RESEARCH ARTICLE

Key Points:

Correspondence to:
P. Laloyaux,
patrick.laloyaux@ecmwf.int

CERA-20C: A Coupled Reanalysis of the Twentieth Century

Patrick Laloyaux[1] , Eric de Boisseson[1] , Magdalena Balmaseda[1] , Jean-Raymond Bidlot[1], Stefan Broennimann[2], Roberto Buizza[1] , Per Dalhgren[1], Dick Dee[1], Leopold Haimberger[3] , Hans Hersbach[1], Yuki Kosaka[4], Matthew Martin[5], Paul Poli[6] , Nick Rayner[5], Elke Rustemeier[7], and Dinand Schepers[1]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK, [2]University of Bern, Bern, Switzerland, [3]University of Vienna, Wien, Austria, [4]Japan Meteorological Agency, Tokyo, Japan, [5]Met Office, Exeter, UK, [6]Meteo France, Paris, France, [7]Deutscher Wetterdienst, Offenbach, Germany

Abstract CERA-20C is a coupled reanalysis of the twentieth century which aims to reconstruct the past weather and climate of the Earth system including the atmosphere, ocean, land, ocean waves, and sea ice. This reanalysis is based on the CERA coupled atmosphere-ocean assimilation system developed at ECMWF. CERA-20C provides a 10 member ensemble of reanalyses to account for errors in the observational record as well as model error. It benefited from the prior experience of the retrospective atmospheric analysis ERA-20C. The dynamical model and the data assimilation systems initially developed for NWP had been modified to take into account the evolution of the radiative forcing and the observing system. To limit the impact of changes in the observing system throughout the century, only conventional surface observations have been used in the atmosphere. CERA-20C improves the specification of the background and the observation errors, two key elements to ensure a consistent weighting of the uncertainties across geophysical variables, space, and time. The quality of CERA-20C has been evaluated against other centennial reanalyses and independent observations. Although CERA-20C inherits some limitations of ERA-20C to represent correctly the tropical cyclones in the first part of the century, it shows significant improvements in the troposphere, compared to ERA-20C and 20CRv2c (the twentieth century reanalysis produced by NOAA/CIRES). A preliminary study of the climate variability in CERA-20C has been carried out. CERA-20C improves on the representation of atmosphere-ocean heat fluxes and mean sea level pressure compared to previous uncoupled ocean and atmospheric historical reanalyses performed at ECMWF.

1. Introduction

Since its creation in 1975, ECMWF has been a key player in the production of reanalyses, which provide a numerical description of the recent climate by combining models with observations. The initial focus was on producing atmospheric reanalyses covering the satellite observing period, from 1979 onward. The first of these reanalyses, FGGE, was produced in the 1980s, followed by ERA-15, ERA-40 (Uppala et al., 2005), and ERA-Interim (Dee et al., 2011). The fifth reanalysis in this series, ERA5, is now in production after years of research since ERA-Interim, and intense technical preparations (Hersbach & Dee, 2016). ECMWF is also producing ocean reanalyses with the replacement in 2016 of ORAS4 (Ocean Reanalysis System 4) by ORAS5 to incorporate the latest improvements in ocean models, data assimilation methods, and forcing fluxes (Balmaseda et al., 2013).

The various reanalysis products have proven to be important resources for weather and climate-related research as well as societal applications at large (Gregow et al., 2016). At ECMWF, reanalyses also support continuous improvement of numerical weather prediction. One example is the construction of a reforecast ensemble based on the current model and initialized from reanalyses. This M-climate set is compared to the operational forecast to compute the Extreme Forecast Index (EFI) which provides specialized forecast guidance for anomalous, extreme, or severe weather events (e.g., heavy precipitation, strong winds, heavy snowfall, extreme temperatures, and unusually high ocean waves). Furthermore, reforecasts started from reanalyses allow a better and more complete estimate of forecast reliability and accuracy. Reanalyses make it possible to study the interannual variability of forecast skill and to test new model versions on past severe weather cases. At ECMWF, this is true for both the medium-range/monthly ensemble (ENS) and for the

seasonal ensemble (SEAS5). ERA-Interim and ORAS5 are the current operational atmospheric and ocean reanalyses at ECMWF, continuing in near real time. They are both created using versions of the data assimilation system and model that are kept unchanged during the reanalysis production, while ingesting a variety of satellite and in situ observations to provide the best state estimate over the data-rich time period since 1979. Several other weather and research centers apply the same methodology to produce recent global atmospheric reanalyses based on their model and data assimilation systems. NASA is now continuing production in near-real-time of the second version of the Modern-Era Retrospective analysis for Research and Applications (MERRA2; Gelaro et al., 2017). The first coupled reanalysis of the global atmosphere, ocean, land surface, and cryosphere was created by NCEP with the Climate Forecast System Reanalysis spanning the period 1979 onward (CFSR; Saha et al., 2010). JMA carried out a reanalysis project known as the Japanese 55-year Reanalysis (JRA-55) to provide time series of a comprehensive set of atmospheric variables which cover the period 1958 to present (Kobayashi et al., 2015). Similar activities are conducted by the operational oceanographic community to continuously produce ocean reanalyses and to coordinate intercomparison projects, such as the Ocean Reanalyses Intercomparison Project ORA-IP (Balmaseda et al., 2015).

Extending these reanalyses further back in time is a tremendous scientific challenge as the observing system is very sparse before the availability of satellite data from the 1970s onward, and especially before the arrival of radiosonde measurements in the 1930s (Stickler et al., 2014). Information about the nature and quality of early instrumentation is often incomplete (Kennedy, 2014). Furthermore, locating and gaining access to early weather observations requires dedicated efforts in data rescue and digitization, especially in parts of the world that are most affected by climate change and variability (Allan et al., 2011). The skill of reanalyses is limited in the Southern Ocean by a lack of historical observations. The ERA-CLIM and ERA-CLIM2 projects have imaged information on surface meteorological variables for the Southern Ocean for the early twentieth century. These documents include dedicated meteorological logbooks as well as common ships' logbooks, meteorological forms, whale catch books, day reports, ice reports, ice charts, and other relevant items. Types and frequency of observations will vary from one document to another but typically there is barometric pressure and air, temperature, wind direction and force, weather, cloud cover, SST, sea state and swell, sometimes salinity and biological observations, sea ice, and icebergs (Bronnimann et al., 2018). To tackle the unavoidable issue of the ever-changing observational network, a whitelisting approach to data selection for reanalyses covering the whole twentieth century has been developed. Instead of assimilating the full observing system at any time, only observation types that are reliably available around the globe for the entire century are used. Compo et al. (2006) demonstrated the feasibility of producing a useful global atmospheric reanalysis spanning a century or more with modern data assimilation methods and only surface pressure observations. However, the surface pressure network changes across the century in terms of sampling and data quality which means that inhomogeneities in the reduced observing system cannot be completed avoided. The Twentieth Century Reanalysis (20CRv2c) data set extending back to 1851 was produced by NOAA/CIRES using the NCEP atmosphere/land model to generate first-guess fields with interpolated monthly sea-surface temperature and sea-ice concentration fields from the Hadley Centre Sea Ice and SST data set (HadISST) as prescribed boundary conditions (Compo et al., 2011). The 20CRv2c reanalysis faithfully reconstructs the large-scale tropospheric circulation having a correlation between 300 hPa geopotential height subdaily anomalies and radiosondes larger than 0.8 for 1958–1978 (Compo et al., 2011).

The quality of such historical reanalyses depends on the density of observations and cannot outperform a comparable system that uses all sources of observations, including satellite and upper-air measurements. In fact, a research reanalysis for the 1939–1967 period which assimilates radiosonde data (ERA-preSAT; Hersbach et al., 2017) correlates better with independent observations than 20CRV2c. The whitelisting approach to data selection, where observations are used only if they are known to be suitable for climate applications, reduces the artificial variability and spurious trends generated by the introduction of new instruments. This is what makes it possible to generate valuable climate reconstructions that cover a period of 100 years or more in order to investigate low-frequency climate variability. The overall aim is to improve our ability to produce consistent reanalyses of the climate system, reaching back in time as far as possible given the available instrumental record. Historical reanalyses try to recreate the weather, while remaining faithful to overall climate trends. This allows researchers to study statistics of weather events on climate timescales. In this context, following the example of NCEP (Compo et al., 2011), ECMWF produced the atmospheric reanalysis ERA-20C during the ERA-CLIM project, which covers the period January 1900 to December 2010 (Poli et al.,

Table 1
List of Selected ECMWF Reanalysis Data Sets Showing the Period Covered, the Observing System Used and the Different Earth System Components Included in the Climate Reconstruction

Data set	Period	Type	Ensemble	Atmosphere	Land	Waves	Ocean	Sea ice
ERA-Interim	1979-present	All obs.		✓	✓	✓		
ORAS5	1975-present	All obs.	✓				✓	✓
ERA-20C	1900–2010	Selected obs.		✓	✓	✓		
ORA-20C	1900–2010	Selected obs.	✓				✓	✓
CERA-20C	1901–2010	Selected obs.	✓	✓	✓	✓	✓	✓

2016). ERA-20C assimilates only conventional observations of surface pressure and marine wind, obtained from well-established climate data collections. Model forcings are specified from the Coupled Model Inter-comparison Project Phase 5 (CMIP5; Taylor et al., 2012) recommendations to obtain an appropriate climate reconstruction. The atmospheric lower boundary conditions are prescribed using the HadISST2 monthly analysis product for sea-surface temperature and sea ice. ERA-20C delivered 3 hourly products describing the spatial and temporal evolution of the atmosphere, land surface, and waves. Recognizing the importance of properly accounting for changing background errors, ERA-20C was first devised as a 10 member ensemble of 4D-Var assimilation system to create spatiotemporally varying background errors (Poli et al., 2013). On this basis, a deterministic 4D-Var system had been run using the flow-dependent background errors and fixing several issues, and then publicly released.

The reanalysis capabilities developed in the ERA-CLIM project have been extended to the ocean and sea-ice components in the ERA-CLIM2 project (Buizza et al., 2018). ORA-20C is an uncoupled ocean reanalysis, and reconstructs the ocean and sea-ice state over the twentieth century (de Boisseson et al., 2017). Temperature and salinity profiles are assimilated into the ocean model, which is also constrained by fluxes from ERA-20C and a sea-surface temperature relaxation towards the HadISST2 product. Observations visibly impact the fields from ORA-20C during the full record, but are only able to constrain large ocean climate signals, such as ocean heat content, during the second half of the century. The lack of constraint obviously poses a challenge for coupled reanalyses. On the other hand, the coupled reanalyses may be able provide a more realistic ensemble of realizations, where ocean and atmosphere observations can be exploited in a more dynamically consistent manner. This is the motivation for CERA-20C, a reanalysis of the twentieth century using the coupled data assimilation CERA. The CERA system simultaneously ingests atmospheric and ocean observations in the coupled Earth system model used for ECMWF's ensemble forecasts (Laloyaux et al., 2016a). This approach accounts for interactions between the atmosphere and the ocean during the assimilation process and has the potential to generate a more balanced and consistent Earth system climate reconstruction. CERA-20C is the first 10 member ensemble of coupled reanalyses of the twentieth century. The main features of some of the reanalyses produced at ECMWF are summarized in Table 1.

This article is organized as follows. In section 2, the dynamical coupled model and the observation data sets used to produce CERA-20C are described. The modifications made to the data assimilation system to compute a retrospective coupled analysis over the twentieth century are presented in section 3. The quality of the atmospheric reconstruction of CERA-20C against other centennial reanalyses and independent observations is assessed in section 4. Different weather and climate signals are discussed in section 5 putting an emphasis on the interactions between the different components of the Earth system. Section 6 presents some preliminary results on the uncertainties that can be estimated from the ensemble reconstruction. Further possible directions of research and development in coupled data assimilation are finally described in section 7.

2. Model, Forcings, and Observations

CERA-20C relies on version CY41R2 of ECMWF's Integrated Forecast System (IFS) which is used for medium-range, monthly, and seasonal weather forecasting applications (ECMWF, 2016). Its capability includes not only an atmospheric model but also coupling to several Earth system components: land (HTESSEL; Balsamo et al., 2009), ocean wave (WAM; Komen et al., 2004), ocean (NEMO; Madec, 2008), and sea ice (LIM2; Bouillon et al., 2009). All components are integrated into a single executable with a common time step loop,

sequentially calling each component and regridding fields as needed on the different model grids (Mogensen et al., 2012a). The atmospheric model was modified as reported by (Hersbach et al., 2015a) to use CMIP5 atmospheric forcing data to incorporate a better long-term evolution of climate trends over the twentieth century. These terms include solar forcing, greenhouse gases, ozone, and aerosols. At the air-sea interface, the sea surface temperature (SST) is relaxed toward the HadISST2 monthly ensemble product (Titchner & Rayner, 2014) to limit the model drift while enabling the simulation of coupled processes. The relaxation coefficient is set to $200 \, \mathrm{Wm^{-2}C^{-1}}$ which is equivalent to about a 2–3 day time-scale over a depth of 10 m. The sea-ice concentration is not relaxed toward HadISST2 but its extension is constrained by the SST. The HadISST2 data set is produced by the Met Office Hadley Centre. Observations used in HadISST2 include in situ SST measurements, AVHRR (Advanced Very High Resolution Radiometer) SST retrievals, and retrievals from AATSR (Advanced Along-Track Scanning Radiometer) reprocessing. Both the radiative forcing and the HadISST2 incorporate a proper long-term evolution of climate trends in the twentieth century, and the occurrence of major events, such as El Niño Southern Oscillations and volcanic eruptions. The impact of radiative forcing has been carefully studied in Hersbach et al. (2015a).

Century long coupled or uncoupled integrations with prescribed radiative forcing are the backbone of the CMIP coordinated experimentation in support of IPCC reports. Century long coupled or uncoupled reanalyses go a step further. The main purpose of assimilating observations is to add information about actual weather events without deteriorating the model representation of low-frequency variability and change. This is obviously a tremendous challenge, but an essential one to tackle if we want to progress on the understanding of climate variability, change, and impacts. The ability to produce consistent reanalyses of the climate system, reaching back in time as far as possible, relies then on the availability of past historical observations. Within the ERA-CLIM project and its follow-up ERA-CLIM2, large amounts of historical observations were rescued, compiled, processed, and made available to the reanalysis efforts. The newly rescued data includes millions of historical surface and upper-air observations (Stickler et al., 2014) which are embedded within the international Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative to ensure exchange of information across various projects (Allan et al., 2011).

The CERA-20C reanalysis does not assimilate the full observing system at any time, but only observation types with a good enough spatial and temporal coverage over the entire century are used. As in ERA-20C, surface pressure and marine wind observations are assimilated in the atmospheric model, from the ISPDv3.2.6 (Cram et al., 2015) and ICOADSv2.5.1 (Woodruff et al., 2011) data sets. The International Surface Pressure Databank (ISPD) is the world's largest collection of pressure observations (1768–2012) and consists of observations from land stations, buoys, ships, coastal stations, and tropical cyclone best track pressure reports. The International Comprehensive Ocean-Atmosphere Data Set (ICOADS) provides surface marine data including sea level pressure and wind speed over the past three centuries (1662–2014) from buoys, ships, and coastal stations. These weather observations are available in large numbers throughout the twentieth century. Initially concentrated in the northern hemisphere, the global coverage increases with time. The evolution of the number of observations assimilated per platform for surface pressure and marine wind is plotted on the top and middle plots of Figure 1. "Best Track" observations from the International Best Track Archive for Climate Stewardship (Knapp et al., 2010) are plotted as bogus observations as they are not always actual measurements but interpolated or approximated from wind data. For the ocean component, observed subsurface temperature and salinity profiles from the EN4.0.2 data set (A. Good et al., 2013) with bias correction from Gouretski and Reseghetti (2010) are
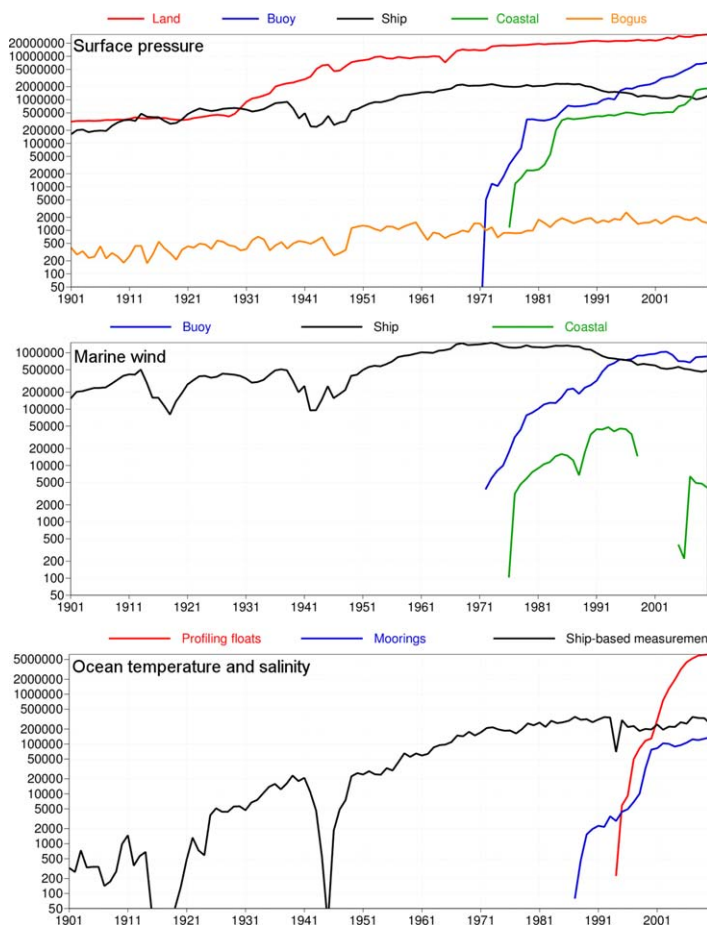
Figure 1. Yearly count of assimilated observations per platform in CERA-20C: (top) surface pressure, (middle) marine wind, and (bottom) ocean temperature and salinity.

assimilated. EN4 is a collection of subsurface profiles obtained across the global oceans from profiling floats, moorings, and ships. The evolution of the number of temperature and salinity observations assimilated in CERA-20C is plotted on the bottom plot of Figure 1. Although the data set is available over the period 1,900 to present, the first part of the century is barely observed. The observation coverage improves considerably with the introduction in the last decade of the Argo global array of about 3,800 free-drifting profiling floats that measures the temperature and salinity of the upper 2,000 m of the ocean. Although no oceanic satellite data is assimilated in CERA-20C, the changes in the in situ component of the ocean observing system has a substantial effect on the temporal variability of the ocean estimate (de Boisséson et al., 2017). No data assimilation is performed for the land, ocean wave, and sea-ice components in CERA-20C. However, the use of the coupled model ensures a dynamically consistent Earth system estimate at any time.

Climate reanalysis can be used to quality control the observations from the databanks as the data assimilation system compares in a systematic way the observations with a physically consistent model. ECMWF climate reanalyses include the production of Observation Feedback Archives (OFA) designed from the outset for user investigations (Hersbach et al., 2015b). This archive is organized by observation report type and source. It contains the observations assimilated (surface pressure and marine wind) and other observations that can be handled by the IFS model. This includes visual observations (present and past weather, visibility, cloud-base height, amounts, and types) and observations that can be exploited quantitatively (surface pressure tendency, air temperature and humidity, and seawater temperature). Upper-air observations have not been processed and they are therefore not available in the OFA. The observation feedback includes in particular the background departure (observation minus background) and the analysis departure (observation minus analysis) for many observations. For example, the ERA-20C feedback information was analyzed in detail prior to CERA-20C, to detect stations reporting suspicious measurements. The CERA-20C feedback information was analyzed similarly. This includes stations or ships with systematically too large background departure, ships with unrealistic speed due to position errors and ships reporting observations over land, or stations reporting constant values for a long-time (stuck sensor). The lists of stations were sent to the data providers to improve the quality control of their data sets (to be used in future reanalyses). This feedback process propagates the benefits of a better model and data assimilation system to better reanalyses, and then to better quality-controlled observation data sets. Improvements in the global observing system, advances in data assimilation methodology, and development of better forecast models yield a strong feedback loop. Ongoing activity in the field of reanalysis ensures that improvements in these areas will also continue.

3. CERA-20C System

To produce the CERA-20C reanalysis, a new assimilation system (CERA) has been developed to simultaneously ingest atmospheric and ocean observations in the coupled Earth system model. The CERA system is based on a variational method with a common 24 h assimilation window shared by the atmospheric and ocean components (Laloyaux et al., 2016a). By extending the 4D-Var analysis window beyond the 6 or 12 h norm of atmospheric Numerical Weather Prediction (NWP), more past and future observations in the atmosphere and the ocean can be included following the idea of a retrospective analysis. The outer loop of the variational algorithm integrates the coupled nonlinear forecast model, producing a four-dimensional state estimate and observation misfits. The ocean and atmospheric increments are computed separately by two different inner loops that minimize a linearized version of the variational formulation for the ocean and the atmosphere. This outer-loop coupling means that air-sea interactions are taken into account when observation misfits are computed and when the increments are applied to the initial condition. By computing two outer iterations in CERA-20C, ocean observations can have a direct impact on the atmospheric analysis and, conversely, atmospheric observations can have an immediate impact on the analyzed state of the ocean within the same assimilation cycle.

The requirements of NWP and reanalysis are different; several key elements of data assimilation algorithms developed for NWP are suboptimal for retrospective analysis. NWP analysis systems are developed for the present global observing network, with many observations that constrain the system in subtle ways, and their absence can cause issues in earlier time periods. In the following, we review the major elements of the

ECMWF operational atmospheric assimilation system and we describe the modifications made to produce a retrospective analysis over the twentieth century.

3.1. Atmospheric Background Error

In any data assimilation scheme, the specification of background and observation errors is crucial to define the relative weights that blend together the information from the measurements and from a prior information of the forecast model (so-called background). In the context of extended climate reanalysis, information about spatially and temporally varying background errors must be supplied. This information should reflect the dependence of errors in atmospheric dynamics and more importantly account for the substantial changes in observation data coverage that take place during the twentieth century.

Poli et al. (2013) demonstrated that it was feasible to use a 10 member Ensemble of 24 h 4D-Var Data Assimilation (EDA; Bonavita et al., 2016) to quantify flow-dependent background errors over a century. The estimates proved to be robust, with a good agreement between overlapping production streams. The EDA consists of an ensemble of data assimilation systems that differ by perturbing observations, using different realizations of the HadISST2 sea surface temperature fields and model physics. Observation errors are represented by perturbations with statistics characterized by the observation error covariance matrix. The sea surface temperature fields are not directly perturbed but different realizations are provided as part of the HadISST2 product. The model physics is perturbed in a stochastic way, adding perturbations to the physical tendencies to simulate the effect of random errors in the physical parameterizations. Flow-dependent ensemble information from the EDA is incorporated into the atmospheric background-error covariance matrix

$$\mathbf{B} = \mathbf{T}^{-1} \Sigma_b^{1/2} \mathbf{C} \Sigma_b^{1/2} \mathbf{T}^{-\mathsf{T}}.$$

The diagonal matrix $\Sigma_b^{1/2}$ is the background-error standard deviation computed from the ensemble. Because of the finite ensemble size and because all the error sources in the EDA are not properly represented, the background error standard deviation is calibrated and smoothed. The background error correlation operator $\mathbf{C}$ is modeled in wavelet space and is determined from a hybrid formulation blending together climatological and current cycle forecast error estimates. The matrix $\mathbf{T}$ maps the control vector used in the minimization of the 4D-Var cost function to the model variables. The other components of the Earth system are not perturbed explicitly with stochastics physics but they receive the perturbed atmospheric forcing fields from the 10 member EDA.

The top plot of Figure 2 shows the background error standard deviation $\Sigma_b^{1/2}$ for vorticity at the surface level computed from the 10 ensemble members for 1 January 1910 overlaid by the background mean-sea-level pressure. Background standard deviation is smaller in the Tropics and larger in active weather systems (e.g., low pressure system over the Atlantic). The bottom plot of 2 shows the same diagnostic for 1 January 2010. The background standard deviation has reduced over time especially over regions with a better observing system. This is similar to results obtained by Poli et al. (2013). However, these same results had also demonstrated that when no upper-air observations were assimilated, the background errors estimated by the EDA method were too large: this allowed spurious increments to be added to the analysis, located very far away from observations. These increments were shown to destroy the natural ability of the model to create the realistic stratospheric trends shown by ERA-20CM (Hersbach et al., 2015a). Learning from this prior experience, the background error standard deviation is reduced in CERA-20C by a factor of 6 in the stratosphere. This prevents the 4D-Var to compute significant corrections in the stratosphere and avoids
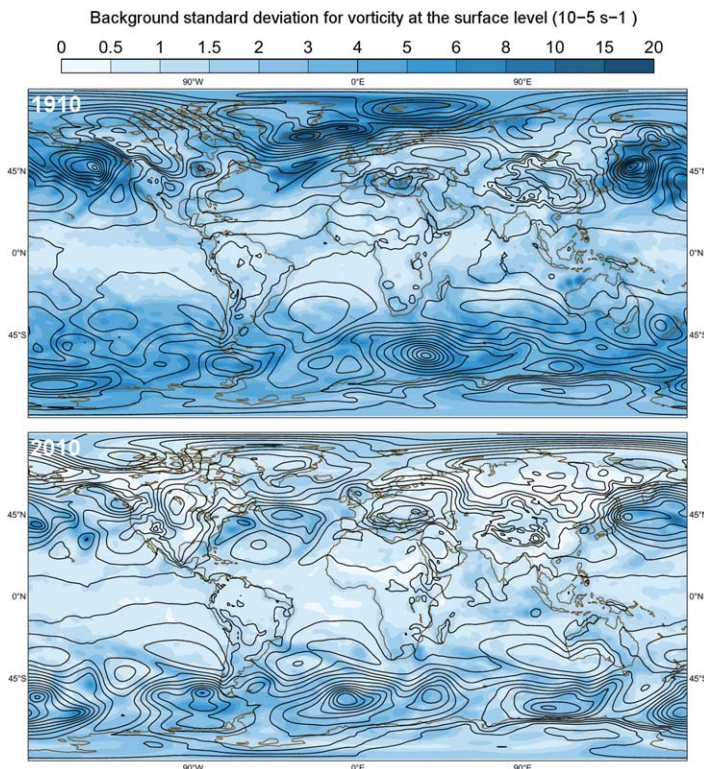
Figure 2. Background standard deviation for vorticity at the surface level $(10^{-5} s^{-1})$ computed from the 10 ensemble members for (top) 1 January 1910 and (bottom) 1 January 2010. The contours represent the background mean sea level pressure.
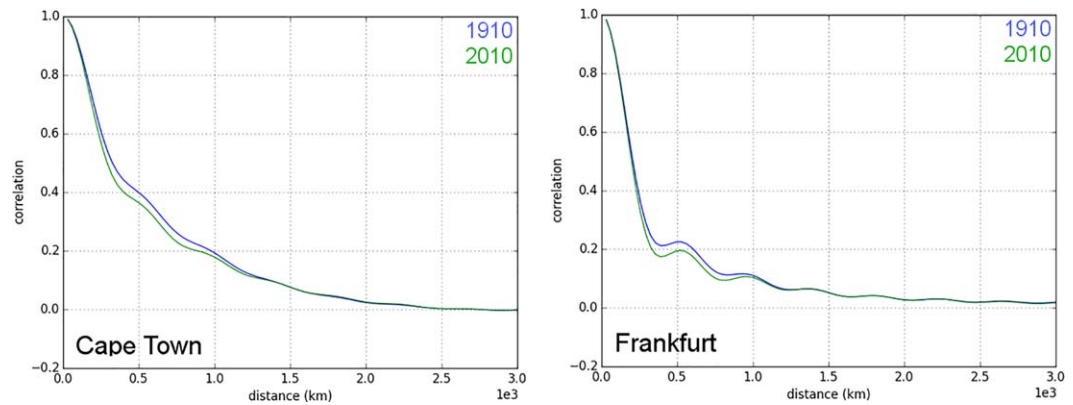
Figure 3. Horizontal correlations for surface pressure for 1 January 1910 (blue) and 1 January 2010 (green) for (left) Frankfurt and for (right) Cape Town.

spurious stratospheric climate signals caused by the vertical propagation of analysis increments related to surface observations.

In the hybrid formulation used for the background error correlation **C**, the climatological forecast error estimates are computed over summer and winter 2014 using a full observing system and get an 85% weight. The 10 daily forecasts produced for each assimilation cycle receive a 15% weight. Giving a larger weight to the small ensemble of 10 daily forecasts would introduce numerical instabilities in the computation of the background error correlation. Figure 3 shows the horizontal correlations for surface pressure for 1 January 1910 (blue) and 1 January 2010 (green) for Frankfurt (left) and for Cape Town (right). Similar to results obtained by Poli et al. (2013), the correlation distances for the background error are changing the synoptic scales. They are decreasing over time as the observing network is getting denser. Consequently, the observations in the year 2010 have a shorter radius of influence on the analysis than in the year 1910. The variational EDA reacts to the observation coverage as follows. In the beginning of the century, a poor observation coverage means that the observations are used to constrain the large-scales features, with little emphasis on small-scale increments. At the end of the century, a denser observation network allows small-scale features to be extracted from the observations (for the correlation value 0.4, the length scale is reduced for Cape Town by 20% over the century). The comparison between the horizontal correlation in Frankfurt and in Cape Town shows the impact of the better observing system over Europe compared to the Southern Hemisphere (for the correlation value 0.4, length scale in Frankfurt is half the size of that in Cape Town). This result is important as it confirms that the EDA is a powerful analysis method to cope with large changes in amounts of observations, allowing observations to have different scales of impact depending on their availability (Poli et al., 2013). This fully flow-dependent background error in CERA-20C is a major improvement as compared to the publicly-released ERA-20C deterministic reanalysis. Indeed, the underlying ensemble product in ERA-20C was not released having spurious stratospheric climate signals caused by the vertical propagation of analysis increments related to surface observations (see Poli et al., 2013, 2015, for more details).

3.2. Atmospheric Observation Error

Specifying a priori the observation errors over the entire twentieth century is not straightforward since measurement processes for the different platforms and representativeness errors (arising from the variability of the observed field at scales smaller than those resolved by the dynamical model) are not well known, and observation practices changed a great deal. Generally, greater automation occurred for all parameters, from reports of time and location to the observed geophysical variable itself, thereby reducing the contribution of human errors in each of them. The initial ensemble realization of ERA-20C used constant, observation errors over time, and the deterministic version used a similar approach, with values for the different platforms are given in Table 2. These error estimates included observation timing error and observation location error. The ensemble production (Poli et al., 2013) demonstrated by analysis of the OFA that these errors should in fact be considered as time-dependent. Indeed, location errors are very likely greater for all marine observations in the early part of the century, as compared to the 1990s when the Global Positioning System (GPS) started to generalize. It is possible to estimate the observation error a posteriori, using the method

Table 2
Specification of the Observation Errors for the Different Platforms in ERA-20C (deterministic) and CERA-20C

Observation type	ERA-20C	CERA-20C
Surface pressure from land stations	1.08 hPa	From 1.6 hPa in 1900 to 0.8 hPa in 2010
Surface pressure from ship	1.46 hPa	From 2.0 hPa in 1900 to 1.2 hPa in 2010
Surface pressure from tropical cyclone bogus	1.56 hPa	2.0 hPa throughout the century
Surface pressure from buoys	0.94 hPa	From 1.0 hPa in 1973 to 0.8 hPa in 2010
Ten-meter wind component from ship	1.5 m/s	2.2 m/s throughout the century
Ten-meter wind component from buoys	1.33 m/s	From 1.7 m/s in 1973 to 1.4 m/s in 2010

proposed by (Desroziers et al., 2005). This method assumes that the background and the observation errors present different correlation structures. This diagnostic was computed from the ERA-20C OFA (Poli et al., 2013). The results could not be applied in the ERA-20C deterministic system, as shown by (Poli et al., 2015), as this would have required reassessing time-varying background errors with an ensemble. However, the results (Poli et al., 2015) were used as in CERA-20C. The linearly time-varying observation errors used in CERA-20C are summarized in Table 2.

Data assimilation diagnostics can be computed from the different statistics stored in the OFA. One can form statistics of the expected departures at the beginning of the analysis window, by computing the square root of the sum of assumed observation error squared and background error squared. They can then be compared with the actual root mean square of background departures at the beginning of the window. Figure 4 shows both quantities for surface pressure observations reported at station level in the first 3 h of the assimilation window. In ERA-20C (left plot) the expected departure is much smaller than the actual departure during the first part of the century, meaning that background and/or observation errors were underestimated. In CERA-20C, the new specification of observation error (1.6 hPa instead of 1.08 hPa in 1900) with a consistent flow-dependent background standard deviation shows a better agreement between the expected and the actual departure for the whole century (right plot). The larger background departure at the beginning of the century in CERA-20C is a consequence of the less confidence in the observations. The drop in the actual background departure due to the sudden large increase of observations in the 1930s' happens at the same time in the predicted departure. As the observation errors are reducing linearly over time, this can only be explained by the EDA system which captures well the changes in the observing system. This illustrates that the specification of the background error in CERA-20C is better and more consistent compared to the deterministic production of ERA-20C.

3.3. Variational Bias Correction

Biases in surface pressure and mean sea level pressure observations are corrected by the Variational Bias Correction (VarBC). In this formalism, observation bias is estimated from a bias model based on a small set of bias parameters and data are stratified into groups where observations share the same set of bias parameters (Dee & Uppala, 2009). In ERA-20C and CERA-20C, the bias model for surface pressure and mean sea level pressure observations depends only on the pressure itself (Poli et al., 2013). Each physical station or platform forms its own bias group in which the bias estimate can evolve independently. This is quite a different situation from operational NWP, where typically thousands of satellite observations all share the same group, but this reflects the issue of metrology and calibration: each barometer poses a distinct problem, similar to a single satellite sensor.

The bias parameters are updated variationally in the 4D-Var assimilation, as additional parts of the analysis control vector. The level of confidence on the bias correction is controlled by an extra term in the 4D-Var cost function that includes a bias error covariance matrix which penalizes for deviations from the background bias parameters. This bias error covariance matrix is kept constant over time and
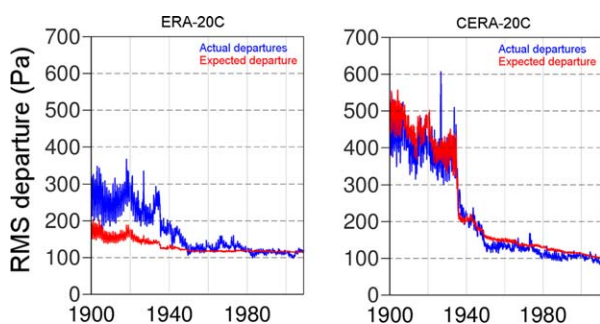
Figure 4. Monthly time series of Root Mean Square of observation minus background actual departures (blue) and predicted departures (red) in (left) ERA-20C and (right) CERA-20C, for assimilated observations of surface pressure reported globally at stations in the first 3 h of the assimilation window.

determines the reaction speed of the variational bias correction. The bias standard deviation error has been tuned to have a fast enough update when there is a sharp change in the observation bias and to have a slow enough update for stations with stable biases. The approach used in ERA-20C was slightly different as the bias computation was performed before the 4D-Var assimilation. This alternative aims to avoid any interaction with the model background and to have a better control on the response time of the observation bias.

3.4. Quality Control

Quality control (QC) of observations is an important component of any data assimilation system (Lorenc & Hammon, 1988). Observations suffer from random measurement errors, and sometimes gross errors due to technical, reading or transmission problems. The goal of QC is to ensure that only correct observations are used, and that erroneous observations are discarded from the analysis process. It has long been recognized that a good quality control process is required because adding erroneous observations to the assimilation can lead to spurious features in the analysis.

The first QC step is the blacklist. It includes a set of predefined rules to reject observations. CERA-20C benefits from an improved blacklist based on information gained from the ERA-20C OFA. From hundreds of stations reporting spurious or constant measurements over time assimilated in ERA-20C were blacklisted in CERA-20C. During the second QC step, preferences are assigned to remove redundant information within a report. The rules, as employed for operational NWP, are as follows: a surface pressure observation is preferred to a surface geopotential observation and a pressure observation reported at the station level is preferred to a pressure observation reported at sea-level (as mean sea level pressure over land is a quantity that is calculated using some assumptions on the temperature profile). The third QC step is the so-called background (or first-guess) check, which rejects any observation whose actual departure from the background value is more than 3 times larger than the expected departure (square root of the sum of observation error squared and background error squared). Finally, a variational quality control (Anderson & Jarvinen, 1999) using the Huber norm adjusts the observation weight within the assimilation (Tavolato & Isaksen, 2015). Data with a large departure are then assigned a smaller weight than data with a small departure. This approach reduces the issues caused by inclusion of outlier observations in the analysis step.

3.5. Ocean Assimilation

No major modification has been made to the ECMWF operational ocean assimilation system to produce CERA-20C, except that the ocean model bias correction scheme has not been applied. This system employs the NEMOVAR three-dimensional variational data assimilation system (3D-Var), using the First-Guess at Appropriate Time (FGAT) approach (Mogensen et al., 2012b). Although CERA-20C has an ensemble of ocean reanalyses, the ensemble information is not used in the background errors. Therefore, there is no need to tune the background errors, apart from accounting for the shorter assimilation window. A substantial difference with respect to the other operational ECMWF ocean reanalyses (ORAS4 and ORAS5) is the lack of ocean model bias correction in CERA-20C. This bias correction scheme aims to correct for temperature and salinity biases in the extra-tropical regions and applies a pressure correction in the tropical regions during the nonlinear model integration (Balmaseda et al., 2007). The bias term is the sum of an offline bias estimated as the monthly climatology of the ocean assimilation increments from a preproduction run over a well-observed recent period (1989–2008) and an online bias which is updated each analysis cycle (Mogensen et al., 2012b). The bias of the unforced ocean model is not necessarily the same as the bias of the coupled model. Since the iterative estimate of the offline bias in coupled mode is prohibitive, the bias correction was not used in CERA-20C for consistency reasons.

3.6. Data Set

The production of climate reanalyses requires large computing and archiving resources. To produce the CERA-20C data set in a reasonable amount of time, the period 1900–2010 was divided into 14 different streams, each covering 10 years. Each production stream was initialized from the uncoupled ERA-20C and ORA-20C reanalyses. Initializing the slow component of the earth system (ocean, sea-ice, and land) to a continuous a priori reanalysis should in principle prevent discontinuities in the ocean temporal evolution due to drifts in the coupled system. The first two years of each production stream were used for spin-up (only 1 year of spin-up for the first stream) to produce the final climate data set for the period 1901–2010. The computation footprint of CERA-20C on ECMWF's high-performance computing facility is significant, with 7

months of production using 20,000 cores, which represents 5% of the total resources. About 500,000 variational problems had to be solved at a pace of one every 30 s and processing up to 5,000,000 observations.

The CERA-20C product describes the spatiotemporal evolution of the atmosphere (125 km horizontal grid with 91 levels in the vertical, between the surface and 0.01 hPa), the land-surface (125 km horizontal grid with four soil layers), the waves, the ocean (110 km horizontal grid with meridional refinement at the equator and 42 vertical levels) and the sea ice. The temporal resolution is 3 h. CERA-20C provides 10 ensemble members and users can access precomputed ensemble mean and spread. Two sets of monthly means are available: synoptic monthly means and monthly means of daily means. In addition, the Observation Feedback Archives (OFA) which contain feedback information is also available. This represents in total more than 1,600 terabytes of data archived in ECMWF's MARS archiving system.

## 4. Synoptic and Upper-Air Evaluation

4.1. Representation of Tropical Cyclones

One potential value of historical reanalyses lies in their ability to study weather extremes, since extremes are rare and satellite-era reanalyses are too short to represent a long enough sample for climate studies. While midlatitude storms, heat waves or cold-air outbreaks are often well-represented in regions with dense observational coverage (Bronnimann & Luterbacher, 2012; Donat et al., 2013; Messori et al., 2016), this is not always the case for tropical cyclones, which are difficult to model and not well-constrained by observations.

In order to study in detail the representation of hurricanes in surface-only reanalyses, we compared two similar hurricanes near the beginning and end of the CERA-20C period: the Galveston hurricane in 1900 and the Katrina hurricane in 2005. The mean sea level pressure analyses from 20CRv2c and CERA-20C have been plotted in Figure 5 overlaid by the analysis departure (analysis - observation) at locations where observations have been assimilated. The Galveston hurricane appears as a deep cyclone in 20CRv2c (top left) while it is clearly missed in CERA-20C (middle left). The study of the observation feedback archive reveals why hurricanes are represented differently in the two reanalyses. The lowest observed pressure values presented to the assimilation systems are mostly those from the so-called "Best Track" observations from the International Best Track Archive for Climate Stewardship (IBTrACS) (Knapp et al., 2010). These observations from the IBTrACS data set are not always actual measurements, some are interpolations or approximations based on wind data (Compo et al., 2011). In CERA-20C and ERA-20C systems, a lot of the Best Track information is rejected by the system in the first guess check since the difference between the background and the observation is too large. All the Best Track observations have been rejected in CERA-20C during the Galveston hurricane. Conversely, in the 20CRv2c system, the Best Track data are allowed to bypass the usual quality control in order to avoid their rejection. The assimilation of the four Best Track observations in 20CRv2c represented by the circle markers in Figure 5 is sufficient to produce an intense hurricane with strong winds and heavy precipitation. The situation is different for Katrina as the hurricane appears in both reanalyses (990 hPa in 20CRv2c and 992 hPa in CERA-20C). The Best Track values (down to 915.2 hPa) are still rejected by the CERA system as they are too far from the background, but there are many other observations that inform the assimilation systems about the hurricane in this case. The 20CRv2c assimilates the Best Track data, but a large amount of the data presented to the 20CRv2c system is not assimilated but thinned-out by the system (Compo et al., 2011). In this procedure, observations that have no strong effect on the ensemble spread are not assimilated, such that the fraction of assimilated observations decreases from 98% in 1891 to 32% in 2005.

Forcing the CERA system to assimilate the Best Track observations has been tested by bypassing the first guess check and the variational quality control, and the results are shown in the bottom row of Figure 5. With this modification, the hurricane Galveston is represented with a lowest mean sea level pressure equal to 974 hPa (968 hPa in 20CRv2c). However, the size of the hurricane is too large and the fit to the other observations in the vicinity of the hurricane is degraded. In a poorly observed period, the background error covariance constrains the large-scales features, spreading too much the information over space with little emphasis on small-scale increments. The forced assimilation of the Best Track observations for Katrina has deepened the hurricane in the CERA system from 992 to 984 hPa (990 hPa in 20CRv2c) improving slightly the fit to the other observations.

This study shows that reproducing hurricanes from the operational NWP 4D-Var in a surface-only reanalysis remains a challenge and that several steps in the procedure (e.g., first-guess checks, variational quality
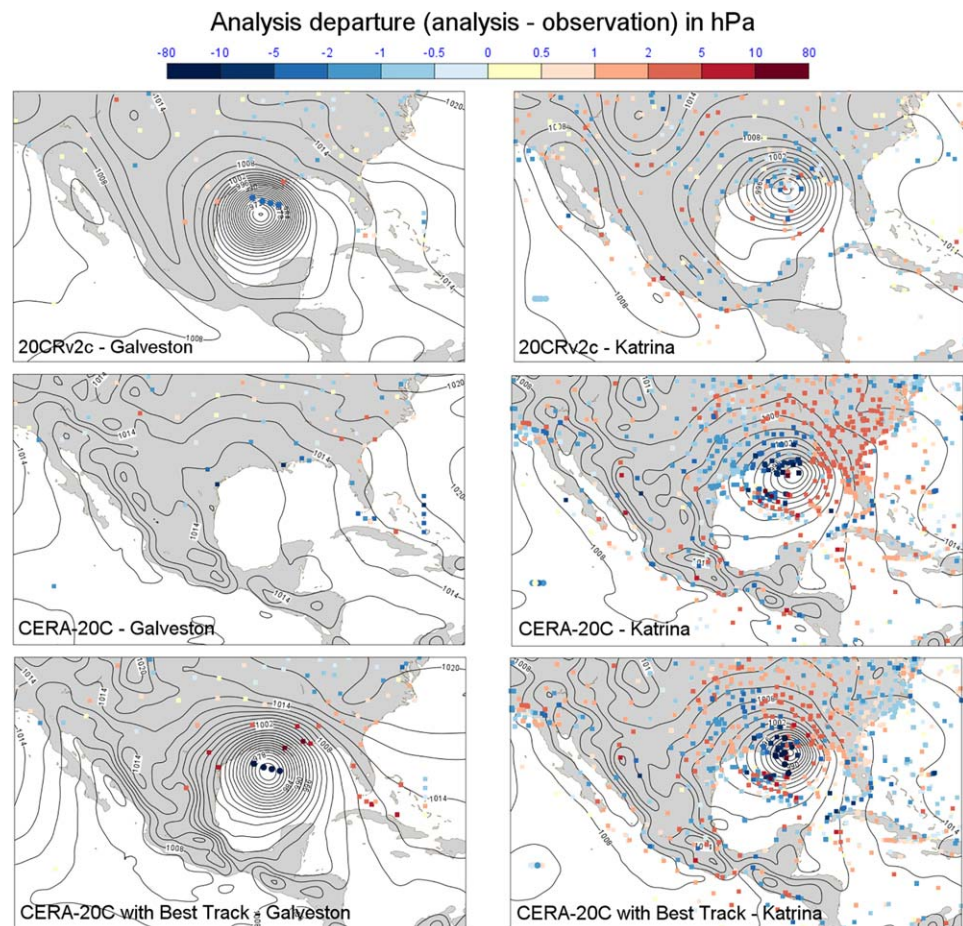
Figure 5. Mean sea level pressure analysis (contours, 2 hPa spacing) from (top) 20CRv2c, (middle) CERA-20C control member, and (bottom) CERA-20C control member where the assimilation of Best Track observations have been forced, for (left) the Galveston hurricane and (right) the Katrina hurricane. Symbols indicate the pressure observations that have been assimilated to produce the analysis (circles for the Best Track observations, squares for the others) and their colors show the analysis departure (analysis - observation) in hPa.

control, assignment of observation error, thinning out of observations) are important. Bypassing the first guess check for Best Track observations does not give satisfactory results in CERA as it degrades the fit to the other observations. Some work is ongoing to improve the way the background error covariance matrix is spreading the information in space, allowing smaller-scale adjustments.

4.2. Upper-Air Evaluation

Since CERA-20C and 20CRv2c only assimilate surface data, any upper-air observations can be used for independent evaluation. The time shortly after the International Geophysical Year (1957–1958) is particularly useful as the global radiosonde network is relatively dense and can provide independent conventional upper-air observations. Figure 6 shows the monthly standard deviation of the analysis departure for CERA-20C and 20CRv2c ensemble means with respect to unadjusted temperature observations from radiosondes at 700 hPa, averaged over years 1959–1960. CERA-20C outperforms 20CRv2c in the northern extratropics (30–60°N) with a standard deviation around 1.6K compared to 2.2K in 20CRv2c. The performance of the two reanalyses looks more similar in the Tropics and over the Southern Ocean region. Similar conclusions can be drawn at 200 hPa. The analysis departure bias in CERA20C are smaller in most regions (not shown), but the picture is less clear in this respect since radiosonde temperature observations can have substantial biases themselves (this is the reason why standard deviation is preferred here over root-mean-square error).

One way to evaluate the quality of the ECMWF reanalyses is to compare the quality of 10 day forecasts initialized from each reanalysis. Figure 7 shows the anomaly correlation coefficient of NH geopotential height
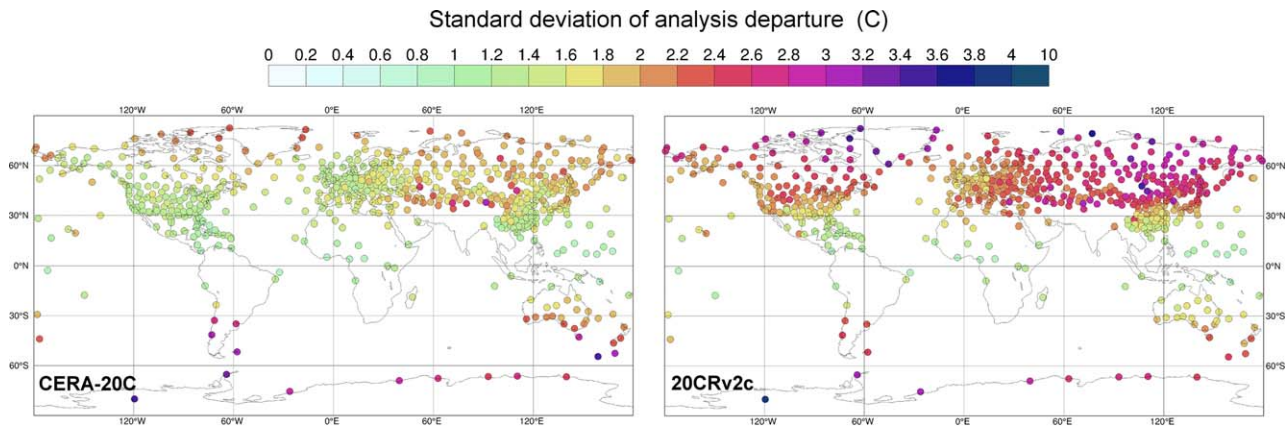
Figure 6. Monthly standard deviation of the analysis departure for (left) CERA-20C and (right) 20CRv2c ensemble means with respect to unadjusted temperature observations from radiosondes at 700 hPa, averaged over years 1959–1960.

at 500 hPa (left), SH geopotential height at 500 hPa (middle), and tropics temperature at 700 hPa (right) in CERA-20C, ERA-20C, and ERA-Interim. The scores are aggregated over 1 March 2010 to 1st June 2010 and the ECMWF operational analysis is used as verifying truth. The 10 day forecasts produced in each reanalyses are all initialized at 00UTC but they are based on different assimilation windows (21UTC to 21UTC in CERA-20C, 9UTC to 9UTC in ERA-20C, 15UTC to 03UTC in ERA-Interim). The forecasts for the CERA-20C system therefore benefit from an additional 12 h of observations compared to ERA-20C and 18 h of observations compared to ERA-Interim. This advantage is taken into account by shifting the ERA-20C and ERA-Interim curves by 12 h and 18 h to get a fair comparison. The improvement in CERA-20C compared to ERA-20C is around 0.5 day at the 70% anomaly correlation threshold and it comes from the use of a newer IFS model, the EDA technique and the ocean coupling in the analysis and forecast. ERA-Interim has better scores since it assimilates upper-air and satellite observations, which improves the geopotential height and temperature analysis.

The impact of the ocean coupling in the CERA-20C reanalysis has been assessed separately by running an extra experiment where the ocean and sea-ice coupling of the CERA-20C system has been switched off (CERA-20C-nocoup). In this uncoupled version, sea surface temperature and sea ice are prescribed using the monthly HadISST2 product. Figure 8 shows the normalized difference in the anomaly correlation coefficient between CERA-20C and CERA-20C-nocoup for NH geopotential height at 500 hPa (left), SH geopotential height at 500 hPa (middle), and the tropics temperature at 700 hPa (right). A positive value means that the ocean coupling increases the anomaly correlation. Ocean and sea-ice coupling improves the anomaly correlation for 500 hPa geopotential height in the Southern hemisphere (mostly covered by oceans). In the Tropics, it improves the anomaly correlation for the temperature analysis and forecast up to day 5. One explanation of the small degradation in the Northern hemisphere is the position of the Gulf stream which is located too North compared to the HadISST2 product. The actual impact of ocean and sea-ice coupling during the assimilation might be difficult to assess using the overall statistics presented in Figure 8. The quality
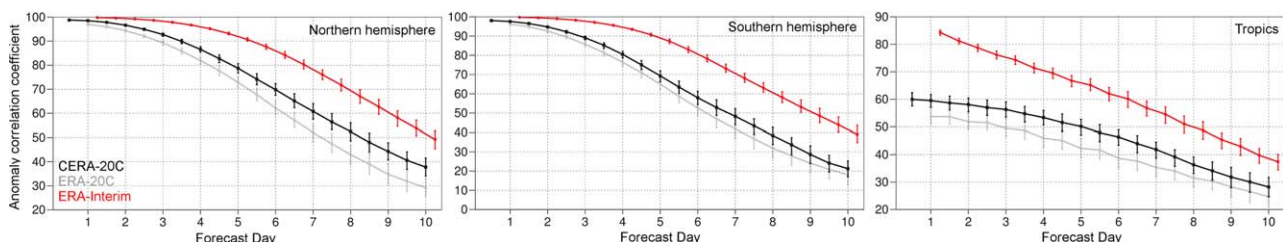
Figure 7. Anomaly correlation coefficient of (left) NH (90° to 30°) geopotential height at 500 hPa, (middle) SH (−30° to −90°) geopotential height at 500h Pa, and (right) TROPICS (30° to −30°) temperature at 700 hPa in ERA-Interim (red), ERA-20C (grey) and CERA-20C (black). Scores are aggregated over 1 March 2010 to 1 June 2010 and the ECMWF operational analysis is used as reference.
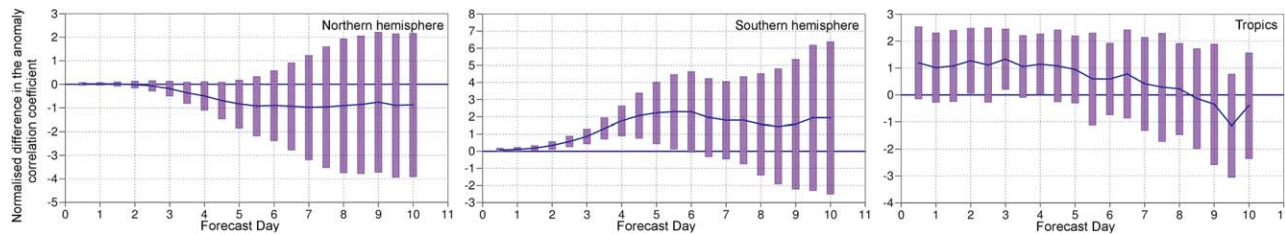
Figure 8. Normalized difference in the anomaly correlation coefficient between CERA-20C and CERA-20C without ocean and sea-ice coupling for (left) NH geopotential height at 500 hPa, (middle) SH geopotential height at 500 hPa, and (right) TROPICS temperature at 700 hPa. A positive value means that the ocean coupling increased the anomaly correlation coefficient. Purple bars represent the 95% confidence interval. Scores are aggregated over 1 March 2010 to 1 June 2010 and the ECMWF operational analysis is used as reference.

of medium-range forecasts in the northern extratropics has generally improved by around 1 day (about 20%) per decade (Magnusson & Kallen, 2013; Simmons & Hollingsworth, 2002). Most of the improvement likely comes from the flow of smaller developments which might have an impact of less than 30 min of forecast skill, or about 0.5%, which is hard to detect with forecast verification. For this reason, statistical evaluation must come from case studies that look at the performance of the ocean coupling for specific weather situations such as Tropical cyclones (Laloyaux et al., 2016b).

5. Weather and Climate Signals

5.1. Ocean Heat Balance

One of the benefits expected from a coupled assimilation system is a more consistent treatment of the air-sea interface. When decoupled, the ocean and atmospheric systems are fed by boundary conditions that do not take into account ocean-atmosphere feedbacks. In ERA-20C, the atmospheric lower boundary conditions come from the HadISST2 sea-surface temperature and sea-ice monthly analysis. Poli et al. (2015) showed that atmospheric temperature increments in ERA-20C are not unbiased and that their impact on the atmospheric heat budget is compensated by energy fluxes at the top of the atmosphere and at the surface (over land and ocean). In ORA-20C, the ocean receives the surface fields from ERA-20C as upper boundary conditions. Those fields are fixed and impacted by the atmospheric energy balance and they cannot adjust to the ocean model behavior. In the early twentieth century, the ocean heat content variability is driven by relatively stable air-sea fluxes. Like in ERA-20C, the ocean temperature increments that work to keep the ocean close to the observed state from the 1950s onward are not unbiased and their impact on the ocean energy budget is compensated by energy fluxes at the air-sea interface. The resulting net heat fluxes over the ocean in ORA-20C show a negative trend from the 1940s onward that responds to a growing positive temperature increment (Figure 9), highlighting the inconsistencies between ocean and atmosphere at the air-sea interface. ERA-20C and ORA-20C are clear examples of the harm that uncoupled analysis system can cause on the energy balance of the climate system. In CERA-20C, the ocean and the atmosphere communicate hourly through the air-sea coupling at the outer-loop level of the variational method. Changes in the state of the atmosphere directly impact the ocean properties and vice versa and both
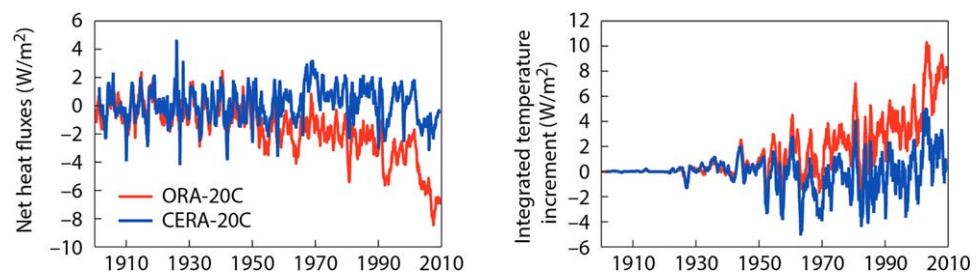
Figure 9. Time series of ORA-20C and CERA-20C control member values of the global average of (left) net air-sea heat fluxes and (right) the integrated ocean temperature increment.

systems adjust to each other. On average, heat flux at the air-sea interface and ocean temperature increments in CERA-20C oscillate around zero (respectively, $0.15 \pm 1.11 \, W/m^2$ and $-0.11 \pm 1.9 \, W/m^2$ in CERA-20C; $-1.62 \pm 1.89 \, W/m^2$ and $1.6 \pm 2.32 \, W/m^2$ in ORA-20C), suggesting a more consistent energy balance between ocean and atmosphere.
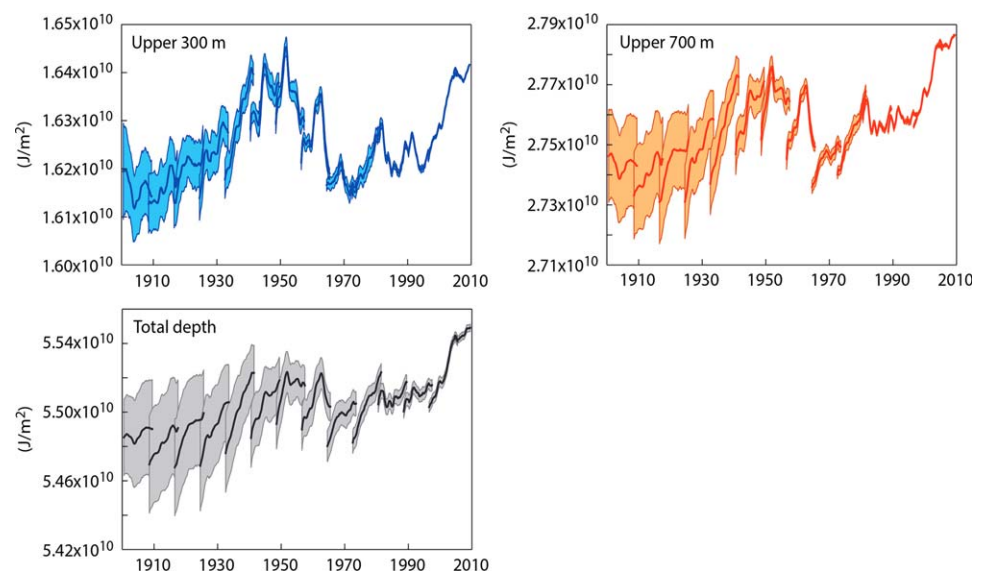
5.2. Ocean Heat Content

The evolution of ocean heat content over the twentieth century is of particular interest as it has been identified in several studies as an indicator of ocean heat uptake, a process that is relevant to climate studies. In CERA-20C, time series of heat content show discontinuities between streams resulting from the model drift from its initial state (Figure 10). The model drift reflects the fact that the initial conditions from ERA-20C and ORA-20C used to initialize the different production streams are inconsistent with the coupled model's natural state. The origin of the drift remains unknown so far. The complexity of the system makes it very difficult to point toward a single explanation and this question remains open to further investigations. In the early twentieth century, when the uncertainty in the state of the ocean is high and the ocean model is poorly constrained by observations, the ocean component of CERA-20C drifts toward its preferred state. As the observing system grows, the uncertainty and the drift are reduced. The relatively well-observed upper ocean adjusts faster than the ocean interior, where the timescales of ocean processes are particularly slow and the observational constraints are very small. Further work is needed to understand and reduce the model drift so that the initial conditions and the ocean model behavior are more realistic in poorly observed periods and areas.
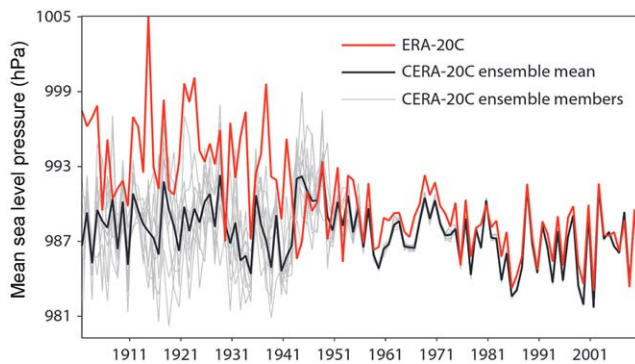
5.3. Mean Sea Level Pressure Trends

An issue with the general circulation in the southern hemisphere has been found in ERA-20C and 20CRv2c (Poli et al., 2013, D. Bromwich, personal communication, 2013). In these reanalysis, the time series of mean sea level pressure (MSLP) decreases significantly between 1900 and 1950 over the Antarctic region, leading to a substantial strengthening of the polar vortex in the first half of the twentieth century (Figure 11). This spurious climate signal has been investigated during the preparatory work for the CERA-20C production.

In any data assimilation scheme, the specification of observation and background errors is crucial as it defines the weights used to blend together the information from the measurements and from the model. It has been found that the observation error for pressure observations specified in ERA-20C (Table 2) are too small at the beginning of the century, giving too much weight to the observations located in the subtropical high pressure belt. Large positive MSLP increments over the unobserved Antarctic region are generated

Figure 10. Time series of the global average ocean heat content (J/m^2) in the CERA-20C ensemble for the (top left) upper 300 m, (top right) the upper 700 m, and (bottom left) the entire water column. The solid lines are the ensemble mean and the shading shows the ensemble standard deviation.

Figure 11. Time series of mean sea level pressure (MSLP) for the latitudes 90°S–60°S averaged over the period September–November each year.

as the assimilation fits the observations by adjusting the flow where it is the least constrained. This is illustrated on the right plot of Figure 12 for the year 1924. The systematic positive increments computed over the years result in too high MSLP values in the analysis. This behavior disappears when the observing system becomes denser in the 1950s, with the first SYNOP stations in Antarctica and more observations from ships in the Antarctic Circle which better constraint the assimilation system. For the production of CERA-20C, observation errors have been reviewed and are now time-varying to make them more realistic (Table 2). The larger observation error at the beginning of the century increases slightly the analysis departure with pressure observations, but prevents large positive increments over the Antarctic region (left plot of Figure 12). As a result, the CERA-20C ensemble mean looks more realistic with better consistency in the climate trends. The larger ensemble spread at the beginning of the century reflects the larger uncertainties in the climate reconstruction as the region is poorly observed before the 1950s.

5.4. Precipitations and Droughts

CERA-20C and ERA-20C precipitation extremes are compared to the gridded GPCC Full Data Daily product (FDD Schamm et al., 2014) which is based on gauge data. The index plotted in Figure 13 is the highest number of Consecutive Dry Days (CDD) over the 1988–2010 period for the three data sets (Tank et al., 2009). A threshold of 1 mm/d is used to define a dry day. The CDD extreme index is an important measure of the aridity of regions and thus an important indicator for droughts. There is a overall good agreement between CERA-20C and ERA-20C (which does not assimilate any precipitation measurement) and the gridded observation product. The longest periods without precipitation occur especially in the desert areas, especially in the Sahara and Atacama deserts (CDD > 7 years). In these dry areas, the differences between the reanalyses are mainly in the extent (e.g., Sahara) or magnitude and position (e.g., North Australia). Differences in rainfall patterns based on the orography are particularly evident in Congo, where CERA-20C outperforms ERA-20C being more consistent with the GPCC-gridded observations.
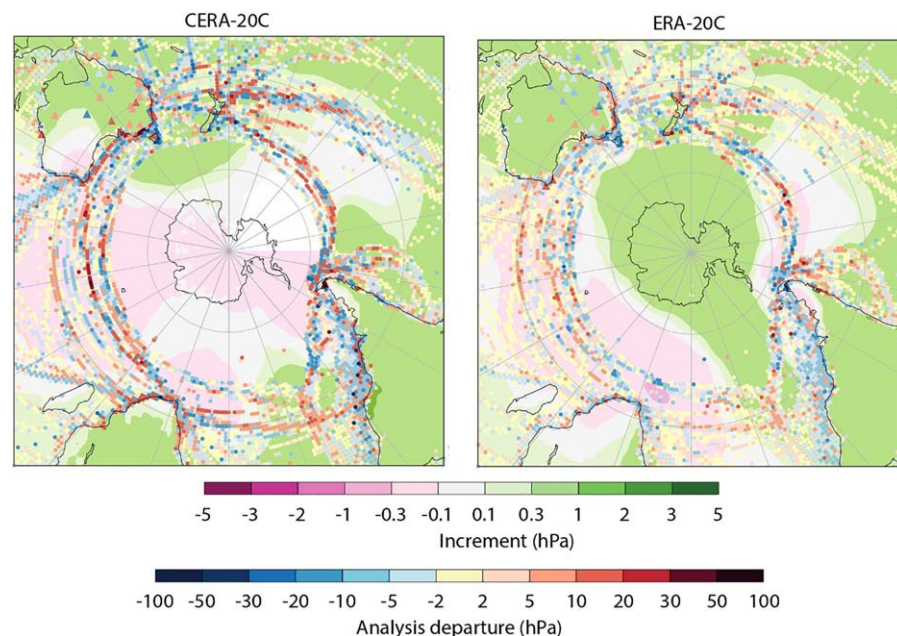
Figure 12. Mean sea level pressure increments (shading) for the year 1924 in (left) ERA-20C and (right) CERA-20C. Dots and triangles represent the analysis departure with mean sea level pressure observations and surface pressure observations, respectively.
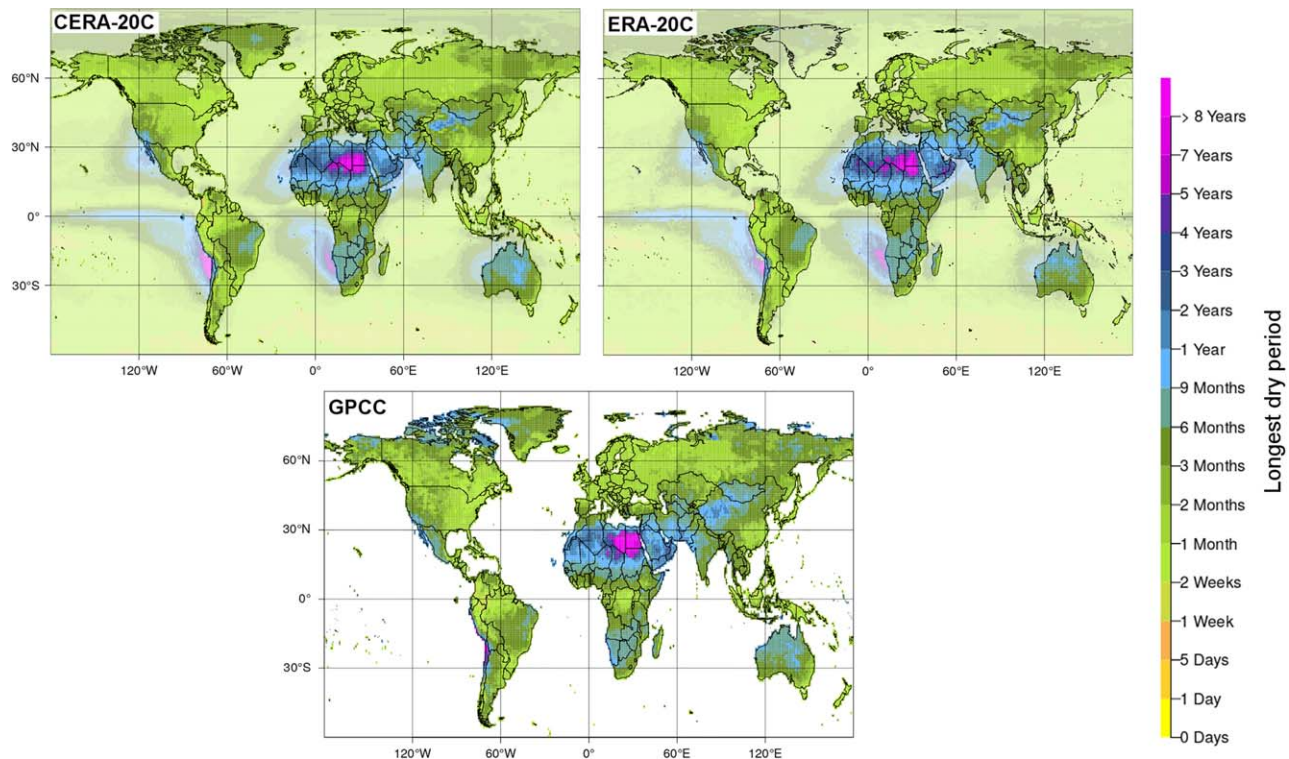
Figure 13. Longest period of consecutive dry days over the 1988–2010 period for (left) CERA-20C, (middle) ERA-20C, and (right) the GPCC gridded observations.
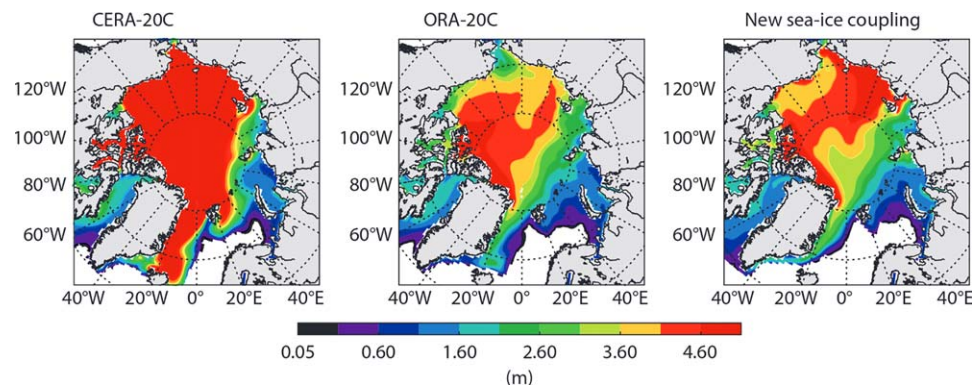
5.5. Sea Ice

Ocean-sea ice interactions through the LIM2 model have only recently been included in ECMWF's coupled model. ORA-20C provides a first record of sea-ice conditions for the twentieth century in ocean-only mode while CERA-20C is the first application allowing these interactions in coupled mode on an interannual time-scale. During the course of the production, CERA-20C sea-ice component showed a lack of summer melting, leading to the accumulation of Arctic sea ice over the years. By the end of a CERA-20C stream, the sea-ice thickness is over 5 m in most of the Arctic basin, more than twice the expected average of 2–2.5 m, as seen in ORA-20C (Figure 14). By constraining the ocean surface temperature toward observations, the SST relaxation was nevertheless able to keep the sea-ice extent under control. Further investigations were conducted using the ECMWF coupled model and it was found that the configuration of the sea-ice coupling with the atmosphere was not optimal. The main issue was the sea-ice temperature information from LIM was not transferred correctly to the atmosphere. This problem was fixed and tested in coupled model experiments that showed a more realistic behavior closer to the ocean-only mode (right plot of Figure 14). Sea-ice inter-actions with the ocean and the atmosphere are highly sensitive processes and will need to be monitored carefully for future reanalysis.

5.6. Seasonal and Subseasonal Coupled Processes

It is important for a product such as CERA-20C to capture seasonal and subseasonal coupled processes as they will have a crucial impact for climate monitoring and predictability studies. The El Niño-Southern Oscilla-tion (ENSO) is the principal mode of variability that impacts both atmosphere and ocean on seasonal to inter-annual timescales. A multivariate ENSO index (MEI) is estimated using sea level pressure and SST from CERA-20C and ERA-20C following Wolter and Timlin (2011). Combined EOFs of SST and MSLP are computed from bimonthly-averaged time series over the Tropical Pacific. The first EOF captures the ENSO signal and the corre-sponding principal component is the MEI (Figure 15). Both ERA-20C and CERA-20C agree very well with the observation-based index from Wolter and Timlin (2011) - with correlation of 0.88 and 0.89, respectively, show-ing that CERA-20C is doing a reasonable job at capturing the ENSO signal at the air-sea interface.

Figure 14. Arctic sea-ice thickness in March 1932 from (left) CERA-20C, (middle) ORA-20C, and (right) a coupled model run with new sea-ice coupling.

Capturing the signature of the ENSO events in the ocean subsurface is also relevant as predictability studies based on coupled reforecasts will use initial conditions from coupled reanalyses such as CERA-20C. Figure 16 shows the upper ocean heat content (0–300 m) in the Equatorial Pacific (5°S–5°N). CERA-20C compares very well with ocean-only reanalyses ORA-20C (de Boisseson et al., 2017) and ORAS4 (Balmaseda et al., 2013) in the well-observed period, capturing the sequence of warming/cooling associated to the major ENSO events of 1972–1973, 1982–1983, and 1997–1998. The early decades do not show clear signals in the ocean subsurface, confirming the general ENSO-weakness during the early to mid-twentieth century described by Wolter and Timlin (2011) with the exception of the 1940–1942 El Niño.

Capturing subseasonal coupled processes also matters for predictability studies. In this paper, we focus on the Tropical Instability Waves (TIW) that are said to play a part in ENSO interannual variability and impact ENSO predictability (An, 2008; Ham & Kang, 2010). TIWs are westward-propagating features of SST mostly visible in the eastern equatorial Pacific in the 1°S–3°N band. These waves have wavelengths of 1,000–2,000 km, periods of 15–30 days, and phase speeds of about 0.5 m/s (Willett et al., 2006). The TIWs are characterized by a very close relationship between SST and wind stress. In the following, SST and surface wind stress fields are spatially high-pass filtered to attenuate wavelength longer than 2,000 km as in Chelton et al. (2001). A temporal band-pass filter (centered on the window 10–60 days) is also applied to focus on the period of the TIWs. As TIWs are more intense in La Niña periods, the filtered SST and wind stress are plotted in Figure 17 at 1°N in the Pacific for the period 1973–1974. The atmosphere-ocean coupling in CERA-20C allows the system to capture the westward propagation of the TIWs. Filtered SST (contours) and wind stress (shading) are in phase. Positive SST values (plain contours) trigger stronger wind stress while negative SST values (dashed contours) coincide with weaker wind stress. By contrast, ERA-20C being forced by a monthly SST analysis cannot represent the atmospheric response to a signal that is not present in its lower boundary conditions. The extent to which the TIW signal propagates in the vertical and how it interacts with the general atmospheric circulation would need further investigations that are beyond the scope of this study.

6. Uncertainties

CERA-20C provides a 10 member ensemble reconstruction for all parameters and levels over a climate period (in this case, the twentieth century). Ensemble generation is based on the EDA system which explicitly accounts for errors in the observational record and in the forecast model. The information from the ten members is used during the assimilation to compute a flow-dependent background error, which determines how to spread the information from observations in space. The ensemble technique also aims to provide an indication of the confidence we can have in the reanalysis product. The
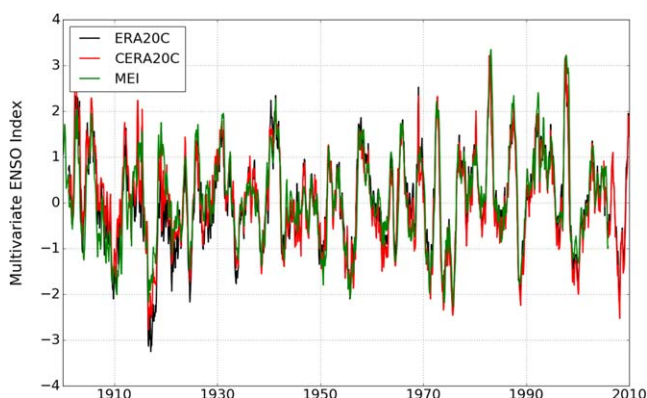
Figure 15. Multivariate ENSO index computed as in Wolter and Timlin (2011) for SST and MSLP. ERA-20C is in black, CERA-20C in red and the index estimated from observational products by (Wolter & Timlin, 2011) is in green (rescaled for comparison).
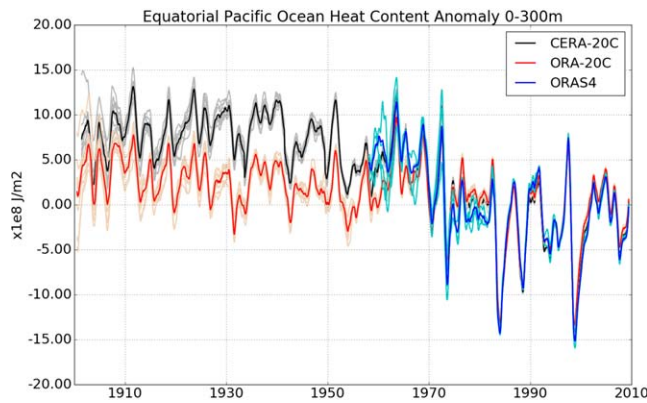
Figure 16. Equatorial Pacific upper 300 m ocean heat content anomalies with respect to 1958–2010. ORA-20C ensemble (10 members) are in light red, with the ensemble mean in red. CERA20C ensemble (10 members) in grey, with the mean in black. An OHC increase of $1 \times 10\,8$ J/m² corresponds to a temperature increase of $0.08°$ C averaged over the top 300 m.

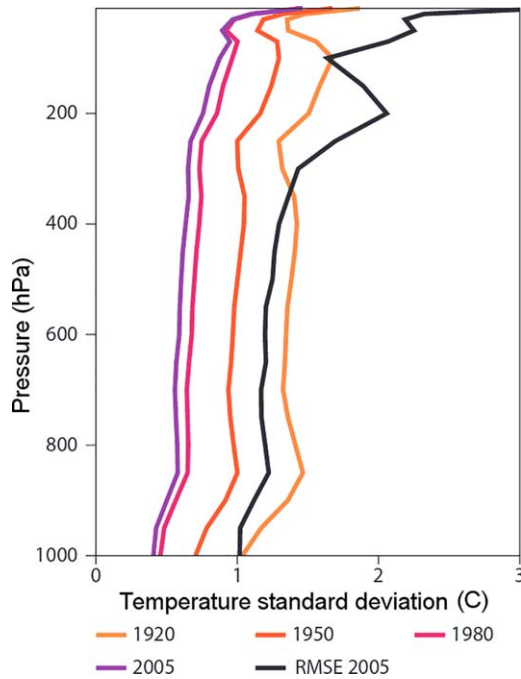ensemble standard deviation for temperature over 60°S–60°N is equal to 1.0°C at 1,000 hPa in 1920 and reduces to 0.4°C in 2005 (Figure 18). The ensemble spread in CERA-20C gradually decreases over time, which implies a higher confidence in the reanalysis as more observations become available. The reanalysis ensemble spread is supposed to represent the error of the reanalysis ensemble mean. It should ideally be equal to the root-mean-square error (RMSE) of the ensemble mean compared to the true atmospheric state. This has been verified for the year 2005, when ERA-Interim provides a good proxy for the truth as it assimilates all types of observation at a higher resolution. Maps of the ensemble spread and of the RMSE show very similar horizontal structures, which means that the EDA correctly captures where the uncertainties are (not shown). However, the RMSE of CERA-20C is about twice as large as the CERA-20C analysis ensemble spread. This global offset between the spread and RMSE can be seen in the vertical profiles (Figure 18). CERA-20C is thus overly confident in the data compared to the actual error. Figure 19 shows a similar study but in observation space using the independent upper-air temperature measurements from radiosondes as another proxy for the truth. One can see a good correspondence between the CERA-20C ensemble spread at the observation locations and the monthly standard deviation of analysis departures averaged over the period 1959–1960, although the ensemble spread is globally too small by a factor of 2. To improve the uncertainty estimations in future reanalysis, the size of the ensemble and the way the perturbations are generated in the different members when the EDA system is used will require further investigation to reduce the current offset between the ensemble spread and the actual error.

The CERA-20C ensemble spread for the wave component has been assessed using the same methodology. The left plot of Figure 20 shows the ensemble mean for significant wave height against all available in situ data for a 2 year period (2005–2006). The observations are mostly from moored buoys, with the exception
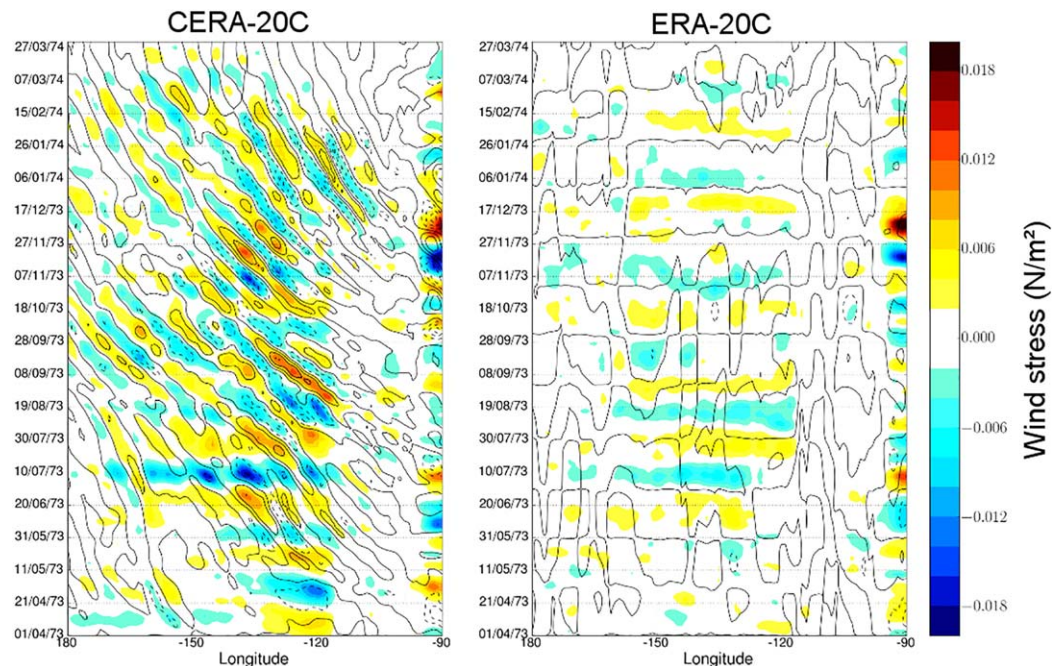
Figure 17. Hovmoller time series of spatially high-passed filtered SST (in C, contours) and wind stress (in N/m², shading) at 1°N in the eastern Pacific from April 1973 to April 1974. SST contours range from −1°C to 1°C every 0.25°C. Positive (negative) contours are plain (dashed). CERA-20C represents the Tropical Instability Waves (TIWs) thanks to the ocean dynamics and the atmosphere is responding accordingly with the surface wind stress sensitive to the ocean TIWs.

Figure 18. Vertical profiles of the standard deviation of the CERA-20C ensemble for temperature over 60°S–60°N for the years 1920, 1950, 1980, and 2005. The root-mean-square error (RMSE) of the CERA-20C ensemble mean compared to ERA-Interim in 2005 has been plotted for comparison.

of platform data from the Gulf of Mexico and the North Sea. The quality control procedure and collocation method are described in Bidlot (2017) and Bidlot et al. (2002). In this validation against observations, there is no special account of the uncertainty and representativeness error of the observations as demonstrated in Saetra and Bidlot (2004). Generally, the quality of CERA-20C wave heights is rather good, with exception in nearshore areas, with a marked degradation, usually associated with a tendency to underestimate (not shown). The coarse resolution of the wave model component as well of the forcing wind fields are the main cause. The right plot of Figure 20 shows the mean of the ensemble spread for significant wave height at the same locations as the in situ observations. As found in the atmosphere, the spread of the CERA-20C ensemble as an estimate of the error of the ensemble mean tends to be too small (note the change in color scale by a factor of 4). The mean spread is at least a factor of 4 lower and fairly uniform. A similar analysis was also carried out for the surface 10 m winds. It was found that the mean spread is too low by a factor of 2–3, more in line with what was found in the atmosphere. It is worth noting that the spread in the waves is only generated from the impact of the spread in the forcing wind as there is no specific uncertainty modeling in the wave model. Too low spread in the winds will yield too low spread in the waves, knowing that at first order, the wind waves generation is proportional to the square of the wind speed. The significant wave height is an integral parameter that includes all contributions from the different wave components, including the swell components generated from storms across the ocean basins days earlier. The CERA system does not necessarily propagate that information in time and space. For these reasons, it is not surprising that the wave height spread is underestimated.

Instead of only measuring uncertainties, the ensemble spread can also be used to directly estimate the confidence in the data relative to climatology. One possibility is to define the confidence as

$$1 - \frac{\text{ensemble spread}}{\text{climatological spread}}$$

where the ensemble spread is the square root of time-averaged analysis ensemble variance, and the climatological spread is the temporal standard deviation of the CERA-20C ensemble mean over a longer reference period. Positive confidence implies the ensemble provides more information than climatology. The
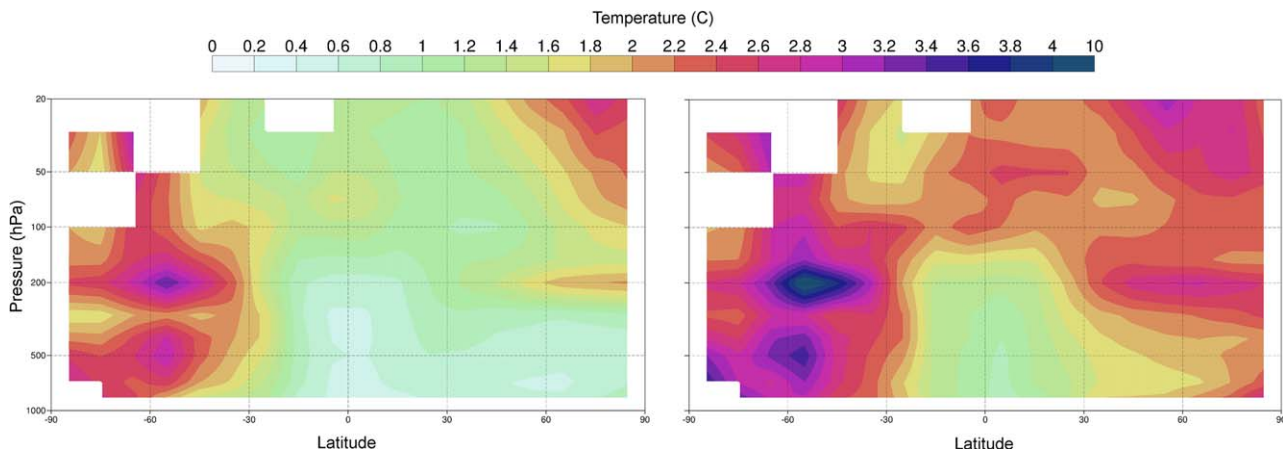
Figure 19. Zonal mean of temperature ensemble spread at (left) radiosondes locations and (right) monthly standard deviation of analysis departure averaged over the period 1959–1960.
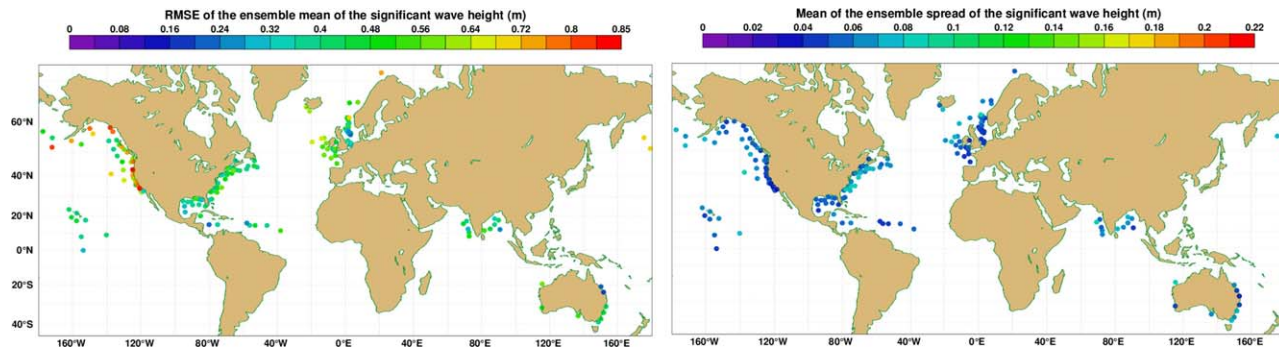
Figure 20. Root-mean-square error of the CERA-20C ensemble mean for significant wave height against all available in situ data (left) and mean of the ensemble spread (right), for a 2 year period (2005–2006).

left plot of Figure 21 shows the 2 m temperature confidence for the JJA period of the year 1959 where the climatology has been computed for the JJA periods over the years 1981–2010. High confidence is located in the northern extra-tropics hemisphere, as well as over Australia and New Zealand. This can be linked to the large number of observations assimilated over these areas (right plot of the Figure 21). Not having observations over the southern ocean produces a confidence close to zero which means that the ensemble spread is equivalent to the climatological spread. The few stations located in Antarctica mainly along the coast play an important role as they are able to increase the reanalysis confidence in their vicinity. This diagnostic identifies regions that have been constrained by observations and where specific weather situations are expected to be recreated. This does not imply necessarily that the true state is contained in the ensemble spread (Figure 18 shows that the ensemble spread is globally too small by a factor of 2 in the atmosphere). The specification of uncertainties in reanalysis will require more work in the future to identify clearly over which regions the users can trust more the data.

7. Future Evolution of the CERA System

The coupled atmosphere-ocean assimilation system used to produce CERA-20C has been recently upgraded to demonstrate the feasibility of a coupled reanalysis over the satellite-era. In this context, the resolution of the CERA system has been increased to a 55 km horizontal grid with 137 levels in the atmosphere and to a quarter of a degree in the ocean. All types of observations including upper-air and satellite measurements are assimilated. Data assimilation is no longer restricted to the ocean and the atmosphere but is also
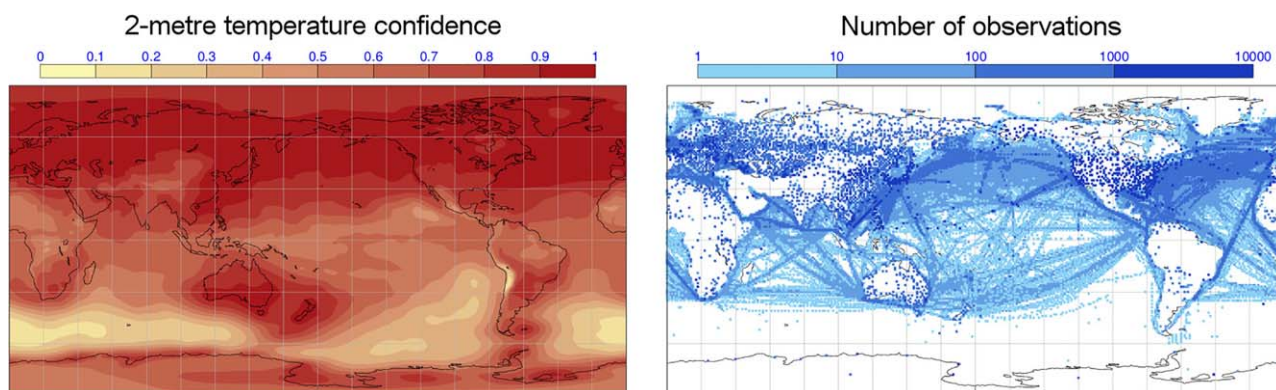
Figure 21. Two meter temperature confidence computed from (left) the CERA-20C ensemble and (right) number of observations assimilated for the JJA period of the year 1959.

performed in the land, wave, and sea-ice components. This proof of concept, called CERA-SAT, will deliver a reanalysis over the period 2008–2016 which will be publicly accessible. The configuration of the atmosphere and ocean waves is based on that of the ensemble component of the ERA5 reanalysis, which at the time of writing was being produced at ECMWF as part of the Copernicus Climate Change Service. Further research and development in coupled data assimilation is being investigated to improve the quality of subsequent versions of the CERA system.

In CERA-20C, the sea surface temperature is nudged toward the HadISST2 analysis. The ocean surface temperature is therefore tightly constrained to this preexisting analysis, and the air-sea interactions in the coupled model are not allowed to evolve freely. This has the advantage of preventing drifts in the coupled model, and makes use of the large amount of effort and expertise that was put into the development of the HadISST2 analysis, but has a number of disadvantages. For instance, the HadISST2 objective analysis has spatially and temporally varying errors which are larger in areas of large data voids, but it is difficult to take these into account in a nudging scheme. The ability to directly assimilate in situ and satellite SST observations within the same framework as the assimilation of the subsurface profile data would address some issues discussed above.

A number of developments would be required to do so however. The bias correction of in situ and satellite data is very important in order not to introduce spurious climate trends due to changes in observing systems. The large amount of work done to provide bias corrected data within the HadISST development could provide the information needed to deal with the historical in situ SST data biases. A variational bias correction scheme for satellite SSTs is also being developed, building on the work done by (Roberts-Jones et al., 2012), in order to make best use of consistently reprocessed satellite data sets such as those described by Merchant et al. (2014). The direct assimilation of such in situ and satellite SST data sets requires careful thought about the diurnal SST signal (which is partially represented in the ocean component of the coupled model) in order to avoid errors in the model's diurnal cycle being spread to the atmosphere. A first approach could be to assimilate data which are representative of the foundation SST (free of diurnal variability). The vertical projection of information from SST and profile data should be made consistent within the profile data assimilation, perhaps using the approach described in Waters et al. (2015) which projects information vertically with length-scales based on the forecast mixed layer depth. Preliminary studies looking at the vertical correlations diagnosed from the ORAS5 ensemble support this approach.

The horizontal spreading of information in both the surface and subsurface of the ocean also requires further development if SST observations are to be assimilated directly. A scheme which could allow spreading of information over large spatial scales when the data are sparse, while making good use of the higher-resolution measurements during the satellite era is the ultimate aim. Developments based on similar ideas to those of Karspeck et al. (2012) whereby empirical orthogonal functions (EOFs) are used to model the large-scale error covariances, applied to surface and sub-surface data are being investigated. These can be used in conjunction with the existing local-scale error covariance model currently used in CERA-20C for sub-surface ocean data.

The use of the EDA technique in the atmospheric assimilation system showed some clear benefits by changing dynamically the weight given to the background depending on the observation density. It is currently being investigated how to extend the EDA technique to the ocean assimilation component. This involves estimating the ocean background error variances from the ensemble and developing methods to deal with sampling limitations. The current diffusion operator used to represent error correlations in the ocean is also being modified to exploit ensemble-based correlation information, taking into account anisotropic and inhomogeneous correlations

The outer-loop coupling implemented in the CERA system is a way to synchronize the information between two otherwise uncoupled DA systems. In this context, observations in one component can correct for errors in another component which provides a consistent state estimate for the coupled model. A set of single observation experiments have been carried out to investigate the timescales at which the synchronization between two components occurs in the CERA system. To look at the potential of a fully coupled assimilation system, this has been compared to a coupled ensemble Kalman filter which uses explicit cross-correlation between the atmosphere and the ocean. The outer loop coupling is effective when the window is long enough that original imbalances in the atmospheric and ocean increments to synchronize within the length

of the assimilation window. On the other hand, we suggest that explicit coupling is preferable for data assimilation systems with short assimilation windows (e.g., 6 h or less).

8. Conclusions

The ERA-CLIM and ERA-CLIM2 projects have initiated an important European program to extend the current global reanalysis capability including data rescue, in order to meet the challenging requirements for climate monitoring, climate research, and the development of climate services. Built upon the experience from the ERA-20C atmospheric climate reconstruction, CERA-20C is the first coupled reanalysis of the twentieth century which includes the atmosphere, ocean, land, waves, and sea ice.

Making optimal use of available observations is key to any successful reanalysis. To this end, a new coupled atmosphere-ocean assimilation system has been implemented to simultaneously ingest atmospheric and ocean observations in the ECMWF coupled Earth system model. Particular attention has been given to the specification of the observation and background error to ensure a consistent analysis across geophysical variables, space and time.

Although CERA-20C struggles to represent tropical cyclones at the beginning of the century, it shows significant improvements in the troposphere compared to ERA-20C and 20CRv2c. CERA-20C improves on the representation of atmosphere-ocean heat fluxes and of mean sea level pressure compared to previous reanalyses. At the same time, there are undesirable discontinuities in ocean heat content and an excessive accumulation of Arctic sea ice. The spread of the ensemble decreases over time as the observational record improves. However, verification suggests that the spread should be larger to give a better indication of the confidence we can have in the reanalysis data.

The observation feedback archive is available for external users and includes in particular the background and analysis departures for each observation processed by CERA-20C. This provides a unique source for detection and correction of biases or other errors in the assimilated observations. The scientific community and data set users also provide feedback and raise important issues that need to be addressed in future reanalyses. This is why reanalysis is an ongoing activity that should never be regarded as completed. New climate reanalyses need to be produced periodically to benefit from the latest updates in the models and data assimilation systems developed for numerical weather prediction. Historical reanalyses, in particular, also benefit from more (and more accurate) observations as historical observations are rescued, digitized, and quality-controlled.

The coupled atmosphere-ocean assimilation system used to produce CERA-20C has been recently upgraded to demonstrate the feasibility of a coupled reanalysis over the satellite-era. This proof of concept called CERA-SAT has delivered a reanalysis over the period 2008–2016 at higher resolution and assimilating all type of observations including upper-air and satellite measurements. The CERA system, originally designed for climate reanalysis, could pave the way for more advanced data assimilation to be used in weather forecasting. CERA-20C enables users to develop weather and climate-sensitive applications in a wide range of fields by providing a longer data set that extends back to the early instrumental record.

Data Access

ECMWF provides access to CERA-20C and other public data sets via a dedicated data portal (http://apps.ecmwf.int/datasets). This consists of a point-and-click web interface used by users to discover the data sets available. A WebAPI is provided so users can programmatically download large amounts of data, which would otherwise require tedious interactions with the web interface. For each data set, there are:

1. instructions on how to download the data efficiently, following the organization of the data set in the MARS archive,
2. documentation explaining how the data set was produced, spatial and temporal resolution and detailed product description, and
3. a link to the conditions of use, which users have to acknowledge before any data download.

References

Allan, R., Brohan, P., Compo, G. P., Stone, R., Luterbacher, J., & Bronnimann, S. (2011). The international atmospheric circulation reconstructions over the earth (ACRE) initiative. *Bulletin of the American Meteorological Society*, *92*(11), 1421–1425.

An, S.-I. (2008). Interannual variations of the tropical ocean instability wave and ENSO. *Journal of Climate*, *21*(15), 3680–3686.

Anderson, E., & Jarvinen, H. (1999). Variational quality control. *Quarterly Journal of the Royal Meteorological Society*, *125*(554), 697–722.

Balmaseda, M. A., Dee, D., Vidard, A., & Anderson, D. L. T. (2007). A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Quarterly Journal of the Royal Meteorological Society*, *133*(622), 167–179.

Balmaseda, M. A., Hernandez, F., Storto, A., Palmer, M., Alves, O., Shi, L., et al. (2015). The ocean reanalyses intercomparison project (ORA-IP). *Journal of Operational Oceanography*, *8*(Supp. 1), s80–s97.

Balmaseda, M. A., Mogensen, K., & Weaver, A. (2013). Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, *139*, 1132–1161.

Balsamo, G., Viterbo, P., Beljaars, A., van den Hurk, B., Betts, M. H. A., & Scipal, K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the ECMWF-IFS. *Journal of Hydrometeorology*, *10*, 623–643.

Bidlot, J.-R. (2017). Twenty-one years of wave forecast verification. *ECMWF Newsletter*, *150*, Reading, UK: ECMWF.

Bidlot, J.-R., Holmes, D. J., Wittmann, P. A., Lalbeharry, R., & Chen, H. S. (2002). Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. *Weather and Forecasting*, *17*(2), 287–310.

Bonavita, M., Holm, E., Isaksen, L., & Fisher, M. (2016). The evolution of the ECMWF hybrid data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 287–303.

Bouillon, S., Maqueda, M. A. M., Legat, V., & Fichefet, T. (2009). An elastic-viscous-plastic sea ice model formulated on arakawa b and c grids. *Ocean Modelling*, *27*(3), 174–184.

Bronnimann, S., Allan, R., Atkinson, C., Buizza, R., Bulygina, O., Dahlgren, P., et al. (2018). Observations for reanalyses. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-17-0229.1

Bronnimann, S., & Luterbacher, J. (2012). Weather and climate extremes during the past 100 years. *Meteorologische Zeitschrift*, *21*(1), 9–11.

Buizza, R., Bronnimann, S., Haimberger, L., Laloyaux, P., Martin, M. J., Fuentes, M., et al. (2018). The EU-FP7 ERA-CLIM2 project contribution to advancing science and production of earth-system climate reanalyses. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-17-0199.1

Chelton, D. B., Esbensen, S. K., Schlax, M. G., Thum, N., Freilich, M. H., Wentz, F. J., et al. (2001). Observations of coupling between surface wind stress and sea surface temperature in the eastern tropical pacific. *Journal of Climate*, *14*(7), 1479–1498.

Compo, G. P., Whitaker, J. S., & Sardeshmukh, P. D. (2006). Feasibility of a 100-year reanalysis using only surface pressure data. *Bulletin of the American Meteorological Society*, *87*(2), 175–190.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28.

Cram, T. A., G. P., Compo, X., Yin, R. J., Allan, C., McColl, R. S., Vose, J. S., et al. (2015). The international surface pressure databank version 2. *Geoscience Data Journal*, *2*(1), 31–46.

de Boisseson, E., Balmaseda, M. A., & Mayer, M. (2017). Ocean heat content variability in an ensemble of twentieth century ocean reanalyses. *Climate Dynamics*, *50*, 3783–3783.

Dee, D. P., & Uppala, S. (2009). Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *135*(644), 1830–1841.

Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, U. A. S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*, 553–597.

Desroziers, G., Berre, L., Chapnik, B., & Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3385–3396.

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., Dunn, R. J. H., et al. (2013). Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *Journal of Geophysical Research: Atmospheres*, *118*, 2098–2118. https://doi.org/10.1002/jgrd.50150

ECMWF (2016). IFS cycle 41R2, *Official documentation*, Reading, UK: ECMWF.

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454.

Good, S. A., Martin, M. J., & Rayner, N. A. (2013). En4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, *118*, 6704–6716.

Gouretski, V., & Reseghetti, F. (2010). On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*(6), 812–833.

Gregow, H., Jylha, K., Makela, H. M., Aalto, J., Manninen, T., Karlsson, P., et al. (2016). Worldwide survey of awareness and needs concerning reanalyses and respondents views on climate services. *Bulletin of the American Meteorological Society*, *97*(8), 1461–1473.

Ham, Y.-G., & Kang, I.-S. (2010). Improvement of seasonal forecasts with inclusion of tropical instability waves on initial conditions. *Climate Dynamics*, *36*, 1277–1290.

Hersbach, H., Bronnimann, S., Haimberger, L., Mayer, M., Villiger, L., Comeaux, J., et al. (2017). The potential value of early (1939–1967) upper-air data in atmospheric climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *143*(704), 1197–1210.

Hersbach, H., & Dee, D. (2016). ERA5 reanalysis is in production, *ECMWF Newsletter 147*, Reading, UK: ECMWF.

Hersbach, H., Peubey, C., Simmons, A., Berrisford, P., Poli, P., & Dee, D. (2015a). ERA-20CM: A twentieth-century atmospheric model ensemble. *Quarterly Journal of the Royal Meteorological Society*, *141*(691), 2350–2375.

Hersbach, H., Poli, P., & Dee, D. (2015b). The observation feedback archive for the ICOADS and ISPD data sets, *ERA Report Series 18*, Reading, UK: ECMWF.

Karspeck, A. R., Kaplan, A., & Sain, S. R. (2012). Bayesian modelling and ensemble reconstruction of mid-scale spatial variability in north Atlantic sea-surface temperatures for 1850–2008. *Quarterly Journal of the Royal Meteorological Society*, *138*(662), 234–248.

Kennedy, J. J. (2014). A review of uncertainty in in situ measurements and data sets of sea surface temperature. *Reviews of Geophysics*, *52*(1), 1–32.

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., & Neumann, C. J. (2010). The international best track archive for climate stewardship (IBTrACS). *Bulletin of the American Meteorological Society*, *91*(3), 363–376.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan Series II*, *93*(1), 5–48.

Komen, G., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., & Janssen, P. (2004). *Dynamics and modelling of ocean waves*. Cambridge, UK: Cambridge University Press.

Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., & Janssen, P. (2016a). A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 65–78.

Laloyaux, P., Thepaut, J.-N., & Dee, D. (2016b). Impact of scatterometer surface wind data in the ECMWF coupled assimilation system. *Monthly Weather Review*, *144*(3), 1203–1217.

Lorenc, A. C., & Hammon, O. (1988). Objective quality control of observations using bayesian methods. theory, and a practical implementation. *Quarterly Journal of the Royal Meteorological Society*, *114*(480), 515–543.

Madec, G. (2008). NEMO ocean engine, *Note du Pole de modélisation 27 ISSN No 1288–1619*, Paris, France: Institut Pierre-Simon Laplace.

Magnusson, L., & Kallen, E. (2013). Factors influencing skill improvements in the ECMWF forecasting system. *Monthly Weather Review*, *141*(9), 3142–3153.

Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from phase 1 of the European space agency climate change initiative (SST CCI). *Geoscience Data Journal*, *1*(2), 179–191.

Messori, G., Caballero, R., & Gaetani, M. (2016). On cold spells in north america and storminess in western Europe. *Geophysical Research Letters*, *43*, 6620–6628. https://doi.org/10.1002/2016GL069392

Mogensen, K., Balmaseda, M., & Weaver, A. (2012b). The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean. Technical Memoranda, 668. Reading, UK: ECMWF.

Mogensen, K., Keeley, S., & Towers, P. (2012a). Coupling of the NEMO and IFS models in a single executable. Technical Memoranda, 673. Reading, UK: ECMWF.

Poli, P., Hersbach, H., Berrisford, P., Dee, D., Simmons, A., & Laloyaux, P. (2015). ERA-20C deterministic, *ERA Report Series 20*, Reading, UK: ECMWF.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., et al. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083–4097.

Poli, P., Hersbach, H., Tan, D., Dee, D., Thepaut, J.-J., Simmons, A., et al. (2013). The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C), *ERA Report Series 14*, Reading, UK: ECMWF.

Roberts-Jones, J., Fiedler, E. K., & Martin, M. J. (2012). Daily, global, high-resolution SST and sea ice reanalysis for 1985–2007 using the Ostia system. *Journal of Climate*, *25*(18), 6215–6232.

Saetra, O., & Bidlot, J.-R. (2004). Potential benefits of using probabilistic forecasts for waves and marine winds based on the ECMWF ensemble prediction system. *Weather and Forecasting*, *19*(4), 673–689.

Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*, 1015–1057.

Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U., et al. (2014). Global gridded precipitation over land: A description of the new gpcc first guess daily product. *Earth System Science Data*, *6*(1), 49–60.

Simmons, A. J., & Hollingsworth, A. (2002). Some aspects of the improvement in skill of numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, *128*(580), 647–677.

Stickler, A., Bronnimann, S., Valente, M. A., Bethke, J., Sterin, A., Jourdain, S., et al. (2014). ERA-CLIM: Historical surface and upper-air data for future reanalyses. *Bulletin of the American Meteorological Society*, *95*(9), 1419–1430.

Tank, A., Zwiers, F., & Zhang, X. (2009). Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. Technical Memoranda WCDMP, *No. 72*, WMO.

Tavolato, C., & Isaksen, L. (2015). On the use of a huber norm for observation quality control in the ECMWF 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1514–1527.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498.

Titchner, H. A., & Rayner, N. A. (2014). The Met Office Hadley centre sea ice and sea surface temperature data set, version 2: 1. sea ice concentrations. *Journal of Geophysical Research: Atmospheres*, *119*, 2864–2889. https://doi.org/10.1002/2013JD020316

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., et al. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, *131*, 2961–3012.

Waters, J., Lea, D. J., Martin, M. J., Mirouze, I., Weaver, A., & While, J. (2015). Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 333–349.

Willett, C. S., Leben, R. R., & Lavin, M. F. (2006). Eddies and tropical instability waves in the eastern tropical pacific: A review. *Progress in Oceanography*, *69*(2), 218–238.

Wolter, K., & Timlin, M. S. (2011). El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index. *International Journal of Climatology*, *31*(7), 1074–1087.

Woodruff, S. D., Worley, S. J., Lubker, S. J., Ji, Z., Eric Freeman, J., Berry, D. I., et al. (2011). ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *International Journal of Climatology*, *31*(7), 951–967.