Northern Hemisphere PIP Dataset README

Last updated: April, 2024

- Uploaded additional collocated surface meteorologic observations for each site
- Included a link to the public API developed for working with this data
- Updated sorting of 2015/2016 OLY yearly data

What is this?

Microphysical observations of precipitating particles are crucial for numerical weather prediction models and remote sensing retrieval algorithms. This dataset provides a unified, comprehensive collection of particle microphysical observations from the Precipitation Imaging Package (PIP) over the Northern Hemisphere. Data spans from 2014-2023 across 10 measurement sites and encompasses over 1 million precipitating minutes. Within this dataset, users will find a range of microphysical attributes for rain and snow, along with higher-order products.

Data Packaging & Conversion:

Data originally in a proprietary ASCII format has been converted to the more universally recognized NetCDF-4 format for ease of sharing and compatibility within the academic community. The conversion process, undertaken using a combination of bash and Python, ensures broader compatibility with various data analysis tools and platforms.

Location Details:

- International Collaborative Experiments for Pyeongchang 2018 Olympic and Paralympic Winter Games (ICE-POP or ICP)
- Olympic Mountains Experiment (OLYMPEX or OLY)
- Haukeliseter (HAK)
- Kiruna (KIS)
- Marquette (MQT)
- Gaylord (APX)
- Finland (FIN)
- North Slope Alaska (NSA)
- NASA Investigation of Microphysics and Precipitation for Atlantic Coast-Threatening Snowstorms (IMPACTS or IMP)
- Iqaluit (YFB)

Data Coverage:

Please see the below image and table for additional site details including the spatiotemporal coverage of each site. Note that days without precipitation are not included in this dataset and there exist other gaps from instrument outages and as a consequence of the quality assurance processing step.
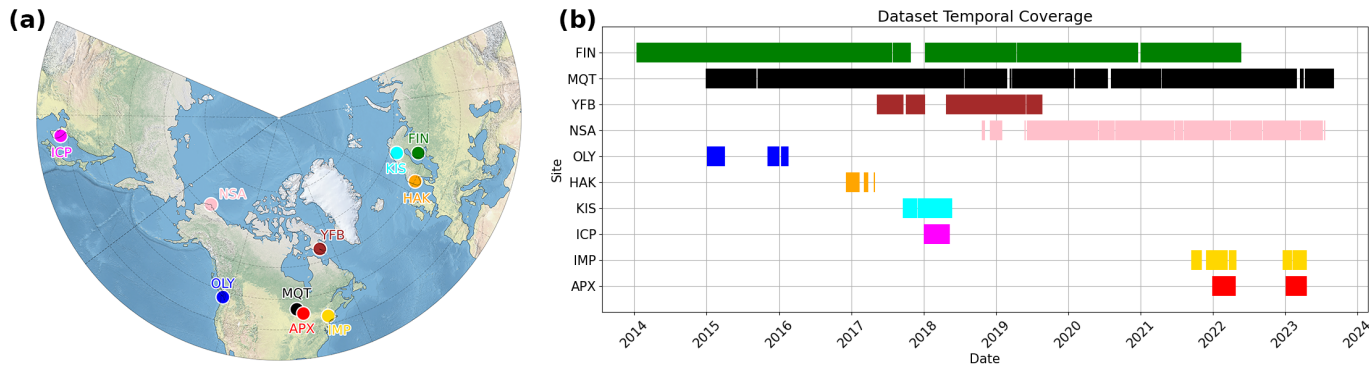
Name	ID	Location	Lat (°N)	Lon (°E)	Elev. (m)	Coverage	Precip. Mins.	Size (Mb)
SMEAR-ii	FIN	Hyytiälä, Finland	61.845	24.287	170	2014/01/14 - 2022/05/19	353700	687
Marquette	MQT	Michigan, USA	46.532	-87.548	426	2015/01/01 - 2023/08/31	363001	804.3
Iqaluit	YFB	Nunavut, Canada	63.747	-68.542	12	2017/05/11 - 2019/08/19	75718	166.1
North Slope	NSA	Alaska, USA	71.323	-156.612	8	2018/10/24 - 2023/07/18	152091	305.9
OLYMPEX	OLY	Washington, USA	47.97	-123.58	1603	2015/02/03 - 2016/02/14	17816	57.2
HiLAMS	HAK	Haukeliseter, Norway	59.81	7.21	991	2016/12/07 - 2017/04/26	9039	23.9
HiLAMS	KIS	Kiruna, Sweden	67.84	20.41	425	2017/09/21 - 2018/05/21	25676	66.7
ICE-POP	ICP	South Korea	37.665	128.7	789	2018/01/07 - 2018/04/24	19781	39.7
IMPACTS	IMP	Connecticut, USA	41.807	-72.294	78	2021/09/15 - 2023/04/16	27248	47.4
Gaylord	APX	Michigan, USA	44.908	-84.719	446	2022/01/01 - 2023/04/17	25807	74.6

Internal Structure of NetCDF Files:

Spatial & Temporal Variables: Lat/Lon and Time Data Variable: Contains one of the L3/L4 PIP products Bin Size Information: bin_centers, bin_edges for different particle diameter bins Note: Each daily file has exactly 1440 time steps with up to 131 bins for 2D variables. Missing data is marked as NaN. Not all variables exist for all days.

Data Levels & Description:

Level 1 (L1): Raw video data with compressed 8-bit grayscale frames (.pvi format) for 10-minute intervals.

Level 2 (L2): Time-stamped particle tables comprising 36 individual particle characteristics for each hydrometeor.

Level 3 (L3): Derived vertical velocity and particle size distribution tables for each minute.

Level 4 (L4): Estimates of effective density, phase classification, and precipitation rate.

Quality Assurance:

A quality assurance (QA) procedure has been undertaken to ensure the integrity of the data. Post QA, the data is transformed into daily NetCDF-4 files following the Climate and Forecast (CF) conventions (version 1.10) and compressed with a level 2 deflation for optimized file size.

Filename Convention:

The naming convention for the NetCDF files is structured as follows: XXXYYYYMMDD_**product**.nc

Where:

- XXX: PIP instrument number
- YYYYMMDD: Date (YearMonthDay)

product can be one of the following:

- min: L4 precipitation product
- rho: Effective density distributions
- psd: Particle size distributions
- vvd: Vertical velocity distributions

Directory Structure:

- SITE_YEAR/
  - netCDF/
    - adjusted_edensity_lwe_rate/
      - XXXYYYYMMDD_min.nc
    - edensity_distributions/
      - XXXYYYYMMDD_rho.nc
    - particle_size_distributions/
      - XXXYYYYMMDD_psd.nc
    - velocity_distributions/
      - XXXYYYYMMDD_vvd.nc

Surface MET:

Daily observations of surface meteorologic observations have also been collected, processed into NetCDF files and stored in the MET directory to provide additional environmental context to the PIP observations. Variables include surface temperature, relative humidity, pressure, wind speed and wind direction. Note that due to different instrumentation steups at different sites, not all locations include observations of all aforementioned variables.

API:

An API for working with the PIP data has also been developed and made available online (https://pipdb.readthedocs.io). This package is a simple query API for parsing, visualizing and performing particle size distribution calculations for PIP data.

Contact:

For any further questions or assistance with the dataset, please reach out to the corresponding data author (Fraser King) or Claire Pettersen via email: kingfr@umich.edu and pettersc@umich.edu