



RESEARCH LETTER

10.1002/2016GL071567

Key Points:

- We use large ensembles of simulations to determine whether anthropogenic aerosols were responsible for the U.S. warming hole
- The large ensembles suggest that the warming hole was a rare event that contained a large contribution from unforced internal variability
- Anthropogenic aerosols increased the likelihood of the warming hole, but their effects were canceled to a large extent by greenhouse gases

Supporting Information:

- Supporting Information S1

Correspondence to:

A. Banerjee,
ab4283@columbia.edu

Citation:

Banerjee, A., L. M. Polvani, and J. C. Fyfe (2017), The United States “warming hole”: Quantifying the forced aerosol response given large internal variability, *Geophys. Res. Lett.*, *44*, 1928–1937, doi:10.1002/2016GL071567.

Received 24 OCT 2016

Accepted 4 FEB 2017

Accepted article online 7 FEB 2017

Published online 22 FEB 2017

The United States “warming hole”: Quantifying the forced aerosol response given large internal variability

A. Banerjee¹ , L. M. Polvani^{1,2} , and J. C. Fyfe³

¹Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, USA, ²Department of Earth and Environmental Sciences, Lamont Doherty Earth Observatory, Palisades, New York, USA, ³Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, British Columbia, Canada

Abstract Twenty-five years of large summer cooling over the southeastern United States ending in the mid-1970s coincided with rapidly increasing anthropogenic aerosol emissions. Here we assess the claim that the cooling in that period was predominantly due to such aerosols. We utilize two 50-member sets of coupled climate model simulations, one with only anthropogenic aerosol forcings and another with all known natural and anthropogenic forcings, together with a long control integration. We show that, in the absence of aerosol forcing, none of the model simulations capture the observed surface cooling rate ($\sim 0.56^\circ\text{C decade}^{-1}$), whereas with increasing aerosol emissions 2 (of 50) of the simulations do. More importantly, however, we find that the cooling from aerosols ($0.20^\circ\text{C decade}^{-1}$) is insufficient to explain the observation. Our results therefore suggest that, while aerosols may have played a role, the observed cooling was a rare event that contained a large contribution from unforced internal variability.

1. Introduction

Understanding the human influence on past climate change is of utmost importance in informing future climate policies. However, signals of external forcing on climate can be obscured by the presence of internal variability, which is inherent to the climate system [e.g., *Deser et al.*, 2012]. Quantifying the magnitude of internal variability is, therefore, a crucial step toward robustly determining whether an external forcing caused an observation.

In this study, we apply the above concepts in the context of the twentieth century United States (U.S.) “warming hole” [e.g., *Robinson et al.*, 2002; *Pan et al.*, 2004; *Kunkel et al.*, 2006]. Here the warming hole is defined as the large cooling in near surface air temperatures that occurred over the southeastern (SE) U.S. in summer (June–August, JJA) between 1951 and 1975. Figure 1 illustrates the observed cooling trends within the boxed region of the SE U.S. and shows that they stand in contrast to warming trends over the western U.S. Averaging over the SE U.S., Figure 2a shows the summertime temperature time series over the entire twentieth century. The linear trendline highlights the large rate of cooling ($0.56^\circ\text{C decade}^{-1}$) between 1951 and 1975. In order to place this cooling within the context of twentieth century temperature trends, Figure 2b shows consecutive, overlapping 25 year temperature trends over the century for the SE U.S. Clearly, the magnitude of the warming hole was uniquely anomalous for that century.

In the present work, we focus on the claim by *Leibensperger et al.* [2012b, henceforth, L12b] that the warming hole was predominantly driven by increasing emissions of anthropogenic aerosols, in particular, of sulfate. Figure 2a highlights the rationale for this hypothesis and the time period (1951–1975, JJA) that we choose to investigate its validity: the observed surface cooling over the SE U.S. during this time (red line) coincides with a large increase in the U.S. sulfate aerosol burden (black line) as simulated by our climate model (see section 2). Both effects are most pronounced in JJA than in other seasons or the annual mean (Figure S1 in the supporting information).

Using the Goddard Institute for Space Studies General Circulation Model 3 (GISS GCM 3), L12b contrasted averages of five model ensemble members in two experiments performed between 1950 and 2050: one with and one without U.S. anthropogenic aerosol sources. They found that peak U.S. aerosol loadings (1970–1990 average) cause a cooling of $0.5\text{--}1.0^\circ\text{C}$ over the central and eastern U.S. during summer/autumn in their model. That response was ascribed to a combination of the aerosol direct and indirect radiative effects,

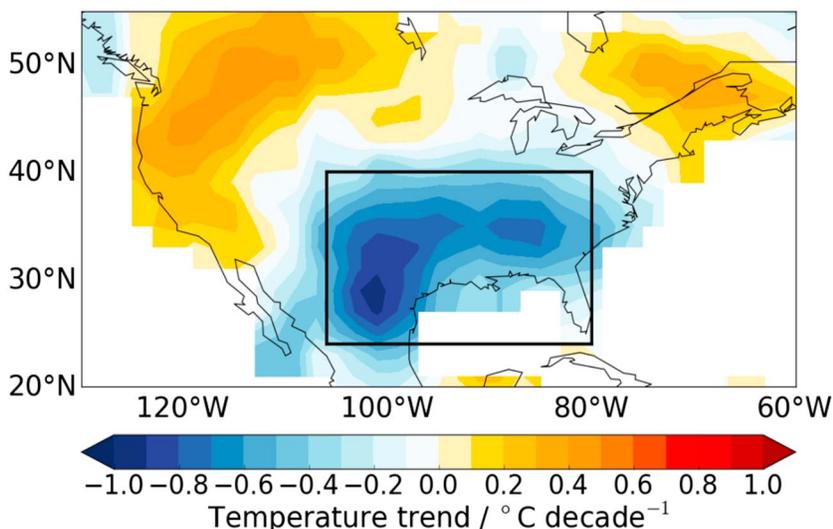


Figure 1. Map of the linear trend in near surface air temperature ($^{\circ}\text{C decade}^{-1}$) in GISS Surface Temperature Analysis (GISTEMP) observations [Hansen et al., 2010] over the U.S. between 1951 and 1975 (JJA). The boxed region represents the SE U.S. and encompasses the area within $24\text{--}40^{\circ}\text{N}$, $80\text{--}106^{\circ}\text{W}$.

modified by changes in atmospheric circulation and the hydrological cycle. The modeled cooling from aerosols was proposed as the main cause of the linear rate of cooling found between 1930 and 1990 in the observations, which occurred over a similar region. However, contrasting the two experiments as averages of model ensemble members only isolates the forced aerosol response but neglects to account for internal variability, which motivates our study.

The aim of this work is to determine whether changes in anthropogenic aerosol emissions can indeed robustly explain the warming hole, given the presence of internal variability. We here answer this question using two large initial condition ensembles of simulations forced by (i) anthropogenic aerosols alone and (ii) all historical forcings. The simulations are performed with the Canadian Earth System Model (CanESM2). These are supplemented by an all forcing large ensemble from the Community Earth System Model (CESM1). As highlighted in Deser et al. [2012], by removing uncertainties due to both intermodel and forcing differences, the use of large ensembles allows for a clean quantification of the relative magnitudes of the modeled forced response and

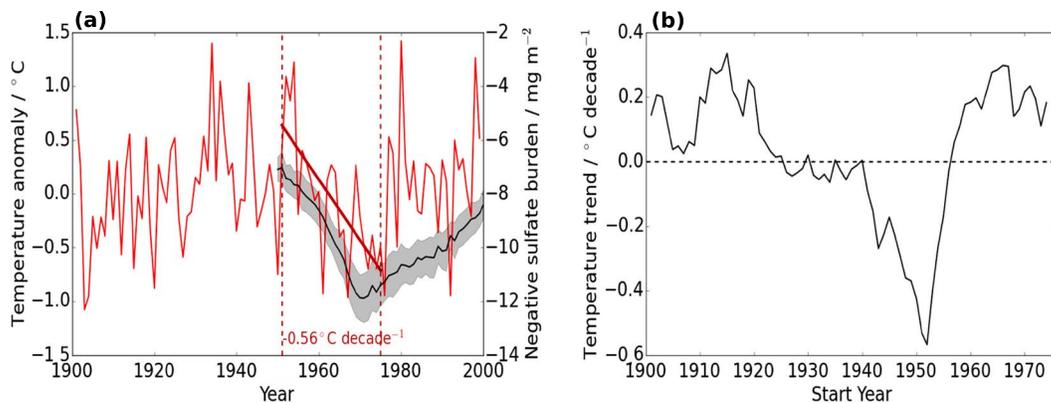


Figure 2. (a) Time series of JJA near surface air temperature ($^{\circ}\text{C}$, anomaly from 1951 to 1990) over the SE U.S. in the observations over the twentieth century, shown by the red line. Overlaid is the negative time series of the sulfate burden, averaged over the U.S. ($24\text{--}50^{\circ}\text{N}$, $70\text{--}124^{\circ}\text{W}$), in the historical all forcing ensemble of the Canadian Earth System Model (CanESM2) (see section 2 for details); this is shown as the ensemble mean (black line) $\pm 1\sigma$ (shading). (b) Consecutive, overlapping 25 year temperature trends in JJA near surface air temperature ($^{\circ}\text{C decade}^{-1}$) over the SE U.S. in the observations over the twentieth century. This is plotted as a function of the trend start year.

internal climate variability; this is in contrast to previous multimodel approaches that have explored the role of internal variability in driving the warming hole [Kunkel *et al.*, 2006; Kumar *et al.*, 2013]. Both CanESM2 and CESM1 also performed long unforced control runs under constant 1850 conditions. We utilize these control runs to cleanly quantify the magnitude of internal variability.

2. Data and Methods

2.1. Model Description

This study uses simulations performed by two state-of-the-art, ocean-coupled climate models, CanESM2 and CESM1, both of which participated in the Coupled Model Intercomparison Project phase 5 (CMIP5) [Taylor *et al.*, 2012]. The analysis will primarily focus on data from CanESM2, which conducted all the simulations needed here (discussed below), while results from CESM1 are provided as supporting information. The treatment of atmospheric anthropogenic aerosols (in particular, of sulfate) is now briefly described for each model, since this is the main forcing of interest in this study.

The atmospheric component of CanESM2 is the Canadian Atmospheric Global Climate model (CanAM4) at T63 horizontal resolution ($\sim 2.8^\circ$) [von Salzen *et al.*, 2013]. Sulfate aerosol is generated following clear-sky and in-cloud oxidation of SO_2 by 3-D climatological, monthly oxidant fields (OH, NO_3 , O_3 , and H_2O_2). CanAM4 computes the direct and cloud albedo (first indirect) effect of aerosols but the cloud lifetime (second indirect) effect is not included. The total 1850–2000 effective radiative forcing (ERF) from aerosol direct and indirect effects has been calculated to be -0.87 W m^{-2} (with a contribution of -0.90 W m^{-2} from sulfate aerosol) for CanESM2 [Myhre *et al.*, 2013]. This lies well within 1σ of the reported multimodel mean of a set of CMIP5 and Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) models ($-1.08 \pm 0.32 \text{ W m}^{-2}$).

The atmospheric component of CESM1 is the Community Atmosphere Model version 5 (CAM5), with a horizontal resolution of 0.9° latitude by 1.25° longitude [Kay *et al.*, 2015; Hurrell *et al.*, 2013]. CAM5 employs a modal aerosol scheme to calculate aerosol size distributions [Liu *et al.*, 2012] and computes both the direct and cloud mediated (first and second indirect) effects of sulfate aerosols [Ghan *et al.*, 2012]. The total aerosol ERF for CESM1-CAM5 between 1850 and 2000 is -1.44 W m^{-2} [Shindell *et al.*, 2013], which is greater in magnitude than for the value within CanESM2 reported above but still within 1σ of the CMIP5/ACCMIP multimodel mean.

We emphasize that both models contain some representation of aerosol indirect effects, which have been suggested as important in driving decadal variability in global mean climate [Wilcox *et al.*, 2013], as well as the warming hole [L12b; Yu *et al.*, 2014]. In particular, L12b find that the magnitude of aerosol-driven cooling over the U.S. is approximately doubled when including aerosol indirect effects, in addition to direct effects.

2.2. Experiments

Two sets of forcings are considered in this study: (i) anthropogenic aerosols only (AA) and (ii) all historical anthropogenic and natural forcings (ALL). Historical, monthly mean emissions of anthropogenic aerosols are obtained from the CMIP5 data set [Lamarque *et al.*, 2010] for both models. Each experiment consists of $N=50$ (42) ensemble members beginning in 1950 (1920) for CanESM2 (CESM1) [Arora *et al.*, 2011; Kay *et al.*, 2015]. Within each of these “large ensembles,” the individual members differ only in their atmospheric initial conditions. This results in a spread of climate responses across ensemble members to a given forcing due to atmospheric internal variability, while the ensemble mean represents the forced response [Deser *et al.*, 2012]. In this work, the ensemble mean trends (μ) in near surface air temperatures that are computed are considered to be significantly different from zero at the 95% confidence level using a two-sided Student's t test, i.e., when $\mu > 2\sigma/\sqrt{N}$, where σ is the standard deviation across N ensemble members.

In addition, we analyze a fully coupled preindustrial control (PI control) run, which was performed under constant 1850 conditions and is 996 (1800) years long in CanESM2 (CESM1). A long control run under constant forcing allows for a clean quantification of internal variability. By comparing the large ensembles forced by anthropogenic aerosols and all forcings to the PI control, our primary aim is to determine how these forcings affected the likelihood of the observed warming hole.

2.3. Observations

We use monthly mean near surface air temperature observations from the NASA GISS Surface Temperature Analysis (GISTEMP) [Hansen *et al.*, 2010]. These are provided on a regular $2^\circ \times 2^\circ$ grid, to which all model data are regridded for comparison.

2.4. Time Period and Region of Interest

We now discuss the time period and region chosen to define the observed warming hole in this study. Within the literature, the warming hole is commonly defined as the region of cooling found in the linear trends in near surface air temperatures between ~1950 and ~2000 or ~2010 [Misra *et al.*, 2012; Robinson *et al.*, 2002; Meehl *et al.*, 2012; Pan *et al.*, 2013; Weaver, 2013; Yu *et al.*, 2014]. However, rather than selecting this 50–60 year period, we here take linear trends over the 25 year period of 1951–1975 for the following reasons: (i) we suggest that taking a linear trend over the last half of the twentieth century is unsuitable since temperatures over the region of interest do not show a linear decrease during this time (for example, see Figure 2a), (ii) the 1951–1975 period is suitable to test the hypothesis of an anthropogenic aerosol influence on the warming hole since the pronounced observed cooling coincides with a large increase in the modeled U.S. sulfate burden (Figure 2a), and (iii) the specific period chosen (1951–1975) is consistent with at least one other study [Pan *et al.*, 2013]. The effects of small changes in the start date of the trend will be investigated and conclusions for the 1951–2000 period will also be briefly discussed. Finally, note that we are only able to investigate periods after 1950 when the CanESM2 large ensemble simulations begin.

The region over which we define the warming hole is the SE U.S. (24–40°N, 80–106°W; boxed in Figure 1). This region is displaced somewhat to the south and west of the region that is investigated in L12b although there is considerable overlap. Here we explicitly choose the region that captures the greatest observed cooling during our selected time period. (As a result, we make a similar choice to the “southeastern” region defined in Pan *et al.* [2013] who investigated the same time period).

3. Results

We first discuss the CanESM2 forced (ensemble mean) responses to anthropogenic aerosols (AA) and all forcings (ALL) in near surface air temperature trends. Figure 3a shows that the effect of increasing anthropogenic aerosols in the AA experiment is a small but statistically significant, summertime cooling over the contiguous U.S. between 1951 and 1975. This temperature response is fairly uniform over the U.S. despite the localized structure of the trend in sulfate aerosol burden (see Figure S2 in the supporting information). This is in contrast to the cooling caused by peak U.S. anthropogenic aerosol loadings in L12b, which is a localized response over the central/eastern U.S. However, we note that the spatial uniformity of the CanESM2 response is in good agreement with the CMIP5 multimodel mean response for the AA experiment (11 model mean; Figure S3a in the supporting information), and in agreement with several previous studies that show non-localized temperature effects resulting from localized aerosol forcings [Ming and Ramaswamy, 2009; Shindell *et al.*, 2010; Kasoar *et al.*, 2016].

Besides the spatial structure, it is also worth roughly comparing the *magnitude* of the temperature response to anthropogenic aerosols in CanESM2 to the results in L12b. For this comparison, we take into account the corresponding change in U.S. SO₂ emissions in each case [Westervelt *et al.*, 2015; Leibensperger *et al.*, 2012a]. L12b showed a cooling of 0.5–1°C for the effects of an additional 13 Tg(S) yr⁻¹ emitted, while CanESM2 simulates a cooling of around 0.5°C (derived from the decadal trend in Figure 3a) for an additional 4 Tg(S) yr⁻¹ emitted. Thus, CanESM2 shows a somewhat stronger temperature sensitivity to anthropogenic aerosol emissions than the model in L12b. Note that our study considers the effects of global aerosols as opposed to just U.S. aerosol sources in L12b. Thus, the remote climate effects of non-U.S. aerosols might partially explain the stronger sensitivity in our model.

For the ALL experiment, the ensemble mean response is found to be statistically insignificant over much of the U.S. (Figure 3b). Thus, the aerosol-driven cooling is canceled to a large extent by the combined impact of all other applied forcings, in particular, of increasing greenhouse gas emissions. Averaging over the SE U.S., we find ensemble mean (and 5–95% confidence interval) responses of $-0.20 \pm 0.08^\circ\text{C decade}^{-1}$ and $-0.08 \pm 0.07^\circ\text{C decade}^{-1}$ in AA and ALL, respectively. Compared to CanESM2, the ALL experiment of CESM1 shows a similarly small response across the U.S. (Figure S4 in the supporting information) including the SE region ($-0.12 \pm 0.08^\circ\text{C decade}^{-1}$). The CMIP5 multimodel mean suggests that the effects of anthropogenic aerosols alone and all forcings were small and comparable (Figure S3 in the supporting information). Responses of $-0.07 \pm 0.12^\circ\text{C decade}^{-1}$ (AA; 11 model mean) and $-0.07 \pm 0.06^\circ\text{C decade}^{-1}$ (ALL; 46 model mean) are calculated for the SE U.S., although this comparison is limited by the small number of models that performed the AA experiment.

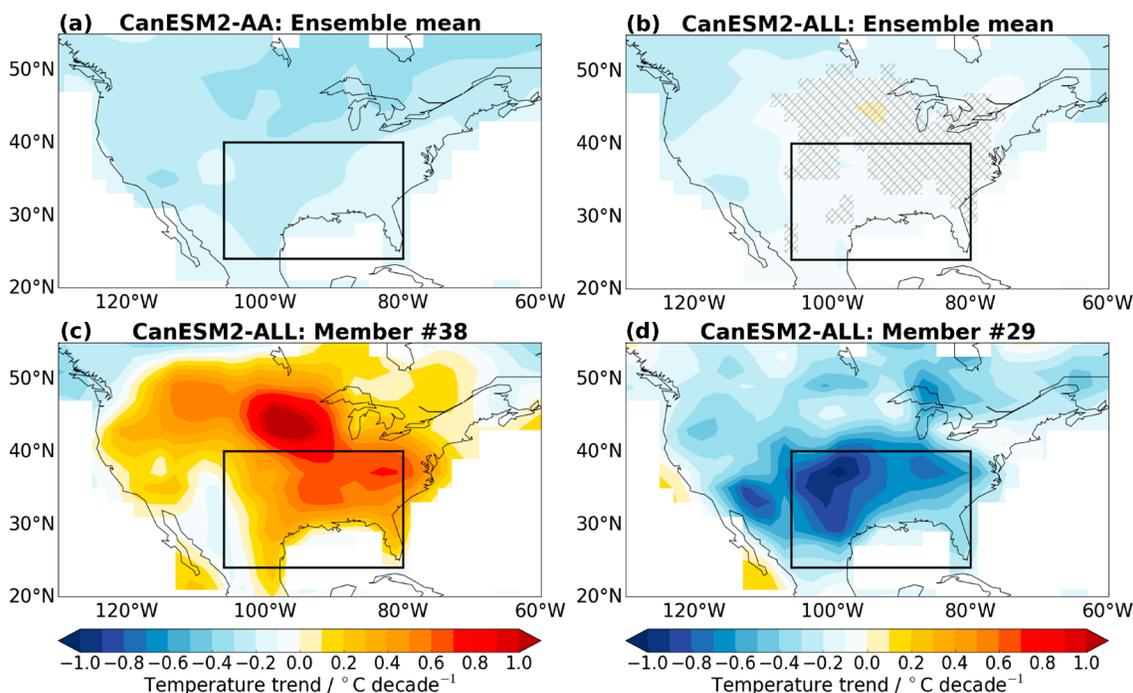


Figure 3. Modeled ensemble mean trend in near surface air temperatures between 1951 and 1975 (JJA) in the (a) AA and (b) ALL experiments of CanESM2. Hatching denotes the areas that are not significant at the 95% level, calculated using the standard deviation across the 50 ensemble members in each experiment. The following selected members are also shown: (c) the member with the strongest warming trend over the contiguous U.S. and (d) a member with strong cooling over the SE U.S., which is similar in location and magnitude to the observed warming hole (Figure 1). Modeled data is masked where there are no observations.

The key aspect that has not been explored in previous studies of the aerosol influence on the warming hole is how trends within individual members compare to the externally forced response. Here it is found that individual realizations in AA and ALL display much larger magnitude trends, of either sign (Figures S5–S7 in the supporting information) than the ensemble mean response, which indicates a larger influence of internal variability than external forcing. This is demonstrated in Figure 3 with two contrasting members of the CanESM2 (ALL) ensemble. The most extreme warming is found in Member 38, for which the trend exceeds $1^{\circ}\text{C decade}^{-1}$ in the north-central U.S. (Figure 3c). In contrast, Member 29 shows cooling over the SE U.S. of similar magnitude to that observed (compare Figures 1 and 3d). It is especially noteworthy that warming over the U.S. can even be found for several members within the AA ensemble (e.g., see Member 11 in Figure S5 in the supporting information). Therefore, simple visual inspection of individual ensemble members clearly suggests that the region and magnitude of the warming hole was greatly influenced by internal variability.

The large range of trends found within the forced ensembles motivates a more rigorous quantification of the relative influences of forcing and internal variability on the likelihood of obtaining the warming hole. To this end, Figure 4 shows summertime temperature trends over the SE U.S. between 1951 and 1975. In each panel, the observed value is shown as the red line. We first draw the reader's attention to Figures 4a and 4b, where values for each individual member of the forced ensembles of CanESM2 are shown by the blue lines (AA) and green lines (ALL); the ensemble mean response is shown as the extended vertical line of the same color. Finally, trends in consecutive, overlapping 25 year periods of the PI control run are shown by the grey shaded bars. For the PI control, the mean is shown by the black vertical line, which, expectedly, lies virtually on zero since this run contains constant external forcing.

Using Figures 4a and 4b, we first determine the influence of external forcing on the likelihood of the warming hole. The observation lies outside the PI control distribution (grey shaded bars), with a p value of 0%. This suggests that, under no external forcing, the likelihood of obtaining a cooling as extreme as the warming hole was zero. Relative to PI control, the cooling due to anthropogenic aerosols and all forcings are manifested as leftward mean shifts in the AA ($\mu = -0.20^{\circ}\text{C decade}^{-1}$, Figure 4a) and ALL ($\mu = -0.08^{\circ}\text{C decade}^{-1}$, Figure 4b) distributions, respectively. However, these forcings only slightly increase the likelihood of obtaining

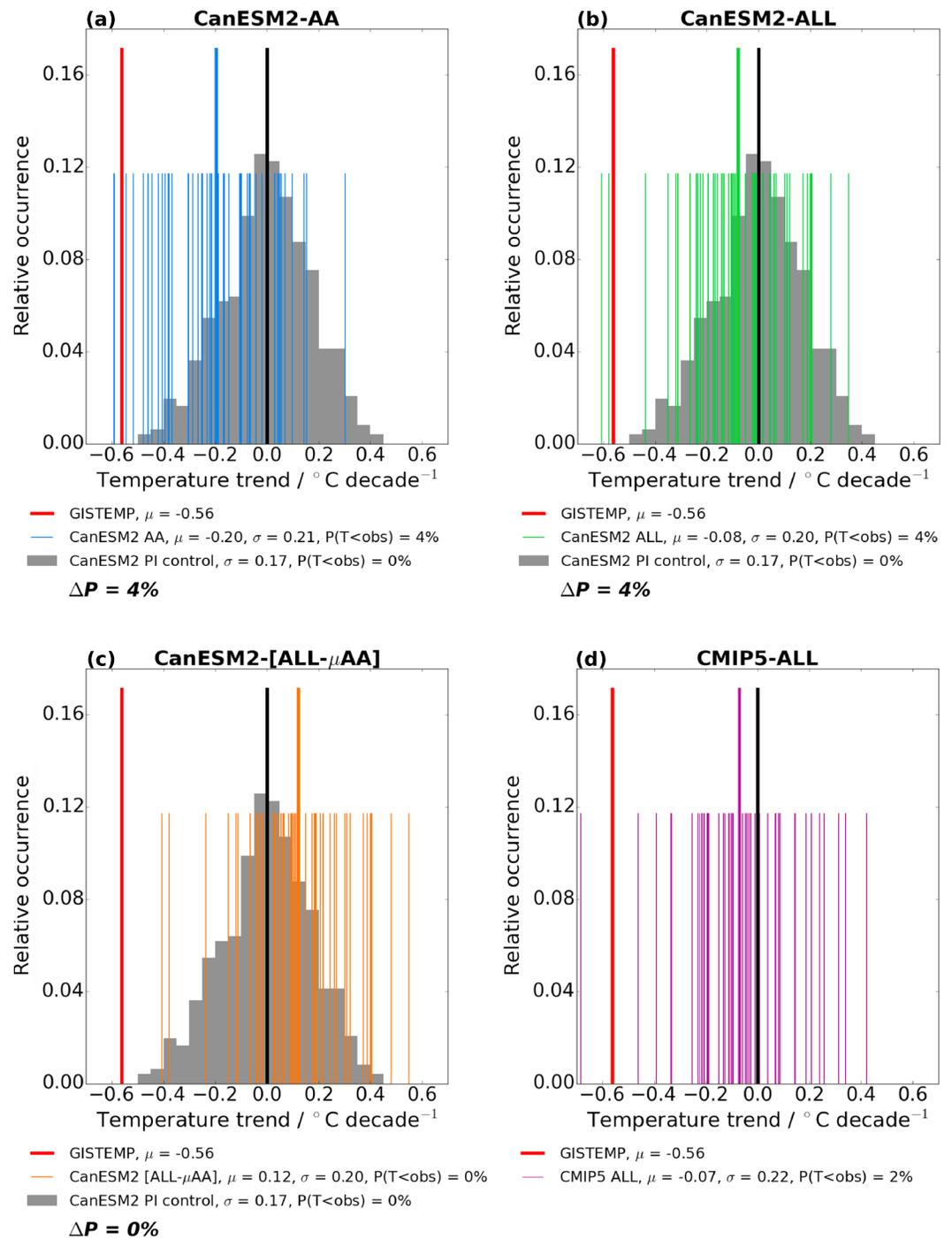


Figure 4. Trends in near surface air temperature ($^{\circ}\text{C decade}^{-1}$) over the SE U.S. between 1951 and 1975 (JJA) as modeled and observed. The observation is shown as the red vertical line in each panel. The remaining thin vertical lines show the trends for each ensemble member within the CanESM2 experiments: (a) AA (blue), (b) ALL (green), (c) ALL minus the ensemble mean of AA (brown), and (d) for each CMIP5 model (first ensemble member) within the ALL experiment (purple). Extended bold lines denote the ensemble or multimodel mean. Also shown in Figures 4a–4c is the histogram of consecutive, overlapping 25 year (JJA) trends in the PI control run of CanESM2. The mean of the PI control distribution (bold black line) lies effectively on zero. Below each panel, we report the mean (μ) and standard deviation (σ) of each distribution, the p value of the observation ($P(T < \text{obs})$), and the change in likelihood due to external forcing relative to PI control (ΔP).

Table 1. Dependence of Observed Trend in Near Surface Air Temperature Over the SE U.S. (JJA) and Corresponding Model-derived Results on the 25-yr Time Window Considered^a

Start Year	Observed / °C decade ⁻¹	CanESM2				
		μ_{AA} / °C decade ⁻¹	μ_{ALL} / °C decade ⁻¹	p_{PI} / %	$p_{AA} (\Delta p_{AA})$ / %	$p_{ALL} (\Delta p_{ALL})$ / %
1950	-0.45	-0.22	-0.09	1	12(11)	4(3)
1951	-0.56	-0.20	-0.08	0	4(4)	4(4)
1952	-0.60	-0.20	-0.07	0	2(2)	0(0)
1953	-0.43	-0.21	-0.06	1	16(15)	4(3)
1954	-0.28	-0.19	-0.04	6	28 (22)	8 (2)
1955	-0.18	-0.18	-0.01	16	48 (32)	16 (0)
1956	-0.03	-0.15	-0.02	42	76 (34)	40 (-2)

^aResults are shown for all negative observed trends with start dates from 1950 onward. Rows for observed trends that are significant at the 95% confidence level are highlighted in bold. The model values tabulated are the ensemble mean of the AA and ALL experiments (μ_{AA} and μ_{ALL}); the p value of the observation within PI control, AA, and ALL (p_{PI} , p_{AA} , and p_{ALL}); and the change in p value in AA and ALL, relative to PI control (Δp_{AA} and Δp_{ALL}).

the warming hole, by 4%, in both cases. The impact of anthropogenic aerosols as compared to all forcings is more clearly highlighted in Figure 4c, where the brown lines show the AA ensemble mean subtracted from each member of the ALL ensemble. Here no members exceed the observed cooling, in contrast to ALL where two members do. Thus, our model results suggest that anthropogenic aerosols played some role in driving the warming hole but were a minor contributor compared to internal variability.

A further conclusion that can be drawn from the above CanESM2 results is that the warming hole was a very rare event. Importantly, this conclusion is dependent on CanESM2 correctly simulating the magnitude of internal variability in near surface air temperatures over our region of focus. To determine whether this is the case, as a measure of decadal timescale internal variability, we compute the standard deviation in the 11 year smoothed temperature time series (following *Kunkel et al.* [2006]) of the twentieth century (JJA) over the SE U.S. in the GISTEMP observations and in the five members of the CanESM2 (ALL) ensemble that cover the entire historical period. We find that the standard deviation ranges between 0.22°C and 0.29°C in CanESM2, which matches well the value of 0.24°C calculated for the observations. This conclusion holds under shorter (longer) smoothing windows, for example, 5 (25) years which show the modeled standard deviation to range between 0.33 and 0.40°C (0.10–0.24°C) and observed values of 0.33°C (0.18°C). These results are in line with other studies [*Kunkel et al.*, 2006; *Deser et al.*, 2012, 2016] which, by similar measures, find models to capture the observed internal variability in surface temperatures over North America; this does not preclude suggestions of modeled underestimation in low-frequency climate variability in other regions (such as the North Atlantic) and/or deficiencies in their remote teleconnections to the warming hole region [*Kunkel et al.*, 2006; *Kumar et al.*, 2013]. Here further confidence in the CanESM2 results is gained from CESM1 and the CMIP5 models. All the models considered suggest that unforced internal variability leads to a spread of around 0.2°C decade⁻¹ in the various distributions shown in Figures 4 and S8. For the all historical forcing scenario, only one member within the CESM1 large ensemble and CMIP5 model runs produces cooling that matches or exceeds that of the warming hole, which is also in agreement with the results of CanESM2.

It is of course possible that the small likelihood of the observation suggested by all the models might reflect some deficiency that is common between them. However, Figure 2b, which shows overlapping 25 year trends over the twentieth century in the observations, highlights that the large cooling between 1951 and 1975 was unique for the century ($p \approx 1\%$). This suggests that the warming hole is rightly captured in a similarly small proportion of model ensemble members and was indeed a rare event.

Given the sharp changes in the observed 25 year trend around ~1950 (Figure 2b), it is worth exploring the sensitivity of our conclusions to small changes in the trend window considered. We consider the four 25 year trends with start dates between 1950 and 1953, after which trends become insignificant at the 95% confidence level and are positive from 1957 onward. Results are summarized in Table 1 for the observations and CanESM2. It is evident that, regardless of the time period considered, the observed value is extremely unlikely

or impossible within the PI control run. The ensemble mean responses in AA (μ_{AA}) and ALL (μ_{ALL}) are also similar between the different time periods, as are the distribution spreads (not shown). The impact of anthropogenic aerosols on the likelihood of the observed trend does show sensitivity to the time period, ranging from 2 to 15%, with the larger values found for the smaller trends between 1950 and 1974, and 1953 and 1977. At most, μ_{AA} accounts for up to half of the observed trend. However, the changes in likelihood due to all forcings reaches only 4%. In summary, we conclude from the modeling evidence that, although anthropogenic aerosols and all forcings increased the likelihood of the warming hole, it was driven to a large extent by internal variability.

Finally, although we have defined the warming hole over a relatively short 25 year period in order to match the aerosol forcing to the observed cooling signal, we note that the relative roles of forcing and internal variability, i.e., the signal-to-noise ratio, can change according to the length of the trend examined [e.g., *Deser et al.*, 2012]. As a demonstrative example, we take the 50 year period between 1951 and 2000 (JJA), which shows a nonsignificant observed trend of $-0.04^{\circ}\text{C decade}^{-1}$ over the central U.S. For this time period, the modeled signal-to-noise ratio to anthropogenic aerosols (as given by the ratio of the AA ensemble mean to the ensemble spread, normalized by the aerosol burden change) is increased by over threefold compared to the 1951-1975 period; however, the observed value remains unlikely ($p=2\%$) in the ALL ensemble.

4. Conclusions

In this study, we have exploited large ensembles of initial condition simulations performed by the CanESM2 model to test the claim that the summertime U.S. warming hole was primarily caused by increasing emissions of anthropogenic aerosols [L12b; *Yu et al.*, 2014].

Here the warming hole has been defined primarily as the marked decline in near surface air temperature ($-0.56^{\circ}\text{C decade}^{-1}$) over the southeastern U.S. between 1951 and 1975 (JJA). It was found that the modeled responses to anthropogenic aerosols ($-0.20 \pm 0.08^{\circ}\text{C decade}^{-1}$) and all forcings ($-0.08 \pm 0.06^{\circ}\text{C decade}^{-1}$) do not alone account for the magnitude of the observed cooling rate but do increase the likelihood of the warming hole occurring. Taking into account different start dates of the temperature trend, we found that the effects of anthropogenic aerosols increased the likelihood of the warming hole by up to 15% (explaining up to half of the observed trend). However, the warming hole was no more than 4% more likely due to all historical forcings for all time periods considered. This suggests a large compensation between the effects of increasing emissions of anthropogenic aerosols and greenhouse gases. Furthermore, there is a substantial spread in spatial patterns and magnitudes ($\sigma \sim 0.2^{\circ}\text{C decade}^{-1}$) in the trends across ensemble members in the forced simulations and a long control run, which is driven by unforced internal variability. Thus, we conclude from the CanESM2 results that the warming hole was not predominantly caused by increasing anthropogenic aerosol emissions or all forcings. Rather, there was a large contribution from unforced internal variability, which resulted in an unusually large rate of cooling. The validity of these results crucially depend on the model correctly simulating the magnitude of internal variability and the response to external forcings. Similar estimates are computed from the CMIP5 models, and a large forced ensemble and long control run performed with the CESM1 model, which reinforce the conclusions drawn from CanESM2.

There remain fundamental uncertainties in the magnitude and pattern of the climate response to changing aerosol emissions, particularly through aerosol-cloud interactions, and the trade-off between the effects of increased greenhouse gas and anthropogenic aerosol emissions [*Stott et al.*, 2006; *Wilcox et al.*, 2015]. Even though we have shown considerable intermodel agreement in our conclusions, it is conceivable that some flaw might be common to all the models, which would change the relative importance of forcing and internal variability. Notwithstanding these uncertainties, this work highlights the potentially large influence of internal variability on the warming hole, which, therefore, is important to consider in its future studies.

Finally, we note that although this study has focused on the role of anthropogenic aerosols in causing the warming hole, several other drivers have also been proposed, partly, due to the different time periods, regions, and seasons investigated in different studies. Besides anthropogenic aerosols, land use changes have been suggested as an external driver [*Misra et al.*, 2012], while a contender for internal variability is the interdecadal Pacific oscillation [*Meehl et al.*, 2012]. Several other studies have associated the warming hole to variations in Pacific and/or Atlantic sea surface temperatures and in the hydrological cycle over the U.S.

[Robinson *et al.*, 2002; Pan *et al.*, 2004; Kunkel *et al.*, 2006; Kumar *et al.*, 2013; Pan *et al.*, 2013; Weaver, 2013]. These other potential drivers might also be, in large measure, manifestations of internal variability (on different timescales, originating in different regions, with different mechanisms, etc.) rather than of anthropogenic forcing.

Acknowledgments

This work was funded, in part, by a grant from the U.S. National Science Foundation (NSF) to Columbia University. The large ensemble simulations used were performed by the Canadian Centre for Climate Modelling and Analysis and the National Centre for Atmospheric Research. The data used here are listed in the text and references. We thank Arlene Fiore, Nora Mascioli, Deepti Singh, and Dan Westervelt for useful discussions regarding this work.

References

- Arora, V. K., J. F. Scinocca, G. J. Boer, J. R. Christian, K. L. Denman, G. M. Flato, V. V. Kharin, W. G. Lee, and W. J. Merryfield (2011), Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, *Geophys. Res. Lett.*, *38*, L05805, doi:10.1029/2010GL046270.
- Deser, C., R. Knutti, S. Solomon, and A. S. Phillips (2012), Communication of the role of natural variability in future North American climate, *Nat. Clim. Change*, *2*(11), 775–779, doi:10.1038/nclimate1562.
- Deser, C., L. Terray, and A. S. Phillips (2016), Forced and internal components of winter air temperature trends over North America during the past 50 years: Mechanisms and implications, *J. Clim.*, *29*(6), 2237–2258, doi:10.1175/JCLI-D-15-0304.1.
- Ghan, S. J., X. Liu, R. C. Easter, R. Zaveri, P. J. Rasch, J.-H. Yoon, and B. Eaton (2012), Toward a minimal representation of aerosols in climate models: Comparative decomposition of aerosol direct, semidirect, and indirect radiative forcing, *J. Clim.*, *25*(19), 6461–6476, doi:10.1175/JCLI-D-11-00650.1.
- Hansen, J., R. Ruedy, M. Sato, and K. Lo (2010), Global surface temperature change, *Rev. Geophys.*, *48*, RG4004, doi:10.1029/2010RG000345.
- Hurrell, J. W., et al. (2013), The community earth system model: A framework for collaborative research, *Bull. Am. Meteorol. Soc.*, *94*(9), 1339–1360, doi:10.1175/BAMS-D-12-00121.1.
- Kasoar, M., A. Voulgarakis, J.-F. Lamarque, D. T. Shindell, N. Bellouin, W. J. Collins, G. Faluvegi, and K. Tsigaridis (2016), Regional and global temperature response to anthropogenic SO₂ emissions from China in three climate models, *Atmos. Chem. Phys.*, *16*(15), 9785–9804, doi:10.5194/acp-16-9785-2016.
- Kay, J. E., et al. (2015), The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability, *Bull. Am. Meteorol. Soc.*, *96*(8), 1333–1349, doi:10.1175/BAMS-D-13-00255.1.
- Kumar, S., J. Kinter, P. A. Dirmeyer, Z. Pan, and J. Adams (2013), Multidecadal climate variability and the warming hole in North America: Results from CMIP5 twentieth- and twenty-first-century climate simulations, *J. Clim.*, *26*(11), 3511–3527, doi:10.1175/JCLI-D-12-00535.1.
- Kunkel, K. E., X.-Z. Liang, J. Zhu, and Y. Lin (2006), Can CGCMs simulate the twentieth-century warming hole in the Central United States?, *J. Clim.*, *19*(17), 4137–4153, doi:10.1175/JCLI3848.1.
- Lamarque, J.-F., et al. (2010), Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application, *Atmos. Chem. Phys.*, *10*(15), 7017–7039, doi:10.5194/acp-10-7017-2010.
- Leibensperger, E. M., L. J. Mickley, D. J. Jacob, W.-T. Chen, J. H. Seinfeld, A. Nenes, P. J. Adams, D. G. Streets, N. Kumar, and D. Rind (2012a), Climatic effects of 1950–2050 changes in US anthropogenic aerosols. Part 1: Aerosol trends and radiative forcing, *Atmos. Chem. Phys.*, *12*(7), 3333–3348, doi:10.5194/acp-12-3333-2012.
- Leibensperger, E. M., L. J. Mickley, D. J. Jacob, W.-T. Chen, J. H. Seinfeld, A. Nenes, P. J. Adams, D. G. Streets, N. Kumar, and D. Rind (2012b), Climatic effects of 1950–2050 changes in US anthropogenic aerosols. Part 2: Climate response, *Atmos. Chem. Phys.*, *12*(7), 3349–3362, doi:10.5194/acp-12-3349-2012.
- Liu, X., et al. (2012), Toward a minimal representation of aerosols in climate models: Description and evaluation in the Community Atmosphere Model CAM5, *Geosci. Model Dev.*, *5*(3), 709–739, doi:10.5194/gmd-5-709-2012.
- 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. Clim.*, *25*(18), 6394–6408, doi:10.1175/JCLI-D-11-00655.1.
- Ming, Y., and V. Ramaswamy (2009), Nonlinear climate and hydrological responses to aerosol effects, *J. Clim.*, *22*(6), 1329–1339, doi:10.1175/2008JCLI2362.1.
- Misra, V., J.-P. Michael, R. Boyles, E. P. Chassignet, M. Griffin, and J. J. O'Brien (2012), Reconciling the spatial distribution of the surface temperature trends in the southeastern United States, *J. Clim.*, *25*(10), 3610–3618, doi:10.1175/JCLI-D-11-00170.1.
- Myhre, G., et al. (2013), Anthropogenic and Natural Radiative Forcing, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K., and New York.
- Pan, Z., R. W. Arritt, E. S. Takle, W. J. Gutowski, 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.
- Pan, Z., X. Liu, S. Kumar, Z. Gao, and J. Kinter (2013), Intermodel variability and mechanism attribution of central and southeastern U.S. anomalous cooling in the twentieth century as simulated by CMIP5 models, *J. Clim.*, *26*(17), 6215–6237, doi:10.1175/JCLI-D-12-00559.1.
- Robinson, W. A., R. Reudy, and J. E. Hansen (2002), General circulation model simulations of recent cooling in the east-central United States, *J. Geophys. Res.*, *107*(D24), 4748, doi:10.1029/2001JD001577.
- Shindell, D., M. Schulz, Y. Ming, T. Takemura, G. Faluvegi, and V. Ramaswamy (2010), Spatial scales of climate response to inhomogeneous radiative forcing, *J. Geophys. Res.*, *115*, D19110, doi:10.1029/2010JD014108.
- Shindell, D. T., et al. (2013), Radiative forcing in the ACCMIP historical and future climate simulations, *Atmos. Chem. Phys.*, *13*(6), 2939–2974, doi:10.5194/acp-13-2939-2013.
- Stott, P. A., J. F. B. Mitchell, M. R. Allen, T. L. Delworth, J. M. Gregory, G. A. Meehl, and B. D. Santer (2006), Observational constraints on past attributable warming and predictions of future global warming, *J. Clim.*, *19*(13), 3055–3069, doi:10.1175/JCLI3802.1.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.
- von Salzen, K., et al. (2013), The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of physical processes, *Atmos. Ocean*, *51*(1), 104–125, doi:10.1080/07055900.2012.755610.
- Weaver, S. J. (2013), Factors associated with decadal variability in great plains summertime surface temperatures, *J. Clim.*, *26*(1), 343–350, doi:10.1175/JCLI-D-11-00713.1.
- Westervelt, D. M., L. W. Horowitz, V. Naik, J.-C. Golaz, and D. L. Mauzerall (2015), Radiative forcing and climate response to projected 21st century aerosol decreases, *Atmos. Chem. Phys.*, *15*(22), 12681–12703, doi:10.5194/acp-15-12681-2015.

- Wilcox, L. J., E. J. Highwood, and N. J. Dunstone (2013), The influence of anthropogenic aerosol on multi-decadal variations of historical global climate, *Environ. Res. Lett.*, *8*(2), 024033, doi:10.1088/1748-9326/8/2/024033.
- Wilcox, L. J., E. J. Highwood, B. B. Booth, and K. S. Carslaw (2015), Quantifying sources of inter-model diversity in the cloud albedo effect, *Geophys. Res. Letts.*, *42*, 1568–1575, doi:10.1002/2015GL063301.
- Yu, S., K. Alapaty, R. Mathur, J. Pleim, Y. Zhang, C. Nolte, B. Eder, K. Foley, and T. Nagashima (2014), Attribution of the United States “warming hole”: Aerosol indirect effect and precipitable water vapor, *Sci. Rep.*, *4*, 6929, doi:10.1038/srep06929.