Evaluation of Historical Diurnal Temperature Range Trends in CMIP5 Models

SOPHIE C. LEWIS AND DAVID J. KAROLY

School of Earth Sciences, and ARC Centre of Excellence for Climate System Science, University of Melbourne, Melbourne, Victoria, Australia

(Manuscript received 10 January 2013, in final form 20 May 2013)

ABSTRACT

Diurnal temperature range (DTR) is a useful index of climatic change in addition to mean temperature changes. Observational records indicate that DTR has decreased over the last 50 yr because of differential changes in minimum and maximum temperatures. However, modeled changes in DTR in previous climate model simulations of this period are smaller than those observed, primarily because of an overestimate of changes in maximum temperatures. This present study examines DTR trends using the latest generation of global climate models participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5) and utilizes the novel CMIP5 detection and attribution experimental design of variously forced historical simulations (natural-only, greenhouse gas–only, and all anthropogenic and natural forcings). Comparison of observed and modeled changes in DTR over the period of 1951–2005 again reveals that global DTR trends are lower in model simulations than observed across the 27-member multimodel ensemble analyzed here. Modeled DTR trends are similar for both experiments incorporating all forcings and for the historical experiment with greenhouse gases only, while no DTR trend is discernible in the naturally forced historical experiment. The persistent underestimate of DTR changes in this latest multimodel evaluation appears to be related to ubiquitous model deficiencies in cloud cover and land surface processes that impact the accurate simulation of regional minimum or maximum temperatures changes observed during this period. Different model processes are likely responsible for subdued simulated DTR trends over the various analyzed regions.

1. Introduction

Changes in diurnal temperature range (DTR) are an identifiable characteristic of recent climatic change and also provide a useful diagnostic index for evaluating global climate models (GCMs) (Braganza et al. 2004). The global mean warming trend over land has been associated with asymmetric minimum and maximum temperature changes. Hence, DTR has been decreasing on a global scale over the last 50 yr because of the relatively stronger increases in daily minimum temperatures (Tmin) than daily maximum temperatures (Tmax) (Karl et al. 1993; Easterling et al. 1997; Vose et al. 2005). As Tmin is closely related to the longwave radiative flux, whereas Tmax is largely determined by shortwave radiative forcing, DTR trends are useful for diagnosing recent climatic change and particularly anthropogenic forcing components.

However, the precise causes of DTR changes and their spatial and seasonal characteristics remain poorly understood (Stone and Weaver 2003). Recent studies investigating regional-scale DTR trends highlight the importance of multiple process in determining DTR, including cloud cover, soil moisture, and precipitation variability (Zhou et al. 2008; Lauritsen and Rogers 2012; Wang et al. 2013). In addition, previous comparisons of GCM simulations with observations indicate a significant underestimation of modeled DTR trends, which in some instances are linked to a strong simulated increase in Tmax (Braganza et al. 2004). Previous differences between model and observational DTR trends are largely attributed to increases in cloudiness and land surface processes that are not well represented in models (Stone and Weaver 2003; Braganza et al. 2004; Lobell et al. 2007). Variations in cloud cover, in particular, are strongly linked to DTR changes: increased cloud cover decreases downward solar radiation and therefore reduces Tmax, whereas cloud cover increases net longwave radiation and consequently Tmin, thereby narrowing the diurnal temperature range (Karl et al. 1993; Dai and Trenberth 1999). DTR is highly sensitive to small changes in both Tmax and Tmin

Corresponding author address: Sophie C. Lewis, School of Earth Sciences, University of Melbourne, McCoy Building, Corner Swantson and Elgin Streets, Parkville VIC 3010, Australia.
E-mail: sophie.lewis@unimelb.edu.au

DOI: 10.1175/JCLI-D-13-00032.1

and changes in land surface processes also influence DTR. Ultimately, modeled representations of cloud and land surface processes can impact the accurate simulation of Tmax and Tmin trends.

In this study, as a basis for comparison with observational datasets, we investigate simulated DTR changes in nine GCMs participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5) archived by the Program for Climate Model Diagnosis and Intercomparison (PCMDI; http://cmip-pcmdi.llnl.gov). We examine whether the accuracy of modeled DTR trends has improved with this latest generation of GCMs participating in CMIP5. Furthermore, to understand DTR trends in relation to forcings more thoroughly, temperature trends are also examined in CMIP5 detection and attribution experiments of the historical period using only a subset of known forcings, including natural-only and greenhouse gas–only forcings (Taylor et al. 2012). In particular, we use these new CMIP5 detection and attribution type experiments to investigate the possible role of aerosols in determining historical changes in DTR.

2. Model and observational datasets

Data for this study were compiled from both model and observational sources. Monthly mean observed land surface temperatures from the Climatic Research Unit (CRU) surface temperature, version 4 (CRUTEM4), dataset (Braganza et al. 2004; Jones et al. 2012) were used, together with monthly averages of observed daily maximum temperatures (TX) and minimum temperatures (TN) from the Hadley Centre Global Historical Climatology Network (GHCN)-Daily (HadGHCND) dataset (Caesar et al. 2006). Diurnal temperature ranges are calculated from HadGHCND as the difference between monthly averages of Tmax and Tmin. In addition, DTR values were used from both the GHCN-Daily climate extremes (GHCNDEX; Donat et al. 2013a) and Hadley Centre Global Climate Extremes Index 2 (HadEX2; Donat et al. 2013b) datasets. It is noted that the partly different data sources and requirements regarding homogeneity and record length of the station observations utilized in creating gridded observational datasets can lead to differences in climate parameters. As HadEX2 stations data are assessed more rigorously for both quality and homogeneity prior to gridding (Donat et al. 2013b), the observed DTR values quoted hereafter are those calculated from HadEX2, unless otherwise indicated.

Observational precipitation data were also used from both the CRU time series, version 3.1 (TS3.1), monthly gridded precipitation dataset [for an earlier version, see Mitchell and Jones (2005)] and GHCN (version 2) gridded land-only precipitation anomalies (Jones and Moberg 2003), in order to investigate temperature–rainfall relationships. Precipitation in both modeled and observational datasets are also utilized as a simple diagnostic for cloud cover changes, as high-quality observations of cloud cover changes are not available for the period of interest. This approach is discussed further in section 5.

For the model experiments, results are detailed from nine global coupled ocean–atmosphere climate models participating in CMIP5 (see Table 1 for model version and institutional details). Each historical period ensemble is initialized from an arbitrary point in a multicentury preindustrial control integration conducted for each model, providing a range of different climatic trajectories across ensembles. Models were selected where mean near-surface air temperature (Tmean), Tmin, Tmax, and precipitation fields were available for at least three ensemble members, which were all included in analyses here. Monthly mean modeled DTR values were calculated as the difference between Tmax and Tmin. Ensemble averages were calculated for each model analyzed and a multimodel mean was determined as average conditions across all realizations, using all analyzed models.

First, we used the historical all-forcings experiment, simulating the climate of 1850–2005. In this experiment, changing atmospheric composition (including CO_2) resulting from anthropogenic and volcanic influences, solar forcing, and emissions of short-lived species and natural and anthropogenic aerosols are imposed to reproduce climate evolution over the twentieth century as accurately as possible. Additional detection and attribution simulations, with natural forcing only (historicalNat) and greenhouse gas forcing only (historicalGHG), were utilized in order to separate anthropogenic and natural influences on twentieth century climate. The major distinction between the historical and historicalGHG experiments is the inclusion of aerosols. Fewer detection and attribution simulations have been performed than for the core historical experiment, and hence a subset of five of the nine selected models with three ensemble members was available for this part of the analysis.

Pre-industrial control experiments (piControl) were also examined to assess potential long-term control drifts that may impact surface temperature trends in the historical experiments and obscure the comparatively small DTR trends. We compared trends in the contemporaneous preindustrial control segment (following historical branching) with the internal variability determined by the spread in ensemble members. Assessments were made for each temperature variable, in each defined region. Drifts in surface properties were found to be small and had a negligible impact on simulated trends in the historical experiment.

TABLE 1. List of climate models and ensembles used in this study. Three realizations were used for each analyzed model, for each experiment. The abbreviated institutional name is used as an abbreviation for each model in most cases. Further details of individual models can be found from the PCMDI.

Model	Model expansion	Institution	Historical ensembles	HistoricalGHG ensembles	HistoricalNat ensembles	Land-use changes
NorESM1-M	Norwegian Earth System Model, version 1 (intermediate resolution)	Norwegian Climate Centre (NCC)	r1i1p1, r2i1p1, r3i1p1	No	No	No
CCSM4	Community Climate System Model, version 4	National Center for Atmospheric Research (NCAR)	r1i1p1, r2i1p1, r3i1p1	No	No	Yes
BCC-CSM1.1	Beijing Climate Center, Climate System Model, version 1.1	Beijing Climate Center (BCC)	r1i1p1, r2i1p1, r3i1p1	No	No	No
MIROC4h	Model for Interdisciplinary Research on Climate, version 4 (high resolution)	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine–Earth Science and Technology	r1i1p1, r2i1p1, r3i1p1	No	No	Yes
GFDL CM3	Geophysical Fluid Dynamics Laboratory Climate Model, version 3	Geophysical Fluid Dynamics Laboratory (GFDL)	r1i1p1, r3i1p1, r5i1p1	r1i1p1, r3i1p1, r5i1p1	r1i1p1, r3i1p1, r5i1p1	Yes
GISS-E2-R	Goddard Institute for Space Studies (GISS) Model E, version 2, coupled with the Russell ocean model	National Aeronautics and Space Administration (NASA) GISS	r1i1p1, r2i1p1, r3i1p1	r1i1p1, r2i1p1, r3i1p1	r1i1p1, r2i1p1, r3i1p1	Yes
CanESM2	Second Generation Canadian Earth System Model	Canadian Centre for Climate Modelling and Analysis (CCCma)	r1i1p1, r2i1p1, r3i1p1	r1i1p1, r2i1p1, r3i1p1	r1i1p1, r2i1p1, r3i1p1	Yes
CNRM-CM5	Centre National de Recherches Météorologiques (CNRM) Coupled Global Climate Model, version 5	CNRM–Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	r1i1p1, r2i1p1, r3i1p1	r1i1p1, r2i1p1, r3i1p1	r2i1p1, r3i1p1, r4i1p1	No
CSIRO Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation (CSIRO) Mark, version 3.6.0	CSIRO–Queensland Climate Change Centre of Excellence	r1i1p1, r2i1p1, r3i1p1	r1i1p1, r2i1p1, r3i1p1	r1i1p1, r2i1p1, r3i1p1	No

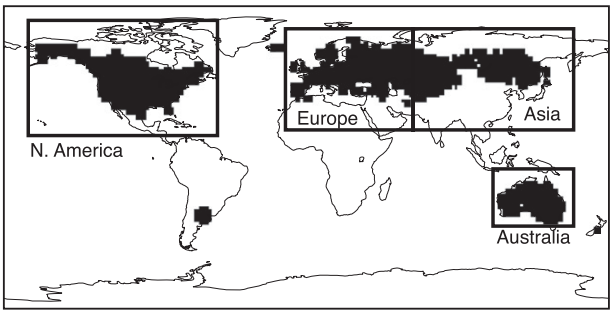

FIG. 1. Land surface mask derived from temporal coverage of GHCNDEX observed temperature datasets and used to mask model data. Shaded areas represent grid boxes included in this study, with greater than 80% observed monthly coverage during 1951–2005 and at least 60% coverage in the first and last 5 yr of the temperature record. Four land surface regions (North America, Europe, midlatitude Asia, and Australia) are also defined.

Model data were regridded onto a common 2.5° by 2.5° horizontal grid, then seasonal [June–August (JJA) and December–February (DJF)] and annual (ANN) averages were calculated and anomalies taken from the 1961–90 climatological mean, so that trends were commensurable with observations. Results are presented from 1951 to 2005, covering the period of overlap between the observed data and model simulations. Model fields were then masked to exclude areas of poor observational data coverage, which were identified from the observed GHCNDEX dataset, which had sparser coverage than either HadGHCND or HadEX2. Specifically, we used a land surface data mask including only grid boxes with at least 80% temporal coverage from 1951 to 2005 and also containing at least 60% of temperature records in the first and last five of these years (Fig. 1). This mask was applied to all model and observational datasets. Finally, four large land surface regions were defined (North America, Europe, midlatitude Asia, and Australia; Fig. 1) to facilitate the examination of spatial temperature trend variability over distinct regional land surface areas.

3. Evaluation of model performance

Before temperature trends were assessed against observations for the suite of models, we conducted an evaluation of model performance in simulating variability in Tmean, Tmax, and Tmin over the period of 1951–2005. First, the standard deviations of interannual variations of detrended regional average anomalies were calculated for each temperature variable and then compared to the equivalent observational standard deviation (Fig. 2). In this comparison, the regional average anomalies were linearly detrended to remove any externally forced monotonic climate signal that would potentially obscure variability on shorter time scales (Karoly and Braganza 2005). Detrending climatic time series using

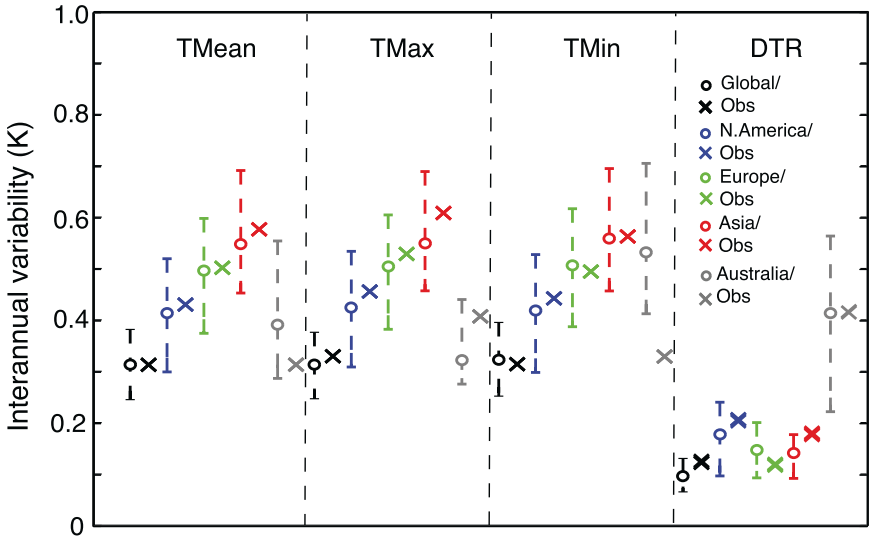

FIG. 2. Standard deviations of interannual variations of linearly detrended regional average Tmean, Tmax, Tmin, and DTR for historical simulations (circles) and observations (crosses). As the HadGHCND, GHCNDEX, and HadEX2 observational data points are effectively identical, a single observational point is shown, representing all analyzed gridded date products. Plot circles represent model ensemble mean values and ranges indicate the 5th–95th percentile values simulated across all realizations for global (black), North American (blue), Europe (green), midlatitude Asian (red), and Australian (gray) area averages.

any polynomial trend may introduce artifacts, particularly for local or regional time series (Sen Gupta et al. 2012). Nonetheless, the multimodel historical mean interannual standard deviations for global Tmean, Tmax, and Tmin are similar to those observed, indicating that the model suite captures the observed temperature variability. Regionally, simulated Tmin standard deviations over Australia are notably larger than those observed ($\Delta\sigma = -0.20$ K). The variability was relatively insensitive to the order of the polynomial used to detrend the time series.

We also compared distributions of global land average detrended annual mean temperatures across the suite of model ensembles, with observations using probability density functions (PDFs) estimated using a kernel smoothing function in order to identify any potential differences between modeled and observed temperatures (Fig. 3). Comparison of the model and observational PDFs for Tmax indicates that there are some differences in temperature distributions that are not apparent in comparisons between the simulated and observed interannual variability of Tmax. In particular, there is an increased (decreased) representation of high (low) Tmax values in the model realizations compared with observations (Fig. 3b). In addition, although the interannual variability of Tmin was similar for both observational data and the multimodel mean, the model Tmin PDFs are somewhat different to the range of values represented in the observational record (Fig. 3c). In particular, there are decreased representations of low Tmin values relative to observations, although the observational Tmin PDF lies in the middle of the simulated Tmin PDFs for warmer values.

Lastly, to assess the similarity of statistical associations between analyzed variables in the historical experiments and those in the observed datasets, the correlation between DTR and Tmean, Tmin, Tmin, Tmax, and precipitation (PR) was calculated for linearly detrended time series for both the historical model experiments and observations (Table 2). The correlation between DTR and each of Tmin, Tmax, Tmean, and precipitation values is similar for model experiments and observations, although it should be noted that correlations are sensitive to both the sampling period and the detrending method (Braganza et al. 2004) and so statistical associations are only broadly insightful.

Overall, the relationship between the temperatures and precipitation fields, and the interannual temperature variability within the models investigated here is similar to that observed. However, as DTR is highly sensitive to changes in both Tmax and Tmin, the overestimate of warm Tmax values and underestimate of low Tmin values may be important for accurately modeling changes in DTR over the historical period.

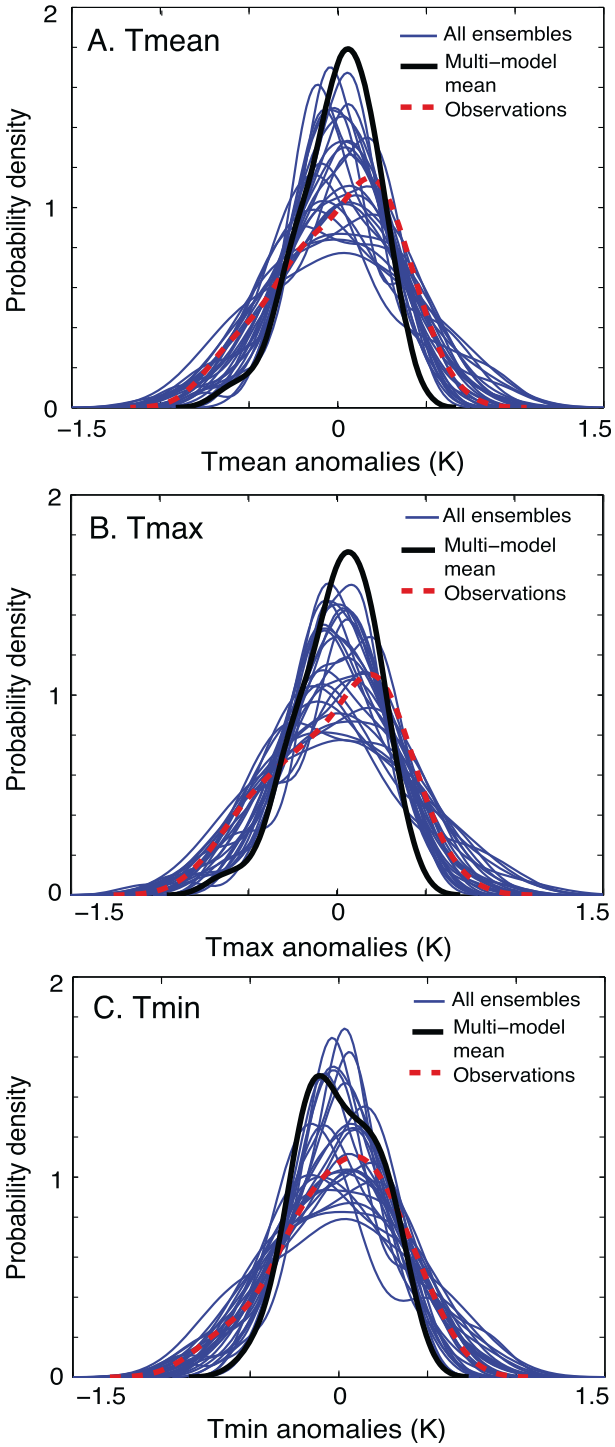

FIG. 3. Probability density functions for global land average detrended annual temperature anomalies for (a) Tmean, (b) Tmax, and (c) Tmin for observations (dashed red) and historical simulations. The multimodel mean is shown in black and values for each of the 27 model realizations are shown in blue, based on the three relevant realizations.

TABLE 2. Correlation *R* matrices of Tmean, Tmax, Tmin, DTR, and PR for observations and historical experiment multimodel mean. Observed Tmean data are derived from CRUTEM4; Tmin and Tmax are from HadGHCND; DTR is the mean *R* of HadGHCND, GHCNDEX, and HadEX2; and precipitation is the mean *R* of the CRU TS3.1 and GHCN monthly gridded precipitation datasets.

	Tmin	Tmax	Tmean	DTR	PR
Obs					
Tmin	1	—	—	—	—
Tmax	0.79	1	—	—	—
Tmean	0.88	0.90	1	—	—
DTR	−0.16	0.37	0.12	1	—
PR	0.04	−0.28	−0.15	−0.50	1
Historical					
Tmin	1	—	—	—	—
Tmax	0.84	1	—	—	—
Tmean	0.94	0.96	1	—	—
DTR	−0.13	0.36	0.15	1	—
PR	−0.01	−0.28	−0.16	−0.54	1

4. Land surface temperature changes

During the 1951–2005 period, the observed near-surface Tmean increases over land are similar to the simulated historical multimodel mean (Fig. 4). The equivalent temperature changes for the historicalNat experiment are notably outside the 5th–95th percentile window of model values across the 27-member historical ensemble for Tmean, Tmax, and Tmin from around 1990 to 2005, highlighting the role of anthropogenic forcings in observed warming in the late twentieth century. Over this period, the simulated evolution of DTR for the historical, historicalGHG, and historicalNat experiments is dissimilar to that observed: the decrease in observed DTR is not captured by the multimodel mean of any experiment. Furthermore, observed DTR anomalies since around 1980 are beyond or near the 5th–95th percentile window of the historical experiment ensemble (Fig. 4d). From 1951 to 2005, the observed annual mean temperature increases over land surfaces (0.21 K decade^{-1}) are similar to multimodel mean changes in the historical simulation (0.17 K decade^{-1}) (Fig. 5). The equivalent multimodel Tmean trend in the historicalGHG experiment over this period is higher (0.29 K decade^{-1}) but is near zero for historicalNat (−0.02 K decade^{-1}), indicating both a dominant role for anthropogenic forcings in determining temperature

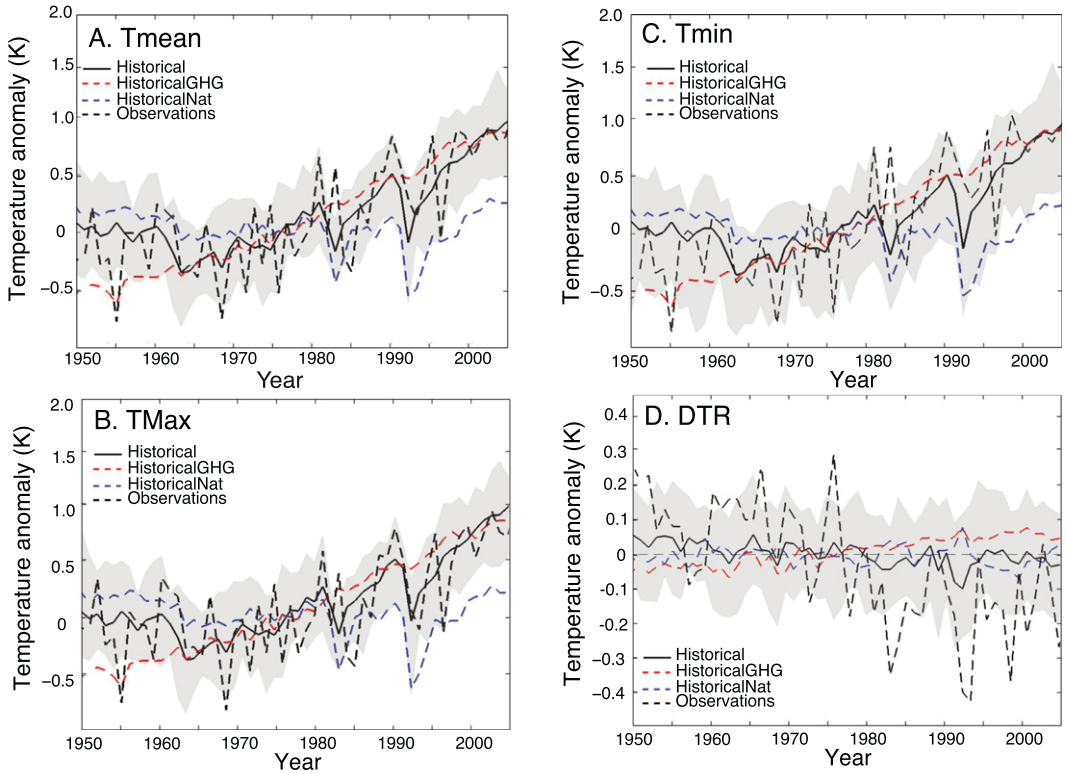

FIG. 4. Time series of global land surface mean annual (a) Tmean, (b) Tmax, (c) Tmin, and (d) DTR for model experiments (historical, historicalGHG, and historicalNat) and observational values. Average HadGHCND, GHCNDEX, and HadEX2 observed DTR values are shown. Gray plumes indicate the 5th and 95th percentile values of each temperature variable simulated in the 27-member historical experiment suite.

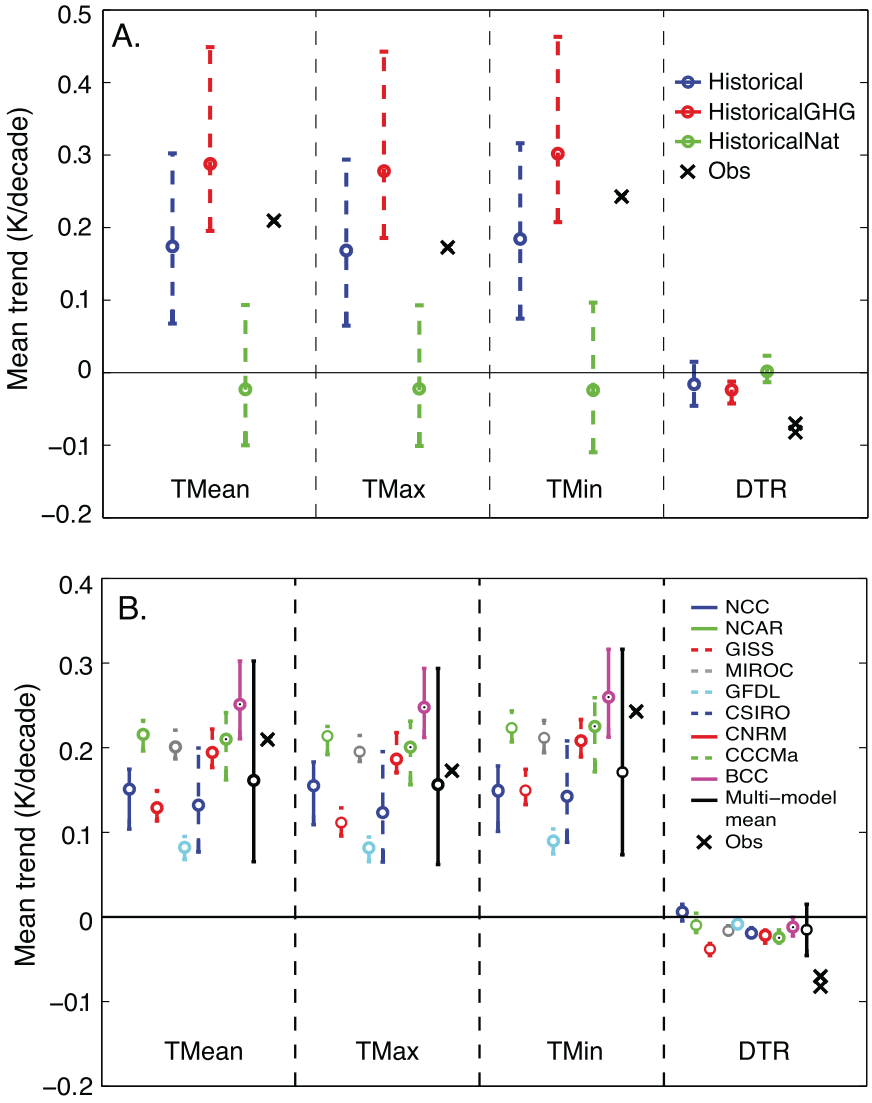

FIG. 5. Global mean trends (K decade^{-1}) in Tmean, Tmax, Tmin, and DTR for observations and for (a) historical, historicalGHG, and historicalNat experiment multimodel means and (b) all models for the historical experiment. Plot circles represent model ensemble mean trends and ranges indicate the minimum and maximum trends simulated across all 27 historical realizations. The observed DTR data are derived from the HadGHCND, GHCNDEX, and HadEX2 datasets, whereas the Tmean trend is calculated the CRUTEM4 gridded dataset and Tmin and Tmax are from HadGHCND. Note that the HadGHCND and GHCNDEX observed DTR trends are plotted together, giving only two distinct data points.

increases throughout this period and a cooling effect from aerosols.

Throughout this period, differences between the historical experiment and observational datasets in Tmin trends, in particular, result in discernible disparities in DTR trends. The multimodel historical mean Tmin trend (0.18 K decade^{-1}) is lower than the average observed trend (0.24 K decade^{-1}), while the mean simulated and observed land surface Tmax trends are the same (0.17 K decade^{-1}). Hence, the multimodel mean simulated DTR trend (-0.02 K decade^{-1}) is substantially smaller than observed (-0.07 K decade^{-1}). The DTR trend in the historical simulations is similar to that simulated in the historicalGHG experiment (-0.02 K decade^{-1}), although no DTR trend is discernible in the historicalNat simulation. The similarity of DTR trends in the historical and historicalGHG experiments suggests a potentially small role for aerosols in driving DTR changes, and previous studies have also indicated a negligible impact of sulfate aerosols on DTR (Mitchell et al. 1995;

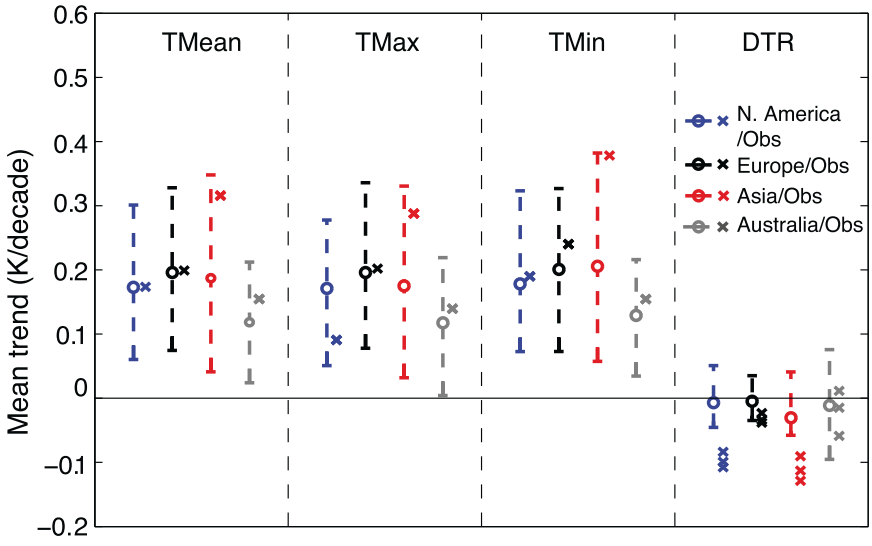

FIG. 6. Regional average trends (K decade^{-1}) in Tmean, Tmax, Tmin, and DTR for observations and historical experiment multimodel mean for North America (blue), Europe (black), midlatitude Asia (red), and Australia (gray). Plot circles represent model ensemble mean trends and ranges indicate the minimum and maximum trends simulated across all 27 historical realizations. The observed DTR data are derived from the HadGHCND, GHCNDEX, and HadEX2 datasets.

Stone and Weaver 2003; Braganza et al. 2004; Lobell et al. 2007). However, in these cases, indirect aerosol effects have not been represented in models. Possible aerosol effects on DTR trends in CMIP5 models are discussed further in section 5.

DTR trends in ensemble means from both the historical and historicalGHG experiments are temporally uniform over the analyzed period, although DTR changes in the observed record are not temporally consistent: large decreases in DTR occur over about 1961–80 and substantially smaller decreases occur subsequent to these decades. For example, the average observed land surface DTR change for 1961–80 is −0.10 K decade^{-1} compared with −0.03 K decade^{-1} over 1981–2000. A similar reduction in the magnitude of DTR trends since around 1990 has been observed in a network of European stations (Makowski et al. 2008).

Furthermore, temperature trends are neither seasonally nor spatially uniform in both model simulations and observations. Globally, larger DTR trends are identifiable from December to February in both the historical simulation (−0.05 K decade^{-1}) and observations (−0.11 K decade^{-1}) than either annually or during JJA (−0.06 K decade^{-1} for observations and 0.005 K decade^{-1} for historical). There are also evident differences in DTR trends between the four regions analyzed here (Fig. 6). Distinct Tmax, Tmin, and DTR responses have been noted previously for different regions in both observational station data (Easterling et al. 1997; Taylor et al. 2012) and in model simulations of future emission scenarios (Lobell et al. 2007).

In both the historical multimodel mean and observational data, the largest DTR trends occur for the midlatitude Asian region (−0.11 K decade^{-1} for observations and −0.03 K decade^{-1} for historical), followed by North America (−0.10 K decade^{-1} for observations and −0.01 K decade^{-1} for historical). However, the DTR trends in these regions, like the global equivalents, are underestimated in model simulations. For the North American region, although the variability in Tmean, Tmax, and Tmin is simulated accurately, there is a general overestimation of the modeled Tmax trends relative to observed (0.09 K decade^{-1} for observations and 0.17 K decade^{-1} for historical) and also an underestimate of simulated Tmin trends over midlatitude Asia (0.21 K decade^{-1}) compared with the average observed trend (0.38 K decade^{-1}). Over the Australian region, although simulated DTR interannual variability is substantially overestimated, the spread of model trends captures the observed changes (Fig. 6).

There is also a notable range of simulated temperature trends across the analyzed models and in individual model realizations. In general, modeled Tmin trends are generally lower than observed, with similar trends to those observed only occurring in ensembles from NCAR, CCCma, and BCC models. Of the various models, GISS-E2-R produces the ensemble member with the most realistic global DTR trend (−0.04 K decade^{-1}). Despite the

accuracy of Tmax or Tmin trends simulated in some realizations, capturing a realistic DTR trend in historical experiments requires faithful simulation of both Tmax and Tmin changes and none of the analyzed historical model realizations reproduces the DTR trends of the observational datasets. Hence, the previously identified underestimate of global DTR trends throughout the twentieth century in GCMs compared with observations (Braganza et al. 2004) is also a ubiquitous feature of the CMIP5 models considered here.

5. Factors influencing modeled DTR trends

Changes in both cloudiness and soil moisture have been identified as important factors determining observed decreases in DTR (Easterling et al. 1997; Dai and Trenberth 1999). Cloud cover influences DTR through the differential impact of clouds on radiative balances during the day and night times, and soil moisture influences DTR through variations in evaporative cooling, surface albedo and the ground heat capacity (Stone and Weaver 2003; Zhou et al. 2009; Jackson and Forster 2013). However, clouds represent one of the greatest uncertainties and are a large source of model spread (Dufresne and Bony 2008). Given the established importance of cloud cover on DTR, potential model inadequacies in accurately representing precipitation may obscure simulated temperature trends and cause the subdued modeled global DTR response relative to observed.

In this study, we diagnose potential changes in cloud cover by investigating changes in simulated precipitation. Although precipitation provides a much-simplified metric of cloudiness that may not be equally applicable in all regions, high-quality precipitation rather than cloud cover datasets are available for analysis and hence may provide some qualitative insight into potential model deficiencies. The influence of precipitation variability on temperature trends was examined by considering ''residual'' temperature variations after removing the part of the temperature signal that is linearly associated with rainfall variations (Nicholls 2003, 2004; Karoly and Braganza 2005). However, the comparison of residual modeled and observed global temperatures trends shows the underestimate of modeled global Tmin trends ($0.14\,\mathrm{K\,decade^{-1}}$) remains, when compared with observed trends ($0.19\,\mathrm{K\,decade^{-1}}$) and hence the modeled residual global DTR trend ($-0.01\,\mathrm{K\,decade^{-1}}$) is still weaker than observed ($-0.05\,\mathrm{K\,decade^{-1}}$).

We then extended this analysis and compared modeled temperature and precipitation responses in two versions of the Australian Community Climate and Earth-System Simulator (ACCESS) model (Bi et al. 2012) participating in CMIP5, using the same experiment suite for two available historical realizations. ACCESS versions 1.0 and 1.3 use distinct land surface models and cloud schemes, with ACCESS1.0 using the Hadley Centre Global Environment Model, version 2 (HadGEM2), atmospheric configuration and ACCESS1.3 coupled to the HadGEM3 atmospheric model with the prognostic cloud, prognostic condensate (PC2) cloud scheme for prognostic cloud fraction and condensate (Hewitt et al. 2011). Regardless of these model differences, temperature trends and their spatial characteristics over this period are found to be similar in both model versions ($-0.06\,\mathrm{K\,decade^{-1}}$ for observations; $0\,\mathrm{K\,decade^{-1}}$ for ACCESS1.0; and $-0.1\,\mathrm{K\,decade^{-1}}$ for ACCESS1.3). When residual temperatures were recalculated, with the influence of precipitation linearly removed, both model versions again showed similar trends and the subdued modeled DTR trend remained pervasive.

The largest regional discrepancies in model data temperature trends occur for midlatitude Asia (Fig. 6), where simulated historical Tmean, Tmax, and Tmin trends are smaller than observed and the observed DTR trend is not captured in any historical realizations. In this region, the underestimate of observed Tmin trends in the historical simulations was not improved by removing rainfall influences on temperature changes. Previous multimodel studies attribute the smaller change in modeled DTR than that observed to a stronger simulated Tmax warming than is realistic over the historical period (Braganza et al. 2004). However, in our study of CMIP5 models, while the multimodel mean global Tmean and Tmax trends are similar to those observed, simulated Tmin trends are noticeably subdued, particularly in the boreal winter.

The persistent residual discrepancies between modeled and observed Tmin trends, both globally and for midlatitude Asia, suggest that model–observational DTR trend discrepancies may result from underlying model problems beyond simulated cloud processes. For example, difficulties in modeling a suite of surface energy balances or controlling factors, such as soil moisture, albedo, and low-level clouds, would impact the ability of models to capture observed temperature trends (Zhou et al. 2009; Zwiers et al. 2011; Jackson and Forster 2013). At the land surface, deficiencies in capturing realistic freezing and thawing cycles of soil in GCMs have been noted previously and are known to impact simulated diurnal temperature ranges through an underestimate of Tmin values (Huth et al. 2003; Saito 2008). In addition, shallow nighttime boundary layer processes may be beyond the capacity of current models because of comparatively coarse model resolutions (McNider et al. 2012).

Our analysis does not indicate that aerosols are a contributing factor to errors in simulated Tmin trend over midlatitude Asia, particularly as simulated temperature trends are similar in the historical all-forcings and greenhouse gas–only simulations. However, aerosols remain a large uncertainty within climate models, and the simulation of aerosol direct and indirect effects remains problematic, particularly in regions such as Asia with high aerosol loadings (Zhou et al. 2009). The regional bias in simulated DTR trends and temporally variable nature of the observed DTR trend, ceasing around the time of peak aerosol emissions, may indicate a role of aerosols in determining DTR changes (Wild 2009). Observational based regional studies from China also highlight a possible role of aerosols in driving DTR trends (Wang et al. 2013).

Previous studies identify that DTR trends and their causes are likely to be regionally variable and connected to changes in cloud cover, precipitation, and soil moisture (Zhou et al. 2009). In our study, the cause of discrepancies between simulated and observed DTR trends over midlatitude Asia, resulting from problematic modeled Tmin changes, appear distinct from over North America, which derive from strong modeled Tmax trends. Although the global multimodel mean Tmax trend is similar to that observed trend, simulated Tmax trends for North America are overly large. After rainfall-associated variations are removed in this region, the accuracy of modeled Tmax trends is somewhat improved, although simulated historical Tmax trends ($0.16\,\mathrm{K\,decade^{-1}}$) remain larger ($0.12\,\mathrm{K\,decade^{-1}}$). The accurate simulation of maximum temperatures requires that numerous climatic processes beyond precipitation be captured with accuracy, including cloud interactions with incoming solar radiation. Maximum temperature values may be affected by the partitioning of radiation into sensible and latent heat fluxes by land surface processes that may be systematically poorly represented in climate models (Perkins et al. 2007). These processes, particularly, are regionally variable and dependent on the characteristics of soil moisture transitioning between limiting to evaporation and nonlimiting (Koster et al. 2004; Perkins et al. 2007).

The representation of probability density functions of daily maximum temperatures, as well as climatic variability, has been improved in some region using more sophisticated land surface model schemes (Lorenz et al. 2012). The relationship between DTR and soil moisture is notably most significant at the warmer end of the Tmax distribution, where, as soil get drier, hot days warm to a greater degree than cooler days (Durre et al. 2000). Hence, strong Tmax trends in North America relative to observed might relate to both model land surface and precipitation processes that impact energy balances. Finally, land-use change has been previously identified as a contributor to observed regional Tmax increases (Narisma and Pitman 2003). Several of the models analyzed here incorporate variable land-use changes over the historical period (Table 1); however, these land models confer variable skill in capturing observed historical DTR trends (Fig. 5b), compared with models utilizing static land surfaces.

Diagnosing the basis for differences in temperature trends in individual models is not attempted here, although intermodel differences in land surface processes and cloud feedbacks may be critical for accurately simulating temperature variability and may account for the spread in simulated temperature trends across the suite of models. Previous assessments of model spread in temperature responses to greenhouse gas forcing indicate that cloud feedback processes are responsible for the largest fraction of CMIP3 model response differences (Dufresne and Bony 2008). Given that the soil moisture, land surface processes, and cloud cover changes that influence both Tmin and Tmax are regionally dependent, the regional variability in accurately modeling DTR trends is also not surprising.

Observed DTR trends

Despite the widely identified inadequacies in the simulation of temperature and precipitation fields in GCMs that may limit the accurate determination of DTR trends, there are also uncertainties in observed trends that may affect DTR. The varying degrees of quality control, homogenization methods, and other observational data processing steps utilized in generating gridded observed data record can produce a spread of observed values across different datasets.

Nonetheless, the global observed DTR trends determined using all datasets are very similar ($-0.07\,\mathrm{K\,decade^{-1}}$ for HadGHCND; $-0.08\,\mathrm{K\,decade^{-1}}$ for GHCNDEX; and $-0.07\,\mathrm{K\,decade^{-1}}$ for HadEX2), despite the more rigorous quality control of the HadEX2 data that explicitly accounts for biases such as changes in time of observation and observational practice. In addition, the average observed DTR trend for this period is relatively robust to the criteria utilized for data masking. For example, in the instance where the criteria are relaxed and data coverage in the first and last years of the historical period are no longer necessary, observed DTR trends are essentially the same ($-0.08\,\mathrm{K\,decade^{-1}}$ for HadGHCND; $-0.08\,\mathrm{K\,decade^{-1}}$ for GHCNDEX; $-0.08\,\mathrm{K\,decade^{-1}}$ for HadEX2; and $-0.02\,\mathrm{K\,decade^{-1}}$ for model historical).

Furthermore, while observed DTR may be affected by changes in land use relating to urbanization, studies

utilizing observational datasets that comprise rural networks or data adjusted for urbanization show the observed DTR trend remains significant (Karl et al. 1993). Similarly, significant observed DTR trends were evident in China at both rural and urban locations (Wang et al. 2013). Here, the largest effect of urbanization was not related to energy balance changes resulting from the urban heat island but rather from anthropogenic aerosol changes associated with increased urbanization. In addition, the largest impact of urbanization in this area was on maximum, rather than minimum trends (Li et al. 2010). Finally, it is possible that the observed DTR trend was a result of natural climate variability and hence a large model ensemble size, beyond the 27-member ensemble utilized here, may be required to capture the observed trend (Meehl et al. 2012). However, in both our study and previously (Zhou et al. 2009), changes in Tmin, Tmax, and DTR are best represented where both natural and anthropogenic forcings are utilized, indicating some anthropogenic role in the observed DTR trends over the last 50 yr.

6. Conclusions

The observed global decrease in diurnal temperature range over the period of 1951–2005 resulting from the relatively larger increase in minimum temperatures than maximum temperatures is not captured in historical experiments of CMIP5 participating models. The multimodel mean DTR trend ($-0.02\,\mathrm{K\,decade^{-1}}$), calculated from 27 realizations from nine analyzed climate models, is lower than the average observed trend ($-0.07\,\mathrm{K\,decade^{-1}}$). Conversely, the trend in Tmean over this period identified in observations is similar to multimodel mean historical trends. The subdued decrease in DTR compared with observations is a robust feature across historical model ensembles and has been identified previously in multimodel DTR evaluation studies (Braganza et al. 2004).

The degree of underestimation of modeled DTR changes in the historical experiment is similar to DTR trends in a parallel historicalGHG experiment, which does not include forcing by aerosol changes over the historical period. Also, the experiment incorporating only natural forcings (historicalNat) does not produce a DTR trend and therefore highlights a discernible greenhouse gas signal in DTR changes over this period. The temporal changes in observed DTR, characterized by stronger trends in the early part of the 1951–2005 period than more recently, are also not reflected in modeled DTR trends in these experiments.

Simulated historical DTR trends are regionally variable, with the largest model–observational discrepancies occurring for midlatitude Asia and North America. For these regions, although the interannual variability in Tmean, Tmax, and Tmin and the relationship between temperature and precipitation is simulated accurately, the simulated DTR trend is underestimated. In the Asia region, the largest DTR discrepancies, associated with subdued Tmin trends, occur in the boreal winter, when shallow boundary layer and soil freezing and thawing cycles are likely difficult to simulate realistically. Conversely, over North America, simulated Tmin trends are realistic, although Tmax trends are overestimated relative to observed, resulting in subdued DTR trends associated with problematic model land surface and cloud processes. Hence, the causes of unrealistically small DTR trends in each region are distinct, relating variously to stable (Tmin) and unstable (Tmax) boundary layer processes.

Overall, diurnal temperature ranges are highly sensitive to small changes in both minimum and maximum temperatures. Anthropogenic forcings result in a decrease in diurnal temperature ranges, through changes in cloudiness, radiative fluxes, and land surface moisture balances. These are poorly understood processes that are still not well represented in models and hence model data inconsistencies in DTR changes remain within these analyzed models. Further analysis of CMIP5 participating models exploiting the full suite of available detection and attribution experiments that incorporate subsets of forcings may assist in identifying model strengths and weaknesses in simulating these processes. Also, explicit evaluation of radiative balances together with the use of a more sophisticated metric for cloudiness (beyond precipitation) may improve our understanding of DTR influences. Finally, although this study does not identify a role for aerosols in DTR trends, they could not be discounted. A more complete understanding of aerosols effects on DTR will require more sophisticated representations of aerosol forcings.

Acknowledgments. This research was supported by funding from the Australian Research Council Centre of Excellence for Climate System Science (Grant CE 110001028). This work was also supported by the NCI National Facility at the ANU. We also acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank climate modeling groups for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank Markus Donat and

Lisa Alexander at the University of New South Wales for providing access to GHCNDEX and HadEX2 observational datasets. Finally, we thank the three reviewers for their useful comments on the manuscript.

REFERENCES

Bi, D., and Coauthors, 2012: The ACCESS coupled model: Description, control climate and evaluation. *Aust. Meteor. Oceanogr. J.,* **63,** 41–64.

Braganza, K., D. J. Karoly, and J. M. Arblaster, 2004: Diurnal temperature range as an index of global climate change during the twentieth century. *Geophys. Res. Lett.,* **31,** L13217, doi:10.1029/2004GL019998.

Caesar, J., L. Alexander, and R. Vose, 2006: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *J. Geophys. Res.,* **111,** D05101, doi:10.1029/2005JD006280.

Dai, A., and K. Trenberth, 1999: Effects of clouds, soil moisture, precipitation, and water vapor on diurnal temperature range. *J. Climate,* **12,** 2451–2473.

Donat, M. G., L. V. Alexander, H. Yang, I. Durre, R. Vose, and J. Caesar, 2013a: Global land-based datasets for monitoring climatic extremes. *Bull. Amer. Meteor. Soc.,* **94,** 997–1006.

——, and Coauthors, 2013b: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *J. Geophys. Res.,* **118,** 2098–2118, doi:10.1002/jgrd.50150.

Dufresne, J.-L., and S. Bony, 2008: An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models. *J. Climate,* **21,** 5135–5144.

Durre, I., J. M. Wallace, and D. P. Lettenmaier, 2000: Dependence of extreme daily maximum temperatures on antecedent soil moisture in the contiguous United States during summer. *J. Climate,* **13,** 2641–2651.

Easterling, D. R., and Coauthors, 1997: Maximum and minimum temperature trends for the globe. *Science,* **277,** 364–367, doi:10.1126/science.277.5324.364.

Hewitt, H. T., D. Copsey, I. D. Culverwell, C. M. Harris, R. S. R. Hill, A. B. Keen, A. J. McLaren, and E. C. Hunke, 2011: Design and implementation of the infrastructure of HadGEM3: The next-generation Met Office climate modelling system. *Geosci. Model Dev.,* **4,** 223–253, doi:10.5194/gmd-4-223-2011.

Huth, R., J. Kyselý, and M. Dubrovský, 2003: Simulation of surface air temperature by GCMs, statistical downscaling and weather generator: Higher-order statistical moments. *Stud. Geophys. Geod.,* **47,** 203–216.

Jackson, L. S., and P. M. Forster, 2013: Modeled rapid adjustments in diurnal temperature range response to CO2 and solar forcings. *J. Geophys. Res.,* **118,** 2229–2240, doi:10.1002/jgrd.50243.

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

——, 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 Coauthors, 1993: Asymmetric trends of daily maximum and minimum temperature. *Bull. Amer. Meteor. Soc.,* **74,** 1007–1023.

Karoly, D. J., and K. Braganza, 2005: A new approach to detection of anthropogenic temperature changes in the Australian region. *Meteor. Atmos. Phys.,* **89** (1–4), 57–67, doi:10.1007/s00703-005-0121-3.

Koster, R. D., and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science,* **305,** 1138–1140, doi:10.1126/science.1100217.

Lauritsen, R. G., and J. C. Rogers, 2012: U.S. diurnal temperature range variability and regional causal mechanisms, 1901–2002. *J. Climate,* **25,** 7216–7231.

Li, X., F. Chen, H. Ye, Y. Xiong, L. Shi, L. Pan, and K. Wang, 2010: Trends of maximum temperature, minimum temperature and diurnal temperature range and their correlations with urbanisation in Xiamen, China. *Int. J. Sustainable Dev. World Ecol.,* **17,** 299–303, doi:10.1080/13504509.2010.488879.

Lobell, D. B., C. Bonfils, and P. B. Duffy, 2007: Climate change uncertainty for daily minimum and maximum temperatures: A model inter-comparison. *Geophys. Res. Lett.,* **34,** L05715, doi:10.1029/2006GL028726.

Lorenz, R., E. L. Davin, and S. I. Seneviratne, 2012: Modeling land–climate coupling in Europe: Impact of land surface representation on climate variability and extremes. *J. Geophys. Res.,* **117,** D20109, doi:10.1029/2012JD017755.

Makowski, K., M. Wild, and A. Ohmura, 2008: Diurnal temperature range over Europe between 1950 and 2005. *Atmos. Chem. Phys.,* **8,** 6483–6498, doi:10.5194/acp-8-6483-2008.

McNider, R. T., and Coauthors, 2012: Response and sensitivity of the nocturnal boundary layer over land to added longwave radiative forcing. *J. Geophys. Res.,* **117,** D14106, doi:10.1029/2012JD017578.

Meehl, G. A., J. M. Arblaster, and G. Branstator, 2012: Mechanisms contributing to the warming hole and the consequent U.S. east–west differential of heat extremes. *J. Climate,* **25,** 6394–6408.

Mitchell, J. F. B., R. A. Davis, W. J. Ingram, and C. A. Senior, 1995: On surface temperature, greenhouse gases, and aerosols: Models and observations. *J. Climate,* **8,** 2364–2386.

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

Narisma, G., and A. Pitman, 2003: The impact of 200 years of land cover change on the Australian near-surface climate. *J. Hydrometeor.,* **4,** 424–436.

Nicholls, N., 2003: Continued anomalous warming in Australia. *Geophys. Res. Lett.,* **30,** 1370, doi:10.1029/2003GL017037.

——, 2004: The changing nature of Australian droughts. *Climatic Change,* **63,** 323–336, doi:10.1023/B:CLIM.0000018515.46344.6d.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J. Climate,* **20,** 4356–4376.

Saito, K., 2008: Arctic land hydrothermal sensitivity under warming: Idealized off-line evaluation of a physical terrestrial scheme in a global climate model. *J. Geophys. Res.,* **113,** D21106, doi:10.1029/2008JD009880.

Sen Gupta, A., L. C. Muir, J. N. Brown, S. J. Phipps, P. J. Durack, D. Monselesan, and S. E. Wijffels, 2012: Climate drift in the CMIP3 models. *J. Climate,* **25,** 4621–4640.

Stone, D., and A. Weaver, 2003: Factors contributing to diurnal temperature range trends in twentieth and twenty-first century simulations of the CCCma coupled model. *Climate Dyn.,* **20,** 435–445.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.,* **93,** 485–498.

Vose, R. S., D. R. Easterling, and B. Gleason, 2005: Maximum and minimum temperature trends for the globe: An update through 2004. *Geophys. Res. Lett.,* **32,** L23822, doi:10.1029/2005GL024379.

Wang, F., C. Zhang, Y. Peng, and H. Zhou, 2013: Diurnal temperature range variation and its causes in a semiarid region from 1957 to 2006. *Int. J. Climatol.,* doi:10.1002/joc.3690, in press.

Wild, M., 2009: Global dimming and brightening: A review. *J. Geophys. Res.,* **114,** D00D16, doi:10.1029/2008JD01147.

Zhou, L., A. Dai, Y. Dai, R. S. Vose, C.-Z. Zou, Y. Tian, and H. Chen, 2008: Spatial dependence of diurnal temperature range trends on precipitation from 1950 to 2004. *Climate Dyn.,* **32** (2–3), 429–440, doi:10.1007/s00382-008-0387-5.

——, R. E. Dickinson, A. Dai, and P. Dirmeyer, 2009: Detection and attribution of anthropogenic forcing to diurnal temperature range changes from 1950 to 1999: Comparing multi-model simulations with observations. *Climate Dyn.,* **35** (7–8), 1289–1307, doi:10.1007/s00382-009-0644-2.

Zwiers, F. W., X. Zhang, and Y. Feng, 2011: Anthropogenic influence on long return period daily temperature extremes at regional scales. *J. Climate,* **24,** 881–892.