

PACE-PAX Field Data Repository dataID Registration and Data Upload

**Michael Shook, Gao Chen, Ali Aknan, and
Morgan Silverman**

Introduction

- <https://www-air.larc.nasa.gov/missions/pacepax/index.html>
- A repository for all PACE-PAX relevant observational and ancillary data products and documentation/reports to facilitate science team data exchange and data processing
- File types include binary (HDF/netCDF), text (ICARTT), and image (jpg/png) files
- Password-protected document and file sharing
- Publication quality (final) data will be transferred to ASDC

Airborne Science Data for Atmospheric Composition

HOME TOOLS MISSIONS DATA CONTACT US

Welcome to the NASA LaRC Suborbital Science Data for Atmospheric Composition

Promoting a better understanding of air quality and climate change issues by archiving, managing, and sharing tropospheric composition data from airborne field studies.

Current / Upcoming Missions

- FireSense 2023-2028
- PACE-PAX 2024
- TEMPO Validation Group
- ASIA-AQ
- STAQS 2021
- DCOTSS
- ACTIVATE 2022 - 2020
- CPEX-CV 2022
- ACCLIP 2022
- SARP 2023 - 2009
- CAMP2Ex/PISTON
- SASA 2022
- FIREX-AQ 2019
- KORUS-AQ
- TOLNet

MEASURES Projects

Creating a Unified Airborne Database for Model Assessment

TabMOP

An Interagency Interest Group on Applying Aerosol Observations to Earth System Models

Project Description

Interactive map (1997 - 2016)

PACE-PAX - Plankton, Aerosol, Cloud, ocean Ecosystem - Postlaunch Airborne eXperiment

Recent Activities

Data Archive: PACE-PAX 2024

File Sharing (private): Telecons, Meetings, Reports, etc.

Relevant Data / Links

- GMAO FLUID Mission Support
- NASA ARC Meteorological Support

Data Upload Tools

- Steps for submitting data to the Archive
- Data Submittal / Scanning
- Register PI dataIDs
- ICARTT Data Format Document

Useful Tools

- Download HDFView - visual tool for browsing & editing HDF files
- Download File Scanning S/W for Windows
- Download Flight Planning S/W for Windows

PACE-PAX

The PACE-PAX mission will be conducted in September 2024, deploying two aircraft, each flying out of their home base.

CIRPAS Twin Otter (TO)

NASA ER-2


CIRPAS Twin Otter (Marina, CA) Direct (in situ) aerosol, cloud measurements


NASA ER-2 (Palmdale, CA) Remote and PACE Proxy measurements

PACE-PAX Field Data Repository

PACE-PAX – Plankton, Aerosol, Cloud, ocean Ecosystem - Postlaunch Airborne eXperiment

Data Access

[Data Archive: PACE-PAX 2024](#) 

[File Sharing \[private\]](#) 
Telecons, Meetings, Reports, etc.

Relevant Data / Links

[GMAO FLUID Mission Support](#) 

[NASA ARC Meteorological Support](#) 

Data Upload Tools

[Steps for submitting data to the Archive](#)

[Data Submittal / Scanning](#) 


[» Help FScan](#)

[Register PI dataIDs](#) 

[ICARTT Data Format Document](#)

Useful Tools

[Download HDFView -- visual tool for browsing & editing HDF files](#) 

[Download FileScanning S/W for Windows](#) (Requires IE) 
[» What's New](#)

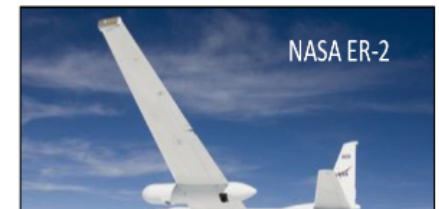
[Download Flight Planning S/W for Windows](#) (Requires Google Earth) 

Recent Activities

The Plankton, Aerosol, Cloud, ocean Ecosystem Postlaunch Airborne eXperiment (PACE-PAX) is a field campaign to support validation of the [PACE mission](#).



The PACE-PAX mission will be conducted in September 2024, deploying two aircraft, each flying out of their home base.



NASA ER-2

Document Upload/Download

dataID Registration and Data Upload

Points of Contact

- **Field Repository (www-air.larc.nasa.gov)**
 - Michael Shook, NASA Langley Research Center, michael.a.shook@nasa.gov, 757-864-5793
 - Gao Chen, NASA Langley Research Center, gao.chen@nasa.gov, 757-759-5642 (cell)
 - Ali Aknan, AMA/NASA Langley Research Center, ali.a.aknan@nasa.gov (website and file scanner)
 - Morgan Silverman, AMA/NASA Langley Research Center, morgan.l.silverman@nasa.gov (standard name issues)

Please free feel to reach out for data download, upload, format, and ACVSNC standard name issues

Data Submission Steps

- DataID Registration (one-time process):
 - dataID is part of the filename and will be used to organize PI files on the data repository, i.e., Linking data files to PI (see file naming convention slide)
 - PI or file creator will need to first register dataID(s) before files can be submitted
- Data Submittal and Scanning:
 - File submission is through a scanning tool (FScan) for checking filenames and content
 - All incoming files are scanned:
 - ICARTT Files: file header, including keywords and data flags, as well as time stamps
 - HDF and netCDF files: data variable dimensions and attributes
 - Other files: file names
 - Support zipped multi-file upload
 - Script-based batch upload and download available
- Username/password: contact POCs

The screenshot displays a web interface for a data archive. At the top, there is a link for 'Data Archive: PACE-PAX 2024' with an ArcView icon. Below this is a 'File Sharing [private]' section with a lock icon and the text 'Telecons, Meetings, Reports, etc.'. The main content area is divided into sections: 'Relevant Data / Links' with links for 'GMAO FLUID Mission Support' and 'NASA ARC Meteorological Support'; 'Data Upload Tools' which is circled in red and contains links for 'Steps for submitting data to the Archive', 'Data Submittal / Scanning' (with an ICARTT logo and a 'Help FScan' link), and 'Register PI dataIDs' (with a document icon); and 'ICARTT Data Format Document'. The 'Useful Tools' section at the bottom lists links for downloading 'HDFView', 'FileScanning S/W for Windows' (with an 'What's New' link), and 'Flight Planning S/W for Windows'.


DataID Registration Page

See file naming convention slide

The PI's data directory will be created from LastName.FirstName. Please enter PI name correctly. Each dataID represents a [separate] group of files in PI's data directory.

PI Last Name :	PI First Name :	Platform (<u>LocationID</u>) :
<input type="text"/>	<input type="text"/>	NASA ER-2 Aircraft (ER2)

dataID: (max 45 chars) <i>Prefix with "pacepax-"</i> e.g., pacepax-CO2	Data Description: (max 380 chars) <i>Describe your measurements; e.g., Carbon Dioxide Mixing Ratio</i>	Instrument(s): (max 190 chars) <i>List Instruments; e.g., LI-COR 6252</i>
<input type="button" value="Reset"/> <input type="text" value="pacepax-"/>	<input type="text" value="Enter one description per line AND press Enter"/>	<input type="text" value="Enter one instrument per line"/>
<input type="button" value="+ Add Another dataID"/>	<input type="button" value="- Remove Last Entry"/>	

Upload Your Instrument(s) Description Document: No file chosen (Select your file) 

*****(NOTE: Any new file/document upload WILL OVERWRITE your previously uploaded document)*****
*****(If you have more than 1 file (document) to upload, please email the files to: gao.chen@nasa.gov and/or ali.a.aknan@nasa.gov. Thank you!)*****

OVERWRITE my previous record (i.e., ALL previously registered dataIDs for this mission will be removed).


Link to PI website, instrument, experiment description, etc
 Optional: to display on LaRC Archive webpage

Text describing PI experiment or measurements (e.g., NASA LaRC DIAL - Troposphere O3, Aerosols, and Clouds Profiles):
 Optional: to display on LaRC Archive webpage

Registered DataID Example

PI Last Name :	PI First Name :	Platform (LocationID) :
<input type="text" value="HAIR"/>	<input type="text" value="JOHN"/>	<input type="text" value="LaRC GIII Aircraft (LARC-G3)"/>

	dataID: (max 45 chars) <i>Prefix with "asiaaq-" e.g., asiaaq-CO2</i>	Data Description: (max 380 chars) <i>Describe your measurements; e.g., Carbon Dioxide Mixing Ratio</i>	Instrument(s): (max 190 chars) <i>List Instruments; e.g., LI-COR 6252</i>
<input type="button" value="Reset"/>	<input type="text" value="asiaaq-HSRL2"/>	HSRL measurements of clouds and aerosols, as well as DIAL measurements of Ozone Concentration	<input type="text" value="HSRL2"/>
<input type="button" value="Reset"/>	<input type="text" value="asiaaq-HSRL2-images"/>	Full flight images of the lidar measurements of clouds, aerosols, and ozone, plus raster-specific images within a flight, designated as Ln (n can be 1,2,3,or 4)	<input type="text" value="HSRL2"/>
<input type="button" value="Reset"/>	<input type="text" value="asiaaq-HSRL2-mlh"/>	Aerosol derived mixed layer heights and mean backscatter and extinction within several layers from HSRL2	<input type="text" value="HSRL2"/>
<input type="button" value="Reset"/>	<input type="text" value="asiaaq-HSRL2-NearSurface"/>	High horizontal resolution measurements of HSRL2 atmospheric aerosol and ozone with lower vertical resolution. Several combinations of horizontal and	<input type="text" value="HSRL2"/>

Upload Your Instrument(s) Description Document: No file chosen (Select your file) 

*****(NOTE: Any new file/document upload WILL OVERWRITE your previously uploaded document)*****
*****(If you have more than 1 file (document) to upload, please email the files to: gao.chen@nasa.gov and/or ali.a.aknan@nasa.gov. Thank you!)*****

OVERWRITE my previous record (i.e., ALL previously registered dataIDs for this mission will be removed).

Link to PI website, instrument, experiment description, etc
 Optional: to display on LaRC Archive webpage

Text describing PI experiment or measurements (e.g., NASA LaRC DIAL - Troposphere O3, Aerosols, and Clouds Profiles):
 Optional: to display on LaRC Archive webpage

Data Upload (FScan) Page Example

SARPEAST File Scanning and Submittal
[ICARTT File Format Document](#)

IMPORTANT: In order to archive files, dataID(s) must be registered first because the PI Data Directory on the Server will be created from the registration.

[Help](#)

What to do: Scan Only Scan and Archive (see dataIDs note above)

File Upload: No file chosen (Select your file)

This is Final (i.e., NOT Field) data -- File Revision MUST be set to R0, R1, etc.

[How to upload multiple files in one step.](#) | [How to archive *.zip ICARTT files.](#)

This application scans three (3) file types -- defined as ICARTT: [FFI 1001](#), [FFI 2110](#), and [FFI 2310](#)

Select Data Time-Interval Type : Start, Stop, and Mid-point

Display FScan results (report) : Detailed Brief

Pressing this button will Scan *[and Archive]* the selected file(s).

Please scan your files on your machine first using [FScanBrowser \(Windows\)](#)
*** Your Browser should allow FScan to open a new Window (to display FScan's results) ***

By: Ali Aknan -- Ver 3.7.8 (June 02, 2022)

*** **Scanning Failed. See [Error Messages Below](#)** ***

Please scroll down to the end of this page to view the results.
NOTE: error messages are displayed in RED; warnings and other info in ORANGE

Function Selected: [Scan and Archive](#)
FScan Results Report: [Detailed](#)
Time-interval Type Selected: [Start, Stop, and Mid-point](#)
Filename Submitted: [DCOTSS-MMS-1HZ_ER2_20210817_RA.ict](#)
File Size: 3066538 bytes
File Content: application/octet-stream

Receiving file was successful - now onto scanning ...

✓ **ALL DONE.**

Please scroll down to the end of this page to view the results.
NOTE: error messages are displayed in RED; warnings and other info in ORANGE.

Function Selected: [Scan and Archive](#)
FScan Results Report: [Detailed](#)
Time-interval Type Selected: [Start, Stop, and Mid-point](#)
Filename Submitted: [ACTIVATE-LARGE-SMPS_HU25_20220111_RA_L2.ict](#)
File Size: 55987 bytes
File Content: application/octet-stream

Receiving file was successful - now onto scanning ...

Note: results open on a separate page, may need to change browser permissions

File Naming Convention for PACE-PAX

DataID_LocationID_DATETIME_R# [_Description].extension

- **DataID:** a short description of measured parameter/species, instrument, or model prefixed by “PACEPAX-”
- **LocationID:** a controlled identifier of data platform, provided on the dataID registration website in a drop-down box.
 - PACE-PAX locationIDs: ER2, CIRPAS-TO, GROUND, MODEL, ANALYSIS, SATELLITE, TRAJECTORY, and OTHER
- **DATETIME = YYYYMMDD** for all other data: **UTC** date of takeoff for flight data or the beginning of the measurement for ground sites.
- **DATETIME = YYYYMMDDTHHMMSS** for ER2 L1C data files, “T” between date and time.
- **R#:** Revision identifier. Typically, RA, RB, RC, ... for field data and R0, R1, R2, ... for the publication quality data. *Note: archived files cannot be overwritten, only replaced with subsequent revisions*
- **Description:** optional additional description of the file if necessary
- **Extension:** *.ict, *.nc, *.cdf, *.hdf, *.h4, *.h5, *.hdf4, *.hdf5, *.he5, *.kmz, *.kml, *.htm, *.html, *.txt, *.jpg, *.jpeg, *.gif, *.png, *.bmp, *.pdf, *.xls, *.xlsx, *.doc, *.docx, *.ppt, *.pptx
- **Extension upon request:** *.zip, *.tar, *.gz, or others
- The underscore, “_”, is used ONLY to separate the different fields of the filename
- Examples: the filename for HSRL cloud and aerosol observations on August 15th, 2024:
 - PACEPAX-HSRL_ER2_20240815_RA.h5 (for draft/field data)
 - PACEPAX-HSRL_ER2_20240815_R0.h5 (for publication quality data)

PACE-PAX Data Submission Schedule

Mission Phase	Data Type	Submission Deadline	Access Control
Field Deployment	Field Data	December 31, 2024	Science team and Partners
Post-Deployment	Publication-quality or “Final” Data	March 31, 2025	Public

- Field data submission deadline may vary depending on field operation constraints
- Field data are **not** citable
- The final data should be of publication quality and time synced to the time standard for each platform

PACE-PAX Data Format Requirements (I)

ER-2 Passive Instruments

- The PACE-PAX Polarimeters L1C products will follow the PACE L1C data format standards:
https://pace.oceansciences.org/docs/NASA_TM2024219027v12_Level1C.pdf
- Global attributes
 - A subset of the PACE L1C global attributes will be required
 - Additional attributes are specific to airborne field study operation, e.g., platform
- Groups and dimensions
 - Adopt the same group names and structure as the PACE L1C format
 - Use the same dimension names as listed in table 1
- Data Variables
 - use the variable names, dimensions, and units specified in table 7
 - long_name and units are required, _FillValues is required if used. Since Panoply uses long_name as axis title, consider put information in “comment”
 - Units should be “1” for unitless variables, based on SI unit system

PACE-PAX Data Format Requirements (II)

ER-2 Passive Instruments (cont.)

- Flag variables, e.g., quality indicators
 - long_name is required, consider to use flag_values and flag_meaning attributes
 - Should not have the “units” attribute
- File checker will be developed to ensure consistency with format standards for the final/publication quality data:
 - Will tailor to each instrument reflecting inherent data product differences
 - Will check mandatory global attributes, dimension names, and group structure and group names
 - Will check data product variable names and dimensions
 - Will check variable attributes: long_name, units, _FillValues attributes if applicable
 - Will provide timely support to resolve format issues
- Need sample files from each polarimeter team to develop and test the file checker

PACE-PAX Data Format Requirements (III)

in-situ measurements and HSRL

- The PACE-PAX data will conform to **ICARTT, netCDF, or HDF** format standards. This supports NASA's Open-Source Science and Open Data initiatives by making PACE-PAX data FAIR
- All in-situ measurements are required to report data in ICARTT format (<http://www-air.larc.nasa.gov/missions/etc/IcarttDataFormat.htm>). The ACVSNC variable standard names should be used for all data product variables
- ICARTT files will be scanned to ensure compliance with the format requirements
- HSRL data will be reported in HDF files following the template for used in CPEX-CV and ACTIVATE, which has the standardized global and variable attributes, including ACVSNC standard names, and properly dimensioned data product variables

PACE-PAX Data Reporting Best Practices for in-situ measurements

- Use the same number, names, and order of variables throughout the mission for files within the same dataID and revision. This prevents issues with the online merge tool, and FScan now checks the variable list between files
- Measurement Time Reporting:
 - Fixed variable name(s): Time_Start, Time_Stop, and Time_Mid
 - Report start, stop, and mid times if integration interval larger than 1 sec
 - Can use one time stamp (e.g., Time_Start or Time_Stop) for data at ≥ 1 Hz
- Use file header (e.g., DATA_INFO) or metadata to indicate whether the measurement time is synced to the time standard. Final/publication quality data should be synchronized to science team pre-agreed time standard
- Report absolute concentrations and aerosol and cloud extensive properties at STP: 273.15K and 1013.25 hPa (i.e., 0°C and 1atm)
- Variable short name should not start with a number or contain “-”
- Recommend standard unit notation: [WMO Codes Registry : wmdr/unit](https://wmo.codes/registry/wmdr/unit)

ACVSNC Variable Standard Names

- Atmospheric Composition Variable Standard Name Convention (ACVSNC) is a NASA Earth Science Data Systems convention, intended to make data more findable and interoperable, and (re)usable:
<https://www-air.larc.nasa.gov/missions/etc/AtmosphericCompositionVariableStandardNames.pdf>
- Constructed from controlled vocabulary
- Tags, **NOT** short names or variable names:
 - **For ICARTT files: short name, unit, *standard name*, long name**
CH2O_pptv, pptv, *Gas_CH2O_insitu_S_AVMR*, mixing ratio by volume
CH2O_LOD_pptv, pptv, *Gas_CH2O_insitu_S_AVMR*, Limit of Detection
NOy_pptv, pptv, *Gas_NOyasNO_insitu_M_AVMR*, Total Reactive Nitrogen Mixing Ratio
Sc700_total, Mm-1, *AER_Scattering_insitu_red_RHd_Total_AMB*, Dry Scattering at 700nm (Total Aerosols)
 - **For netCDF and HDF files: use ACVSN_standard_name attribute**
- Contact Morgan, Gao, or Michael for questions or adding new standard names

PACE-PAX Science Data Policy

All participants are requested to accept the following responsibilities:

- Submit data in ICARTT, netCDF or HDF format no later than the deadlines
- If unexpected events lead to any delay in data submission, the PI is required to notify the project leadership as soon as issues are known
- **Publication-quality or “final” data should be submitted to the archive prior to any presentation at scientific conferences (e.g. AGU, AMS) or manuscript preparation**
- **Use of none public data at public conferences requires explicit authorization is obtained from the program managers and clear marking to indicate that data are not citable**
- Consult with PIs when using their data in conference/data workshop presentations and/or manuscript
- Invite PIs of any data used to be co-authors (particularly during post-deployment research phase)
- PIs should be available to answer questions about their data