Tellus SERIES A DYNAMIC METEOROLOGY AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2008), 60A, 598–603

TELLUS

Uncertainty in atmospheric temperature analyses

By ROLF H. LANGLAND[1]*, RYAN N. MAUE[2] and CRAIG H. BISHOP[1], [1]*Marine Meteorology Division, Naval Research Laboratory, Monterey, CA 93943, USA;* [2]*Department of Meteorology, Florida State University, Tallahassee, FL 32306, USA*

(Manuscript received 21 February 2008; in final form 28 April 2008)

ABSTRACT

This report illustrates and quantifies the unanticipated large uncertainty and differences in tropospheric temperature analyses within current global operational forecast systems and historical re-analysis products. Results reveal that regional patterns of uncertainty in seasonally averaged and daily atmospheric upper-air temperature analyses are related to the irregular distribution of in situ and satellite observations. There is *less* uncertainty in analysed temperature where in situ radiosonde observations are plentiful, primarily over developed nations in the Northern Hemisphere, and more uncertainty over regions that are observed primarily by satellites with fewer in situ temperature observations, including oceanic areas, the cryosphere, and developing nations. The results suggest that operational weather forecasting and climate monitoring would benefit from an improved global observing network, including additional in situ components. There is also a need for progress in data assimilation to extract more information from the wealth of current and future satellite observations.

1. Introduction

Accurate observation and analysis of atmospheric temperature is essential for numerical weather prediction (NWP) and the monitoring of regional and global climate change. Multivariate atmospheric models provide the background forecasts and dynamical context in which analyses of atmospheric temperatures, winds, and other variables are produced. The observations ingested by data assimilation procedures supply new information that is needed to continuously adjust the model trajectory toward a new analysis, which is an *estimate* of the true atmospheric state. Many factors affect the quality of atmospheric temperature analyses, including the variable quality of satellite observations (Christy et al., 2003; Santer et al., 2003a), the irregular distribution of important observation types, including radiosondes (Lanzante et al., 2003) and limitations of data assimilation procedures (Rabier, 2005).

For NWP, multivariate atmospheric analyses provide initial conditions and verification for deterministic and probabilistic forecasts over the global domain. To the extent that the means and covariances of forecast and observation errors are accurately specified in the data assimilation systems used to produce analyses, these analyses represent state estimates with minimum error variance or maximum likelihood—depending on the type of scheme employed (Daley, 1991). However, because the data assimilation schemes and models used to produce high-resolution estimates of the atmospheric state are constantly evolving and do not typically provide estimates of their own uncertainty, analyses have not found much use in the detection of climate change.

Instead, climate temperature trends have more typically been inferred from the observational record. This has limited such studies to regions (or vertical levels) where observations are sufficient in number and accuracy to provide good estimates of the 'true temperature.' Traditionally, the network of radiosonde upper air observations (Fig. 1) and in situ surface observations have provided the best foundation for such evaluation of climate trends. Limitations of radiosonde observation quality in the context of climate studies are discussed by Sherwood et al. (2005), Randel and Wu (2006) and Trenberth et al. (2007).

However, over much of the globe, in situ temperature observations are sparse, and satellite-based instruments provide the majority of surface and upper-air temperature observations, primarily from microwave and infrared radiometers and sounders aboard polar-orbiting and geostationary platforms. An important issue is that with satellite-derived temperature observations, the problems of bias correction, observation error and quality control are generally more complex than with in situ temperature observations such as those provided by radiosondes.

Bias in satellite-derived observations can vary with time (diurnally and seasonally), geographic location, the underlying surface (land, sea or ice) and scan position of the satellite (Aulign

*Corresponding author.
e-mail: rolf.langland@nrlmry.navy.mil
DOI: 10.1111/j.1600-0870.2008.00336.x

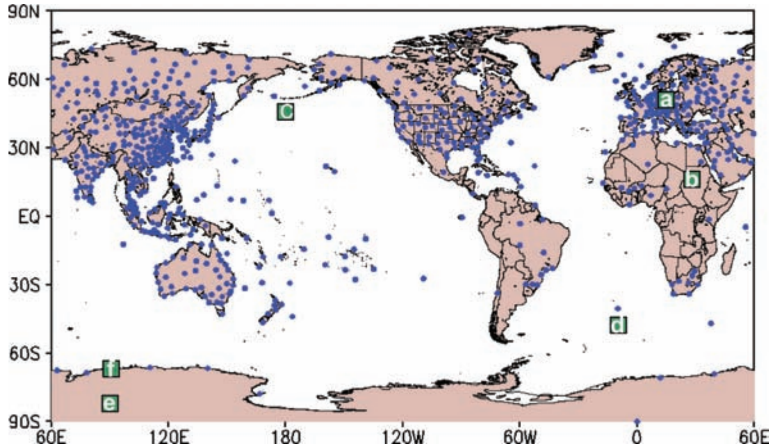

Fig. 1. Radiosonde coverage at all observation times for approximately 700 station locations, from data ingested into the ERA-40 the week of January 1, 2001. Letters a-f indicate locations of vertical profiles in Fig. 4.

et al., 2007). Also, bias in satellite data cannot generally be assumed as constant over time due to orbital decay, drift, and other issues (Christy et al., 2003; Trenberth et al., 2006). In most cases, satellite observations, including radiances, include a dependence on background information from the model, which is a source of additional bias (Dee, 2005). Current multivariate operational data assimilation schemes utilize sophisticated statistical techniques to identify and reject suspect observations and to detect and correct satellite observation bias (Harris and Kelly, 2001).

It is a stated objective that future climate studies should use atmospheric temperature reanalyses (Folland et al., 2006). Reanalysis data sets do not contain spurious trends due to changes in forecast models or data assimilation technique, because these remain the same for the entire re-analysis period. However, reanalysis products can include biases caused by changes in the global observing system over time, which can reduce the reliability of inferred climate trends (Bengtsson et al., 2004; Santer et al., 2004; Sterl, 2004; Trenberth and Smith, 2006). A disadvantage of state estimates provided by reanalyses is that they are produced at coarser horizontal and vertical resolution than analyses produced for short to medium-range NWP. In this report, we examine the differences and uncertainty that exist in temperature analyses from several operational and re-analysis products.

2. Comparison of operational temperature analyses

How much uncertainty exists in current operational analyses of upper-air temperature? We can investigate this question by comparing the analyses of temperature produced by independent numerical weather prediction centres, including the National Centers for Environmental Prediction (NCEP), the United Kingdom Meteorological Office (UKMO), and the Fleet Numerical Meteorology and Oceanography Center (FNMOC). These centres generate atmospheric analyses several times each day and maintain state-of-the-art data assimilation procedures and high-resolution forecast models.

Each forecast centre assimilates several million or more surface and upper-air observations (including satellite radiances) every day to produce the analyses, which are used to initialize their respective deterministic and ensemble forecast systems, and are used as a primary tool for model verification. The temperature analyses produced by each assimilation and forecast system are independent estimates of the ‘true temperature.’ We do not expect these various temperature analyses to be identical, because of differences in data assimilation methods and in the forecast models that provide the background for the analyses. Each forecast centre also uses somewhat different procedures for observation quality control, bias correction, and assimilates a somewhat different selection of observation types.

In Figure 2A we illustrate the difference of the 6-month mean (DM) analysed 500 hPa temperature produced by the FNMOC Navy Operational Global Atmospheric Prediction System (NOGAPS) and the NCEP Global Forecast System (GFS). These results are based on daily analyses at 0000 UTC and 1200 UTC from January 1 to June 30, 2007. The root mean of the squared daily differences (RMSD) in analysed 500 hPa temperature from NOGAPS and GFS is shown in Fig. 2B. The RMSD indicates day-to-day variance between the two temperature analyses. Where the RMSD is large, there is greater uncertainty in our ability to quantify the ‘true temperature.’

From comparison of Figs. 2A and B with Fig. 1, it is clear that differences in analysed temperatures (e.g. the uncertainties in our knowledge of the ‘true temperature’) are closely related to the distribution of radiosonde observations. Thus, the uncertainty in temperature analyses is significantly less over Northern Hemispheric continental regions, Australia, and various island locations, such as Hawaii and New Zealand where radiosonde observations are plentiful. For example, the RMSD of mean analysed 500 hPa temperature (Fig. 2B) is less than 0.9 K over most of North America, Europe and northern Asia, but is generally more than 1.3 K (and in places as large as 2.5 K) over the

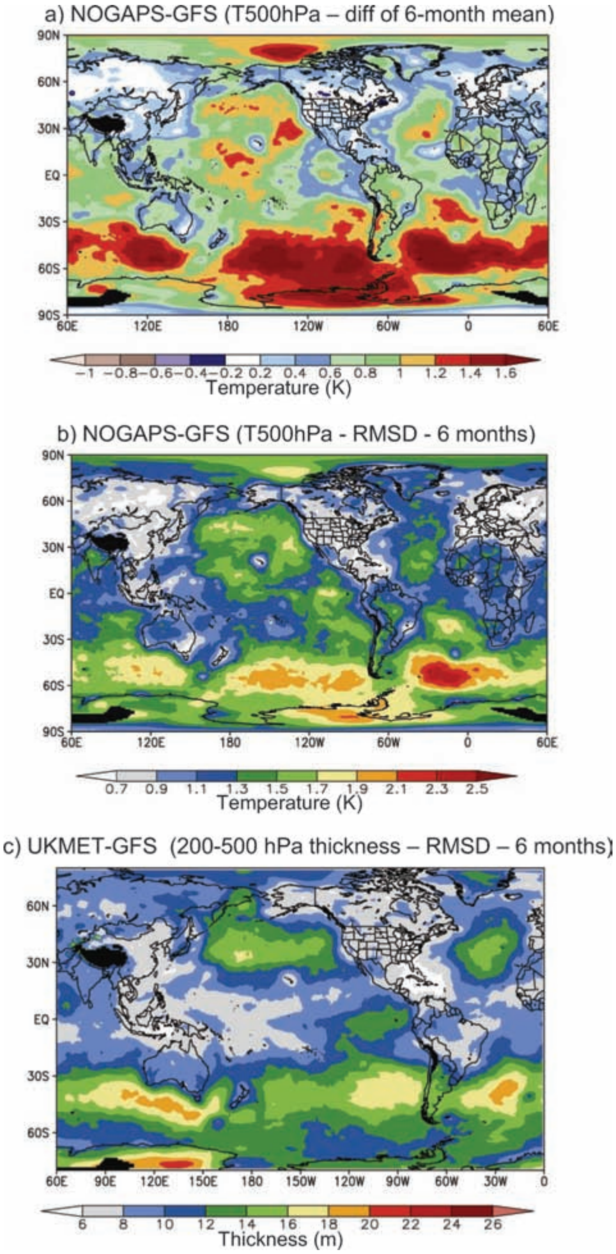

Fig. 2. (A) Difference (DM) of 6-month mean analysed 500 hPa temperature, K (B) root mean square difference (RMSD) of daily 500 hPa temperature, K (C) RMSD of daily 200–500 thickness (m). All panels based on analyses at 0000 UTC and 1200 UTC from January 1 to June 30, 2007.

Pacific and Atlantic oceans and polar regions where satellite observations are the primary source of temperature observations. Temperature analyses over most developing nations, including those in South America and Africa, are generally more uncertain than those over developed nations in Europe, Asia and North America.

Figure 2C illustrates the RMSD of the analysed thickness between the 200 and 500 hPa pressure levels in the UKMET global model and the GFS. This quantity is relevant for studies of tropopause height as discussed, for example, by Santer et al. (2003b). Again, we see that the two analyses are in much closer agreement over regions with radiosonde coverage. Differences between the two analyses are much larger over the southern hemisphere, and in mid-Atlantic and mid-Pacific.

3. Comparison of temperature reanalyses

It is also instructive to examine differences in temperature analyses provided by re-analysis products, including the 'second-generation' Japanese 25-yr Re-analysis Project (JRA-25, Onogi et al., 2007), and the European Center for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40, Uppala et al., 2005). Re-analysis products, similar to those produced by operational center data assimilation systems, require a sequence of short-range forecasts (usually 6–12 h) to generate background information, which is combined with new observations to generate an analysis. Because they are produced outside the constraints of operational NWP, reanalyses generally include a somewhat larger number of observations than the routine daily analyses. However, the same concerns related to correction of bias in satellite observations that apply in operational NWP also affect the quality of re-analysis products.

Figure 3 illustrates the RMSD of analysed temperatures during calendar year 2001 (the ERA-40 data ends in August 2002), on the standard pressure levels of (a) 500 hPa, (b) 300 hPa and (c) 100 hPa. It is evident from Fig. 3 that re-analysis temperatures also have substantially larger uncertainty in regions without plentiful radiosonde observations, as was shown in Fig. 2 for operational temperature analyses. For example, the RMSD of analysed temperature at 300 hPa (Fig. 3B) over parts of Antarctica is as large as 3.5 K, more than three times larger than values typical over North America, Europe, Australia and eastern Asia.

It is important to consider the uncertainty of analysed temperatures at different levels in the atmosphere, since differences in tropospheric and stratospheric temperature trends may be a fingerprint of anthropogenic climate change (Santer et al., 2004). Vertical profiles of the RMSD of analysed temperature at six selected global locations are shown in Fig. 4, using the JRA-25, ERA-40 and NCEP/NCAR (National Center for Atmospheric Research) re-analysis products. Note that for these reanalyses, the ERA40 and JRA-25 assimilated satellite radiances, while NCEP/NCAR used retrievals of satellite brightness temperature.

Comparatively small uncertainty in analysed temperature, about 0.5 K through most of the troposphere, is found at Berlin (Fig. 4A), which is representative of regions with dense radiosonde and aircraft observations (Fig. 1). In locations that have sparse radiosonde or other in situ observations, there is substantially more uncertainty in analysed temperature throughout

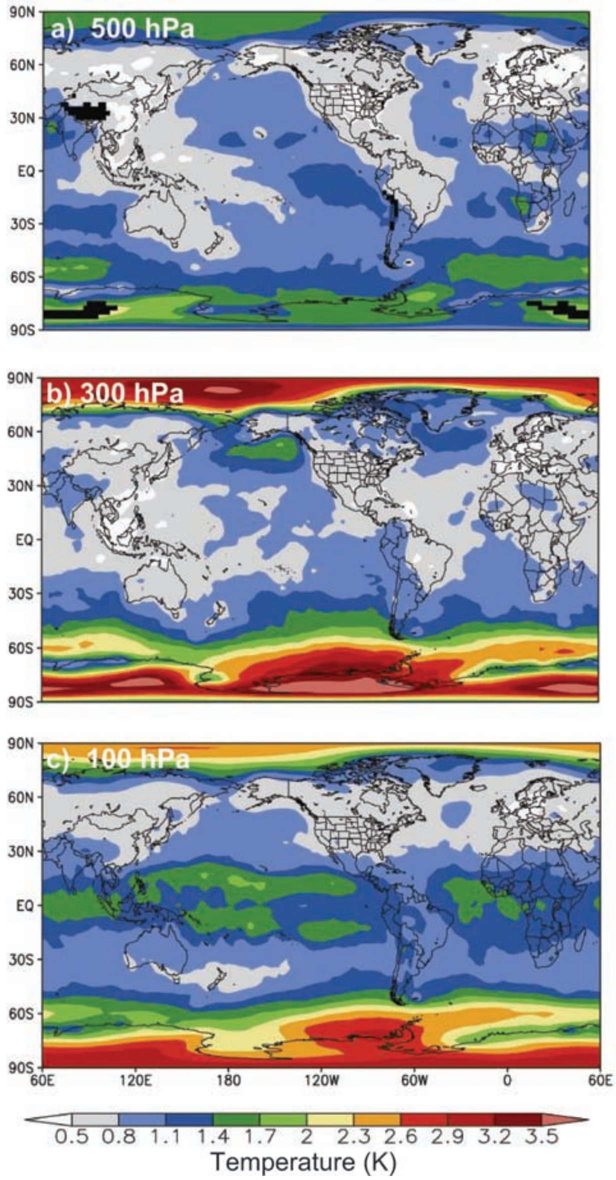

Fig. 3. Root mean square differences of analysed temperatures from the JRA25 and ERA40 reanalyses. (A) 500 hPa, (B) 300 hPa and (C) 100 hPa. All panels based on analyses at 0000 UTC and 1200 UTC for year-2001.

the atmospheric column. For example, the RMSD of lower-tropospheric temperature at a location in northeastern Africa (Fig. 4B) is about 1K, or twice that found at Berlin, and even larger below 700 hPa. At typical ocean locations, there are few or no radiosondes, and the uncertainty of analysed tropospheric temperature at two selected locations (Figs. 4C and D) is 1.5–2.0 K through most of the troposphere. Two profiles are included to illustrate uncertainty in analysed temperature at Antarctic locations: one is an interior location without a radiosonde station (Fig. 4E, note extension of the temperature scale in this panel) and the other colocated with a radiosonde (Fig. 4F). The RMSD is 2–3 K at the interior location (Fig. 4E) and 1–1.5 K at the location near a radiosonde (Fig. 4F). Thus substantial differences in the uncertainty of upper-air temperature analyses are found over various regions of Antarctica.

4. Discussion

The goal of this report is to illustrate several issues related to analyses of atmospheric temperature, which may be used to verify operational weather forecasts or in climate change research. Each analysis is an independent *estimate* of the 'true temperature', and it is assumed that each forecast or re-analysis centre produces temperature analyses that are equally likely representations of the 'true temperature'. There is no definitive way to determine which analysis is closer to the truth, or to construct an analysis of the 'true temperature', because we do not have exact information about observation or analysis error. The actual uncertainty in current temperature analyses is likely to be even *larger* than the differences among available atmospheric temperature analyses shown here, because the various operational forecast centres use shared sets of observations, and the data assimilation procedures are based on similar methods.

We have documented statistics of differences between temperature analyses, since these differences provide estimates of the uncertainty in our knowledge of the 'true temperature'. Based upon examination of several well-known products, we find that uncertainty in current analyses of atmospheric temperature has substantial regional variability. The least reliable temperature analyses are generally located over regions which have fewer radiosonde observations, including polar areas, oceanic regions, and developing nations, where accurate temperature monitoring is critical to quantify the effects of climate change. These results are consistent with Swanson and Roebber (2008), who found that differences in NCEP and ECMWF re-analysis 500 hPa heights are relatively large over the North Pacific Ocean, with effects on forecast skill into the medium-range.

In a global context, satellite-based observations now provide the majority of information for tropospheric and stratospheric temperature. There are no regions of the globe that can today be considered 'data-void' or even 'data-sparse.' While there is no question that satellite observations have improved the quality of atmospheric analyses over previously data-sparse regions, it is evident that greater uncertainty in upper-air temperature analyses exists where satellite observations predominate, compared to regions with plentiful radiosonde observations.

In regions where accurate in situ temperature observations are not available, efforts to quantify fingerprints of climate change should rely on the best available estimates of the 'true temperature,' which are likely to be analyses produced by state-of-the-art data multivariate data assimilation systems. However, in some regions, the uncertainty in these analyses may be too large to permit the detection of relatively small temperature signals

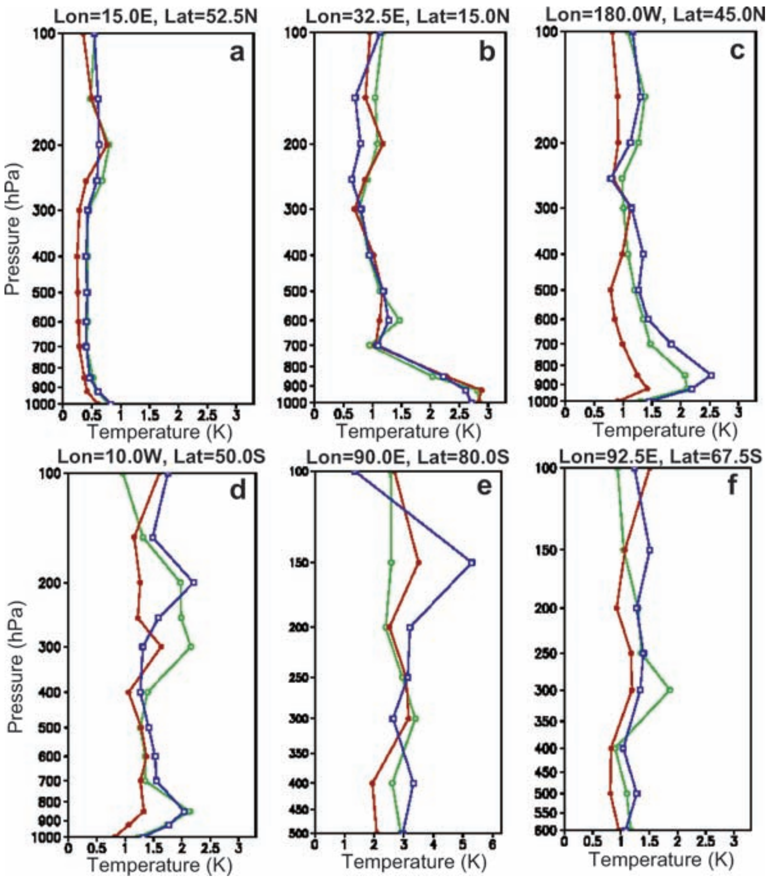

Fig. 4. Vertical profiles of RMSD temperature in reanalysis products for year-2001 at selected locations (shown in Fig. 1): (A) near Berlin, Germany, (B) northeast Africa, (C) north Pacific Ocean, (D) south Atlantic Ocean, (E) interior Antarctica and (F) coastal Antarctica (at radiosonde location). The three lines show the RMSD for: ERA-40 and NCEP/NCAR (blue), JRA-25 and ERA-40 (red), and JRA-5 and NCEP/NCAR (green). The *x*-axis is temperature (K) and the *y*-axis is pressure in hPa. Note extended temperature scale (to 6 K) on panel e. Panels e and f extend to respectively, 500 hPa and 600 hPa, which approximates the surface pressure. Obtained from 2.5° resolution reanalysis fields, each profile is an average for a 5° × 5° latitude–longitude square centred on the given location.

associated with climate change. It is problematic to assume that sets of satellite observations can be adequately analysed for use in climate studies outside the context of these multivariate data assimilation procedures. For example, accurate temperature data may not necessarily be obtained by simple averaging of millions of raw satellite observations (Wentz and Schabel, 2000) because this will not optimally account for observation bias and quality control.

The results also illustrate the need for an improved global atmospheric observing system, as suggested by Trenberth et al. (2002) and Seidel et al. (2004). The future global observing system should include an optimal mix of in situ and satellite observations, and support the needs of routine numerical weather prediction as well as the monitoring of climate change. Additional resources should also be directed towards more optimal use of current and future satellite observations and improvements in data assimilation procedures.

5. Acknowledgments

RL and CB are supported by the Naval Research Laboratory and the Office of Naval Research. RM is supported by a NASA Earth Science Systems Fellowship and acknowledges the Center for Ocean-Atmospheric Prediction Studies (COAPS) at Florida State University. RM collaborated with RL and CB during a visit to the Naval Research Laboratory (NRL), Marine Meteorology Division in Monterey, CA, funded by the ONR summer intern program [2007].

6. Appendix

A.1. Sources of temperature analyses

The NCEP-GFS temperature analyses were obtained from NOAA at: nomad5.ncep.noaa.gov/. The FNMOC-NOGAPS temperature analyses were obtained from the Naval Research Laboratory (the corresponding author can be contacted for additional information). The temperatures have been interpolated from the original model grids to surfaces of constant pressure and 1° latitude–longitude output grids, at 0000 UTC and 1200 UTC. Information about forecast systems and data assimilation procedures is available at: www.emc.ncep.noaa.gov/modelinfo/ (NCEP-GFS), and www.nrlmry.navy.mil/nogaps_his.htm (FNMOC-NOGAPS). The temperatures used to produce Figs. 3 and 4 were obtained from the NCAR mass storage system: dss.ucar.edu/datasets/ (NCEP/NCAR, ERA-40) and the Japanese 25-yr Re-analysis Project archive: www.jreap.org/indexe.html (JRA-25) on constant pressure surfaces and 2.5° latitude–longitude grids.

A.2. Statistical procedures

The DM is the difference of mean temperatures, defined as: $\frac{1}{n}\left(\sum T_1 - \sum T_2\right)$, where T_1 and T_2 are analysed temperatures provided on a discrete grid from two different assimilation/forecast systems, and n is the number of analyses in the summation. The RMSD is the root mean of the sum of the squared differences between two temperatures, defined as: $\sqrt{\frac{1}{n}\sum (T_1 - T_2)^2}$, where T_1 and T_2 are analysed temperatures from two different assimilation/forecast systems and n is the number of analyses in the summation.

References

Auligné, T., McNally, A. P. and Dee, D. P. 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.* **133**, 631–642.

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

Christy, J. R., Spencer, R. W., Norris, W. B. and Braswell, W. D. 2003. Error estimates of Version 5.0 of MSU–AMSU bulk atmospheric temperatures. *J. Atmos.Ocean Tech.* **20**, 613–629.

Daley, R. 1991. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, UK , 457 pp.

Dee, D. P. 2005. Bias and data assimilation. *Q. J. R. Meteorol. Soc.* **131**, 3323–3343.

Folland, C. K., Parker, D.E., Reynolds, R.W., Sherwood, S.C. and Thorne, P.W. 2006. What measures can be taken to improve the understanding of observed changes? In: *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences* (eds T. R. Karl, S. J. Hassol, C. D. Miller and W. L. Murray). A report by the Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC , Chapter 6, 119–127.

Harris, B.A. and Kelly, G. 2001. A satellite radiance-bias correction scheme for data assimilation. *Q. J. Roy. Meteorol. Soc.* **127**, 1453–1468.

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

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., and co-authors. 2007. The JRA-25 reanalysis. *J. Meteorol. Soc. Jpn.* **85**, 369–432.

Rabier, F. 2005. Overview of global data assimilation developments in numerical weather prediction centres. *Q. J. R. Meteorol. Soc.* **131**, 3213–3233.

Randel, W. J. and Wu, F. 2006. Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *J. Climate* **19**, 2094–2104.

Santer, B. D. Wigley, T.M.L., Meehl, G.A., Wehner, M.F., Mears, C., and co-authors. 2003a. Influence of satellite data uncertainties on the detection of externally forced climate change. *Science* doi:10.1126/science.1082393.

Santer, B. D. Wehner, M.F., Wigley, T.M.L., Sausen, R., Meehl, G.A., and co-authors. 2003b. Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science* **301**, 479–483, doi:10.1126/science.1084123.

Santer, B. D., Wigley, T.M.L., Simmons, A.J., Källberg, P.W., Kelley, G.A., and co-authors. 2004. Identification of anthropogenic climate change using a second-generation reanalysis. *J. Geophys. Res.* **109**, D21104. doi:10.1029/2004JD005075.

Seidel, D. J., Angell, J.K., Christy, J., Free, M., Klein, S.A., and co-authors. 2004. Uncertainty in signals of large-scale climate variations in radiosonde and satellite upper-air temperature datasets. *J. Climate* **17**, 2225–2240.

Sherwood, S., Lanzante, J. and Meyer, C. 2005. Radiosonde daytime biases and late-20th century warming. *Science* doi:10.1126/science.1115640.

Sterl, A. J. 2004. On the (in)homogeneity of reanalysis products. *J. Climate* **17**, 3866–3873.

Swanson, K. L. and Roebber, P. J. 2008. The impact of analysis error on medium range weather forecasts. *Mon. Wea. Rev.* doi:10.1175/2008MWR2475.

Trenberth, K. E. and Smith, L. 2006. The vertical structure of temperature in the tropics: different flavors of El Niño. *J. Climate* **19**, 4956–4973.

Trenberth, K. E., Karl, T. R. and Spence, T. W. 2002. The need for a systems approach to climate observations. *Bull. Am. Meteorol. Soc.* **83**, 1593–1602.

Trenberth, K.E., Moore, B., Karl, T. R. and Nobre, C. 2006 Monitoring and prediction of the Earth's climate: a future perspective. *J. Climate* **19**, 5001–5008.

Trenberth, K.E., Jones, P.D., Ambenje, P., Bojariu, R., Easterling, D., and co-authors. 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. Cambridge University Press, Cambridge, UK .

Uppala, S. M., Källberg, P.W., Simmons, A.J., Andrae, U., da Costa Bechtold, V., and co-authors. 2005. The ERA-40 reanalysis. *Q. J. R. Meteorol. Soc.* **131**, 2961–3012.

Wentz, F. J. and Schabel, M. 2000. Precise climate monitoring using complementary satellite data sets. *Nature* **403**, 414–416.