BAMS Article

Independent Quality Assessment of Essential Climate Variables

Lessons Learned from the Copernicus Climate Change Service

Chunxue Yang, Chiara Cagnazzo, Vincenzo Artale, Bruno Buongiorno Nardelli, Carlo Buontempo, Jacopo Busatto, Luca Caporaso, Claudia Cesarini, Irene Cionni, John Coll, Bas Crezee, Paolo Cristofanelli, Vincenzo de Toma, Yassmin Hesham Essa, Veronika Eyring, Federico Fierli, Luke Grant, Birgit Hassler, Martin Hirschi, Philippe Huybrechts, Eva Le Merle, Francesca Elisa Leonelli, Xia Lin, Fabio Madonna, Evan Mason, François Massonnet, Marta Marcos, Salvatore Marullo, Benjamin Müller, Andre Obregon, Emanuele Organelli, Artur Palacz, Ananda Pascual, Andrea Pisano, Davide Putero, Arun Rana, Antonio Sánchez-Román, Sonia I. Seneviratne, Federico Serva, Andrea Storto, Wim Thiery, Peter Throne, Lander Van Tricht, Yoni Verhaegen, Gianluca Volpe, and Rosalia Santoleri

ABSTRACT: If climate services are to lead to effective use of climate information in decision-making to enable the transition to a climate-smart, climate-ready world, then the question of trust in the products and services is of paramount importance. The Copernicus Climate Change Service (C3S) has been actively grappling with how to build such trust: provision of demonstrably independent assessments of the quality of products, which was deemed an important element in such trust-building processes. C3S provides access to essential climate variables (ECVs) from multiple sources to a broad set of users ranging from scientists to private companies and decision-makers. Here we outline the approach undertaken to coherently assess the quality of a suite of observation- and reanalysis-based ECV products covering the atmosphere, ocean, land, and cryosphere. The assessment is based on four pillars: basic data checks, maturity of the datasets, fitness for purpose (scientific use cases and climate studies), and guidance to users. It is undertaken independently by scientific experts and presented alongside the datasets in a fully traceable, replicable, and transparent manner. The methodology deployed is detailed, and example assessments are given. These independent scientific quality assessments are intended to guide users to ensure they use tools and datasets that are fit for purpose to answer their specific needs rather than simply use the first product they alight on. This is the first such effort to develop and apply an assessment framework consistently to all ECVs. Lessons learned and future perspectives are outlined to potentially improve future assessment activities and thus climate services.

KEYWORDS: Climate change; Climate records; Quality assurance/control; Satellite observations; Reanalysis data; Climate services

https://doi.org/10.1175/BAMS-D-21-0109.1
Corresponding author: Chunxue Yang, chunxue.yang@cnr.it
In final form 28 March 2022

AFFILIATIONS: Yang, Artale, Busatto, de Toma, Essa, Le Merle, Leonelli, Organelli, Pisano, Serva, Storto, Volpe, and Santoleri—Institute of Marine Sciences, National Research Council of Italy, Rome, Italy; **Cagnazzo, Buontempo, and Obregon**—European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom; **Buongiorno Nardelli**—Institute of Marine Sciences, National Research Council of Italy, Naples, Italy; **Caporaso**—Institute of Marine Sciences, National Research Council of Italy, Rome, and European Commission Joint Research Centre, Ispra, Italy; **Cesarini**—CLU Ltd., Castelfranco Emilia, Italy; **Cionni**—Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Frascati, Italy; **Coll and Throne**—Irish Climate Analysis and Research Units, Maynooth University, Maynooth, County Kildare, Ireland; **Crezee, Hirschi, and Seneviratne**—Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland; **Cristofanelli and Putero**—Institute of Atmospheric Sciences and Climate, National Research Council of Italy, Bologna, Italy; **Eyring**—Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, and University of Bremen, Institute of Environmental Physics (IUP), Bremen, Germany; **Fierli**—European Organisation for Exploitation of Meteorological Satellites, Darmstadt, Germany; **Grant and Thiery**—Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium; **Hassler**—Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany; **Huybrechts, Van Tricht, and Verhaegen**—Earth System Science, and Departement Geografie, Vrije Universiteit Brussel, Brussels, Belgium; **Lin, Massonnet, and Rana**—Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium; **Madonna**—Institute of Methodologies for Environmental, National Research Council of Italy, Rome, Italy; **Mason, Pascual, and Sánchez-Román**—Instituto Mediterráneo de Estudios Avanzados, Esporles, Spain; **Marcos**—Instituto Mediterráneo de Estudios Avanzados, Esporles, and Departament de Física, Universitat de les Illes Balears, Palma, Spain; **Marullo**—Institute of Marine Sciences, National Research Council of Italy, Rome, and Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Frascati, Italy; **Müller**—Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany; **Palacz**—Institute of Oceanography, Polish Academy of Sciences, Sopot, Poland

The awareness of the risks posed by the changing climate has grown rapidly in the last decade, involving decision-makers, politicians, private sectors, scientists, and the public, as demonstrated by the recently released Intergovernmental Panel on Climate Change climate assessment reports (IPCC 2022) that have drawn tremendous attention worldwide. To effectively manage risks and mitigate impacts associated with the changing climate, delivering quality-assured climate services is essential. To respond to these needs, the European Commission launched the Copernicus Climate Change Service (C3S) in 2014 to provide climate data and authoritative information about climate change to support European Union climate adaptation and mitigation policies. The C3S is implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) and became operational in 2018. The backbone of the C3S is the climate data freely distributed through the Climate Data Store (CDS; https://cds.climate.copernicus.eu), which is a single-entry point to a catalogue of climate datasets and derived products including essential climate variables (ECVs) from satellite and in situ observations, climate reanalysis, seasonal forecasts, and climate projections. The C3S has launched a coordinated effort to perform an overarching quality assessment of the CDS components, referred to as the evaluation and quality control (EQC) function, to ensure that the climate data, applications, and tools are reliable and suitable to meet users' varied needs.

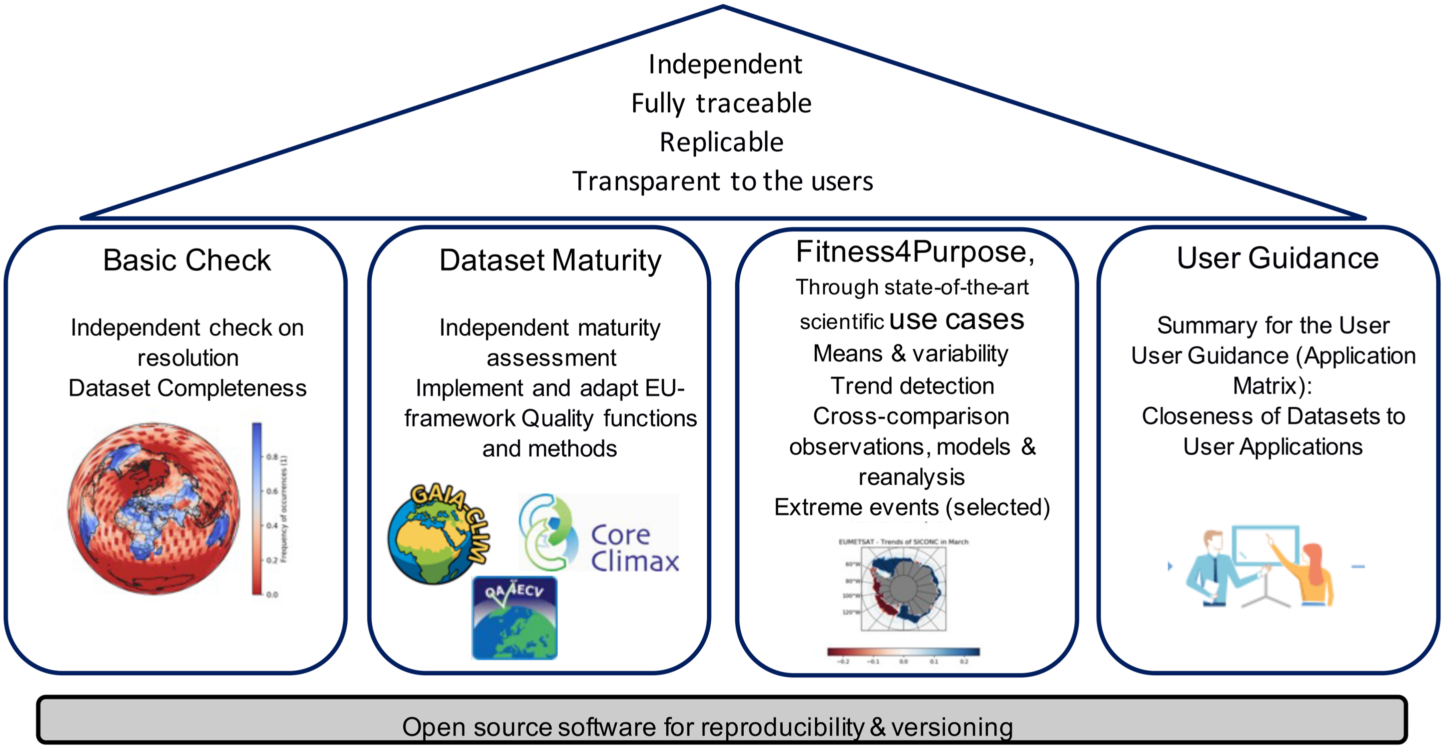

Fig. 1. The independent assessment framework developed in IQA-ECVs.

One essential component of C3S EQC functions is the expert-based Independent Quality Assessment of ECVs (C3S_511, hereafter IQA-ECVs) launched in 2017. The newly developed assessment framework (Fig. 1) is based on previous C3S EQC activities such as C3S_51 Lot 2[1] and precursor EU-FP7 projects like QA4ECV,[2] CORE-CLIMAX,[3] and GAIA-CLIM.[4] The goal of IQA-ECVs is to deliver timely independent assessments of the quality and fitness for purpose of ECV products as defined by the Global Climate Observing System (GCOS) (Table 1) from reanalyses, satellite, and in situ observations covering atmosphere, ocean, and terrestrial domains. It is a step forward to have as much as possible a consistent and overarching assessment approach for ECVs from different domains and sources.

[1] https://climate.copernicus.eu/c3s51-evaluation-and-quality-control-function-climate-data-store
[2] www.qa4ecv.eu/
[3] www.coreclimax.eu/
[4] www.gaia-clim.eu/

Although during the production of these ECVs datasets data providers themselves have, to varying extents and by a variety of means, validated the quality of the climate datasets, the provision of a further consistent tier of independent assessment by scientific experts can assess each dataset across several aspects such as traceability and documentation, provision of expert-level guidance and, ultimately, complement the evaluation from data providers (Su et al. 2018). The main difference from the data providers evaluation is that the IQA-ECVs framework addresses the quality of these climate datasets in terms of the fitness for purpose for climate studies and applications and provides C3S users with guidance on the reliability,

Table 1. Essential climate variables (ECVs) assessed in IQA-ECVs.

Domain	Variables
Surface atmosphere	Temperature, wind, humidity, pressure, radiation, precipitation
Atmosphere	Ozone, temperature, winds, aerosol, carbon dioxide, methane and other greenhouse gases, clouds, moisture
Ocean	Sea surface temperature, sea ice, sea level, surface currents, surface heat fluxes, surface stress, surface salinity, subsurface currents, subsurface temperature, subsurface salinity, inorganic carbon, ocean color, sea state
Land	Soil moisture, albedo, snow, glaciers, land cover, fraction of absorbed photosynthetically active radiation, leaf area index, ice sheets, ice caps, lakes, fires, permafrost, runoff

usability, and suitability of climate datasets in a consistent manner. Such independent assessments can also provide key feedback to data providers on how to improve their datasets in terms of overall quality, documentation, usability, and accessibility, thus contributing to increasing the maturity of the assessed dataset.

The core function of IQA-ECVs is deliver in-depth independent quality assessments and summarize key information in a scientific and informative format and provide user guidance on the reliability and suitability of ECV datasets for climate studies and applications. To achieve this goal, the IQA-ECVs service has gathered top-level EU institutions with practitioners who can supply the required expertise on all the target ECVs from different disciplines.

The aim of this paper is to outline the newly developed evaluation framework (IQA-ECVs) benefiting from previous climate service activities, which could potentially be applied to similar EQC activities beyond C3S. Example assessments are given to show how the assessment was performed for selected ECVs. Lessons learned and perspectives are outlined to provide suggestions to design future assessment activities.

Independent assessment framework

The independent assessment framework (Fig. 1) developed in IQA-ECVs comprises four fundamental pillars: basic data check, data maturity, fitness for purpose, and guidance to users. In addition, two key elements are included to support the whole assessment process: 1) source codes used for evaluations are open to users to ensure the traceability of the assessment, and 2) external (experts outside of IQA-ECVs team) feedback is collected to enhance the quality of the assessment.

The following sections describe the rationale behind the design of each element of the assessment framework and examples are given to illustrate the assessment procedure.

Basic data check. The first step is to assess the technical quality of ECV datasets by providing users basic characteristics including the metadata, data format, and the spatiotemporal coverage and resolution. This basic information helps users to understand the potential suitability of a specific ECV dataset for their usage at a glance based upon data availability and format.

Table 2 presents an example of basic check of satellite ozone data.[5] Detailed evaluation of the spatial and temporal coverage identifies any spatiotemporal gaps within the datasets. For example, Fig. 2 shows that data coverage of satellite ozone over the Arabian and Sahara Deserts decrease during spring and summer seasons likely due to dust storm events that affect the data retrieval as detailed in the evaluation report.

[5] https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-ozone-v1?tab=form

Table 2. Basic characteristics of satellite (sensor IASI-B) total column ozone data (available online at https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-ozone-v1?tab=form). The first column presents basic characteristics of the data and the second column shows modifications (if any) done to the original data by evaluators who may need to convert the units or take full years of the temporal coverage to follow conventional analysis.

	Provided on the CDS	For evaluation report
Horizontal resolution	1° × 1°	1° × 1°
Temporal resolution	Monthly	Monthly
Spatial coverage	Global	Global
Temporal coverage	May 2013–May 2019	The evaluated dataset is January 2014–December 2018.
Variables	ozone_mole_content (mol m^{-2})	For easier-to-understand figures, the units were adjusted to DU. Data are given in mol m^{-2}. Multiplication factors for the conversion to DU are given in the original file.
Format	NetCDF—CF compliant	The dataset was slightly modified. Single monthly files have been aggregated along the time dimension. The units were converted to DU.

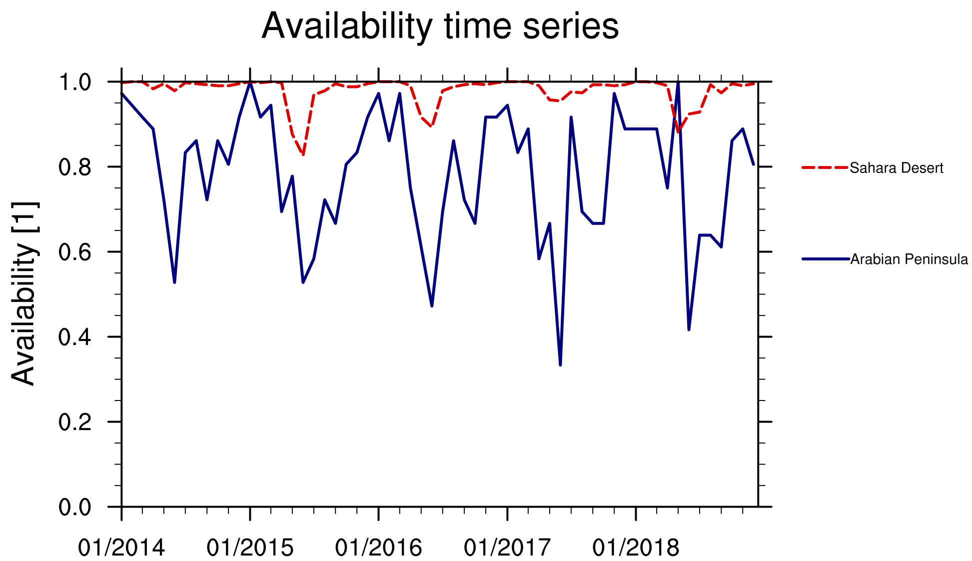

Fig. 2. Data availability time series (aggregated over longitude and latitude) of tropospheric column of ozone for the dataset 06TC_IASIB (2014–18) over part of the Arabian Peninsula (19°–25°N, 50°–56°E) and Sahara Desert (16°–30°N, 10°–30°E). Daytime values only. Figure produced on 26 May 2020 with CDS data downloaded on 26 May 2020 (refer to quality assessment available in https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-ozone-v1?tab=form).

Data maturity. The second step is to evaluate the maturity of each ECV dataset to help users answer questions such as, 1) Is the dataset served well documented, stored, and disseminated? 2) Has the uncertainty been systematically assessed? 3) Have these datasets been widely used and user feedback systems been implemented to improve confidence in the product?

To evaluate these aspects, the maturity evaluation system used was developed based on the system maturity matrix (SMM) established by the Coordinating Earth Observation Data Validation for Reanalysis for Climate Services project (CORE-CLIMAX) (Su et al. 2018) to assess the maturity of ECV datasets in terms of documentation, storage, and dissemination (Peng et al. 2019). The maturity assessments are made based on five major categories of any data product assessed: 1) metadata; 2) user documentation; 3) uncertainty characterization; 4) public access, feedback, and update; and 5) usage. A matrix of scores ranging from 1 to 6 are assigned to reflect the maturity of each dataset with respect to specific criteria. Scores may evolve in time during the life cycle of a product. High scores indicate high maturity for a dataset in a specific category. Low scores might be given in some categories, especially during the earliest period of the life cycle of a product. However, low scores do not necessarily indicate low scientific quality of the dataset. An evolving guidance document (Coll et al. 2020a) has been developed and maintained to enable evaluators to give consistent scores to specific datasets such as reanalyses where these do not differ by ECV.

The outcome of the data maturity evaluation is presented as colorblind-friendly tables to highlight the maturity of the ECV (Fig. 3) and defensible traces (appendix) to justify the evaluation. The evaluation provides users with a synthesis of the status of selected ECV datasets, which requires a substantive supporting literature review on the part of the evaluators. Additionally, the maturity evaluation also helps data providers to identify weaknesses of their dataset in terms of the documentation, storage, and dissemination and encourages them to take actions to improve it for users' convenience.

Fitness for purpose. In addition to assessments of the technical quality, the assessment of scientific quality is essential to quantitatively evaluate to which extent the ECV data meet users' specific needs, referred to as fitness for purpose. The following scientific aspects are

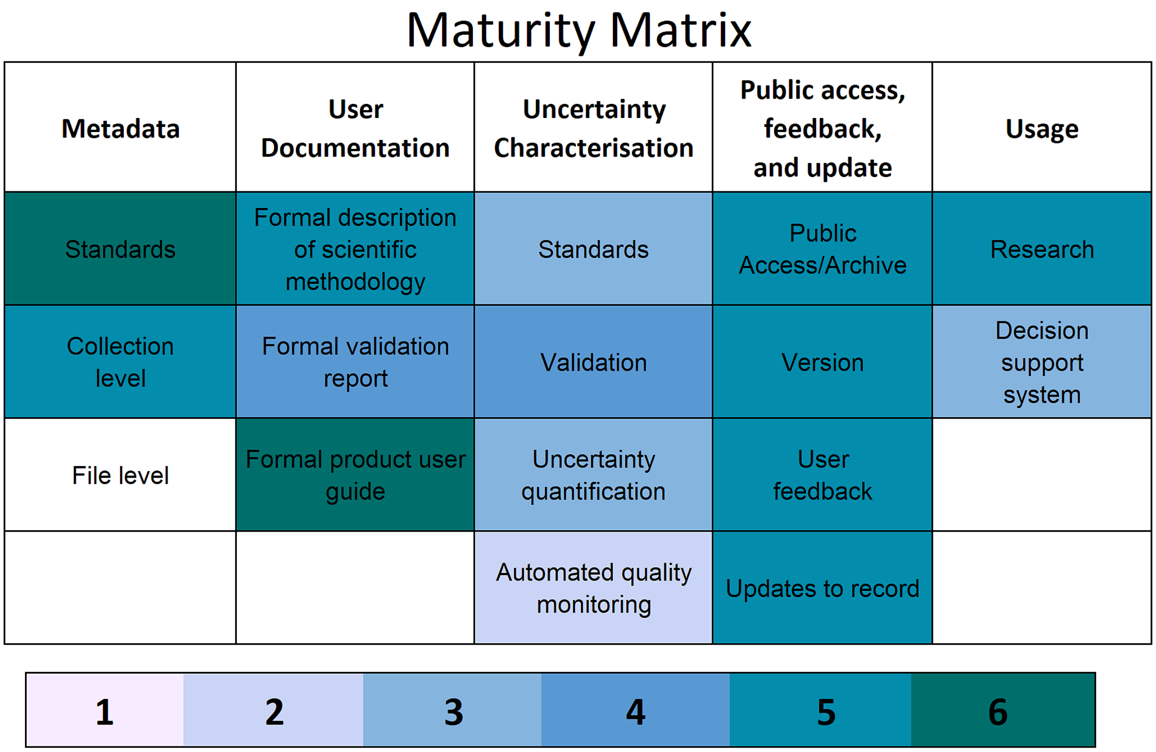

Fig. 3. Maturity matrix for the ERA5 winds on pressure levels (1979–2020). The different colors indicate different stages of maturity for the categories indicated and subcategories evaluated, whereby a score of 1 (light purple) denotes a less mature category than 6 (dark green). The respective meaning and details for these scores can be found in the accompanying document (https://confluence.ecmwf.int/display/CKB/Guidance+document+on+applying+the+Maturity+Matrix+as+part+of+the+Evaluation+and+Quality+Control).

assessed: 1) temporal and spatial homogeneity and potential discontinuities, 2) spatiotemporal variability at different time scales, 3) long-term homogeneity, 4) capability to reproduce known climate extremes (where appropriate), 5) suitability for climate applications, 6) consistencies and discrepancies with products from different sources (e.g., satellite, in situ, and reanalyses), and 7) physical consistency and coherency of climate signals across different ECVs.

To assess these scientific aspects, the following measures as a minimum set of diagnostics are quantified: 1) means and variability, 2) trends characterization, and 3) applicability for Earth system model (ESM) evaluation, 4) detection of climate extremes (where applicable), 5) application performance metrics, 6) cross comparison of observations and reanalyses for a specific ECV, and 7) integrative assessment across ECVs for climate studies.

The outcome of the evaluation is introduced by a general description of the dataset explaining the origin and production process based on the relevant literature and technical reports, followed by the assessments detailed below.

Means and variability. As a starting point, calculations of means and variability for the available time period, globally and for selected regions, are performed. Figure 4, for instance, presents the seasonal mean of ERA5 total cloud cover percentage (Hersbach et al. 2020), showing a prominent seasonal cycle. The global mean and standard deviation are provided for further indications of the scientific quality and compared, wherever possible, to values from the scientific literature.

Diagnostics such as the zonal or meridional means, as well as anomaly plots provide a more in-depth assessment to identify relevant climate events and potential discontinuities. Figure 5 provides an example of the assessment of the ERA5 upper-air temperatures (Hersbach et al. 2020). The variability associated with El Niño (e.g., the 1997/98 event) and

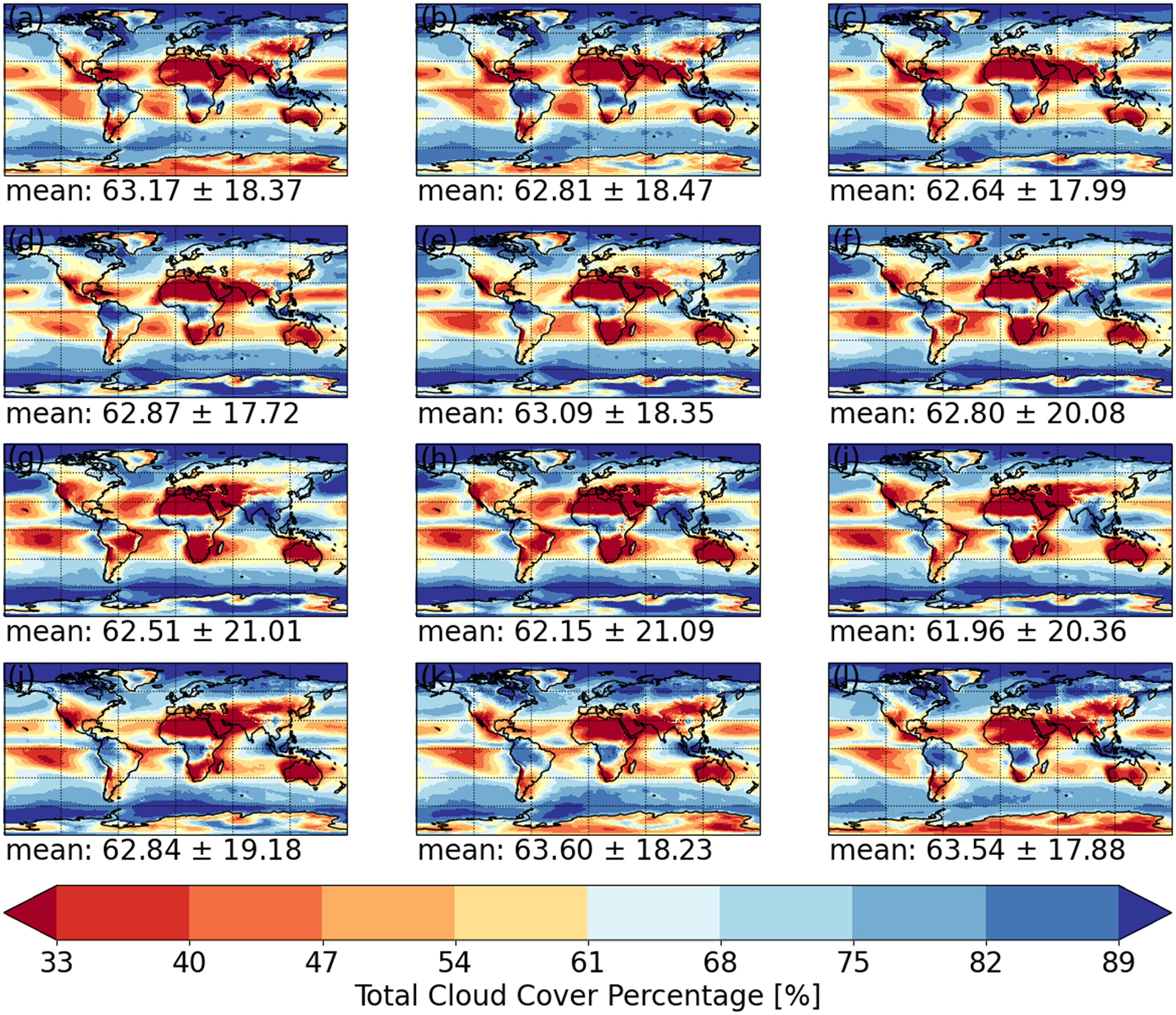

Fig. 4. Global monthly climatologies of ERA5 (1979–2020) total cloud-cover percentage: (a) January–(l) December with global mean value and standard deviation (in %) given below each panel.

large volcanic eruptions (e.g., El Chichón in 1982) is evident in this Hovmöller diagram. Potential spurious variability and discontinuities are also identified, e.g., in 1985 and 1998. Based on their scientific knowledge and literature review, evaluators have suggested an explanation for these discontinuities, in this case arising due to transitions in the assimilated observation system components.

Trends characterization. The detection of trends can be affected by data inhomogeneities and sudden changes in the time series due, for example, to changes in processing methods and observing networks. Depending on the ECV, modern statistical homogenization methods can deliver different breakpoint assessments (Coll et al. 2020b; Madonna et al. 2022). This is particularly important for merged satellite datasets produced by concatenation of successive, and sometimes overlapping, satellite platforms (Weatherhead et al. 2017).

To evaluate the suitability and reliability of products for trend analysis in climate studies and applications, the IQA-ECVs assessment framework first quantifies the trends and their robustness. Based on the scientific knowledge of evaluators and the review of

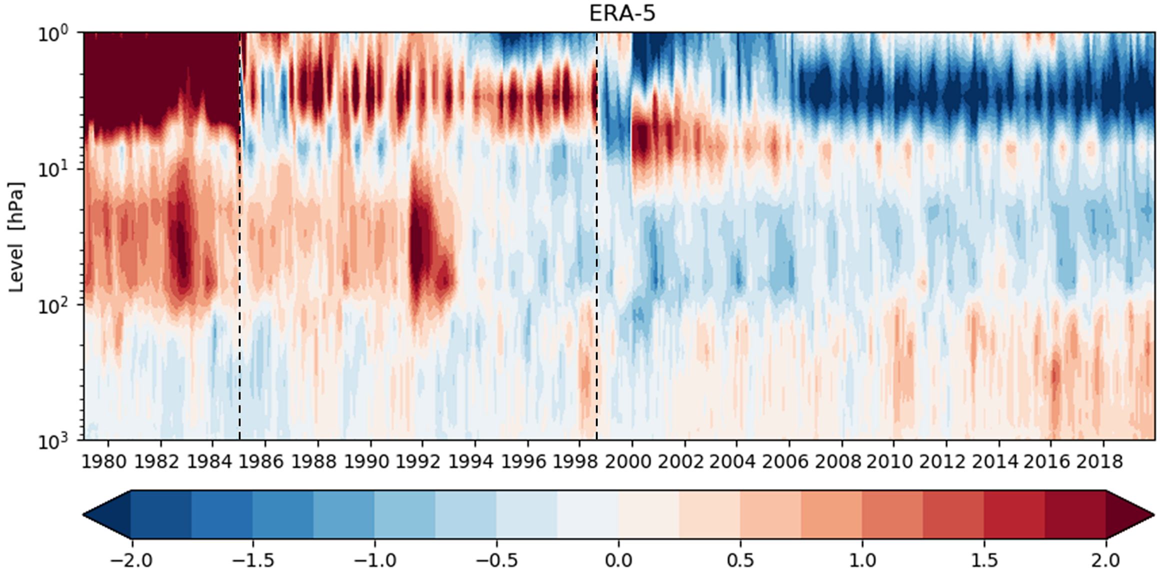

Fig. 5. ERA5 global upper-air temperature anomaly (K) over the period (1979–2019) as a function of time and pressure (in logarithmic representation on *y* axis). Anomaly is calculated by subtracting its monthly climatology from the monthly mean temperature. Dashed lines highlight the discontinuities at years 1985 and 1998.

relevant literature, evaluators indicate limitations in the usage of estimated trends arising from the dataset.

An example of a linear trend assessment of in situ surface air temperature (from E-OBS) is provided in Fig. 6, showing linear trends consistent with results derived from reanalyses (van der Schrier et al. 2013). However, potential limitations of the E-OBS air temperature[6] linear trends analysis were detected in some regions. Particularly, the relative sparsity of the available stations in North Africa (Fig. 6) used for E-OBS may require caution in interpretation and use as the substantive warming is physically unlikely and unsupported by other estimates from other sources. Such potential (in)homogeneities in the ECV records are highlighted in the evaluation report in order to help users to understand the potential limitations of use of the evaluated product in trend analysis.

[6] https://cds.climate.copernicus.eu/cdsapp#!/dataset/insitu-gridded-observations-europe?tab=overview

Applicability for ESM evaluation. Facilitating the complex evaluation of ESMs through the development of evaluation tools was a key goal of the Coupled Model Intercomparison Project phase 6 (CMIP6; Eyring et al. 2016). Climate data from different sources (e.g., reanalysis, satellite observations, in situ) are indispensable for an objective ESM evaluation. Current datasets for ESM evaluation come from a heterogeneous set of sources (Flato et al. 2013; Lauer et al. 2017; Eyring et al. 2021). CDS provides an excellent resource of high-quality climate ECV records that in principle should be suitable for ESM evaluation. Therefore, an independent assessment to help determine the applicability of ECV records for model evaluation is important for the climate science community. In this context, the assessment produced by IQA-ECVs provides an essential tier of information to judge the suitability of an ECV record for ESM evaluation. As a first step we analyzed to what extent each data product has been used in model evaluation exercises in terms of the mean, variability, and trends based on a thorough literature review. Next, the potential future use and existing caveats relating to the dataset for ESM evaluation are considered. For example, the evaluation of the ECV soil moisture satellite data based on the European Space Agency Climate Change Initiative (ESA CCI) soil moisture[7] highlights that this dataset was used for evaluating ESMs in terms of climatology, variability, and trends to

[7] https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-soil-moisture?tab=overview

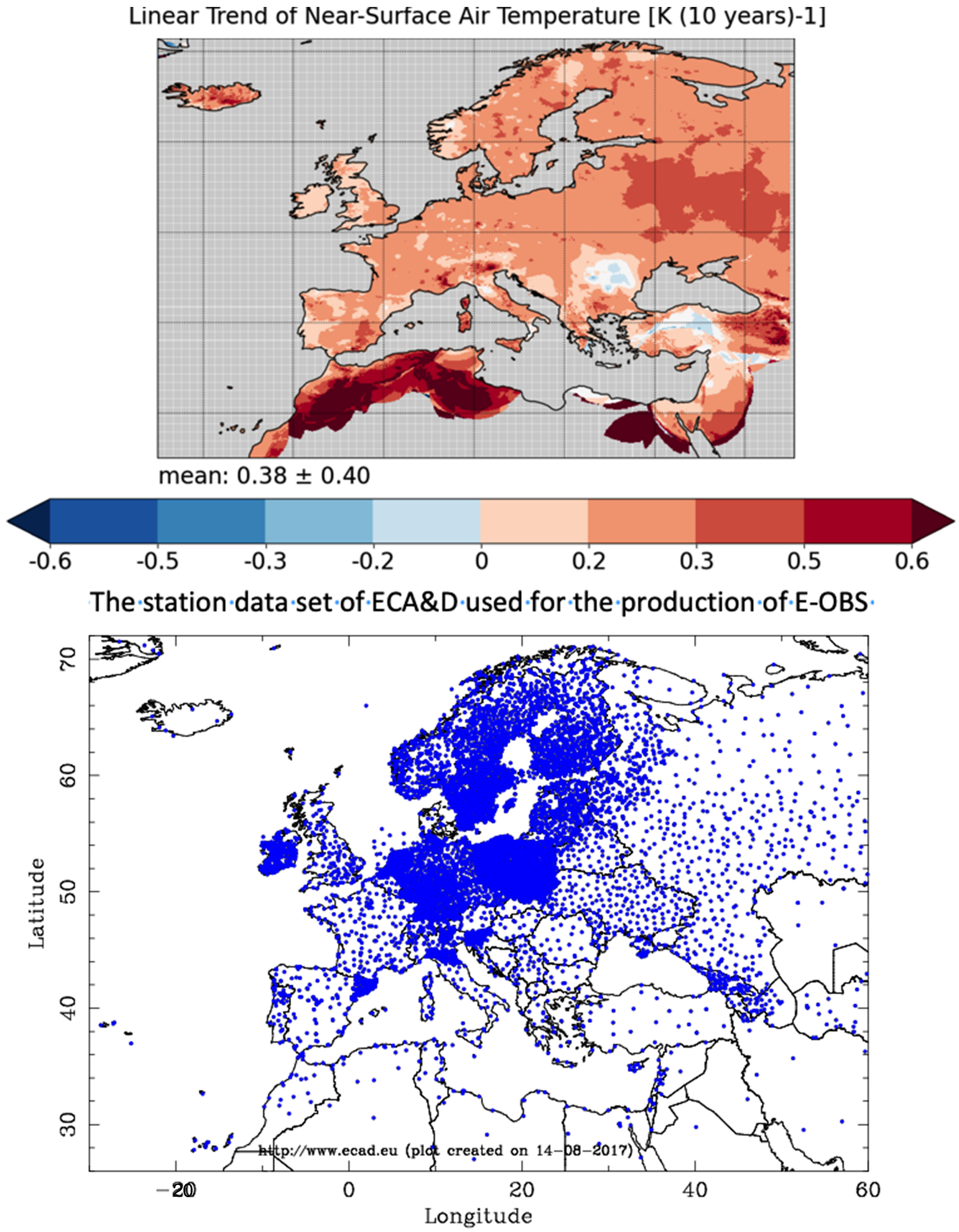

Fig. 6. (top) Spatial trends per decade of daily E-OBS v21.0 (1950–2019) for mean air temperature and (bottom) the station dataset of ECA&D used for the production of E-OBS (source: van der Schrier et al. 2019). Note that the spatial gradients of availability in part pertains to data policy and sharing policies rather than necessarily true gradients in potential data availability.

a lesser extent (e.g., Lauer et al. 2017) suggesting that this dataset is suitable for ESM evaluation. However, the assessment also emphasizes that caution should be taken in the interpretation of long-term trends due to inherent inhomogeneities which may result from the merging of different satellites in space and time (Su et al. 2016). Moreover, the quantitative comparison of absolute soil moisture values is impractical due to differences in how soil layer thickness is represented in remote sensing products and among ESMs.

Detection of climate extremes. ECV records should capture extremes of climate such as droughts, floods, and heatwaves, which are essential for many climate service applications. It is therefore important to evaluate how well extreme events are captured within assessed ECV records.

IQA-ECVs assesses the ability of ECV records to reproduce well-documented climate anomalies, in particular in temperature (surface and marine heatwaves), precipitation (floods and

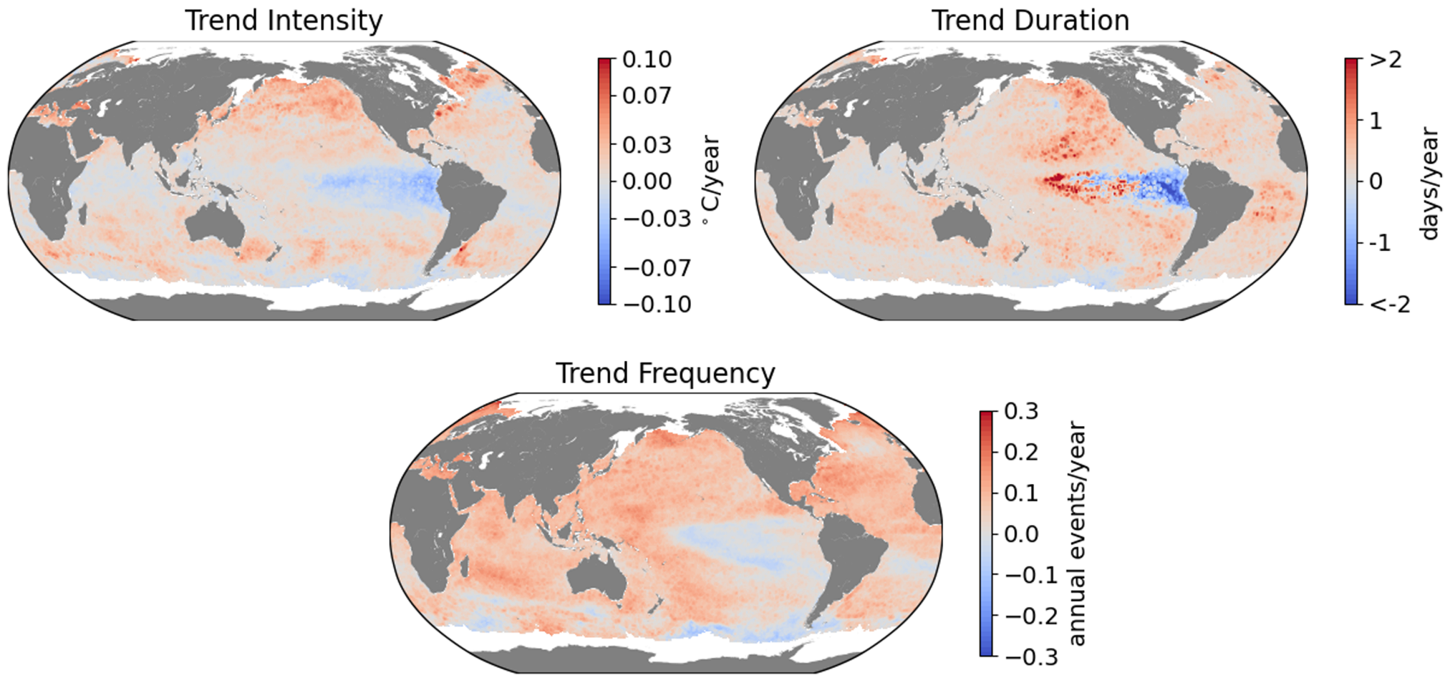

Fig. 7. Spatial distribution of average properties of marine heatwaves based on the ESA CCI SST level-4 dataset v2.0: (top left) annual mean intensity (°C); (top right) duration (days); and (bottom) frequency (events per year).

droughts), and soil moisture. This specific analysis focuses on well-documented and defined extreme events, such as droughts and heatwaves (e.g., Russia, 2010; China, 2010; Europe, 2003, 2006; Sahel 1980s) and floods (e.g., Europe, 2016). IQA-ECVs assembled a catalogue of well-documented extreme events (Crezee et al. 2019), which are also supported by the scientific literature, and defined a common matrix of fit criteria (magnitude, duration, extent) following by WMO guidelines (WMO 2018) to quantify the extreme events where possible. Figure 7 illustrates an ECV assessment of the representation of global marine heatwaves (Hobday et al. 2016), examining their intensity, duration, and frequency using the ESA CCI SST product (Merchant et al. 2019). Our results yield signals for the events, which are in agreement with previous studies (Hobday et al. 2018), and hence give confidence to users that this particular dataset can be used for extreme marine heatwave events analysis.

Application performance metrics. Assessments of the applicability of ECV data records to different climate applications are defined. A large number of dedicated services addressing specific applications are growing within C3S, which requires a clear enhancement of guidance for users. Expert evaluators in IQA-ECVs have used a combination of literature, technical reports, and their scientific expertise to select applications related to agriculture, environmental monitoring, etc., for each ECV dataset, and have assessed to what extent the records are likely to meet users' requirements. Building on this synthesis of information, a guidance document has been produced to assist the evaluation of suitability for specific climate applications (Muller et al. 2020). Tables 3 and 4 provide examples of potential climate applications and the evaluation of the suitability of carbon dioxide datasets for one specific application (CO_2 emission monitoring). We first describe the climate application and demonstrate its importance for a certain sector (Table 3). Then, based on the assessment of the characteristics of this dataset, the outcome of the assessment (Table 4) for users is that this dataset does not fulfill all the requirements due to its low spatial resolution and insufficient temporal coverage.

Cross comparison of observations and reanalyses. With the emergence of large multiobservational and multimodel ensembles, novel analyses become possible by comparing one

Table 3. Application use case description for ECV Carbon dioxide dataset (Buchwitz et al. 2018).

Use case title	Monitoring of carbon dioxide emissions reductions in megacities
Description	Carbon (dioxide) emissions from cities represent the single largest human contribution to climate change, and they are projected to undergo rapid change over the next two decades with growth in developing countries and stabilization in developed countries. Despite slow progress in reaching emissions stabilization agreements between nations, many of the largest cities are now taking actions that should have detectable impacts on emission trends. Usually, carbon dioxide emissions are reported to regional or national inventories for record keeping. However, there are discrepancies between different inventories and independent measurements. Monitoring the long-term CO_2 emissions from megacities with satellite instruments could help tracking the progress of emission reductions (Duren and Miller 2012).
ECV	Surface CO_2
Period (number of years)	—
Start time	—
Stop time	Today
Spatial coverage	Continents
Vertical coverage	Surface to ~3 km
Temporal resolution	Annually
Spatial resolution	~20 km
Vertical resolution	—
Data completeness	100% in regions of interest
Uncertainty estimates	Total uncertainty
Sectors of relevance	Policy

ECV from different sources (e.g., satellite and reanalyses). However, from an EQC perspective this presents additional challenges in terms of performing robust comparisons and guiding users to the most appropriate set of products to meet their needs. In IQA-ECVs, the evaluation aims to assess if products for one specific ECV from different sources measure the same signal and provide guidance on the potential use of a set of products available for this single ECV. We have implemented a suite of basic metrics (e.g., bias, correlation, trends) to process an arbitrary number of time series, and display in an integrated manner their consistency and highlight any discrepancies in terms of the mean state and variability. We confined our evaluation metrics to these basic diagnostics in order to have an operational climate service, which is crucial to climate monitoring. This approach/framework can be expanded by adding more diagnostics and metrics based on the needs. Nonetheless, these metrics (time series, trends, etc.) could efficiently assess the quality of the climate datasets in climate applications and climate monitoring (Van Den Besselaar et al. 2015).

Figure 8 highlights the intercomparison of sea ice concentration from satellite and reanalysis datasets, showing the difference between each dataset with an arbitrarily chosen benchmark dataset. The choice of a benchmark dataset does not imply that the product is correct but is based upon a clear rationale according to the evaluator's expertise on specific ECVs and based on routinely applied international standards. In this particular case the OSF-450[8] is chosen as it is a well-used product with the highest resolution at the time of the evaluation. We also extend the analysis to datasets available from outside the CDS to provide a more comprehensive evaluation of the performance and to better support users in choosing a suitable dataset for their applications.

[8] https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-sea-ice-concentration?tab=overview

Here we give sea ice as an example for cross evaluation of observations and reanalysis. Other selected ECVs is available in the CDS portal. Additionally, we also have produced peer-reviewed articles based on the outcome of these evaluations to further elaborate the quality of the climate dataset for specific ECVs [Yang et al. (2021) for SST; Hassler and Lauer (2021) for precipitation].

Table 4. Application table for the use cases of carbon dioxide datasets. The right column describes the dataset characteristics. Green (red) cells refer to the dataset matching (not matching) the use-case requirements.

Use case	Monitoring of carbon dioxide emissions reductions in megacities	Carbon dioxide data from 2002 to the present derived from satellite observations
Period	—	16 years
Start time	—	2003
Stop time	Today	2018
Spatial coverage	Continents	Approximately between 70°N and 70°S
Vertical coverage	—	Total atmospheric column
Temporal resolution	Annually	Monthly
Spatial resolution	~20 km	5° × 5°
Vertical resolution	—	Single layer
Data completeness	100% in the regions of interest	Some data missing over the oceans and high latitudes
Uncertainty		Total uncertainty
Total uncertainty		
Random uncertainty		
Systematic uncertainty		
Closeness		

Integrative assessment across ECVs for climate studies. Physical consistency across ECVs is assessed through thematic assessments. In IQA-ECVs, the following themes were considered: 1) ocean circulation and ocean heat content, 2) Mediterranean heat and freshwater budget closure, 3) global energy fluxes, 4) land–climate coupling, 5) the Arctic freshwater and heat budget and interactions with the North Atlantic, and 6) global phytoplankton community structure in response to changing environmental conditions. The proposed thematic assessments are closely following the GCOS initiatives to improve observations of energy, water, and carbon cycles (GCOS 2016). The thematic assessments, which combine different ECVs, evaluate the consistency (Popp et al. 2020) of the climate signals among these ECV datasets and assess their compatibility in the representation of energy, water, and carbon cycles. Our assessment, therefore, provides users with an evaluation of the climate state as represented by these ECV datasets and the suitability for conducting climate studies.

Guidance to users. After the overall independent assessment for one ECV in a single dataset, one ECV in multiple datasets, and across ECVs for climate studies, guidance to users is provided including a general description of the evaluation and essential information, such as key strengths, weaknesses, and known issues with the dataset. The suitability of ECV records depends on many factors, such as length of the data records, homogeneity of the time series, and the extent of missing data. Based on the evaluation, the assessment provides users with an independent expert assessment of the suitability of the ECV dataset for particular climate applications and/or climate studies.

Tools and external user feedback. The assessment carried out in our service is based on free software, such as custom python scripts and the Earth System Model Evaluation Tool (ESMValTool) v2.0 (Eyring et al. 2020; Lauer et al. 2020; Righi et al. 2020; Weigel et al. 2021), which are widely used in the climate community. All the code developed in IQA-ECVs is provided as open access to make sure the assessment is findable, accessible, interoperable, and reusable (FAIR). A few prototype diagnostics have been elaborated within the CDS Toolbox (https://cds.climate.copernicus.eu/cdsapp#!/toolbox) ecosystem. Future

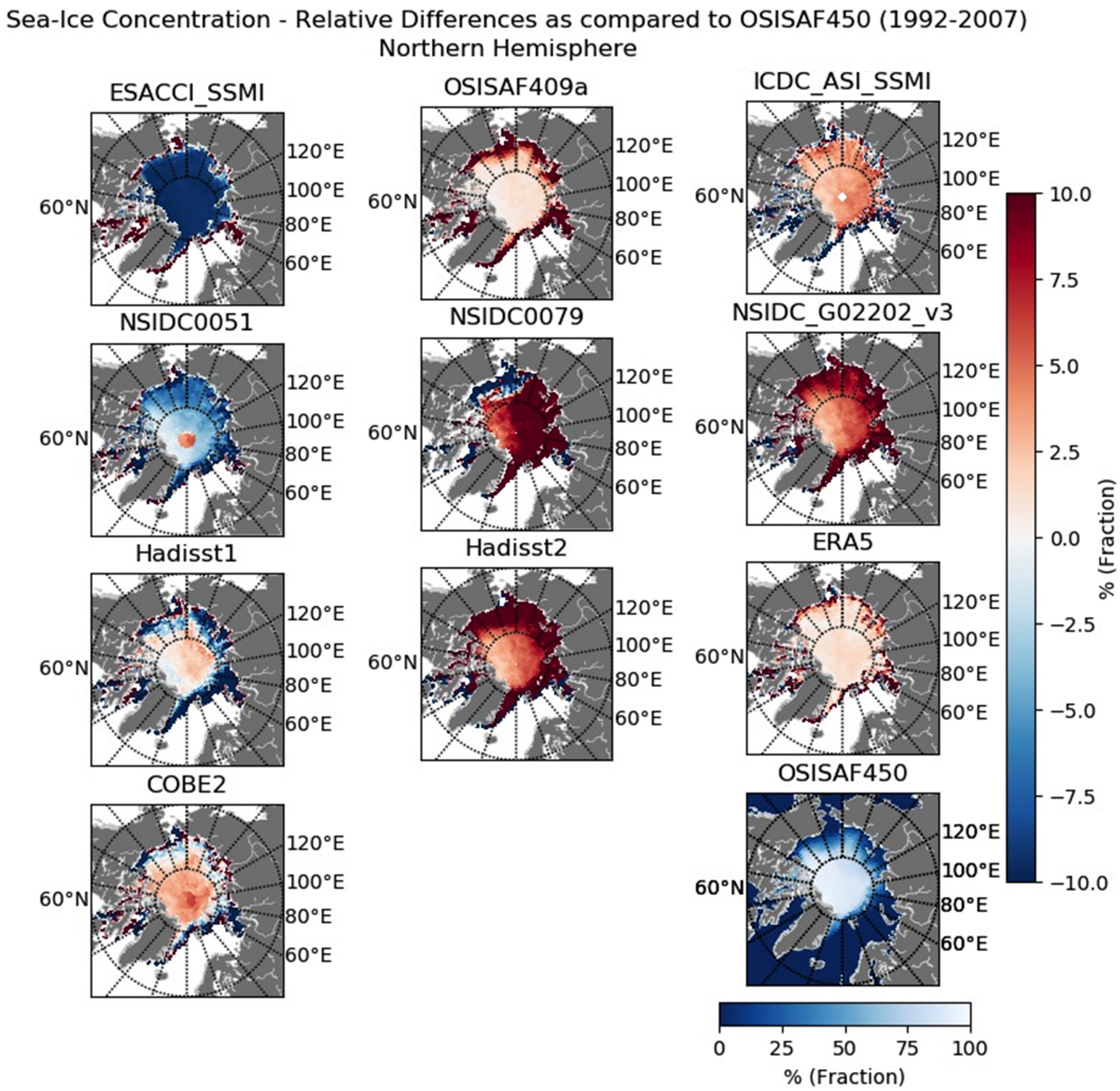

Fig. 8. Relative differences (Obs-RefObs/RefObs) of SICONC (%) as represented by various observation products over the Northern Hemisphere in September. (bottom right) Here RefObs is the OSI-SAF450 dataset.

developments of the Toolbox functionalities (e.g., support for external software and libraries) and compatibility with an increased number of datasets would allow a fully CDS-based reproducible assessment.

In IQA-ECVs, a user feedback methodology was designed and implemented to gather comments from users to enhance the independent assessment. The feedback gathered feeds into the C3S systematic user requirement database to benefit the whole C3S activity. The evaluation reports are first made available to external (experts who are not involved in the evaluation) users for collecting feedback. Then, when/if the feedback is received (no feedback is received for some reports), the user's comments are addressed in the final evaluation reports. This external review process strengthens the quality of the evaluation report from the users' perspective.

Lessons learned and future perspectives

The independent assessment procedure involves the following elements: expert evaluators, data providers, users, and the assessment framework. The following sections provide lessons learned and future perspectives for each element.

Expert evaluators. The assessment of ECVs from different domains in a homogeneous and consistent manner requires a team of scientific experts who, based on their fields of research, have extensive experience with and a solid scientific knowledge of each of the ECVs. Such a structure and approach offers considerable advantages, since each expert evaluator has the scientific expertise to 1) design appropriate approaches in the assessment of the scientific quality of the dataset; 2) identify possible spurious variability and provide feedback to data providers; 3) detect climate signals in the ECV based on their knowledge, and understand the reliability and suitability of the dataset, and check ECVs present in a physically consistent manner; and 4) look for the most relevant literature and documentation to support evaluations of specific ECVs.

Although, the IQA-ECVs framework is designed to apply to all ECVs in all the domains, for particular ECVs (e.g., glaciers) evaluators perform additional/diverse diagnostics to assess the quality based on their scientific expertise for these particular ECVs.

Such detailed assessments incorporating scientific experts could not be achieved in a purely automatic approach. Within this framework, the scientific expertise for each ECV is essential to critically interpret the outcome of scientific use cases conducted as part of the assessment, and to identify potential issues presented in the dataset with respect to expected climate signals.

Therefore, the expertise of scientific evaluators is essential and required to provide in-depth scientific assessments in the future assessment activities and help to further tailor the assessment framework to meet users' needs.

Data providers. Despite our evaluations being independent of the data providers, the workflow of our assessment has involved an iterative interaction with data providers: their feedback is only considered if reasonable in the final stage of evaluation reports publication. This process guarantees the consistency of the evaluation and provides the final quality monitoring of the assessment, while also minimizing the potential undue influence from the product originators. The interaction also includes feedback from evaluators to data providers, especially when issues are identified in ECV datasets which can strengthen subsequent data product releases.

The engagement of data providers in the assessment processes proved to be an essential iterative process in providing a consolidated assessment of ECVs, which is critical for the future assessment. There already exist notable examples of close interaction between the scientific evaluators and data providers (e.g., Yang et al. 2021) that can inspire future evolutions.

Users. To further improve the independent assessment, the engagement of potential users is desirable. The interaction with as wide a range of users as possible is important in helping shape the assessments in both structure and level of detail. First, since our scientific experts are research scientists engaged in producing scientific publications, there is a natural tendency for the language and style in the reports to be oriented toward their peer community. Second, the target users who are going to read the reports are likely to have a diverse range of needs. Without eliciting user feedback, the language in the evaluation reports can become too technical and hence off-putting for some users; indeed, this is the most frequently raised point in the feedback we received. Expanded engagement with users therefore would help scientific evaluators better understand the extent to which their language should be simplified for less specialist readers.

The engagement of users helped us to evaluate the fitness for purpose of each dataset. For example, one component of the independent assessment is the application performance matrix, which assesses whether a particular dataset is suitable for a specific application. Currently, the applications we have adopted tend to be from the scientific literature or website-based information that demonstrates that the relevant variable has been used for a specific application. However, up-to-date requirements of applications from users and linking closely these applications and underpinning data are required to understand how the datasets are currently used and whether or not they are truly fit for purpose.

One interesting feedback from users to date on our evaluation reports is that assessments of the datasets at regional scales are required (e.g., Europe). The current assessment framework implemented in IQA-ECVs mainly focuses on the global scale, with a few case studies at regional scales. However, a more refined assessment of ECVs at regional and national scales may help users to understand the suitability of the ECV products for their intended applications. This would improve the assessment of the potential utility of the dataset for monitoring regional climates for appropriate climate mitigation and adaptation responses, and hence their potential use for climate risk management.

Overall, promoting greater interaction between evaluators and users to support explicitly codesign and codevelopment processes in the future is desirable to improve the usability and applicability of evaluation reports to meet users' needs. Interactions with users by means of call desks tickets or user forums, and access to statistical analysis of the C3S central user database could help gather valuable user requirements in the future.

Framework improvement and adaptation. The prototype independent assessment framework presented here has not yet covered the evaluation of uncertainty estimates, which is essential to enhancing trust in the climate data for climate services. Consequently, in the next stage of the service, greater emphasis should be placed on uncertainty assessment to consider adequacy and usability to inform robust decision-making. The assessment may consider the comprehensiveness and applicability of the provided uncertainty information at a range of spatiotemporal scales to aid users in making informed decisions under uncertainty (Otto et al. 2016).

Currently, the independent assessment framework applies to reanalyses, satellite, and (gridded) in situ observation data. To merge assessment frameworks for different types of datasets (e.g., seasonal forecasts and climate projections) proved to be difficult due to specific characteristics of each type of dataset. Further development is needed in order to permit application to these data types. Expert evaluators have to design dedicated analysis and scientific use cases for different types of datasets. However, the data checking, data maturity, and application performance protocols developed as part of IQA-ECVs could in principle be transferred to assess the seasonal forecasts datasets or climate projections following appropriate refinements.

Particularly for climate projections, one scientific use case in the current framework is dedicated to assessing the suitability of the ECV dataset for ESM evaluation. In a next step, the direct evaluation of climate models using reanalysis, satellite observations within the CDS could be added. This would serve not only to improve understanding of the different reanalysis, in situ, and satellite datasets, but also to evaluate the climate projections available in the CDS directly.

In general, the assessment approach developed in IQA-ECVs could be applied to similar data assessment activities far beyond Copernicus services. That is to say that the techniques developed and deployed could in principle add value to any similar climate data portal.

Summary

In conclusion, the assessment carried out by dedicated scientific experts for specific ECVs provides users with a scientific and in-depth evaluation which could not be performed by a more automated method. These evaluation reports enhance the trust of users in relation to climate datasets provided by the Copernicus Climate Change Service, and guide users as to the usability and suitability of the datasets. However, the limited resources of scientific experts require user feedback to prioritize assessment of the most desired datasets and regions in order to make the operational service sustainable. Furthermore, in such an expert driven approach, scientific experts (now with service-specific experiences) can further tailor the assessment framework based on lessons learned and future perspectives mentioned above to meet users' needs; continue to guide the analysis and assessment of datasets served on the CDS; and recommend the reliability and suitability of datasets to users in an independent, transparent, and traceable manner.

Acknowledgments. This work is funded by European Copernicus Climate Change Service (C3S) implemented by European Centre for Medium-Range Weather Forecasts (ECMWF) under the service contract Independent Assessment on ECVs led by the National Research Council of Italy (CNR) with the service contract number as ECMWF/Copernicus/2017/C3S_511_CNR.

Data availability statement. All the dataset and assessments reports mentioned in this article are available online at https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset and particularly assessments reports can be found in the tab "quality assessment" of each dataset catalogue entry.

Appendix: Defensible traces for maturity evaluation of ERA5 winds presented in Fig. 3

Metadata.

Standards (score: 6/6). The ERA5 data present on the CDS can be downloaded in both GRIB and netCDF formats. The netCDF files follow the standards defined within the metadata. These standards are systematically applied and international standards are met. Furthermore, the compliance is coherent for the dataset.

Collection level (score: 5/6). The standardized attributes on the collection level of the dataset are sufficient to understand the data's origins without further documents, including standardized information on how to obtain raw data and its preprocessing procedures.

Note: The collection level is including the information available from confluence wiki.

User documentation.

Formal description of scientific methodology (score 5/6). The scientific description is comprehensive and publicly available in the form of a scientific report/ATBD. The description is kept up to date with the updated dataset. There is also a peer-reviewed methodological journal paper and an evaluation activity/result of ERA5 is published (Hersbach et al. 2020).

Note: In this case the confluence wiki is considered as the scientific report/ATBD.

Formal validation report (score: 4/6). The ERA5 dataset is still missing comprehensive details on the validation activities that have been performed to assess the fidelity of the data record.

Note: An intercomparison with observations (radiosondes, dropsondes, PILOT) and with ERA-Interim is available for upper-air zonal wind (Hersbach et al. 2020). However, no evaluation of uncertainty is reported.

PUG (score: 6/6). There is a regularly updated comprehensive formal product user guide (PUG) for the dataset publicly available.

Note: In this case the confluence wiki is regarded as the PUG.

Uncertainty characterization.

Standards (score: 3/6). Uncertainty information follows standard nomenclature.

Note: In this case the ensemble members are regarded as uncertainty measures.

Validation (score: 4/6). Information available is not fully clear. More details and inter(comparison) results against other CDRs would be very useful. Hersbach et al. (2020) provided a comparison of upper-air zonal wind with observations (radiosondes, dropsondes, PILOT) and with ERA-Interim.

Uncertainty quantification (score: 3/6). A comprehensive uncertainty quantification of systematic and random effects is available.

Note: In this case the ensemble members are regarded as uncertainty measures.

Automated quality monitoring (score: 2/6). There is no automated quality monitoring documented for the dataset. Albeit an operational production monitoring is existing for ERA5 (Hersbach et al. 2018), no documents are available about the specific definition and the implementation of an automated quality monitoring system.

Note: Automated quality monitoring exists, but it is not made publicly.

Public access, feedback, and update.

Access and archive (score: 5/6). The dataset is publicly available. The different versions of data including documentation and source code are archived by the data provider.

Version control (score: 5/6). There is full information on version control of documentation, data and/or metadata available for the dataset. The documented version control information is fully traceable from the files.

Note: In this case the version control is referring to the confluence wiki.

User feedback (score: 5/6). There is a public reach-out/feedback form/contact point for collecting feedback for the dataset. There are regular events, groups, two-way feedback mechanisms, etc., organized by the data provider. The established feedback fed back into data production is documented, including third party international data quality assessment results.

Updates to record (score: 5/6). There are regular operational updates available for the dataset, depending on the availability of input data and including improved methodology.

Usage.

Research (score: 5/6). The ERA5 dataset is designed to be used also for research purposes. Peer-reviewed publications that describe the usage of the product in a research application are available; peer-reviewed papers exist about the usage of ERA5 as reference for ESM evaluation and for global climate analysis (Blunden and Arndt 2019; Swart et al. 2019).

Decision support system (score: 3/6). The product can be used for certain applications, i.e., impact assessment studies and preliminary studies that show the benefit of this dataset. At the current stage the relevant publications (e.g., Kalverla et al. 2019) are limited.

References

Blunden, J., and D. S. Arndt, Eds., 2019: State of the Climate in 2018. *Bull. Amer. Meteor. Soc.*, **100**, S1–S306, https://doi.org/10.1175/2019BAMSStateoftheClimate.1.

Buchwitz, M., and Coauthors, 2018: Copernicus Climate Change Service (C3S) global satellite observations of atmospheric carbon dioxide and methane. *Adv. Astronaut. Sci. Technol.*, **1**, 57–60, https://doi.org/10.1007/s42423-018-0004-6.

Coll, J., P. Domonkos, J. Guijarro, M. Curley, E. Rustemeier, E. Aguilar, S. Walsh, and J. Sweeney, 2020a: Application of homogenization methods for Ireland's monthly precipitation records: Comparison of break detection results. *Int. J. Climatol.*, **40**, 6169–6188, https://doi.org/10.1002/joc.6575.

——, B. Muller, B. Crezze, C. Cagnazzo, and C. Yang, 2020b: Guidance document on applying the maturity matrix as part of the evaluation and quality control. ECMWF, https://confluence.ecmwf.int/display/CKB/Guidance+document+on+applying+the+Maturity+Matrix+as+part+of+the+Evaluation+and+Quality+Control.

Crezee, B., M. Hirschi, P. Thorne, J. Coll, B. Müller, C. Yang, A. Pisano, and B. Hassler, 2019: Data catalogue for extreme events and extreme events metrics. Copernicus Climate Change Service Doc. D511.1.5, 52 pp., https://drive.google.com/file/d/13KlOW-1xq82LuMKLvbS2MmOtCDMMpYWO/view?usp=sharing.

Duren, R., and C. Miller, 2012: Measuring the carbon emissions of megacities. *Nat. Climate Change*, **2**, 560–562, https://doi.org/10.1038/nclimate1629.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

——, and Coauthors, 2020: Earth System Model Evaluation Tool (ESMValTool) v2.0—An extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP. *Geosci. Model Dev.*, **13**, 3383–3438, https://doi.org/10.5194/gmd-13-3383-2020.

——, and Coauthors, 2021: Human influence on the climate system. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 423–551, www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter03.pdf.

Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.

GCOS, 2016: The Global Observing System for Climate: Implementation needs. WMO Doc. GCOS-200, 341 pp., https://library.wmo.int/doc_num.php?explnum_id=3417.

Hassler, B., and A. Lauer, 2021: Comparison of reanalysis and observational precipitation datasets including ERA5 and WFDE5. *Atmosphere*, **12**, 1462, https://doi.org/10.3390/atmos12111462.

Hersbach, H., and Coauthors, 2018: Operational global reanalysis: Progress, future directions and synergies with NWP. ERA Rep. Series 27, 65 pp., www.ecmwf.int/sites/default/files/elibrary/2018/18765-operational-global-reanalysis-progress-future-directions-and-synergies-nwp.pdf.

Hersbach, H, and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hobday, A. J., and Coauthors, 2016: A hierarchical approach to defining marine heatwaves. *Prog. Oceanogr.*, **141**, 227–238, https://doi.org/10.1016/j.pocean.2015.12.014.

——, and Coauthors, 2018: Categorizing and naming marine heatwaves. *Oceanography*, **31** (2), 162–173, https://doi.org/10.5670/oceanog.2018.205.

IPCC, 2022: *Climate Change 2022: Impacts, Adaptation, and Vulnerability.* H.-O. Pörtner et al., Eds., Cambridge University Press, 3676 pp.

Kalverla, C., B. Duncan, G. Steeneveld, and A. M. Holtslag, 2019: Low-level jets over the North Sea based on ERA5 and observations: Together they do better. *Wind Energy Sci.*, **4**, 193–209, https://doi.org/10.5194/wes-4-193-2019.

Lauer, A., and Coauthors, 2017: Benchmarking CMIP5 models with a subset of ESA CCI phase 2 data using the ESMValTool. *Remote Sens. Environ.*, **203**, 9–39, https://doi.org/10.1016/j.rse.2017.01.007.

——, and Coauthors, 2020: Earth System Model Evaluation Tool (ESMValTool) v2.0—Diagnostics for emergent constraints and future projections from Earth system models in CMIP. *Geosci. Model Dev.*, **13**, 4205–4228, https://doi.org/10.5194/gmd-13-4205-2020.

Madonna, F., and Coauthors, 2022: The new Radiosounding Harmonization (RHARM) data set of homogenized radiosounding temperature, humidity, and wind profiles with uncertainties. *J. Geophys. Res. Atmos.*, **127**, e2021JD035220, https://doi.org/10.1029/2021JD035220.

Merchant, C. J., and Coauthors, 2019: Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data*, **6**, 223, https://doi.org/10.1038/s41597-019-0236-x.

Müller, B., F. Madonna, C. Yang, C. Caganzzo, 2020, Guidance document on applying the application table or application matrix (AT/AM) to C3S_511 ECV dataset assessments. Copernicus Climate Change Service Doc., 14 pp., https://drive.google.com/file/d/12TbHpz9nyAZic87_1aDdcoy4RGw_EgXh/view?usp=sharing.

Otto, J., and Coauthors, 2016: Uncertainty: Lessons learned for climate services. *Bull. Amer. Meteor. Soc.*, **97**, ES265–ES269, https://doi.org/10.1175/BAMS-D-16-0173.1.

Peng, G., and Coauthors, 2019: Practical application of a Data Stewardship Maturity Matrix for the NOAA *OneStop* Project. *Data Sci. J.*, **18**, 41, https://doi.org/10.5334/dsj-2019-041.

Popp, T., and Coauthors 2020: Consistency of satellite climate data records for Earth system monitoring. *Bull. Amer. Meteor. Soc.*, **101**, E1948–E1971, https://doi.org/10.1175/BAMS-D-19-0127.1.

Righi, M., and Coauthors, 2020: Earth System Model Evaluation Tool (ESMValTool) v2.0—Technical overview. *Geosci. Model Dev.*, **13**, 1179–1199, https://doi.org/10.5194/gmd-13-1179-2020.

Su, C.-H., D. Ryu, W. Dorigo, S. Zwieback, A. Gruber, C. Albergel, R. H. Reichle, and W. Wagner, 2016: Homogeneity of a global multisatellite soil moisture climate data record. *Geophys. Res. Lett.*, **43**, 11 245–11 252, https://doi.org/10.1002/2016GL070458.

Su, Z., and Coauthors, 2018: An overview of European efforts in generating climate data records. *Bull. Amer. Meteor. Soc.*, **99**, 349–359, https://doi.org/10.1175/BAMS-D-16-0074.1.

Swart, C., and Coauthors, 2019: The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci. Model Dev.*, **12**, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019.

Van Den Besselaar, E. J. M., and Coauthors, 2015: International Climate Assessment & Dataset: Climate services across borders. *Bull. Amer. Meteor. Soc.*, **96**, 16–21, https://doi.org/10.1175/BAMS-D-13-00249.1.

van der Schrier, G., E. J. M. van den Besselaar, A. M. G. Klein Tank, and G. Verver, 2013: Monitoring European average temperature based on the E-OBS gridded dataset. *J. Geophys. Res. Atmos.*, **118**, 5120–5135, https://doi.org/10.1002/jgrd.50444.

——, and Coauthors, Eds., 2019: The Climate Data Guide: E-OBS: High-resolution gridded mean/max/min temperature, precipitation and sea level pressure for Europe & Northern Africa. NCAR, accessed 20 May 2020, https://climatedataguide.ucar.edu/climate-data/e-obs-high-resolution-gridded-meanmaxmin-temperature-precipitation-and-sea-level.

Weatherhead, E. C., and Coauthors, 2017: How long do satellites need to overlap? Evaluation of climate data stability from overlapping satellite records. *Atmos. Chem. Phys.*, **17**, 15 069–15 093, https://doi.org/10.5194/acp-17-15069-2017.

Weigel, K., and Coauthors, 2021: Earth System Model Evaluation Tool (ESMValTool) v2.0—Diagnostics for extreme events, regional and impact evaluation, and analysis of Earth system models in CMIP. *Geosci. Model Dev.*, **14**, 3159–3184, https://doi.org/10.5194/gmd-14-3159-2021.

WMO, 2018: Guidelines on the definition and monitoring of extreme weather and climate events. Accessed 23 February 2022, https://wmoomm.sharepoint.com/:b:/s/wmocpdb/EYGx0kanmIhLi8bt4pOoQZoBGm9p-lgySZ44hmVCDdywJQ.

Yang, C., and Coauthors, 2021: Sea surface temperature intercomparison in the framework of the Copernicus Climate Change Service (C3S). *J. Climate*, **34**, 5257–5283, https://doi.org/10.1175/JCLI-D-20-0793.1.