Independent confirmation of global land warming without the use of station temperatures

Gilbert P. Compo,[1,2] Prashant D. Sardeshmukh,[1,2] Jeffrey S. Whitaker,[2] Philip Brohan,[3] Philip D. Jones,[4,5] and Chesley McColl[1,2]

Received 11 February 2013; revised 27 March 2013; accepted 28 March 2013; published 17 June 2013.

[1] Confidence in estimates of anthropogenic climate change is limited by known issues with air temperature observations from land stations. Station siting, instrument changes, changing observing practices, urban effects, land cover, land use variations, and statistical processing have all been hypothesized as affecting the trends presented by the Intergovernmental Panel on Climate Change and others. Any artifacts in the observed decadal and centennial variations associated with these issues could have important consequences for scientific understanding and climate policy. We use a completely different approach to investigate global land warming over the 20th century. We have ignored all air temperature observations and instead inferred them from observations of barometric pressure, sea surface temperature, and sea-ice concentration using a physically based data assimilation system called the 20th Century Reanalysis. This independent data set reproduces both annual variations and centennial trends in the temperature data sets, demonstrating the robustness of previous conclusions regarding global warming. **Citation:** Compo, G. P., P. D. Sardeshmukh, J. S. Whitaker, P. Brohan, P. D. Jones, and C. McColl (2013), Independent confirmation of global land warming without the use of station temperatures, *Geophys. Res. Lett.*, *40*, 3170–3174, doi:10.1002/grl.50425.

1. Introduction

[2] The observed increase in near-surface air temperature over land (2 m air temperature, hereafter TL_{2m}) is a core indicator of global warming [*Trenberth et al.*, 2007]. However, the accuracy of data sets documenting the increase continues to be debated [*Pielke et al.*, 2007; *Fall et al.*, 2011; *Montandon et al.*, 2011; *Christy*, 2012], mainly because the TL_{2m} record consists of observations taken irregularly in space and time using a variety of instruments and measurement techniques [*Peterson et al.*, 1998; *Hansen et al.*, 2010; *Vose et al.*, 2012a; *Brohan et al.*, 2006; *Parker*, 2011; *Karl et al.*, 1986; *Jones and Wigley*, 2010]. The observation sites are not fully representative of the global land cover, tending to be weighted more toward urban and crop settings [*Montandon et al.*, 2011], and represent their larger scale environment to varying degrees [*Pielke et al.*, 2007; *Christy*, 2012; *Fall et al.*, 2011]. Instrument relocations introduce further inhomogeneities in the record [*Peterson et al.*, 1998; *Brohan et al.*, 2006]. Changes in some sites resulting, for example, from conversion to cropland [*Montandon et al.*, 2011], from the construction of buildings [*Christy*, 2012], or other urbanization effects [*Hausfather et al.*, 2013; *Christy*, 2012; *Parker*, 2011; *Jones and Wigley*, 2010; *Brohan et al.*, 2006; *Peterson et al.*, 1998] are also a concern [*Pielke et al.*, 2007; *Christy*, 2012; *Fall et al.*, 2011; *Montandon et al.*, 2011]. These and other heterogeneities affect the accuracy of climate variability deduced from such observations [*Pielke et al.*, 2007; *Christy*, 2012; *Peterson et al.*, 1998; *Brohan et al.*, 2006; *Karl and Williams*, 1987; *Karl et al.*, 1986; *Ellis*, 1890; *Jones and Wigley*, 2010; *Hausfather et al.*, 2013]. Any artifacts in observed climate variations associated with these issues could have important consequences for scientific understanding [*Trenberth et al.*, 2007] and climate policy [*IPCC*, 2007].

[3] Previous global analyses of TL_{2m} have addressed many of these issues [*Jones and Wigley*, 2010; *Trewin*, 2010]. Adjustments to correct for potential biases have been developed [*Peterson et al.*, 1998; *Hansen et al.*, 2010; *Vose et al.*, 2012a; *Brohan et al.*, 2006; *Karl and Williams*, 1987; *Karl et al.*, 1986]. Uncertainties in these adjustments, as well as in the observations and their aggregation into gridded analyses, have been estimated and included in the Climatic Research Unit gridded Temperature version 3 (CRUTEM3) [*Brohan et al.*, 2006] and version 4 (CRUTEM4) [*Jones et al.*, 2012] data sets developed by the University of East Anglia Climatic Research Unit and Met Office Hadley Centre and the Merged Land Ocean Surface Temperature version 3.5.2 (MLOST) data set [*Vose et al.*, 2012a] developed by the National Oceanic and Atmospheric Administration (NOAA). Despite these efforts, debate continues as to the reliability of the global temperature record, particularly for assessing trends and decadal variability [*Pielke et al.*, 2007; *Fall et al.*, 2011; *Montandon et al.*, 2011; *Jones and Wigley*, 2010], as highlighted recently before the United States Congress [*Christy*, 2012].

[4] In light of these issues, we have taken a completely different approach to deduce global 20th century TL_{2m} warming without using any TL_{2m} observations. We use the 20th Century Reanalysis (20CR [*Compo et al.*, 2011], see supporting information), a physically based state-of-the-art data assimilation system, to infer TL_{2m} given only CO_2, solar, and volcanic radiative forcing agents; monthly averaged

Additional supporting information may be found in the online version of this article.

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.
[2]Physical Sciences Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA.
[3]Met Office Hadley Centre, Exeter, UK.
[4]Climatic Research Unit, University of East Anglia, Norwich, UK.
[5]Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia.

Corresponding author: G. P. Compo, CIRES University of Colorado, NOAA Physical Sciences Division, Earth System Research Laboratory, 325 Broadway R/PSD1, Boulder, CO 80305, USA. (compo@colorado.edu)

©2013. American Geophysical Union. All Rights Reserved.
0094-8276/13/10.1002/grl.50425

sea surface temperature (SST) and sea-ice concentration fields (both from the HadISST1.1 [*Rayner et al.*, 2003]); and hourly and synoptic barometric pressure observations (from the International Surface Pressure Databank [*Compo et al.*, 2011]). In our analysis, we find that TL_{2m} warming is reproduced using the independent 20CR data, confirming that the observed warming is not an artifact of deficiencies in station-temperature measurements.

2. Data

[5] Details of data processing are given in the supporting information. Data access is found in Table S1.

2.1. The 20th Century Reanalysis and Simulation

[6] The 20CR is based on an Ensemble Kalman Filter technique [*Whitaker and Hamill*, 2002] and uses a time-varying weighting between pressure observations and an ensemble of 56 nine-hour forecasts made with a NOAA atmosphere/land general circulation model (AGCM) to estimate the three-dimensional state of the atmosphere every 6 h (see supporting information). A detailed assessment of the overall quality has been reported [*Compo et al.*, 2011], and the data set has been used in a wide range of climate and weather applications [reanalyses.org, 2013], including examining the United States [*Vose et al.*, 2012b] and global [*Parker*, 2011] TL_{2m} trend since 1979.

[7] A 138 year set of 56 simulations using the same AGCM, boundary forcing, and radiative forcing agents as 20CR (AMIP20C hereafter) was also generated as an additional independent estimate of TL_{2m} (see supporting information).

2.2. TL_{2m} Observational Data Sets

[8] Eight different near-global data sets constructed from observations of TL_{2m} are used here (Table S1). We have used Climatic Research Unit (CRU) and Met Office Hadley Centre grids of CRUTEM3 [*Brohan et al.*, 2006] and CRUTEM4 [*Jones et al.*, 2012]; the CRU in-filled time series grids (CRU_TS3.10) [*Mitchell and Jones*, 2005; *Jones and Harris*, 2011; *Harris et al.*, 2013]; the NASA Goddard Institute for Space Studies grids at 250 km and 1200 km smoothing (GISTEMP250 and GISTEMP1200) [*Hansen et al.*, 2010]; the NOAA MLOST grids [*Vose et al.*, 2012a]; the Japan Meteorological Agency temperature grids (JMATEMP, JMA, unpublished data 2012); and the University of Delaware temperature grids (UDELv3.01) [*Willmott and Robeson*, 1995].

3. Results

[9] As the 20CR does not use temperature observations from land stations, it is entirely independent of those observations. Nevertheless, the time variations of TL_{2m} in the 20CR are very similar to those previously reported in the station-based data sets [*Brohan et al.*, 2006; *Hansen et al.*, 2010; *Vose et al.*, 2012a; *Jones et al.*, 2012], both over the 1901 to 2010 period and the more rapidly warming 1952 to 2010 period (Figure 1 and Table S2). We have focused on the TL_{2m} anomalies averaged over 90°N–60°S in these two periods because this coverage is common to all the data sets considered (see supporting information). The temporal (Pearson) correlation between the 20CR and station-based estimates of the 1901–2010 globally averaged annual

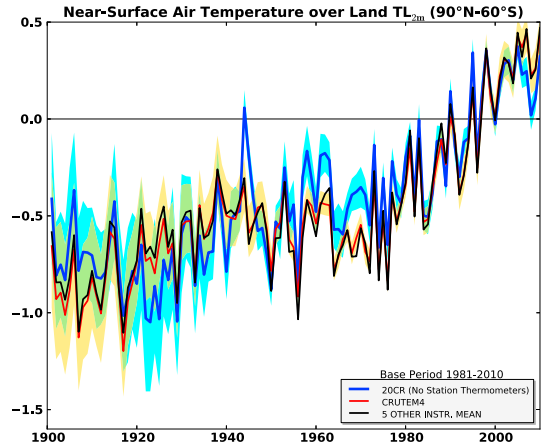

Figure 1. Temporal comparison of near-global land (90°N–60°S) 2 m air temperature anomalies (TL_{2m}) between 20CR and station-temperature based estimates. Red curve: global TL_{2m} anomaly series from CRUTEM4, black curve: the average of five additional station-temperature data sets (marked with asterisk in Table S2), and blue curve: the 20CR. 95% uncertainty ranges are shown for CRUTEM4 (yellow fill) and 20CR (blue fill) and their overlap (green fill).

TL_{2m} range from 0.84 ($P = 0.006$) with the UDELv3.01 data set to 0.92 ($P = 0.001$) with the MLOST data set. These and all temporal significance tests take into account the reduction in temporal degrees of freedom (dofs) arising from the autocorrelation in all series [*Livezey and Chen*, 1983].

[10] A good agreement is seen even after removing the long-term trend and multidecadal variability (Table S2). The high-pass filtered data, constructed by removing 7 year running means from the monthly anomalies, shows correlations over the 1901–2010 period from 0.74 ($P < 0.0001$) with JMATEMP to 0.81 ($P < 0.0001$) with four different data sets (Table S2). The correlations are higher for the shorter 1952–2010 period. All P are 0.0001 or smaller.

[11] This consistency of TL_{2m} from land stations and the 20CR strongly suggests that the determinations of TL_{2m} variations from monthly to centennial scales using the station temperatures are robust and reliable. The comparison is insensitive to the choice of data set (Figure 1 and Table S2) and the precise near-global region over which it is performed (Figure S1 and Table S3).

[12] Despite being based on an independent set of observations, the uncertainty estimates from 20CR are comparable to those from CRUTEM4 (Figure 1) and also from CRUTEM3 and MLOST (Figure S1 and Table S3). The uncertainties in both the 20CR and station estimates are smaller if the comparison is restricted to the 60°N–60°S region (Figure S1 and Table S3). This avoids the sparsely observed Arctic and regions of low sea-ice concentrations that have a warm bias in 20CR [*Compo et al.*, 2011; *Brönnimann et al.*, 2012]. The data sets agree even better over the 60°N–60°S domain (Table S3). For example, over the 1901–2010 period, four data sets correlate with 20CR at 0.91 ($P < 0.0012$).

[13] Although the agreement between 20CR and the station-based data sets is strong, the mean square differences between them are somewhat larger than expected from their respective confidence intervals (Table S3). This suggests

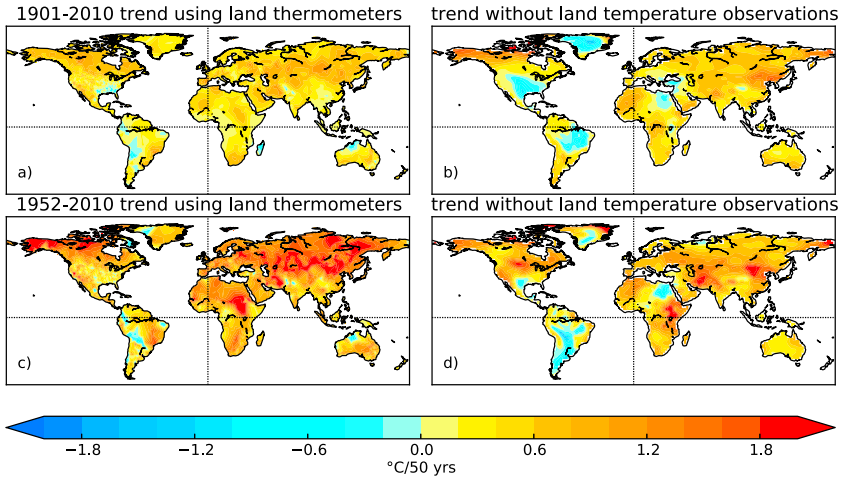

Figure 2. Spatial comparison of 2 m air temperature (TL_{2m}) trends between 20CR and the station-temperature based CRU_TS3.10. Local linear trends estimated using least-squares from (a, c) CRU_TS3.10 and (b, d) 20CR shown as °C change per 50 years over the 1901–2010 (a,b) and 1952–2010 (c,d) periods.

that the data sets underestimate their uncertainty, particularly 20CR during the periods of disagreement in 1944–45, and the 1960s and 1970s (Figure S1).

[14] The same general agreement, with differences in details, is evident in the spatial patterns of the least-squares linear trends (Figures 2, 3, S2, and S3 and Table S2). The pattern correlations [*Miyakoda et al.*, 1972] for trends over 1901–2010 range from 0.67 ($P = 0.035$) with JMATEMP to 0.78 ($P = 0.011$) with MLOST (Table S2). All P values are smaller than 0.035 assuming 8 spatial dofs [*Jones et al.*, 1997]. A few regions, however, do deviate from this general agreement, notably the midwestern United States, eastern Brazil, and Argentina (Figures 2, S2, and S3). If the differences come from random effects, local linear trends (Figures 2, S2, and S3) will be evenly divided between being larger or smaller than the 20CR trend, with an expected 50% areal coverage of larger trends and a binomial sampling distribution. If the differences are systematic, indicating biases in either data set, a different distribution is likely. None of the areal percentages (Table S2) for either period are statistically distinguishable from the 50% expected for a binomial distribution if the spatial dofs of the TL_{2m} trend field are less than 13. Estimates of the spatial dofs range from 3 to 8 [*Jones et al.*, 1997]. Assuming 8 degrees of freedom, all P values are greater than 0.144.

[15] The global average trend of TL_{2m}, computed as the area-weighted average of the local trends (Figures 2, S2, and S3), also shows quantitative agreement (Table S2). The 20CR trend is 0.45°C/50 years for 1901–2010 and 0.67°C/50 years for 1952–2010. None of the trends in Table S2 is significantly different from the corresponding 20CR trend assuming 8 spatial dofs ($P \geq 0.11$).

[16] We have also inferred TL_{2m} from a parallel ensemble of AGCM runs without using any pressure observations (AMIP20C). Though the SSTs are the dominant contributor to land warming, both regionally [e.g., *Compo and Sardeshmukh*, 2009; *Dommenget*, 2009] and globally [*Compo and Sardeshmukh*, 2009; *Dommenget*, 2009; *Hoerling et al.*, 2008], the 20CR agrees better with the station data sets than it does with even this AGCM ensemble (Figure 3), particularly on the monthly timescale (Table S2), confirming the important influence of the pressure observations in 20CR.

[17] Unlike 20CR, the AMIP20C simulation is not intended to and cannot reproduce the observed day-to-day (or month-to-month) variations, i.e., the climate system's particular chaotic "sample path" over the 20th century, and thus cannot be regarded as providing independent *observational* confirmation of the long-term variations (just as successful simulation of 20th century global warming using climate models with prescribed radiative forcings cannot be viewed as providing *observational* confirmation of that warming). This is consistent with the much lower correlation (~0.35) in Table S2 of

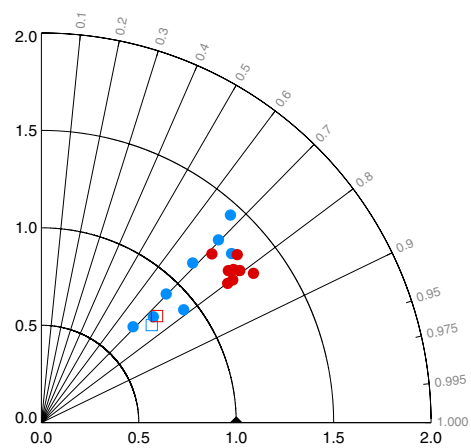

Figure 3. Taylor diagram [*Taylor*, 2001] comparing the spatial patterns and amplitudes of local trends of TL_{2m} from 20CR with eight station-temperature based data sets (filled circles, Table S1) and from SST-forced simulation AMIP20C (open squares). Trends are computed over 1901–2010 (blue symbols) and 1952–2010 (red symbols). The cosine of the angle at which each symbol is plotted is the correlation between the two trend patterns being compared. The radial distance is the ratio of the average amplitudes of the patterns. Perfect correspondence with 20CR would align with the black triangle symbol at a unit radius and an angle of 0°.

20CR with the AMIP20C high-pass monthly TL_{2m} time series than with all the thermometer-based series. On the other hand, that 20CR represents not just the low-frequency but even high-frequency aspects of the observed "sample path" is reflected in its good match with independent estimates of subdaily and monthly tropospheric-average temperatures [*Compo et al.*, 2011], daily temperatures [*Parker*, 2011], monthly high-pass filtered TL_{2m} (Table S2), and also in 24 h forecast skill when used as initial conditions [*Compo et al.*, 2011].

4. Conclusions

[18] The independent estimate of TL_{2m} from 20CR demonstrates that, in spite of recently published [*Pielke et al.*, 2007; *Fall et al.*, 2011; *Montandon et al.*, 2011] and public [*Christy*, 2012] concerns with the station temperature record, the temperature analyses [*Brohan et al.*, 2006; *Hansen et al.*, 2010; *Vose et al.*, 2012a; *Jones et al.*, 2012; *Japan Meteorological Agency*, unpublished data, 2012; *Harris et al.*, 2013] used by the Intergovernmental Panel on Climate Change (IPCC) [*Trenberth et al.*, 2007] and many others for climate science and for input to climate policy [*IPCC*, 2007] are reliable and robust estimates of large-scale TL_{2m} variability and change. Still, while more than 80% of the temporal variation of global land-average temperature (Tables S2 and S3) and 60% of the spatial variation in the century and recent half-century trends are captured (Table S2 and Figure 3), 20CR does show interesting differences in some regions such as the midwestern United States, Argentina, and eastern Brazil (Figures 2, S2, and S3) and during some time periods such as 1944–1945, the 1960s, and 1970s (Figures 1 and S1). Resolving these differences may include addressing previously unrecognized issues with the pressure observations, time-varying land use and land cover [*Nuñez et al.*, 2008; *Pielke et al.*, 2007; *Brohan et al.*, 2006], aerosols [*Parker*, 2011; *Jones et al.*, 2012], remaining inhomogeneities in the station records [*Jones and Wigley*, 2010; *Trewin*, 2010], or the prescribed boundary conditions in the 20CR system.

[19] **Acknowledgments.** All data are publicly available (Table S1). We thank the NOAA ESRL/PSD IT and Data group for support. 20CR used resources at NERSC (DE-AC02-05CH11231) and OLCF (DE-AC0500OR22725) funded by the Department of Energy (DoE) Office of Science. This work was supported by NOAA Climate Program Office and the Office of Science (BER), U.S. Department of Energy. P.B. was supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101). P.D.J. has been supported by the USDoE (Grant DE-SC0005689). We also thank two anonymous reviewers for their comments.

[20] The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *J. Geophys. Res.*, *111*, D12106, doi:10.1029/2005JD006548.

Brönnimann, S., A. N. Grant, G. P. Compo, T. Ewen, T. Griesser, A. M. Fischer, M. Schraner, and A. Stickler (2012), A multi-data set comparison of the vertical structure of temperature variability and change over the Arctic during the past 100 years, *Cli. Dyn.*, *39*, 1577–1598, doi:10.1007/s00382-012-1291-6.

Christy, J. R. (2012), Congressional testimony to the Senate Environment and Public Works Committee 1 August 2012. 22 pages. Online at http://epw.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=66585975-a507-4d81-b750-def3ec74913d.

Compo, G. P., and P. D. Sardeshmukh (2009), Oceanic influences on recent continental warming, *Clim. Dyn.*, *32*, 333–342, doi:10.1007/s00382-008-0448-9.

Compo, G. P., et al. (2011), The Twentieth Century Reanalysis Project, *Q. J. R. Meteorol. Soc.*, *137*, 1–28, doi:10.1002/qj.776.

Dommenget, D. (2009), The ocean's role in continental climate variability and change, *J. Climate*, *22*, 4939–4952, doi:10.1175/2009JCLI2778.1.

Ellis, W. (1890), On the difference produced in the mean temperature derived from daily maximum and minimum readings, as depending on the time at which the thermometers are read, *Q. J. R. Meteorol. Soc.*, *16*, 213–220, doi:10.1002/qj.4970167605.

Fall, S., A. Watts, J. Nielsen-Gammon, E. Jones, D. Niyogi, J. R. Christy, and R. A. Pielke Sr. (2011), Analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends, *J. Geophys. Res.*, *116*, D14120, doi:10.1029/2010JD015146.

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

Harris, I., P. S. Jones, T. J. Osborn, and D. H. Lister (2013), Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset, *Int. J. Climatol.*, in press.

Hausfather, Z., M. J. Menne, C. N. Williams, T. Masters, R. Broberg, and D. Jones (2013), Quantifying the effect of urbanization on U.S. Historical Climatology Network temperature records, *J. Geophys. Res. Atmos.*, *118*, 481–494, doi:10.1029/2012JD018509.

Hoerling, M., A. Kumar, J. Eischeid, and B. Jha (2008), What is causing the variability in global mean land temperature?, *Geophys. Res. Lett.*, *35*, L23712, doi:10.1029/2008GL035984.

IPCC (2007), *Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, L. A. Meyer, (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA).

Jones, P. D., and T. M. Wigley (2010), Estimation of global temperature trends: What's important and what isn't, *Clim. Chang.*, *100*, 59–69, doi:10.1007/s10584-010-9836-3.

Jones, P. D., and I. Harris (2011), University of East Anglia Climatic Research Unit (CRU) Time Series (TS) high resolution gridded data version 3.10, [Internet]. NCAS British Atmospheric Data Center, 27 March 2013. Available from http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__ACTIVITY_fe67d66a-5b02-11e0-88c9-00e081470265

Jones, P. D., T. J. Osborn, and K. R. Briffa (1997), Estimating sampling errors in large-scale temperature averages, *J. Climate*, *10*, 2548–2568, doi:10.1175/1520-0442(1997)010 < 2548:ESEILS > 2.0.CO;2.

Jones, P. D., D. H. Lister, T. J. Osborn, C. Harpham, M. Salmon, and C. P. Morice (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, T. R., and C. N. Williams (1987), An approach to adjusting climatological time series for discontinuous inhomogeneities, *J. Climate and Applied Meteorol.*, *26*, 1744–1763, doi:10.1175/1520-0450(1987)026 < 1744:AATACT > 2.0.CO;2.

Karl, T. R., C. N. Williams, P. J. Young, and W. M. Wendland (1986), A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the United States, *J. Climate and Applied Meteorol.*, *25*, 145–160, doi:10.1175/1520-0450(1986)025 < 0145:AMTETT > 2.0.CO;2.

Livezey, R. E., and W. Y. Chen (1983), Statistical field significance and its determination by Monte Carlo techniques, *Mon. Wea. Rev.*, *111*, 46–59, doi:10.1175/1520-0493(1983)111 < 0046:SFSAID > 2.0.CO;2.

Mitchell, T., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatology*, *25*, 693–712, doi:10.1002/joc.1181.

Miyakoda, K., G. D. Hembree, R. F. Strickler, and I. Shulman (1972), Cumulative results of extended forecast experiments I: Model performance for winter cases, *Mon. Wea. Rev.*, *100*, 836–855, doi: 10.1175/1520-0493(1972)100<0836:CROEFE>2.3.CO;2.

Montandon, L. M., S. Fall, R. A. Sr. Pielke, and D. Niyogi (2011), Distribution of landscape types in the Global Historical Climatology Network, *Earth Interact*, *15*, 1–24, doi:10.1175/2010EI371.1.

Nuñez, M. N., H. H. Ciapessoni, A. Rolla, E. Kalnay, and M. Cai (2008), Impact of land use and precipitation changes on surface temperature trends in Argentina, *J. Geophys. Res.*, *113*, D06111, doi: 10.1029/2007JD008638.

Parker, D. E. (2011), Recent land surface air temperature trends assessed using the 20th Century Reanalysis, *J. Geophys. Res.*, *116*, D20125, doi: 10.1029/2011JD016438.

Peterson, T. C., et al. (1998), Homogeneity adjustments of in situ atmospheric climate data: A review, *Int. J. Climatol.*, *18*, 1493–1517, doi:10.1002/(SICI)1097-0088(19981115)18:13 < 1493::AID-JOC329 > 3.0.CO;2-T.

Pielke, R. A. Sr., et al. (2007), Unresolved issues with the assessment of multidecadal global land surface temperature trends, *J. Geophys. Res.*, *112*, D24S08, doi:10.1029/2006JD008229.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (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, D14.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*(D7), 7183–7192, doi:10.1029/2000JD900719.

Trenberth, K. E., et al., (2007), Observations: Surface and atmospheric climate change, in *Climate Change 2007*: *The Physical Science Basis*, edited by Solomon, S., et al., pp. 235–336, Cambridge University Press, New York.

Trewin, B. (2010), Exposure, instrumentation, and observing practice effects on land temperature measurements. *WIREs Clim Change*, *1*, 490–506, doi: 10.1002/wcc.46.

Vose, R. S., et al. (2012a), NOAA's merged land–ocean surface temperature analysis, *Bull. Amer. Meteor. Soc.*, *93*, 1677–1685, doi:10.1175/BAMS-D-11-00241.1.

Vose, R. S., S. Applequist, M. J. Menne, C. N. Williams Jr., and P. Thorne (2012b), An intercomparison of temperature trends in the U.S. Historical Climatology Network and recent atmospheric reanalyses, *Geophys. Res. Lett.*, *39*, L10703, doi:10.1029/2012GL051387.

Whitaker, J. S., and T. M. Hamill (2002), Ensemble data assimilation without perturbed observations, *Mon. Wea. Rev.*, *130*, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Willmott, C. J., and S. M. Robeson (1995), Climatologically aided interpolation (CAI) of terrestrial air temperature, *Int. J. Climatology*, *15*, 221–229, doi:10.1002/joc.3370150207.