

h5fortran: object-oriented polymorphic Fortran interface for HDF5 file IO

Michael Hirsch¹

¹ Boston University

DOI: [10.21105/joss.0XXXX](https://doi.org/10.21105/joss.0XXXX)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Editor: [Editor Name](#) ↗

Submitted: 01 January 1900

Published: 01 January 3030

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

Fortran has only raw file input-output (IO) built in to the language. To support reproducibility of work done in any programming language and long-term usefulness of the data generated or processed, it is beneficial to use self-describing data file formats like HDF5 (Folk, Heber, Koziol, Pourmal, & Robinson, 2011). Many popular languages and libraries used for simulation and data science use the HDF5 (The HDF Group, 1997) file IO library. Most programs and libraries intended for use by practitioners such as modelers and data scientists themselves use an object-oriented HDF5 interface like h5py (Collette, 2012) (Collette, 2013).

h5fortran (Hirsch, 2018) is a Fortran 2008 interface to HDF5 that abstracts away most details of a frequently-used subset of HDF5 operations. h5fortran makes HDF5 use from Fortran as easy as in high-level scripting languages. h5fortran is known to work for Gfortran 6 or newer and Intel Fortran 19.0 or newer for Linux, MacOS, Windows on Intel / AMD, ARM and IBM POWER systems. The Fortran 2008 standard is adhered to by h5fortran, so the main limit to Fortran compiler support is the compiler supporting Fortran 2008 standard.

h5fortran has general applicability to projects needing to do any of:

- writing variables to HDF5: scalar to 7-D, of type real32, real64 or integer
- reading variables from HDF5: scalar to 7-D, of type real32, real64 or integer
- reading or writing character variables to / from HDF5 file
- writing variable attributes to disk
- getting the shape of a disk variable, for example to allocate a memory variable to read that variable

In addition to the object-oriented interface, h5fortran provides single-command read / write procedures. Array slicing on read allows reading a portion of a large disk variable into memory. If the user has HDF5 with SZIP or ZLIB compression enabled, h5fortran is capable of reading and writing compressed variables, which can save over 50% disk space depending on the data lossless compressibility. Data shuffling and Fletcher32 checksums provide better compression and a check of file integrity respectively. h5fortran was designed for use by individual users on their laptops or embedded devices, as well as for use in HPC applications where parallel tasks need read only part of a milestone or shared HDF5 variable.

h5fortran was originally developed for the GEMINI (Zettergren, Grubbs, & Hirsch, 2017) (Zettergren & Snively, 2019) ionospheric model, funded in part by NASA ROSES #19-HDEE19_2-0007.

Other programs

While other HDF5 interfaces exist, h5fortran presents a broad set of commonly used features, comprehensive test coverage and robustness across compilers and computing systems. We have written a companion library for NetCDF4 called nc4fortran (Hirsch, 2019), which by design has a nearly identical user-facing API. Other Fortran HDF5 interfaces such as (Irwin, 2017) use a functional interface mimicking the HDF5 LT functions, which require the end user to keep track of extra variables versus the single object used by h5fortran.

Collette, A. (2012). H5py: HDF5 python interface. Retrieved from <https://github.com/h5py/h5py>

Collette, A. (2013). *Python and hdf5: Unlocking scientific data*. " O'Reilly Media, Inc."

Folk, M., Heber, G., Koziol, Q., Pourmal, E., & Robinson, D. (2011). An overview of the hdf5 technology suite and its applications. In *Proceedings of the edbt/icdt 2011 workshop on array databases*, AD '11 (pp. 36–47). New York, NY, USA: Association for Computing Machinery. doi:[10.1145/1966895.1966900](https://doi.org/10.1145/1966895.1966900)

Hirsch, M. (2019). Nc4fortran: Object-oriented fortran netcdf4 interface. doi:[10.5281/zenodo.3718418](https://doi.org/10.5281/zenodo.3718418)

Hirsch, M. (2018, December). H5fortran: Object-oriented fortran hdf5 interface. doi:[10.5281/zenodo.3733569](https://doi.org/10.5281/zenodo.3733569)

Irwin, J. (2017). HDF5 utils. Retrieved from https://github.com/tiasus/HDF5_utils

The HDF Group. (1997). Hierarchical Data Format, version 5. Retrieved from <http://www.hdfgroup.org/HDF5/>

Zettergren, M. D., & Snively, J. B. (2019). Latitude and longitude dependence of ionospheric tec and magnetic perturbations from infrasonic-acoustic waves generated by strong seismic events. *Geophysical Research Letters*, 46(3), 1132–1140. doi:[10.1029/2018GL081569](https://doi.org/10.1029/2018GL081569)

Zettergren, M., Grubbs, G., & Hirsch, M. (2017). GEMINI3D: Time-dependent ionospheric model. doi:[10.5281/zenodo.3731716](https://doi.org/10.5281/zenodo.3731716)