

RESEARCH ARTICLE

The ERA5 global reanalysis

Hans Hersbach¹  | Bill Bell¹ | Paul Berrisford¹ | Shoji Hirahara² | Andr s Hor nyi¹ | Joaqu n Mu oz-Sabater¹  | Julien Nicolas¹  | Carole Peubey¹ | Raluca Radu¹ | Dinand Schepers¹  | Adrian Simmons¹  | Cornel Soci¹ | Saleh Abdalla¹  | Xavier Abellan¹  | Gianpaolo Balsamo¹  | Peter Bechtold¹  | Gionata Biavati¹  | Jean Bidlot¹  | Massimo Bonavita¹ | Giovanna De Chiara¹  | Per Dahlgren³ | Dick Dee¹  | Michail Diamantakis¹  | Rossana Dragani¹ | Johannes Flemming¹  | Richard Forbes¹  | Manuel Fuentes¹  | Alan Geer¹ | Leo Haimberger⁴  | Sean Healy¹  | Robin J. Hogan¹  | El as H lm¹ | Marta Janiskov ¹  | Sarah Keeley¹  | Patrick Laloyaux¹  | Philippe Lopez¹ | Cristina Lupu¹  | Gabor Radnoti¹ | Patricia de Rosnay¹  | Iryna Rozum¹ | Freja Vamborg¹  | Sebastien Villaume¹  | Jean-No l Th paut¹ 

¹European Centre for Medium-Range
Weather Forecasts, Reading, UK

²Japan Meteorological Agency, Tokyo,
Japan

³The Norwegian Meteorological Institute,
Oslo, Norway

⁴Department of Meteorology and
Geophysics, Universit t Wien, Vienna,
Austria

Correspondence

H. Hersbach, ECMWF, Shinfield Park,
Reading RG2 9AX, Reading, UK.
Email: hans.hersbach@ecmwf.int

Funding information

European Union through the Copernicus
Climate Change Service

Abstract

Within the Copernicus Climate Change Service (C3S), ECMWF is producing the ERA5 reanalysis which, once completed, will embody a detailed record of the global atmosphere, land surface and ocean waves from 1950 onwards. This new reanalysis replaces the ERA-Interim reanalysis (spanning 1979 onwards) which was started in 2006. ERA5 is based on the Integrated Forecasting System (IFS) Cy41r2 which was operational in 2016. ERA5 thus benefits from a decade of developments in model physics, core dynamics and data assimilation. In addition to a significantly enhanced horizontal resolution of 31 km, compared to 80 km for ERA-Interim, ERA5 has hourly output throughout, and an uncertainty estimate from an ensemble (3-hourly at half the horizontal resolution). This paper describes the general set-up of ERA5, as well as a basic evaluation of characteristics and performance, with a focus on the dataset from 1979 onwards which is currently publicly available. Re-forecasts from ERA5 analyses show a gain of up to one day in skill with respect to ERA-Interim. Comparison with radiosonde and PILOT data prior to assimilation shows an improved fit for temperature, wind and humidity in the troposphere, but not the stratosphere. A comparison with independent buoy data shows a much improved fit for ocean wave height. The uncertainty estimate reflects the evolution of the observing systems used in ERA5. The enhanced temporal and spatial resolution allows for a detailed evolution of weather systems. For precipitation, global-mean correlation with monthly-mean GPCP data is increased from 67%

to 77%. In general, low-frequency variability is found to be well represented and from 10 hPa downwards general patterns of anomalies in temperature match those from the ERA-Interim, MERRA-2 and JRA-55 reanalyses.

KEYWORDS

climate reanalysis, Copernicus Climate Change Service, data assimilation, ERA5, historical observations

1 | INTRODUCTION

The role of reanalyses in climate monitoring applications is now widely recognized. ECMWF's ERA-Interim reanalysis (Dee *et al.*, 2011) has been routinely used, together with other datasets, as input to the WMO annual assessment of the State of the Climate and in the assessments carried out by the IPCC¹. ERA-Interim and the earlier ERA-40 reanalysis (Uppala *et al.*, 2005) have also been a resource for the production of ECVs (Bojinski *et al.*, 2014)) and Climate Indicators recommended by the GCOS. By optimally combining observations and models, reanalyses indeed provide consistent “maps without gaps” of ECVs and strive to ensure integrity and coherence in the representation of the main Earth system cycles (e.g., water, energy).

Reanalyses have found a wide application in atmospheric sciences, not least in operational weather centres where, for example, reanalyses are used to assess the impact of observing system changes, to gauge progress in modelling and assimilation capabilities, and to obtain state-of-the-art climatologies to evaluate forecast-error anomalies.

ECMWF has a long history with reanalysis, as indicated in Table 1, and ERA5 is the fifth generation of atmospheric reanalysis to be produced. Activities on atmospheric reanalysis started in 1979 with the FGGE project, followed by the production of ERA-15 in the mid 1990s, ERA-40 from 2001 to 2003, and ERA-Interim from 2006 to 2019. Successive atmospheric reanalyses have typically offered higher horizontal resolution, more sophisticated DA schemes and have benefitted from the continuous development of forecast models. All include a land component and, from ERA-40 onwards, ocean surface wave and atmospheric ozone products have also been included. Beyond ECMWF, several groups produce global atmospheric reanalyses and the most recent products are the MERRA-2 reanalysis (Gelaro *et al.*, 2017) from the NASA GMAO, JRA-55 (Kobayashi *et al.*, 2015) produced by the JMA and CFSR (version 2) produced by NCEP (Saha *et al.*, 2014).

In recent years, ECMWF has also undertaken the systematic production of ocean reanalyses culminating in the most recent, ORAS5. Although originally these ocean reanalyses were passive applications of the atmospheric reanalyses in which the latter provided the atmospheric forcing for the former, ORAS5 went a step further. For the first time, the selection of consistent records for SST and SIC took into account the needs for atmospheric reanalyses (Hirahara *et al.*, 2016). Hence, the SST records are common to both ORAS5 and ERA5.

Within the European Commission-funded GEMS and MACC projects, and now the Copernicus Atmosphere Monitoring Service, ECMWF has also produced reanalyses of atmospheric composition in 2008, 2010 and 2018, respectively (Table 1).

Preceding and in parallel with the development of ERA5, ECMWF has produced centennial reanalyses within the EC-funded ERA-CLIM and ERA-CLIM2 research projects, involving international consortia and coordinated by ECMWF. One aspect was the provision of boundary datasets over the oceans (SST and SIC provided by the Met Office Hadley Centre) and forcing terms in radiation (from CMIP5) to provide a good representation of their evolution over the 20th century. This was successfully implemented in ERA-20CM, a century-long ten-member ensemble model-only integration. The next step was the production of a century-long reanalysis, ERA-20C, using surface pressure and marine wind observations only (Poli *et al.* (2016)). The production of such a model-based reanalysis that extends back more than one century was first pursued in the 20CR Project (Compo *et al.*, 2006) at NOAA's Earth Systems Research Laboratory, where a reanalysis spanning the period 1871–2010 was conducted (Compo *et al.*, 2011). An experimental reanalysis from 1939 to 1967 using in addition upper-air temperature and wind was also produced (Hersbach *et al.*, 2017). In the ERA-CLIM2 project, research towards coupling with the ocean culminated in a century-long ten-member reanalysis (CERA-20C) which is based on outer-loop coupling between the ocean and atmosphere, and an

¹All acronyms are defined in the Glossary (Table 8).

TABLE 1 ECMWF reanalyses

Reanalysis	Period covered	Grid resolution	Assimilation scheme	IFS model cycle (year)	Reference
<i>Atmospheric reanalyses</i>					
FGGE	1979	208 km	OI	(1980)	Bengtsson <i>et al.</i> (1982)
ERA-15	1979–1994	125 km	OI	13r4 (1995)	Gibson <i>et al.</i> (1999)
ERA-40	1957–2002	125 km	3D-Var	23r4 (2001)	Uppala <i>et al.</i> (2005)
ERA-Interim	1979–2019	80 km	4D-Var	31r2 (2006)	Dee <i>et al.</i> (2011)
ERA5	1950–present	31 km	4D-Var	41r2 (2016)	This paper
<i>Ocean reanalyses</i>					
ORAS3	1959–2012	1.0°	OI	(2006)	Balmaseda <i>et al.</i> (2008)
ORAS4	1959–2018	1.0°	3DVar-FGAT	(2010)	Balmaseda <i>et al.</i> (2013)
ORAS5	1979–present	0.25°	3DVar-FGAT	(2016)	Zuo <i>et al.</i> (2018)
<i>Atmospheric composition reanalyses</i>					
MACC	2003–2012	80 km	4D-Var	36r1 (2010)	Inness <i>et al.</i> (2013)
CAMS-Interim	2003–2018	110 km	4D-Var	40r2 (2014)	Flemming <i>et al.</i> (2017)
CAMS	2003–present	80 km	4D-Var	42r1 (2016)	Inness <i>et al.</i> (2019)
<i>Centennial reanalyses and model-only climate integrations</i>					
ERA-20CM	1899–2010	125 km	4D-Var	38r1 (2012)	Hersbach <i>et al.</i> (2015)
ERA-20C	1900–2010	125 km	4D-Var	38r1 (2012)	Poli <i>et al.</i> (2016)
CERA-20C	1901–2010	125 km	4D-Var	41r2 (2016)	Laloyaux <i>et al.</i> (2018)
CERA-SAT	2008–2016	65 km	4D-Var	42r1 (2016)	Schepers <i>et al.</i> (2018)

eight-year reanalysis (CERA-SAT) for the current-day full observing system with the same resolution as the ensemble component of ERA5. Collectively, these developments laid the foundations for future coupled reanalyses at ECMWF.

The developments outlined above also allow the extension of ERA5 further back in time (to 1950) than ERA-Interim, though the focus for this paper is the period common to both, 1979–2019. ERA5 replaces the very popular ERA-Interim reanalysis, which was progressively becoming outdated and was stopped at the end of August 2019. ERA5 is a highly visible activity within the Copernicus Climate Change Service (C3S; Thépaut *et al.*, 2018), where it provides an improved and consistent record for a large number of ECVs for the C3S Climate Data Store (CDS; Raoult *et al.*, 2017). Besides a considerable increase in resolution (both in the horizontal and vertical) and the benefit of 10 years of model and data assimilation (DA) developments, ERA5 provides an enhanced number of output parameters (such as the 100 m wind product), hourly high-resolution output throughout and 3-hourly

uncertainty information. This uncertainty information is obtained from the underlying ten-member ensemble 4D-Var DA system.

Another innovative aspect is a timely, preliminary product that is available within 5 days of real time. It is replaced by a more thoroughly quality-checked final product two months later. In practice, though, it is expected that both products will rarely differ, and in case they do (due to considerable errors found in the early release), users will be notified.

The step forward with ERA5 is illustrated by Figure 1, which shows a gain of up to one day in skill of re-forecasts started from ERA5 (thick lines) analyses using the ERA5 model, compared to the re-forecasts run using the ERA-Interim system (thin lines). The distinct improvement originates from a better forecast model (which is an integral part of the assimilation system as well) and in particular the improved analyses from which these forecasts are started.

This paper provides an overview of the configuration of ERA5 and a basic description of its characteristics

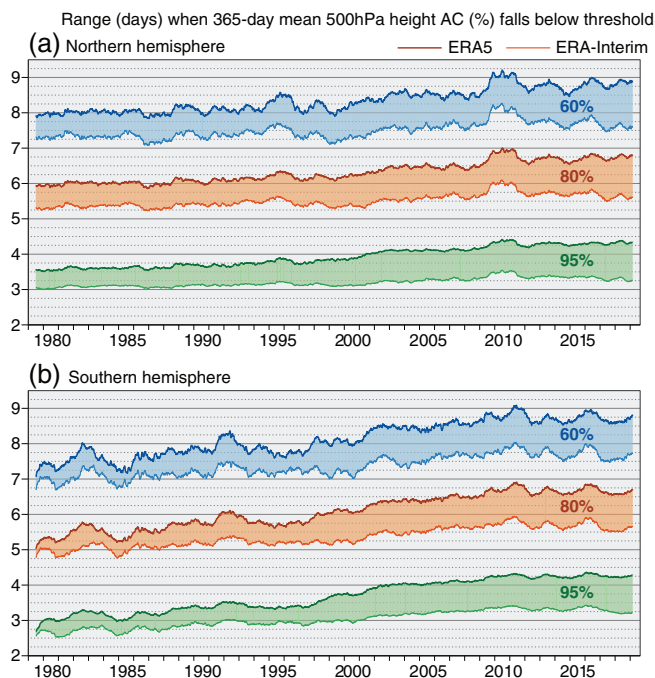


FIGURE 1 Range (days) at which running 365-day mean anomaly correlations of 500 hPa height forecasts from 0000 and 1200 UTC reach 95% (green), 80% (orange) and 60% (blue), for (a) the extratropical Northern and (b) Southern Hemispheres, from 1979 onwards. The heaviest lines denote ERA5, and the thin lines ERA-Interim. Shading denotes the difference between ERA5 and ERA-Interim. The improvement of ERA5 with respect to ERA-Interim is thought to be dominated by a better analysis at short forecast ranges, while towards longer ranges the improved forecast model will also contribute

and performance. In general, results cover January 1979 to August 2019, which corresponds with the period for which ERA-Interim is available for comparison. The structure is as follows. The general set-up is described in Section 2. Details on the ERA5 production and data access are provided in Section 3. A concise summary of the model improvements that took place between the ECMWF IFS cycle releases between ERA-Interim and ERA5 is given in Section 4. Section 5 details the observations which were assimilated in ERA5, a considerable number of which originate from reprocessed datasets. Section 6 provides information on the ingested SST, sea ice, and external forcing applied in the radiation scheme. The paper continues with some diagnostics from ensemble spread, departure statistics and analysis increments (Section 7). A basic assessment of characteristics is made in Section 8 (weather) and Section 9 (climate). This includes an intercomparison with the global reanalyses MERRA-2 and JRA-55. Section 10 summarizes strengths and weaknesses of ERA5 and ends with concluding remarks.

2 | ERA5 CONFIGURATION

2.1 | General overview

This section provides an overview of the configuration of ERA5. A summary is displayed in Table 2 which includes a comparison with ERA-Interim.

The starting point for ERA5 is IFS Cy41r2, which was used in the ECMWF operational medium-range forecasting system from 8 March to 21 November 2016. With respect to Cy31r2 on which ERA-Interim is based, this incorporates 10 years of R&D for all its components (atmosphere, land, ocean waves, observation operators and additional observations; Sections 4 and 5) and improvements in the DA methodology, which is now based on a hybrid incremental 4D-Var system (Bonavita *et al.*, 2016). ERA5 contains an ensemble component (EDA; Isaksen *et al.*, 2010) of one control and nine perturbed members which provide background-error estimates for the deterministic HRES DA system. The EDA system provides estimates of analysis and short-range forecast uncertainty which are considered to represent the evolution of the errors in the HRES system. This allows for the estimation of uncertainties in the reanalysis products. A concise assessment of this innovative feature is provided in Section 7.1. Details of the EDA system are deferred to Section 2.4. Before then, focus is on the HRES system.

Although ozone is part of the Earth-system component atmosphere, for technical reasons in this paper the term atmosphere will be mainly used to denote the dynamical and moist thermodynamical state of the atmosphere. The specifics of the ozone analysis are discussed separately.

The assimilation system makes use of 12-hourly windows in which observations are used from 0900 to 2100 UTC (inclusive) and from 2100 to 0900 UTC (inclusive) the next day. The resulting analysis fields follow the time evolution within the window and are stored hourly. Information gathered within each analysis window is transported by a short forecast initiated from the analysis fields 9 hr into the window, that is, from 1800 and 0600 UTC, where it provides the starting point (first guess) for the next assimilation. In these forecasts (which are also archived hourly) all ERA5 components are coupled. The atmosphere generates ocean waves through the surface wind stress, while the waves influence the atmospheric boundary layer via sea-state dependencies in the surface roughness. A two-way interaction also exists between the atmosphere and land (Section 4.3).

The interaction between ozone and the (rest of the) atmosphere is one-way. Ozone is advected by the atmospheric flow. In addition, the prognostic ozone model includes the representation of the stratospheric ozone chemistry based on the Cariolle and Teyss  dre (2007)

TABLE 2 Overview of characteristics of ERA5

	ERA-Interim	ERA5
Publicly available now	1979 – August 2019	1979 onwards
Expected in 2020		1950 – 1978
Availability behind real time	2–3 months	2–3 months (final product)
		5 days (preliminary product)
Model cycle (year)	Cy31r2 (2006)	Cy41r2 (2016)
Atmospheric DA	12 hr 4D-Var	12 hr 4D-Var ensemble
Window for 0000, 1200 UTC	(15 _{day-1} , 03], (03,15]	(21 _{day-1} , 09], (09, 21]
Model input	As in ERA-40,	Appropriate for climate,
(radiation and surface)	inconsistent SST	e.g., evolution greenhouse gases,
	and sea ice	aerosols, SST and sea ice
Spatial resolution	79 kms (TL255)	31 km (TL639), HRES
	60 levels to 10 Pa	137 levels to 1 Pa
Ocean waves	1°	0.36°
Inner-loop resolution	TL95, TL159	TL95, TL159, TL255
Land-surface model	TESSEL	HTESSEL
Soil moisture DA	1D-OI	SEKF
Snow DA	Cressman	2D-OI
Uncertainty estimate	None	From the 4D-Var ensemble, EDA
		10 members at 63 km (TL319),
		ocean waves 1°,
		TL127, TL159 inner loops
Output frequency	6-hourly for analyses,	Hourly throughout
	3-hourly for forecasts	(uncertainty 3-hourly)
Output parameters	84 (sfc), 25 (wave), 27 (ua)	205 (sfc), 46 (wave), 30 (ua)
Extra observations	Following ERA-40, GTS	In addition, latest instruments
Reprocessed FCDRs	Some	Many more (Table 4)
Radiative transfer model	RTTOV v7	RTTOV v11
	Clear-sky assimilation	Partly all-sky assimilation
VarBC	Radiances only	Also ozone, ground-based radar–gauge composites,
		aircraft temperature, surface pressure
Radiosonde corrections	RAOBCORE	RICH
Other corrections	scatterometer, altimeter	scatterometer, altimeter

Note: DA=Data Assimilation; sfc=surface; ua=upper air.

parametrization scheme in which the time evolution is expressed as a linear expansion with respect to the photochemical equilibrium for the local value of the ozone mass mixing ratio, the local overhead ozone column, the local temperature and an additional term for the rapid depletion associated with the emergence of the ozone hole. However, the ERA5 prognostic ozone has no feedback on the

atmosphere via the radiation scheme. Diagnostic ozone is used instead (Section 6.1).

The HRES assimilation system contains two main components: incremental 4D-Var (Courtier *et al.*, 1994) for the atmosphere plus ozone, and the land DA (LDAS). The interaction between these is an example of weak coupling (Penny *et al.*, 2017), where the influence from the

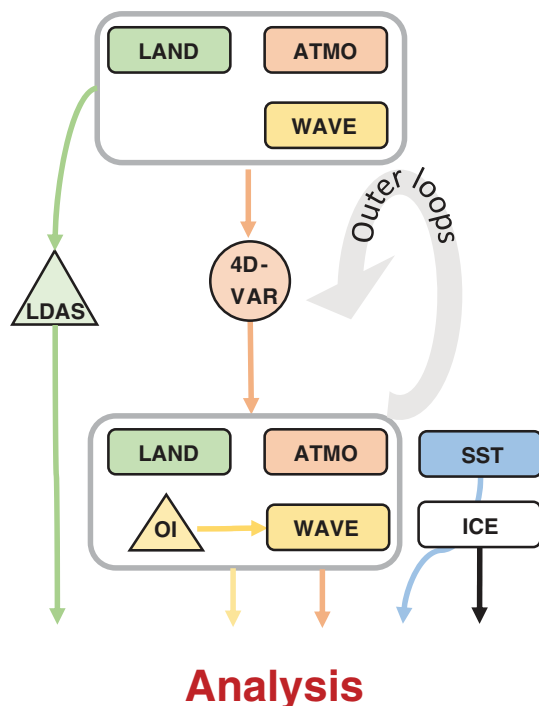


FIGURE 2 Assimilation diagram for ERA-Interim and ERA5 regarding the atmosphere including ozone (ATMO), land surface (LAND), ocean waves (WAVE), sea surface temperature (SST) and sea ice (ICE). Large boxes represent outer-loop integrations (trajectories) where the indicated domains are coupled. Triangles represent the land-data assimilation (LDAS) and ocean wave optimal interpolation (OI), while circles correspond to 4D-Var inner loops. The ocean wave OI assimilation is performed only inside the final trajectory. For ERA5 the LDAS assimilation is an example of weak coupling (Penny *et al.*, 2017) where the influence from land surface and other observations is only mixed in the next analysis window via the coupled short forecast from the current analysis

land surface and other observations is only mixed in the next analysis via the coupled short forecast from the resulting sub-analyses. A graphical representation is provided in Figure 2. The incremental formulation involves the step-wise minimization of a linearized quadratic 4D-Var cost function at reduced resolution in inner loops, with nonlinear updates at full resolution in outer loops (trajectories). These latter also involve the integration of the coupled model over the length of the assimilation window. ERA5 HRES uses three inner loops, and its ensemble component two inner loops. The ocean wave analysis, which is based on optimal interpolation (OI) is performed in the final trajectory of 4D-Var. In order to match the hourly output of ERA5, this OI is now performed hourly, rather than 6-hourly in ERA-Interim.

The LDAS consists of a number of sub-steps which are detailed in Section 2.3. Information from SST and SIC is obtained from external level-4 sources (i.e., gridded and

complete). Their ingestion includes a minor interpolation step which involves regridding onto the ERA5 model grid, subject to its land-sea mask and some cross-checks between these two quantities.

2.2 | 4D-Var including variational bias correction (VarBC)

The objective of 4D-Var is to find the best estimate of the state of the atmosphere within an assimilation time window, given a background forecast $\mathbf{x}^b$ valid at the start of the window and observations $\mathbf{y}^o$ falling within that window. Here, approximately following the unified notation proposed by Ide *et al.* (1997), any $\mathbf{x}$ contains the model data at one particular time across all locations, levels and variables, and any $\mathbf{y}$ contains all the available observations in the window. The aim is to reduce the misfit $\mathbf{d}$ between the observations and their modelled equivalents $\mathbf{y}$, that is,

$$\mathbf{d} = \mathbf{y}^o - \mathbf{y}, \quad (1)$$

consistent with the estimated uncertainty of the background and observations. This is done by adjusting the state of the atmosphere at the start of the assimilation window $\mathbf{x}$ and, recognising the possibility of bias in observations, by adjusting a vector of parameters $\boldsymbol{\beta}$ that describe such biases b in observation space. Hence the vector of simulated observations is computed as

$$\mathbf{y} = H(\mathbf{x}) + b(\mathbf{x}, \boldsymbol{\beta}). \quad (2)$$

In 4D-Var the observation operator $H()$ and bias model $b()$ include the model integration from the start of the window to the time of the observation, as well as interpolation to the observation location and simulating of the observed quantity (such as radiance) from the model state. The bias parameter vector consists of a large number of small subsets (containing between 1 and 12 elements), each of which determines the bias estimate for a particular sub-group of observations, a ‘bias group’. Bias groups contain anything from a handful to tens of thousands of observations, collecting for example all the data from one aircraft flight or from one channel of one satellite instrument on one satellite. In ERA5 all bias models are linear (Dee, 2005), that is,

$$b_i(\mathbf{x}, \boldsymbol{\beta}) = \sum_{j \in S} p_j(\mathbf{x}) \beta_j, \quad (3)$$

with p_j the linear predictors and where the sum is limited to the small subset S of bias parameters that relates to the bias group to which one observation i belongs.

Each observation enters either exactly one bias group, or none at all, in which case $S = \emptyset$ and no bias correction is applied. Such latter observations are also called anchors. The simplest predictor is a constant, while others can depend on characteristics of the observations and/or the model state at the observation location and time. ERA-Interim was the first ECMWF reanalysis to include variational bias corrections (Dee and Uppala, 2009). It was applied to radiance data. In ERA5 this has been extended to ground-based radar–gauge composites, total column ozone, aircraft temperature and surface pressure observations. The bias correction for radiosonde temperatures, scatterometer backscatter and altimeter wave height is prescribed independently of the DA.

4D-Var minimizes the cost function

$$J(\delta\mathbf{x}, \delta\boldsymbol{\beta}) = J_b(\delta\mathbf{x}) + J_p(\delta\boldsymbol{\beta}) + J_o(\mathbf{d}) = \frac{1}{2}\delta\mathbf{x}^T \mathbf{B}^{-1} \delta\mathbf{x} + \frac{1}{2}\delta\boldsymbol{\beta}^T \mathbf{B}_\beta^{-1} \delta\boldsymbol{\beta} + \frac{1}{2}\mathbf{d}^T \mathbf{R}^{-1} \mathbf{d}. \quad (4)$$

Here $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b$ and $\delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}^b$ are the deviation, or increment from the model and bias parameter backgrounds $\mathbf{x}^b$ and $\boldsymbol{\beta}^b$. T is the transpose operator. While $\mathbf{x}^b$ is the result of a model integration starting from the previous analysis, there is no bias evolution model, so $\boldsymbol{\beta}^b$ are simply the final values from the previous analysis. Covariance matrices $\mathbf{B}$, $\mathbf{B}_\beta$ and $\mathbf{R}$ express the second-order moment error characteristics in the background model state, bias parameter and (bias-adjusted) observations, respectively. For simplicity in Equation (4), one term with minor effect, J_c , has been omitted; this is a digital filter for reducing gravity waves in the increments (Gauthier and Thépaut, 2001).

With the exception of some satellite radiances, correlations between observation errors are neglected, that is, $\mathbf{R}$ is mostly diagonal. In general, ERA5 uses the observation errors described in Part I of ECMWF R&D (2016). Correlations between bias parameters are also neglected, that is, $\mathbf{B}_\beta$ is diagonal. The diagonal elements model the rate at which bias parameters are allowed to change in one assimilation cycle. For most satellite data, where each bias group typically contains thousands of radiances, the weight is chosen to be that of 5,000 observations. Weights for other bias parameters vary. Off-diagonal terms in $\mathbf{B}$ describe the correlation length-scales in the model background-error estimate, as well as correlations across variables (such as between wind and temperature). At ECMWF, $\mathbf{B}$ is modelled using the wavelet formulation described by Fisher (2003), where background-error correlations are localized in both the spectral and the spatial domains. It is evolved dynamically by the underlying ERA5 ensemble (Section 2.4).

As mentioned above, the incremental formulation of 4D-Var solves Equation (4) through a series of linearised

quadratic cost functions known as inner loops. Here minimizations are performed where in Equation (4), $\mathbf{d}$ is approximated by the Taylor expansion:

$$\mathbf{d} \simeq \mathbf{d}^n - \mathbf{H}(\mathbf{x} - \mathbf{x}^n) - \mathbf{P}\delta\boldsymbol{\beta}, \quad (5)$$

with $\mathbf{d}^n = \mathbf{y}^o - \mathbf{H}(\mathbf{x}^n) - \mathbf{b}(\mathbf{x}^n, \boldsymbol{\beta}^b)$, the nonlinear departure based on the previous trajectory started at $\mathbf{x}^n, \boldsymbol{\beta}^b$ where $\mathbf{x}^n$ is the updated estimate from the previous minimization. Matrix $\mathbf{H} = \partial\mathbf{H}/\partial\mathbf{x}$ represents a linearization of the observation operator and model equations around the most recent trajectory (indexed n) started from $\mathbf{x}^n$, performed at lower inner-loop resolution (Table 2). $\mathbf{P} = \partial\mathbf{b}/\partial\boldsymbol{\beta}$ is a matrix representation of the VarBC predictors from Equation (3). In expansion (5), the dependency of $\mathbf{x}$ on $\mathbf{b}$ has been neglected. Note that, in contrast to the model state, the formulation for the bias parameters is not incremental, that is, the optimized increments are only added in the final trajectory which is conducted after the final minimization.

The first-guess departure is $\mathbf{d}^b$ from the background trajectory, while the analysis departure $\mathbf{d}^a$ is that from the final trajectory. In some implementations of 4D-Var, the terms background and first guess have different meanings, but here they are interchangeable. So in the remainder of the paper we refer to $\mathbf{d}^b$ as the first-guess departure. Note that both $\mathbf{d}^b$ and $\mathbf{d}^a$ include the estimated bias corrections.

Observations in 4D-Var are subject to a range of quality controls. This includes the *a priori* blacklisting of data known to be of poor quality, and a check on the size of the first-guess departure $\mathbf{d}^b$ (relative to its expected value, which is estimated from the observation and background errors), both of which are applied in the first outer loop prior to the assimilation. In addition, observations can be downgraded to have a much reduced weight inside the minimization, using the method of variational quality control (VarQC). This procedure is applied from the second minimization. In ERA5 it is based on the more robust Huber norm (Tavolato and Isaksen, 2015) for conventional data, while for satellite radiances the method of gross error (Anderson and Järvinen, 1999), as used for all observations in ERA-Interim, is still applied.

The blacklist contains rules that vary from the exclusion of certain channels on certain satellites, particular data over specific regions like land, sea or ice, entire instruments for periods of known anomalies, and particular observables from *in situ* observations (like 10 m wind over land). For ERA5 this is a blend between the blacklists of ERA-Interim (details in Dee *et al.*, 2011) and the ECMWF operational system (for more recent periods), augmented with rules for newly ingested reprocessed data. In addition, some extra exclusions were added, such as the blacklisting of stuck pressure sensors (section 3.4.2 of Hersbach *et al.*,

2018) and, based on experience in ERA-Interim, the exclusion of the HIRS instrument on NOAA-18 (as shown later in Figure 5).

The VarBC coefficients β are initialized by either zero or based on the most populated bin of first-guess departures $\mathbf{d}^b$. In addition, there is the possibility to spin up their estimates over a desired period, prior to actively using the data within their corresponding bias groups. This facility is used for a large number of satellite data.

The 4D-Var cost function (4) is minimized for the atmosphere and ozone simultaneously. However, the background term $\mathbf{B}$ is univariate for ozone, that is, there are no cross-correlations with any other parts of the control vector. In addition, changes in ozone advection due to wind increments at the beginning of the assimilation window or changes in ozone due to temperature perturbations are neglected in the inner-loop minimizations. The zeroing of the first dependency, which was already in place for ERA-Interim, prevents anomalously large wind increments in the stratosphere due to systematic model bias. Note that updates in wind and temperature in the outer loops do have an effect on ozone. Therefore, although reduced, the atmospheric analysis does influence the ozone analysis. On the other hand, ozone influences the assimilation of other atmospheric quantities, via the observation operator for a number of satellite channels that are sensitive to both ozone and temperature or humidity. Examples of these are the HIRS instrument (channel 9), and specific channels for the hyperspectral radiances from IASI, AIRS and CrIS. The provision of improved estimates for these observation operators was one of the drivers to implement a prognostic ozone scheme at ECMWF. Only when no such observations are available (as in the pre-satellite era before the early 1970s) does ozone have no effect on the rest of the atmospheric reanalysis.

2.3 | Land data assimilation (LDAS)

ERA5 includes an advanced land DA system to analyze land surface prognostic variables (de Rosnay *et al.*, 2014). It is weakly coupled with 4D-Var (Figure 2). First-guess values are provided by the short forecasts that originate from the previous LDAS and upper-air 4D-Var assimilation. However, the LDAS and 4D-Var assimilation are produced separately and their influence is only combined towards the next assimilation. It consists of the following components:

1. A two-dimensional optimal interpolation (2D-OI) scheme for the analysis of screen-level 2 m temperature

and relative humidity, and for snow (depth and density),

2. A point-wise Simplified Extended Kalman Filter (SEKF; de Rosnay *et al.*, 2013) for three soil moisture layers in the top 1 m of soil, and
3. A one-dimensional OI for soil, ice and snow temperature.

These analyses only use data where the land-sea mask is 50% or higher (i.e., including many islands). Over sea, where 2 m temperature and humidity are purely diagnostic fields in the LDAS (i.e., they do not influence any subsequent results), increments are set to zero, so the analysis equals the first-guess. This is an improvement on ERA-Interim and avoids the incorrect spill-over of increments from local land observations into the oceans. Over land, these parameters are fed into the soil-moisture analysis, and do cycle information forward in time.

For soil moisture, the ERA5 hourly analysis products follow the temporal evolution of the SEKF within the 12 hr window. Each OI analysis (components 1. and 3. above) is based on 6-hr sub-windows along the 12 hr assimilation windows of 4D-Var, each only providing an analysis at the central time of its ± 3 hr window, without making any corrections for misfits in timing. To accommodate an hourly ERA5 product, these OI analyses are now performed hourly, that is, 12 times in the LDAS analysis step, each retaining a ± 3 hr observation window. Many observations are three-hourly (at 0000, 0300, ..., 2100 UTC). As a result of this, analyses at other hours of the day may suffer from systematic biases. On the other hand, observations available at other times, such as some observations from Australia, can introduce systematic biases for analyses at, for example, 0000 and 1200 UTC. Note that a particular observation can be used in up to six OI analyses. There is no conflict here, since up to only one of them (the 0600 or 1800 UTC analysis) is used for the transport of information towards the next analysis window.

The SST and SIC products are stored hourly as well, although their values only change once daily in line with the temporal resolution of the ingested level-4 products.

2.4 | The Ensemble of Data Assimilation (EDA) systems

The ensemble component of ERA5 is an EDA of 10 members which provides background-error estimates for the deterministic HRES 4D-Var DA system.

The analysis method is the same for each EDA member and follows that of the HRES as displayed in Figure 2. In particular, each member makes use of the flow-dependent

B matrix as determined by the ensemble as a whole. However, resolution is lower (TL319, compared to TL639 for the HRES), and two inner loops are performed, rather than three. Details are given in Table 2. In addition, each member (except the control) is run with different random perturbations added to the observations. The perturbations of observations are sampled from a zero-mean Gaussian distribution with variance equal to the expected variances of the observation errors. Likewise, the model physical tendencies are perturbed (Leutbecher *et al.*, 2017) in the short forecasts that link subsequent analysis windows. Perturbations in SST and SIC are taken from the spread within the range of available products (Hirahara *et al.*, 2016). The perturbations applied to the observations, the SST, SIC and the model imply that the resulting background (i.e., short-range forecast) of each member is implicitly perturbed, thus avoiding the need for explicitly perturbing the background fields. It can be shown (Isaksen *et al.*, 2010) that in a weakly nonlinear environment the combination of samples of perturbed short-range forecasts provides a good flow-dependent error estimate for **B** when the ensemble size is sufficiently large.

However, the ensemble size in ERA5 is quite limited. Partly for this reason, the EDA and the HRES 4D-Var make use of a hybrid **B** formulation (Bonavita *et al.*, 2016). This means that, as for the operational medium-range system, a static, climatological background-error covariance matrix (**B**_{cli}) is combined with a dynamic one computed using short-range forecasts from the EDA:

$$\mathbf{B} = (1 - \alpha)\mathbf{B}_{\text{cli}} + \alpha\mathbf{B}_{\text{EDA}}. \quad (6)$$

This latter brings in flow-dependent correlation structures to the resulting background-error covariance matrix. The weight α increases with wavenumber, from 0.15 for the largest to 0.74 for the smallest scales. These weights are lower than used in the Cy41r2 operational NWP system (0.3 to 0.93), in order to limit sampling errors from the smaller ensemble size (10 versus 25). The vertical profiles of the global average ensemble standard deviation for model state variables are determined by the forecast errors of the day from the ensemble members.

For data produced prior to March 2017, α had been kept constant (0.15) inadvertently for all scales and a similar weighting for standard deviation was applied as well. This affected ERA5 data from January 2000 to June 2005, and from January 2010 to October 2014. As will be seen in Section 7.1, this change had a minor effect on the HRES reanalysis products. It did have a larger effect on the spread in the EDA ensemble and standard deviation of analysis increments.

The static part **B**_{cli}, that is, the long-term average error characteristics of the model first-guess, depends on the

state of maturity of the observing system as a whole. Therefore, it should follow the evolution of the observing system. In ERA5 this is handled in a rather crude way. One single estimate is used for the period from 1979 to the end of 1999 (called 1979-**B**_{cli}) which represents the early satellite era, and one from 2000 onwards for the modern satellite era, which follows what is used by default in Cy41r2 (called 41r2-**B**_{cli}). The effect of the abrupt transition on 1 January 2000 on the resulting reanalysis products is described in Section 9. The narrower correlations of the 41r2-**B**_{cli} allow for more accurate, local adjustments in the model analysis around observations, while 1979-**B**_{cli} distributes the information of sparser observations over larger distances. The static **B**_{cli} matrices were determined from the spread in a large number of short-range forecasts from dedicated hybrid EDA runs, spanning several months and combining winter and summer cases in the period of interest.

The EDA also benefits from VarBC (Dee, 2005). As for the model initial state **x**, bias parameters β are not perturbed. Analysis updates for β are only estimated from the control, and are propagated as estimates for β to all perturbed members. The reason for this is to avoid artificially long correlations in the background-error statistics as derived from the short-range ensemble forecasts (Isaksen *et al.*, 2010).

3 | PRODUCTION AND AVAILABILITY

ERA5 is produced on the ECMWF high-performance computing facility. Typically 6 to 9 days of reanalysis can be completed per day. To produce 70 or so years of reanalysis in a few years, ERA5 has been split into a number of parallel streams which are later merged into one consolidated public dataset. The original target for the 40 years discussed in this paper was to have one stream per decade starting from ERA-Interim initial conditions, with the exception of the stream for the 1980s and mid-2010s which were initialized from ERA-40 and ECMWF operational (Cy40r1 for 30 May 2014) analyses, respectively. A one-year overlap between streams allows for a long spin-up period. In practice, the final consolidated product from 1979 onwards was built up from more streams. Details are listed in Table 3. This deviation was driven by practical solutions which were implemented to handle issues that were discovered and resolved during production. The most serious of these were: the initial appearance of anomalously high ozone during the polar night because of issues with the assimilation of ozone; an incomplete response to the Pinatubo eruption due to the usage of a sub-optimal **B**_{cli} in the stratosphere; and the persistent appearance of anomalous sea ice over parts of the Baltic

TABLE 3 Start months in the consolidated public ERA5 dataset of the collation of production streams from January 1979 to end of 2019

HRES	Jan 1979 [†] , Jul 1981, Apr 1986 [†] , Oct 1988, Aug 1993 [†] , Sep 1995, Jan 2000 [†] , Oct 2000 [†] , Oct 2001 [†] , Oct 2002 [†] , Oct 2003 [†] , Oct 2004 [†] , Oct 2005 [†] , Oct 2006 [†] , Jan 2008, Jan 2010, Jan 2015, Mar 2019
EDA	Jan 1979, Apr 1986, Aug 1993, Jan 2000, Jan 2010, Jan 2015, Mar 2019

Note: The [†] labels repair runs, while the stream starting March 2019 is effectively a continuation of (i.e., initialized by) its predecessor and is also used for the timely updates.

during summer. By extending certain streams sufficiently long into the affected subsequent stream, the final consolidated public dataset was not affected by these issues. Details are provided in Hersbach *et al.* (2018). Another solution, which was used to clear the sea ice issue which was discovered quite late in the production, was to conduct a number of relatively short ‘repair’ runs (11 in total) which replace data from affected periods. Much care was taken to ensure that the mean state of such re-runs only deviates significantly from the original products where required. This was accomplished by warm-starting such streams from the initial production with a two-week spin-up. In addition, the VarBC bias parameter estimates β (Equation (3)) were imposed from the original production; this method is also used to avoid systematic differences between the EDA members (Section 2.4). Such re-runs were not conducted for the ensemble component, which means that the sea ice problem was not resolved for the uncertainty estimate; the reason for this was the prohibitive cost of so doing. Some details are provided in Section 6.2.

As shown later in Section 9, the mean state of the resulting consolidated dataset exhibits very few significant jumps in the troposphere or stratosphere at transition points between streams. The most serious one is a discontinuity in stratospheric and upper-tropospheric temperature on 1 January 2000, which is the result of the switch of the climatological part of the model background error from 1979- $\mathbf{B}_{\text{cli}}$ to 41r2- $\mathbf{B}_{\text{cli}}$. Stratospheric humidity has a marked discontinuity at both this transition point and early in 1986, and smaller discontinuities occur at other times. In the observation-free mesosphere and deep soil (where spin-up can take several years), at the seams discontinuities in time are in general observed, with the exception of repair-runs (the streams denoted by a [†] in Table 3), since these were warm-started.

The consolidated dataset is archived in the ECMWF MARS. Its volume from 1979 onwards totals about 5 petabytes. To ensure fast access to ERA5 data, a post-processed product (around 1 petabyte), is available on the CDS cloud server (<https://cds.climate.copernicus.eu/>; accessed 11 April 2020). This includes upper-air parameters on 37 pressure levels from 1,000 to 1 hPa,

and a large number of near-surface parameters and other two-dimensional fields. The CDS data have been converted from the native reduced-Gaussian grid to a regular latitude–longitude grid (0.25° for the high-resolution deterministic reanalysis and 0.5° for the ensemble products; for ocean wave products, 0.5 and 1°, respectively). Several parameters, such as precipitation, surface fluxes and minimum and maximum temperatures, are provided on the CDS as hourly timeseries which combine hourly analysis fields with short-range forecasts as needed. This simplifies many technical difficulties that users have encountered in the past when retrieving ECMWF reanalysis data.

For convenience, ensemble spread and mean data are directly available. Monthly-mean averages (both for a particular hour of the day and averaged over all hours in the entire month) have been pre-computed as well. Monthly-mean values for ensemble mean and spread were not pre-computed.

The ERA5 online data documentation (available via <https://confluence.ecmwf.int/display/CKB>; accessed 11 April 2020) provides a detailed description of the various products and a list of all available geophysical parameters.

Timely updates are provided. Each day, one day of reanalysis products is added with a delay of 5 days. Once a month, one month of this preliminary product is replaced by a more thoroughly quality-checked final product with a delay of 2–3 months. In practice, both products will rarely differ, and if they do (due to considerable errors found in the daily release), users will be notified.

At the time of writing, the back extension from 1950 to 1978 had been completed, though was not yet publicly available. The stream consolidated up to December 1999 has been continued up to end of 2006 since, as will be discussed in Section 7, its usage of the 1979- $\mathbf{B}_{\text{cli}}$ has a beneficial effect on stratospheric temperature and the uncertainty estimate for ozone. It differs very little in the troposphere. A more in-depth assessment is provided in Simmons *et al.* (2020). This stream, which is labeled ERA5.1 in this document, was made available as a supplement to the existing ERA5 dataset in May 2020.

4 | BENEFITS FROM A DECADE OF IFS IMPROVEMENTS AND PREVIOUS REANALYSES

In the ten-year period between ERA-Interim (Cy31r2) and ERA5 (Cy41r2), many significant improvements have been made to the representation of model processes (and the treatment of their adjoint and tangent-linear approximations, where applicable) and to the DA methodology in the IFS. This section provides for each component (atmosphere, ozone, land and ocean waves) a summary of such changes, other than those already mentioned in Section 2. It also provides a summary of the major changes in observation handling.

4.1 | Improvements for the atmosphere

The radiation scheme used in ERA5, McRad, described by Morcrette *et al.* (2008), is a major upgrade from the scheme used by ERA-Interim. It incorporates the MCICA (Pincus *et al.*, 2003) for representing subgrid cloud structure and overlap, and the short-wave RRTMG (Iacono *et al.*, 2008), consistent with the existing use of RRTMG in the long-wave. The radiation scheme is called every hour on a grid 2.5 times coarser in each horizontal direction. To mitigate erroneous temperatures at coastlines caused by the coarser grid, approximate updates to the fluxes are performed every time step and gridpoint (Hogan and Bozzo, 2015). Infrequent radiation calls can lead to a warm stratosphere bias through accumulated numerical error in sunrise and sunset times, but this has been mitigated by the Hogan and Hirahara (2016) scheme for computing effective solar zenith angle.

The large-scale cloud and precipitation scheme, based on Tiedtke (1993), was upgraded with an improved representation of mixed-phase clouds (Forbes and Ahlgrimm, 2014), and prognostic variables for precipitating rain and snow (Forbes and Tompkins, 2011; Forbes *et al.*, 2011). In addition, there were numerous improvements to the parametrization of the microphysics, particularly for warm-rain processes (Ahlgrimm and Forbes, 2014) but also ice-phase processes and ice supersaturation.

Changes to the parametrization of convection, originally based on Tiedtke (1989), include a thorough revision of the entrainment and the coupling with the large scale, leading to a large redistribution of rainfall from the Hadley Cell to the Walker Cell, a large improvement in the distribution of rain rate versus TRMM and an improved representation of tropical variability (Bechtold *et al.*, 2008; Hirons *et al.*, 2013). Improvements in the diurnal cycle of convection, shifting the rainfall peak over land from noon to late afternoon, have been achieved by use of

a modified convective available potential energy (CAPE) closure (Bechtold *et al.*, 2014).

There were changes to the parametrizations of orographic drag, subgrid turbulent mixing and interactions with the surface in unstable and stable conditions (Sandu *et al.*, 2011; Sandu *et al.*, 2014).

A non-orographic gravity wave drag parametrization was introduced to represent the effects of upward propagating gravity waves from tropospheric sources such as deep convection, frontal disturbances, and shear zones. The parametrization uses a globally uniform wave spectrum, and propagates it vertically through changing horizontal winds and air density, thereby representing the wave breaking effects and associated drag due to critical level filtering and nonlinear dissipation in the stratosphere and mesosphere (Orr *et al.*, 2010).

An improvement in the wind extrapolation scheme SETTLS used for the departure point calculation, described in Diamantakis (2014), reduced numerical noise in the upper stratosphere typically occurring during SSW events. The practical benefits of this modification was a large reduction of both analysis and forecast temperature error and an overall enhanced medium-range predictability of SSW events.

Improved de-aliasing of the pressure gradient term (ECMWF R&D, 2016), reduced numerical noise in the adiabatic tendencies, allowing a reduction of the horizontal diffusion used in the forecast.

Regarding the tangent-linear and adjoint physics that are used in the inner loops of the 4D-Var DA system (Janisková and Lopez, 2013), improvements include:

1. Inclusion of the freezing of rain in the moist physics;
2. Substantial revision of the moist physics to match the nonlinear reference large-scale cloud and convection schemes;
3. Replacement of the old long-wave radiation parametrization (neural network) by the more elaborate scheme of Morcrette (1991);
4. Linearized version of the new non-orographic gravity wave drag;
5. Added simplified linearized parametrization scheme for surface processes to represent the evolution of the top soil layer, snow and sea ice temperatures; and
6. Revision of the linearized vertical diffusion to match the changes of the exchange coefficients in the nonlinear scheme.

4.2 | Improvements for ozone

The heterogeneous ozone chemistry, which is based on a modified version of the Cariolle and Déqué (1986) scheme,

was updated. A number of changes were introduced in the assimilation system, such as the extension of VarBC to observations of total column ozone, which are described in Sections 4.5 and 5.10.

4.3 | Improvements for the land component

In ERA5 the HTESSEL land surface scheme is used. Balsamo *et al.* (2015) documented the revised hydrology of the HTESSEL scheme as used in ERA-Interim/Land compared with the TESSEL scheme (van den Hurk *et al.*, 2000) used in ERA-Interim. Some of the most significant changes from ERA-Interim to ERA5 are related to (a) the introduction of the soil texture map (Balsamo *et al.*, 2009), and (b) an improved representation of bare soil evaporation (Albergel *et al.*, 2012). The new scheme also accounts for seasonally varying monthly vegetation maps specified from a MODIS-based satellite dataset (Boussetta *et al.*, 2013). In addition, an enhanced snowpack parametrization allows a more realistic timing of runoff and terrestrial water storage variations and a better match of the albedo to satellite products (Dutra *et al.*, 2010). Balsamo *et al.* (2012) introduced the capacity of forecasting of inland water bodies and evaluated the impact when coupled to the atmosphere, following a previous offline evaluation of sensitivity to lakes (Dutra *et al.*, 2009). The chosen parametrization for lakes (FLake; Mironov *et al.*, 2010), allows consideration of both subgrid and resolved water bodies (Manrique-Suñén *et al.*, 2013). This series of changes contributes to significant improvements in the soil moisture and land surface fluxes consistency, which allowed for the usage of satellite data in ERA5 to analyse soil moisture as described below.

Regarding the LDAS, the snow analysis and the SEKF for soil moisture are an improvement on ERA-Interim in which a Cressman interpolation and 1D-OI method were used, respectively (de Rosnay *et al.*, 2013).

4.4 | Improvements for ocean waves

The model bathymetry was updated to use a more recent version of ETOPO2 (NOAA, 2006). A new wave advection scheme was introduced with a revised unresolved bathymetry scheme to better account for the propagation along coastlines and to better model the impact of unresolved islands (Bidlot, 2012). The slow attenuation of long-period swell as well as the impact of shallow water on the wind input was introduced with an overall retuning of the level of dissipation due to white-capping (Bidlot, 2012). Extra output parameters were introduced to better

characterise freak waves (Janssen and Bidlot, 2009), swell systems and wave-modified fluxes to the oceans.

4.5 | Improved handling of observations

ERA5 benefits from many improvements in the observation operators and in the handling of observations implemented in the IFS since the start of ERA-Interim (based on Cy31r2, with RTTOV-7 as the radiative transfer model). ERA5 uses RTTOV-11 (Saunders *et al.*, 2018) as the observation operator for radiance data (Lupu and Geer, 2015), which incorporates improvements in the underlying spectroscopy in both the microwave and infrared as well as improvements in the optical depth predictor model relative to RTTOV-7.

ERA5 also benefits from the ongoing development of all-sky assimilation at ECMWF (Geer *et al.*, 2017). Initially implemented for the microwave imagers, the scheme was successfully extended to microwave humidity sounding data. The approach exploits the capability of RTTOV to model radiative transfer in cloudy and precipitating atmospheres, as well as a linearised moist physics scheme (Janisková and Lopez, 2013), to assimilate microwave observations in all-sky conditions. All-sky assimilation improves analyses both through the improved analysis of moist variables, as well as through improved analyses of dynamical fields resulting from the ability of 4D-Var to extract wind information from the advection of tracers, in this case humidity, cloud and rain. The scheme rectified a problem with the earlier 1D+4D-Var assimilation of rain-affected radiances which, in ERA-Interim, resulted in an underestimation of global rainfall (Dee *et al.*, 2011).

Several other developments have enhanced the exploitation of observations since ERA-Interim. The diagnosis and modelling of several types of observation error has advanced significantly since 2007 (Bormann *et al.*, 2009). For example, for AMSU-A improvements were made allowing for increased weight in the assimilation and improved handling of observation errors in cloudy scenes and over orography. Situation-dependent observation errors for AMVs were introduced and the observation errors for GNSS-RO bending angle data were re-tuned, with more weight given to bending angles in the middle and upper stratosphere. Advances have also been made in extending the use of microwave data over land and sea ice surfaces (Bormann *et al.*, 2017).

For GNSS-RO data, allowance for tangent point drift was introduced (Poli *et al.*, 2009) and a 2D observation operator implemented (Healy *et al.*, 2007). The refractive index coefficients used in the ray-path computation were also revised, including non-ideal gas effects (Healy, 2011).

For observations of total column ozone, quality control was made stricter, rejecting all observations where first-guess departures exceed 30 Dobson units. Assimilation of IR ozone-sensitive radiances was introduced (Dragani and McNally, 2013), exploiting a part of the spectrum never used before from HIRS (channel 9) and the hyperspectral IR sounders IASI, AIRS and CrIS.

Regarding dedicated efforts for the ERA5 reanalysis, several aspects of the observation operators for key datasets have been improved for the early satellite era. Observations from IR sounding instruments provide important information on upper-air temperatures. This information is mainly extracted from channels at wavelengths around the $15\ \mu\text{m}$ CO_2 band. For these channels, the form of their weighting functions, and in particular the heights of the peaks in the weighting functions, is determined by the form of the CO_2 concentration profile which exhibits seasonal variability as well as long-term trends. Shine *et al.* (2008) have shown that the long-term trend in CO_2 , if unaccounted for, gives rise to spurious atmospheric temperature trends from SSU observations which range from $-0.4\ \text{K decade}^{-1}$ to $+0.4\ \text{K decade}^{-1}$, depending on altitude. In ERA5 the variability in CO_2 is taken into account for the early IR sensors assimilated in the reanalysis for the duration of these missions: HIRS, SSU and VTPR. CO_2 profiles used in RTTOV-11 are estimated from zonal fields as used in the ERA5 radiation scheme (Section 6.1). For the advanced IR sounders (AIRS, IASI and CrIS) the evolution of atmospheric CO_2 concentrations is not taken into account and the effect on the estimated bias corrections is as expected and is illustrated in Figures S1 and S2.

In a separate development, an improved observation operator for the assimilation of SSU observations has been incorporated in ERA5. The SSU instruments, forming part of the TOVS suite of instruments and operational from late-1978 until mid-2006, provide valuable information on mid-upper stratospheric temperatures in the pre-ATOVS era (i.e., pre-1998). The SSU instrument is an infrared radiometer employing a detection technique based on pressure modulation. Leaks in the pressure modulator cells have led to complex time-dependent biases in the observations (Nash and Saunders, 2015). A parametrized correction scheme has been developed and tested, based on measured cell pressures, which improves the simulation of brightness temperatures for NOAA-7 and NOAA-11.

5 | OBSERVATIONS

This section provides a detailed account of the observing system as used in ERA5. It starts with a general overview

in Section 5.1, followed by more details in dedicated subsections.

5.1 | The evolving observing system

The number of observations assimilated in ERA5 increases from approximately 0.75 million per day on average in 1979 to around 24 million per day by January 2019. In the 40 years from 1979 to 2019 inclusive, 94.6 billion observations were actively assimilated in 4D-Var, 65 million in the OI ocean-wave component and about one billion observations each of surface air temperature and relative humidity were processed by the LDAS.

The 4D-Var component of ERA5 uses observations from over 200 satellite instruments or types of conventional data. It extracts information from *in situ* observations of 10 m wind over sea and 2 m humidity over land, and pressure over land and sea. Upper-air observations of wind, temperature and humidity are obtained from PILOT, radiosonde, dropsonde and aircraft measurements. Upper-air wind is also obtained from AMV winds from a number of polar and geostationary satellites. In addition, ERA5 uses information on rain rate from ground-based radar-gauge composite observations from 2009. ERA5 uses measurements from many satellite platforms. These include radiances sensitive to upper-air temperature, humidity and ozone from (clear-sky) HIRS, MSU, SSU, AMSU-A, AMSU-B, ATMS, MWHS and (all-sky) MWHS2, MHS, and from microwave imagers (also all-sky) SSMI, SSMIS, TMI, AMSR-2, AMSRE, GMI, and (hyperspectral infrared radiances) from IASI, AIRS, CRIS and (mixed clear-sky and all-sky) from geostationary satellites. ERA5 uses level 2 ozone from a range of instruments. ERA5 also benefits from GNSS-RO bending angles (from 2001, but large quantities from 2006), providing information on upper-air temperature and humidity. Information on ocean vector wind (4D-Var) and land soil moisture (LDAS) is obtained from scatterometers, while information on ocean-wave height (OI) is obtained from altimeters (both types of instruments from 1991). In addition, the LDAS uses *in situ* observations of the global SYNOP network for temperature and humidity at screen level, soil moisture and snow depth. From 2004 onwards it also uses information on snow cover over the Northern Hemisphere from the multi-sensor IMS system.

Figure 3 shows daily counts of those observations per assimilated variable, on a logarithmic scale. Radiances are the dominant and growing source of measurements throughout the period. Major developments for this class of observations have included the transition from the TOVS (HIRS-2, MSU, SSU) to the ATOVS suite

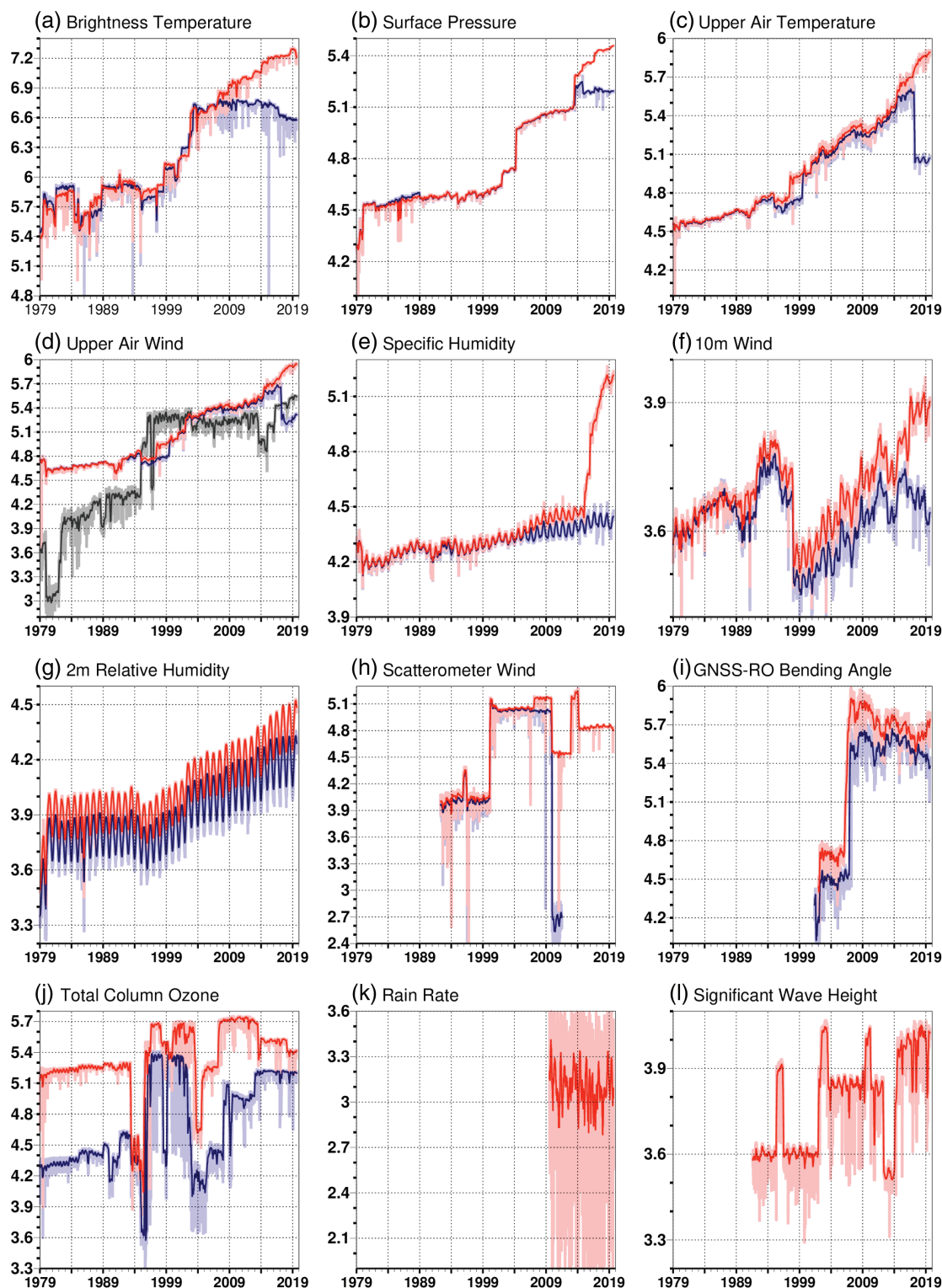


FIGURE 3 Number of daily actively assimilated observations (weak colours) and 30-day means (strong colours), both in log10 scale, for ERA-Interim (blue) and ERA5 (red) for the 12 observables for which observations are assimilated in 4D-Var and ocean wave assimilation. Numbers for (d) upper-air wind exclude assimilated AMV wind. Numbers of these latter are shown for ERA5 (black), while no statistics are available for ERA-Interim. Rain (k) is assimilated only in ERA5. No statistics are available for (l) ERA-Interim significant wave height, while information on data usage in LDAS is incomplete for both reanalyses. Each 0.3 tick on the vertical axis corresponds to a factor of 2, and minor ticks (0.1) to a difference of 26%

(HIRS-3/-4, AMSU-A, AMSU-B) of sounding instruments, the introduction of hyperspectral infrared radiances and the increasing availability of observations from a growing constellation of microwave imagers. There has been a marked increase in the number of other satellite observations assimilated, notably GNSS-RO, scatterometer wind observations and Level-2 ozone products. The volume of conventional observations has increased steadily throughout the period. An analysis of the observation impact, based on the Degrees of Freedom for Signal diagnostic (DFS; Cardinali *et al.*, 2004) shows that, as expected, satellite observations play a progressively more important role through the period (Horányi, 2017).

Figure 3 also shows the observation volumes assimilated in ERA-Interim. Generally, more observations are assimilated in ERA5. Several discrepancies are apparent. The divergence in the volume of radiance measurements assimilated post-2007 (Figure 3a) is due to the assimilation of many new observations in ERA5, such as the hyperspectral observations from IASI and CrIS, which are not assimilated in ERA-Interim, together with the gradual decline in the numbers assimilated in ERA-Interim, as instruments and channels gradually fail. ERA5 uses a revised cloud-detection scheme for HIRS (Krzeminski *et al.*, 2009) which appears to remove more observations and this is why, initially, fewer radiance measurements are used in ERA5. The initially smaller gain in forecast skill for ERA5 in Figure 1 seems to be unrelated to this scheme, though, since ERA-Interim type experiments with the revised scheme performed comparably to ERA-Interim itself. ERA-Interim also shows a sharp decline in the number of surface pressure and upper-air winds and temperatures following the transition by data providers to a BUFR format from 2013 onwards which the system cannot handle. In addition, ERA5 makes use of several new and improved reprocessed datasets.

5.2 | Reprocessed and new datasets

Improvements in the characterisation, inter-calibration and processing of conventional and satellite measurements have enabled data providers to progressively refine the quality of historical observations, in terms of coverage and accuracy. ERA5 has made use of several reprocessed satellite datasets, which were acquired from a number of space agencies and institutes, as listed in Table 4. The EU-funded precursor project, ERA-CLIM, provided AMVs from Meteosat-8 and -9 and Metop-A AVHRR. Also shown in Table 4 are some new datasets not used in earlier ECMWF reanalyses. Details are provided in the following subsections.

5.3 | Conventional observations and ground-based radar–gauge composites

Conventional meteorological measurements used as input in ERA5 sample the troposphere and the lower and middle stratosphere, and come from observations made near the surface on land and over oceans, upper-air soundings, and atmospheric measurements from instruments on board aircraft operating on air routes. Their spatial coverage and temporal resolution vary in time, from sparse observations in the 1950s particularly over the Southern Hemisphere, to the current dense observing network. ERA5 uses conventional observations prepared initially for ERA-40, spanning September 1957 to December 2001, and from the operational ECMWF data archive, received through the GTS, to cover the period from 2002 to the present. For the period prior to 1979, ERA5 also benefits from the assimilation of improved reprocessed conventional datasets, including the ISPD v 3.2.6 dataset (Cram *et al.*, 2015), the ICOADS v2.5.1 (Woodruff *et al.*, 2011) and data collections from the NCEP. Figure 4 provides a detailed overview of the observation usage in ERA5 (and compared to ERA-Interim).

Data selection rules for observations are predefined in blacklists that cover the entire reanalysis period, and are applied in the same way as in ERA-Interim (Dee *et al.*, 2011). Conventional observations are classified into five types:

- *SYNOP*, consisting of measurements made near the surface at land stations (surface pressure, relative humidity in daytime in 4D-Var and 2 m temperature, relative humidity and snow depth in the land component) including airport weather reports (surface pressure), and on ships (surface pressure, wind components at 10 m);
- *DRIBU*, comprising drifting and moored buoys (surface pressure, wind components at 10 m);
- *TEMP*, which includes radiosondes and dropsondes (temperature, wind components, specific humidity);
- *PILOT* balloon observations (wind components), and wind profilers (wind components), and
- *AIRCRAFT*-based atmospheric observations (temperature, wind components and specific humidity).

In addition to those observations assimilated in ERA-Interim, ERA5 assimilates the NCEP stage IV quantitative precipitation estimates produced over the USA by combining precipitation estimates from the NEXRAD with gauge measurements. An overview of the method used to assimilate this product is provided by Lopez (2011). The usage of these measurements benefits from VarBC

TABLE 4 Reprocessed and new(*) satellite data assimilated in ERA5

Instrument / Satellite	Period covered	Agency
<i>Atmospheric motion vectors</i>		
Meteosat 1st Gen.(M-2 to -7)	1982–2000	EUMETSAT
Meteosat 2nd Gen.(M-8,-9)	2004–2012	EUMETSAT
GMS (-1,-3,-4,-5)	1979, 1987–2003	JMA
MTSAT-1R	2005–2009	JMA
GOES-9	2003–2009	NOAA
GOES GVAR (-8 to 13,-15)	1995–2013	CIMSS
AVHRR (NOAA-7 to -18)	1981–2014	CIMSS
AVHRR (MetOp-A)	2007–2012	EUMETSAT
<i>Radiances</i>		
DMSP SSMI (F-08 to -15)	1987–2008	EUMETSAT CM SAF
Meteosat Second Gen. ASRs	2003–2012	EUMETSAT
IASI* (Metop-A,-B)	2006–present	EUMETSAT
CrIS* (S-NPP/ NOAA-20)	2012–present	NOAA
MWHS*/MWHS-2* (FY-3B,-3C)	2012/2014–present	CMA
TMI*/SSMIS*/AMSR-2*/GMI*	2005/2009/2012/2015-	NASA/DMSP/
	2015/(3)present	JAXA/NASA
Ozone channels *(HIRS, AIRS, IASI and CrIS)	1979–present	NOAA, NASA, EUMETSAT and NOAA
<i>Radio occultation</i>		
Blackjack	2001–2014	UCAR
(GRACE-A, CHAMP, SAC-C)		
IGOR	2006–2014	UCAR
(TerraSAR-X, COSMIC-1 to -6)		
<i>Scatterometer wind</i>		
ASCAT* (MetOp-A,-B)	2007–2014	EUMETSAT
Oceansat*	2012–2014	ISRO
<i>Ozone retrievals</i>		
GOME-2 (Metop-A,-B)	2007–2013	ESA/EUMETSAT
GOME (ERS-2)	1996–2002	ESA
MIPAS (ENVISAT)	2005–2012	ESA
MLS (EOS-AURA)	2004–2014	NASA
OMI (EOS-AURA)	2004–2015	NASA
BUV (Nimbus-4)*	1970–1977	NASA
SBUV and SBUV-2 (Nimbus-7, NOAA-9,-11,-14,-16,-17,-18,-19)	1978–2013	NOAA
SCHIAMACHY (ENVISAT)	2002–2012	ESA
TOMS (NIMBUS-7, Earth Probe ADEOS-1)	1978–2006	NASA
	1996–1997	NASDA

TABLE 4 Continued

Instrument / Satellite	Period covered	Agency
<i>Scatterometer soil moisture</i>		
AMI on ERS-1,-2	1991–2006	TU Wien
MetOp-A,-B ASCAT	2007–2014	EUMETSAT H SAF
<i>Altimeter wave height</i>		
RA on ERS-1,-2	1991–2003	ESA
AltiKa on SARAL*	2014 onwards	CNES/ISRO
SIRAL on CryoSat-2*	2014 onwards	ESA
Poseidon-2 on Jason-1	2001–2010	NASA / CNES
RA-2 on Envisat	2002–2012	ESA

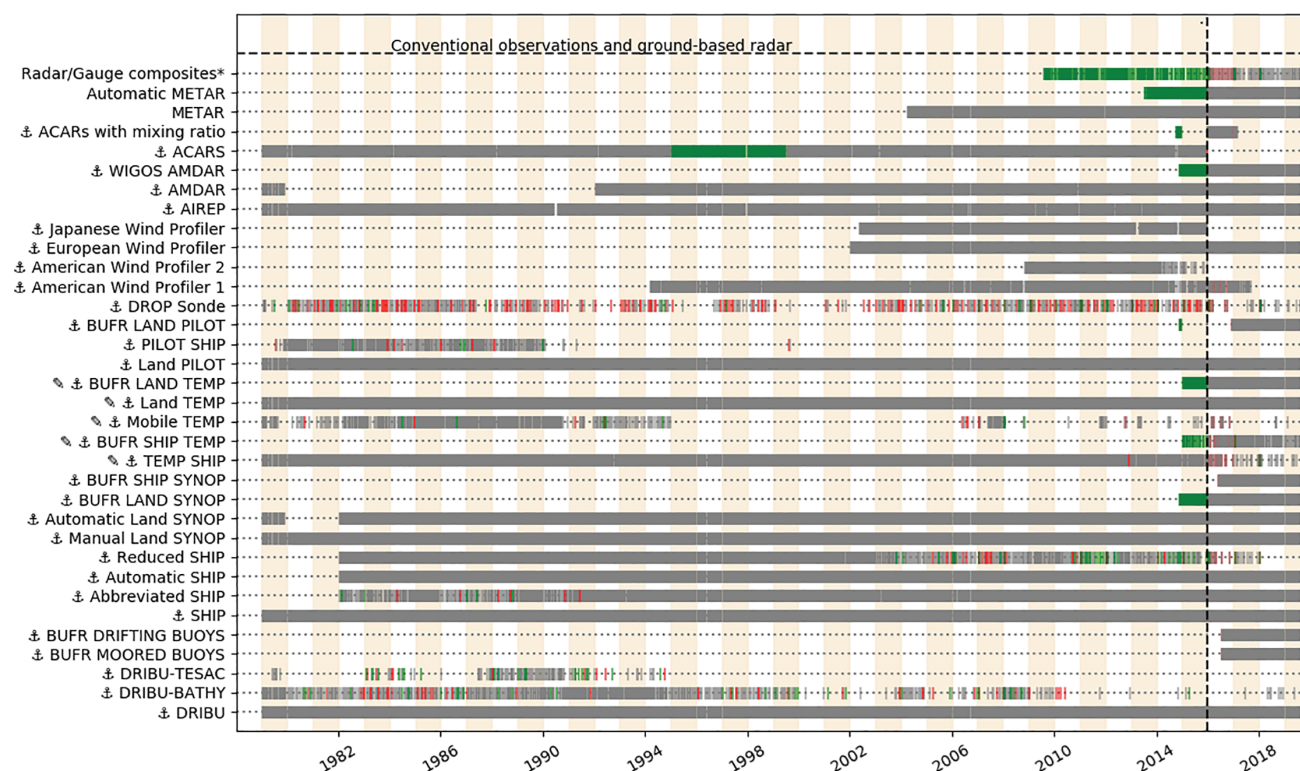


FIGURE 4 Conventional observations assimilated per day (2100 to 2100 UTC) in ERA5 during the period 1979–2018. Grey bars indicate types of observation used in both ERA5 and in the reference system (ERA-Interim, from 1500 to 1500 UTC, until 2016; ECMWF operations from 2100 to 2100 UTC thereafter; the transition is indicated by the vertical dashed line). Red bars indicate observations assimilated in the reference system but not in ERA5, and green bars indicate observations that are used in ERA5, but not the reference system. A pencil symbol preceding an observation type indicates that a prescribed bias correction is applied to at least one assimilated variable (or and/or channel for satellite observations) provided by that observation type. Similarly, an anchor symbol indicates that at least one variable or channel provided by the observation type is used to anchor the analysis, i.e., it is assimilated without applying a bias correction. Details on anchored variables and channels as well as on prescribed bias corrections are given in the sections covering the respective observations. The NCEP stage IV quantitative precipitation estimates observation type is indicated as radar/gauge composites

where all observations are grouped into one bias group subject to a six-parameter bias model.

In order to better represent past extreme weather events, the ERA5 segment prior to 1979 will benefit from the assimilation of the tropical cyclone best-track

minimum sea level pressure data contained in the ISPD dataset. These are not used in the period from 1979.

At cruise level (around 200 hPa), aircraft observations are on average biased warm by about 0.2 K, which had affected temperature analyses in ERA-Interim when their

numbers significantly increased around 1999 (Dee and Uppala, 2009). In ERA5 this has been alleviated by extending VarBC to aircraft temperatures with one bias group per type of aircraft. Only one predictor (a constant) is used, rather than three (also ascent and descent speed) in the ECMWF operational system at the time since, during the preparations for ERA5, that implementation was discovered to be flawed. This has since been corrected (Ingleby *et al.*, 2018) in Cy45r1 (5 June 2018).

For surface pressure, bias estimates are updated by VarBC as it was developed for the ERA-20C reanalysis (Poli *et al.*, 2016), using one group per platform with one predictor (a constant) and a background B_β term that corresponds to a response time of 60 days. For these observations, VarBC is performed in the screening task, that is, before the minimization in 4D-Var. This is achieved by solving Equation (4) for $\delta\beta$ with $\delta x = 0$, which is then decoupled into a set of low-dimensional linear equations, one per bias group. For a one-parameter bias model, its solution is trivial. The reason for this choice is to avoid an undesired (and understood) interaction with the applied Huber norm (Tavolato and Isaksen, 2015) which would lead to a far too slow response.

In the ERA5 troposphere, about a 10% higher weight is assigned to radiosonde temperatures than in ERA-Interim (Tavolato and Isaksen, 2015). In the stratosphere the opposite is true, where prescribed observation errors were inflated by 20–30%. Details are provided in Table 5. At all heights the weight assigned to PILOT and radiosonde wind was increased by 10%, while the weight assigned to radiosonde humidity (assimilated below 100 hPa only) is unaltered. Regarding bias corrections for radiosonde temperature, an update of the method of a pre-calculated RAOBCORE (Haimberger *et al.*, 2008) homogenization as in ERA-Interim is used, where estimates are now also based on comparison between neighbouring stations, rather than from departure statistics alone (RICH; Haimberger *et al.*, 2012). An additional solar-elevation-dependent correction is applied as in ERA-Interim. From 1 January 2015 onwards, such estimates are not available and ERA5 instead follows the bias-correction scheme in the operational medium-range forecast system.

5.4 | Microwave radiances

5.4.1 | Clear-sky assimilation of temperature sounder radiances

Measurements from the microwave sounders MSU, AMSU-A and ATMS (Bormann *et al.*, 2012), assimilated

TABLE 5 Prescribed observation errors (K) for radiosonde temperature in ERA5 and ERA-Interim as a function of height

Pressure (hPa)	ERA-Interim	ERA5
10	1.47	1.66
20	1.05	1.34
30	0.98	1.28
70	0.98	1.28
100	0.91	1.09
150	0.88	0.70
250	0.81	0.73
300	0.70	0.64
400	0.63	0.57
500	0.66	0.61
700	0.77	0.70
1000	0.98	0.89

Note: At other heights, observation error is a linear interpolation between the values at the listed pressures, while it is constant from 10 hPa upwards. In the troposphere/stratosphere in ERA5, a higher/lower weight than in ERA-Interim is assigned, with a crossover point at about 132 hPa.

as brightness temperatures (as is the case for all radiance observations), provide information on temperature throughout the troposphere and stratosphere. ATMS additionally provides humidity sounding capability. This component of the observing system evolved considerably throughout the 1979–2019 period (Figure 5). For example in the mid-1990s, typically observations from two microwave temperature sounding instruments were assimilated. By early 2005 this had increased to four, and by early 2019 observations from nine instruments (seven AMSU-A and two ATMS) were assimilated. This improves the resilience of the system to discontinuities resulting from outages of any single instrument, but also reduces analysis errors through the effect of averaging independent errors in the observations, as well as through improved sampling in time through the analysis window. The radiances are bias corrected using VarBC (Auligné *et al.*, 2007). The bias correction model for channels assimilated from these sensors employ a constant term, a scan-angle-dependent correction (based on a third-order polynomial in scan-angle) and four airmass predictors, represented by the thickness of layers 1000–300 hPa, 200–50 hPa, 50–5 hPa and 10–1 hPa. The exception here is AMSU-A channel 14 for which no bias correction is applied, in order to anchor the temperature analysis in the upper stratosphere.

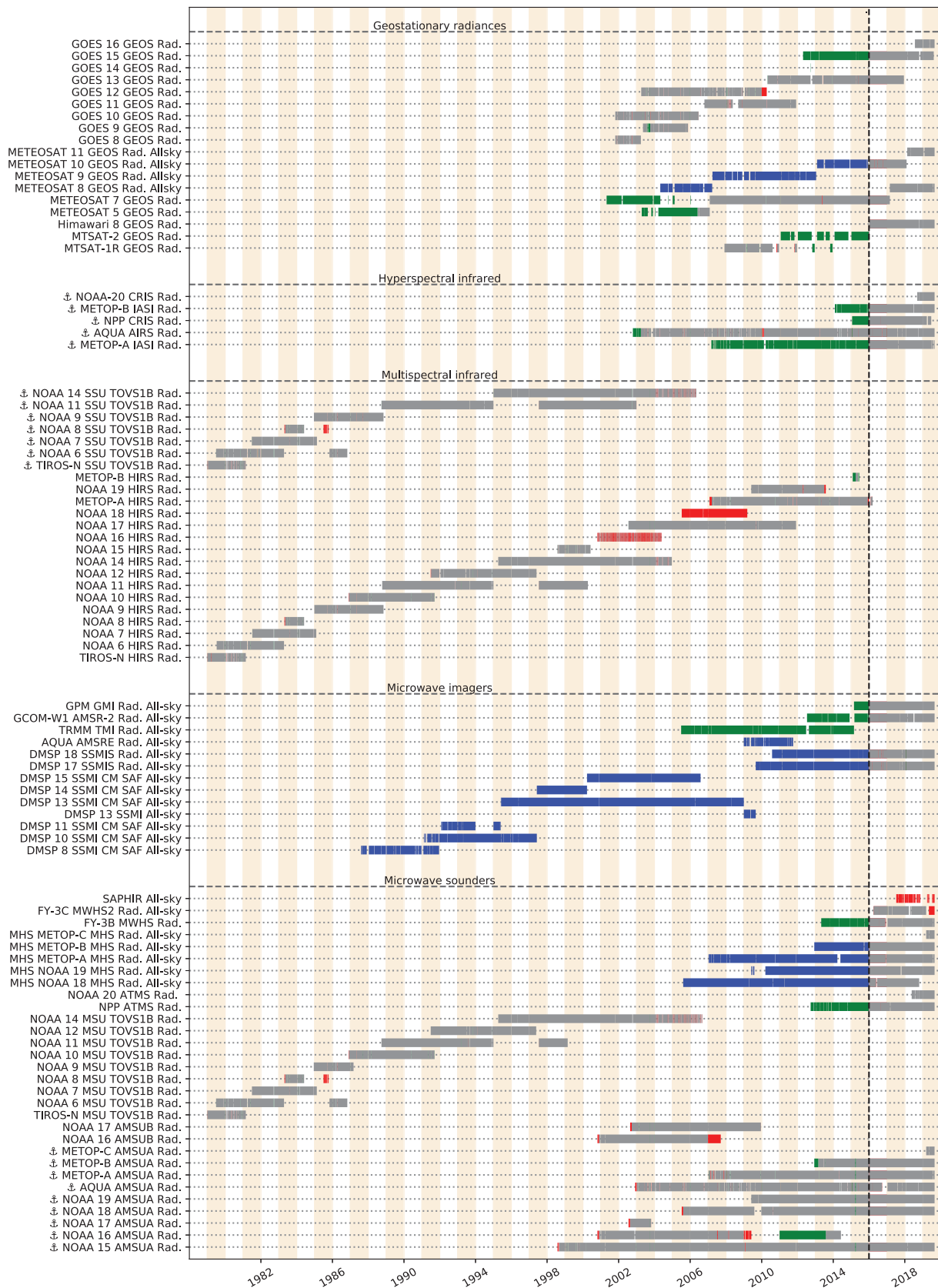


FIGURE 5 Radiance observations assimilated in ERA5. The colour scheme is as for Figure 4. Additionally, blue bars indicate observations that are reprocessed relative to those assimilated in ERA-Interim, or for which the processing has changed significantly since ERA-Interim

Similar bias models are used for the other radiance datasets described below in Sections 5.4.3, 5.5.1 and 5.5.2.

5.4.2 | All-sky assimilation of humidity sounder radiances

Humidity information throughout the troposphere is obtained from passive microwave observations, for most sensors using the all-sky approach (Geer *et al.*, 2017). Humidity sounding radiances from MHS instruments on board NOAA-18 and -19 and the MetOp satellites (Figure 5) are assimilated in all-sky conditions along with the FY-3C humidity sounder MWHS-2 (Lawrence *et al.*, 2018). However the MWHS-1 sensor on FY-3B (Chen *et al.* 2014) and AMSU-B sensors on NOAA-16 and -17 were assimilated using clear-sky scenes only, as an all-sky configuration was not available in the Cy41r2 at the start of ERA5.

5.4.3 | All-sky assimilation of microwave imager radiances

Microwave imagers provide radiance observations which, over ice-free ocean surfaces, improve the analysis of lower-tropospheric humidity, cloud liquid water and ocean surface wind speed. An overview is given in Figure 5. ERA5 assimilates EUMETSAT CM SAF SSM/I FCDRs (Fennig *et al.*, 2017) during the period August 1987–December 2008. This aspect of the observing system has grown considerably since the 1980s and radiances from TMI and AMSR-E were used starting in 2005 and 2009 (Geer *et al.*, 2010), SSMI/S on F-17 starting in 2009, and AMSR2 and GMI in 2012 and 2015 (Kazumori *et al.*, 2016; Geer *et al.*, 2017). However, to avoid possible bias problems, usage has been limited through the blacklist to a maximum of three imagers at any one time. In addition, the humidity sounding channels of SSMI/S have been assimilated, including those from the F-18 satellite, which is not otherwise used.

5.5 | Infrared sounder radiances

5.5.1 | Multispectral IR radiances

The HIRS instruments are 20-channel infrared radiometers providing information on temperature throughout the troposphere and lower-mid stratosphere and humidity information in the troposphere. An overview is given in Figure 5. HIRS radiances have been used from all NOAA

and MetOp platforms (with the exception of NOAA-16 and NOAA-18) since the TIROS-N satellite.

The SSU instruments provide information on stratospheric temperatures, and have three channels with peak sensitivities in the range 1–15 hPa. ERA5 assimilates SSU observations from NOAA platforms during the period 1979–2006, and the treatment of SSU includes several improvements since ERA-Interim (Section 4.5 and Kobayashi *et al.*, 2009). SSU channel 3 is used to anchor the analysis. In contrast to ERA-Interim, it continued to be used as an anchor with the advent of AMSU-A observations in 1998.

5.5.2 | Hyperspectral IR radiances

Radiances, assimilated as brightness temperatures, have been assimilated from the hyperspectral IR instruments AIRS, IASI and CrIS. An overview is given in Figure 5. AIRS radiances were assimilated from October 2002 (McNally *et al.*, 2006) and by June 2019 radiances from two CrIS instruments (S-NPP and NOAA-20; Eresmaa *et al.*, 2017) and two IASI instruments (Metop-A and -B; Collard and McNally 2009) were also assimilated. The hyperspectral IR sounders, measuring in the thermal IR region of the spectrum, provide information on temperature throughout the troposphere and lower-mid stratosphere and tropospheric humidity. Those channels that provide sufficient information are assimilated using a cloud-detection technique based on McNally and Watts (2003). Observation errors, including inter-channel correlations, have been estimated using observation space diagnostics (Bormann *et al.*, 2015).

Variational bias correction is applied to all assimilated channels with the exception of ozone-sensitive channels at 9.6 μm for AIRS (channel 1088), CrIS (channel 626) and IASI (channel 1585).

5.6 | Geostationary radiances

An overview is given in Figure 5. Infrared radiances from geostationary satellites are first assimilated in ERA5 in May 2001 (Meteosat-7). Since that time radiances from the US GOES series of satellites (specifically GOES-8 to -16), covering America and neighbouring oceanic longitudes, with sub-satellite longitudes of 135°W and 60–75°W, from the Meteosat satellites (-5 to -11) covering much of Europe and Africa at longitudes of 0°–10°E, and MTSAT-1R and -2 as well as Himawari-8 covering the Western Pacific, East Asia and Australia at 140°E have been assimilated. The IR imager instruments on board these satellites measure radiances in the mid-IR water vapour band, in one

to three channels depending on instrument, and provide information on tropospheric humidity. Details of the operational implementation of these radiances can be found in Köpken *et al.* (2003), Munro *et al.* (2004), Lupu and McNally (2012), Letertre-Danczak (2016) and Burrows (2018). These observations have also been shown to improve analyses of wind fields through the ability of 4D-Var to extract dynamical information from the observed advection of water vapour features (Peubey and McNally, 2009).

5.7 | Atmospheric motion vectors

AMVs are winds derived by tracking clouds or water vapour features in a sequence of images obtained from geostationary satellites, or from pairs of images obtained from polar orbiting satellites. AMVs therefore provide information on vector winds throughout the troposphere. Heights are assigned to the derived AMV winds using infrared window channel radiances. These observations are not bias corrected.

ERA5 assimilates winds throughout the period from 1979 to the present (Figure 6). AMVs from the GOES series of satellites, typically located at 135°W and 60–75°W, cover the American continent from 1979. At 140°E the GMS/MTSAT/Himawari series of satellites, supplemented by repositioned GOES satellites, provide near-continuous coverage of the Western Pacific and East Asia since 1979. The Meteosat series of satellites, located at 0–10°E provide coverage over Europe, the Middle East and Africa since 1982. The status of AMV assimilation at Cy41r2 is summarised in (Salonen and Bormann, 2016).

ERA5 assimilates a number of reprocessed AMV datasets (Table 4) from the three major providers and some new datasets (i.e., datasets which did not exist previously, e.g., recently generated polar winds from NOAA LEO satellites operating in the early 1980s). In pre-production testing, significant benefit was obtained by assimilating reprocessed GOES observations (–8 to –13, covering the period 1995–2013). Model background fits to low-level winds (below 400 hPa) were improved by 10–30% relative to those obtained in experiments assimilating the original near-real-time operational datasets.

5.8 | GNSS-RO bending angles

The GNSS-RO bending angles provide high-quality temperature information in the upper troposphere and lower/middle stratosphere. They complement the information provided by satellite radiances, because they have good vertical resolution. GNSS-RO measures an accurate time

delay. The physical retrieval process is well known, so they can be assimilated without bias correction. The latter means that GNSS-RO bending angles are anchor measurements in the VarBC system, and therefore they constrain the bias corrections applied to the radiances. It has been shown that the consistency of lower/middle stratospheric temperatures amongst the global reanalyses has improved since the assimilation of COSMIC GNSS-RO measurements in 2006 (Long *et al.*, 2017).

The reprocessed COSMIC GNSS-RO dataset, provided by UCAR, incorporates improved filtering of the measured phase delays. The improvement was implemented in the near-real-time operational COSMIC RO dataset in November 2009 and resulted in mean departure statistics more consistent with those of Metop-A GRAS. An overview is presented in Figure 6.

5.9 | Scatterometer wind and soil moisture

Backscatter (level 1B) from scatterometers provides information on near-surface vector wind over the global oceans and soil moisture over land. ERA5 is the first ECMWF reanalysis to include remotely sensed observations in a soil moisture analysis. An overview of the usage of scatterometer data is presented in Figure 6.

In addition to ERS-1/-2 and QuikSCAT, ERA5 makes use of wind information from the scatterometer on board Oceansat-2 (2013), and the ASCAT scatterometers on Metop-A/-B (2007–present), which were not used in ERA-Interim. To improve their usage for ocean vector wind, the observation operator now acts on model equivalent-neutral wind at 10 m height rather than on 10 m wind itself (ERA-Interim). For ASCAT and ERS-2 measurements from 22 August 2003 onwards, the relation between wind and backscatter is provided by the CMOD5.n geophysical model function (Hersbach, 2010), while for ERS-2 and (all) ERS-1 observations prior to that date the (bias-corrected) CMOD4 model (Stoffelen and Anderson, 1997) as also used in ERA-Interim, was used by mistake (however, acting on neutral wind). Although departure statistics suggests that this had a limited effect on the surface-wind mean state, it likely had a negative impact on the usage of extreme scatterometer winds for that period.

Soil moisture information is only extracted from C-band, and not from Ku-band scatterometers, for which the shorter wavelength penetrates less deeply into the surface. The usage of a reprocessed product from ERS-1 and -2 brings the entire time series from 1992 to 2006 into consistency with the soil moisture products from MetOp-A/-B ASCAT.

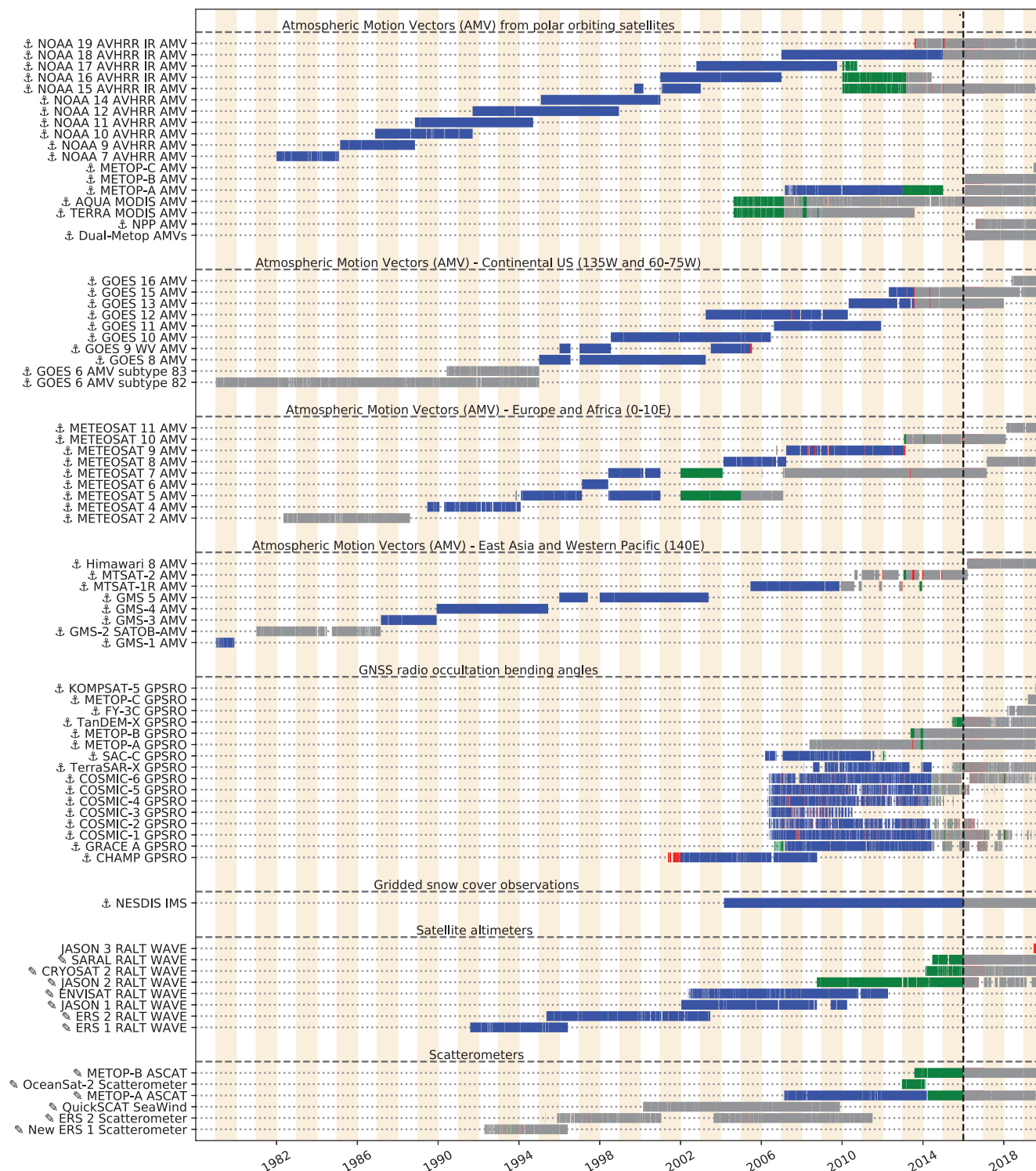


FIGURE 6 Non-radiance satellite observations assimilated in ERA5. The colour scheme is as for Figure 5

Reprocessed ASCAT observations from Metop-A prior to 2014 (both wind and soil moisture) correct inconsistencies in backscatter and brings observations into agreement with the operational data stream from 2014 onwards.

5.10 | Ozone

An overview is provided in Figure 7. All Level-2 ozone products assimilated in ERA5 except METOP-B GOME-2, METEOR-3 and ADEOS-1 TOMS have been improved

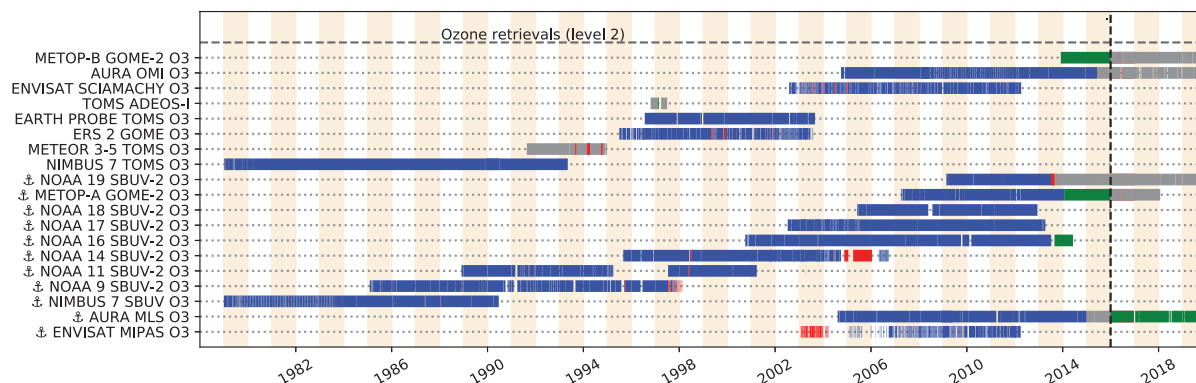


FIGURE 7 Ozone observations assimilated in ERA5. The colour scheme is as for Figure 5

through recent reprocessing efforts. Each reprocessed dataset is based on a single, temporally consistent retrieval algorithm used throughout the period of availability. Reprocessed BUV/SBUV/SBUV-2 datasets (version 8.6; McPeters *et al.*, 2013), from NOAA/NASA, offer an increased number of levels in the vertical relative to earlier data releases, as well as improved consistency over the period 1970–2013 while near-real-time observations were used afterwards. Total column ozone observations from Nimbus-7, and Earth Probe TOMS (v8.0) were used in conjunction with an earlier version for METEOR-3 and ADEOS-1 TOMS observations. The reprocessed limb ozone profiles from MLS (v3) were assimilated until December 2014 when the assimilation was switched to the near-real-time product.

For European instruments, several algorithms and datasets were available at the start of ERA5. Initiatives such as the European Space Agency–Climate Change Initiative (ESA-CCI) were instrumental in providing long-term data records with improved inter-satellite consistency and uncertainty characterisation. Dedicated round-robin assimilation experiments were performed for datasets for which multiple algorithms were available (i.e., OMI, SCIAMACHY, MIPAS, GOME, and GOME-2) aiming at identifying the best candidates for ERA5. Detailed assessment studies were also performed to evaluate possible synergies from using ozone products derived from instruments with different characteristics (Dragani, 2016).

Additional information on ozone in ERA5 is provided by ozone-sensitive channels of the nadir-viewing infrared sounders (HIRS, AIRS, IASI and CrIS; Dragani and McNally 2013). ERA5 uses the available operational NWP datasets for these sensors.

Total column ozone observations are subject to VarBC (Dragani, 2009). The two-parameter bias model corrects for a global bias and temporal-spatial systematic biases that vary with solar elevation angle. Observations of partial ozone layers from nadir- and limb-viewing instruments are not bias corrected and act as anchors.

Ozone observations from SBUV are not assimilated in the EDA prior to 2000. Their initial usage in this component had been shown to lead to an unexplained high ensemble spread over the polar night, which largely reduced the weight assigned to the model first guess, and led to anomalously high values of (total column) ozone (Hersbach *et al.*, 2018).

5.11 | Altimeter wave height

Altimeter measurements provide information on significant wave height over the ocean. Observations have been used since the advent of ERS-1 in 1991 (Figure 6). Observations from SARAL/AltiKa, CryoSat-2 and Jason-2 are based on the operational stream, while those from ERS-1/-2, Jason-1 and Envisat are based on reprocessed products. Altimeter observations are subject to prescribed bias corrections such that wave height estimates emerge unbiased with respect to the ERA5 model, rather than with respect to independent *in situ* measurements. These wave-height-dependent tables were based on a comparison with ocean waves from ERA5-type test runs without using altimeter data. The test period (January to May 2003) focused on Envisat data, while results for other altimeters were determined by inter-calibration of overlaps between the various instruments.

6 | OBSERVATION-BASED GRIDDED FORCINGS AND BOUNDARY CONDITIONS

Besides information from sub-daily observations, the IFS relies on climatological information, such as those which influence forcing from the radiation scheme and the prescription of SST and sea ice over the global oceans. For ERA5 a special effort was made to include state-of-the-art

TABLE 6 Global mean energy budgets ($\text{W} \cdot \text{m}^{-2}$) according to Trenberth *et al.*, (2009), ERA-Interim and the ensemble mean of ERA-20CM, averaged from March 2000 to May 2004 for Trenberth *et al.*, (2009), and from 1989 to 2008 for ERA5, ERA-Interim (both based on the first 12-hr forecasts) and ERA-20CM

Model	Trenberth <i>et al.</i> (2009)	ERA-20CM	ERA-Interim	ERA5
Incoming solar radiation (TSI/4)	341.3	340.4	344.2	340.4
Net absorbed solar radiation (ASR)	239.4	240.9	244.3	242.7
Outgoing long-wave radiation (OLR)	238.5	240.6	245.5	242.2
TOA net radiation in ($\mathbf{R}_T$)	0.9	0.3	-1.2	0.4
Net energy absorbed by surface ($\mathbf{F}_S$)	0.9	1.9	6.9	6.1
Atmosphere net ($\text{TEI} = \mathbf{R}_T - \mathbf{F}_S$)	0.0	-1.6	-8.1	-5.6

datasets that describe well the low-frequency variability of the climate system. A large part of this work was prepared during the ERA-CLIM project.

6.1 | CMIP5 radiation forcing terms

The provision of the total solar irradiance (TSI) is very important, as well as the provision of fields of aerosols, greenhouse gases and ozone. A reanalysis spanning several decades requires that such fields follow the observed 20th and 21st century evolution. Within the ERA-CLIM project, state-of-the-art standardized sets of such long-term forcing fields from the WCRP initiative CMIP5 were implemented as options in the IFS. They were first tested in an ensemble of century-long model integrations (ERA-20CM). These modifications are shared in ERA5. Details may be found in (Hersbach *et al.*, 2015). This is an improvement on ERA-Interim, which, for example, omitted the occurrence of stratospheric sulphate due to major volcanic eruptions.

The average effect of these forcings on global radiation budgets averaged from 1989 to 2008 is displayed in Table 6 where, as in Berrisford *et al.* (2011), values are compared with Trenberth *et al.* (2009). From this it directly emerges that the lower value of TSI for ERA5 (based on rescaling to match the Total Irradiance Monitor instrument (Lean *et al.*, 2005), first used at ECMWF in the Seasonal System 4 implementation (Molteni *et al.*, 2011)) compares considerably better with the estimates from Trenberth *et al.* (2009) than ERA-Interim does, which by mistake used values which were too high. The net energy input at the TOA, which results in a net global warming, agrees within known uncertainties (Allan *et al.*, 2014; Trenberth *et al.*, 2014); this is in contrast to ERA-Interim which has the wrong sign. However, as for ERA-Interim, the net energy absorbed by the surface is far too large. This leads to a net energy loss of the atmosphere of about $5.6 \text{ W} \cdot \text{m}^{-2}$ ($8.1 \text{ W} \cdot \text{m}^{-2}$ for ERA-Interim). Apparently the assimilation system

systematically adds energy, which is then deposited into the surface during the connecting short forecasts, which emerges as the diagnosed loss (Mayer and Haimberger, 2012). The loss of $-1.6 \text{ W} \cdot \text{m}^{-2}$ for ERA-20CM results from an unknown error in the calculation of the post-processed 2D fields for energy budgets within the IFS, since for these model-only runs the energy in the atmosphere should rise by only about $0.01 \text{ W} \cdot \text{m}^{-2}$ (associated with global warming). Although the magnitude of this deficit varies with model cycle, this would suggest that the actual energy imbalance in ERA5 is in the order of $4 \text{ W} \cdot \text{m}^{-2}$.

A more detailed picture is presented in Figure 8, which shows the evolution from 1979. Figure 8a shows that the response from the El Chichón and Pinatubo eruptions is clearly captured by ERA5 and ERA-20CM, but missed by ERA-Interim. Responses from El Niño events are captured by all. At the TOA there is no obvious and significant long-term change. This is in sharp contrast to the surface (Figure 8b), and the resulting net loss in energy (Figure 8c) is worse when going further back in time. This could be the result of larger systematic increments. Around 2010, ERA5 for a brief time almost reaches the 'energy-neutral' state of ERA-20CM. A detailed study on the origins and evolution of sinks (and sources) is required. For ERA-20C and CERA-20C the Total Energy Input (TEI) is remarkably good and comparable to that for ERA-20CM.

6.2 | Sea-surface boundary conditions

In ERA5, conditions for SST and SIC are provided by existing level-4 (i.e., gap-less) gridded datasets. ERA-Interim had used partly what was used in ERA-40 (Fiorino, 2004) and subsequently what was used in the operational medium-range forecasting system at the time. Details may be found in table 1 of Dee *et al.* (2011). For ERA5 a careful selection procedure was conducted (Hirahara *et al.*, 2016).

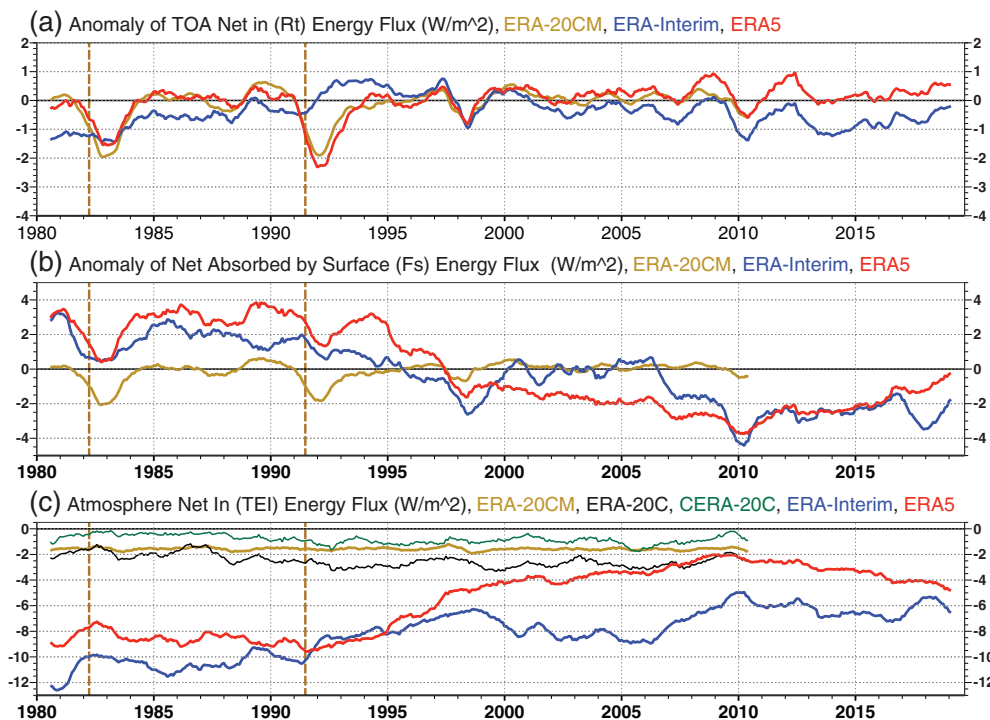


FIGURE 8 Evolution of one-year moving average of energy budgets in ERA5 (red), ERA-Interim (blue) and ERA-20CM (gold), for (a) the TOA Net-in radiation (relative to 1989–2008), (b) Net Absorbed Surface radiation (relative to 1989–2008) and for (c) the Atmosphere Net flux. The latter includes ERA-20C (black) and CERA-20C (green). The vertical ochre dashed lines indicate the eruption dates of El Chichón and Pinatubo

The goal was to compile a dataset from 1950 onwards that is

- (a) as accurate as possible at each moment in time,
- (b) has quality suitable for climate applications, for example exhibiting no noticeable breaks at transitions between datasets, and
- (c) is able to provide timely data for the ERA5 continuation close to real time.

For SST various flavours of the Met Office Hadley Centre HadISST2 product (J.J. Kennedy, 2016, personal communication) were considered (as developed within the ERA-CLIM project and used in the ERA-20CM, ERA-20C and CERA-20C centennial products), as well as the Climate Change Initiative (ESA CCI) SST v1.1 (Merchant *et al.*, 2014), to be combined with the Met Office OSTIA product (Donlon *et al.*, 2012) used in the ECMWF medium-range forecasting system since 2007. For SIC the EUMETSAT OSI SAF reanalysis product (v409a; Eastwood *et al.*, 2014) and various flavours of the HadISST2 sea ice product (Titchner and Rayner, 2014) were considered, to be combined with the operational OSI SAF product that is also part of the OSTIA product.

As a result of this study, the choices for ERA5 are displayed in Table 7. As mentioned in Section 2.4, the SST and SIC for the ERA5 EDA follow a perturbation method as described in Hirahara *et al.* (2016). ERA5 uses the OSTIA product at the appropriate validity date. This is in contrast to the ECMWF operational forecasting system, where, due

to the stricter analysis cut-off time, only the product from the previous day is available. The long-term evolution of SST and SIC as used in the ERA5 HRES is displayed in Figure 9. The global-mean SST shows the impact of global warming from the mid-1970s, as well as the influence from El Niño events and major volcanic eruptions. Arctic sea ice shows a general decline over time, especially during summer.

Compared to ERA-Interim, ERA5 uses enhanced quality control to deal with spurious coastal sea ice in the Northern Hemisphere. The limit of clearing ice when SST exceeds 1 °C (ERA-Interim) was raised to 3 °C. It is very effective and not overactive. For the modern period, where OSTIA is used, this check is not required. Unfortunately this enhanced check was only introduced quite late in the production, to counteract the lack of any quality control in the initial part of the production, which appeared to give rise to spurious ice over the Gulf of Finland and several other locations during each summer prior to 2008. For the HRES final production, this was resolved through repair runs as indicated in Table 3. For the EDA no repair runs were conducted. As a result the EDA does contain spurious ice during summer months in the periods from January 1979 to June 1981, April 1986 to September 1988, August 1993 to August 1995, and from January 2000 to August 2007 (the end of usage of the OSI SAF sea ice product). The effect on the ERA5 ensemble spread is found to be relatively minor. An example is provided in Figure 10 which confirms (a) a good-quality ice product for the HRES, (b) a degraded estimate for the EDA, but (c) with a mild effect on

TABLE 7 SST and SIC products as used in ERA5

Time period	Sea Surface Temperature	Sea Ice Concentration	Grid (deg)
January 1949–December 1960	HadISST2.1.0.0 (monthly)	HadISST2.0.0.0	0.25 x 0.25
January 1961–December 1978	HadISST2.1.1.0 (pentad)	HadISST2.0.0.0	0.25 x 0.25
January 1979–August 2007	HadISST2.1.1.0 (pentad)	OSI SAF (409a)	0.25 x 0.25
September 2007 onwards	OSTIA	OSI SAF oper	0.05 x 0.05

Note: All products are daily, although ‘pentad’ is based on 5-daily and ‘monthly’ (and all HadISST2 ice datasets) on one-monthly analysis windows. HadISST2 sea ice is gridded on 0.25°, although the native resolution is 1°. The OSI SAF (409a) 10 km polar stereographic grid is regridded in-house to facilitate its usage. OSTIA is used for the correct date (see text for details).

FIGURE 9 Time series of (a) global sea surface temperature (°C) and (b) Arctic sea ice cover (percent) as used in the ERA5 HRES assimilation for data that have been released at the time of writing (from 1979 onwards), and produced but not yet released (1950–1978), for monthly (blue) and yearly (red) running-mean averages

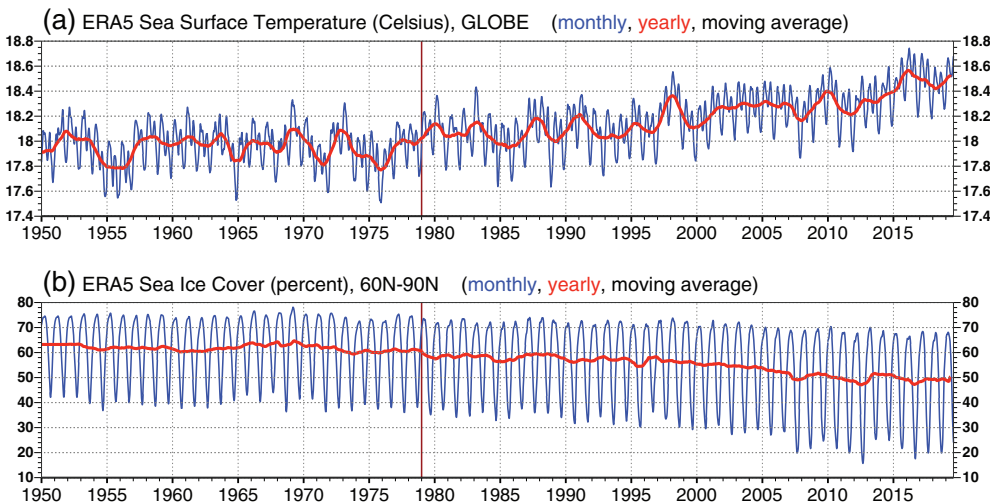
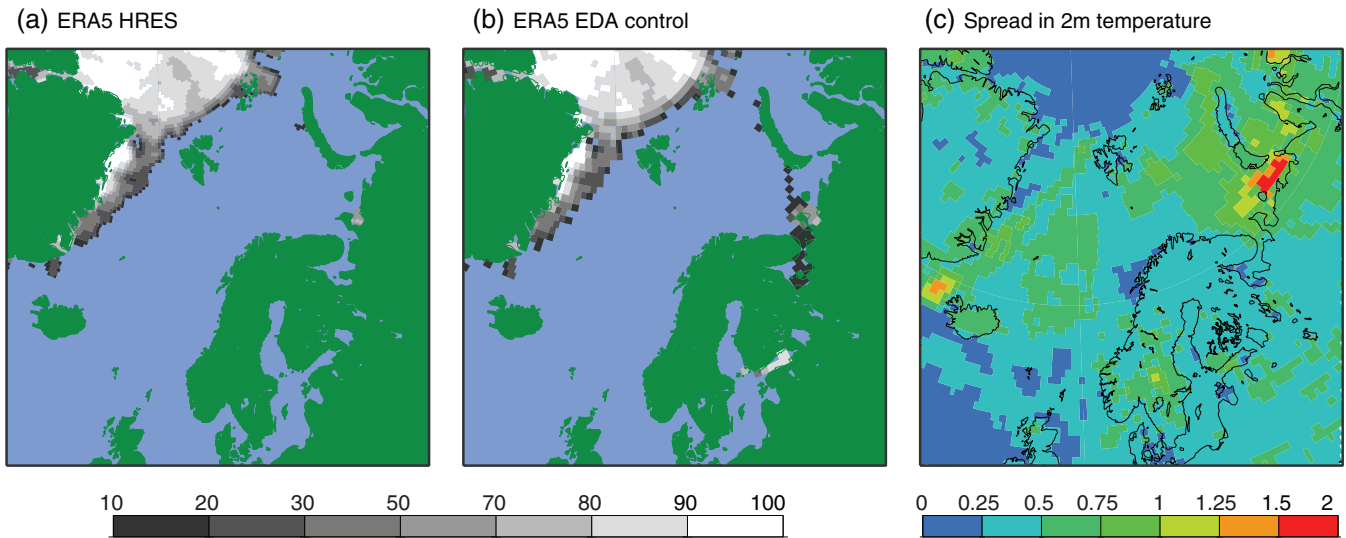


FIGURE 10 Sea ice cover (percent) on 27 July 2006 from (a) ERA5 HRES, (b) the ERA5 EDA control and (c) the daily-mean ensemble spread of 2 m temperature (K)

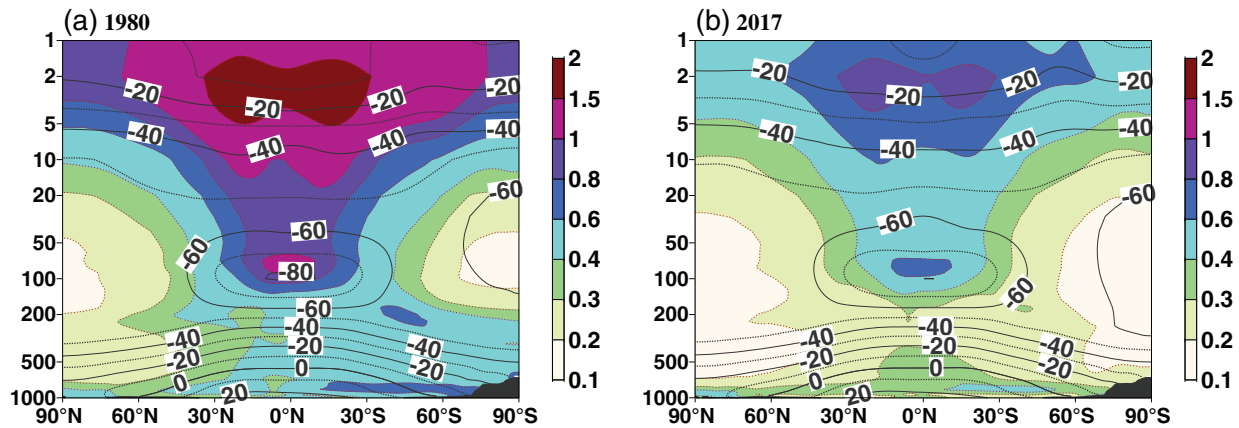


FIGURE 11 Zonal-mean cross-section on a logarithmic pressure scale (hPa) of ERA5 control (contours) and ensemble spread (colour shading) of temperature ($^{\circ}\text{C}$) averaged over (a) 1980 and (b) 2017. The top of the black areas at the bottom mark the zonally and yearly averaged surface pressure (hPa)

the ensemble spread for 2 m temperature. The large spread south of Novaya Zemlya appears for a number of days in July 2006 and indicates a particular sensitive location for that month.

ERA5 does not impose the 100% ice concentration north of 82.5°N as was applied to ERA-Interim between January 1989 and February 2009 (Figure 10c), which was particularly poor in September 2007 when sea ice retreated beyond that perimeter. In addition, in ERA5 the minimum non-zero sea ice fraction was lowered from 20% (ERA-Interim) to 15%, since the latter coincides with the usual threshold for defining ice edge.

7 | DATA ASSIMILATION DIAGNOSTICS

7.1 | Evolution of the ERA5 uncertainty estimate

The ERA5 EDA spread among the ten ensemble members can be interpreted as a measure for the uncertainty in the HRES estimates. It mainly samples random errors, although perturbations in the HadISST2.2 SST dataset do contain long time correlations. The perturbed model tendencies can also lead to small systematic differences in model climate for perturbed members with respect to the unperturbed control. However, systematic differences between perturbed members are small. For example, in 2018 the mean difference (globally) between the nine ensemble members and the control member, for the variables temperature, relative humidity and u -component of wind at the 500 hPa level were 0.006 K, 0.3% and 0.4 cm s^{-1} respectively. Values of a similar magnitude were found for 1980.

For this reason the ensemble spread should mainly be used as a guide for the quality of representing the correct synoptic situation at a given time, rather than for long-term and/or large-scale averages, such as the global mean 2 m temperature. For such quantities any systematic errors in the ERA5 mean state may become significant and are not represented by the ensemble.

The magnitude of the ensemble spread is closely related to the quality of the observing system. An example is provided by Figure 11, which shows cross-sections of the one-yearly and zonally averaged (synoptic) ensemble spread in temperature for (a) 1980 and (b) 2017. The year 1980 falls in the early-satellite era with upper-air sensitive data predominantly from the TOVS satellites and radiosondes. This explains why ensemble spread is the lowest over the Northern Hemisphere troposphere and lower stratosphere where radiosondes are mostly available. The spread in the tropical upper stratosphere is quite large. In 2017 the observing system is much more comprehensive; about 30 times more observations are assimilated. As a result ensemble spread has tightened up almost everywhere.

The 40-year evolution of monthly and globally averaged spread is displayed in Figure 12. The reduction over time of temperature (Figure 12a) at 3 hPa is quite large. Major improvements seem to coincide with the advent of ATOVS data in 1998, and increasing numbers of GNSS-RO data in 2006. At other heights improvement of temperature estimates is more gradual. In the troposphere, ensemble spread is smallest in the mid to upper part and maximal at around 850 hPa. For zonal wind (Figure 12b), improvement over time is also largest for the upper part of the stratosphere, with the most radical changes marked around 1998 and 2006 as well. Spread is lowest near to the surface.

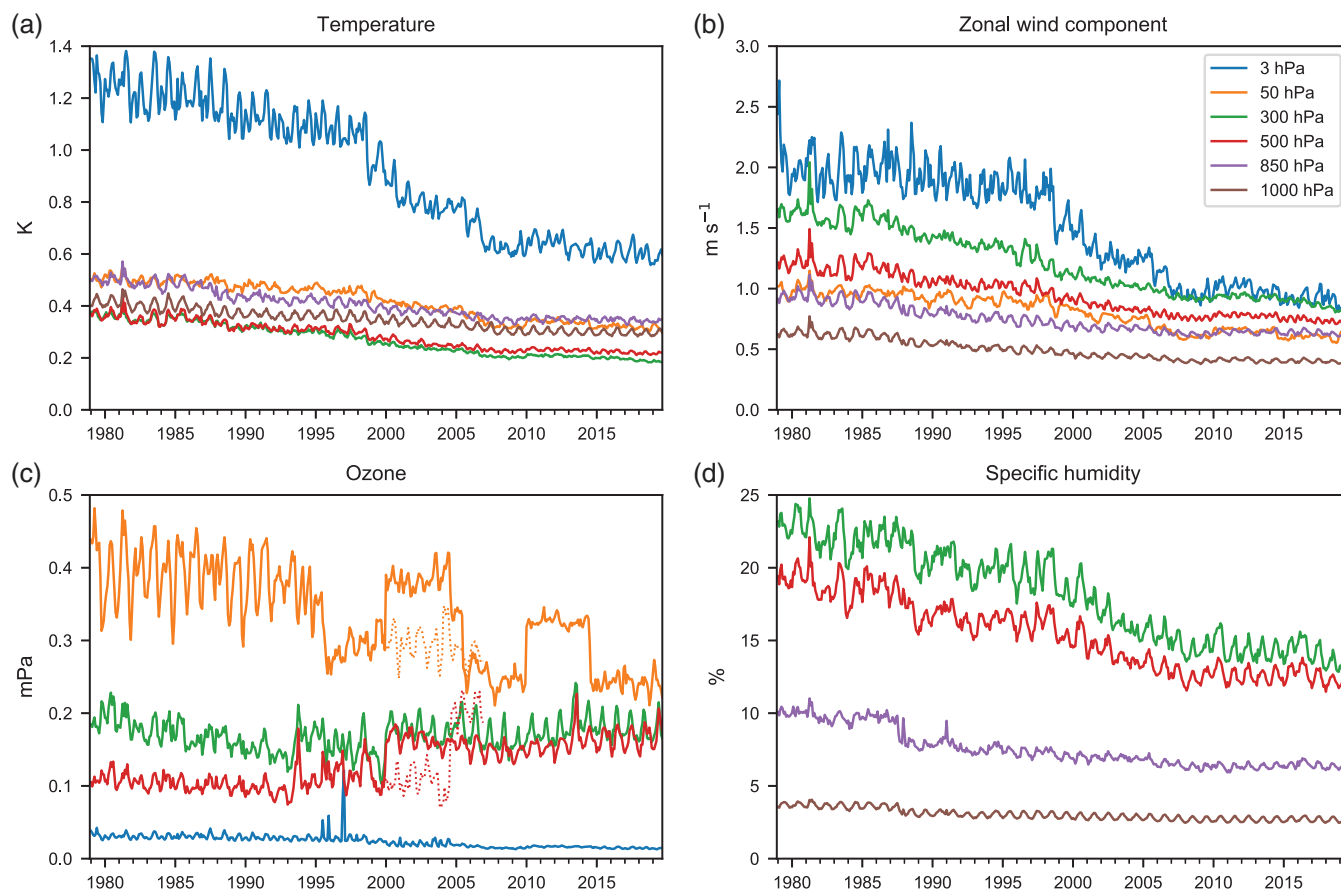


FIGURE 12 Time series of monthly and globally averaged ERA5 ensemble spread from 1979 to 2018 at indicated pressure levels for (a) temperature (K), (b) zonal wind (m s^{-1}), (c) ozone (partial pressure in mPa) and (d) specific humidity (in percent of the 1981 to 2010 mean value at the pressure level in question). For (c), ozone estimates from ERA5.1 (dotted lines) have been included for 50 and 300 hPa

Ozone (Figure 12c) shows two enhanced periods for spread at 50 hPa which start at the consolidation seams of January 2000 and January 2010. This inflation is not related to differences in data usage. Instead, these are the result of a configuration change of the EDA configuration in March 2017, as explained in Section 2.4. These segments originate from the production streams that had been produced prior to that date. For the extension ERA5.1 (also displayed in Figure 12c) this increase is not observed.

The spikes in spread at 3 hPa in 1995 and 1997 are related to anomalously high values of ozone in the polar night at those heights and were created by erroneous analysis increments. These stem from the same mechanism as described in Hersbach *et al.* (2018), although in a much milder form, and hardly affect estimates of total column ozone. Somewhat surprisingly, the spread for tropospheric ozone increases over time.

For tropospheric humidity (Figure 12d) the spread decreases over time, with a sharp drop in late 1987 at 850 hPa coinciding with the start of microwave imager assimilation with the first SSM/I. Relative spread is lowest near the surface.

7.2 | Fit to observations

The increasing confidence of ERA5 HRES estimates over time, as apparent from the EDA ensemble spread, can be verified by a comparison with observations. An example is provided in Figure 13 for ERA5 analysis ocean wave height versus independent buoy observations. Compared to ERA-Interim, the scatter index (normalized standard deviation) is in addition much lower (i.e., improved), where both are verified against the same dataset. One can also infer such evolution from the statistics of departures (defined in Equations (1) and (2)) that are readily available from the assimilation system. Analysis departures are not useful here, since these express the extent to which the analysis has drawn to the observations, rather than providing an independent assessment of performance. As an alternative, first-guess departures are much more informative, since these represent the comparison with observations just prior to their assimilation. Therefore these are more (but not totally) independent. The 40-year evolution of the standard deviation of such first-guess departures is displayed in Figure 14 by red curves for

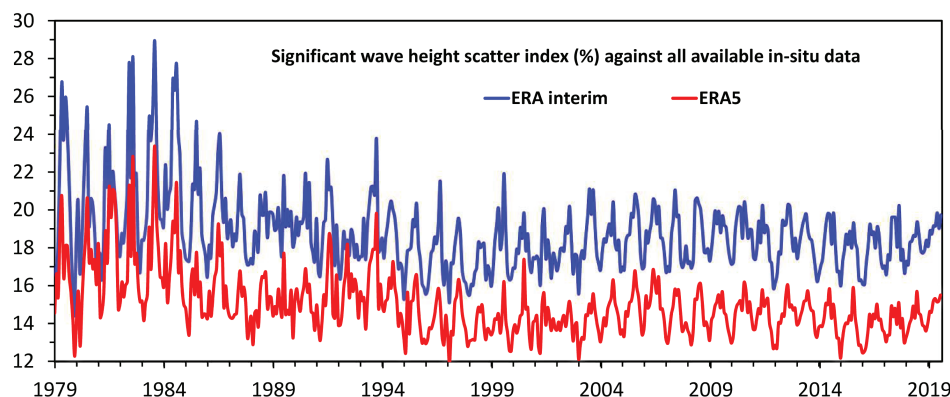


FIGURE 13 Scatter index (% , lower is better) with respect to independent buoy wave height observations for ERA-Interim (blue) and ERA5 (red) analyses

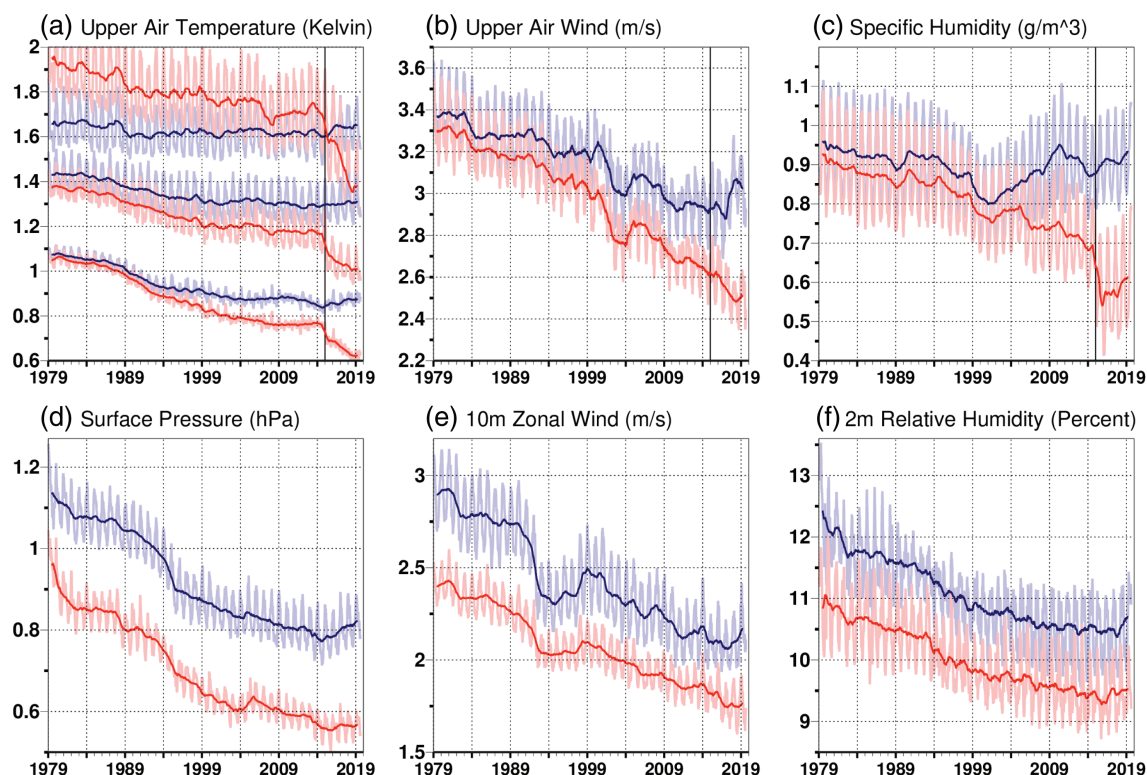


FIGURE 14 30-day mean (weak colours) of the standard deviation of first-guess departures and 360-day mean (strong colours), for used data in ERA-Interim (blue) and ERA5 (red) for (a) upper-air temperature from radiosondes (from top to bottom in the panel) within ± 25 hPa of 50, 850 and 400 hPa; (b) upper-air zonal wind from radiosondes, dropsondes and PILOTs; (c) upper-air humidity from radiosondes and dropsondes; (d) surface pressure from SYNOP, buoys, ships and METAR; (e) 10 m zonal wind over sea from various in-situ sources; and (f) 2 m relative humidity from SYNOP. The statistics for (b, c) are number-weighted averages of the standard deviation over pressure bands of 50 hPa throughout the vertical, without making corrections for any pressure-dependent biases. The vertical black line marks the start date of the usage of BUFR TEMP data in ERA5 from 1 January 2015 (Figure 4)

upper-air data from assimilated radiosonde, PILOT and dropsonde data (a–c) and for near-surface observations from SYNOP, buoys, ships and METAR data (d–f). The fits do improve considerably over time. However, it should be realized that these statistics measure the joint (random) error between the reanalysis short forecast and observations, and the latter, generally, also improve over time. In addition, changes in sampling can also have an effect. A clear example is the apparent massive improvement in

statistics for upper-air temperature once data from BUFR radiosondes are assimilated from 1 January 2015. This is partly explained in Ingleby *et al.* (2016). Where previously observations were more commonly used near significant levels which are typically more difficult to represent, the BUFR radiosonde data sample the vertical more densely and homogeneously (figure SB1 of that paper). Figure S4 here shows the degree to which the spread (expressed as a standard deviation) in the ERA5 first-guess departures,

for radiosonde temperatures at 500 hPa, is consistent with the spread in the EDA and the assumed observation errors.

For ERA-Interim, corresponding fits are displayed by the blue curves in Figure 14. Note that the comparison between ERA5 and ERA-Interim is not pure, since the verification sets (the used datasets) differ, as well as their bias adjustments. However, for the report types considered here, the differences are reasonably small prior to January 2015 (when ERA5 starts using BUFR data). For ERA5 the fit to observations is better over the entire troposphere, in particular for the near-surface parameters. The stratosphere is an exception, where the fit for ERA5 is actually worse for temperature (Figure 14a) and wind (not shown). The transition point is around 100 hPa. One reason for this is that, despite 10 years of model development, Cy41r2 exhibits a larger lower stratosphere cold bias. Therefore, the analysis system is required to apply larger increments. In addition, in Cy41r2, approximately 20–30% lower weight (Table 5) is given to radiosonde stratospheric temperatures, so the larger misfit is weighted less. Thirdly, and this applies from 2000, the smaller correlation lengths in **B** are less able to spread the initially sparse information horizontally. This may have prevented part of the reduction for temperature at 50 hPa around 2006 (when large amounts of GNSS-RO data become available) to occur earlier.

The improvement over time is in line with the evolution of the skill of re-forecasts, as displayed in Figure 1.

7.3 | Mean observation departures

In general the time evolution of departure statistics provide a sensitive health check for the ingested observations. Spikes in either their mean, standard deviation or bias estimates typically indicate data problems. Alternatively, these can also relate to problems in the analysis products themselves. For this reason such departure statistics are closely monitored in the ERA5 production streams. The amount of statistics is immense (several thousands): one time series per satellite channel, and a range of heights for upper-air data, each stratified with respect to the globe, and sub-areas. A comparison with similar statistics from a reference (ERA-40 prior to 1979, ERA-Interim, and for the more recent period, the ECMWF operational NWP assimilation system), is made as well. Identified problems are investigated. When a simple solution is available (typically a temporal blacklisting, or re-initialized bias estimates; Hersbach *et al.*, 2018), the production suite in question is rewound to a date prior to the event (typically a month or two). This procedure has been applied several times during the production, and contributes to the quality assurance of the final product.

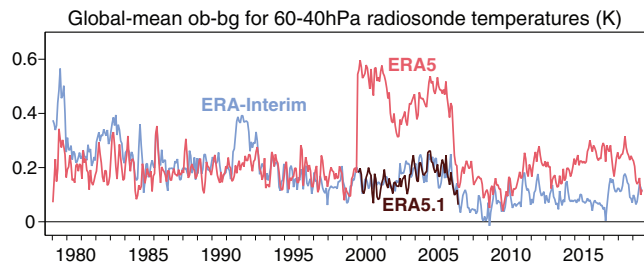


FIGURE 15 Monthly average observation–background differences from 1979 onwards for all assimilated bias-adjusted radiosonde temperature data (K) between 40 and 60 hPa, for ERA-Interim, ERA5 (based on 1979-**B**_{cli} before 2000 and 41r2-**B**_{cli} afterwards) and ERA5.1 (using 1979-**B**_{cli} from 2000–2006)

An example of a misfit that is related to the reanalysis product, rather than to the ingested observations, is displayed in Figure 15. This shows the mean bias-adjusted departures of radiosonde data from the ERA5 (red) and ERA-Interim (blue) background forecasts, averaged over all assimilated radiosonde data from 40 to 60 hPa. Departures for ERA5 suffer from a large jump at the transition point at January 2000 (Table 3). As mentioned above, this is the result of the usage of 41r2-**B**_{cli} which is less able to correct the model lower-stratospheric cold bias. The resulting (cold) model background leads to more positive observation minus background departures. The ERA5 fit to the radiosonde data considerably improves in 2006 once substantially more GNSS-RO data are assimilated. As mentioned in Section 3, the stream that was consolidated prior to January 2000 has been extended up to the end of 2006 (as ERA5.1). It continues the usage of 1979-**B**_{cli} and the positive effect on departure statistics (brown curve in Figure 15) is striking, providing a more-or-less seamless transition when the GNSS-RO counts considerably increase.

The upward spike in the ERA-Interim radiosonde fit in mid-1979 is due to a slow adaptation to a change in calibration of the MSU radiances from TIROS-N. This is an example where one type of observation has had an adverse effect on the reanalysis product which then emerges in the departure statistic of another type of observation. Guided by the departure-based monitoring, the radiance bias correction was adjusted for ERA5 to account for this special case. The positive departures seen for ERA-Interim, which are related to a partial response to the Pinatubo eruption in 1991, are not evident in ERA5.

7.4 | Analysis increments

After observation bias adjustments, the assimilation system implicitly assumes unbiased error characteristics. As

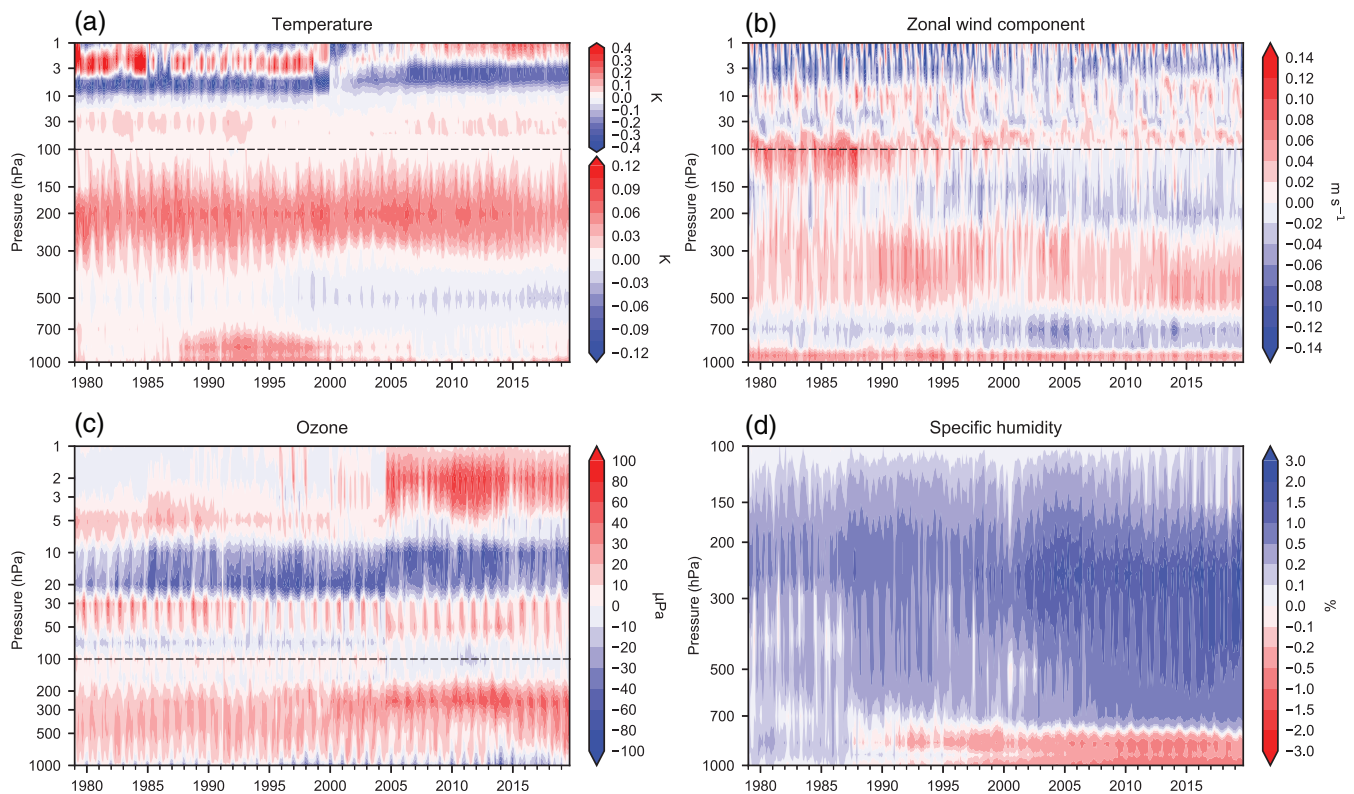


FIGURE 16 Profiles of monthly and globally averaged increments between ERA5 analyses and first-guess fields at 0000 and 1200 UTC for (a) temperature (K), (b) zonal wind ($\text{m}\cdot\text{s}^{-1}$) (c) ozone (partial pressure) and (d) specific humidity (percent), on a logarithmic pressure scale from 1,000 to 100 hPa, and (a–c) on an additional scale from 100 to 1 hPa

such, long-term averages of analysis increments δx from the minimum of Equation (4) should be small. Systematic deviations typically indicate an average conflict between observations and the forecast model that is used both within the 4D-Var trajectories and the short forecast linking to the next assimilation cycle. VarBC can alleviate this incompatibility, but only as long as relative biases can be traced back to the observations, rather than to the model. Abrupt changes in mean analysis increments can usually be related to changes in the observing system. The standard deviation of increments indicates the work done by the assimilation, and is less informative about the health of the system.

Pressure–time diagrams for monthly and globally averaged mean increments are displayed in Figure 16. In general, these are significantly lower than for ERA-Interim (shown in Figure 17). Exceptions are ozone and mid-tropospheric humidity where magnitudes are comparable. For temperature (Figure 16a) the positive mean ERA5 increment in the lower stratosphere is related to the model cold bias in this region.

Opposite increments at the top of the stratosphere indicate a bias with respect to anchoring satellite observations which peak at those heights. In this respect, the effect of the introduction of the anchoring AMSU-A channel 14

observations in September 1998 is smaller than it was for ERA-Interim. The reason for this is that ERA5 continues to use SSU channel 3 as an anchor (from its first assimilation in July 1979 on NOAA 6), while it is bias-corrected in ERA-Interim once AMSU-A emerges. The jump in January 2000 at these heights is the consequence of the change in B_{cli} . This is in line with the time series for radiosonde departures, as shown in Figure 15. The emerging positive temperature increments in the lower troposphere between 1988 and 2007 are over the oceans and a drying tendency is seen starting in 1988 for humidity and continuing to the present day (Figure 16d). These issues coincide with the advent of microwave imagers, which are known to warm and dry the analysis at 850 hPa over the ocean (Geer *et al.*, 2017). The exact mechanisms causing these biases remain unclear. The mid-troposphere negative temperature increments are mainly over land (Figure 16b). Analysis increments for specific humidity are in general smaller over land (not shown) than over oceans, while for zonal wind the opposite applies. For upper-air wind (Figure 16b), the reason for the reduction of positive increments around 100 hPa around 1987 is unclear. For ozone the sharp transition in the upper stratosphere is related to the start of using AURA MLS and OMI data. Pressure–time diagrams for monthly and globally averaged

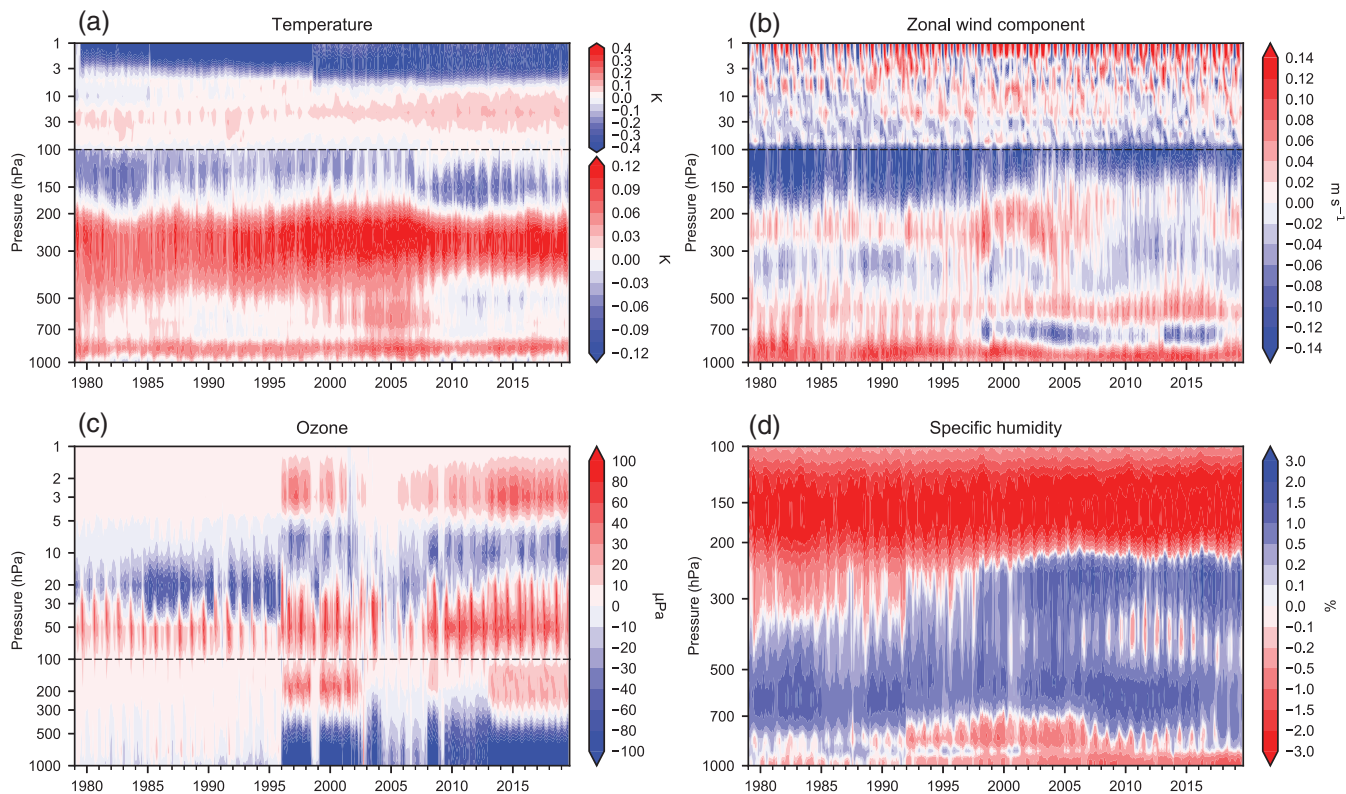


FIGURE 17 As Figure 16, but for ERA-Interim

standard deviations of analysis increments are shown in Figure S3.

8 | IMPROVED RESOLUTION FOR SYNOPTIC VARIABILITY

The considerable increase in resolution of ERA5 allows much more detail to be represented both in space and in time (hourly output). Many examples can be provided or have already been reported in the literature. For example, the positive effect on the quality of near-surface wind is described in Olauson (2018). Re-forecasts from ERA5 reanalysis fields are more skilful than those from ERA-Interim, as was demonstrated in Figure 1. ERA5 improves the representation of tropical cyclones. Central pressures are lower, and closer to those of the ECMWF operational HRES analysis than is the case for ERA-Interim (F. Prates, 2018, personal communication). An illustration of this is provided in Hersbach (2019) for hurricane *Florence*, which hit the east coast of the USA on 15 September 2018.

The benefit of hourly temporal resolution is illustrated in Figure 18 which shows the evolution of storm *Lothar* at 0900, 1000 and 1100 UTC on 26 December 1999 when it swept rapidly through Western Europe after its landfall in Brittany, early in the morning of that day (Wernli *et al.*,

2002). *Lothar* was the first of two December 1999 storms and was followed by *Martin* 36 hr later. Both of these storms caused considerable damage. The ERA5 hourly resolution presents a detailed view of this rapidly evolving storm, as is apparent from Figure 18a, b, c. Although at 0900 UTC the ERA5 minimum pressure is about 5 hPa higher (ensemble spread 2.1 hPa) than the reported minimum (Hewson and Neu, 2015) (which also has an associated uncertainty) and also somewhat misplaced, the match is within 1 hPa at 1000 and 1100 UTC (spread 0.5 hPa at 1200 UTC), and the analysed position is highly accurate at 1100 UTC. ERA5 provides a detailed view of wind gust (one of the available parameters, originating from the short forecasts linking analysis windows), with maximum values up to 42 m s^{-1} in the Black Forest and French Alpine area. Maximum observed gusts in the Black Forest area were 59 m s^{-1} (DWD, 2000).

A synoptic example for the stratosphere is displayed in Figure 19. It maps the exceptional break-up of the southern polar vortex in late September and early October 2002. Water vapour is close to a conserved variable at the level shown, and no humidity observations are assimilated at this level. It thus serves as a convenient tracer to illustrate the short-term dynamics of the polar stratospheric vortex, despite quantitative limitations. Prior to break-down, the vortex is characterised by a relatively high abundance of water vapour. The performance of the operational

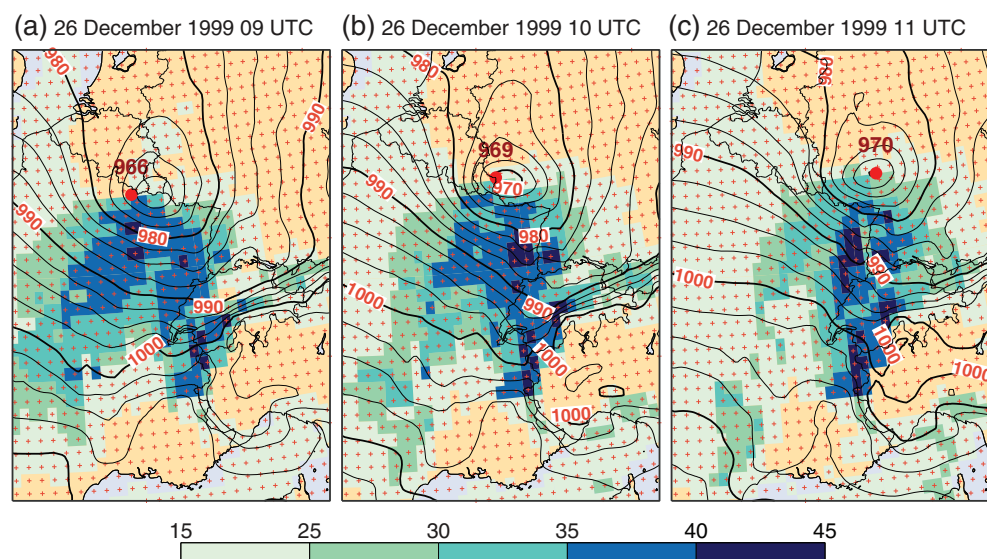


FIGURE 18 ERA5 analysis of mean sea level pressure (contours, in steps of 2.5 hPa) and wind gust within the preceding hour (colours, $\text{m}\cdot\text{s}^{-1}$) for storm *Lothar* on 26 December 1999 at (a) 0900, (b) 1000, and (c) 1100 UTC. Pressure minima for ERA5 are 971/969/970 hPa and maximum gusts are 42/41/42 $\text{m}\cdot\text{s}^{-1}$, for 0900/1000/1100 UTC, respectively. Ensemble spread at minimum pressure location is 2.1/0.5 hPa at 0900/1200 UTC, respectively. Red dots and values mark the position and pressure (hPa) of the reported low. Red pluses indicate the ERA5 native grid

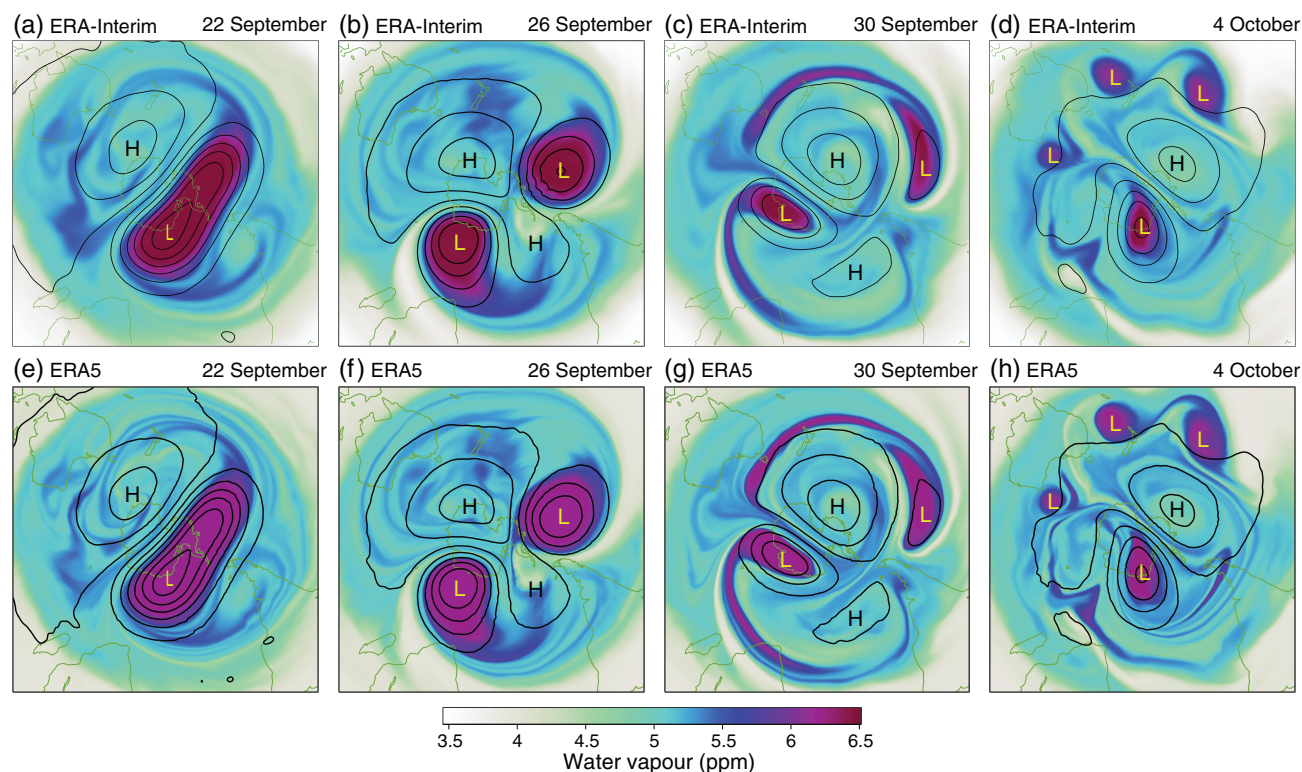


FIGURE 19 Analyses over the Southern Hemisphere of the abundance of water vapour (mole fraction in ppm; shading) and Montgomery potential (contour interval $4,000 \text{ m}^2 \text{ s}^{-2}$) on the 850 K isentropic surface, for (a, e) 22, (b, f) 26 and (c, g) 30 September and (d, h) 4 October 2002, from (a–d) ERA-Interim and (e–h) ERA5

ECMWF system at the time is discussed by Simmons *et al.* (2005).

On the first day shown, 22 September, the polar vortex is already elongated, and has been reduced to a smaller size than usual by a succession of events that strip material

from it. Remnants of this material can be seen within the flanking anticyclones south of Australia and over southern South America. By 26 September the vortex has split completely into two similarly sized parts, and the dominant anticyclone in the Australian sector extends to the

South Pole. Thereafter, the vortex in the Pacific sector is stretched around the northern flank of this dominant anticyclone, and breaks into several smaller vortices due to dynamical instability of the ambient easterly flow. Three of these vortices are evident in the maps for 4 October. The corresponding vortex in the Indian Ocean sector loses less material in this way, and is eventually re-established as the primary polar vortex, albeit a much weaker one than is usual for October. This succession of events was also seen in the operational analyses at the time. Simmons *et al.* (2005) note that the operational analyses were consistent with the 10 hPa temperature and wind measurements from Australian and neighbouring radiosondes during the passage of the smaller vortices around the anticyclone.

ERA5 and ERA-Interim are very similar in their synoptic evolution. The filaments of air drawn from the vortices or entrained around them are sharper and richer in structure in ERA5, as expected given its higher resolution, and the small vortices in the Australian sector are a little stronger. ERA5 benefits from further refinements of the assimilating model's semi-Lagrangian advection scheme (Diamantakis, 2014; Diamantakis and Magnusson, 2016) and is accordingly free of the noise that occurs in the easterly flow on the southern flank of the Pacific sector vortex in the ERA-Interim analysis for 26 September.

9 | PERFORMANCE OF LOW-FREQUENCY VARIABILITY AND CLIMATE TRENDS

9.1 | Surface air temperature

Since April 2019, ERA5 analyses of 2 m temperature are used as input to monthly summaries published by the C3S. Before that they had been based on ERA-Interim for almost 4 years. Such analyses from ERA-Interim, and earlier ones from ERA-40, have been shown to be of reasonable quality and complementary to the products of conventional analyses of climatological station data (Simmons *et al.*, 2017, and references).

A number of regional and local improvements are found in ERA5. An example is displayed in Figure 20 which shows a considerably improved fit to observations for a location in the Arctic. Annual-mean analysis increments are smaller in most regions. This is largely due to smaller mean background errors, although some differences have been traced to differences in data quality control: ERA5 rejects fewer data as it has fewer instances of large differences from background values.

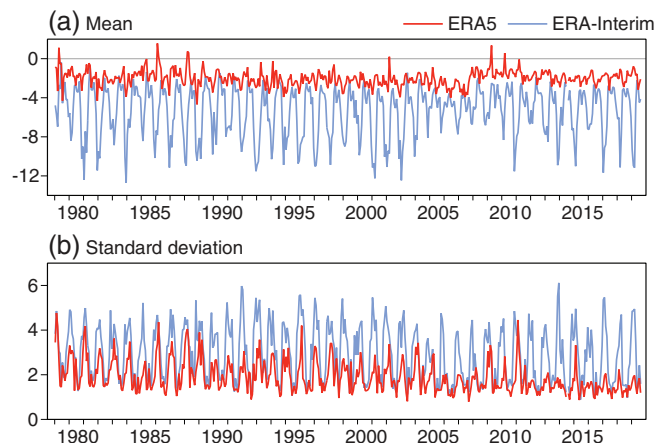


FIGURE 20 (a) Monthly means and (b) standard deviations of 4D-Var (background minus observation) departures of 2 m temperature (K) at Longyearbyen, Svalbard, Norway (78.2°N, 15.5°E) for ERA-Interim (blue) and ERA5 (red)

The agreement found previously among various datasets including ERA-Interim led to the expectation that time series of global-mean temperature from ERA5 would not be substantially different to those from ERA-Interim. This is confirmed by Figure 21. The largest differences between ERA5 and ERA-Interim occur in 2005 and 2006, a period when the differences among various datasets are relatively large. This is also a period in which differences in SST analysis are quite large and in which the reanalyses have large anomalies in polar regions that are not sampled well by the conventional analyses. Differences in the relationship between SST and marine air temperature also play a part. Relatively large differences among datasets have happened again recently, but in this case ERA5 and ERA-Interim (which at this time use common SST and sea ice analyses) give very similar results.

The global fit of the analysis to observations shows little drift over time (not shown). The background forecasts have a global cold bias that is largely removed by the optimal interpolation scheme: the monthly-mean analysis fit varies between zero and -0.16 K, and its annual range decreases over time. These variations are small compared with the rise in mean temperature over land, which for the past four decades has been rather larger than the rise in global-mean temperature shown in Figure 21.

9.2 | Global balance

The extent to which the sequence of ERA analyses achieves global balance of quantities such as mass and water provides measures of the consistency of these analyses, and of general progress made in DA and in observational quality and coverage. The approach adopted for

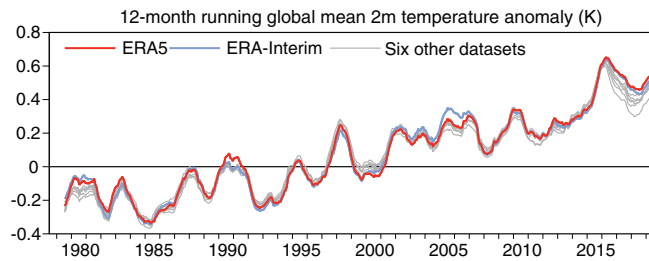


FIGURE 21 Twelve-month running averages from 1979 onwards of global-mean surface air temperature anomalies (K) relative to 1981–2010 for ERA5 (red) and ERA-Interim (blue). Grey lines denote the spread from six other datasets: JRA-55 (Kobayashi *et al.*, 2015); GISTEMP version 4 (Hansen *et al.*, 2010); HadCRUT4 (Morice *et al.*, 2012); NOAA GlobalTemp version 5 (Karl *et al.*, 2015); an infilled version of HadCRUT4 from Cowtan and Way (2014); and a dataset from the Berkeley Earth Surface Temperature Project. For ERA-Interim, values over sea were taken from the first guess rather than the analysis, and prior to 2002 sea points are further adjusted by subtracting 0.1 K

ERA5 and its predecessors contrasts with that adopted by the producers of the MERRA-2 reanalysis (Gelaro *et al.*, 2017), for which the assimilation method preserves, to first order, the physical global constraint of conservation of dry-air mass (Takacs *et al.*, 2016).

Berrisford *et al.* (2011) examined the global atmospheric budgets from ERA-Interim and made comparisons with ERA-40. Although most measures indicated improvement of ERA-Interim over ERA-40, this was not the case for the global budget of dry mass for the years compared (1989–2002). Neglecting the effects of fossil fuel burning, which contribute a variation of approximately 0.01 hPa, dry air mass is expected to be approximately conserved in the atmosphere (Trenberth and Smith, 2005). Given that there are no constraints on the global dry mass in the DA system, the analysed variations of the global dry mass provide a simple measure of the quality of the reanalysis. As shown earlier by Trenberth and Smith (2005), ERA-40 performed much more poorly prior to the early 1970s. This was found by Uppala *et al.* (2005) to be associated with higher analysed surface pressure, particularly over the data-sparse oceans of the Southern Hemisphere, and with lower analysed water vapour prior to assimilation of IR soundings, which began in 1973.

The dry mass of the atmosphere is estimated from the global-mean surface pressure by subtracting the contribution from the water content of the atmosphere. Figure 22 shows the dry mass for ERA5 and ERA-Interim. Neither reanalysis conserves the contribution of dry air to surface pressure to within 0.3 hPa over the whole period. However, they differ in behaviour; ERA-Interim has similar values at the beginning and end of the period, but a spurious

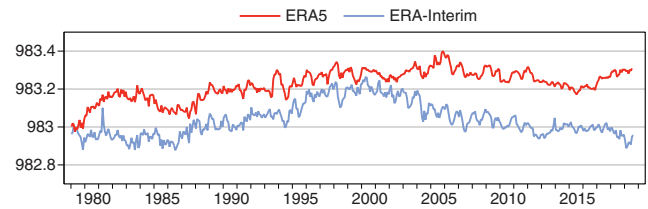


FIGURE 22 Monthly estimates from 1979 onwards of the contribution of dry air to the global-mean surface pressure (hPa) from ERA5 (red) and ERA-Interim (blue), computed by subtracting the contribution from the total water content of the atmosphere from the global-mean surface pressure

rise and fall in dry mass centred around the year 2000. In contrast, dry mass increases quite sharply in the early years of ERA5, but is reasonably uniform after 1990. The range of values is a little larger in ERA5 than ERA-Interim. The ERA5 rise in dry mass in the early and late 1980s is due to rises in the global mean of the analysed surface pressure that are not accompanied by rises in analysed moisture content (not shown). The variations in dry mass in ERA-Interim are likewise due mainly to variations in global-mean surface pressure that are not matched by variations in the contribution from moisture. Reasons for the different variations in surface pressure analyses have yet to be identified. However, given that the spurious variations occur in different periods in the two reanalyses, these problems may well be due to separate causes in each reanalysis.

Aspects of the global hydrological budget are presented in Figure 23. Variations over time and imbalance are generally larger in ERA-Interim than in ERA5. The degree of global annual balance between precipitation and evaporation in ERA5 changes over time. Balance is quite good for a period of twenty or so years from the mid-1990s. The change that brings better balance at that time appears to be in values over sea, where evaporation increases more sharply than precipitation. An earlier gradual decrease in marine precipitation is not matched by a corresponding decrease in evaporation, which also improves the balance.

In the period when ERA5 is in good balance, both precipitation and evaporation increase over sea, but not over land. An increase over sea has been inferred from salinity observations (Durack and Wijffels, 2010), but Figure 23 shows a much smaller increase in marine precipitation from the GPCP (Adler *et al.*, 2003) than from ERA5. Inter-annual variations in net precipitation over land from ERA5 agree quite well with values from GPCP and the underlying data from GPCC² (Becker *et al.*, 2013). Although

²Data downloaded from <https://www.dwd.de/EN/ourservices/gpcp/gpcp.html>; accessed 13 April 2020



FIGURE 23 Twelve-month running averages from 1979 onwards of global mean precipitation (blue) and evaporation (red) rates ($\text{mm}\cdot\text{day}^{-1}$) from (a) ERA-Interim and (b) ERA5. The precipitation estimates from version 2.3 of GPCP are also shown (green). The corresponding contributions to these global averages from (c, d) sea and (e, f) land are also shown. Contributions from land are also shown in (e, f) for estimates from GPCC (brown). The latter are based on GPCC's v.2018 monthly full-data product until the end of 2016, version 6 of its monitoring product for most of the following period, and its first-guess monthly product for the latest 2 months

an improvement over ERA-Interim in this respect, ERA5 exhibits a larger decline in precipitation over land from the 1980s and 1990s to the 2000s than is the case for ERA-Interim (discussed by Simmons *et al.*, 2014). Such a decline is not seen in the GPCC and GPCP data, and was not expected in ERA5 as it addressed issues believed to be responsible for this behaviour in ERA-Interim. Further effort is needed to understand these findings.

9.3 | Comparison of long-term and monthly average precipitation rates

In addition, long-term and monthly average precipitation rates from ERA-Interim and ERA5 have been evaluated by comparing them with values from other datasets. Figure 24 presents an example. Version 7 of NASA's TRMM Multi-satellite Precipitation Analysis (TMPA)

3B43 dataset (Huffman *et al.*, 2010) is used in this evaluation. The dataset covers the period from 1998 onwards and the region from 50°S to 50°N . It utilizes data from TRMM until April 2015 and from several other satellite instruments measuring in the microwave or the infrared. Analyses of direct precipitation measurements by gauges over land from GPCC (Becker *et al.*, 2013) are also used in producing the TRMM/3B43 dataset. The dataset has 0.25° spatial resolution, and the ERA data are interpolated to this grid to make the comparisons. The dataset is not entirely independent of ERA, as ERA-Interim, ERA5 and TRMM/3B43 all make use of precipitation information from microwave imagery, albeit in different ways, and the estimates of rainfall rate over the USA assimilated in ERA5 from mid-2009 onwards make use of US gauge data that are also used in forming the GPCC datasets.

Figure 24 shows maps of the mean differences (reanalysis minus TRMM/3B43) for ERA-Interim and ERA5 for

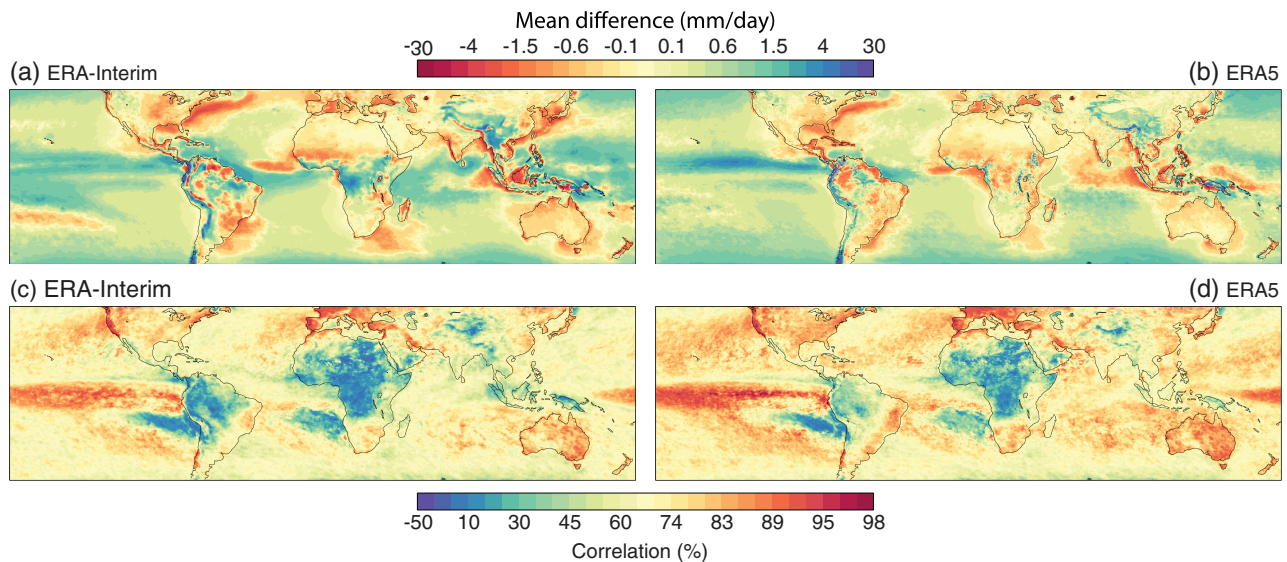


FIGURE 24 Mean difference between (a) ERA-Interim and TRMM/3B43, and between (b) ERA5 and TRMM/3B43 precipitation rates ($\text{mm}\cdot\text{day}^{-1}$) for 1998–2018. (c, d) show the corresponding correlations (%) of the sequence of monthly values from the two pairs of datasets. The mean annual cycle is removed from each dataset prior to calculating the correlations

the period 1998–2018, and the corresponding correlations of monthly anomalies relative to the 1998–2018 means for each month. Differences from TRMM/3B43 are in most respects smaller for ERA5 than ERA-Interim, including along mountain ranges and coastlines, offshore of eastern North America and Asia, and over the Congo Basin. Differences are larger over the ITCZ in the eastern Pacific, and over the extratropical Pacific, South Atlantic and Indian Oceans. Overall, the mean absolute difference over the domain from 50°S to 50°N is $0.58\text{ mm}\cdot\text{day}^{-1}$ for ERA5 and $0.65\text{ mm}\cdot\text{day}^{-1}$ for ERA-Interim.

ERA5 is also closer than ERA-Interim is to TRMM/3B43 in its representation of the temporal variability of monthly precipitation. The correlation maps in Figure 24 show this to be widespread, including regions such as the eastern Pacific ITCZ where mean differences are larger, as well as regions where mean differences are smaller. Improvement occurs both over the extratropical regions and tropical oceanic zones where correlations are already quite high in ERA-Interim, and over the tropical land masses where ERA-Interim has particularly low correlations. Correlations computed over the whole domain are 70% for ERA5 and 63% for ERA-Interim.

Comparisons (shown in Figure S5) have also been made with the GPCC and GPCP datasets and with JRA-55, for 12-month continental averages over the full period from 1979. Interpretation of mean differences is complicated by differences between GPCC and GPCP due in part to the adjustments for gauge undercatch made in GPCP but not in GPCC, as already found for ERA-Interim (Simmons *et al.*, 2010). Long-term variations from ERA5 are generally closer to those from GPCC and GPCP than are

those from ERA-Interim and JRA-55. Nevertheless ERA5 exhibits shifts over time compared with GPCC and GPCP. In particular, differences between ERA5 and both GPCC and GPCP exhibit a distinct decline for a few years centred on the year 2000 over several regions, especially over the Congo Basin and southeastern China. Differences are steadier thereafter. This can be seen in the net contributions to global precipitation from all land areas that are shown for ERA5, GPCC and GPCP in Figure 23. Overall, the 1979–2018 whole-globe correlations for the 2.5° resolution GPCP dataset are 77% for ERA5 and 67% for ERA-Interim. The correlations for this period with the 1° resolution GPCC dataset, which covers all land except Antarctica, are 63% for ERA5 and 50% for ERA-Interim.

9.4 | Intercomparison of ERA5 upper-air fields

This section examines the temporal evolution of upper-air temperature and ozone fields from ERA5 and includes some comparison with three other reanalyses (ERA-Interim, JRA-55, and MERRA-2) for selected pressure levels. Note that a comparison with observation-based datasets, such as radiosondes or satellite retrievals, is not included here but can be found in the recent literature, in particular Davis *et al.* (2017) and Long *et al.* (2017)³.

³Note that these two papers do not use data from ERA5. However, as of writing, an updated version of Davis *et al.* (2017) that evaluates ERA5 is in preparation (M.I Hegglin and S.M Davis, 2019, personal communication)

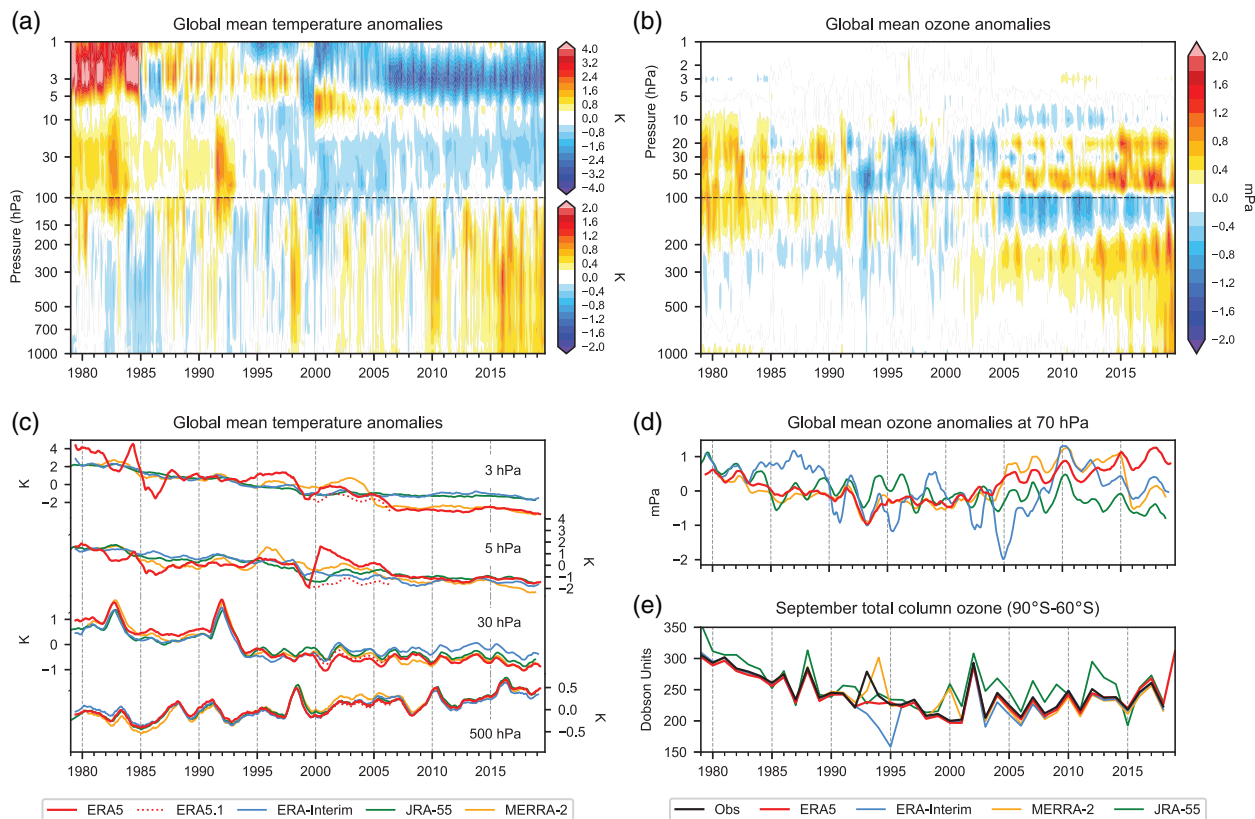


FIGURE 25 Height-time evolution of monthly and globally averaged anomalies in (a) temperature and (b) ozone partial pressure from ERA5. The anomalies are calculated by removing the 1981–2010 monthly climatology of ERA5. Time series of 12-month running mean (c) global-average temperature anomalies at 3, 5, 30, and 500 hPa; (d) global-average ozone partial pressure anomalies at 70 hPa; and (e) September mean total column ozone (TCO3) averaged over Antarctica (90°S–60°S). The time series are based on data from ERA5 (red), ERA-Interim (blue), JRA-55 (green), and MERRA-2 (orange). In (c), the time series of ERA5.1 (Section 3 gives details) are also shown for the 3 and 5 hPa levels. (e) includes TCO3 time series from a C3S merged satellite product (Copernicus, 2019) labeled ‘Obs’ (black) in the legend. All anomalies are calculated with respect to the 1981–2010 monthly climatology of each dataset. In (a) and (b), the vertical space is partitioned equally between pressure levels below and above 100 hPa. In (c), note the different intervals used in the vertical for each pressure level

Furthermore, time series of several upper-air fields from ERA5 were included for the first time in the 2019 annual “State of the Climate” report (Blunden and Arndt, 2019).

Figure 25a displays the time–height evolution of globally averaged upper-air temperature anomalies from ERA5 while Figure 25c compares time series of these anomalies with those from ERA-Interim, JRA-55, and MERRA-2 at four pressure levels. These levels are representative of the mid-troposphere (500 hPa), lower stratosphere (30 hPa), and upper stratosphere (5 and 3 hPa). Both figures highlight transient increases in the global mean temperature related to well-known climatic events such as El Niño events (1983, 1987, 1998, 2010, 2016) in the troposphere, and the volcanic eruptions of El Chichón (1982) and Mount Pinatubo (1991) in the lower stratosphere. ERA5 also captures the observed cooling of the latter and its levelling-off since the late 1990s (Randel *et al.*, 2016; Maycock *et al.*, 2018).

In Figure 25c, the temperature anomalies at 500 hPa reveal close overall agreement between the four

reanalyses, although MERRA-2 anomalies remain slightly lower (higher) during the 1980s (1990s). At 30 hPa, the four reanalyses exhibit very similar interannual variability but reveal some differences in the overall trend, with ERA5 (ERA-Interim) suggesting larger (smaller) cooling. The behaviour of ERA-Interim may be related to a cold bias in the lower stratosphere which persists through the early 2000s but is then largely corrected through the assimilation of GNSS-RO bending angles (Simmons *et al.*, 2014). This explanation would thus lend more credence to the trends of the other three reanalyses.

Spurious variations and inhomogeneities in ERA5 temperatures are more clearly apparent in the upper stratosphere (above 10 hPa) and are generally concurrent with changes seen in the analysis increments (Figure 16 and Section 7.4). The large interannual variability of ERA5 temperatures at 5 and 3 hPa (Figure 25c) greatly exceeds that seen in the other three reanalyses. The degraded skill of ERA5 relative to ERA-Interim may come as a surprise

given the improved treatment of SSU observations implemented in ERA5 (Section 4.5). The largest anomalies in ERA5 stratospheric temperatures occur in the early 1980s above 5 hPa (Figure 25a) and clearly exceed the positive anomalies seen in the other reanalyses during the same period (Long *et al.*, 2017). In this instance, the peculiar behaviour of ERA5 can be ascribed to a combination of three factors: (a) the use of smaller correlation length-scales in the model background-error estimates in ERA5 compared to ERA-Interim, (b) the known warm bias of the IFS model in the upper stratosphere (Shepherd *et al.*, 2018), and (c) the more sporadic availability of SSU observations in the early 1980s. The subsequent sharp drop in ERA5 temperature in early 1985 coincides with the transition between NOAA-7 SSU and NOAA-9 SSU already cited as a problem in ERA-Interim by Long *et al.* (2017), albeit at higher levels.

Two other inhomogeneities in upper-stratospheric temperatures occur in 1998 and 2000 and can be linked, respectively, to the beginning of the anchoring of AMSU-A channel 14 observations (Section 7.4) and to the change in B_{cli} (Section 2.4). The impact of the latter is most visible in the vicinity of 5 hPa. Figure 25c shows that the marked discontinuity in 2000 at this pressure level all but disappears in ERA5.1 (dashed red curve in Figure 25c), underscoring the significantly improved temporal consistency provided by this stream. At 3 hPa, JRA-55 and ERA-Interim show remarkable agreement throughout the period while the marked drop in MERRA-2 temperature in 2004 is an artefact known to be related to the assimilation of Aura MSL temperature profiles (Gelaro *et al.*, 2017; Long *et al.*, 2017). The markedly reduced temperature variations from 2006 onwards in all four reanalyses denote the greater constraint on stratospheric temperatures provided by the assimilation of GNSS-RO observations.

Figure 25b shows the time–height evolution of ERA5 global mean anomalies in ozone partial pressure while Figure 25d compares the time series of these anomalies with those from the other three reanalyses at 70 hPa. Above 100 hPa, ERA5 captures the observed significant decline in stratospheric ozone during the 1980s and early 1990s, followed by a gradual recovery in subsequent years (Steinbrecht *et al.*, 2018). This recovery is exaggerated in ERA5 by a spurious increase in ozone in 2004 at 40–90 hPa (concurrent with a decrease at 90–150 hPa), which can be traced to the beginning of the assimilation of ozone profiles from Aura MLS. These observations were shown to cause inhomogeneities in the ozone field of other reanalyses as well (Davis *et al.*, 2017), most notably in MERRA-2 (Figure 25d). This figure also highlights large interannual variability in ERA-Interim ozone time series, which Davis *et al.* (2017) found to exceed that seen in the observations. The distinct behaviour of JRA-55 relative

to ERA5 and MERRA-2 in Figure 25d can be explained by the fact that this reanalysis did not directly assimilate ozone observations but rather used ozone fields generated offline by a chemistry climate model (Kobayashi *et al.*, 2015; Davis *et al.*, 2017). The growing discrepancy between ERA5 and the other reanalyses from around 2015 onwards, which is associated with larger positive anomalies within 100–10 hPa in Figure 25b, is not fully understood at this stage.

Stratospheric ozone depletion during austral spring over Antarctica has been an important driver of atmospheric changes in high southern latitudes in recent decades. Therefore, we conclude this section by comparing, in Figure 25e, the time series of September mean total column ozone (TCO3) averaged over 90°S–60°S from the four reanalyses previously used. Additionally, we include output from a Level-4 merged satellite ECV product (Copernicus, 2019) spanning the 1979–2018 period. One must keep in mind that this dataset is not independent from the reanalyses since to a large extent both rely on the same satellite observations. There is excellent agreement overall between ERA5 and the ECV product, with the exception of September 1993 and (to a lesser extent) 1994. These two months fall within a period of reduced TCO3 observation coverage, between the end of the NIMBUS 7 TOMS record in mid-1993 and the beginning of the ERS-2 GOME record in late 1995. The lesser observational constraint during these two months has likely had some effect on the uncertainty of TCO3 estimates from the two products, making it difficult to tell which one should be regarded as more reliable. However, it is noteworthy that ERA5 estimates for September 1993 are very much in line with the previous and following years, which is not the case of the ECV product. Furthermore, the larger discrepancies between MERRA-2 and ERA-Interim (on the one hand) and the ECV product and ERA5 (on the other hand) in 1993–1995 can similarly be explained by the reduced observation availability (Davis *et al.*, 2017). For example, these authors found that MERRA-2 did not produce an Antarctic ozone hole in 1994. Agreement is lower overall between JRA-55 and the four other datasets for the same reason already mentioned in the previous paragraph (no direct ozone DA).

10 | CONCLUDING REMARKS AND FUTURE DIRECTIONS

10.1 | ERA5 strengths

As stated above, a major strength of ERA5 is the much higher temporal and spatial resolutions than

those of previous global reanalyses. The hourly output, 31 km horizontal resolution and 137 levels spanning the surface of the Earth to 0.01 hPa capture much finer details of atmospheric phenomena than in previous, lower-resolution, global reanalyses. The assimilation of a much larger number of reprocessed datasets has also improved the reanalysis products. As shown in this paper, the improvement with respect to ERA-Interim is considerable in the troposphere.

The provision of the accompanying ensemble, which was not available for ERA-Interim, provides an uncertainty estimate for the reanalysis products, which is much sought after by users of the data.

Another advantage of ERA5 compared to ERA-Interim is a much shorter latency of 5 days rather than 2–3 months. Based on experience of the production of ERA5 so far, it is expected that this preliminary product will only rarely deviate from the fully quality-checked final product that is released 2 months later. This timely product serves important classes of user needing up-to-date climate information in combination with a long consistent climate record.

In ERA-Interim, several different datasets for prescribed SST and SIC were required to cover the 40-plus years of the reanalysis, which made these two fields quite inhomogeneous. ERA5 has only one such dataset change, in 2007, so SST and sea ice are more homogeneous in ERA5 than in ERA-Interim.

In ERA5, CMIP5 specifications provide more realistic input to the model radiative forcing than in ERA-Interim. As a result, ERA5 has an improved response to major volcanic eruptions.

Following the initial release of ERA5, several independent studies have evaluated its performance. ERA5 performs well in the Arctic (Graham *et al.*, 2019) and Antarctic (Tetzner *et al.*, 2019) in representing winds, temperature and humidity. Mayer *et al.* (2019) found excellent closure of the Arctic energy budget using ERA5 atmospheric data. The representation of irradiance fields has been compared with other reanalyses (Trolliet *et al.*, 2018) and with ground-based and satellite observations (Urraca *et al.*, 2018). The representation of precipitation over the continental USA, in support of hydrological applications, has been evaluated relative to other modern reanalyses (Xu *et al.*, 2019) and observations (Tarek *et al.*, 2019). Finally, the characteristics of ERA5 surface and low-level winds over the ocean, relative to observations and other reanalyses, has been the subject of several studies (Belmonte Rivas and Stoffelen, 2019; Olauson, 2018; Kalverla *et al.*, 2019). Generally, ERA5 performs well in these comparisons.

10.2 | Known issues

An up-to-date list of known issues with more background information is provided in the ERA5 online data documentation (<https://confluence.ecmwf.int/display/CKB>; accessed 13 April 2020).

Although, compared to ERA-Interim, ERA5 benefits from a decade of research and development at ECMWF, some aspects did not improve. The main example is the larger cold bias in the lower stratosphere and a larger warm bias near the stratopause. As a result, in the stratosphere the fit to radiosonde data is worse, and above 10 hPa the temporal consistency of the ERA5 product is compromised due to the time-evolving competition between model bias and sparse observations. In addition, in the mesosphere, where there are no observations to control it, Cy41r2 can suffer from an overly strong tropical westerly jet, which particularly affects the transition seasons. The resulting volatility of this jet is one of the reasons why there can be large discontinuities in the mesosphere at the transition points between different production streams. Another example of discontinuities at these transition points is that of tropical stratospheric humidity, which is a slowly evolving quantity.

Although within the 12-hr assimilation windows the model constraint ensures a smooth hourly product, analysis increments can introduce systematic jumps at the transition points between the windows. For ERA5 this has been observed for wind in the boundary layer. For example over Paris, France at 1000 UTC, ERA5 exhibits on average a jump of about 0.25 m s^{-1} in 10 m wind speed. This decrease in wind speed is small, but systematic, so can be seen in climatologies.

At specific locations, ERA5 occasionally produces amounts of precipitation that are unrealistically high. These “rain bombs” became apparent in the IFS since cycle 40r1 (which became operational in November 2013), at which time changes to its convection scheme were introduced to capture better the diurnal cycle of deep convection. Under very special conditions (moist air, low-level orographic forcing with converging low-intensity winds) there is potential for “explicit-convection” at isolated grid-points (Malardel and Ricard, 2015). These features occur infrequently, of the order of ten episodes for an entire year and mostly concentrated in Africa, at isolated grid-points mostly in orographic areas. This feature was later resolved in the IFS with the introduction of the octahedral reduced Gaussian grid.

In mountainous regions above about 1,500 m, the snow depth is unrealistically large. This is due to the IFS's representation of the snow pack with a single layer of snow which does not produce enough melting.

In the northern winter of 1996/1997, the ozone values in the upper stratosphere at high northern latitudes are many times larger than normal. This problem does not significantly affect total column ozone.

Due to the very long spin-up time and non-optimal initialization, some soil parameters, such as root-zone soil moisture, show discontinuities at the transition points of some of the production streams.

Besides these limitations, there were several production oversights that could have been prevented. From 1979 to 2013, the SST was not used over the Great Lakes to nudge the lake model, as it should have been. Consequently, the 2 m temperature over these lakes has an annual cycle that is too strong, with temperatures being too cold in winter and too warm in summer. This problem is particularly pronounced over Lake Superior. Over other confined areas, like over the northern Caspian Sea and the Gulf of Finland, systematic differences between the SST products used before (HadISST2) and after 2007 (OSTIA) are known to exist.

The prescribed sea ice field tends to exhibit sea ice in the Baltic Sea in summertime, which is not there in reality. Such occurrences of sea ice have been removed in the HRES ERA5 reanalysis but it is still partly present in the ten-member ensemble. However, its impact on the ensemble spread is found to be tempered.

Up to once or twice per year, the analysed near-surface (e.g., 10 m) winds in ERA5 suffer from a problem of extremely large wind speeds; the largest speeds seen so far are of order 300 m s^{-1} . Typically, these occur in the last two or three hours of the assimilation window and only at one of several preferred locations around the globe, most of which are near to orographic features. A more thorough quality check could have prevented such cases by using the solution that is followed when instabilities occasionally occur in the 4D-Var tangent-linear physics. For the ERA5 timely updates, that check was implemented from 19 February 2020 onwards to prevent further occurrences, while a practical solution for existing cases is provided in the online documentation.

10.3 | Related products

In conjunction with ERA5, a down-scaled land product has been made available. This ERA5-Land product, at 9 km horizontal resolution (Muñoz-Sabater, 2019), was produced through a single simulation driven by near-surface atmospheric fields from ERA5, with thermodynamical orographic adjustment of temperature. In addition to improving the quality of near-surface quantities, it provides a more homogeneous dataset for soil parameters between ERA5 production streams.

In 2020 the period from 1950 to 1978 will be made available. This will extend the hourly ERA5 record to 70 years. Details on this extension and its characteristics will be reported elsewhere. The improved segment from 2000 to 2006 (ERA5.1) was recently made available as a separate product. It provides better global-mean temperatures in the stratosphere and uppermost troposphere, with very similar performance in the lower and middle troposphere.

Also in 2020 access will be provided to 12-hourly averaged statistics of the ERA5 observation usage. This will include information on first-guess and analysis departures, bias estimates, and applied observation errors and quality control. This will, for example, allow the reproducibility of ERA5 curves in most of the observation-based Figures of this paper.

10.4 | Future directions

The interaction between model bias and an evolving observing system (which comprises components that generally exhibit non-zero, and often significant, biases) is a concern in climate reanalysis, since it can affect the temporal consistency and accuracy of the product. For future reanalysis activities at ECMWF, research is needed to tailor the latest available formulation of weak-constraint 4D-Var for reanalysis as well as the optimization of the use of anchor data in VarBC. Regarding weak-constraint 4D-Var, at ECMWF recently significant progress has been made in handling the model bias in the stratosphere (Laloyaux *et al.*, 2020). Potentially, large-scale model bias as estimated from the recent well-observed era (i.e., including the availability of anchoring GNSS-RO data) can be used as a forcing term to temper the adverse effect of model drift. This would make the mean state of reanalysis products more resilient with respect to changes in the observing system.

The non-closure of energy budgets, and particularly its evolution over time, need to be better understood and improved in future reanalysis, as well as the lack of conservation of the global dry mass and hydrological balance.

The hybrid incremental 4D-Var formulation in ERA5 allows for flow-dependent estimates of the background-error covariance matrix. For future C3S reanalysis, a more dynamic system is to be put in place that allows for an improved response to the major changes in the observing system. For the ocean (and also land), a similar system would be required, such as to account for the enormous change since the advent of Argo floats in the early 2000s. In addition, prescribed observation errors should evolve over time to reflect improvements in instrumentation, rather than to keep those constant over time in ERA5.

TABLE 8 Glossary

Acronym	Description
ADEOS-1	Advanced Earth Observing Satellite 1
AIRS	Advanced Infrared Sounder
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
AMSR-2	Advanced Microwave Scanning Radiometer 2
AMSU-A	Advanced Microwave Sounding Unit-A
AMSU-B	Advanced Microwave Sounding Unit-B
AMV	Atmospheric Motion Vector
ASCAT	Advanced Scatterometer
ATOVS	Advanced TIROS Operational Vertical Sounder
AVHRR	Advanced Very High Resolution Radiometer
BUFR	Binary Universal Form for the Representation of meteorological data
CAMS	Copernicus Atmosphere Monitoring Service
CCI	(ESA) Climate Change Initiative
CDS	C3S Climate Data Store
(C)ERA-20C	(Coupled) ECMWF Reanalysis of the 20th Century
CERA-SAT	Coupled ECMWF Reanalysis for the modern satellite era
CFSR	Climate Forecast System Reanalysis
CHAMP	Challenging Mini-Satellite Payload
CIMSS	Cooperative Institute for Meteorological Satellite Studies
CKB	Copernicus Knowledge Base
CMIP5	Coupled Model Intercomparison Project, Phase 5
CMOD	C-band model
CM SAF	Climate Monitoring Satellite Applications Facility
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate
CRIS	Cross-Track Infrared Sounder
C3S	Copernicus Climate Change Service
DMSP	Defense Meteorological Satellite Program
DRIBU	Report from Drifting and moored Buoy
EC	European Commission
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
EDA	Ensemble of Data Assimilations
ENVISAT	Environmental Satellite
EOS	Earth Observing System
ERA	ECMWF Reanalysis

TABLE 8 Glossary

Acronym	Description
ERA5	70-year ERA starting from January 1950 onwards with timely updates
ERA-15	15-year ERA starting from January 1979 to February 1994
ERA-40	45-year ERA from September 1957 to August 2002
ERA-20CM	20th Century ECMWF Model integration
ERA-Interim	40-year ERA from January 1979 to August 2019
ERA-CLIM	European Reanalysis of Global Climate Observations
ERA-CLIM2	European Reanalysis of Global Climate Observations 2
ERS	European Remote Sensing Satellite
ESA	European Space Agency
ETOPO2	2-Minute Gridded Global Relief Data
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
FGGE	First GARP Global Experiment
GARP	Global Atmospheric Research Program
GCM	Global Circulation Model
GCOS	Global Climate Observing System
GEMS	Global and regional Earth system Monitoring using Satellite and <i>in situ</i> data
GMAO	Global Modeling and Assimilation Office
GMI	Global precipitation monitoring mission Microwave Imager
GMS	Geostationary Meteorological Satellite
GNSS-RO	Global Navigation Satellite System – Radio Occultation
GOES	Geostationary Operational Environmental Satellite
GOME	Global Ozone Monitoring Experiment
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GRAS	Global Navigation Satellite Systems Receiver for Atmospheric Sounding
GTS	Global Telecommunication System
HadISST2	Hadley Centre Sea Ice and Sea Surface Temperature dataset
HRES	High Resolution component (of ERA5)
HIRS	High-Resolution Infrared Sounder
HTESSEL	Revised Tiled ECMWF Scheme for Surface Exchanges over Land

TABLE 8 Glossary

Acronym	Description
IASI	Infrared Atmospheric Sounding Interferometer
ICOADS	International Comprehensive Ocean and Atmosphere Data Set
IFS	Integrated Forecast System
IMS	Interactive Multisensor Snow and Ice Mapping System
IPCC	Intergovernmental Panel on Climate Change
IR	Infrared Radiation
IRAS	Infrared Atmospheric Sounder
ISPD	International Surface Pressure Databank
ITCZ	InterTropical Convergence Zone
JAXA	Japan Aerospace Exploration Agency
JMA	Japan Meteorological Agency
JRA-55	Japanese 55-year Reanalysis Project
LDAS	Land Data Assimilation System
LEO	Low Earth Orbit
MACC	Monitoring Atmospheric Composition and Climate
MARS	ECMWF Meteorological Archival and Retrieval System
McICA	Monte Carlo Independent Column Approximation
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2-2
METAR	Meteorological Aerodrome Reports
METEOSAT	Meteorological Satellite
METOP	Meteorological Operational Satellite
MHS	Microwave Humidity Sounder
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLS	Microwave Limb Sounder
MODIS	Moderate Resolution Imaging Spectroradiometer
MSU	Microwave Sounding Unit
MTSAT	Multifunctional Transport Satellite
MWHS	Microwave Humidity Sounder
MWHS2	Microwave Humidity Sounder 2
NASA	National Aeronautics and Space Administration (USA)
NASDA	National Space Development Agency (Japan)
NCEP	National Centers for Environmental Prediction (USA)
NESDIS	National Environmental Satellite, Data, and Information Service (USA)
NEXRAD	Next-Generation Radar Network

TABLE 8 Glossary

Acronym	Description
NOAA	National Oceanic and Atmospheric Administration (USA)
OI	Optimal interpolation
OLR	Outgoing Long-wave Radiation
OMI	Ozone Monitoring Instrument
ORASN	Nth ECMWF Ocean Reanalysis System
OSI SAF	Ocean and Sea Ice SAF
OSTIA	Operational Sea Surface Temperature and Sea Ice Analysis
PILOT	Wind report from pilot balloon
QuikSCAT	Quick Scatterometer
RAOBCORE	Radiosonde Observation Correction using Reanalyses
RICH	Radiosonde Innovation Composite Homogenization
RISE	Radiosonde adjustments with solar elevation dependence
RRTMG	Rapid Radiative Transfer Model for GCMs
RTTOV	Radiative Transfer for TOVS
SARAL	Satellite with ARGOS and AltiKa
SBUV	Solar Backscattered UltraViolet
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Cartography
SEKF	Simplified Extended Kalman Filter
SETTLS	Stable Extrapolation Two-Time-Level Scheme
SIC	Sea Ice Concentration
SSM/I	Special Sensor Microwave/Imager
SSM/I/S	Special Sensor Microwave Imager/Sounder
SST	Sea Surface Temperature
SSU	Stratospheric Sounding Unit
SSW	Sudden Stratospheric Warming
SYNOP	Surface Synoptic Report
TEI	Total Energy Input
TEMP	Report from radiosounding
TESSEL	Tiled ECMWF Scheme for Surface Exchanges over Land
TIROS	Television Infrared Observation Satellite
TMI	TRMM Microwave Imager
TOA	Top Of Atmosphere
TOMS	Total Ozone Mapping Spectrometer
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission

TABLE 8 Glossary

Acronym	Description
TSI	Total Solar Irradiance
UCAR	University Corporation for Atmospheric Research (USA)
UNFCCC	United Nations Framework Convention on Climate Change
VarBC	Variational bias correction
VTPR	Vertical Temperature Profiling Radiometer
WCRP	World Climate Research Programme
WMO	World Meteorological Organisation
20CR	Twentieth Century Reanalysis
4D-Var	Four-Dimensional Variational data assimilation

ERA5 uncertainty estimates are based on the spread in the EDA component which mainly samples random error (although the perturbed HadISST2 realizations do introduce long-term correlations near the surface). More information is desirable on systematic errors which, for example, capture and explain large-scale and long-term systematic differences with respect to other datasets.

Following user requirement analyses from communities, a large part of the preparation for future C3S reanalysis are being addressed at ECMWF, while part of the work takes place via external C3S contracts with other European organisations and includes wider international collaborations. This includes a considerable effort by EUMETSAT on the reprocessing of a large number of satellite datasets for usage in future reanalysis. Support for climate reanalysis regarding satellite data rescue has been initiated, as is the development and maintenance of quality-controlled global databases containing all known digitised *in situ* upper-air weather observations. These will contain meta-data and information needed for DA such as bias adjustments and uncertainty estimates. In addition, a set of services to improve access to available *in situ* instrumental data records and data streams from observing networks is in place, as needed for monitoring climate change and to support climate science.

All these developments and data will feed into the next generation of global reanalysis (ERA6), which is to be based on a coupled atmosphere–ocean system. It is foreseen that, until completion of this new full-observing-system reanalysis, ERA5 will be updated into the mid-2020s.

ACKNOWLEDGMENTS

Reanalysis touches on a large number of activities at ECMWF and its success relies on the efforts from many,

many people across ECMWF and from many collaborations. ECMWF implements the Copernicus Climate Change Service on behalf of the European Union, and ERA5 was produced with funding from this Service. The EU, through the 7th Framework Programme, supported the ERA-CLIM and ERA-CLIM2 projects which served as precursors to the ERA5 reanalysis. The efforts of the many scientists involved in these projects is gratefully acknowledged. Preparation of ERA5 has been supported by ECMWF staff, three of which were supported by the ERA-CLIM2 project, funded by the European Union's Seventh Framework Program under grant agreement 607029, two of which were funded by EUMETSAT, and by staff secondments from NCAS, a NERC Collaborative Centre, and the JMA. ERA5 benefits from the usage of a large number of reprocessed datasets that were prepared by CIMMS, ESA, EUMETSAT Satellite Application Facilities (SAF), JMA, NASA, NASDA, NOAA, TU Wien and UCAR. The Copernicus (2019) ozone data were downloaded from the C3S CDS. In addition, we thank three reviewers for their valuable feedback.

ORCID

Hans Hersbach  <https://orcid.org/0000-0001-5330-7071>

Joaquín Muñoz-Sabater  <https://orcid.org/0000-0002-5997-290X>

Julien Nicolas  <https://orcid.org/0000-0003-0518-100X>

Dinand Schepers  <https://orcid.org/0000-0002-2611-487X>

Adrian Simmons  <https://orcid.org/0000-0002-7327-6310>

Saleh Abdalla  <https://orcid.org/0000-0003-4469-7949>

Xavier Abellan  <https://orcid.org/0000-0002-1999-1823>

Gianpaolo Balsamo  <https://orcid.org/0000-0002-1745-3634>

Peter Bechtold  <https://orcid.org/0000-0002-1967-3382>

Gionata Biavati  <https://orcid.org/0000-0002-1675-6967>

Jean Bidlot  <https://orcid.org/0000-0001-7423-5118>

Giovanna De Chiara  <https://orcid.org/0000-0002-4540-0687>

Dick Dee  <https://orcid.org/0000-0002-8321-9125>

Michail Diamantakis  <https://orcid.org/0000-0003-2279-9717>

Johannes Flemming  <https://orcid.org/0000-0003-4880-5329>

Richard Forbes  <https://orcid.org/0000-0002-3596-8287>

Manuel Fuentes  <https://orcid.org/0000-0003-1544-9612>

Leo Haimberger  <https://orcid.org/0000-0002-0379-6353>

Sean Healy  <https://orcid.org/0000-0003-4810-9593>

Robin J. Hogan  <https://orcid.org/0000-0002-3180-5157>

Marta Janisková  <https://orcid.org/0000-0003-2644-8068>

Sarah Keeley  <https://orcid.org/0000-0002-8046-765X>
 Patrick Laloyaux  <https://orcid.org/0000-0003-2808-0463>
 Cristina Lupu  <https://orcid.org/0000-0003-3530-8104>
 Patricia de Rosnay  <https://orcid.org/0000-0002-7374-3820>
 Freja Vamborg  <https://orcid.org/0000-0003-3092-0775>
 Sebastien Villaume  <https://orcid.org/0000-0002-3041-076X>
 Jean-Noël Thépaut  <https://orcid.org/0000-0003-3214-5266>

REFERENCES

- Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P. and Nelkin, E. (2003) The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). *Journal of Hydrometeorology*, 4(6), 1147–1167.
- Ahlgrimm, M. and Forbes, R. (2014) Improving the representation of low clouds and drizzle in the ECMWF model based on ARM observations from the Azores. *Monthly Weather Review*, 142(2), 668–685.
- Albergel, C., Balsamo, G., de Rosnay, P., Muñoz Sabater, J. and Boussetta, S. (2012) A bare ground evaporation revision in the ECMWF land-surface scheme: evaluation of its impact using ground soil moisture and satellite microwave data. *Hydrology and Earth System Sciences*, 16, 3607–3620. <https://doi.org/10.5194/hess-16-3607-2012>
- Allan, R.P., Liu, C., Loeb, N.G., Palmer, M.D., Roberts, M., Smith, D. and Vidale, P.-L. (2014) Changes in global net radiative imbalance 1985–2012. *Geophysical Research Letters*, 41(15), 5588–5597. <https://doi.org/10.1002/2014GL060962>
- Anderson, E. and Järvinen, H. (1999) Variational quality control. *Quarterly Journal of the Royal Meteorological Society*, 125, 697–722.
- Auligné, T., McNally, A.P. and Dee, D.P. (2007) Adaptive bias correction for satellite data in a numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 133, 631–642.
- Balmaseda, M.A., Mogensen, K. and Weaver, A.T. (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, 139, 1132–1161.
- Balmaseda, M.A., Vidard, A. and Anderson, D. (2008) The ECMWF Ocean Analysis System: ORA-S3. *Monthly Weather Review*, 136, 3018–3034.
- Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D.P., Dutra, E., Muñoz Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T. and Vitart, F. (2015) ERA-Interim/Land: a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, 19(1), 389–407.
- Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T. and Potes, M. (2012) On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. *Tellus A*, 64(1). <https://doi.org/10.3402/tellusa.v64i0.15829>
- Balsamo, G., Viterbo, P., Beljaars, A., van den Hurk, B., Hirschi, M., Betts, A. and Scipal, K. (2009) A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *Journal of Hydrometeorology*, 10, 623–643. <https://doi.org/10.1175/2008JHM1068.1>
- Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M.J., Vitart, F. and Balsamo, G. (2008) Advances in simulating atmospheric variability with the ECMWF model: from synoptic to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, 134, 1337–1351.
- Bechtold, P., Semane, N., Lopez, P., Chaboureaud, J.-P., Beljaars, A. and Bormann, N. (2014) Representing equilibrium and nonequilibrium convection in large-scale models. *Journal of the Atmospheric Sciences*, 71(2), 734–753.
- Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U. and Ziese, M. (2013) A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, 5(1), 71–99.
- Belmonte Rivas, M. and Stoffelen, A. (2019) Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT. *Ocean Science*, 15(3), 831–852. <https://doi.org/10.5194/os-15-831-2019>
- Bengtsson, L., Kanamitsu, M., Källberg, P.W. and Uppala, S.M. (1982) FGGE 4-dimensional data assimilation at ECMWF. *Bulletin of the American Meteorological Society*, 63, 29–43.
- Berrisford, P., Källberg, P., Kobayashi, S., Dee, D.P., Uppala, S.M., Simmons, A.J., Poli, P. and Sato, H. (2011) Atmospheric conservation properties in ERA-Interim. *Quarterly Journal of the Royal Meteorological Society*, 137, 1381–1399.
- Bidlot, J.-R. (2012). Present status of wave forecasting at ECMWF, pp. 25–27 in Workshop on Ocean Waves, ECMWF, Reading, UK.
- Blunden, J. and Arndt, D.S. (2019) State of the Climate in 2018. *Bulletin of the American Meteorological Society*, 100(9), S1–S305. <https://doi.org/10.1175/2019BAMSStateoftheClimate.1>
- Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A. and Zemp, M. (2014) The concept of essential climate variables in support of climate research, applications, and policy. *Bulletin of the American Meteorological Society*, 95(9), 1431–1443.
- Bonavita, M., Hólm, E.V., Isaksen, I. and Fisher, M. (2016) The evolution of the ECMWF hybrid data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 142, 287–303.
- Bormann, N., Fouilloux, A. and Bell, W. (2012). Evaluation and assimilation of ATMS data in the ECMWF system. Technical Memorandum 689, ECMWF, Reading, UK.
- Bormann, N., Bonavita, M., Dragani, R., Eresmaa, R., Matricardi, M. and McNally, A.P. (2015). Enhancing the impact of IASI observations through an updated observation error covariance matrix. Technical Memorandum 756, ECMWF, Reading, UK.
- Bormann, N., Collard, A.D. and Bauer, P. (2009). Estimates of spatial and interchannel observation error characteristics for current sounder radiances for NWP. Technical Memorandum 600, ECMWF, Reading.
- Bormann, N., Lupu, C., Geer, A.J., Lawrence, H., Weston, P. and English, S.J. (2017). Assessment of the forecast impact of surface-sensitive microwave radiances over land and sea-ice. Technical Memorandum 804, ECMWF, Reading, UK.
- Boussetta, S., Balsamo, G., Beljaars, A., Kral, T. and Jarlan, L. (2013) Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model. *International Journal of Remote Sensing*, 34(9–10), 3520–3542.

- Burrows, C. (2018). Assimilation of radiance observations from geostationary satellites: first year report. EUMETSAT/ECMWF Fellowship Programme Research Report 47, ECMWF, Reading, UK.
- Cardinali, C., Pezzulli, S. and Anderson, E. (2004) Influence-matrix diagnostic of a data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 130, 2767–2786.
- Cariolle, D. and Déqué, M. (1986) Southern Hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. *Journal of Geophysical Research*, 91, 10825–10846.
- Cariolle, D. and Teyssède, H. (2007) A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations. *Atmospheric Chemistry and Physics Discussion*, 7, 1655–1697.
- Chen, K., English, S.J., Bormann, N. and Zhu, J. (2014). Assessment of FY-3A and FY-3B MWS observations. ECMWF Technical Memorandum 734, ECMWF, Reading.
- Collard, A.D. and McNally, A.P. (2009) The assimilation of Infrared Atmospheric Sounding Interferometer radiances at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 135, 1044–1058. <https://doi.org/10.1002/qj.410>
- Compo, G.P., Whitaker, J.S. and 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., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R.I., Grant, A.N., Groisman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, Ø., Ross, T.F., Trigo, R.M., Wang, X.L., Woodruff, S.D. and Worley, S.J. (2011) The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, 137, 1–28.
- Copernicus (2019). Ozone monthly gridded data from 1970 to present, Level 4, version 0021. Available at <https://cds.climate.copernicus.eu/portfolio/dataset/satellite-ozone>; accessed 11 April 2020.
- Courtier, P., Thépaut, J.-N. and Hollingsworth, A. (1994) A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120, 1367–1388.
- Cowan, K. and Way, R.G. (2014) Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society*, 140, 1935–1944.
- Cram, T.A., Compo, G.P., Yin, X., Allan, R.J., McColl, C., Vose, R.S., Whitaker, J.S., Matsui, N., Ashcroft, L., Auchmann, R., Bessemoulin, P., Brandsma, T., Brohan, P., Brunet, M., Comeaux, J., Crouthamel, R., Gleason Jr, B.E., Groisman, P.Y., Hersbach, H., Jones, P.D., Jónsson, T., Jourdain, S., Kelly, G.A., Knapp, K.R., Kruger, A., Kubota, H., Lentini, G., Lorrey, A., Lott, N., Lubker, S.J., Luterbacher, J., Marshall, G.J., Maugeri, M., Mock, C.J., Mok, H.Y., Nordli, Ø., Rodwell, M.J., Ross, T.F., Schuster, D., Srncic, L., Valente, M.A., Vizi, Z., Wang, X.L., Westcott, N., Woollen, J.S. and Worley, S.J. (2015) The international surface pressure databank version 2. *Geoscience Data Journal*, 2(1), 31–46. <https://doi.org/10.1002/gdj3.25>
- Davis, S.M., Hegglin, M.I., Fujiwara, M., Dragani, R., Harada, Y., Kobayashi, C., Long, C., Manney, G.L., Nash, E.R., Potter, G.L., Tegtmeier, S., Wang, T., Wargan, K. and Wright, J.S. (2017) Assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses as part of S-RIP. *Atmospheric Chemistry and Physics*, 17(20), 12743–12778. <https://doi.org/10.5194/acp-17-12743-2017>
- de Rosnay, P., Balsamo, G., Albergel, C., Muñoz Sabater, J. and Isaksen, I. (2014) Initialisation of land surface variables for numerical weather prediction. *Surveys in Geophysics*, 35(3), 607–621. <https://doi.org/10.1007/s10712-012-9207-x>
- de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C. and Isaksen, I. (2013) A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 139, 1199–1213. <https://doi.org/10.1002/qj.2023>
- Dee, D.P. (2005) Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131, 3323–3343.
- Dee, D.P. and Uppala, S.M. (2009) Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 135, 1830–1841.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.
- Diamantakis, M. (2014) Improving ECMWF forecasts of sudden stratospheric warmings. *ECMWF Newsletter*, 141, 30–36.
- Diamantakis, M. and Magnusson, L. (2016) Sensitivity of the ECMWF model to semi-Lagrangian departure point iterations. *Monthly Weather Review*, 144(9), 3233–3250.
- Donlon, C.J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W. (2012) The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, 116, 140–158.
- Dragani, R. (2009). Variational bias correction of satellite ozone data. Technical Report R43.8/RD/0934, ECMWF, Reading, UK.
- Dragani, R. (2016) A comparative analysis of UV nadir-backscatter and infrared limb-emission ozone data assimilation. *Atmospheric Chemistry and Physics*, 16(13), 8539–8557.
- Dragani, R. and McNally, A.P. (2013) Operational assimilation of ozone-sensitive infrared radiances at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 139, 2068–2080.
- Durack, P.J. and Wijffels, S.E. (2010) Fifty-year trends in global ocean salinities and their relationship to broad-scale warming. *Journal of Climate*, 23(16), 4342–4362.
- Dutra, E., Balsamo, G., Viterbo, P., Miranda, P.M.A., Beljaars, A., Schär, C. and Elder, K. (2010) An improved snow scheme for the ECMWF land surface model: description and offline validation. *Journal of Hydrometeorology*, 11(4), 899–916. <https://doi.org/10.1175/2010JHM1249.1>
- Dutra, E., Stepanenko, V., Balsamo, G., Viterbo, P., Miranda, P., Mironov, D. and Schär, C. (2009). Impact of lakes on the ECMWF surface scheme. ECMWF Technical Memorandum 608, ECMWF, Reading, UK.

- DWD (2000). Bewertung der Orkanwetterlage am 26.12.1999 aus klimatologischer Sicht. Available at: http://www.wetter-extrem.de/stuerme/lothar/orkan_lothar.pdf; accessed 11 April 2020.
- Eastwood, S., Laverne, T. and Tonboe, R. (2014). Algorithm Theoretical Basis Document for the OSI SAF global reprocessed sea ice concentration product, version 1.1. EUMETSAT Satellite Application Facilities, Darmstadt, Germany.
- ECMWF R&D (2016). IFS documentation cy41r2. <https://www.ecmwf.int/en/publications/ifs-documentation>; accessed 11 April 2020.
- Eresmaa, R., Letertre-Danczak, J., Lupu, C., Bormann, N. and McNally, A.P. (2017) The assimilation of Cross-track Infrared Sounder radiances at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 143, 3177–3188. <https://doi.org/10.1002/qj.3171>
- Fennig, K., Schroeder, M. and Hollmann, R. (2017). Fundamental climate data record of microwave imager radiances, edition 3. Technical Report, Satellite Application Facility on Climate Monitoring, EUMETSAT, Darmstadt, Germany, DOI 10.5676/EUM_SAF_CM/FCDR_MWI/V003, (to appear in print).
- Fiorino, M. (2004). A multi-decadal daily sea surface temperature and sea ice concentration data set for the ERA-40 reanalysis. Technical Report 12, ECMWF, Reading, UK.
- Fisher, M. (2003). Background error covariance modelling. pp. 45–63 in Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean. ECMWF, Reading, UK.
- Flemming, J., Benedetti, A., Inness, A., Engelen, R.J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A., Peuch, V.-H., Akritidis, D. and Katragkou, E. (2017) The CAMS interim reanalysis of carbon monoxide, ozone and aerosol for 2003–2015. *Atmospheric Chemistry and Physics*, 17(3), 1945–1983.
- Forbes, R.M. and Tompkins, A.M. (2011) An improved representation of cloud and precipitation. *ECMWF Newsletter*, 129, 13–18.
- Forbes, R.M. and Ahlgrim, M. (2014) On the representation of high-latitude boundary-layer mixed-phase cloud in the ECMWF global model. *Monthly Weather Review*, 142(9), 3425–3445.
- Forbes, R.M., Tompkins, A.M. and Untch, A. (2011). A new prognostic bulk microphysics scheme for the IFS. ECMWF Technical Memorandum 649, ECMWF, Reading, UK.
- Gauthier, P. and Thépaut, J.-N. (2001) Impact of the digital filter as a weak constraint in the preoperational 4D-VAR assimilation system of Météo-France. *Monthly Weather Review*, 129(8), 2089–2102.
- Geer, A.J., Baordo, F., Bormann, N., English, S.J., Kazumori, M., Lawrence, H., Lean, P., Lonitz, K. and Lupu, C. (2017) The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Quarterly Journal of the Royal Meteorological Society*, 143, 3189–3206. <https://doi.org/10.1002/qj.3172>
- Geer, A.J., Bauer, P. and Bormann, N. (2010) Solar biases in microwave imager observations assimilated at ECMWF. *IEEE Transactions on Geoscience and Remote Sensing*, 48(6), 2660–2669. June
- Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M. and Zhao, B. (2017) The Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2). *Journal of Climate*, 30(14), 5419–5454.
- Gibson, J.K., Kållberg, P., Uppala, S.M., Hernandez, A., Nomura, A. and Serrano, E. (1999). ECMWF re-analysis project report 1, ERA-15 description (version 2). Technical Report, ECMWF, Reading, UK.
- Graham, R.M., Hudson, S.R. and Maturilli, M. (2019) Improved Performance of ERA5 in Arctic Gateway Relative to Four Global Atmospheric Reanalyses. *Geophysical Research Letters*, 46(11), 6138–6147. <https://doi.org/10.1029/2019GL082781>
- Haimberger, L., Tavolato, C. and Sperka, S. (2008) Toward elimination of the warm bias in historic radiosonde temperature records: some new results from a comprehensive intercomparison of upper-air data. *Journal of Climate*, 21, 4587–4606.
- Haimberger, L., Tavolato, C. and Sperka, S. (2012) Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *Journal of Climate*, 25, 8108–8131.
- Hansen, J., Ruedy, R., Sato, M. and Lo, K. (2010) Global surface temperature change. *Reviews of Geophysics*, 48(4)
- Healy, S.B., Eyre, J., Hamrud, M. and Thépaut, J.-N. (2007) Assimilating GPS radio occultation measurements with two-dimensional bending angle observation operators. *Quarterly Journal of the Royal Meteorological Society*, 133, 1213–1227. <https://doi.org/10.1002/qj.63>
- Healy, S.B. (2011) Refractivity coefficients used in the assimilation of GPS radio occultation measurements. *Journal of Geophysical Research*, 116. <https://doi.org/10.1029/2010JD014013>
- Hersbach, H. (2010) Comparison of C-band scatterometer CMOD5.N equivalent neutral winds with ECMWF. *Journal of Atmospheric and Oceanic Technology*, 27(4), 721–736.
- Hersbach, H. (2019) ECMWF's ERA5 reanalysis extends back to 1979. *ECMWF Newsletter*, 158, 1
- Hersbach, H., Peubey, C., Simmons, A.J., Berrisford, P., Poli, P. and Dee, D.P. (2015) ERA-20CM: a twentieth-century atmospheric model ensemble. *Quarterly Journal of the Royal Meteorological Society*, 141, 2350–2375.
- Hersbach, H., Brönnimann, S., Haimberger, L., Mayer, M., Villiger, L., Comeaux, J., Simmons, A.J., Dee, D.P., Jourdain, S., Peubey, C., Poli, P., Rayner, N.A., Sterin, A.M., Stickler, A., Valente, M.A. and Worley, S.J. (2017) The potential value of early (1939–1967) upper-air data in atmospheric climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 143, 1197–1210.
- Hersbach, H., de Rosnay, P., Bell, B., Schepers, D., Simmons, A.J., Soci, C., Abdalla, S., Balmaseda, M.A., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J., de Boissés, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P., Dee, D.P., Dragani, R., Diamantaki, M., Flemming, J., Forbes, R., Geer, A.J., Haiden, T., Hólm, E.V., Haimberger, L., Hogan, R., Horányi, A., Janisková, M., Laloyaux, P., Lopez, P., Muñoz Sabater, J., Peubey, C., Radu, R., Richardson, D., Thépaut, J.-N., Vitart, F., Yang, X., Zsótér, E. and Zuo, H. (2018). Operational global reanalysis: progress, future directions and synergies with NWP. ERA Report Series no. 27, ECMWF, Reading, UK.
- Hewson, T.D. and Neu, U. (2015) Cyclones, windstorms and the IMILAST project. *Tellus A*, 67(1), 27128
- Hirahara, S., Balmaseda, M.A., de Boissés, E. and Hersbach, H. (2016). Sea surface temperature and sea ice concentration for ERA5. ERA Report Series no. 26, ECMWF, Reading, UK.

- Hirons, L., Inness, P., Vitart, F. and Bechtold, P. (2013) Understanding advances in the simulation of intraseasonal variability in the ECMWF model. Part II: the application of process-based diagnostics. *Quarterly Journal of the Royal Meteorological Society*, 139, 1427–1444.
- Hogan, R.J. and Bozzo, A. (2015) Mitigating errors in surface temperature forecasts using approximate radiation updates. *Journal of Advances in Modeling Earth Systems*, 7, 836–853.
- Hogan, R.J. and Hirahara, S. (2016) Effect of solar zenith angle specification in models on mean shortwave fluxes and stratospheric temperatures. *Geophysical Research Letters*, 43, 482–488.
- Horányi, A. (2017) Some aspects on the use and impact of observations in the ERA5 Copernicus Climate Change Service reanalysis. *Idojaras*, 121(4), 329–344.
- Huffman, G.J., Adler, R.F., Bolvin, D.T. and Nelkin, E.J. (2010). The TRMM Multi-satellite Precipitation Analysis (TMPA), Chapter 1 in *Satellite Rainfall Applications for Surface Hydrology*. Gebremichael, M., Hossain, F. (eds) Springer, Berlin.
- Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A. and Collins, W.D. (2008) Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13). <https://doi.org/10.1029/2008JD009944>
- Ide, K., Courtier, P., Ghil, M. and Lorenc, A.C. (1997) Unified notation for data assimilation: operational, sequential and variational. *Journal of the Meteorological Society of Japan*, 39, 2038–2052.
- Ingleby, B., Isaksen, L., Kral, T., Haiden, T. and Dahoui, M. (2018) Improved use of atmospheric *in situ* data. *ECMWF Newsletter*, 155, 20–25.
- Ingleby, B., Pauley, P., Kats, A., Ator, J., Keyser, D., Doerenbecher, A., Fucile, E., Hasegawa, J., Toyoda, E., Kleinert, T., Qu, W., St. James, J., Tennant, W. and Weedon, R. (2016) Progress toward high-resolution, real-time radiosonde reports. *Bulletin of the American Meteorological Society*, 97(11), 2149–2161.
- Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J.J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M. and Suttie, M. (2019) The CAMS reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics*, 19(6), 3515–3556. <https://doi.org/10.5194/acp-19-3515-2019>
- Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J.W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M.G., Simmons, A.J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M. and Zerefos, C. (2013) The MACC reanalysis: an 8-year data set of atmospheric composition. *Atmospheric Chemistry and Physics*, 13, 4073–4109.
- Isaksen, L., Bonavita, M., Buizza, R., Fisher, M., Haseler, J., Leutbecher, M. and Raynaud, L. (2010). Ensemble of Data Assimilations at ECMWF. Technical Memorandum 636, ECMWF, Reading, UK.
- Janisková, M. and Lopez, P. (2013). Linearized physics for data assimilation at ECMWF, pp. 251–286 in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, Vol. II. Park, S.K., Xu, L. (eds), Springer, Berlin.
- Janssen, P.A.E.M. and Bidlot, J. (2009). On the extension of the freak wave warning system and its verification. Technical Memorandum 588, ECMWF, Reading, UK.
- Kalverla, P.C., Duncan, J.B.J., Steeneveld, G.-J. and Holtslag, A.A.M. (2019) Low-level jets over the North Sea based on ERA5 and observations: together they do better. *Wind Energy Science*, 4(2), 193–209.
- Karl, T.R., Arguez, A., Huang, B., Lawrimore, J.H., McMahon, J.R., Menne, M.J., Peterson, T.C., Vose, R.S. and Zhang, H.-M. (2015) Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, 348, 1469–1472. <https://doi.org/10.1126/science.aaa5632>
- Kazumori, M., Geer, A.J. and English, S.J. (2016) Effects of all-sky assimilation of GCOM-W/AMS2 radiances in the ECMWF numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 142, 721–737. <https://doi.org/10.1002/qj.2669>
- Kobayashi, S., Matricardi, M., Dee, D.P. and Uppala, S.M. (2009) Toward a consistent reanalysis of the upper stratosphere based on radiance measurements from SSU and AMSU-A. *Quarterly Journal of the Royal Meteorological Society*, 135, 2086–2099. <https://doi.org/10.1002/qj.514>
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K. and Takahashi, K. (2015) The JRA-55 reanalysis: general specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Series II*, 93(1), 5–48. <https://doi.org/10.2151/jmsj.2015-001>
- Köopken, C., Thépaut, J.-N. and Kelly, G.A. (2003). Assimilation of Geostationary WV Radiances from GOES and Meteosat at ECMWF. EUMETSAT/ECMWF Fellowship Programme Research Report 14, ECMWF, Reading, UK.
- Krzeminski, B., Bormann, N., Kelly, G.A., McNally, A.P. and Bauer, P. (2009). Revision of the HIRS cloud detection at ECMWF. EUMETSAT/ECMWF Fellowship Programme Research Report 19, ECMWF, Reading, UK.
- Lalouaux, P., de Boisseson, E., Balmaseda, M.A., Bidlot, J.-R., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D.P., Haimberger, L., Hersbach, H., Kosaka, Y., Martin, M., Poli, P., Rayner, N., Rustemeier, E. and Schepers, D. (2018) CERA-20C: a coupled reanalysis of the twentieth century. *Journal of Advances in Modeling Earth Systems*, 10(5), 1172–1195. <https://doi.org/10.1029/2018MS001273>
- Lalouaux, P., Bonavita, M., Dahoui, M., Farnan, J., Healy, S.B., Hölm, E.V. and Lang, S.T.K. (2020) Towards an unbiased stratospheric analysis. *Quarterly Journal of the Royal Meteorological Society*. <https://doi.org/10.1002/qj.3798>
- Lawrence, H., Bormann, N., Geer, A.J., Lu, Q. and English, S.J. (2018) Evaluation and assimilation of the Microwave Sounder MWS-2 onboard FY-3C in the ECMWF numerical weather prediction system. *IEEE Transactions on Geoscience and Remote Sensing*, 56, 3333–3349.
- Lean, J., Rottman, G., Harder, J. and Kopp, G. (2005) Sorce contributions to new understanding of global change and solar variability. *Solar Physics*, 230, 27–53.
- Letertre-Danczak, J. (2016). Monitoring and operational assimilation of Himawari-9 clear-sky geostationary radiances. Research Department Memorandum RD16-029, ECMWF, Reading, UK.

- Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S.T., Balsamo, G., Bechtold, P., Bonavita, M., Christensen, H.M., Diamantakis, M., Dutra, E., English, S., Fisher, M., Forbes, R.M., Goddard, J., Haiden, T., Hogan, R.J., Juricke, S., Lawrence, H., MacLeod, D., Magnusson, L., Malardel, S., Massart, S., Sandu, I., Smolarkiewicz, P.K., Subramanian, A., Vitart, F., Wedi, N. and Weisheimer, A. (2017) Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, 143, 2315–2339. <https://doi.org/10.1002/qj.3094>
- Long, C., Fujiwara, M., Davis, S.M., Mitchell, D.M. and Wright, C.J. (2017) Climatology and interannual variability of dynamic variables in multiple reanalyses evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP). *Atmospheric Chemistry and Physics*, 17(23)
- Lopez, P. (2011) Direct 4D-Var assimilation of NCEP Stage IV radar and gauge precipitation data at ECMWF. *Monthly Weather Review*, 139, 2098–2116.
- Lupu, C. and Geer, A.J. (2015). Operational Implementation of RTTOV-11 in the IFS. Technical Memorandum 636, ECMWF, Reading, UK.
- Lupu, C. and McNally, A.P. (2012). Assimilation of cloud-affected radiances from Meteosat-9 at ECMWF. EUMETSAT/ECMWF Fellowship Programme Research Report 25, ECMWF, Reading, UK.
- Malardel, S. and Ricard, D. (2015) An alternative cell-averaged departure point reconstruction for pointwise semi-Lagrangian transport schemes. *Quarterly Journal of the Royal Meteorological Society*, 141, 2114–2126.
- Manrique-Suñén, A., Nordbo, A., Balsamo, G., Beljaars, A. and Mammarella, I. (2013) Representing land surface heterogeneity: offline analysis of the tiling method. *Journal of Hydrometeorology*, 14(3), 850–867. <https://doi.org/10.1175/JHM-D-12-0108.1>
- Maycock, A.C., Randel, W.J., Steiner, A.K., Karpechko, A.Y., Christy, J., Saunders, R., Thompson, D.W.J., Zou, C.-Z., Chrysanthou, A., Abraham, N.L., Akiyoshi, H., Archibald, A.T., Butchart, N., Chipperfield, M., Dameris, M., Deushi, M., Dhomse, S., Genova, G.D., Jackel, P., Kinnison, D.E., Kirner, O., Ladstetter, F., Michou, M., Morgenstern, O., O'Connor, F., Oman, L., Pitari, G., Plummer, D.A., Revell, L.E., Rozanov, E., Stenke, A., Visioni, D., Yamashita, Y. and Zeng, G. (2018) Revisiting the mystery of recent stratospheric temperature trends. *Geophysical Research Letters*, 45(18), 9919–9933. <https://doi.org/10.1029/2018GL078035>
- Mayer, M. and Haimberger, L. (2012) Poleward atmospheric energy transports and their variability as evaluated from ECMWF reanalysis data. *Journal of Climate*, 25(2), 734–752.
- Mayer, M., Tietsche, S., Haimberger, L., Tsubouchi, T., Mayer, J. and Zuo, H. (2019) An improved estimate of the coupled Arctic energy budget. *Journal of Climate*, 32, 7915–7934.
- McNally, A.P. and Watts, P.D. (2003) A cloud detection algorithm for high-spectral-resolution infrared sounders. *Quarterly Journal of the Royal Meteorological Society*, 129, 3411–3423. <https://doi.org/10.1256/qj.02.208>
- McNally, A.P., Watts, P.D., Smith, J.A., Engelen, R., Kelly, G.A., Thpaut, J.-N. and Matricardi, M. (2006) The assimilation of AIRS radiance data at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 132, 935–957. <https://doi.org/10.1256/qj.04.171>
- McPeters, R.D., Bhartia, P., Haffner, D., Labow, G. and Flynn, L. (2013) The version 8.6 SBUV ozone data record: an overview. *Journal of Geophysical Research: Atmospheres*, 118(14), 8032–8039. <https://doi.org/10.1002/jgrd.50597>
- Merchant, C.J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C.E., Corlett, G.K., Good, S., McLaren, A., Rayner, N.A., Morak-Bozzo, S. and Donlon, C. (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. <https://doi.org/10.1002/gdj3.20>
- Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N. and Terzhevik, A. (2010) Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Environment Research*, 15(2)
- Molteni, F., Stockdale, T., Balmaseda, M.A., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T.N. and Vitart, F. (2011). The new ECMWF Seasonal Forecast System (System 4). Technical Memorandum 656, ECMWF, Reading, UK.
- Morcrette, J.-J. (1991) Radiation and cloud radiative properties in the European Centre for Medium Range Weather Forecasts forecasting system. *Journal of Geophysical Research: Atmospheres*, 96(D5), 9121–9132.
- Morcrette, J.-J., Barker, H.W., Cole, J.N.S., Iacono, M.J. and Pinus, R. (2008) Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System. *Monthly Weather Review*, 136, 4773–4798.
- Morice, C.P., Kennedy, J.J., Rayner, N.A. and Jones, P.D. (2012) Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. *Journal of Geophysical Research: Atmospheres*, 117(D8). <https://doi.org/10.1029/2011JD017187>
- Muñoz-Sabater, J. (2019) First ERA5-Land dataset to be released this spring. *ECMWF Newsletter*, 159 8–9.
- Munro, R., Kopken, C., Kelly, G.A., Thépaut, J.-N. and Saunders, R. (2004) Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context. *Quarterly Journal of the Royal Meteorological Society*, 130, 2293–2313.
- Nash, J. and Saunders, R. (2015) A review of Stratospheric Sounding Unit radiance observations for climate trends and reanalyses. *Quarterly Journal of the Royal Meteorological Society*, 141, 2103–2113.
- NOAA (2006). 2-Minute Gridded Global Relief Data (ETOPO2) v2. National Geophysical Data Center, NOAA, Asheville, NC.
- Olauson, J. (2018) ERA5: the new champion of wind power modelling?. *Renewable Energy*, 126, 322–331.
- Orr, A., Bechtold, P., Scinocca, J., Ern, M. and Janisková, M. (2010) Improved middle atmosphere climate and forecasts in the ECMWF model through a non-orographic gravity wave drag parameterization. *Journal of Climate*, 23(22), 5905–5926.
- Penny, S.G., Akella, S., Buehner, M., Chevallier, M., Counillon, F., Draper, C., Frolov, S., Fujii, Y., Karspeck, A., Kumar, A., Laloyaux, P., Mahfouf, J.-F., Martin, M., Peña, M., de Rosnay, P., Subramanian, A., Tardif, R., Wang, Y. and Wu, X. (2017). Coupled data assimilation for Integrated Earth System Analysis and Prediction: Goals, challenges, and recommendations. WWRP report 2017-3, WMO, Geneva, Switzerland.
- Peubey, C. and McNally, A.P. (2009) Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context. *Quarterly Journal of the Royal Meteorological Society*, 135, 1863–1873.

- Pincus, R., Barker, H.W. and Morcrette, J.-J. (2003) A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous clouds. *Journal of Geophysical Research: Atmospheres*, 103(4376). <https://doi.org/10.1029/2002JD003322>
- Poli, P., Hersbach, H., Dee, D.P., Berrisford, P., Simmons, A.J., Vitart, F., Laloyaux, P., Tan, D.G.H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E.V., Bonavita, M., Isaksen, I. and Fisher, M. (2016) ERA-20C: an atmospheric reanalysis of the twentieth century. *Journal of Climate*, 29(11), 4083–4097.
- Poli, P., Moll, P., Puech, D., Rabier, F. and Healy, S.B. (2009) Quality control, error analysis, and impact assessment of FORMOSAT-3/COSMIC in numerical weather prediction. *Terrestrial Atmospheric and Oceanic Sciences*, 20, 101–113. [https://doi.org/10.3319/TAO.2008.01.21.02\(F3C\)](https://doi.org/10.3319/TAO.2008.01.21.02(F3C))
- Randel, W.J., Smith, A.K., Wu, F., Zou, C.-Z. and Qian, H. (2016) Stratospheric temperature trends over 1979–2015 derived from combined SSU, MLS, and SABER satellite observations. *Journal of Climate*, 29(13), 4843–4859. <https://doi.org/10.1175/JCLI-D-15-0629.1>
- Raoult, B., Bergeron, C., Alós, A.L., Thépaut, J.-N. and Dee, D.P. (2017) Climate service develops user-friendly data store. *ECMWF Newsletter*, 151, 22–27.
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M. and Becker, E. (2014) The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208.
- Salonen, K. and Bormann, N. (2016). Atmospheric Motion Vector observations in the ECMWF system: Fifth year report 41. EUMETSAT/ECMWF Fellowship Programme Research Report 41, ECMWF, Reading, UK.
- Sandu, I., Beljaars, A., Balsamo, G. and Ghelli, A. (2011) Revision of the surface roughness length table. *ECMWF Newsletter*, 130, 8–10.
- Sandu, I., Beljaars, A., Bechtold, P., Mauritsen, T. and Balsamo, G. (2014) Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models?. *Journal of Advances in Modeling Earth Systems*, 5(2), 117–133.
- Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A.J., Bormann, N. and Lupu, C. (2018) An update on the RTTOV fast radiative transfer model (currently at version 12). *Geoscientific Model Development*, 11(7), 2717–2737.
- Schepers, D., de Boissésou, E., Eresmaa, R., Lupu, C. and de Rosnay, P. (2018) CERA-SAT: A coupled satellite-era reanalysis, ECMWF Newsletter 155, 32–37
- Shepherd, T.G., Polichtchouk, I., Hogan, R.J. and Simmons, A.J. (2018). Report on Stratosphere Task Force. Technical Memorandum 824, ECMWF, Reading, UK.
- Shine, K.P., Barnett, J.J. and Randel, W.J. (2008) Temperature trends derived from Stratospheric Sounding Unit radiances: The effect of increasing CO₂ on the weighting function. *Geophysical Research Letters*, 35(L02710)
- Simmons, A.J., Hortal, M., Kelly, G.A., McNally, A.P., Untch, A. and Uppala, S.M. (2005) ECMWF analyses and forecasts of stratospheric winter polar vortex breakup: September 2002 in the Southern Hemisphere and related events. *Journal of the Atmospheric Sciences*, 62(3), 668–689.
- Simmons, A.J., Willett, K.M., Jones, P.D., Thorne, P.W. and Dee, D.P. (2010) Low-frequency variations in surface atmospheric humidity, temperature and precipitation: inferences from reanalyses and monthly gridded observational datasets. *Journal of Geophysical Research*, 115, 22987–22994.
- Simmons, A.J., Poli, P., Dee, D.P., Berrisford, P., Hersbach, H., Kobayashi, H. and Peubey, C. (2014) Estimating low-frequency variability and trends in atmospheric temperature from the ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 140, 329–353.
- Simmons, A.J., Berrisford, P., Dee, D.P., Hersbach, H., Hirahara, S. and Thépaut, J.-N. (2017) A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets. *Quarterly Journal of the Royal Meteorological Society*, 143, 101–119.
- Simmons, A.J., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., Flemming, J., Haimberger, L., Healey, S.B., Hersbach, H., Horányi, A., Inness, A., Muñoz-Sabater, J., Radu, R. and Schepers, D. (2020). Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1. Technical Memorandum 859, ECMWF, Reading, UK.
- Steinbrecht, W., Hegglin, M.L., Harris, N. and Weber, M. (2018) Is global ozone recovering?. *Comptes Rendus Geoscience*, 350(7), 368–375.
- Stoffelen, A. and Anderson, D. (1997) Scatterometer data interpretation: estimation and validation of the transfer function CMOD4. *Journal of Geophysical Research: Oceans*, 102(C3), 5767–5780.
- Takacs, L.L., Suárez, M.J. and Todling, R. (2016) Maintaining atmospheric mass and water balance in reanalyses. *Quarterly Journal of the Royal Meteorological Society*, 142, 1565–1573.
- Tarek, M., Brissette, F.P. and Arseneault, R. (2019) Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modeling over North America. *Hydrology and Earth System Sciences Discussion*, 2019, 1–35. <https://doi.org/10.5194/hess-2019-316>
- Tavolato, C. and Isaksen, I. (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, 1514–1527.
- Tetzner, D., Thomas, E. and Allen, C. (2019) A validation of ERA5 reanalysis data in the Southern Antarctic peninsula – Ellsworth Land region, and its implications for ice core studies. *Geosciences*, 9. <https://doi.org/10.3390/geosciences9070289>
- Thépaut, J.-N., Dee, D.P., Engelen, R. and Pinty, B. (2018). The Copernicus programme and its climate change service. pp. 1591–1593 in IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain.
- Tiedtke, M. (1989) A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, 117(8), 1779–1800.
- Tiedtke, M. (1993) Representation of clouds in large-scale models. *Monthly Weather Review*, 121(11), 3040–3061.
- Titchner, H.A. and 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(6), 2864–2889.
- Trenberth, K.E., Fasullo, J.T. and Balmaseda, M.A. (2014) Earth's energy imbalance. *Journal of Climate*, 27(9), 3129–3144. <https://doi.org/10.1175/JCLI-D-13-00294.1>
- Trenberth, K.E., Fasullo, J.T. and Kiehl, J. (2009) Earth's global energy budget. *Bulletin of the American Meteorological Society*, 90, 311–323.

- Trenberth, K.E. and Smith, L. (2005) The mass of the atmosphere: a constraint on global analyses. *Journal of Climate*, 18(6), 864–875.
- Trollet, M., Walawender, J.P., Bourlès, B., Boilley, A., Trentmann, J., Blanc, P., Lefèvre, M. and Wald, L. (2018) Downwelling surface solar irradiance in the tropical Atlantic Ocean: a comparison of re-analyses and satellite-derived data sets to PIRATA measurements. *Ocean Science*, 14(5), 1021–1056.
- Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E.V., Hoskins, B.J., Isaksen, I., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P. and Woollen, J. (2005) The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131, 2961–3012.
- Urraca, R., Huld, T., Gracia-Amillo, A., Martinez-de Pison, F.J., Kaspar, F. and Sanz-Garcia, A. (2018) Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Solar Energy*, 164, 339–354.
- van den Hurk, B.J., Viterbo, P., Beljaars, A. and Betts, A. (2000). Offline validation of the ERA40 surface scheme. Technical Memorandum 295, ECMWF, Reading, UK.
- Wernli, H., Dirren, S., Liniger, M.A. and Zillig, M. (2002) Dynamical aspects of the life cycle of the winter storm *Lothar* (24–26 December 1999). *Quarterly Journal of the Royal Meteorological Society*, 128, 405–429.
- Woodruff, S.D., Worley, S.J., Lubker, S.J., Ji, Z., Freeman, J.E., Berry, D.I., Brohan, P., Kent, E.C., Reynolds, R.W., Smith, S.R. and Wilkinson, C. (2011) ICOADS Release 2.5: extensions and enhancements to the surface marine meteorological archive. *International Journal of Climatology*, 31(7), 951–967. <https://doi.org/10.1002/joc.2103>
- Xu, X., Frey, S.K., Boluwade, A., Erler, A.R., Khader, O., Lapen, D.R. and Sudicky, E. (2019) Evaluation of variability among different precipitation products in the Northern Great Plains. *Journal of Hydrology: Regional Studies*, 24(100608). <https://doi.org/10.1016/j.ejrh.2019.100608>
- Zuo, H., Balmaseda, M.A., Mogensen, K. and Tietsche, S. (2018). OCEAN5: The ECMWF Ocean Reanalysis System and its real-time analysis component. Technical Report 823, ECMWF, Reading, UK.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Hersbach H, Bell B, Berrisford P, *et al.* The ERA5 global reanalysis. *Q J R Meteorol Soc.* 2020;146:1999–2049. <https://doi.org/10.1002/qj.3803>