3

Supplementary Materials

Observations: Surface and Atmospheric Climate Change

Coordinating Lead Authors:

Kevin E. Trenberth (USA), Philip D. Jones (UK)

Lead Authors:

Peter Ambenje (Kenya), Roxana Bojariu (Romania), David Easterling (USA), Albert Klein Tank (Netherlands), David Parker (UK), Fatemeh Rahimzadeh (Iran), James A. Renwick (New Zealand), Matilde Rusticucci (Argentina), Brian Soden (USA), Panmao Zhai (China)

Contributing Authors:

R. Adler (USA), L. Alexander (UK, Australia, Ireland), H. Alexandersson (Sweden), R. Allan (UK), M.P. Baldwin (USA), M. Beniston (Switzerland), D. Bromwich (USA), I. Camilloni (Argentina), C. Cassou (France), D.R. Cayan (USA), E.K.M. Chang (USA), J. Christy (USA), A. Dai (USA), C. Deser (USA), N. Dotzek (Germany), J. Fasullo (USA), R. Fogt (USA), C. Folland (UK), P. Forster (UK), M. Free (USA), C. Frei (Switzerland), B. Gleason (USA), J. Grieser (Germany), P. Groisman (USA, Russian Federation), S. Gulev (Russian Federation), J. Hurrell (USA), M. Ishii (Japan), S. Josey (UK), P. Kållberg (ECMWF), J. Kennedy (UK), G. Kiladis (USA), R. Kripalani (India), K. Kunkel (USA), C.-Y. Lam (China), J. Lanzante (USA), J. Lawrimore (USA), D. Levinson (USA), B. Liepert (USA), G. Marshall (UK), C. Mears (USA), P. Mote (USA), H. Nakamura (Japan), N. Nicholls (Australia), J. Norris (USA), T. Oki (Japan), F.R. Robertson (USA), K. Rosenlof (USA), F.H. Semazzi (USA), D. Shea (USA), J.M. Shepherd (USA), T.G. Shepherd (Canada), S. Sherwood (USA), P. Siegmund (Netherlands), I. Simmonds (Australia), A. Simmons (ECMWF, UK), C. Thorncroft (USA, UK), P. Thorne (UK), S. Uppala (ECMWF), R. Vose (USA), B. Wang (USA), S. Warren (USA), R. Washington (UK, South Africa), M. Wheeler (Australia), B. Wielicki (USA), T. Wong (USA), D. Wuertz (USA)

Review Editors:

Brian J. Hoskins (UK), Thomas R. Karl (USA), Bubu Jallow (The Gambia)

This chapter should be cited as:

Trenberth, K.E., P.D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F. Rahimzadeh, J.A. Renwick, M. Rusticucci, B. Soden and P. Zhai, 2007: Observations: Surface and Atmospheric Climate Change. In: *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., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Appendix 3.B: Techniques, Error Estimation and Measurement Systems

Appendix 3.B: Techniques, Error Estimation and Measurement Systems

3.B.1 Methods of Temperature Analysis: Global Fields and Averages

The first step in creating representative global gridded datasets is to take account of the number and error-characteristics of the observations within individual grid-boxes, reducing the variance of grid-box values if they are based on sparse or unreliable data, and yielding uncertainty estimates for each grid-box value (Jones et al., 2001; Rayner et al., 2006; Smith and Reynolds, 2005; Brohan et al., 2006). Grid-box values have been (a) used to create maps of trends over specified periods and (b) combined with areal weighting to derive regional, continental, hemispheric and global time series. A number of maps and time series are shown in Section 3.2, all with temperatures expressed as anomalies or departures from 1961–1990. Estimates of actual temperatures can be retrieved by adding back the climatologies to the anomaly data (Jones et al., 1999). Estimates of uncertainties of time series values must involve an estimate of the number of spatial degrees-of-freedom, as only a fraction of all the observations are statistically independent (see Jones et al., 1997, 2001; Rayner et al., 2006; Brohan et al., 2006). Vinnikov et al. (2004) have also presented a new technique for analysis of diurnal and seasonal cycles and trends, in which anomalies are calculated only implicitly.

The effects of changes in coverage over the instrumental period (now since 1850 for global scales) were first assessed by the 'frozen grid' and theoretical approaches (see Jones et al., 1997, 1999, 2001). Subsequently, Reduced-Space Optimal Interpolation (RSOI) has been used to infill incomplete and noisy fields and to provide local error estimates (Kaplan et al., 1997; Rayner et al., 2003). Optimal averaging (OA) yields large-area averages with error-bars (Folland et al., 2001). Global estimates are less reliable before 1900 (by a factor of two) than since 1951, but this is principally expressed on the interannual timescale. The sparser grids of the late-19th century estimate decadal and longer-timescale averages for periods since 1940 very reliably. RSOI and OA use the major patterns of variability (such as that associated with El Niño), to account for areas with no observations. The patterns are derived using data for recent, well-sampled years, and the technique relies on the assumption that the same patterns occurred throughout the record. Hence it depends on the stationarity of the record and this is a questionable assumption given known climate change. If the regions affected by a pattern are sparsely sampled, the pattern is accorded reduced weight in the analysis and error estimates are augmented. Neither RSOI nor OA can reproduce trends reliably (Hurrell and Trenberth, 1999); the data must therefore first be detrended by, for example, using the covariance matrix to estimate the temperature anomaly pattern associated with global warming, and removing the projection of this pattern from the data. After the techniques have been applied to the residuals, the trend component is restored.

Vose et al. (2005) show that estimates of global land surface air temperature trends are affected less by local data coverage than by the choice between a weighted grid-box average for the globe and the average of the weighted grid-box averages for the two hemispheres. This underscores the value of the OA technique which takes optimal account of unsampled regions. However, because OA assumes zero anomalies in the absence of information, it yields global anomalies of smaller magnitude than other techniques when data are sparse (Hurrell and Trenberth, 1999).

In addition to errors from changing coverage and from random measurement and sampling errors, errors arise from biases (Section 3.B.2). Major efforts have been made to adjust for known systematic biases, but some adjustments nonetheless are quite uncertain. Nevertheless, recent studies have estimated all the known errors and biases to develop error bars (Brohan et al., 2006). For example, for SSTs, the transition from taking temperatures from water samples from uninsulated or partially-insulated buckets to engine intakes near or during World War II is adjusted for, even though details are not certain (Rayner et al., 2006).

3.B.2 Adjustments to Homogenize Land Temperature Observations

Long-term temperature data from individual climate stations almost always suffer from inhomogeneities, owing to non-climatic factors. These include sudden changes in station location, instruments, thermometer housing, observing time, or algorithms to calculate daily means; and gradual changes arising from instrumental drifts or from changes in the environment due to urban development or land use. Most abrupt changes tend to produce random effects on regional and global trends, and instrument drifts are corrected by routine thermometer calibration. However, changes in observation time (Vose et al., 2004) and urban development are likely to produce widespread systematic biases; for example, relocation may be to a cooler site out of town (Böhm et al., 2001). Urbanisation usually produces warming, although examples exist of cooling in arid areas where irrigation effects dominate.

When dates for discontinuities are known, a widely used approach is to compare the data for a target station with neighbouring sites, and the change in the temperature data due to the non-climatic change can be calculated and applied to the pre-move data to account for the change, if the discontinuity is statistically significant. However, often the change is not documented, and its date must be determined by statistical tests. The procedure moves through the time series checking the data before and after each value in the time series (Easterling and Peterson, 1995; Vincent, 1998; Menne and Williams, 2005): this works for monthly or longer means, but not daily values owing to greater noise at weather timescales. An extensive review is given by Aguilar et al. (2003).

The impact of random discontinuities on area-averaged values typically becomes smaller as the area or region becomes larger, and is negligible on hemispheric scales (Easterling et al., 1996). However, trends averaged over small regions, in particular, may be biased by systematic heterogeneities in the data (Böhm et al., 2001), and the impact of non-random discontinuities can be important even with large averaging areas. The time-of-observation bias documented by Karl et al. (1986) shows a significant impact even with time series derived for the entire contiguous United States. Adjustments for this problem remove an artificial cooling that occurs due to a switch from afternoon to morning observation times for the U.S. Cooperative Observer Network (Vose et al., 2004).

Estimates of urban impacts on temperature data have included approaches such as linear regression against population (Karl et al., 1988), and analysis of differences between urban and rural sites defined by vegetation (Gallo et al., 2002) or night lights (Peterson, 2003) as seen from satellites. Urbanisation impacts on global and hemispheric temperature trends (Karl et al., 1988; Jones et al., 1990; Easterling et al., 1997; Peterson, 2003; Parker, 2004, 2006) have been found to be small. Furthermore, once the landscape around a station becomes urbanized, long-term trends for that station are consistent with nearby rural stations (Böhm, 1998; Easterling et al., 2005, Peterson and Owen, 2005). However, individual stations may suffer marked biases and require treatment on a case-by-case basis (e.g., Davey and Pielke, 2005); the influence of urban development and other heterogeneities on temperature depends on local geography and climate so that adjustment algorithms developed for one region may not be applicable in other parts of the world (Hansen et al., 2001; Peterson, 2003).

Homogenization of daily temperature series requires much more metadata than monthly assessment (see the extensive discussion in Camuffo and Jones, 2002) and only a few series can be classed as totally homogeneous. Daily minima and maxima, and consequently also DTR and analysis of extremes, are particularly sensitive to non-climatic heterogeneities, including changes in height above ground, housing and ventilation of instruments (Auer et al., 2001; Brunet et al., 2006). The ongoing automation of measuring networks is typically accompanied by a change from large and unventilated screens to small and continuously ventilated ones. Assessment of potential homogeneity problems in a network of 60 daily maximum and minimum temperature series, for Europe for the 20th century by Wijngaard et al. (2003), suggests that 94% of series should be classed as of doubtful homogeneity. The percent of doubtful series reduces to 61% when considering 158 series for 1946–1999. Vincent et al. (2002) in a Canadian study of over 200 daily temperature series, develop daily adjustments by smooth interpolation of monthly adjustments. But a new technique adjusts higher order daily statistics (Della Marta and Wanner, 2006).

3.B.3 Adjustments to Homogenize Marine Temperature Observations

Owing to changes in instrumentation, observing environment and procedure, SSTs measured from modern ships and buoys are not consistent with those measured before the early 1940s using canvas or wooden buckets. SST measured by canvas buckets, in particular, generally cooled during the sampling process. So systematic adjustments are necessary (Folland and Parker, 1995; Smith and Reynolds, 2002; Rayner et al., 2006) to make the early data consistent with modern observations that have come from a mixture of buoys, engine inlets, hull sensors and insulated buckets. A combined physical-empirical method (Folland and Parker, 1995) is mainly used, as reported in the TAR, to estimate adjustments to ship SST data obtained up to 1941 to compensate for the heat losses from uninsulated (mainly canvas) or partly insulated (mainly wooden) buckets. The adjustments are based on the physics of heat-transfer from the buckets and are independent of land-surface air temperature or night marine air temperature (NMAT) data measured by ships. The adjustments increased between the 1850s and 1940 because the fraction of canvas buckets increased and because ships moved faster, increasing the ventilation. By 1940 the adjustments were 0.4°C for the global average and approached 1°C in winter over the Gulf Stream and Kuroshio where surface heat fluxes are greatest. Folland (2005) verified these spatially and temporally complex adjustments by comparing the Jones and Moberg (2003) land-surface air temperature anomalies on global and continental scales with simulations using the Hadley Centre Atmospheric Climate Model HadAM3 forced with observed SST and sea-ice extents since 1871. The simulated decadal and longer-term variations of land surface air temperatures on global and continental scales were much better when the model was driven with adjusted than with unadjusted SSTs, providing strong support to the SST adjustments globally, regionally and seasonally (Folland, 2005). Smith and Reynolds (2002) have independently bias-adjusted updated COADS (Slutz et al., 1985) SST anomalies to agree with COADS NMAT anomalies before 1942, using historical variations in the pattern of the annual amplitude of air-sea temperature differences in unadjusted data, and derive rather similar spatiotemporal adjustments to Folland and Parker (1995), although there are seasonal differences. Overall, they recommend use of the Folland and Parker (1995) adjustments as these are independent of any changes in NMAT data and more fully take into account evaporation errors in uninsulated buckets, especially in the tropics. Smith and Reynolds (2004) analysis of ICOADS (formerly COADS Release 2.0, Woodruff et al., 1998) requires SST bias adjustments before 1942 similar to those of Smith and Reynolds (2002), except in 1939–1941 when ICOADS contains a new data source which clearly has many more engine intake data that do not need adjustment. Rayner et al. (2006), in a new analysis of the ICOADS data with no interpolation, adapt the Folland and Parker (1995) adjustments in 1939–1941 in a similar way to Smith and Reynolds (2004) but, unlike Smith and Reynolds (2005), do not widen the error bars because the new adjustments are compatible with well-understood changes in the data.

There are smaller biases between modern SSTs taken separately by engine inlets and insulated buckets (Kent and Kaplan, 2006) and between overall ship and buoy observations (Rayner et al., 2006). These biases may arise from the different measurement depths (buckets, typically 30 cm; buoys, typically 1 m; engine inlets, typically 10 m) and from heat inputs from the ship near engine inlets. Biases can also vary by nation. The biases are not large enough to prejudice conclusions about recent warming. The increasing amount of buoy data, although in principle more accurate than most ship measurements, introduces further inhomogeneities (Kent and Challenor, 2006; Kent and Kaplan, 2006), which may have caused an underestimate of recent warming (Rayner et al., 2006). The exact effect on trends of the changes in the methods of measurement in recent decades has not yet been assessed.

Modern observations of SST made in situ have been supplemented by satellite-based data since about 1980 giving much better geographical coverage. However, satellite estimates are of skin (infrared) or sub-skin (typically 1 cm, microwave) temperatures and the infrared data are also affected by biases, especially owing to dust aerosol and to misinterpretation of thin clouds and volcanic aerosols as cool water. Also, instruments on successive satellites are not identical, and instruments in orbit can degrade slowly or show spurious jumps. In situ observations have been used to provide calibration for the satellite measurements, which can then been used to fill in the spatial patterns for areas where there are few ships or buoys (Reynolds et al., 2002).

Some efforts have been made to monitor SST from satellite data alone. Lawrence et al. (2004) have compared SSTs from the Pathfinder dataset, which uses the Advanced Very High Resolution Radiometer (AVHRR) with SSTs from the ATSR. The analysed data are not truly global because of problems in distinguishing SST from cloud top temperatures in many regions. Also, the Pathfinder data have time varying biases (Reynolds et al., 2002), and the method for combining data from two different ATSR instruments may need more scrutiny. Nevertheless, the Pathfinder dataset shows similar rates of warming to in situ data over 1985–2000. These rates are insignificantly different from the global trend over 1979–2005 from in situ data (0.13ºC per decade) (see Table 3.2). ATSR data also show warming but the period available (1991–2004 with some gaps) is too short to assess a reliable trend (O'Carroll et al., 2006). In future, satellite SST data may be improved by combining infrared and microwave data to provide global coverage where clouds make infrared data unreliable (Wentz et al., 2000; Donlon et al., 2002; Reynolds et al., 2004). The new Global Ocean Data Assimilation Experiment (GODAE) high-resolution SST pilot project (GHRSST-PP) will establish uncertainty estimates (bias and standard deviation) for all satellite SST measurements by careful reference to in situ SST observations, accounting for the mixed layer and differences in different bulk and skin temperatures.

Air temperatures taken on board ship have also been biased, mainly because ships have become larger, so that the temperatures were measured typically 6m above the sea in the late-19th century, 15m in the mid-20th century, and over 20m today. In addition, observing practices were irregular during the Second World War and in the 19th century. The data have been adjusted by Rayner et al. (2006). Owing to biases arising from solar heating of the ships' fabric, marine air temperature analyses have so far been based on night-time data (Rayner et al., 2003), though Berry et al. (2004) have developed a model for correcting the daytime data. Note that surface air temperatures do not bear a fixed relation to SST: thus, surface heat fluxes in the tropics change with the phase of ENSO, and surface fluxes in the N Atlantic vary with the NAO. However, in many parts of the world oceans and on larger space scales, air temperature and SST anomalies follow each other closely on seasonal and longer time scales (Bottomley et al., 1990).

Many historical in situ marine data still remain to be digitized and incorporated into ICOADS (Worley et al., 2005). These will improve coverage and reduce the uncertainties in our estimates of past marine climatic variations, but progress has been made since the TAR. The CLIWOC project (Garcia et al., 2005) has digitized an additional 40,000 marine air temperature (MAT) and SST data for the period before 1850. These data, and those of Chenoweth (2000) which have had quality control and bias adjustment, might allow NMAT to be extended back usefully to the early 19th century. Coverage would also be improved if daytime values could be corrected for time-varying daytime biases consistently through the whole dataset (Berry et al., 2004).

3.B.4 Solid/Liquid Precipitation: Undercatch and Adjustments for Homogeneity

3.B.4.1 Precipitation Undercatch (Snow and Rain)

Studies of biases in precipitation measurements by in-situ rain gauges (Poncelet, 1959; Sevruk, 1982; Sevruk and Hamon, 1984; Legates and Wilmott, 1990; Goodison et al., 1998; Golubev et al., 1995, 1999; Bogdanova et al., 2002a,b) find that (a) light rainfall and snowfall are strongly underestimated owing to wind-induced acceleration and vertical motion over the rain gauge orifice (for snowfall, the resulting biases can be as high as 100% of "ground truth" precipitation). The main physical reasons for the observed systematic undercatch of conventional raingauges when exposed to the wind, including the considerably more severe losses of snowfall, were modelled and compared to field observations by Folland (1988). To fix this deficiency, wind-scale correction factors have been developed (cf. Sevruk, 1982; Goodison et al., 1998); (b) most precipitation gauges have trouble reporting the full amount of precipitation that reaches the gauge owing to gauge precision problems (traces), losses (retention, evaporation) and accumulation (condensation) of water in/from the gauge; to fix these deficiencies additive corrections have been developed (cf. Sevruk, 1982; Golubev et al., 1995, 1999) and (c) in windy conditions with snow on the ground, blowing snow enters the gauges causing "false" precipitation; only recently has this factor started to be taken into account in major precipitation datasets in high latitudes (Bryazgin and Dement'ev, 1996; Bogdanova et al., 2002a,b).

After the completion of the International Solid Precipitation Intercomparison Project (Goodison et al., 1998), several attempts to adjust precipitation in high latitudes and create new regional climatologies (Mekis and Hogg, 1999; Yang, 1999; Yang et al., 1999; Yang and Ohata, 2001) and global datasets (Adam and Lettenmaier, 2003) accounted for problems identified in items (a) and (b). However, when this approach was applied to high latitudes (e.g., Yang, 1999; Mekis and Hogg, 1999; Yang and Ohata, 2001), unrealistically high precipitation estimates caused confusion among hydrologists. Critical reassessment of the problem was conducted by Golubev et al. (1995, 1999), Golubev and Bogdanova (1996), and Bogdanova et al. (2002a,b). Universal adjustments have emerged from their studies using parameters of wind speed, gauge and precipitation types, wetting and evaporation adjustments, and flurry and blow-in adjustments. Measured precipitation values are ignored when wind at 10 meters above the snow-covered ground reaches a 10 m s^{-1} threshold and are replaced with estimates of mean regional snowfall intensity and its duration, although this could introduce biases if snowfall rate is correlated with wind strength. A precipitation climatology over the Arctic Ocean (Bogdanova et al., 2002a) using this approach replaces measured annual totals of 128 mm with adjusted annual totals of 165 mm, an increase of 28% over measured values. This climatology corresponds broadly with independent estimates over the Arctic Ocean from aerological and snow cover measurements but is much less than proposed by Yang (1999) for the same region using the same data.

All correction routines suggest higher (in relative terms) adjustments for frozen than for liquid precipitation undercatch. If rising temperature increases the chances for rainfall rather than snowfall, then unadjusted gauges will show precipitation increases owing to the better catch of liquid precipitation. This mechanism was shown to be a major cause of artificially inflated trends in precipitation over the Norwegian Arctic (Førland and Hanssen-Bauer, 2000) but it is estimated to have a small effect on the measured precipitation trends in the European Alps (Schmidli et al., 2002).

3.B.4.2 Homogeneity Adjustments

Precipitation series are affected by the same sort of homogeneity issues as temperature: random changes owing to relocations (both in position and height above the ground) and local gauge changes, and more spatially consistent effects such as nationwide improvements to gauges and observation practices (Auer et al., 2005). Adjustment of precipitation series at the monthly, seasonal and annual timescale is much more demanding than for temperature, as the spatial correlation of precipitation fields is much weaker. Similar approaches have been tried as for land temperatures, looking at time series of the ratio of the catches at a candidate station to those of neighbours. In many regions, however, the networks are not dense enough to find many statistically significant differences. Auer et al. (2005) for the Greater Alpine Region give typical distances above which adjustments are not possible, these being timescale and season dependent, and range from 150 km separation at the monthly to 40 km at the daily timescale. Only a few networks are, therefore, dense enough to consider homogeneity assessment of daily precipitation totals and large-scale studies have rarely been undertaken. In the Wijngaard et al. (2003) study for Europe, the quality of daily precipitation series appears higher than for temperature, perhaps because there were fewer tests that could be applied than for temperature owing to larger natural variability. Only 25% of 88 stations with near-complete records for the 20th century were classed as doubtful, falling to 13% (of 180) for 1946–1999. The reliability of estimated trends in daily extreme precipitation depends on the completeness as well as the homogeneity of the record and is seriously degraded if more than about one third of the daily data are missing (Zolina et al., 2005).

3.B.5 The Climate Quality of Free-Atmosphere and Reanalysis Datasets

3.B.5.1 Evolution of the Observing System: Radiosondes

Radiosondes measure temperature, humidity and wind speed as they ascend, generally reaching the lower stratosphere before balloons burst. The quality of radiosonde measurements has improved over the past 5 decades, but oceanic coverage has declined owing to the demise of ocean weather ships: coverage over land has also declined in the 1990s. Counts of standard-level (e.g., 50 hPa) stratospheric measurements have risen, likely due to better balloons, but there may be remaining biases as balloon bursts still occur more frequently when cold (Parker and Cox, 1995). Many stations changed from twice daily to once daily reporting thereby potentially affecting trends, and only a subset of current stations has sufficiently long records to be directly useful for climate monitoring, except through reanalysis. There have been many changes to instrument design and observing practices to improve the accuracy of weather forecasts, and many manufacturers have released multiple radiosonde models. There have also been changes in the radiation corrections applied to account for insolation, in ground equipment, and in calculation methods. Only some of these changes have been documented (Gaffen, 1996 and subsequent updates), and rarely have simultaneous measurements been made to accurately quantify their effects. Developers of Climate Data Records (CDRs) from radiosondes have, therefore, to cope with a highly heterogeneous and poorly documented raw database. Since the TAR, efforts have been made to improve global digital databases incorporating more thorough homogeneity and outlier checks (e.g., Durre et al., 2006). Two major efforts to form homogeneous temperature CDRs from these records illustrate the range of possible approaches. Lanzante et al. (2003a, b) (LKS) homogenised data from 87 well-spaced stations using a manually intensive method. They used indicators from the raw data and metadata to try to identify the times of artificial jumps resulting from non-climatic influences. The resulting homogenised station data series were closer to the only available satellite-based MSU time series at the time (Christy et al., 1998). Thorne et al. (2005), in contrast, created a global database, HadAT2, containing 676 stations. They used LKS and the GCOS Upper Air Network (GUAN) to define an initial set of 477 adequate stations and then a neighbour comparison technique and metadata to homogenise their data. Subsequently the data from the remaining stations were incorporated in a similar way. The quality control identified an average of about 6 breakpoints per station that required adjustments, 70% of which were not identified with any known change in procedures, while about 29% were identified with changes in sonde or equipment. Most recently, effort has focused on reduction of possibly time-varying biases in daytime radiosonde data (Sherwood et al., 2005). Moisture data from radiosondes generally contain even more complex problems, and no climate quality homogenised databases have yet been produced.

3.B.5.2 Evolution of the Observing System: Buoys, Aircraft and Satellite Data

Other types of observations have compensated for the decline in radiosonde coverage. New data by 1979 included MSU, HIRS and SSU soundings from satellites. In 1979, winds derived by tracking features observed from geostationary satellites first became available in significant numbers and there were substantial increases in buoy and aircraft data. Overall observation counts declined for a while after 1979, but recovered during the 1980s. The frequency and coverage of wind and temperature measurements from aircraft increased substantially in the 1990s. The launch of the Earth Radiation Budget Experiment (ERBE) in 1984 began a series of satellite instruments that provided the first climate quality record of top-of-atmosphere radiative fluxes. Beginning in 1987, newer satellite-based data from microwave instruments provided improved observations of total water-vapour content, surface wind speed, rain rate, and atmospheric soundings (Uppala et al., 2005).

3.B.5.3 Analysis of Tropospheric and Stratospheric Temperature using Microwave Radiances

The MSU that has been used for climate monitoring as well as in reanalyses, has been flown continuously since late 1978 (AMSU since 1998) on polar orbiting satellites. Two retrieval channels have been used to create CDRs. MSU channel 2 and its AMSU near-equivalent measure a thick layer of the atmosphere, with approximately 75−80% of the signal coming from the troposphere, 15% from the lower stratosphere, and the remaining 5–10% from the surface. MSU Channel 4 and its AMSU sequel receive their signal almost entirely from the lower stratosphere (see Figure 3.16 of Chapter 3). Each satellite has lasted several years, and usually at least two satellites have been monitoring at roughly 6-hour intervals. Although the instruments are designed to the same specifications for each satellite, MSU instruments have had relative biases of the order 1–2°C. As the orbits have tended to drift, MSU instruments measure at systematically later local times over a satellite's lifetime requiring adjustments to be made for the diurnal cycle, a procedure accommodated automatically in ERA-40 by inserting the observation at the appropriate time. Satellite orbits also tend to decay, affecting the limb soundings of Channel 2 used by UAH to gain a lower tropospheric retrieval (Spencer and Christy, 1992; Wentz and Schabel, 1998). Finally, there is a suspected, time-varying systematic effect of the instrument body temperature upon the retrievals.

The original set of MSU data records produced by UAH has undergone improvement of the correction for diurnal drift, although the effect on trends was small; an error analysis was made and the record was extended to include AMSU measurements (Christy et al., 2003),. A new set of data records for channel 2 was constructed by RSS (Mears et al., 2003). Despite starting with identical raw satellite radiances, differences arise between RSS and UAH from the choice of data used to determine the parameters of the calibration target effect. RSS utilizes pentad-mean intersatellite-difference data without further averaging for calculation of the target temperature coefficients. UAH averages daily data into periods of at least 60 days and focuses on reducing low-frequency differences. RSS employs all difference-data, i.e., data from all co-orbiting, overlapping spacecraft, seeking the statistically best consensus for intersatellite bias determination. UAH omitted very small segments (e.g., 45 days or so) which occur at the tail-ends of the satellites' operational periods, to avoid the use of data segments which are too short for the averaging technique and are near the end of a satellite lifetime when its biases may be unrepresentative of its full span. The resulting parameters from the UAH procedure for NOAA-9 (1985–1987) were reported to be outside of the physical bounds expected (Mears et al., 2003). Hence the large difference in the calibration parameters for the single instrument mounted on the NOAA-9 satellite accounts for a substantial part (~50%) of the global trend difference between the UAH and RSS results. The rest arises from differences in merging parameters for other satellites, differences in the correction for the drift in measurement time (Mears et al., 2003; Christy and Norris, 2004; Mears and Wentz, 2005), and ways the hot point temperature is corrected for (Grody et al., 2004; Fu and Johanson, 2005). In the tropics, these account for T2 trend differences of order 0.1°C per decade after 1987 and discontinuities are also present in 1992 and 1995 at times of satellite transitions (Fu and Johanson, 2005).

The T2 data record of Grody et al. (2004) and Vinnikov et al. (2006) (VG2) uses a zonal mean latitude-dependent analysis that allows for errors that depend on both the calibration target temperature and the atmospheric temperature being measured. However, because temporal averaging is used to reduce noise in overlapping satellite measurements, issues remain in accounting for temporal variations in calibration target temperatures on individual satellites. The need to account for the target effect as a function of latitude is related to the diurnal cycle correction. The

VG2 method does not, however, fully address the correction for diurnal drift before merging and does not produce maps, so that differences between land and ocean remain to be evaluated.

A new benchmark method for measuring atmospheric temperatures is based on a time measurement using Radio Occultation (RO) from Global Positioning System (GPS) satellites. The promise of this method is revealed by Schroder et al. (2003) who found that UAH T4 retrievals in the Arctic lower stratosphere in winter were biased high relative to temperatures derived from GPS RO measurements.

3.B.5.4 Reanalysis and Climate Trends

Comprehensive reanalyses from NRA (Kalnay et al., 1996; Kistler et al., 2001), NCEP-2 reanalysis (Kanamitsu et al., 2002) and ERA-15/ERA-40 (Uppala et al., 2005) derived by processing multi-decadal sequences of past meteorological observations using modern data assimilation techniques have found widespread application in many branches of meteorological and climatological research. Care is needed, however, in using them to document and understand climatic trends and low-frequency variations. Atmospheric data assimilation comprises a sequence of analysis steps in which background information for a short period, typically of 6 or 12 hours duration, is combined with observations for the period to produce an estimate of the state of the atmosphere (the 'analysis') at a particular time. The background information comes from a short-range forecast initiated from the most-recent preceding analysis in the sequence. Problems for climate studies arise partly because the atmospheric models used to produce these "background forecasts" are prone to biases. If observations are abundant and unbiased, they can correct the biases in background forecasts when assimilated. In reality, however, observational coverage varies over time, observations are themselves prone to bias, either instrumental or through not being representative of their wider surroundings, and these observational biases can change over time. This introduces trends and low-frequency variations in analyses that are mixed with the true climatic signals, making long-timescale trends over the full length of the reanalyses potentially unreliable (Bengtsson et al., 2004; Simmons et al., 2004; see also http://wwwt.emc.ncep.noaa.gov/mmb/rreanl/). Better representation of trends by reanalysis systems requires progress on identifying and correcting model and observational biases, assimilating as complete a set of past observations as possible (particularly more sub-daily surface data before about 1970), and general improvements to the methods of data assimilation: in this regard the second-generation ERA-40 reanalysis represents a significant improvement over the earlier first generation analyses produced in Europe and the United States.

3.B.5.5 Bias Correction for Reanalysis

Reliable depiction of temperature trends by a reanalysis requires that changes over time in the biases of the assimilated observations be taken into account, just as they have to be when deriving trend information from radiosonde or MSU data alone. For satellite data, trends in the ERA-40 reanalysis have been affected adversely by difficulties in radiance bias adjustment for the early satellite data. Correcting older radiosonde data for reanalysis is also demanding owing to large, spatially and temporally variable biases and a lack of metadata. In ERA-40 no corrections were applied prior to 1980, but statistics of the difference between the observations and background forecasts are now being used to estimate corrections for application both in future reanalyses and in direct trend analysis (Haimberger, 2005).

References

Adam, J.C., and D.P. Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias, J. Geophys. Res., 108, 4257, doi:10.1029/2002JD002499.

Aguilar, E., et al., 2003: Guidelines on Climate Metadata and Homogenization. WCDMP-No. 53, WMO-TD No. 1186, World Meteorological Organization, Geneva, Switzerland, 55 pp.

Auer, I., R. Böhm and W. Schöner, 2001: Austrian long-term climate - Multiple instrumental climate time series from Central Europe. Österr. Beitr. Meteor. Geophys, 25, 147 pp.

Auer, I., et al., 2005: A new instrumental precipitation dataset in the greater alpine region for the period 1800–2002. Int. J. Climatol., 24, 139–166.

Bengtsson, L., S. Hagemann, and K.L. Hodges, 2004: Can climate trends be calculated from reanalysis data? J. Geophys. Res., 109, D11111, doi:10.1029/2004JD004536.

Berry, D.I., E.C. Kent, and P.K. Taylor, 2004: An analytical model of heating errors in marine air temperatures from ships. J. Atmos. Oceanic Technol., 21, 1198–1215.

Bogdanova, E.G., B.M. Ilyin, and I.V. Dragomilova, 2002a: Application of a comprehensive bias correction model to precipitation measured at Russian North Pole drifting stations. J. Hydrometeorol., 3, 700–713.

Bogdanova, E.G., et al., 2002b: A new model for bias correction of precipitation measurements, and its application to polar regions of Russia. Russian Meteorol. Hydrol., 10, 68–94.

Böhm, R., 1998: Urban bias in temperature series – a case study for the city of Vienna. Climatic Change, 38, 113–128.

Böhm R., et al., 2001: Regional temperature variability in the European Alps 1760-1998 from homogenised instrumental time series. Int. J. Climatol., 21, 1779-1801.

Bottomley, M., et al., 1990: Global Ocean Surface Temperature Atlas "GOSTA"., HMSO, London, 20 pp.+iv, 313 plates.

Brohan, P., et al., 2006: Uncertainty estimates in regional and global observed temperature changes: A new dataset from 1850 , J. Geophys. Res., 111, D12106, doi:10.1029/2005JD006548.

Brunet, M., et al., 2006: The development of a new dataset of Spanish daily adjusted temperature series (SDATS) (1850-2003). Int. J. Climatol., 25, 1777-1802.

Bryazgin, N.N., and A.A. Dement'ev, 1996: Dangerous meteorological events in Russian Arctic (in Russian), Gidrometeoizdat, St. Petersburg, 156 pp.

Camuffo, D., and Jones, P.D. (eds.), 2002: Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources. Kluwer Academic Publishers, Dordrecht, 392 pp..

Chenoweth, M., 2000: A new methodology for homogenisation of 19 th century marine air temperature data. J. Geophys. Res., 105, 29145–29154.

Christy, J.R., and W.B. Norris, 2004: What may we conclude about tropospheric temperature trends? Geophys. Res. Lett., 31, L0621, doi:10.1029/2003GL019361.

Christy, J.R., R.W. Spencer, and E.S. Lobl, 1998: Analysis of the merging procedure for the MSU daily temperature time series. J. Climate, 11, 2016–2041.

Christy, J.R., et al., 2003: Error estimates of version 5.0 of MSU/AMSU bulk atmospheric temperatures. J. Atmos. Oceanic Tech., 20, 613–629.

Davey, C.A., and R.A. Pielke Sr., 2005: Microclimate exposures of surface-based weather stations: implications for the assessment of long-term temperature trends. Bull. Amer. Meteor. Soc., 86, 497–504.

Della-Marta, P.M., and H. Wanner, 2006: A method for homogenising the extremes and mean of daily temperature measurements. J. Climate, 19, 4179−4197.

Donlon, C.J., et al., 2002: Toward improved validation of satellite sea surface temperature measurements for climate research. J. Climate, 15, 353–369.

Durre, I., R.S. Vose, and D.B. Wuertz, 2006: Overview of the integrated global radiosonde archive. J. Climate., 19, 53–68.

Easterling, D.R., and T.C. Peterson, 1995: A new method of detecting undocumented discontinuities in climatological time series. Int. J. Climatol., 15, 369–377.

Easterling, D.R., T.C. Peterson, and T.R. Karl, 1996: Notes and correspondence: On the development and use of homogenized climate datasets. J. Climate, 9, 1429–1434.

Easterling, D.R., et al., 1997: Maximum and minimum temperature trends for the globe. Science, 277, 364–367.

Easterling, D.R., et al., 2005: A comparison of local and regional trends in surface and lower tropospheric temperatures in western North Carolina. Earth Interactions, 9 (28), 1–9.

Folland, C.K., 1988: Numerical models of the raingauge exposure problem, field experiments and an improved collector design. Quart. J. Roy. Meteorol. Soc., 114, 1485–1516.

Folland, C.K., 2005: Tests of bias corrections to sea surface temperature using a climate model. Int. J. Climatol., 25, 895–911.

Folland, C.K. and D.E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. Quart. J. Roy. Meteorol. Soc., 121, 319–367.

Folland, C.K., et al., 2001: Global temperature change and its uncertainties since 1861. Geophys. Res. Lett., 28, 2621–2624.

Førland, E.J., and I. Hanssen-Bauer, 2000: Increased precipitation in the Norwegian Arctic: True or false? Climatic Change, 46, 485–509.

Fu, Q., and C.M. Johanson, 2005: Satellite-derived vertical dependence of tropical tropospheric temperature trends. Geophys. Res. Lett., 32, L10703, doi:10.1029/2004GL022266.

Gaffen, D.J., 1996: A digitized metadata set of global upper-air station histories. NOAA Technical Memorandum ERL ARL-211.

Gallo, K.P. et al., 2002: Satellite-based detection of global urban heat-island temperature difference. J. Geophys. Res., 107, D24, 4776, doi:10.1029/2002JD002588.

García, R., et al., 2005: CLIWOC: A climatological database for the world's oceans 1750–1854. Climatic Change, 73, 1-12.

Golubev, V.S., and E.G. Bogdanova, 1996: Accounting of blowing snow events in precipitation measurement in Russia. In: Proc. ACSYS Solid Precipitation Climatology Project Wkshp, Reston, VA, U.S., 12–15 September 1995, World Meteorol. Org., WCRP No. 93, WMO/TD No. 739.

Golubev, V.S., V.V. Koknayeva, and A.Yu. Simonenko, 1995: Results of the atmospheric precipitation measurements with national standard instruments of Canada, Russia, and USA. Meteozologiya i Gidzologiya Meteorology Hydrology, 2, 102–110.

Golubev, V.S., et al., 1999: Precipitation measurement correction and quality of corrected data from Valdai hydrological station. Russian Meteorol. Hydrol., 1, 65–74.

Goodison, B.E., P.Y.T. Louie, and D. Yang, 1998: WMO solid precipitation intercomparison, Final Report. World Meteorol. Org., Instruments and Observing Methods Rep. 67, WMO/TD 872, 87 pp. + Annexes.

Grody, N.C., et al., 2004: Calibration of multi-satellite observations for climatic studies: Microwave sounding unit (MSU). J. Geophys. Res., 109, D24104, doi:10.1029/2004JD005079.

Haimberger, L., 2005: Homogenization of radiosonde temperature time series using ERA-40 analysis feedback information. ERA-40 Project Report Series 23, ECMWF, Reading, U.K., 68 pp.

Hansen, J., et al., 2001: A closer look at United States and global surface temperature change. J. Geophys. Res., 106, 23947–23963.

Hurrell, J.W., and K.E. Trenberth, 1999: Global sea surface temperature analyses: multiple problems and their implications for climate analysis, modeling and reanalysis, Bull. Amer. Meteor. Soc., 80, 2661–2678.

Jones, P.D., and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and update to 2001. J. Climate, 16, 206–223.

Jones, P.D., et al., 1990: Assessment of urbanization effects in time series of surface air temperature over land. Nature, 347, 169–172.

Jones, P.D., T.J. Osborn and K.R. Briffa, 1997: Estimating sampling errors in large-scale temperature averages. J. Climate, 10, 2548-2568.

Jones, P.D., et al., 1999: Surface air temperature and its changes over the past 150 years. Rev. Geophys., 37, 173–199.

Jones, P.D., et al., 2001: Adjusting for sampling density in grid-box land and ocean surface temperature time series. J. Geophys. Res., 106, 3371–3380.

Kalnay, E., et al., 1996: The NCEP/NCAR Reanalysis Project. Bull. Amer. Meteor. Soc., 77, 437–471.

Kanamitsu, M., et al., 2002: An overview of NCEP-DOE AMIP-II Reanalysis (R-2). Bull. Amer. Meteor. Soc., 83, 1631–1643.

Kaplan, A., et al., 1997: Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. J. Geophys. Res., 102, 27835–27860.

Karl, T.R., et al., 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum, and mean temperature for the United States. J. Clim. Appl. Meteor., 25, 145–160.

Karl, T.R., H.F. Diaz, and G. Kukla, 1988: Urbanization: its detection and effect in the United States climate record. J. Climate, 1, 1099–1123.

Kent, E.C. and P.G. Challenor, 2006: Toward estimating climatic trends in SST data, part 2: Random errors. J. Atmos. Oceanic Technol., 23, 476–486.

Kent, E.C. and A. Kaplan, 2006: Toward estimating climatic trends in SST data, part 3: Systematic biases. J. Atmos. Oceanic Technol., 23, 487-500.

Kistler, R., et al., 2001: The NCEP-NCAR 50-year reanalysis: Month means CD-ROM and documentation. Bull. Amer. Meteor. Soc., 82, 247–268.

Lanzante, J.R., S.A. Klein, and D.J. Seidel, 2003a: Temporal homogenization of monthly radiosonde temperature data. Pt I: Methodology. J. Climate, 16, 224–240.

Lanzante, J.R., S.A. Klein, and D.J. Seidel, 2003b: Temporal homogenization of monthly radiosonde temperature data. Pt II: Trends, sensitivities, and MSU comparison. J. Climate, 16, 241–262.

Lawrence, S.P., D.T. Llewellyn Jones, and S.J. Smith, 2004: The measurement of climate change using data from the Advanced Very High Resolution and Along Track Scanning Radiometers. J. Geophys. Res., 109, C08017, doi:10.1029/2003JC002104.

Legates, D.R., and C.J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected, global precipitation, Int. J. Climatol., 10, 111–127.

Mears, C.A., M.C. Schabel, and F.J. Wentz, 2003: A reanalysis of the MSU channel 2 tropospheric temperature record. J. Climate, 16, 3650–3664.

Mears, C.A. and F.J. Wentz, 2005: The effect of diurnal correction on satellite-derived lower tropospheric temperature. Science, 309, 1548-1551.

Mekis, E., and W.D. Hogg, 1999: Rehabilitation and analysis of Canadian daily precipitation time series, Atmosphere-Ocean, 37, 53−85.

Menne, M.J. and C.N. Williams, Jr., 2005: Detection of undocumented change points: On the use of multiple test statistics and composite reference series. J. Climate, 18, 4271−4286.

O'Carroll, A.G., R.W. Saunders, and J.G. Watts, 2006: The measurement of the sea surface temperature climatology by satellites from 1991 to 2005. J. Atmos. Oceanic Tech.,23, 1573-1582.

Parker, D.E., 2004: Large-scale warming is not urban. Nature, 432, 290–290.

Parker, D.E., 2006: A demonstration that large-scale warming is not urban. J. Climate, 19, 2882-2895.

Parker, D.E. and D.I. Cox, 1995: Towards a consistent global climatological rawinsonde data-base. Int. J. Climatol., 15, 473–496.

Peterson, T.C., 2003: Assessment of urban versus rural in situ surface temperatures in the contiguous United States: no difference found. J. Climate, 16, 2941–2959.

Peterson, T.C. and T.W. Owen, 2005: Urban heat island assessment: Metadata are important. J. Climate, 18, 2637–2646.

Poncelet, L., 1959: Comparison of raingauges. WMO Bull., 8(4), 186–190.

Rayner, N.A., et al., 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J. Geophys. Res., 108, 4407, doi:10.1029/2002JD002670.

Rayner, N.A., et al., 2006: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset. J. Climate, 19, 446-469.

Reynolds, R.W., et al., 2002: An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609–1625.

Reynolds, R.W., C.L. Gentemann, and F. Wentz, 2004: Impact of TRMM SSTs on a climate-scale SST analysis. J. Climate, 17, 2938–2952.

Schmidli, J., et al., 2002: Mesoscale precipitation in the Alps during the 20th century. Int. J. Climatol., 22, 1049–1074.

Schroder, T., et al., 2003: Validating the microwave sounding unit stratospheric record using GPS occultation. Geophys. Res. Lett., 30, 1734, doi:10.1029/2003GL017588.

Sevruk, B., 1982: Methods of correction for systematic error in point precipitation measurement for operational use. WMO Oper. Hydrol. Rep., 21, Publ. 589, World Meteorol. Org., Geneva, Switzerland, 91 pp.

Sevruk, B., and W.R.. Hamon, 1984: International comparison of national precipitation gages with a reference pit gage. WMO Instruments and Observing Methods Report No. 17, World Meteorol. Org., Geneva, 111 pp.

Sherwood, S., J. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late 20 th century warming. Science, 309, 1556-1559.

Simmons, A.J., et al., 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.

Slutz, R.J., et al., 1985: COADS: Comprehensive Ocean-Atmosphere Data Set. Release 1. NOAA Climate Res. Progr., Env. Res. Lab., Boulder, CO, 262 pp.

Smith, T.M., and R.W. Reynolds, 2002: Bias corrections for historical sea surface temperatures based on marine air temperatures. J. Climate, 15, 73–87.

Smith, T.M., and R.W. Reynolds, 2004: Improved extended reconstruction of SST (1854–1997). J. Climate, 17, 2466–2477.

Smith, T.M., and R.W. Reynolds, 2005: A global merged land and sea surface temperature reconstruction based on historical observations (1880–1997). J. Climate, 18, 2021–2036.

Spencer, R.W., and J.R. Christy, 1992: Precision and radiosonde validation of satellite gridpoint temperature anomalies, Part II: A tropospheric retrieval and trends, 1979-90. J. Climate, 5, 858-866.

Thorne P.W., et al., 2005: Revisiting radiosonde upper air temperatures from 1958 to 2002, J. Geophys. Res., 110, D18105, doi:10.1029/2004JD005753.

Uppala, S.M., et al., 2005: The ERA-40 reanalysis. Quart. J. Roy. Meteor. Soc., 131, 2961-3012.

Vincent, L.A., 1998: A technique for the identification of inhomogeneities in Canadian temperature series. J. Climate, 11, 1094–1104.

Vincent, L.A., et al., 2002: Homogenization of daily temperatures over Canada. J. Climate, 15, 1322–1334.

Vinnikov, K.Y., et al., 2004: Analysis of diurnal and seasonal cycles and trends in climatic records with arbitrary observation times. Geophys. Res. Lett., 31, L06205, doi:10.1029/2003GL019196.

Vinnikov, K.Y., et al., 2006: Temperature trends at the surface and in the troposphere. J. Geophys. Res., 111, D03106, doi:10.1029/2005JD006392.

Vose, R.S., et al., 2004: Impact of land-use change on climate. Nature, 427, 213–214.

Vose, R.S., et al., 2005: An intercomparison of surface air temperature analyses at the global, hemispheric and grid-box scale. Geophys. Res. Lett., 32, L18718, doi:10.1029/200GL023502.

Wentz, F.J., and M. Schabel, 1998: Effects of satellite orbital decay on MSU lower tropospheric temperature trends. Nature, 394, 661–664.

Wentz, F.J., et al., 2000: Satellite measurements of sea surface temperature through clouds. Science, 288, 847–850.

Wijngaard, J.B., A.M.G. Klein Tank, and G.P. Können, 2003: Homogeneity of 20th century European daily temperature and precipitation series. Int. J. Climatol, 23, 679–692.

Woodruff, S.D., et al., 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. Phys. Chem. Earth, 23, 517–527.

Worley, S.J., et al., 2005: ICOADS release 2.1 data and products. Int. J. Climatol., 25, 823–842.

Yang, D., 1999: An improved precipitation climatology for the arctic ocean. Geophys. Res. Lett., 26, 1625–1628.

Yang, D., and T. Ohata, 2001: A bias corrected Siberian regional precipitation climatology. J. Hydrometeorol., 2, 122–139.

Yang, D., et al., 1999: Bias correction of precipitation data for Greenland. J. Geophys. Res.-Atmospheres, 105, 6171–6182.

Zolina, O., et al., 2005: On the robustness of the estimates of centennial-scale variability in heavy precipitation from station data over Europe. Geophys. Res. Lett., 32, doi:10.1029/2005GL023231.