



Quality Management Framework for Climate Datasets

RESEARCH PAPER

CARLO LACAGNINA

FRANCISCO DOBLAS-REYES

GILLES LARNICOL

CARLO BUONTEMPO

ANDRÉ OBREGÓN

MONTSERRAT COSTA-SURÓS

DANIEL SAN-MARTÍN

PIERRE-ANTOINE BRETONNIÈRE

SURAJ D. POLADE

VANYA ROMANOVA

DAVIDE PUTERO

FEDERICO SERVA

ALBA LLABRÉS-BRUSTENGA

ANTONIO PÉREZ

DAVIDE CAVALIERE

OLIVIER MEMBRIVE

CHRISTIAN STEGER

NÚRIA PÉREZ-ZANÓN

PAOLO CRISTOFANELLI

FABIO MADONNA

MARCO ROSOLDI

AKU RIIHELÄ

MARKEL GARCÍA DÍEZ

][ubiquity press

**Author affiliations can be found in the back matter of this article*

CORRESPONDING AUTHOR:

Carlo Lacagnina

Barcelona Supercomputing
Center, Barcelona, Spain

carlo.lacagnina@bsc.es

ABSTRACT

Data from a variety of research programmes are increasingly used by policy makers, researchers, and private sectors to make data-driven decisions related to climate change and variability. Climate services are emerging as the link to narrow the gap between climate science and downstream users. The Global Framework for Climate Services (GFCS) of the World Meteorological Organization (WMO) offers an umbrella for the development of climate services and has identified the quality assessment, along with its use in user guidance, as a key aspect of the service provision. This offers an extra stimulus for discussing what type of quality information to focus on and how to present it to downstream users. Quality has become an important keyword for those working on data in both the private and public sectors and significant resources are now devoted to quality management of processes and products. Quality management guarantees reliability and usability of the product served, it is a key element to build trust between consumers and suppliers. Untrustworthy data could lead to a negative economic impact at best and a safety hazard at worst. In a progressive commitment to establish this relation of trust, as well as providing sufficient guidance for users, the Copernicus Climate Change Service (C3S) has made significant investments in the development of an Evaluation and Quality Control (EQC) function. This function offers a homogeneous user-driven service for the quality of the C3S Climate Data Store (CDS). Here we focus on the EQC component targeting the assessment of the

KEYWORDS:

data quality; trust; FAIR;
climate services; scalability

TO CITE THIS ARTICLE:

Lacagnina, C, Doblas-Reyes, F, Larnicol, G, Buontempo, C, Obregón, A, Costa-Surós, M, San-Martin, D, Bretonnière, P-A, Polade, SD, Romanova, V, Putero, D, Serva, F, Llabrés-Brustenga, A, Pérez, A, Cavaliere, D, Membrive, O, Steger, C, Pérez-Zanón, N, Cristofanelli, P, Madonna, F, Rosoldi, M, Riihelä, A, Díez, MG. 2022. Quality Management Framework for Climate Datasets. *Data Science Journal*, 21: 10, pp. 1–25. DOI: <https://doi.org/10.5334/dsj-2022-010>

CDS datasets, which include satellite and in-situ observations, reanalysis, climate projections, and seasonal forecasts. The EQC function is characterised by a two-tier review system designed to guarantee the quality of the dataset information. While the need of assessing the quality of climate data is well recognised, the methodologies, the metrics, the evaluation framework, and how to present all this information to the users have never been developed before in an operational service, encompassing all the main climate dataset categories. Building the underlying technical solutions poses unprecedented challenges and makes the C3S EQC approach unique. This paper describes the development and the implementation of the operational EQC function providing an overarching quality management service for the whole CDS data.

1. INTRODUCTION

Climate change and variability pose an unprecedented challenge to the overall society, which requires mitigation and adaptation responses to reduce the threats and maximise the opportunities presented to organisations of all kinds. The impacts of climate variability and change can take various forms such as physical, social, financial, or political, and as such climate change adaptation has a very broad scope. Both business and public administrations are vulnerable to potentially disruptive risks and are key actors in the creation of a climate-resilient future (ISO 14090:2019; ISO 14091:2021).

Both monitoring and modelling of the Earth system can provide the information and guidance necessary for the policy and decision makers to deal with climate-related challenges. This has led to the establishment of various initiatives designed to better understand the Earth system through an improvement in both observational capabilities and modeling tools. As a result, an increasing amount of environmental data about past, present, and future climate is becoming available. Unfortunately, these data often come with inconsistent or missing metadata, inhomogeneous documentation, and sometimes sparse evidence concerning their uncertainty and validation. A variety of data streams is generated independently and from multiple sources, adhering to different definitions and assumptions, often not standardised across communities, and, at times, with overlapping but disconnected objectives. As a consequence, the users can feel disoriented when it comes to identifying the most appropriate dataset for an intended application (Nightingale et al. 2019).

Given the ever more prominent role that climate products are assuming in decision making, it is unavoidable that the quality of these data will come under increasing scrutiny in the future. Climate services are emerging as the link to narrow down the gap between upstream climate science and downstream users. Climate services form the backbone of the process that translates climate knowledge and data into bespoke products for decision making in diverse sectors of society, ranging from public administrations to private business (Hewitt et al. 2020; Medri et al. 2012). The Global Framework for Climate Services (GFCS) of the World Meteorological Organization (WMO) stresses the increasing need for robust climate information based on observations and simulations covering future periods, ranging from several months up to centuries for economic, industrial, and political planning. Moreover, climate services play a crucial role in disseminating relevant standards (GFCS WMO) fostering adoption of common data models and formats and with sufficient metadata uniformly stored. The GFCS offers an umbrella for the development of climate services and has identified the quality assessment, along with its use in user guidance, as a key aspect of the service. The services, and the quality assessments in particular, need to be provided to users in a seamless manner and need to respond to user requirements.¹ The ultimate goal is building trust between data providers and users, as well as maximising usage uptake (Callahan et al. 2017; Rfl 2020). Thus, the questions of what type of quality information to provide and how to present it to users are receiving sustained attention.

1 <https://public.wmo.int/en/bulletin/what-do-we-mean-climate-services>.

A relatively young operational climate service is the Copernicus Climate Change Service (C3S, Thépaut et al. 2018), one of the six operational thematic services established by the European Commission within the Copernicus Earth Observation Programme (EC 2020). The C3S, implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF), aims to be an authoritative source of climate information for a wide variety of downstream users, ranging from policy makers to industrial sectors. The backbone of C3S is a cloud-based Climate Data Store (CDS), designed to be a single point of access to a catalogue of climate datasets of different categories, including in-situ and satellite observations, seasonal forecasts, reanalysis, and climate projections. In a progressive commitment to establish relations of trust between data providers and downstream users, as well as providing sufficient guidance for users to address their specific needs, C3S has made significant investments in the development of an Evaluation and Quality Control (EQC) function. By being transparent and characterising data quality attributes in a traceable and reproducible way, C3S is setting the basis for the inclusion of reliable climate data into policies and actions.

The establishment of an EQC function has three main advantages: i) it guides the users into the documentation to understand properly the dataset, it simplifies the comparison across datasets and builds trust in the products available, ii) it helps the data providers to understand which information they need to deliver to be compliant with standards, increases data uptake from the users, and clarifies how to provide standardised dataset quality information, and iii) it triggers actions for the service improvement leading to the provision of the most relevant dataset information and ensuring that published datasets are mature enough to support the authoritative character of C3S. The relation of trust between data providers and downstream users generated by these advantages is a necessary condition to foster a flourishing market for climate services (Zeng et al. 2019).

While the need for a quality assessment of the data available on the CDS is well recognised, the methodologies, the metrics, the assessment framework, and the way to present all this information have never been developed before in an operational service that disseminates all the main climate dataset categories. The proof-of-concept framework described in Nightingale et al. (2019) focused on observations only. Instead, here we implement the operational delivery of homogeneous EQC information across all the CDS climate dataset categories mentioned before. This objective, along with the task of building the underlying technical solutions, posed an unprecedented design challenge and makes the C3S EQC unique. The framework for the operational EQC function of the CDS builds on past and present international research initiatives. A variety of concepts were developed in previous projects (mainly the EU FP7 funded QA4ECV² and the C3S_51 Lots 2-4³), most of these concepts were integrated while moving the EQC function towards its operational phase.

This paper describes the challenges, the development and the implementation of the operational EQC function providing an overarching quality assurance service for the whole CDS. The results of the dataset assessments are publicly available in the CDS Catalogue.⁴ The manuscript focuses on the framework implemented for the CDS datasets only, leaving out other aspects of the EQC function, such as the assessment of the software made available by the CDS to explore the datasets (i.e. Toolbox), the assessment of the CDS infrastructure and the assessment of the user's satisfaction and requirements.

2. THE EQC FRAMEWORK IN A NUTSHELL

Commitment to quality is fundamental to build trust among stakeholders and the EQC function is the key element of C3S devoted to reach this goal. The main purpose of the EQC framework is to provide the C3S with a consistent, structured, and pragmatic approach to enhance the dataset reliability and usability. The design of this framework adopts an iterative approach integrating continual learning and improvement.

² <http://www.qa4ecv.eu/>.

³ <https://climate.copernicus.eu/c3s51-evaluation-and-quality-control-function-climate-data-store>.

⁴ <https://cds.climate.copernicus.eu>.

The EQC function regularly informs and recommends C3S about drawbacks, shortcomings and limitations related to the CDS datasets. These analyses are completed by a continuous user-engagement process to identify the user expectations that need to be addressed. The EQC team provides technical and scientific quality information of the CDS datasets via a set of homogeneous Quality Assurance Reports (QARs), helping to set the minimum requirements and baseline criteria for including new datasets in the CDS Catalogue. The QARs are filled templates called Quality Assurance Templates (QATs). Consistency across the QATs is obtained through the adoption of a vocabulary of homogeneous concepts and common practices.

The general strategy for assessing the CDS datasets consists of five steps:

- designing QATs for all the dataset categories with a consistent terminology and a structure as similar as possible;
- interacting with the data providers, who are the ones with the best knowledge of their datasets, and encouraging them to fill in the QATs or, in case the data provider is not available, fill in the templates using the documentation publicly available;
- evaluating the content collected in the QATs, paying attention to ensure that the content is understandable for the users, the level of detail is similar across datasets, and the type of information is complete, correct and consistent with the template requests;
- performing an independent quality assessment of the dataset, looking at aspects like (meta)data completion and integrity, scientific soundness and other characteristics that illustrate the multi-faceted nature of data quality; and
- publishing the information in the CDS dataset catalogue, once the QAR is approved by the corresponding authority, which in this case is the C3S governance board.

The production of the QARs calls for setting the procedures to initiate, develop and update the QARs (e.g., workflow), developing the software tools to support the assessments (e.g., data checker), and engaging with a wide range of stakeholders to choose the most adequate options for the QARs. These steps lead to the creation of QARs, which provide users with comprehensive information about the technical and scientific quality of the datasets. The different sections of the QARs are made accessible to the users in the CDS web portal through a synthesis table. The synthesis table is devised as a tool to organise and homogenise the EQC information, which is made of atomic elements corresponding to the different entries of this table. These entries contain links leading the user to the respective subsection of the QAR, where the user can find the EQC information of interest.

The overall EQC framework is guided by homogeneity and scalability approaches. The former leads to consistency of the EQC information across the CDS dataset categories, the latter leads to the integration of automatic tools to produce timely and sustainable data assessments in an operational environment. In particular, the EQC framework is driven by:

- a modular and flexible system able to consider new data/information sources and new actors involved;
- as much as possible automation of information acquisition (e.g., variable description, metadata checks) and its update in order to reduce human errors, speed up the QAR production, and make the system sustainable in the long-term;
- an iterative and reproducible approach permeable to the evolving requirements of both users and C3S to ensure continuous improvement;
- a user-friendly presentation of the quality information provided, clustered to facilitate its consultation and uptake to facilitate users in making their own decisions about climate data;
- consistent provision of the CDS dataset quality information, recognising the existence of inherent differences across the dataset categories;
- transparency and traceability of the quality assessments;
- FAIR (Wilkinson et al. 2016) and TRUST (Lin et al. 2020) principles and ISO 19157:2013;
- service management practices to make the EQC activities resilient in an operational environment.

Finally, the guidelines provided in Peng et al. (2021) have been followed when developing the EQC framework, as shown in [Table 1](#).

FAIR-DQI GUIDELINES (PENG ET AL. 2021)	EQC FRAMEWORK DESCRIBED IN THIS PAPER
Guideline 1: dataset	The dataset is described with a comprehensive online page providing various information that includes DOI, rich metadata, and licence.
Guideline 2: assessment model	The assessment method is available online together with the quality information. This paper itself details further the assessment model used. The assessment model is versioned and publicly retrievable.
Guideline 3: quality metadata	The assessments are captured into a structured schema/template (QAT). The quality information is standardised in a machine-readable (in our case using the CMS) and reusable form.
Guideline 4: assessment report	The quality information is structured in a template and is accessible online, versioned, and human-readable.
Guideline 5: reporting	The assessments are disseminated in an organized way via a web interface including the quality aspects assessed, the evaluation method, and how to understand and use the quality information.

Table 1 Mapping of the Peng et al. (2021) guidelines to the EQC framework characteristics described here.

3. DATASET QUALITY INFORMATION AND ITS DISSEMINATION

The QAT is the tool used to gather information on the most relevant aspects of the CDS datasets, informing the user in a quicker way rather than accessing and reading several documents (e.g., user guides, peer-reviewed papers, dataset descriptions). The QAT includes all the relevant quality information, in a concise and standardised form, with references and links leading to further details.

The general strategy is to provide seamless QATs, which are as homogenous as possible across all dataset categories. The QATs for each dataset category (i.e., satellite and in-situ observations, reanalysis, seasonal forecasts, climate projections) are available as supplementary material. The QATs are regularly reviewed to gradually converge towards harmonisation. Much improvement has been achieved by adopting a common terminology (see section 6) and common minimum requirement fields. The homogenisation of the QATs of different dataset categories ideally tends towards adopting one single QAT for all datasets. However, this goal is not feasible due to the diverse nature of the CDS dataset categories (concepts like ‘processing level’ or ‘quality flag’ are relevant for observational datasets, but not for other categories; along the same lines, the concept of ‘ensemble size of the hindcast’ is mostly relevant for seasonal forecasts). This homogenisation effort was pragmatically addressed by mapping the different QATs, one for each category, onto a general table agnostic of the dataset type.

In practice, all the QAT fields were grouped under main sections and subsections with common names for all dataset categories. An excerpt of the resulting QAT is reported in [Figure 1](#). Having common names for sections and subsections to all the QATs allows to organise and homogenise the EQC information in a general table named synthesis table ([Figure 2](#)).

Figure 1 QAT excerpt: the cells of the synthesis table (Figure 2) correspond to the subsections of the QAT (yellow text), while the column titles of the table correspond to the QAT sections (white text). The fields with an asterisk indicate the minimum requirements (see section 3.1). The QAT questions are in the grey area (left column). The middle column defines the data type, the rightmost column reports guidance about the type of content expected. Text in cyan appears as a tooltip when hovering the mouse over the web-form of the synthesis table.

INTRODUCTION A quick overview of the data characteristics (e.g., name, provider, time resolution)		
DATASET OVERVIEW Basic information about the data (e.g., format, name)		
Data format*	Drop-down box of: GRIB, NetCDF, ASCII, csv, txt, HDF, other (with area to write their ‘other’ option) (Allow multiple options to be selected, data can come in more than one format)	Format of the data as served by the CDS, no version
Physical quantity name*	Free text	It is consistent with the CDS Catalogue overview

Introduction	User documentation	Access	Independent assessment
Dataset overview	User guide	Toolbox compatibility	Data check
Temporal and spatial coverage and resolution	Scientific methodology	Archive	Expert evaluation
Providers	Uncertainty quantification		Dataset maturity
Dataset version	Validation		Key strengths and limitations
Data update	Inter-comparison		

Figure 2 Synthesis table, conceived as a tool to organise and homogenise the EQC information, as well as to guide users through the documentation. The table fields, each identifying an aspect of the dataset, are grouped into columns. Note the correspondence between the field ‘dataset overview’ of the column ‘Introduction’ and Figure 1.

The synthesis table entries contain links leading the user to the respective subsection of the QAR (i.e., filled QAT), where the user can find the EQC information of interest. Therefore, the structure of the synthesis table is agnostic of the dataset category, while the QAT fields, within each subsection (the information that is displayed when clicking on a cell in the synthesis table), depend on the dataset category.

The synthesis table offers an effective approach to guide the users into the documentation and homogenise the access to it. It addresses a typical user requirement: ‘most of the time, the problems with the documentation are not due to the lack of it, but to the difficulty in finding it’ (extracted from the C3S User Requirement Analysis Document 09/2019) or ‘all documentation should be easy-to-access’ (Nightingale et al. 2019). For instance, a non-expert user might not know the meaning of ATBD (Algorithm Theoretical Baseline Document) and the synthesis table overcomes this complication by guiding the user through questions (i.e., QAT fields), answered with high-level information further detailed in the complete document referred (ATBD in this case). Moreover, the synthesis table offers an extra level of assurance through independent assessments and guarantees the user that all the information made available through this table is traceable and quality controlled, because the information given by the provider is double checked by the EQC team and is versioned in the CMS. An extra advantage of the synthesis table is the possibility to track which EQC material the user is interested in by recording the user’s actions in the table. These actions can be analysed in a later stage to steer the future decisions of the EQC function and C3S in general.

The information accessible through the synthesis table may be grouped into two categories:

- Descriptive data information. Documentation has been selected to tackle data provenance, showing the origin, history, and methodology used to create the data. This information is available prominently in the column ‘user documentation’ of the synthesis table (Figure 2), in particular in the scientific methodology part. The documentation is completed by references to more detailed material, such as uncertainty characterisation, license, citation, and the like, for further user queries. Information here is the result of the documentation and accessibility assessments. In general, content is filled in by the data provider and reviewed by the EQC team.
- Independent assessments. An analysis of the dataset quality is performed independently of the provider, with the advantage of using the same metrics and tools nonpartisan of the source where the dataset was generated. This guarantees a uniform and impartial basic evaluation across datasets. Information here is the result of the technical and scientific assessments. See Appendix III for a definition of ‘technical’ and ‘scientific’ in this context.

The table is characterized by fields grouped into columns (Figure 2). The column with the header ‘introduction’ gives a quick overview of the data characteristics (e.g., name, provider, time resolution), as inspired by the WIGOS guide on metadata standards (WMO/WIGOS 2017). The column ‘user documentation’ provides the essential documentation for the effective use and understanding of the dataset (e.g., user guide). The column ‘access’ describes whether the dataset variable can be served by the CDS Toolbox and which are the archiving practices followed for this dataset. Finally, the column ‘independent assessment’, being more articulated, is explained in detail in Appendix I.

3.1 MINIMUM REQUIREMENTS FOR PUBLICATION OF CDS DATASETS

Some of the QAT entries are considered mandatory and some optional in the EQC framework. The content of the mandatory entries is considered so fundamental that, when missing, the dataset is not usable/understandable and thus unservable. These mandatory fields define the minimum requirements (MRs) for a dataset variable to be published (or withdrawn) by the CDS. The identification of these fields probably represents the first systematic effort towards the inclusion and development of an operational check of MRs, encompassing a wide range of dataset categories. Indeed, the identification of a suitable set of MRs was indicated among the ‘Science Gaps’ in assessing climate data quality by Nightingale et al. (2019).

The list of MRs is specifically thought to facilitate a timely publication of a dataset in the CDS, ensuring, at the same time, a sufficient (but not necessarily optimal) quality of the dataset. The MRs cover several aspects ranging from the dataset documentation to the compliance of metadata with community standards. The fields were identified as a result of the interaction between the EQC team, data providers, C3S, and users. The analysis of the MRs leads to recommendations to the C3S governance board on whether a dataset shall be made public on (or withdrawn from) the CDS.

To guarantee the maintainability of the MRs, they are an integral part of the QATs and are updated with the same frequency. See the supplementary material for a complete list of the MRs, indicated by an asterisk next to the QAT entry. Typical examples of MRs are ‘data format’, ‘physical quantity name’, ‘user guide documentation’ or ‘validation activity description’. Beyond the mandatory text necessary to fill in the QAT entries, a number of documents are also requested to be linked in the QATs as minimum requirements. These are:

- dataset user guide (e.g., seasonal forecasts SEAS5 user guide⁵). In the satellite observations community, this is usually referred to as PUG (Product User Guide);
- documentation describing processing of the dataset or a model/system technical documentation, including the description of the different components. In the QAT this document answers the question labelled ‘model/system technical documentation’ (e.g., reanalysis UERRA⁶). In the satellite observations community, this is usually referred to as ATBD (Algorithm Theoretical Baseline Document);
- product traceability chain. Only mandatory for reference datasets (e.g., in-situ observations GRUAN humidity⁷). Definition of reference in this context is given by GCOS⁸; and
- uncertainty characterisation and validations reports. In the QAT these documents answer questions about ‘validation or inter-comparison or uncertainty characterisation activities performed’ (e.g. climate projections CORDEX-CCLM⁹). In the satellite observations community, this is usually referred to as PQAD (Product Quality Assurance Document) and PQAR (Product Quality Assessment Report).

Before publication in the CDS, it is essential that the documents listed above are made available alongside the datasets they refer to.

The current version of the MRs fits the existing technology infrastructure as well as available human resources. Ideally, the list shall be extended to include basic technical checks about the data and metadata, such as time and space consistency and completeness, physical plausibility. However, it would require a technical infrastructure that was not available on the CDS at the time. In particular, it requires setting up automatic tools (a data checker software available for all the dataset categories) and tackling technical challenges (downloading and queuing time, memory disk space, enforcement of common metadata standards). Solving these technical limitations will help to extend the MRs list homogeneously across all the dataset categories.

5 https://www.ecmwf.int/sites/default/files/medialibrary/2017-10/System5_guide.pdf.

6 <https://confluence.ecmwf.int/display/UEER>.

7 http://datastore.copernicus-climate.eu/documents/in-situ/GRUAN_product_traceability_chain_GAIA_CLIM_Humidity.pdf.

8 <https://gcos.wmo.int/en/>.

9 <https://gmd.copernicus.org/articles/7/1297/2014/>.

Building the framework of the EQC requires designing protocols, software tools, QATs, and workflows for the QAR production as well as following common vocabularies and practices (e.g., TRUST Principles for Data Repositories, Lin et al. 2020). Among these, we focus now on the technical solutions underlying the EQC framework. Substantial technical developments have been undertaken during the onset of the operational phase of the EQC and more will be needed while it gets more mature over time:

- a Content Management System (CMS) and its maintenance, more details below;
- a Drupal-based module, inspired by the shiny-app R package,¹⁰ to show dynamic plots resulting from the scientific assessments;
- the integration of the EQC tab into the CDS infrastructure and its synchronisation with the other catalogue elements (e.g., download tab);
- software packages adapted to perform the scientific assessment tailored to the CDS characteristics (e.g., the ESMValTool¹¹ was adapted for climate projections analyses);
- a data-checker software to scrutinise CDS data and metadata;
- compatibility tests to check whether the data variables can be served through the CDS Toolbox; and
- setting up the software infrastructure, such as a network of virtual machines with the right environment and a Git manager for software repository, data flow architecture, and so on.

4.1 CONTENT MANAGEMENT SYSTEM (CMS)

At the heart of the EQC assessments is the Content Management System (CMS), an application used to manage content stored in a database and displayed in a presentation layer based on a set of templates, i.e. the QATs. Its objective is to ease the collaborative definition of the QAT structure and to facilitate and manage the creation of the QAT content. Creation of the QAT content is partially automated, as detailed in section 5.2.

The CMS facilitates the QAR production following a workflow that involves several roles, described in [Table 2](#), that access the CMS sequentially.

ROLE	RESPONSIBILITY
EQC main contact	One EQC team member who acts as the main contact of a specific dataset category. As QAR production is a multi-actor process, it is important that there is a central person, the EQC main contact, to coordinate the QAR production. This member contacts the data providers to agree on when they are available to fill in the QAT. Once the link with the providers is established, the EQC main contact defines the QAR name, fills in the QAT entries that identify the QAR uniquely and selects the team involved in the QAR production. Eventually, the EQC team member lets the actors involved know where there is a potential issue before it impacts the production.
Data provider	Typically, a member of the team that provided the CDS with the dataset under evaluation. The providers fill in the information requested in the QAT, because they are considered the best source to fully describe their datasets and so are the preferential choice for this task.
Evaluator	An EQC member who vets the QAR content and fills in the independent assessment fields. This role interacts with the provider for guidance about the amount and type of content expected and for any clarification needed.
Reviewer	An EQC member who scrutinises the whole QAR content for completeness and understandability. The reviewer is fundamental, because of her/his work in checking and verifying the correctness and consistency of all information introduced, while interacting with the evaluator to address any issue encountered.
Approver	Role covered by one C3S governance board member, who makes decisions about the publication of the dataset, (also) based on the QAR, and conducts a final check of the QAR before making it public together with the dataset. If the QAR requires further review, it will be sent back to the EQC team, commenting about what is still needed. Otherwise, it will be published in the CDS.

Table 2 Roles involved in the QAR production workflow.

¹⁰ <https://shiny.rstudio.com/>.

¹¹ <https://www.esmvaltool.org/>.

Figure 3 sketches the interaction between the roles involved in the generation of the QARs. The implementation of this workflow into the QAR production needs more consideration; it needs to distinguish between fast and in-depth assessment cycles as well as between common and non-common QAT fields. Details are provided in the next section.

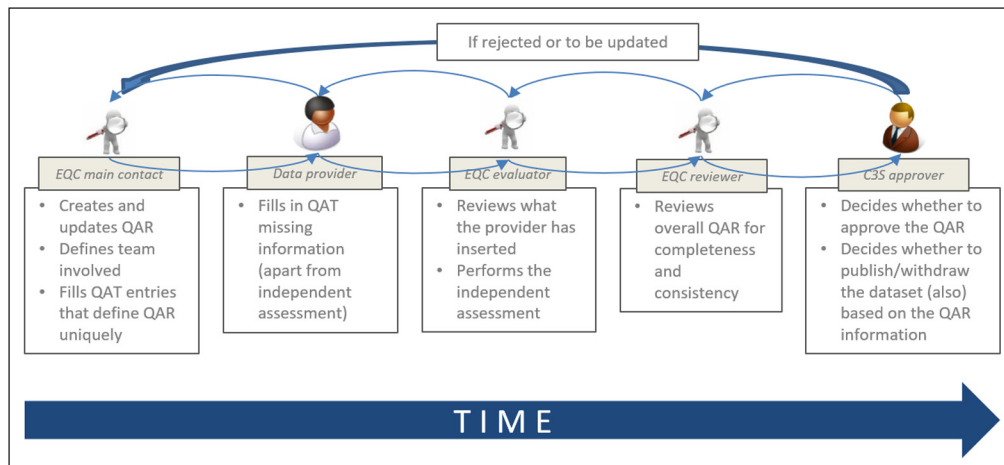


Figure 3 Sketch showing the basic roles, their interactions, and their responsibilities during the QAR production within the CMS. Note the iteration loop between roles to allow refinement of the content. The sketch gives a grasp of the more complex workflow shown in the next figures.

To complete the list of roles involved in the CMS, two additional roles are considered but they are not directly part of the QAR production workflow:

- QAR support role: can edit any part of published QARs to fix simple issues like typos or broken links or update technical data checks performed on new parts of a dataset regularly extended over time. Depending on the task, this role is covered by either an EQC member or a CMS automatic functionality.
- Observer: can read and comment in the CMS but is not granted the right to insert any content. This role, typically covered by a C3S technical officer, can intervene in case of blocking issues with the provider or whenever some aspects in the QARs need improvement.

5. WORKFLOW AND PROCEDURES FOR THE QAR PRODUCTION

The trade-off between timely and detailed assessment is tackled by splitting the QAR production workflow into two phases:

- A first phase (the fast assessment), mostly focused on verification of the minimum requirement fields. The dataset stewardship is scrutinised in terms of documentation, accessibility, and compliance with metadata standards.
- If the C3S governance board decides to make the fast assessment part of the QAR public together with the associated dataset, a second phase (the in-depth assessment) starts. During this phase, the complete independent assessment is performed, and the other fields are updated in case of need. The in-depth assessment focuses on the technical, scientific, and maturity data evaluation.

An additional element that makes the QAR production sustainable is the identification of QAT fields associated with common content across several QARs. More details are reported in Appendix II. In the following section, it is shown how these two elements, fast/in-depth assessment and common/non-common fields, come into play during the QAR production.

5.1 WORKFLOW IN A NUTSHELL

In a nutshell, the process for the QAR production for both already published and submitted for publication datasets may be summarised as follows:

- **Start:**
 - triggered by a QAR release calendar or new datasets available in the CDS
- **Input:**
 - data provider and EQC team fill in the QAT according to the workflow in the CMS

- **Output:**
 - QARs released to support the C3S governance board while deciding whether to publish or reject a CDS dataset
- **End:**
 - C3S governance board approves/rejects the QAR in the CMS
- **Manage updates:**
 - caused by issues identified by users, indications by providers, new data available from operational datasets extending over time, regular review of new dataset documentation, additional independent assessment analyses

The trigger for a new assessment is associated with two main causes:

- A completely new dataset appears or a new version of an already existing dataset is made available (e.g. ERA5 reanalysis replaced by ERA6). As a consequence, new QARs are produced. It corresponds to 'start' in the steps described above.
- Correction of QAR content (e.g., broken links), new dataset documentation is made available (e.g., validation reports), a dataset is extended backward or forward in time (e.g. a new month of seasonal forecasts), additional independent analyses are performed (e.g., inter-comparison with other datasets). In this case, the dataset version remains the same, but some information in the EQC scope is updated and the QAR needs maintenance. It corresponds to 'manage updates' in the steps described above.

Throughout the QAR production process, the *user engagement* team of the EQC iterates with the users to harvest their feedback about the different steps or improvements taken in a co-production process, making sure to advance in the direction of fulfilling the user's needs. These user's requests will result in reports to be discussed at regular EQC meetings, where they shall be further investigated and eventually trigger framework and QAR updates. User requirements also help to refine the QATs and to prioritise the performance metrics to be employed during the independent assessment. User engagement outcomes are thus the basis to conduct a gap analysis of the information made available to users and to steer the EQC design evolution in terms of framework and dissemination activities. This virtuous feedback loop is crucial for a user-oriented service as C3S.

5.2 DATASETS ALREADY PUBLISHED

The EQC function has been implemented after many datasets were already published in the CDS. As a consequence, a workflow needed to be envisaged to produce the necessary QARs. In this case, the trigger of the QAR initiation is a QAR release calendar, defined by the EQC team together with the C3S governance board. Once the QARs are triggered, the workflow is managed in the CMS, as shown in [Figure 4](#).

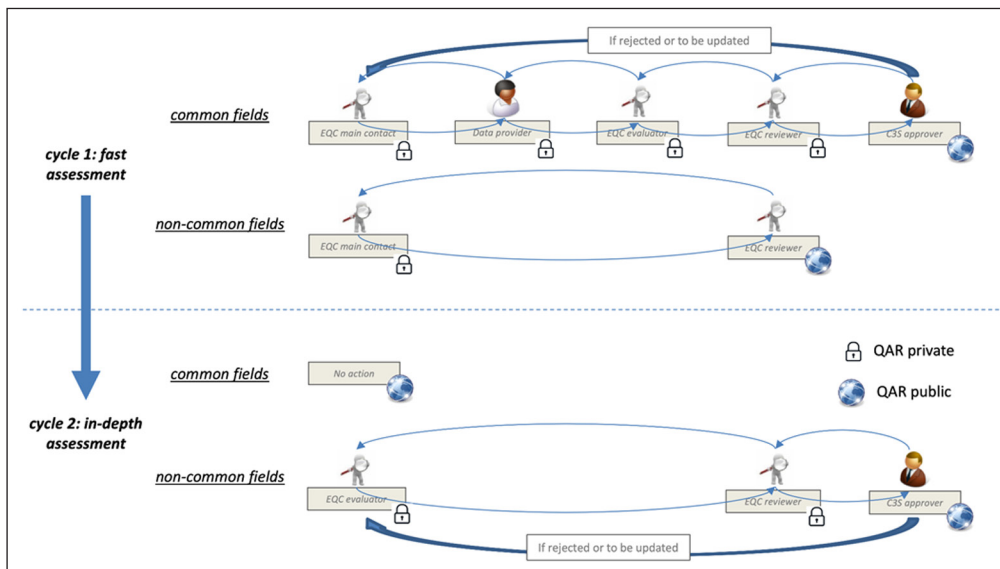


Figure 4 Sketch showing the interaction across roles during the QAR production within the CMS. Compared to Figure 3, here the distinctions between fast/in-depth assessment cycles and common/non-common fields are explicit and it is clarified at which stage the QAR is published.

Given the same roles identified in section 4, the QAT is filled in the private domain during the fast assessment cycle and then published. Once public, the QAR is completed with the independent assessment during the in-depth cycle and finally updated in the public domain. Each assessment cycle distinguishes between common and non-common fields:

- The part associated with the common fields requires the data provider expertise and does not include the independent assessment, as such this part involves many members, but it does not need to go through the in-depth cycle.
- The part associated with the non-common fields is unique to each QAR, because it is tailored to each variable. During the fast assessment, almost¹² any non-common field (e.g., variable description) can be extracted from precompiled tables validated with the data provider and extracted automatically to fill in the unique QAR. This part of the QAR production is nearly automatic, so it can involve fewer members, the EQC main contact to initiate the QAR and a reviewer to guarantee that the content is meeting expectations. Once the fast assessment cycle is complete, the non-common fields follow the in-depth cycle, where the evaluator includes the independent assessment material.

Once published, one QAR might need to be updated. More details about the procedure in this case are given in Appendix II.

5.3 DATASETS READY TO BE PUBLISHED

In the future, the evolution of EQC will need to consider a workflow for datasets ready to be published. So far, this workflow has not been implemented. Several options are considered based on the lessons learned during the ramp-up phase of EQC. Here we give some recommendations.

A new dataset is a dataset the provider considers ready to be served through the CDS. At this stage, the provider and C3S officers iterate to ingest data information, like documentation or the location where data are stored. Much of this information could be collected in the CMS (or a tool connected with the EQC CMS), which facilitates the completion of part of the QARs shortly after. Instead of EQC asking for similar information again, we can leverage the content already stored in the CMS to streamline the flow of information exchanged among the various authors involved, by introducing a workflow starting with the fast assessment explained in the previous section. Once the fast assessment cycle is complete, the dataset could be either rejected because, for instance, the minimum documentation required is not complete, or accepted for publication. When the dataset and the initial QARs are public, the in-depth assessment cycle starts. The logical flow, illustrated in [Figure 5](#), may be summarised as follows:

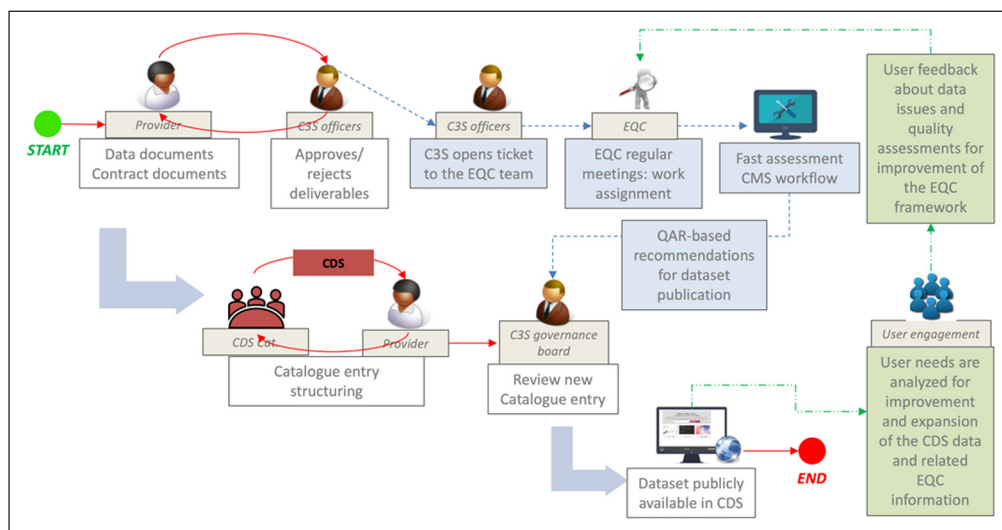


Figure 5 Scheme of the preliminary workflow about the way EQC may be engaged in the fast assessment of datasets not yet published in the CDS. Once a dataset is published (identified with ‘end’ in the figure), the in-depth assessment cycle starts as usual. The part associated with the EQC is in light blue, while the part associated with the EQC user engagement team is in green.

¹² The results of the metadata checks, i.e., data format, Toolbox compliance and standard convention compliance cannot be structured in precompiled tables, but the checks are produced by automatic software validated in advance.

- When the dataset is ready to be evaluated, the C3S governance board opens a ticket addressed to the EQC team.
- The EQC team meets regularly to assign the work for the QAR production based on, among other sources, the tickets received.
- Once the related QARs are completed of the fast assessment and approved in the CMS, the CMS closes automatically the ticket considered.
- The C3S governance board decides whether to publish, postpone the publication or discard the dataset using, among other input, the QARs made available.
- Once the dataset is published along with the preliminary QARs, the in-depth assessment to complete the QARs can start, as described in the previous section.

The independent assessment is part of the in-depth cycle that always starts after the dataset is published with its QAR. However, for new datasets it would be convenient that the data provider performs the technical assessment, that is, data checks, and reports the evidence logs to the EQC team. The EQC team then controls that evidence is available for the entire dataset and performs random checks autonomously on a subset of the entire dataset. The reason for this logical flow is that there are technical limitations that make the data checks timely only when done by the provider. Indeed, the downloading and queuing time and the disk storage requirements are technical limitations that would demand more resources for EQC, while these resources are likely already allocated on the provider's side. The strategy described would also reduce duplication of efforts and optimise resources, while guaranteeing independent checks.

6. PROTOCOLS AND PRACTICES COMPLEMENTING THE IMPLEMENTATION OF THE EQC FUNCTION

Besides the QAR production, the EQC function for the CDS is completed by additional protocols that make it a solid building block of C3S. Here follows a brief list of the protocols and practices considered.

6.1 PROTOCOLS FOR GAP ANALYSIS

Communication channels to inform C3S with recommendations to avoid gaps, address drawbacks and shortcomings and identify limitations have been established. These issues are reported via the EQC communication channels in the form of tickets sent to the rest of C3S. The tickets are sorted by resolution timing and priority as follows:

- *Short-term issues (<1 month)* requiring quick attention, critical/blocking issues. These answer questions along the lines of 'Is it something that hampers the EQC work?', 'Is it something limiting the user experience significantly?', 'Is it an obvious bug, an error on the website?', such as a Catalogue entry lands to page-not-found.
- *Mid-term issues (1 to 6 months)* identified problems and recommendations about the CDS data. These answer questions along the lines of 'Is it a problem that requires extensive analyses or impacts several aspects of the CDS?', such as unclear data licences.
- *Long-term issues (>6 months)* based on user requirements, non-blocking issues. These answer questions along the lines of 'Is it a user need that the EQC team is constantly facing when engaging with the users?', 'Is it a requirement coming from the EQC acting as a user?', such as the entry point to the Catalogue could be more efficient shifting from dataset category-based to variable-based. These requirements are inserted in a user requirement database and then analysed to become tickets. Usually, these tickets steer the evolution of the service over time.

The different issues are analysed by C3S and may trigger internal processes to deal with them. In this respect, the EQC team supports the evolution of C3S through gap analysis of the current capabilities of the CDS and formulates recommendations.

6.2 COMMON PRACTICES TO ENSURE CONSISTENCY ACROSS DATASET CATEGORIES

One key common practice to ensure consistency across dataset categories is to define a common vocabulary. The definition of shared vocabularies and common practices provides a foundation

for interoperability, reduces interpretation ambiguities, and boosts an efficient communication exchange. The efforts to harmonise existing terminologies in a structured vocabulary aim to facilitate the C3S products usage by the downstream and upstream users, and it is also beneficial for the coordination with the rest of the C3S activities ensuring consistency when referring to specific CDS elements. It shall be noted that the lack of an overarching consistent EQC vocabulary was identified as one of the priority gaps in climate data quality (Nightingale et al. 2019).

Agreeing on a common terminology is by no means a simple task, as it is time-consuming and comes with a variety of challenges, especially in the case of C3S where datasets come from different communities adopting different conventions. For instance, numerous terms are interpreted differently across data communities. What is defined as ‘product’ in the satellite observations community¹³ is way different to what the seasonal forecasts community¹⁴ refers to or to what the ECMWF MARS (Meteorological Archival and Retrieval System) archive considers it to mean.¹⁵ Some terms are very general (e.g., ‘observation’) and lead to long discussions to reach an agreement. For these cases, a practical solution has been to include mostly CDS-related terms, leaving out general terms as much as possible. The definitions are continuously monitored and improved.

According to the FAIR principles (Wilkinson et al. 2016), it is critical to use controlled vocabularies to describe and structure (meta)data in order to ensure findability and interoperability. A common vocabulary also refers to a set of common standards for data and metadata (formats and conventions) to be enforced by C3S. Indeed, gathering the metadata in a single system, as the CDS, with a common format, needs standardisation, as it is necessary to encourage data providers to convert their metadata inventories into formatted inventories that can be transferred to the C3S service. Including metadata in a consolidated and centralised system requires and/or encourages providers to agree to share the information with the community at large (Aguilar et al. 2003; Brunet et al. 2020). Having a common metadata standard for the many communities gathered by C3S is not realistic, because these communities have their own standards. Thus, a first practical approach is that each dataset category follows a community-recognised standard, as identified by the EQC team. Examples are CMIP, CORDEX, ESA-CCI, and obs4MIPs metadata conventions.

Finally, consistency of the EQC framework is also achieved by commitment to transparency following the TRUST Principles for Data Repositories (Lin et al. 2020). In particular, the methodology, the software and the assessments are made available to the users through the QARs public in the CDS. This helps to increase transparency and verifiability of the assessment, as well as resilience of the processes considered, which is open to further improvement.

A few more practices have been identified to ensure consistency across the QATs, among these are the following:

- Engagement with the data providers to guarantee that the QAT entries entail similar understanding. For most QAT entries, an exhaustive explanation is added to clarify the type of information requested. Explanations appear as tooltips in the CDS (see [Figure 1](#)).
- Engagement with the users. Feedback from a focus group (i.e. a users sample consulted on a regular basis to provide feedback on the new EQC releases) helps to reconsider terms that are not of immediate understanding for non-experts.
- Production of a quick-start guide to support the data providers in navigating and filling the QATs.
- Production of guidelines for the EQC team to reduce subjective or wrong interpretations of the QAT requests. Beyond fostering a common platform of understanding for the whole team, it gives resilience to the EQC function in case members leave and new ones join the team. All guides are continuously updated, leveraging the team experience.

13 <https://www.temis.nl/airpollution/absaai/>.

14 https://climate.copernicus.eu/charts/c3s_seasonal/.

15 <https://confluence.ecmwf.int/display/ECC/WMO%3D14+element+table>.

- Standardised style of the plots and references introduced in the QARs and harmonised QAR titles and filenames to deliver independent assessment results as much uniform as possible across the dataset categories assessed.

7. LESSONS LEARNED

The evolution of the EQC function would benefit from optimising the protocols, the templates and the workflow implemented so far, while tackling the gaps identified. The main issues encountered and more general considerations follow:

- **EQC workflow is a multi-actor process: collaborative iteration is key.** Several lessons have been learned while working on the QAR production and made clear that the EQC framework is a multi-actor process that requires collaborative iterations with different stakeholders. Tied collaboration between the C3S contracts and approver, user engagement, and data provider is extremely important for sustaining the delivery of EQC information. Clear responsibilities and timelines have been defined. Perhaps, the most important interaction is with the data providers/producers for filling the QATs. Responses from the providers are sometimes sparse or non-existent. Data providers/producers are considered the best source to fully describe their datasets and are the preferential choice to contribute to the QARs. This gap has been narrowed by:
 - o including the officer in charge of the relationship with the data provider contractor (a technical officer in the case of C3S), who facilitates the interaction with the provider and ensures that specialistic knowledge about the data under scrutiny is fully accounted for;
 - o ensuring that the EQC takes place earlier in the data ingestion process than now where the EQC work is done on already published data; and
 - o making all EQC-related tasks a contractual obligation in the data provider commitments. For brokered datasets (i.e., pre-existing dataset, not subject to the Copernicus license, to which C3S only acquires a license for the purpose of making it available in the CDS), the situation is not so different, because the contact point with EQC is the broker.
- **Technical constraints limit the extension of the minimum requirements.** The current version of the minimum requirements (MRs) fits the existing technological infrastructure as well as available human resources. The most important constraints to be faced are downloading and queueing time, memory disk space needs, lack of metadata information (that the provider should make available) about valid ranges for some variables, need for (land/sea) mask, and lack of automatic tools for data checks of each dataset category. To favor the development of data checker tools, common standards for data and metadata (formats and conventions) shall be enforced by C3S. Overall, climate services play a crucial role in disseminating relevant standards, in this case metadata standards (GFCS WMO¹⁶). A service aiming at providing seamless products necessarily needs to disseminate data in a common format with files structured similarly and with sufficient metadata uniformly stored. This supports interoperability of data files and archives for automated data processing through improved and extended standards and metadata. As a first step, datasets served in NetCDF should comply with the CF¹⁷ convention, but this covers only units, dimensions, and a few variables' metadata attributes. It is beneficial to ingest input data following a more constrained and controlled common vocabulary, covering variable naming, file names, grid descriptions, and global attributes. Examples of domain-specific conventions are CMIP¹⁸ and ESA-CCI.¹⁹ For the latter, C3S and ESA are coordinating to homogenise their metadata standards for satellite observation datasets. In addition, the ACDD²⁰ conventions cover the global attributes

16 <https://gfcs.wmo.int/>.

17 <http://cfconventions.org/cf-conventions/v1.6.0/cf-conventions.html>.

18 <https://pcmdi.llnl.gov/mips/cmip5/requirements.html>.

19 https://climate.esa.int/sites/default/files/CCIDataStandards_v2-1_CCI-PRGM-EOPS-TN-13-0009.pdf.

20 https://wiki.esipfed.org/Attribute_Convention_for_Data_Discovery_1-3.

for easier discovery and interoperability of the data (e.g., spatial and time coverage description, keywords).

- **Capacity building both in terms of human resources and technologies.** The implementation of the EQC requires cross-disciplinary knowledge (e.g., science, data management, computer engineering) to design protocols, software, and workflows following best practices. A constraint emerged during the work described in this paper is the importance of designing solutions sustainable for a large number of datasets. Otherwise, it is not possible to guarantee the necessary throughput with the available resources. As a consequence, some choices that seem to be straightforward and better than what has been implemented could not be considered because they do not scale for the large number of datasets under scrutiny (e.g., writing individual QARs for each variable). It was necessary to automatise as many parts of the workflow as possible to guarantee a timely production of the QARs. Based on our experience, it is also challenging to find a sufficient number of experts willing to regularly review in-depth all data streams. Considering that on demand requests for review do not necessarily guarantee that the same level of expertise could be kept over time, it would be appropriate to identify suitable funding mechanisms and contractual arrangements to keep the experts engaged for a longer period. The EQC framework must be tailored to the service infrastructure to be successful both in terms of human resources, coordination, and technology capabilities. This should translate into an increased effort towards capacity building, the need of which was also highlighted by Hewitt et al. (2020) for climate services in general. Capacity building shall be closely taken into account during the implementation of a sustainable EQC framework.
- **Optimise the production of more insightful independent assessments.** The scientific assessment needs to expand towards the diagnostics and standard metrics considered most insightful for the users. It would be beneficial to engage with the data providers to identify common baseline metrics to be applied independently. This will help to converge towards stronger provider engagement and satisfaction with the assessments performed and to reduce iterations during the in-depth cycle of the QAR production process. The huge and increasing amount of CDS datasets will need to consider pragmatic approaches to streamline the production of the independent assessment. While the fast cycle of the EQC framework has been made very efficient, more work will be needed to establish sustainable mechanisms for the detailed in-depth cycle of the EQC framework.
- **Consistent and shared vocabulary.** A common terminology should be shared and kept improving across the various C3S components. This facilitates consistency within the C3S (across the Catalogue entries for instance), it backs the user support desk when answering frequent questions by users about terminology, it gives a reference for the users to consult when jargon or acronyms are found, and it supports the project management to avoid ambiguities across C3S contracts. As such, the common vocabulary is considered a fundamental guidance document to be integrated in the C3S portal to benefit both users and service.
- **Benchmarking and cross-service coordination.** Consolidation of the EQC function also passes through the investigation of the most recent approaches. Given the challenges and opportunities that arose while implementing the EQC framework of the CDS C3S, exploration of the existing literature and coordination with other Earth data services, and in particular the other Copernicus services, is a key task to build a state-of-art EQC framework (i.e., benchmarking). It is important to investigate the standards and best practices implemented in similar services operating around the world (e.g., Leadbetter et al. 2020, RfII 2020). This helps to assess the applicability of the different approaches to the EQC function.
- **Scientific gaps warrant further research.** There are clear scientific gaps that hinder smooth development of protocols for data quality assessments. Further research investments shall be considered by major funding bodies (e.g., Horizon Europe) to fill the current scientific gaps. For instance,
 - o the system maturity matrix CORE-CLIMAX (EUMETSAT 2014) was identified as a tool for the maturity assessments, but it exhibits limitations of scalability in an operational environment. An example, the in-situ observations GRUAN dataset

required three months of work to complete the maturity assessment. In an environment like C3S, with increasing datasets and new versions, this effort is not practical. Scalability requires to detail the guidelines of the assessment (e.g., specify the metadata standards to check against or the source considered for citation to score usage). Moreover, the current scoring rules of the maturity matrix might require a peer-review process, possibly involving the data provider, to reduce subjectivity of the judgments. This poses once more challenges to produce scalable and timely assessments. In addition, the literature does not offer system maturity matrices for climate projections and seasonal forecasts, which opens intriguing questions about the different role maturity and verification play in modeling and observational communities.

- o Another case that would benefit from further research is the development of a metadata standard convention for seasonal forecasts and ERA5 data served in GRIB2 format.
- o Another example is the lack of well-defined ranges of physical plausibility for all the CDS variables, as well as a list of variable names and descriptions consistent across dataset categories. For instance, so far it is not clear what reference to use to name and describe the surface temperature, it ranges from '(surface) temperature' in GCOS²¹ to 'near-surface air temperature' in CMIP tables.²²
- **User engagement as an integral component of the EQC.** User engagement needs to be an integral part of the EQC evolution towards user-endorsed practices. It is inevitable that the users will drive the requirement for the provision of bespoke dataset quality information. The FP7 EUPORIAS project (Buontempo et al. 2018) defined a set of principles for a successful climate service, which are particularly relevant for the design of an EQC framework in a user-driven context. The next phase of EQC would benefit from identifying the type of QAR content considered most useful from the user's perspective. At present, our understanding of the usage of EQC assessments by users is still limited, but more attention to uncertainty characterisation, dataset stewardship, and ways the information is presented seem good candidates to start with. Thanks to the first QARs made available online, it will be insightful to test the efficacy of the communication strategies to ensure that appropriate and accurate quality assessments reach the users and are interpreted correctly.
- **Central repository of information.** The tool used for the QAR production (the Content Management System in our case) may need to expand to consider the several data ingestion processes happening in C3S beyond EQC and may need to upgrade enhancing functionalities in relation to data import and synchronisation with the rest of the CDS information. The tool may be better integrated into a single system for data ingestion and repository of information within the service to avoid duplication of effort and duplication of content describing the datasets. Better integration would also make the flow of information smoother across the actors involved in the service. Given the granularity of the datasets and their quality assessments, it may be convenient to make the quality information accessible through a structured API.
- **New challenges for the next phase.** The next phase of EQC will need to evolve to tackle some new challenges. For instance, it remains to define the details of a workflow dealing with new datasets ready but not yet public in the CDS. Another example is defining how to trigger QAR updates and their maintenance due to dataset new versions, due to datasets extending over time and due to new documentation available. Some practical directions have been suggested, but before adopting them operationally they would need to be assessed and tested. Evolution on how the quality information is disseminated (e.g., synthesis table) and development of an advisory service about dataset robustness (e.g., scoring scheme) will also deserve further exploration. A scoring system can be introduced so that users can quickly see which atomic elements of the EQC information have a good amount of detail. The scoring scheme would not aim at determining whether one dataset is better than another comparable dataset in an absolute sense, but only

21 <https://gcos.wmo.int/en/essential-climate-variables/surface-temperature>.

22 https://github.com/PCMDI/cmip6-cmor-tables/blob/master/Tables/CMIP6_3hr.json.

indicate the amount of quality information available. The scheme has to be simple and can be based on levels of increasing amount of detail/justification provided, as inspired by common practices in the literature (e.g., Nightingale et al. 2018, GEO Label²³). While working on the EQC framework, some ideas have been already put forward, the levels of appraisal may depend on the fulfillment of the minimum requirements described in this paper, making the scoring objective and prone to automation.

8. SUMMARY AND CONCLUSIONS

The current framework developed for the ramp-up operational phase of the Evaluation and Quality control (EQC) for the C3S CDS was presented. The framework considers the tasks, protocols and tools required to ensure that data are reliable and usable. It is inspired by the WMO GFCS guidelines, the ISOs 14090/1 and previous EU FP7/C3S projects. The framework is driven by a holistic approach aiming at homogenising the type of information made available across different climate datasets in a way that is both human and machine readable. It is characterised by a two-tier review system to assure the quality of dataset information released to the public. On the one hand, this approach enables fair and consistent comparison across datasets and facilitates guidance on the best use of data for the intended user's application; on the other hand, this approach makes the assessments sustainable and maintainable in an operational environment. In doing so, the framework explored optimal mechanisms (e.g., fast vs in-depth assessment and common vs non-common dataset information) for setting up sustaining delivery of EQC information, meeting the guidelines of the European Roadmap for Climate Services (EC 2015) and Peng et al. (2021).

The establishment of a quality management framework demonstrated benefits to the many actors involved:

- the users: easy access and guidance to quality assurance information;
- the data providers: feedback on data quality and incentive for improvement to increase data uptake and usability. This is in line with the good practice to connect the quality information with the dataset before release to reduce data misuse since this may also result in damage to the reputation of the data provider (Peng et al. 2021);
- the service itself: delivery of trusted and authoritative climate information, commitment to a user-driven evolution of the service. The service benefits from an established vehicle that triggers actions to improve the service itself; and
- the funding agencies: to get a measure of how compliant the funded datasets or services are with specific requirements.

As mentioned in the introduction, it is the first time that the methodologies, the metrics, the evaluation framework and the way to present all this information are being developed in an operational service that disseminates the majority of the climate dataset categories (including in-situ and satellite observations, seasonal forecasts, reanalysis, and climate projections). Building the underlying technical solutions makes the C3S EQC unique and requires pragmatic decisions for its implementation. The first part of the EQC framework design focused on ensuring the robustness of the baselines and processes to collect the information required for the QATs, keeping their coherence and their comparability across all the datasets available in the CDS. Having a set of QATs covering consistently all the dataset categories is a unique endeavour. These activities needed continuous improvement to fine-tune the EQC framework, benefitting from the operational assessments (QARs) and user engagement process. During the second part of the EQC framework design, activities focused on defining the level of content required in the QARs and its homogeneity across. Optimisation of the QAR update by means of automation tools and workflow streamlining has also played an increasing role.

The EQC framework was developed and implemented for all datasets published in the CDS at the beginning of 2020. QARs have been generated at the granularity of the variable for each dataset and made available to the users via the CDS web platform. The implementation of the EQC function addressed most of the recommendations that arose during the pre-operational phase (see Nightingale et al. 2019): (i) designing of dataset category specific

23 <https://geolabel.net>.

QATs, (ii) enhancement of the CMS functionalities, (iii) identification of minimum requirements to publish a CDS dataset, (iv) establishment of an overarching consistent EQC vocabulary, (v) creation of guidance documents for evaluators and reviewers to guarantee consistency in the QAR production, (vi) regular benchmarking activities brought into the operational process, (vii) ability to track changes in the QAR content.

Several constructive pieces of feedback from data providers, downstream users and C3S officers made the dissemination of the EQC information more robust over time. Orchestrating the different elements involved requires considerable coordination efforts and a continuous improvement approach to integrate the inputs regularly emerging from stakeholders and technical constraints. A number of lessons learned and science knowledge gaps were identified during the development of the EQC function and are detailed in the paper. These warrant further investment to comprehensively address the quality dimension of climate datasets in an operational environment.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **QATs.** Consistent set of QAT designs for a variety of dataset categories supported by the CDS: in-situ observations, satellite observations, seasonal forecasts, global and regional climate projections, global and regional reanalyses. DOI: <https://doi.org/https://doi.org/10.5334/dsj-2022-010.s1>

APPENDIX I: CHARACTERISTICS OF THE INDEPENDENT ASSESSMENT

The independent assessment is a fundamental piece of the quality assessments. Indeed, evidence that the dataset has been independently validated is key criteria for most data users (Nightingale et al. 2019). Applying the same approach and tools, nonpartisan of the supplier source, guarantees a uniform and impartial auditing. The independent assessment is part of the QAT and is designed to accommodate the information on the following topics:

- Data and metadata checks, performed to verify whether a reported data value is representative of what was intended to be measured or simulated and has not been contaminated by unrelated factors. Lawrence et al. (2011) postulated a generic checklist for technical quality assessments within a data review procedure. Building on the Lawrence et al.'s (2011) checklist, the EQC developed a data checker software to detect whether the CDS data respond to the data models defined for the specific dataset category (i.e. community metadata standards, e.g., ESA-CCI Data Standards V2.1²⁴), have the expected format and metadata with no unforeseen gaps, and no suspicious outliers (physical plausibility).
- Basic metrics (e.g., bias, correlation, linear trends), appropriate for each dataset category, to check the scientific soundness and performance of the CDS datasets. Results are available through the synthesis table cell named 'expert evaluation'. These standard diagnostics represent a first step for more insightful scientific assessments to be developed over time. For instance, climate projections related metrics could include performance analyses of future climate projections simulations or more reference datasets could be considered. Given the many different analyses options, priority shall be given according to the user needs. Based on our experience, it is anyway recommended to engage with the data providers to identify, together with EQC, common baseline metrics to be applied independently. This will help to converge towards stronger provider engagement and satisfaction with the assessments performed.
- System maturity matrix (SMM) assessment, which is performed for these six clusters: metadata, user documentation, uncertainty characterisation, access/feedback/update, archive, and usage. The SMM model used is based on the CORE-CLIMAX approach and is intended to extend these assessments to all dataset categories in the CDS. However, the

²⁴ https://climate.esa.int/sites/default/files/CCIDataStandards_v2-1_CCI-PRGM-EOPS-TN-13-0009.pdf.

literature does not offer a SMM for climate projections and seasonal forecasts datasets, which makes this extension exercise fitting better a research project rather than an operational EQC service. It was therefore not possible to expand the SMM to all CDS categories.

- Key strengths and limitations, reporting the concluding remarks about the independent assessment and offering guidance with not-to-miss information that helps the user to understand whether the dataset fits the specific user's application. In the future, we see this document improving its guidance facilitating the translation of complex, often technical, information into a format that can be easily understood by users focusing on the most relevant aspects (e.g., known issues, dataset highlights, uncertainty).

APPENDIX II: DETAILS ABOUT THE WORKFLOW

THE CONCEPT OF COMMON AND NON-COMMON QAT FIELDS

An additional element that makes the QAR production sustainable is the identification of QAT fields associated with content common across several QARs. It shall be taken into account that the QAR granularity is at a variable level. Thus, given for instance the same model, version, model run and Catalogue entry, all the variables share some information that is exactly the same. The concept is clarified by the following example: for a given Catalogue entry = 'CMIP5 monthly data on single levels', model = 'inmcm4', experiment = 'amip', some answers to the QAT for climate projections have the same content, for instance 'description of the model' or 'horizontal resolution'. In this respect, instead of repeating the same answers, these are written once in the common part of the QAR, while the specific answers, unique to the specific variable, enter the non-common part of the QAR. Continuing with the example above, let us suppose that we want to produce one QAR about

- variable '2m temperature';
- model 'inmcm4';
- experiment 'amip'; and
- available in the Catalogue entry distributing monthly fields, that is, 'CMIP5 monthly data on single levels'.

When creating the QAR, the EQC main contact fills these four fields that make the QAR unique and selects the common fields associated with model 'inmcm4', like data format or model components description. These fields are filled once in the CMS and then propagated automatically to all variables and experiments for the same model. One extra consideration that reduces the manual intervention is recognising that the non-common fields are typically the variable name, units, and description, and these are the same independent of the specific model given the same Catalogue entry. Thus, the associated information can be extracted from precompiled tables to fit each QAR. For instance, the definition of '2m temperature' is the same for model 'inmcm4' or for model 'access1-3', but the common fields (e.g., model components description) are very different. As a result, the CMS makes it possible to fill in each QAR by extracting the non-common information from precompiled tables. Benefitting from this functionality, any change to these tables can be quickly propagated to hundreds of QARs. These tables need to be validated by the data provider. Note that all this also makes the approval process of thousands of QARs sustainable in the fast assessment cycle. Indeed, while the common part of a QAR needs the usual CMS workflow, the remaining non-common part fast assessment would require only automatic extraction of content from tables agreed with the data provider. The merging of agreed common fields and agreed tables for non-common fields avoids the need for the approval stage of each single QAR in the fast assessment. The EQC team spent time identifying which QAT fields are common and which are non-common, results are shown in the supplementary material.

TRIGGERS FOR THE QAR UPDATE

Once published, one QAR might need to be updated till a new version of the same dataset is published in the CDS. Moreover, an update can also be caused by:

- users detecting and reporting via the helpdesk deficiencies or shortcomings with the dataset or with the QAR published. For simple actions, like typo correction, the QAR support role in the CMS can apply changes to the published QARs, whereas for more complex editing, the EQC team shall withdraw the QAR for updating. In this case, the EQC main contact restarts the QAR;
- regular updates that require manual intervention to take into account possible novelties in the documentation, for instance. The regular update depends on the QAT field and on the dataset category:
 - o the general rule is that it is usually done once a year manually and only for the QAT common fields;
 - o variable definitions might improve over time, for instance. In this case, C3S warns the EQC team about the improvement and the non-common fields are updated by changing manually the tables containing these fields. The manual update of the tables triggers an automatic update of the QARs affected. In the future, both the non-common fields tables and the Catalogue would benefit from automatic synchronisation preventing manual intervention; and
 - o seasonal forecasts QARs need a more frequent manual update because new versions of the systems are upgraded frequently, often yearly. The update is done by restarting the common fields once every three months, focusing on a few QAT fields preselected (e.g., operational status of the system);
- regular updates of the assessments that do not require human intervention, that is, data checker and Toolbox compatibility software. This is particularly valuable for datasets that are regularly extended in time, such as seasonal forecasts and reanalysis (a.k.a. near real-time datasets). The appropriate parts of the QARs are automatically updated once a month, considering the last month of data available; and
- additional independent assessment analyses. This can happen during the in-depth cycle for non-common fields only. If new analyses are available (e.g., inter-comparison assessments), the in-depth cycle restarts to update the appropriate QAR.

It shall be noted that in case a new version of a dataset is available, a completely new QAR has to be produced. In this respect, a new dataset version is not considered a trigger for QAR updates, but the workflow follows the usual QAR production.

APPENDIX III: TECHNICAL, SCIENTIFIC AND DOCUMENTATION ASSESSMENTS

Throughout the paper, the authors use the concepts of ‘scientific’ and ‘technical’ assessments. It might be worth avoiding any misunderstanding clarifying the notion in this context. The scientific assessment consists of data content and cross-data content checks, as opposed to the technical assessment that regards the data and metadata files checks (Stockhouse et al. 2012). As an example, when the evaluator plots the dataset variable and checks for reproducibility of El Niño events against skill metrics, here the scientific soundness of the data content is considered (e.g., Haiden et al. 2019). When the evaluator looks for the metadata standard compliance (e.g., ACDD convention), here the evaluator is checking that the attributes describing the files are according to a set of community-recognised metadata characteristics (e.g., specific date-time format). There is no check of the data content, no scientific evaluation in this case, but a metadata conformity check; it is a purely technical assessment (Evans et al. 2017; Stockhouse et al. 2012). The physical consistency checks are borderline in this distinction. On the one hand, these are technical checks when analyses regard the match between information in the valid ranges reported in the metadata and the max/min values in the file content. On the other hand, the check becomes more complex if a plot associated with analyses is necessary to identify the suspicious outliers. This operation becomes necessary when the metadata does not contain the valid ranges, which is a good practice to include but at times it is very challenging for exotic variables. In this paper, the physical plausibility checks are considered part of

the technical assessment, albeit automatic statistical analyses about the data content is necessary.

As far as the scientific assessment is concerned, this refers to scientific analyses of the physical content described by the dataset to check for its scientific soundness. Given the nature of this assessment, it is typically carried out by domain experts. Analyses may include uncertainty estimation, validation against reference datasets, and reproducibility of temporal/spatial patterns. Once clear what is considered as technical and what is considered as scientific, part of the assessments described in this paper are about documentation (availability and completeness) and file accessibility, archiving and compatibility within the service (i.e., with the Toolbox). All these analyses are neither purely technical nor scientific because they are not about the files per se, but about the associated material needed to access and understand the dataset. These analyses enter the group of stewardship assessment. It is true that stewardship regards all the aspects about distribution of the dataset and so potentially also technical and scientific assessments enter in this category. However, with stewardship assessments here we consider any assessment that guarantees accessibility and understandability of the dataset distributed, and so anything of relevance in dataset quality that is not associated with the data and metadata file content, such as documents accompanying the dataset describing how to use it. Typical examples regard the description of the algorithms or models used to produce and process the data, provision of the DOI and licence of use, grid description, verified network address to access the data, and information about the archiving procedures. The goal is to ensure that the dataset is well documented, the processing chain is visible, and the data is readily obtainable and usable.

At times, the assessments described above are accompanied by maturity assessment models. These are formal approaches to support compliance verification, usually defined in discrete stages to evaluate practices applied in organisations, services, or products. Maturity is meant as a desired or anticipated evolution from a more ad hoc approach to a more managed process (Peng 2018). Datasets associated with high maturity are produced following best practices of the community and in a more managed fashion, increasing user trust in the data record provided. It should be noted that a low maturity rating does not necessarily imply a low scientific value for a dataset. It can happen especially for datasets managed by a single investigator that may be flagged to have low maturity due to poor quality in metadata, documentation, and accessibility.

ABBREVIATIONS

ACDD – Attribute Convention for Data Discovery

ATBD – Algorithm Theoretical Baseline Document

C3S – Copernicus Climate Change Service

CAMS – Copernicus Atmosphere Monitoring Service

CDS – Climate Data Store

CF – Climate and Forecast

CMEMS – Copernicus Marine Environment Monitoring Service

CMIP – Coupled Model Intercomparison Project

CMS – Content Management System

CORDEX – COordinated Regional climate Downscaling EXperiment

CORE-CLIMAX – COordinating Earth observation data validation for RE-analysis for CLIMAtE Services

DOI – Digital Object Identifier

ECMWF – European Centre for Medium-Range Weather Forecasts

ECV – Essential Climate Variables

EQC – Evaluation and Quality Control

ERA5/6 – Fifth/Sixth generation of ECMWF atmospheric Re-Analyses
ESA-CCI – European Space Agency Climate Change Initiative
ESMValTool – Earth System Model Evaluation Tool
FAIR – Findable, Accessible, Interoperable, Reusable
FP7 – Seventh Framework Programme
GCOS – Global Climate Observing System
GRUAN – GCOS Reference Upper-Air Network
ISO – International Organization for Standardization
MARS – Meteorological Archival and Retrieval System
MR – Minimum Requirement
Obs4MIPs – Observations for Model Intercomparisons Project
PQAD – Product Quality Assurance Document
PQAR – Product Quality Assessment Report
PUG – Product User Guide
QA4ECV – Quality Assurance for Essential Climate Variables
QAR – Quality Assurance Report
QAT – Quality Assurance Template
SEAS5 – Fifth generation of the ECMWF seasonal forecasting system
SMM – System Maturity Matrix
TRUST – Transparency, Responsibility, User focus, Sustainability and Technology
UERRA – Uncertainties in Ensembles of Regional ReAnalysis
WIGOS – WMO Integrated Global Observing System
WMO – World Meteorological Organization

ACKNOWLEDGEMENTS

This study is based on work carried out in the C3S_512 contract funded by Copernicus Programme and operated by ECMWF on behalf of the European Commission (Service Contract number: ECMWF/COPERNICUS720187C3S_512_BSC). We would like to acknowledge the work of colleagues from several European institutions, the data providers and C3S, who contributed to the development of the EQC framework as well as to the QAR production. We would also like to acknowledge the focus group users, who took time to review and provide valuable feedback on the QARs, QATs, minimum requirements and the CDS quality assessment tab. The authors are grateful to the anonymous reviewers for their constructive comments that have helped for the improvement of this paper.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

CL is the main contributor to conceptualization, project planning, and writing of the original draft. FD contributed significantly to conceptualization, project planning, reviewing, and editing. GL contributed significantly to conceptualization, writing of the original draft, reviewing, and editing. CB contributed significantly to conceptualization, reviewing, and editing. AO and CS contributed significantly to conceptualization, project planning, reviewing, and editing. MC, DS, PB, SP, VR, DP, FS, AL, AP, DC, OM, PC, NP, FM, MR, AR, and MG contributed to conceptualization, reviewing, and editing.

Carlo Lacagnina  orcid.org/0000-0001-9434-9809

Barcelona Supercomputing Center, Barcelona, Spain

Francisco Doblas-Reyes  orcid.org/0000-0002-6622-4280

Barcelona Supercomputing Center, Barcelona, Spain; Institutió Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Gilles Larnicol  orcid.org/0000-0002-6694-1458

Magellium, Ramonville-Saint-Agne, France

Carlo Buontempo  orcid.org/0000-0002-1874-5380

ECMWF, Shinfield Park, Reading, UK

André Obregón  orcid.org/0000-0001-8759-4451

ECMWF, Shinfield Park, Reading, UK

Montserrat Costa-Surós  orcid.org/0000-0002-3319-4513

Barcelona Supercomputing Center, Barcelona, Spain

Daniel San-Martín  orcid.org/0000-0002-2862-2704

Predictia Intelligent Data Solutions SL, Santander, Spain

Pierre-Antoine Bretonnière  orcid.org/0000-0002-3066-6685

Barcelona Supercomputing Center, Barcelona, Spain

Suraj D. Polade  orcid.org/0000-0002-3892-1433

Finnish Meteorological Institute, Helsinki, Finland

Vanya Romanova  orcid.org/0000-0001-8059-108X

Deutscher Wetterdienst, Offenbach, Germany

Davide Putero  orcid.org/0000-0002-9721-1036

National Research Council of Italy (CNR) – Institute of Atmospheric Sciences and Climate (ISAC), Turin – Bologna, Italy

Federico Serva  orcid.org/0000-0002-7118-0817

National Research Council of Italy (CNR) – Institute of Marine Sciences (ISMAR), Rome, Italy

Alba Llabrés-Brustenga  orcid.org/0000-0003-2144-675X

Barcelona Supercomputing Center, Barcelona, Spain

Antonio Pérez  orcid.org/0000-0002-0250-1461

Predictia Intelligent Data Solutions SL, Santander, Spain

Davide Cavaliere  orcid.org/0000-0002-7466-8893

National Research Council of Italy (CNR) – Institute of Marine Sciences (ISMAR), Rome, Italy

Olivier Membrane  orcid.org/0000-0001-9773-6150

Météo-France – Centre de Météorologie Spatiale, Lannion, France

Christian Steger  orcid.org/0000-0002-8244-8751

Deutscher Wetterdienst, Offenbach, Germany

Núria Pérez-Zanón  orcid.org/0000-0001-8568-3071

Barcelona Supercomputing Center, Barcelona, Spain

Paolo Cristofanelli  orcid.org/0000-0001-5666-9131

National Research Council of Italy (CNR) – Institute of Atmospheric Sciences and Climate (ISAC), Turin – Bologna, Italy

Fabio Madonna  orcid.org/0000-0001-7628-8870

National Research Council of Italy (CNR) – Institute of Methodologies for Environmental Analysis (IMAA), Potenza, Italy

Marco Rosoldi  orcid.org/0000-0001-8688-481X

National Research Council of Italy (CNR) – Institute of Methodologies for Environmental Analysis (IMAA), Potenza, Italy

Aku Riihelä  orcid.org/0000-0001-6581-8792

Finnish Meteorological Institute, Helsinki, Finland

Markel García Díez  orcid.org/0000-0002-4760-0527

Predictia Intelligent Data Solutions SL, Santander, Spain

REFERENCES

- Aguilar, E, Auer, I, Brunet, M, Peterson, TC and Wieringa, J.** 2003. Guidelines on climate metadata and homogenization, WCDMP-No. 53, WMO-TD No. 1186. Geneva: World Meteorological Organization.
- Brunet, M, Brugnara, Y, Noone, S, Stephens, A, Valente, M A, Ventura, C, Jones, P, Gilabert, A, Brönnimann, S, Luterbacher, J, Allan, R, Brohan, P and Compo, GP.** 2020. Best Practice Guidelines for Climate Data and Metadata Formatting, Quality Control and Submission. Reading, UK: Copernicus Climate Change Service.

- Buontempo, C, Hanlon, HM, Bruno Soares, M, Christel, I, Soubeyroux, J-M, Viel, C, Calmanti, S, Bosi, L, Falloon, P, Palin, EJ, Vanvyve, E, Torralba, V, Gonzalez-Reviriego, N, Doblaz-Reyes, F, Pope, ECD, Newton, P and Liggins, F.** 2018. What have we learnt from EUPORIAS climate service prototypes? *Climate Services*, 9: 21–32. DOI: <https://doi.org/10.1016/j.cliser.2017.06.003>
- Callahan, T, Barnard, J, Helmkamp, L, Maertens, J and Kahn, M.** 2017. Reporting data quality assessment results: Identifying individual and organizational barriers and solutions. *eGEMs*, 5(1). DOI: <https://doi.org/10.5334/egems.214>
- European Commission (EC).** 2020. Copernicus and earth observation in support of eu policies. Part I, Copernicus uptake in the european commission. DOI: <https://doi.org/10.2760/024084>
- European Commission (EC), Directorate-General for Research and Innovation.** 2015. European Union: A European Research and Innovation Roadmap for Climate Services. DOI: <https://doi.org/10.2777/702151>
- European Organization for the Exploitation of Meteorological Satellites (EUMETSAT).** 2014. CORE-CLIMAX System Maturity Matrix Instruction Manual (Doc. No. CC/EUM/MAN/13/002). Available at https://masif.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_CORE_CLIMAX_MANUAL&RevisionSelectionMethod=LatestReleased&Rendition=Web.
- Evans, B, Druken, K, Wang, J, Yang, R, Richards, C and Wyborn, L.** 2017. A data quality strategy to enable fair, programmatic access across large, diverse data collections for high performance data analysis. *Informatics*, 4(4): 45. DOI: <https://doi.org/10.3390/informatics4040045>
- Hewitt, CD, Allis, E, Mason, SJ, Muth, M, Pulwarty, R, Shumake-Guillemot, J, Bucher, A, Brunet, M, Fischer, AM, Hama, AM, Kolli, RK, Lucio, F, Ndiaye, O and Tapia, B.** 2020. Making society climate resilient: International progress under the global framework for climate services. *Bulletin of the American Meteorological Society*, 101(2): E237–E252. DOI: <https://doi.org/10.1175/BAMS-D-18-0211.1>
- ISO 14090:2019.** Adaptation to climate change — Principles, requirements and guidelines. Geneva, Switzerland. <https://www.iso.org/standard/68507.html>.
- ISO 14091:2021.** Adaptation to climate change — Guidelines on vulnerability, impacts and risk assessment. Geneva, Switzerland. <https://www.iso.org/standard/68508.html>.
- ISO 19157:2013.** Geographic information — Data quality. Geneva, Switzerland. <https://www.iso.org/standard/32575.html>.
- Lawrence, B, Jones, C, Matthews, B, Pepler, S and Callaghan, S.** 2011. Citation and peer review of data: Moving towards formal data publication. *International Journal of Digital Curation*, 6(2): 4–37. DOI: <https://doi.org/10.2218/ijdc.v6i2.205>
- Leadbetter, A, Carr, R, Flynn, S, Meaney, W, Moran, S, Bogan, Y, Brophy, L, Lyons, K, Stokes, D and Thomas, R.** 2020. Implementation of a data management quality management framework at the marine institute, Ireland. *Earth Science Informatics*, 13(2): 509–521. DOI: <https://doi.org/10.1007/s12145-019-00432-w>
- Lin, D, Crabtree, J, Dillo, I, Downs, RR, Edmunds, R, Giarretta, D, De Giusti, M, L’Hours, H, Hugo, W, Jenkyns, R, Khodiyar, V, Martone, ME, Mokrane, M, Navale, V, Petters, J, Sierman, B, Sokolova, DV, Stockhause, M and Westbrook, J.** 2020. The TRUST Principles for digital repositories. *Scientific Data*, 7(1): 144. DOI: <https://doi.org/10.1038/s41597-020-0486-7>
- Medri, S, Banos de Guisasaola, E and Gualdi, S.** 2012. Overview of the main international climate services. Social Science Research Network. SSRN Scholarly Paper ID 2194841. DOI: <https://doi.org/10.2139/ssrn.2194841>
- Nightingale, J, Boersma, KF, Muller, J-P, Compernelle, S, Lambert, J-C, Blessing, S, Giering, R, Gobron, N, De Smedt, I, Coheur, P, George, M, Schulz, J and Wood, A.** 2018. Quality assurance framework development based on six new ecv data products to enhance user confidence for climate applications. *Remote Sensing*, 10(8): 1254. DOI: <https://doi.org/10.3390/rs10081254>
- Nightingale, J, Mittaz, JPD, Douglas, S, Dee, D, Ryder, J, Taylor, M, Old, C, Dieval, C, Fouron, C, Duveau, G and Merchant, C.** 2019. Ten priority science gaps in assessing climate data record quality. *Remote Sensing*, 11(8): 986. DOI: <https://doi.org/10.3390/rs11080986>
- Peng, G.** 2018. The state of assessing data stewardship maturity – an overview. *Data Science Journal*, 17: 7. DOI: <https://doi.org/10.5334/dsj-2018-007>
- Peng, G, Lacagnina, C, Downs, RR, Ramapriyan, H, Ivánová, I, Ganske, A, le Roux, J, et al.** 16 Apr. 2021. International Community Guidelines for Sharing and Reusing Quality Information of Individual Earth Science Datasets, OSF Preprints. DOI: <https://doi.org/10.31219/osf.io/xsu4p>
- Rfl, German Council for Scientific Information Infrastructures.** 2020. The Data Quality Challenge. Recommendations for Sustainable Research in the Digital Turn. Göttingen.
- Stockhause, M, Höck, H, Toussaint, F and Lautenschlager, M.** 2012. Quality assessment concept of the World Data Center for Climate and its application to CMIP5 data. *Geoscientific Model Development*, 5(4): 1023–1032. DOI: <https://doi.org/10.5194/gmd-5-1023-2012>
- Thépaut, J, Dee, D, Engelen, R and Pinty, B.** 2018. The Copernicus Programme and its Climate Change Service. *IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium*, 1591–1593. DOI: <https://doi.org/10.1109/IGARSS.2018.8518067>

- Wilkinson, MD, Dumontier, M, Aalbersberg, IJ, Appleton, G, Axton, M, Baak, A, Blomberg, N, Boiten, J-W, da Silva Santos, LB, Bourne, PE, Bouwman, J, Brookes, AJ, Clark, T, Crosas, M, Dillo, I, Dumon, O, Edmunds, S, Evelo, CT, Finkers, R, Mons, B, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1): 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>
- WMO/WIGOS. 2017. WIGOS Metadata Standard. Geneva: World Meteorological Organization, WMO no. 1192.
- Zeng, Y, Su, Z, Barmpadimos, I, Perrels, A, Poli, P, Boersma, F, Frey, A, Ma, X, Bruin, K de, Goosen, H, John, VO, Roebeling, R, Schulz, J and Timmermans, WJ. 2019. Towards a traceable climate service: Assessment of quality and usability of essential climate variables. *Remote Sensing*, 11(10): 1–28. DOI: <https://doi.org/10.3390/rs11101186>

TO CITE THIS ARTICLE:

Lacagnina, C, Doblás-Reyes, F, Larnicol, G, Buontempo, C, Obregón, A, Costa-Surós, M, San-Martín, D, Bretonnière, P-A, Polade, SD, Romanova, V, Putero, D, Serva, F, Llabrés-Brustenga, A, Pérez, A, Cavaliere, D, Membrive, O, Steger, C, Pérez-Zanón, N, Cristofanelli, P, Madonna, F, Rosoldi, M, Riihelä, A, Diez, MG. 2022. Quality Management Framework for Climate Datasets. *Data Science Journal*, 21: 10, pp. 1–25. DOI: <https://doi.org/10.5334/dsj-2022-010>

Submitted: 25 November 2021

Accepted: 17 March 2022

Published: 04 April 2022

COPYRIGHT:

© 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.