Discussion of Long Action 4

Action D3: Improving discovery and access to data and metadata in Global Climate Data Centres

Imke Durre - 19 September 2024

Review of Activities and Implementation

Activities

1. Support the development and maintenance of software tools that assist users in utilizing publicly available data sets.

2. Compile, maintain, and update a list of datatypes for which single global databases and improved web access are needed for a broad range of applications.

3. Based on the identified priorities, facilitate and support the creation of global databases and systems for accessing various atmospheric, terrestrial, and oceanic parameters for which such archives of interoperable, easily accessible data holdings currently do not exist. The emphasis should be on parameters, data sources, and compilations for ECVs and/or for which a broad set of uses across multiple sectors has been identified. Examples include the following:

a. a global database of current and historic observational field campaign and research data, particularly those that are used for calibration and validation of satellite-based instruments or that fill gaps in the historical records;

b. improve systems to store and access ocean biogeochemical, ocean oxygen, and groundwater data

c. a global database of thunder reports that can be used to extend the record of lightning observations backward in time.

Implementers:

Global Data Centres, Funding Agencies, Organizations engaged in developing open-source software

Means of Assessing Progress:

1. Availability of open-source software tools for accessing data in common data formats
2. Prioritized list of datatypes for which improved access to a single global integrated, harmonized database is clearly needed
3. Inventories, integrated databases with consistent data formats and processing procedures, and efficient web access for observational field campaign and research data; ocean oxygen data; groundwater observations; and thunder reports.
4. Prioritized list of other data types for which such integrated datasets are needed.

Connections to other actions we have discussed:

D1, D2: Global Data Centre requirements and existing global repositories
C3: In-situ data records

Considerations

Consideration 1: ECVs in global collections

Prerequisite?: Outcome of Action D2 to determine which ECVs are currently included in well-known global collections and which are not?

Linkages: global data repositories for raw ECV observations -> integrated data sets -> gridded datasets

Consideration 2: Interconnections

Global repositories, golobal data collections, ECVs, and data discovery/access tools are all interconnected.
Can we consolidate our information collection efforts to inform multiple IP Actions?

Consideration 3: Tools for discovery and access

What do we think of as discovery/access tools?
What are some examples of such tools?

Are we aware of some examples of gaps that hinder access?
Based on these examples, how should we inventory these tools and associated gaps?

Consideration 4: Specific data types listed in activity 3

Field campaign data: To what extent can/should AOPC act on this specifically? If so, how?
Ocean biological data and groundwater data: not in AOPC's purview
Global thunder day database: Addition action by AOPC needed at this point?

Overarching questions

How does AOPC/GCOS meaningfully "support" the activities in this Action?

How do we best track the measures of progress in this Action?
1. Availability of open-source software tools for accessing data in common data formats
2. Prioritized list of datatypes for which improved access to a single global integrated, harmonized database is clearly needed
3. Inventories, integrated databases with consistent data formats and processing procedures, and efficient web access for observational field campaign and research data; ocean oxygen data; groundwater observations; and thunder reports.
4. Prioritized list of other data types for which such integrated datasets are needed.

Next steps?