ERA report series

25 Estimates of variations and trends of global surface temperature

Adrian Simmons, Paul Berrisford, Dick Dee, Hans Hersbach,
Shoji Hirahara and Jean-Noël Thépaut

Submitted for publication in the Quarterly Journal of the Royal Meteorological Society

Series: ERA Report Series

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

http://old.ecmwf.int/publications/

Contact: library@ecmwf.int

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

Abstract

The ERA-Interim and JRA-55 reanalyses of synoptic data and several conventional analyses of monthly climatological data provide similar estimates of global-mean surface warming since 1979. They broadly agree on the character of interannual variability and the extremity of the recent warm spell to which a strong El Niño and low Arctic sea-ice cover contribute. Global and regional averages nevertheless differ on various timescales due to differences in data coverage and sea-surface temperature analyses; averages from those conventional datasets that infill where they lack direct observations show better agreement with the averages from the reanalyses. Warming from 1998 to 2012 is larger than indicated by earlier versions of the conventional datasets used to characterize what the Fifth Assessment Report of the Intergovernmental Panel on Climate Change termed a hiatus in global warming. None of the datasets exhibit net warming over the Antarctic since 1979.

Centennial trends from the conventional datasets, HadCRUT4 on one hand and GISTEMP and NOAAGlobalTemp on the other, differ mainly because sea-surface temperatures differ. Infilling of values where direct observations are lacking is more questionable for the data-sparse earlier decades. Change since the18th century is inevitably more uncertain than change over and after a modern baseline period. The latter is arguably best estimated separately for taking stock of actions to limit climate change, exploiting reanalyses and using satellite data to refine the conventional approach. Nevertheless, early in 2016, however briefly, global temperature appears to have first reached or breached a level 1.5°C above that early in the Industrial Revolution, having touched the 1.0°C level briefly in 1998 during a previous El Niño.

Atmospheric energy is an alternative metric for tracking change. It gives more weight to tropical than high-latitude variability, due to the greater vertical penetration of the thermal signal and importance of latent energy in the tropics.

1 Introduction

The latest two assessment reports of the Intergovernmental Panel on Climate Change (IPCC) have stated that warming of the climate system is unequivocal, citing among other evidence the increases in global average surface air and ocean temperatures inferred from observations (IPCC, 2007; 2013). Differences in estimates of short-term trends in global-mean surface temperature have nevertheless been sufficiently large to prompt debate within the scientific community over reference to a "hiatus" or "slowdown" in warming over the fifteen or so years following the 1997/98 El Niño event (Lewandowsky *et al.*, 2015; Fyfe *et al.*, 2016). Differences between datasets in their rankings of individual years and months in terms of warmth also hamper clear public communication of reliable information concerning extreme values.

These issues have come to the fore recently for several reasons. Newer versions and a wider range of datasets show a higher rate of warming from 1998 to 2012 than indicated in IPCC (2013). Atmospheric temperatures have reached levels in a recent spell that are by a considerable margin unprecedented over the period of instrumental record. And the aim of "holding the increase in the global average temperature to well below 2°C above pre-industrial levels and pursuing efforts to limit the temperature increase to 1.5°C above pre-industrial levels" was agreed by nations meeting in Paris late in 2015 (UNFCCC, 2015). The Paris Agreement includes an undertaking to take stock periodically of progress towards achieving its purpose and long-term goals, starting in 2023 and continuing at intervals of five years unless subsequently decided otherwise. This points to a continuing need to reduce uncertainties in estimates of temperature and to improve the interpretation of the sub-decadal variations in a temperature record that combines interacting effects of anthropogenic and natural external forcings and internal variability of the climate system.

Aside from setting its target for limiting the rise in global temperature, the Paris Agreement also established a global goal of enhancing capacity to adapt to climate change. In doing so it recognised that the challenge of adaptation had local, subnational, national, regional and wider international dimensions, with requirements for strengthened research, systematic observation and early warning systems. This in turn sets requirements for the observation, analysis and prediction of key impact variables such as surface air temperature, including needs for spatial and temporal resolution and for estimation of uncertainty. There are also requirements for services to deliver data and derived information.

The term reanalysis is used in a specific context to describe the use of a fixed modern data assimilation system to synthesize states of the atmosphere and interacting components of the climate system from past and present observations. This type of reanalysis has a unique contribution to make in that it provides globally complete estimates of many of the key variables with a frequency and resolution that are becoming increasingly high in newer products, yet also provides global and regional averages that are competitive with those from conventional monthly temperature and humidity products for studying trends and low-frequency variability over recent decades. For example, ECMWF's ERA reanalyses have shown up well in comparisons with several other reanalyses and conventional products, and directly in comparison with observations (Simmons *et al.*, 2004, 2010; Jones *et al.*, 2012; Simmons and Poli, 2015). ECMWF's latest comprehensive atmospheric reanalysis, ERA5, has recently entered production as a contribution to Europe's Copernicus Climate Change Service (http://climate.copernicus.eu), and is generating hourly products using a background assimilating model with close to 30km horizontal resolution. It is expected to address a number of the remaining issues related to the consistency over time of temperature and humidity analyses both near the surface and in the free atmosphere, as discussed in particular by Dee *et al.* (2011) and Simmons *et al.* (2014).

This report extends the comparisons of large-scale temperature trends and variability to new versions of the three most widely used conventional datasets and a new reanalysis, and includes a documentation of the representations by the various datasets of the extreme warm temperature anomalies that have occurred in recent months. It explores the extent to which values created by infilling or extrapolation in what are observation-void areas for the conventional datasets agree with the values produced in such regions by reanalyses, which infer values there from additional synoptic observations of surface air temperature, from *in situ* and satellite observations of other variables and from modelling. Sources of differences among datasets are identified. A particular aim is to establish more firmly the credentials of reanalysis for monitoring global and regional temperature, especially with regard to use of ERA products to provide information delivered by the Copernicus Climate Change Service. This necessitates placing recent decades in the centennial-scale context for which the conventional datasets still play the primary role. Results from reanalysis are also used to place the surface air temperature record in the context of changes in atmospheric energy and upper-air temperature.

Time series and maps of surface air temperature (temperature nominally at a height of two metres above the ground) are compared for four datasets that are based on different analyses of sea-surface temperature (SST), and for other datasets that differ in the completeness of their global coverage. Two are reanalyses that include an analysis of synoptic surface air temperature observations: ECMWF's ERA-Interim (Dee *et al.*, 2011) and the Japan Meteorological Agency's JRA-55 (Kobayashi *et al.*, 2015). The other two datasets based on differing SSTs are the latest versions of "surface temperature" products that have been conventionally used to characterize the long-term warming that has occurred since the 19th century: HadCRUT4 (Morice *et al.*, 2012), produced by the Met Office Hadley Centre in collaboration with the Climatic Research Unit of the University of East Anglia, and NOAAGlobalTemp (Karl *et al.*, 2015), produced by the US National Oceanic and Atmospheric Administration (NOAA). These datasets combine analyses of climatological reports of monthly-mean surface air temperature from stations over land with monthly analyses of SST. SST is used rather than marine surface air temperature as the latter is more difficult to analyse reliably directly from observations. HadCRUT4 is an ensemble of 100 possible realizations of past temperature change that sample some of the uncertainties in estimating multi-decadal variability; results presented here are primarily for the medians of the values from the ensemble, although the spread of the ensemble is also examined.

Comparisons are also made with the GISTEMP (Hansen *et al.*, 2010) dataset produced by the US National Aeronautics and Space Administration (NASA). Various versions of GISTEMP and the Met Office and NOAA datasets have commonly been used to produce summary results on global temperature, not least by the IPCC and in the annual statements of the World Meteorological Organization (WMO) on the status of the global climate. The latest version of GISTEMP studied here uses the same SST analyses and input database of monthly climatological station data as NOAAGlobalTemp, but differs in its data processing, particularly in the extent to which values are spread to regions distant from where there are station observations or analysed values of SST. Two datasets that spatially extend the HadCRUT4 median (Cowtan and Way, 2014) are also examined.

The outline of the report is as follows. Further information on the datasets and their processing is given in the following section. Time series of anomalies in global- and European-average temperatures from 1979 onwards are compared in section 3, and global trends are discussed in section 4. Differences in individual monthly values are examined in section 5. Maps illustrating geographical coverage and variations are presented in section 6, and the contributions of the polar regions and middle and low latitudes to global trends and variability form the topic of section 7. The period from 1979 is placed in longer-term context in section 8, while section 9 discusses variations in atmospheric energy and upper-air temperature over the period. Concluding discussion is provided in section 10.

2 Datasets

The ERA-Interim and JRA-55 reanalyses cover the time ranges to the present from January 1979 and January 1958 respectively. Monthly averages are used for comparison with the conventional datasets. Both reanalyses provide a two-metre temperature product derived from an analysis of synoptic screen-level observations that uses the background fields from their primary 4D-Var data assimilation schemes. Except where stated otherwise, the analysis fields are used over land and the background fields over sea. More detail and discussion in the case of ERA-Interim is given by Simmons *et al.* (2004, 2010) and Simmons and Poli (2015). Kobayashi *et al.* (2015) give an outline of the screen-level analysis used in JRA-55.

A further adjustment of the standard ERA-Interim output is made. The two-metre temperatures over ice-free sea, and the SSTs when used alternatively, are reduced by 0.1°C for all months prior to January 2002. ERA-Interim, like other atmospheric reanalyses, uses externally produced analyses of SST. A new source for these analyses had to be used from the beginning of 2002, and this resulted in a widespread and largely uniform reduction by around 0.1°C in the ERA-Interim SST, to which two-metre temperature is closely linked. Further information can be found in Simmons and Poli (2015) and the earlier references cited above. Other technical aspects of the processing of data from ERA-Interim and JRA-55 are as described by Simmons *et al.* (2010, 2014).

Previous comparisons of ERA-Interim and JRA-55 for upper-air temperature (Simmons *et al.*, 2014) and Arctic surface air temperature (Simmons and Poli, 2015) included evaluations of the MERRA reanalysis (Rienecker *et al.*, 2011). Production of MERRA has since been discontinued, following availability of MERRA-2 (Bosilovich *et al.*, 2015), which runs from 1980 onwards. Monthly-mean two-metre temperature fields[1] from MERRA-2 have been processed and compared with the other datasets evaluated in this report. Results are not discussed at length, however, as several issues render this reanalysis a clear outlier in terms of trends. Some discussion is given later in the report.

HadCRUT4 covers the period from January 1850. Data for the latest month are added regularly, but values for previous months are updated only intermittently. Version 4.4.0.0 is used in this study. NOAAGlobalTemp runs from January 1880 onwards. Past values in this dataset may change with each monthly release; version 4.0.1.201604 is used here. Both datasets provide values for 5°x5° grid squares, and for both there are gaps in global coverage. As illustrated later, these gaps are larger in the case of HadCRUT4, for which no infilling of data is performed to construct values for grid squares for which a direct calculation cannot be made. A gap can occur either because land-station or marine temperature data are lacking for the month in question or because there are insufficient data to establish a background climatological value for the location.

The SST analyses used by HadCRUT4 (HadSST3; Kennedy *et al.*, 2011), JRA-55 (COBE; Ishii *et al.*, 2005), and NOAAGlobalTemp (ERSSTv4; Huang *et al.*, 2015) are based only on *in situ* measurements, whereas the sequence of SST analyses used by ERA-Interim (Dee *et al.*, 2011) benefits from observations from satellites as well as direct measurements. HadCRUT4 and NOAAGlobalTemp analyse monthly data records from a largely similar set of land stations, but as HadCRUT4 provides data for some locations where NOAAGlobalTemp does not, there are evidently some differences in data input or quality control. The reanalyses assimilate many types of observation in their 4D-Var schemes, and the assimilated data are largely the same for ERA-Interim and JRA-55. The synoptic surface-air temperature data analysed by both come from many more stations than provide the monthly data used by HadCRUT4 and NOAAGlobalTemp, although a few of the stations that report monthly do not provide synoptic data that regularly reach ECMWF via the WMO Global Telecommunications System. This includes some stations in data-sparse parts of Africa. Maps and statistics can be found in GCOS (2015).

[1] Downloaded from the MERRA-2 instM_2d_asm_Nx data stream, version 5.12.4, held by the NASA Goddard Space Flight Center Distributed Active Archive Center.

GISTEMP covers the period since 1880 and its current version uses the same input land-station observations and ERSSTv4 dataset as NOAAGlobalTemp. Analysed data are provided on a 2°x2° grid, and global coverage for recent years is close to but not quite complete. This coverage results from an analysis method that exploits the correlation of temperature change found for stations separated by up to 1200km (Hansen and Lebedeff, 1987) and data are available generally with 1200km smoothing. 2°x2° grid values over many continental land areas and islands are also supplied with 250km smoothing. The GISTEMP results for global means presented here use downloaded mean values. Regional means and maps are produced from a merged dataset that uses 2°x2° values from the 250km smoothed dataset where available and 1200km smoothed values otherwise. The GISTEMP datasets for past months may change when a new monthly release is made: the datasets used here were downloaded on 16 May 2016 from http://data.giss.nasa.gov/gistemp/.

Cowtan and Way (2014) reported the construction of two globally complete datasets that extend HadCRUT4. The first uses solely the spatial interpolation method known as kriging, while the second is a hybrid approach that adds information from the record of tropospheric temperature derived at the University of Alabama in Huntsville (UAH; Christy *et al.*, 2007) from space-based microwave sounding data. Updated versions of the Cowtan and Way datasets downloaded from (and documented at) http://www-users.york.ac.uk/~kdc3/papers/coverage2013/series.html are used in the present study: the hybrid dataset Had4_UAH_v2 that is restricted to the period from 1979 when the satellite data are available and the Had4_krig_v2 dataset covering the period from 1850 onwards.

The HadCRUT4 and related datasets comprise values that are anomalies relative to the reference period 1961-1990, NOAAGlobalTemp provides anomalies relative to 1971-2000 and GISTEMP anomalies relative to 1951-1980. A common reference period 1981-2010 is used here. Global or regional averages shown as time series are first calculated for the conventional dataset using their original data values, and the time series are then adjusted to be relative to 1981-2010 averages. The times series for the reanalyses are calculated directly from anomalies relative to 1981-2010.

For maps of monthly and annual anomalies, values adjusted to be relative to 1981-2010 are shown only for grid squares where there are data for at least 90% of the months from 1981 to 2010, and only for grid squares at which a value is provided for every month of the year in question in the case of the annual averages. This results in a poorer geographical coverage than if values from the conventional datasets had been plotted relative to their native reference periods. This is particularly so for the monthly maps from HadCRUT4. The grid squares for which coverage is lost are indicated in the maps.

3 Time series of global and European average temperatures

Figure 1 shows time series of twelve-month running averages of the estimated global-mean temperature from 1979 onwards for ERA-Interim, JRA-55, the HadCRUT4 median and NOAAGlobalTemp. Each dataset provides a similar overall picture: the general warming since the late 1970s is not in doubt, nor is the occurrence of warmer and colder spells linked with El Niño events, volcanic eruptions, variations in sea-ice cover and other sources of variability. In each case, temperatures for the calendar year of 2015 are higher than for any earlier twelve-month period, but increasingly exceeded by the twelve-month averages ending in each of the first four months of 2016. Values for 2015 are around 0.45°C warmer than the 1981-2010 average, with slightly lower values from the reanalyses: 0.43°C from JRA-55 and 0.44°C from ERA-Interim, compared with 0.47°C from both HadCRUT4 and NOAAGlobalTemp. All datasets show above-average values from 2001 onwards, and for each of them the warmest sixteen calendar years are 1998 and 2001-2015. The averages for the twelve months to April 2016 are 0.57°C, for all four datasets.

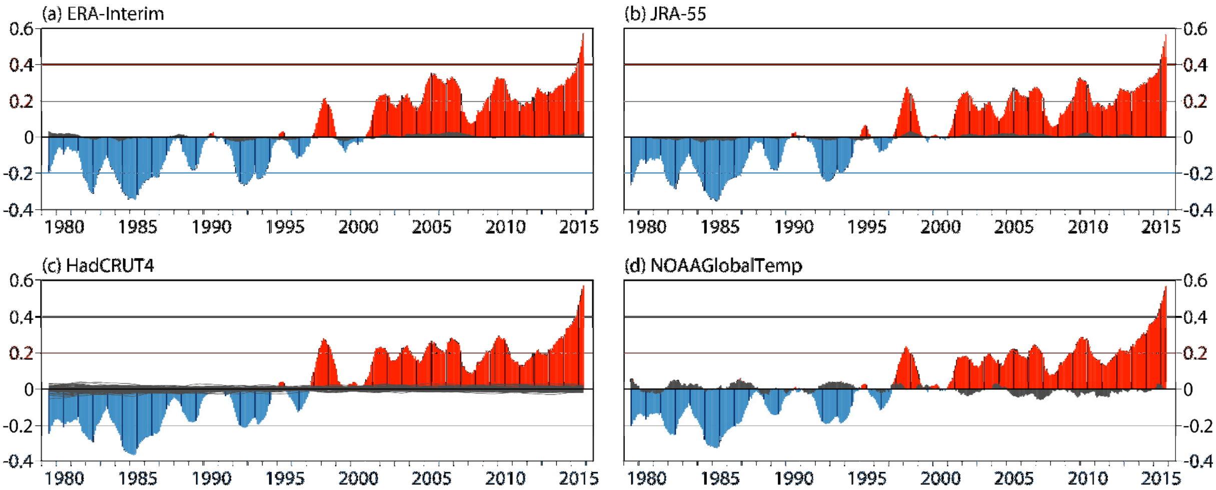

Figure 1 Twelve-month running means of anomalies in globally-averaged surface temperature (°C) relative to 1981-2010, for (a) ERA-Interim, (b) JRA-55, (c) the HadCRUT4 median and (d) NOAAGlobalTemp, based on data from January 1979 to April 2016. Red denotes above-average values and blue denotes below-average values. The darker coloured bars are the calendar-year means. ERA-Interim and JRA-55 values are based on air-surface (two-metre) temperature over sea; the small differences between these values and those using sea-surface temperature (as in HadCRUT4 and NOAAGlobalTemp) are shown in dark grey in panels (a) and (b). The overlapping dark grey lines in panel (c) denote the differences between the values of the 100 HadCRUT4 ensemble members and the HadCRUT4 median. The differences between NOAAGlobalTemp and GISTEMP are shown in grey in panel (d).

The datasets differ more considerably in their estimates of the magnitudes of individual warm and cold spells, and accordingly in their rankings of the set of warmest calendar years. ERA-Interim shows the largest peaks. Its average for the year 2014 is exceeded by averages for twelve-month periods within 2005–2006 and 2009–2010, including the calendar years of 2005, 2006 and 2010. JRA-55 has 2014 as the second warmest calendar year by a narrow margin, but has a slightly warmer twelve-month spell in 2009–2010. NOAAGlobalTemp shows the lowest maxima in the period from 1999 to 2013; 2014 is clearly the warmest calendar year prior to 2015 for this particular dataset. The temperature anomaly for 2005 ranges from 0.23°C for NOAAGlobalTemp to 0.35°C for ERA-Interim. The spread among datasets for this year is the largest for any calendar year in the period. The largest spread in the sets of twelve-month means is 0.14°C, for the mean from May 2005 to April 2006. Spreads above 0.1°C occur only in the early 1980s and the mid-2000s. Factors behind these differences are discussed in subsequent sections.

HadCRUT4, NOAAGlobalTemp and the related GISTEMP dataset comprise a combination of surface air temperature data over land and sea-surface temperature data. Most reanalysis results presented here are for surface air temperature over both land and sea, but the upper two panels of Figure 1 show in dark grey the differences between these global means and global means based on surface air temperature over land and SST otherwise. Differences are evidently small, but have a systematic component: they are slightly negative over most of the first half of the period and slightly positive over most of the second half, for both reanalyses. They can also shift the ordering of years according to their warmth: using SST rather than air temperature makes 2006 slightly cooler rather than warmer than 2014 for ERA-Interim. The somewhat larger trend in air than sea temperature is such as to reduce air-sea differences over time, and is consistent with the climate model results reported by Cowtan *et al.* (2015). Although small, the differences do need to be kept in mind when comparing reanalyses with the conventional surface-temperature datasets, as is the case also when models are compared with such datasets (Cowtan *et al.*, 2015).

The differences between the HadCRUT4 median and the 100 individual ensemble members (each expressed as an anomaly with respect to its own 1981-2010 average) are plotted as a set of largely overlapping grey lines in panel (c) of Figure 1. There is little variation over time in the spread of the ensemble for the period shown. The means for 2015 range from 0.45°C to 0.49°C; those for 2005 range from 0.25°C to 0.29°C. This is despite considerable differences between the two years in the level of agreement between all datasets. The spread of the HadCRUT4 ensemble is generally smaller than the spread of the alternative datasets over the period from January 1979. The root-mean-square spread of monthly values for this period is 0.05°C between ensemble members and 0.10°C between GISTEMP, the HadCRUT4 median, NOAAGlobalTemp and the two reanalyses. Morice *et al.* (2012) acknowledge that HadCRUT4 does not provide a full description of uncertainties, and advocate that users of the dataset test the robustness of their results by comparing with other datasets. In particular, all members of the HadCRUT4 ensemble have the same limited data coverage, so by themselves provide no information on uncertainties in global averages that arise from this. Morice *et al.* (2012) provide estimates of additional uncertainty, arising from measurement uncertainty, under-sampling within a grid box and lack of coverage, for their calculations of temporal and spatial averages. No such additional estimates are included in the results presented here.

Panel (d) of Figure 1 includes the differences between NOAAGlobalTemp and GISTEMP values, shown in grey. Although generally small compared to the variations of either, there is a systematic component to the differences, GISTEMP showing generally larger negative anomalies earlier in the period and larger positive anomalies later in the period. 2015 is an exception in that GISTEMP has a smaller anomaly of 0.44°C, the same as ERA-Interim. GISTEMP is generally closer to the reanalyses than HadCRUT4 and NOAAGlobalTemp are. This suggests that the additional geographical spreading of values that it provides is reasonably consistent with what is provided by the reanalyses, as will be seen in specific cases discussed later. The same holds for the extensions of HadCRUT4 provided by Had4_UAH_v2 and Had4_krig_v2, also illustrated later.

Figure 2 shows time series of temperatures averaged over European land areas. Variability is much higher for this considerably smaller domain, but the region is well observed, and all datasets are in good agreement. The reanalyses exhibit very slightly larger maxima and minima than the other datasets, as do HadCRUT4 and GISTEMP compared with NOAAGlobalTemp. For Europe, 2014 and 2015 are the warmest two calendar years on record, with little to separate them. Neither calendar year is quite as warm as the twelve months from the middle of 2006 to the middle of 2007, however.

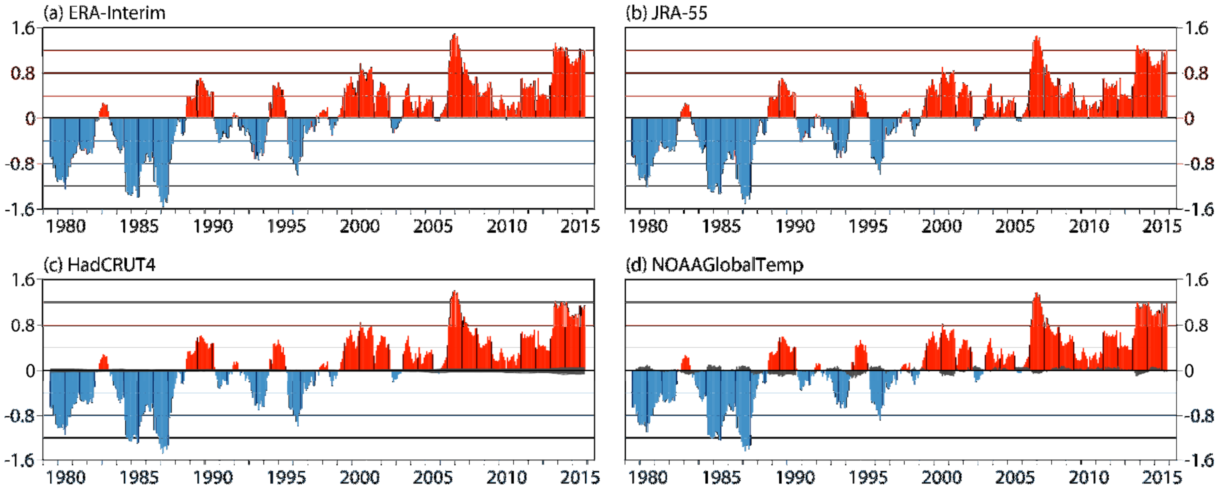

Figure 2 Twelve-month running-means of anomalies in surface air temperature over Europe (°C), relative to 1981-2010, for (a) ERA-Interim, (b) JRA-55, (c) the HadCRUT4 median, and (d) NOAAGlobalTemp, based on data from January 1979 to April 2016. Dark grey lines are plotted in panel (c) to denote the differences between the values of the 100 HadCRUT4 ensemble members and the HadCRUT4 median; they all lie close to the zero line. Differences between NOAAGlobalTemp and GISTEMP values are shown in grey in panel (d). Values are averages over land areas located between 20°W and 40°E, and 35°N and 80°N. The ERA-Interim land-sea mask is used to partition coastal grid-box values between land and sea.

4 Global temperature trends

Linear trends in global-mean surface temperature are presented in Figure 3. They have been computed by a least squares fit to monthly anomalies relative to 1981-2010; these anomalies are also shown in the figure. The recent warm spell is exceptional in the extent to which it deviates from the linear trend over the full period from January 1979 to April 2016, for all datasets. The deviation exceeds 0.5°C in February 2016 for ERA-Interim, JRA-55 and GISTEMP.

The full-period trends differ little among the datasets. ERA-Interim, JRA-55 and GISTEMP give warming rates of 0.17°C/decade, HadCRUT4 gives 0.18°C/decade while NOAAGlobalTemp gives 0.16°C/decade. Trends range from 0.16 to 0.19°C/decade for the HadCRUT4 ensemble, and Had4_UAH_v2 and Had4_krig_v2 both give 0.18°C/decade. The trends from both reanalyses are slightly smaller if sea-surface temperature rather than marine air temperature is used, rounding to 0.17°C/decade again for ERA-Interim but to 0.16°C/decade for JRA-55. Trends are slightly smaller still if background values of air temperature are used over land rather than the values from analysing screen-level observations: in this case the trends from both reanalyses round to 0.16°C/decade. Although the full period concludes at a time of extreme values, the largest 30-year trend within the period runs from 1982 to 2011 in five of the six datasets, and from 1981 to 2010 in the other.

The Fifth Assessment Report of the Intergovernmental Panel on Climate Change (AR5; IPCC, 2013) reported a warming rate of 0.05°C/decade with a 90% uncertainty range from -0.05 to 0.15°/decade for the fifteen-year period 1998-2012, compared with a rate of 0.12°C/decade with uncertainty range from 0.08 to 0.14°C/decade for the period 1951-2012. AR5 refers to this as the hiatus in global mean surface warming of the past fifteen years, although the use of words such as “hiatus” or “slowdown” has caused debate (Lewandowsky *et al.*, 2015; Fyfe *et al.*, 2016). Results from reanalyses were not used in computing these trends, and newer versions or replacements are now available for the three datasets (HadCRUT4, NOAA's MLOST (Smith *et al.*, 2008) and GISTEMP) that were used in the AR5 calculation.

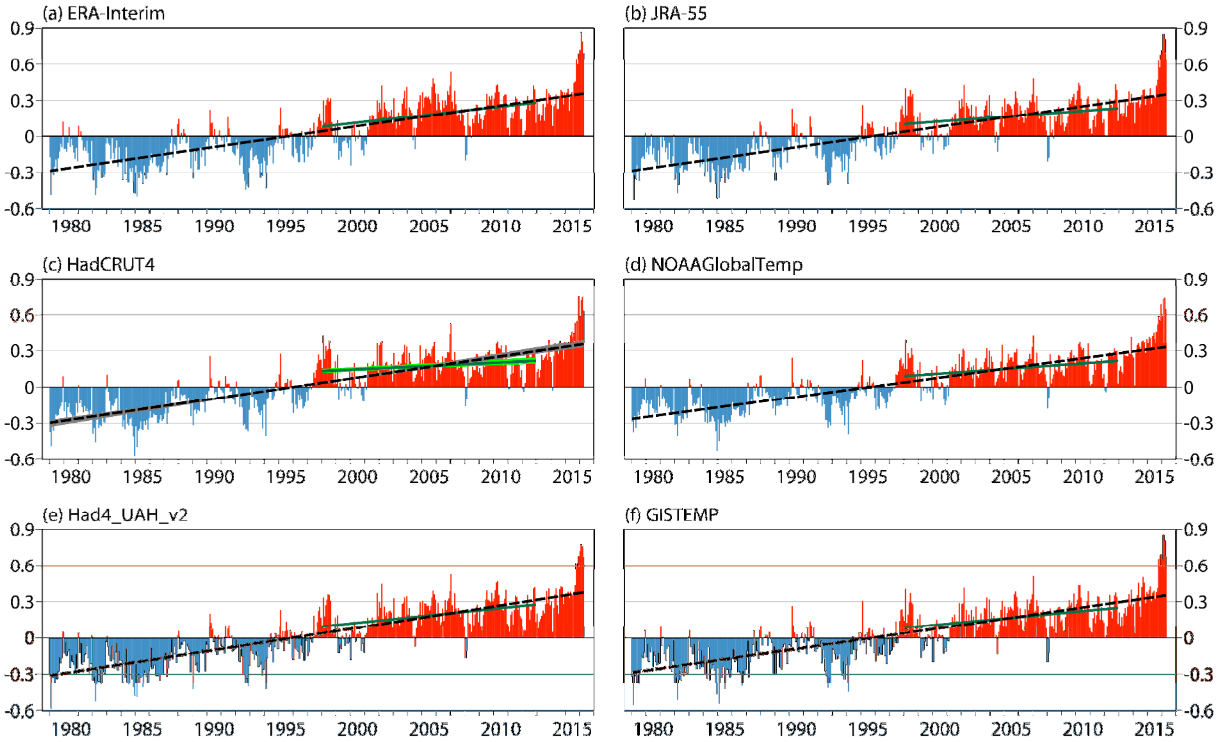

Figure 3 Monthly anomalies in globally-averaged surface temperature (°C) relative to 1981-2010, for (a) ERA-Interim, (b) JRA-55, (c) the HadCRUT4 median, (d) NOAAGlobalTemp, (e) Had4_UAH_v2 and (f) GISTEMP, from January 1979 to April 2016. Also shown are least-squares linear fits to the monthly values computed for the full period (black, dashed lines) and for 1998-2012 (dark green, solid line). In the case of HadCRUT4, the corresponding linear fits for each ensemble member are plotted as sets of (overlapping) grey and lighter green lines.

The datasets examined here unsurprisingly differ much more in their trends estimated for 1998-2012 than they do in their trends for the full period. Apart from some HadCRUT4 ensemble members, all datasets give 1998-2012 trends that are higher than AR5's central estimate, although all lie within its uncertainty interval. All are lower than the trend values for the full period, but most are within the range of the 1951-2012 trend quoted in AR5. The 1998-2012 warming rate from the HadCRUT4 median is 0.06°C/decade, close to the central estimate from the IPCC report, while the HadCRUT4 ensemble members range from 0.04°C/decade to 0.08°C/decade. MLOST gave a rate of 0.04°C/decade, but its NOAAGlobalTemp replacement provides a higher value of 0.08°C/decade. Karl *et al.* (2015) discuss the reasons for this higher estimate. The fifteen-year warming rates from the other datasets are higher still: 0.09°C/decade for JRA-55, 0.11°C/decade for GISTEMP and Had4_krig_v2, 0.12°C/decade for Had4_UAH_v2 and 0.14°C/decade for ERA-Interim. The ERA-Interim trend may be overestimated due to its use of an SST analysis that underestimated the warmth of the 1997/98 El Niño (Simmons and Poli, 2015), but the overestimate from this cause is likely to be by no more than 0.01°C/decade. This is based on repeating the calculations using SST rather than marine air temperature, and employing the more recent HadISST2 (version 2.1.1.0; Kennedy *et al.*, 2016, unpublished manuscript) and OSTIA (Donlon *et al.*, 2012) SST analyses that are being used in ERA5.

5 Differences in monthly values

Variations from month to month in global averages can be seen in Figure 3, along with the longer-term variability depicted in Figure 1. Most notable is the extremity of the warm temperature anomalies in late 2015 and early 2016. Values from the reanalyses and GISTEMP reach not far short of 0.9°C above the 1981-2010 norm in February 2016.

For all datasets, the warmest month (relative to its 1981-2010 norm) preceding the recent spell was January 2007, whose anomaly ranges from 0.43°C for NOAAGlobalTemp to 0.54°C for ERA-Interim. In each case, this earlier maximum was exceeded by the monthly temperature anomalies for all months from October 2015 to April 2016, by more than 0.3°C in the case of February 2016 for all but HadCRUT4 and its derivatives.

There are nevertheless quite pronounced variations between some of the datasets for particular months within the recent extreme spell. Table 1 documents values from October 2015 to April 2016 for all datasets. In this spell the HadCRUT4 dataset is the evident outlier. Its median and almost all its ensemble values are lower than the values from all other datasets for all months from October to February except December, when all ensemble values are above those from the other datasets. There is little doubt that this outlying behaviour is due to the dataset's limited spatial sampling, as the two spatially extended versions of HadCRUT4 are much closer to the other datasets, particularly the two reanalyses, which in turn are close to each other. It is also noteworthy that the more spatially complete GISTEMP dataset is closer to the reanalyses than is NOAAGlobalTemp.

	October 2015	November 2015	December 2015	January 2016	February 2016	March 2016	April 2016
ERA-Interim	0.64	0.60	0.69	0.72	0.86	0.78	0.69
JRA-55	0.63	0.57	0.68	0.72	0.85	0.81	0.70
HadCRUT4 median	0.54	0.55	0.75	0.57	0.72	0.75	0.64
HadCRUT4 ensemble	0.52-0.57	0.53-0.58	0.73-0.78	0.55-0.60	0.70-0.75	0.72-0.77	0.61-0.66
Had4_UAH_v2	0.62	0.59	0.70	0.72	0.77	0.75	0.67
Had4_krig_v2	0.63	0.61	0.71	0.72	0.78	0.76	0.65
NOAAGlobalTemp	0.59	0.56	0.69	0.58	0.73	0.74	0.65
GISTEMP	0.66	0.60	0.69	0.65	0.86	0.81	0.67

Table 1 Anomalies in global-mean surface temperature (°C) relative to 1981-2010 for the months of October 2015 to April 2016 from the datasets listed in the left column.

Results for May 2016 are available from ERA-Interim and JRA-55 at the time of writing. Anomalies for the month are 0.59°C and 0.56°C respectively. Twelve-month running means from June 2015 are

0.60°C and 0.59°C. Results for May 2016 are included in section 9, which is based only on reanalysis data.

6 Geographical coverage

Maps of the annual-mean temperature anomalies relative to 1981-2010 from six datasets are presented in Figure 4 for 2011 and 2015. The year 2011 was chosen in addition to the latest calendar year for several reasons. It was a year with an SST anomaly over the Pacific opposite to that in 2015, as can be clearly seen for all datasets in Figure 4. It was also the year with the most anomalously warm temperature averaged north of 60°N in the datasets with complete coverage, and the year with the largest spread in estimates of global means since the 2005-7 period that is discussed further in the following section.

The Arctic void in data from HadCRUT4 and NOAAGlobalTemp covers only a small part of the globe, but it is the region where the largest temperature anomalies occur in the reanalyses, associated with anomalous winter sea-ice conditions, as illustrated by Simmons and Poli (2015). Arctic temperatures much above normal occurred early and late in the year 2011. Summer temperatures were also relatively warm, though not as warm as in 2007 and 2012, the only years with lower minimum Arctic sea-ice extent according to the Sea Ice Index of the US National Snow and Ice Data Center (NSIDC; http://nsidc.org/arcticseaicenews). The Arctic warmth of 2011 as a whole is shown clearly by the maps for the reanalyses in Figure 4. Warm temperatures are also indicated by HadCRUT4 where values are available. NOAAGlobalTemp provides fewer data values over north-western Russia: for the island of Novaya Zemlya and around the Kara Sea to the east. ERA-Interim has been shown by Simmons and Poli (2015) to fit well the wintertime synoptic data from stations in this region. Nearby values from NOAAGlobalTemp are less anomalous than the values from the reanalyses and HadCRUT4. These differences carry over into the corresponding spatially extended datasets: GISTEMP has less anomalously warm temperatures than Had4_UAH_v2 over the Arctic. The same is seen for 2015.

Elsewhere the patterns and amplitudes of the temperature anomalies shown in Figure 4 are in generally good agreement where observational coverage is good. ERA-Interim and JRA-55 differ most over western and southern Africa, over South America (where it is ERA-Interim that is the more consistent with the conventional datasets) and over Antarctica. HadCRUT4 and NOAAGlobalTemp do not provide values over sea-ice off the coast of Antarctica, and the spatially more-extensive Had4_UAH_v2 and GISTEMP datasets produce weaker anomalies in this region than the reanalyses, which are reasonably consistent in their depictions of temperature anomalies that can be linked to anomalies in sea-ice cover.

The datasets also differ in their resolution of SST anomalies, most evidently that over the tropical Pacific Ocean associated with the El Niño in 2015. The two reanalyses provide a much sharper picture than NOAAGlobalTemp and GISTEMP, which as noted earlier use the same SST analysis. Although supplied on the same 5°x5° grid as NOAAGlobalTemp, HadCRUT4 (where it provides data) and the extended Had4_UAH_v2 provide a more-detailed depiction of SST anomalies, closer to that of the reanalyses.

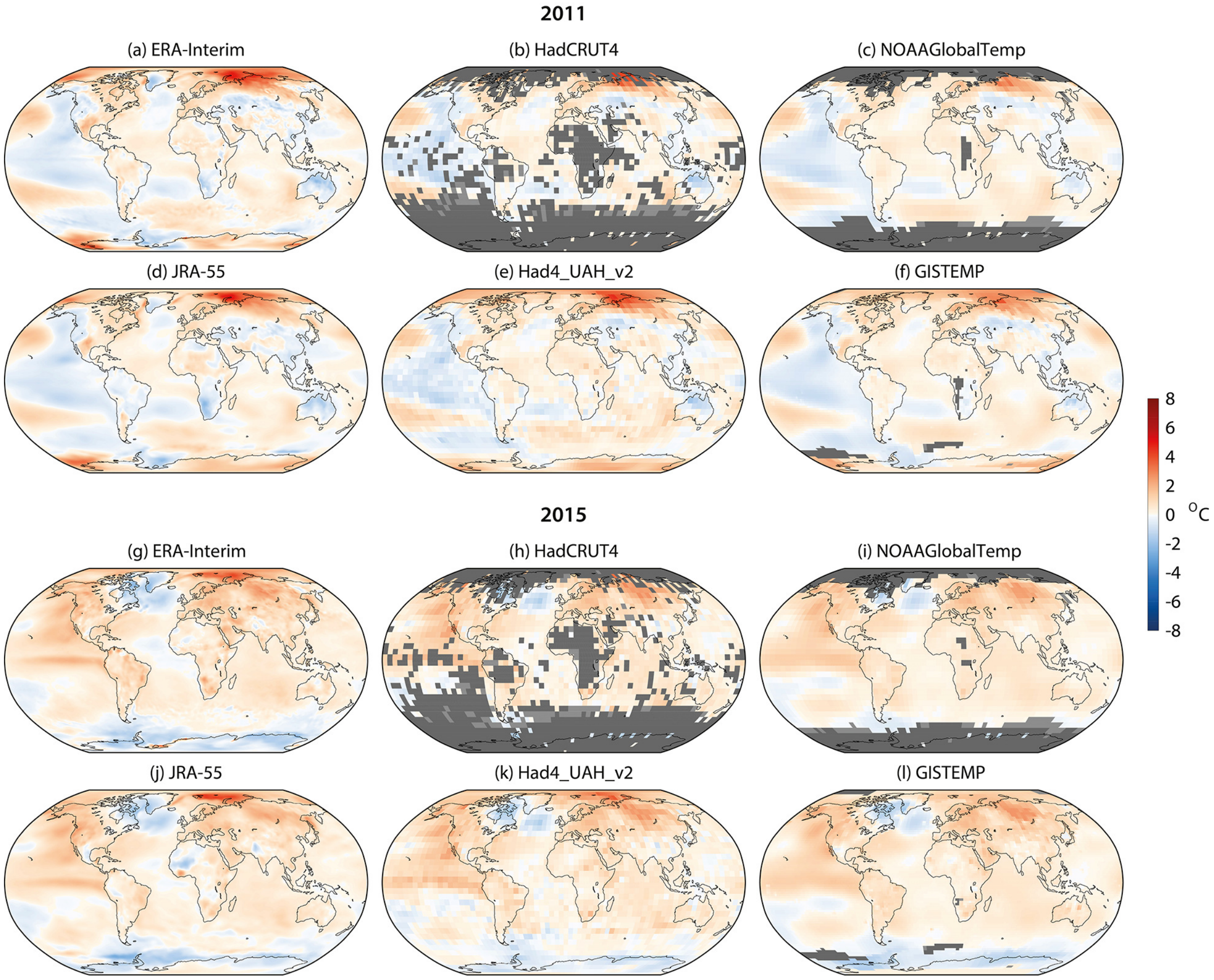

Figure 4 Surface temperature anomalies (°C) relative to 1981-2010 for 2011 (upper) and 2015 (lower), from (a, g) ERA-Interim, (b, h) HadCRUT4, (c, i) NOAAGlobalTemp, (d, j) JRA-55. (e, k) Had4_UAH_v2 and (f, l) GISTEMP. Grid boxes where values are missing are coloured grey. Lighter grey colouring indicates boxes that would have had values had maps been presented as anomalies relative to their standard reference periods.

Differences from monthly climatological averages for 1981-2010 are illustrated in Figure 5 for December 2015 and January and February 2016. The datasets are in overall agreement for the three months, showing relatively warm conditions throughout over the tropical and sub-tropical oceans, though with declining amplitude over the Pacific as time progresses, and persistent warm conditions over most of South America and southern Africa. Relatively warm conditions also persist over the Barents and Kara Seas and over the adjacent Arctic Ocean to the north, where winter sea-ice cover was unusually low, as indicated either by the datasets used by ERA-Interim and JRA-55 or by the NSIDC Index. Elsewhere, although the winter is predominantly less cold than average at middle and high northern latitudes, exceptionally so for some regions and months, there is also pronounced month-to-month variability. Temperatures shift from above to below and then again to above average over eastern Europe for example, and below-average temperatures over parts of the Arctic in December give way to above-average temperatures in January.

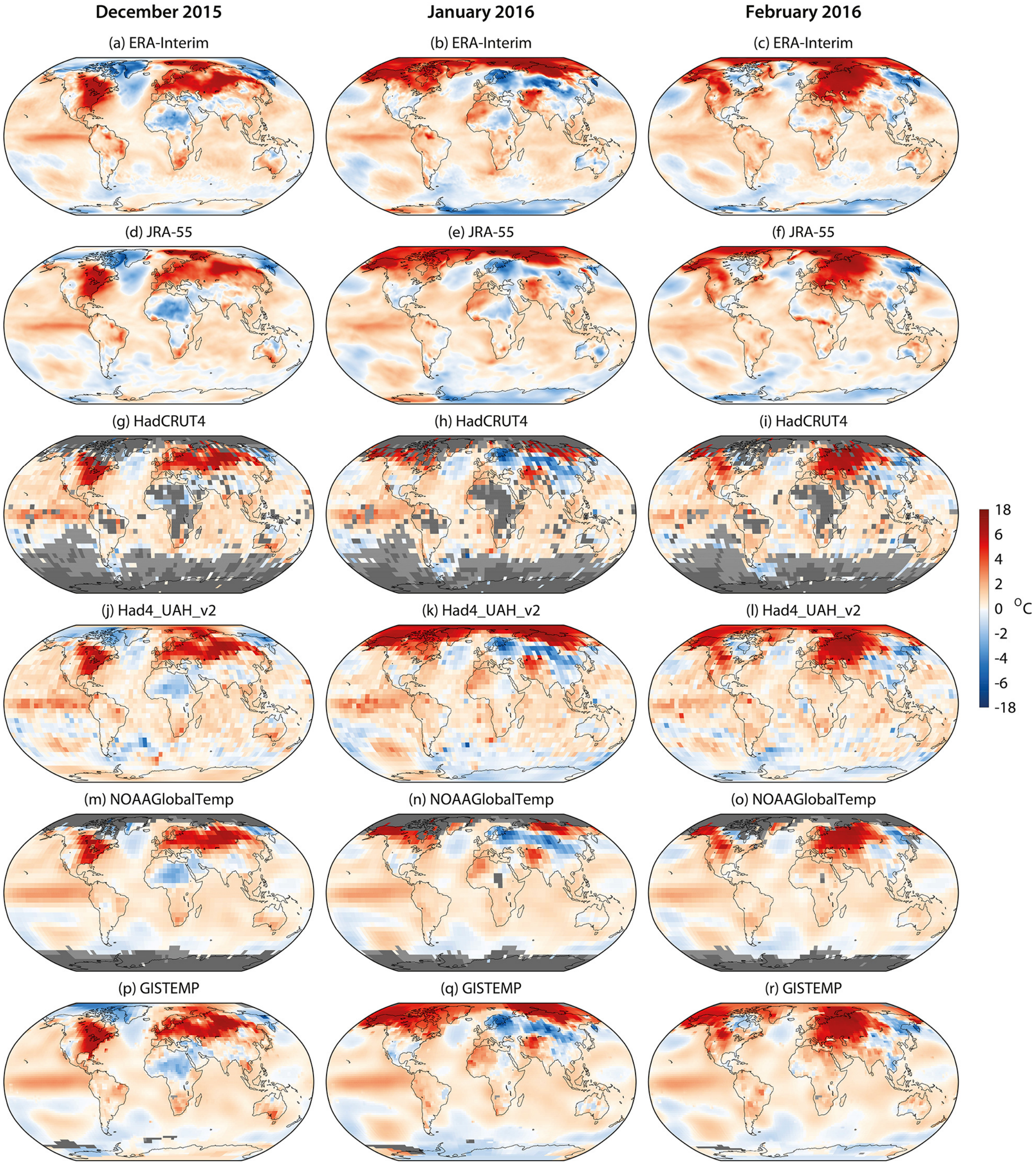

Figure 5 Surface temperature anomalies (°C) relative to 1981-2010 for December 2015 (left), January 2016 (centre) and February 2016 (right), from (a, b, c) ERA-Interim, (d, e, f) JRA-55, (g, h, i) HadCRUT4, (j, k, l) Had4_UAH_v2, (m, n, o) NOAAGlobalTemp and (p, q, r) GISTEMP. Grid boxes where values are missing are coloured grey. Lighter grey colouring indicates boxes that would have had values had maps been presented as anomalies relative to their standard reference periods.

The comments made on the resolution provided by the various datasets with regard to the annual-mean maps apply also to the monthly maps. The HadSST3 analysis used in HadCRUT4 is evidently prone to produce local values for grid squares that stand out from neighbouring values, and these are inherited by Had4_UAH_v2. These generally do not have counterparts in the fields from the reanalyses. It is beyond the scope of this report to evaluate comprehensively the local detail provided by the reanalyses over land, which is not always in good agreement between the two, but the spatial and temporal variations that are sharper in the reanalyses than in the other datasets over Australia compare reasonably with the anomalies in mean monthly temperatures reported routinely by the Bureau of Meteorology at http://www.bom.gov.au/climate/current/.

The principal differences among datasets in the recent monthly global-mean temperatures shown in Table 1 can be appreciated qualitatively from the differences in spatial coverage of the datasets shown in Figure 5. HadCRUT4's warmer mean values in December are consistent with it missing cold anomalies over northern Africa and a quite substantial part of the Arctic. Colder values from both HadCRUT4 and NOAAGlobalTemp in January and February are consistent with them missing above-average temperatures over much of the Arctic. The spatial extensions provided by GISTEMP and Had4_UAH_v2 appear to work well as judged by comparison with the values provided through the radically different approach of reanalysis.

7 Contributions to global means from polar and other regions

7.1 Averages over the polar regions

A more quantitative identification of differences is provided by time series of temperature anomalies for various regions. Figure 6 shows twelve-month running means for the polar regions north of 60°N and south of 60°S where most of the sea-ice cover occurs. These regions are referred to here simply as the Arctic and Antarctic; results differ little if the boundaries are placed at 65° latitude. The full averages over these domains for ERA-Interim and JRA-55 in the upper panels of Figure 6 show the substantial Arctic warming that was examined for these reanalyses by Simmons and Poli (2015) for the years up to 2013. The averages for HadCRUT4 and NOAAGlobalTemp based on their partial coverage of the Arctic are lower for much of the period from around 2005 onwards. Corresponding values for the Antarctic show little long-term change.

The middle panels of Figure 6 compare instead the full-domain averages of ERA-Interim and JRA-55 with GISTEMP and Had4_UAH_v2. The better coverage of GISTEMP and Had4_UAH_v2 compared with NOAAGlobalTemp and HadCRUT4 brings better agreement with the reanalyses for the Arctic, especially later in the period and more so for Had_UAH_v2 than GISTEMP. GISTEMP and Had4_UAH_v2 also improve agreement for the Antarctic for much of the period, though not for the latest years, when the two are close to each other, but less so to the reanalyses.

The lower panels of Figure 6 provide corresponding results for CRUTEM4 (Jones *et al.*, 2012), the land component of HadCRUT4, and for the reanalyses when sampled only at the Arctic and Antarctic grid squares where CRUTEM4 provides values. CRUTEM4 is chosen because it provides more data values over land south of 60°S than NOAAGlobalTemp, and values only for grid squares that include monthly station data. ERA-Interim and JRA-55 are evidently both in quite good agreement with CRUTEM4 when sampled in this way, more so for the less data-sparse Arctic. ERA-Interim is closer to CRUTEM4 than JRA-55 is for the Antarctic.

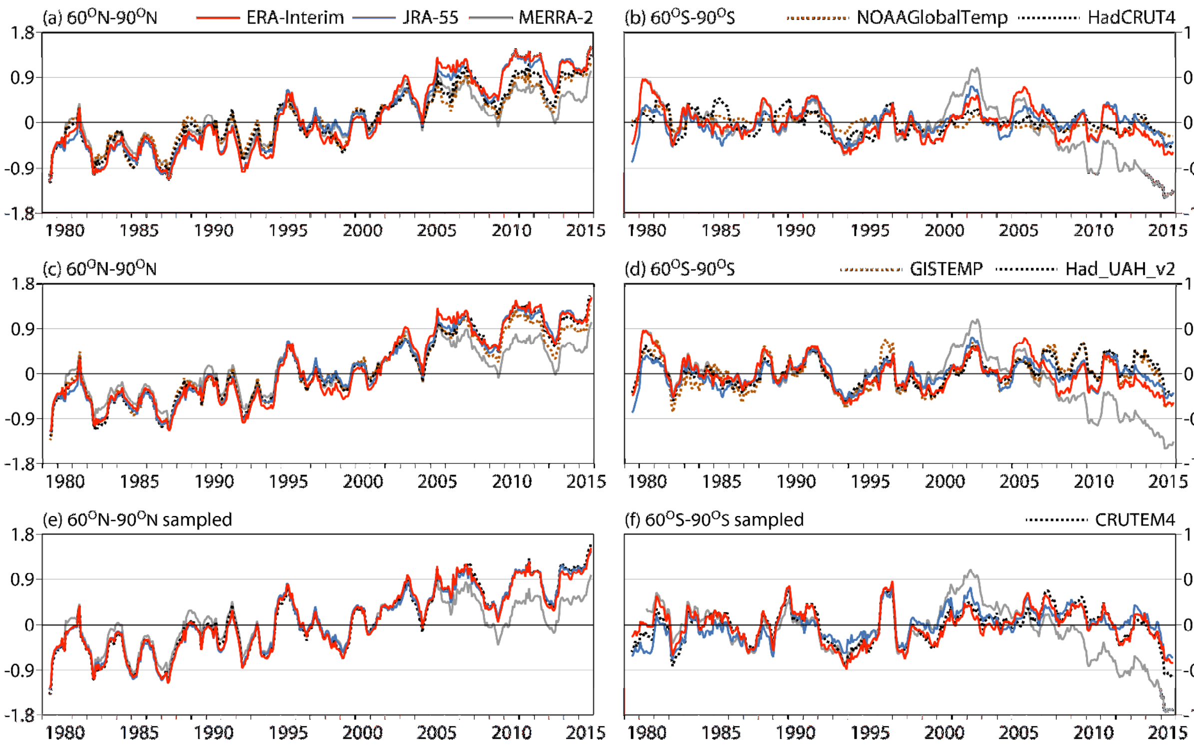

Figure 6 Twelve-month running-mean surface temperature anomalies relative to 1981-2010 (°C) based on data from January 1979 to April 2016, for (a) HadCRUT4 (black, dotted) and NOAAGlobalTemp (orange, dotted) averaged over all grid boxes from 60°N-90°N where they provide values, and for ERA-Interim (red, solid), JRA-55 (blue, solid) and MERRA-2 (grey solid) averaged over the whole 60°N-90°N domain, (b) as (a) but for 60°S-90°S, (c) and (d) as (a) and (b), but showing Had4_UAH_v2 (black, dotted) and GISTEMP (orange, dotted), (e) CRUTEM4 (black, dotted; version 4.4.0.0), ERA-Interim, JRA-55 and MERRA-2 averaged over all grid-boxes from 60°N-90°N where CRUTEM4 provides values, and (f) as (e) but for 60°S-90°S.

Figure 6 also includes results from the MERRA-2 reanalysis. Although the most recent of the reanalyses studied here, it is an evident outlier, providing values that are colder relative to 1981-2010 averages than are provided by all other datasets from around 2005 onwards, for both the Arctic and the Antarctic. Similar behaviour for the Arctic was shown by Simmons and Poli (2015) for the predecessor reanalysis, MERRA. Maps show that the Antarctic cooling in the later years of MERRA-2 is associated with a shift to colder values around the coastline of Antarctica and over the offshore region that is ice-covered in winter. MERRA-2 does not benefit from an analysis of synoptic surface air temperature observations such as is used in ERA-Interim and JRA-55, but this screen-level analysis adds relatively little for these two reanalyses, as their background fields are close to the analysis fields. This is illustrated by Simmons and Poli (2015) for the Arctic in the case of ERA-Interim.

7.2 Comparison of ERA-Interim with Antarctic station values

Figure 7 presents direct comparisons of the ERA-Interim background with observed values from the six Antarctic stations (specifically, stations with WMO identifiers greater than 89000) for which ERA-Interim has access to data for every month from 1979 to 2015. The information comes from ERA-Interim's 4D-Var data assimilation system. Although this system does not assimilate two-metre temperature data, these data are passed through it passively, and a record is kept of how well they are fitted by the background forecast. The 4D-Var operates over a twelve-hour period, and observations are compared with background values to within fifteen minutes of their reported time, although only observations for the standard synoptic hours of 00, 06, 12 and 18UTC are used to produce the averages shown in Figure 7. This reduces a possible effect of changes in frequency of reporting, which in general has increased over time (GCOS, 2015).

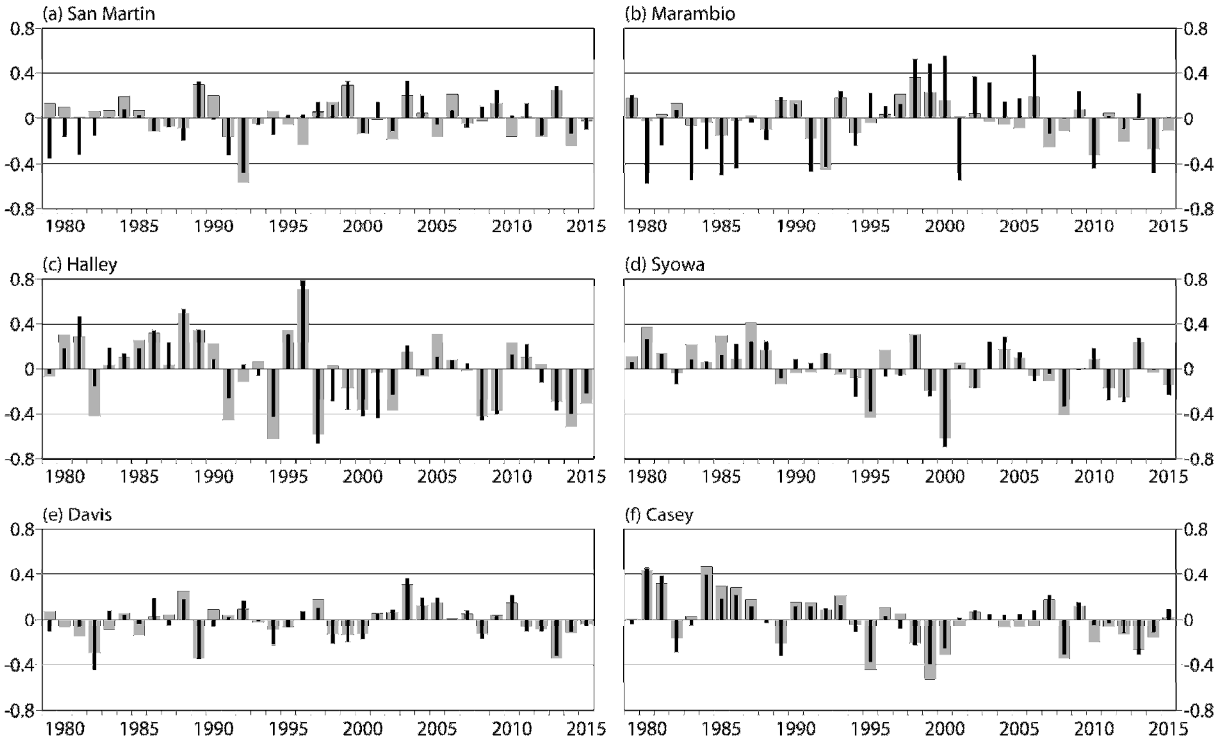

Figure 7 Annual-mean surface air temperature anomalies (°C) relative to 1981-2010, at (a) San Martin (68.1°S, 67.1°W), (b) Marambio (64.2°S, 56.7°W) (c) Halley (75.6°S, 26.7°W), (d) Syowa (69.0°S, 39.6°E), (e) Davis (68.6°S, 78.0°E) and (f) Casey (66.3°S, 110.5°E). Narrow, darker bars denote observed station values and broader, lighter bars denote corresponding ERA-Interim background forecast values.

Annual-mean anomalies relative to 1981-2010 are what is shown in Figure 7. This masks the mean error of the background, which ranges from 3.6 to -6.2°C for these stations, all of which are close to the Antarctic coast. Fréville *et al.* (2014) discuss a warm bias in ERA-Interim land- and air-surface temperatures over the Antarctic Plateau.

ERA-Interim has an understandable problem in reproducing the observations from one of the six stations, Marambio. This station is located on an island close to and east of the northern limit of the Antarctic Peninsula. The topography of the region is not resolved by the assimilating model used by ERA-Interim, which treats the location as a sea point that generally has only partial ice cover. As a result, the background temperature lacks the interannual variation that occurs in the observations, as is evident in Figure 7.

The ERA-Interim background captures the observed interannual variability quite well for the other stations. This provides some confidence in the variability described by the background elsewhere around the continent, as the surface air temperature observations provide essentially independent validating data for the region, as discussed for the Arctic by Simmons and Poli (2015). Trends over the period are mixed for these stations, and smaller in magnitude than the warming trends that predominate in the Arctic. Maps of the ERA-Interim trend, which are similar for the background and for the analysis, nevertheless show regions of strong warming, most notably over central West Antarctica, for which discussion is given by Bromwich *et al.* (2013) using observations available intermittently from Byrd Station. A cooling trend offshore of East Antarctica is consistent with an increase of sea-ice concentration there over recent decades, which is discussed in IPCC (2013).

7.3 Averages from 60^{o}N to 60^{o}S

Figure 8 complements the information provided for the Arctic and Antarctic by presenting information for the region from 60°N to 60°S, in this case separated into the contributions from land and sea to the average temperature for the zone as a whole. The datasets are in clear agreement as to the recent warmth of both land and sea, particularly so for land, for which the major discrepancy occurs for MERRA-2. This contributes together with the polar regions to cause MERRA-2 to underestimate substantially the recent warming seen in all the other datasets considered in this report. The only other difference worthy of note for the land is the provision of slightly cooler values for recent years by HadCRUT4, and by implication CRUTEM4. This comes mainly from limitations in spatial coverage, although a hint of it can be seen in the time series that are based on a common spatial sampling of the domain.

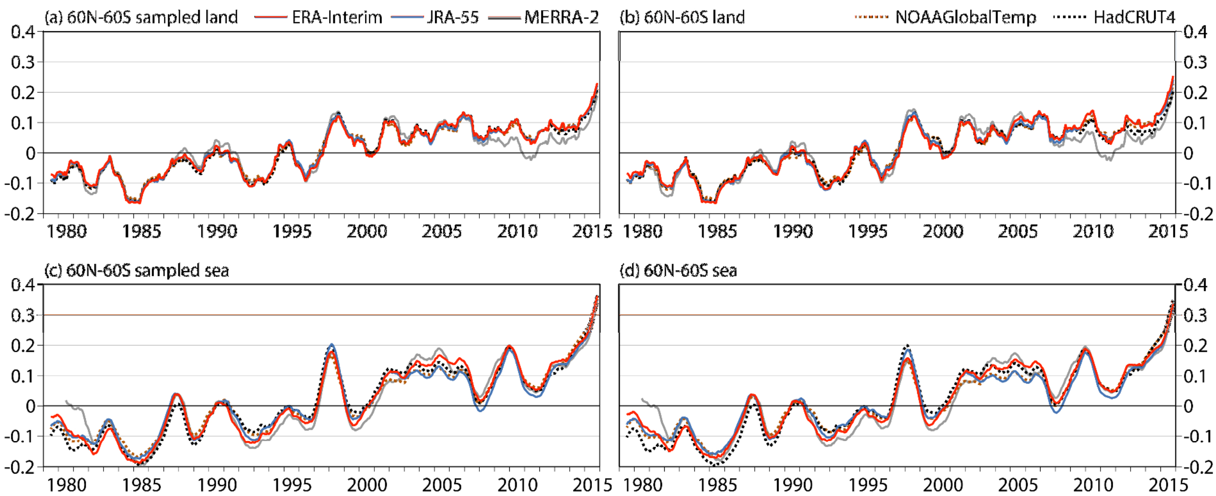

Figure 8 Contributions from land (a, b) and sea (c, d) to twelve-month running-mean surface temperature anomalies relative to 1981-2010 (°C) based on data from January 1979 to April 2016, for the region from 60°N-60°S. (a) and (c) are for HadCRUT4 (black, dotted), NOAAGlobalTemp (orange, dotted), ERA-Interim (red, solid), JRA-55 (blue, solid) and MERRA-2 (grey, solid) averaged over the sets of 5° grid boxes where all datasets provide values for land and sea respectively. (b) and (d) are for HadCRUT4 and NOAAGlobalTemp averaged over all grid boxes where each separately provides values, and for ERA-Interim, JRA-55 and MERRA-2 averaged over all land and sea.

Differences are larger over sea than land. MERRA-2 agrees quite well with the other datasets from 2009 onwards, when it uses the same OSTIA SST analysis as ERA-Interim, but it also exhibits shifts associated with changes in SST analysis, which are discussed by Bosilovich *et al.* (2015). Otherwise, HadCRUT4 has a slightly larger trend than ERA-Interim, while the COBE SST and the corresponding JRA-55 marine air temperature analysis have a slightly smaller one. There is a particular spread of marine values from 2003 to 2006, with ERA-Interim distinctly warmer than JRA-55 relative to their respective 1981-2010 means. Maps show the mean 2003-2006 differences in SST anomalies between ERA-Interim and JRA-55 to be geographically widespread and largely of the same sign. The values from NOAAGlobalTemp (and by implication GISTEMP) are also lower than those from ERA-Interim over this period. The differences in SST anomalies are also relatively large from 1979 to 1982.

Twelve-month running mean temperatures over sea continue to rise through to the end of the period shown in Figure 8. Although monthly anomalies peaked in December 2015, values for the first few months of 2016 were considerably higher than those for the corresponding months of 2015, or indeed for any other year. Expressed as a contribution to the complete 60°N-60°S average, these values are at least 0.2°C higher than at the times of preceding El Niño events. This is in contrast to the situation in the tropical eastern Pacific, where peak El Niño temperatures (specifically averages for the region from 5°N to 5°S and 180°W to 80°W) were a little lower in the 2015/16 event than in the 1997/98 event. More-generally warm SSTs, as well as the latest El Niño and low Arctic sea-ice cover, thus appear to contribute to the exceptional recent values of global-mean temperature.

7.4 Contributions to global means

Table 2 presents summary information for the globally complete ERA-Interim, JRA-55 and Had4_UAH_v2 datasets, and for GISTEMP, which has few missing values. It shows annual averages for the years 2005, 2006 and 2010-2015 of the global-mean surface temperature and of the contributions to the global mean from the regions 60°N-90°N, 60°N-60°S and 60°S-90°S. All these years have temperatures that are globally above the 1981-2010 average. The four datasets are in good agreement for 2010-2015, during which the contribution from the relatively warm Arctic to the global mean varies from around 50% in 2011 to 15% in 2015. The contribution from the Antarctic is small.

More difference is seen in 2005 and 2006, when the anomaly for ERA-Interim is larger than for any of the other datasets considered, as shown already in Figure 1. For these years, a significant contribution to the differences between ERA-Interim on the one hand and JRA-55 and GISTEMP on the other stems from the SST differences reported above. ERA-Interim also has larger warm anomalies in both the Arctic and the Antarctic at the time, as can be seen also in Figure 6. Although some further discussion is given in Section 10, further investigation seeking to establish which of the datasets is the most trustworthy in each of the domains is beyond the scope of this report.

8 Longer-term data records

The preceding discussion has focused on the period since 1979, and the recent warm spell in particular. ERA-Interim does not go back earlier. JRA-55 runs from 1958, but neither reanalysis matches the length of record of the established conventional datasets. Century-scale reanalyses that assimilate surface pressure and in some cases wind observations but no other meteorological data are available, but their agreement over land with CRUTEM4 has been shown to be poorer on annual and longer timescales than that of atmospheric model simulations using similar SST analyses and external forcings (Hersbach *et al.*, 2015).

		2005	2006	2010	2011	2012	2013	2014	2015
Global	ERA-Interim	0.35	0.29	0.31	0.19	0.22	0.25	0.29	0.44
	JRA-55	0.27	0.22	0.29	0.18	0.21	0.25	0.30	0.43
	GISTEMP	0.26	0.21	0.29	0.18	0.21	0.22	0.32	0.44
	Had4_UAH_v2	0.29	0.23	0.32	0.20	0.22	0.24	0.31	0.44
60°N-90°N	ERA-Interim	0.09	0.07	0.08	0.09	0.08	0.05	0.08	0.07
	JRA-55	0.08	0.06	0.08	0.09	0.08	0.05	0.08	0.07
	GISTEMP	0.07	0.05	0.07	0.08	0.06	0.03	0.06	0.06
	Had4_UAH_v2	0.07	0.05	0.08	0.09	0.08	0.04	0.07	0.07
60°N-60°S	ERA-Interim	0.22	0.21	0.25	0.08	0.15	0.21	0.24	0.41
	JRA-55	0.18	0.16	0.23	0.06	0.13	0.20	0.23	0.39
	GISTEMP	0.16	0.17	0.22	0.08	0.15	0.17	0.25	0.42
	Had4_UAH_v2	0.20	0.19	0.24	0.08	0.14	0.17	0.23	0.39
60°S-90°S	ERA-Interim	0.04	0.02	-0.02	0.02	-0.01	-0.01	-0.03	-0.04
	JRA-55	0.02	0.01	-0.02	0.03	0.00	0.00	-0.01	-0.03
	GISTEMP	0.02	-0.01	0.00	0.02	-0.01	0.02	0.00	-0.04
	Had4_UAH_v2	0.02	0.00	0.00	0.03	0.00	0.02	0.01	-0.02

Table 2 Anomalies in global-mean surface temperature (°C) relative to 1981-2010 for the years 2005, 2006 and 2010-2015 from ERA-Interim, JRA-55, GISTEMP and Had4_UAH_v2, and the contributions from the zonal bands 60°N-90°N, 60°N-60°S and 60°S-90°S.

The warming of the atmosphere since 1979 is substantially larger than the differences between the estimates provided by the various datasets examined, for both global and European averages. Appeal may thus be made to the behaviour of the longer-term datasets to establish which statements about warm extremes made on the basis of ERA-Interim and other datasets available only for the past few decades can be expected to hold over a much longer period. The HadCRUT4 dataset is particularly useful for this purpose, as it extends the farthest back in time, to 1850, and gives information on uncertainty through its ensemble of 100 possible realizations.

Figure 9 presents the time series to the present day of twelve-month running averages of global- and European-mean surface temperatures for the HadCRUT4 median from 1850 and NOAAGlobalTemp from 1880. Various time series of differences are also shown, relating to the HadCRUT4 ensemble and Had4_krig_v2, and to GISTEMP and JRA-55. It is evident that maxima prior to 1979 in all displayed datasets are at no time as large as the maxima that occur from the late 1980s onwards. The same is true of the model simulations and century-scale reanalyses examined by Hersbach *et al.* (2015). This makes it very likely that statements made concerning the extremity of the recent warmth based on ERA-Interim and other limited-duration data records apply for the whole period since 1850, at least as regards twelve-month averages.

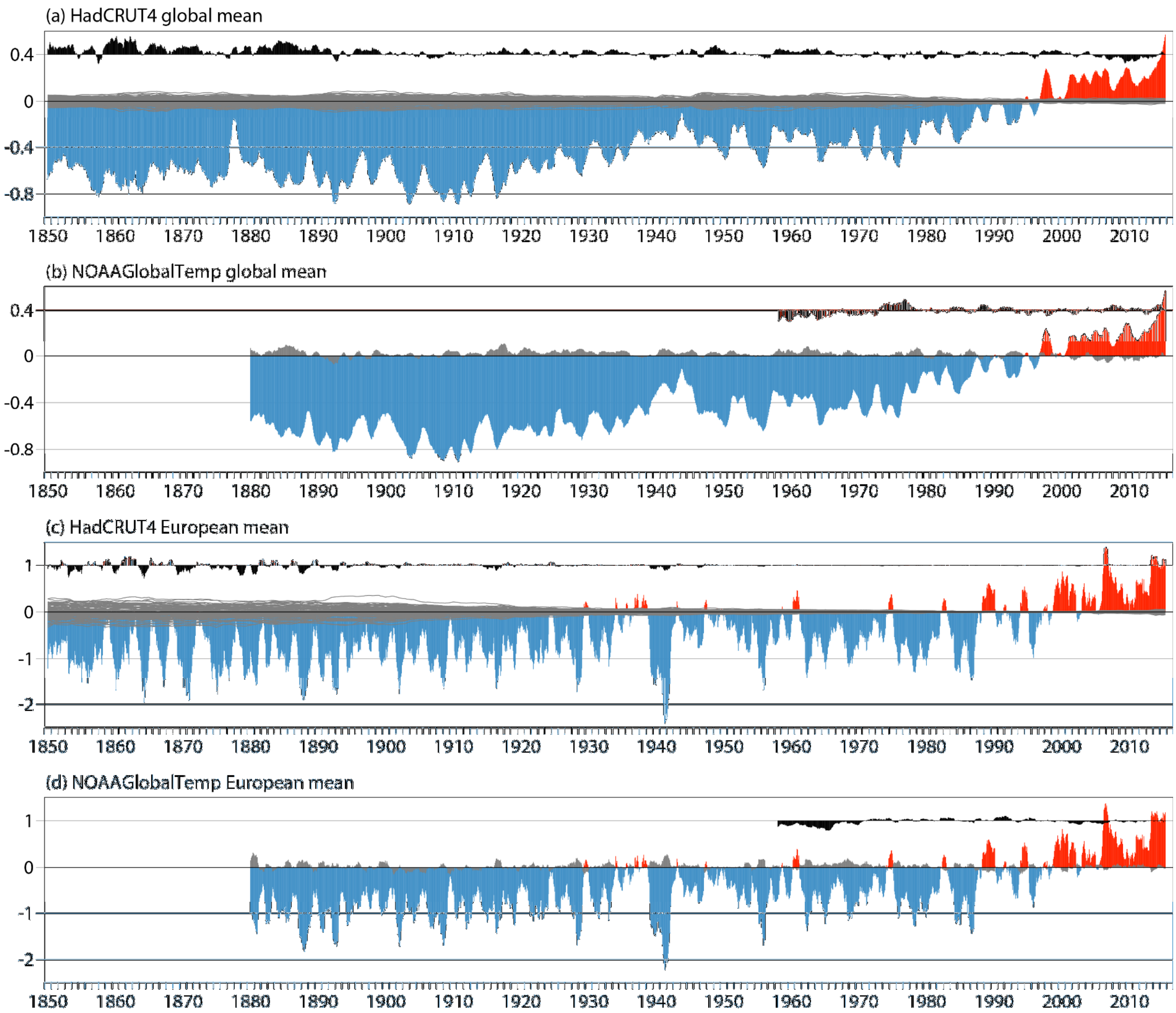

Figure 9 Twelve-month running means of anomalies in global-average (a, b) and European-average (c, d) surface temperatures (°C) relative to 1981-2010, for the full periods of record from HadCRUT4 (a, c) and NOAAGlobalTemp (b, d). The dark grey lines in panels (a) and (c) denote the differences between the values of the 100 HadCRUT4 ensemble members and the HadCRUT4 median. The differences between NOAAGlobalTemp and GISTEMP values are shown in grey in panels (b) and (d). The differences between HadCRUT4 and Had4_krig_v2 are shown in black in panels (a) and (c), shifted by 0.4°C and 1°C respectively for clarity, and differences between GISTEMP and JRA-55 are shown in a similar manner in panels (b) and (d).

NOAAGlobalTemp is generally similar to HadCRUT4 over the period of common record from January 1880 onwards. Both datasets show a spell of anomalously low global-mean temperature early in the 20th century, and relatively warm conditions in the first half of the 1940s, notwithstanding markedly below-average European temperatures during the early years of the Second World War. Relative to their respective 1981-2010 means, NOAAGlobalTemp is on average colder than the HadCRUT4 median for earlier years, and GISTEMP is colder still. The mean differences between 1880-1980 and 1981-2010 averages are -0.49°C for HadCRUT4, -0.50°C for Had4_krig_v2, -0.54°C for NOAAGlobalTemp and -0.56°C for GISTEMP. The global-mean temperature anomalies averaged for 1911-1940 and 1941-1970 from NOAAGlobalTemp and GISTEMP are in each case colder than the coldest corresponding 30-year average from the HadCRUT4 ensemble. Differences are much smaller for 1881-1910.

The global means from the two datasets with increased global coverage, GISTEMP and Had4_krig_v2, are mostly a little colder than their respective NOAAGlobalTemp and HadCRUT4 equivalents. Differences are larger in the earlier years when observational coverage is poorer, especially in the case of Had4_krig_v2 and HadCRUT4. However, in contrast to the quite similar changes they bring to the representation of the short-term variability of recent years, GISTEMP and Had4_krig_v2 remain closest in earlier years to the datasets that use the same SST analyses, respectively NOAAGlobalTemp and HadCRUT4.

Figure 10 illustrates this in maps of averages for the periods 1881-1910, 1911-1940 and 1941-1970 relative to 1981-2010. As data coverage is poorer for the earlier periods, values for a grid square are plotted here if there is at least 80% rather than 90% data availability for each period. The anomalies from NOAAGlobalTemp and GISTEMP for 1911-1940 and 1941-1970 can be seen to be generally slightly colder over the oceans than those from HadCRUT4 and Had4_krig_v2. For both periods, NOAAGlobalTemp is colder than HadCRUT4 when averaged over either the land or the sea areas where both provide data, but differences are larger over sea, even though the anomalies are larger over land. The datasets give more-similar averages for 1881-1910.

GISTEMP and Had4_krig_v2 also differ quite substantially in the earlier years over the regions where they perform substantial infilling. Figure 10 shows pronounced differences in the infilling over Africa and South America, with Had4_krig_v2 predominantly colder than GISTEMP over Africa but warmer over South America. These two datasets also differ over the Atlantic sector of the Arctic Ocean for 1911-1940.

Figure 10 includes corresponding plots for 1911-1940 and 1941-1970 from the mean of the ten-member ERA-20CM ensemble of atmospheric model integrations, which used prescribed time-varying SSTs, sea-ice distributions, solar radiative forcings and radiatively active trace gases and aerosols (Hersbach *et al.*, 2015). The ensemble averaging produces somewhat smoother temperature distributions than those of the observational datasets, but ERA-20CM temperatures over land are generally cooler, more so for 1941-1970. This is shown more fully in the comparison with CRUTEM4 presented by Hersbach *et al.* (2015). Over ice-free sea the ERA-20CM temperature anomalies are constrained by the geographically complete HadISST2 (version 2.1.0.0) analyses. As might be expected, the anomalies are closer to those of the HadSST3-based HadCRUT4 and Had4_krig_v2 datasets than those of the ERSSTv4-based NOAAGlobalTemp and GISTEMP datasets. ERA-20CM has distinctly colder temperatures in earlier than later decades over regions where HADISST2 indicates greater sea-ice concentrations in the earlier years. This includes the Antarctic, notwithstanding the observed increase in sea-ice since the 1970s. Such features in the temperature field are at least qualitatively in line with expectations, but are not reproduced by the methods of geographical extension used in deriving GISTEMP and Had4_krig_v2.

Europe has reasonably complete data coverage throughout the period of record of HadCRUT4. Figure 9 shows that the differences between HadCRUT4 and Had4_krig_v2 are accordingly small. This is particularly so from around 1900, although differences can be seen to be subsequently a little larger at the times of the two World Wars. There are other sources of uncertainty, however, which lead to a quite considerable spread among the HadCRUT4 ensemble for the 19th century and early decades of the 20th century. A single member stands out for almost thirty years from around 1890 in having a much warmer average temperature over Europe.

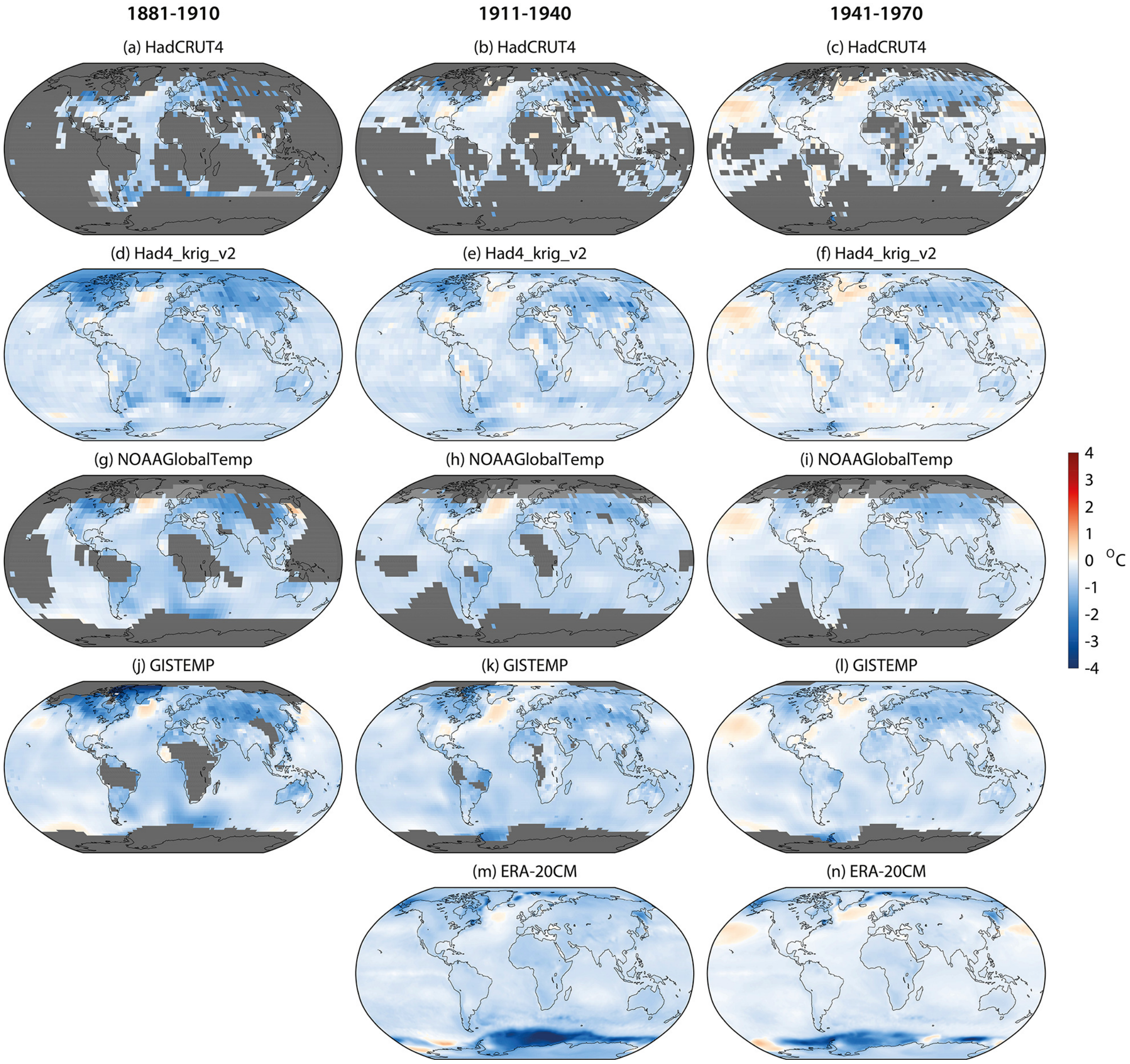

Figure 10 Surface temperature anomalies (°C) relative to 1981-2010 from (a, b, c) HadCRUT4 (d, e, f) Had4_krig_v2, (g, h, i) NOAAGlobalTemp and (j, k, l) GISTEMP for 1881-1910 (left), 1911-1940 (centre) and 1941-1970 (right), and from the ERA-20CM ensemble-mean for (m) 1911-1940 and (n) 1941-1970. Grid boxes where values are missing are coloured grey. Lighter grey colouring indicates boxes that would have had values had maps been presented as anomalies relative to the datasets' standard reference periods. Two small regions around Antarctica where cold anomalies exceed 4°C have been shaded at the 4°C level in panels (m) and (n).

The differences between the JRA-55 and GISTEMP global-mean anomalies shown in Figure 9 are larger before 1979 than afterwards, and the differences in the anomalies for Europe are relatively large in the 1960s. This is likely due both to the generally poorer global observing system available for reanalysis prior to 1979, and to specific gaps in coverage of the surface synoptic observations used by JRA-55 for the 1960s. The pre-1979 data used by JRA-55 were largely based on the collection of data made earlier for ERA-40, which lacked surface synoptic data from several countries prior to 1967, including European ones (Simmons *et al.*, 2004).

Time series of monthly values since 1850 from HadCRUT4 are shown in Figure 11. Only median values are shown, for clarity of display. For the global average it is again likely that statements concerning the extremity of the warmer periods of the last four decades in fact apply for the record back to 1850. However, it is evident for Europe that individual months can be relatively warm right back to 1850, even if the frequency of warm months is higher in recent decades, and the warmest few months occur over the past thirty years.

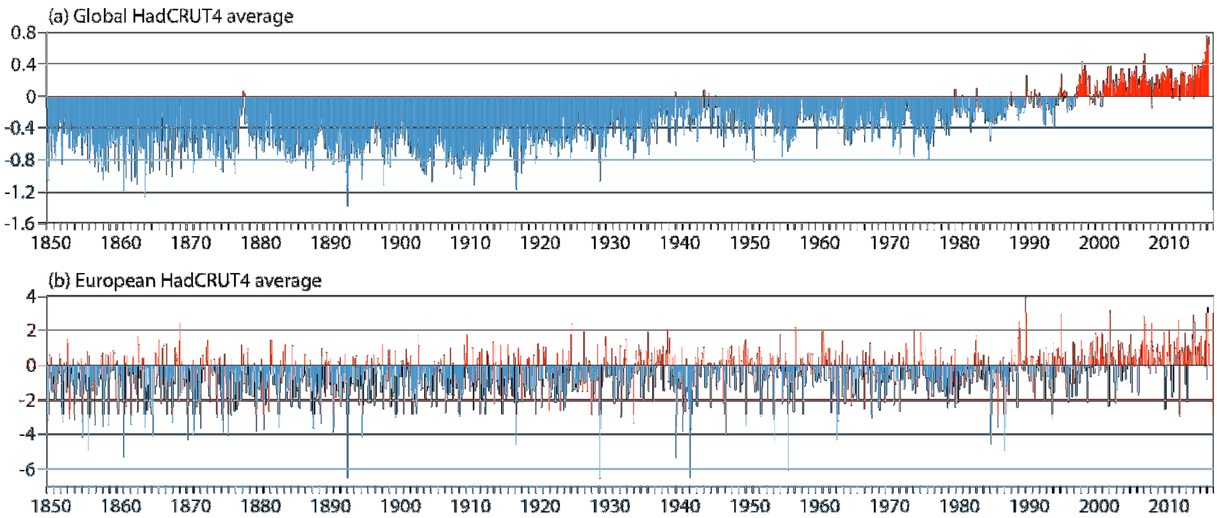

Figure 11 Monthly anomalies in (a) global-average and (b) European-average surface temperature (°C) relative to 1981-2010 from HadCRUT4.

The monthly extremes from HadCRUT4 are larger than those from NOAAGlobalTemp but smaller than those from GISTEMP. For example, the warmest monthly European temperature anomaly prior to 1989 is a little below 2°C in NOAAGlobalTemp, whereas HadCRUT4 identifies five warmer months between 1880 and 1988, and GISTEMP ten. HadCRUT4's warmest pre-1989 value, about 2.4°C for February 1869, is nevertheless larger than any of the 1880-1980 GISTEMP values. These figures help place in context the recent warm European anomalies seen in Figure 5, which are 3.2°C for December 2015 and 3.9°C for February 2016 from ERA-Interim. The corresponding anomalies from JRA-55 are 3.3°C and 3.7°C. HadCRUT4 gives smaller values, 3.0°C and 3.4°C, but its lower spatial resolution has to be taken into account: ERA-Interim anomalies are reduced to 2.9°C and 3.4°C respectively if an equivalent of the 5°x5° HadCRUT4 dataset is first constructed from ERA-Interim, and then averaged over European land areas following what was done for HadCRUT4.

9 Atmospheric energy and upper-air temperature

The use of global-mean surface temperature as a metric of climate change, whether as a target of the Paris Agreement or as the measure of response to a doubling of atmospheric carbon dioxide in the concepts of equilibrium and transient climate sensitivity (IPCC, 2007; 2013), is open to some question because of its sensitivity to changes in winter sea-ice cover, which becomes larger as datasets resolve better the Arctic and Antarctic, and because it is not a direct measure of the energy of the atmosphere. However, with the oceans estimated to have absorbed about 93% of the increase in energy of the climate system between 1971 and 2010, and melting ice and warming land accounting for much of the remainder, the increase in energy of the atmosphere has been estimated to account for only about 1% of the increase in energy of the system as a whole (Box 3.1, Rhein *et al.*, 2013). This estimate for the atmosphere was based on the microwave temperature sounding record of Mears and Wentz (2009), which like the UAH record has limited vertical resolution. It also assumed a fractional increase in water vapour content as temperature increases, and neglected changes in potential and kinetic energy.

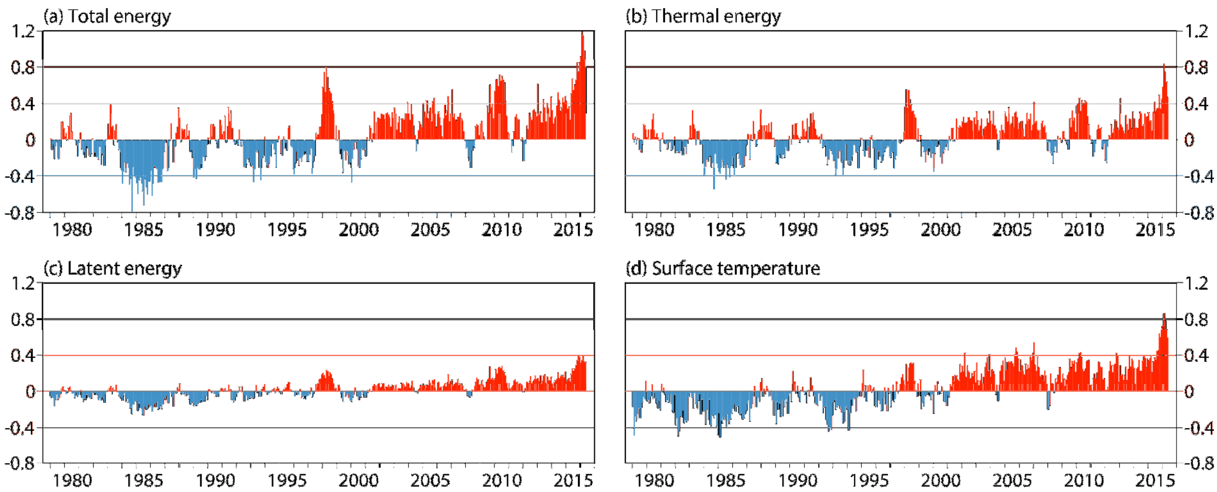

Figure 12 Monthly anomalies of atmospheric (a) total energy, (b) thermal energy and (c) latent energy, and of (d) global-mean surface air temperature (°C), relative to 1981-2010, for January 1979 to May 2016 from ERA-Interim. Latent and total energy, and surface temperature, are adjusted after 1991 and before 2002 respectively, as discussed in the text. The unit for energy is 10^7 Joules divided by the area of the Earth in m^2.

Reanalysis data enable a complete calculation of changes over time in atmospheric energy, though subject to provisos as to the temporal consistency of its estimation of the underlying state variables. Figure 12 shows time series of the anomalies relative to 1981-2010 in total energy and in its dominant thermal and latent energy components, from ERA-Interim (Berrisford *et al.*, 2011). The corresponding time series of global-mean surface air temperature is included for comparison. In addition to the dataset adjustments described previously for surface air temperature, the latent energy and its contribution to total energy have been increased by a factor 2.6% after 1991, based on the shift in ERA-Interim values relative to ERA-20CM caused by problematic assimilation of rain-affected radiances from 1992, a shift that is consistent with the differences in total column water vapour between ERA-Interim and retrievals from microwave imagery (Hersbach *et al.*, 2015; Poli *et al.*, 2016). The chosen units make the values of the anomalies in thermal energy close to those for the global mass-weighted average atmospheric temperature in °C. The change in total energy of 2TW estimated by Rhein *et al.* (2013) for the period from 1979 to 2010 is equivalent to a shift of about 0.4 over this period in the units used in Figure 12.

The anomalous spells seen in the various time series shown in Figure 12 are generally common to all series, but their magnitudes vary considerably. The anomalies in winter surface air temperature tend to be associated with relatively shallow structures in the vertical, and thus do not show as prominently in thermal energy. In contrast, variations in surface air temperature in the tropics tend to be associated with larger variations in upper tropospheric temperature, and thus with more pronounced features in thermal energy. Anomalies in latent energy are most pronounced in the tropics and subtropics. The El Niño events of 1997/98, 2009/10 and 2015/16 are thus much more marked for total energy than for surface air temperature. A tendency for tropospheric cooling due to the 1982 volcanic eruption of El Chichón counters warming due to the 1982/83 El Niño.

Broad peaks in latent energy occur during the 1997/98, 2009/10 and 2015/16 El Niño events. Latent energy can be sustained at anomalous levels in such events because of the greater capacity of an anomalously warm atmosphere to carry water vapour, but is converted to thermal energy in the declining phase of events as anomalous amounts of moisture are removed from the atmosphere by increased precipitation. ERA-Interim's representation of the latter over land for the 1997/98 and 2009/10 events has been shown by Simmons *et al.* (2014) to agree well with the independent analyses of precipitation produced by the Global Precipitation Climatology Centre; it is too early to confirm that the same holds for the latest event.

Variations in latent energy are found more generally to lead variations in thermal energy, in that correlations between the time series shown in Figure 12 are 77% and 73% with latent energy leading thermal energy by one and two months respectively, and 72% and 68% with latent energy lagging by one and two months, using the period from March 1979 to March 2016 for thermal energy. The zero-lag correlation between the time series is 79%. As is the case for surface air temperature, the annual-mean latent energy for 2015 is higher than for any previous twelve-month period, whereas the annual-mean thermal energy for 2015 is lower than for both 1998 and 2010. The twelve-month mean thermal energy first exceeds the earlier peak values in the mean for the twelve months that end in February 2016.

Figure 13 shows time series of monthly anomalies in temperature at 700, 500 and 300hPa from ERA-Interim and JRA-55. The larger amplitudes at higher levels of the perturbations associated with El Niño events are evident. Again in contrast to the situation for surface air temperature, recent 500 and 300hPa temperature anomalies exceed their previous maximum value, reached in 1998, only for two or three months, from February 2016.

Although the two reanalyses tell essentially the same story, interpretation of detail in their time series is hampered by differences between the two. ERA-Interim exhibits a weaker overall warming trend over the period at 700 and 500hPa; reasons why it likely underestimates trends in the lower and middle troposphere are discussed by Simmons *et al.* (2014). Conversely, it is JRA-55 that exhibits the weaker warming trend at 300hPa. This comes in particular from a drop in temperature in July 2006, when assimilation of substantial additional information on upper tropospheric and stratospheric temperatures from GPS radio occultation starts to reduce a warm bias that stems from the background model (Kobayashi *et al.*, 2015).

Figure 14 shows maps of the temperature anomalies at 700, 500 and 300hPa for December 2015 and January and February 2016, from ERA-Interim. The surface air temperature anomaly over the tropical Pacific Ocean reached peak amplitude in November 2015 and was slowly declining over the subsequent three months, as shown earlier in Figure 5. An extensive warm anomaly over the eastern tropical and sub-tropical Pacific can be seen at 300hPa in December, and warming of this region continues into January at all levels shown. Warmth is more widespread geographically by February. This can occur both directly through advection of sensible heat and through advection and remote release of latent heat. The shallower nature of the middle- and high-latitude anomalies that are predominant in the surface maps is also evident in Figure 14.

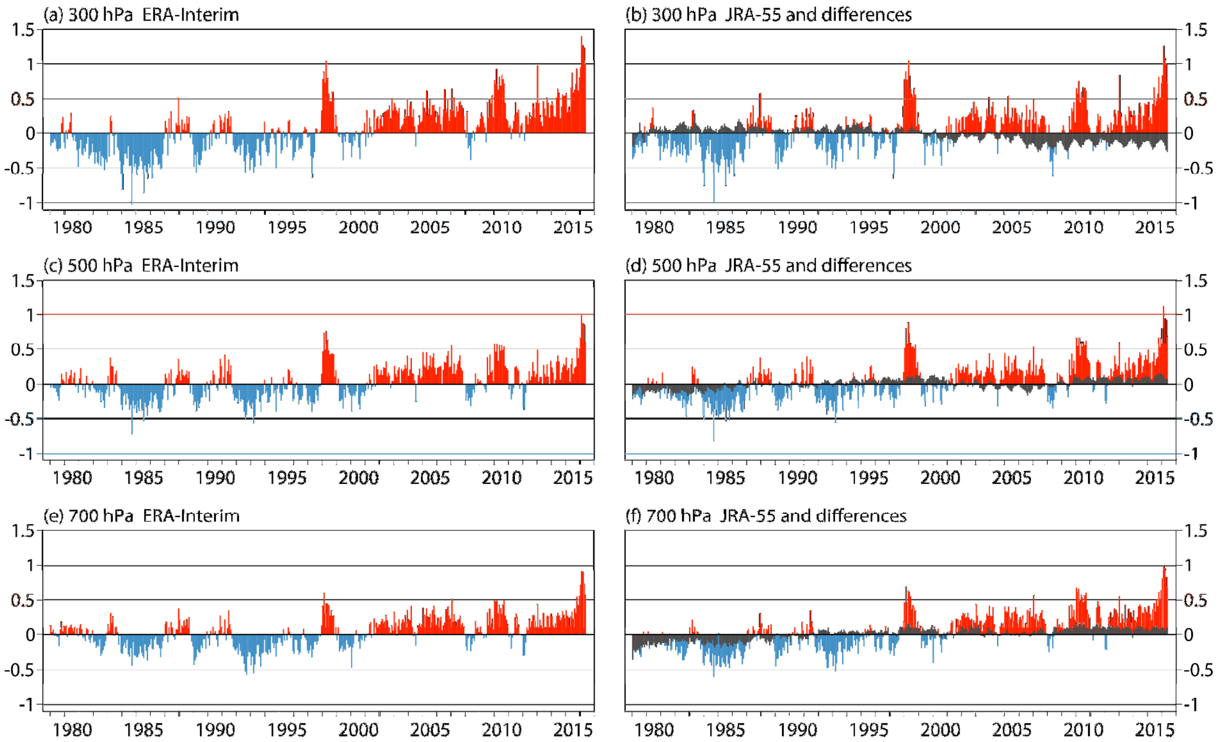

Figure 13 Monthly anomalies in globally-averaged temperature (°C) relative to 1981-1990 at (a, b) 300hPa, (c, d) 500hPa and (e, f) 700hPa for (a, c, e) ERA-Interim and (b, d, f) JRA-55, from January 1979 to May 2016, denoted by red and blue bars. Grey bars in (b, d, f) denote differences between JRA-55 and ERA-Interim.

10 Concluding discussion

The latest versions of several well-established conventional datasets and two recent reanalyses have been shown to agree well in their depiction of the net warming that has taken place at the Earth's surface over the past three to four decades. They agree also on the general character of the variability that has occurred over this period, and on the extremity of global warmth that has occurred recently as a strong El Niño, more-generally warm sea-surface temperatures and low Arctic sea-ice cover have raised atmospheric temperatures during the boreal winter of 2015/16. The versions of the conventional datasets that infill data where they lack direct observational sources agree better with the reanalyses. Giving weight to these, the evidence indicates that based on norms for 1981-2010, twelve-month running averages of global-mean surface temperature anomalies have reached around 0.6°C during the current warm spell, and monthly anomalies peaked in the range 0.8-0.9°C early in 2016.

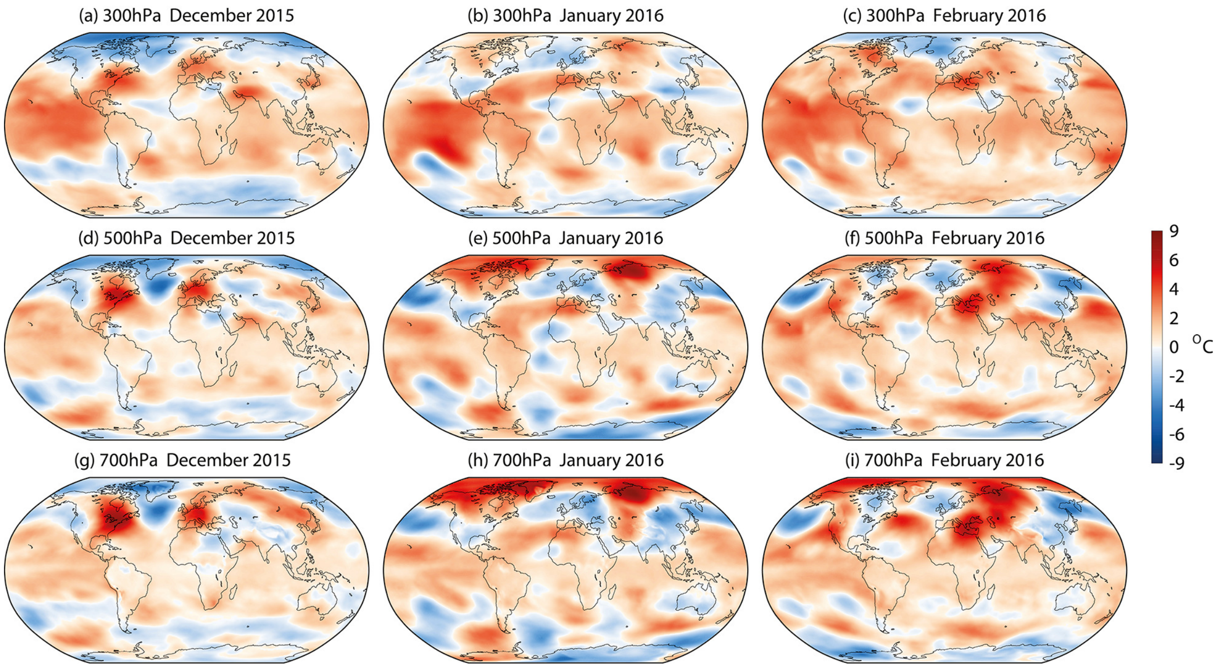

Figure 14 Anomalies in monthly-average ERA-Interim temperatures (°C) relative to 1981-1990, at (a, b, c) 300hPa, (d, e, f) 500hPa and (g, h, i) 700hPa, for (a, d, g) December 2015, (b, e, h) January 2016 and (c, f, i) February 2016.

Placing these numbers in the context of the Paris Agreement requires that an estimate be made of the extent to which the 1981-2010 norms for global temperature lie above the global temperature for a pre-industrial period that is not precisely defined. There are also issues of sparser data coverage and poorer data quality in earlier years. Uncertainty in the amount of warming from the start of the Industrial Revolution, around the middle of the 18th century, to the final decades of the 20th century is considerably larger than uncertainty in the amount of warming over the past three to four decades. Taking the warming since the latter half of the 19th century from the conventional datasets discussed in section 8 and adding 0.1°C for earlier warming based on evidence presented by Masson-Delmotte *et al.* (2013) suggests a value 0.7°C for the warming from the pre-industrial to 1981-2010, with a two-standard-deviation uncertainty range upwards of ±0.1°C (Morice *et al.*, 2012). This in turn suggests that global-mean surface temperatures have been close to or more than 1°C above the mid-1700s level for the past two years, and that they recently peaked at around 1.5°C above. It is salutary to note from Figure 3 that on the same basis it was probably only in 1998 that the 1°C level was first reached. This was in the latter stages of a somewhat stronger El Niño event than the latest one, but one that was accompanied by below- rather than above-average temperatures in the Arctic (Simmons and Poli, 2015) and by sea-surface temperatures that were on average cooler away from the tropical eastern Pacific.

What constitutes damaging climate change cannot be encapsulated in the value of a single metric such as global-mean surface temperature, even if the latter arguably provides the best single quantity for which to express an overall target (Knutti *et al.*, 2015). Nevertheless, whatever metric is used in a particular case, the critical value or range of values must be an absolute one, even if imperfectly known, not one relative to a pre-industrial level of uncertain value that is likely to change as data, modelling and understanding are refined. Targets for limiting future change, to be achieved by limiting anthropogenic disruption of the climate system, would be better framed and monitored in a global stocktake in terms of change relative to the recent period over which the system has been comparatively well observed. This does not mean that work to improve estimation and understanding of change from the pre-industrial to the recent past is not needed, as it serves purposes that include evaluating climate models and determining responsibilities for past change and its impacts. Splitting the calculation of change into two parts would allow the conventional approaches to be improved for application to the recent past, for example to use sea-surface temperature analyses that draw on satellite as well as *in situ* data, or to use satellite data to help fill gaps over land and sea-ice, as already done by Cowtan and Way (2014) in their hybrid approach to spatially extending HadCRUT4.

The absence of infilling and spatial smoothing in HadCRUT4 has the advantage of identifying grid squares where data are most directly tied to conventional climate observations and thus particularly suitable for cross-checking with the results of reanalysis. It does, though, have the downside of making the HadCRUT4 median prone to be an outlier, separated from the consensus of other datasets by more than might be expected from the spread of the underlying ensemble of realizations. This is not unexpected when it applies to the monthly variations examined here for recent years, as the HadCRUT4 ensemble samples uncertainty primarily on multi-decadal timescales. However, if all data are expressed relative to 1981-2010, the ensemble also does not have the spread to encompass the values from GISTEMP and NOAAGlobalTemp in averages for 1911-1940 and 1941-1970, echoing a similar AR5 finding for the decadal averages from previous versions of the datasets (Hartmann *et al.*, 2013). Differences in SST analyses are the main contributor. It is important to recall in this regard that GISTEMP and NOAAGlobalTemp use the same SST analyses, and should not be regarded as independent datasets when combined or compared with HadCRUT4. Also, differences between the infillings provided by GISTEMP and Had4_krig_v2 and the absence of features evidently linked to sea-ice changes indicate uncertainties in these datasets for the earlier periods.

The present results provide more evidence of the value of reanalysis as an important complement to the conventional analyses for depicting variability and change in surface temperature. Reanalysis exploits the richness of the observing system that has been in place over recent decades. ERA-Interim and JRA-55 have set a standard for the period in the sense that newer versions of the conventional datasets, and those datasets that extend data coverage spatially, are closer to the two reanalyses than are earlier versions. This is seen in particular in the estimates of the global temperature trend over the period 1998-2012 discussed only for the conventional datasets in AR5. Reanalysis also provides data with higher spatial and temporal resolution, and linked information on related variables such as surface humidity and precipitation (Simmons *et al.*, 2010; 2014) and atmospheric energy as discussed in this report. The conventional datasets still provide the primary source of information for earlier decades, for which data recovery and improved data assimilation are needed to advance the contribution of centennial-scale reanalysis.

Use of reanalysis to monitor recent and future change nevertheless requires a careful, comparative and selective approach. More is needed than simply displaying multiple time series from all or many of the available products and judging uncertainty from the spread of values. Some reanalyses are more fit-for-purpose than others for a particular application such as surface-temperature trends, whereas others may be competitive for other applications. Although newer reanalyses are expected generally to perform better than older ones, ERA-Interim and JRA-55 have been shown here to perform more consistently with each other and with the conventional surface-temperature datasets than the newer MERRA-2 does. Moreover, it has been necessary even for ERA-Interim and JRA-55 to combine background fields over sea and analysis data over land, and to make adjustments for inconsistencies in SST analysis for ERA-Interim, in order to achieve the reported level of agreement.

The need for high quality SST analyses is common to the conventional datasets and the reanalyses, although in the longer term there is the prospect of using direct assimilation of SST observations in reanalysis systems that couple atmosphere, ocean and sea ice (Dee *et al.*, 2014; Laloyaux *et al.*, 2016). The most significant difference in SST anomalies between ERA-Interim and JRA-55 occurs between 2003 and 2006, when the COBE SST used by JRA-55, in common with the ERSSTv4 used by GISTEMP and NOAAGlobalTemp, gives colder values than the operational SST analysis of the US National Centers for Environmental Prediction (Gemmill *et al.*, 2007) used by ERA-Interim. The difference is just sufficient by itself to account for 2014 being a warmer year than 2005 in JRA-55 but not ERA-Interim. ERA5, the replacement for ERA-Interim, will not resolve the difference, as the ten HadISST2 realizations it will use prior to 2007 in its ensemble data assimilation all differ little from the ERA-Interim SST over the period in question. Moreover, the year 2010 is warmer than 2014 in both the spatially extended Had4_UAH_v2 dataset and ERA-Interim. Unqualified statements that 2014 is ranked the second warmest year globally after 2015 are commonly made, but do not have a strong basis.

ERA-Interim is also relatively warm in the global-mean for 2005 because it has a larger warm anomaly in the Antarctic than the other datasets, including contributions in winter and spring from regions with anomalously low sea-ice concentrations. Direct observational data on surface air temperature is sparse for the Antarctic, but none of the datasets examined provides evidence of net warming south of 60°S since 1979, a period during which sea-ice extent increased a little. Under-sampling of this region when estimating global means from the conventional datasets thus tends to compensate for under-sampling of the Arctic, where warming has been greater than elsewhere. The agreement between datasets concerning the global temperature trend since 1979 may thus be partly fortuitous, as evidence points to warming of the Antarctic over earlier decades and model projections are for future warming, although confidence is low in several aspects of observation, modelling and understanding for the region (IPCC, 2013).

The comparisons presented here support the current use of ERA-Interim as the primary source of information for the temperature summaries issued early each month by the Copernicus Climate Change Service (at http://climate.copernicus.eu). The summaries extend time series and update maps of the type shown in this report. They are based on provisional ERA-Interim analyses for the preceding month, which can be cross-checked with promptly-available JRA-55 analyses and with some of the published national information that the reanalyses help place in regional and global contexts. Observational data supply for the conventional datasets is slower, and results from them are normally available two to four weeks after the end of the month; these datasets serve to provide longer-term context and indications of uncertainties. Complete ERA-Interim data are released only after a delay of around two months, following more comprehensive checks. This allows short periods to be rerun if remediable problems are detected.

Acknowledgment

This study has been partially funded by the Copernicus Climate Change Service. ECMWF implements this Service and the Copernicus Atmosphere Monitoring Service on behalf of the European Commission.

References

Berrisford P, Dee D, Poli P, Brugge R, Fielding K, Fuentes M, Kållberg P, Kobayashi S, Uppala S, Simmons A. 2011. The ERA-Interim archive. Version 2.0. ERA Report Series no. 1, 23pp. Available from www.ecmwf.int

Bosilovich MG, Akella S, Coy L, Cullather R, Draper C, Gelaro R, Kovach R, Liu Q, Molod A, Norris P, Wargan K, Chao W, Reichle R, Takacs L, Vikhliaev Y, Bloom S, Collow A, Firth S, Labow G, Partyka G, Pawson S, Reale O, Schubert SD, Suarez M. 2015. MERRA-2: Initial Evaluation of the Climate. NASA/TM–2015-104606/Vol. 43. Available from http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/

Bromwich DH, Nicolas JP, Monaghan AJ, Lazzara MA, Keller LM, Weidner GA Wilson AB. 2013. Central West Antarctica among the most rapidly warming regions on Earth. Nature Geosci, 6: 139–145, doi:10.1038/ngeo1671

Christy JR, Norris WB, Spencer RW, Hnilo JJ. 2007. Tropospheric temperature change since 1979 from tropical radiosonde and satellite measurements. J. Geophys. Res. 112(D6): D06 102, doi: 10.1029/2005JD006881

Cowtan K, Way RG. 2014. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. Q.J.R. Meteorol. Soc., 140, 1935–1944, doi: 10.1002/qj.2297

Cowtan K, Hausfather Z, Hawkins E, Jacobs P, Mann ME, Miller SK, Steinman BA, Stolpe MB, Way RG. 2015. Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. Geophys. Res. Lett., 42, 6526–6534, doi: 10.1002/2015GL064888

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F. 2011a. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteor. Soc., 137: 553–597, doi: 10.1002/qj.828

Dee DP, Balmaseda M, Balsamo G, Engelen R, Simmons AJ, Thépaut J-N. 2014. Toward a consistent reanalysis of the climate system. Bull. Am. Meteorol. Soc., 95, 1235-1248, doi: 10.1175/BAMS-D-13-00043.1

Donlon CJ, Martin M, Stark J, Roberts-Jones J, Fiedler E, Wimmer W. 2012. The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. Remote Sens. Environ., 116: 140-158, doi: 10.1016/j.rse.2010.10.017

Fréville H, Brun E, Picard G, Tatarinova N, Arnaud L, Lanconelli C, Reijmer C, van den Broeke M. 2014. Using MODIS land surface temperatures and the Crocus snow model to understand the warm bias of ERA-Interim reanalyses at the surface in Antarctica. The Cryosphere, 8, 1361-1373, doi: 10.5194/tc-8-1361-2014

Fyfe JC, Meehl GA, England MH, Mann ME, Santer BD, Flato GM, Hawkins E, Gillett NP, Xie S-P, Kosaka Y, Swart NC. 2016. Making sense of the early-2000s warming slowdown. Nature Clim. Change, 6, 224–228, doi: 10.1038/nclimate2938

GCOS. 2015. Status of the global observing system for climate. GCOS publication no. 195, WMO, Geneva. 353pp. Available from http://www.wmo.int/pages/prog/gcos/

Gemmill W, Katz B, Li X. 2007. Daily Real-Time Global Sea Surface Temperature - High Resolution Analysis at NOAA/NCEP. NOAA/ NWS/ NCEP/ MMAB Office Note 260, 39 pp. Available from http://polar.ncep.noaa.gov/mmab/papers/tn260/

Hansen JE, Lebedeff S. 1987. Global trends of measured surface air temperature. J. Geophys. Res., 92, 13345-13372, doi: 10.1029/JD092iD11p13345

Hansen J, Ruedy R, Sato M, Lo K. 2010. Global surface temperature change. Rev. Geophys., 48, RG4004, doi: 10.1029/2010RG000345

Hartmann D., Klein Tank AMG, Rusticucci M, Alexander LV, Brönnimann S, Charabi Y, Dentener FJ, Dlugokencky EJ, Easterling DR, Kaplan A, Soden BJ, Thorne PW, Wild M, Zhai PM. 2013: Observations: Atmosphere and Surface. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Hersbach H, Peubey C, Simmons A, Berrisford P, Poli P, Dee D. 2015. ERA-20CM: a twentieth-century atmospheric model ensemble. Q.J.R. Meteorol. Soc., 141, 2350–2375, doi: 10.1002/qj.2528

Huang B, Banzon VF, Freeman E, Lawrimore J, Liu W, Peterson TC, Smith TM, Thorne PW, Woodruff SD, Zhang H-M. 2015. Extended Reconstructed Sea Surface Temperature Version 4 (ERSST.v4). Part I: Upgrades and Intercomparisons. J. Cli., 28, 911-930, doi: 10.1175/JCLI-D-14-00006.1

IPCC. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp

IPCC. 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp

Ishii M, Shouji A, Sugimoto S, Matsumoto T. 2005. Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the KOBE collection. Int. J. Climatol., 25, 865–879, doi: 10.1002/joc.1169

Jones PD, Lister DH, Osborn TJ, Harpham C, Salmon M, Morice CP. 2012. Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010. J. Geophys. Res., 117, D05127, doi:10.1029/2011JD017139

Karl TR, Arguez A, Huang B, Lawrimore JH, McMahon JR, Menne MJ, Peterson TC, Vose RS, Zhang H-M. 2015. Possible artifacts of data biases in the recent global surface warming hiatus. Sciencexpress, www.sciencemag.org/content/early/recent, 1-7, doi: 10.1126/science.aaa5632

Kennedy JJ, Rayner NA, Smith RO, Saunby M, Parker DE. 2011. Reassessing biases and other uncertainties in sea-surface temperature observations since 1850, part 2: biases and homogenisation. J. Geophys. Res., 116, D14104, doi:10.1029/2010JD015220

Kobayashi S, Ota Y, Harada Y, Ebita A, Moriya M, Onoda H, Onogi K, Kamahori H, Kobayashi C, Endo H, Miyaoka K, Takahashi K. 2015. The JRA-55 Reanalysis: General Specifications and Basic Characteristics. J. Meteorol. Soc. Japan, 93, 5-48, doi:10.2151/jmsj.2015-001

Knutti R, Rogelj J, Sedlacek J, Fischer EM, 2015. A scientific critique of the two-degree climate change target. Nature Geosci., 9, 13–18, doi:10.1038/ngeo2595

Laloyaux P, Balmaseda M, Dee D, Mogensen K, Janssen P, 2016. A coupled data assimilation system for climate reanalysis. Q. J. R. Meteorol. Soc., 142, 65-78, doi: 10.1002/qj.2629

Lewandowsky S, Risbey J, Oreskes N. 2015. The "Pause" in Global Warming: Turning a Routine Fluctuation into a Problem for Science. Bull. Amer. Meteor. Soc, doi:10.1175/BAMS-D-14-00106.1, in press

Masson-Delmotte V, Schulz M, Abe-Ouchi A, Beer J, Ganopolski A, González Rouco JF, Jansen E, Lambeck K, Luterbacher J, Naish T, Osborn T, Otto-Bliesner B, Quinn T, Ramesh R, Rojas M, Shao X, Timmermann A. 2013. Information from Paleoclimate Archives. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Mears CA, Wentz FJ. 2009. Construction of the Remote Sensing Systems V3.2 Atmospheric Temperature Records from the MSU and AMSU Microwave Sounders. J. Atmos. Oceanic Technol., 26: 1040–1056, doi: 10.1175/2008JTECHA1176.1

Morice CP, Kennedy JJ, Rayner NA, Jones PD. 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset. J. Geophys. Res., 117, D08101, doi:10.1029/2011JD017187

Poli P, Hersbach H, Dee D, Berrisford P, Simmons A, Vitart F, Laloyaux P, Tan D, Peubey C, Thépaut J-N, Trémolet Y, Holm E, Bonavita M, Isaksen L, Fisher M. 2016. ERA-20C: An atmospheric reanalysis of the 20th century. J. Climate, doi:10.1175/JCLI-D-15-0556.1, in press

Rhein M, Rintoul SR, Aoki S, Campos E, Chambers D, Feely RA, Gulev S, Johnson GC, Josey SA, Kostianoy A, Mauritzen C, Roemmich D, Talley LD, Wang F. 2013. Observations: Ocean. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim G-J, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redde, CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen, J. 2011. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. J. Climate, 24: 3624-3648, doi: 10.1175/JCLI-D-11-00015.1

Simmons, AJ, Jones, PD, da Costa Bechtold, V, Beljaars, ACM, Kållberg, PW, Saarinen, S, Uppala, SM, Viterbo, P, and Wedi, N. 2004. Comparison of trends and low-frequency variability in CRU, ERA-40 and NCEP/NCAR analyses of surface air temperature. J. Geophys. Res., 109, D24115, doi: 10.1029/2004JD005306

Simmons AJ, Willett KM, Jones PD, Thorne PW, Dee DP. 2010. Low-frequency variations in surface atmospheric humidity, temperature and precipitation: Inferences from reanalyses and monthly gridded observational datasets. J. Geophys. Res., 115, D01110, doi: 10.1029/2009JD012442

Simmons AJ, Poli P, Dee DP, Berrisford P, Hersbach H, Kobayashi S, Peubey C. 2014. Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. Q.J.R. Meteorol. Soc., 140: 329-353, doi: 10.1002/qj.2317

Simmons AJ, Poli P. 2015. Arctic warming in ERA-Interim and other analyses. Q.J.R. Meteorol. Soc., 141, 1147-1162, doi: 10.1002/qj.2422

Smith TM, Reynolds RW, Peterson TC, Lawrimore J. 2008. Improvements to NOAA's Historical Merged Land–Ocean Surface Temperature Analysis (1880–2006). J. Climate, 21, 2283–2296, doi: 10.1175/2007JCLI2100.1

UNFCCC. 2015. The Paris Agreement reached by Parties to the United Nations Framework Convention on Climate Change. http://unfccc.int/meetings/paris_nov_2015/items/9445.php