A software development to retrieve, process, store and visualize satellite data from CDAAC

Telmo S. Klipp [1], Adriano Petry and José V. Bageston

Southern Regional Space Research Center, National Institute for Space Research

Santa Maria, RS, Brazil

Received on January 01, 2015 / accepted on *****, 2015

Abstract

The use of observational data to validate and improve computational simulation of physical processes is an important issue to be addressed. Usually data may use file formats different from the one desired, and frequently is available through graphical interfaces. This is specially true for data obtained from COSMIC Data Analysis and Archive Center (CDAAC), where data products are available for a number of satellite missions, and provide scientific data access in a variety of geophysical domains. This work presents the development of a software tool to easily login, download, filter, process and save data from CDAAC in text format. Besides that, the measurements available after filtering are shown in a geographic map, and the data values (e.g., temperature) for a pre-defined geographic region and time interval are recorded in text format along other relevant information. This tool can help researchers to get quick access to atmospheric data without applying effort on software implementation, which enhance and speed up data acquisition process, especially when a large amount of data is required. Currently, our software can provide processed data in text format from 15 radio occultation satellite missions and 10 different data files structures, and filtering analysis is based on desired date and time over an geographic region of interest. Successful tests were performed for temperature profiles. Comparisons of the temperature obtained from GPS Radio Occultation (GPS-RO), TIMED/SABER satellite and Radiosondes near Santa Maria/RS, Brazil, from April 29 to May 2^{nd}, 2016, showed good agreement. GPS-RO provided temperature values with a good accuracy (<0.5 K) and high vertical resolution (100 m) as deep in the atmosphere as near the Planetary Boundary Layer (PBL), i.e., around 2.5 km. The developed software proved to be useful to provide easy access to important information on dynamical processes on a variety of geophysical domains.

Keywords: CDAAC, Radio Occultation, atmospheric data, data acquisition, data processing.

[1]E-mail Corresponding Author: voidtsk@gmail.com

1. Introduction

An important step in the scientific research of physical processes based on computer models is to obtain observational data to populate simulations and to analyse, improve and validate the models. The Earth's atmosphere is an environment where several physical processes and dynamical phenomena take place, drawing attention of many research fields. Aligned to this, remarkable efforts have been made over the past two decades to explore Earth's atmosphere employing space-based atmospheric remote sensing. The GPS-RO limb sounding technique proved to be a reliable source of observational data for global atmosphere including accurate and precise profiles of electron density in the ionosphere and information on temperature and water vapor in the stratosphere and troposphere [1]. The first GPS-RO satellite mission (GPS/MET) was launched on April 3rd 1995 for the proof-of-concept that began a revolution in profiling Earth's atmosphere [1] [4]. Since then, several satellite missions have been launched and concluded, establishing periods of activity among the missions and contributing to a wide range of scientific investigation in fields such as meteorology, climate, ionosphere and geodesy/gravity [2]. As a claim of interest, data products derived from satellite soundings (accomplished and active missions) have been made available for the research community through different data files structures by COSMIC Data Analysis and Archive Center (CDAAC) [10].

Classical atmospheric parameters (e.g., temperature, dry air and atmospheric pressure) are derived from the refractive index profile by a process detailed in [6]. The refractive index comes from soundings performed by GPS-RO limb technique, which consists in measuring the signal phase delays of radio waves at L1 and L2 frequencies from GPS satellites as they are occulted by Earth's atmosphere and captured by micro-satellites receivers positioned at Low Earth Orbit (LEO) [2] (see Figure 1).

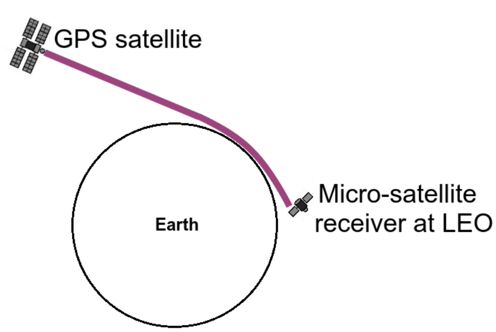

Figure 1 - The GPS-RO limb technique illustration. Adapted from [1].

Although it is a relevant data source, the CDAAC mainly provides profiles in NetCDF format [8], which requires the use of a software tool or a specific software implementation for reading these files and extracting information. Data transfer process also needs to be addressed. Some of the features that could be desired in a software implementation to obtain and process data from CDAAC are: (i) Search and download of available data files, considering desired satellite missions and period of time; (ii) Handle time sequence between data files, considering date and time collection of stored data; (iii) Perform the reading and extraction of data from data files; (iv) Process and manage the data (e.g., considering a closed geographic region or specific atmospheric parameters), making it available for other applications or write it on output files for later use.

The implementation of these requirements can comprise different technologies, which has to provide acquisition and control of large amount of data. Regarding such requirements, this work presents a software development to retrieve, process, store and visualize satellite data from CDAAC.

2. CDAAC Data Source Overview

CDAAC allows the download of different geophysical data file structures (profiles), encompassing atmospheric/ionospheric data from all RO missions since 1995 [10]. It access is granted under a security policy implemented by University Corporation for Atmospheric Research (UCAR) members and the main profile acquisition forms are: (i) graphical user interfaces (GUIs) at CDAAC web site, (ii) ftp and web servers or (iii) Local Data Manager (LDM) [9]. Software implementation expertise is a requirement to get data from ftp and web servers and, server configuration to get data from LDM. Direct download of profiles is achieved through GUIs, however, getting profiles manually has an impact on the data acquisition time, especially when large amount of data is needed.

CDAAC data availability and download methods are also relevant. Currently LDM - a real time event-driven data exchange system developed by Unidata community (one of UCAR's Community Programs) that allows researchers and research institutions to share data almost independent of data unit structure [3] - can provide only few profiles from CDAAC. Although providing more profiles, CDAAC ftp server grants access to only two missions over the fifteen available. One of them is cosmicrt, which is currently operational. The CDAAC web server can provide data from all fifteen available missions and download process is accomplished via wget utility. Instead of ftp requisitions that gets individual profiles, wget requests downloads compacted files, each one having daily generated profiles for a given mission and profile type. Also, download via web server is faster than ftp server.

Profiles are classified on six main data divisions: raw GPS data, orbit determination, tiny ionospheric photometer, atmospheric profiles, scintillations, total electron content and ionospheric profiles. Another classification focus on data maturity and takes into account the current status of data, that can be: real time, post-processed or re-processed data [11]. Real time data is provided by active missions and is recommended to be used on weather forecasting and space weather monitoring applications while re-processed and post-processed data is recommended to be used on climate and atmospheric research. The developed CDAAC data acquisition program (CDAAC-DAP) provides atmospheric, total electron content and ionospheric data profiles. Table 1 shows the current status of CDAAC for missions, their period of activity and profile types, considering CDAAC-DAP processing support.

3. Software Development

CDAAC-DAP was written in C++ and uses curl library and wget tool for files exchange, boost-posix library to deal with date/time comparisons [7], netCDF4 C++ API [5] to read profiles, and Grads software to generate geographic maps for available data. The CDAAC-DAP execute five main functionalities: dowload of profiles, control and management, read/extract of information, write text data and generation of maps. Figure 2 shows the CDAAC-DAP software architecture.

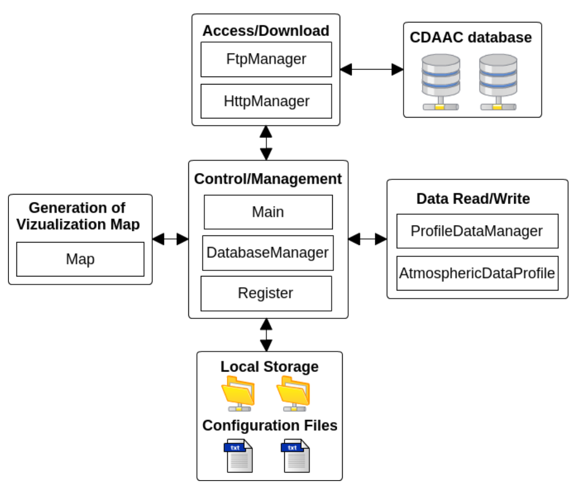

Figure 2 - CDAAC-DAP software architecture. Each CDAAC-DAP functionality is related to specific classes.

Table 1: CDAAC missions, their period of activity and profile types.

Mission	Profile type	Period of Activity
champ2014	atmPrf, ecmPrf, gfsPrf, ionPrf, podTec, sonPrf, wetPrf	2001.138 to 2008.279
cnofs	atmPrf, eraPrf, wetPrf	2010.060 to 2011.365
cnofsrt	atmPrf, avnPrf, wetPrf	2012.001 to 2015.193
cosmic2013	atmPrf, ecmPrf, eraPrf, echPrf, gfsPrf, ionPrf, podTec, sonPrf, wetPrf	2006.112 to 2014.120
cosmic	atmPrf, ecmPrf, eraPrf, echPrf, gfsPrf, ionPrf, podTec, sonPrf, wetPrf	2014.121 to 2016.060
cosmicrt	atmPrf, avnPrf, ionPrf, podTec, wetPrf	2014.181 in activity
gpsmet	atmPrf, ecmPrf, sonPrf, wetPrf	1995.111 to 1997.047
gpsmetas	atmPrf, ecmPrf, sonPrf, wetPrf	1995.237 to 1997.016
grace	atmPrf, ecmPrf, eraPrf, gfsPrf, ionPrf, podTec, sonPrf, wetPrf	2007.059 to 2016.030
kompsat5rt	atmPrf, avnPrf, wetPrf	2015.305 in activity
metopa2016	atmPrf, eraPrf, gfsPrf, podTec, sonPrf, wetPrf	2007.274 to 2015.365
metopb	atmPrf, eraPrf, gfsPrf, podTec, sonPrf, wetPrf	2013.032 to 2015.059
sacc	atmPrf, ecmPrf, eraPrf, gfsPrf, , podTec, sonPrf, wetPrf	2006.068 to 2011.215
saccrt	atmPrf, avnPrf, wetPrf	2011.329 to 2013.226
tsx	atmPrf, ecmPrf, eraPrf, gfsPrf, podTec, sonPrf, wetPrf	2008.041 to 2015.333

3.1 Download and Management of Profiles

Downloading data from CDAAC implies getting profiles of a specific day, mission and profile type. Profiles have in data body classical atmospheric parameters according to CDAAC data classification. CDAAC-DAP can provide processed data from 15 missions and 10 different profile structures (see Table 1). Ftp server gives access to only two missions, while web server grants access to all available missions. The downloading process is executed by FtpManager and HttpManager classes. The former sends an initial profile names request to the ftp server looking at a specific day, mission and profile type and, subsequently, performs individual download requests to each profile name returned. The later, instead, sends directly download requests to the web server for compacted tar files, each one having daily profiles from a given mission and profile type.

Downloaded profiles are stored in the input file folder (InFiles). CDAAC-DAP always performs local search of downloaded profiles and recognition of profiles date and time. So, if a file is found locally it is not downloaded again, while date and time recognition ensures that data processing can be constrained to a time period.

3.2 Data reading and writing

Profiles follow the Network Common Data Form (netCDF) structure, which is a collection of data format libraries with the purpose of data description and storage, apart from operational system and platform. CDAAC-DAP uses the netCDF4 C++ library [5] to read the data that is later written in output files in text format. Each output file has filtered data extracted from a single profile. The read, process and write functions are executed by ProfileDataManager and AtmosphericDataProfile classes.

ProfileDataManager extracts desired data values and support information from profiles, processing one at time. AtmosphericDataProfile manages the downloaded profiles list passing it to ProfileDataManager. Then, extracted data is written to output files, which contains: relevant header information (e.g., satellite identification, data content, data collection time, measurements units and profile origin), variables names and data values. The output files from each user requisition are stored in individual directories inside the output folder (OutFiles), to keep data organized.

3.3 Control and Management

Control and management deals with time operations and data requests considering metadata text files and a configuration file. Metadata files keep knowledge of CDAAC database such as available missions, its period of activity and profiles. Also, maintains knowledge of data variables that each profile contains. Metadata files are unmodifiable, unless new missions and data variables are created or old ones modified. A specific configuration file determines filtering options to data search and process such as time interval to search for files, data variables to be extracted and user-defined region on global map. The classes Main, DatabaseManager and Register are responsible for control and management functionalities.

Main class manages user requests, which are passed by command line parameters that specify mission and profile type. It also controls the CDAAC-DAP execution flow, instantiating other objects and necessary system calls. DatabaseManager verify the consistency of user requests through its registered lists of missions, period of activity and profiles types. Register class performs date and time management, applying libraries [7] to provide: date and time operations (subtraction, addition and time comparisons), time period determination and output information.

3.4 Map Generation

The measurements available after filtering are shown in a geographic map, providing a quick visual verification of density and geographic location of observations. A script based on Grid Analysis and Display System (GrADS) language was used to generate a 2D map with the coordinates (latitude/longitude pairs) where data is available. Figure 3 shows the locations of available temperature observations from latitude -30°S to 50°N and longitude -40°W to 40°E for May 5^{th}, 2014, with time interval between 13:00 and 17:00 UT.

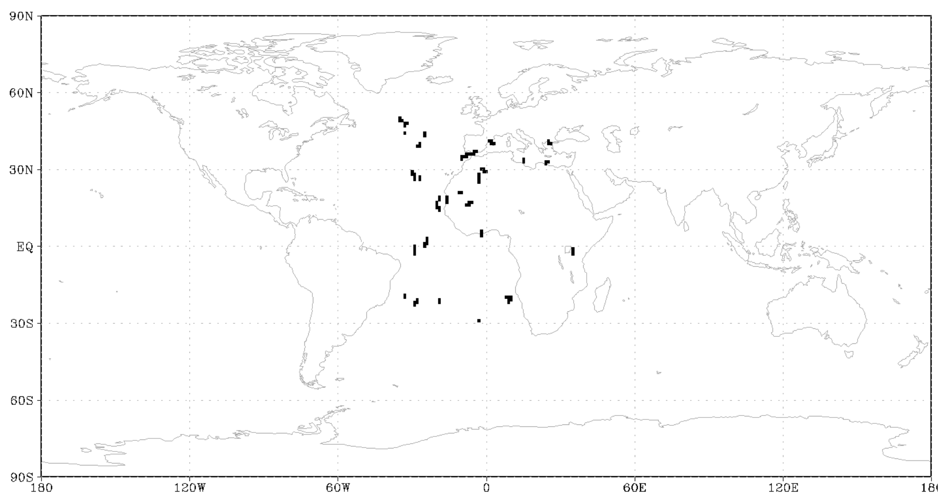

Figure 3 - Locations of available temperature observations.

4 Case Study

Successful tests were performed for atmospheric temperature profiles. Comparisons of the temperature profiles obtained from GPS-RO technique with weather balloon and TIMED/SABER satellite soundings over the central region of Rio Grande do Sul, Brazil, between April, 29 and May 2^{nd}, 2016 were performed. Figure 4 shows temperature profiles obtained by three different techniques, that is, GPS-Radio Occultation (GPS-RO), Meteorological Balloons (Radiosondes) and TIMED/SABER satellite since near the ground up to 80 km height. The time in the legend is in Universal Time (UT:

LT+3 hs) and each sounding location is identified by its respective latitude and longitude (" $-$ ", i.e., negative, means South and West in lat. and lon., respectively). April 30, 2016 was chosen for this study based on the criteria of time and space coincidence, that is, the GPS-RO profile should reach as down as 5 km or less and so high as 60 km height in order to compare with both radiosonde and SABER profiles, and the time of the GPS-RO profile should be as close as possible of the radiosonde measurements. Also, the distance of the satellite profile should be as near as possible to the place of the balloon campaign.

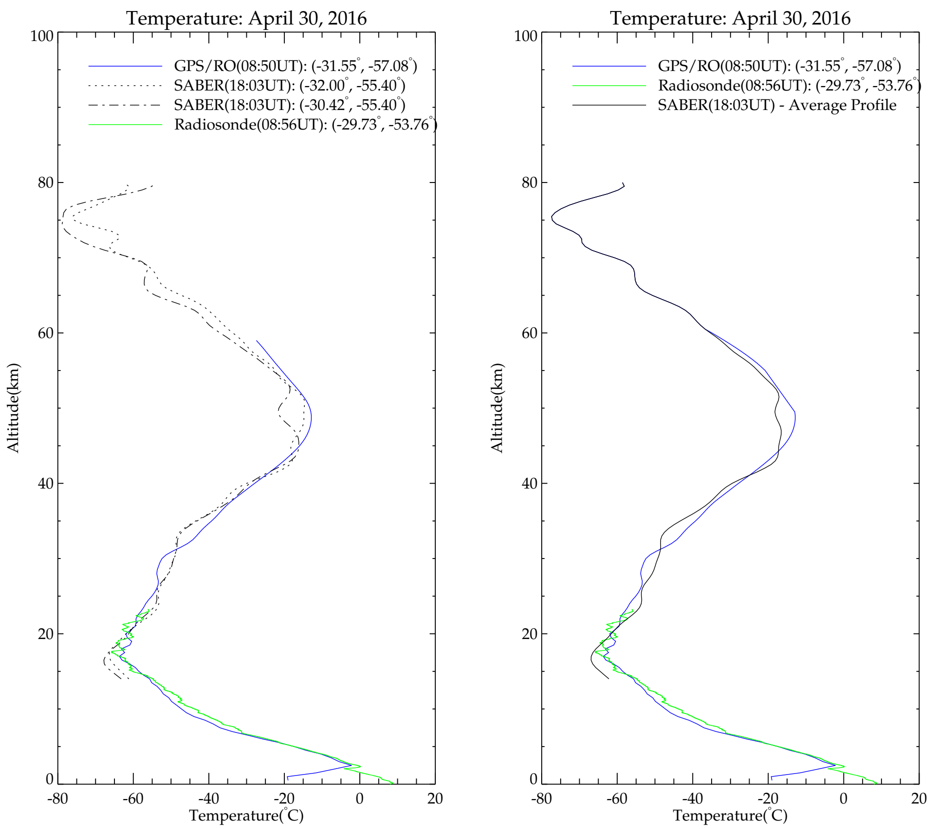

Figure 4 - Temperature profiles obtained by GPS-RO technique for comparison to the Radiosonde and SABER instruments.

The first good coincidence in the chosen GPS-RO sounding profile is that the local time matched very well with the radiosonde balloon preparation and launch, which started 08:15 UT(05:15 UT) and the launch was at 08:

56 UT. However, for this case the geographical distance between the two sounding points was a little large, i.e., ~376 km (the distribution of the sounding points on the map is not shown here), but close enough to see the very good matching of the two profiles. Actually, there was a good agreement between the GPS-RO (blue) and Radiosonde (green) temperature profiles, mainly above 2.5 km height. Bellow this altitude the profiles clearly diverge from one other, and this can be explained by the water vapor near the ground (the balloon was released at around 06:00 LT) that can affect the GPS-RO measurements [12].

5. Conclusions

CDAAC-DAP can be used to download data from 15 missions and 10 different data files. Processed data is written to text output files along relevant information about data identification and origin, with data to be extracted being specified by user through parameters on configuration file. Available data is shown on a geographic map by plotting coordinates where observations occurs, allowing the user a quick verification of existing values. This 2D shape view ignores the altitude profile, however facilitates the verification of data density for a given region and time interval. CDAAC-DAP can help researchers to get easy access to atmospheric data, which enhance and speed up data acquisition process, especially when large amount of data is needed. Successful tests were performed for atmospheric profiles of temperature, where different temperature data sources were compared. The developed software allows researchers to focus on their core problem, instead of data acquisition requirements, software coding or server configuration. Future works may consider the extension of the software capabilities to different databases, and include data from CDAAC on ongoing projects related with ionospheric behaviour modelling and Troposphere/Stratosphere gravity wave dynamics.

References

[1] Anthes, R. A. Exploring Earths atmosphere with radio occultation: Contributions to weather, climate and space weather. Atmos. Meas. Tech, 4: 1077-1103, 2011.

[2] Anthes, R. A., et al. The COSMIC/FORMOSAT-3 mission: Early results. Bull. Am. Meteorol. Soc., 89: 313333, 2008.

[3] Fulker, D., Bates, S., & Jacobs, C. Unidata: A virtual community sharing resources via technological infrastructure. Bull. Amer. Meteor. Soc., 78: 457468, 1997.

[4] Hajj, G. A., Kursinski, E. R., Romans, L. J., Bertiger, W. I., & Leroy, S. S. A technical description of atmospheric sounding by GPS occultation. J. Atmos. Sol.-Terr. Phys., 64(4): 451-469, 2002.

[5] Hartnett, E., & Rew, R.K. Experience with an enhanced netCDF data model and interface for scientific data access. In 24th Conference on IIPS, 2008.

[6] Kursinski, E. R., Hajj, G. A., Schofield, J. T., Linfield, R. P.,& Hardy, K. R. Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. J. Geophys. Res. Atmos., 102: 23429–23465, 1997.

[7] Mukherjee, A. *Learning Boost C++ Libraries*. Packt Publishing Ltd, Birmingham, UK, 2015, 311-334.

[8] Rew, R., & Davis, G. NetCDF: an interface for scientific data access. IEEE Comput. Graph. Appl., 10(4): 76-82, 1990.

[9] Rew, R., & Wilson, A. The Unidata LDM system: Recent improvements for scalability. In Preprints, 17th Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, Albuquerque, NM, Amer. Meteor. Soc. Vol. 4, 2001.

[10] Schreiner, B., Hunt, D., Rocken, C., & Sokolovskiy, S. Radio occultation data processing at the COSMIC Data Analysis and Archival Center (CDAAC). *First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies*. Springer Berlin Heidelberg, 2003, 536-544.

[11] Schreiner, B., Hunt, D., Rocken, C., & Sokolovskiy, S. Approach and Quality Assessment of Precise GPS Data Processing at the UCAR CDAAC. ROCSAT-3/COSMIC Radio Occultation Science Workshop II, Taipei, TPQ, TW, 2004, pp. 5. Available at: `http://nldr.library.ucar.edu/repository/assets/osgc/OSGC-000-000-001-427.pdf`

[12] Wickert, J. and Reigber, C., Beyerle, G., Konig, R., Marquardt, C., Schmidt, T., Grunwaldt, L., Galas, R., Meehan, T. K., Melbourne, W. G., & Hocke, K. Atmosphere sounding by GPS radio occultation: First results from CHAMP. Geophys. Res. Lett., 28(17): 3263-3266, 2001.