The Journal of Open Source Software

pypromice: A Python package for processing automated weather station data

Penelope R. How[1¶], **Patrick J. Wright**[1], **Kenneth D. Mankoff**[1,2,3], **Baptiste Vandecrux**[1], **Robert S. Fausto**[1], and **Andreas P. Ahlstrøm**[1]

1 Department of Glaciology and Climate, Geological Survey of Denmark and Greenland (GEUS), Copenhagen, Denmark **2** Autonomic Integra, New York, NY, USA **3** NASA Goddard Institute for Space Studies, New York, NY, USA ¶ Corresponding author

DOI: 10.21105/joss.05298

Software

- Review
- Repository
- Archive

Editor: Hauke Schulz

Reviewers:

- @jroettenbacher
- @simonrp84

Submitted: 03 March 2023
Published: 22 May 2023

Summary

The `pypromice` Python package is for processing and handling observation datasets from automated weather stations (AWS). It is primarily aimed at users of AWS data from the Geological Survey of Denmark and Greenland (GEUS), which collects and distributes in situ weather station observations to the cryospheric science research community. Functionality in `pypromice` is primarily handled using two key open-source Python packages, `xarray` (Hoyer & Hamman, 2017) and `pandas` (The pandas development team, 2020).

A defined processing workflow is included in `pypromice` for transforming original AWS observations (Level 0, `L0`) to a usable, CF-convention-compliant dataset (Level 3, `L3`) (Figure 1). Intermediary processing levels (`L1`,`L2`) refer to key stages in the workflow, namely the conversion of variables to physical measurements and variable filtering (`L1`), cross-variable corrections and user-defined data flagging and fixing (`L2`), and derived variables (`L3`). Information regarding the station configuration is needed to perform the processing, such as instrument calibration coefficients and station type (one-boom tripod or two-boom mast station design, for example), which are held in a `toml` configuration file. Two example configuration files are provided with `pypromice`, which are also used in the package's unit tests. More detailed documentation of the AWS design, instrumentation, and processing steps are described in (Fausto et al., 2021).

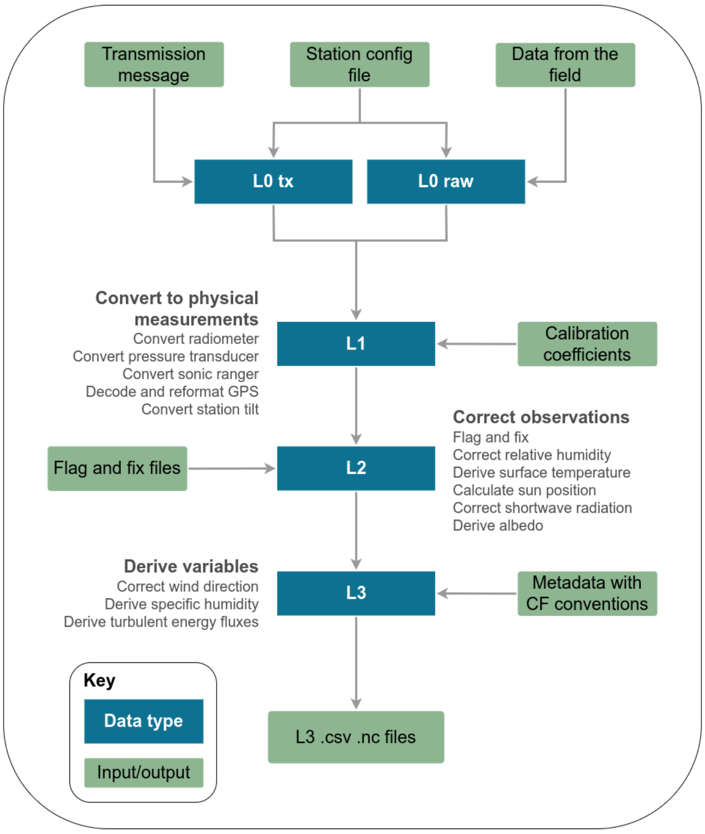

Figure 1: AWS data Level 0 (`L0`) to Level 3 (`L3`) processing steps, where `L0` refers to raw, original data and `L3` is usable data that has been transformed, corrected and filtered

`L0` data is either collected from an AWS during a station visit or is transmitted in near-real-time from each AWS via the Iridium Short Burst Data (SBD) service. An object-oriented workflow for fetching and decoding SBD messages to Level 0 data (`L0 tx`) is included in `pypromice` (Figure 2). Alongside the processing module, this workflow can be deployed for operational uses to produce `L3` AWS data in near-real-time. A post-processing workflow is also included to demonstrate how near-real-time AWS data can be treated after `L3` for submission to global weather forecasting models under the World Meteorological Organisation (WMO).

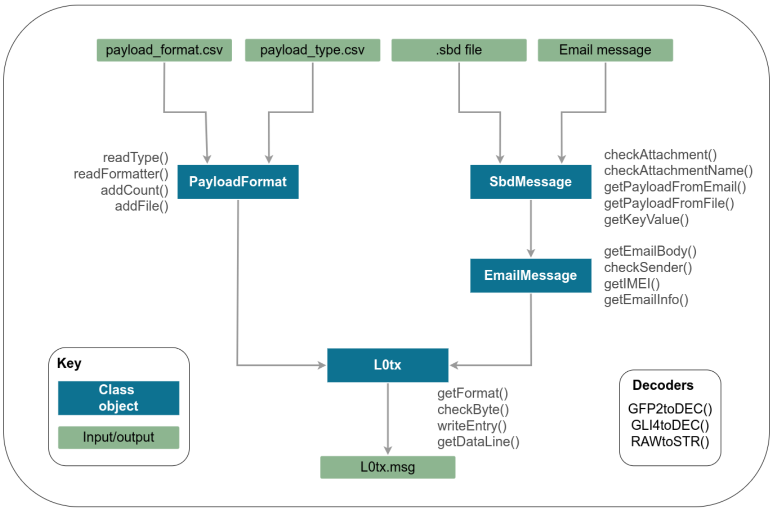

Figure 2: Object-oriented workflow in `pypromice.tx` for fetching and decoding AWS transmission messages to Level 0 (`L0 tx`) data

Statement of need

`pypromice` has four main research purposes:

1. Process and handle AWS observations
2. Document AWS data processing with transparency and reproducibility
3. Supply easy and accessible methods to handle AWS data
4. Provide opportunities to contribute to the processing and handling of AWS data in an open and collaborative manner

The `pypromice` software has been designed to handle and process data from AWSs located in Greenland. The compilation and processing of data from national AWS networks has historically been conducted through un-distributed, oftentimes proprietary software. Similar Python packages to `pypromice` has been developed to process data from historical AWS in Greenland (K. Steffen et al., 2023; Vandecrux et al., 2020 - 2023) or commercial AWS (e.g. pywws (Easterbrook, 2008 - 2023)) or to post-process and harmonize AWS data from different institutions (e.g. JAWS (Zender et al., 2017 - 2019)). Therefore, there was a key need for the development of `pypromice` in order to have a package with a complete and operational `L0` to `L3` workflow.

Usage

The `pypromice` software handles data from 43 AWSs on hourly, daily and monthly time scales. The AWS data products have been used in high impact studies (Box et al., 2022; Machguth et al., 2016; Oehri et al., 2022), and have been crucial for evaluating the effect of climate change on land ice in annual reports such as the Arctic Report Card and "State of the Climate" (Moon, Mankoff, Fausto, Fettweis, Loomis, et al., 2022; Moon, Mankoff, Fausto, Fettweis, Tedesco, et al., 2022). The AWS data originates from three national monitoring programmes - the Programme for Monitoring of the Greenland Ice Sheet (PROMICE), the Greenland Climate Network (GC-Net) and the Greenland Ecosystem Monitoring programme (GEM).

GEUS is responsible for the Programme for Monitoring of the Greenland Ice Sheet (PROMICE), which is now a network of over 21 AWSs installed across the Greenland Ice Sheet (Ahlstrøm & the PROMICE project team, 2008). Launched in 2007, these one-level tripod stations are designed to stand on ice and move with the ice flow close to the ice sheet periphery (Fausto et al., 2021; How, Mankoff, et al., 2023a). The PROMICE stations are designed to monitor the surface melt and its meteorological drivers in the ablation area of the ice sheet.

In 2021, GEUS assumed responsibility of the Greenland Climate Network (GC-Net) AWS locations (C. Steffen et al., 1996), previously maintained by the United States National Science Foundation (NSF), National Aeronautics and Space Administration (NASA) and Swiss Federal Institute for Forest, Snow and Landscape Research (WSL). This expansion added 16 two-level mast stations to GEUS' sites. The data from these stations are intended to monitor conditions on the inner regions of the ice sheet, including snow accumulation and surface conditions (How, Mankoff, et al., 2023b).

The Greenland Ecosystem Monitoring programme (GEM) is an integrated, long-term monitoring effort that examines the effects of climate change on Arctic ecosystems. Established in 1995, GEM includes monitoring at Zackenberg, Kobbefjord, and Disko, Greenland. The program offers access to over 1000 freely-available environmental datasets, including data from 6 GEUS-designed AWS installations (GEM, 2020) which have been used in scientific publications (Messerli et al., 2022).

Documentation

`pypromice` versions accompany releases of GEUS AWS one-boom and two-boom data publications (How, Wright, et al., 2023).

Package documentation is available on the `pypromice` readthedocs.

Guides for general GEUS AWS processing operations under PROMICE and GC-Net are included at the GEUS Glaciology and Climate GitHub pages.

Acknowledgements

This work is funded through the Danish Ministry of Climate, Energy and Utilities via The Programme for Monitoring of the Greenland Ice Sheet (PROMICE) and the Greenland Climate Network (GC-Net).

References

Ahlstrøm, A. P., & the PROMICE project team. (2008). A new programme for monitoring the mass loss of the greenland ice sheet. *GEUS Bulletin*, *15*, 61–64. https://doi.org/10.34194/geusb.v15.5045

Box, J. E., Hubbard, A., Bahr, D. B., Colgan, W. T., Fettweis, X., Mankoff, K. D., Wehrlé, A., Noël, B., Broeke, M. R. van den, Wouters, B., Bjørk, A. A., & S., F. R. (2022). Greenland ice sheet climate disequilibrium and committed sea-level rise. *Nature Climate Change*, *12*, 808–813. https://doi.org/10.1038/s41558-022-01441-2

Easterbrook, J. (2008 - 2023). *Python software for USB wireless weather stations* (Version 23.2.0). https://github.com/jim-easterbrook/pywws

Fausto, R. S., As, D. van, Mankoff, K. D., Vandecrux, B., Citterio, M., Ahlstrøm, A. P., Andersen, S. B., Colgan, W., Karlsson, N. B., Kjeldsen, K. K., Korsgaard, N. J., Larsen, S. H., Nielsen, S., Pedersen, A. Ø., Shields, C. L., Solgaard, A. M., & Box, J. E. (2021). Programme for monitoring of the greenland ice sheet (PROMICE) automatic weather station data. *Earth System Science Data*, *13*(8), 3819–3845. https://doi.org/10.5194/essd-13-3819-2021

GEM. (2020). GlacioBasis zackenberg - snow cover - snow depth radar. *Dataset*. https://doi.org/10.17897/E594-NV64

How, P., Mankoff, K. D., Wright, P. J., Vandecrux, B., Ahlstrøm, A. P., & Fausto, R. S. (2023a). AWS one boom tripod edition 4. *Dataset*. https://doi.org/10.22008/FK2/IW73UU

How, P., Mankoff, K. D., Wright, P. J., Vandecrux, B., Ahlstrøm, A. P., & Fausto, R. S. (2023b). AWS two boom mast edition 1. *Dataset*. https://doi.org/10.22008/FK2/GNYFUK

How, P., Wright, P. J., Mankoff, K. D., & Vandecrux, B. (2023). *Pypromice* (Version 1.0.0). GEUS Dataverse. https://doi.org/10.22008/FK2/IPOHT5

Hoyer, S., & Hamman, J. (2017). Xarray: N-d labeled arrays and datasets in Python. *Journal of Open Research Software*, *5*(1), 10. https://doi.org/10.5334/jors.148

Machguth, H., MacFerrin, M., As, D. van, Box, J. E., Charalampidis, C., Colgan, W., Fausto, R. S., Meijer, H. A. J., Mosley-Thompson, E., & Wal, R. S. W. van de. (2016). Greenland meltwater storage in firn limited by near-surface ice formation. *Nature Climate Change*, *6*, 390–393. https://doi.org/10.1038/nclimate2899

Messerli, A., Arthur, J., Langley, K., How, P., & Abermann, J. (2022). Snow cover evolution at qasigiannguit glacier, southwest greenland: A comparison of time-lapse imagery and mass balance data. *Frontiers in Earth Science*, *10*. https://doi.org/10.3389/feart.2022.970026

Moon, T. A., Mankoff, K. D., Fausto, R. S., Fettweis, X., Loomis, B. D., Mote, T. L., Poinar, K., Tedesco, M., Wehrlé, A., & Jensen, C. D. (2022). Greenland ice sheet. *Arctic Report Card 2022*. https://doi.org/10.25923/c430-hb50

Moon, T. A., Mankoff, K. D., Fausto, R. S., Fettweis, X., Tedesco, M., Wehrlé, A., Loomis, B. D., Mote, T. L., Jensen, K., C. D., Cappelen, J., & Winton, Ø. A. (2022). Greenland ice sheet. In: State of the climate in 2021 - the arctic. *Bulletin of the American Meteorological Society*, *103*, S276–S279. https://doi.org/10.1175/BAMS-D-22-0082.1

Oehri, J., Schaepman-Strub, G., Kim, J.-S., Grysko, R., Kropp, H., Grünberg, I., Zemlianskii, V., Sonnentag, O., Euskirchen, E. S., Reji Chacko, M., Muscari, G., Blanken, P. D., Dean, J. F., Sarra, A. di, Harding, R. J., Sobota, I., Kutzbach, L., Plekhanova, E., Riihelä, A., … Chambers, S. D. (2022). Vegetation type is an important predictor of the arctic summer land surface energy budget. *Nature Communications*, *13*(1), 6379. https://doi.org/10.1038/s41467-022-34049-3

Steffen, C., Box, J., & Abdalati, W. (1996). Greenland climate network: GC-net, in US army cold regions reattach and engineering (CRREL). *CRREL Special Report*, *96*, 98–103.

Steffen, K., Vandecrux, B., Houtz, D., Abdalati, W., Bayou, N., Box, J., Colgan, L., Espona Pernas, L., Griessinger, N., Haas-Artho, D., Heilig, A., Hubert, A., Iosifescu Enescu, I., Johnson-Amin, N., Karlsson, N. B., Kurup, R., McGrath, D., Naderpour, R., Pederson, A. Ø., … Ahlstrøm, A. (2023). *GC-net level 1 automated weather station data* (Version 2). https://doi.org/10.22008/FK2/VVXGUT

The pandas development team. (2020). *Pandas-dec/pandas: pandas*. Zenodo. https://doi.org/10.5281/zenodo.3509134

Vandecrux, B., Box, J., Houtz, D., & Revheim, M. K. (2020 - 2023). *The GC-net level 1 dataset and processing scripts* (Version 1.1). Geological Survey of Denmark; Greenland (GEUS). https://github.com/GEUS-Glaciology-and-Climate/GC-Net-level-1-data-processing

Zender, C., Wang, W., & Saini, A. (2017 - 2019). *JAWS: An extensible toolkit to harmonize and analyze polar automatic weather station datasets, manuscript in preparation for geosci. Model dev.* (Version 1.0). University of California, Irvine. https://github.com/jaws/jaws