AMERICAN METEOROLOGICAL SOCIETY
45 Beacon Street, Boston, Massachusetts 02108-3693 USA
TEL: 617-227-2425
FAX: 617-742-8718
E-mail: amsinfo@ametsoc.org
Web: www.ametsoc.org

Best Practices for Data Management

A Best Practice Statement of the American Meteorological Society

(Adopted by the AMS Council 27 May, 2025)

Why best practices are necessary

Research, education, and subsequent applications for societal use in the atmospheric and related sciences have consistently relied on data. Given the expansion and sophistication of the data networks, increasing volumes of both observations and model data, and the need for analysis, it is critical to consider data management strategies and accepted practices for optimal use. Data management practices can better facilitate collaboration and reuse of data. Furthermore, with public access mandates from federal agencies concerning dissemination and sharing of research results including a data management plan, investigators are expected to share their data with the greater research community, and for that matter, the public. Despite the diversity of data management practices, which depend on factors including project type and data volume, there are few established standards for data management. As such, it would be beneficial to the community to develop accepted practices with a focus on principles of data management and general guidelines. The AMS Board on Open Science, Data, and Software advises and serves the Society on issues related to data stewardship including coordination of activities and services to enhance functional and meaningful access to and use of data. The Board on Best Practices has reviewed the most recent AMS Open Data Statement (2019) to ensure coordination and compatibility with this document. The

Society recognizes the need to provide guidance in the area of data management for the community of researchers, educators, practitioners, and others who either produce or use data in different contexts.

Who should observe these best practices

The best practices outlined below aim to provide guidance for effective data management practices for the community of research scientists and educators, as well as for all that produce data or provide access to data. They are intended for a broad audience and are informed by other organizational guidelines gathered from the community. The purpose of this best practices document is to provide guidance not only for making data more easily understood and more usable, but also for addressing issues related to data sharing and management, such as intellectual property, provenance, privacy, and effective preservation.

Recommended best practices

Data management includes a broad spectrum of activities in the data lifecycle, including data collection, governance, sharing, storage, and archival practices. From a best practice viewpoint, the following principles and guidelines on data management need to be considered in order to adhere to community practices:

1. Data access plan: Devise a plan for how users will be able to access and retrieve data, supporting the access points such as setting up and maintaining a data portal, and the practice and policies regarding the release of data.
2. Data costs: Identify costs for open data access, management, and long-term storage and preservation.
3. Data products: List and describe the types of data, data products, and formats that will be generated, including software and curricular materials.
4. Data formats: Describe the format in which the data or products are stored to indicate the context in which the data were produced. This component should include a mechanism to indicate whether the dataset is complete and when and

how the dataset has changed. It is critical that data formats follow accepted community standard formats, if possible, through a self-documenting format that supports interoperability and reuse.

5. Intellectual property: Describe potential licensing, copyright, restrictions on data access, and tangible research property.
6. Quality assurance (data integrity and source): Describe the processes applied for quality assurance. Identify the initial data source and describe any inclusion of other data sources.
7. Data from other sources and provenance tracking: Identify and describe any data used from other sources that may be a part of the original dataset and clarify whether there are any restrictions related to sharing of this data.
8. Reuse, re-distribution, and production of derivatives: Describe policies regarding the use of data provided via general access or sharing, and conditions for the use of data in other settings, if applicable.
9. Data preservation and archiving: Describe whether, how, and where data will be archived and how preservation will be handled. This should include protection of data in an environment for long-term access and reuse, transfer of data to different format(s) in response to changes in technology, and accessible indexing of data. This component may also contain guidelines for a routine process to guarantee data integrity associated with appropriate metadata to ensure sustained access.
10. Data Governance: Identify which policies (e.g., Public Access to Research Results) from which entities (agencies, professional societies, publishers, etc.) will guide the data management plan and its implementation. Describe the established project timeline for compliance with policy conventions.
11. Data used in publications should be cited appropriately and archived in a repository accessible to the community, either in raw or interim formats or as final publications with a permanently assigned identifier (such as a DOI).

References

American Meteorological Society Statement on Full, Open, and Timely Access to Data (2019): https://www.ametsoc.org/index.cfm/ams/about-ams/ams-statements/statements-of-the-ams-in-force/full-open-and-timely-access-to-data/

American Geophysical Union Geosciences Data Leadership: https://www.agu.org/Learn-About-AGU/About-AGU/Data-Leadership

National Science Foundation GEO Data Management and Sharing Plan Guidance: https://www.nsf.gov/geo/data-management-sharing-plans

Recommendations to Improve Downloads of Large Earth Observation Data (based on contributions from Amazon, Google, NASA, and Microsoft) (2018): https://datascience.codata.org/articles/10.5334/dsj-2018-002/

NASA Data Management Plan: https://www.earthdata.nasa.gov/engage/data-management-guidance

Unidata Data Management Resource Center: https://www.unidata.ucar.edu/data/dmrc/

Unidata Best Practices for Data Management: https://www.unidata.ucar.edu/data/dmrc/#best_practices

[This statement is considered in force until ***2030 DATE*** unless superseded by a new statement issued by the AMS Council before this date]