

¹ pyfive: A pure-Python HDF5 reader

² **Bryan N. Lawrence**  ¹, **Ezequiel Cimadevilla**  ¹, **Wout De Nolf**  ¹, **David**
³ **Hassell**  ¹, **Jonathan Helmus**¹, **Brian Maranville**  ¹, **Benjamin Hodel**¹, **Kai**
⁴ **Mühlbauer**  ¹, and **Valeriu Predoi**  ¹

⁵ 1 National Center for Atmospheric Science (NCAS), United Kingdom. 2 Department of Meteorology,
⁶ University of Reading, Reading, United Kingdom.  3 Instituto de Física de Cantabria (IFCA),
⁷ CSIC-Universidad de Cantabria, Santander, Spain. 4 European Synchrotron Radiation Facility (ESRF),
⁸ Grenoble, France. 5 NIST Center for Neutron Research 6 Institute of Geosciences, Meteorology Section,
⁹ University of Bonn, Germany.

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

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

Editor: [Open Journals](#) 

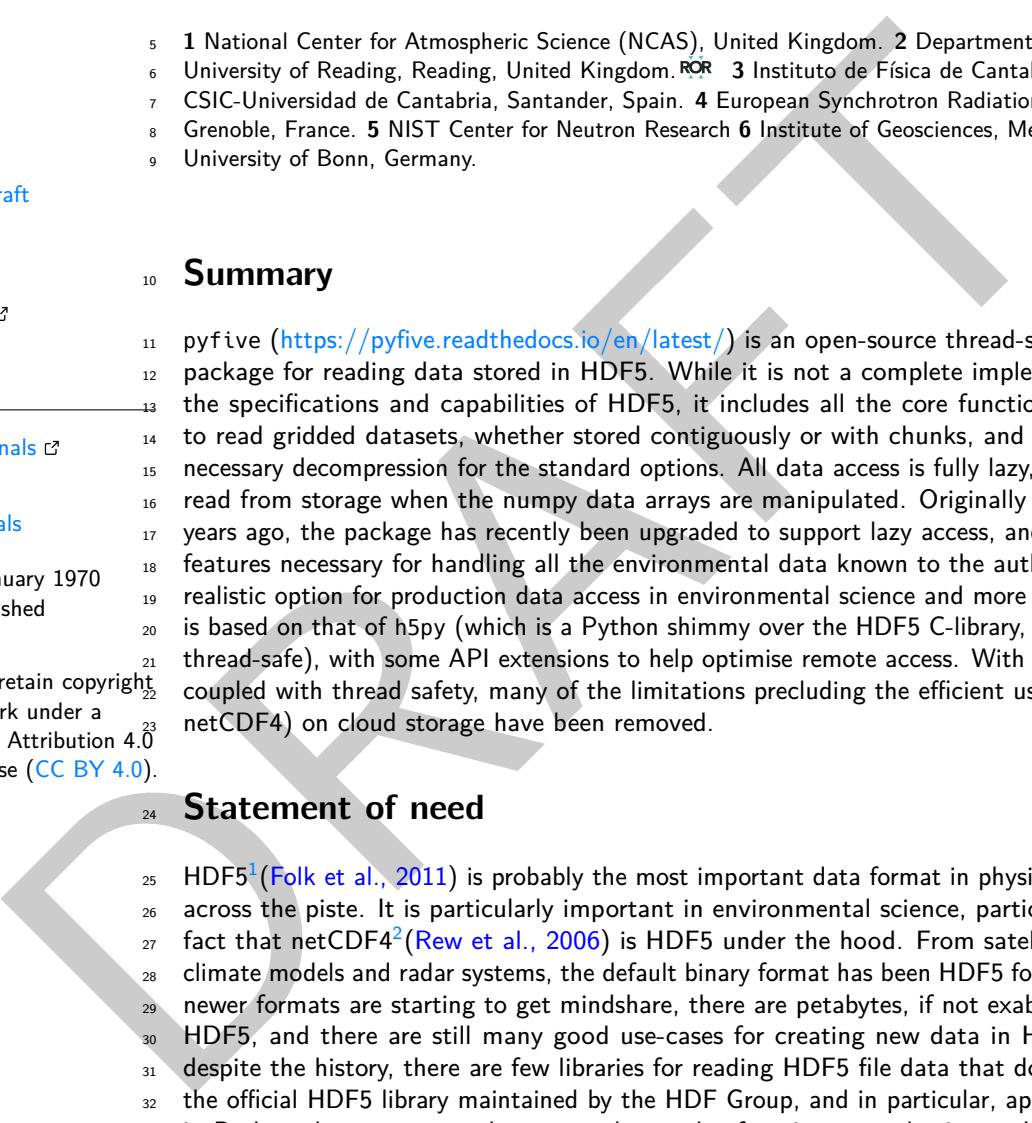
Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

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


²⁴ Statement of need

²⁵ HDF5¹ ([Folk et al., 2011](#)) is probably the most important data format in physical science, used
²⁶ across the piste. It is particularly important in environmental science, particularly given the
²⁷ fact that netCDF4² ([Rew et al., 2006](#)) is HDF5 under the hood. From satellite missions, to
²⁸ climate models and radar systems, the default binary format has been HDF5 for decades. While
²⁹ newer formats are starting to get mindshare, there are petabytes, if not exabytes, of existing
³⁰ HDF5, and there are still many good use-cases for creating new data in HDF5. However,
³¹ despite the history, there are few libraries for reading HDF5 file data that do not depend on
³² the official HDF5 library maintained by the HDF Group, and in particular, apart from pyfive,
³³ in Python there are none that cover the needs of environmental science. While the HDF5
³⁴ c-library is reliable and performant, and battle-tested over decades, there are some caveats to
³⁵ depending upon it: Firstly, it is not thread-safe, secondly, the code is large and complex, and
³⁶ should anything happen to the financial stability of the HDF5 Group, it is not obvious the
³⁷ C-code could be maintained. Finally, the code complexity also meant that it is not suitable
³⁸ for developing bespoke code for data recovery in the case of partially corrupt data. From a
³⁹ long-term curation perspective both of these last two constraints are a concern.

⁴⁰ The original implementation of pyfive (by JH), which included all the low-level functionality

¹<https://www.hdfgroup.org/solutions/hdf5/>

²<https://www.unidata.ucar.edu/software/netcdf>

41 to deal with the internals of an HDF5 file was developed with POSIX access in mind. The
42 recent upgrades were developed with the use-case of performant remote access to curated
43 data as the primary motivation, but with additional motivations of having a lightweight HDF5
44 reader capable of deploying in resource or operating-system constrained environments (such
45 as mobile), and one that could be maintained long-term as a reference reader for curation
46 purposes. The lightweight deployment consequences of a pure-Python HDF5 reader needs no
47 further introduction, but as additional motivation we now expand on the issues around remote
48 access and curation.

49 Thread safety has become a concern given the wide use of Dask³ in Python based analysis
50 workflows, and this, coupled with a lack of user knowledge about how to efficiently use HDF5,
51 has led to a community perception that HDF5 is not fit for remote access (especially on cloud
52 storage). Issues with thread safety arise from the underlying HDF5 c-library, and cannot be
53 resolved in any solution depending on that library, hence the desire for a pure Python solution.
54 Remote access has been bedevilled by the widespread need to access remotely data which has
55 been chunked and compressed, combined with the use of HDF5 data which was left in the
56 state it was when the data was produced - often with default unsuitable chunking (Rew, 2013)
57 and with interleaved chunk indexes and data. Solutions have mainly consisted of reformatting
58 the data (and rechunking it at the same time) or utilising kerchunk mediated direct access
59 to chunked HDF5 data⁴. However, in practice using kerchunk requires the data provider to
60 generate kerchunk indices to support remote users, and it leads to issues of synchronicity
61 between indices and changing datasets.

62 This version of pyfive was developed with these use-cases in mind. There is now full support
63 for lazy loading of chunked data, and methods are provided to give users all the benefits of
64 using kerchunk, but without the need for a priori generation. Because pyfive can access and
65 cache (in the client) the b-tree (index) on a variable-by-variable basis, most of the benefits of
66 kerchunk are gained without any of the constraints. However, the kerchunk index is always a
67 contiguous object accessible with one get transaction, this is not necessarily the case with the
68 b-tree, unless the source data has been repacked to ensure contiguous metadata using a tool
69 like h5repack. Much of the community is unaware of the possibility of repacking the index
70 metadata, and this together with relatively opaque information about the internal structure
71 of files (and hence the necessity or other wise of such repacking), means that repacking is
72 rarely done. To help with this process, pyfive also includes extensions to expose information
73 about how data and indexes are distributed in the files. With these tools, index extraction
74 with pyfive can be comparable in performance to obtaining a kerchunk index, and completely
75 opaque to the user.

76 With the use of pyfive, suitably repacked and rechunked HDF5 data can now be considered
77 "cloud-optimised", insofar as with lazy loading, improved index handling, and thread-safety,
78 there are no "format-induced" constraints on performance during remote access. To aid in
79 discovering whether or not a given HDF5 dataset is cloud-optimised, pyfive also now provides
80 simple methods to expose information about file layout - both in API extensions, and in a
81 new p5dump utility packaged with the pyfive library, which provides (in the default view)
82 functionality similar to ncdump, and when used with p5dump -s, information about storage
83 characteristics.

84 The issues of the dependency on a complex code maintained by one private company in the
85 context of maintaining data access (over decades, and potentially centuries), can only be
86 mitigated by ensuring that the data format is well documented, that data writers use only the
87 documented features, and that public code exists which can be relatively easily maintained.
88 The HDF5group have provided good documentation for the core features of HDF5 which
89 include all those of interest to the weather and climate community who motivated this reboot
90 of pyfive, and while there is a community of developers beyond the HDF5 group (including

³<https://www.dask.org/>

⁴<https://fsspec.github.io/kerchunk/>

some at the publicly funded Unidata institution), recent events suggest that given most of those developers and their existing funding are US based, some spreading of risk would be desirable. To that end, a pure Python code, which is relatively small and maintained by an international constituency, alongside the existing C-code, provides some assurance that the community can maintain HDF5 access for the foreseeable future. A pure Python code also makes it easier to develop scripts which can work around data and metadata damage should they occur.

Examples

We now introduce three aspects of the new functionality that pyfive now provides: remote access, configurable lazy loading, and determining whether files are cloud optimised.

Remote Access

A notable feature of the recent pyfive upgrade is that it was carried out with thread-safety and remote access using fsspec (<https://filesystem-spec.readthedocs.io>) in mind. We provide two examples of using pyfive to access remote data, one in S3, and one behind a modern http web server:

For accessing the data on S3 storage, we will have to set up an s3fs virtual file system, then pass it to pyfive:

```
import pyfive
import s3fs
# storage options for an anon S3 bucket
# there are also caching options for the s3 middleware, not shown here
storage_options = {
    "anon": True,
    "client_kwargs": {"endpoint_url": "https://s3server.ac.uk"}
}
fs = s3fs.S3FileSystem(**storage_options)
file_uri = "s3-bucket/myfile.nc"
with fs.open(file_uri, "rb") as s3_file:
    nc = pyfive.File(s3_file)
    dataset = nc["var"]
```

for an HTTPS data server, the usage is similar:

```
import fsspec
import pyfive
# there are also caching options for the fsspec middleware, not shown here
fs = fsspec.filesystem("http")
with fs.open("https://site.com/myfile.nc", "rb") as http_file:
    nc = pyfive.File(http_file)
    dataset = nc["var"]
```

This is of course exactly the same pattern as remote access using h5py, and that is by design - to make moving to pyfive easy for users!

Lazy Loading

A key tenet of efficient remote access is that variable inspection is quick and involves the minimum of network traffic between storage and the client application. However when this is coupled with the common pattern of using Dask, some flexibility in what is loaded when is beneficial.

116 By default when one inspects the contents of a file using pyfive nothing more is read from
 117 the file than the names of the variables ("datasets" in the language of HDF5): for example:

```
118   with pyfive.File("myfile.h5", "r") as f:  
119     variables_in_file = [v for v in f]
```

118 involves nothing more than getting a set of variable names. When one wishes to inspect these
 119 variables:

```
118   with pyfive.File("myfile.h5", "r") as f:  
119     temp = f["temp"]  
120     print(temp[1:10])
```

120 the default in pyfive is to get only the metadata associated with each variable - but crucially
 121 at this point the b-tree index is also loaded and the variable can now be accessed outside the
 122 context manager. Data loading is now completely lazy and the variable instance (temp) has all
 123 the information needed to extract data as needed. This is done so that in Dask applications,
 124 when one passes each Dask computational chunk a portion of the variable, each such Dask
 125 chunk has already got the index, and when it does want to load data, it can be as efficient as
 126 possible.

127 However, there are situations where loading the b-tree at variable instantiation is not wanted,
 128 and all is wanted is to be able to view all the variable attributes. To support this option pyfive
 129 also offers the `get_lazy_view` file method, so one can do:

```
128   with pyfive.File("myfile.h5", "r") as f:  
129     temp = f.get_lazy_view("temp")  
130     print(temp.attrs())
```

130 It is still possible to access the data and the b-tree index is loaded when data access is first
 131 attempted.

132 This extra lazy view is new functionality. It would obviously have been possible to make the
 133 default not load the b-tree, but in the opinion of the current pyfive maintainers loading the
 134 b-tree at variable instantiation is likely more consistent with user expectations as it is more
 135 equivalent to the behaviour of other packages like h5py.

136 Cloud Optimisation

137 To be fully cloud optimised - as defined by Stern et al. (2022) - an HDF5 file needs to have
 138 a contiguous index for each variable, and the chunks for each variable need to be sensibly
 139 chosen and broadly contiguous within the file. When these criteria are met, indexes can be
 140 read efficiently, and middleware such as fsspec can make sensible use of readahead caching
 141 strategies.

142 HDF5 data files direct from simulations and instruments are often not in this state as information
 143 about the number of variables, the number of chunks per variable, and the compressed size of
 144 those variables is not known as the data is being produced.

145 In such cases the data is also often not chunked along the dimensions being added to as the
 146 file is written (since it would have to be buffered first).

147 Of course, once the file is produced, such information is available. Metadata can be repacked to
 148 the front of the file and variables can be rechunked and made contiguous - which is effectively
 149 the same process undertaken when HDF5 data is reformatted to other cloud optimised formats.

150 The HDF5 library provides a tool "h5repack" which can do this, provided it is driven with
 151 suitable information about required chunk shape and the expected size of metadata fields.
 152 pyfive supports both a method to query whether such repacking is necessary, and to extract
 153 necessary parameters.

154 In the following example we compare and contrast the unpacked and repacked version of a
 155 particularly pathological file, and in doing so showcase some of the pyfive API extensions
 156 which help us understand why it is pathological, and how to address those issues for repacking.

157 If we extract just a piece of the output of p5dump -s on this file (which has surface wind
 158 velocity at three hour intervals for one hundred years):

```

159 float64 time(time) ;
160         time:_standard_name = "time" ;
161         time:_n_chunks = 292192 ;
162         time:_chunk_shape = (1,) ;
163         time:_btree_range = (31808, 19854095942) ;
164         time:_first_chunk = 9094 ;
165
166 float32 uas(time, lat, lon) ;
167         uas:_Storage = "Chunked" ;
168         uas:_n_chunks = 292192 ;
169         uas:_chunk_shape = (1, 143, 144) ;
170         uas:_btree_range = (28672, 19854809382) ;
171         uas:_first_chunk = 36520 ;

```

172 we can immediately see that this will be a problematic file! The b-tree index is clearly interleaved
 173 with the data (compare the first chunk address with last index addresses of the two variables),
 174 and with a chunk dimension of (1,), any effort to use the time-dimension to locate data of
 175 interest will involve a ludicrous number of 1 number reads (all underlying libraries read the
 176 data one chunk at a time). It would feel like waiting for the heat death of the universe if one
 177 was to attempt to manipulate this data stored on an object store!

178 It is relatively easy (albeit slow) to use h5repack to fix this - e.g see Hassell & Cimadevilla
 179 Alvarez (2025) - after which we see:

```

180 float64 time(time) ;
181         time:_Storage = "Chunked" ;
182         time:_n_chunks = 1 ;
183         time:_chunk_shape = (292192,) ;
184         time:_btree_range = (11861, 11861) ;
185         time:_first_chunk = 40989128 ;
186         time:_compression = "gzip(4)" ;
187 float32 uas(time, lat, lon) ;
188         uas:_Storage = "Chunked" ;
189         uas:_n_chunks = 5844 ;
190         uas:_chunk_shape = (50, 143, 144) ;
191         uas:_btree_range = (18663, 347943) ;
192         uas:_first_chunk = 41041196 ;
193         uas:_compression = "gzip(4)" ;

```

194 Now data follows indexes, the time dimension is one chunk, and there is a more sensible
 195 number of actual data chunks. While this file would probably benefit from splitting, with a
 196 contiguous set of indexes, it is now possible to exploit this data via S3.

197 All the metadata shown in this dump output arises from pyfive extensions to the
 198 pyfive.h5t.DatasetID class. pyfive also provides a simple flag: consolidated_metadata
 199 for a File instance, which can take values of True or False for any given file, which simplifies
 200 at least the “is the index packed at the front of the file?” part of the optimisation question -
 201 though inspection of chunking is a key part of the workflow necessary to determine whether or
 202 not a file really is optimised for cloud usage.

203 Author contributions

204 JH designed and implemented the original software library, including all the fundamental
205 infrastructure for working with low-level HDF artifacts. BM, WDN and BH made contributions
206 to earlier versions of the library. BNL led the overall refresh of the library, designed and
207 implemented the support for lazy loading of chunked data, cloud and other optimisations,
208 and wrote the paper (with input from the other authors). EC, DH, BM, KM and VP made
209 contributions to the recent versions of the library.

210 Acknowledgements

211 Most of the recent developments outlined have been supported by the UK Met Office and
212 UKRI via 1) UK Excalibur Exascale programme (project ExcaliWork), 2) the UKRI Digital
213 Research Infrastructure programme (project WacaSoft), and 3) the national capability funding
214 of the UK National Center for Atmospheric Science (NCAS). Ongoing maintenance of pyfive
215 is expected to continue under the auspices of NCAS national capability funding.

216 References

- 217 Folk, M., Heber, G., Koziol, Q., Pourmal, E., & Robinson, D. (2011). An overview of the HDF5
218 technology suite and its applications. *Proceedings of the EDBT/ICDT 2011 Workshop on
219 Array Databases*, 36–47. <https://doi.org/10.1145/1966895.1966900>
- 220 Hassell, D., & Cimadevilla Alvarez, E. (2025). *Cmip7repack: Repack CMIP7 netCDF-4
221 datasets*. Zenodo. <https://doi.org/10.5281/zenodo.17550920>
- 222 Rew, R. (2013). Chunking data: Choosing shapes. In *News @ Unidata*. [https://www.unidata.ucar.edu/blo
veloper/en//entry/chunking_data_choosing_shapes](https://www.unidata.ucar.edu/blo
223 veloper/en//entry/chunking_data_choosing_shapes).
- 224 Rew, R., Hartnett, E., & Caron, J. (2006). NetCDF-4: Software implementing an enhanced
225 data model for the geosciences. *22nd International Conference on Interactive Information
226 Processing Systems for Meteorology, Oceanography and Hydrology*.
- 227 Stern, C., Abernathey, R., Hamman, J., Wegener, R., Lepore, C., Harkins, S., & Merose, A.
228 (2022). Pangeo forge: Crowdsourcing analysis-ready, cloud optimized data production.
229 *Frontiers in Climate*, 3. <https://doi.org/10.3389/fclim.2021.782909>