

# ERA report series



## 20 ERA-20C Deterministic

---

Paul Poli, Hans Hersbach, Paul Berrisford, Dick Dee, Adrian Simmons  
and Patrick Laloyaux

Series: ERA Report Series

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/en/research/publications>

© Copyright 2015

European Centre for Medium Range Weather Forecasts  
Shinfield Park, Reading, Berkshire RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

## Abstract

The ERA-20C deterministic reanalysis resolves several issues identified in the earlier ERA-20C ensemble reanalysis. Foremost, it produces more realistic climate trends, thanks to background error correlations from the ECMWF operational system, instead of ones generated by the ERA-20C ensemble. However, this setup brings also an unwanted consequence, slightly degrading the weather analyses, owing to the use of background error structure functions that were not tuned specifically to the surface-only observing system. The observation quality controls are also generally improved as compared to ERA-20C ensemble, except for the tropical cyclone bogus observations, which are largely rejected in both ERA-20C reanalyses.

The present report shows that the inter-annual variability of total column water vapor over tropical ocean is better represented by ERA-20C than by ERA-Interim and by JRA-55, using HOAPS and RSS observational products as references. The time evolution of the total column ozone over northern high latitudes is also found to be realistic, as compared to the Multi-Sensor Reanalysis (MSR-2) observational product for the years 1979–2008, and of better quality there than the model integration ERA-20CM. Assimilating enough surface pressure observations may be sufficient to properly locate the polar vortex and represent correctly the distribution of stratospheric ozone *via* the model dynamics and radiation. The report also proposes a simple but robust way of quantifying our ability to assess climate change within the reanalysis context, by considering the analysis increments in the space of atmospheric total energy change.

Regarding the general circulation, based on several metrics, ERA-20C deterministic appears adequate in the Northern hemisphere mostly after World War II. In the Southern hemisphere, the quality of the circulation is much more doubtful until the late 1980s (which coincides with a wider deployment of drifting buoys). This is supported by several metrics, as well as the inter-annual variability of total column ozone.

As the ERA-20C deterministic reanalysis only provides a single realisation, the use of the spread between the various members of the earlier ERA-20C ensemble production is explored as a proxy for estimating uncertainties. For two-metre temperature, the ensemble spread may be used to assess uncertainties in synoptic weather charts, but not uncertainties related to climate time-scales (the spread between monthly means is probably too small to be realistic). This is not surprising because, except for the SST forcing, all sources of uncertainties were prescribed with a short memory and no time-correlation. During the course of the century, the ensemble spread in two-metre temperature decreases as the corresponding intra-monthly variability increases. Comparing with similar metrics from the model integration ERA-20CM, this result is explained by the assimilation of synoptic weather observations.

For a future repeat of a 20th century reanalysis at ECMWF, several leads for improvement are proposed in the report. Most are related to the observation data handling. Otherwise, for further improvements, the report indicates that in most cases the barriers to improving our knowledge of the 20th century weather lie with the availability of observations.

## Contents

<b>1</b>	<b>Introduction and rationale for this deterministic production</b>	<b>1</b>
<b>2</b>	<b>Methodology</b>	<b>2</b>
2.1	Model configuration . . . . .	2
2.2	Observation data selection and quality control . . . . .	2
2.3	Data assimilation configuration . . . . .	4
2.4	Observation bias correction . . . . .	7
2.5	Multi-stream production . . . . .	10
<b>3</b>	<b>Products</b>	<b>10</b>
3.1	Fields . . . . .	10
3.2	Observation feedback . . . . .	11
<b>4</b>	<b>Data assimilation performance</b>	<b>11</b>
4.1	Forecast scores . . . . .	11
4.2	background and analysis departures, and bias correction . . . . .	12
4.3	Analysis increments . . . . .	14
4.4	Observations assimilated . . . . .	18
4.5	Observation errors . . . . .	19
<b>5</b>	<b>Climate value</b>	<b>22</b>
5.1	Two-metre temperature . . . . .	22
5.2	Total column water vapor . . . . .	22
5.3	Precipitation . . . . .	24
5.4	Ozone . . . . .	25
5.5	Large-scale and regional circulations . . . . .	26
5.6	Energy budget . . . . .	26
<b>6</b>	<b>Uncertainties</b>	<b>33</b>
6.1	Three-hourly ensemble spreads . . . . .	33
6.2	On the limits of these estimates as measures of uncertainties . . . . .	33
<b>7</b>	<b>Issues</b>	<b>38</b>

7.1	Rejection of tropical cyclone bogus observations . . . . .	38
7.2	Dubious wind observations from some buoys . . . . .	38
7.3	Missing tropical waves . . . . .	38
<b>8</b>	<b>Unknowns: open questions</b>	<b>38</b>
8.1	Impact of observation coverage change on climate trends . . . . .	40
8.2	Providing meaningful estimates of uncertainties relevant for the climate time-scales . . . . .	40
8.3	Suspicious trends in the Southern Hemisphere in the early part of the 20th century . . . . .	40
<b>9</b>	<b>Conclusions and outlook</b>	<b>41</b>

# 1 Introduction and rationale for this deterministic production

In 2011, the European Union funded a project called *European Reanalysis of Global Climate Observations* (ERA-CLIM), lead by D. Dee. This project enabled the first production in Europe of a reanalysis of the 20th century weather. Conceptually similar to that of Compo *et al.* (2011), this reanalysis used an ensemble and only assimilated surface observations, as documented in ERA Report Series 14 (Poli *et al.*, 2013). This effort was part of a comprehensive strategy to develop European capability to monitor all components of the climate, going back as far as possible in time as allowed by the instrumental record, and bringing together observations of the various spheres of the Earth system, including the atmosphere, the ocean, the terrestrial component, the cryosphere, and the biosphere (Dee *et al.*, 2014).

This initial production is hereafter referred to as the ERA-20C ensemble (see ERA Report Series 14 (Poli *et al.*, 2013) for a comprehensive report). The aim was to produce a global dataset covering the 20th century with the following features: (1) representing with fidelity the global weather, (2) reproducing the evolution of the global climate, and (3) providing a measure of uncertainty in the form of a ten-member ensemble. The aforementioned report, referenced as RS14 in the remainder of the present report, indicates that the goals (1) and (3) were reasonably met in the troposphere. However, goal (2) was missed in most of the troposphere and stratosphere. The results proved that the analysis produced increments away from the surface, destroying the otherwise reasonable low-frequency variability of the forecast model (driven by realistic forcings, see Hersbach *et al.*, 2015a, for a climate model integration without data assimilation).

In retrospect, the challenge to represent properly the climate time-scales stemmed from insufficient attention during system development to diagnose in advance the quality of the low-frequency information. The focus on the quality of the weather maps and the ensemble spreads during system development paid off, but similar efforts should have been devoted to the problem of representing the climate scales.

Much was learnt from this initial effort regarding the feasibility of gauging in advance how well a reanalysis may reproduce the climate time-scales. For example, one can use a series of assimilation experiments, spread years apart, or reduce the observing system to represent fairly such years. This can help inform about the level of confidence that one may place upon the low-frequency information. This can be achieved without completing a full and continuous multi-decadal reanalysis record.

From these lessons, and also to address more minor issues identified in RS14, a reanalysis system was assembled in Autumn 2013 for a partial rerun of the initial ERA-20C production, considering only a single member, hence qualified as ‘deterministic’ in the present report. This deterministic production started in December 2013 and was completed in February 2014 (*i.e.*, under two months).

It would have been preferable to re-produce a full ensemble set, but this would have been too costly, while at the same time not essential in order to address the main issue of unrealistic stratospheric climate trends in the ERA-20C ensemble. Unfortunately, the uncertainty estimates could not be re-produced again with a single-member run. However, such estimates may still be formed from the ten-member ensemble spread produced earlier. This possibility is discussed in the present report.

The document outline is the following. Section 2 explains the differences between the ensemble and deterministic productions. Section 3 lists the ERA-20C products available to users. Section 4 is an initial evaluation of the assimilation performance. Section 5 answers partially the question of the ‘climate quality’ of the products, discussing low-frequency variability of a few parameters. Section 6 discusses how uncertainties may be presented to users, and what these uncertainties represent. Section 7 presents known issues in the ERA-20C deterministic product. Section 8 discusses unknowns, or major questions that remain open. Finally, conclusions are presented in Section 9.

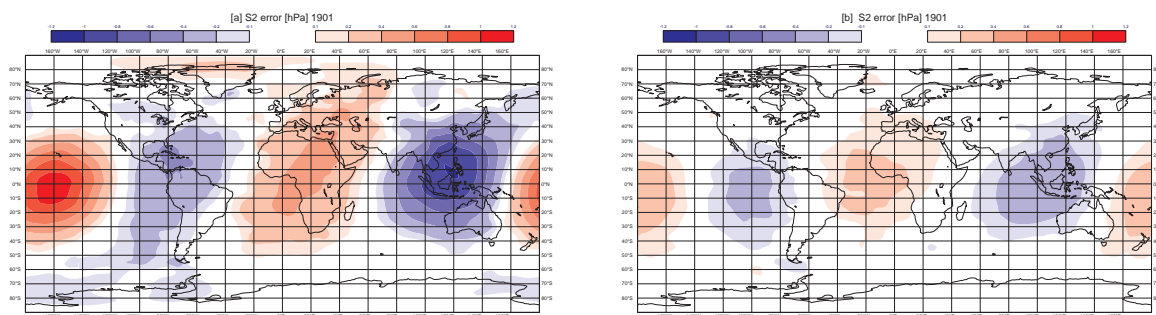


Figure 1: Mean error in semi-diurnal tide ( $S_2$ ) representation of surface pressure in the background, using the analysis as a verification, for 1 January–31 May 1901. Left: (a) 60-minute time-step (ERA-20C ensemble member 0). Right: (b) 30-minute time-step (configuration used in ERA-20C deterministic).

## 2 Methodology

The reanalysis presented in this report builds on the ERA-20C ensemble reanalysis system described in RS14. The present report only lists differences.

Table 1 lists the issues found in the ERA-20C ensemble, as documented in the last section of RS14, and provides a summary of the solutions adopted. Differences between the two ERA-20C systems can be grouped into the following categories: model configuration, data handling (data selection, quality control, and bias correction), data assimilation configuration, and multi-stream production.

### 2.1 Model configuration

The model time-step is reduced from 60 minutes to 30 minutes. RS14 indicates that the 1-hour time-step caused a misrepresentation of the atmospheric tides in the ensemble. Assimilation experiments conducted to test this impact demonstrate that these tidal errors are reduced significantly with a shorter model time-step. Figure 1 shows that the local maxima of 1.2 hPa in the semi-diurnal tide error is reduced to 0.4 hPa with the reduced time-step. Reducing this time-step also has the consequence of reducing the signature of these mis-represented tides in the surface pressure analysis increments (2).

### 2.2 Observation data selection and quality control

The *a priori* quality control is updated to reject observations of surface pressures contained in CTD, XCTD, MBT, and XBT reports. This decision is supported by findings presented in RS14 showing that these observations present larger-than-normal observation errors. It is also further supported by the unclear origin of these measurements; such sondes do not measure atmospheric pressure so pressure measurements reported in such reports probably come from neighboring and possibly less well documented sensors.

Another modification of the quality control is the rejection of observations of geopotential (these were assimilated in the ensemble). These observations of geopotential are computed from observations of surface pressure, and typically assimilated in their place for stations located over high terrain in the IFS (ECMWF, 2013). The reason for their rejection in ERA-20C deterministic is that these observations suffer sometimes from large biases, and the assimilation system does not currently have a bias correction scheme for geopotential observations. In addition, such observations, if biased, have a greater potential to disrupt the large-scale flow, owing to

Ensemble production Table adapted from RS14				Deterministic production	
Issue no.	Nature	Diagnostic	Possible solution	Solution adopted	See section
1	Analysis increments spread vertically introducing spurious trends	Vertical mean analysis increments show slow, but large time variations	Ideally: Remove model bias More practical: Modify background errors	Background error correlations time-invariant from ECMWF Operations	2.3
2	Background check too loose (18 times expected error) and some bad data not rejected	Instances of large obs minus background departures assimilated	Revise bg. check settings and edit blacklist to remove gross errors by inspecting ERA-20C feedback	Reduced bg. check allowance and introduced new QC to remove constant timeseries	2.2
3	Variational bias correction of surface pressure slower than expected	Global RMS of bias correction converges after 2–4 years	Check procedure	VarBC on first-guess departures and correct a bug in procedure	2.4
4	Model bias in representation of atmospheric tides	Geographical patterns in analysis increments of surface pressure	Reduce model time-step	Model time-step reduced from 60 to 30 minutes	2.1
5	Suboptimal observation errors	Desroziers diagnostics and comparison between observed and expected RMS of innovations	Use estimates as documented in this report	Found that using new obs. errors estimate requires new ensemble to derive new backgd. errors	4.5
6	Large error estimates for pressures from CTD/XCTD/MBT/XBT observations	About 2 times that of other observations	Understand the cause, or blacklist	Blacklist pressure obs from (X)CTD, MBT, and XBT	2.2
7	Background error lags seasonally	Maximum of error in observed RMS innovation lags behind the predicted values	Use past and future states from ERA-20C ensemble in future estimation	N/A see point 1. above	

Table 1: Summary of the issues found in the ERA-20C ensemble production as documented in RS14, and solutions adopted in the ERA-20C deterministic production.



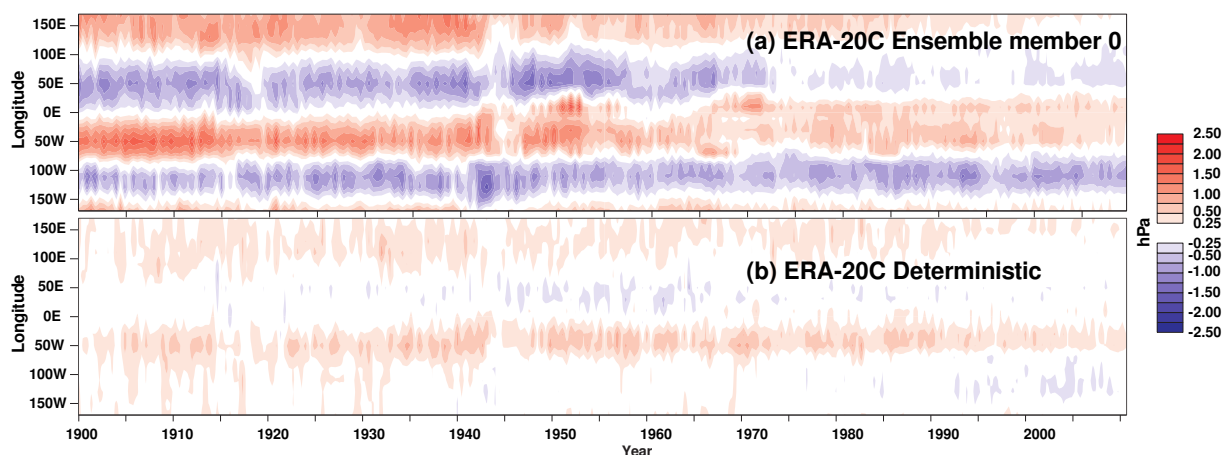


Figure 2: Longitude (vertical axis) versus date (horizontal axis) average surface pressure analysis increment at the beginning of the analysis window, for the latitude band  $10^{\circ}\text{S}$ – $10^{\circ}\text{N}$  in (a) ERA-20C ensemble member 0 and (b) ERA-20C deterministic.

their position high in altitude.

Another problem found in the ERA-20C ensemble was the assimilation of obviously bad observations. This was unfortunately possible because of the very loose rejection criterion of 18 times the expected error standard deviation. After some testing, this limit is reduced for ERA-20C deterministic to 7 times the expected error. This effectively reduces the maximum acceptable observation departure limit from about 310 hPa to about 120 hPa for pressure observations.

Last, the assimilation of bad observations in the ERA-20C ensemble was traced in two cases to constant time-series in the reported pressures. In the early days of automated stations, this could be caused by a stuck pressure sensor and insufficient monitoring and/or maintenance. However, in the earlier times of human observations, this is more likely an error in data transcription, keying, or reporting. To address this issue, a novel quality control procedure is included in ERA-20C deterministic: pressure observations from a given station, at a given date and time, are rejected from the assimilation if all the observations reported within a 5-day time-window centred on the date and time of interest are constant, provided there are at least 3 observations within that 5-day window. Table 2 lists the ten land station with the largest numbers of data detected by this test. Such scheme captures a problem of constant pressures reported continuously, at exactly 1000.10 hPa, 8847 times in 1992–1994, by the station ‘766910’ near the Gulf of Mexico. Implementing similar quality controls could help to improve databases of *in situ* observations.

Finally, a few stations are also blacklisted upon advice from Compo (*pers. comm.*, 2013): 8 stations for which the records of surface pressure and mean-sea-level pressure are interchanged, and 1 ship for which the position is known to be incorrect. We now know of other incorrect ship locations (“D-Day Analyses”, published online <http://www.ecmwf.int/en/research/projects/era-clim/d-day-analyses>, last accessed 2 June 2015).

### 2.3 Data assimilation configuration

The origin of the poor representation of climate trends away from observations in the ERA-20C ensemble was traced to the background error covariances computed on-the-fly. These covariances allowed increments caused by surface observations to propagate to unobserved areas in the upper troposphere and even higher. After the ERA-20C ensemble, several experiments were conducted to introduce vertical localisation of the analysis increments; the goal was to force the background errors to decrease vertically to near-zero at some level. After

Station ID	No. of obs.	Pressure (Pa)	Start date	End date
837210	45791	94620	19510120	19560829
836120	39061	92410	19510101	19560718
835830	34355	100020	19510801	19550801
821930	28464	100710	19510112	19551210
837480	17809	100790	19550103	19570617
765773	13951	100010	19920609	19940630
766753	13924	100010	19920416	19940630
309110	13048	101420	19640102	19681229
765255	11828	100010	19920609	19940630
362310	11696	104210	19640102	19681229

Table 2: Top 10 of the land stations reporting constant pressures for some time periods.

several trials, including using the same localisation settings as Compo *et al.* (2011), the temperature increments in the upper troposphere were indeed suppressed. However, none of the experiments produced fully satisfactory results; new, larger increments, appeared in the middle troposphere, and these proved to reduce dramatically the quality of the synoptic analyses.

The solution employed in ERA-20C deterministic is to revert to the time-invariant background error covariances of ECMWF operations as of 2013, and to rescale the global variances therein to match the time-varying background errors computed by the ERA-20C ensemble. The time-varying rescaling factor is shown in Figure 3. It corroborates features observed in ERA-20C ensemble, such as generally decreasing background errors over the course of the century, punctuated by larger errors around the two World Wars, and a faster rate of improvement after WWII. The intra-year fluctuations summarised in Figure 4 indicate that the global means of background errors are the lowest in March-April, and greatest in December and September. However, the variations between the years are greatest in February-March, and lowest in July to November. Such conclusions are difficult to interpret as they represent the combined effect of variability and predictability over land and ocean, as well as variations in the observing system and the method employed in the ERA-20C ensemble to derive background errors (which were found to be lagging seasonally, see issue number 7 in Table 1)

Using this combination of background error correlations with global variances rescaled by a time-varying factor removes the unwanted temperature analysis increments in the stratosphere. Also, the much-needed time-variation of background variances are preserved (so the assimilation weights the observations accordingly as the background quality improves).

One notes that now the background error correlations are not derived from a poorly observed system, but from ECMWF Operations, and hence may not be adapted to the analysis system at hand. In fact, a quick inspection of synoptic maps produced with this setup confirms a slight degradation of the synoptic analyses produced, as compared to the ensemble reanalysis. For example, Figure 5 shows the analysis for the European Great Storm of 1987, from various systems, including ECMWF operations at the time, as well as analyses and forecasts from the following reanalyses: ERA-15, ERA-40, ERA-Interim, ERA-20C ensemble (ensemble mean analysis, and member 0 forecasts), ERA-20C deterministic, and NOAA 20CR (ensemble mean analysis). The 24-hour forecast from ERA-20C deterministic is quite degraded as compared to that from ERA-20C ensemble. Looking at other historical weather cases indicates usually similar results (not shown). The background errors in ERA-20C deterministic use correlations from ECMWF operations; these correspond to a well-observed analysis system that routinely ingests upper-air in situ and satellite observations. The ERA-20C ensemble used correlations that were fully consistent with the observations assimilated. RS14 indicates that the latter present wider horizontal correlations lengths. Overall, this degradation result confirms the importance of using



Figure 3: Timeseries of the inflation factor applied to the background error covariances from ECMWF operations as of 2013 to match the level of background errors estimated by ERA-20C ensemble.

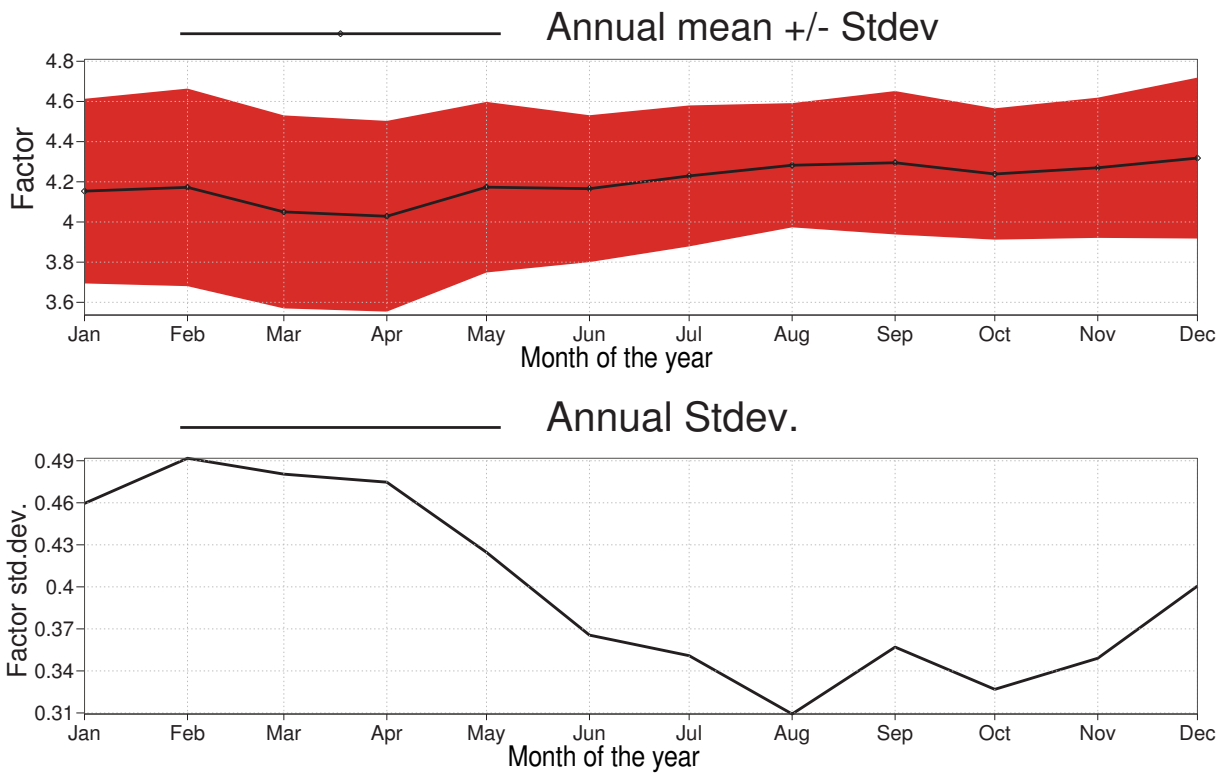


Figure 4: Same as Figure 3, but showing the annual cycle.

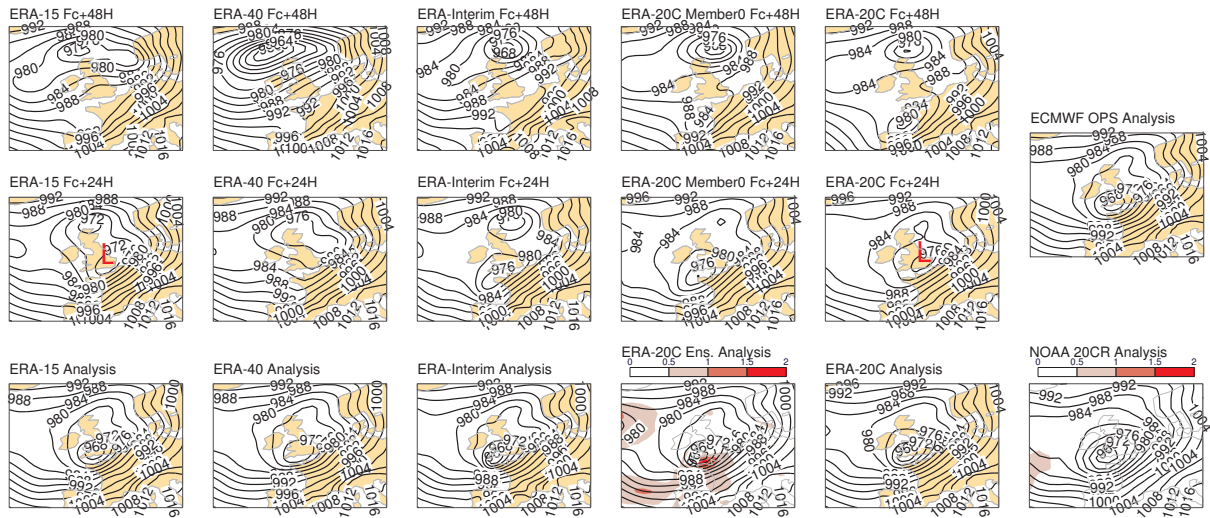


Figure 5: Analyses and forecasts for 16 October 1987, 00 UTC. Color shading shown for ensemble systems (analyses from ERA-20C ensemble and from NOAA 20CR ensemble) is the ensemble spread.

background forecast errors structure functions fit for the observing system employed by the assimilation.

## 2.4 Observation bias correction

In both ERA-20C productions, corrections for surface pressure and mean sea level pressure are based on the method of variational bias correction, VarBC (Dee, 2004). In this formalism, observation bias is estimated from a (linear) bias model based on a small set of bias parameters. Data are stratified into groups where observations share the same set of bias parameters. There is one group per station and reporting practice (sea-level, or station-level). The bias parameters are updated variationally in the 4D-Var assimilation, as additional degrees of freedom, part of the analysis control vector. The level of confidence on the bias correction estimated earlier is controlled by a term in the 4D-Var cost function that penalises for deviations from the background bias parameters. This also avoids potential over-fitting of the bias model.

Changes, slow and/or abrupt, result from *e.g.*, degrading instruments, a change in instrumentation, or a change in the station's environment or its location. For this reason, in ERA-20C each physical station or platform forms its own bias group in which the bias estimate can evolve independently. Each analysis cycle, only a few observations may be found in each bias group (from zero to typically one daily, or many more for Coastal-Marine Automated Network (C-MAN) stations). The number of bias groups per analysis window runs in the tens of thousands, and collected over the entire century more than two million different stations and platforms are identified. This is quite a different situation from operational NWP, where typically thousands of satellite observations all share the same group, and the number of bias groups is much lower.

The bias model in ERA-20C is the station-dependent bias estimate in pressure itself. In this very simple local one-dimensional bias model it is very important that the background term is not too large, since otherwise a large part of the signal would be taken out of the observations. On the other hand, the background term should also not be too stringent, since in such case the response to breaks in bias characteristics would be too slow. To compromise between these two competing requirements, a breakpoint analysis was performed prior to the integration of ERA-20C. It is based on 12-month moving averages of departure statistics from the 20CR reanalysis. Result of this analysis is a quantity, called bias volatility, ranging from 0 (no likely break) to 1 (certainly a break). It is related to the standard normal homogeneity test (SNHT, Haimberger 2007,

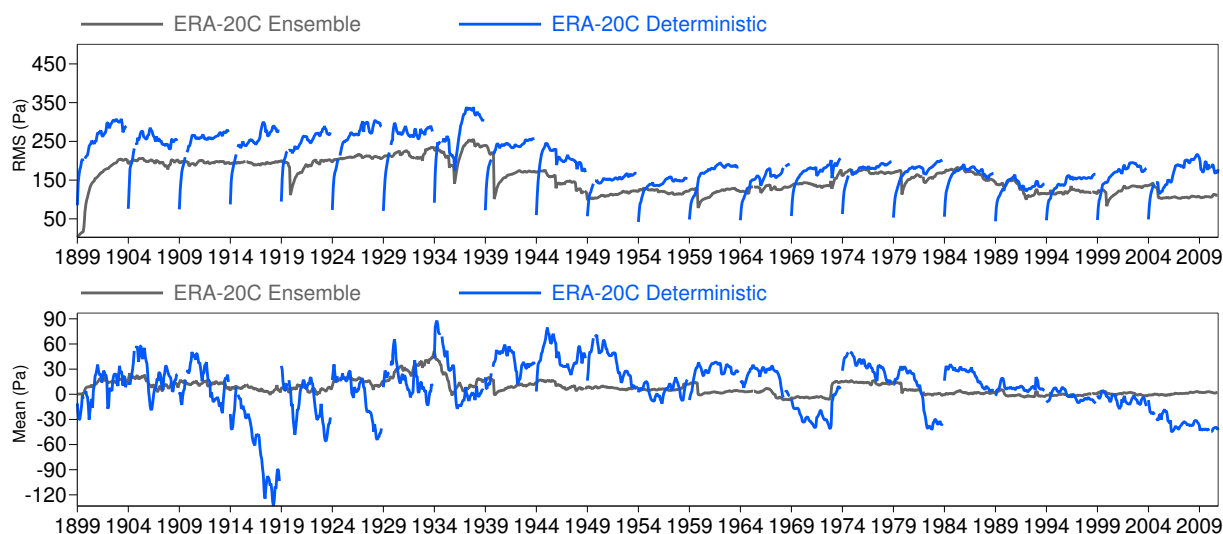


Figure 6: Timeseries of RMS (top) and mean (bottom) bias correction for observations of surface pressure assimilated in the ERA-20C ensemble member 0 and ERA-20C deterministic.

Alexandersson 1986). Details on the calculation of the bias volatility are given by Hersbach *et al.* (2015b), who also address the problem of station identification when limited information on platform name is available (such as ‘SHIP’ for many early maritime observations).

In the ERA-20C ensemble, the initial weight of the bias background term was chosen to be 60 days worth of observations, corresponding to a similar response time. As long as the bias volatility remained below a threshold value (0.15), each day, one day was added to this response time to express an increasing confidence in the bias estimate. However, whenever the bias volatility exceeded this threshold, a break was suspected, and the response time was reset to the initial value of 60 days. After the start of the production of ERA-20C ensemble it soon became apparent that this bias adaptivity was much slower than anticipated. After a detailed analysis it was found that this emerged once the ensemble used its own background error correlations. The reason for this is unclear, but it is suspected that the longer correlation lengths had an adverse effect on the pre-conditioning of VarBC, leading to an incomplete optimisation of the cost function with respect to the bias parameters. This problem was addressed by pragmatically reducing the bias background weight by a factor of ten, after which the adaptivity was found to be largely improved. Later it was found that, for certain classes of stations, a software bug affected the mechanism intended to respond to known bias breaks. As a result for such data the bias became less and less adaptive, and the correction scheme was taking more than a year to converge.

In ERA-20C deterministic these issues disappear: the software bug is corrected and the adaptation of the bias parameters is performed before 4D-Var on background departures, rather than inside the assimilation. In this manner any interaction with the model background is avoided, and the response time is controlled exactly as desired. The response times themselves are also modified, now starting with (and resetting to) 30 days, and capping the response time to 90 days to avoid a too slow adaptivity. An example for one station is presented in Figure 1 of Hersbach *et al.* (2015b).

The final bias corrections are larger as a result (Figure 6). In spite of the 22 5-year streams, the agreements between most mean bias corrections (especially in the recent years) suggest that the deterministic production finds more appropriate mean corrections.

Stream number	Experiment identifier	Start date	End date
1	1852	1 January 1899	31 December 1904
2	1853	1 January 1904	31 December 1909
3	1854	1 January 1909	31 December 1914
4	1855	1 January 1914	31 December 1919
5	1856	1 January 1919	31 December 1924
6	1857	1 January 1924	31 December 1929
7	1858	1 January 1929	31 December 1934
8	1859	1 January 1934	31 December 1939
9	1860	1 January 1939	31 December 1944
10	1861	1 January 1944	31 December 1949
11	1862	1 January 1949	31 December 1954
12	1863	1 January 1954	31 December 1959
13	1864	1 January 1959	31 December 1964
14	1865	1 January 1964	31 December 1969
15	1866	1 January 1969	31 December 1974
16	1867	1 January 1974	31 December 1979
17	1868	1 January 1979	31 December 1984
18	1869	1 January 1984	31 December 1989
19	1870	1 January 1989	31 December 1994
20	1871	1 January 1994	31 December 1999
21	1872	1 January 1999	31 December 2004
22	1873	1 January 2004	31 December 2010

Table 3: Streams employed for ERA-20C deterministic production. Experiment identifier refers to the Meteorological Archive and Retrieval System (MARS).

## 2.5 Multi-stream production

In order to achieve a fast rerun, the deterministic production employs 22 parallel streams, which is more than the 6 streams of the ensemble production. The streams, shown in Table 3, are separated by 5 years, usually covering 6 years (except for the last one). The first year of each stream is for spin-up. RS14 demonstrated that a two-year overlap was not bringing much improved consistency in terms of meteorology between the streams; for that reason, each stream in this second production is stopped as soon as a one-year overlap has been completed with the following stream's spin-up year. The ERA-20C deterministic production started on 17 December 2003. After about two weeks of tuning to improve turnaround and submission of jobs on the supercomputer, the production progressed at about 50 days/day for each stream, and was completed in about six weeks (February 2014), making this production the fastest of all major atmospheric reanalyses produced so far at ECMWF.

Contrary to the problems encountered in the ensemble production and documented in section 2.6 of RS14, this production only encountered 1 minimisation failure that required halving the model time-step (*i.e.*, from 30 to 15 minutes): for the analysis of 21 April 2004, stream number 21, experiment identifier 1872. This more stable behavior is very likely due to the smaller model time-step (30 minutes, instead of 60 minutes previously for the ensemble production).

## 3 Products

The ERA-20C products are available from the Meteorological Archive and Retrieval System (MARS) under *class=e2, expver=0001* and from the ECMWF Public Datasets web interface <sup>1</sup>.

### 3.1 Fields

The ERA-20C fields are available at horizontal truncation T159, or approximately 125 km resolution. There are eight sets of products:

- invariant (constant over time),
- so-called daily (in fact available 3-hourly for most of them),
- monthly means of daily means (available monthly),
- synoptic monthly means (available monthly, with separate means for each 3-hourly synoptic hour),
- ocean wave invariant (constant over time),
- ocean wave daily (available 3-hourly),
- ocean wave monthly means of daily means (available monthly), and
- ocean wave synoptic monthly means (available monthly, with separate means for each 3-hourly synoptic hour).

Two types of products are available:

---

<sup>1</sup>Direct access to ERA-20C data is available from: <http://apps.ecmwf.int/datasets/data/era20c-daily/>

- analysis (MARS keyword *type=an*), and
- forecast (*type=fc*).

The ocean wave products only characterize the ocean surface (on 25 frequencies and 12 directions), whereas the other products are available on various level types:

- surface (MARS keyword *levtype=sfc*),
- 37 pressure levels (*levtype=pl*: 1000, 975, 950, 925, 900, 875, 850, 825, 800, 775, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 225, 200, 175, 150, 125, 100, 70, 50, 30, 20, 10, 7, 5, 3, 2, 1 hPa),
- 91 model levels <sup>2</sup> (*levtype=ml*),
- 16 potential temperature levels (*levtype=pt*: 850, 700, 600, 530, 475, 430, 395, 370, 350, 330, 320, 315, 300, 285, 275, 265 K), and
- the 2 Potential Vorticity Unit (PVU) level (*levtype=pv*).

The parameters available for each combination of product set, type, and level are listed in the ECMWF Public Datasets web interface. The daily or ocean wave daily products are available 3-hourly, except for analyses of the following surface parameters, only available 6-hourly: sea surface temperature and ice cover, volumetric soil water layers 1 to 4, soil temperature levels 1 to 3, 2-metre temperature, 2-metre dewpoint temperature, snow density and depth, and temperature of snow layer.

## 3.2 Observation feedback

The ERA-20C observation feedback (MARS keyword *type=ofb*) contains all the observations input to ERA-20C from ISPD v3.2.6 (Yin *et al.*, 2008) and ICOADS v2.5.1 (Woodruff *et al.*, 2011). The observation feedback attributes available are listed in Table 4. More details about each of these attributes, and their traceability to the input datasets, can be found in the report by Hersbach *et al.* (2015b). The native format of the archive is as described by Kuchta (2009), but data can also be retrieved in text format.

# 4 Data assimilation performance

## 4.1 Forecast scores

Both ERA-20C productions include the daily integration of the model for 10-day forecasts. The initial conditions are analyses at 00 UTC. The accuracy of those forecasts can be computed by comparing to verifying analyses, assuming these analyses are representative of the truth. However, when few observations are assimilated, *e.g.*, in remote areas at the beginning of the century, the analysis is essentially unchanged from the previous background; in such case the forecasts can appear to be very good, because there are no observations to pull the analyses away from the model trajectory. Figure 7 shows the forecast scores at day 3 between January 1900 and December 2010. The scores in the Northern Hemisphere increase in two phases, around 1945 and around 1975. By the argument given earlier, the scores in the Southern Hemisphere extratropics are not

<sup>2</sup>Model level definitions can be found at the following URL: <http://www.ecmwf.int/en/where-can-i-find-ecmwf-model-level-definitions>



MARS attributes
expver (always '0001'), class (always 22), stream (always 1025), type (always 263), andate, antime (always 0), reportype, groupid (always 17)
Observation type attributes
obstype@hdr, codetype@hdr, station_type@conv
Observation time and location attributes
date@hdr, time@hdr, lat@hdr, lon@hdr
Observation vertical location attributes
stalt@hdr, vertco_type@body, vertco_reference_1@body, ppcode@conv_body
Observation values attributes
varno@body, obsvalue@body
Observation identification attributes
source@hdr, collection_identifier@conv, unique_identifier@conv, timeseries_index@conv, statid@hdr, seqno@hdr
Observation environment attributes (background estimates)
lsm@modsurf, orography@modsurf, solar_elevation@conv, seaice@modsurf, t2m@modsurf, windspeed10m@modsurf, u10m@modsurf, v10m@modsurf
Observation quality control attributes
datum_qcflag@conv_body, datum_event1@body, datum_status@body
Data assimilation feedback attributes
biascorr@body, varbc_ix@body, bias_volatility@body, qc_pge@body, eda_spread@errstat, an_depar@body, fg_depar@body, an_sens_obs@body, obs_error@errstat, fg_error@errstat

Table 4: List of observation feedback attributes available from ERA-20C. See Hersbach et al. (2015b) for more details.

meaningful until the 1960s. They are nevertheless interesting to show, because they appear to be relevant after the 1960s. This means that before this date the quality of the reanalysis is likely to be poor for most of the synoptic weather in the Southern Hemisphere extratropics. In spite of this apparent relevance, it is unclear why the scores would continue to decrease from 1960 to 2000; either the observations are not used properly, or the analysis quality keeps improving, thereby giving a more accurate reference. During the 2000s we observe a sharp increase in the scores, which is very likely due to the increase in the number of observations from buoys.

The figure also shows scores from the ERA-20C ensemble. These appear to be superior, but the difference between the two is small. One may question the quality of using own analysis as verification of each reanalysis. Using then ERA-Interim analyses as a reference, different scores are shown for a 3-month time period in Figure 8. They confirm that ERA-20C ensemble represents better the synoptic weather than ERA-20C deterministic. This may be due to the use of background error correlations in the ensemble that were tuned to the observing system coverage, whereas the deterministic production employed background error correlations that were relevant for a 2013-type of observing system.

## 4.2 background and analysis departures, and bias correction

Figure 9 shows that ERA-20C deterministic has overall a slightly smaller RMS of background departures after bias correction than ERA-20C ensemble. This comes partly from smaller mean departures. The origin of this improvement can be found by comparing Figures 10 and 11: the bias corrections are larger in the deterministic run. In the ensemble (Figure 10), the bias correction is much smaller, and leaves a residual non-zero background departure, which is slowly varying and much larger than in ERA-20C deterministic. In ERA-20C ensemble,

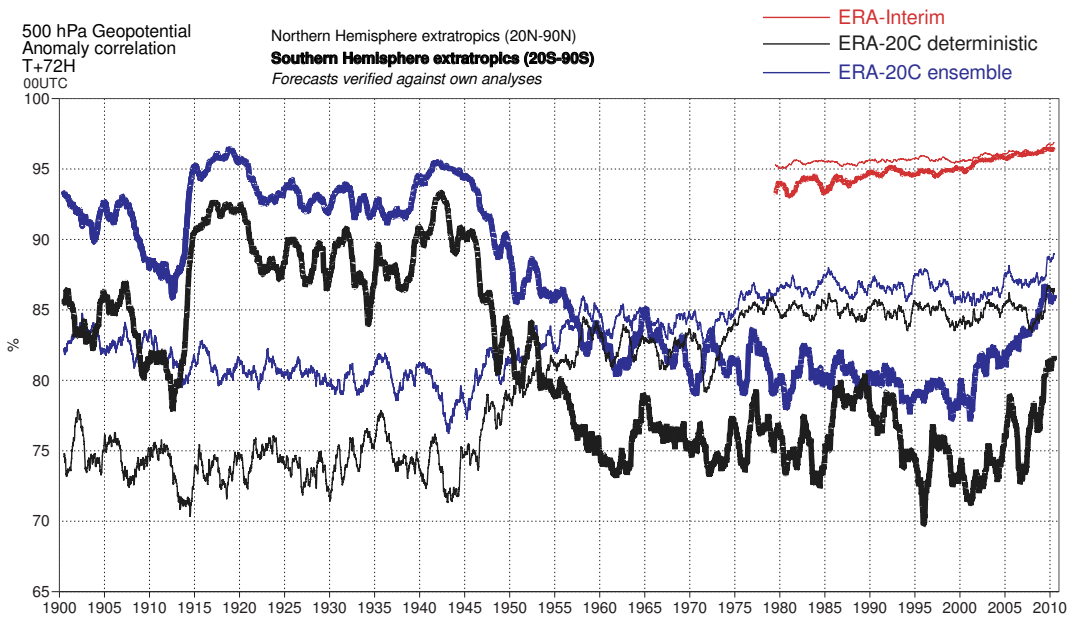


Figure 7: One-year running mean of anomaly correlations for 3-day forecasts of geopotential at 500 hPa issued at 00 UTC every day from 3 different reanalysis productions: ERA-Interim, ERA-20C ensemble member 0, and ERA-20C deterministic. Forecasts are verified against own analyses.

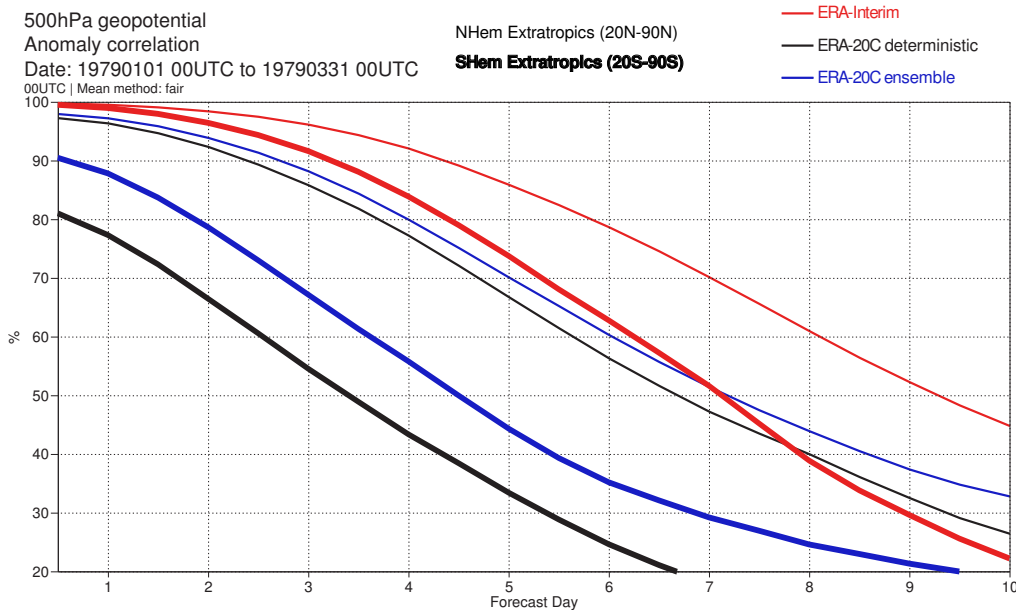


Figure 8: Anomaly correlation up to 10 days for forecasts of geopotential at 500 hPa issued at 00 UTC every day from 3 different reanalysis productions (ERA-Interim, ERA-20C ensemble member 0, and ERA-20C deterministic). Forecast is verified against ERA-Interim analysis.

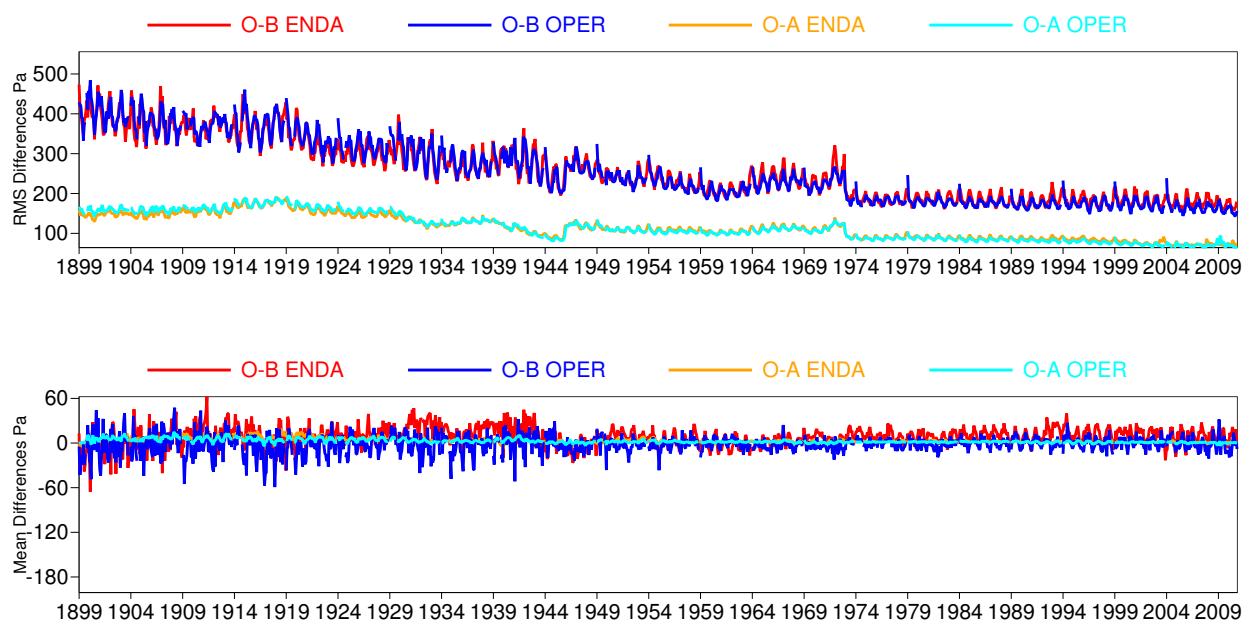


Figure 9: Observation minus background (O-B) and observation minus analysis (O-A) RMS (top) and mean (bottom) for observations of mean-sea-level pressure assimilated in ERA-20C ensemble member 0 (ENDA) and ERA-20C deterministic (OPER). All statistics are after bias correction.

smaller corrections *via* the observation bias mean that the assimilation has to fit the observations with larger mean analysis increments. The larger bias corrections seen in Figure 11 reflect the choices made in the design of the deterministic production.

### 4.3 Analysis increments

Figure 12 shows mean global increments for temperature as a function of model level (from the surface at the bottom to 0.01 hPa at the top), for (a) ERA-20C ensemble and (c) ERA-20C deterministic. The problems of increments ‘leaking’ in the stratosphere in the ensemble are not found in the deterministic production. The mean increments in the deterministic production are mostly negative near the surface (around -0.2 K on average) for model levels 91–81 (surface to about 925 hPa) and slightly positive (but below 0.1 K) in the rest of the troposphere (model levels 81–51 or pressure levels 925–225 hPa). The RMS presented in (b) and (d) show that the deterministic production confines more the increments in the vertical, restricting them mostly to the lower half of the troposphere (the increments are usually below 0.4 K RMS for levels 91–61 or surface to 420 hPa). The information found in the ERA-20C fields above that level comes mostly from the forecast model and the forcings, with little direct influence from the data assimilation, except *via* the modified model trajectory through updated atmospheric dynamics and radiation that reflect the conditions analysed at the surface. The level of maximum impact of the assimilation on temperatures seems to be around the 850 hPa pressure level (Figure 12d).

After the year 1979, the mean analysis increments can be compared with those of ERA-Interim. Figure 13 shows that, except near the surface, the mean increments of ERA-20C deterministic are generally much smaller than ERA-Interim. One may interpret this result as ERA-20C containing fewer jumps and spurious trends than ERA-Interim. However, one must remember that this plot first reflects the absence of the assimilation of upper-air observations in ERA-20C. The figure also shows a strong seasonality in the negative temperature increments near the surface (this is more difficult to see on the century-long time axis of Figure 12).

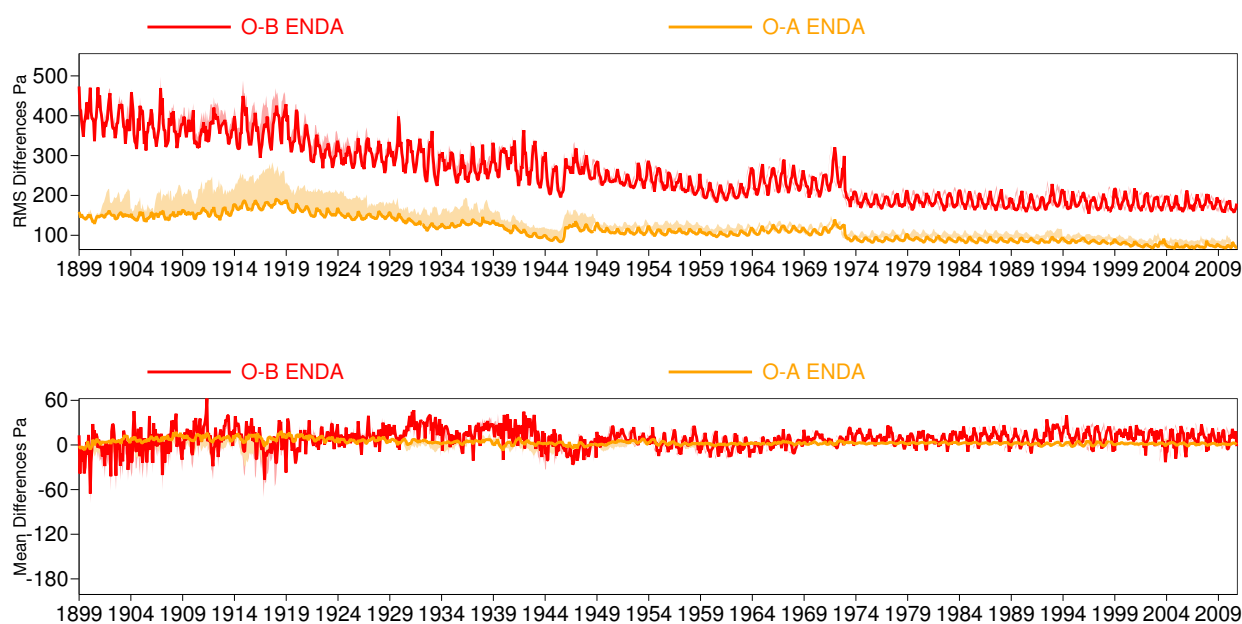


Figure 10: Observation minus background (O-B) and observation minus analysis (O-A) RMS (top) and mean (bottom) for observations of mean-sea-level pressure assimilated in ERA-20C ensemble member 0. Shading shows area between the statistics before bias correction and after bias correction.

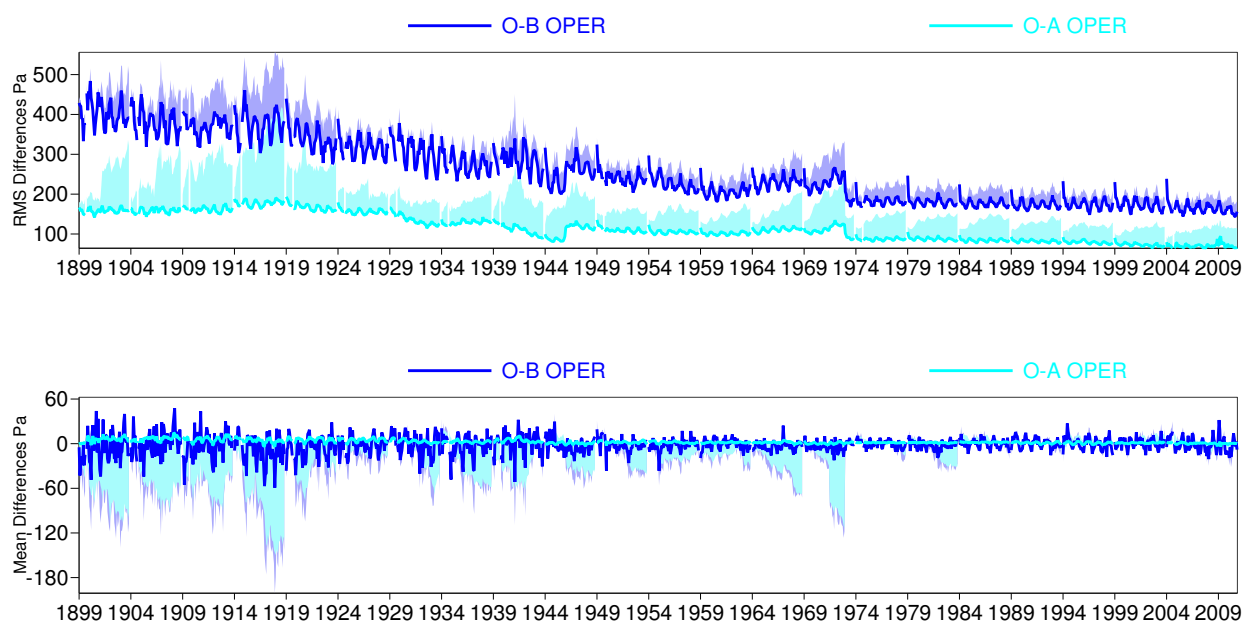


Figure 11: Same as Figure 10, but for ERA-20C deterministic.

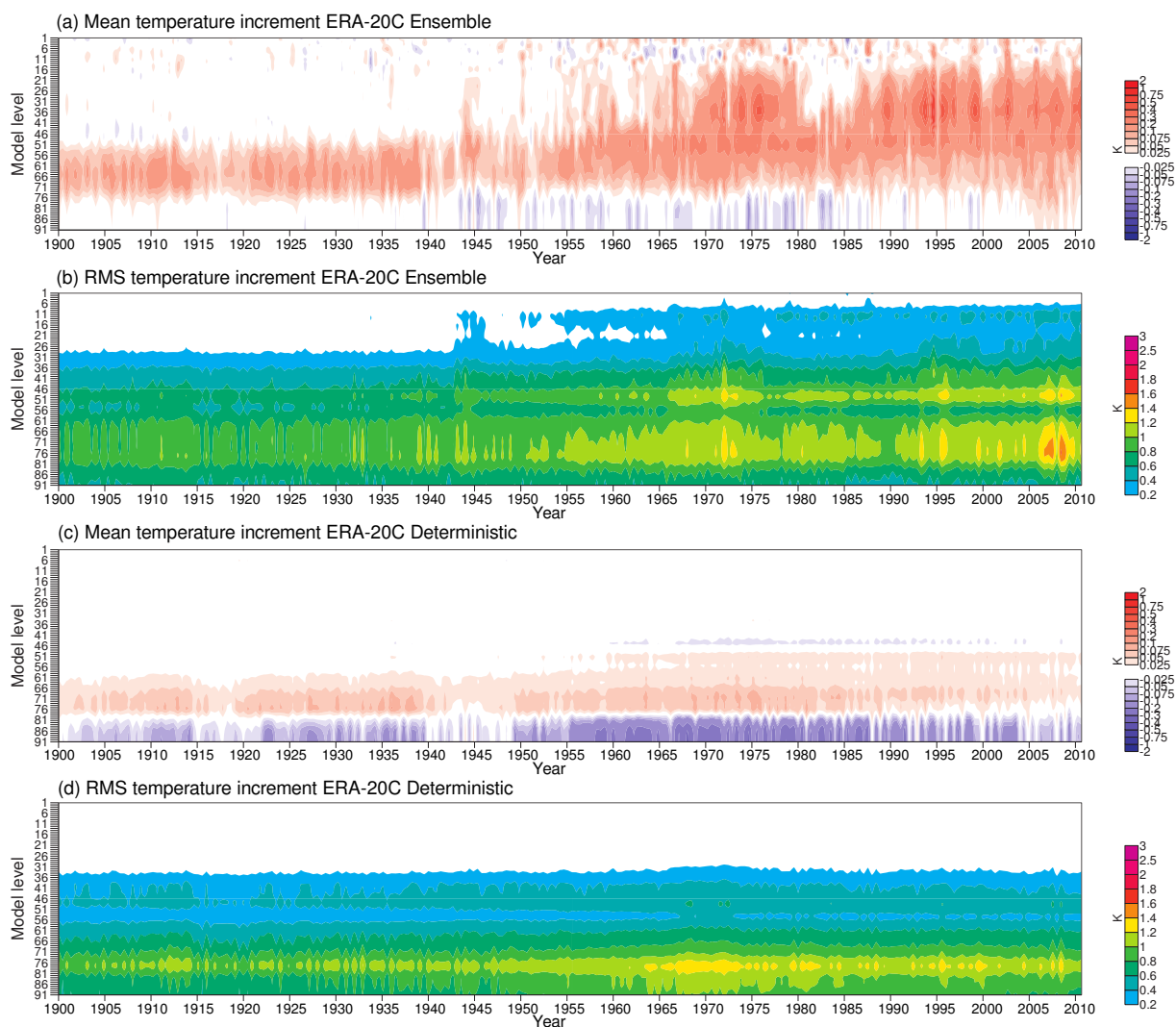


Figure 12: Mean (a,c) and RMS (b,d) analysis increments at the beginning of the analysis window for temperature, throughout the vertical and averaged monthly, for ERA-20C ensemble member 0 and ERA-20C deterministic.

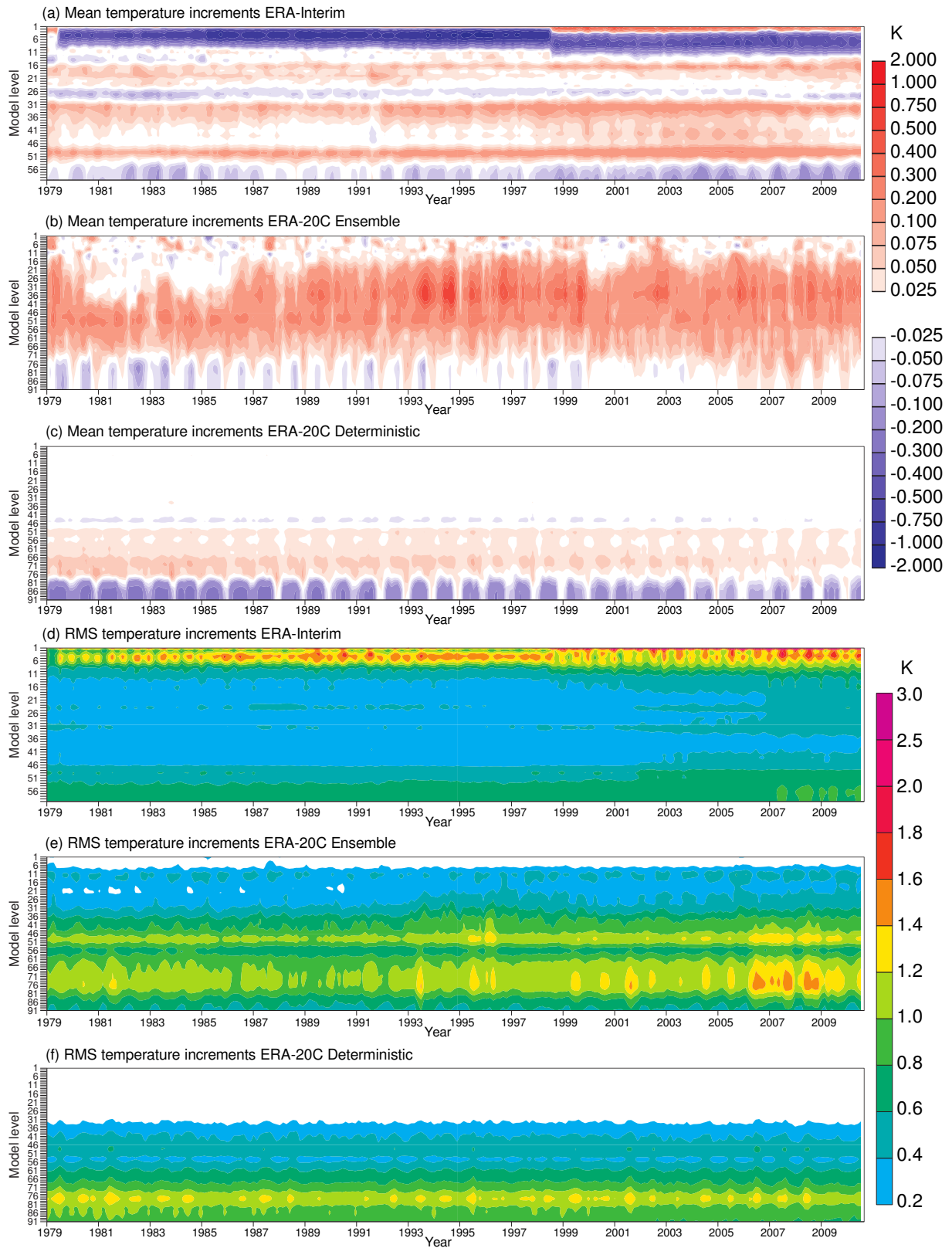


Figure 13: For the years 1979-2010, mean temperature analysis increments at the beginning of the analysis window, throughout the vertical and averaged monthly, for (a) ERA-Interim and (b,c) the two ERA-20C productions. (d,e,f) show RMS.

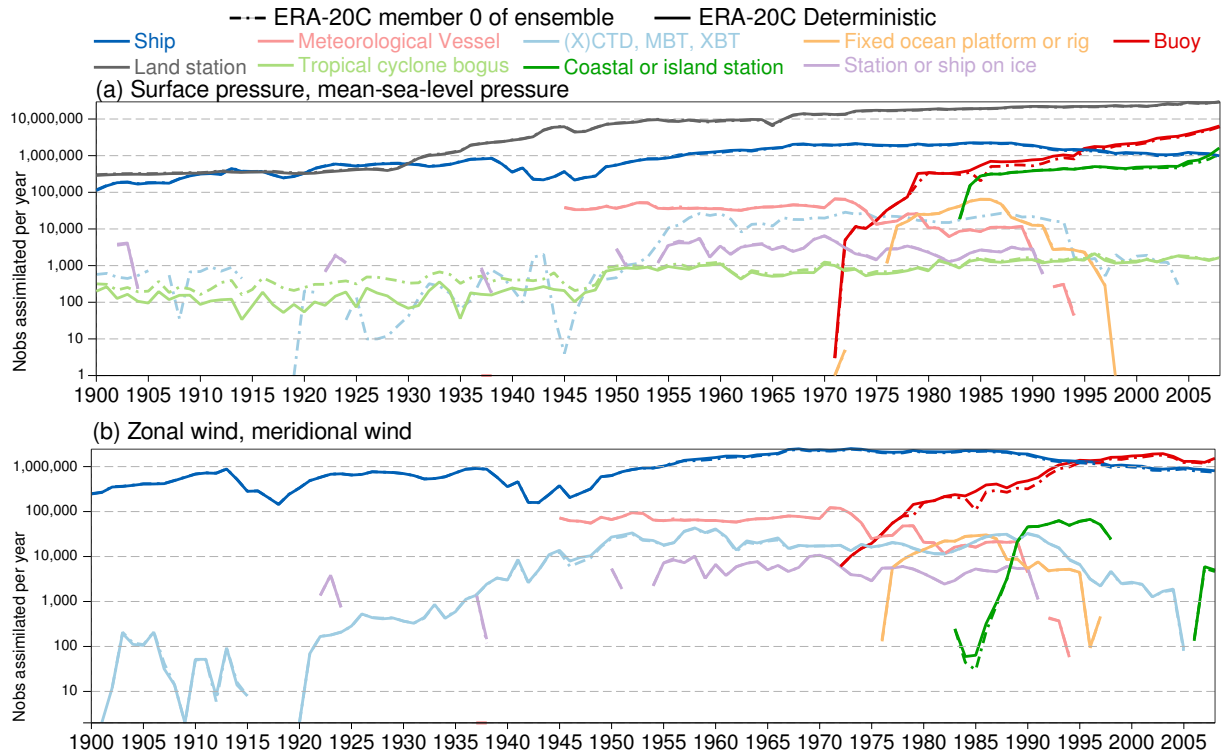


Figure 14: Number of observations assimilated per year in ERA-20C ensemble member 0 (chain lines) and ERA-20C deterministic (solid lines), for (a) surface and mean-sea-level pressure observations and (b) wind components.

#### 4.4 Observations assimilated

Figure 14 shows the number of observations assimilated in the ensemble and deterministic productions. The numbers are nearly identical except for the following: surface pressure from (X)CTD, MBT, and XBT reports are no longer assimilated in ERA-20C deterministic (result of an explicit decision, explained in section 2.2), and the number of tropical cyclone bogus data assimilated in ERA-20C deterministic is severely reduced as compared to ERA-20C ensemble before 1950.

The effect of the larger rate of rejections of tropical cyclone (TC) observations of pressure in ERA-20C deterministic as compared to ERA-20C ensemble is illustrated in Figure 15. The fit to observations after the analysis is degraded in ERA-20C deterministic as compared to ERA-20C ensemble. For comparison, the figure also includes results from NOAA/CIRES 20CR. Note that the observation inputs for TC from IBTrACS (Knapp *et al.*, 2010) differ between 20CR (which used ISPD v2.2) and ERA-20C (which used ISPD v3.2.6). Searching for the causes of these rejections is required before conducting again a similar reanalysis. In the first half of the reanalysis, for years 1900–1949 (respectively, years 1950–2010), ERA-20C deterministic assimilates 29% (53%) of these observations, rejects 24% (33%) because of background departures above the maximum admissible threshold, rejects 12% (11%) by variational quality control, rejects 34% (2%) because the observations belong to a constant timeseries, and rejects under 1% (*idem*) for other reasons. The seemingly useful test of checking for constant observations from sensors detailed in section 2.2 has the unforeseen effect of discarding the bogus data derived from TC tracks. The reason for the perfect stationarity of these data is found in the greater uncertainty associated with early cyclones reported in IBTrACS, and for which the lowest pressure is only very often reported to the nearest hecto-Pascal (such reduced precision makes it more likely to find constant timeseries). Another likely explanation is that the lowest pressure for early cyclones may have been estimated only at a few days’ interval, and left unchanged whereas the tropical cyclone position was updated at

several hours' intervals. By allowing for larger observation minus background departures and not including any check for constant timeseries, ERA-20C ensemble assimilates more tropical cyclone data. The lesson learnt from this for a similar future reanalysis is to not apply the constant timeseries check to bogus observations, and also possibly to inflate the assumed observation error estimate for these observations.

## 4.5 Observation errors

Table 1 indicates that one of the lessons learnt from ERA-20C is that observation errors varied over time, with various observing systems seemingly improving (see also presentation by Poli, 2013). Initial intentions for ERA-20C deterministic were to use observation errors adapted over time, learning from the statistics gathered in the ERA-20C ensemble. **Trial experiments** were hence set-up to adapt the observation errors as follows:

**Surface pressures from land stations** (IFS obstype 1, codetype 11) modified from about 1.08 hPa previously to decrease linearly with time from 1.6 hPa in 1900 to 0.8 hPa in 2009

**Surface pressures from ship** (IFS obstype 1, codetype 21) modified from about 1.46 hPa previously to decrease linearly with time from 2.0 hPa in 1900 to 1.2 hPa in 2009

**Surface pressures from tropical cyclone bogus** (IFS obstype 1, codetype 23) modified from 1.56 hPa previously to decrease linearly with time from 4.5 hPa in 1900 to 4.0 hPa in 2009

**Surface pressures from buoys** (IFS obstype 4) modified from 0.94 hPa previously to decrease linearly with time from 1.0 hPa in 1973 to 0.8 hPa in 2009

**Ten-metre wind component from ship** (IFS obstype 1) modified from  $1.5 \text{ ms}^{-1}$  previously to  $2.2 \text{ ms}^{-1}$  throughout the whole century

**Ten-metre wind component from buoys** (IFS obstype 4) modified from about  $1.33 \text{ ms}^{-1}$  previously to decrease linearly with time from  $1.7 \text{ ms}^{-1}$  in 1973 to  $1.4 \text{ ms}^{-1}$  in 2009

As expected, this experimental setup, using generally larger observation errors, leads to a looser fit to observations. Subsequently, the distance of the background to observations is also increased in the following assimilation window. This would require the background errors to be re-assessed. Figure 16 illustrates the effect of using inflated observation errors without modifying the background errors, for a short timeseries of surface pressure observations and assimilation feedback from Montreal, Québec, analysed with two ensemble systems. The first system or control is the ERA-20C ensemble, and the second one differs only by the assumed observation errors and the model time-step (30 minutes instead of 60 minutes, though this has little bearing on the results shown here). Both systems use the same background error global variances and correlations. The figure indicates that, in the presence of larger observation errors, the spread between the analysis members increases, and so does the spread between the subsequent short-term forecasts or backgrounds. In such case, one should re-compute the background errors to capture this increased ensemble spread. Such approach would allow to keep a proper balance in the assimilation error budget between observation errors and background errors. Not recomputing the background errors when increasing the assumed observation errors leads to under-weighting of the observations, and suboptimal use in the assimilation. Consequently, in ERA-20C deterministic, without re-running an ensemble that would allow to estimate updated background error variances, the solution adopted is to keep the observation errors identical to those of ERA-20C ensemble.

After ERA-20C deterministic production, the diagnostics from Desroziers *et al.* (2005) are computed from the observation feedback. The results are shown in Figure 17. They are very similar to the results of the ERA-20C ensemble, confirming that updated observation error estimates should be used in future ensemble reanalyses of the 20th century.



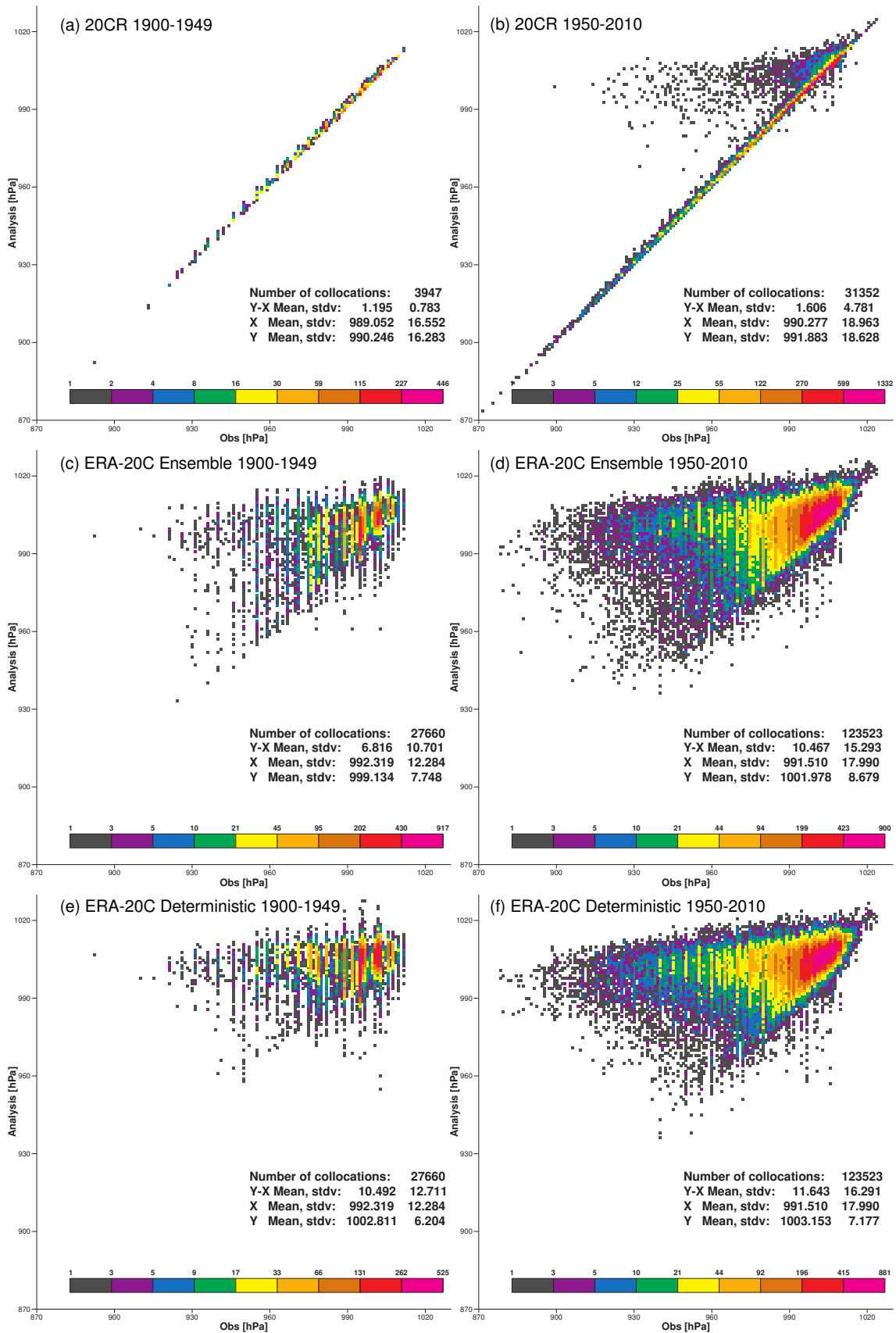
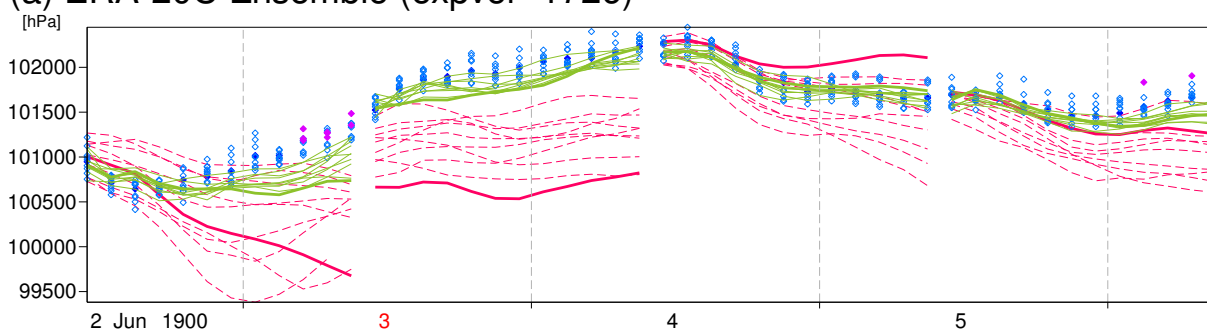


Figure 15: Density plot of observations (horizontal axis) versus analysis (vertical axis) for tropical cyclone bogus in (a,b) NOAA/CIRES 20CR, (c,d) ERA-20C ensemble member 0, and (e,f) ERA-20C deterministic, for years 1900–1949 (left column) and 1950–2010 (right column).

(a) ERA-20C Ensemble (expver=1726)



(b) Inflated observation errors (expver=1814)

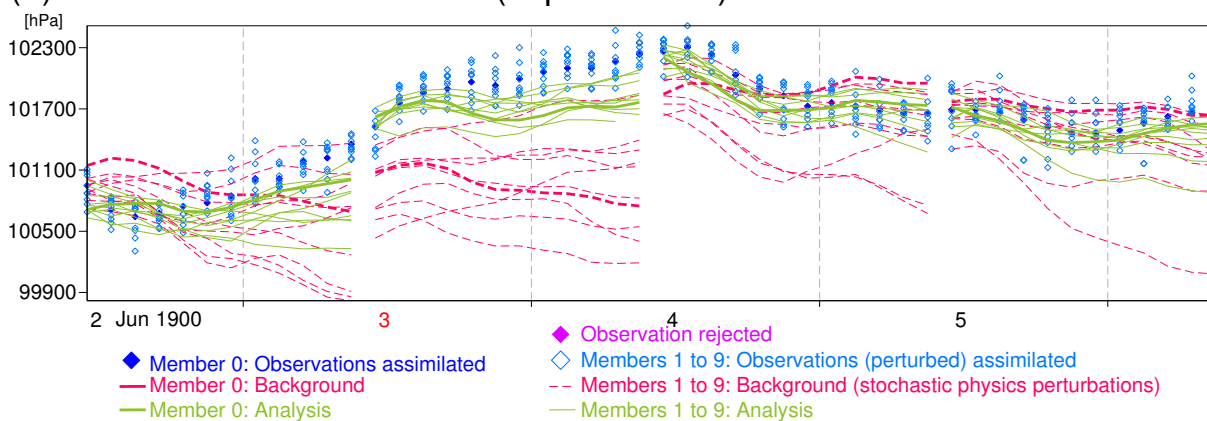


Figure 16: Timeseries of observations, backgrounds, and analyses of surface pressure at Montréal, Québec, for (a) the 10 members of the ERA-20C ensemble and (b) an ensemble experiment which uses the same background errors but larger observation errors.

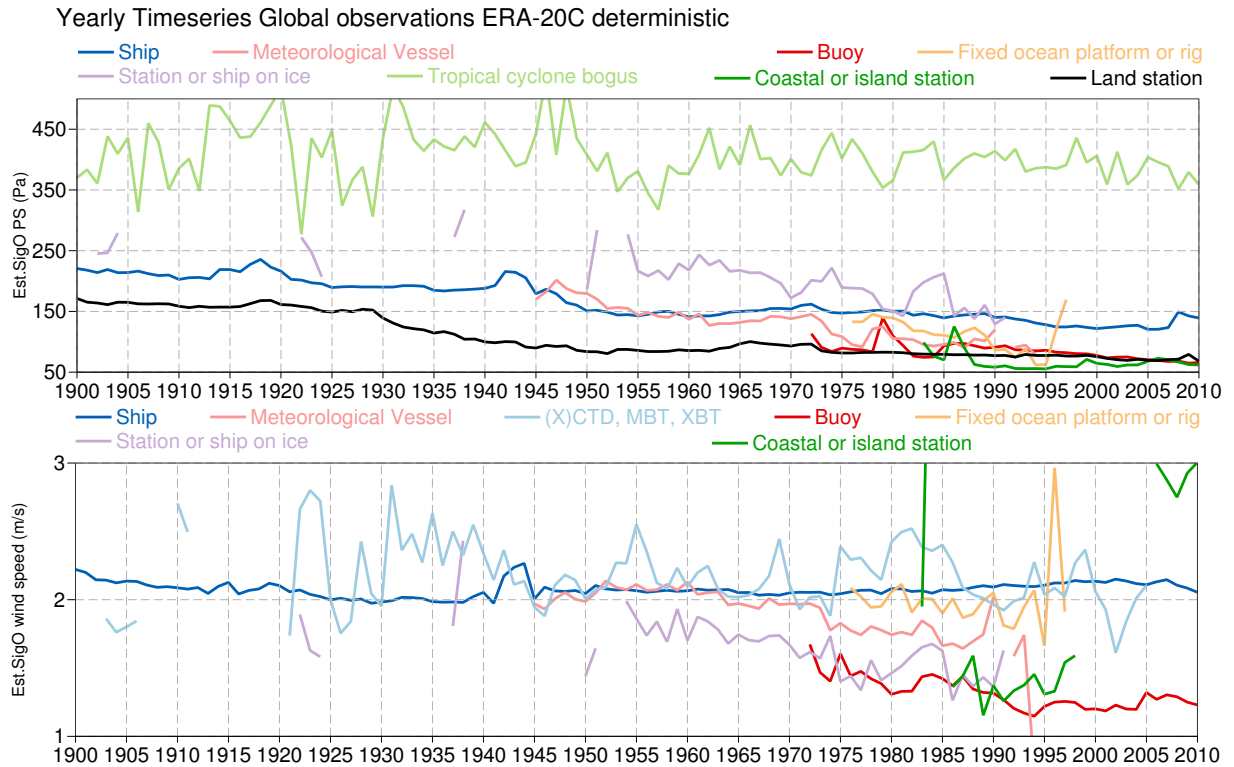


Figure 17: Observation error standard deviations estimated using Desroziers’ diagnostics applied to ERA-20C deterministic innovation and residual statistics.

## 5 Climate value

The following section assesses low-frequency variability for several parameters in ERA-20C deterministic, which is hereafter simply referred to as ERA-20C (the present section does not show results from ERA-20C ensemble, unless explicitly indicated).

### 5.1 Two-metre temperature

The low-frequency variability of two-metre temperature is assessed here by comparison with a model integration. Hersbach *et al.* (2013) describe such product, which was later repeated with updated Sea-Surface Temperature (SST) forcing (Hersbach *et al.*, 2015a), identical to that used in ERA-20C. As expected, because ERA-20C and ERA-20CM both use the same SST forcing, they feature very similar climate trends over oceans (Figure 18j, showing two-metre temperature anomalies with respect to years 1961–1990). Over land (Figure 18c), the differences are most pronounced before the 1940s. There are also significant differences for the recent, well-observed times. For example, Figure 18e shows that over South America over land ERA-20C finds larger maxima than ERA-20CM for the strong El Niño events of 1982/3, 1997/8.

### 5.2 Total column water vapor

Regarding the water cycle, Figure 19a shows timeseries of total column water vapor over tropical oceans (defined here as latitudes within 20° of the Equator). Over the whole century, increasing water vapor amounts are

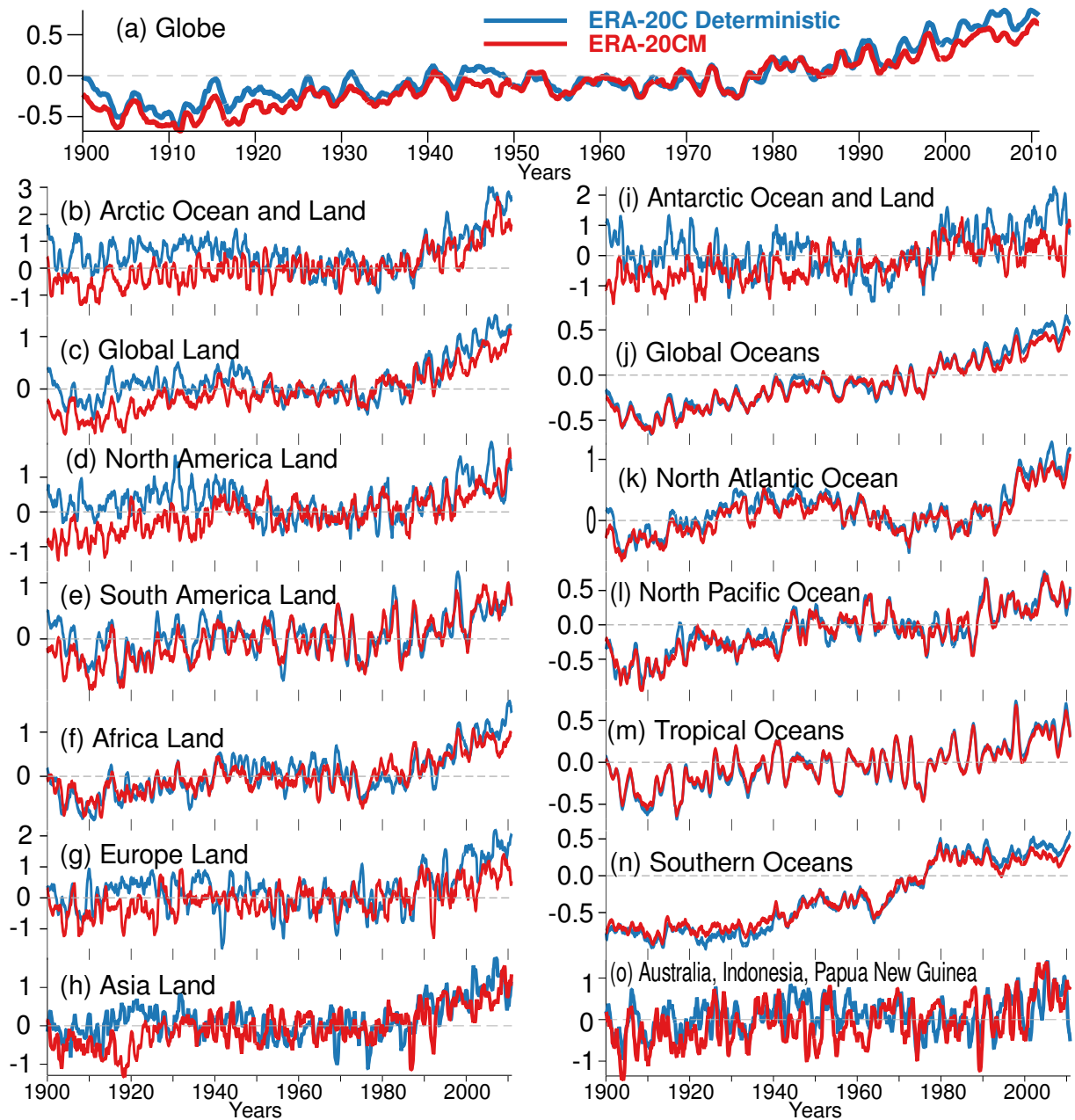


Figure 18: Monthly anomalies for two-metre temperature from ERA-20C analyses (blue) and ERA-20CM (red) for (a) the Globe, and (b–o) various regional averages. Reference years for the climatologies are 1961–1990. 12-month moving average is applied.

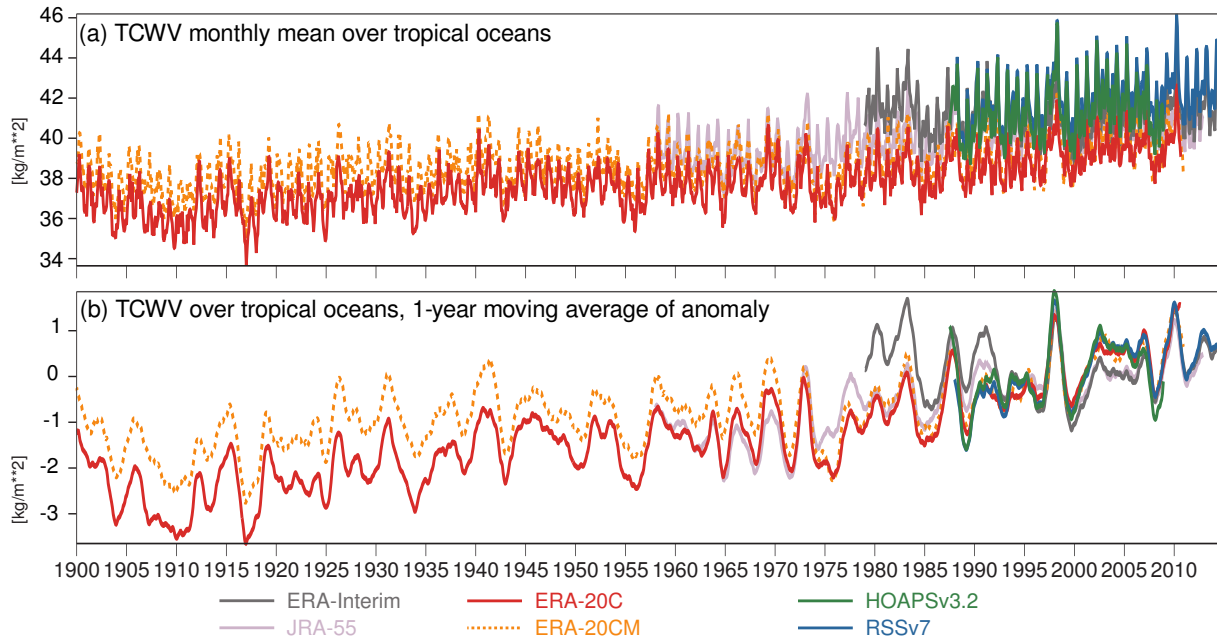
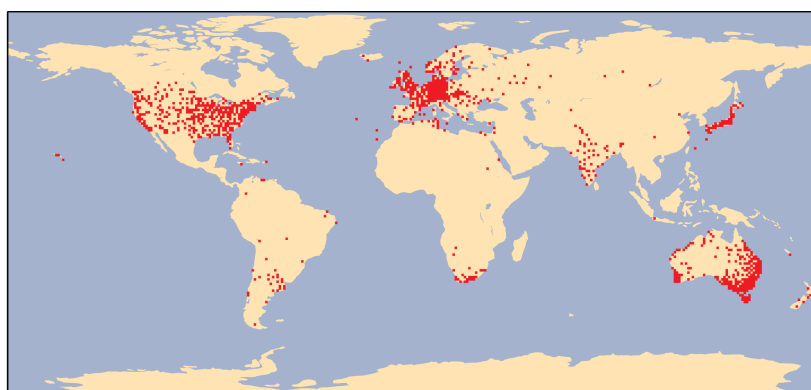


Figure 19: (a) Monthly total column water vapor over oceans in latitude band 20S–20N according to various reanalysis and observational products, between January 1900 and June 2014; (b) 12-month moving average of monthly anomalies, obtained by removing each product's monthly climatology over the time period common to all products (years 1988–2008).

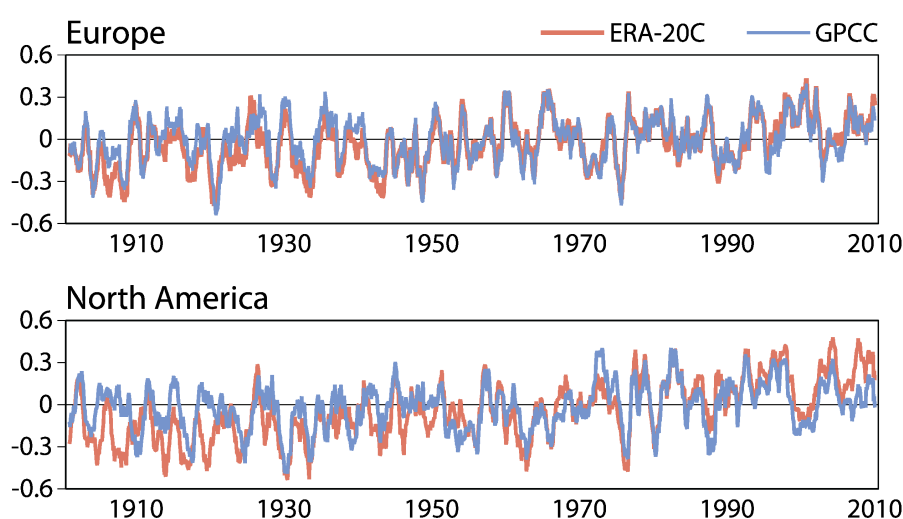
visible in ERA-20C and ERA-20CM (Hersbach *et al.*, 2015a), though trends appear larger for ERA-20C. Over recent years, ERA-20CM and ERA-20C converge, so the difference over the early part of the century comes from ERA-20C being noticeably drier than ERA-20CM then. Since the version of the IFS model employed in ERA-20C(M) features a dry bias – ERA-20CM being about 2 kg/m<sup>2</sup> smaller than observational datasets such as HOAPS and RSS – and since the trends in two-metre temperature over the same area are about similar in ERA-20C and ERA-20CM (Figure 18m), there may be a component of the assimilation in the early years that is affecting the water cycle. An obvious one to investigate would be the assimilation of near-surface wind observations, which could affect the trade winds representation and hence the width of the intertropical convergence zone. This could impact the evaporation, the extent of the region of large-scale vertical ascent, and the convective rainfall. However, Figure 19b suggests that ERA-20C agrees better with JRA-55 than with ERA-20CM in terms of anomalies between 1960 and 1975. For that time period, in the absence of an independent observational product to compare with, it is unclear which one of the three products, between ERA-20C, ERA-20CM, and JRA-55, is the most accurate in terms of inter-annual variability. For years 1988–2008, Figure 19b suggests that ERA-20C and ERA-20CM are closer than ERA-Interim and JRA-55 to the observational products HOAPS and RSS.

### 5.3 Precipitation

Related to the water cycle representation, Figure 20 shows a comparison of ERA-20C with processed rain gauge measurements from the Global Precipitation Climatology Centre (Becker *et al.*, 2013). Most regions present insufficient amounts of rain gauge data for validation over the whole period of ERA-20C, so time-series are shown only for Europe and North America. Over Europe, ERA-20C seems to present a fair ability to represent the inter-annual fluctuations for the whole century, with a noticeable improvement from 1945 onwards. Over North America, a similar improvement comes in later, around 1960; in the early years (until about 1925) ERA-



1°x1° grid squares in which GPCC has at least one station for each month from 1901 to 2010



12-month running mean precipitation averaged over European and North American subsets of these grid squares, expressed in mm/day as anomalies relative to 1961-1990

Figure 20: (top) Map of rain gauges used for comparison with ERA-20C, and (middle and bottom) time-series over Europe and North America for the rain gauges in these regions.

20C presents in fact little or no realism in terms of precipitation anomalies in that region.

## 5.4 Ozone

In ERA-20C, ozone is a time-varying three-dimensional prognostic variable. The ozone mass mixing ratio is determined by a prognostic equation where sources and sinks are parametrised as relaxing towards a photochemical equilibrium depending on the time-varying equivalent chlorine content of the stratosphere (ECMWF, 2013). For radiative transfer calculations, a time-varying ozone monthly climatology reconstruction is used as forcing input, as done in the ERA-20CM model integration (Hersbach *et al.*, 2015a). The resulting radiation budget affects the thermal vertical structure of the atmosphere, which in turn affects ozone concentration *via* the model dynamics and the ozone parametrisation. Also, the model dynamics affects ozone content through advection. Using as reference the Multi-Sensor Reanalysis-2 (van der A *et al.*, 2015), Figure 21 shows the evolution of total column ozone in January over the Northern Hemisphere for ERA-Interim, ERA-20CM, and ERA-20C between 1979 and 2008. MSR-2 and ERA-Interim feature very similar spatial patterns, which is to be expected because (1) both used similar satellite instrument observations as input, and (2) MSR-2 uses transport from ERA-Interim dynamical fields.

We note nonetheless that ERA-Interim features suspicious low ozone amounts in January 1995 and 2003 as

compared to MSR-2. Inspecting closely the January maps for MSR-2, ERA-20CM, and ERA-20C, especially for the areas and years of anomalous ozone concentrations in MSR-2, ERA-20C appears usually to be a better fit than ERA-20CM to MSR-2 (*e.g.*, 1982, 2006), with the exception of a handful of odd years (*e.g.*, 1991, 1997). Time-series of equal-area monthly averages for the regions poleward of 60° latitude are shown in Figure 22, for January 1979 until December 2010. Spikes are visible for ERA-Interim in end 1995 and end 2002, coinciding with (respectively) the beginning and the end of Global Ozone Monitoring Experiment (GOME) ozone profile assimilation in ERA-Interim. ERA-20C reproduces better than ERA-20CM the inter-annual variability of MSR-2 in the northern high latitudes for the whole period. A similar statement applies in the southern high latitudes but only after the 2000s in the Southern Hemisphere. Assuming MSR-2 as a trusted reference, this indicates that when a reanalysis like ERA-20C produces realistic weather maps, it also produces realistic stratospheric ozone, or at least more realistically than a model integration which does not necessarily have the right timing of the weather patterns. When the quality of the synoptic weather maps (and hence location of the polar vortex) is not so good, as shown earlier with time-series of scores before the 1990s, the ozone representation also suffers. These results would hence tend to also suggest the existence of a tight connection between total ozone concentration and synoptic weather, as observed from the surface by pressure observations. This result is not inconsistent with recent findings from Brönnimann and Compo (2012) who looked at dynamical links between total ozone, the flow near the tropopause, and tropospheric circulation.

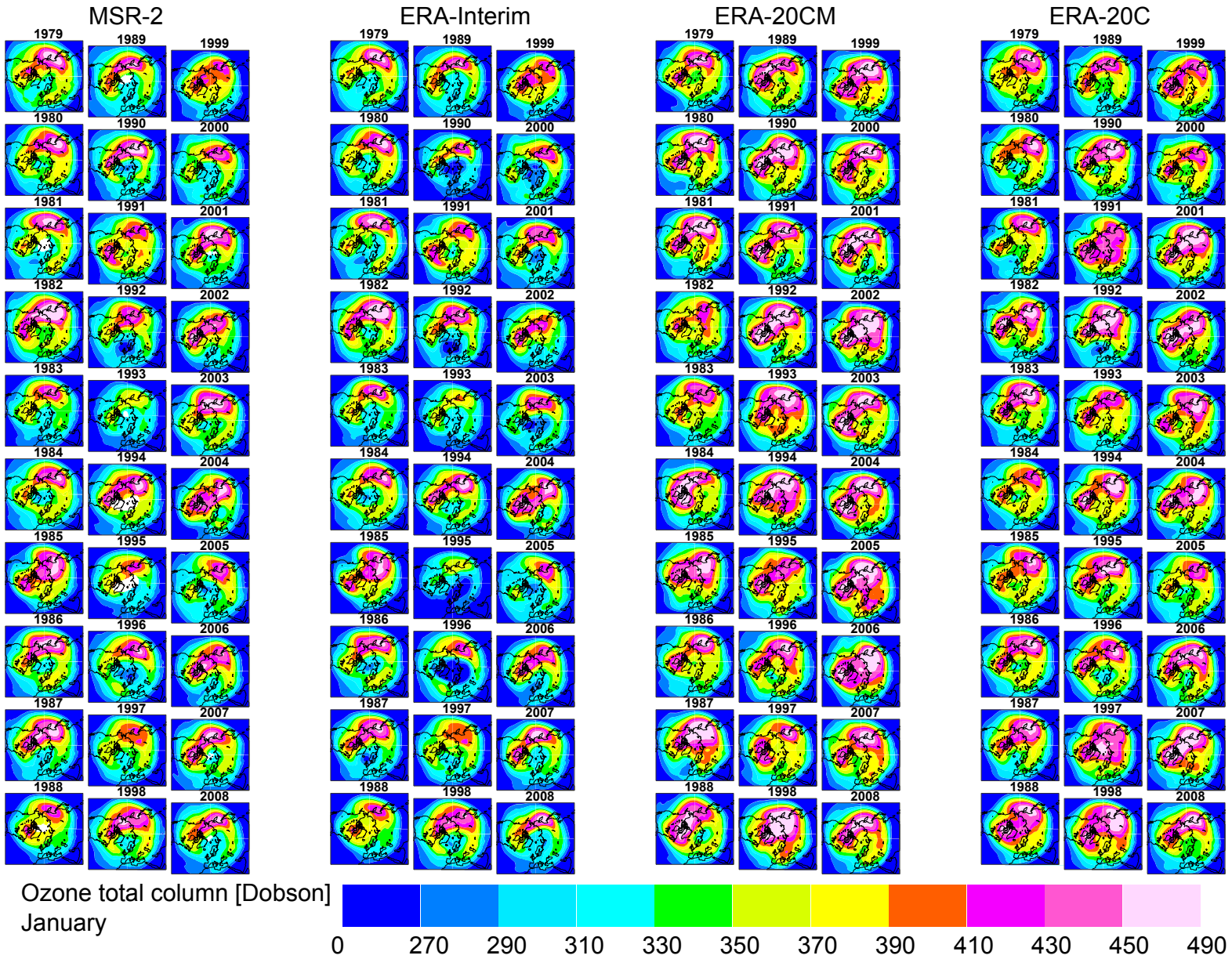
## 5.5 Large-scale and regional circulations

In order to assess the large-scale circulation, Figure 23 shows where the 30-year average of the zonal wind is close to zero for the months of January to March. Showing only these areas (and differing by the sign) is equivalent to highlighting where the surface winds change direction, and hence to delineate the edges of persistent zonal circulation cells. The shaded areas of near-zero zonal wind on this map are related to the two outer edges of the Hadley cells in both hemispheres, but also highlight the warm pool in the Indo-Pacific region. Regional circulations related to upwelling along the coasts of California and Chile are also visible. Other regional circulations such as the Tehuano wind near Panama also appear. The rationale for this map is to be able to use the same diagnostics in observation space, to find out whether the locations of zonal circulation cells in the reanalysis agrees with observations. The match between the analyses (color contours) and the observations (dots) suggest that for the season shown here the edges of the zonal circulation cells agree generally well for the three 30-year periods shown here. However, we also find that for the first 30-year period, the Northern edge of the Hadley cell in the Pacific is probably located too far South in ERA-20C as compared to the observations. For the years 1940-1969, the Eastern side of the Northern edge of the Hadley cell in the Atlantic is also probably misplaced North of the area suggested by observations. Close inspection of these edges in observation space (not shown) suggests that these edges are quite stable but undergo multi-decadal oscillations, which are somewhat weaker in ERA-20C than in the surface marine observations. Whether this is related to structural functions in the forcing fields of SST (Kennedy *et al.*, 2015) that may be mostly governed by current-day satellite observations remains to be investigated. Looking at another season (July to September), Figure 24 seems to confirm the conclusion regarding the Hadley cell mis-location in the Atlantic in 1940-1969.

## 5.6 Energy budget

Figure 25 shows one-year running means for anomalies in globally averaged energy budgets relative to years 1900–1909 for ERA-20CM ensemble (top), ERA-20C ensemble (middle), and ERA-20C deterministic (bottom). As discussed by Hersbach *et al.* (2015a) for the century-long model integration, the thermal (plus heat) fluxes and short wave-energy fluxes are rather stable up to about 1970, both at the top of the atmosphere (TOA) and at the surface. After this period, the net downward flux at both the TOA and surface increases by about

Figure 21: Maps of total column ozone over the Northern Hemisphere for the month of January, years 1979 to 2008.





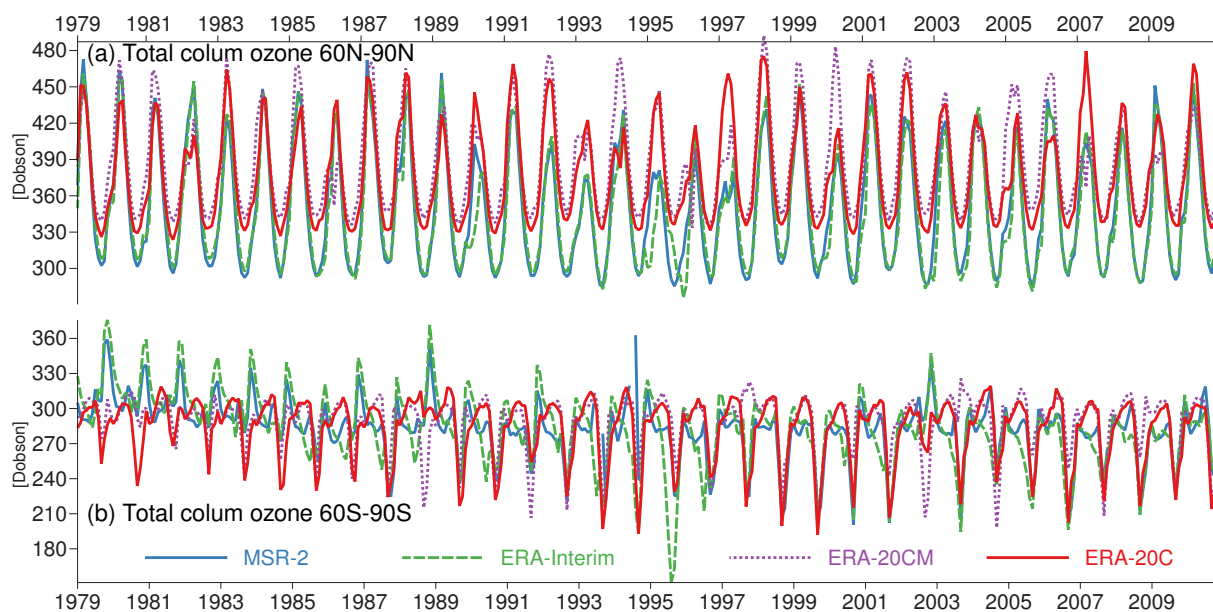


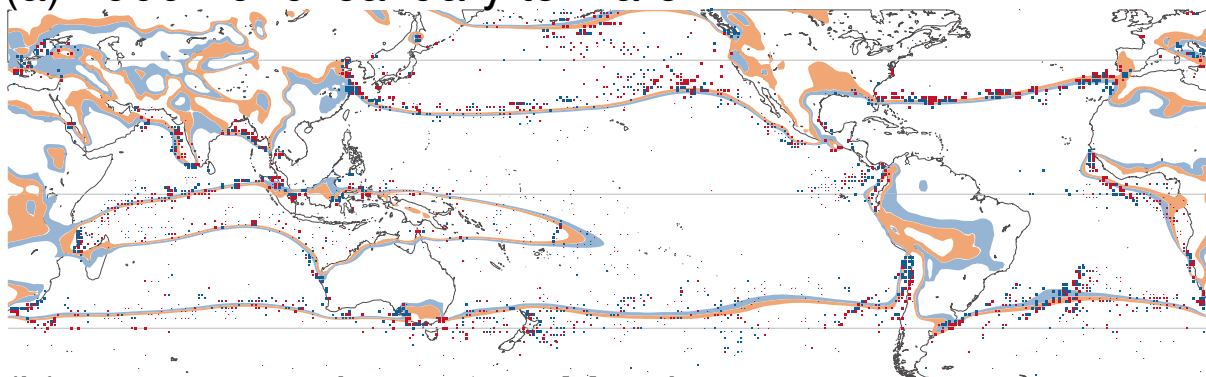
Figure 22: Time-series of total column ozone monthly averages between January 1979 and December 2010, over (a) latitudes 60N–90N and (b) latitudes 60S–90S.

$1 \text{ Wm}^{-2}$ , which reflects the warming global climate. Fluctuations in the net fluxes are influenced by El Niño–Southern Oscillations and volcanic eruptions, as enforced by the prescribed SST and radiative CMIP5 forcings.

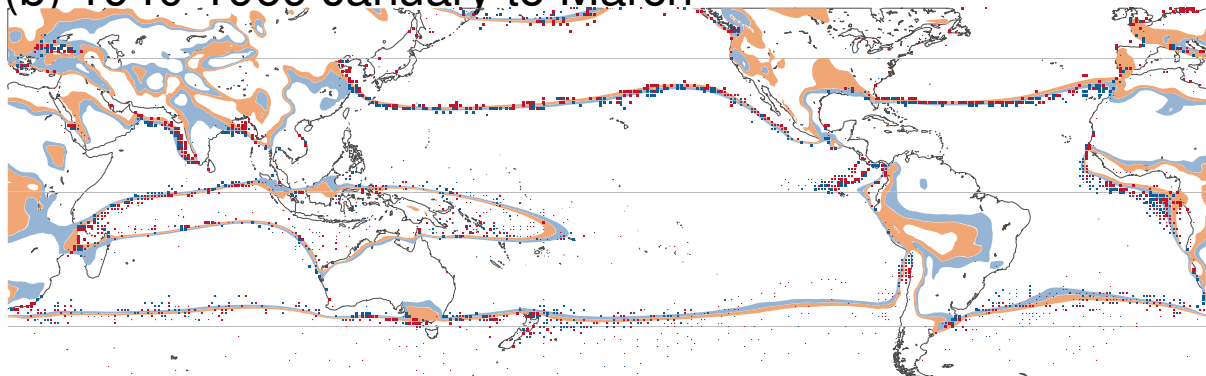
In contrast to ERA-20CM, the ERA-20C reanalyses are not plain model integrations, since they involve data assimilation. The daily increments as shown in Figure 12 are not unbiased and, therefore, will systematically change the energy content of the atmosphere. In order for the system to be stable over time a surplus/deficit of energy has to leave/enter the boundaries of the atmosphere accordingly (apart from a small net average influx of about  $0.01 \text{ Wm}^{-2}$  to account for global warming). This is what is probably observed in the middle and lower panels of Figure 25. In ERA-20C ensemble, the increasingly positive temperature increments over time seem to add energy to the atmosphere which, compared to ERA-20CM, is released at both the surface and the TOA. ERA-20C deterministic is more stable in time, though the larger negative temperature increments over the lower troposphere between the 1950s and 1980s seem to be compensated by a relative increase of the net influx of energy into the atmosphere during the model integrations between the daily analyses (note that in absolute terms the net surface and TOA fluxes still indicate a loss of energy from the atmosphere, not shown). Unfortunately, the magnitude of the effect of the biased increments is not small compared to the climate signal. Nevertheless, the energy budget appears much improved in ERA-20C deterministic as compared to ERA-20C ensemble.

The (vertically integrated) total energy of the atmosphere is available for ERA-20C deterministic analyses from the publicly available products. The formula used for computation is given by Berrisford *et al.* (2011a) and the total energy is the sum of the potential, kinetic, internal and latent energies. Figure 26a shows that in ERA-20C deterministic the energy in the atmosphere (integrated globally, weighted by area) increases over the course of the century, and most particularly in the recent decades. This simply reflects an uptake of energy by the atmosphere with climate change. This increase is more visible by considering the 12-month moving average of total energy anomaly (relative to years 1900–1909) in Figure 26c. Figure 26d shows that the 12-month moving average anomaly of the time derivative of total energy fluctuates around zero, but on average becomes very slightly positive in recent decades, consistent with the accelerated warming occurring at this time.

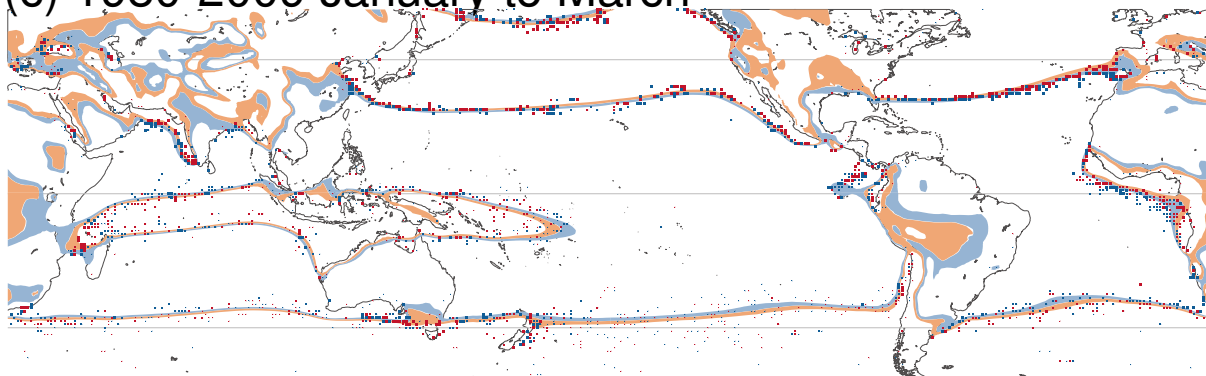
(a) 1900-1929 January to March





(b) 1940-1969 January to March



(c) 1980-2009 January to March



ERA-20C average zonal wind  
 between -0.5 and -0.01 m/s  
 between 0.01 and 0.5 m/s



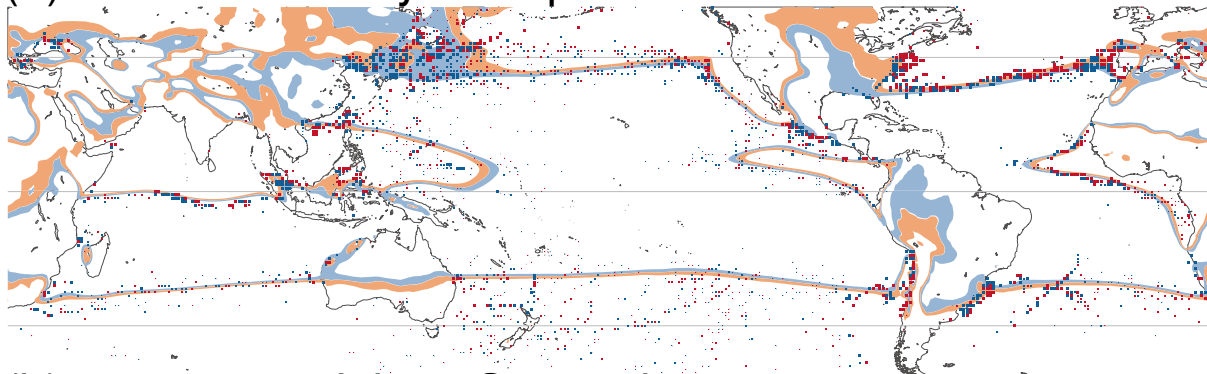
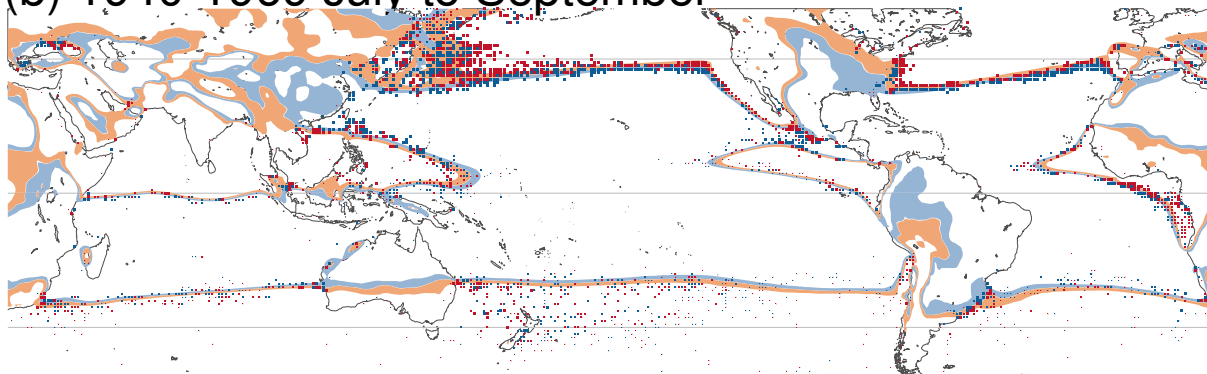
ICOADSv2.5.1 average zonal wind  
 between -0.5 and -0.01 m/s  
 between 0.01 and 0.5 m/s

Figure 23: 30-year averages of zonal wind in ERA-20C for January to March, with a choice of contours that only show zonal wind around  $0 \text{ m s}^{-1}$ . Dark-colored dots show the same estimates but from the wind observations in ICOADS, with a variable dot size that indicates relatively greater amounts of observations within a given map.

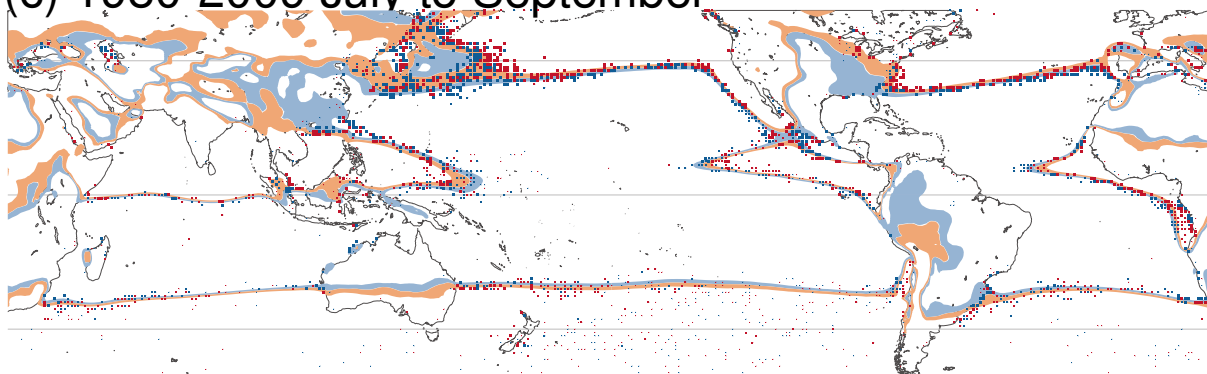
(a) 1900-1929 July to September



(b) 1940-1969 July to September



(c) 1980-2009 July to September



ERA-20C average zonal wind  
 between -0.5 and -0.01 m/s  
 between 0.01 and 0.5 m/s

ICOADSv2.5.1 average zonal wind  
 between -0.5 and -0.01 m/s  
 between 0.01 and 0.5 m/s

Figure 24: Same as Figure 23 but for July to September.

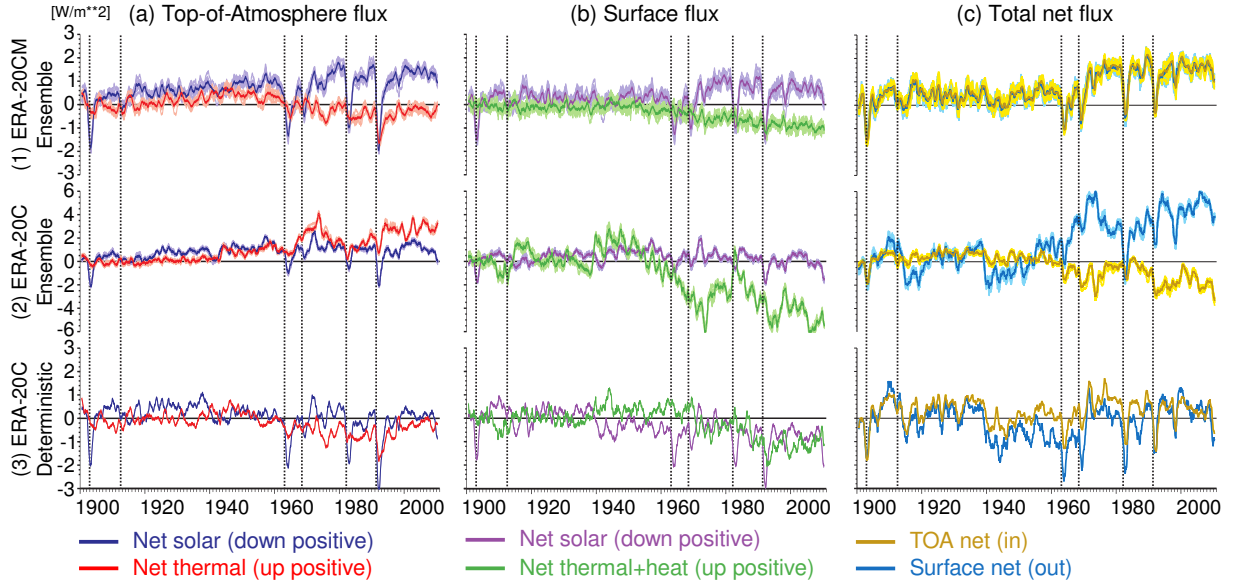


Figure 25: Evolution of the 12-month moving average of anomalies of energy budgets for (1) ERA-20CM ensemble, (2) ERA-20C ensemble and (3) ERA-20C deterministic, relative to years 1900–1909. Column (a) shows TOA net radiation fluxes for solar and thermal, column (b) shows net surface fluxes for solar radiation, and thermal radiation plus latent and sensible heat, and column (c) shows TOA net total radiation flux ( $R_T$  in gold) and surface net total flux ( $F_S$  in blue). In (1) and (2), both ensemble products, dark (light) colors represent the ensemble mean (individual members, respectively). Note the vertical scale for ERA-20C ensemble is twice as large as for the other two products. The vertical black dash lines indicate major volcanic eruptions (in chronological order): Santa Maria, Novarupta, Mount Agung, Fernandina Island, El Chichón, and Pinatubo.

The time derivative in global total energy,  $dE/dt$ , in (re)analyses has two sources:

$$\frac{dE}{dt} = F_E + A_E \quad (1)$$

where  $F_E$  is the influence of the model in total energy space, which represents the effects of various processes and forcings (e.g. SST, sea-ice, aerosols, ozone, and greenhouse gases) and  $A_E$  is the analysis increment in total energy space. Furthermore,

$$F_E = R_T - F_S + C_E \quad (2)$$

where  $R_T$  and  $F_S$  are the net downward fluxes at the TOA and surface respectively.  $C_E$  represents model errors due to lack of global total energy conservation.

The analysis increment  $A_E$  gives an absolute measure in total atmospheric energy space (in units of  $\text{Wm}^{-2}$ ) of the overall discrepancy between observations on the one hand and atmospheric model and forcings on the other. It thus yields a simple metric to trace progress in climate research and reanalysis. Unfortunately, at the time of writing, the analysis increments in total energy space are not available from the ERA-20C deterministic products. These increments require the additional computation of the total energy from the ERA-20C deterministic forecasts.

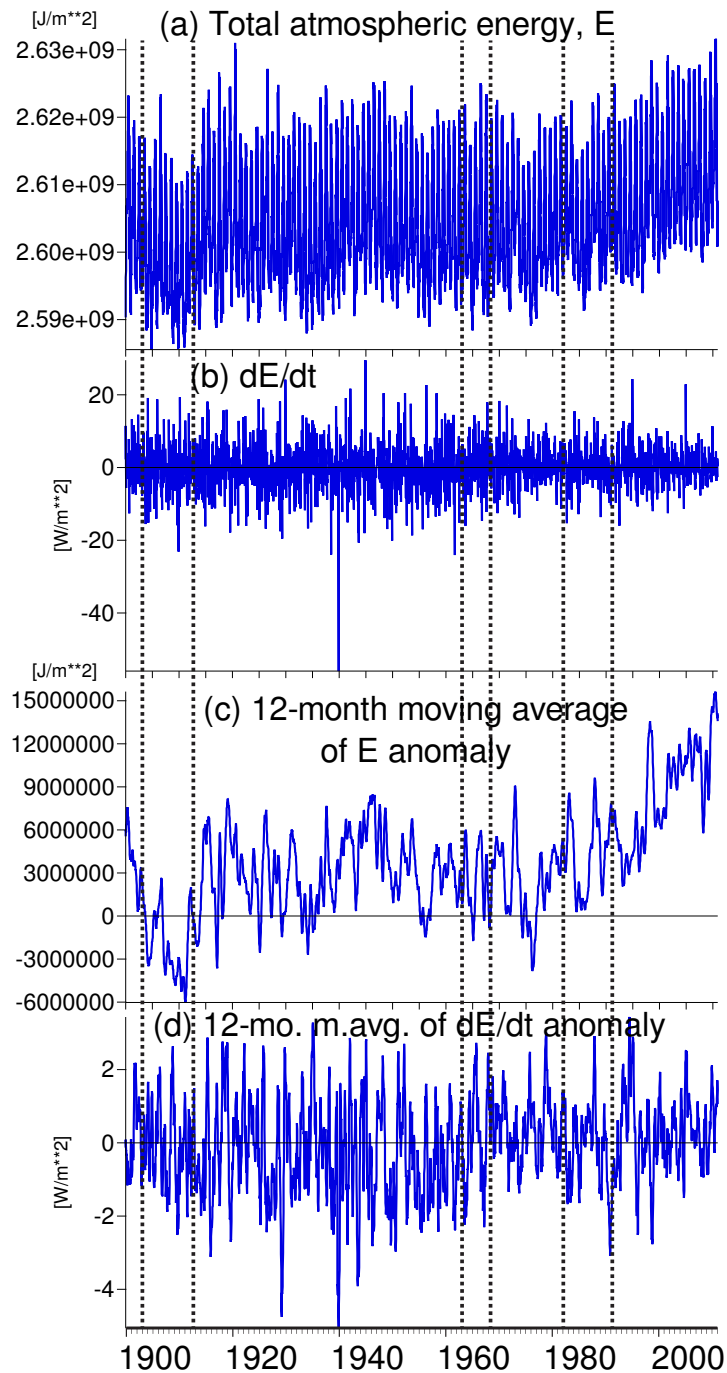


Figure 26: Evolution of the total energy in the atmosphere in ERA-20C deterministic in (a), and of the time derivative in (b). Anomalies in (c,d) are with respect to years 1900–1909. Vertical dash lines are major volcanic eruptions as described in Figure 25.

## 6 Uncertainties

### 6.1 Three-hourly ensemble spreads

Providing users with uncertainties was one of the underlying motivations for an ensemble approach in ERA-20C. In practice such approach was also required in order to estimate time-varying background errors for the data assimilation, owing to time variations in the observing system. The ERA-20C deterministic reanalysis makes use of the ensemble produced earlier in order to estimate such errors (see section 2.3). However, this reanalysis no longer provides an ensemble to users. This section is a preliminary exploration of the various types of ensemble spreads that may be computed from the ERA-20C ensemble and provided to users.

The time evolution of the temperature spread between the different members of the ERA-20C ensemble at forecast step 3 is shown for the troposphere, for five latitude bands, in Figure 27. Over the course of the century, the sharpest drop occurs in the mid-latitudes. In the Northern mid-latitudes, the ensemble spread reduces mostly after 1950, whereas in the Southern mid-latitudes this reduction mostly occurs after the 1980s. In both cases these reductions coincide with major observing system improvements (respectively: more ships and land stations, and advent of drifting buoy observations). Judging by the ensemble spread, the quality of the Southern high latitudes analyses is of similar quality throughout the century until the late 1990s. The Northern high latitudes feature a decreasing ensemble spread from the 1930s onward, as the Soviets set up an observing network covering the largest land mass in the Northern Hemisphere, largely unobserved on a regular basis until then. Except for the Southern high latitudes, the spread between the ensemble members is below 0.8 K in the lower troposphere at the end of the 20th century.

### 6.2 On the limits of these estimates as measures of uncertainties

These ensemble spreads are incomplete estimates of uncertainties. They represent the combined effect of the uncertainties that were explicitly specified: in the SST (HadISST 2.1.0.0), in the observations (Isaksen *et al.*, 2010), and in the atmospheric model *via* stochastic physics (Palmer *et al.*, 2009). All other inputs to the reanalysis systems are assumed perfect. The impact of this approximation is illustrated on maps of the various estimates of ensemble spreads, in Figure 28, for two-metre temperature in January 1900. The analysed state of two-metre temperature is noted  $\mathbf{x}_i^j$ , where index  $i$  denotes the time (1 for 00 UTC on the first day of the month, 2 for 06 UTC on the first day of the month, until 18 UTC on the last day of the month, so up to  $N=124$  for 31-day-long months) and index  $j$  is the ensemble member number (from 0 to 9). The following quantities are estimated:

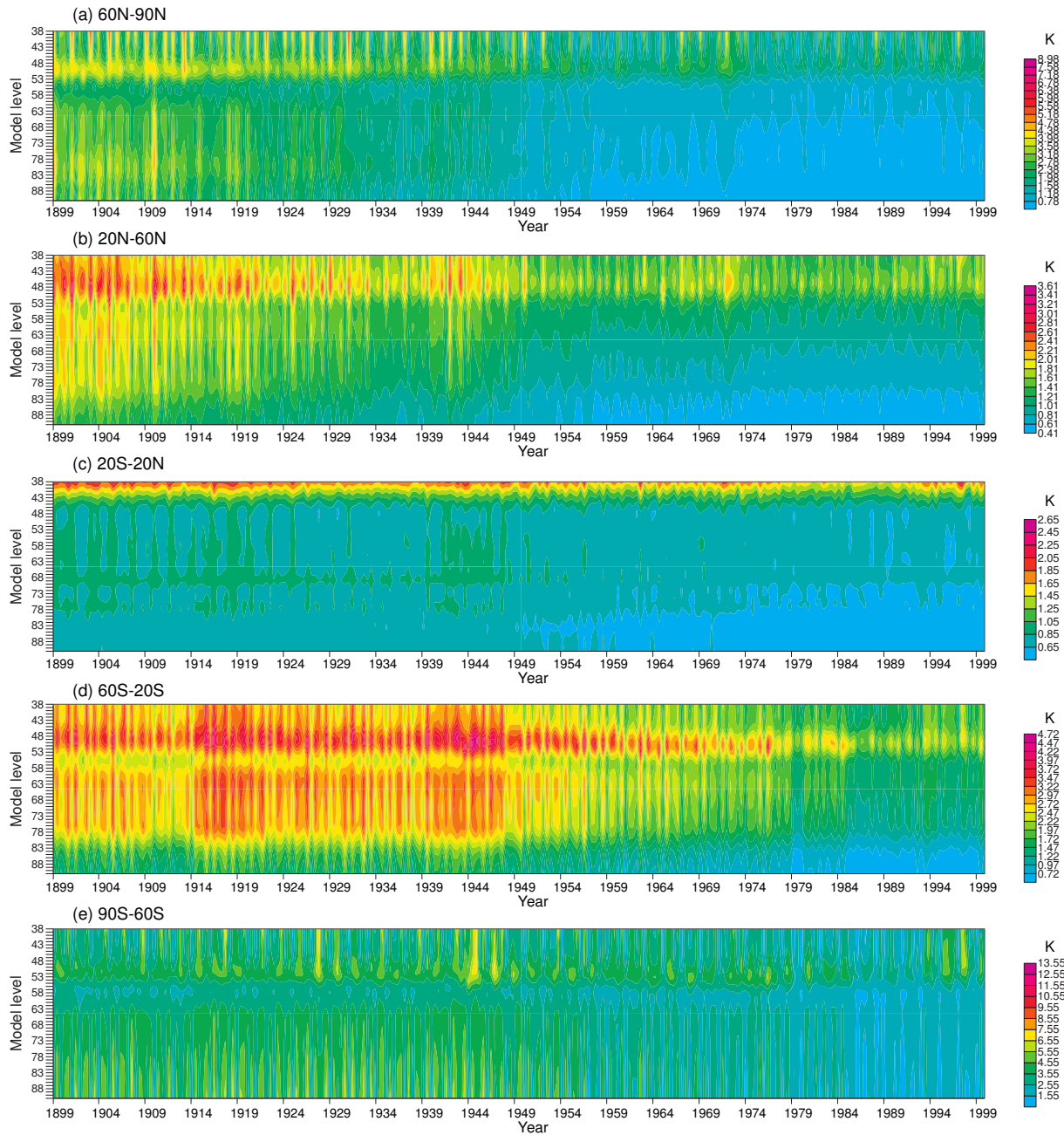


Figure 27: Mean ensemble spread for background (3-hour) forecasts in the ERA-20C Ensemble, for (a–e) five different latitude bands and vertical model levels 91–38 (surface to about 100 hPa). Plots show monthly averages.

$$\mathbf{M} = \frac{1}{N} \frac{1}{10} \sum_{i=1}^N \sum_{j=0}^9 \mathbf{x}_i^j \quad (3)$$

$$\hat{\mathbf{x}}_i = \frac{1}{10} \sum_{j=0}^9 \mathbf{x}_i^j \quad (4)$$

$$\mathbf{V}_1 = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\hat{\mathbf{x}}_i - \mathbf{M})^2} \quad (5)$$

$$\bar{\mathbf{x}}^j = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i^j \quad (6)$$

$$\mathbf{V}_2 = \frac{1}{10} \sum_{j=0}^9 \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i^j - \bar{\mathbf{x}}^j)^2} \quad (7)$$

$$\mathbf{V}_3 = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{9} \sum_{j=0}^9 (\mathbf{x}_i^j - \hat{\mathbf{x}}_i)^2} \quad (8)$$

$$\mathbf{V}_4 = \sqrt{\frac{1}{9} \sum_{j=0}^9 (\bar{\mathbf{x}}^j - \mathbf{M})^2} \quad (9)$$

$$\mathbf{V}_5 = \sqrt{\frac{1}{N-1} \frac{1}{9} (\mathbf{x}_i^j - \mathbf{M})^2} \quad (10)$$

where Eq. 3 evaluates the ensemble and monthly mean  $\mathbf{M}$  (shown in Figure 28a), Eq. 4 evaluates the ensemble mean  $\hat{\mathbf{x}}_i$  at time  $i$ , Eq. 5 evaluates the variability within the month of the ensemble means (noted  $\mathbf{V}_1$ , shown in Figure 28b), Eq. 6 evaluates the monthly mean  $\bar{\mathbf{x}}^j$  of ensemble member  $j$ , Eq. 7 evaluates the ensemble mean of the monthly variability (noted  $\mathbf{V}_2$ , shown in Figure 28c), Eq. 8 evaluates the monthly mean of the ensemble spreads (shown in Figure 28d), Eq. 9 evaluates the ensemble spread of the monthly means (shown in Figure 28e), and Eq. 10 evaluates the total spread across the month and ensemble members (shown in Figure 28f).

Some of the quantities presented here make little sense for a statistician interested in summing up total uncertainties (in which case only RMS should be considered), but bear some physical sense to separate the sources of variability by time-scale. Comparing Figures 28b and 28c shows that the intra-month variability is generally smaller in the ensemble mean than when averaged between individual members, particularly over the oceans, for the month shown. In other terms, the ensemble mean is smoother than individual members over oceans in January 1900; this is likely because the SST perturbations trigger additional activity in the simulation of synoptic systems over oceans (and this activity averages out in the mean between ensemble members). The same result is found over land in poorly observed areas, such as South America or Northern Siberia. However, in well-observed regions such as North America and Europe, Figures 28b and 28c are similar: in these regions the individual ensemble members are constrained by observations to represent about the same weather.

Comparing Figures 28d with 28e shows that an ensemble spread is only present at the synoptic time-scales (every 6 hours), but not on the monthly time-scale (the ensemble spread between the 10 monthly means is very small everywhere). Such determinism over monthly time-scales is clearly undesirable, for it would suggest that monthly means are near-perfect, although they are clearly not. The absence of spread on such time-scales stems from the design the ensemble system, where two sources of perturbations are of short memory: observations are perturbed independently of one another, and the model stochastic physics acts at each model



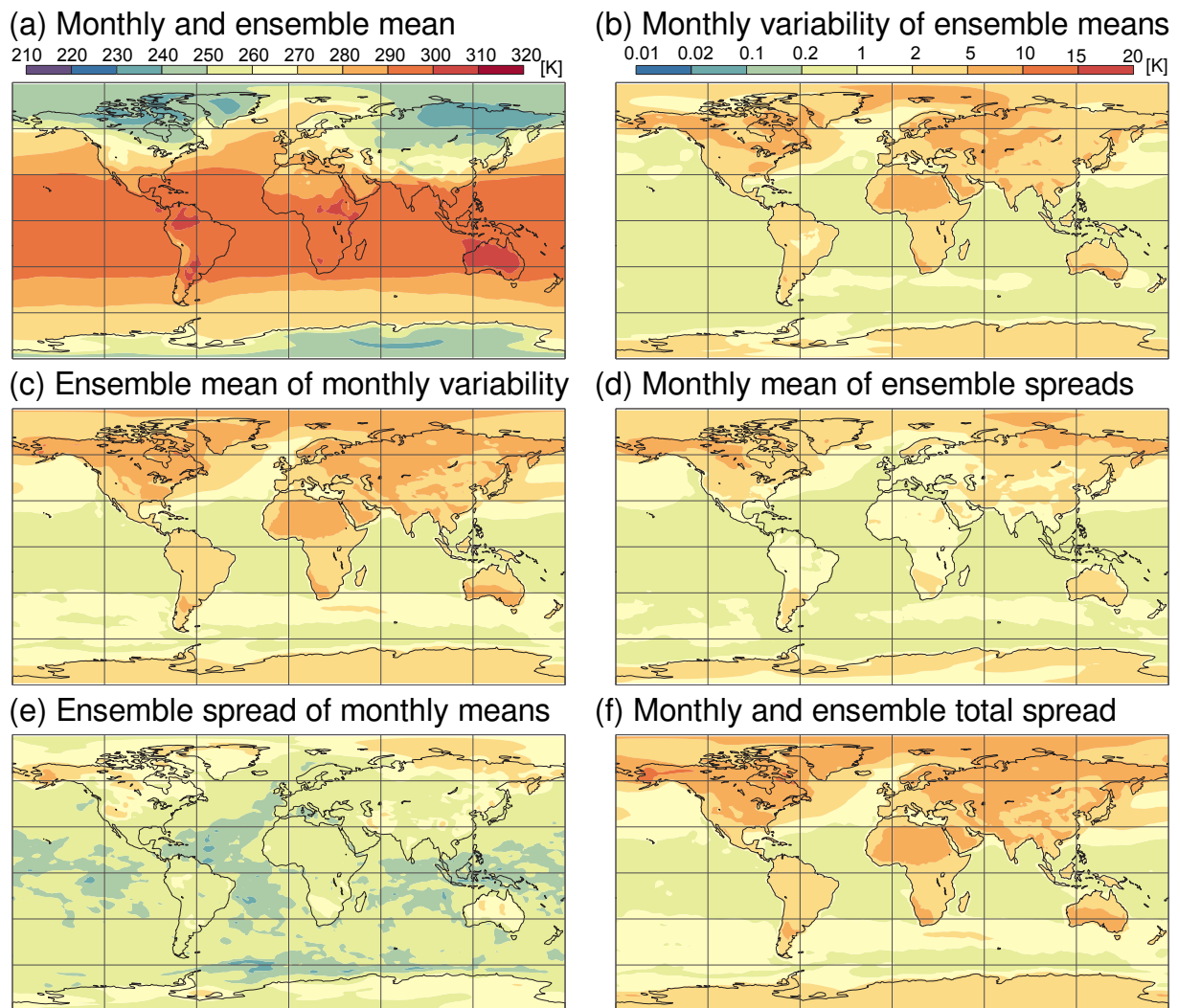


Figure 28: Two-metre temperature for January 1900 (a) monthly and ensemble mean, (b–f) various estimates of variability and spread between ensemble members and analysis times over the month. Maps (b–f) use the same color scale.

time-step, without long time correlations. The third source of perturbation (the SST forcing) includes longer time correlations. However, this is an incomplete attempt to represent uncertainties on the climate time-scales: a more complete approach would expand the ensemble to capture uncertainties in all sources of information, including the atmospheric model forcings (aerosols, ozone, solar radiation, greenhouse gases, land-use), as well as time-correlated choices in model parameterisations (*e.g.*, vertical diffusion, radiation). The second part is quite difficult to put into practice with a given model, and may be more easily tackled by considering an ensemble of atmospheric models. This may be too impractical, because for each reanalysis data assimilation system is usually built around a given atmospheric model.

As a consequence, the total monthly and ensemble spread (Figure 28f) is very close to the monthly variability (Figure 28c), because the ensemble spreads only contribute marginally to the spread on the monthly time-scale. This suggests that in spite of all the efforts to represent uncertainties in a meaningful manner, an estimator for uncertainties on the monthly time-scales may simply be to consider the intra-month time variability (a.k.a. poor man's ensemble).

The conclusions drawn from Figure 28 apply for January 1900. Timeseries of area averages for each of the

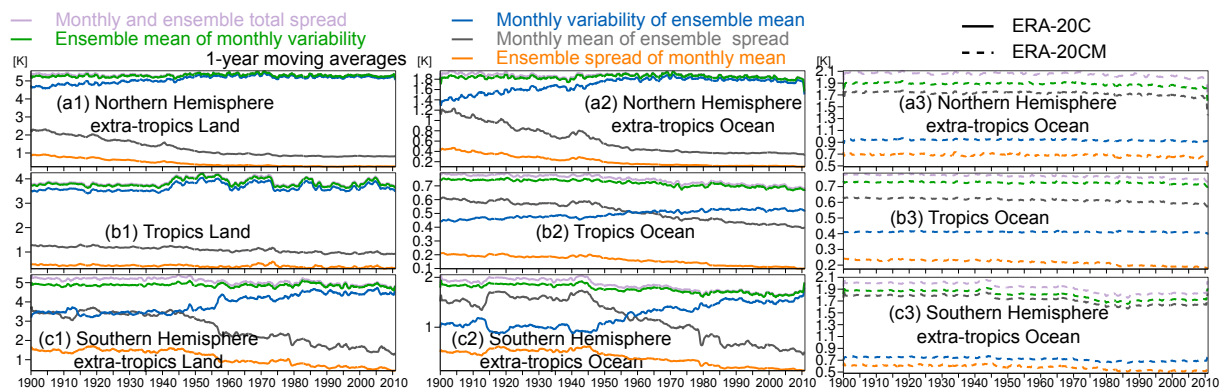


Figure 29: Timeseries of various measures of ensemble spreads for two-metre temperature shown in Figure 28. (a1–c2) show estimates from ERA-20C ensemble, (a3,b3,c3) show estimates from ERA-20CM ensemble of model integrations.

quantities shown in these maps (except for the ensemble and monthly mean) are summarized in Figure 29. The maximum in Figures 29(a1–c2) is the total monthly and ensemble spread, as expected. It is mostly made up of the ensemble mean of monthly variability. The latter is very close to monthly variability of the ensemble mean for years and areas well constrained by observations. As the observing system improves, these two metrics converge, indicating that the synoptic systems are similarly represented in each member. The monthly mean of the ensemble spread reduces over time, as seen earlier with upper-air temperatures, with a more drastic reduction in the extra-tropics than in the tropics, probably because assimilating only surface observations is insufficient to constrain the tropical analyses. The ensemble spread of monthly means is always the smallest quantity, typically below 1 K, as small as 0.1 K in the tropical oceans in the 2000s.

The timeseries suggest a few break-points. In the Northern hemisphere extratropics over ocean (in the Southern hemisphere extratropics over land), the change point appears to be right after WWII; thereafter, the monthly mean of ensemble spread is about stable (starts to decrease, respectively). In the tropics, over ocean (land), there are large episodes of suspiciously larger spread (larger variability, respectively) around both World Wars (1940–1960, respectively), which could be related to issues with the observations. Over tropical oceans, improvements become significant from 1950 onwards. For this region, as well as in the Southern hemisphere extratropics over ocean, one can nearly spot the first International Geophysical Year of 1957; from that point onwards the meteorological observations were much more regularly exchanged (and hence available today in numerical form, ready for assimilation). Extending the numerical observation record backward in time, by rescuing or digitising observations currently only on paper records, would allow to push to earlier years this limit of uncertainty. In the ERA-20C ensemble, this year is at the crossing point of the time-series of monthly variability of ensemble means and monthly mean of ensemble spreads. In fact, closer inspection suggests that both time-series mirror each other; when one increases, the other typically decreases (this is also visible in the Southern hemisphere extratropics over land).

Looking for a possible explanation of this compensating effect in the SST, Figures 29(a3,b3,c3) show similar quantities but computed from the ERA-20CM model integration. Such metrics indicate the reduction in monthly mean of ensemble spreads, but does not feature significant changes in the monthly variability of ensemble means. This indicates that the increase of monthly variability and the decrease of ensemble spread clearly come from the assimilation of observations. Why the variations of the two appear to compensate each other remains somewhat of a mystery. Even if this result appears intuitively valid, it is still possible that the assimilation method may have a role in this compensation effect.

## 7 Issues

The following issues are found in the ERA-20C deterministic production.

### 7.1 Rejection of tropical cyclone bogus observations

Many tropical cyclone bogus observations are rejected as documented in section 4.4.

### 7.2 Dubious wind observations from some buoys

Inspection of the assimilation statistics by report type, further split by platform type, shows that the ten-metre wind from drifting buoys is of lower quality than from moored buoys. Figure 30 shows density plots of observed zonal and meridional wind (for decades when buoy observations are available) versus background equivalents at the observation times and locations. In the 1970s, wind data from buoys are only available from moored platforms. From the 1980s onwards, wind is also available from platforms identified as drifting buoys. The quality of the correspondence between observations and background suggests that only moored buoy winds should probably be assimilated. Note that in the current operational implementation of observation processing prior to assimilation by IFS, there is no high-level distinction between moored and drifting buoys (although a blacklist is generated on a monthly basis by station identifier, thereby allowing expert judgment to be applied).

In addition, for both buoy types, besides the main diagonal axis, one suspects a secondary axis, perpendicular to the main one. This axis suggests that some observations report wind directions reversed by  $180^\circ$  for some time. Both aspects would require further investigations before repeating a similar reanalysis.

### 7.3 Missing tropical waves

Comparing brightness temperatures measured by SSM/I sensors with clear-sky radiative transfer calculations for scenes believed to be free of rain contamination, Poli *et al.* (2015) show that ERA-20C is able to reproduce 90% of the  $(21 \text{ K})^2$  variance within the observations for channel 3 (22 GHz), sensitive to water vapor. The remaining 10% of the variance, about  $(7 \text{ K})^2$ , cannot be solely explained by the observation measurement error. Looking for a possible source of missing water vapor variability, the authors show that anomalies over the tropical Eastern Pacific in the 30–50 day component of brightness temperatures for SSM/I channel 2 (mostly sensitive to SST) feature waves propagating westward, possible signatures of tropical waves. However, the calculations from ERA-20C miss these features completely. The SST in ERA-20C is a monthly dataset. The representation of tropical waves may be improved in a future reanalysis with the use of a SST that contains this signal (*i.e.*, pentad or daily SST product). This conjecture could be tested first with a model integration.

## 8 Unknowns: open questions

After conducting two ERA-20C productions (initial ensemble, and deterministic rerun), several questions remain unanswered.

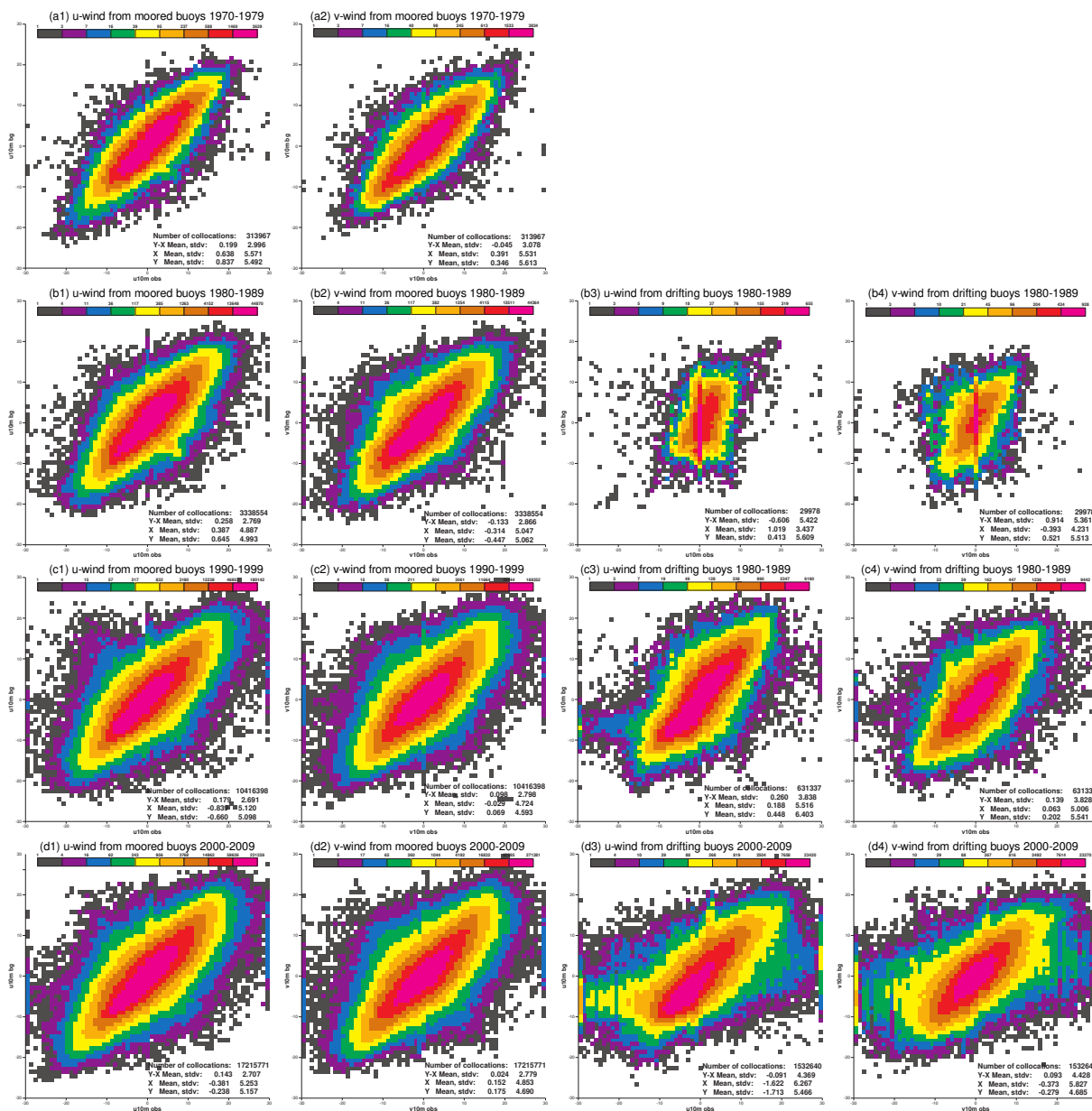


Figure 30: Density plot of observation versus ERA-20C background zonal and meridional wind (u- and v-wind, respectively) for latitudes 60S–60N, for several decades and buoy types (moored buoys in columns 1,2; drifting buoys in columns 3,4). All plots use the same axes, between  $-30 \text{ m s}^{-1}$  and  $30 \text{ m s}^{-1}$ .

## 8.1 Impact of observation coverage change on climate trends

First, having conducted in sequence a model integration only (ERA-20CM) and then a reanalysis with data assimilation (ERA-20C), one still has to quantify the impact of the changing (increasing) number of observations on the representation of climate trends. One way to assess this could be to conduct a data withholding experiment, by sub-sampling the year with most observations to the network of earlier years. However, this will not provide complete answers because natural variability and forcings change over time. Also, the forcing sea-surface conditions are based on observations; they may thus also suffer from artificial changes related to the changing number of observations over time.

## 8.2 Providing meaningful estimates of uncertainties relevant for the climate time-scales

Second, it is unclear how to present users with uncertainty estimates that characterize more than the short-term background errors (see section 6.2), but characterize also the climate time-scales. Spread between the members of the ERA-20C ensemble seem to mostly correspond to weather time-scales and not to climatic, systematic differences between the members, which would be required in order to derive reliable climate trend uncertainties.

## 8.3 Suspicious trends in the Southern Hemisphere in the early part of the 20th century

Last, D. Bromwich and J. Nicolas reported a suspicious feature in ERA-20C and NOAA 20CR at the first General Assembly of the ERA-CLIM2 project in November 2014: the general circulation in the Southern Hemisphere in both these 20th century reanalyses changes significantly over the course of the 20th century, whereas it does not in ERA-20CM. This problem is summarised in Figure 31. The timeseries of mean-sea-level pressure (MSLP) decrease significantly between the 1920s and the 2000s over a wide area. This change seems unrealistic but is featured in all seasons by both reanalyses, except for 20CR in DJF. However, the model integration (ERA-20CM) does not feature such decrease in pressure. Also noticeable is the very similar high-frequency behavior between both reanalyses, which stems from the fact that they both used nearly the same observation input, and that this input modifies the mass field at each analysis cycle, when the analysis increments are applied. In ERA-20CM, the mass field is only modified by the model integration, although this includes a mechanism to keep the total dry mass of the air constant globally.

The maps in Figure 32 show the September 1920 monthly mean MSLP analyses in the top row (Surface Pressure, SP, is shown in the bottom row). Overlaid on each map are dots showing the monthly mean of observations minus analyses, for the observations found in the ERA-20C feedback archive and that were assimilated in ERA-20C. The left two columns both show ERA-20C analyses; the only difference between them is how the observation minus analysis difference was estimated: Left-hand-side is directly from the ERA-20C observation assimilation feedback, so-called online, whereas right-hand-side is using a collocation tool to the observation location, and time within 6 hours, so-called offline. The second column is only shown here to demonstrate that the collocation procedure works correctly. The third column shows ERA-20CM, the model integration, which did not assimilate any observations of surface pressure or wind. The collocation tool is employed to calculate differences with the observations assimilated in ERA-20C. The fourth column is NOAA 20CR, where the observations shown are also those assimilated by ERA-20C, though most of these may have been also assimilated by NOAA 20CR since it used ISPD v2.2 (ERA-20C used ISPD v3.2.6). The departures from NOAA 20CR are also from the collocation tool.

As inferred from the timeseries shown in Figure 31, these maps confirm that ERA-20C and 20CR appear

more similar between one another, than they are with ERA-20CM. The dipole deep purple/deep green is much stronger in ERA-20CM than in both reanalyses, suggesting that the vortex may have been strong in ERA-20CM early on, whereas both reanalyses built it up over time (see timeseries). Also, the two reanalyses are most different from ERA-20CM where there are no observations, that is over the Antarctic plateau and the nearby seas. It is counter-intuitive that, without observations, the reanalyses would be most different from the model integration, for one would expect instead that the ERA-20C reanalysis without observation would look like the ERA-20CM model integration. The differences between observations and ERA-20CM (in the third column) are generally consistent with what ERA-20C and NOAA 20CR did with the observations. This can be checked by adding the value indicated by the dots (large swathes of red, means positive departures, large swathes of blue means negative departures) to the ERA-20CM contours, and finding more or less the values analysed by ERA-20C and 20CR where there are observations. However, this does not explain why the maps differ where there are no observations.

The overall effect of assimilated observations is to generate analyses with displaced latitudinal pressure structures. A reduced gradient results, which may be consistent with the reduced vortex intensity and the higher pressure seen over the South Pole. The very centre of what may have been the polar vortex is also quite different between the model integration and both reanalyses, until observations start to be available (especially after the 1950s). Over time (not shown), with more observations over Antarctica, the latitudinal pressure gradients in both reanalyses become more similar with the model integration. There lies a problematic trend in accelerating circulation, because this accrued gradient accelerates the zonal wind in the 40–60°S latitude band.

The inspection of the average observation departures from ERA-20CM by observation collection and by year, for MSLP and SP separately, does not reveal any collection that would be obviously biased by as much as about 10 hPa as compared to ERA-20CM southwards of 45°S. One exception is maybe the World Ocean Database collection (collection number 780 in ICOADS v2.5.1), which features biases upwards of 20 hPa for some months. However, this collection contains only few data in the region and none before the 1910s or the map shown in Figure 32.

One may conclude then that the pressure observations are not the actual problem. Also, ERA-20C and NOAA 20CR used different data assimilation systems and models, seemingly also excluding data assimilation or models as the problem. However, both reanalyses used one forcing component that was fairly similar, the sea-ice (Titchner and Rayner, 2014). One possibility is that this sea-ice forcing is inconsistent with the reality at the time. In the presence of a systematic error or bias in the background, assimilation of observations would have only corrected this problem where and when they are present, but with the constraint of dynamics being still constrained by sea-ice. Such possibility could have resulted in damaging the large-scale features such as latitudinal pressure gradients.

The recent reanalysis 20CRv2c produces results similar to the earlier version, 20CR (Compo, *pers. comm.*). However, 20CRv2c reanalysis uses a different sea-ice product as forcing, thereby suggesting that sea-ice may not be the actual issue. Consequently, the problem remains hence unresolved at the time of writing.

## 9 Conclusions and outlook

The ERA-20C deterministic reanalysis solves several issues in the ERA-20C ensemble, and produces more realistic climate trends. The present report indicates that the inter-annual variability of total column water vapor in ERA-20C deterministic for years 1988–2008 is closer than ERA-Interim and JRA-55 to observational products such as HOAPS and RSS. The time evolution of the total column ozone over northern high latitudes in ERA-20C deterministic is also found to be realistic as compared to MSR-2 for the years 1979–2008, and of better quality than the model integration ERA-20CM. This indicates that assimilating surface pressure observa-

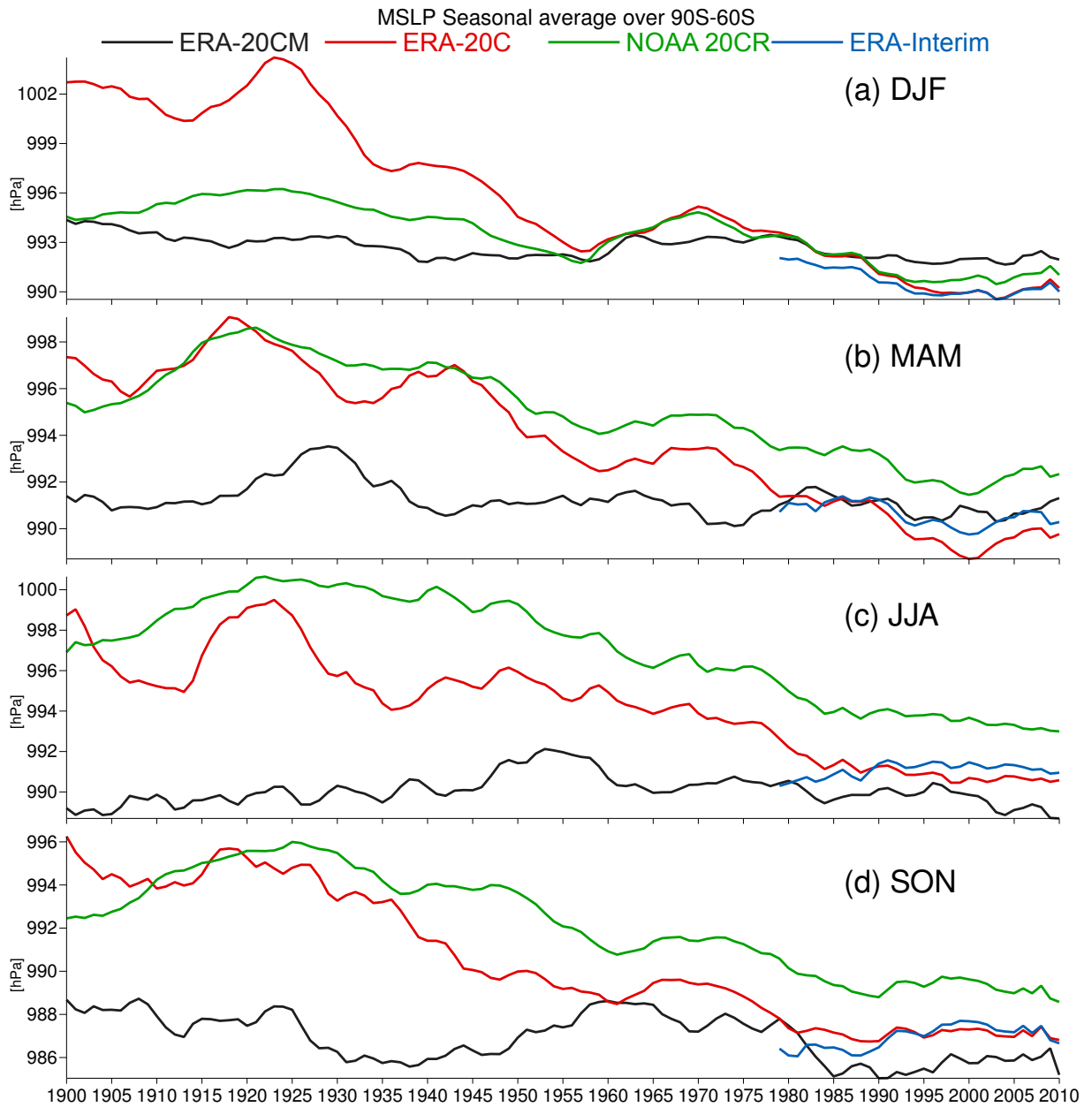


Figure 31: Timeseries of mean-sea-level pressure (MSLP) averaged over the latitudes 90S-60S in one model integration (ERA-20CM) and various reanalyses (ERA-20C, NOAA 20CR, ERA-Interim).

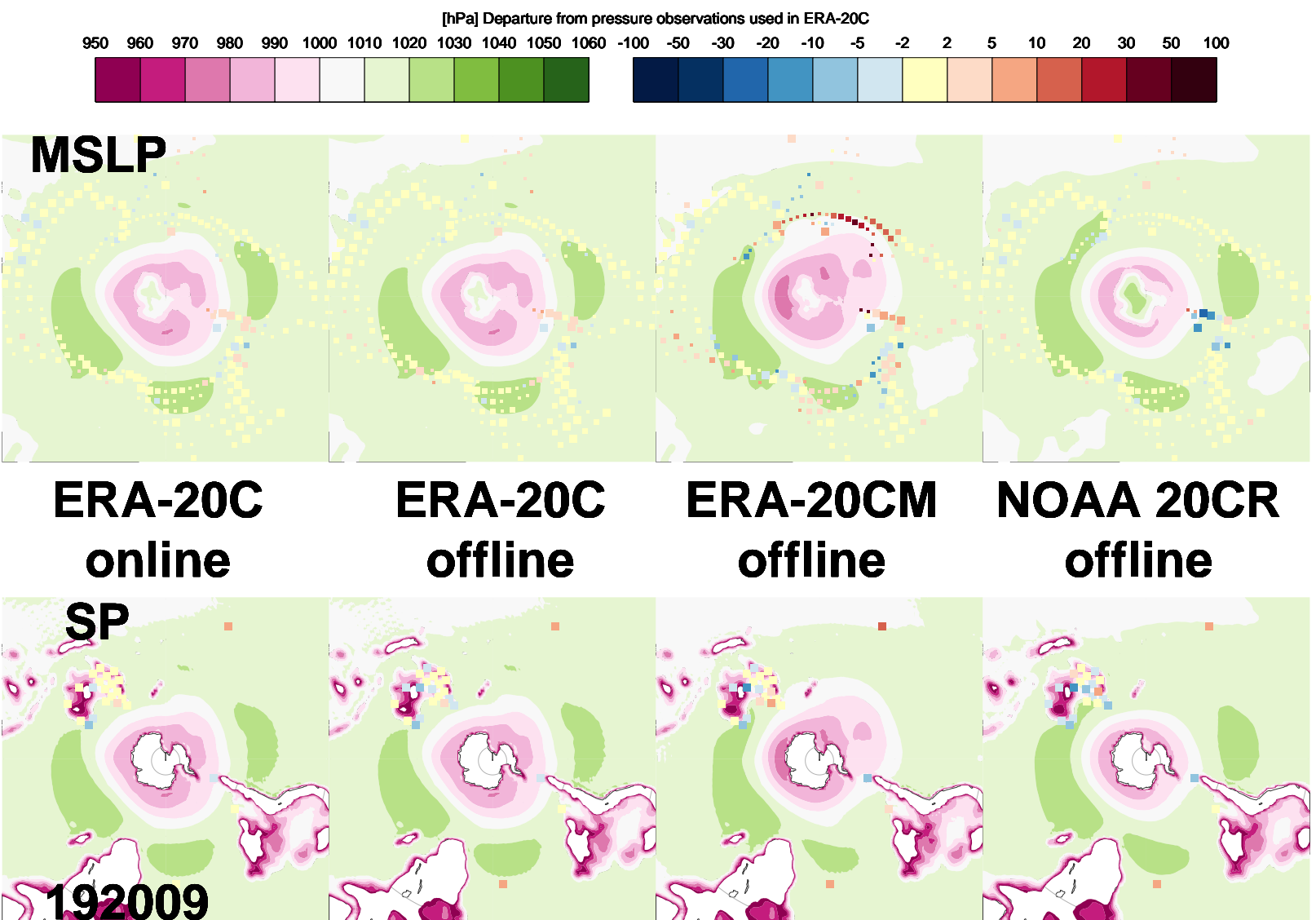


Figure 32: Maps of MSLP (top row) and SP (bottom row) analyses from two reanalyses (ERA-20C and NOAA 20CR) and a model integration (ERA-20CM), for September 1920, colored using the palette on the left (from purple to green). Dots show monthly mean of observations minus analyses departures, considering only the MSLP or SP observations assimilated in ERA-20C (color scale on the top right, from blue to red, yellow suggesting agreement between the analyses and the pressure observations). All maps use the same polar stereographic projection. The sizes of the dots are relative to the number of observations within the map. See text for details.



tions may help locate the polar vortex in a way that benefits the representation of ozone *via* the model dynamics and radiation. Based on several metrics, the quality of the general circulation in ERA-20C deterministic in the Northern hemisphere seems adequate mostly after the 1950s, and marked by two phases of clear improvements after each World War. In the Southern hemisphere, the quality of the circulation is much more doubtful, and several metrics suggest that it may only be realistic after the deployment of many drifting buoys in the 1980s. Various indicators point to the reanalysis quality being limited by the availability of input observations.

In terms of energy budgets, ERA-20C deterministic features a much improved overall balance as compared to ERA-20C ensemble. However, the budget is still quite far from the near-perfect balance obtained with ERA-20CM, where there are no observations to contradict the model's interpretation of the forcings. The present report proposes to compute analysis increments in the space of total atmospheric energy change as a way of estimating this discrepancy. This would give an absolute measure, in  $W/m^2$ , of how much global disagreement exists as a function of time between the model and its forcings on the one hand, and the observations on the other hand. Figure 10d by Berrisford *et al.* (2011b) shows a time series of budget residual as a proxy for the analysis increment in terms of total energy for ERA-Interim, but that estimate also contains some of the model error. The approach proposed in the present report could result in a benchmark metric comparable across reanalyses, enabling to track progress in understanding of climate change, and quantifying the value of data rescue to help improve our historical record.

The weather maps produced by ERA-20C deterministic are of slightly lower quality than ERA-20C ensemble. This is caused by background error structure functions that do not reflect the poor observing system, and locate the increments too close to the observation locations. As intended, this modification has the benefit of preventing analysis increments in the stratosphere and preserving low-frequency trends created by the model and its forcings. However, since ERA-20C productions, developments in the IFS system have been carried out (Bonavita, 2015), whereby flow-dependent correlations are derived through a linear combination of climatological structure functions and the ensemble information (The ERA-20C ensemble only relied on the ensemble information to derive flow-dependent structure functions, resulting in vertical correlations reaching from the surface to the stratosphere). Figure 33(b) shows that in the new IFS ensemble system (version CY41r2) it is possible to compute background errors structure functions that produce analysis increments that remain localized. For completeness, Figure 33(a) shows the situation for IFS ensemble version CY41r1, which precedes these latest developments, and where the results resemble the ERA-20C ensemble (CY38r1). One may thus expect that a future repeat of ERA-20C ensemble with IFS version CY41r2 would produce climate trends similar to ERA-20C deterministic, while still using background error structure functions that correspond to the poor observing system.

The present report explores the use of the spread from the earlier ERA-20C ensemble production as a proxy for estimating uncertainties. Results suggest that these metrics can be used for uncertainties in synoptic weather charts, but not to assess uncertainties related to climate time-scales. For these to be represented, one option that would require developments is to use an ensemble of model forcings, or better yet to use an ensemble of models in the assimilation. Why the increase in intra-monthly variability compensates quantitatively the decrease in ensemble spread of two-metre temperatures remains unexplained, even if one would qualitatively expect such result from the assimilation of synoptic weather observations.

In addition, several leads for improvement are proposed for a future repeat of a similar reanalysis: (1) detect buoys reporting bad wind observations, or observations with a  $180^\circ$  rotation in the wind direction, (2) exclude from the assimilation wind observations associated with XBT, (X)CTD, and MBT reports, (3) do not subject bogus observations (such as tropical cyclone tracks) to the constant time-series observation quality control, (4) augment the observation errors assumed for such observations, and revise more generally the observation error estimates for all reports, according to figures given in the present report (and use these in an ensemble framework, so the subsequent background errors are appropriately updated), and (5) assess the impact of a

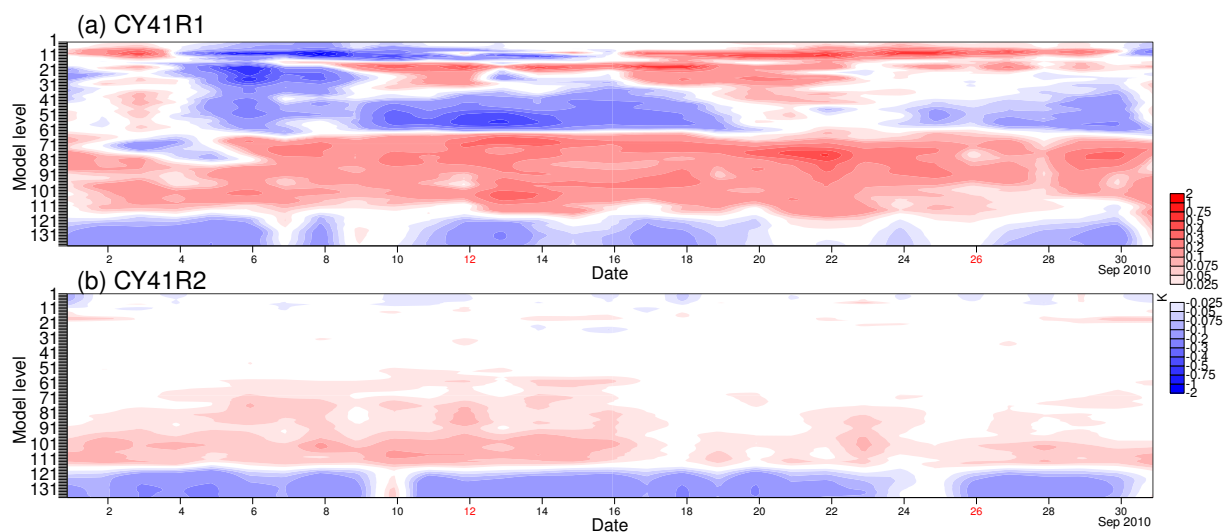


Figure 33: Mean analysis increments at the beginning of the analysis window for temperature, for a data assimilation system similar to ERA-20C but in more recent versions of the IFS.

sub-monthly SST forcing on the representation tropical waves.

Points requiring further research concern how to present users with meaningful estimates of uncertainty for the monthly time-scale, and how to quantify for users the impact of the observation coverage change on the climate trends.

The next 20th century reanalysis at ECMWF will include coupling with ocean (Laloyaux *et al.*, 2014). In doing so, it will bring about new challenges. Yet, it will be valuable to integrate in it as many as possible of the findings of the ERA-20C ensemble and deterministic reanalyses. Other developments include the addition of upper-air observations, of which a trial was conducted in parallel with ERA-20C deterministic production, and will be reported on separately. Adding such observations is expected to add realism, but does not change the fundamental problem that the main barriers to improving our knowledge of the past requires data rescue, so as to bring to digital form observations currently held on paper or other form.

## Acknowledgement

The European reanalysis of global climate observations project (ERA-CLIM) is funded by EU research FP7 Environment under grant agreement no. 265229. In addition, ERA-CLIM benefits from significant contribution from ECMWF, for high-performance computing, archive, data services, and supplementary staff support. We thank NOAA and NCAR for providing the ICOADS v2.5.1 and ISPD v3.2.6 observational datasets. We are grateful to Rob Allan and his collaborators for their efforts in setting up and coordinating the Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative, which facilitates and undertakes data rescue activities essential for the continuous improvement of ISPD and ICOADS. The HadISST2.1.0.0 dataset was provided by the Met Office Hadley Centre with partial funding from ERA-CLIM. We thank the ERA-CLIM advisory board for fruitful comments and suggestions during the preparation of ERA-20C. Support for the International Surface Pressure Databank is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. MSR-2 data used for validation were retrieved from

[http://www.temis.nl/macc/index.php?link=o3\\_msr\\_intro.html](http://www.temis.nl/macc/index.php?link=o3_msr_intro.html) on 12 May 2015.

## References

Alexandersson H. 1986. A homogeneity test applied to precipitation data. *J. Climatol.*, 6: 661–675. DOI: 10.1002/joc.3370060607

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, 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 1901present. *Earth Syst. Sci. Data*, 5: 71–99. DOI: 10.5194/essd-5-71-2013

Berrisford P, Dee D, Poli P, Brugge R, Fielding K, Fuentes M, Kållberg P, Kobayashi S, Uppala S, Simmons A. 2011a. The ERA-Interim archive Version 2.0. ERA Report Series 1 version 2. Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom, and on the web from [http://old.ecmwf.int/publications/library/ecpublications/\\_pdf/era/era\\_report\\_series/RS\\_1\\_v2.pdf](http://old.ecmwf.int/publications/library/ecpublications/_pdf/era/era_report_series/RS_1_v2.pdf) (last accessed 20 May 2015)

Berrisford P, Kållberg P, Kobayashi S, Dee D, Uppala S, Simmons AJ, Poli P, Sato H. 2011b. Atmospheric conservation properties in ERA-Interim. *Q.J.R. Meteorol. Soc.*, 137: 1381–1399. DOI: 10.1002/qj.864

Bonavita M. 2015. Ensemble of Data Assimilations and Hybrid Gain EnDA. ECMWF Data Assimilation Training Course 2015. Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom, and on the web from [https://software.ecmwf.int/wiki/download/attachments/30048347/Bonavita\\_EDA\\_TC2015.pdf](https://software.ecmwf.int/wiki/download/attachments/30048347/Bonavita_EDA_TC2015.pdf) (last accessed 17 June 2015)

Brönnimann S, Compo GP. 2012. Ozone highs and associated flow features in the first half of the twentieth century in different datasets. *Meteorol. Zeit.*, 21: 49–59. DOI: 10.1127/0941-2948/2012/0284

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS, Rutledge G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli O, Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ. 2011. The Twentieth Century Reanalysis Project. *Q.J.R. Meteorol. Soc.*, 137: 1–28. DOI: 10.1002/qj.776

Dee DP, Balsameda M, Balsamo G, Engelen R, Simmons AJ, Thépaut J-N. 2014. Toward a Consistent Reanalysis of the Climate System. *Bull. Amer. Meteor. Soc.*, 95: 1235–1248. DOI: <http://dx.doi.org/10.1175/BAMS-D-13-00043.1>

Dee DP. 2004. Variational bias correction of radiance data in the ECMWF system. Proc. ECMWF workshop on assimilation of high spectral resolution sounders in NWP. Reading, United Kingdom, 28 June - 1 July 2004

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Q.J.R. Meteorol. Soc.*, 131: 3385–3396. DOI: 10.1256/qj.05.108

ECMWF. 2013. IFS documentation CY38R1, available from <http://old.ecmwf.int/research/ifsdocs/CY38r1> (last accessed 12 May 2015)

Haimberger L. 2007. Homogenization of Radiosonde Temperature Time Series Using Innovation Statistics. *J. Climate*, 20: 1377–1403. DOI: <http://dx.doi.org/10.1175/JCLI4050.1>

Hersbach H, Peubey C, Simmons A, Berrisford P, Poli P, Dee D. 2015a. ERA-20CM: a twentieth century

atmospheric model ensemble. Q.J.R. Meteorol. Soc., DOI: 10.1002/qj.2528

Hersbach H, Poli P, Dee D. 2015b. The observation feedback archive for the ICOADS and ISPD data sets. ERA Report Series 18. Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom, and on the web from

[http://old.ecmwf.int/publications/library/ecpublications/\\_pdf/era/era\\_report\\_series/RS\\_18.pdf](http://old.ecmwf.int/publications/library/ecpublications/_pdf/era/era_report_series/RS_18.pdf) (last accessed 20 May 2015)

Hersbach, H, Peubey C, Simmons A, Poli P, Dee D, Berrisford P. 2013. ERA-20CM: a twentieth century atmospheric model ensemble. ERA Report Series 13. Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom, and on the web from

[http://old.ecmwf.int/publications/library/ecpublications/\\_pdf/era/era\\_report\\_series/RS\\_16.pdf](http://old.ecmwf.int/publications/library/ecpublications/_pdf/era/era_report_series/RS_16.pdf) (last accessed 20 May 2015)

Isaksen L, Bonavita M, Buizza R, Fisher M, Haseler J, Leutbecher M, Raynaud L. 2010. Ensemble of data assimilations at ECMWF. TM 636. December 2010

Kennedy J, Rayner NA, Millington SC, Saunby M. 2015. The Met Office Hadley centre sea ice and sea-surface temperature data set, version 2: 2.sea-surface temperature analysis. In preparation

Knapp KR, Kruk MC, Levinson DH, Diamond HJ, Neumann CJ. 2010. The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone best track data. Bull. Amer. Meteor. Soc. 91: 363–376. DOI: 10.1175/2009BAMS2755.1

Kuchta. P. 2009. Observation feedback archiving in MARS. 12th Workshop on Meteorological Operational Systems, 2-6 November 2009. ECMWF, Reading, United Kingdom.

[http://www.ecmwf.int/newsevents/meetings/workshops/2009/MOS\\_12/Presentations/Kuchta.pdf](http://www.ecmwf.int/newsevents/meetings/workshops/2009/MOS_12/Presentations/Kuchta.pdf)

Laloyaux P, de Boissesson E, Balmaseda M, Mogensen K. 2014. The Coupled ECMWF ReAnalysis system. The World Weather Open Science Conference, Montréal, Canada, 16-21 August 2014. Presentation available from

<https://www.wmo.int/pages/prog/arep/wwrp/new/wwosc/documents/laloyaux.pdf> (last accessed 19 May 2015)

Palmer TN, Buizza R, Doblas-Reyes F, Jung T, Leutbecher M, Shutts GJ, Steinheimer M, Weisheimer A (2009): Stochastic parametrization and model uncertainty. ECMWF Research Department Technical Memorandum, 598, ECMWF, Shinfield Park, Reading RG2-9AX, UK, pp. 42

Poli, P, Peubey C, Fennig K, Schröder M, Roebeling R, Geer A. 2015. Pre-assimilation feedback on a Fundamental Climate Data Record of brightness temperatures from Special Sensor Microwave Imagers: A step towards MIPs4Obs? ERA Report Series 19 Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom, and on the web from

[http://old.ecmwf.int/publications/library/ecpublications/\\_pdf/era/era\\_report\\_series/RS\\_19.pdf](http://old.ecmwf.int/publications/library/ecpublications/_pdf/era/era_report_series/RS_19.pdf) (last accessed ?????? 2015)

Poli P, Hersbach H, Tan T, Dee D, Thépaut J-N, Simmons A, Peubey C, Laloyaux P, Komori T, Berrisford P, Dragani R, Trémolet Y, Holm E, Bonavita M, Isaksen L, Fisher M. 2013. The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C). ERA Report Series 14. Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom, and on the web from

[http://old.ecmwf.int/publications/library/ecpublications/\\_pdf/era/era\\_report\\_series/RS\\_14.pdf](http://old.ecmwf.int/publications/library/ecpublications/_pdf/era/era_report_series/RS_14.pdf) (last accessed 20 May 2015)

Poli, P. 2013. Estimation of surface pressure and wind observation errors from ERA-20C observation feedback.

13th European Meteorological Society (EMS) Annual Meeting, 9-13 September 2013, Reading, United Kingdom. Presentation EMS2013-173 Available from

[http://presentations.copernicus.org/EMS2013-173\\_presentation.pptx](http://presentations.copernicus.org/EMS2013-173_presentation.pptx) (last accessed 27 August 2014)

Titchner HA, Rayner NA. 2014. The Met Office Hadley Centre sea ice and sea-surface temperature data set, version 2: 1. sea ice concentrations. *J. Geophys. Res.* 119:2864–2889. DOI: 10.1002/2013JD020316

van der A RJ, Allaart MAF, Eskes HJ. 2015. Extended and refined multi sensor reanalysis of total ozone for the period 1970-2012. *Atmos. Meas. Tech. Discuss.* 8: 3283–3319, DOI: 10.5194/amtd-8-3283-2015

Woodruff SD, Worley SJ, Lubker SJ, Ji Z, Freeman JE, Berry DI, Brohan P, Kent EC, Reynolds RW, Smith SR, Wilkinson C. 2011. ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, 31, 951-967. DOI: 10.1002/joc.2103

Yin X, Gleason BE, Compo GP, Matsui N, Vose RS. 2008. The International Surface Pressure Databank (ISPD) land component version 2.2. National Climatic Data Center, Asheville, North Carolina, United States of America. 12 pp.