Proposed definition and guidance regarding Climate Data Records and Interim Climate Data Records

# Preamble

To support the generation of valuable environmental data records in support of Earth system science, services and governmental obligations, space agencies may undertake, commission or otherwise support reprocessing of their holdings of level 1 data to derive environment information.

In the domain of climate applications, the outcome of such reprocessing and improvement activities is creation of a **Climate Data Record** (CDR), created in support of climate science, services and decision making. (The term environmental data record has been used as a broader category that would include CDR, but is not defined here.)

Extending a CDR routinely in time with full consistency may be possible, but often extension in time will require provision of an **Interim Climate Data Record** (ICDR).

The concept of the **Fundamental Climate Data Record** (FCDR) is used here. See {ADD CROSS REFERENCE}.

# Climate Data Record

## Short Definition:

A climate data record consists of a long, stabilised record of uncertainty-quantified retrieved values of a geophysical variable relevant to characterise Earth’s climate, together with all ancillary data used in retrieval and uncertainty estimation. The CDR is linked to (an) underlying fundamental climate data record(s).

## Explanation and further guidance:

1. This CDR definition is intended to define the target objective (goal) for CDR activities. In the following points, “should” is used to indicate what is necessary to achieve this CDR definition. Importantly, initiatives that deliver societal and scientific benefit from reprocessing and improvement activities but meet this goal only partially should not for that reason be discouraged or inhibited.
2. A CDR is a record that is specifically useful for quantifying aspects of Earth’s climate system (but may also support non-climate applications).
3. The CDR may be provided at more than one level of processing: at the full resolution and with the geolocation of the underlying FCDR (level 2), gridded and/or merged (level 3), and/or gap-filled (level 4). The higher levels should be derived from and consistent with the lower-level CDR.
4. The CDR should cover a duration long enough to quantify climate variability and change. No single duration can be specified, since this depends on the variable, the uncertainty and stability of observation, and the climate science context. Usually, the label CDR would not be applied for satellite records shorter than 10 years. During periods of overlaps of sensors constituting a measurement series, all overlapping data should be included in the CDR at level 2, since overlaps should be exploited to maximise stability, including in derived higher-level CDRs.
5. The CDR should be stabilised, which means it should be harmonised or otherwise improved so as to maximise observational stability and/or meet climate requirements for observational stability.
6. The complete CDR should include unambiguous traceability to the underlying lower-level data (FCDRs) and the ancillary data from which geophysical values and their estimated uncertainty derive.
7. The data in the CDR should be located in time and space using a clearly defined co-ordinate system.
8. Quantitative uncertainty information should be provided with the CDR observations. Examples of uncertainty information include: standard uncertainty, fractional uncertainty, and error covariance matrices. Uncertainty may be provided as data arrays or (where valid, in order to minimise data volumes) as parametric expressions to calculate uncertainty from other data in the CDR. Uncertainty information should be provided per file or per datum as necessary to represent any significant variability of uncertainty, while taking account of data volume. Uncertainty estimates should cover all important sources of error, and be validated, documented and traceable.
9. Flags relating to assessment of data quality should be made available in the CDR together with the observations and their uncertainty.

# Interim Climate Data Record

## Short Definition:

An interim climate data record is a seamless extension in time of a climate data record, in the processing of which a stated level of inconsistency is present as necessary to meet the required timeliness.

**Definition:**

*An Interim Climate Data Record (ICDR) regularly extends in time a Fundamental or Thematic Climate Data Record using a system having optimum consistency with and lower latency than the system used to generate the FCDR or TCDR.*

## Explanation and further guidance:

1. The ICDR extends in time the CDR with no temporal gap.
2. The scientific principles, algorithms and software implementation of the processing of the ICDR are identical to or demonstrably compatible with those that generated CDR.
3. The presentation of the ICDR states its origin and interim status relative to an identified CDR.
4. The substantive data content and formatting of CDR and ICDR are compatible, such that users can seamlessly use both data sets in their applications.
5. Various sorts of inconsistencies between CDR and ICDR may arise, such as the need to use near-real time level 1 data, updated software environments, interim values of stabilisation parameters, etc. Information is made available to users describing these inconsistencies and their potential impact on the ICDR compared to CDR.