

NextG-Climate Science EUREC4A_OA Data Policy

Document history:

date	comment	name
29. July 2020	Draft document created	Johannes Karstensen

NOTE: This policy is part of Milestone 13: “Data Policy agreed”

Introduction

EUREC4A_OA, a global multidisciplinary climate research project, requires a wide range of data and needs data centres to ingest, quality control, archive and distribute these data. The EUREC4A_OA data policy provides guidelines for how these data should be handled in a consistent manner so as to achieve the project’s scientific objectives. The policy aims to ensure widespread access through the free and unrestricted sharing and exchange of EUREC4A_OA data and metadata. EUREC4A_OA data policy is intended to be fully compatible with IOC [1], WMO [2], GCOS [3] and GEOSS [4] data principles.

Definitions used in the EUREC4A_OA Data Policy

1. EUREC4A_OA data

“EUREC4A_OA Data” consists of directly observed data, derived data, gridded fields, and other data products generated and/or used within EUREC4A_OA to further its scientific goals. EUREC4A_OA should strive to ensure that all data relevant to EUREC4A_OA are accessible freely and without restriction, including those collected by other projects and programmes.

2. Metadata

Metadata is defined as the descriptive information such as content, quality, condition that characterizes a set of measurements.

EUREC4A_OA Data Policy and Principles

For EUREC4A_OA to succeed, high-quality data and metadata need to be collected, processed and exchanged without significant delay in a free and unrestricted manner. This was recognized

in the EUREC4A_OA Proposal. The EUREC4A_OA data policy is enshrined in the EUREC4A_OA data principles below:

1. Free and unrestricted exchange

All EUREC4A_OA data should be made available freely and without restriction. “Freely” means at no more than the cost of reproduction and delivery, without charge for the data itself. “Without restriction” means without discrimination against, for example, individuals, research groups, or nationality.

2. Timely exchange

EUREC4A_OA investigators should make data available voluntarily and with minimal delay, to maximize their value to EUREC4A_OA. In cases where extensive post-processing of delayed mode data is needed before a final research quality data set can be generated, early release of a preliminary version of the data is required.

3. Quality control

EUREC4A_OA investigators retain the primary responsibility for the quality of the data they produce and distribute. Data originators and those generating climate data products are required to ensure that their data meet international quality standards wherever possible. Appropriate reference for quality control procedures applied to the data should be provided.

4. Metadata

Metadata are required to enable the use of data without ambiguity or uncertainty. Metadata for EUREC4A_OA data sets will be developed and managed in accordance with international standards.

5. Preservation of data

Long-term survival, integrity, and access to EUREC4A_OA data must be preserved for future generations. Internationally agreed standards should be used for the acquisition, processing, and final archival of data and metadata. Data distributed in real and near-real time should, wherever possible, be replaced in a delayed mode after it has undergone quality control and full documentation. Use of accredited Global Data Assembly Centres (GDACs) is strongly encouraged for long-term storage of observational data from the EUREC4A campaign. Specific repositories for model data shall be used.

6. Plan for reuse in reanalysis

While datasets will be used individually and in combination for research purposes, the sum total of all EUREC4A_OA and EUREC4A_OA-relevant data will have great value in reanalysis activities. To aid this, uniformity of data format and quality should be a high priority.

7. Easy access

EUREC4A_OA encourages the use of the most recent advances in communication to ensure widespread access to data collected under auspices of the programme.

8. Use of existing national and international mechanisms and centres

Where feasible, EUREC4A_OA will use existing national and international mechanisms for the exchange and storage of oceanic and atmospheric data, and build on the data management structure of existing programmes. Specifically the AERIS website is one such data centre. In this way, the effectiveness of the data system will be improved by reducing redundancy and duplication and identifying opportunities and system economies, with financial costs minimized.

9. Reporting Requirements

Data and metadata should be submitted to recognized data assembly centers as well as to appropriate national and international archival institutions so that the collected information may be safeguarded for future analysis. Inventories of data and related information should be readily accessible and updated as needed on a routine basis.

References

- [1] IOC Data Policy (<http://ioc3.unesco.org/iode/contents.php?id=200>)
- [2] WMO Resolution 40 (Cg-XII; see <http://www.nws.noaa.gov/im/wmor40.htm>)
- [3] Implementation plan for the Global Observing System for Climate in support of the UNFCCC, 2004; GCOS – 92, WMO/TD No.1219.
- [4] Global Earth Observation System of Systems GEOSS 10-Year Implementation Plan Reference Document (Final Draft) 2005. GEO 204. February 2005.

EUREC⁴A Policy (see <http://eurec4a.eu/data-simulations/data>)

EUREC⁴A intends to follow the open data policy established for NARVAL. This means that we do not intend to police the use of, or access to, the data. Instead we rely on the good will and scientific integrity of those interested in using the data to respect the efforts of those whose were responsible for its collection by adhering to the two guidelines below:

1. No redistribution: Groups who are using the data are asked not to redistribute this data, but rather refer those interested in accessing the data to the original source, and these guidelines for the data use.

2. Coordinated Use and Acknowledgement: Unlike in some other contexts the EUREC⁴A data is made available through the individual efforts and particular funding of a number of independent research groups. Often the data that is collected represents the fruits of many years of effort to develop, certify and operate an instrument. This effort is often connected to individual research projects, and even PhD projects. For this reason we consider it a matter of scientific integrity for users to respect the priority of those who collected the data, by coordinating your planned use with the instrument PIs (listed below), and making an effort to explicitly and clearly acknowledge the origin of the data when it is presented to a broader community in the form of publications or presentations.

3. For publications that include EUREC⁴A data, include the following in your acknowledgements: The data used in this publication was gathered in the EUREC⁴A field campaign and is made available through [insert name of the institute]. EUREC⁴A is funded with support of the European Research Council (ERC), the Max Planck Society (MPG), the German Research Foundation (DFG), the German Meteorological Weather Service (DWD) and the German Aerospace Center (DLR).

4. In addition specific data may be bound by agreements associated with the platforms from which they were collected, for instance the data policies of the German Research Vessels, or HALO