CORE-CLIMAX

Structured Process for the Generation of Climate Data Records

Deliverable D2.26

Doc.No.	:	CC/EUM/REP/15/004
Issue	:	v1
Date	:	26 July 2015
WBSDBS:		

EUMETSAT
Eumetsat-Allee 1, D-64295 Darmstadt,
Germany
Tel: +49 6151 807-7

Document History

Contributors	*Organisation*	*Action*	*Date*
Jörg Schulz	EUMETSAT	Draft outline created	2 April 2015
Jörg Schulz	EUMETSAT	Initial version	1 May 2015
Jörg Schulz	EUMETSAT	Final Version	24 July 2015

Table of Contents

1 INTRODUCTION

1.1 Purpose and Scope

The purpose of this report is to document a structured process to generate Climate Data Records (CDR) including GCOS Essential Climate Variable (ECV). The intended readership is data record producers that may use this document to improve their own process to create CDRs.

The report addresses the need outlining a structured process for CDR generation that should enable and ensure high quality of CDRs used in Climate Services and research.

1.2 Reference Documents

RD.1	Dowell, M., P. Lecomte, R. Husband, J. Schulz, T. Mohr, Y. Tahara, R. Eckman, E. Lindstrom, C. Wooldridge, S. Hilding, J. J. Bates, B. Ryan, J. Lafeuille, and S. Bojinski (2013): Strategy towards and architecture for climate monitoring from space.	[available from: www.ceos.org, www.wmo.int/sat, www.cgms-info.org].
RD.2	Core-Climax System Maturity Matrix Instruction Manual (2014)	CC/EUM/MAN/13/002 [available at www.eumetsat.int]
RD.3	GCOS-143, 2010: Guideline for the Generation of Datasets and Products Meeting GCOS Requirements, An update of the “Guideline for the Generation of Satellite-based Datasets and Products meeting GCOS Requirements” (GCOS-128, WMO/TD-No. 1488), including in situ datasets and amendments, 12 pp.	Available at: http://www.wmo.int/pages/prog/gcos/Publications/gcos-143.pdf.

2 EUROPEAN AND INTERNATIONAL CONTEXT

Developing ECV climate data records poses many challenges because of the varied use of climate data, the complexities of data record generation, and the difficulties in sustaining the activities over extended periods of time. To ensure an efficient use of resources to produce climate data records it is important to have a structured process for generating them.

A global group of space agencies has developed the architecture for climate monitoring from space [RD.1] that considers the whole value adding chain from making measurements to the development of policy and decision making. The report by Dowell et al. (2013) [RD.1] defined a so called logical view that serves the promotion of a common understanding of the implementation implications of meeting the various climate monitoring requirements. It represents the functional and data-flow implications of the requirements baseline as a set of interlinked functions and associated data-flows. Leaving aside performance considerations (e.g. accuracy, uncertainty, stability, coverage etc.), the logical view could be considered as the "target" for a climate monitoring system and, in the sense that it is applicable to all ECVs, this representation is generic. As this view is intimately tied to the requirements baseline (and not to the physical implementation of a climate monitoring system) this view is as stable as the requirements baseline and, once established, should only need to be updated when the functional aspects of the requirements change. The logical view is "end-to-end" and, as a result, a four-pillar logical architecture was proposed (Figure 1).

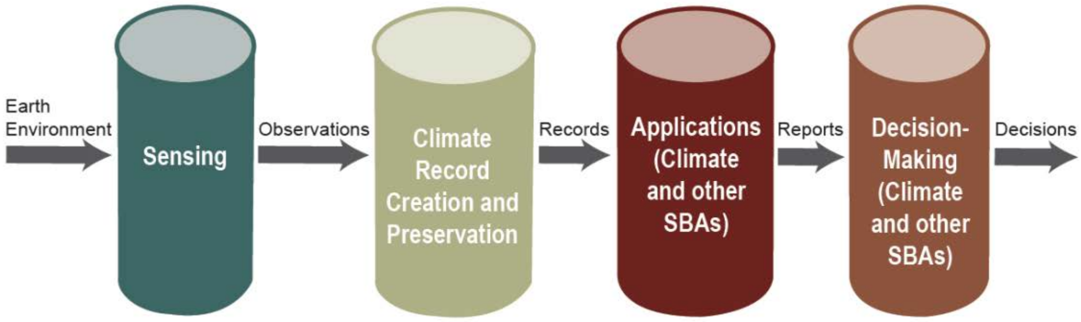

Figure 1: Main components of the logical view from [RD.1].

The information flow in [RD.1] starts with the sensing of the Earth environment (by EO satellites). The resultant observations are then assembled, processed and converted to climate records. These records are then used by the relevant applications to generate reports that are, in turn, used by decision-making entities (including policy-makers) to decide on a course of action.

This high-level, conceptual representation has been specifically generated to highlight the main structural elements of the logical view. However, in order to support the two identified usage scenarios, it is necessary to "drill down" within each of the pillars in order to expose their constituent elements.

To support the international activities described above and the establishment of the Copernicus Climate Change Service one major objective of the CORE-CLIMAX project is to provide a deeper view into the second pillar of the logical view, in particular, for climate data records derived from space based observing systems.

3 METHODOLOGY TO DISPLAY THE STRUCTURED PROCESS

3.1 The approach for displaying the process

The resulting decompositions starting with Figure 1 can be quite complex, with many potential data-flows, methodologies and associated tools. In order to ensure consistency of approach, and to be able to make use of off-the-shelf tools that are essential to manage the complexity, it was decided to adopt the IDEFØ (Integration Definition Method) standard (http://www.idef.com/IDEF0.htm) for functional modelling for the further development of this logical view following the approach used in [RD.1]. IDEFØ is a method designed to model the decisions, actions, and activities of a system. The box and arrow graphics of an IDEFØ diagram show the function as a box and the interfaces to or from the function as arrows entering or leaving the box. To express functions, boxes operate simultaneously with other boxes, with the interface arrows "constraining" when and how operations are triggered and controlled. The basic syntax for an IDEFØ model is shown in Figure 2. IDEFØ starts with a top level view which is then expanded in a cascade of diagrams that describe more detail with each level. The resulting diagrammatic representations differ somewhat in format compared to Figure 1 but nevertheless, the same high-level components are still evident.

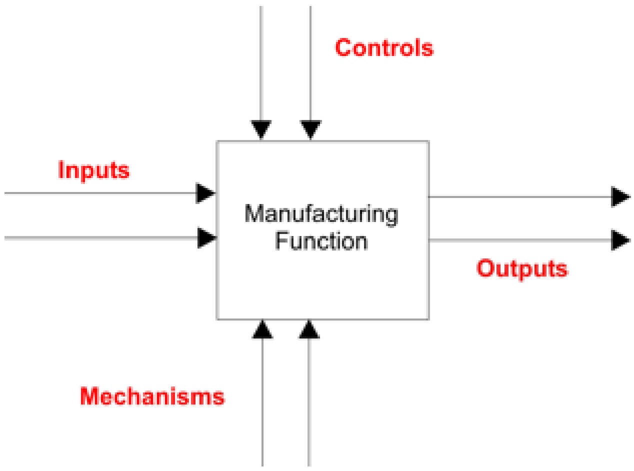

Figure 2: Basic IDEFØ boxes and arrows (from http://www.idef.com/IDEF0.htm)

3.2 The connect to the System Maturity Matrix (SMM)

The structured process outlined in this document refers to columns and sub-columns of the System Maturity Matrix [RD.2] to define needed sub diagrams within the IDEFØ representation of pillar two of the architecture for climate monitoring from space. In particular, there are individual diagrams for the documentation, archive and access as well as user feedback aspects. Within the uncertainty characterisation area the process of validation versus user requirements is specifically displayed in the following sections.

4 THE STRUCTURED PROCESS

4.1 Decomposition of the architecture for climate monitoring from space

Figure 3 shows the decomposition that exposes some of the main generic functional elements of the architecture pillars displayed in Figure 1. Pillar 2 is expanded into three independent activities that create data records for different applications. The box A3 is the classical way of creating long-term climate data records mostly used to study decadal variability and trends which is reflected in by box A7. This application is also supplied with so called climate information records, e.g., frequency of storms in a certain region or other climate index information which is often derived using model-based reanalysis.

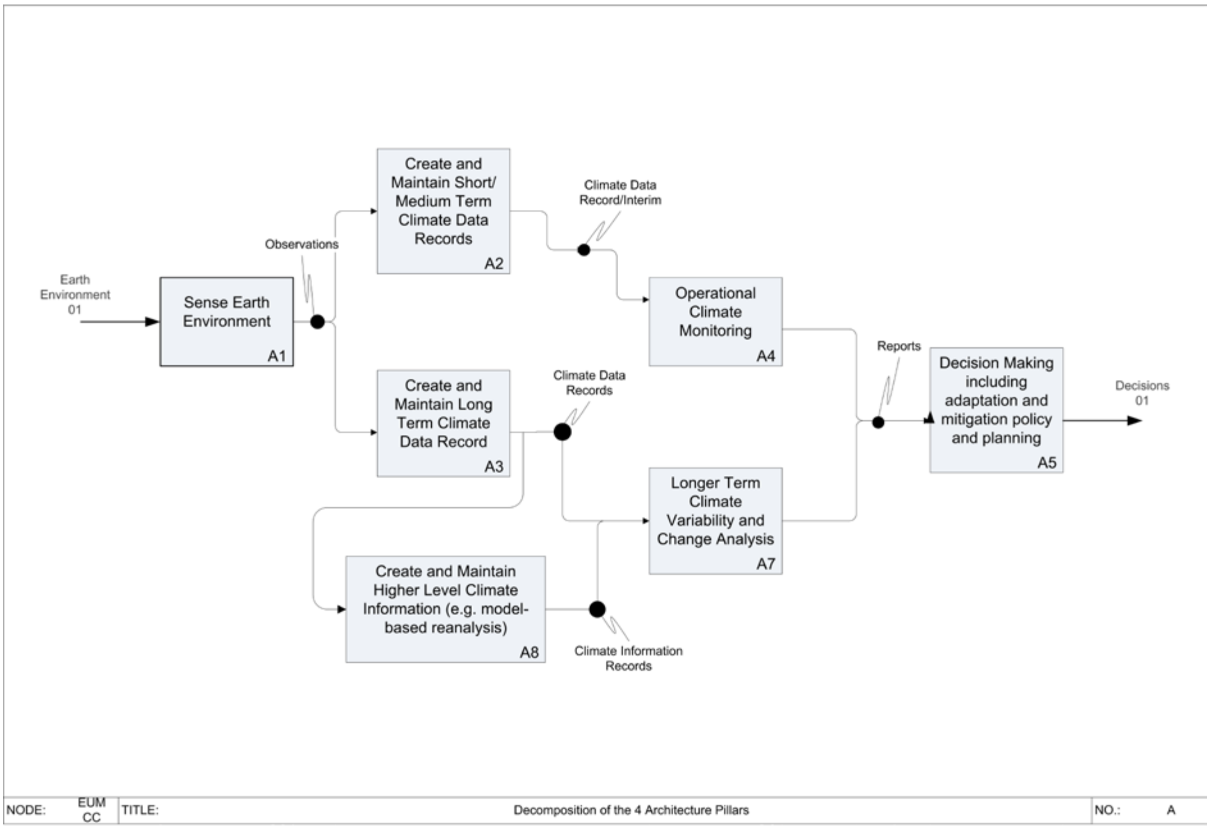

Figure 3: Decomposition of the 4 pillars of the architecture shown in Figure 1.

The application of operational climate monitoring sketched in box A4 that for instance delivers information about how the last month's weather relates to the long-term climate needs a specific provision of so called interim climate data records. The activity (function A2) produces data records, with relative high timeliness (a couple of days after observation), that are in an ideal case consistent with the long-term records produced under box A3. This goal is not completely reached today but climate data record providers in Europe are working towards it. Finally, the application boxes A4 and A5 lead to findings often presented as written reports that are used by policy and decision makers in the public but also private sectors to inform decisions impacted by climate variability and change.

The structure of this logical view has the major advantage that for each of the boxes, responsible organisations can be identified to fulfil that specific function. For instance, the creation and maintenance of climate data records (functions A2 and A3 in Figure 3) involve many activities, as shown in Figure 5, including the creation of Fundamental Climate Data Records that are best performed by space agencies operating the specific satellite sensor. In contrast, the generation of Higher-Level Climate Information Records such as climate indices, or the number of storms that make landfall in a certain region, often need the combination of both CDRs originating from space-borne and ground-based systems, as well as modelling components. Thus, such an activity might be best placed in an organisation that combines information, such as reanalysis centres, climate service centres or environmental agencies. Even at a lower level, the logical view can be employed to organise tasks within internationally-operating groups that create CDRs in a distributed way, e.g., the implementation of one retrieval algorithm at different agencies.

In the following sections this report will concentrate only on the box A3 and presents the next two levels of details downstream from Figure 3. Considering the other activities related to the architecture for climate monitoring form space is beyond the CORE-CLIMAX remit and resources. The consideration will also concentrate on satellite data products but some activities are also clearly applicable to in situ data. It will be pointed out in the descriptive text when this is the case.

4.2 Create and maintain Climate Data Records

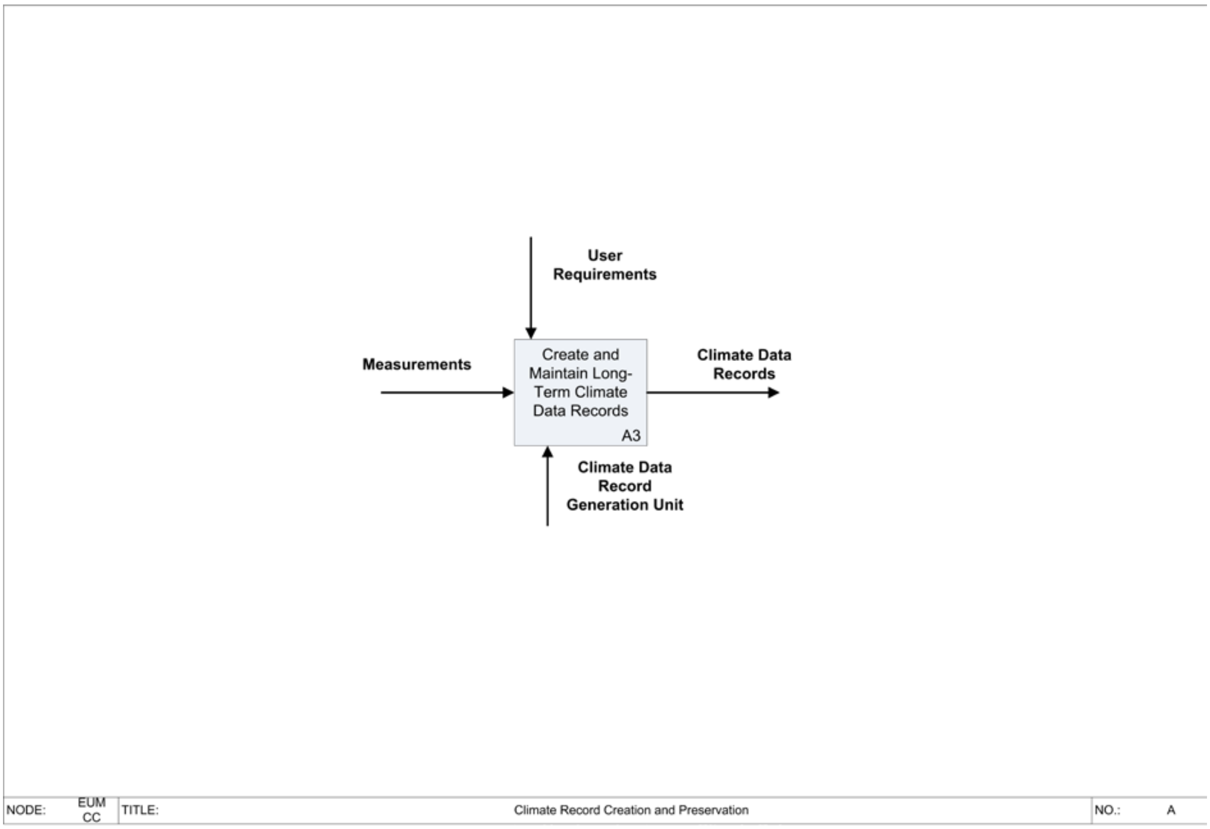

Figure 4: Top Level IDEF0 diagram for Box A3 from Figure 3.

Figure 4 shows the top level IDEFØ diagram for the second architecture pillar Climate Record Creation and Preservation which is the only one that is considered in the CORE-CLIMAX project. This pillar has a major input the measurements delivered by satellite and in situ observing systems and the major output are climate data records. The task is controlled by user requirements, e.g. those provided by GCOS but also more general requirements related to all applications relevant to climate services and climate research. The generation of climate data records is performed by specific units, i.e., operational and research institutions around the world. Although the IDEFØ representation is very suitable to display the responsibilities for the needed activities this report will not further consider this topic and it addresses more a management aspect and not the process to generate climate data records.

The process of creating Climate Data Records displayed in Figure 5 starts with the availability of measurements. The measurements are corrected, geo-located and calibrated (see function A31) to produce a product that is termed a Fundamental Climate Data Record, with the calibration parameters applied or attached being derived from a combination of internal data, in situ data and external satellite data (see function A32).

The Fundamental Climate Data Record is then converted (see Function A33) to an individual or a set of geophysical parameter(s) which are termed a Thematic Climate Data Record. Depending on the particular GCOS ECV under consideration, this TCDR may correspond directly to an ECV or, if the ECV in question is broadly defined (e.g. "Cloud Properties") the TCDR may form just one component of an ECV (see [RD.1] for further explanations of terms).

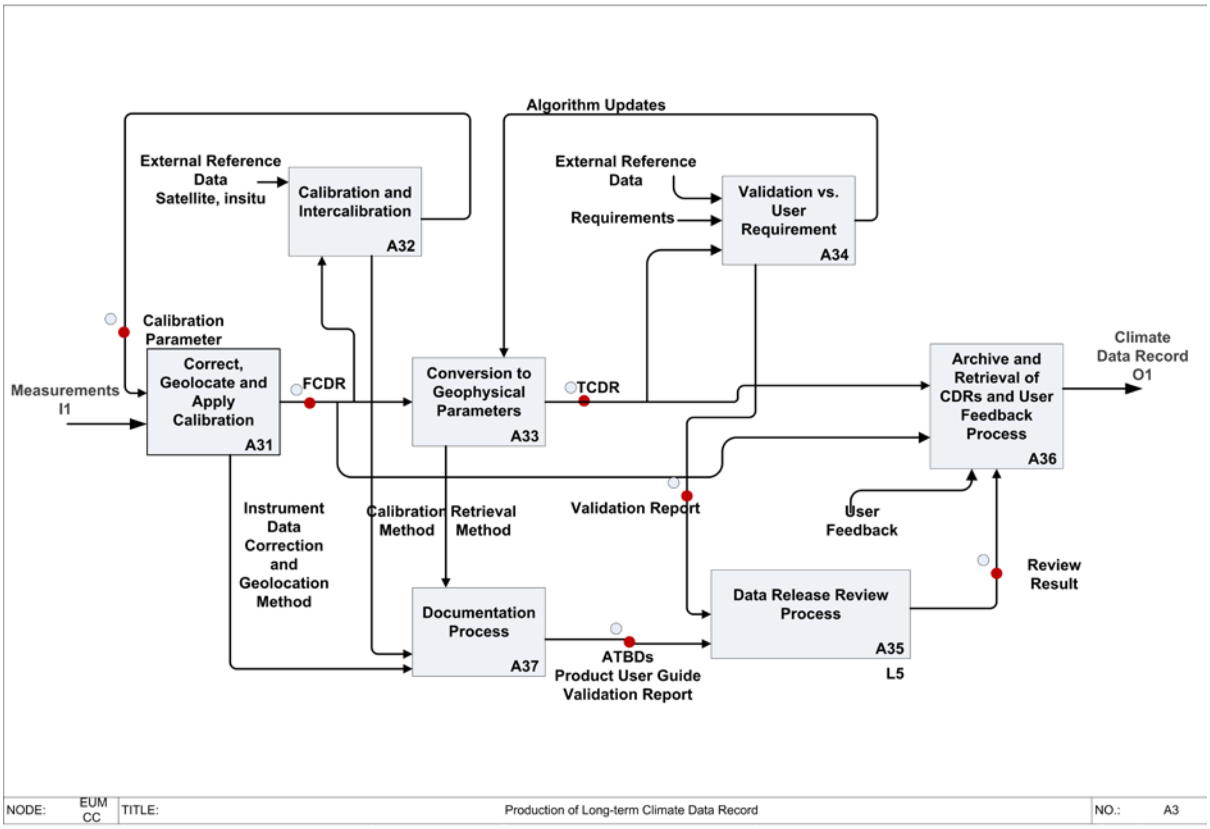

Figure 5: Decomposition of "Create and Maintain Long-term Climate Data Records" A3 shown in Figure 4.

The TCDRs and FCDRs are then validated against user requirements (see function A34) to attest fitness for purpose. If successful this is followed by archiving the data records (see function A36) together with other relevant information, such as the results of a comparison with the requirements baseline and the validation results for the data records and the results of the Peer Review described in [RD.3] – see function A35. Function A35 has attached a function L5 where L stands for Life cycle. Although not explicitly shown in Figure 5 other parts of the process creating a climate data record are subject of a review cycle being an important part of the climate data record life cycle. Section 4.3 will address the life cycle more detailed in specific set of IDEFØ diagrams.

The archived information is collectively termed Climate Data Record. When access to these Climate Data Records is required, a request is submitted and the appropriate records are then retrieved and dispatched to the requesting entity (see function A36).

It is noted that, in general, the functions depicted in Figure 5 are recursive because, when improved information becomes available (e.g. better algorithms for the generation of FCDRs/TCDRs, or improved calibration information) the measurements will be processed again to generate improved climate data records. Such processing activities are typically synchronised with the impending use of the particular climate data records in major climate-related projects, e.g., model-based re-analysis, climate model evaluation, or climate analysis for periodically appearing reports such as the annual State of the Climate in the USA. Implicit in Figure 5 is also the need to have in place a comprehensive configuration management system, to provide full document traceability of the processes and data used to derive the climate data records. It is also emphasised that, even at this lower level of decomposition, the process description is generic, as it is applicable to all GCOS ECVs. Although Figure 5 is specifically made for satellite data most of the functions (with the exception of A33 because in situ measurements are more direct) are also applicable for in situ data.

Figure 5 also shows that many requirements and guidelines such as stated in [RD.3] are embedded in activities, or inputs to activities, in the creation of climate data records. For instance, a peer review process is a mandatory part of the creation of a climate data record and an independent assessment can be seen as an input to the validation activity that assesses the quality of a data record versus the user requirements.

The following subsections provide in depth consideration for all functions of Figure 5.

4.2.1 Sensor Corrections and Geolocation

The function A31 on sensor correction and geolocation is very specific for satellite data. Sensor corrections are also developed and applied for in situ measurements but the diagram shown here only applies to satellite data. Geolocation is only needed for observing systems flying over the Earth and may also apply to sub-orbital observation systems.

The processing of raw satellite data involves computing the ortho-geolocation of each instrument pixel as well as radiometric calibration coefficients and appending them to numerical counts for later application. The process starts with raw input data and the source packet processing (function A311) unpacks and validates the source packet data and auxiliary data- It also converts data into engineering units for higher processing levels that generate the products and geolocate the data.

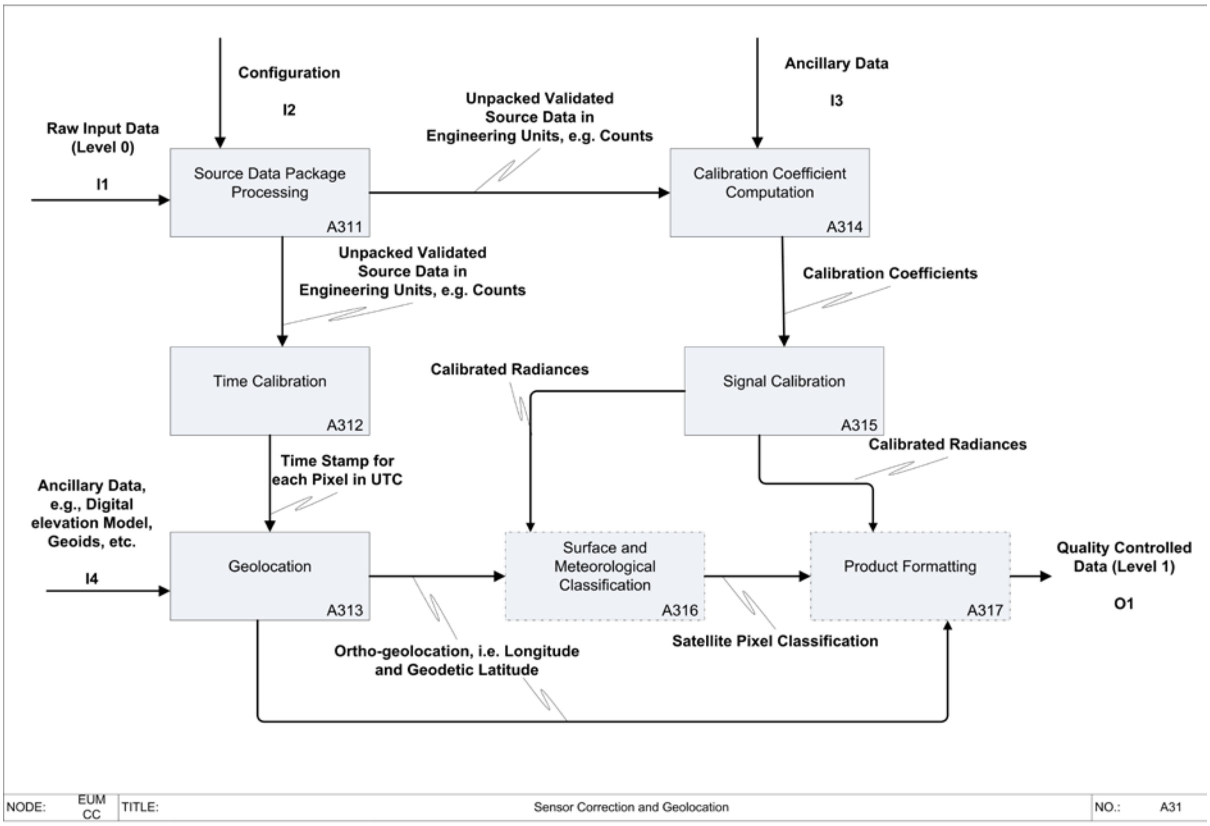

Figure 6: Decomposition of function A31 Sensor Correction and Geolocation.

Function A312 called time calibration performs the derivation of the scan time from the time stamps of each Instrument Source Package and the computation of the time stamp of each pixel (i.e. the time associated to each scan position) in UTC format. This is followed by the geolocation (function A313) computing the ortho-geolocation (i.e. longitude and geodetic latitude, both corrected for the Digital Elevation Model (DEM) and altitude with respect to the reference ellipsoid WGS84). For some satellite instruments data are also transferred to Cartesian grids with various spatial sampling.

Function A314 calculates the calibration coefficients needed to convert from engineering units into radiances. For infrared channels this is the calculation of offset and slope that describes the linear relationship between pixel count and radiance for the thermal IR channels. The parameters are determined from the black body pixel counts and the black body temperatures, and the process makes often use of look-up tables for the conversion of temperature to radiance.

Further processing of the data often consists of application of the measurement calibration, pixel classification, e.g., surface properties and cloud detection of all pixels.

Function A315 aims to calibrate numerical counts associated with each pixel using calibration coefficients derived in function A314 for all instrument channels. Function A316 often uses the image pixel latitude and longitude to derive the surface type for each image pixel. Often also a cloud detection scheme is applied to ingest this information into Level 1 data for easier application later. Finally function A317 takes all information created and packs the data into a prescribed sensor specific format. This step can include interpolation to other sensor resolutions and also integration over time into full satellite orbits.

4.2.2 Calibration and Inter-Calibration

Function A32 performing calibration and inter-satellite calibration is of high importance for long-term climate data records as it is the activity that ensures a high stability of the data record that is stitched together from a couple of similar sensors records over time each having its own behaviour that needs to be carefully studied and eventual corrections being developed. It should be noted that also observing systems exist, e.g., the radio occultation technique that does not need calibration and inter-satellite calibration activities because every single measurement is tied to a measurement of time that can be referenced to an atomic clock on Earth. However, most of the existing satellite data records, in particular, those that reach back to the 1960s and 1970s of the 20th Century need the activities described in Figure 7 to convert them into useful climate data records.

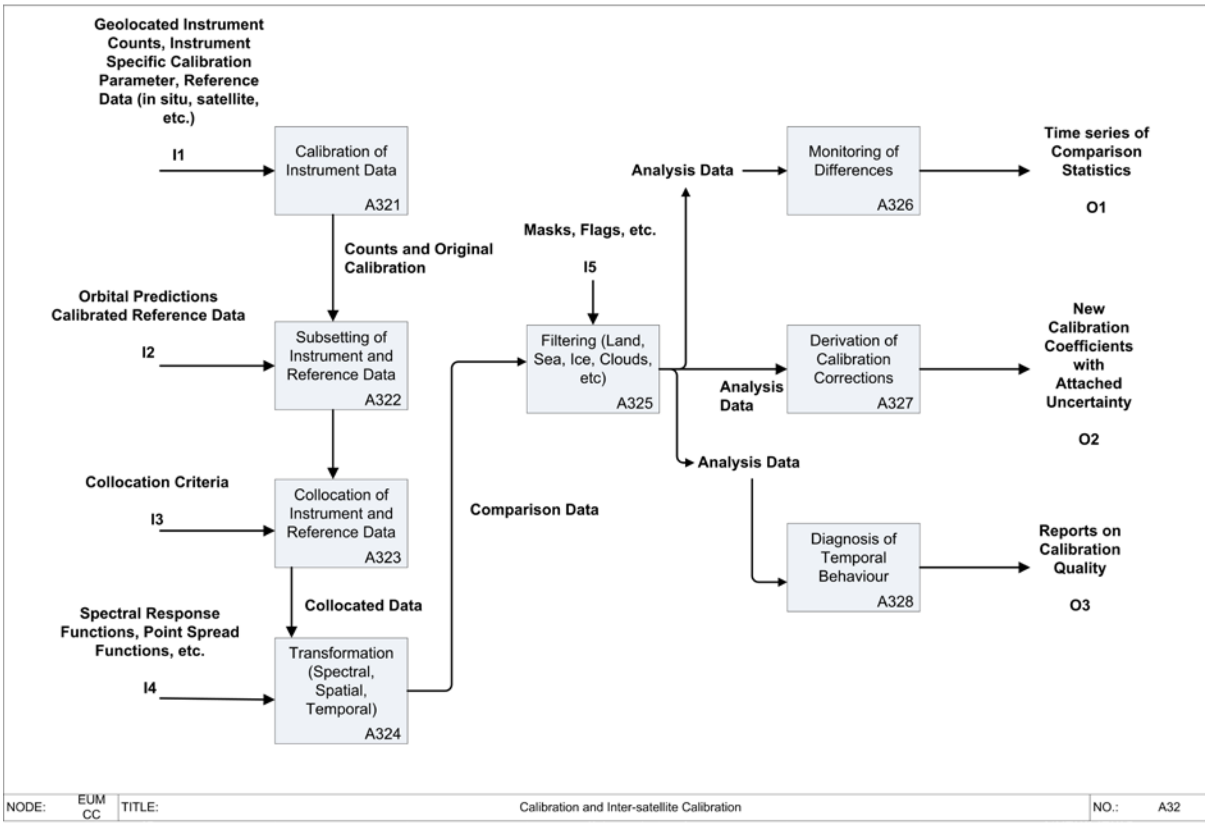

Figure 7: Decomposition of function A32 Calibration and Inter-Satellite Calibration.

The output of function A31 provides the input to function A32. It starts with the application of calibration coefficients if not already done in the function before (function A321). To perform inter-satellite calibration reference satellite data are needed where the measurements used can be inter-calibrated to. Function A322 is sub-setting data from both the reference and the actual measurements to reduce the data volume that needs to be searched for real collocations (function A323) in the next step. The collocation process is informed by collocation criteria which are specific differences in time and space in terms of underlying variability of the Earth System that can be tolerated for the calibration process. The collocated data from reference and target instruments need to be spectrally adjusted (function A324) to become comparable. Deriving a useful inter-satellite calibration can require the filtering of clouds and specific land surface properties such as ice and snow surfaces (function A325) that may represent unfavourable conditions for the inter-satellite calibration. The finally adjusted collocated data set is then used in three functions to provide a simple monitoring of differences between reference and target instrument (function A326), to derive calibration corrections for the target instrument (function A327) and to diagnose the temporal behaviour of the differences and corrections over time (function A328).

4.2.3 Conversion to Geophysical Parameter

Figure 8 shows the standard set up for the retrieval of a geophysical parameter from satellite data. It starts with the development of a processing configuration (function A331) that is controlled by the product requirements developed early on in the development life cycle (see Section 4.3.1). The result of this function is a set of processing instructions that are first used in a general input data processing (function A332) that can consist of singling out static information such as observation geometry (angles), pre-computed look up tables using radiative transfer models, but also mapping of input data to a specific grid on which the retrieval shall be performed. It is expected that the input to function A332 does also contain uncertainty information.

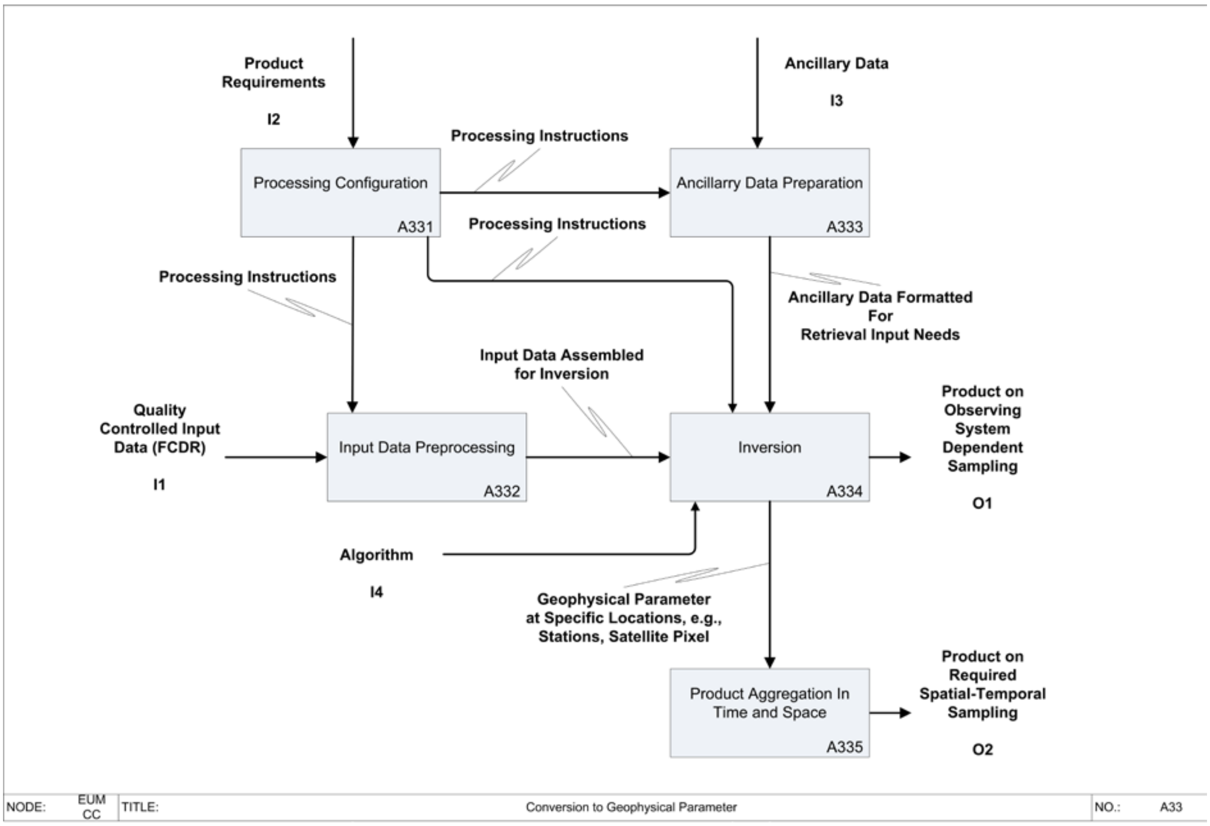

Figure 8: Decomposition of function A33 Conversion to Geophysical Parameter.

Processing instructions are also used for the preparation of ancillary data needed in the retrieval process (function A333), e.g., information needed on surface properties or for atmospheric correction. This information may also be mapped to the grid used for the retrieval which still can be the original satellite pixel representation. Function A334 is only doing the inversion process from radiance into geophysical parameter space applying a suitable algorithm. The algorithm type used depends on the combination of parameter to be retrieved, spectral range of the measurements and the type of sensor. In an ideal case the outcome of the inversion should not only contain the retrieved parameter but also the associated uncertainty that is also a result of the propagated uncertainty from the Fundamental Climate Data Record that serves as input to the process.

Depending on the product requirements the last function (A335) performs aggregation of the product in time and space and has as output the so called Thematic Climate Data Record.

4.2.4 Validation versus Requirements

Function A34, the validation against user requirements should be applied for the outputs of function A31-A33, i.e., at all levels of the data record generation. The schematic shown in Figure 9 displays a part of the whole validation process that was described in every detail in the outcome of work package 2 of CORE-CLIMAX. Here validation is understood as a comparison to a reference data which need to be available at the right spatio-temporal scales and as a statistical analysis of the resulting time series that delivers for instance information on the variability of the parameter considered.

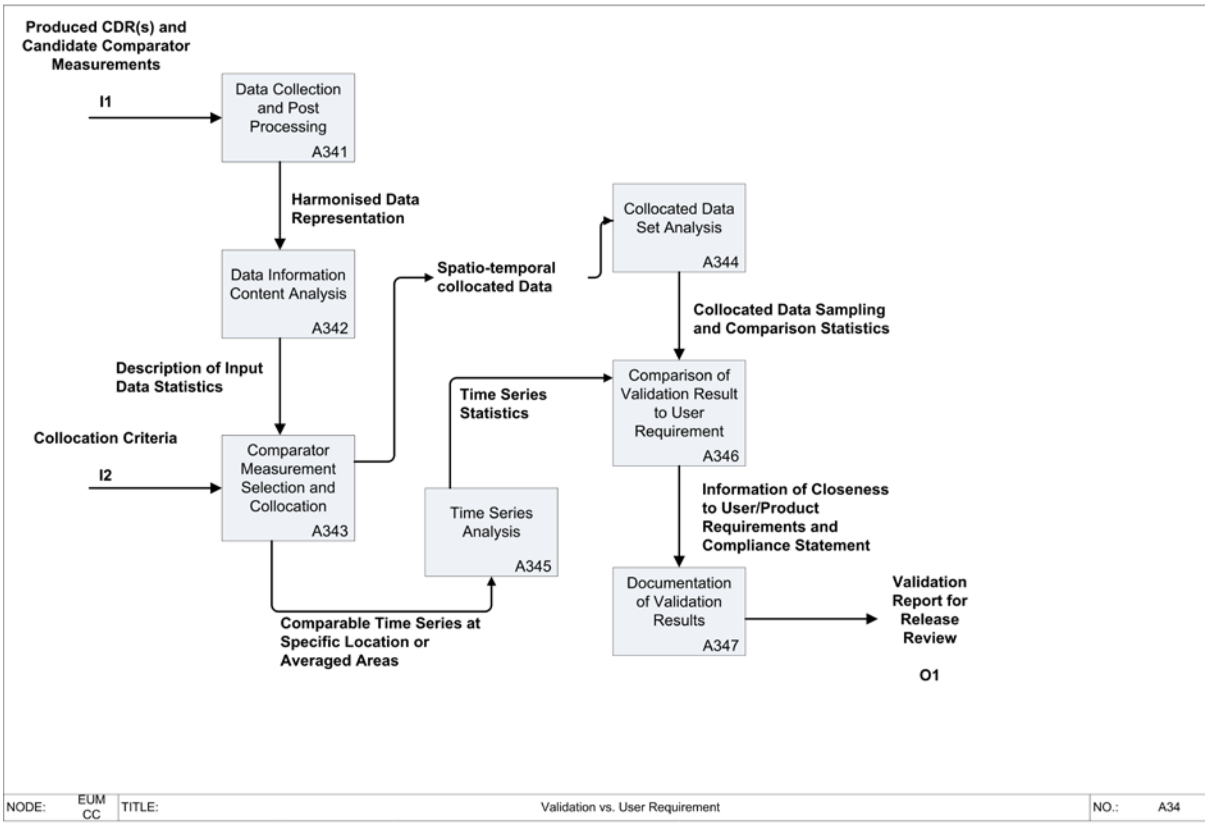

Figure 9: Decomposition of function A34 Validation versus User Requirements.

Almost always the validation process starts with the collection of the products to be validated and suitable comparator measurements. The post processing indicated in function A341 is mostly quality control of the various data sets which is of high importance to not include data that may mask comparison statistics later. However, great care needs to be taken in removing data, it is always better to just flag but retain all data. A second step of the post processing is often aiming at a harmonised data representation, i.e., comparable meta data and formats.

Within Function A342 a statistical analysis of the data is performed to better know what information the data really contain at what time and space scales and if they may systematically are leaving out certain situations, e.g., ship measurements have rather a fair weather bias. Function A343 then brings the data together following collocation criteria that can range from being very similar to the inter-satellite calibration to convert data onto monthly mean grids for comparison. In the end this depends on the defined metric for the comparison. Functions A344 and A345 then analyse all aspects of the collocated data producing statistics for the comparison data that can be used to address to what level the quality of the produced climate data record matches the user requirement (function A346). The last box (A347) summarises all findings in a specific document called validation report that is the major output of the validation activity. This report is of paramount importance to clearly communicate the validation results also to non experts.

4.2.5 Archive and Access to Data

Following a successful Data Release Review (see Section 4.3.3) function A36 (Figure 5) unfolds into what is shown in Figure 10. The product technical specifications including data volume and granularity are input to the process of product preparation for the archive. This process is preparing the data structures for the new product in a data archive. Often archives use specific data formats which may creates the need for reformatting the produced data records and also prescribes the availability of on the fly format converters to produce the format required by users which may differ from the archived format. Function A362 is then starting with a technical quality control of all data files that should be ingested into the archive that sometimes result in smaller repairs of the created data records, in particular, the meta data. This is then followed by the actual ingestion of the data which is monitored for being performed without error (function A363).

During or after the data ingestion sometimes also creates graphical representation of the data are created (function A364) that are used in data portals during data ordering. The status information from the ingestion process A363 is also used to perform the task of long-term data preservation (function A365). Depending on the requirements for data preservations this task secures the availability and contains activities such as regularly moving the data to new media.

Parallel to the ingestion of data into the archive preparation for the public access are undertaken. The new data record needs to become part of the data catalogue of the data provider and assuming the data record is accompanied with a digital object identifier a so called landing page for the access via the internet needs to be established (function A366). Function A367 then links the ingested data with the access mechanism which is most often tested offline. Function A368 opens the landing page and archived data for public access.

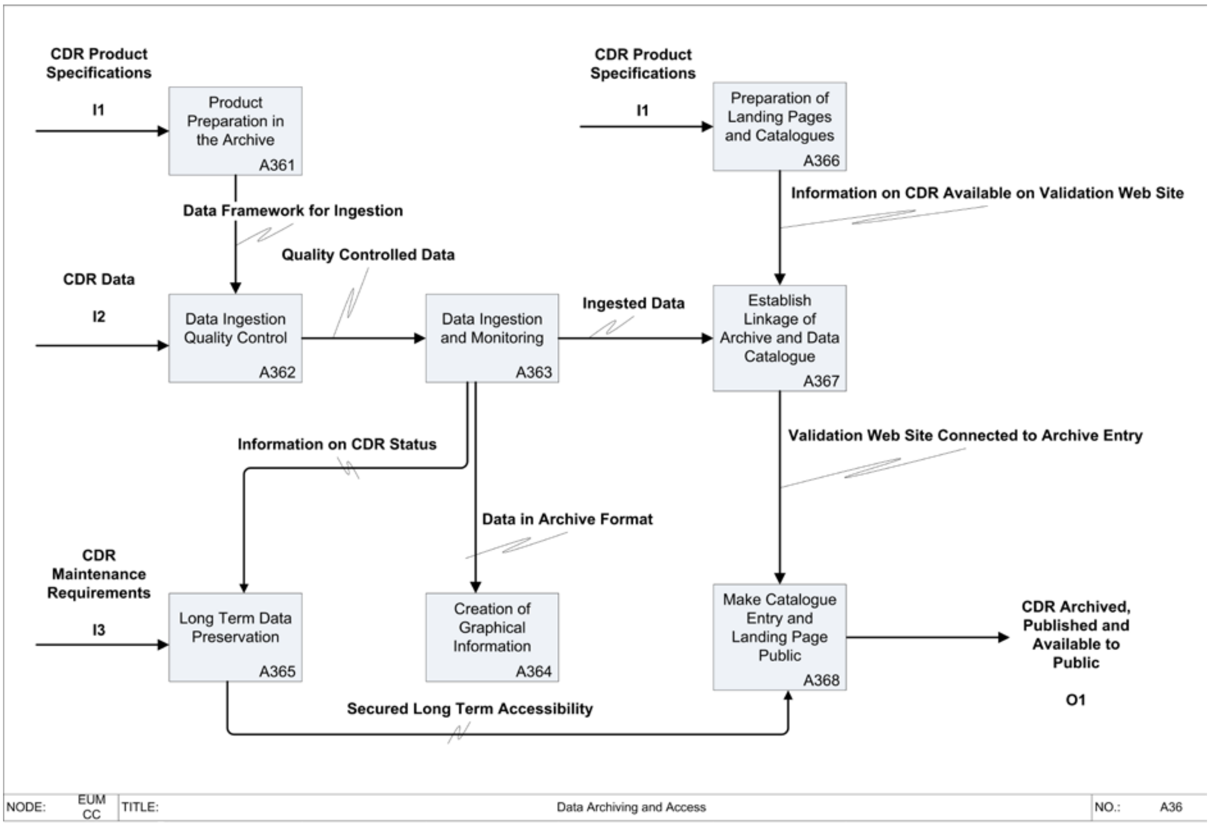

Figure 10: Decomposition of function A36 Data Archiving and Access.

4.2.6 User Feedback Mechanism

The ability to gather feedback for provided climate data records is essential for a data provider. Figure 11 displays three different and rather independent strands of gathering feedback which collect different kind of feedback and also have degree of organisation. These are:

- Statistics derived from the process of product ordering;
- A help desk function that collects comments from users on existing products;
- A product feedback mechanism organised by the data provider.

The first two are driven by the users and the data provider has little influence on what will come. The data provider needs to be able to gather, store, categorise and analyse the results of this type of feedback and can eventually use it to further improve services on data and product information provision. Eventually, some user requests and comments have a quality to inform product improvements.

The third mechanism is clearly driven by the data provider and needs preparation. Figure 11 shows a user feedback workshop as an example for such a process but there are others such as user surveys, structured interviews, or standing advisory groups. The choice of the mean depends on what information the data provider wants to gather. For instance a user survey could be effective in gathering needs on the technical side, e.g., on data formats, additional meta data contents, etc. It may also be successfully used to support enhancements of tools to work with the data either provided with support or implemented as web based application. However, if the data provider is seeking information on the development of a new or enhanced product in a specific application area it might be much more effective to interact directly with the corresponding community, e.g., in a workshop.

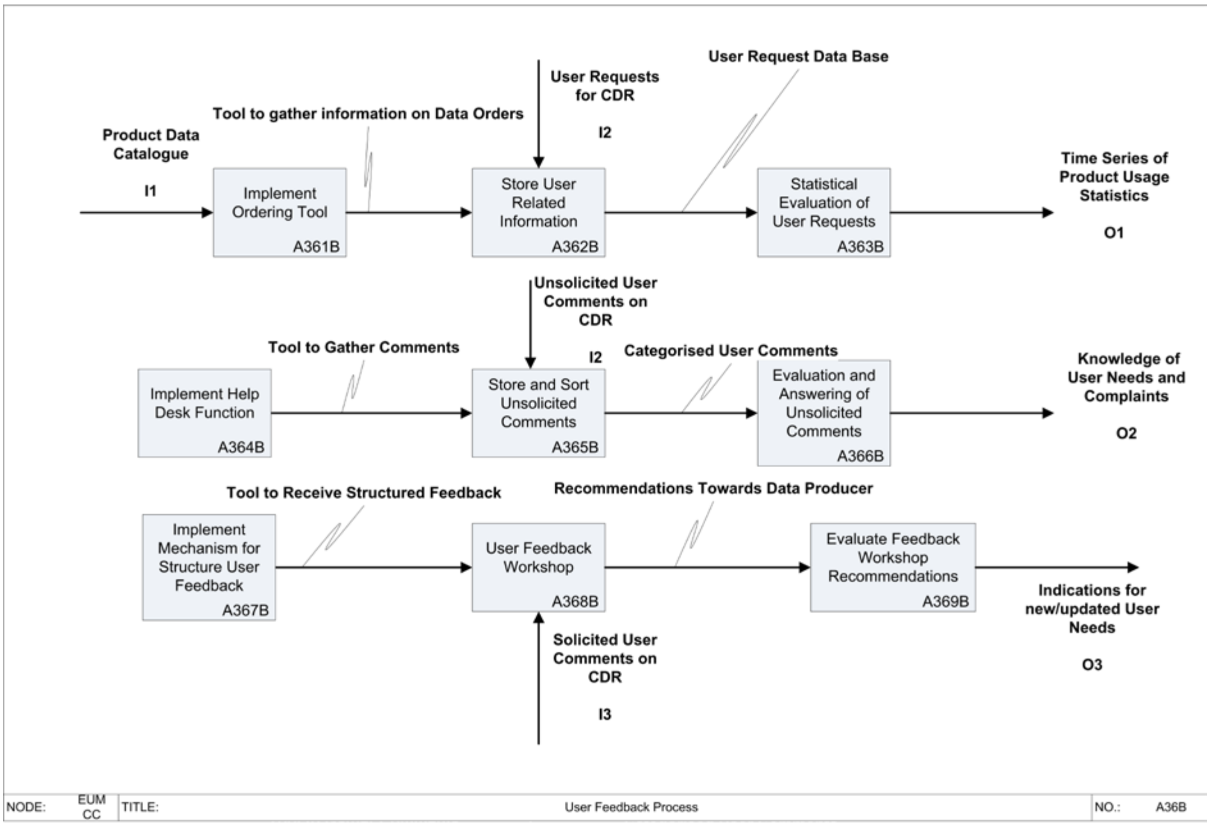

Figure 11: Decomposition of function A36 User Feedback Process.

4.3 The Life Cycle of CDR Development

The process to generate a climate data record has an underlying life cycle. Figure 5 depicts one important element, the Data Release Review that helps in the quality assurance of the released product. The following subsections provide a simplified overview of the life cycle and associated reviews that are independent from the source of data and should be applicable in satellite and in-situ data as well as reanalysis communities. It is important that such a life cycle is not burdened with an excessive formal process because this may hamper rapid progress. However, custom-made formalities help very much guiding developments of products, secure state of the art products and support the quality assurance for the finally delivered data records.

Figure 13 shows the decomposition of the life cycle for climate data records that start with the availability of a set of user requirements. The user requirements should describe what product needs to be available in what form and in what quality. In particular, the quality requirement very much depends on the application the product will be used for. For long-term climate data records it is at least necessary to state accuracy and stability requirements that are consistent with the required spatial and temporal coverage and sampling. As an example Figure 13 has the GCOS requirements as an input but principally such requirements can also be gained and continuously updated in the above mentioned user feedback mechanism.

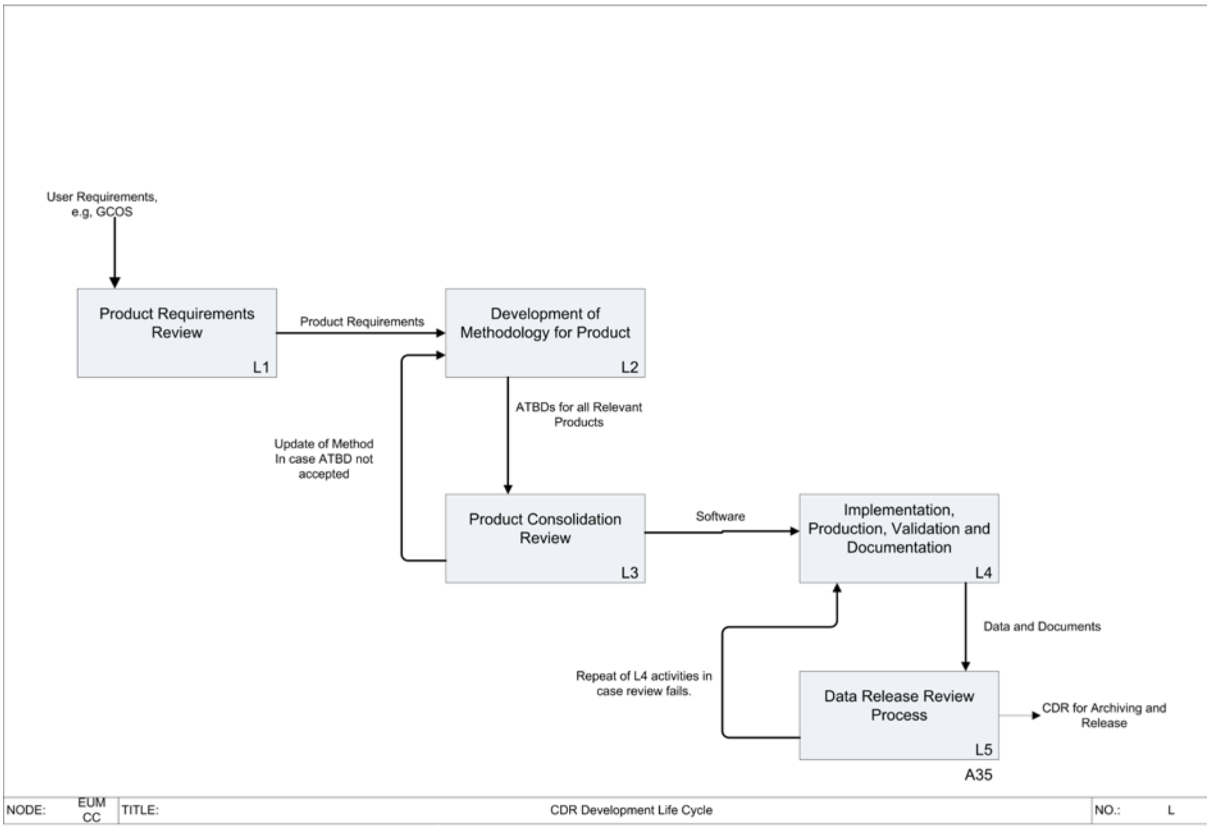

Figure 12: Decomposition of the Life Cycle for climate data record generation. This diagram is attached to the overall process at the level of the Data Release Review that has the co-registration for function A35.

At the side where the climate data record will be developed and/or produced the user requirements need to be analysed and sometimes also be transferred into so called product requirements that are an extended version of user requirements that may contain a much more specific technical description of the data record to be produced. Depending on the quality of the available user requirement product requirements sometimes also contain a de-scoping of the user requirements. For instance in a case where a user requirement was formulated with the vision of a future observing system such requirements appear not in reach for a climate data record based on historical measurements. In such a case the historic data record may only reach a quality to analyse inter-annual variability but not long-term climate trends which may only become possible with the future observing system. Nevertheless, the generation of the record from the historical data still makes sense to analyse climate variability at shorter time scales. From this short consideration it becomes already clear that a so called Requirements Review is a very useful instrument to review user and product requirements which supports the decision if a specific data record should be developed or not. The endorsed requirement documents after such a review provide a clear baseline for the development of a climate data record and its later validation.

The second step in the life cycle (function L2) is the development of the methodology to derive the climate data record. This can consist of all the steps described in earlier sections on calibration, inter-calibration, conversion to geophysical parameter space, etc. To ensure that the development has produced a method that is suitable to be used the so called Product

Consolidation Review (function L3) reviews all descriptions of the methodology including material that clearly demonstrates that the endorsed requirements are likely to be fulfilled. If successful the associated review board recommends the start of the technical implementation, production and scientific evaluation of the climate data record. As described above the scientific evaluation needs to address the closeness of the climate data record to the agreed user/product requirements. The final step in the life cycle is then the Data Release Review that reviews user related documents and also sample data to become able to endorse the publication of the produced climate data record.

4.3.1 Review Mechanics

The above mentioned three reviews (Requirements, Product Consolidation and Data Release) are conducted following the same simple system that is outlined in Figures 14, 15 and 16. The only differences between the reviews are the inputs and outputs of the review and the type of recommendation given by the review board to a decision body.

The review process starts with the preparation of the review inputs, i.e., the material needed to enable a review board to give a recommendation on a next step in the life cycle. An important element needed for all reviews is the so called Organisational Note that described the review to be conducted including its objectives, who participates with what role in the review and a schedule that includes submission of documents, correspondence and a review meeting that works out the final recommendations.

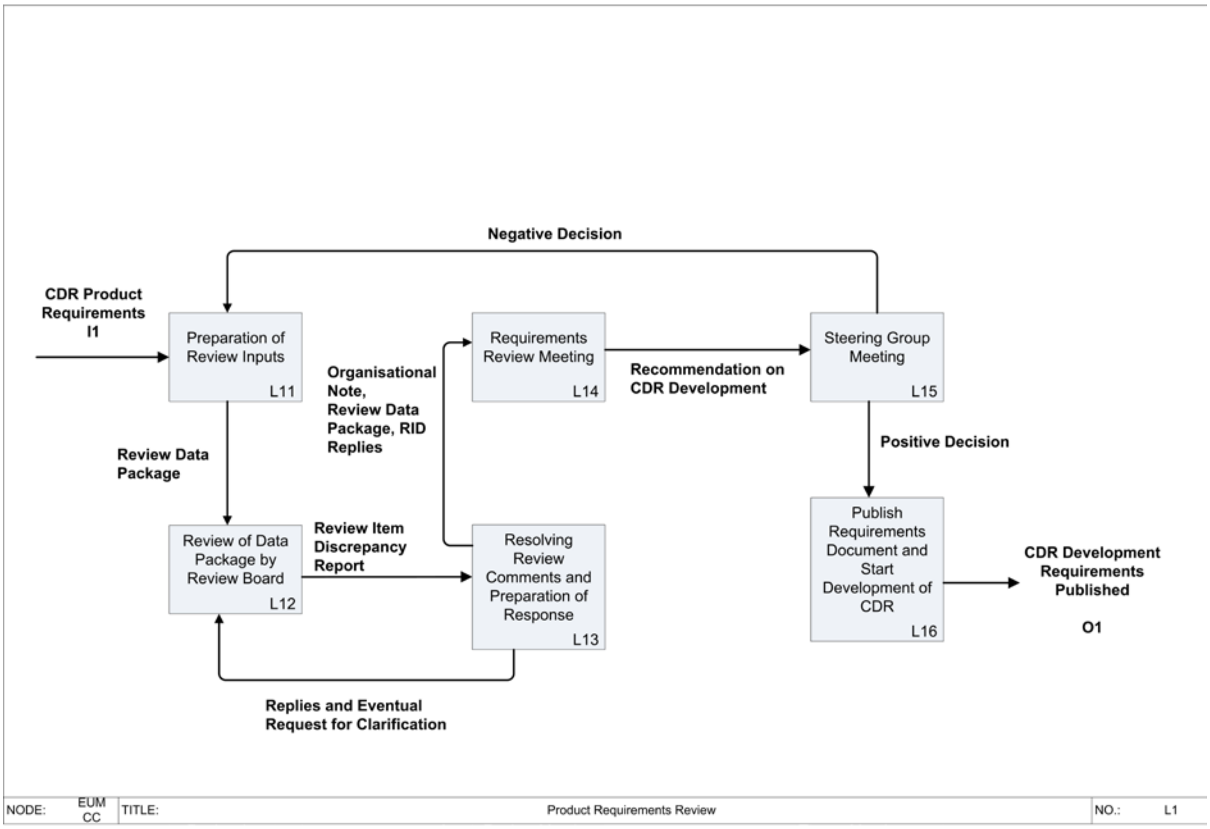

Figure 13: Decomposition of life cycle function L1 Product Requirements Review process.

In the case of the Requirements Review the applicable user and product requirements including all technical specifications and justified deviations between user and product requirements needs to the prepared for the review. The so called review data package is submitted to the review board that provides written comments and questions so called Review Item Discrepancy (RID) Report to the group that develops the climate data record. The RIDs are often categorised to enable all review participants to concentrate on the important issues discovered during the review. The categories also help to distinguish between issues that would block a further development and those that need to be resolved which may be done during the following steps of the life cycle. The developers have the opportunity to request clarification if RIDs were not completely understood and provide written answers to all those that were understood.

The Review Data Package and the RIDs including answers are then used in a specific meeting between the reviewers and the producers that goes through all RIDs and takes a decision how to move forward with it. For instance if the answer/explanation given to the RID is sufficient they may be closed by answer, in other cases specific changes in the documentation provided or updates in the specifications can be the results. If many things need to be changed a so called close out of the review is performed that simply checks if all actions agreed were successfully performed. In case of a successful review meeting the review board formulates recommendations towards a decision body on each of the review objectives, one of them certainly being to move to the next step of the life cycle. The decision body here labelled as a Steering Group takes its decision based on the recommendation which most often has the consequence that resources for the next step become available. In a case that a Steering does not accept the recommendation of the review board it may also order to start the process from the beginning.

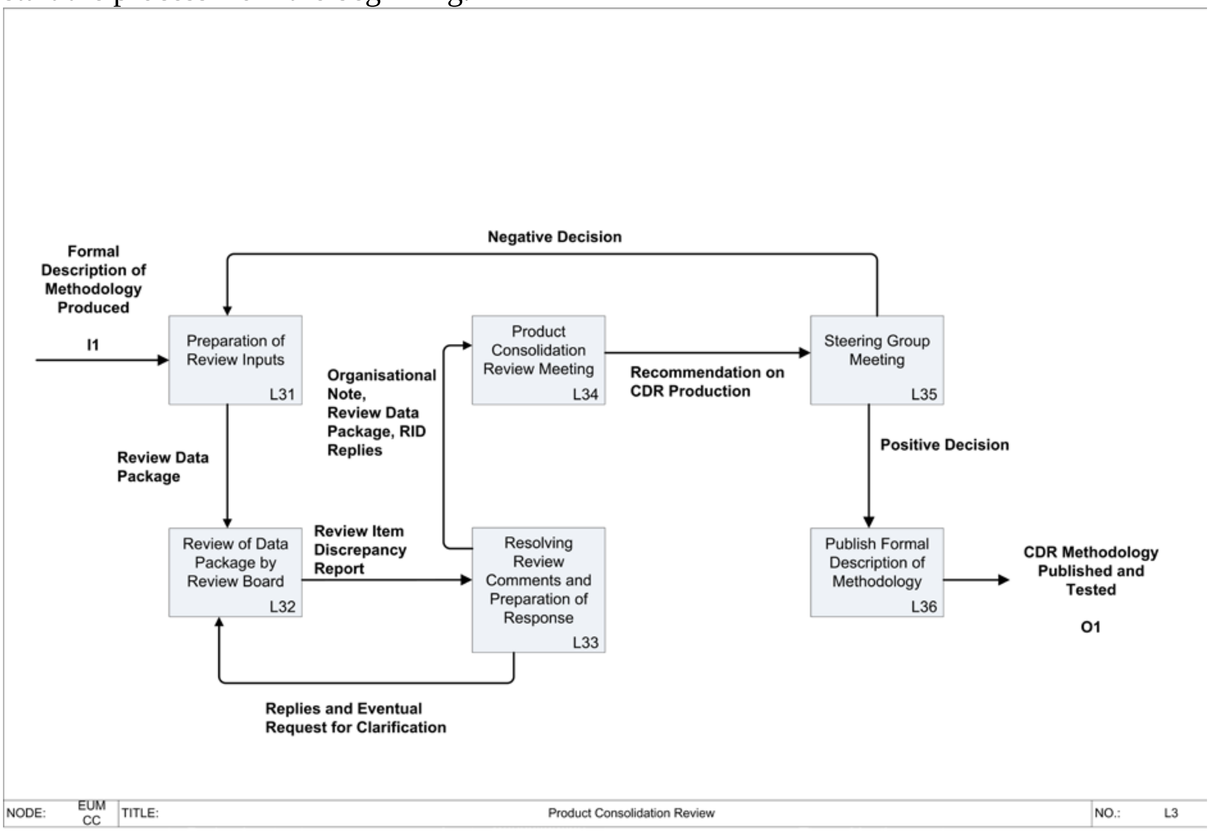

Figure 14: Decomposition of life cycle function L3 Product Consolidation Review process.

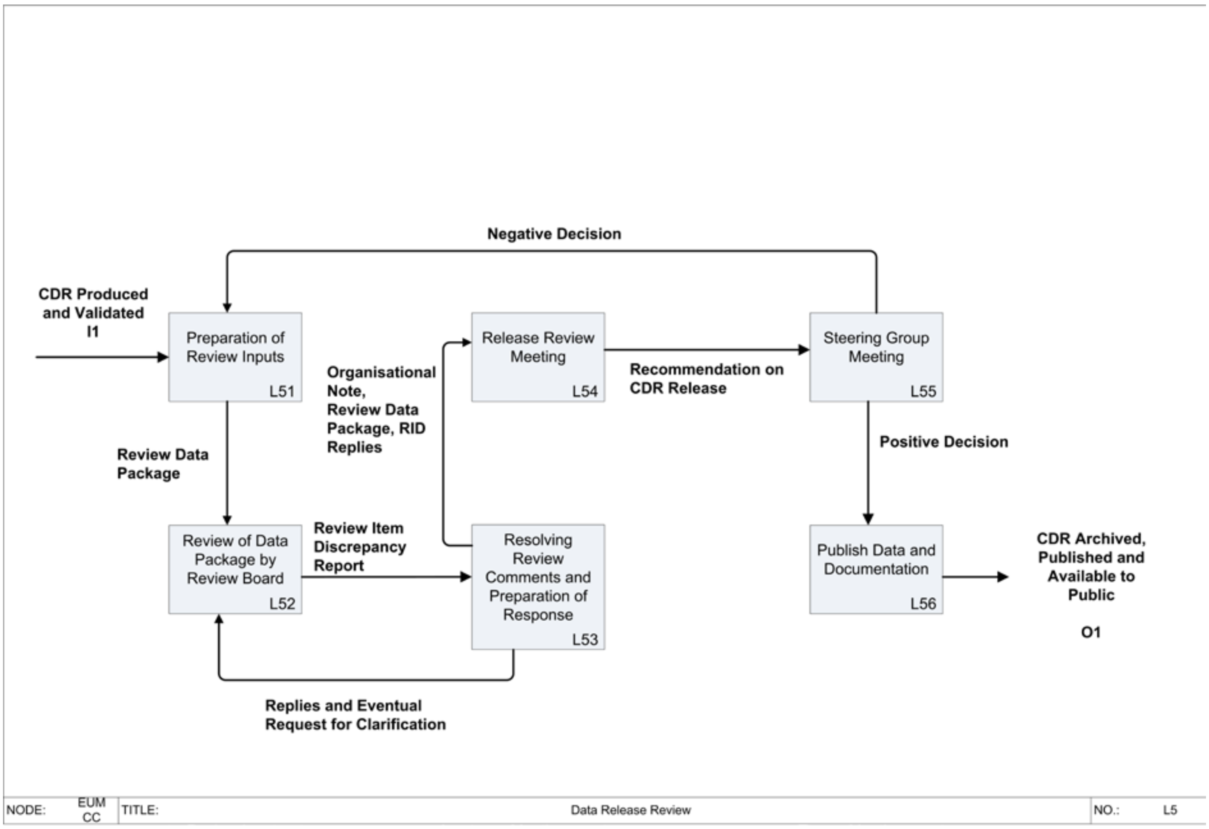

Figure 15: Decomposition of life cycle function L5 Data Release Review process.

4.3.2 Documentation

As the final part of the description of the structured process to generate a climate data record, Figure 16 shows a more detailed picture of the documentation task (function A37 in Figure 5) and how the various parts of the documentation are linked to the life cycle elements described above. The different reviews during the life cycle endorse the documentation prepared. Four major components of the user related documentation are a product requirements document, a formal description of the methodology, a scientific quality evaluation report and a product user guide.

The requirements document provides the baseline for the development of the climate data record and it important for the developer but also for the user because it enables always an assessment if the development is addressing the user requirements. The formal description of the methodology is essential to maintain a full traceability on how data are mathematically manipulated until the final product has been produced. It contains all trade-offs on calibration, a description how geophysical parameters are derived and how uncertainties are determined and propagated down the processing chain, and finally how data are eventually (re)-sampled, averaged, mapped etc. to arrive at the final product delivered to the user. The scientific quality evaluation report provides important information about the closeness of the climate data record to its own product and the user requirements that formed the baseline for its generation. From this document the educated user can immediately judge about the applicability of the data record for an intended application. The product user guide is an essential document for the user to learn how to obtain the data, to learn about the product content, how to work with the data and where the limits for the applications using the climate data record are.

There are many additional documents that are created during the structured process, mostly on engineering related to the generation and implementation of software as well as the review of the implementation that are not particularly considered here because they have limited direct relevance for the users of the climate data records. However, they are important to establish a generation of a climate data record that can be understood and repeated by different people which is establishing a real sustainable generation of climate data records in an operational environment.

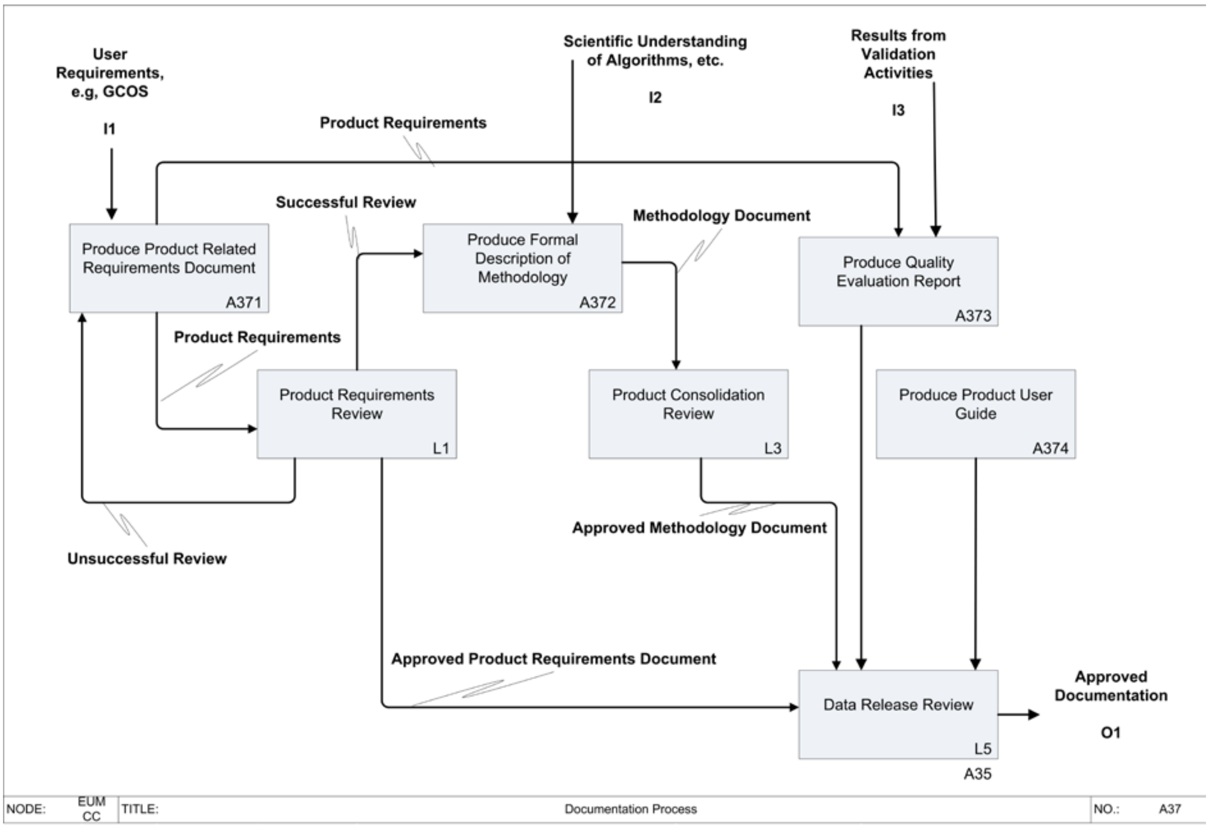

Figure 16: Decomposition of function A37 Documentation.

5 CONCLUSIONS

The document summarises the structured process for the generation of climate data records. The structured process is displayed employing the IDEFØ standard for functional modelling that is suited to manage the complexity of the process. The structured process has been developed on the basis of the architecture for climate monitoring from space. It addresses all relevant parts of the process in terms of activities needed and also connects the activities to the life cycle of the climate data record from the establishment of user requirements to the publication of the climate data record, hence its public availability to users.

The life cycle elements including the cascade of reviews is applicable to the use of data from any observing system, in situ and satellite as well as global and regional reanalysis. The life cycle provides a balanced approach to create a climate data record in a controlled fashion that enables production according to user requirements, ensures the traceability of the methods by providing full documentation and supports the quality assurance of the final products by a specific release review.

The practical steps outlined in Section 4 are mostly inspired by the treatment of satellite data but many parts in particular on quality evaluation, data archiving, public access and user feedback seem also applicable to in situ data. It should be noted that with every level of more detail it becomes more difficult to maintain a view applicable to all possible variants of the process. Thus, the sketched diagrams in this report in some cases need to be understood as baseline depiction of a process that may vary depending on the observing system used and maybe also on the climate system variable considered.

Finally the described structured process is in line with the described validation practises from work package 3 of the project and with the elements of the System Maturity Matrix that gives suggestions what the best practices for the different process elements is. The described process is certainly suitable to be considered for improving existing and stabling new instances that produce in climate data records.

APPENDIX A LIST OF ACRONYMS

CDR	Climate Data Record
CEOS	Committee on Earth Observation Satellite
CGMS	Coordination Group for Meteorological Satellites
DEM	Digital Elevation Model
ECV	Essential Climate Variable
EO	Earth Observation
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
GCOS	Global Climate Observing System
ICDR	Interim Climate Data Record
IDEFØ	Integration Definition Method
RID	Review Item Discrepancy
SBA	Social Benefit Area
SMM	System Maturity Matrix
TCDR	Thematic Climate Data Record
USA	United States of America
WGS84	World Geodetic System 1984
WMO	World Meteorological Organization