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

Can we simplify this by making the FDR a single instrument single satellite data record? This would constitute the best Level-1 achievable for each sensor. FCDR would then add the dimension “sufficient length” including the harmonisation or homogenisation (if needed). There might be a few exceptions, e.g., MODIS that constitute a single sensor/satellite FCDR.

# Preamble

To support the generation of valuable environmental data records in support of Earth system science, services and governmental obligations, space agencies reprocess mission and multi-mission archives to improve the characteristics of the level 1 data.

The outcome of such reprocessing and improvement activities is creation of a **Fundamental Data Record** (FDR) from which higher-level data records may be derived.

A specific case is the creation of **Fundamental Climate Data Records** (FCDRs), which is an FDR meeting some additional criteria, that can be used as a baseline for long-term **Climate Data Records** (CDRs) in support of climate science, services and decision making. The FCDR definition therefore includes and is more specific than the FDR definition.

# Fundamental Data Record

## Short Definition:

An FDR consists of a consistently [re]processed single sensor record of [ideally uncertainty-quantified] observations that are calibrated to physical units and located in time and space. [, together with all ancillary and lower-level instrument data used to calibrate and locate the observations and to estimate uncertainty.]

## Explanation and further guidance:

1. This FDR definition is intended to define the target objective (goal) for FDR activities. In the following points, “should” is used to indicate what is necessary to achieve this FDR 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. This in particular applies to historical satellite data for which information to characterise uncertainty may not exist.
2. The FDR should be consistently [re]processed or otherwise improved so as to maximise observational temporal consistency, in support of more consistent derived data records.
3. The complete FDR should include lower-level observations with the means to derive calibrated quantities from these (such as Earth view counts together with the appropriate parameters for transformation to physical units). This enables efficient re-estimation of calibrated quantities should further research improve understanding of the sensors’ calibration. Users may additionally require a consistent version of the FDR comprising calibrated measured values in their physical units, particularly if the conversion to calibrated values has some complexity. The FDR should include all further ancillary data, telemetry, etc., used in the process of calibration (or in deriving calibration parameters). “Together with” does not necessarily mean “in the same file”, but the FDR should provide as far as possible unambiguous traceability between the lower-level data, the calibrated data and the corresponding ancillary data.
4. The data in the FDR should be located in time and space using a clearly defined co-ordinate system, exploiting the best available estimates of orbital elements (where relevant).
5. Quantitative uncertainty information should be provided with the FDR observations. The form of uncertainty information should reflect the nature of the sensor and its error characteristics, as well as the requirements for quantifying uncertainty in derived higher-level data records. 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 FDR. Uncertainty information should be provided per orbit/file, per scan 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. Data underlying or ancillary to the uncertainty estimates should also be provided together with the uncertainty information in the FDR.
6. Flags relating to assessment of data quality and instrument status should be made available in the FDR together with the observations and their uncertainty. To increase usability, summary flags may be provided that give simple guidance to users about the quality status of pixels based on expert understanding of which of the flags and instrument conditions indicate questionable or invalid data.
7. The FDR format should be consistent,in particular if various similar sensors form a time series. Versioning, a unique identifier and informative metadata should be implemented, including links from calibrated observations to ancillary data (if not in the same files).

# Fundamental Climate Data Record

## Short Definition:

An FCDR consists of a sufficiently long, stabilised in most cases multi-instrument record of [ideally uncertainty-quantified] sensor observations that are calibrated to physical units and located in time and space. [, together with all ancillary and lower-level instrument data used to calibrate and locate the observations and to estimate uncertainty.]

## Explanation and further guidance:

1. This FCDR definition is intended to define a target objective (goal) for FCDR activities. In the following points, “should” is used to indicate what is necessary to achieve this FCDR 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. An FCDR is a record that is specifically useful for quantifying aspects of Earth’s climate system (but may also support non-climate applications).
3. The FCDR should cover a duration long enough to quantify long-term climate variability (minimum interannual variability) and change. No single duration can be specified, since this depends on the variable, the uncertainty and stability of observation, signal size in the climate system, and the climate science context. Usually, the label FCDR 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 FCDR, since overlaps should be exploited for harmonisation and when deriving CDRs to maximise stability.
4. The FCDR should be stabilised, which means it should be harmonised or otherwise improved so as to maximise observational stability. Harmonisation is the recalibration of the sensor observations, using a stated reference and/or overlaps with other sensors in the series. The purpose of the harmonisation [recalibration] is to bring consistency between sensors given the known (best estimate) differences in instrument characteristics (e.g., spectral response functions) between sensors. A harmonised (or otherwise stabilised) FCDR should enable more stable CDRs to be derived. The word stabilised in the definition should be considered to encompass this technical definition. Other approaches to stabilisation may be valid, such as homogenisation.
5. The complete FCDR should include lower-level observations with the means to derive calibrated quantities from these (such as Earth view counts together with the appropriate parameters for transformation to physical units). This enables efficient re-estimation of calibrated quantities should further research improve understanding of the sensors’ calibration. Users may additionally require a version of the FCDR comprising calibrated measured values in their physical units, particularly if the conversion to calibrated values has some complexity. In both cases, the observations should be provided together with all further ancillary data, telemetry, etc., used in the process of calibration (or in deriving calibration parameters). “Together with” does not necessarily mean “in the same file”, but the FCDR should provide as far as possible unambiguous traceability between the lower-level data, the calibrated data and the corresponding ancillary data.
6. The data in the FCDR should be located in time and space using a clearly defined co-ordinate system, exploiting the best available estimates of orbital elements (where relevant).
7. Quantitative uncertainty information should be provided with the FCDR observations. The form of uncertainty information should reflect the nature of the sensor and its error characteristics, as well as the requirements for quantifying uncertainty in derived CDRs. 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 FCDR. Uncertainty information should be provided per orbit/file, per scan 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. Data underlying or ancillary to the uncertainty estimates should also be provided together with the uncertainty information in the FCDR.
8. Flags relating to assessment of data quality and instrument status should be made available in the FCDR together with the observations and their uncertainty. To increase usability, summary flags may be provided that give simple guidance to users about the quality status of pixels based on expert understanding of which of the flags and instrument conditions indicate questionable or invalid data.
9. The FCDR format should be consistent across the various sensors of the series. Versioning, a unique identifier, and informative metadata should be implemented, including links from calibrated observations to ancillary data (if not in the same files).

# FDR/FCDR distinction

An FCDR is an FDR that is intended to support the generation of climate data records and that meets additional criteria:

1. An FCDR is of sufficient length to usefully quantify climate variability and change. An FDR may have suitable properties, but be too short to be informative about climate variability and change.
2. The contents of an FCDR are relevant to climate applications (directly or indirectly). A record containing data relevant only to non-climate applications is an FDR, irrespective of length.
3. Whereas an FDR must be a “consistently reprocessed” record, an FCDR is defined as a “stabilised record”. Stabilisation implies consistent reprocessing, and adds more specific requirements on recalibration or harmonisation of the sensor series, as noted in the explanatory guidance. The purpose is to maximise the observational stability of products derived from the FCDR in support of securely quantifying long-term climate changes.