Can CGCMs Simulate the Twentieth-Century "Warming Hole" in the Central United States?

KENNETH E. KUNKEL, XIN-ZHONG LIANG, JINHONG ZHU, AND YIRUO LIN

Illinois State Water Survey, Champaign, Illinois

(Manuscript received 1 May 2005, in final form 12 December 2005)

ABSTRACT

The observed lack of twentieth-century warming in the central United States (CUS), denoted here as the "warming hole," was examined in 55 simulations driven by external historical forcings and in 19 preindustrial control (unforced) simulations from 18 coupled general circulation models (CGCMs). Twentieth-century CUS trends were positive for the great majority of simulations, but were negative, as observed, for seven simulations. Only a few simulations exhibited the observed rapid rate of warming (cooling) during 1901–40 (1940–79). Those models with multiple runs (identical forcing but different initial conditions) showed considerable intramodel variability with trends varying by up to 1.8°C century^{-1}, suggesting that internal dynamic variability played a major role at the regional scale. The wide range of trend outcomes, particularly for those models with multiple runs, and the small number of simulations similar to observations in both the forced and unforced experiments suggest that the warming hole is not a robust response of contemporary CGCMs to the estimated external forcings. A more likely explanation based on these models is that the observed warming hole involves external forcings combined with internal dynamic variability that is much larger than typically simulated.

The observed CUS temperature variations are positively correlated with North Atlantic (NA) sea surface temperatures (SSTs), and both NA SSTs and CUS temperature are negatively correlated with central equatorial Pacific (CEP) SSTs. Most models simulate rather well the connection between CUS temperature and NA SSTs. However, the teleconnections between NA and CEP SSTS and between CEP SSTs and CUS temperature are poorly simulated and the models produce substantially less NA SST variability than observed, perhaps hampering their ability to reproduce the warming hole.

1. Introduction

An interesting regional feature of the spatial pattern of temperature trends is the lack of twentieth-century warming in portions of the United States. Folland et al. (2001) show that in the central and southeastern United States there was warming from the early 1900s to the 1940s, followed by cooling into the 1970s, and a resumption of warming thereafter. The net trends over the entire twentieth century are near zero. This contrasts with detectable upward trends in most other land areas during this same period. Analyses of different periods illustrate geographic variations. Trends in temperature during the period 1976–2000 for the summer season only (Folland et al. 2001) show an area of cooling in the central United States, centered somewhat to the north and west of the center of the area of annual cooling found by Folland et al. (2001) for the entire twentieth century; this area of summer cooling was termed a "warming hole" by Pan et al. (2004). Robinson et al. (2002) analyzed the 1951–97 period and found annual cooling in the south-central United States, centered somewhat to the west of the twentieth-century annual cooling area and to the south of the 1976–2000 summer cooling area, although overlapping both. The term warming hole will be adopted here to refer to the general phenomenon found in all of these studies while the region to be studied will overlap all of the above areas and will be defined based on both physical and societal considerations. In addition to the lack of warming on a centennial time scale, the multidecadal variations are an interesting and integral aspect and will be examined along with the century-scale trends.

There are potentially important implications of this lack of warming. If this reflects a specific regional surface climate response to anthropogenic and natural ex-

Corresponding author address: Dr. Kenneth Kunkel, Illinois State Water Survey, 2204 Griffith Dr., Champaign, IL 61820-7495.
E-mail: kkunkel@uiuc.edu

ternal forcing that may continue into the future, this should then influence the assessment of impacts and long-term decisions that attempt to incorporate future climate change. However, this may simply reflect internal variability of the climate system or a temporary forcing response that may change as the forcing magnitude changes. In this case, there is more uncertainty about the future path of climate; this region might "catch up" to globally averaged changes, or vary in other ways.

The causes of various aspects of the warming hole have been the subject of recent study. Robinson et al. (2002) used the Goddard Institute for Space Studies (GISS) GCM to investigate the potential role of sea surface temperature (SST) as a forcing agent for the lack of annual warming for the period 1951–97. They found that this GCM, driven by observed SST variations, could reproduce the general temporal features of observed temperature variations. Their study did not answer the question of whether the observed SST variations were a result of internal variability of the coupled ocean–atmosphere system, or were driven by external forcing. Pan et al. (2004) used a regional climate model to show that local land surface feedback may be an alternate cause for the lack of warming in the summer, a process that occurs at scales perhaps too small to be simulated by coupled GCMs (CGCMs).

The process of preparation of the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC) has engaged many global modeling groups to produce new simulations of historical and future climate. These newly available simulations provide an opportunity to investigate the central United States warming hole. This study has the following two objectives: 1) to determine whether the present generation of CGCMs, when driven by modern estimates of time-varying forcing, can reproduce the observed warming hole and associated multidecadal features and 2) to use the models to provide insight into potential causal mechanisms for this feature. This analysis concentrated on annual temperature (although selected seasonal results are also presented) and focused on multidecadal variability and centennial trends. The ultimate motivation for this work is to ascertain the degree to which future climate projections are reliable at a regional scale.

2. Data and methods

Simulations were produced by 18 separate CGCMs from 13 different modeling groups. Basic information about the models is given in Table 1 (with more details online at http://www-pcmdi.llnl.gov/ipcc/model_documentation/ipcc_model_documentation.php). The modeling groups undertook a number of experiments, including preindustrial control, historical forcing, and several future scenarios. The primary focus of this analysis was on a set of CGCM simulations of the historical late-nineteenth- and twentieth-century climate. In this experiment, called the "climate of the twentieth century" and denoted as "20C3M," the CGCMs were driven by various estimated historical forcings, including greenhouse gases, solar variations, volcanic eruptions, ozone variations, halocarbons, land use, sulfate emissions, and other anthropogenic aerosol forcings. The modeling groups were free to use their own best estimates of these forcings (G. Meehl 2005, personal communication). As such, the forcings were not identical among models: although all simulations with documentation included forcings from greenhouse gases and sulfate emissions, other forcings were not common to all models; the direct and indirect effects of sulfate aerosols were not consistently incorporated either. The starting and ending dates of the simulations varied somewhat among modeling groups. A time period common to all model simulations was 1901–99 and this was used in the present analysis. The preindustrial (no external forcing) control (PICNTRL) experiment was also analyzed in this study.

For the 20C3M experiment, multiple runs with the same forcing but different initial conditions were performed for some of the individual models, resulting in a total of 55 separate simulations. Each of these multiple realizations will be initially treated as if it were an independent model simulation, an assumption that will be assessed. The large number of available simulations is unprecedented and opens the way for a thorough examination of internal climate variability, at least as produced by models.

The central United States (CUS) region chosen for study (Fig. 1) is one of the most agriculturally productive regions of the world and roughly defined around what is known as the "Corn Belt." Its climate is characterized by ample precipitation sufficient to support nonirrigated production of summer crops while irrigation for production of summer crops is widespread just to the west. It contains a number of large metropolitan areas, including Chicago, the third largest in the United States, and is the same region used in a recent study by Kunkel and Liang (2005). A set of 252 surface climate stations with less than 10% missing temperature data during 1901–99 and located within the CUS region were identified and used to construct observed time series. The CGCM surface air temperature data were mapped through bilinear spatial interpolation to a 30-km grid (a resolution similar to the observed data) used for re-

TABLE 1. Model information, including abbreviation, modeling group, country, resolution, and forcings: well-mixed greenhouse gases (CHG), halocarbons (H), sulfate tropospheric aerosols (Su), land use (L), volcanic aerosols (V), ozone (O_3), other aerosols, such as black carbon, sea salt, dust (B), and solar irradiance (So).

Name	Modeling group	Country	AGCM resolution	20C3M forcing
CCSM3	National Center for Atmospheric Research	United States	T85(~1.4°), 26 levels	GHG, So, Su, V, O_3, H, B
CGCM3.1	Canadian Centre for Climate Modelling and Analysis	Canada	T47(~2.5°), 32 levels	CHG, Su
CNRM-CM3	Meteo-France/Centre National de Resherches Meteorologiques	France	T63(~1.9°), 45 levels	GHG, Su
CSIRO-Mk3.0	CSIRO Atmospheric Research	Australia	T63(~1.9°), 18 levels	No documentation
ECHAM5/MPI-OM	Max Planck Institute for Meteorology	Germany	T63(~1.9°), 31 levels	GHG, Su, H, O_3
FGOALS-g1.0	LASG/Institute of Atmospheric Physics	China	2.8 lat × 2.8 lon, 26 levels	GHG, Su, H
GFDL-CM2.0 GFDL-CM2.1	Geophysical Fluid Dynamics Laboratory	United States	2.0 lat × 2.5 lon, 24 levels	GHG, O_3, Su, V, So, L
GISS-AOM	Goddard Institute for Space Studies	United States	3.0 lat × 4.0 lon, 12 levels	GHG, So, Su, V, O_3, H, L, B
GISS-EH GISS-ER			4.0 lat × 5.0 lon, 12 levels	
INM-CM3.0	Institute for Numerical Mathematics	Russia	4.0 lat × 5.0 lon, 21 levels	GHG, So, Su, V
IPSL-CM4	L'Institut Pierre-Simon LaPlace	France	2.5 lat × 3.75 lon, 19 levels	GHG, Su, H
MIROC3.2(hires)	Center for Climate System Research, National Institute for Environmental Studies, and Frontier Research Center for Global Change	Japan	T106(~1.1°), 56 levels	GHG, So, V, O_3, L, Su, B, H
MIROC3.2(medes)			T42(~2.8°), 20 levels	
MRI-CGCM2.3.2	Meteorological Research Institute	Japan	T42(~2.8°), 30 levels	GHG, H, Su, V, So
PCM	National Center for Atmospheric Research	United States	T42(~2.8°), 26 levels	GHG, So, Su, V, O_3
UKMO-HadCM3	Hadley Centre for Climate Prediction and Research/Met Office	United Kingdom	2.5 lat × 3.75 lon, 19 levels	No documentation

gional climate model simulations (Liang et al. 2004); those grids within the Fig. 1 box were arithmetically averaged to produce model time series. In addition, the observed SSTs were adopted from the global analysis of the First Hadley Centre for Sea Ice and Sea Surface Temperature (HadISST1) dataset (Rayner et al. 1996, 2003), while global surface air temperature data were obtained from the GISS (Hansen et al. 2001). Finally, certain analyses utilized monthly and annual data for U.S. climate divisions produced by the National Climatic Data Center.

As noted above, the location of a warming hole varies with the period and season chosen. The CUS, defined primarily based on considerations of the nature of agricultural production, does not correspond exactly with any of the warming hole regions used in other studies, although overlapping all of them. It is possible that model simulations could produce a warming hole but offset from the observed locations. The assessment of model performance could then be sensitive to the exact region chosen. One analysis will explore this possibility.

The focus of this study is on variability at decadal and longer time scales. The time series of all variables (surface air temperature, SSTs, wind components) were filtered with a moving average window of 11-yr length.

FIG. 1. Location of climate stations with at least 90% available temperature data for the period 1901–99. The box outlines the central U.S. area examined in this study.

The filtered time series were used in all analyses, including the calculation of trends, correlation coefficients, and standard deviations. Thus, variations associated with higher frequency phenomena such as El Niño are not considered here. Linear trends of the time series were calculated for 1901–40, 1940–79, 1980–99, and the overall period 1901–99. Teleconnection patterns were examined between the time series of CUS temperature and global pointwise SSTs. Statistical significance of correlation coefficients was determined through the following Monte Carlo analysis. In the case of CUS temperature correlations between a model and observations, the model time series was randomly reordered and then used to calculate a correlation coefficient. In the case of correlations between observed CUS temperature and observed pointwise SSTs, the CUS temperature time series was randomly reordered and then used to calculate a correlation coefficient. This was repeated 5000 times and the resulting distribution of correlation coefficients determined the 95% level of confidence thresholds. Values exceeding such thresholds would be considered statistically significant if only a single simulation were available to compare with observations. Since the analysis here involves a large number of simulations, it is expected that a few may exceed such thresholds just by chance.

3. Results

As suggested by Robinson et al. (2002), teleconnections between CUS temperature and SST patterns are a possible, even likely, factor associated with the warming hole, due to the observed multidecadal time scales. However, whether SST teleconnections, or for that matter land surface feedback, are integral features, we speculate that the warming hole could arise with or without external forcing, or possibly a combination of both internal and external forcing. If the warming hole is a result of external forcing, if the responsible forcings are the same or very similar in all models, and if model limitations do not compromise the models' ability to reflect the forcing response, then the warming hole should be common in these simulations. Furthermore, this outcome should be relatively insensitive to the initial conditions of the model simulation. On the other hand, if the warming hole arises from internal variability of the coupled climate system, then the warming hole might be less common in the forced simulations and might also be present in the control simulations. In this case, the simulation outcomes might be sensitively dependent on initial conditions.

To investigate these issues, the forced simulations (20C3M) are first examined to determine whether the warming hole is a common outcome. A more detailed assessment of multidecadal variations within the overall centennial trend is also undertaken to provide additional information for understanding model behavior. Then, the control simulations (PICNTRL) are analyzed to assess the possibility that the warming hole can arise from internal variability alone. Also, those forced simulations with multiple runs are explored to determine the sensitivity to initial conditions. Finally, SST teleconnections may mediate the response, whether forced or unforced, and the models are evaluated to determine if they can reproduce observed teleconnection patterns.

a. Externally forced variability

The observed global temperature time series (Fig. 2) shows the well-known pattern of a substantial rise from 1901 to about 1940, a slight decline from 1940 to the 1970s, followed by a rapid rise for the remainder of the century. CUS temperature follows a similar qualitative pattern but with some critical quantitative differences. There is substantial rise from 1901 to the 1930s, followed by a decrease of similar magnitude from the 1930s to a minimum in the late 1970s. There has been a slight rise since 1980, but the net linear trend, derived from least squares linear regression, over the entire century is $-0.2°C\ century^{-1}$ while the global trend is $+0.6°C\ century^{-1}$.

There are substantial seasonal variations in trends (Table 2). For 1901–40, there are upward trends in all seasons, but the trend in spring ($+0.6°C\ century^{-1}$) is quite small compared to large trends in winter (+5.5) and summer (+3.9). For 1940–79, the downward trend in winter is very large (−6.3), while there is a slight upward trend in spring (+0.5). For 1980–99, the annual upward trend is due entirely to large warming in winter (+9.1) while there are downward trends in the other

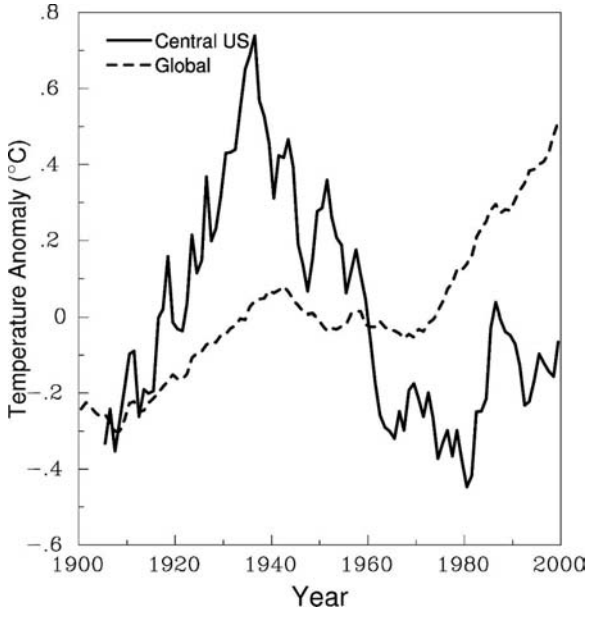

FIG. 2. Time series of observed annual temperature (expressed as the deviation from the 1901–99 average) for the central United States (thick solid line) and the globe (thin solid line). Both time series were smoothed with an 11-yr moving average filter.

three seasons, particularly in spring (−3.2) and summer (−2.7). The net (1901–99) trends are negative and relatively small in all seasons except for no trend in spring.

Correlation coefficients (Fig. 3) for CUS annual temperature were computed between each model time series for the 20C3M experiment and the observed time series. Values vary widely from −0.67 to +0.55. Only one value is above +0.55, the 95% significance level derived from the Monte Carlo analysis if only a single run were available to compare with observations. The model mean is slightly negative (−0.07). In general, correlation coefficients vary widely between individual runs of the same model with multiple realizations. Time series (Fig. 4) illustrate the behavior of selected simulations. The time series for the model with the highest correlation of +0.55 (GISS-ER run 8) shows warming until about 1950, cooling into the 1970s, and then slight warming into the 1990s. The time series for the model

TABLE 2. Temperature trends (°C century^{-1}) in the central United States for selected periods.

	1901–40	1940–79	1980–99	1901–99
Winter	5.3	−6.3	9.1	−0.1
Spring	0.6	0.5	−3.2	0.0
Summer	3.9	−1.7	−2.7	−0.1
Fall	1.7	−1.8	−0.6	−0.7
Annual	2.9	−2.3	0.4	−0.2

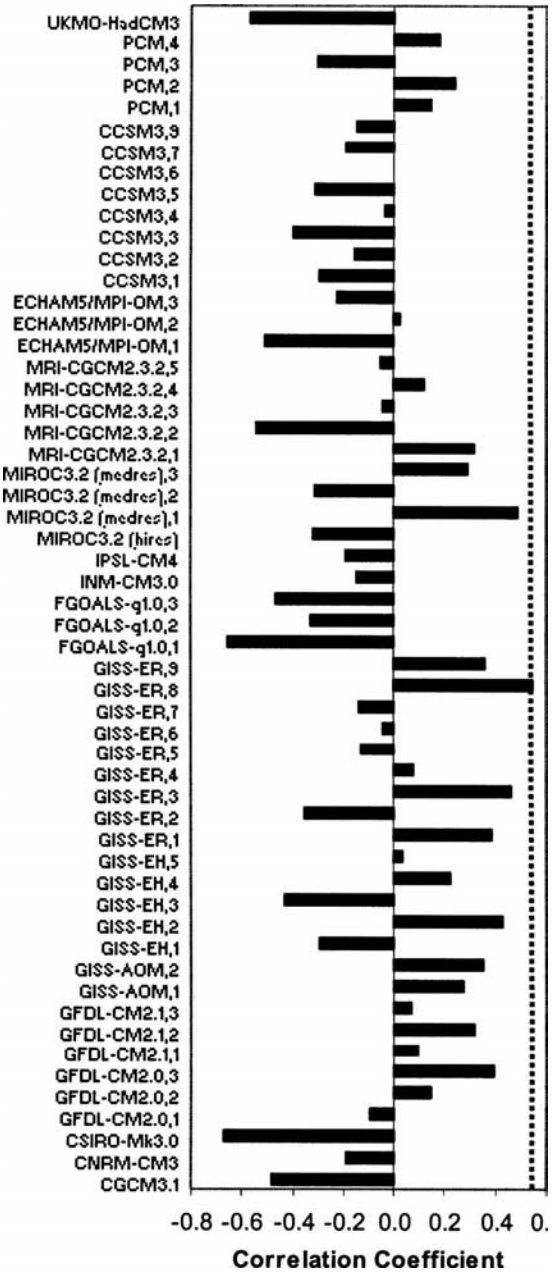

FIG. 3. Correlation coefficients between annual CUS temperature time series for 20C3M model experiments and observations during 1901–99. Time series were smoothed with an 11-yr moving average filter before computing correlations. The number after the comma in the label indicates the run number for those models with multiple simulations. The dashed vertical line indicates the value of apparent statistical significance at the 95% level of confidence.

with the second highest correlation of +0.48 [the Model for Interdisciplinary Research on Climate version 3.2 (MIROC3.2 medres run 1) exhibits warming from the early 1900s to the 1940s, cooling until the 1970s, and

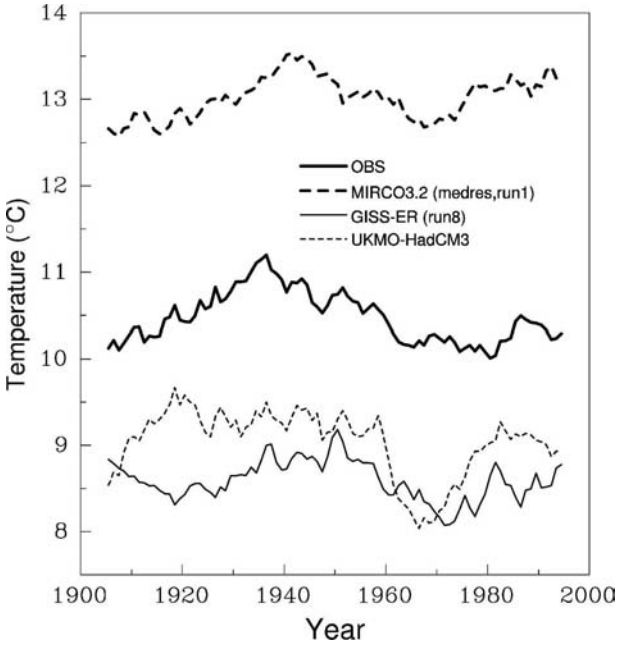

FIG. 4. Selected CUS temperature time series for 1901–99 including observations, two 20C3M (GISS-ER run 8 and MIROC3.2 medres run 1) simulations, and one preindustrical control (UKMO-HadCM3) simulation. All time series have been smoothed with an 11-yr moving average filter.

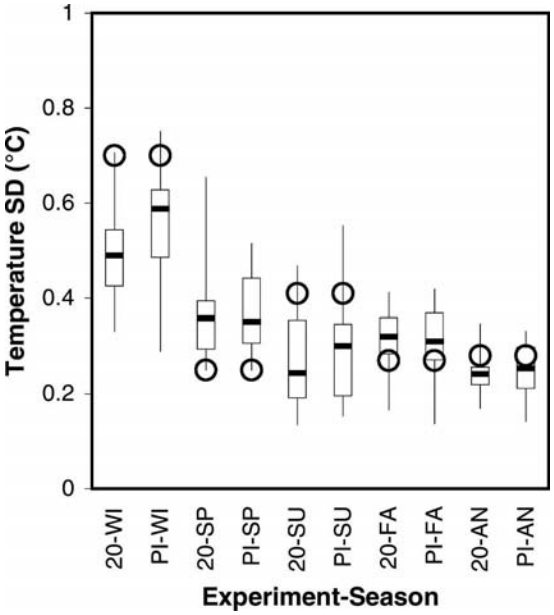

FIG. 5. Box and whisker plot illustrating the distribution of the standard deviation of CUS temperature time series for the 20C3M ("20") and PICNTRL ("PI") experiments for the four seasons ("WI," "SP," "SU," and "FA") and annual ("AN") time periods. The upper and lower limits of the box indicate the 75 and 25 percentile values, the solid horizontal line in the box indicates the median, and the ends of the whiskers indicate the highest and lowest value. The open circle indicates the observed value.

warming thereafter. Both of these simulations manifest features similar to observations.

There is a substantial model range in CUS mean annual temperature, as partially illustrated in Fig. 4. While the GISS-ER (run 8) is about 2°C cooler than observed, the MIROC3.2 (medres run 1) is about 3°C warmer than observed. The entire range (not shown) from the coolest to warmest model is about −3° to +3°C relative to observed. Although the biases for some models are rather large, a comparison of mean biases with trends and correlation coefficients did not reveal any systematic relationships. This is a similar finding to that of Kunkel and Liang (2005) who showed that, for this same region, various types of comparisons with observations for nine CGCMs were not related to mean biases.

The model time series in Fig. 4 appear visually to be characterized by a level of variability similar to that of the observed time series. The standard deviation of the time series was computed for each simulation. An examination of the results indicated that there were very small differences among members of an ensemble and, subsequently, the ensemble values were averaged to produce a single value for each model. Values for the PICNTRL simulations were produced by averaging values for overlapping segments of 99 years in length. A box and whiskers presentation (Fig. 5) of the results for both the 20C3M and PICNTRL experiments and for annual and seasonal time periods shows that there are relatively small differences between the two experiments, although winter PICNTRL values are noticeably higher than 20C3M values and the summer median value for PICTNRL is higher than for 20C3M. The observed values are within the range of the model simulations. However, the observed winter (spring) values are at the very upper (lower) end of the model distributions and observed values are higher than the inner-quartile ranges for summer and 20C3M annual and lower than the inner-quartile range for 20C3M fall. The absolute difference between the observed and model mean values is small for the annual time period. Stouffer et al. (2000) examined annual surface air temperature variability for three models and found generally good agreement with observations in this region, similar to the findings here for the annual period. An examination of spatial patterns of variance indicated that, in both models and observations, the primary feature is a zonal pattern with increasing values from south to north. The CUS region is not one of unusually high or low variance relative to surrounding regions.

Trends for the entire 1901–99 period (Fig. 6) vary from −0.5°C century^{-1} for GISS-ER run 1° to +1.9°C

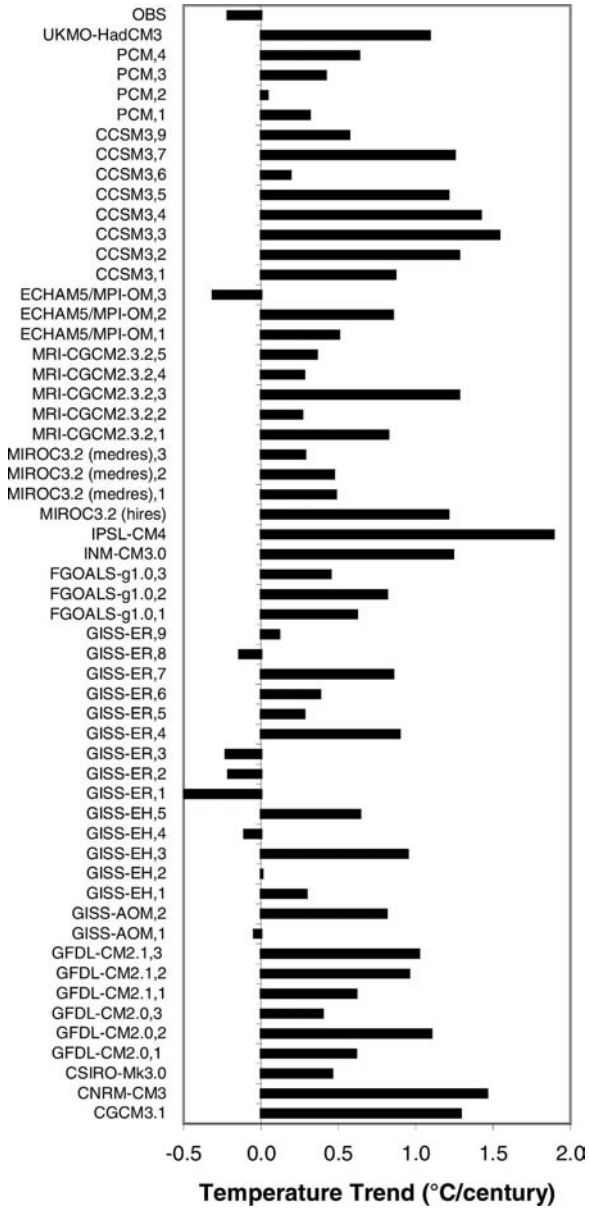

FIG. 6. Annual CUS temperature trends (°C century^{-1}) during 1901–99 for observations ("OBS") and 20C3M model simulations. The number after the comma in the label indicates the run number for those models with multiple simulations.

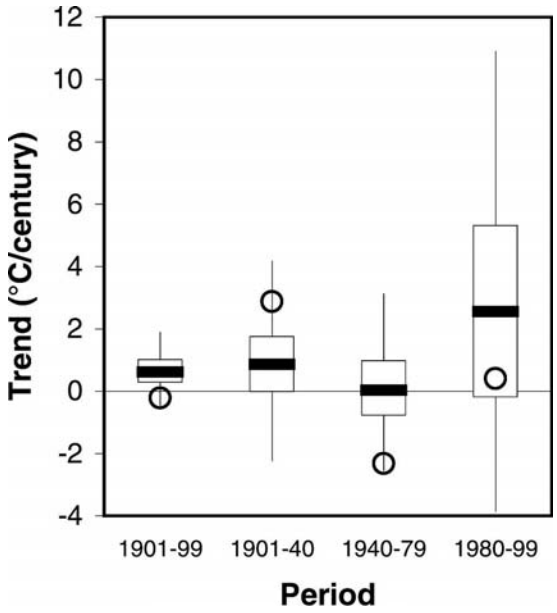

FIG. 7. Box and whisker plot illustrating the distribution of annual CUS temperature trends from the 20C3M model simulations for various periods. The upper and lower limits of the box indicate the 75 and 25 percentile values, the solid horizontal line in the box indicates the median, and the ends of the whiskers indicate the highest and lowest values. The open circle indicates the observed value.

century^{-1} for the L'Institut Pierre-Simon Laplace Coupled Model version 4 (IPSL-CM4), compared to the observed value of -0.2°C century^{-1}. Only 7 of the 55 model simulations exhibit negative trends. A box and whisker presentation of trends for this period and the three subperiods (Fig. 7) shows some interesting aspects of the ensemble of model simulations. For the 1901–40 and 1940–79 subperiods, the observed trends are within the envelopes, but near the upper and lower extremes, respectively. For the 1901–40 period, only 9 model simulations exhibit warming greater than $+2.0$°C century^{-1} while the observed trend is $+2.9$°C century^{-1}. For 1940–79, only 3 model simulations exhibit cooling less than -2.0°C century^{-1}, compared to the observed value of -2.3°C century^{-1}. For 1980–99, the observed trend is within the inner-quartile range of the model trends. This period exhibits a much larger model range than that of the 1901–40 and 1940–79 subperiods, but this may be due in part to the short length, making the trend estimates more susceptible to decadal scale variability. While most models exhibit positive trends (several larger than $+6$°C century^{-1}) for 1980–99, 13 simulations exhibit negative trends, compared to the observed value of $+0.4$°C century^{-1}.

A comparison of seasonal model trends with observations (Fig. 8) reveals that, in general, there is much less seasonal variation in model trends than exhibited by observations. For 1901–40, the observed strong upward trends in winter and summer are well above the inner-quartile range, although within the entire range, of the model simulations. For 1940–79, the strong winter cooling is more negative than any single model simulation, while in the other three seasons the observed values are within or near the inner-quartile range. For 1980–99, the observed trends are higher than the inner-quartile range for winter and lower for the

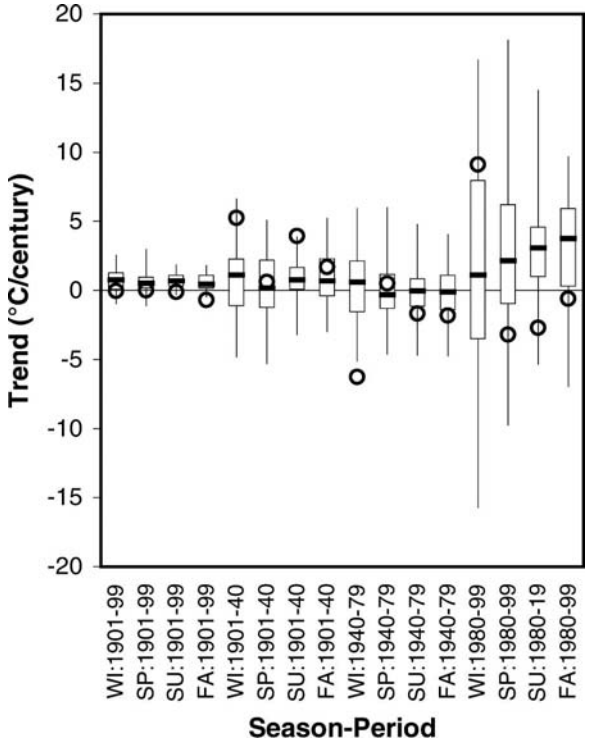

FIG. 8. As in Fig. 7 but for seasonal CUS temperature trends.

other three seasons. For the net trend over 1901–99, the observed values are lower than the inner-quartile range in all seasons.

It is possible that the models could produce features similar to observed, but offset slightly in time. In that case, the comparison of trends for fixed model periods would provide a misleading picture of model performance. To examine this possibility, trends were computed for each overlapping 20- (1901–20, 1902–21, etc.) and 40-yr (1901–40, 1902–41, etc.) period. The distributions of these model trends were plotted (not shown) as a function of starting date and compared to observed trends. The model trend distributions (particularly the inner-quartile range) varied quite slowly with starting date, indicating that the model trend distributions shown in Figs. 7 and 8, exactly matching the observation periods, are representative of the distributions for small lags between the observation and model periods. Of further significance, there are no model periods for which the observed trends of 1901–40 and 1940–79 are within the inner-quartile range of model trends. Thus, the display of matching periods in Figs. 7 and 8 provides an essentially accurate picture of model performance.

If the observed warming hole is a robust regional response to anthropogenic forcing, then it is reasonable to expect similar behavior in many of the simulations, unless external forcing uncertainties and/or model deficiencies are so great as to not provide a realistic representation of reality. However, only a very few of the simulations produce behavior similar to observed. This suggests either that the limitations in the present generation of models and/or estimates of external forcings substantially affect multidecadal regional responses to external forcing or that the observed behavior reflects large internal variability.

Another possibility is that the models simulate a warming hole, but offset from the CUS region. This was explored by producing trend maps for each season and each period and visually examining them. A number of the simulations exhibited warming holes somewhere in the United States; however, there was no preferred location, even among ensemble runs from the same model. To quantify the similarity of the entire pattern, a spatial correlation coefficient for the contiguous United States was calculated between model and observed trends. The results for 1980–99 for summer and annual periods (Fig. 9) shows that most models have rather high correlation coefficients for the annual period. However, the correlation coefficients for summer (when the warming hole is strongest) vary widely, from −0.56 to +0.60. If there were a common shift of the warming hole in the models, the correlation coefficients would be more clustered. The wide variations illustrate what inspection of the individual maps indicates: that there is no systematic spatial pattern of trends in the models. Thus, the specific definition of the region is unlikely to affect the general results.

b. *Internal variability*

The PICNTRL simulations provide a source of data to evaluate internal model variability. Since there is no external forcing in these simulations, climate variations must be generated through internal processes of the coupled ocean–land–atmosphere modeling system. Correlation coefficients for CUS temperature were calculated between the observed time series and the PICNTRL time series for those models with simulations of at least 150 years in length. In this case, correlation coefficients were calculated for each overlapping 99-yr period (i.e., years 1–99, 2–100, 3–101, etc.) and the maximum value in each PICNTRL run was determined. Since these simulations should exhibit no preference for any particular variation, the maximum value is a measure of how closely model internal variability can produce temporal variations matching observed patterns. The maximum correlation coefficients (Fig. 10) vary from +0.22 to +0.49. None of the maximum correlations reach the +0.55 level that would be significant at the 95% confidence level if only a single run was available to compare with observations. One sample

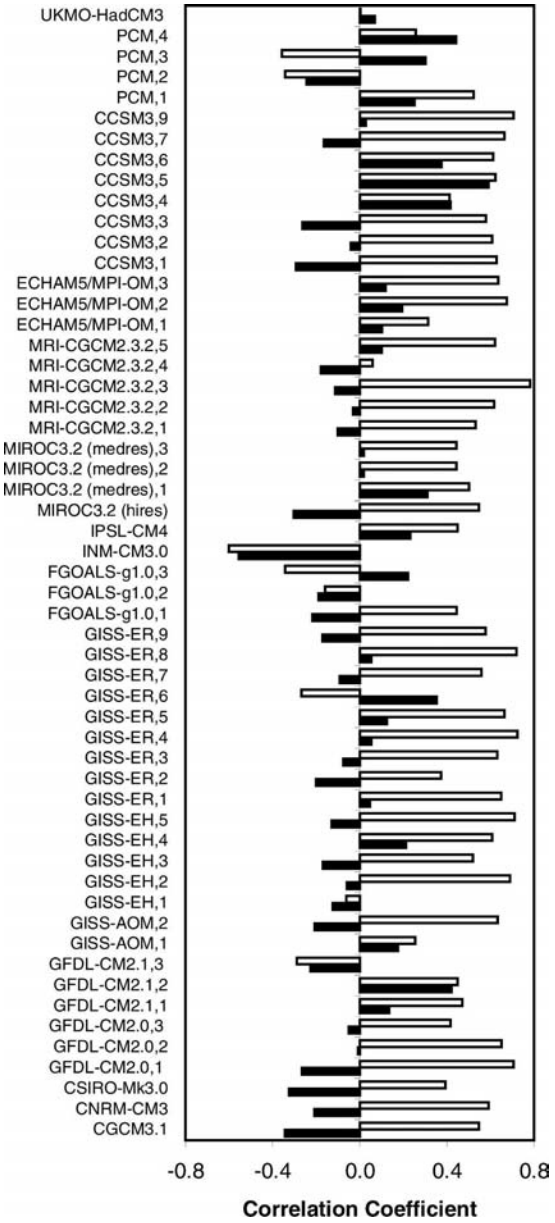

FIG. 9. Spatial correlation coefficients (calculated over the contiguous United States between the patterns of observed and 20C3M model temperature trends during 1980–99. Time series were smoothed with an 11-yr moving average filter before computing the correlations. Open (solid) bars indicate values for annual (summer) trends.

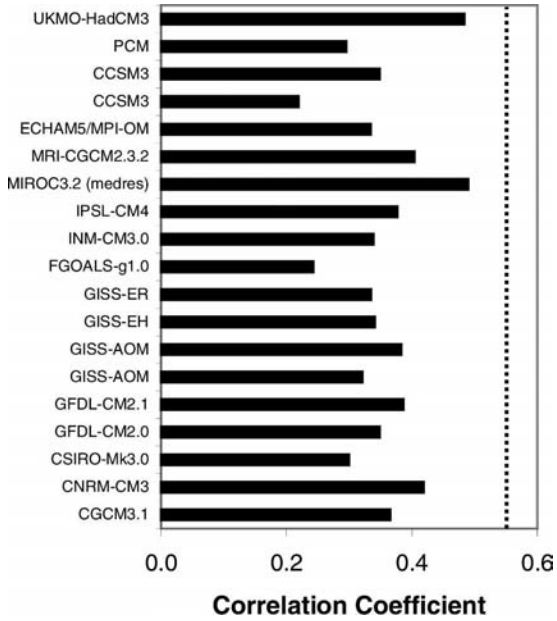

FIG. 10. Correlation coefficients between CUS temperature time series for observations (the 1901–99 period) and 99-yr segments of preindustrial control simulations. These are the maximum values obtaining by moving the model simulation segment by 1 year and repeating for every possible segment. Time series were smoothed with an 11-yr moving average filter before computing the correlations. The dashed vertical line indicates the value of apparent statistical significance at the 95% level of confidence.

time series (Fig. 4) for the Met Office (UKMO) Third Hadley Centre Couped Model (HadCM3), tied with the MIROC3.2 (medres) for the highest correlation coefficient of +0.49, exhibits a broad peak in the first half of the period, followed by a sequence of a sharp drop and sharp rise of more than 1°C. This illustrates the magnitude of multidecadal scale changes possible from internal variability alone. However, considering the large number of sequences examined (~4900), the maximum variance explained (<25%) for those sequences closest to observed is quite small. One interpretation of this finding is that internal variability alone is an unlikely cause of the observed behavior and that both internal and external forcing are required to fully explain the observations. Delworth and Knutson (2000) investigated the early twentieth-century rise in global surface temperature using an ensemble of model simulations and found that a combination of unusually high internal variability and external forcing was required to reproduce this feature. However, an analysis of the lagged autocorrelation function of the CUS temperature time series shows that the model spectra are not red enough. For example, the 1-decade lag correlation coefficient is 0.42 for observations, compared to averages of only 0.05 and 0.14 for the PICNTRL and 20C3M simulations, respectively. Thus, model deficiencies introduce some uncertainty about the above interpretation.

Several of the modeling groups performed multiple simulations with identical twentieth-century forcing but different initial conditions, providing an additional opportunity to investigate internal model variability. The

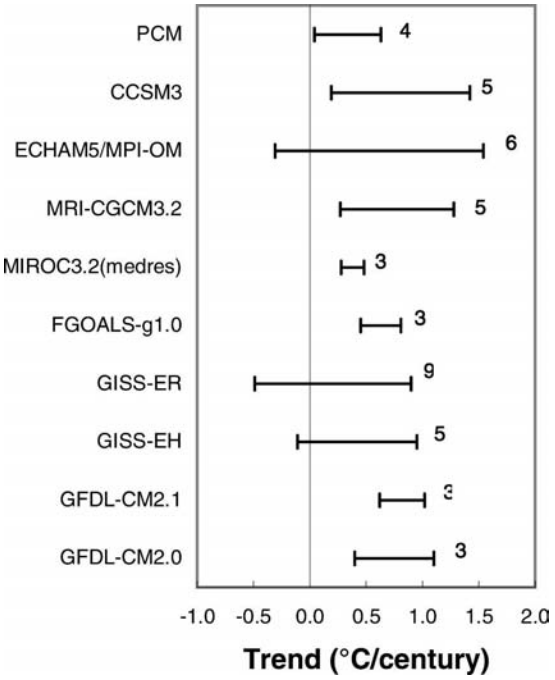

FIG. 11. Range of annual CUS temperature trends during 1901–99 for models with at least three simulations for the 20C3M experiment. The integer to the right of the range is the number of simulations.

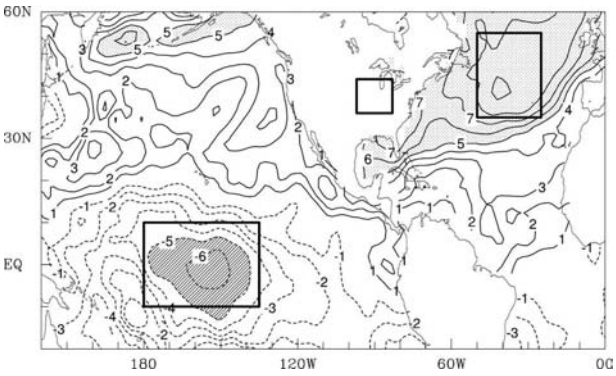

FIG. 12. Distribution of correlation coefficients between CUS air temperature and SSTs observed during 1901–99. Both time series were smoothed with an 11-yr moving average filter before computing correlations. The boxes show the outlines of the central United States, North Atlantic, and central equatorial Pacific regions. The shading highlights areas with absolute values of correlations greater than 0.5. Contour line labels are values times 10.

range of 1901–99 trends for those 10 models performing at least three simulations (Fig. 11) is quite large in some cases. The range for the fifth-generation ECHAM/Max Planck Institute for Meteorology Ocean Model (ECHAM5/MPI-OM) is −0.3° to +1.5°C century^{-1}. The range is greater than 1°C century^{-1} for four other models [the Community Climate System Model version 3 (CCSM3), the Meteorological Research Institute coupled GCM version 3.2 (MRI-CGCM3.2), GISS-ER, and GISS-EH]. The models with smaller ranges all have four or fewer simulations and the smaller range may, in part, reflect the small sample size. These ranges of regional trends are much larger than the ranges of global trends. Nine of the 10 models with multiple simulations have global temperature trend ranges of less than 0.3°C century^{-1} (not shown); the exception is the Geophysical Fluid Dynamics Laboratory Coupled Model version 2.0 (GFDL-CM2.0), which has a range of 0.4°C century^{-1}. The large range of regional trends provides support for the initial assumption to treat multiple runs as independent model simulations.

The wide range of trends in the multiple historical forcing simulations of the 20C3M experiment points to the existence of large internally generated climate variability in the CUS. This finding (large variability) combined with the result that behavior similar to observed occurs in a few of the forced and control simulations supports the possibility that internal variability is an important contributor to the twentieth-century lack of warming, plus the multidecadal features of an early century rapid warming, midcentury cooling, and small changes in the late century.

c. SST teleconnections

The observed correlations of CUS temperature with pointwise SSTs for 1901–99 (Fig. 12) are very high (>+0.80) in the North Atlantic. There is a large area of negative correlations (<−0.50) in the central equatorial Pacific. Two latitude–longitude boxes were defined for evaluating the model simulations with respect to these observed features: North Atlantic (NA: 35°–55°N, 25°–50°W) and central equatorial Pacific (CEP: 10°S–10°N, 135°W–180°); the boxes are outlined in Fig. 12 along with the CUS box. Correlation coefficients between CUS temperature and SSTs averaged in these boxes were calculated for observations and individual simulations. The results (Fig. 13) for the NA region indicate that nearly all simulations have positive coefficients, 18 have coefficients that exceed the apparent statistical significance level of 95% confidence, and 3 [the Institute for Numerical Mathematics Coupled Model version 3.0 (INM-CM3.0), CCSM3 run 5, and UKMO-HadCM3] exceed +0.80, compared to the observed value of +0.85. By contrast, for the CEP region where the observed correlation is −0.55, only seven simulations have negative correlations and only one (GISS-ER run 1) has an apparent statistically significant negative correlation (−0.62). For six simulations,

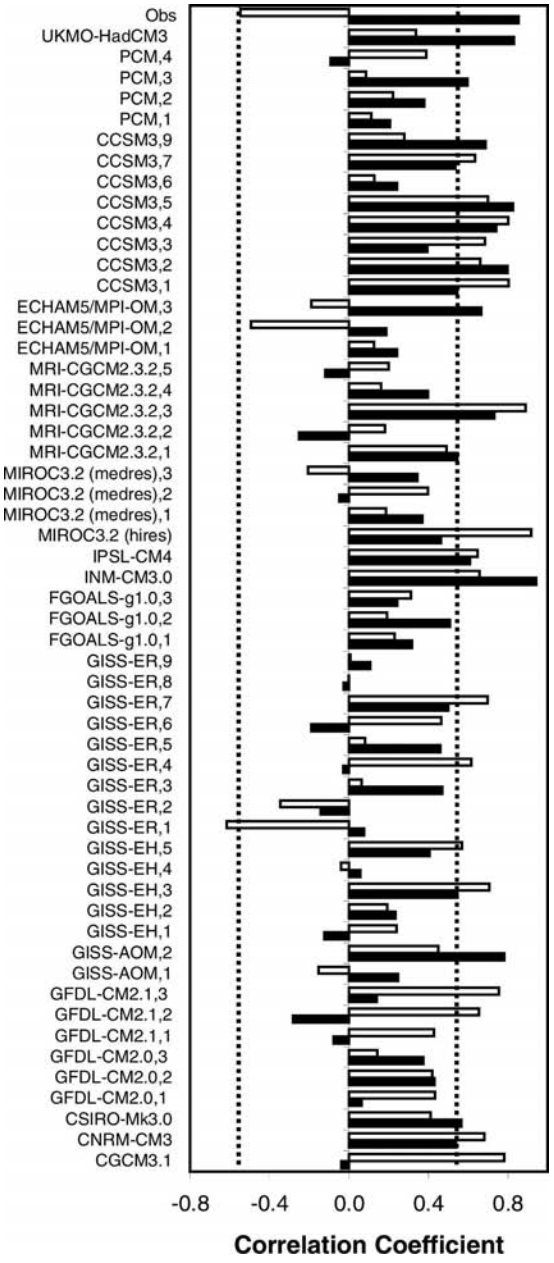

FIG. 13. Correlation coefficients between the 1901 and 1999 time series of CUS temperature and SSTs in the NA (solid bars) and the CEP (open bars) region. Time series were smoothed with an 11-yr moving average filter before computing correlations. The integer after the comma in the label indicates the run number for those models with multiple simulations. The dashed vertical lines indicate the values of apparent statistical significance at the 95% level of confidence.

the correlation coefficients are positive for NA and negative for CEP, similar to observed. As found in the trend analysis, there is substantial variability among multiple runs of the same model. For example, the CEP

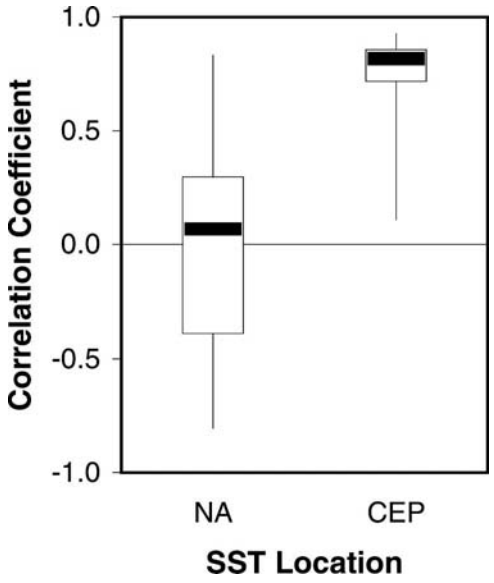

FIG. 14. Box and whisker plot illustrating the distribution of correlation coefficients between observed and 20C3M model SSTs for the NA and CEP regions. The upper and lower limits of the box indicate the 75 and 25 percentile values, the solid horizontal line in the box indicates the median, and the ends of the whiskers indicate the highest and lowest value.

region correlation coefficient for GISS-ER varies from −0.62 to +0.70.

The SST analysis is interesting because the agreement with observations in the NA region is considerably better overall than for trends, indicating a qualitatively correct teleconnection relationship between the model NA SSTs and model CUS temperature. However, the poor agreement in the CEP region suggests possible deficiencies with tropical Pacific teleconnections, which is explored later. Since a number of the simulations produce the correct sign of the correlations between model NA SSTs and model CUS temperature, but negative correlations between model and observed CUS temperature, this suggests that the temporal evolution of NA SSTs is different than observed. Correlations between observed and 20C3M model SST time series (Fig. 14), in a box and whiskers presentation, indicate rather high correlations in the CEP region. However, in the NA region the correlations vary widely, the model median being near 0. Thus, for the NA region the temporal evolution of SSTs is highly variable among the model simulations and generally not in good agreement with the observed evolution. This could reflect model deficiencies or sensitivity of the temporal evolution of SSTs to initial conditions in the models.

An analysis of the SST variability (Fig. 15), computed as in Fig. 5, indicates that the majority of models pro-

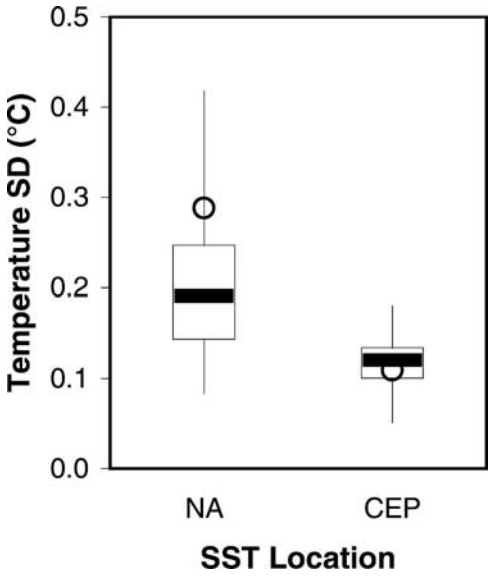

FIG. 15. As in Fig. 14 but for the standard deviation of the SST time series for the 20C3M experiment for the NA and CEP regions.

duce too little variability in the NA region and the range among models is very large. By contrast, in the CEP region, the model values are much closer to observed levels.

4. Discussion

A recent study (Sutton and Hodson 2005) found a positive and statistically significant relationship between the Atlantic multidecadal oscillation (AMO) index and summer temperature in portions of North America, with model simulations indicating that the AMO is the driver of the variations. The entire CUS region defined here is included in their area of positive response. The analysis presented herein (Fig. 12) shows very high positive correlations between NA SSTs and CUS annual temperature. A seasonal analysis (not shown) indicates a similar positive correlation for all seasons except spring. While the NA region defined here for calculating an SST index is smaller than that used for calculating the AMO index, it does include the area where local SST variations are most highly correlated with the AMO index (Sutton and Hodson 2005). In fact, the correlation coefficient between CUS annual temperature and the AMO index is +0.63, statistically significant at the 95% level of confidence. Thus, the AMO appears to be perhaps a very important factor involved in the observed multidecadal variations of CUS temperature. Variations in CEP SSTs are also strongly related (Fig. 12), but correlations are not quite as high as for NA SSTs.

The model simulations of these relationships are mixed. Most models do simulate the right sign of the relationship between NA SSTs and CUS temperature, and correlations reach levels that would be apparently significant at the 95% level of confidence for about a fourth of the simulations. However, the correlations between model simulations and observations for the twentieth-century time series of NA SSTs are in most cases small, well below apparent levels of statistical significance. Thus, few models simulate the actual observed evolution of NA SSTs. Furthermore, most models produce too little variance of NA SSTs.

Differences in simulating the temporal evolution of NA SSTs could result, at least in part, from sensitivity to initial conditions. The wide range of correlation values from multiple runs of the same model suggests that there is considerable sensitivity. However, the low SST variance simulated by models points to a more fundamental problem, perhaps related to correct simulation of fluctuations in the thermohaline circulation and/or of teleconnections with Pacific SSTs. Lower variance would lead to lesser anomalous forcing on the atmosphere and perhaps reduce the strong coupling between NA SSTs and CUS temperature variations.

The picture with CEP SSTs is different. There are high correlations between observed and model SSTs in the CEP region and variance levels are similar to observed, but the correlations between model CEP SSTs and CUS temperature are generally of the wrong sign. Robinson et al. (2002) contended that warm tropical SSTs result in increased atmospheric water vapor content, cloud cover, and precipitation and decreased temperature in the central and eastern United States Their model simulations indicated increased precipitation in the central United States in response to warm tropical Pacific SSTs and an analysis of observed data showed upward trends in cloud cover and atmospheric water vapor content and downward trends in temperature during 1951–97. Although century-long datasets of atmospheric water vapor content and cloud cover are not available to examine their response to observed multidecadal CEP SST variations, the response of U.S. precipitation patterns to CEP SSTs was examined. This analysis, covering the period 1901–99, revealed that annual CUS precipitation is positively correlated with CEP SSTs and negatively correlated with CUS temperature (Fig. 16), in general agreement with the Robinson et al. (2002) model results covering the shorter period 1951–97. The area of highest correlations includes the northwest part of the CUS region. On a seasonal basis (not shown), the relationships are similar in spring, summer, and fall, but opposite in winter. A majority of the models show a similar positive correla-

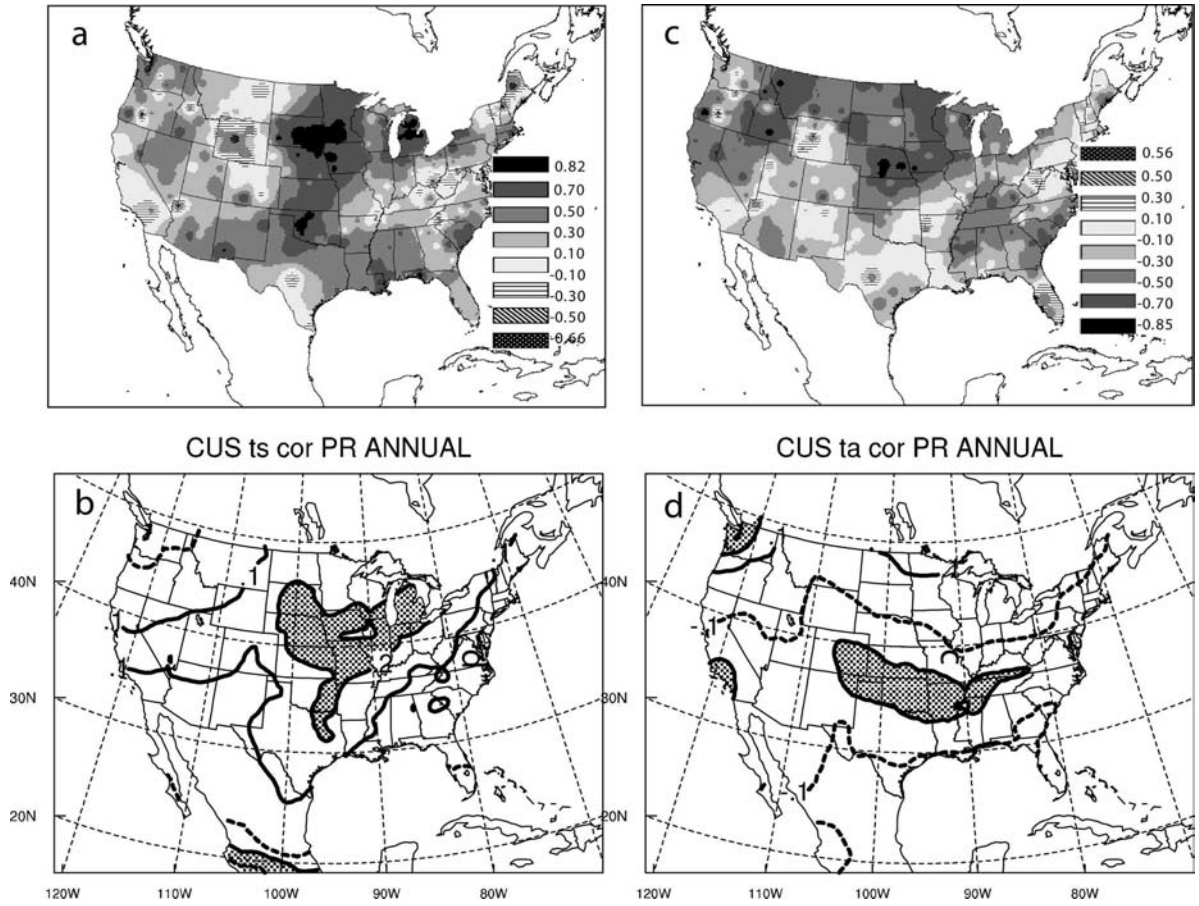

FIG. 16. Correlation coefficients between (a) observed CEP SSTs and annual precipitation for U.S. climate divisions, (b) model CEP SSTs and annual precipitation, (c) observed CUS annual temperature and annual precipitation for U.S. climate divisions, and (d) model CUS annual temperature and annual precipitation. In (b) and (d) positive (negative) contours are solid (dashed) and the contour interval is 0.1. Values above +0.20 [below −0.20] are shaded in (b) [(d)].

tion between CEP SSTs and CUS annual precipitation. A model mean map of correlations (Fig. 16) indicates the highest values are largely contained within the CUS region although the values are not particularly high, reflecting considerable model variability. For the link between precipitation and CUS temperature, model mean correlations (Fig. 16) are negative over large areas of the United States, of the same sign as observed, but the absolute values are small and the area of largest negative values is to the south of the observed area.

Past studies (e.g., Lau and Nath 1996, 2001) proposed that equatorial Pacific SST variations induce an atmospheric response that in turn forces SST variations in the North Atlantic, this atmospheric "bridge" applying to ENSO events. Although the present study focuses only on multidecadal time scales, the correlation coefficient between observed CEP and NA SSTs is a statistically significant −0.53, suggesting a physical connection of some type. Correlation coefficients between CEP and NA SSTs for each 20C3M model simulation (Fig. 17) do not agree well with the observed value. A minority (15) of the 55 simulations have negative values, only 2 of which reach levels of apparent statistical significance. This poor representation of the observed linkage may partially explain the low NA SST variance in the models. Since the models represent the connection between NA SSTs and CUS temperature rather well, the general lack of agreement between observed and model CUS temperature trends may arise partially from model deficiencies related to representation of the Pacific–Atlantic teleconnections.

In the Pan et al. (2004) study using a single CGCM, a high-resolution regional climate model (RCM) was required to produce a warming hole through a positive feedback mechanism involving an increase in low-level jet frequency and associated changes in moisture convergence and soil moisture, mesoscale processes that may not be well simulated in lower-resolution global

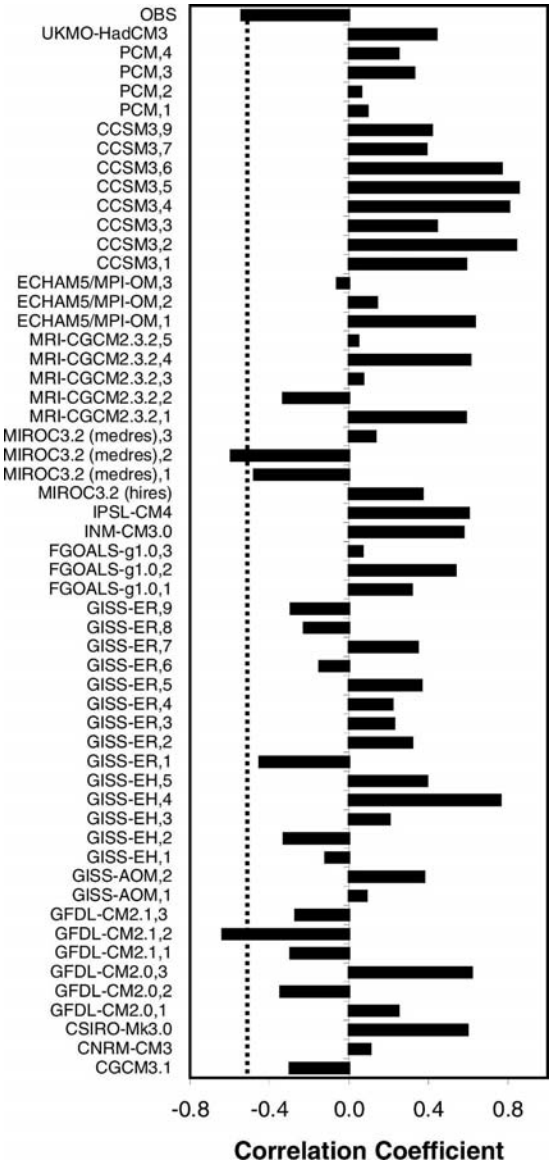

FIG. 17. Correlation coefficients for the 1901–99 time series between NA SSTs and CEP SSTs. Time series were smoothed with an 11-yr moving average filter before computing correlations. The integer after the comma in the label indicates the run number for those models with multiple simulations. The dashed vertical line indicates the value of apparent statistical significance at the 95% level of confidence.

models. Another recent study (Liang et al. 2006), also using an RCM, found a warming hole in the central United States relative to the driving CGCM. However, in the latter study, the warming hole appeared to arise from better representation of cumulus convection compared to the CGCM. The high temporal resolution data needed to investigate low-level jet frequencies was not available from the modeling groups, precluding an investigation of this mechanism. However, an analysis of monthly wind data provides insights into the flow changes associated with temperature variations in the models. Model mean maps of the correlation coefficient between CUS temperature and point-wise meridional wind component at 850 hPa and the zonal wind component at 200 hPa calculated over the period 1901–19 (Fig. 18) revealed very direct relationships. At 850 hPa, positive correlations are found in the CUS in all seasons such that warm periods are associated with enhanced southerly flow, indicating the importance of decadal-scale changes in advection in the models. At 200 hPa, correlations are positive to the north and negative to the south, implying that warm periods are associated with a northward displacement of the westerlies, an expected relationship from thermal wind considerations.

5. Conclusions

A few of the model simulations exhibit trends and multidecadal variations similar to observed CUS temperature, both in the PICNTRL and 20C3M experiments. Interestingly, the observed early century warming and midcentury cooling occur in only a very few models, while a low rate of late-century warming, similar to observed, occurs in a sizeable number of models, with many even exhibiting cooling (although the majority of models exhibit warming greater than observed). Over the entire twentieth century, a few of the simulations exhibit cooling, as observed. Another very interesting result is the wide range of trends and correlation coefficients among different twentieth-century simulations from the same model. This is indicative of the considerable internal variability produced by the models at a regional scale.

Although the cause(s) of the observed variations has not been definitively determined, there are high observed correlations between CUS temperature and SSTs in the CEP and NA. The relationship to North Atlantic variations is particularly strong, with the index used here essentially reflecting the Atlantic multidecadal oscillation. Thus, correct model simulation of this coupling (CUS temperature versus NA SSTs) and the coupling between CEP and NA SSTs is presumably crucial to reproduction of the warming hole. These results support the following conclusions:

1) The models appear to simulate rather well the link between CUS temperature and NA SSTs and also the observed variations of CEP SSTs, but not the apparently critical connection between NA and CEP SSTs. This may partially explain the low NA

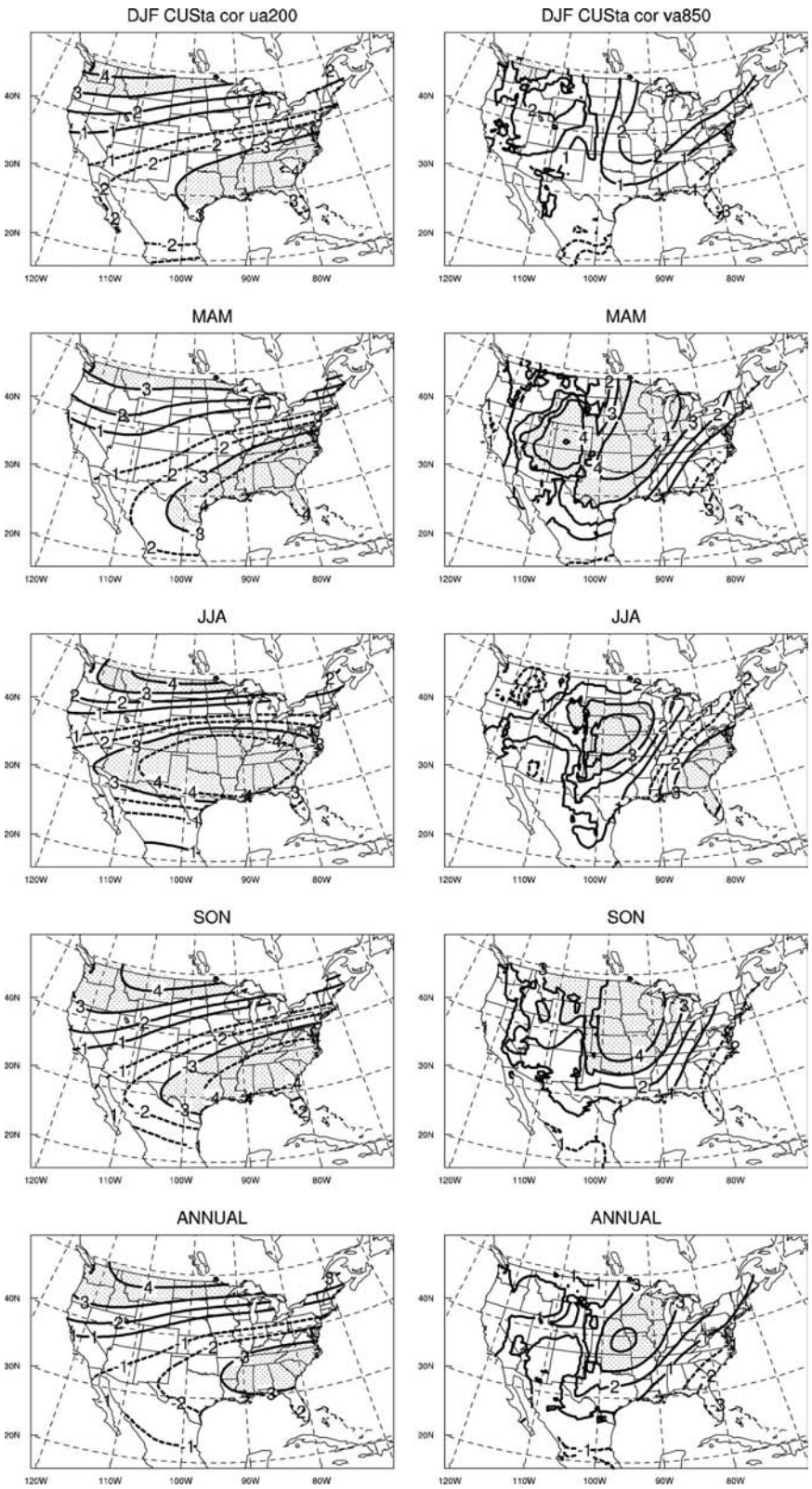

FIG. 18. Model mean maps of correlation coefficients between CEP temperature and (left) meridional wind speed at 850 hPa and (right) zonal wind speed at 200 hPa for each season and annual means. The monthly data were aggregated to seasonal or annual averages before performing the correlations. Solid (dashed) lines indicate positive (negative) values. The shading highlights areas with absolute values of correlations greater than 0.3. Contour line labels are the values times 10.

SST variance in the models and compromise their ability to produce the CUS warming hole. Also, the possible direct influence of CEP SSTs on CUS temperature is not reflected in most model simulations.

2) The wide range of trend outcomes, particularly for those models with multiple runs of the twentieth century, is striking. Even considering that models may be biased in their simulation of the mean and variability and do not capture all of the important SST linkages, the range is sufficiently large to suggest that large multidecadal regional variability in the CUS is an internal dynamic feature inherent to the climate system.
3) The presence of selected simulations in both the PICNTRL and 20C3M historical forcing simulations that exhibit similarity to the observed behavior indicate that internal variability is a strong candidate as a hypothesis to explain the CUS warming hole and the associated multidecadal features including the substantial warming (cooling) in the early (middle) part of the twentieth century.
4) The small number of 20C3M simulations that are similar to observed suggests that the warming hole is not a robust response to twentieth-century external forcing. This conclusion is tempered by the possibility that model limitations and uncertainties in external forcing may contribute to the general lack of similarity to the observed behavior. Nevertheless, support for this conclusion derives from the large number of models included in this analysis and the absence of a consistent response among multiple runs from single models.
5) The small explained variance (~25%) in the PICNTRL sequences most highly correlated with the observed CUS temperature time series suggests that internal variability alone is not adequate to explain the observed fluctuations; rather, they are likely due to a combination of unusually large realization of internal variability and external forcing, the same conclusion reached by Delworth and Knutson (2000) for early twentieth-century global warming. However, the spectra of the model time series are not red enough, pointing to model deficiencies that introduce uncertainty in this interpretation.

The warming hole indicates that anthropogenic forcing of the climate system can be accompanied by a regional temperature response different than expected; this has important implications for impacts assessments. For example, certain aspects of air quality are sensitively dependent on temperature and important chemical processes occur on a regional scale. This analysis suggests that internal variability at a regional scale is large and thus there is a need for multiple simulations of the future to provide probabilistic assessments. In the aforementioned case of air quality, air quality modeling simulations of future changes are quite computationally expensive, but multiple simulations may be necessary to adequately capture the range of potential future outcomes.

Acknowledgments. We acknowledge the international modeling groups for providing their data for analysis: the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modelling (WGCM) and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model data analysis activity, the IPCC WG1 TSU for technical support, and NCSA/UIUC and NOAA/FSL for supercomputer support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy. Partial support for this project was provided by the Illinois Environmental Protection Trust Fund, the U.S. Environmental Protection Agency (EPA) under award EPA RD-83096301-0, and the National Oceanic and Atmospheric Administration (NOAA) under Contract EA133E-02-CN-0027. The views expressed herein are those of the authors and do not necessarily represent those of the EPTF, NOAA, USEPA, or the Illinois State Water Survey.

REFERENCES

Delworth, T. L., and T. R. Knutson, 2000: Simulation of early 20th century global warming. *Science,* **287,** 2246–2250.

Folland, C. K., and Coauthors, 2001: Observed climate variability and change. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 99–181.

Hansen, J., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl, 2001: A closer look at United States and global surface temperature change. *J. Geophys. Res.,* **106,** 23 947–23 963.

Kunkel, K. E., and X.-L. Liang, 2005: CMIP simulations of the climate in the central United States. *J. Climate,* **18,** 1016–1031.

Lau, N.-C., and M. J. Nath, 1996: The role of the "atmospheric bridge" in linking tropical Pacific ENSO events to extratropical SST anomalies. *J. Climate,* **9,** 2036–2057.

——, and ——, 2001: Impact of ENSO on SST variability in the North Pacific and North Atlantic: Seasonal dependence and role of extratropical sea–air coupling. *J. Climate,* **14,** 2846–2866.

Liang, X.-Z., L. Li, K. E. Kunkel, M. Ting, and J. X. L. Wang, 2004: Regional climate model simulation of U.S. precipitation during 1982–2002. Part I: Annual cycle. *J. Climate,* **17,** 3510–3528.

——, J. Pan, J. Zhu, K. E. Kunkel, J. X. L. Wang, and A. Dai, 2006: Regional climate model downscaling of the U.S. sum-

mer climate and future change. *J. Geophys. Res.,* **111,** D10108, doi:10.1029/2005JD006685.

Pan, Z., R. W. Arritt, E. S. Takle, W. J. Gutowski Jr., C. J. Anderson, and M. Segal, 2004: Altered hydrologic feedback in a warming climate introduces a "warming hole." *Geophys. Res. Lett.,* **31,** L17109, doi:10.1029/2004GL020528.

Rayner, N. A., E. B. Horton, D. E. Parker, C. K. Folland, and R. B. Hackett, 1996: Version 2.2 of the global sea-ice and sea surface temperature data set, 1903–1994. Climate Tech. Note 74, Hadley Centre for Climate Prediction and Research, U.K. Meteorological Office, 43 pp.

——, 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.

Robinson, W. A., R. Reudy, and J. E. Hansen, 2002: On the recent cooling in the east-central United States. *J. Geophys. Res.,* **107,** 4748, doi:10.1029/2001JD001577.

Stouffer, R. J., G. Hegerl, and S. Tett, 2000: A comparison of surface air temperature variability in three 1000-yr coupled ocean–atmosphere model integrations. *J. Climate,* **13,** 513–537.

Sutton, R. T., and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science,* **309,** 115–118.