

# Inferring Aerosol Cooling from Hydrological Sensitivity

TIMOTHY DELSOLE AND XIAOQIN YAN

*George Mason University, Fairfax, Virginia, and Center for Ocean–Land–Atmosphere Studies, Calverton, Maryland*

MICHAEL K. TIPPETT

*Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, and  
Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz  
University, Jeddah, Saudi Arabia*

(Manuscript received 20 May 2015, in final form 31 May 2016)

## ABSTRACT

Hydrological sensitivity is the change in global-mean precipitation per degree of global-mean temperature change. This paper shows that the hydrological sensitivity of the response to anthropogenic aerosol forcing is distinct from that of the combined response to all other forcings and that this difference is sufficient to infer the associated cooling in global-mean temperature. This result is demonstrated using temperature and precipitation data generated by climate models and is robust across different climate models. Remarkably, greenhouse gas warming and aerosol cooling can be estimated in a model without using any spatial or temporal gradient information in the response, provided temperature data are augmented by precipitation data. Over the late twentieth century, the hydrological sensitivities of climate models differ significantly from that of observations. Whether this discrepancy can be attributed to observational error, which is substantial as different estimates of global-mean precipitation are not even significantly correlated with each other, or to model error is unclear. The results highlight the urgency to construct accurate estimates of global precipitation from past observations and for reducing model uncertainty in hydrological sensitivity. This paper also clarifies that previous estimates of hydrological sensitivity are limited in that standard regression methods neglect temperature–precipitation relations that occur through internal variability. An alternative method for estimating hydrological sensitivity that overcomes this limitation is presented.

## 1. Introduction

A key parameter in predicting future climate change and in explaining past climate changes is climate sensitivity. The equilibrium climate sensitivity, for example, is defined as the global-mean surface warming in response to a doubling of CO<sub>2</sub> after the system has reached a new steady state (Knutti and Hegerl 2008). Recent, authoritative estimates suggest that the equilibrium climate sensitivity is likely to be in the range 1.5°–4.5°C (Bindoff et al. 2013; Knutti and Hegerl 2008). This uncertainty range has remained virtually unchanged despite decades of research (Mitchell et al. 1990; Knutti and Hegerl 2008). Estimates of climate sensitivity based on instrumental records are limited by

significant uncertainties in observations, especially in ocean heat uptake and the magnitude of aerosol radiative forcing (Knutti et al. 2002; Hansen et al. 2005).

A significant challenge in inferring climate sensitivity from observations is to disentangle greenhouse gas warming from internal variability and from the response to other forcings, particularly anthropogenic aerosols. For instance, some aerosols directly cool Earth by enhancing scattering of solar radiation to space, whereas others, such as black carbon, absorb solar radiation and thereby warm the atmosphere, altering the general circulation. Aerosols also can serve as cloud condensation nuclei and thereby alter cloud properties. In particular, aerosols increase the number of small liquid cloud droplets, causing an increase in total droplet surface area and an increase in the albedo of liquid clouds, leading to a cooling effect (Boucher et al. 2013). Most studies find that, overall, anthropogenic aerosols have cooled the planet (Boucher et al. 2013), implying that they have masked some of the global warming that would have occurred in their absence.

---

*Corresponding author address:* Timothy DelSole, George Mason University, 112 Research Hall, Mail Stop 2B3, 4400 University Drive, Fairfax, VA 22030.  
E-mail: delsole@cola.iges.org

Stott et al. (2006) applied optimal fingerprinting techniques to estimate the cooling effects of aerosols and the warming effects of anthropogenic greenhouse gases. Both spatial and temporal aspects of the observed temperature change were necessary to constrain the relative roles of greenhouse gas warming and sulfate cooling over the twentieth century. In particular, the analysis exploited distinctive differences between the different forcings, especially the differential warming rates between the hemispheres, between land and ocean, and between mid- and low latitudes.

Recent studies have revealed another difference between greenhouse gas warming and sulfate cooling—namely, their hydrological sensitivity (Hansen et al. 1997; Allen and Ingram 2002; Lambert and Faull 2007; Andrews et al. 2009; Bala et al. 2010; O’Gorman et al. 2012). Hydrological sensitivity is defined as the change in global-mean precipitation per degree of global-mean temperature. Climate models suggest that the hydrological sensitivity for greenhouse gas forcing is less than that of sulfate aerosols. This difference has been explained in terms of differences in the response on fast time scales (Bala et al. 2010; Andrews et al. 2009, 2010). Specifically, numerical experiments reveal that precipitation initially decreases in response to an instantaneous change in CO<sub>2</sub>. The initial decrease can be explained by the fact that an increase in greenhouse gas forcing reduces radiative cooling in the troposphere, and since on monthly time scales energy is balanced above the atmospheric boundary layer, a decrease in radiative cooling requires a reduction in condensational heating from convection (Yang et al. 2003; Andrews et al. 2009; Takahashi 2009). After this initial “fast” adjustment, precipitation then increases with increasing sea surface temperatures on “slow” multiyear time scales. Thus, the fast and slow time-scale responses are in opposite directions for warming due to greenhouse gases. In contrast, an increase in sulfate aerosol forcing enhances solar reflectivity, which only modestly impacts the atmospheric energy balance on fast time scales, leaving the slow response to dominate the total hydrological sensitivity. Interestingly, the hydrological sensitivity on slow time scales tends to be independent of forcing (Andrews et al. 2010). Thus, the total hydrological sensitivity, because of the combined effect of fast and slow responses, is less for greenhouse gas warming than for sulfate cooling. Because different forcings share a common slow response, their hydrological sensitivities tend to be similar and hence difficult to distinguish.

The purpose of this paper is to demonstrate that anthropogenic aerosol cooling can be separated from the response to all other forcings using just their hydrological sensitivities. This result is demonstrated using

temperature and precipitation data generated by climate models, rather than from observations, because global precipitation observations are too uncertain (as will be shown). We further show that this result is reproducible across different climate models and demonstrate that aerosol cooling can be estimated without using any spatial or temporal gradient information in the response, provided temperature data are augmented by precipitation data. In addition, this paper clarifies that previous regression methods for estimating hydrological sensitivities do not account for temperature–precipitation ( $T$ – $P$ ) relations that occur through internal variability and are not optimized to detect specific forcings. This paper proposes an alternative method for estimating hydrological sensitivity that overcomes these limitations. Finally, this paper presents a stark contrast between hydrological sensitivities between models and observations but also presents strong arguments that these discrepancies are very likely due to observational errors. The data used in this paper are discussed in section 2. Our methodology is explained in section 3, with mathematical details relegated to an appendix. The results of our analysis on model simulations and observations are discussed in section 4. We conclude with a summary and discussion of our results.

## 2. Data

Four types of simulations are used in our analysis: historical, AA, noAA, and control. Historical runs refer to simulations that include both natural and anthropogenic forcings. The AA runs refer to simulations that include only anthropogenic aerosols. The noAA runs refer to simulations that include all forcings except anthropogenic aerosol forcing. Control runs refer to simulations in which natural and anthropogenic forcing does not vary from year to year. The historical, AA, and noAA runs were analyzed over the period 1900–2004. We only consider models if they have at least 500 years of control simulations. We use only the last 500 years of each control run.

Only two models from the CMIP5 archive have all four runs—namely, CSIRO Mk3.6.0 and IPSL-CM5A-LR, denoted CSIRO and IPSL, respectively. We focus on results from these two models in the main text. We also consider four additional models that have AA runs (but not noAA runs) and at least 500 years of control simulations. These models are GFDL-ESM2M, NorESM1-M, GFDL CM3, and CCSM4. Results from these models are discussed in the appendix and confirm results from CSIRO and IPSL. Table 1 also indicates the models that include the first indirect effect of aerosols [as documented in Wilcox et al. (2013)], which is the

TABLE 1. Summary of CMIP5 model attributes used in this study. The AA and noAA response parameters for the CMIP5 models are based on global-mean temperature and precipitation. Sensitivities are expressed as ensemble mean plus or minus interensemble standard deviation (zero standard deviation implies only one member was available). Models that include the first indirect effect of aerosols are indicated by “Y” in the “first indirect” column.

Model	First indirect	Pattern correlation	Condition number	$\gamma_{AA}$ (% K <sup>-1</sup> )	$\gamma_{noAA}$ (% K <sup>-1</sup> )
CSIRO Mk3.6.0	Y	0.85	3.5	4.0 ± 0.38	1.8 ± 0.057
GFDL-ESM2M	N	0.36	1.5	5.4 ± 0.00	1.4 ± 0.000
GFDL CM3	Y	0.46	1.6	3.0 ± 0.15	0.95 ± 0.085
IPSL-CM5A-LR	Y	0.94	5.7	3.8 ± 0.00	2.6 ± 0.043
CCSM4	N	0.68	2.3	3.4 ± 0.41	1.9 ± 0.081
NorESM1-M	Y	0.77	2.8	3.5 ± 0.00	1.8 ± 0.000

tendency for aerosols to increase droplet concentration and decrease droplet size, making clouds more reflective.

Observational estimates of global precipitation were obtained from the Global Precipitation Climatology Project (GPCP), version 2 (<http://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html>) and from the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997; <http://www.esrl.noaa.gov/psd/data/gridded/data.cmap.html>).

Observational estimates of global temperature were obtained from the Hadley Centre/Climatic Research Unit, version 4 (HadCRUT4; downloaded from <http://www.cru.uea.ac.uk/cru/data/temperature/>). Annual means were computed when at least 10 months per calendar year were available. A grid point is included only when annual means are available for every year in the period 1979–2014. The resulting map of available grid points covers 75% of the surface area of the globe and was used to analyze both observations and model output of temperature.

### 3. Methodology

#### a. Estimating hydrological sensitivity

Hydrological sensitivity is usually estimated by fitting the following equation:

$$P = \gamma' T + \varepsilon, \quad (1)$$

where  $T$  and  $P$  are global-mean temperature and precipitation, respectively;  $\varepsilon$  is a regression error; and  $\gamma'$  is a slope parameter. The above equation, however, is not derived from a physical law. A more physically realistic model would allow temperature and precipitation to covary owing to both internal processes and external forcing. A widely used paradigm in climate studies is to model climate variability as a linear combination of responses to external forcings plus random internal variability. To the extent that the response can be

characterized by a single time series  $a(t)$ , called the response time series, temperature and precipitation would be modeled as follows:

$$\begin{bmatrix} T(t) \\ P(t) \end{bmatrix} = \begin{pmatrix} 1 \\ \gamma \end{pmatrix} a(t) + \begin{bmatrix} u_T(t) \\ u_P(t) \end{bmatrix}, \quad (2)$$

where  $\gamma$  is the hydrological sensitivity of the response to a particular forcing, and  $u_T$  and  $u_P$  are (correlated) random variables representing internal variability. The vector  $(1, \gamma)^T$  is called the response vector. Ideally, the model parameters would be constrained by physical laws, but no universally acceptable approach to these constraints has emerged. Accordingly, the parameters will be estimated from climate model simulations.

Importantly, the hydrological sensitivity defined in (2) does not coincide with the slope parameter in (1). To show this, note that the least squares estimate of  $\gamma'$  is the following:

$$\hat{\gamma}' = \frac{\text{cov}(P, T)}{\text{var}(T)} = \frac{\gamma \text{var}(a)}{\text{var}(a) + \text{var}(u_T)}, \quad (3)$$

assuming the true model is (2). This equation shows that  $\gamma' \approx \gamma$  only when  $\text{var}(a) \gg \text{var}(u_T)$ . This latter regime is appropriate when using simulations with “strong” changes in forcing to estimate sensitivity. However, the two parameters differ for weak forcing [i.e.,  $\text{var}(a) \leq \text{var}(u_T)$ ].

We estimate hydrological sensitivity (i.e.,  $\gamma$ ) and response time series  $a(t)$  using the method of maximum likelihood, which, as we show, is equivalent to finding the response vector that is maximally detectable (see appendix). The resulting estimates tend to be more sensitive to sampling error than ordinary least squares (OLS), so it is not recommended as a more robust substitute for OLS. Rather, its main virtue is that it converges to the correct sensitivity in the limit of infinite samples.

One may question whether hydrological sensitivity can be treated as a constant. Indeed, noAA represents a

mix of forcings (e.g., greenhouse gases, ozone, and volcanoes), so assuming a constant sensitivity effectively assumes that the temperature–precipitation response to these different forcings follows a constant ratio. If the constant-ratio assumption is not appropriate, then the regression model should be expanded to include additional sensitivities ( $\gamma_1, \gamma_2, \dots$ ) and corresponding time series [ $a_1(t), a_2(t), \dots$ ]. An important aspect of the maximum likelihood method discussed in the [appendix](#) is that it leads to an eigenvalue problem with multiple solutions. These solutions have the following property: if the eigenvectors are ordered by decreasing eigenvalue, then the first eigenvector maximizes detectability, the second eigenvector maximizes detectability subject to being uncorrelated with the first, and so on. The statistical significance of detectability can be tested for each component to determine the number of components that should be included in the regression model. For all single-forcing runs, the leading eigenvector is the only component that has statistically significant detectability. This result means that the second component should not be included in the regression model because it cannot be distinguished from internal variability. Physically, this result implies that responses to different noAA forcings either are weak or have nearly the same hydrological sensitivity (i.e., they are nearly collinear).

#### b. Estimating the response time series

Our goal is to infer the role of different forcings based on time series  $T(t)$  and  $P(t)$  that contain a mixture of different forcings (plus internal variability). Assuming the forced responses can be modeled as (2), the appropriate model for superposed forcings is the following:

$$\begin{bmatrix} T(t) \\ P(t) \end{bmatrix} = \begin{pmatrix} 1 \\ \gamma_{AA} \end{pmatrix} a_{AA}(t) + \begin{pmatrix} 1 \\ \gamma_{noAA} \end{pmatrix} a_{noAA}(t) + \begin{bmatrix} u_T(t) \\ u_P(t) \end{bmatrix}. \quad (4)$$

In principle, the response time series  $a_{noAA}(t)$  could be subdivided further into other forcings such as those due to greenhouse gases, volcanic aerosols, solar insolation, land-use change, and ozone. Such partitioning of  $a_{noAA}$  leads to more quantities to estimate from data, leading to larger uncertainties and overfitting. On the other hand, some of the forcings grouped under noAA may have hydrological sensitivities similar to those of AA (e.g., solar and ozone; [Andrews et al. 2010](#)). Nevertheless, we expect the response of anthropogenic greenhouse gases to dominate the noAA response over the twentieth century. Consistent with this, we find that the noAA response can be characterized by a single, detectable time series, as discussed in [section 3a](#),

suggesting that the proposed model is reasonable. Alternative groupings are certainly worth exploring, but we pursue the AA–noAA grouping because we are interested specifically in estimating the response to anthropogenic aerosols.

A standard approach to estimating the response to external forcing using observations is to introduce scaling factors for each forcing and then choose the scaling factors to best fit observations using generalized least squares methods. This approach, called optimal fingerprinting, uses the full spatiotemporal structure of the response to inform the estimation. Instead, we directly estimate the time series  $a_{AA}(t)$  and  $a_{noAA}(t)$  (which subsume the associated scaling factors). This step frees us from assuming that the model correctly simulates the temporal evolution of the response to individual forcings. Unfortunately, the uncertainties become unacceptably large when the time series are estimated at each year independently of others. Consequently, we assume that the response evolves slowly in time, which allows the estimation method to pool data across years to reduce uncertainty. Specifically, the temporal variation over 1900–2004 is represented by a third-order polynomial in time. This representation is most questionable during volcanic eruptions, in which the response is dominated by a decrease in global average temperature and precipitation over a few years ([Iles et al. 2013](#)). Since the dominant response to volcanic eruptions spans only a few years, smoothing over periods of major eruptions leads to only minor errors on multidecadal time scales. Mathematical details of our methodology are discussed in the [appendix](#).

## 4. Results

The estimated sensitivities for AA and noAA forcing are shown in [Fig. 1](#) (and listed in [Table 1](#)). These estimates generally are larger than those reported previously ([Andrews et al. 2010](#)), owing partly to differences in estimation method [e.g., (3) implies our estimate will be larger than the least squares estimate; see also the [appendix](#), especially [Fig. A1](#)]. As anticipated in the introduction, the sensitivity to noAA forcing is consistently less than that of AA in the same model. Moreover, for each model, the sensitivity of AA forcing is clearly separated from that of noAA forcing.

The estimated sensitivities are distinct in each model but nevertheless may be so close as to cause multicollinearity problems. As discussed in the [appendix](#), the angle between the AA and noAA response vectors emerges naturally as a measure of collinearity from generalized least squares. This measure is effectively the angle between the vectors in “whitened space” and can

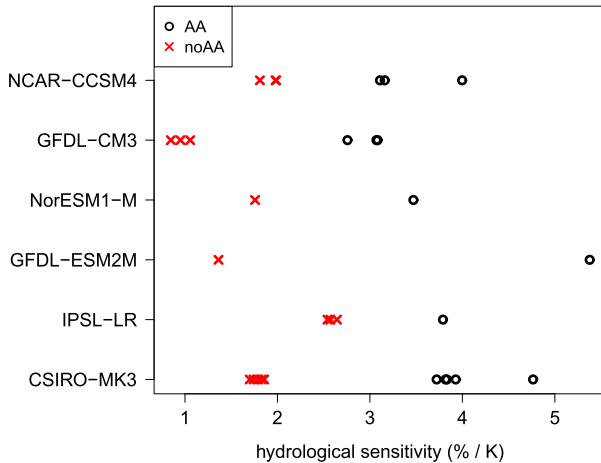


FIG. 1. Hydrological sensitivity associated with anthropogenic aerosol forcing (AA) and with all forcings except anthropogenic aerosols (noAA), estimated by ML methods from CMIP5 simulations. Each estimate is derived from a single ensemble member of a single-forcing simulation for 1900–2004 and 500 years of a control run from each model. Different points for the same model indicate estimates from different ensemble members.

be interpreted as a kind of “pattern correlation” between the two response vectors. These pattern correlations, as well as their corresponding condition numbers, are listed in Table 1. The largest condition number is 5.7, which is less than 10 and hence implies that multicollinearity is not a serious problem.

We test our methodology by applying it to historical simulations containing both anthropogenic and natural forcings. The response time series estimated from the historical simulations are shown in Fig. 2 as green and red curves, while those derived from the AA and noAA runs are shown in blue and black curves. The latter time series were estimated for each year separately, with no temporal smoothing, which can be done accurately because only one type of forcing exists (i.e., no separation of responses is required). Different curves of the same color show results for different ensemble members. The response time series estimated from the historical runs are in excellent agreement with each other and with response time series estimated from the AA and noAA runs. This result demonstrates the remarkable fact that the hydrological sensitivity parameter is sufficient to recover the correct (multidecadal) time history of the response to both AA and noAA forcings simultaneously. Moreover, time series estimated from different ensemble members of the same forcing are close to each other relative to the secular changes, indicating that the estimates are not sensitive to different realizations of internal variability.

Performing the same fingerprinting procedure on historical simulations that lack a corresponding noAA

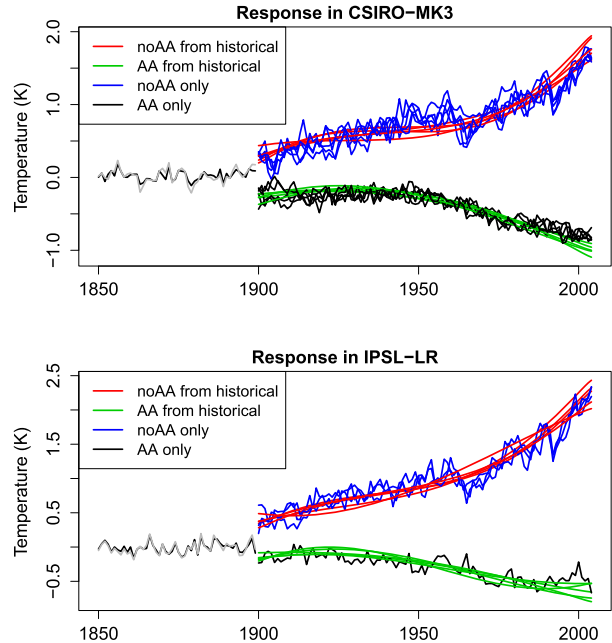


FIG. 2. Changes in global-average temperature due to AA and noAA, estimated from historical simulations containing both anthropogenic and natural forcings (red and green curves) and from individual-forcing runs containing only those forcings (blue and black curves) in (top) CSIRO and (bottom) IPSL. The temperature changes are estimated using the technique described in the appendix, which uses the joint covariability between temperature and precipitation. Estimates from AA and noAA simulations are obtained for each year separately, while those from the historical simulations are obtained from a polynomial fit. Different curves of the same color show estimates based on different ensemble members. The black and gray curves before 1900 show the projection of the AA and noAA response vectors in the preindustrial control run, respectively.

simulation gives similar results (see appendix). Also, when the response time series for AA and noAA are regressed out from those respective runs, the residual contains no statistically significant detectable component in most cases, indicating that the two hydrological sensitivities are sufficient to capture practically the whole climate change signal (not shown). We also considered the “imperfect model” case in which response vectors from one model are used to detect AA and noAA responses in another model. For these two models, temperature changes derived from the perfect and imperfect cases are close to each other (not shown). This robustness is not surprising because the only information used in fingerprinting are the hydrological sensitivities, which are nearly the same for the two models (see Fig. 1). (Technically, fingerprinting also depends on the covariances of internal variability, but swapping covariances from different models produces nearly identical results.) Thus, differences between



inferred time series can be attributed solely to differences in hydrological sensitivities. Of course, greater disagreement can be found by choosing models with greater differences in hydrological sensitivities. Also, this approach does not tell the whole story because models differ not only in their response to forcing but also in the nature and magnitude of the forcing itself. Nevertheless, these issues do not affect our main conclusion, which is that hydrological sensitivities for AA and noAA forcings are sufficiently distinct in each model that they can be used to infer time series for anthropogenic cooling and greenhouse gas warming in that model.

Having established that the methodology works in climate models, we now consider observations. Unfortunately, global precipitation data are limited to the satellite era (i.e., since the 1970s). Repeating the above analysis on climate models but for the shorter period 1979–2004 (not shown) yields time series consistent with those shown above but with larger uncertainties because of the smaller sample size. It should be recognized that global anthropogenic aerosol forcing varies weakly over the period 1979–2004 (Myhre et al. 2013, their Fig. 8.18), so the response to aerosols over this period is likely to have small trend (as can be seen in Fig. 2).

To visualize the situation, we construct a scatterplot of all the data in temperature–precipitation space, as in Fig. 3. The AA and noAA simulations (blue and red circles, respectively) are seen to be displaced from each other. However, this displacement is not relevant for distinguishing the two forcings because the time series will be centered prior to analysis (to remove model bias). Instead, the relevant feature is the slope of the variability, which corresponds to hydrological sensitivity. For reference, the hydrological sensitivities of the AA and noAA simulations determined by maximum likelihood are indicated in the lower-right corner of each panel. These values differ slightly from those in Table 1 because the sensitivities were computed from model data whose spatial grid was masked in such a way to mimic missing data in the temperature observations.

The historical and control runs also are shown in Fig. 3. The historical runs (gold circles) trace a path that cannot be explained by internal variability (gray circles). Geometrically, optimal fingerprinting represents each year of the historical run as a linear combination of the response vectors shown in the lower-right corner. The observations are shown as green dots and green crosses (after aligning their mean with that of the historical run over the same period). The figure reveals that observations vary over a region in  $T$ – $P$  space outside that of any of the climate simulations. To highlight differences between models and observations further, we show in

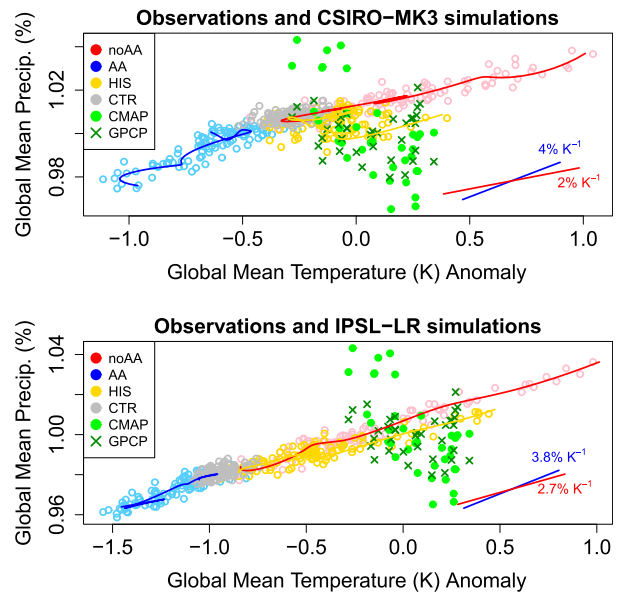


FIG. 3. Scatterplot of annual-mean, global-mean temperature and precipitation in simulations from (top) CSIRO and (bottom) IPSL and observations (green symbols). Simulations are shown with AA (blue), with noAA (red), with both natural and anthropogenic forcings (gold symbols), and with no natural and anthropogenic forcing (gray symbols). Also shown are observations based on GPCP (green filled circles) and CMAP (green crosses). Temperature in the historical run and observations are centered to have zero mean over the common period 1979–2004, while other simulations are shown relative to the mean of the historical run over this period. Precipitation in the historical run and observations are normalized by their respective means over the period 1979–2004, while other simulations are normalized by the mean of the historical run over the same period. The solid curves show the projection of a low-pass time series of the respective runs. The two lines in the lower-right corner indicate the hydrological sensitivities estimated from the AA and noAA simulations.

Fig. 4 (see also Fig. A3) the sensitivity estimated from observations against a histogram of hydrological sensitivities computed for every 36-yr period in historical simulations. The sensitivity estimated from observations are very unlikely relative to the sensitivities found in the historical simulations.

Note that the two observational datasets differ even between themselves. These differences are due entirely to differences in precipitation datasets [e.g., global-mean, annual-mean temperature shows little sensitivity to estimation method; see Fig. 2.14 in Hartmann et al. (2013)]. Indeed, the correlation between the annual-mean, global-mean precipitation time series for CMAP and GPCP over 1979–2014 is 0.24, which is not statistically significant at the 5% level. Previous attempts to estimate hydrological sensitivity from observations have yielded negative values (Arkin et al. 2010), values around  $2\% \text{ K}^{-1}$  (Adler et al. 2008), and values around  $7\% \text{ K}^{-1}$  (Wentz et al. 2007; Liepert and Previdi 2009),

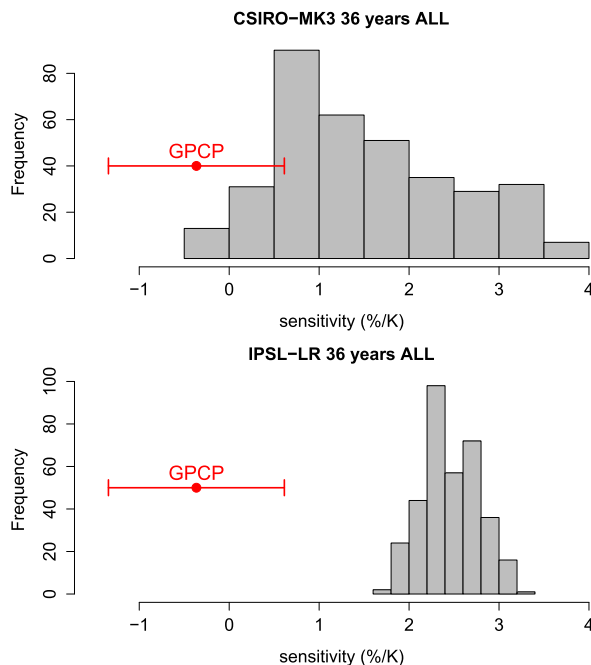


FIG. 4. Histogram of the hydrological sensitivity computed for every 36-yr period in historical simulations over the period 1900–2004 (gray bars) in four CMIP5 models. The sensitivity computed from GPCP in the period 1979–2014 is indicated by the red dot; the error bar indicates one standard error. All sensitivities were computed from OLS.

clearly indicating sensitivity to data source, estimation method, and time period. Yin et al. (2004) document additional differences between GPCP and CMAP precipitation datasets and show that certain trends are clear artifacts of changes in satellite data input and sampling of atoll data.

## 5. Summary and discussion

This paper shows that, in climate models, differences in hydrological sensitivity can be used to separate the cooling effects of aerosols from the effects of other forcings. The proposed methodology for doing this uses both temperature and precipitation data and requires estimates of the hydrological sensitivity associated with different forcings. We introduce a new estimate for hydrological sensitivity that accounts for temperature–precipitation covariability due to internal variability and converges to the “true sensitivity” [the sensitivity in the model in (2)] in the limit of large sample sizes, in contrast to the linear regression estimate. This method is formally equivalent to finding the response vector that maximizes detectability (Jia and DelSole 2012). Finally, the fingerprinting method is generalized to infer the multidecadal evolution due to different forcings. Applying this methodology to historical simulations

containing both anthropogenic and natural forcings yields response time series that closely match those obtained from individual forcing runs. This result demonstrates that greenhouse gas warming and aerosol cooling can be estimated without using any spatial or temporal gradient information in the response, provided temperature data are augmented by precipitation data. This result also highlights the fact that hydrological sensitivity is important not only for predicting future climates but also has implications for attributing past climate changes to man.

This study also reveals striking inconsistencies between models and observations. Specifically, over the late twentieth century, climate models predict a robust positive relation between global-mean, annual-mean temperature and precipitation that differs significantly from that of observations. Whether this discrepancy can be attributed to observational error, which is substantial as different estimates of global-mean precipitation are not even significantly correlated with each other, or to model error is not clear. Because global-mean temperature and precipitation are fundamental indices of climate change, this clear discrepancy between models and observations seems serious. Previous attempts to estimate hydrological sensitivity from global observational datasets yield results that are sensitive to data source, estimation method, and time period. Moreover, precipitation estimates can differ by as much as 20% in the tropics and as much as 50% in midlatitudes (Adler et al. 2012). Such differences are not surprising given well-known sensitivity of precipitation estimates to changes in satellite and gauge data input. Precipitation estimates over land are considered to be more accurate owing to the availability of rain gauge measurements for calibration purposes. Using land-average precipitation in our model-only analysis does not yield accurate estimates of aerosol-forced signals (not shown). A more comprehensive analysis that includes spatial variations in land precipitation will be discussed in future work.

*Acknowledgments.* This work was sponsored by National Science Foundation Grant ATM1338427 (TD and XY), National Aeronautics and Space Administration Grant NNX14AM19G (TD and XY), National Oceanic and Atmospheric Administration Grant NA14OAR4310160 (TD and XY), Department of Energy Grant ER65095 (TD and XY), and Office of Naval Research Grant N00014-16-1-2073 (MKT). We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software

infrastructure in partnership with the Global Organization for Earth System Science Portals.

## APPENDIX

### Calculation Details and Additional Information

#### a. Maximum likelihood estimates of the signal parameters

The regression model in (2) can be written in the following form:

$$[T(t) \ P(t)] = a(t)(1 \ \gamma) + [u_T(t) \ u_P(t)], \quad (\text{A1})$$

which, when expanded in the time dimension with  $N$  discrete steps along rows, can be written as

$$\begin{matrix} \mathbf{y} \\ N \times 2 \end{matrix} = \begin{matrix} \mathbf{a} \\ N \times 1 \end{matrix} \begin{matrix} \mathbf{p}^T \\ 1 \times 2 \end{matrix} + \begin{matrix} \mathbf{u} \\ N \times 2 \end{matrix}, \quad (\text{A2})$$

where  $\mathbf{u}$  is white noise in which each time step is drawn from a Gaussian distribution with zero mean and  $2 \times 2$  covariance matrix  $\Sigma_U$ ,  $\mathbf{y}$  is the temperature–precipitation time series from a single ensemble member of an AA or noAA run, and the dimension of each matrix is indicated below that matrix. The vector  $\mathbf{a}$  will be called the “response time series” and the vector  $\mathbf{p}$  will be called the “response vector.” One might question whether the (temporal) white noise assumption for internal variability is appropriate. We have found that the above model (including the white noise assumption) gives reasonable estimates of uncertainty, in the sense that the results derived from separate ensemble members frequently lie (at the expected rate) within the confidence interval derived from any one of the members. This consistency suggests that the above model is reasonable.

The model in (A1) is known as an errors-in-variables model (Fuller 1980). There exist a variety of methods for estimating parameters of such models. Here, we estimate  $\mathbf{a}$  and  $\mathbf{p}$  by maximum likelihood (ML) methods assuming  $\Sigma_U$  is known, as discussed below. One reason for choosing this method is that it clarifies the connection between this estimation problem and the problem of finding response vectors that are maximally detectable, as proposed by Jia and DelSole (2012).

Technically, the response vector has one unknown element  $\gamma$ :

$$\mathbf{p} = \begin{pmatrix} 1 \\ \gamma \end{pmatrix}. \quad (\text{A3})$$

It is convenient to allow the entire vector  $\mathbf{p}$  to be unknown and then determine sensitivity by taking the ratio

of the appropriate elements of  $\mathbf{p}$ . The likelihood function for the above model is

$$L = \frac{1}{2\pi|\Sigma_U|^{1/2}} \exp\{-\text{tr}[(\mathbf{y} - \mathbf{a}\mathbf{p}^T)\Sigma_U^{-1}(\mathbf{y} - \mathbf{a}\mathbf{p}^T)^T/2]\}, \quad (\text{A4})$$

where  $\text{tr}[\cdot]$  denotes the trace of a matrix. Following standard procedure, we differentiate  $-2\log L$  with respect to  $\mathbf{a}$  and  $\mathbf{p}$ :

$$\frac{\partial(-2\log L)}{\partial \mathbf{a}} = (-2)(\mathbf{y}\Sigma_U^{-1}\mathbf{p} - \mathbf{a}\mathbf{p}^T\Sigma_U^{-1}\mathbf{p}) \quad \text{and} \quad (\text{A5})$$

$$\frac{\partial(-2\log L)}{\partial \mathbf{p}} = (-2)\Sigma_U^{-1}(\mathbf{y}^T\mathbf{a} - \mathbf{p}\mathbf{a}^T\mathbf{a}). \quad (\text{A6})$$

Setting each of these equations to zero and solving gives

$$\mathbf{a} = \mathbf{y}\Sigma_U^{-1}\mathbf{p}(\mathbf{p}^T\Sigma_U^{-1}\mathbf{p})^{-1} \quad \text{and} \quad (\text{A7})$$

$$\mathbf{y}^T\mathbf{a} = \mathbf{p}(\mathbf{a}^T\mathbf{a}). \quad (\text{A8})$$

Substituting (A7) into (A8) gives

$$\begin{aligned} \mathbf{y}^T\mathbf{y}\Sigma_U^{-1}\mathbf{p}(\mathbf{p}^T\Sigma_U^{-1}\mathbf{p})^{-1} \\ = \mathbf{p}(\mathbf{p}^T\Sigma_U^{-1}\mathbf{p})^{-1}\mathbf{p}^T\Sigma_U^{-1}\mathbf{y}^T\mathbf{y}\Sigma_U^{-1}\mathbf{p}(\mathbf{p}^T\Sigma_U^{-1}\mathbf{p})^{-1}, \end{aligned} \quad (\text{A9})$$

$$\mathbf{y}^T\mathbf{y}\Sigma_U^{-1}\mathbf{p} = \mathbf{p}(\mathbf{p}^T\Sigma_U^{-1}\mathbf{p})^{-1}\mathbf{p}^T\Sigma_U^{-1}\mathbf{y}^T\mathbf{y}\Sigma_U^{-1}\mathbf{p}, \quad \text{and} \quad (\text{A10})$$

$$\mathbf{y}^T\mathbf{y}\Sigma_U^{-1}\mathbf{p} = \lambda\mathbf{p}, \quad (\text{A11})$$

where

$$\lambda = \frac{\mathbf{p}^T\Sigma_U^{-1}\mathbf{y}^T\mathbf{y}\Sigma_U^{-1}\mathbf{p}}{\mathbf{p}^T\Sigma_U^{-1}\mathbf{p}}. \quad (\text{A12})$$

An eigenvalue problem is defined in (A11). The eigenvalue  $\lambda$  is related to the likelihood function, as can be seen by substituting (A7) into the likelihood function:

$$L = \frac{1}{2\pi|\Sigma_U|^{1/2}} \exp\{(\lambda - \text{tr}[\mathbf{y}\Sigma_U^{-1}\mathbf{y}^T])/2\}. \quad (\text{A13})$$

Since we want to maximize the likelihood function, we choose the largest eigenvalue  $\lambda$ . The eigenvector is unique up to a multiplicative constant. This non-uniqueness is a consequence of the fact that the regression model depends only on the product  $\mathbf{a}\mathbf{p}^T$ , so  $\mathbf{p}$  can be multiplied by any nonzero factor as long as  $\mathbf{a}$  also is divided by the same factor. To specify a unique solution, it is natural to normalize the eigenvector such that

$$\mathbf{p}^T\Sigma_U^{-1}\mathbf{p} = 1. \quad (\text{A14})$$



Substituting the resulting eigenvector solution into (A7) gives the time series  $\mathbf{a}$ . The hydrological sensitivity  $\gamma$  is determined by taking the ratio of the second element of  $\mathbf{p}$  to the first.

We now show that the maximum likelihood estimate of the response vector  $\mathbf{p}$  is identical to the response vector that maximizes detectability (Jia and DelSole 2012). This can be seen by defining the projection vector

$$\mathbf{q} = \Sigma_U^{-1} \mathbf{p}, \tag{A15}$$

in which case the eigenvalue problem in (A11) can be written as

$$\Sigma_Y \mathbf{q} = \frac{\lambda}{N} \Sigma_U \mathbf{q}, \tag{A16}$$

where

$$\Sigma_Y = \frac{1}{N} \mathbf{y}^T \mathbf{y}. \tag{A17}$$

The eigenvalue problem in (A16) is identical to the eigenvalue problem that is solved to obtain the most detectable component (Jia and DelSole 2012). This equivalence allows us to invoke properties of the solution as documented in Jia and DelSole (2012). In particular, if the eigenvectors are ordered by decreasing eigenvalue, then the first eigenvector maximizes detectability, the second maximizes detectability subject to being uncorrelated with the first, and so on. In addition, the statistical significance of individual components can be tested, which in turn defines the number of components that should be included in the regression model.

In practice, the covariance matrix  $\Sigma_U$  is estimated from preindustrial control simulations. A comparison between the OLS estimate of  $\gamma$  and the maximum likelihood estimate of  $\gamma$ , derived from the AA and noAA runs, is shown in Fig. A1. Figure A1 reveals a tendency for ML estimates to exceed the OLS estimates, a fact that follows from (3).

*b. Model for combined forcings*

The historical run is modeled as follows:

$$\mathbf{y} = \mathbf{a}_{AA} \mathbf{p}_{AA}^T + \mathbf{a}_{noAA} \mathbf{p}_{noAA}^T + \mathbf{u}, \tag{A18}$$

where  $\mathbf{p}_{AA}$  and  $\mathbf{p}_{noAA}$  are the response vectors derived from the AA and noAA runs, respectively, and  $\mathbf{a}_{AA}$  and  $\mathbf{a}_{noAA}$  are time series to be estimated. All time series are centered, so that the above regression model does not include a climatological mean term. We further assume that the time series evolve slowly, or more precisely, can be expressed as a cubic polynomial of time. A low-order

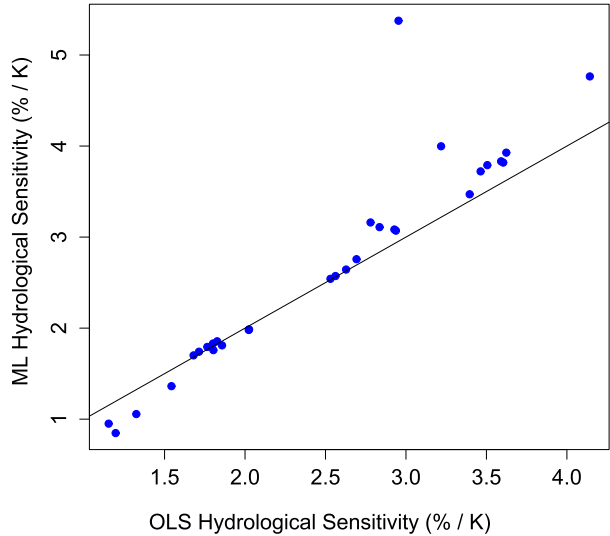


FIG. A1. Comparison of the OLS and ML estimates of hydrological sensitivity for AA and noAA forcing computed from simulations from the CMIP5 archive. The solid line shows  $y = x$ .

polynomial is desirable since the uncertainties become large for high orders. These assumptions lead to the time series models:

$$\mathbf{a}_{AA} = \mathbf{L} \mathbf{c}_{AA} \quad \text{and} \quad \mathbf{a}_{noAA} = \mathbf{L} \mathbf{c}_{noAA}, \tag{A19}$$

where the three columns of  $\mathbf{L}$  are linear, squared, and cubic functions of time, each with zero mean, and  $\mathbf{c}_{AA}$  and  $\mathbf{c}_{noAA}$  are vectors, of dimension three, containing the coefficients for the polynomial. The coefficients are estimated by the method of maximum likelihood, which is equivalent to minimizing the objective function:

$$\varepsilon = \text{tr}[(\mathbf{y} - \mathbf{L} \mathbf{C} \mathbf{P}^T) \Sigma_U^{-1} (\mathbf{y} - \mathbf{L} \mathbf{C} \mathbf{P}^T)^T], \tag{A20}$$

where

$$\mathbf{C} = (\mathbf{c}_{AA} \quad \mathbf{c}_{noAA}) \quad \text{and} \quad \mathbf{P} = (\mathbf{p}_{AA} \quad \mathbf{p}_{noAA}). \tag{A21}$$

The solution is

$$\mathbf{C} = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{y} \Sigma_U^{-1} \mathbf{P} (\mathbf{P}^T \Sigma_U^{-1} \mathbf{P})^{-1}. \tag{A22}$$

As a result of the normalization in (A14), the matrix  $\mathbf{P}^T \Sigma_U^{-1} \mathbf{P}$  has ones along the diagonal. Thus, the off-diagonal element corresponds to the cosine of the angle between the response vectors in “whitened” space:

$$\cos \theta = \frac{\mathbf{p}_{AA}^T \Sigma_U^{-1} \mathbf{p}_{noAA}}{\sqrt{\mathbf{p}_{AA}^T \Sigma_U^{-1} \mathbf{p}_{AA}} \sqrt{\mathbf{p}_{noAA}^T \Sigma_U^{-1} \mathbf{p}_{noAA}}}. \tag{A23}$$

In essence, the off-diagonal elements can be interpreted as the “pattern correlation” between the

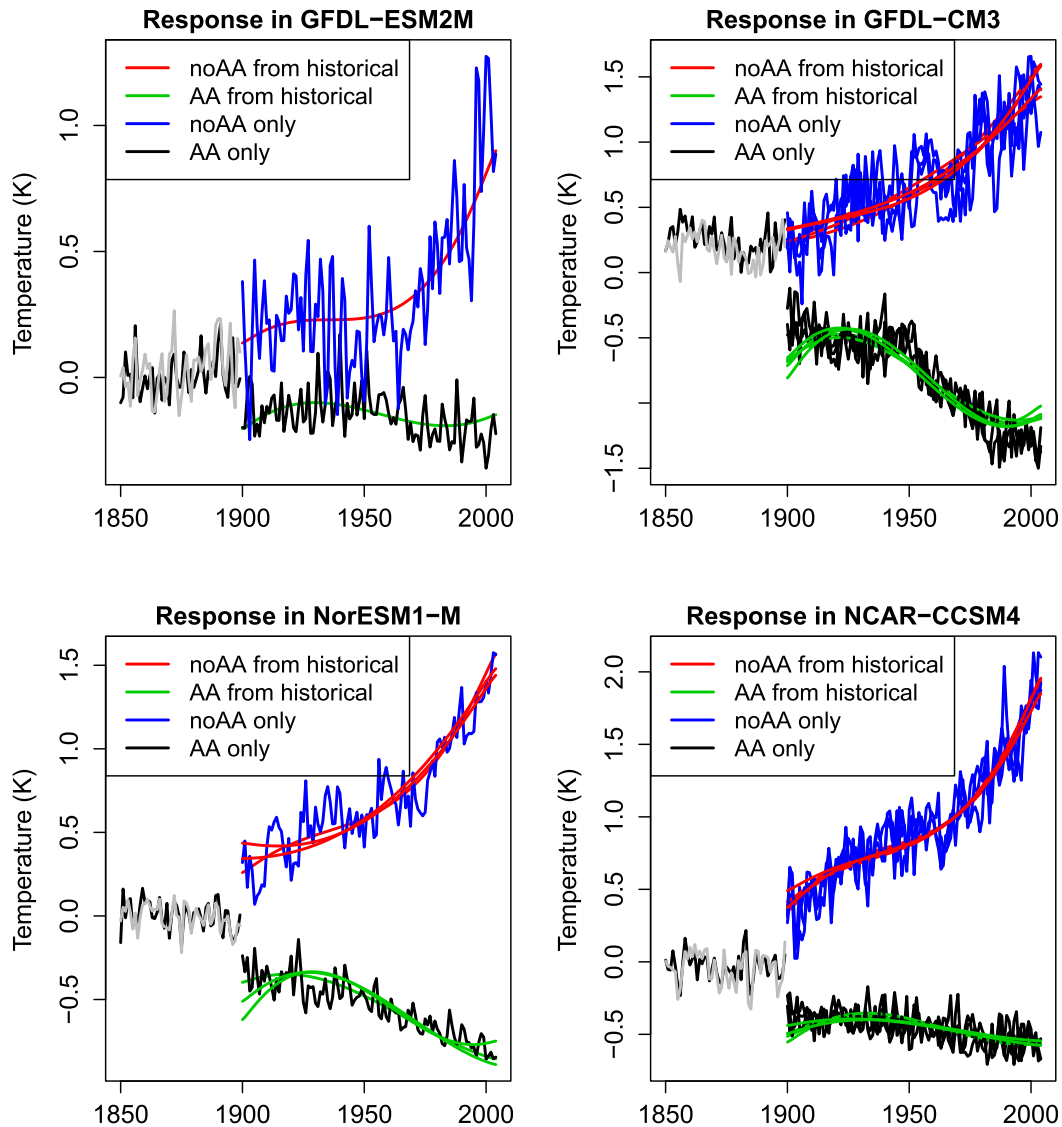


FIG. A2. Results of applying optimal fingerprinting to historical simulations of four different models, but with the noAA parameters determined from the residual between historical and AA runs. The noAA response time series is shown as the blue curve, and the AA response time series (determined from the AA run) is shown as the black curve. The green and red curves show time series of AA and noAA response time series estimated from each ensemble member of the historical simulations. Different curves of the same color show results from different ensemble members. The black and gray curves before the year 1900 show estimates of the amplitude of the AA and noAA signal vector, respectively, in the control run. Each time series is obtained as anomalies with respect to the mean of the full period but is offset by a constant for display purposes.

response vectors for AA and noAA forcing. Numerically, the off-diagonal elements are 0.87 and 0.94 for CSIRO and IPSL, respectively. The corresponding condition numbers (i.e., the square root of the ratio of the largest to the smallest eigenvalues) are obtained from the equation

$$\kappa = \sqrt{\frac{1 + \cos\theta}{1 - \cos\theta}}. \quad (\text{A24})$$

The pattern correlation and condition numbers are listed in Table 1.

### c. Analysis of models without noAA runs

In the CMIP5 archive, at least four models have AA runs without a corresponding noAA run. In such cases, the noAA signal can be estimated by subtracting one ensemble member of the AA run from one ensemble member of the historical run, and then determining the

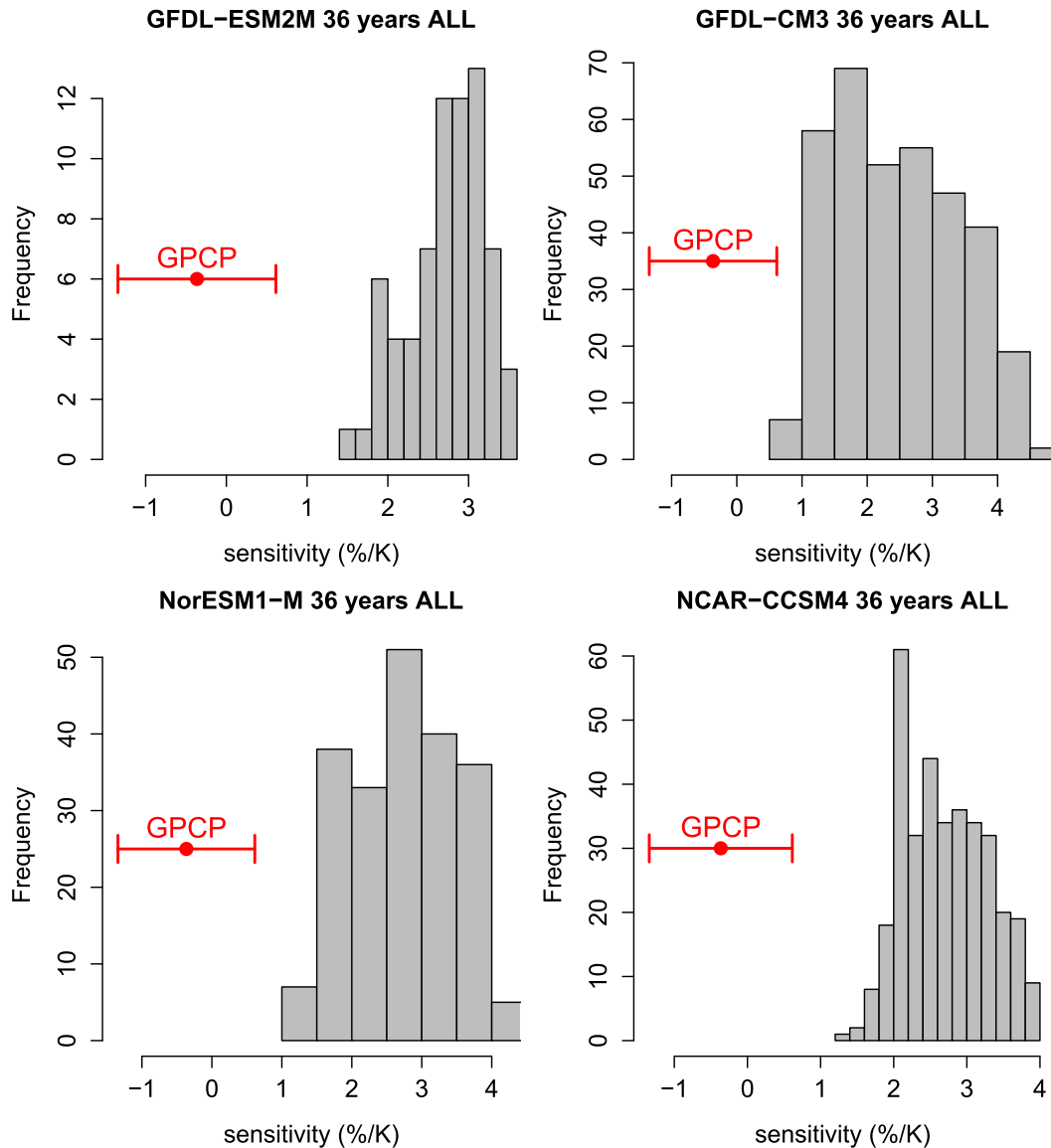


FIG. A3. Histogram of the hydrological sensitivity computed for every 36-yr period in historical simulations over the period 1900–2004 (gray bars) in four CMIP5 models. The sensitivity computed from GPCP in the period 1979–2014 is indicated by the red dot; the error bar indicates one standard error. All sensitivities were computed from OLS.

maximum likelihood estimates of the signal parameters of the noAA signal from (A7) and (A11) (where  $\mathbf{y}$  is the residual). To avoid using the same data to derive the hydrological sensitivity and to test fingerprinting analysis, only one ensemble member from each simulation is used to compute the residual, leaving other ensemble members of the historical run to be used as an independent check on the accuracy of isolating the AA signal. The result of this analysis is shown in Fig. A2. It is seen that response time series estimated from the historical run are in excellent agreement with each other

and with response time series estimated from the individual-forcing runs. Only one ensemble member of the AA and historical run is available for GFDL-ESM2M and NorESM1-M, so independent checks are not possible for these models. Nevertheless, in all cases for which independent ensemble members exist, the response time series are estimated very accurately relative to the standard error of internal variability. It is worth noting that models that include the indirect aerosol effect exhibit a much stronger response to AA forcing than those that do not (e.g., the AA temperature

changes in CCSM4 and GFDL-ESM2M seen in Fig. A2 are much weaker than the AA temperature changes in other models).

For completeness, we show in Fig. A3 the sensitivity estimated from observations against a histogram of hydrological sensitivities computed for every 36-yr period in historical simulations.

#### REFERENCES

- Adler, R. F., G. Gu, J.-J. Wang, G. J. Huffman, S. Curtis, and D. Bolvin, 2008: Relationships between global precipitation and surface temperature on interannual and longer timescales (1979–2006). *J. Geophys. Res.*, **113**, D22104, doi:10.1029/2008JD010536.
- , —, and G. J. Huffman, 2012: Estimating climatological bias errors for the Global Precipitation Climatology Project (GPCP). *J. Appl. Meteor. Climatol.*, **51**, 84–99, doi:10.1175/JAMC-D-11-052.1.
- Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 224–232, doi:10.1038/nature01092.
- Andrews, T., P. M. Forster, and J. M. Gregory, 2009: A surface energy perspective on climate change. *J. Climate*, **22**, 2557–2570, doi:10.1175/2008JCLI2759.1.
- , —, O. Boucher, N. Bellouin, and A. Jones, 2010: Precipitation, radiative forcing and global temperature change. *Geophys. Res. Lett.*, **37**, L14701, doi:10.1029/2010GL043991.
- Arkin, P. A., T. M. Smith, M. R. P. Sapiano, and J. Janowiak, 2010: The observed sensitivity of the global hydrological cycle to changes in surface temperature. *Environ. Res. Lett.*, **5**, 035201, doi:10.1088/1748-9326/5/3/035201.
- Bala, G., K. Caldeira, and R. Nemani, 2010: Fast versus slow response in climate change: Implications for the global hydrological cycle. *Climate Dyn.*, **35**, 423–434, doi:10.1007/s00382-009-0583-y.
- Bindoff, N. L., and Coauthors, 2013: Detection and attribution of climate change: From global to regional. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 867–952.
- Boucher, O., and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 571–657.
- Fuller, W. A., 1980: Properties of some estimators for the errors-in-variables model. *Ann. Stat.*, **8**, 407–422, doi:10.1214/aos/1176344961.
- Hansen, J., M. Sato, and R. Ruedy, 1997: Radiative forcing and climate response. *J. Geophys. Res.*, **102**, 6831–6864, doi:10.1029/96JD03436.
- , and Coauthors, 2005: Earth's energy imbalance: Confirmation and implications. *Science*, **308**, 1431–1435, doi:10.1126/science.1110252.
- Hartmann, D., and Coauthors, 2013: Observations: Atmosphere and surface. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 159–254.
- Iles, C. E., G. C. Hegerl, A. P. Schurer, and X. Zhang, 2013: The effect of volcanic eruptions on global precipitation. *J. Geophys. Res. Atmos.*, **118**, 8770–8786, doi:10.1002/jgrd.50678.
- Jia, L., and T. DelSole, 2012: Optimal determination of time-varying climate change signals. *J. Climate*, **25**, 7122–7137, doi:10.1175/JCLI-D-11-00434.1.
- Knutti, R., and G. C. Hegerl, 2008: The equilibrium sensitivity of the earth's temperature to radiation changes. *Nat. Geosci.*, **1**, 735–743, doi:10.1038/ngeo337.
- , T. F. Stocker, F. Joos, and G.-K. Plattner, 2002: Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature*, **416**, 719–723, doi:10.1038/416719a.
- Lambert, F. H., and N. E. Faull, 2007: Tropospheric adjustment: The response of two general circulation models to a change in insolation. *Geophys. Res. Lett.*, **34**, L03701, doi:10.1029/2006GL028124.
- Liepert, B. G., and M. Previdi, 2009: Do models and observations disagree on the rainfall response to global warming? *J. Climate*, **22**, 3156–3166, doi:10.1175/2008JCLI2472.1.
- Mitchell, J. F. B., S. Manabe, V. Meleshko, and T. Tokioka, 1990: Equilibrium climate change - and its implications for the future. *Climate Change: The IPCC Scientific Assessment*, J. Houghton, G. Jenkins, and J. Ephraums, Eds., Cambridge University Press, 131–172.
- Myhre, G., and Coauthors, 2013: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 659–740.
- O'Gorman, P., R. Allan, M. Byrne, and M. Previdi, 2012: Energetic constraints on precipitation under climate change. *Surv. Geophys.*, **33**, 585–608, doi:10.1007/s10712-011-9159-6.
- 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. Climate*, **19**, 3055–3069, doi:10.1175/JCLI3802.1.
- Takahashi, K., 2009: Radiative constraints on the hydrological cycle in an idealized radiative-convective equilibrium model. *J. Atmos. Sci.*, **66**, 77–91, doi:10.1175/2008JAS2797.1.
- Wentz, F. J., L. Ricciardulli, K. Hilburn, and C. Mears, 2007: How much more rain will global warming bring? *Science*, **317**, 233–235, doi:10.1126/science.1140746.
- 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**, 024033, doi:10.1088/1748-9326/8/2/024033.
- Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.
- Yang, F., A. Kumar, M. E. Schlesinger, and W. Wang, 2003: Intensity of hydrological cycles in warmer climates. *J. Climate*, **16**, 2419–2423, doi:10.1175/2779.1.
- Yin, X., A. Gruber, and P. Arkin, 2004: Comparison of the GPCP and CMAP merged gauge-satellite monthly precipitation products for the period 1979–2001. *J. Hydrometeorol.*, **5**, 1207–1222, doi:10.1175/JHM-392.1.