data processed by SPIRE. The first issue is that only dry temperature data and no humidity data are available in the BUFR files processed by SPIRE. Since the temperature data from ECMWF is wet temperature, the O-B comparison can only provide data quality screening and won't be able to quantitatively evaluate the relative bias of the L2 temperature data processed by SPIRE.

Another issue is that the pressure coordinate data in the L2 BUFR file from SPIRE seems to have staircase structure, *i.e.* several layers have the same pressure level (Figure A-43a). This causes the staircase structure in the O-B temperature bias at high altitude (low pressure) region (Figure A-43b).

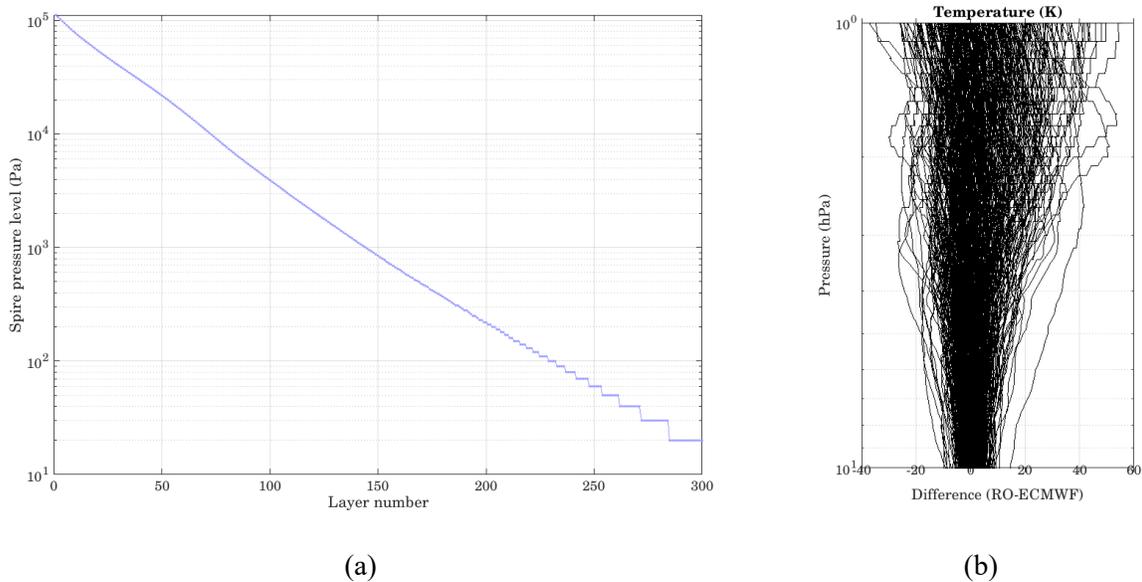


Figure A-43: Illustration of the issue with the pressure coordinate data in SPIRE L2 data (BUFR format) processed by SPIRE. (a) Staircase structure of pressure coordinate in SPIRE L2 data; (b) Consequent large staircase structure in the O-B temperature bias at high altitude region.

Figure A-44 shows O-B (ECMWF) temperature bias of SPIRE-90 and -99 RO data. It can be seen from Figure A-44a that both SPIRE -90 and SPIRE -99 dry temperature data show similar behavior. Large dry temperature bias can be seen near surface (600-1000 hPa) which is mainly due to the difference between dry and wet temperature. At low pressure region (0-10 hPa), *i.e.* high altitude, the large bias should not be impacted by the lack of humidity considered in retrieval and may be related to the inaccurate pressure level coordinates in the dry temperature retrieval.

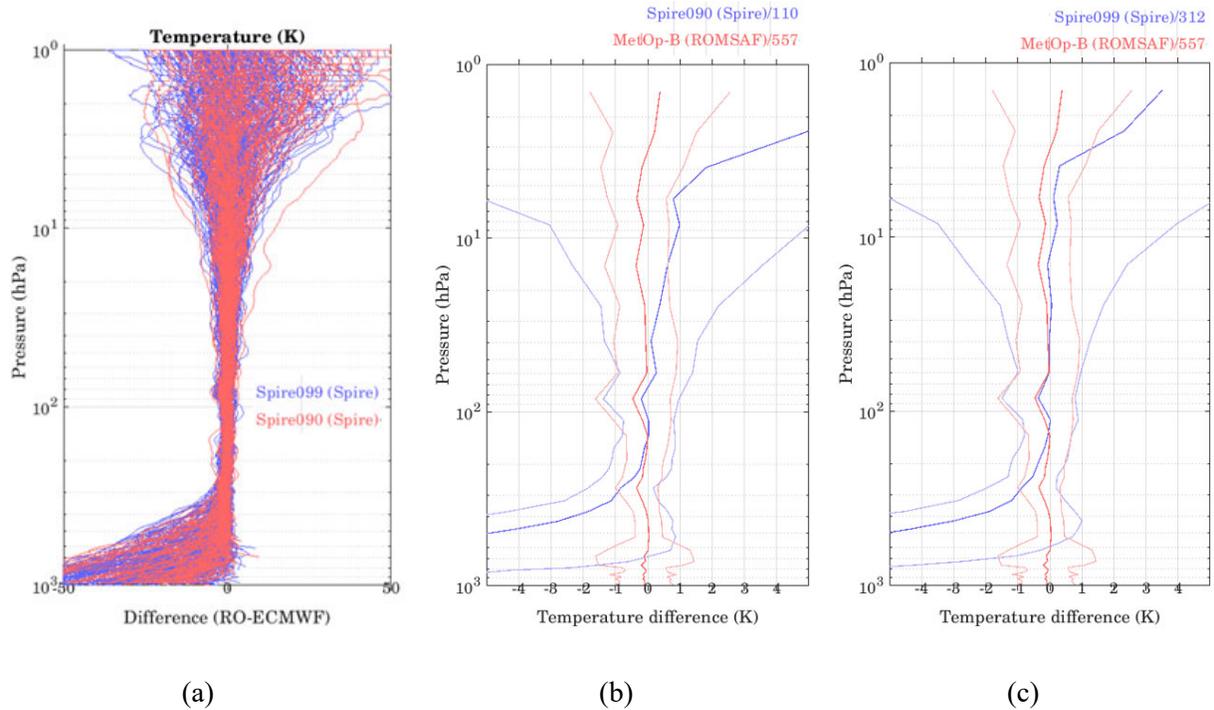


Figure A-44: O-B bias analysis for SPIRE -90/99 L2 RO data (SPIRE Version): (a) O-B (ECMWF) temperature difference of SPIRE -90/99 (SPIRE Version) L2 RO data; (b) Mean and standard deviation of O-B temperature and humidity difference of SPIRE-90 L2 RO data and (c), mean and standard deviation of O-b bias for MetOp-B RO data are also plotted. Note that SPIRE-90/99 L2 RO data (SPIRE version) only have dry temperature data.

### O-B Analysis of SPIRE -90/99 (UCAR Version) L2 Data

The SPIRE -90/99 L2 RO data processed by UCAR are compared with ECMWF background data in Figure A-45. Figure A-45a, b and c show the O-B (ECMWF) temperature and humidity difference of SPIRE -90/99 (UCAR Version), mean and standard deviation of O-B temperature and humidity difference of SPIRE -90 and of SPIRE -99, respectively. Comparing mean O-B temperature bias of SPIRE-99 with that of SPIRE -90 from UCAR, it can be immediately seen that there is a large negative temperature bias in the region with pressure < 100 hPa in SPIRE-99 data. The SPIRE -99 and SPIRE-90 temperature data from UCAR are inconsistent. Given the consistency in the dry temperature comparison between SPIRE -90 and SPIRE -99 of SPIRE version (Figure A-45) and the growing deviation of O-B bias in SPIRE-99 UCAR version data, the inconsistency in UCAR version could be due to an anomaly in the processing of SPIRE-99 RO data into L2 temperature data by UCAR. Such anomaly in SPIRE -99 UCAR data also prevents further quantitative evaluation of the overall quality of the SPIRE-99 temperature data processed by UCAR.

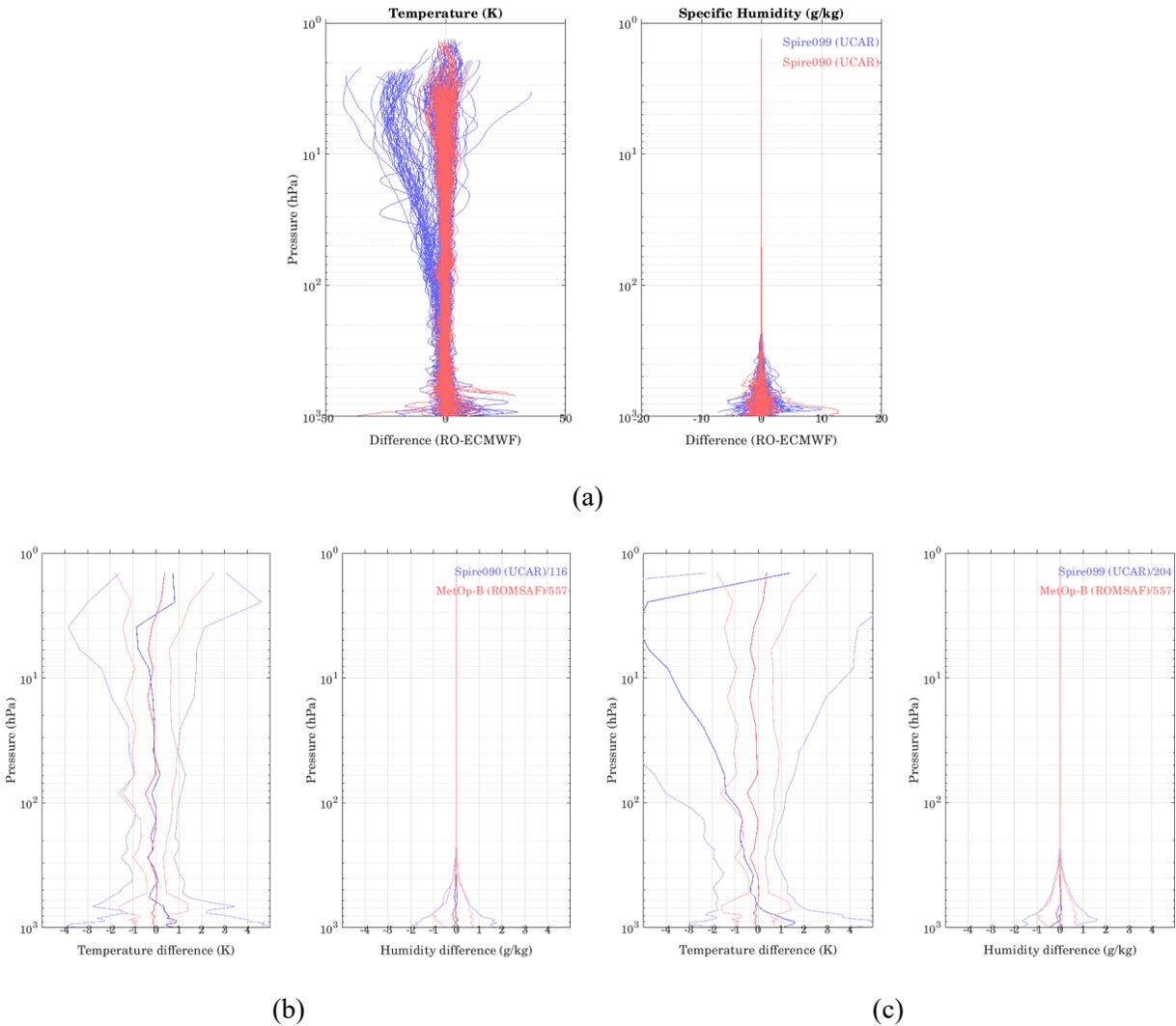


Figure A-45: O-B bias analysis for SPIRE -90/99 L2 RO data (UCAR Version): (a) O-B (ECMWF) temperature and humidity difference of SPIRE -90/99 (UCAR Version) L2 RO data; (b) Mean and standard deviation of O-B temperature and humidity difference of SPIRE-90 L2 RO data. In (b) and (c), mean and standard deviation of O-B bias for MetOp-B RO data are also plotted.

In Figure A-45b, the mean O-B temperature and humidity biases of SPIRE-90 are consistent with MetOp-B over the pressure range between 10 and 700 hPa. For pressure levels greater than 700 hPa (especially near surface pressure level) and less than ~10 hPa, there are significant differences in O-B temperature bias between SPIRE-90 and MetOp-B. The standard deviations of O-B temperature bias of SPIRE-90 are larger than MetOp-B for pressures greater than 200 hPa and pressures less than 20 hPa. The largest O-B temperature bias reaches ~1 K for SPIRE-90 L2 data processed by UCAR.

### Summary of O-B (ECMWF) Analysis for SPIRE RO Data

To quantitatively evaluate the SPIRE RO data quality in different atmospheric pressure regions in terms of O-B bias and its standard deviation, Table A-9 summarizes the O-B (ECMWF)

temperature bias and standard deviation (K) at different atmospheric pressure level regions for SPIRE-90 and -99 RO L2 data processed by SPIRE and UCAR. Table A-10 contains the summary of O-B (ECMWF) humidity bias and standard deviation (g/kg) at different atmospheric pressure regions for SPIRE-90 and -99 RO data processed by SPIRE and UCAR.

*Table A-9: Summary of O-B (ECMWF) temperature bias and its standard deviation (K) at different atmospheric pressure levels for SPIRE-90 and -99 RO L2 data processed by SPIRE and UCAR. Derived from June 8, 2019 data for all RO data. ERA5 data from ECMWF are used as background.*

RO Instrument	L2 Data Processor	Pressure Levels (hPa)				# of Valid Profiles	Note
		600.0--1000.0	90.0-- 600.0	10.0-- 90.0	0.1-- 10.0		
MetOp-A	ROMSAF	-0.093 ± 1.025	-0.108 ± 0.621	-0.203 ± 0.988	-0.040 ± 1.382	572	As reference; Wet T.
MetOp-B	ROMSAF	-0.098 ± 1.054	-0.097 ± 0.627	-0.204 ± 1.001	-0.009 ± 1.403	557	As reference; Wet T.
Spire-090	Spire	-16.752 ± 15.316	-1.433 ± 3.610	0.302 ± 2.082	4.050 ± 11.336	110	Dry Temperature; Pressure data anomaly
Spire-099	Spire	-17.675 ± 14.494	-1.541 ± 3.685	-0.067 ± 1.784	1.675 ± 9.745	312	Dry Temperature; Pressure data anomaly
Spire-090	UCAR	0.435 ± 2.938	-0.114 ± 1.146	-0.081 ± 1.398	-0.495 ± 2.782	116	Wet Temperature
Spire-099	UCAR	0.603 ± 2.873	-0.556 ± 1.626	-2.329 ± 4.751	-4.538 ± 9.159	204	Wet Temperature

Since the SPIRE version L2 data contain only dry temperature data, it prevents an equitable comparison with the background wet temperature from ECMWF. However, the consistency between SPIRE version SPIRE-90 and -99 dry temperature data can still be seen from the O-B bias summary in Table A-9 and Table A-10. The mean O-B temperature bias of SPIRE-90 and -99 (SPIRE version) is close to MetOp-B (reference) only in the pressure region of 10-90 hPa. The large dry temperature bias in pressure 90-600 hPa and 600-1000 hPa is mainly due to the difference between dry and wet temperature. The large bias in low pressure region (0.1-10 hPa) may be due to the inaccurate pressure level coordinates in the dry temperature retrieval by SPIRE. The issue with pressure data in SPIRE version data has direct impact on the O-B analysis.

For UCAR L2 SPIRE data, there is apparent inconsistency between mean O-B temperature bias of SPIRE-90 and SPIRE-99. The growing deviation of O-B bias in SPIRE-99 UCAR version data can be due to an anomaly in the processing SPIRE-99 RO data into L2 temperature data by UCAR. For UCAR L2 data of SPIRE-90, the overall mean O-B temperature and humidity bias is close to MetOp-B (reference) except larger O-B bias in pressure levels 800-1000 hPa in comparison with MetOp-B. The overall standard deviation of O-B bias of UCAR L2 temperature data for SPIRE-90 are larger than MetOp-B.

Since SPIRE L2 data processed by SPIRE only contains dry temperature data, it prevents an equitable comparison with the RO data processed by UCAR. It is also worth pointing out that UCAR L2 SPIRE-90 data has similar number of profiles as those processed by SPIRE. On the other hand, UCAR L2 SPIRE-99 data has fewer profiles than those processed by SPIRE.

*Table A-10: Summary of O-B (ECMWF) humidity bias and its standard deviation (g/kg) at different atmospheric pressure levels for SPIRE-90 and -99 RO L2 data processed by SPIRE and*

UCAR. Derived from June 8, 2019 data for all RO data. ERA5 data from ECMWF are used as background.

RO Instrument	L2 Data Processor	Pressure Levels (hPa)			# of Valid Profiles	Note
		800.0--1000.0	500.0-- 800.0	200.0-- 500.0		
MetOp-A	ROMSAF	$-0.0492 \pm 0.6815$	$-0.0659 \pm 0.6380$	$-0.0067 \pm 0.1187$	572	As reference; Wet T.
MetOp-B	ROMSAF	$0.0007 \pm 0.6993$	$-0.0353 \pm 0.6055$	$-0.0021 \pm 0.1324$	557	As reference; Wet T.
Spire-090	Spire	N/A	N/A	N/A	110	Dry Temperature; Pressure data anomaly
Spire-099	Spire	N/A	N/A	N/A	312	Dry Temperature; Pressure data anomaly
Spire-090	UCAR	$-0.0359 \pm 1.5674$	$-0.0502 \pm 0.6710$	$-0.0157 \pm 0.1236$	116	Wet Temperature
Spire-099	UCAR	$-0.1291 \pm 1.3747$	$0.0440 \pm 0.7165$	$-0.0049 \pm 0.1688$	204	Wet Temperature

#### A.4.2.3 Data Penetration

We investigate the lowest penetration height of the SPIRE data compared to other RO missions. (i.e., Metop-A, -B, -C, COSMIC, COSMIC-2 and KOMPSAT5, also see Table A-11). The lowest penetration height of RO tracking is usually related to the signal to noise ratio (SNR) of the data. In general, observations with higher SNR usually penetrate deeper than those observations with lower SNR. Figure A-46 shows the histograms of SNR distribution of SPIRE, COSMIC, and KOMPSAT-5.

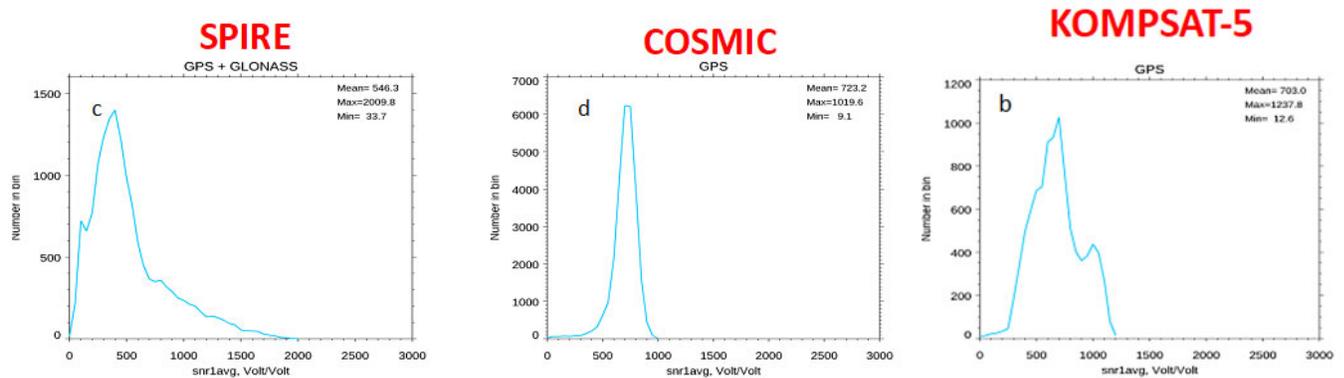


Figure A-46: The histogram of SNR distribution of (a) SPIRE, (b) COSMIC, and (c) KOMPSAT-5.

Table A-11 shows the lowest penetration height of 80% of the total data for different RO missions at different latitudinal zones using UCAR processed L2 data. It shows that the penetration depth of SPIRE data is comparable to COSMIC2 data at all latitudes where COSMIC2 data are available.

Table A-11: The lowest penetration height of 80% of the total data for different RO missions at different latitudinal zones. All the data are from July 2019.

	10N-10S	10N-30N	10S-30S	30N-45N	30S-45S	45N-60N	45S-60S	60N-90N	60S-90S
Metopa	3.2	7.2	4.0	4.9	2.1	3.2	1.2	3.0	3.6
Metopb	2.6	4.5	3.7	4.0	2.0	2.6	1.3	2.6	3.6
Metopc	2.8	4.7	4.0	4.9	1.8	3.2	1.4	3.0	3.5
Cosmic	1.5	1.2	1.4	1.9	0.6	1.9	0.4	1.5	2.2
Cosmic2	1.2	0.5	1.1	1.8	0.6				
spire	1.1	2.2	1.1	2.1	0.5	1.3	0.3	1.2	2.7
kompsat5	1.7	1.4	1.4	1.8	0.6	1.2	0.3	1.1	2.8

#### A.4.2.4 Monitoring of Long-term Stability

Figure A-47a and b depict the fractional bending angle and fractional refractivity differences between Spire and MetOp-B GRAS, respectively. The GRAS and Spire pairs are within 300 km and 3 hours. Result shows that the SPIRE N biased negatively relative to that of GARS. UCAR SPIRE retrieved N profiles contain negative biases above 15 km altitude.

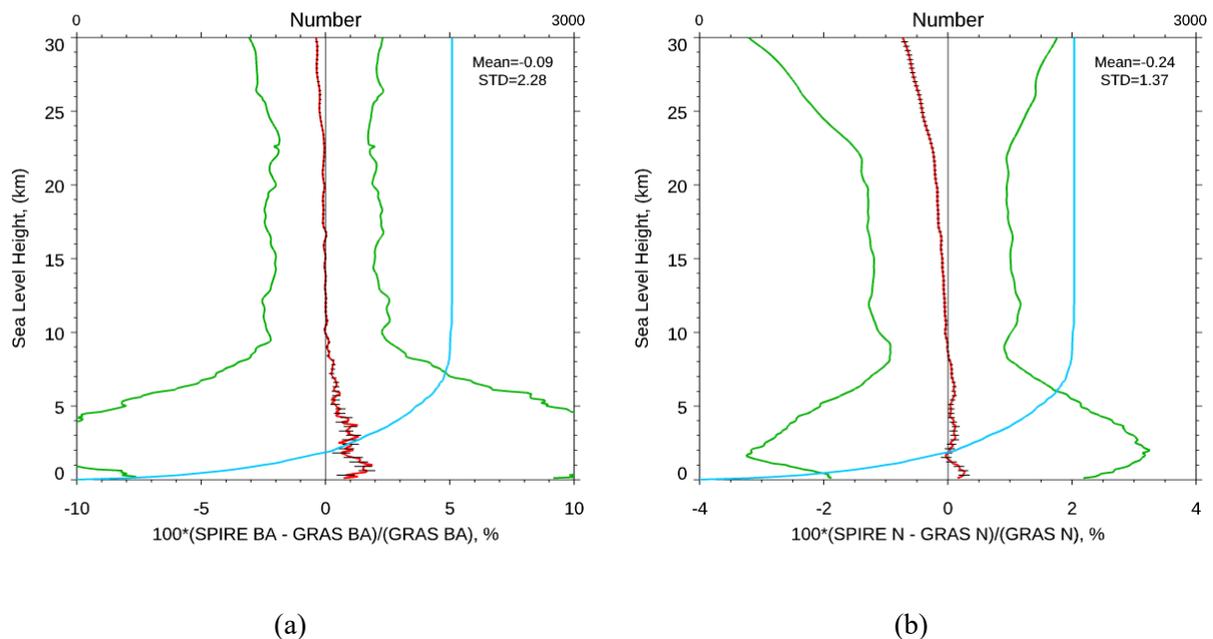
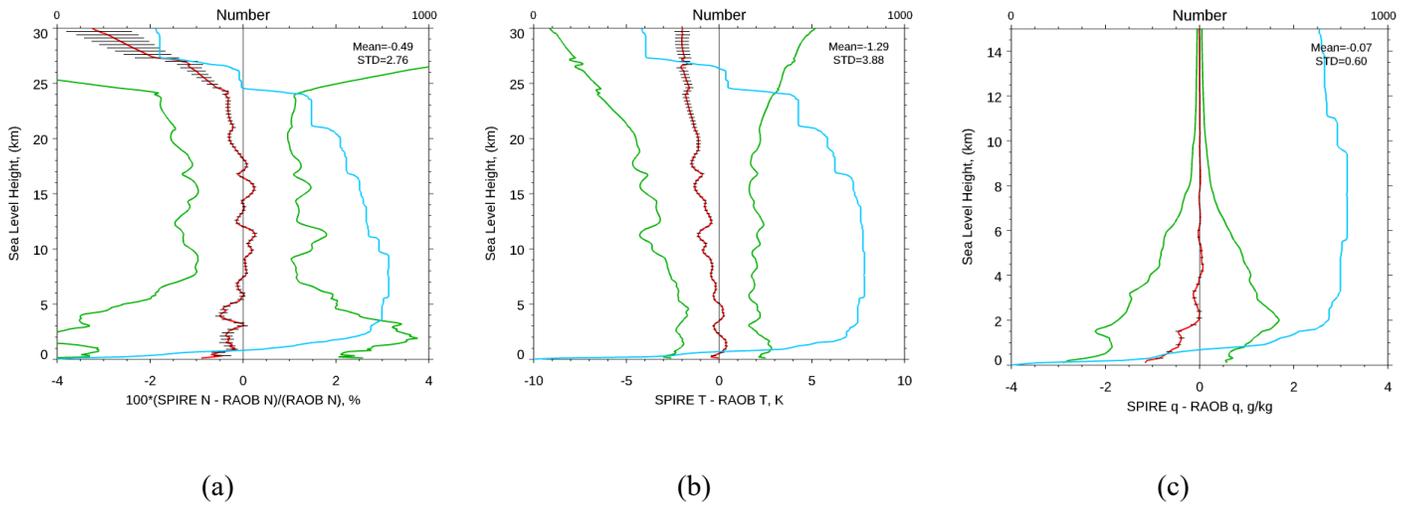


Figure A-47: The (a) fractional bending angle and (b) fractional refractivity differences between Spire and MetOp-B GRAS where the GRAS and Spire pairs are within 300 km and 3 hours.

#### A.4.2.5 Uncertainty of RO Water Vapor Retrievals

One month of SPIRE and RAOB pairs are collected from July 1 to 31, 2019. The comparison results for fractional refractivity difference, the temperature difference, and water vapor mixing

ratio difference are shown in Figure A-48a-c, respectively. The SPIRE 1D-var retrieved N and T are negatively biased to those of RAOB measurements.



*Figure A-48: (a) The fractional refractivity difference, (b) temperature difference, and (c) water vapor mixing ratio difference between Spire and RAOB measurements where the SPIRE and RAOB pairs are within 300 km and 3 hours.*

The CWDP Spire dataset consists of 2 months of data, June-July 2019, and includes 12 instruments marked as 046, 061, 075, 080, 084, 085, 086, 090, 091, 099, 100, and 101. For analysis, all instruments have been combined for monthly statistics.

SPIRE bending angles and refractivity have been compared with collocated KOMPSAT5 observations ( $\Delta t \leq 2$  hours and  $\Delta D \leq 300$  km). Results for June of 2019 are presented on Figure A-49.

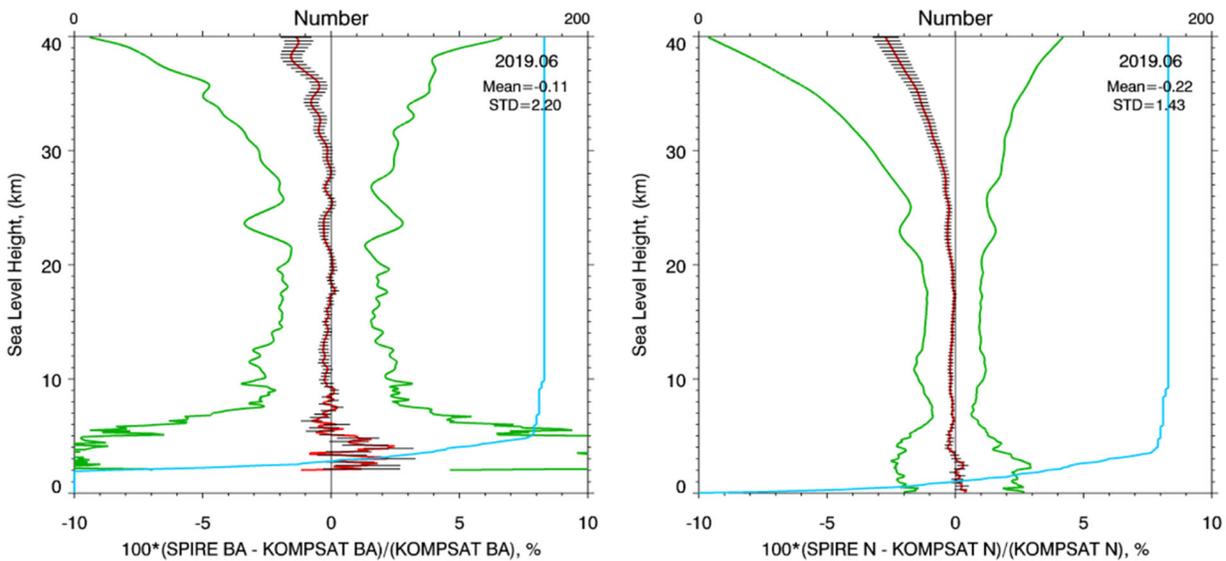


Figure A-49: Bias (red line) and StDv (green line) of bending angle (left panel) and refractivity (right panel) for (SPIRE-KOMPSAT5)/KOMPSAT5 differences, June 2019.

UCAR/SPIRE obtained N-values have been processed with the STAR 1D-Var algorithm and then compared with RAOBAs described in Section 0. Since UCAR-processed N-values have been used as input, it was found that genuine UCAR retrievals are close to STAR retrievals, at least in terms of monthly averages and between both when compared to RAOB (e.g. STAR-UCAR, STAR-RAOB, and UCAR-RAOB). Because of that, the following figures present only the STAR-RAOB comparison for SPIRE data. Figure A-50 shows N-residual between N-values, simulated from retrieved state and RAOB, in 4 latitude zones for SPIRE July 2019 set. Negative bias about -2% (usual for RO technique versus in-situ observations) is found in all latitude zones except Southern mid-latitudes 20S-60S. N-residual variability does not exceed 5% on monthly average. Note that N-comparison is performed on fine altitude grid with increment of 0.1 km.

**N-Residual  $RO_{RTR} - RAOB_{SIM}$ : SPIRE  $RTR_{STAR}(N_{UCAR})$ ; July 2019**

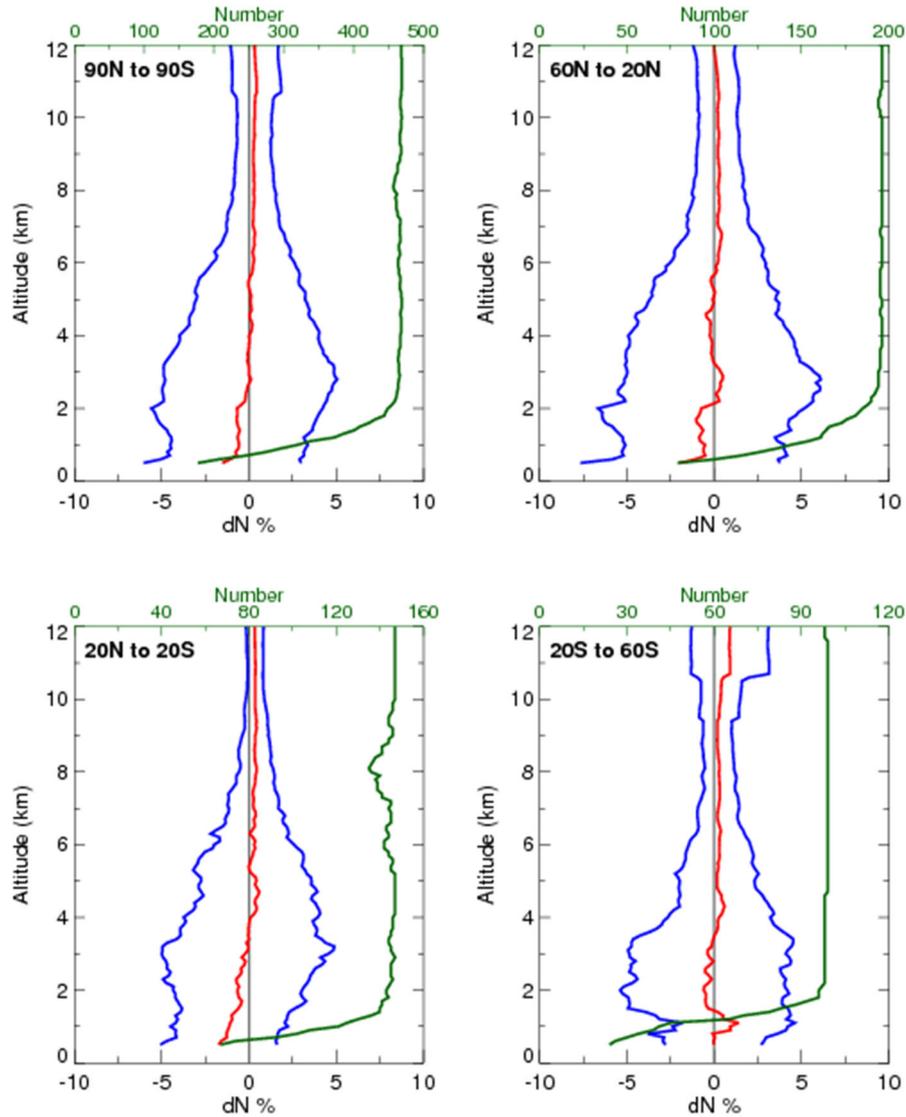


Figure A-50: SPIRE N-residual statistics for N-values, simulated from retrieved states and RAOB: bias (red line) and StDv (blue line) in 4 latitude zones. Green line is a sample size; July 2019.

Comparison of STAR retrieved humidity with RAOB for SPIRE July 2019 dataset is presented on Figure A-51. In terms of monthly averaging, water vapor bias does not exceed few tenth of 1 g/kg in the lower troposphere below 3 km with variability of less/about 2 g/kg.

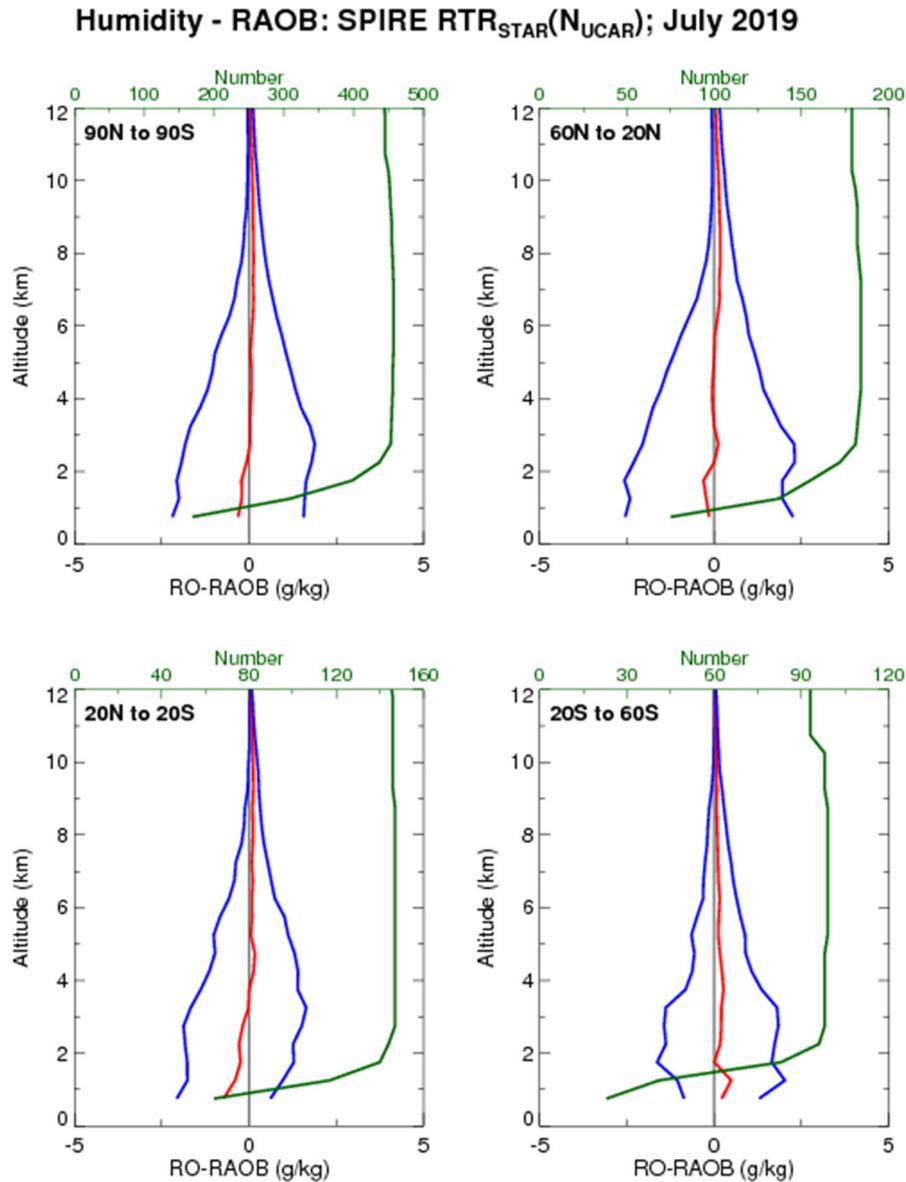
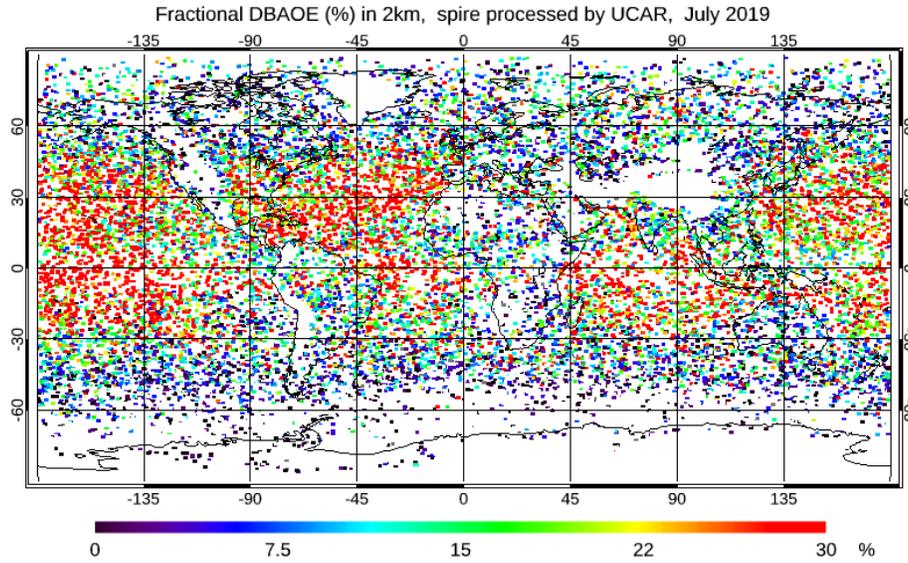


Figure A-51: SPIRE/STAR-RAOB comparison for water vapor in 4 latitude zones, July 2019.

#### A.4.2.6 Observation Uncertainty Estimates

The observation uncertainty of SPIRE is quantified in this section.

Figure A-52 depicts the distribution of fractional DBAOE (defined as  $100\% \times \text{DBAOE}/\text{bending angle}$  at the same altitude) at 2 km MSL on July 2019 for SPIRES. It is obvious that the fractional DBAOE is larger over tropical regions over oceans than that over mid-/high-latitudes over lands. In general, the distributions for fractional DBAOE is highly correlated with the water vapor amount in the atmosphere.



*Figure A-52: Distribution of fractional DBAOE (unit: %) at 2 km MSL on July 2019 for Spire.*

Figure A-53 depicts that the fractional DBAOE for several RO missions including SPIRE at different latitudinal zones. Figure A-53 depicts that the fractional DBAOE for Spire is a little bit higher comparing with other RO missions except for the north-hemisphere mid-latitude.

### Mean Fractional DBAOE, 2019 July , setting, GPS

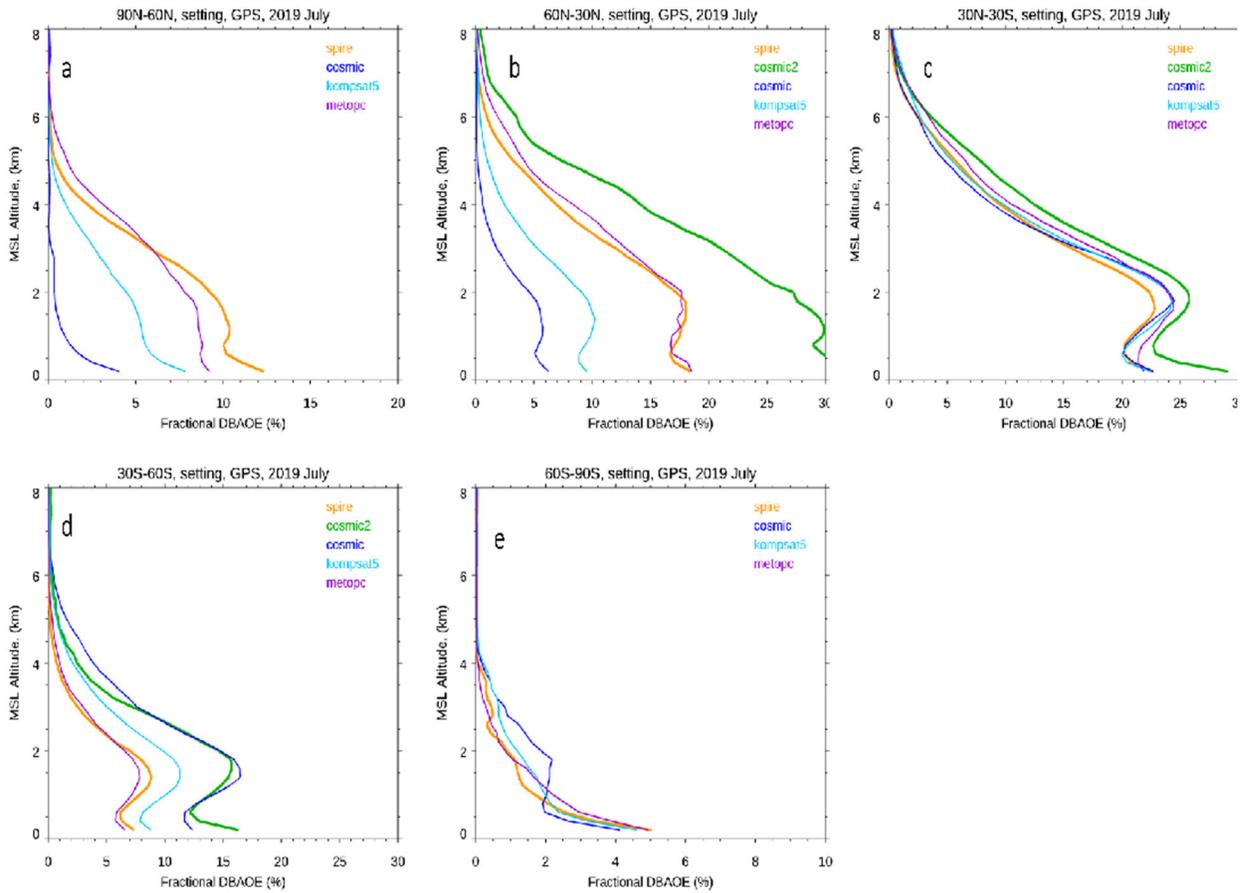


Figure A-53: Comparison of the fractional DBAOE for several RO missions over (a) North Hemisphere high-latitude, (b) North Hemisphere mid-latitude, (c) Tropical region, (d) South Hemisphere mid-latitude, and (e) South Hemisphere high-latitude.

The zonal mean of the fractional bending angle observation error for Spire, COSMIC, and KOMPSAT5 are shown in Figure A-54a-c, respectively.

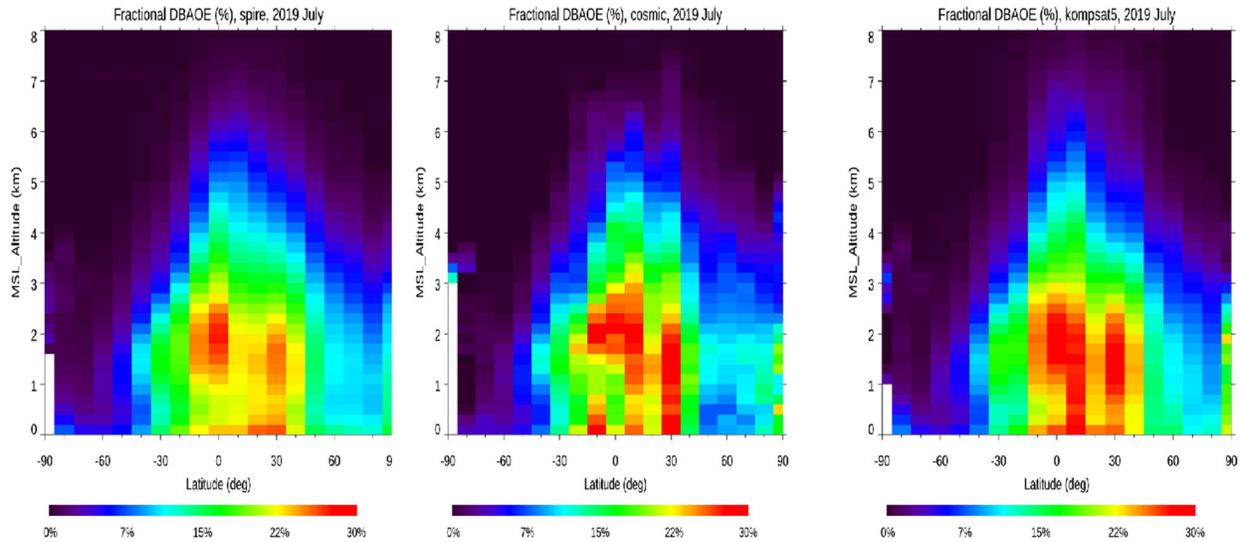


Figure A-54: The zonal mean of the fractional bending angle observation error for (a) GeoOptics, (b) COSMIC, and (c) KOMPSAT5. The contours are the total RO sounding counts (unit: x1000) in an interval of 0.5. RO soundings are sorted in latitude bins of 10o and altitude bins of 200 m.

#### A.4.2.7 Comparison to MW and IR Sounders

For simulation with SPIRE-90/99 (SPIRE and UCAR version) RO data, a detailed description of input variables and parameters for simulations with CRTM is provided in Table A-12. SPIRE version data provides dry temperature, and humidity is provided by ECMWF. UCAR version provides wet temperature and humidity. The three-dimensional variables of ozone mixing ratio, as well as the two-dimensional variables of surface skin temperature, surface wind speed and surface wind direction, are obtained from ECMWF analyses, which are available at a 6-h interval.

Table A-12: Input variables and parameters for SPIRE-90/99 RO data evaluation with CRTM.

SPIRE-90/99 (SPIRE version)		SPIRE-90/99 (UCAR version)	
Category	Variable	Data source	
Atmosphere	Level and layer pressure temperature	<b>SPIRE (dry temperature)</b>	<b>SPIRE</b>
	Specific humidity	<b>ECMWF</b>	<b>SPIRE</b>
	Ozone mass mixing ratio	ECMWF	ECMWF
Surface	Water type	1 (sea water)	1 (sea water)
	Skin temperature	ECMWF	ECMWF

	Wind speed	ECMWF	ECMWF
	Wind direction	ECMWF	ECMWF
Geometry	Altitude	Satellite data	Satellite data
	Satellite zenith angle	Satellite data	Satellite data
	Satellite azimuth angle	Satellite data	Satellite data
	Solar zenith angle	Satellite data	Satellite data
	Solar azimuth angle	Satellite data	Satellite data
	Latitude/longitude	Satellite data	Satellite data
Parameters	Climatology	U.S. standard profile	U.S. standard profile
	Water coverage	1 for ocean	1 for ocean

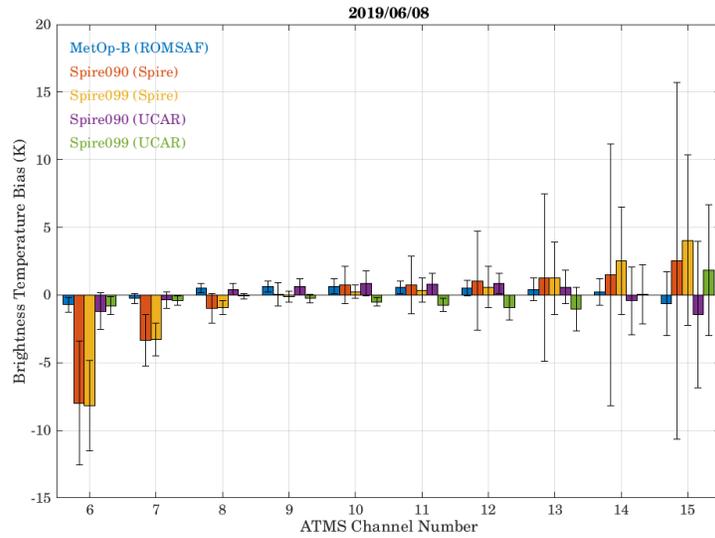
#### A.4.2.7.1 RO vs. MW Comparison

Figure A-55a shows mean and STD for RO vs. MW comparison for ATMS channel 6-15 using SPIRE-90/99 (SPIRE and UCAR) RO data and (b) shows its enlarged results. Table A-13 shows the summary of RO vs. MW data points. MetOp-B is used as reference. Note that for SPIRE version, the humidity is from ECMWF.

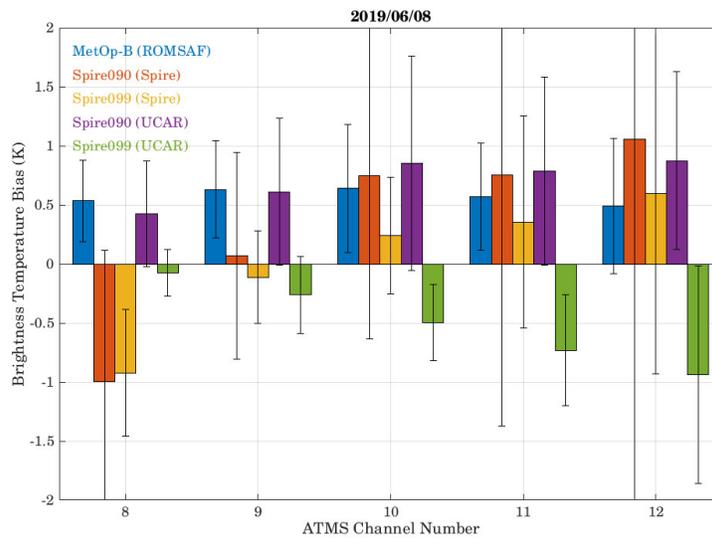
For the BT comparison with SPIRE (SPIRE versions) and UCAR SPIRE-90, the differences for ATMS channel 8-13 are with  $\pm 1.1$  K, which is consistent with the O-B comparison results that the temperature and humidity biases between SPIRE and ECMWF during 270 to 10 hPa are relatively small (see Table A-9 and Table A-10). However, UCAR version SPIRE-99 shows constantly opposite results for channel 9-13, this is because of the abnormal large temperature bias at pressure less than 100hPa (Table A-9).

For SPIRE version, large bias and STD can be noticed for channel 6-8 and channel 13-14. This is because of the large bias at pressure larger than 200 hPa, which is dry temperature, and the large bias at pressure less than 20 hPa for SPIRE version (Table A-10), respectively.

For UCAR Version, on the other hand, SPIRE-90/99 for channel 6 also shows relatively large bias, which dues to the SPIRE-90/99 temperature bias for 700-1000 hPa (Table A-11). And the large BT bias for SPIRE-90/99 (UCAR version) of channel 15 is because of the cut-off at  $\sim 5$  hPa (Table A-11).



(a)



(b)

Figure A-55: Mean and STD for RO vs. MW comparison: SPIRE-90/99 (SPIRE and UCAR)

Table A-13: Summary of RO vs. MW data points. Derived from June 8, 2019 data for all SPIRE RO data.

RO Instrument	L2 Data Processor	# of Valid Profiles	Note
MetOp-B	ROMSAF	15	As reference; Wet T.

SPIRE-090	SPIRE	7	Dry Temperature; Pressure data anomaly
SPIRE-099	SPIRE	4	Dry Temperature; Pressure data anomaly
SPIRE-090	UCAR	8	Wet Temperature
SPIRE-099	UCAR	4	Wet Temperature

#### A.4.2.7.2 RO vs. IR Comparison

Figure A-56 and Figure A-57 show mean and STD for RO vs. IR comparison using SPIRE and UCAR versions of SPIRE-90/99 RO data derived from June 8, 2019 to June 13, 2019, respectively. ERA5 data from ECMWF are used as background and MetOp-B is used as reference. Table A-14 shows the summary of SPIRE vs. IR data points. Note that for SPIRE version, the humidity is from ECMWF reanalysis.

To screen out the cloud effect, the BT difference profiles whose maximum BT difference larger than 4 K for ground channels (860 to 880 wavenumber) have been removed for SPIRE (UCAR version). However, 4 K screening based on ground channel is not suitable for SPIRE version given that SPIRE (SPIRE version) provides dry temperature and the temperature difference with ECMWF can be as large as ~50 K for 400~1000 hPa (Table A-10). Considering the relative small temperature difference for 50~100 hPa (Table A-10), 4 K screening for 680-690 wavenumbers is applied for SPIRE version.

For SPIRE-90/99 (SPIRE version), the BT differences for surface sensitive bands (750 to 1000 wavenumber) are as large as to ~10 K. The reasons for large difference are the large temperature difference for 400~1000 hPa (Table A-10) due to dry temperature provided and the cloud effect cannot be removed.

Large BT bias for 1000 to or 1000 to 1080 wavenumber mainly due to the large temperature difference for 400~1000 hPa (Table A-10). And the relatively large bias for the 670 to 750 wavenumber region is also due to the lack of CO<sub>2</sub> profile. Large bias and STD can be noticed for 670 wavenumber given the large temperature bias at 0~10 hPa for SPIRE-90/99 (SPIRE version).

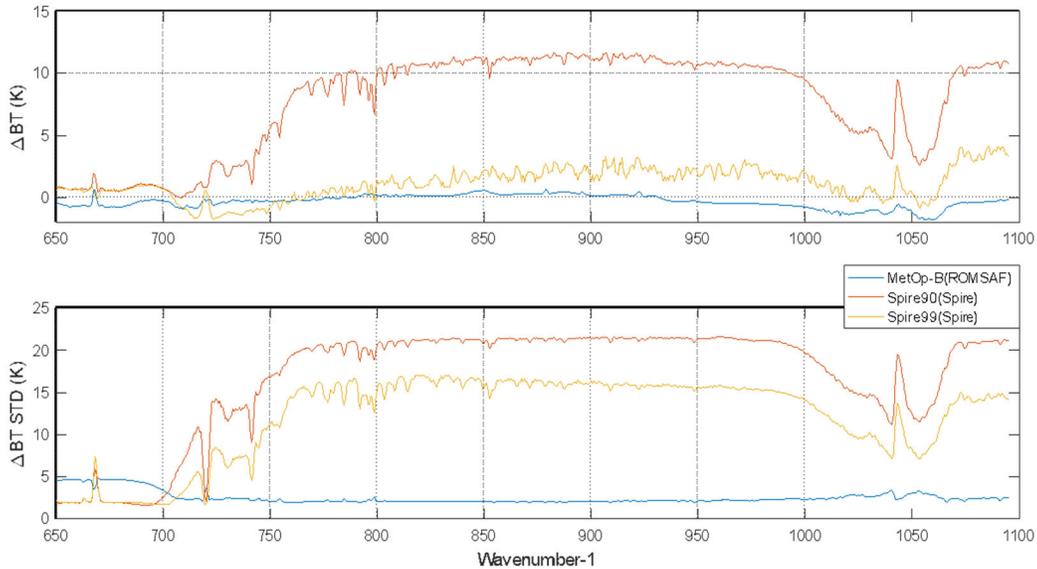


Figure A-56: Mean and STD for RO vs. IR comparison: SPIRE-90/99 (SPIRE Version)

Table A-14: Summary of RO vs. IR data points. Derived from June 8, 2019 to June 13, 2019 data for all SPIRE RO data. ERA5 data from ECMWF are used as background.

RO Instrument	L2 Data Processor	# of Valid Profiles	Note
MetOp-B	ROMSAF	121	As reference; Wet T.
Spire-90	Spire	23	Dry Temperature; Pressure data anomaly
Spire-99	Spire	36	Dry Temperature; Pressure data anomaly
Spire-90	UCAR	25	Wet Temperature
Spire-99	UCAR	35	Wet Temperature

For SPIRE-90/99 (UCAR version) on the other hand, note that for 710 to 1100 wavenumber (corresponding to 200 to 1000 hPa), the BT differences for both SPIRE-099 and SPIRE-099 (UCAR version) are consistent due to their similar performances of temperature and humidity (Table A-10). For SPIRE-099 (UCAR version), however, the temperature shows gradually large difference ranges from 1 to 5 K from 200 up to 3 hPa.

Therefore, BT differences of SPIRE-099 and SPIRE-099 (UCAR version) for 710 to 1100 wavenumber are consistent. BT difference for surface sensitive bands (750 to 990 wavenumber) are relatively stable but shows constant  $\sim 1$  K bias due to the temperature bias ( $\sim 1$  K) for

800~1000 hPa (Table A-10). The large BT difference between 1000 to 1080 wavenumber are likely due to inaccuracies in the reanalysis O<sub>3</sub> profile.

As for 650 to 710 wavenumber, large bias also can be noticed for 670 wavenumber given the temperature and humidity cut-off at ~3 hPa for SPIRE-90/99 (UCAR version). The large BT bias from 650 to 700 wavenumber of SPIRE-099 (UCAR version) is due to large temperature bias at pressure less than 100 hPa for SPIRE-099 (UCAR version) (Table A-10). And the relatively large bias for the 670 to 750 wavenumber region is also due to the lack of accurate CO<sub>2</sub> input profile.

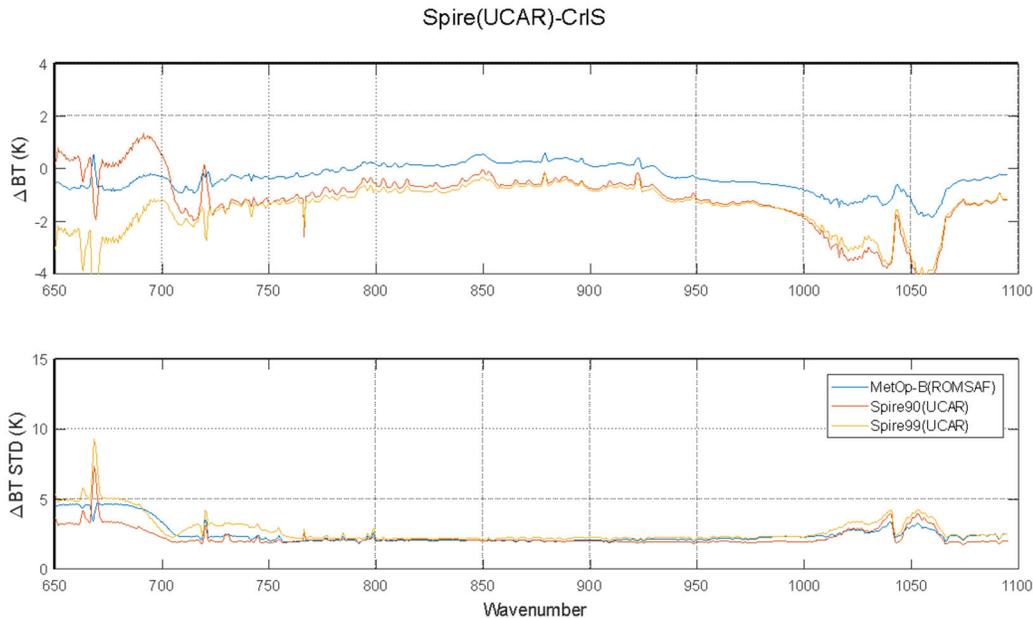


Figure A-57: Mean and STD for RO vs. IR comparison: SPIRE-90/99 (UCAR Version)

## A.5 Summary, Recommendation, and Lessons Learned

- 1) The stability of the L1b positions for GeoOptics and SPIRE are of inferior quality compared to COSMIC2. The LEO positions for GeoOptics have similar stability as COSMIC2 positions (mean absolute deviation of velocity variation ~ 2 – 3 mm/s), but the GNSS positions have larger variations (10 -70 mm/s). For SPIRE, both the LEO and GNSS positions show larger variations compared to COSMIC2.
- 2) The phase data analyzed using Doppler also revealed that the excess phase variations in SPIRE is larger than both COSMIC2 and GeoOptics. COSMIC2 and GeoOptics have larger noise in the L2 excess phase than in L1 excess phase, whereas SPIRE data shows a larger noise in the L1 excess phase. The standard deviation of the Doppler frequency used to analyze the excess phase noise shows that above impact height of –10 km, the L1 standard deviation is < 2 Hz, and increases with decreasing impact height. The corresponding L1 standard deviation for SPIRE data is 8 – 10 Hz from 50 to -50 km impact height. The standard deviation indicates greater noise in the excess phase data. It indicates that the Level L1b data excess phase for SPIRE has greater noise than GeoOptics and COSMIC excess phase.

- 3) The signal to noise ratio for both GeoOptics and SPIRE are on average, smaller than for COSMIC2. However, the penetration depths remain comparable. SPIRE bending angles and refractivity have better agreement with KOMPSAT5 in altitude range of 8-20 km than GeoOptics. Above 20 km, two commercial products demonstrate different signs of bias: it becomes negative for SPIRE, while GeoOptics is positively biased versus KOMPSAT5.

## Appendix A.A CWDP Data Delivery to NOAA/STAR

### A.A.1 Data Pipeline to NOAA/STAR

Vendors push data to the Amazon Web Services (AWS) secure ingest S3 bucket where it goes through a series of security verifications and if found clean is then mirrored to the distribution bucket. NOAA/STAR produces log entries every few minutes by surveying the contents of the distribution bucket and comparing it to the local storage at NOAA/STAR. Log entries are made for new files and files that have changed with respect to the files stored locally at NOAA/STAR.

Files with newly generated log entries are synchronized to the local storage at NOAA/STAR. Files are stored at STAR indefinitely.

From here, new files are mirrored on the STAR SFTP server where they are made available for download by centers participating in the CWDP study for a period of 30 days. Files containing checksums generated by the vendors were provided with the corresponding tar archive. Each center procures credentials for accessing the server from NOAA/STAR IT so that the CWDP data is not available publicly.

Upon request, files older than 30 days, or replacement files, are made available to CWDP participants via the STAR web server for download from preapproved IP addresses.

### A.A.2 GeoOptics

#### A.A.2.1 Data Delivery

GeoOptics data delivery began on November 1, 2018. Initially, data from a single satellite, #85, was provided. On November 10, 2018, a second GeoOptics satellite, #86, was launched and data delivery from #86 began on November 15, 2018. This satellite was not in a sun-synchronous orbit leading to interference during the portions of the orbit entirely exposed to the sun and its eventual failure in April 2019. During its lifetime, #86 was operated at a lower duty cycle than other GeoOptics satellites, producing fewer occultations overall. The third GeoOptics satellite was launched on November 29, 2018 and provided data from December 3, 2018 through the end of the delivery period. Satellite #87 had a similar orbit and produced a comparable yield of occultations to satellite #85.

While the extended data delivery period ended on September 30, 2019, GeoOptics data delivery of the L1a files continued through October 31, 2019, providing one year of data. However, the L2 data delivery of the latest processing version (version 5; see Table A-19) continued only through September 30, 2019, with a previous version covering October 1-25, 2019.

*Table A-15: GeoOptics Satellites*

Number	NORAD Name	Inclination
85	CICERO-7	97.4

86	CICERO-10	85.0
87	CICERO-8	97.4

GeoOptics L1a data was delivered in data dumps corresponding to data downlinks throughout the day with one tar files per coming in every ~2 hours for a total of ~12 files per day per satellite. The L2 data was processed by JPL and provided daily in single tar files with a latency of more than one day. L2 data was initially provided only in netCDF4 format and eventually in BUFR format starting with the version 3 L2 processing (see Table A-19).

*Table A-16: GeoOptics L1a data types and file formats*

Data type	Format	Sample File Name	Notes
Attitude	ASCII in CHAMP-like format	CHatt_2019.300.085.07_txt	
Precise orbit determination	RINEX 3.02	rinex_2019.300.085.07.00_rnx	
Occultations	opnGns 2.1	opngns_2019.300.085.07.00_bnx	High rate data at 100 Hz from GPS and Glonass
Scintillation	RINEX	scn_2019.300.085.07_txt	
Orbits	SP3-D	sp3_2019.300.085.07_sp3	Calculated on-board by the POD receiver. Data is also provided in the attitude files, thus this file type is not always provided.
podTec	NetCDF4	podtec_2019.300.085.07_nc	
occTec	NetCDF4	occtec_2019.300.085.07_nc	

*Table A-17: GeoOptics L2 data types and file formats*

Data type	Format	Sample File Name	Notes
-----------	--------	------------------	-------

L2 bending angle, refractivity, temperature, and water vapor profiles	netCDF4	cicPrf_2019.150.085.00.g14.005_nc	Each netCDF4 file contains multiple occultations, each with its own group. Each occultation group contains atmo_phase, atm_profiles, and wet_profiles groups.
L2 bending angle, refractivity, temperature, and water vapor profiles	BUFR	bfrPrf_2019.150.07.53.085.R24.005_bf	Provided in version 3 and after (see table). A correction to the undulation was applied in version 5.
Low-Earth-orbiter precise orbit determination product	NetCDF4	cicPod_2019.150.085.00.005_nc	

### Versions

GeoOptics launched two of their three satellites during the data delivery period and the third a short time prior to the start. While L1a processes is done in-house, L2 processing is performed by JPL. Both processing chains underwent several changes during the data delivery period. Additionally, flight software updates to the instrument were performed.

Only one L1a software update was significant enough to warrant a reprocessing of the entire data record to that point (indicated by an asterisk in Table A-18), so no one L1a version covers the entire data delivery period. Other changes, outlined in Table A-18, were minor or involved the addition of ionosphere data products to the L1a distribution.

*Table A-18: GeoOptics L1a data versions. Those denoted with an asterisk involved reprocessing the prior mission data.*

Version Name	Description	Coverage Range
V0	Sample data provided prior to start of data delivery on 1 Nov. 2018	08/30/2018 – 09/06/2018 10/28/2018 – 10/31/2018
V1	A portion of the L1a data covered by this version lack S4 data due to use of 1 Hz Novatel range messages.	11/01/2018 – 11/28/2018
V2018-11-27-afx	New baseline adding naming convention for versioning	11/01/2018 – 11/27/2018
V2018-11-28-acqq*	Corrected Cion high-rate data ground processing to maintain precision	11/29/2018 – 11/29/2019
V2018-11-29-aadr	Improved efficiency of ground processing	11/29/2018 – 05/15/2019
V2019-05-01-axig	Capability for producing podTec from Novatel pseudorange and phase implemented; however, files were not immediately distributed while bugs were fixed	04/30/2019 – 05/29/2019
2019-05-28-anhu	Refactoring of ground reprocessing	05/30/2019 – 08/20/2019
V2019-07-30-ahzk	Capability to produce occTec profiles from Cion ionospheric observation implemented; however, files not immediately released while data was assessed	08/21/2019 – 08/28/2019
V2019-08-28-ajln	Updates to occTec files while evaluation continued	08/29/2019 – 09/27/2019

2019-09-27-awbf:	Release of occTec files and fixes for bugs in podTec processing that yielded empty files	09/16/2019 – 10/31/2019
------------------	--	-------------------------

JPL also implemented a number of changes to the L2 processing to make continuous improvements to the Glonass processing and provide BUFR data products. Several version updates involved reprocessing the entire data delivery to that point. The final version, version 5, covers the entire data delivery period until the end of September when data delivery officially ended. Although L1a data delivery continued throughout October, only version 3 L2 data was delivered during that month. The update to version 5 replaced the undulation variable in the version 4 processing, and by October production of the version 4 processing had ceased in favor of a newer update that was not delivered as part of the CWDP program.

*Table A-19: GeoOptics L2 processing versions. Those denoted with an asterisk involved reprocessing the prior mission data.*

Version Name	Description	Coverage Range
Version 1 (V2018-11-01-astb)	Initial version (files not labeled as such)	11/01/2018 – 01/30/2019
Version 2 (V2019-02-01-aqdb)*	Format revisions	11/01/2018 – 05/27/2019
Version 3 (V2019-06-11-apzs)*	Correction for small negative bias in Glonass processing leading to rejection of profiles corrected. BUFR files produced.	11/01/2018 – 10/25/2019
Version 4 (V2019-09-03-apdd)*	Applied canonical transform (i.e. wave optics) processing to Glonass profiles	11/01/2018 – 09/30/2019
Version 5 (V2019-11-04-amed)*	Applied a patch to the undulation value in BUFR files to improve altitudes calculated using impact parameter (Note: the undulation in the L2 netCDF files was not updated)	11/01/2018 – 09/30/2019

### **A.A.2.2 Daily Counts**

GeoOptics set out to provide 1200 occultations per day, however, following the loss of #86, it was difficult to meet this target. To compensate the data delivery period was extended by two additional months. Both #85 and #87 performed consistently throughout the data delivery period providing ~400 occultations each per day from which L2 data could be retrieved using version 5 of the JPL processing. Just over 60% of L2 occultations come from GPS signals, and nearly 40% come from Glonass signals. Of the L2 profiles provided in BUFR format, ~70% are flagged as nominal.

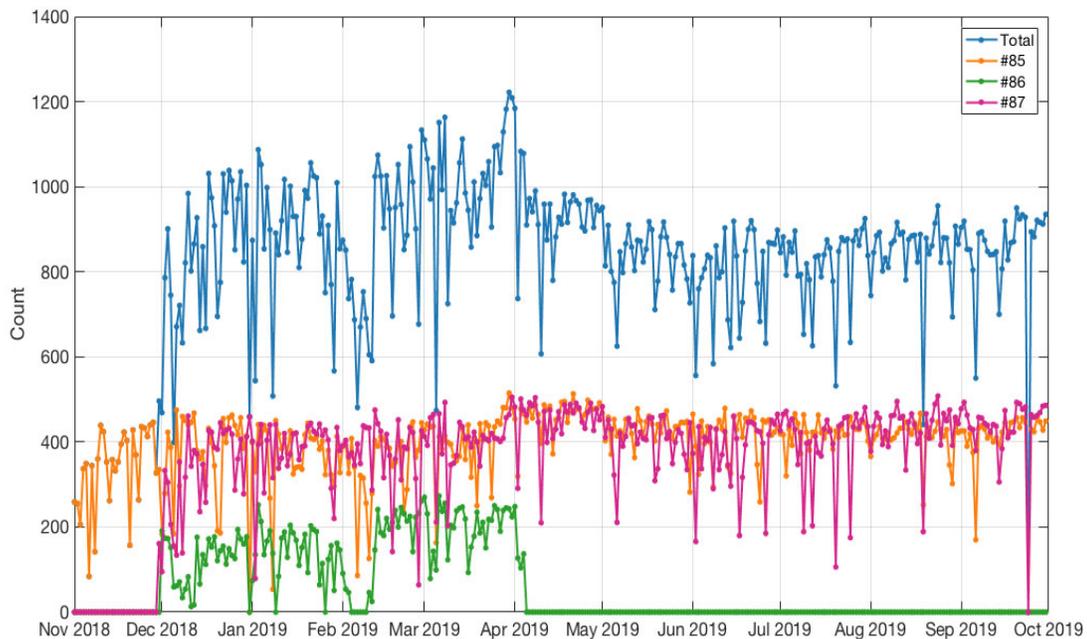
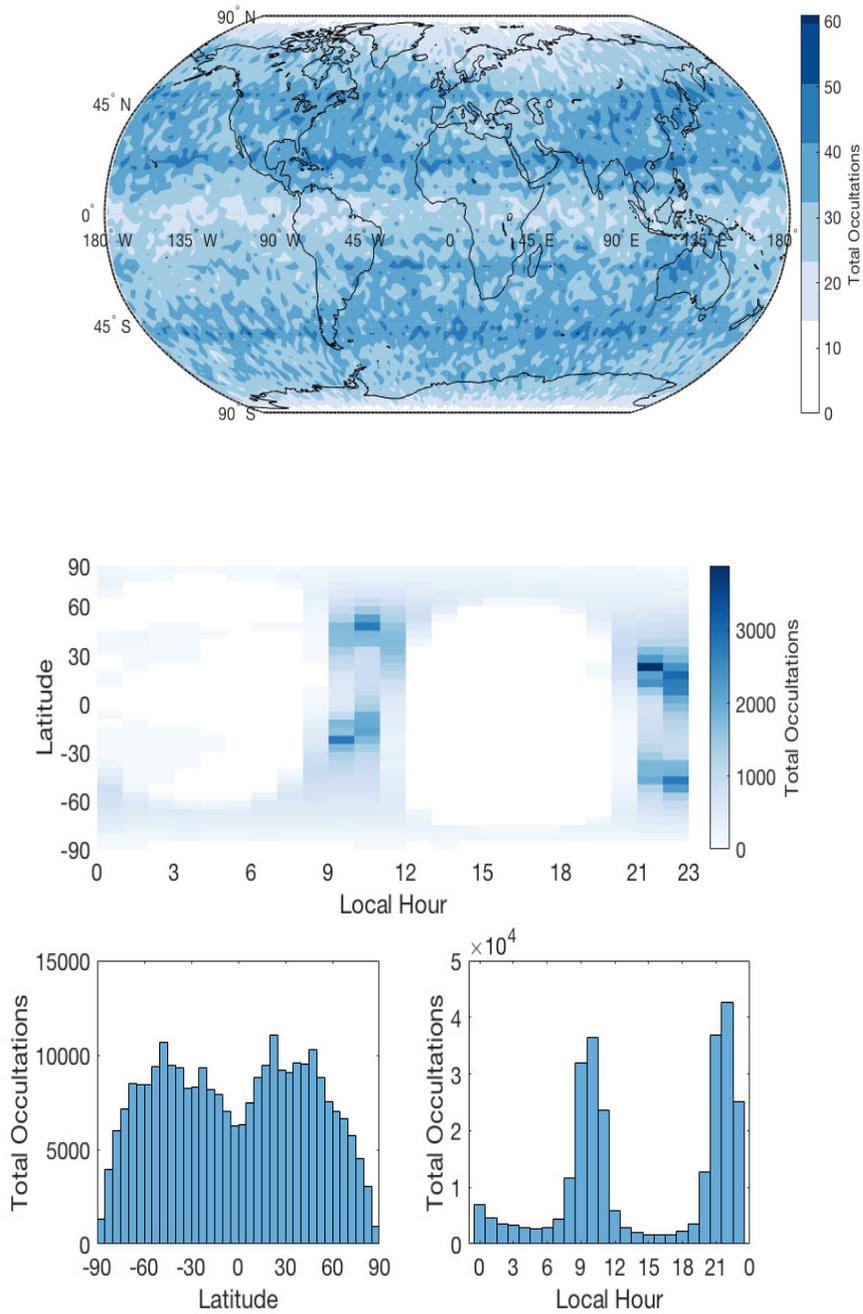


Figure A-58: GeoOptics Daily Counts of L2 profiles with retrieved wet profile variables for version 5.

### A.A.2.3 Spatial and Temporal Characteristics

GeoOptics satellites #85 and #87 are both in high inclination sun-synchronous orbits allowing for good spatial coverage of the Earth as can be seen in the top panel Figure A-59. However, local time coverage is restricted as can be seen in the middle panel of Figure A-59. At each latitude, there is the potential for twice-daily observations that allows for the capture of the diurnal cycle at each latitude.

Satellite #86 was not in a sun-synchronous orbit and offered a more extensive coverage of local time with latitude; however, it only provided a smaller amount of data for a portion of the data delivery period.



*Figure A-59: The upper panel shows distribution of GeoOptics L2 occultations (version 5) distributed globally over the entire data delivery period from Nov. 2018 through Sept. 2019. The middle panel shows the counts over the whole period as a function of latitude and local hour. The lower two panels show the distribution of the occultations with respect to latitude and local hour respectively.*

### **A.A.2.4 Lessons Learned**

The data delivery period for GeoOptics coincided roughly with the launch of their satellites. Thus, the data processing chain had not reached a stable configuration and several processing versions were required to identify and correct issues with both the L1a and the L2 data. Early L2 netCDF4 files contained improperly formatted data and missing variables. Ionospheric data was not produced for the entirety of the data delivery period.

NWP data assimilation systems are designed to ingest observations in BUFR format, standardized by the WMO. Early versions of GeoOptics L2 processing did not include BUFR files. STAR provided the GeoOptics data contained in the netCDF4 files in BUFR format to the JCSDA so they could begin their analysis; however, missing variables still posed a problem for use with the GSI and other data assimilation systems.

For future rounds, it is important that a provider has a stable processing chain that can reliably produce the various levels of data and data types in the common file formats that conform to specifications.

The GeoOptics L1a data was delivered throughout the day of collection and the L2 data the following day. For each file delivered, SHA256 checksums were provided in separate files to verify the contents. Occasionally the need will arise to redeliver a file to correct for an incomplete delivery or processing issue. Redelivered files need to be distinguishable from the original delivery and include an updated checksum so that all provided files can be retained and verified.

The naming convention used by GeoOptics for tar files did not easily accommodate renaming redelivered files. A four-character code that signifies an internal process was added to the files, with the intention of incrementing the first character to signify a redelivered file. In practice, this option was not used. Occasionally redelivered files would replace the original delivery because the file names were identical. Redeliveries often lacked the accompanying updated checksum file leading to the appearance of corrupted data when compared against that from the original delivery.

Future efforts could avoid this scenario by choosing a file naming convention that can easily differentiate replacement files, e.g. including a creation date. Greater care is needed to ensure that all checksum files are present for verification of contents. Stricter screening to reject redelivered files of the same name and to guard against checksum mismatches should also be implemented.

Tar files delivered by GeoOptics for both the L1a and L2 data contain a directory structure that unpacks files of a given type into a corresponding directory. Internal directories in tar archives are undesirable. They make inventorying the contents of the file challenging for the purpose of archiving the data. They also can be cumbersome for users working with the data as they do not allow the user to decide how data will be stored. Future deliveries should not have directories in the tar archives.

### **A.A.3 SPIRE**

#### **A.A.3.1 Data Delivery and Counts**

SPIRE data delivery commenced on June 1, 2019 and ran through July 31, 2019. While contracted to provide 500 occultations per day, SPIRE elected to deliver 600 occultations per

day starting on June 11, 2019. Following the end of data delivery in July, SPIRE redelivered both June and July datasets. The June redelivery only provided updates to the podTec files; however, the July redelivery included an additional 600 profiles per day bringing the total daily count to 1200 in addition to the new podTec files. Both June and July redeliveries replaced the original occultations, although the overlap with the original dataset is not complete.

*Table A-20: SPIRE satellites that collected data provided during the data delivery period*

FM#	NORAD Name	RO Antenna Configuration	Altitude (km)	Inclination	Number of Occultations
46	LEMUR-2-LYNSEY-SYMO	BRO	500	SSO (LTDN 10:30)	69
61	LEMUR-2-BROWNCOW	BRO	505	SSO (LTAN 9:30)	2831
75	LEMUR-2-KADI	BRO	585	SSO (LTAN 11:15)	491
80	LEMUR-2-URAMCHANSOL	BRO	585	SSO (LTAN 11:15)	996
84	LEMUR-2-ALEXANDER	BRO	445	51.6	5709
85	LEMUR-2-VU	BRO	445	51.6	131
86	LEMUR-2-ZUPANSKI	BRO	500	83	41
90	LEMUR-2-DULY	BRO	505	SSO (LTDN 10:00)	4019
91	LEMUR-2-REMY-COLTON	BRO & FRO	585	SSO (LTAN 11:23)	193
99	LEMUR-2-JOHANLORAN	BRO & FRO	505	SSO (LTAN 09:30)	19106
100	LEMUR-2-BEAUDACIOUS	BRO & FRO	505	SSO (LTAN 09:30)	13766
101	LEMUR-2-ELHAM	BRO & FRO	505	SSO (LTAN 09:30)	6833

SPIRE occultations all have good quality control flag. SPIRE operates many satellites, from which a subset of 12 satellites was used to collect the data provided for this effort (see Table A-20). SPIRE receivers collect occultations from GPS and Glonass GNSS constellations, as well as Galileo and QZSS. Relatively few occultations were collected from Galileo and QZSS as can be seen in Figure A3.

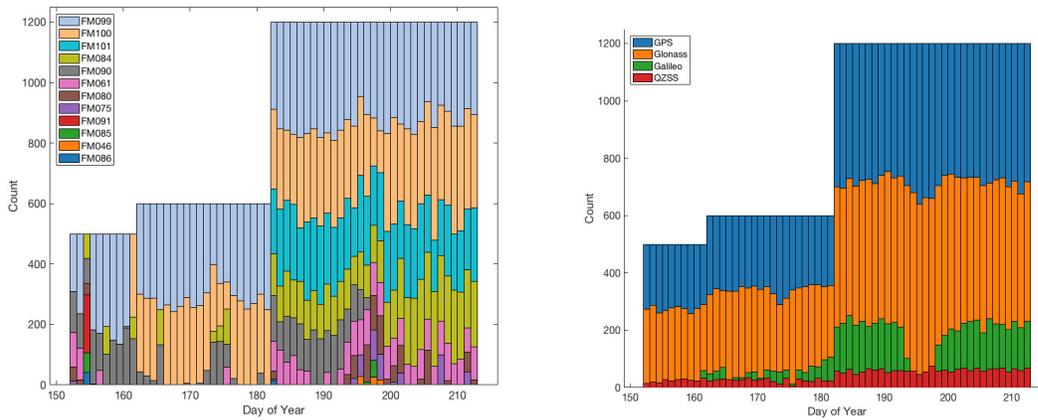


Figure A-60: SPIRE counts over the data delivery period from June through July 2019 broken down by LEO satellite (left panel) and GNSS constellation (right panel).

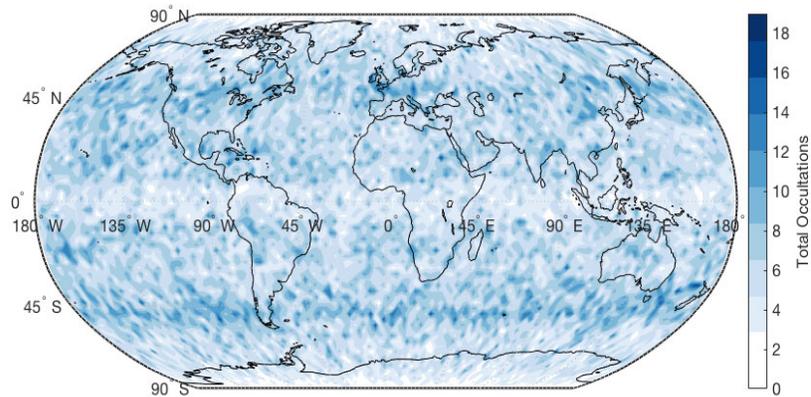
Table A-21: File types delivered by SPIRE.

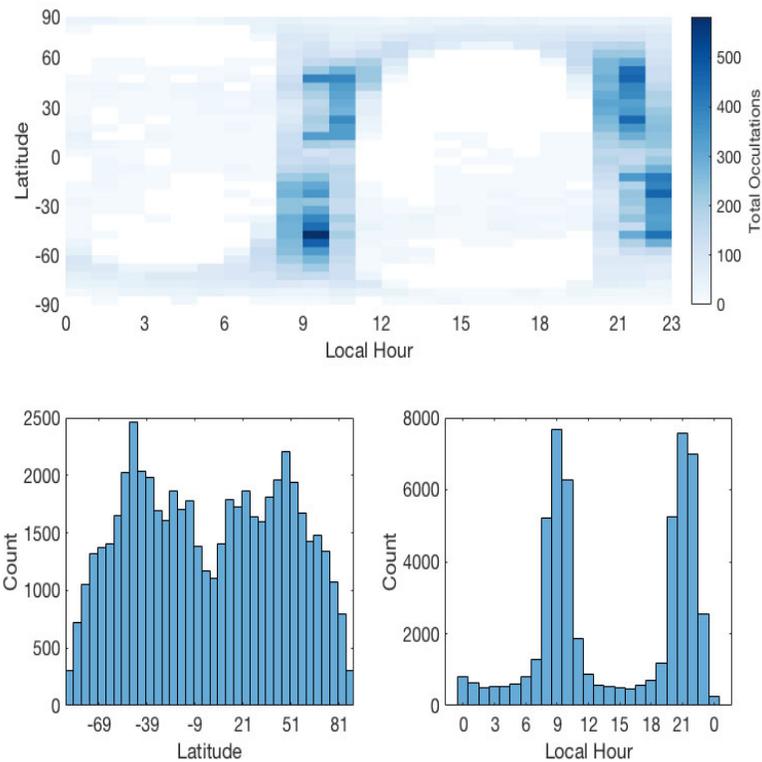
Data type	Format	Sample File Name
Attitude	ASCII in CHAMP-like format	leoAtt_2019-07-01T08-39-19Z.2546245.090.log
Precise orbit determination	RINEX 3.02	podObs_2019-07-01T08-39-19Z.2546246.090.antPOD.rnx
Occultations	opnGns 2.1	opnGns_2019-07-01T13-58-15Z.2547788.090.G12.antPOD_2_GPS_L1_CA_15.rst.bin
Coarse orbit position and velocity	SP3-D	leoOrb_2019-07-01T08-39-19Z.2546246.090.sp3
Precise orbit position and velocity	SP3-D	podObs_2019-07-01T08-39-19Z.2546246.090.nav_sp3
Ionosphere total electron content	NetCDF	podTec_2019-07-01T08-39-19Z.2647799.090.G09.antPOD_seg0.nc

Ionosphere scintillation	NetCDF	scnLv1_2019-07-01T08-39-19Z.2647799.090.G12.POD.seg0.nc
Electron density	NetCDF	ionPrf_2019-07-01T08-39-19Z.2647799.090.G09.antPOD.seg0.nc
Excess phase	NetCDF	atmPhs_2019-07-01T09-05-59Z.2546318.090.E12.nc
L2 bending angle, refractivity, temperature	BUFR	bfrPrf_2019-07-01T09-38-06Z.2546332.090.G06.bufr

### A.A.3.2 Spatial and Temporal Characteristics of Delivered Data

Most of the SPIRE satellites that contributed to the data collection provided for this study are in sun synchronous orbits. This restricts to local time of observations as can be seen in the middle panel of Figure A4. The satellites in low elevation orbits (see Table A-20) provided relatively few occultations.





*Figure A-61: The upper panel shows distribution of SPIRE L2 occultations (BUFR) distributed globally over the entire data delivery period from June 2019 through July 2019. The middle panel shows the counts over the whole period as a function of latitude and local hour. The lower two panels show the distribution of the occultations with respect to latitude and local hour respectively.*

### A.A.3.3 Lessons Learned

1. While SPIRE has a robust processing chain and provides data in the common formats, the section of the BUFR files reserved for L2 data contains only the dry temperature rather than the specified L2 variables of “wet” temperature and specific humidity. The L1b portion of the BUFR files conform to the specification, containing raw bending angles that were unoptimized, i.e. foregoing the use of climatology or model data to statistically correct the bending angle profile, and were thus suitable for use in NWP. However, the lack of the expected L2 data made it difficult to perform other validation studies. This data would be desirable to other users performing studies beyond NWP applications where only the bending angle is used.
2. SPIRE’s processing throughout the data delivery period was stable. However, redeliveries of data were performed to provide additional profiles and later to update leoOrb files that were reprocessed to provide improved POD solutions (Note: these update POD solutions were not used for the analysis in this report).

When files are redelivered it is important that the filenames be distinguishable from the original delivery. SPIRE’s file naming convention for their tar files did not easily accommodate

this requirement. SPIRE tar file names contain a seemingly random string of characters that denote an internal process related to the production of L2 BUFR files. If the BUFR files produced for the corresponding dump were different, this string of characters differed between the original and subsequent delivery. However, in many instances, this was not the case and redelivered files meant to replace the original delivery had the same name. Furthermore, in some instances redelivered tar files containing only leoOrb files would also have the same name as the original delivery they were meant to supplement, not replace.

In the future, efforts should be made to implement a strategy that makes replacement files and files with different contents easily distinguishable. The inclusion of a creation date will satisfy the first criteria and a prefix denoting contents will take care of the second.

3. SPIRE delivered tar files correspond to a particular data collection range of times for a given satellite/receiver, and the files were named accordingly. Files within the tar file included a directory structure also named according to this collection range. Internal directory hierarchies are undesirable as they make inventorying the contents of the archive a challenge. In addition, unpacking the data grouped by collection time and satellite is cumbersome for users working with the data, particularly when the date of the data itself does not correspond to the collection window, i.e. when it falls on the cusp between days.

In the future, requirements for TAR files without internal directories should be implemented, both for the purpose of archiving the data where inventories are important and to allow users to decide how to store the contents.

## **Appendix A.B Bending Angle Retrieval from Level1a Dataset**

In this appendix, we provide the steps converting the phase observation (opnGns files) to bending angle (cicPhs files) and compare our solution with GeoOptics and UCAR to understand the differences and errors in excess phase and bending angle products.

### **A.B.1. Extracting the opnGns phase and Rinex file phase**

The OpnGns of GeoOptics has 100 HZ observations, approximately 0.02 seconds intervals. Usually the file provides the observations (pseudo range: C1C,L1 Phase: L1C,L2 Phase: L2L, and SNR: S1C,S2C ) and modeled phase data (L1C(M) and L2L(M)). One important step in this opnGNS processing is to extract all the information and concatenate into one formatted file (easily readable into matlab) and determine each Radio Occultation events and their associated start and end time. Based on the each RO event start and end time, we look into the same day RINEX observations from POD antennas. For each continuous of a single GNSS observation by the POD antenna, we called them a reference link event. The pair between the an RO event and an reference link event are determined by looking at the reference link data SNR and its time range covering the RO event.

### **A.B.2. Orbital Determination for LOE/GNSS position and velocity**

The LEO (GeoOptics) POD has been provided as SP3-D format with 1HZ interval. At this time, due to the lengthy processing in Bernese, we decided to use the GeoOptics provided L1a POD information. We extract the GPS time, Position and velocity in ECEF and Clock bias from all the SP3-D files in one day and forms a formatted data file into matlab. This step is relatively easy since it only involves the format change. However, we found that the SP3-D orbit has large clock bias. The bias in the time derivatives of the positioning can cause wavy structure in the bending angle profiles. We also looked into the level-2 POD file, cicPOD, which has complete 1HZ POD information. However, the POD position of the LEO satellite is given in ECI coordinate system and we have to do the coordinate transformation.

To derive the GNSS position/velocity and clock information, we have to rely on the CODE/IGS products. The 15 minutes CODE GNSS solution and Earth Orientation Parameters are downloaded/reformatted and feed into the Bernese software. The orbit and clock are reproduced into 30 seconds products using SP3-C format. The SP3-C formats are then used to form formatted inputs to excess phase model to provide the GNSS pos/vel/clk data. For both LEO and GNSS orbit, a high order polynomial interpolation (9<sup>th</sup> order) has been used in the interpolation of orbital time to observational time.

### **A.B.3 Extracting the Attitude Information**

The GeoOptics attitude file adopts the Champ convention. The quaternions are defined from space craft coordinate to ECI (J2000). We extract the GPS time and the Quaternions into formatted file ready into matlab. Also, the GeoOptics attitude files include the POD and antenna offset information. From here we can derive the antenna offset for POD and excess phase calculation. Together with the SP3/cicPOD file provided GNSS/LEO mass center position, the antenna offset and attitude information can be used to derive the LEO antenna position/velocity in ECI coordinate system.

### **A.B.4 Earth Coordinate System Conversion**

The GNSS/LEO satellite positions are given in ECEF coordinate system with the IGS convention. However, the excess phase calculation involves terms must be corrected in ECI coordinate system. Thus, we need convert the Pos/Vel data sets from ECEF to ECI. While the general coordinate transformation can be done using a generalized matrix considering the earth rotation, the accurate conversion from ECEF to ECI (vice versa) needs a well-defined equatorial plane and an earth's pole, where the earth's rotation, , precession, nutation and Polar wander must be considered. Here, we use True of Date coordinate system, one of the ECI, the same as UCAR used for our processing. Before doing the temporal interpolation from POD time (30 seconds interval) to the high rate phase observation time (100 Hz), we carried out the coordinate transformation first, which means the interpolation is done in ECI coordinate system.

### A.B.5 Cycle Slip Detection

The cycle slip happens when connecting the observed phase in a range of  $[-\pi, +\pi]$  to a continuous, unwrapped phase time series. In the discontinuity, an integer number of the (half) wave lengths must be added into the time series; otherwise the time derivative will be changed abruptly and hence affects the bending angle calculation. Generally, the residual phase between the observed L1/L2 phase and phase model are looked. During the phase-locked loop (close) stage, the difference can be easily identified since the change of residual phase between two observational points are close to zero, hardly exceed quarter of wave length (0.19cm for L1). The integer number of  $\pi$  can be added by minimizing the difference between the previous observation and current. However, during open loop stage, we may also need to rely on the navigation bit time series or an internal NDM correction. Once the phase has been reconnected, it will be added back to the phase model and then provided for excess phase calculation.

### A.B.6 Calculation of Excess Phase

We designed an excess phase model (mainly in matlab) to incorporate all steps from coordinate conversion, polynomial interpolation, excess phase calculation and netCDF data output.

The excess phase model can be expressed as the following equation (Figure A-62) (Shreiner et al. 2009):

$$\begin{aligned}
 \text{Carrier Phase Measured } L1_r^s(t_r) &= \underbrace{c \cdot \delta t_r(t_r)}_{\text{Leo Clock Error}} + \underbrace{c \cdot \delta t_{r,\text{rel}}(t_r)}_{\text{Range of GNSS/LEO}} + \underbrace{\rho_r^s(t_r)}_{\text{GNSS Clock Error}} + \underbrace{\delta \rho_{r,\text{rel}}^s(t_r)}_{\text{Excess Phase } (\Delta S) \text{ Wanted}} \\
 &+ \underbrace{c \cdot \delta t_r^s(t_r - \tau_r^s)}_{\text{GNSS Clock Error}} + \underbrace{c \cdot \delta t_{\text{rel}}^s(t_r - \tau_r^s)}_{\text{GNSS Clock Error}} + \underbrace{\delta \rho_{r,\text{ion}}^s(t_r)}_{\text{GNSS Clock Error}} + \underbrace{\delta \rho_{r,\text{trop}}^s(t_r)}_{\text{GNSS Clock Error}} + \underbrace{\lambda_1 \cdot N_{\text{amb}}}_{\text{Phase Ambiguity}} + \varepsilon \\
 &\quad \text{Relativity Effects}
 \end{aligned}$$

Figure A-62: The calculation of excess phase involves different types of phase corrections

From the equation, the dominate part is the LEO clock error. For COSMIC-1, the high rate POD observations can be paired with OCC high rate observations to remove this clock error completely (only increase noise level due to smoothing procedure of L1/L2 phases). For the GNSS clock error, since the GNSS clock is usually very stable (or drift rate is very stable), the error from polynomial interpolation from 30 seconds (or 1HZ GNSS ground station

observations) is negligible, a zero differencing can be applied to these parts usually within a few millimeter differences in excess phase. However, we need to estimate the time propagation  $\frac{s}{c}$  from the transmitter to receiver. This can be roughly calculated using the direct distance divided by speed of light and recursively calculated using the GNSS orbit and LEO receive position in ECI coordinate system. The range between transmitter and receiver,  $s_r(t^r)$ , refers to the antenna phase center position. Thus, it combines the POD mass center position, the antenna offset and the antenna phase center variation. However, to connect these three, the attitude information must be taken into accounts since the antenna offset and PCVs are given in instrument coordinate system.

It recognized that the POD with accuracy of 10 cm level does not have significant impact on the bending angle calculations, such as COSMIC-1. By examining the POD error in the GeoOptics POD netCDF file, the uncertainty of the POD is on the order of 0.5 cm level (as  $1\sigma$  value). Similar to the CLOCK drift, the position error can be easily more than centimeter level. To reduce error, a 10<sup>th</sup> order polynomial interpolation scheme has been applied to the position and velocity interpolation. The general relativity effects consider time and distance difference when referencing different clocks on different satellites. Those corrections are small but must be done in a inertial system, that's why the excess phase is always associated with position/velocity in the ECI coordinate system.

The basic steps in calculation of excess phase:

- 1) Determine the high rate OCC time, in 0.01 second interval.
- 2) NDM removal for residual phase and added back to phase model.
- 3) Calculate the phase center position/velocity of the LEO satellite receiver and GNSS transmitter at OCC time 10<sup>th</sup> order polynomial interpolation.
- 4) Calculate the range of the transmitter and receiver.
- 5) Subtract the range from both L1/L2 phase.
- 6) Calculate general relativity terms and subtract these corrections from phase.
- 7) Calculate the referenced link 1HZ L1/L2 observations from RENIX file and pair with the open
- 8) Remove the GNSS clock bias using zero differencing (4<sup>th</sup> order polynomial interpolation).
- 9) Apply the same algorithm to 1HZ POD observations and interpolate the results into OCC observation time.
- 10) Linearly combine L1 and L2 phase from 7) to form L3, the single differencing clock error.
- 11) Subtract L3 in 6) from residual phase in 4.) and subtract a constant to make the first excess phase value as zero.
- 12) Format the datasets and output as netcdf as function of GPS time (with bias corrected). The netCDF variables include excess phase, phase model, position and velocity of transmitter and receiver, SNR as function of the GPS time (with bias corrected).

While the steps look straight forward, there are many caveats and subtle steps needing additional care. Such as dealing with the LEO receiver POD/OCC antenna offsets needs the LEO attitude information, dealing with GNSS antenna offsets needs the GNSS attitude information. The accuracy of observational time is also important. With receiver clock bias on the order of  $\sim 0.1$ ms, the range error between transmitter and receiver can be 10 centimeters difference (think of the LEO/GNSS relative movements) if not subtracted from the receiving time.

### A.B.7 Calculation of Bending Angle

Once the excess phase and SNR as well as associated position/velocity in ECI have been determined out of excess phase model, we can use the Radio Occultation Processing Package for next step, bending angle conversion. The inputs are solely a netCDF file containing all the excess phase related information and a parameter control file. Figure A-63 shows one case of the bending angle profiles using the SP3-D and cicPOD files. We notice that the SP3-D files can cause large errors in the BA, while using cicPOD can result in better solution in bending angle, which is very close to GeoOptics bending angle.

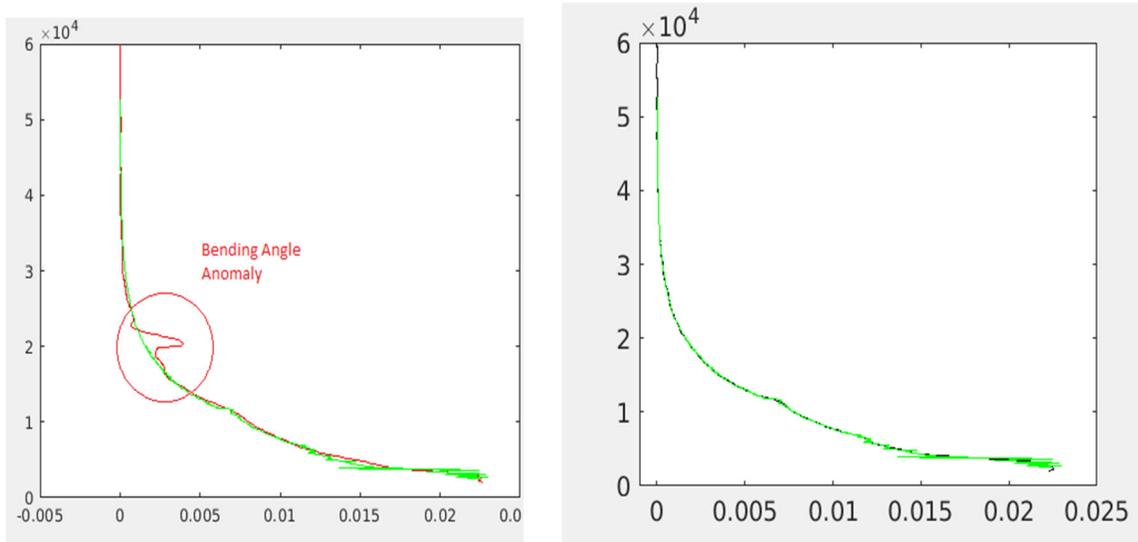


Figure A-63: Bending Angle Profiles using GeoOptics SP3-D (left) and cicPOD orbit products.

**References**

- Bean, B. R., and E. J. Dutton, 1966: *Radio Meteorology. National Bureau of Standards Monogr.*, No. 92, U.S. Government Printing Office, 435 pp.
- Culverwell, I. D., H. W. Lewis, D. Offiler, C. Marquardt, and C. P. Burrows, 2015: The radio occultation processing package, ROPP. *Atmos. Meas. Tech.*, **8**, 1887–1899, doi:10.5194/amt-8-1887-2015.
- Kursinski, E.R., G.A. Hajj, J.T. Schofield, R.P. Linfield and K.R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, **112**, No. D19, 23,429-23,465.
- Rodgers, C., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Series on Atmospheric, Oceanic, and Planetary Physics, Vol. 2, World Scientific, 238 pp.
- Schreiner, W. and Rocken, C. 2010: Quality assessment of COSMIC/FORMOSAT-3 GPS radio occultation data derived from single-and double-difference atmospheric excess phase procession, *GPS Solut.*, 14, 13–22.
- Xia, P., Ye, S., Jiang, K., & Chen, D., 2016: Estimation and evaluation of COSMIC radio occultation excess phase using undifferenced measurements.

## Abbreviations and Acronyms

1D-VAR	1 Dimensional Variation
ATMS	Advanced Technology Microwave Sounder
BT	Brightness Temperature
CICS	Cooperative Institute for Climate and Satellites
COSMIC	Constellation Observing System for Meteorology, Ionosphere, and Climate
CrIS	Cross-track Infrared Sounder
CRTM	Community Radiative Transfer Model
CWDP	Commercial Weather Data Pilot
DA	Data Assimilation
DBAOE	Dynamic Bending Angle Observation Error
ECEF	Earth Centered Earth Fixed
ECMWF	European Center for Medium-range Weather Forecasts
ERA	ECMWF Reanalysis
FFT	Fast Fourier Transform
FSI	Full Spectrum Inversion
GFS	Global Forecast System
GLONASS	GLObal NAVigation Satellite System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GRAS	GNSS Receiver for Atmospheric Sounding
GSI	Gridpoint Statistical Interpolation
ICVS	Integrated Calibration Validation System
IR	Infrared
JCSDA	Joint Center for Satellite Data Assimilation
JPL	Jet Propulsion Laboratory

L0	Level 0
L1a	Level 1a
L2	Level 2
LEO	Low Earth Orbit
LSW	Local Spectral Width
LWIR	Longwave Infrared
MAD	Mean Absolute Deviation
MW	Microwave
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
NWS	National Weather Service
POD	Precise Orbit Determination
QC	Quality Control
RAOB	Radiosonde Observation
RO	Radio Occultation
ROM SAF	Radio Occultation Meteorology Satellite Application Facility
ROPP	Radio Occultation Processing Package
SNR	Signal to Noise Ratio
STAR	Center for Satellite Applications and Research
TOE	True of Epoch
UCAR	University Corporation for Atmospheric Research