Data Release Notes for
Airborne/Balloon-borne Radio Occultation Data

Bing Cao, Jennifer S. Haase, Michael J. Murphy and Eric K. Wang

Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography,
University of California San Diego
La Jolla, CA

2021-06-25

Contents

1   Data Format Version Notes

- First version: 2019-04-19

- Second version: 2020-05-04

- Third version: 2020-09-30

- Fourth version: 2021-06-25

2   Use of the Data

The Strateole-2 data is freely and openly available to all users, with the requirement that the following acknowledgement is included in any use or publication of the data or derived products, and the following publications are cited:

"Development of the ROC instrument and data collection were supported by National Science Foundation award 1642650. The data were collected by Principal Investigator Dr. Jennifer S. Haase and her research team at Scripps Institution of Oceanography. Infrastructure for the balloon campaign was provided by NSF and the National Centre for Space Studies (CNES) and the National Centre for Scientific Research, Dynamic Meteorology Laboratory (LMD). CNES provided the balloons, flight control systems, flight control gondola 'Euros' and Flight Control Center. LMD provided the science payload gondola 'Zephyr' which carried the scientific instruments including ROC, as well as the science data transmission equipment, and provided the associated Mission Control Center (MCC) for monitoring and control of the instruments and the dataset collection."

Cao, B., J. S. Haase, M. J. Murphy, Alexander, M. Bramberger, A. Hertzog, and W. Zhang: Equatorial waves resolved by balloon-borne Global Navigation Satellite System Radio Occultation in the Strateole-2 Campaign, 2021

Haase, J. S., M. J. Alexander, A. Hertzog, L. Kalnajs, T. Deshler, S. M. Davis, R. Plougonven, P. Cocquerez, and S. Venel (2018), Around the world in 84 days - Strateole-2 investigates the tropical atmosphere with long duration superpressure balloons, EOS, 99.

The AR Recon Airborne Radio Occultation data is freely and openly available to all users, with the requirement that the following acknowledgement is included in any use or publication of the data or derived products, and the following publications are cited:

Haase, J. S., Murphy, M. J., Cao, B., Ralph, F. M., Zheng, M., and Monachel, L. D.: Multi-GNSS Airborne Radio Occultation Observations as a Complement to Dropsondes in Atmospheric River Reconnaissance, Journal of Geophysical Research: Atmospheres, in publication, 2021.

Ralph, F. M., F. Cannon, V. Tallapragada, C. A. Davis, J. D. Doyle, F. Pappenberger, A. Subramanian, A. M. Wilson, D. A. Lavers, C. A. Reynolds, J. S. Haase, L. Centurioni, B. Ingleby, J. J. Rutz, J. M. Cordeira, M. Zheng, C. Hecht, B. Kawzenuk, and L. Delle Monache (2020), West Coast Forecast Challenges and Development of Atmospheric River Reconnaissance, Bulletin of the American Meteorological Society, 101, E1357.

3   Acronym

- **GNSS** Global Navigation Satellite System

- **GPS** Global Positioning System

- **ARO/RO** Airborne Radio Occultation/Radio Occultation

- **GO** Geometric Optics

- **PM** Phase Matching

- **FSI** Full Spectrum Inversion

- **PPP** Precise Point Positioning

- **IMU** Inertial Measurement Unit

- **CDAAC** COSMIC Data Analysis and Archive Center

4   Version Control

Format Version (SSSS)
0001 = Murphy format for Predict data set
0002 = Format version of data release from 2019-04-19

0003 = Format version of data release from 2020-05-05

Software version (VVVV)

0001 = Brian Murphy, closed_loop, do_aro (provide date)

0011 = Jeff Sussmen, closed_loop, do_aro

0204 = Brian Murphy, open_loop, geometric optics (provide date)

0206 = Brian Murphy, open_loop, phase matching (2013-09-01)

0207 = Feiqin Xie, open_loop, full spectrum inversion (provide date)

0021 = Bing Cao of 2019-04-15 closed loop, do_aro, doppler correction, geometric optics

0022 = Bing Cao of 2020-05-04, closed loop, do_aro, doppler correction, geometric optics

0023 = Bing Cao of 2020-09-30, closed loop, pppar, doppler correction, geometric optics

0024 = Bing Cao of 2021-06-25, closed loop pppar, doppler correction, geometric optics

5   Data Releases

2021-06-25 Strateole2 balloon-borne RO data release for dates 2019.340-356, 2021.001, 2021.009-013 processing version V.0003.0024.

2020-09-30 Strateole2 balloon-borne RO data release for date 2019.346

processing version V.0003.0023

2020-05-05 Strateole2 balloon-borne RO preliminary data release for date 2019.346

processing version V.0002.0022

2019-04-15 ARRecon2018 airborne data release for dates 2018.026 to 2018.027

processing version V.0002.0021

6   Introduction

This release note is specifically for airborne and balloon-borne RO. In this context, "receivers" means GNSS receivers onboard either an aircraft or a balloon. The "satellites" are solely used in reference to GNSS satellites, not LEO satellites as is the case for other spaceborne RO.

This release note covers the basic information about how the meteorological parameters such as pressure and temperature, and refractivity profiles are retrieved from airborne and balloon-borne RO measurements. The data format of the final products in netCDF format are described in detail. The format follows conventions based on the standard CDAAC data format, but some changes are made to accomodate specific features of airborne and balloon-borne RO profiles.

6.1   Retrieval Procedures

The retrieval processes start from the high-accuracy aircraft/balloon trajectory determination. The trajectory is calculated using either the Applanix POS-Pac software (desired equipment used with or without IMU data) or the PANDA PPP software developed at Wuhan University GNSS Research Center. Then, using the raw GNSS carrier phase observation data, the receiver trajectory and the satellite orbit ephemerides, the excess phase/Doppler residuals are calculated using either our internal **do_aro** (algorithm version 0002) or **pppar** (algorithm version 0003) methods. The radius and center of local curvature are determined to approximate the ellipsoidal Earth surface so that the positions of satellites and receivers can be transformed to a locally spherically symmetric coordinate system. The excess Doppler observations are used to determine the bending angle in this shifted geometry using geometric optics or phase matching (algorithm version 0004 above). Then, positive and negative bending angles are calculated and partial bending angle is obtained. The inverse-Abel transform is used to derive the refractivity. Finally, meteorological parameters such as dry pressure and dry temperature, with water vapor neglected, are derived using the hydrostatic equilibrium assumption and ideal gas law. An alternate way of retrieving pressure, temperature and humidity is to use 1D-var method with model profiles as the first guess, however this has not yet been implemented.

Users should understand that dry temperature is not a good representation of actual temperature below 9 km altitude, and by preference should use the refractivity profiles for analysis. For data recovered from tropical area in Strateole2 campaign, the dry temperature deviates from true temperature from $\sim$12 km below which moisture is not negligible.

6.2   Horizontal Location of Drifting Tangent Points

The refractivity profile is derived using the Inverse-Abel transform for each tangent point height and corresponding impact parameter. Because of the interpolation of impact parameter during this process, an individual value of refractivity retrieved at certain tangent point altitude does not have an associated time of observation, nor does it have a horizontal location. However, the positions of the satellite and receiver, bending angle and impact parameter are all a function of time. With the known satellite and receiver positions and a climatological refractivity profile (CIRA-Q), we run a ray-tracing forward model (ROSAP) to simulate the bending angle at each time, and in the process establish the relationship between time and the tangent point location and height. In this way, we can assign a tangent point height and location and time to each retrieved value of refractivity in the level2 **atmPrf** file.

However, CDAAC explains that they do not mix time and height together as independent variables. "It was decided to leave these data (time information) in the atmPhs file as it is difficult to accurately and unambiguously assign a correspondence between time (the independent variable in the atmPhs files) and height(the independent variable in **atmPrf** files)." We point out a significant difference with the spaceborne RO in which it is possible to determine geometrically the horizontal location of the tangent point without establishing the time correspondence. Because we associate

time with each tangent point, can determine the relationship between spatial and temporal resolution. In this case, the resolution in altitude is set as 100m, and the corresponding temporal resolution is roughly 3–8 sec.

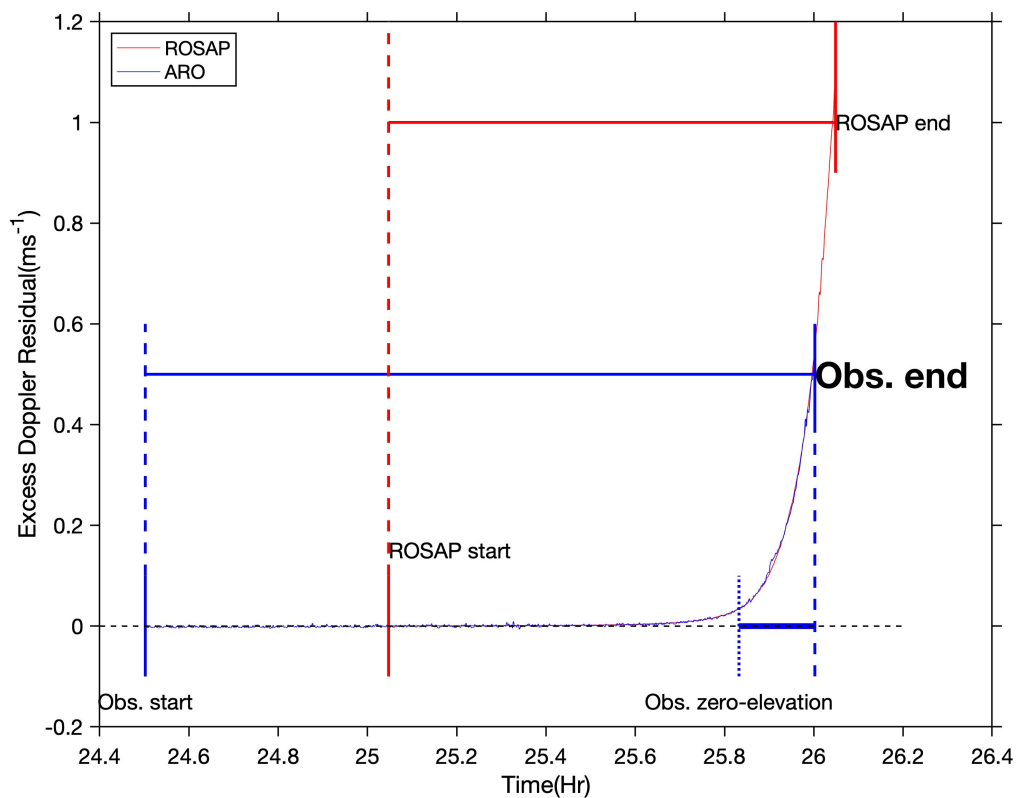## 6.3 Time Marks for an Occultation

Figure 1: Time marks of a typical setting occultation with corresponding ROSAP forward simulation results.

All the discussions in this section will be first described for the case of a typical SETTING occultation. Differences for the case of a RISING occultation will be described at the end. For a typical occultation, it is clear and straightforward to label an occultation in time using the ending time (**Obs.end** in Figure 1) when a GNSS receiver loses track of the signal, for example due to low SNR. This time mark is the unique for an occultation and defined in the atmPrf data files. In the current data processing algorithm, measurements beginning from 1 to 2 hours before the **Obs.end** are selected for processing and saved in a standard dop2alp_in format. This staring point shown as **Obs.start** in Figure 1 (**Obs.start** – **Obs.end** =1–2 hr). This **Obs.start** time mark has no physical meaning, however, it is used in the name of all intermediate files in the processing and saved as **old_fileStamp** in atmPrf files for the purpose of internal reference.

Another important time mark for an occultation is the time when the GNSS satellite reaches

zero-elevation angle with respect to the receiver. This time (**Obs. zero-elevation**) marks the transition from positive elevation angle bending to the negative elevation angle bending. At this time, the tangent point is at the location of the aircraft GNSS antenna and it is the starting of the slant profile where the tangent point starts to drift away and downward from the GNSS antenna. It is determined by calculating the straight-line elevation angle between GNSS satellite and receiver, without considering the bending of the ray path. The time period between **Obs.zero-elevation** and **Obs.end** properly and unambiguously labels this occultation in the time domain. In the final atmPrf products, these two time points are ulitized as **start_time(=Obs.zero-elevation)** and **end_time(=Obs.end)**. As stated in the previous section, the time is not an independent variable for atmPrf files but we interpolate the [**time, tangent_point_height**] relation determined in the forward ray tracing to specify the time at each of the tangent points. The interpolated time range closely matches the [**start_time, end_time**], but not exactly because the time of the actual zero-elevation angle is slightly different from the straight line **Obs.zero-elevation** due to ray path bending.

An oblate correction is necessary prior to the sphericlly symmetric Abel-transform step in the refractivity retrieval. The global attribute **occpt_offset** is introduced to label the time offset from the start of the occultation **(Obs.start)** to the time point at which the oblate correction is determined. In our dataset, the oblate correction is always applied when the tangent point reaches its minimum height, which corresponds to the end of the occultation **end_time(=Obs.end)**. So this **occpt_offset** is calcualted as **end_time – start_time**, which also corresponds to the duration of an occultation.

For a RISING occultation, the **Obs.start** is the time the receiver starts to track of GNSS signal. **Obs.end** is obtained by extending 1–2 hr from the **Obs.start**. In this case, the occultation is label as **start_time(=Obs.start)** and **end_time(=Obs.zero-elevation)**. The global attribute **occpt_offset** is still defined as **end_time – starting_time**.

7   "atmPrf" Data Format (netCDF and Ascii)

7.1   Directory Name Format

Strateole2 Directory Name Format:
strateole2/data/YYYY-MM-DD_CAID/data_release_YYYY-MM-DD/level2/atmPrf/YYYY.DOY

First **YYYY-MM-DD**	Start date of flight/campaign
CAID	Campaign ID or Balloon ID
Second **YYYY-MM-DD**	Date of data release
YYYY.DOY	Date of daily data subdirectory

Example: strateole2/data/2019-12-05_str2/data_release_2020-09-30/level2/atmPrf/2019.346

AR Recon. Mission Directory Name Format:
airborne_ro/data/YYYY.DOY_CAID/data_release_YYYY-MM-DD/level2/atmPrf/YYYY.DOY_FLID

Example: airborne_ro/2018.023_ar2018/data_release_2019-04-15/level2/atmPrf/2018.027_rf01

France ISPL Data Sever Directory Name Normat:
/DATA/CampaignID/InstrID/ProjID_CampaignID_BallonID_ConfigID_InstrID_YYYYDOY_vYYYYMMDD
Example: /DATA/C0/ROC/ST2_C0_06_STR1_ROC_2019346_v20210625
Notes: YYYYDOY is the flight date, vYYYYMMDD is the data release date.

7.2 File Name Format

File Name Template:
/atmPrf/YYYY.DDD_RFID/atmPrf_IIII.YYYY.DDD.HH.MM.GXXS_GYYH.SSSS.VVVV.nc

IIII	Receiver/Site ID (SEPT, ROC2, N49T, STR2)
YYYY	Year
DDD	Day of year
RFID	Research Flight or IOP ID if any
HH	Hour
MM	Minute
GXX	Low elevation GNSS satellite PRN
	G=GPS, E=Galileo, R=GLONASS
S/R	S=Setting or R=Rising occultation
GYYH	High elevation GNSS satellite PRN.
SSSS	Data format version
VVVV	Software algorithm version

7.3 Global Attributes

1. **year, month, day, hour, minute, second**

   - **Description:** Date/time of the occultation
   - **Date Type:** double(8)
   - **Units:** NaN

2. **GPS_week**

   - **Description:** GPS week of the occultation
   - **Date Type:** double(8)
   - **Units:** NaN

3. **start_time**

   - **Description:** Starting time for the occultation

- **Date Type:** double(8)
- **Units:** GPS epoch in second

4. **end_time**

   - **Description:** Ending time for the occultation
   - **Date Type:** double(8)
   - **Units:** GPS epoch in second

5. **curv**

   - **Description:** The X, Y and Z offset of the center of sphericity used in the inversion from the ECEF Earth center
   - **Date Type:** 3*double(8)
   - **Units:** km
   - **Valid Range:** -200, 200

6. **occpt_offset**

   - **Description:** The time offset from the start of the occultation to the lowest occultation point, at which **curv** is measured
   - **Date Type:** double(8)
   - **Units:** sec
   - **Valid Range:** -1000, 1000

7. **rgeoid**

   - **Description:** The height of the geoid ABOVE the reference ellipsoid at the lat/lon position of the lowest occultation point. MSL altitude = Ellip altitude – geoid height.
   - **Date Type:** double(8)
   - **Units:** km
   - **Valid Range:** -150, 150

8. **flict**

   - **Description:** Local curvature radius of the reference ellipsoid for the occultation point
   - **Date Type:** double(8)
   - **Units:** km
   - **Valid Range:** 6200, 6600

9. **lat**

- **Description:** Latitude of perigee point at occultation point
- **Date Type:** double(8)
- **Units:** decimal degrees
- **Valid Range:** -90, 90

10. **lon**

- **Description:** Longitude of perigee point at occultation point
- **Date Type:** double(8)
- **Units:** decimal degrees
- **Valid Range:** -180, 180

11. **azim**

- **Description:** Azimuth angle of the occultation plane at tangent point measured clockwise with respect to North at 0 degrees. The angle is measured between North and the direction of the raypath from aircraft/balloon antenna towards GNSS satellite.
- **Date Type:** double(8)
- **Units:** decimal degrees
- **Valid Range:** -180, 180

12. **fileStamp**

- **Description:** The complete ID for this occultation (IIII.YYYY.DDD.HH.MM.GXXS_GYYH).
- **Date Type:** char

13. **old_fileStamp**

- **Description:** The internal reference ID for this occultation (prn_gXXs_gXX_secofweek).
- **Date Type:** char
- **Note:** This is used in all the intermediate files during the data processing.

14. **bad**

- **Description:** Badness flag. 1 = Profile flunked quality control, 0 = Profile passed QC
- **Date Type:** int

15. **errstr**

- **Description:** Error string. If bad = 1, contains the description of the problem found, else contains the null string
- **Date Type:** string

7.4   Profile Data (12 variables)

1. **Time**

   - **Description:** Corresponding time of each ray path
   - **Date Type:** float
   - **Units:** GPS second of week
   - **Valid Range:** 0, 999999
   - **Missing Value:** -999
   - **Note:** See previous section for the explaination of the time.

2. **MSL_alt**

   - **Description:** Mean sea level altitude of perigee point
   - **Date Type:** float
   - **Units:** km
   - **Valid Range:** 0, 9999
   - **Missing Value:** -999

3. **Ellip_alt**

   - **Description:** Ellipsoidal altitude of perigee point above WGS-84
   - **Date Type:** float
   - **Units:** km
   - **Valid Range:** 0, 9999
   - **Missing Value:** -999

4. **Ref**

   - **Description:** Refractivity
   - **Date Type:** float
   - **Units:** N
   - **Valid Range:** 0, 9999
   - **Missing Value:** -999

5. **Pres**

   - **Description:** Dry pressure
   - **Date Type:** float
   - **Units:** mb or hPa

- **Valid Range:** 0, 1500

- **Missing Value:** -999

6. **Temp**

   - **Description:** Dry temperature

   - **Date Type:** float

   - **Units:** C

   - **Valid Range:** 0, 9999

   - **Missing Value:** -999

7. **Bend_ang**

   - **Description:** Partial bending angle = negative elevation angle bending angle – positive elevation angle bending angle

   - **Date Type:** float

   - **Units:** rad

   - **Valid Range:** 0, π

   - **Missing Value:** -999

8. **Opt_bend_ang**

   - **Description:** Top-fixed/optimized partial bending angle

   - **Date Type:** float

   - **Units:** rad

   - **Valid Range:** 0, π

   - **Missing Value:** -999

   - **Notes:** At high altitudes, the retrieved bending angle is too noisy to be useful. To reduce the propagation of this noise into retrieved refractivity at lower altitudes in the inverse abel calculation, the retrieved bending angle is replaced for the section of profile around the maximum impact parameter. The noisy section of the bending angle profile is replaced with smoother values of bending calculated using a smooth curve fitted to the first (noisy) bending angle profile. A new refractivity profile is then calculated from this optimized bending angle profile. This version will have the entire positive elevation section, which should have very little variation in bending angle, also replaced with a smooth fit of original profile.

9. **Impact_para**

   - **Description:** Impact parameter

- **Date Type:** float
- **Units:** km
- **Valid Range:** 0, 9999
- **Missing Value:** -999

10. **Lat**

- **Description:** Latitude of perigee point
- **Date Type:** float
- **Units:** degree
- **Valid Range:** -90, 90
- **Missing Value:** -999

11. **Lon**

- **Description:** Longitude of perigee point
- **Date Type:** float
- **Units:** degree
- **Valid Range:** -180, 180
- **Missing Value:** -999

12. **Azim**

- **Description:** zimuth angle of the occultation plane at tangent point measured clockwise with respect to North at 0 degrees. The angle is measured between North and the direction of the raypath from aircraft antenna towards GNSS satellite.
- **Date Type:** float
- **Units:** degree
- **Valid Range:** -180, 180
- **Missing Value:** -999

8   Known Issues

1. Different methods of calculating geodetic coordinates (latitude, longitude, altitude) from Cartesian coordinates (ECEF_xyz). Some corrections are needed to compensate the difference. Matlab uses a function **ecef2geodetic** while ROSAP model uses a simplified method to do the conversion.

2. Consistent precision of the measurements, especially the coordinates.

3. Definition of the time, zero-elevation time is not a good option but works sufficiently well.

9   Contacts

If any errors are noticed in the data, please contact Bing Cao (bic020@ucsd.edu) and Jennifer Haase (jhaase@ucsd.edu).