Climatic Change
https://doi.org/10.1007/s10584-018-2341-9

The role of aerosols and greenhouse gases in Sahel drought and recovery

Alessandra Giannini[1] · Alexey Kaplan[2]

Received: 27 May 2018 / Accepted: 12 November 2018/Published online: 07 December 2018

Abstract

We exploit the multi-model ensemble produced by phase 5 of the Coupled Model Intercomparison Project (CMIP5) to synthesize current understanding of external forcing of Sahel rainfall change, past and future, through the lens of oceanic influence. The CMIP5 multi-model mean simulates the twentieth century evolution of Sahel rainfall, including the mid-century decline toward the driest years in the early 1980s and the partial recovery since. We exploit a physical argument linking anthropogenic emissions to the change in the temperature of the sub-tropical North Atlantic Ocean relative to the global tropical oceans to demonstrate indirect attribution of late twentieth century Sahel drought to the unique combination of aerosols and greenhouse gases that characterized the post-World War II period. The subsequent reduction in aerosol emissions around the North Atlantic that resulted from environmental legislation to curb acid rain, occurring as global tropical warming continued unabated, is consistent with the current partial recovery and with projections of future wetting. Singular Value Decomposition (SVD) applied to the above-mentioned sea surface temperature (SST) indices provides a succinct description of oceanic influence on Sahel rainfall and reveals the near-orthogonality in the influence of emissions between twentieth and twenty-first centuries: the independent effects of aerosols and greenhouse gases project on the difference of SST indices and explain past variation, while the dominance of greenhouse gases projects on their sum and explains future projection. This result challenges the assumption that because anthropogenic warming had a hand in past Sahel drought, continued warming will result in further drying. In fact, the twenty-first century dominance of greenhouse gases, unchallenged by aerosols, results in projections consistent with warming-induced strengthening of the monsoon, a response that has gained in coherence in CMIP5 compared to prior multi-model exercises.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10584-018-2341-9) contains supplementary material, which is available to authorized users.

✉ Alessandra Giannini
alesall@iri.columbia.edu

[1] International Research Institute for Climate and Society, The Earth Institute at Columbia University, 61 Route 9W, Palisades, NY 10964, USA

[2] Lamont-Doherty Earth Observatory, Columbia University, 61 Route 9W, Palisades, NY 10964, USA

1 Introduction

The Sahel witnessed an outstanding climatic shift with the abrupt onset of drought in the late 1960s. The magnitude and spatial extent of the downward trend in precipitation over the twentieth century (Greene et al. 2009) and the persistence of years of deficient rainfall through the 1970s and 1980s (Lamb 1982; Nicholson 1983; Dai et al. 2004; Ali and Lebel 2009) led scientists to hypothesize that Sahelian drought may be a sign of anthropogenic change. Some pointed to the local pressure of rapid population growth on vegetation cover and the consequent impact of an increase in albedo on the atmospheric energy budget (Charney 1975). Others linked drought to the nascent preoccupation with the impact of global anthropogenic emissions, of greenhouse gases (GHGs), and other pollutants, on the general circulation of the atmosphere (Bryson 1973).

We reconcile the just-mentioned arguments about the role of anthropogenic emissions in Sahel drought, and more recent arguments attributing global precipitation changes to aerosols, natural and anthropogenic (Robock and Liu 1994; Gillett et al. 2004; Liepert et al. 2004; Lambert et al. 2004; Zhang et al. 2007; Haywood et al. 2013) with demonstration that the dominant cause of Sahel drought lies in changes in the surface temperatures of the global oceans (Folland et al. 1986; Giannini et al. 2003). We do so by linking the unique post-World War II combination of aerosol and greenhouse gas emissions to the patterns of sea surface temperature (SST) known to affect Sahel rainfall.

On interannual to millennial time scales, Sahel rainfall variations have been ascribed to interhemispheric differences in SST, whether global, restricted to the Atlantic, or to the tropical Atlantic Ocean (Lough 1986; Shanahan et al. 2009; Park et al. 2015; Lindzen and Nigam 1987; Kang et al. 2008; Schneider et al. 2014). Recent work attempting to explain the disagreement in the sign of projected twenty-first century rainfall change (Biasutti and Giannini 2006) with differences in projected SST change added the explicit consideration of oceanic warming, resulting in a linear regression model that describes Sahel rainfall as a function of two regional SST indices: the difference of tropical North and tropical South Atlantic SST averages and the average of tropical Indo-Pacific sector SSTs (Biasutti et al. 2008). This regression model was further distilled into a single predictor, the North Atlantic Relative Index (NARI; Giannini et al. 2013), computed as the difference between SST averages over the sub-tropical North Atlantic (10°N–40°N, 75°W–15°W) and the global tropical oceans (20°S–20°N). Here, we denote these as NA and GT, respectively, so that NARI = NA − GT. In this difference, the tropical SST average (GT) captures global vertical stability (Chou and Neelin 2004): the stabilization effect of warming SSTs is communicated vertically through deep convection and is spread laterally by upper-atmosphere wave dynamics, given the weak temperature gradient constraint (Chiang and Sobel 2002). The sub-tropical North Atlantic SST average (NA) counters stabilization through moisture supply (Giannini et al. 2008; Seth et al. 2013). To the extent that Sahel rainfall variations on different time scales are the result of the underlying oceanic forcing's preferred time scales of variation, our physical interpretation of NARI follows the paradigm that convection is in quasi-equilibrium with its environment (Emanuel et al. 1994). This assumption eliminates the need to consider time scales of variation separately, as long as these are longer than the time scales of convective adjustment (including the adjustment in the coupled ocean-atmosphere system described in Chiang and Sobel (2002)). In addition, it should be noted that a "relative index" effectively detrends variations (Vecchi et al. 2008), in the sense that insofar as both SST time series express GHG-induced warming, taking their difference largely removes it. In

observations, NARI explains ~50% of the variance in twentieth century Sahel rainfall variability, all time scales included (Giannini et al. 2013).

In the "CMIP5 simulation of Sahel rainfall" section, we show that the CMIP5 multi-model mean reproduces the twentieth century evolution of Sahel rainfall, rendering discussion of its attribution to external forcing possible. In the "Climate change in the Sahel through the lens of oceanic influence" section, we propose an explanation that reconciles studies of variability, which typically seek to relate oceanic forcing and regional rainfall response, with studies of change, which typically seek to attribute regional response to external forcing, natural or anthropogenic, by developing an argument for indirect attribution of Sahel rainfall, through the influence of anthropogenic emissions on sea surface temperatures. In the "Transformation of predictors" sub-section, we exploit Singular Value Decomposition (SVD) to describe the linearly independent linear combinations of the two SST predictors defined above, i.e., the averages of sub-tropical North Atlantic and global tropical SSTs. In the "Linear regressions of Sahel rainfall" sub-section, we use the resulting singular vectors in bivariate linear regressions that describe their influence on Sahel rainfall. We present explicit formulas for these SVD-based predictors in the bivariate case; although, as noted by Preisendorfer (1988), such formulas had been available since a publication by Pearson (1901), their presentation in terms of non-dimensional parameters, introduced here, helps to visualize the underlying relationships and facilitates the interpretation of results. In particular, we show that when the ratio of standard deviations of the original predictors is close to one, the bivariate SVD rotation results in the sum and difference of these original predictor time series. In our case, this means that NARI, the difference of NA and GT, which we defined above based on a physical argument, is nearly proportional to one of the SVD predictors and is exhaustively complemented by the other, which is nearly proportional to the sum of the same SST indices, and represents global warming. The derivation of all necessary formulas is relegated to the Appendix. In the "Conclusions: Past is not prologue" section, we conclude with a synthesis that finds coherence in attributing past Sahel drought to anthropogenic emissions, while lending credence to projections of a wetter future.

2 CMIP5 simulation of Sahel rainfall

The multi-model ensemble produced by the Coupled Model Intercomparison Project in support of the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), referred to as CMIP5 (Taylor et al. 2012), reproduces the observed twentieth century evolution of Sahel rainfall and projects wetter end of twenty-first century conditions more coherently than its CMIP3 predecessor (e.g., compare Biasutti and Giannini (2006) with Biasutti (2013)).

We restrict our analysis to 29 models participating in CMIP5, because only these have all the thermodynamic and dynamical variables necessary to evaluate their moisture budget (the subject of a parallel study). The models used are named in the tables in Supplementary Material. We focus on ensemble means, because we are interested in attribution of observed twentieth century Sahel rainfall variability, i.e., in describing its externally forced component, whether the external forcing is natural (i.e., due to variations in incoming top of the atmosphere insolation and spikes in aerosol concentration from volcanic eruptions) or anthropogenic (most notably, emissions of aerosols and GHGs from fossil fuel burning). Single-model ensemble sizes range between 1 and 10 in the historical simulations of the twentieth century and in the

RCP8.5 scenario simulations of the twenty-first century. (In the case of the pre-Industrial control, typically, one simulation is run per model, albeit of varying length.) For each model and type of simulation, when more than one is available, we average realizations in the model's ensemble mean. We then compute the multi-model mean as the average of the single models' ensemble mean, a procedure which further suppresses the manifestation of internal variability.

The top panels in Fig. 1 display standardized Sahel rainfall (the precipitation average for land points in 10°N–20°N, 20°W–40°E, in each model's original resolution, over July–September (Nicholson 1983; Giannini et al. 2003)): twentieth century (1900–1999) time series are depicted on the left-hand side and twenty-first century (2006–2099) time series on the right-hand side. To facilitate comparison, we standardize all the time series: each model's

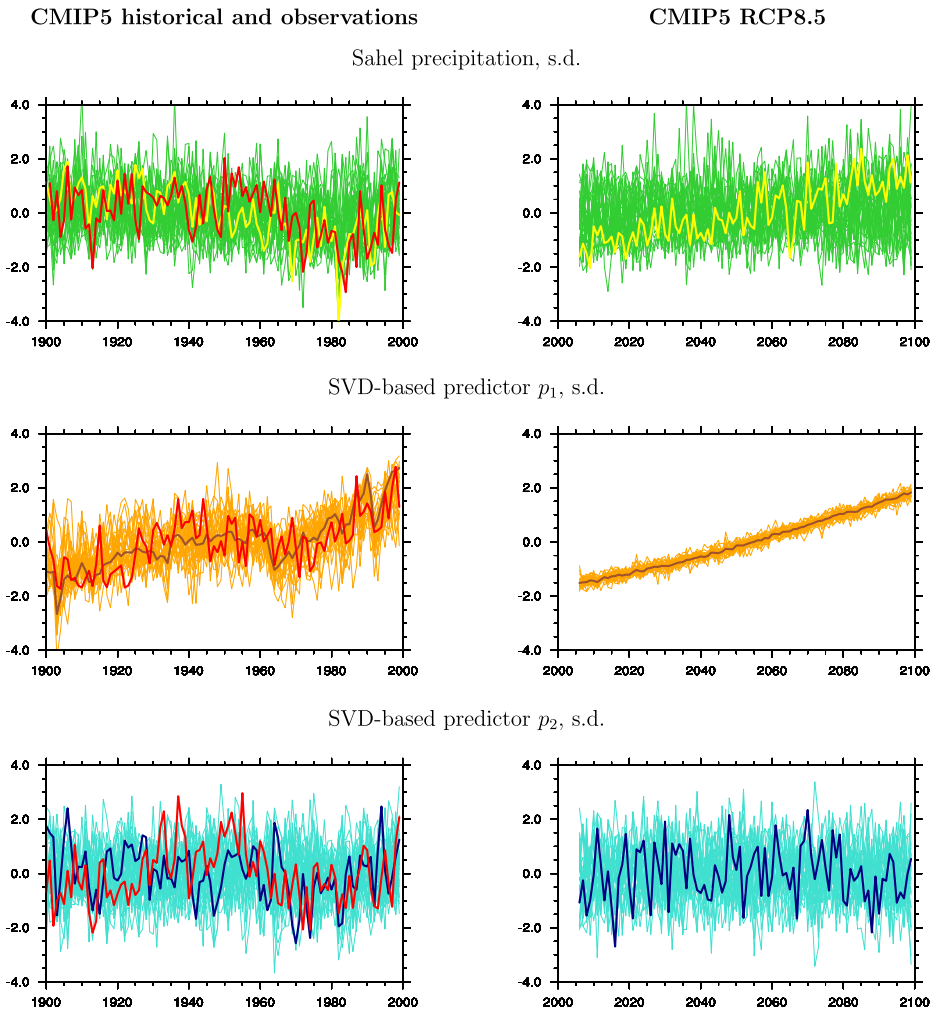

Fig. 1 Time series of standardized Sahel rainfall (*top*) and time series of the dominant and trailing SVD modes of the predictors' space p_1 (*middle*) and p_2 (*bottom*) for the single-model ensemble means of 29 CMIP5 models (*thin green*, *turquoise*, and *orange lines*) and for the multi-model ensemble mean (*thick yellow*, *brown*, and *blue lines*), from the historical simulations (1900–1999) in the *left panels*, and from the RCP8.5 simulations (2006–2099) in the *right panels*. The *red lines* represent observations

ensemble mean, in the thin green lines, is standardized by its own mean and standard deviation. The multi-model mean, in the thick yellow line, is first computed as the average of the original (non-standardized) single-model ensemble mean precipitation time series and then standardized. The reason for proceeding in this order is that we wish to characterize the externally forced component of Sahel rainfall. Since the models' ability to capture the relevant physical processes varies, if we computed the multi-model mean on the standardized single-model ensemble means, we would give all models, "good" and "bad," equal weight. Conversely, if the multi-model mean is taken on rainfall anomalies, the internal variability that dominates models that do not respond to external forcing is weighted out, because by definition it does not have a preferred temporal sequencing. This allows the temporally coherent response to external forcing to emerge. For a comparison to non-standardized time series, i.e., Sahel rainfall anomalies in units of mm/day, the reader is referred to Fig. 2 in

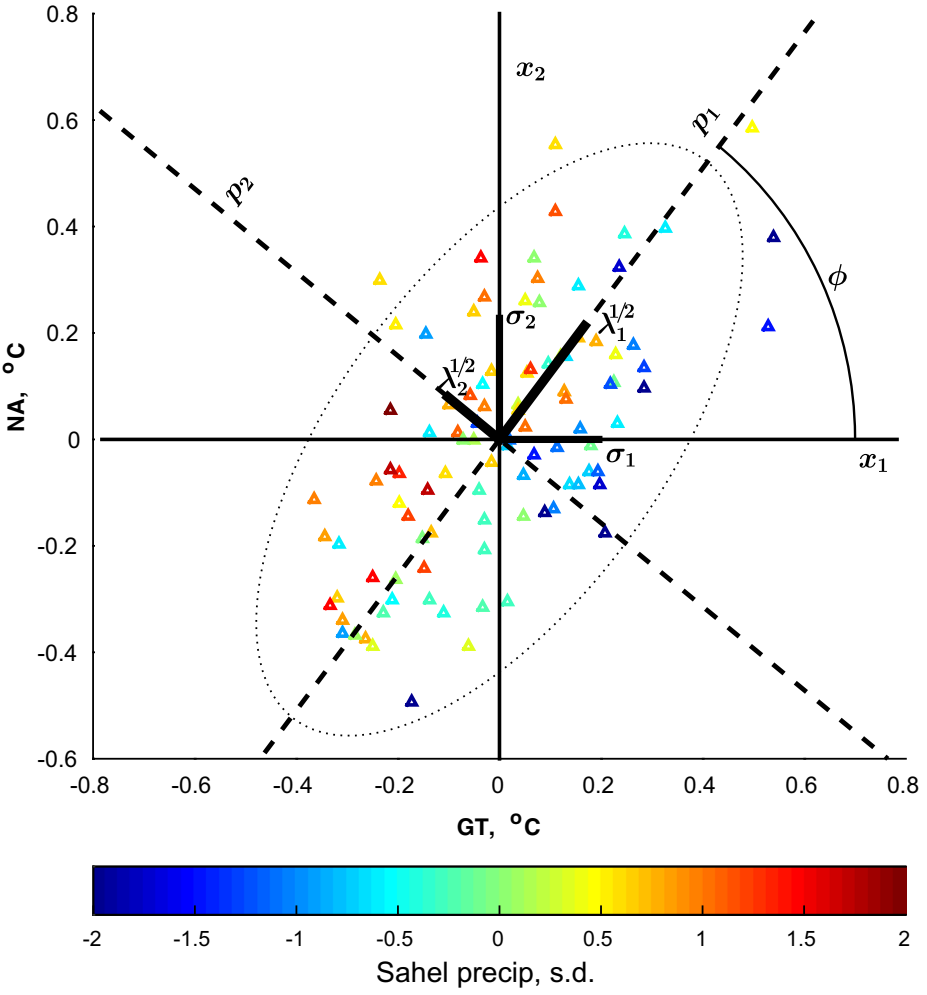

Fig. 2 Scatterplots of sub-tropical North Atlantic and global tropical ocean temperatures in observations (1901–1999) colored by standardized values of Sahel rainfall, with depictions of the angle of rotation ϕ, standard deviations of the original (σ_1, σ_2) and SVD-based predictors ($\lambda_1^{1/2}$, $\lambda_2^{1/2}$) and the concentration ellipse (*dotted*): see text for details

Biasutti (2013): there it can be seen that (i) the amplitude of variations in the multi-model mean is smaller than that in observations, as expected, given that the single real-world realization expresses the combination of internal variability and externally forced change. However, (ii) the observed evolution falls within the envelope defined by individual ensemble member simulations. The thick red line in the left panel of our Fig. 1 represents observed Sahel precipitation (Becker et al. 2013). Its correlation with the multi-model mean is 0.37 and is corroborated by a visible qualitative match: the first half of the century is characterized by variations around the mean, and is followed by a precipitous decline that lasts until the mid-1980s, and a partial recovery since. Standardization, because it normalizes the amplitude of variation, makes it possible to clearly discern the minimum in the multi-model mean, which occurs in 1982 and coincides with the first of three driest consecutive years in observations, 1982–1984, and with the eruption of El Chichón (Robock and Liu 1994). In contrast, the twenty-first century multi-model mean (top, right panel of Fig. 1) shows an upward trend, albeit with periods characterized by interannual-to-decadal variations that mask the trend, most notably at the end of the century.

3 Climate change in the Sahel through the lens of oceanic influence

We build an argument for indirect attribution based on the influence of external forcing—external to the coupled ocean-atmosphere system—on the SST patterns that affect Sahel rainfall. In the "Transformation of predictors" sub-section, we concisely describe oceanic influence using SVD applied to our set of two original predictors for Sahel rainfall, sub-tropical North Atlantic and global tropical SSTs. In the "Linear regressions of Sahel rainfall" sub-section, we use the resulting singular vectors as predictors in bivariate regression. Results of the analyses of the multi-model means and observations are reported in Tables 1 and 2. Results of the analyses of the single-model ensemble means are collected in tables in Supplementary Material. The Appendices provide a brief review of the nomenclature and definitions for SVD in the general *m*-dimensional case (Singular Value Decomposition of an m-dimensional predictors' space) and for linear regression (Use of time-dimensional singular vectors as predictors), as well as the derivation of the explicit form for the parameters of SVD in the bivariate case (Singular Value Decomposition of a two-dimensional predictors' space; derivation of Eqs. (1)–(5)).

Table 1 Predictors' space characteristics for the multi-model means from CMIP5 simulations and for twentieth century observations

Century	CMIP5 simulation	σ_1, °C GT	σ_1, °C NA	ϱ	k	ϕ, arc°	λ_1 °C^2	λ_2 °C^2	$\lambda_1/(\lambda_1+\lambda_2)$ %	$\lambda_2/(\lambda_1+\lambda_2)$ %
19th	PI control	0.03234	0.04195	0.369	0.771	62.75	0.0020	0.0008	71.90	28.10
20th	Historical	0.13150	0.12475	0.955	1.054	43.42	0.0321	0.0007	97.74	2.26
21st	RCP8.5	0.86711	0.83118	0.999	1.043	43.79	1.4422	0.0005	99.96	0.04
1901–1999	Observations	0.19565	0.22748	0.619	0.860	51.86	0.0734	0.0166	81.52	18.48

Columns from the left to right show the simulated century and the type of CMIP simulation (or observational reference period), the standard deviations of the original predictors, σ_1 and σ_2, in °C (for GT and NA, respectively), their correlation coefficient, ρ, the ratio of their standard deviations, $k = \sigma_1/\sigma_2$, angle ϕ from Eq. (3), in arc degree, variances λ_1 and λ_2, in $(°C)^2$, respectively, explained by the leading and trailing modes, corresponding to time series p_1 and p_2, and the same in percent of the total variance $\lambda_1 + \lambda_2$

Table 2 Regression coefficients and skill for Sahel rainfall in CMIP5 multi-model means and in twentieth century observations

Century	CMIP5 simulation	a	b	Correlation coefficient of y, $\hat{y}$
19th	PI-control	0.054	0.312	0.317
20th	Historical	−0.320	0.477	0.574
21st	RCP8.5	0.783	0.171	0.801
1901–1999	Observations	−0.182	0.607	0.633

Columns from the left to right show the simulated century and the type of CMIP simulation (or observational reference period), regression coefficients a and b of Sahel rainfall on the time series p_1 and p_2, respectively, and the correlation coefficient between y, Sahel rainfall simulated in CMIP5, or observed, and its predicted values from regression model (6): $\hat{y} = ap_1 + bp_2$

3.1 Transformation of predictors

We relate differences between our predictors' covariance structures to differences in the external forcing applied to the pre-Industrial control and twentieth and twenty-first century simulations. The latter two types of simulations are run with time-varying external forcing, natural and anthropogenic: variations in each realization of a single model's ensemble are a combination of internal variability and externally forced change, and their ensemble mean is taken to filter out internal variability. The former are run with constant external forcing, including CO_2 concentrations held fixed at pre-Industrial levels (280 ppm), the intent being to provide a truthful estimation of each model's internal variability. We repeat analyses on each model's first 100 years of the pre-Industrial control simulation (typically, there is only one such simulation per model, of variable length) and on each model's ensemble mean (with ensemble sizes varying between one and ten) for the twentieth and twenty-first century simulations.

When there are only two predictors—in our case, $x_1 = \mathrm{GT}$ and $x_2 = \mathrm{NA}$—standardized and mutually uncorrelated SVD-based predictors p_1, p_2 are obtained from the original predictors x_1, x_2 by orthogonal rotation in the x_1x_2 plane followed by rescaling:

$$p_1 = (x_1\cos\phi + x_2\sin\phi)/\sqrt{\lambda_1} \tag{1}$$

and

$$p_2 = (x_2\cos\phi - x_1\sin\phi)/\sqrt{\lambda_2} \tag{2}$$

where the rotation angle ϕ and squared rescaling coefficient λ_1, λ_2 are given by:

$$\phi = \begin{cases} \frac{1}{2}\mathrm{arccot}\left[\left(k - k^{-1}\right)/2\rho\right] & : \rho > 0 \\ \pi/2 & : \rho = 0, 0 < k < 1 \\ 0 & : \rho = 0, k \geq 1 \\ \frac{1}{2}\mathrm{arccot}\left[\left(k - k^{-1}\right)/2\rho\right] + \pi/2 & : \rho < 0 \end{cases} \tag{3}$$

and

$$\lambda_1 = \frac{1}{2}\left(\sigma_1^2 + \sigma_2^2 + \mathcal{D}\right), \lambda_2 = \frac{1}{2}\left(\sigma_1^2 + \sigma_2^2 - \mathcal{D}\right), \mathcal{D} = \left[\left(\sigma_1^2 - \sigma_2^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2\right]^{1/2}. \tag{4}$$

In Eqs. (3)–(4), $k = \sigma_1/\sigma_2$ is the ratio of standard deviations σ_1 and σ_2 of x_1 and x_2, respectively, and ρ is their correlation coefficient. Incidentally, k and ρ are the two non-dimensional parameters, characterizing the variance-covariance matrix of x_1 and x_2 up to a constant factor; to represent the same in terms of p_1 and p_2, one such parameter is ϕ, while the relative value of either of λ_1 or λ_2 where $\lambda_1 > \lambda_2$, can be another:

$$\frac{\lambda_2}{\lambda_1 + \lambda_2} = \frac{1}{2} - \frac{1}{2}\left[\left(\frac{k - k^{-1}}{k + k^{-1}}\right)^2 + \left(\frac{2\rho}{k + k^{-1}}\right)^2\right]^{1/2}. \quad (5)$$

Derivation of Eqs. (1)–(5) is given in "Singular Value Decomposition of a two-dimensional predictors' space; derivation of Eqs. (1)–(5)".

Figure 2 illustrates these relationships in twentieth century observations (see the last row of Table 1 for the values of relevant parameters). The observed values of the original predictors are the coordinates of the color triangles on the (x_1, x_2) plane, with the corresponding standardized values of Sahel rainfall shown by color. The thick segments shown on the x_1, x_2 axes indicate sample standard deviations, σ_1 and σ_2, of these original predictors. Their values are approximately 0.20 and 0.23 °C, respectively, and their ratio is $k = \sigma_1/\sigma_2 = 0.86$. The correlation coefficient of the original predictors is $\rho = 0.62$. SVD-based predictors p_1, p_2, whose corresponding axes are shown by dash lines, are obtained by orthogonal rotation of the original predictors x_1, x_2 with rotation angle about $\phi = 52°$, indicated by the arc connecting the x_1 and p_1 axes. The thick segments on the p_1 and p_2 axes show standard deviations (in our notation, $\lambda_1^{1/2} = 0.27$ °C and $\lambda_2^{1/2} = 0.13$ °C) of the data projections on these axes. The variance of projections is maximized for p_1, among all possible directions on the x_1, x_2 plane, therefore $\lambda_1 > \lambda_2$, necessarily. Also by construction, the in-sample correlation coefficient between p_1 and p_2 is zero; since there are only two predictors, p_2, being the "last" of them, minimizes the variance among all directions on the x_1, x_2 plane. Therefore, p_2 only explains $\lambda_2/(\lambda_1 + \lambda_2) = 18.5\%$ of variance in the sample (while p_1 explains the other 81.5%). These features of the data and predictors' pairs are illustrated by the shape of "concentration ellipses" (von Storch and Zwiers 1999; see their example 2.8.12, p. 43). The 95% concentration ellipse is shown by the dotted line in Fig. 2; its major semiaxes are $2.45\lambda_1^{1/2}$ and $2.45\lambda_2^{1/2}$. It should be noted that λ_1 and λ_2 do not necessarily determine the relative importance of the predictors p_1 and p_2 in modeling other variables: as color changes among triangles in Fig. 2 demonstrate, p_2 is much more important than p_1 in describing Sahel rainfall changes over 1901–1999. This is confirmed by the calculation of the regression coefficients in the last row of Table 2.

In models, since the ratio, k, of standard deviations of global tropical and sub-tropical North Atlantic SSTs is also close to 1 in the twentieth and twenty-first centuries, and the correlation coefficient of the SST indices, ρ, is positive (Table 1), the angle ϕ in Eq. (3) is close to 45°. That the diagonal is the direction along which is expressed the greatest covariance between $x_1 = \mathrm{GT}$ and $x_2 = \mathrm{NA}$ can also be clearly seen in Fig. 3. The remaining variance, orthogonal to it by construction, is larger in the twentieth (blue dots in Fig. 3) than in the twenty-first century (red squares). Since $\sin\phi = \cos\phi$ for $\phi = 45^o$, p_1 and p_2, respectively, in Eqs. (1) and (2), become nearly proportional to the sum and difference of the original predictors, and specifically, p_2 approximates NARI. The time series of p_1 are depicted in the middle panels of Fig. 1: indeed, p_1 represents a "global warming" mode, evident in both sub-tropical North Atlantic and global tropical ocean temperatures. p_2 (~NARI) explains more variance in the twentieth than in the twenty-first century. Its standardized form is depicted in the bottom panels of Fig. 1. In the middle and bottom left-hand side panels, thick red lines represent the principal components in observations (Kaplan et al. 1998).

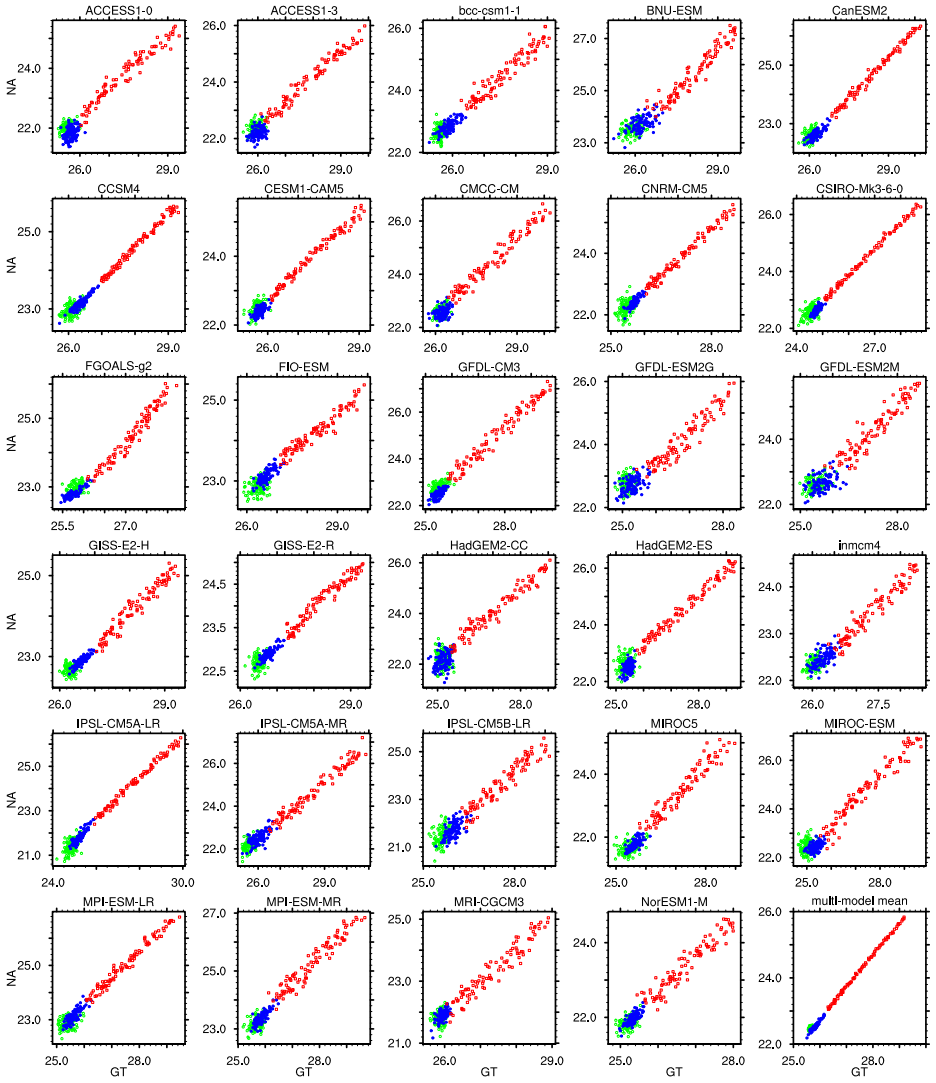

Fig. 3 Scatterplots of sub-tropical North Atlantic and global tropical ocean temperatures in 100 years of pre-Industrial control (*green circles*), in the twentieth century/historical (*blue dots*), and in the twenty-first century/RCP8.5 (*red squares*) simulations. The first 29 panels correspond to single-model ensemble means, and the last panel (in the lower right corner) corresponds to the multi-model mean

Our original predictors, GT and NA, are highly correlated in the twentieth and twenty-first century simulations, less so in the pre-Industrial control simulations (Fig. 3 and Table 1). As a consequence, the percent of variance explained by the leading mode, λ_1 in Eq. (4), increases in the simulations with time-dependent external forcing, as compared to those with constant forcing. The correlation further increases from twentieth to twenty-first century and consequently so does the relative variance of the leading mode, $\lambda_1/(\lambda_1+\lambda_2) = 1-\lambda_2/(\lambda_1+\lambda_2)$, see Eq. (5). This behavior represents one prominent way in which external forcing expresses itself: it forces a positive

correlation between the two predictors, through their common, GHG-induced warming. In addition, the relative cooling of the North Atlantic associated with aerosols contributes to the co-variability of indices with a second, preferred direction along which variance is expressed, which by construction is orthogonal to the first. This direction expresses more variance, visible in the scatter plots in Fig. 3, in the twentieth (blue dots) than in the twenty-first century (red squares).

3.2 Linear regressions of Sahel rainfall

Our SVD analysis of the original predictors confines any trend entirely to the leading mode time series, p_1, i.e., the "sum" or "warming" mode, while the residual mode, p_2, i.e., the "difference" or NARI, is completely devoid of any significant trend. We model standardized Sahel rainfall y, our predictand, by linear regression on p_1 and p_2 (see Section "Use of time-dimensional singular vectors as predictors" in the Appendix):

$$\hat{y} = ap_1 + bp_2 \tag{6}$$

Regression coefficients a and b and the correlation coefficients between simulated and predicted Sahel rainfall for the multi-model mean and for observations are reported in Table 2. Since the SVD-based predictor time series are standardized and mutually uncorrelated, and predictand time series are standardized, the regression coefficients, a and b, are precisely the correlation coefficients of the predictand and corresponding predictors. In the twentieth century, correlation of the multi-model mean simulation of Sahel rainfall with its SST-based regression is 0.574 and correlation in observations is 0.633. Both terms in the regression contribute, but p_2 contributes more (since $b > a$): late twentieth century Sahel drought is explained by the absence of North Atlantic warming, relative to warming of the global tropical oceans. The (positive) sign of b, the regression coefficient multiplying p_2 (NARI), is consistent across centuries. However, its magnitude and explanatory power vis à vis Sahel rainfall are drastically diminished in the twenty-first century. Over the twenty-first century, the correlation between the simulated Sahel rainfall and its in-sample prediction is 0.801. The multi-model ensemble achieves this high a correlation essentially through a positive correlation between Sahel rainfall and warming, captured by p_1. This happens because the coherence among models in projections of twenty-first century precipitation change has increased considerably from CMIP3 to CMIP5 (Biasutti 2013): a majority of models in CMIP5 project a future increase in Sahel rainfal to go along with warming (Schewe and Levermann 2017), in analogy to paleoclimate reconstructions of the "Green Sahara", when summer insolation was stronger than now (de Menocal et al. 2000). Among the models with a twenty-first century correlation coefficient between p_1 and Sahel rainfall larger than 0.45 (thus explaining ~20% of the variance with the single predictor p_1), seven of nine project future wetting (Table S3 in Supplementary Materials). In sum, regression models between centuries are nearly orthogonal: p_2, analogous to the difference of predictors or NARI, explains a larger fraction of twentieth century variation, while p_1, analogous to the sum of predictors or warming, alone explains the externally forced twenty-first century change.

4 Conclusions: Past is not prologue

Our analysis of the multi-model mean by design emphasizes the externally forced response and leads us to propose the following indirect attribution argument for Sahelian climate change, past and future. In the second half of the twentieth century, as global dimming (Stanhill and Cohen 2001; Liepert 2002) opposed global warming in the northern hemisphere, a unique combination of anthropogenic emissions contributed to the late twentieth century drying of the Sahel through their effect on sea surface temperatures: aerosols by cooling the North Atlantic and greenhouse gases by warming the tropical oceans, especially the Indian Ocean. Warming of the global tropical oceans "upped the ante" for deep convection (Chou and Neelin 2004; Held et al. 2005), while the absence of warming in the North Atlantic reduced the moisture supply to the monsoon and thus its potential to meet the "upped ante" and to trigger vertical instability (Giannini et al. 2008). Our argument is distinct from others previously proposed, which attributed late twentieth century Sahel drought solely to aerosols, whether through cooling of the North Atlantic or of the entire northern hemisphere (Rotstayn and Lohmann 2002; Kawase et al. 2010; Ackerley et al. 2011; Booth et al. 2012; Hwang et al. 2013; Park et al. 2015; Wang et al. 2016), in three ways. First of all, we argue for an *indirect effect* of anthropogenic emissions, i.e., mediated by sea surface temperatures. Linking emissions to Sahel rainfall through SSTs has the added advantage that it allows the synthesis in a single physical explanation of *natural* variability and *anthropogenic* change. Secondly, our argument is not about the role of interhemispheric gradients in SST shifting the latitudinal location of the Intertropical Convergence Zone and by extension rainfall in the Sahel. Rather, it is about *quasi-equilibrium in convection* as the world warms. The observed late twentieth century drying of the Sahel was more profound—longer lasting and of greater amplitude—than, e.g., the early twentieth century drought around 1910, despite the North Atlantic being cooler during the early period. This observation justifies the search for additional factors in drought, which we identify in greenhouse gas-induced tropical warming, which in the twentieth century occurred simultaneously with North Atlantic cooling. Therefore, thirdly, we argue for the *combined drying effect* of the two contrasting anthropogenic influences in twentieth century: drought was not caused by aerosols alone or by the cooling of the North Atlantic alone. Neither could it have been caused by greenhouse gases alone, as is made evident in simulations which include only the latter forcing (Biasutti 2013; Dong and Sutton 2015; Gaetani et al. 2017; Hill et al. 2017). Rather, drought resulted from the combination of aerosols and greenhouse gases. One influence cooled and reduced the moisture supply, while the other, warming, raised the threshold for convection—a double jeopardy. In the twenty-first century, external influence is dominated by GHG-induced warming, of both the North Atlantic and global tropical oceans. NARI, their difference, tends to zero, leaving all response to external forcing to be explained by their sum.

SVD rotation of the original predictors to their weighted sum and difference and the relative sizes of their regression coefficients (Table 2) expose the near-orthogonality of the twentieth and twenty-first centuries in the response of SSTs and rainfall to external forcing. The oceans' translation of external forcing of predominantly anthropogenic nature leads to contrasting but not inconsistent outcomes in the past and the future. Late twentieth century drying and twenty-first century projections of wetting are consistent with different balances in the influences of aerosol and greenhouse gas emissions on regional energy and moisture budgets, underlined by one coherent physical explanation rooted in the dynamics of quasi-equilibrium in convection. The fact that emissions need to be considered holistically, not additively, explains how it is

possible that while anthropogenic warming had a role in past drought, continued warming in the twenty-first century gives rise to a strengthening of the monsoon. As aerosol emissions have abated, especially around the North Atlantic, it is the combination of warmings of both North Atlantic and global tropical oceans that explains the strengthening of the monsoon: the upped ante can now be met. This situation is consistent with understanding of monsoon response to warming on paleoclimate time scales: when the ocean warms enough to contribute to an "upped ante" with increased moisture supply but not enough to shift the favored locus of convection from land to ocean (Chou et al. 2001), it reinforces the monsoon, as is the case for West Africa (Braconnot et al. 2012). With the usual caveats related to uncertainty in multi-model mean projections, that further warming may result in wetting of the Sahel is a conclusion worth reinforcing, in equal parts because it has gained in coherence between CMIP3 and CMIP5 (Biasutti 2013), and because the recovery of the rains is being experienced on the ground, in a fashion that is remarkably consistent with expectation from anthropogenic warming (West et al. 2008; Lodoun et al. 2013).

Acknowledgements This article is dedicated to the memory of professor A. Ben Mohamed. The authors acknowledge the technical support of Naomi Henderson and Haibo Liu, in the Ocean and Climate Physics division of LDEO. Through partial processing and local storage, they facilitated the access to the CMIP simulations for the IRI & LDEO community. The authors also acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and thank the climate modeling groups listed in the tables in Supplementary Material for producing and making available their model output. They acknowledge the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison, in partnership with the Global Organization for Earth System Science Portals, for their coordination to support and develop the software infrastructure to distribute CMIP simulations. AG was supported by National Science Foundation grant AGS-0955372 (CAREER), by National Aeronautics and Space Administration grant NNX16AN29G (SERVIR), and by Columbia University's World Project ACToday. AK was partially supported by Office of Naval Research Multidisciplinary University Research Initiatives grant N00014-12-1-0911 and by National Oceanic and Atmospheric Administration award NA17OAR4310156.

Appendix. Methods

Singular Value Decomposition of an *m*-dimensional predictors' space

Suppose an $n \times m$ matrix $\mathbf{X}$ contains, by columns, centralized samples of m variables x_1, x_2, ..., x_m (our *original predictors*), each observed at the same $n > m$ times. The well-known *thin* Singular Value Decomposition (SVD) presents $\mathbf{X}$ as:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{E}^{\mathbf{T}}$$

where $\mathbf{U}$ is an $n \times m$ matrix of orthonormal columns, $\mathbf{E}$ is an $m \times m$ orthogonal matrix, and $\mathbf{\Sigma}$ is a diagonal matrix with non-negative diagonal elements in non-ascending order, called singular values (Golub and Van Loan 1996). Superscript ^{T} denotes matrix transposition. The columns of $\mathbf{U}$ and $\mathbf{E}$ are called left and right singular vectors. Rescaling the left singular vectors by $\mathbf{P} = \sqrt{n-1}\mathbf{U}$ and introducing $\mathbf{\Lambda} = \mathbf{\Sigma}^2/(n-1)$, we rewrite $\mathbf{X}$ as:

$$\mathbf{X} = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{E}^{\mathrm{T}} \tag{7}$$

By construction, the columns of $\mathbf{P}$ are standardized, mutually uncorrelated time series. Each column is associated with a pattern of *loadings* (from the corresponding column of matrix $\mathbf{E}$)

on the original m variables and with the amount of variance (given by the corresponding diagonal element in $\mathbf{\Lambda}$) that it *explains* within the total variance of the m original variables. The latter is easy to see from observing that the variance-covariance matrix of the original variables:

$$\mathbf{C} = \mathbf{X}^{\mathrm{T}}\mathbf{X}/(n-1) = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^{\mathrm{T}} \tag{8}$$

has its eigenvalues and eigenvectors equal to diagonal elements of $\mathbf{\Lambda}$ and to columns of $\mathbf{E}$, respectively. The columns of $\mathbf{E}$ are also known as Empirical Orthogonal Functions (EOFs), while the columns of $\mathbf{V} = \mathbf{P}\mathbf{\Lambda}^{1/2}$ are known as Principal Components (PC) of the data set contained in matrix $\mathbf{X}$. Decomposition of the data set into EOFs and PCs:

$$\mathbf{X} = \mathbf{V}\mathbf{E}^{\mathrm{T}}$$

is often used in climate science to represent the overall co-variability of the original variables as the superposition of linearly independent modes, with several leading modes usually explaining most of the total variance in the system (von Storch and Zwiers 1999).

Singular Value Decomposition of a bivariate predictors' space; derivation of Eqs. (1)–(5)

We centralize our predictors, denote them x_1 and x_2, respectively, arrange them column-wise in matrix $\mathbf{X} = [\mathbf{x}_1\mathbf{x}_2]$, and perform the SVD of $\mathbf{X}$ in the form of Eq. (7). The columns of matrix $\mathbf{P} = [\mathbf{p}_1\ \mathbf{p}_2]$, called singular vectors, are obtained from the columns of $\mathbf{X}$ by an orthogonal rotation with 2x2 matrix $\mathbf{E}$, followed by their normalization by the square roots of diagonal elements of 2x2 matrix $\mathbf{\Lambda}$:

$$\mathbf{P} = \mathbf{X}\mathbf{E}\mathbf{\Lambda}^{-1/2} \tag{9}$$

By construction, $\mathbf{p}_1$ and $\mathbf{p}_2$ are standardized, mutually uncorrelated time series.

When there are only two predictors, x_1 and x_2, it is possible to present the elements of matrices $\mathbf{E}$ and $\mathbf{\Lambda}$ of the canonical decomposition of the predictors' variance-covariance matrix $\mathbf{C} = \mathbf{X}^{\mathrm{T}}\mathbf{X}/(n-1) = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^{\mathrm{T}}$ as functions of its elements.

Let standard deviations of x_1 and x_2 be σ_1 and σ_2, respectively, and their correlation coefficient be ρ, then their variance-covariance matrix is:

$$\mathbf{C} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}. \tag{10}$$

Note that:

$$\mathbf{C} = \frac{\sigma_1^2 + \sigma_2^2}{2}\,\mathbf{I} + \left[\left(\frac{\sigma_1^2 + \sigma_2^2}{2}\right)^2 + \rho^2\sigma_1^2\sigma_2^2\right]^{1/2} \begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix},$$

where $\mathbf{I}$ is a 2x2 identity matrix, and ϕ from the interval $[0, \pi)$ is determined by:

$$\cot(2\phi) = \left(\sigma_1^2 - \sigma_2^2\right)/(2\rho\sigma_1\sigma_2). \tag{11}$$

Introducing a non-dimensional parameter $k = \sigma_1/\sigma_2$, and expressing ϕ from Eq. (11) as a function of k and ρ, we obtain Eq. (3).

Using the identity:

$$\begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}^T$$

we rewrite:

$$\mathbf{C} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}^T$$

with λ_1 and λ_2 defined by Eqs. (4). Therefore, the canonical decomposition of $\mathbf{C}$ is performed with eigenvector and eigenvalue matrices, $\mathbf{E}$ and $\mathbf{\Lambda}$, as follows:

$$\mathrm{E} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \quad \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{12}$$

where ϕ is given by Eq. (3) and λ_1 and λ_2 by Eq. (4).

Inserting expressions for $\mathbf{E}$ and $\mathbf{\Lambda}$ from Eq. (12) into Eq. (9), we write them out in the component form for an arbitrary row of matrix P to obtain Eqs. (1)–(2).

Preisendorfer (1988) noted that Eqs. (11) and (4) are known since Pearson (1901). Equation (3) derived above gives an explicit form for the angle ϕ, expressing it through the non-dimensional parameters k and ρ of the original predictors' variance-covariance matrix, making it more convenient for interpretation. A traditional non-dimensional parameter of such matrix for SVD-based predictors p_1, p_2, the portion of variance explained by one of the predictors, follows from Eq. (4). For p_1, it is expressed through k and ρ as:

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{1}{2} + \frac{1}{2}\left[\left(\frac{k - k^{-1}}{k + k^{-1}}\right)^2 + \left(\frac{2\rho}{k + k^{-1}}\right)^2\right]^{1/2}. \tag{13}$$

For the predictor p_2 with the smaller variance, λ_2, since $\lambda_2/(\lambda_1 + \lambda_2) = 1 - \lambda_1/(\lambda_1 + \lambda_2)$, using (13) results in Eq. (5). The general structure of the dependence of the angle ϕ, characterizing the rotation of the original predictors toward uncorrelated singular vectors, and of ratio $\lambda_2/(\lambda_1 + \lambda_2)$, characterizing the relative variance of the second (trailing) predictors' mode, on non-dimensional parameters of the original predictors' variance-covariance matrix, ρ, the correlation coefficient of the original predictors x_1 and x_2, and k, their ratio of standard deviations, are illustrated in Fig. 4, with points corresponding to the data samples of multi-model means from CMIP5 simulations and from observations, as listed in Table 1.

Finally, the relationships between parameters of variance-covariance matrices in the (x_1, x_2) and (p_1, p_2) predictor planes are illustrated by the shape of "concentration ellipses," which are described by equation:

$$x_1^2/\sigma_1^2 - 2\rho x_1 x_2/(\sigma_1\sigma_2) + x_2^2/\sigma_2^2 = \left(1 - \rho^2\right) q(\alpha)$$

in (x_1, x_2) coordinates and by equation:

$$p_1^2/\lambda_1 + p_2^2/\lambda_2 = q(\alpha)$$

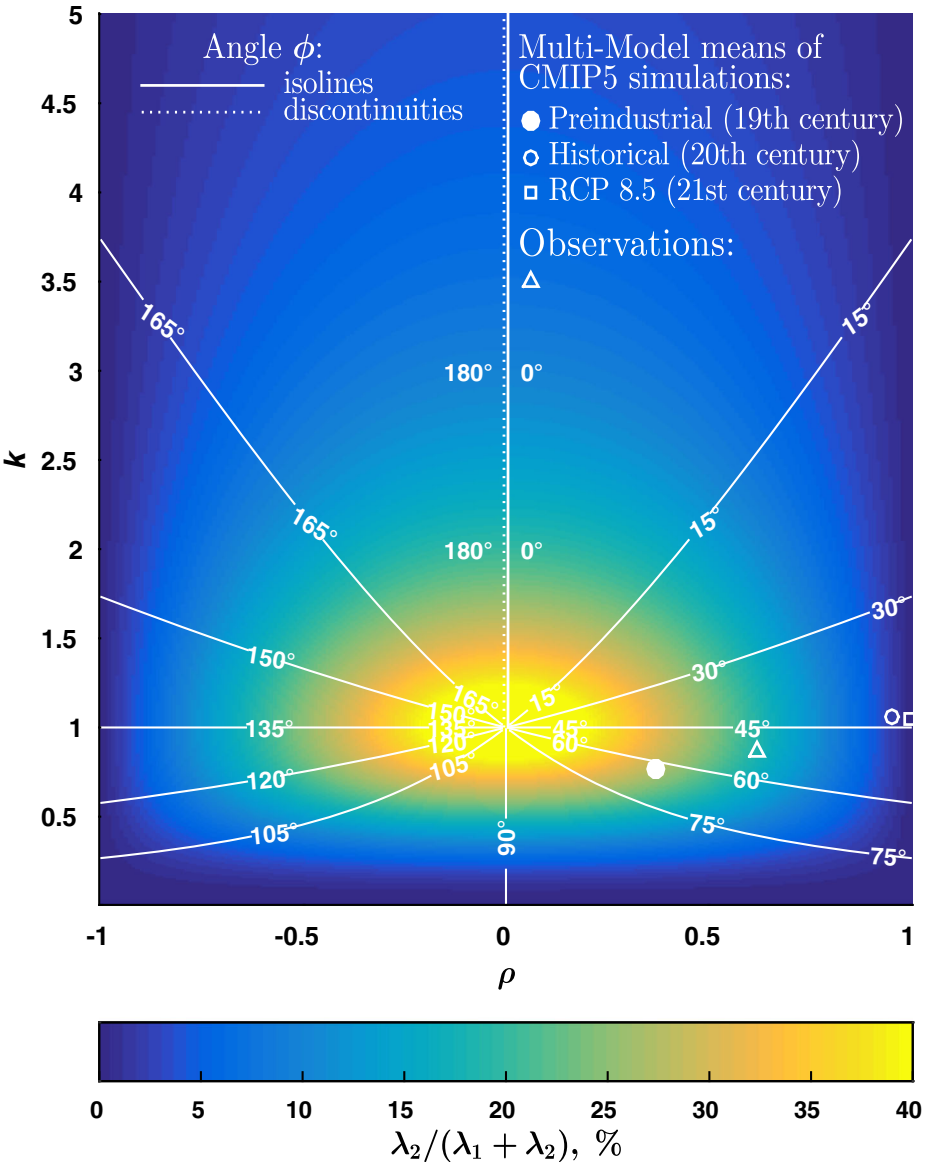

Fig. 4 Dependence of the angle ϕ (*white contours*) and of ratio $\lambda_2/(\lambda_1+\lambda_2)$, in percent (*color shading*) on the correlation coefficient ρ and ratio of standard deviations k of GT and NA. White dot, circle, square, and triangle show points on the (ρ, k) plane corresponding to the multi-model means of CMIP5 pre-Industrial control, historical, and RCP8.5 scenario simulations and of observations, respectively. Note that $\phi(\rho, k)$ is discontinuous at all points of the ray $\{(\rho, k): \rho = 0, k \geq 1\}$. In particular, while the definition in Eq. (3) made ϕ right-continuous at $\rho = 0$ for all $k > 1$, since $\phi(0, k) = \phi(0+, k) = 0$, the left limit of ϕ in these points is different: $\phi(0-, k) = \pi$

in (p_1, p_2) coordinates. Here, $q(\alpha)$ is the $\alpha\%$ quantile of the $\chi^2(2)$ distribution with two degrees of freedom, so that if (x_1, x_2) observations obey a bivariate normal distribution, $\alpha\%$ of the data points is contained within this concentration ellipse (von Storch and Zwiers 1999, see their

example 2.8.12, p. 43). An example of such an ellipse, expected to contain (for normally distributed data) $\alpha = 95\%$ of the data, is shown in Fig. 2.

Use of time-dimensional singular vectors as predictors

In regression studies, it often proves advantageous to use as predictors a set of standardized and mutually uncorrelated variables, like $p_1, p_2, ..., p_m$ (their time series $\mathbf{p}_i$, $i = 1, 2, ..., m$, are given by columns of matrix **P** and represent rescaled time-dimensional (left) singular vectors of **X**), instead of the original variables that lack these properties, leading to more tractable and easier to interpret results. Indeed, consider the regression model:

$$y = a_1 p_1 + a_2 p_2 + \ldots + a_m p_m + \varepsilon \tag{14}$$

where ε is a random error and, for simplicity, the predictand y is assumed standardized. The vector of unknown regression coefficients $\mathbf{a} = (a_1, a_2, \ldots, a_m)^T$ has to be found by the least-squares best fit to the system

$$\mathbf{y} = \mathbf{P}\mathbf{a}$$

where **y** is an n-dimensional vector—a column that contains the standardized data sample for variable y. Dividing normal equations

$$\left(\mathbf{P}^{\mathrm{T}}\mathbf{P}\right)\mathbf{a} = \mathbf{P}^{\mathrm{T}}\mathbf{y}$$

by n–1, we find, since $\mathbf{P}^{\mathrm{T}}\mathbf{P}/(n-1) = \mathbf{I}$ (where **I** is an $n \times n$ identity matrix),

$$\mathbf{a} = \mathbf{P}^{\mathrm{T}}\mathbf{y} \Big/ (n-1) = (\rho_1, \rho_2, \ldots, \rho_m)^T,$$

where ρ_i, for each $i = 1, 2, ..., m$, is the sample correlation coefficient between p_i and y ($\rho_i = p_i^T y/(n-1)$, since **y** and columns of **P** are all standardized). Therefore, regression model Eq. (14) results in the prediction:

$$\hat{y} = \rho_1 p_1 + \rho_2 p_2 + \ldots + \rho_m p_m \tag{15}$$

for which the sample estimate $\hat{\rho}$ of its correlation coefficient with the actual value of y satisfies the relationship:

$$\hat{\rho}^2 = \rho_1^2 + \rho_2^2 + \ldots \rho_m^2,$$

making it easy to see how the inclusion of time series p_i of each individual mode i into the regression model contributes to the overall skill of prediction. Another measure of in-sample skill estimate for prediction Eq. (15), the standard deviation $\hat{e}$, of the difference $\hat{y}$–y, can be expressed through $\hat{\rho}$ as $\hat{e} = \sqrt{1-\hat{\rho}^2}$.

References

Ackerley D, Booth BBB, Knight SHE, Highwood EJ, Frame DJ, Allen MR, Rowell DP (2011) Sensitivity of 20th century Sahel rainfall to sulfate aerosol and CO2 forcing. J Climate 24:4999–5014

Ali A, Lebel T (2009) The Sahelian standardized rainfall index revisited. Int J Climatol 29:1705–1714

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, Ziese M (2013) A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. Earth Syst Sci Data 5:71–99. https://doi.org/10.5194/essd-5-71-2013

Biasutti M (2013) Forced Sahel rainfall trends in the CMIP5 archive. J Geophys Res 118:1613–1623

Biasutti M, Giannini A (2006) Robust Sahel drying in response to late 20th century forcings. Geophys Res Lett 11:L11706. https://doi.org/10.1029/2006GL026067

Biasutti M, Held IM, Sobel AH, Giannini A (2008) SST forcings and Sahel rainfall variability in simulations of 20th and 21st centuries. J Clim 21:3471–3486

Booth BBB, Dunstone NJ, Halloran PR, Andrews T, Bellouin N (2012) Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. Nature 484:228–232

Braconnot P, Harrison S, Kageyama M, Masson-Delmotte V, Abe-Ouchi A, Otto- Bliesner B, Zhao Y (2012) Evaluation of climate models using palaeoclimatic data. Nat Clim Chang 2:417–424

Bryson RA (1973) Drought in Sahelia who or what is to blame? Ecologist 3:366–371

Charney JG (1975) Dynamics of deserts and drought in the Sahel. Q J Roy Meteor Soc 101:193–202

Chiang JCH, Sobel AH (2002) Tropical tropospheric temperature variations caused by ENSO and their influence on the remote tropical climate. J Clim 15:2616–2631

Chou C, Neelin JD (2004) Mechanisms of global warming impacts on regional tropical precipitation. J Clim 17(13):2688–2701

Chou C, Neelin JD, Su H (2001) Ocean-atmosphere feedbacks in an idealized monsoon. Q J R Meteorol Soc 127:1869–1891

Dai A, Lamb PJ, Trenberth KE, Hulme M, Jones PD, Xie P (2004) The recent Sahel drought is real. Int J Climatol 24:1323–1331

de Menocal P, Ortiz J, Guilderson T, Adkins J, Sarnthein M, Baker L, Yarusinskya M (2000) Abrupt onset and termination of the African Humid Period: rapid climate responses to gradual insolation forcing. Quat Sci Rev 19:347–361

Dong B, Sutton R (2015) Dominant role of greenhouse-gas forcing in the recovery of Sahel rainfall. Nat Clim Chang 5:757–760

Emanuel KA, Neelin JD, Bretherton CS (1994) On large-scale circulations in convecting atmospheres. Q J R Meteorol Soc 120:1111–1143

Folland CK, Palmer TN, Parker DE (1986) Sahel rainfall and worldwide sea temperatures, 1901-85. Nature 320: 602–607

Gaetani M, Flamant C, Bastin S, Janicot S, Lavaysse C, Hourdin F, Braconnot P, Bony S (2017) West African monsoon dynamics and precipitation: the competition between global SST warming and CO_2 increase in CMIP5 idealized simulations. Clim Dyn 48:1353–1373

Giannini A, Biasutti M, Held IM, Sobel AH (2008) A global perspective on African climate. Clim Chang 90: 359–383. https://doi.org/10.1007/s10584-008-9396-y

Giannini A, Salack S, Lodoun T, Ali A, Gaye AT, Ndiaye O (2013) A unifying view of climate change in the Sahel linking intra-seasonal, interannual and longer time scales. Environ Res Lett 8:024010

Giannini A, Saravanan R, Chang P (2003) Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. Science 302:1027–1030. https://doi.org/10.1126/science.1089357

Gillett NP, Weaver AJ, Zwiers FW, Wehner MF (2004) Detection of volcanic influence on global precipitation. Geophys Res Lett 31:L12217. https://doi.org/10.1029/2004GL020044

Golub GH, Van Loan CF (1996) Matrix computations, 3rd edn. Johns Hopkins University Press, Baltimore, MD, U.S.A., p 694

Greene AM, Giannini A, Zebiak SE (2009) Drought return times in the Sahel: a question of attribution. Geophys Res Lett 36:L12701. https://doi.org/10.1029/2009GL038868

Haywood JM, Jones A, Bellouin N, Stephenson D (2013) Asymmetric forcing from stratospheric aerosols impacts Sahelian rainfall. Nature Climate Change, 3:660–665. https://doi.org/10.1038/nclimate1857

Held IM, Delworth TL, Lu J, Findell K, Knutson TR (2005) Simulation of Sahel drought in the 20th and 21st centuries. Proc Natl Acad Sci 102:17891–17896. https://doi.org/10.1073/pnas.0509057102

Hill SA, Ming Y, Held IM, Zhao M (2017) A moist static energy budget-based analysis of the Sahel rainfall response to uniform oceanic warming. J Clim 30:5637–5660

Hwang Y-T, Frierson DMW, Kang SM (2013) Anthropogenic sulfate aerosol and the southward shift of tropical precipitation in the late 20th century. Geophys Res Lett 40:2845–2850. https://doi.org/10.1002/grl.50502

Kang SM, Held IM, Frierson DMW, Zhao M (2008) The response of the ITCZ to extratropical thermal forcing: idealized slab-ocean experiments with a GCM. J Clim 21(14):3521–3532
Kaplan A, Cane MA, Kushnir Y, Clement AC, Blumenthal MB, Rajagopalan B (1998) Analyses of global sea surface temperature 1856-1991. J Geophys Res 103:18,567–18,589
Kawase H, Abe M, Yamada Y, Takemura T, Yokohata T, Nozawa T (2010) Physical mechanism of long-term drying trend over tropical North Africa. Geophys Res Lett 37:L09706
Lamb PJ (1982) Persistence of Subsaharan drought. Nature 299:46–48
Lambert FH, Stott PA, Allen MR, Palmer MA (2004) Detection and attribution of changes in 20th century land precipitation. Geophys Res Lett 31:L10203. https://doi.org/10.1029/2004GL019545
Liepert BG 2002 Observed reductions in surface solar radiation in the United States and worldwide from 1961 to 1990. Geophys Res Lett doi:https://doi.org/10.1029/2002GL014910
Liepert BG, Feichter J, Lohmann U, Roeckner E (2004) Can aerosols spin down the water cycle in a warmer and moister world? Geophys Res Lett 31:L06207. https://doi.org/10.1029/2003GL019060
Lindzen RS, Nigam S (1987) On the role of sea surface temperature gradients in forcing low-level winds and convergence in the tropics. J Atmos Sci 44:2418–2436
Lodoun T, Giannini A, Traoré PS, Somé L, Sanon M, Vaksmann M, Millogo-Rasolodimby J (2013) Changes in the character of precipitation in Burkina Faso associated with late-20th century drought and recovery in the Sahel. Environ Dev 5:96–108
Lough JM (1986) Tropical Atlantic Sea surface temperatures and rainfall variations in sub-Saharan Africa. Mon Weather Rev 114:561–570
Nicholson SE (1983) Sub-Saharan rainfall in the years 1976-1980: evidence of continued drought. Mon Weather Rev 111:1646–1654
Park J-Y, Bader J, Matei D (2015) Northern-hemispheric differential warming is the key to understanding the discrepancies in the projected Sahel rainfall. Nat Commun 6:5985. https://doi.org/10.1038/ncomms6985
Pearson K (1901) On lines and planes of closest fit to systems of points in space. Philos Mag, 6th Series 2:559–572
Preisendorfer RW (1988) *Principal component analysis in meteorology and oceanography*, posthumously compiled. In: Mobley CD (ed) Developments in atmospheric science, vol 17. Elsevier, Amsterdam, New York, pp xviii–xv425
Robock A, Liu Y (1994) The volcanic signal in Goddard Institute for Space Studies three-dimensional model simulations. J Clim 7:44–55
Rotstayn L, Lohmann U (2002) Tropical rainfall trends and the indirect aerosol effect. J Clim 15:2103–2116
Schewe J, Levermann A (2017) Non-linear intensification of Sahel rainfall as a possible dynamic response to future warming. Earth System Dynamics 8:495–505
Schneider T, Bischoff T, Haug GH (2014) Migrations and dynamics of the intertropical convergence zone. Nature 513:45–53. https://doi.org/10.1038/nature13636
Seth A, Rauscher SA, Biasutti M, Giannini A, Camargo SJ, Rojas M (2013) CMIP5 projected changes in the annual cycle of precipitation in monsoon regions. J Clim 26:7328–7351
Shanahan TM, Overpeck JT, Anchukaitis KJ, Beck JW, Cole JE, Dettman DL, Peck JA, Scholz CA, King JW (2009) Atlantic forcing of persistent drought in West Africa. Science 324:377–380
Stanhill G, Cohen S (2001) Global dimming: a review of the evidence for a widespread and significant reduction in global radiation with discussion of its probable causes and possible agricultural consequences. Agric For Meteorol 107:255–278
Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Amer Meteor Soc 93:485–498
Vecchi GA, Swanson KL, Soden BJ (2008) Whither hurricane activity? Science 322:687–689
von Storch H and Zwiers FW 1999 Statistical analysis in climate research. Cambridge University Press pp. 484
Wang H, Xie S-P, Tokinaga H, Liu Q, Kosaka Y (2016) Detecting cross-equatorial wind change as a fingerprint of climate response to anthropogenic aerosol forcing. Geophys Res Lett 43:3444–3450. https://doi.org/10.1002/2016GL068521
West CT, Roncoli C, Ouattara F (2008) Local perceptions and regional climate trends of the Central Plateau of Burkina Faso. Land Degrad Dev 19:289–304
Zhang X, Zwiers FW, Hegerl GC, Gillett NP, Solomon S, Stott PA, Nozawa T (2007) Detection of human influence on twentieth-century precipitation trends. Nature 448:461–466. https://doi.org/10.1038/nature06025

Climatic Change manuscript No.
(will be inserted by the editor)

The role of aerosols and greenhouse gases in Sahel drought and recovery: Supplementary material

Alessandra Giannini · Alexey Kaplan

Received: 27 May 2018 / Accepted: 12 November 2018

A. Giannini
International Research Institute for Climate and Society, Earth Institute at Columbia University, 61 Rt 9W, Palisades NY 10964, U.S.A.

A. Kaplan
Lamont-Doherty Earth Observatory, Earth Institute at Columbia University, 61 Rt 9W, Palisades NY 10964, U.S.A.

Supplementary Material

Table S1: Predictors' space characteristics and linear regression results for Sahel rainfall from each individual model's ensemble mean in CMIP5 pre-Industrial control simulations. Columns from left to right show model name, the correlation coefficient, ρ, of the original predictors (x_1 =GT, global tropical SST mean, and x_2 =NA, sub-tropical North Atlantic SST mean—see Section 1 for the exact definition of these indices), the ratio of their standard deviations, $k = \sigma_1/\sigma_2$, angle ϕ from Eq. (3), in arc degree, the relative variance of the trailing mode, $\lambda_2/(\lambda_1+\lambda_2)$, bivariate regression coefficients *a* and *b* of Sahel rainfall *y* (see Section 2 for an exact definition) on the SVD-based predictors p_1 and p_2, obtained from the original predictors x_1 and x_2, as described in Eqs. (1)-(4) (see Sections 5.1-5.2 in the Appendix for an explanation and derivations), and the correlation coefficient between *y*, Sahel rainfall as simulated in CMIP5, and its predicted values from regression model (6): $\hat{y} = ap_1 + bp_2$. Models highlighted in bold font are those whose value in the rightmost column is greater than 0.45, representing 20% of the total variance. Multi-model mean results are included in the last line of the table.

Table S2: same as in Table S1, but for the CMIP5 twentieth century/historical simulations. The second column from the left reports the ensemble size.

Table S3: same as in Table S1, but for the CMIP5 twenty-first century/RCP8.5 simulations. The second column from the left reports the ensemble size.

Table S1

CMIP5/PI-control pre-industrial	ρ	k	ϕ, arc $^\circ$	$\lambda_2/(\lambda_1+\lambda_2)$, %	a	b	Correlation coefficient of y, $\hat{y}$
ACCESS1-0	0.188	0.946	53.18	40.23	-0.235	0.164	0.286
ACCESS1-3	0.228	1.027	41.66	38.55	-0.312	0.283	0.421
bcc-csm1-1	0.173	0.589	81.31	24.62	-0.004	-0.168	0.168
BNU-ESM	**0.317**	**1.078**	**38.33**	**33.75**	**-0.404**	**0.244**	**0.472**
CanESM2	0.357	0.939	49.99	31.92	-0.289	0.264	0.391
CCSM4	**0.479**	**1.162**	**36.25**	**25.17**	**-0.360**	**0.380**	**0.523**
CESM1-CAM5	**0.380**	**0.873**	**54.88**	**30.01**	**-0.207**	**0.610**	**0.644**
CMCC-CM	0.316	0.778	64.35	30.37	0.141	-0.047	0.149
CNRM-CM5	0.320	0.599	74.57	22.48	0.136	-0.083	0.159
CSIRO-Mk3-6-0	0.440	0.922	50.23	27.69	0.138	0.073	0.156
FGOALS-g2	**0.177**	**0.952**	**52.81**	**40.81**	**-0.244**	**0.448**	**0.511**
FIO-ESM	0.266	1.292	22.90	32.01	-0.051	-0.166	0.173
GFDL-CM3	0.460	0.951	48.10	26.87	0.358	0.067	0.364
GFDL-ESM2G	0.132	0.824	72.91	38.47	0.185	0.215	0.284
GFDL-ESM2M	0.348	1.156	33.66	31.33	-0.202	0.141	0.246
GISS-E2-H	-0.067	0.989	130.18	46.62	-0.155	0.042	0.160
GISS-E2-R	0.141	0.745	77.38	34.16	0.180	-0.188	0.261
HadGEM2-CC	**0.007**	**0.580**	**89.63**	**25.20**	**0.175**	**0.533**	**0.561**
HadGEM2-ES	**0.107**	**0.688**	**82.14**	**31.47**	**0.062**	**0.458**	**0.462**
inmcm4	0.300	0.888	55.85	33.95	-0.163	0.301	0.342
IPSL-CM5A-LR	0.423	0.666	67.32	22.59	0.078	0.250	0.262
IPSL-CM5A-MR	0.452	0.785	59.21	25.02	-0.097	0.426	0.437
IPSL-CM5B-LR	0.398	0.467	77.29	14.47	0.134	-0.264	0.297
MIROC-ESM	0.180	0.668	78.23	29.15	-0.040	0.229	0.232
MIROC5	**0.393**	**1.302**	**27.90**	**27.04**	**-0.533**	**0.229**	**0.580**
MPI-ESM-LR	0.357	1.074	39.36	31.82	-0.369	0.253	0.447
MPI-ESM-MR	0.342	0.857	57.17	31.45	-0.241	0.312	0.394
MRI-CGCM3	0.115	0.519	85.34	20.85	0.152	0.043	0.158
NorESM1-M	0.352	0.964	47.99	32.30	-0.110	0.197	0.226
Multi-model mean	*0.369*	*0.771*	*62.75*	*28.10*	*0.054*	*0.312*	*0.317*

Table S2

CMIP5/historical [1900-1999]	ens size	ρ	k	ϕ, arc $^\circ$	$\lambda_2/(\lambda_1+\lambda_2)$, %	a	b	Correlation coefficient of y, $\hat{y}$
ACCESS1-0	1	0.303	0.735	67.97	29.22	-0.040	0.190	0.194
ACCESS1-3	1	0.299	0.908	53.96	34.36	-0.087	0.223	0.240
bcc-csm1-1	3	0.759	1.054	43.00	11.99	-0.015	0.289	0.290
BNU-ESM	1	0.553	1.144	38.13	21.80	-0.213	0.257	0.334
CanESM2	**5**	**0.758**	**1.186**	**38.63**	**11.72**	**-0.188**	**0.442**	**0.480**
CCSM4	**6**	**0.920**	**1.211**	**39.10**	**3.85**	**-0.367**	**0.488**	**0.610**
CESM1-CAM5	**3**	**0.621**	**1.086**	**41.19**	**18.78**	**-0.214**	**0.473**	**0.520**
CMCC-CM	1	0.322	1.027	42.60	33.85	-0.084	-0.112	0.140
CNRM-CM5	10	0.817	0.779	53.57	8.54	0.221	0.242	0.328
CSIRO-Mk3-6-0	10	0.765	0.884	49.59	11.56	-0.051	0.098	0.111
FGOALS-g2	5	0.881	1.150	40.48	5.83	-0.190	0.284	0.341
FIO-ESM	3	0.749	1.276	35.91	11.72	0.040	0.267	0.270
GFDL-CM3	4	0.732	0.847	51.41	13.00	0.297	0.334	0.447
GFDL-ESM2G	1	0.488	0.952	47.90	25.50	-0.244	0.300	0.386
GFDL-ESM2M	1	0.389	1.203	32.24	28.82	-0.192	-0.006	0.192
GISS-E2-H	**5**	**0.856**	**1.278**	**36.92**	**6.77**	**-0.334**	**0.422**	**0.538**
GISS-E2-R	**5**	**0.826**	**1.197**	**38.83**	**8.39**	**-0.316**	**0.360**	**0.479**
HadGEM2-CC	1	0.304	0.640	73.29	24.94	0.142	0.323	0.353
HadGEM2-ES	4	0.439	0.554	72.46	17.61	0.297	0.285	0.411
inmcm4	1	0.621	1.020	44.08	18.96	-0.180	0.135	0.225
IPSL-CM5A-LR	**5**	**0.901**	**0.806**	**51.80**	**4.72**	**-0.289**	**0.395**	**0.489**
IPSL-CM5A-MR	**1**	**0.675**	**0.982**	**45.77**	**16.23**	**-0.249**	**0.431**	**0.498**
IPSL-CM5B-LR	1	0.592	0.616	65.20	15.28	0.223	-0.180	0.286
MIROC-ESM	**3**	**0.462**	**1.065**	**41.11**	**26.75**	**-0.295**	**0.467**	**0.552**
MIROC5	**5**	**0.718**	**1.027**	**43.95**	**14.11**	**-0.228**	**0.593**	**0.635**
MPI-ESM-LR	3	0.826	0.833	51.29	8.40	0.072	0.373	0.380
MPI-ESM-MR	3	0.810	0.890	49.12	9.38	0.284	0.137	0.316
MRI-CGCM3	**3**	**0.615**	**0.670**	**61.89**	**15.81**	**0.403**	**0.365**	**0.544**
NorESM1-M	3	0.747	0.936	47.53	12.58	-0.162	0.248	0.296
Multi-model mean	***29***	***0.955***	***1.054***	***43.42***	***2.26***	***-0.320***	***0.477***	***0.574***

Table S3

CMIP5/RCP8.5 [2006-2099]	ens size	ρ	k	ϕ, arc $^\circ$	$\lambda_2/(\lambda_1+\lambda_2)$, %	a	b	Correlation coefficient of y, $\hat{y}$
ACCESS1-0	1	0.974	1.104	42.09	1.27	0.186	0.196	0.270
ACCESS1-3	1	0.982	1.112	41.91	0.87	0.088	0.070	0.112
bcc-csm1-1	1	0.959	1.051	43.51	2.07	-0.101	0.142	0.174
BNU-ESM	**1**	**0.960**	**0.911**	**47.78**	**2.00**	**0.639**	**0.309**	**0.709**
CanESM2	5	0.994	1.069	43.07	0.29	0.173	0.236	0.292
CCSM4	**6**	**0.994**	**1.150**	**40.98**	**0.32**	**0.613**	**0.081**	**0.619**
CESM1-CAM5	3	0.991	1.173	40.42	0.41	0.188	0.392	0.435
CMCC-CM	1	0.978	1.031	44.12	1.08	-0.147	0.216	0.262
CNRM-CM5	**5**	**0.990**	**0.978**	**45.63**	**0.49**	**0.520**	**0.105**	**0.531**
CSIRO-Mk3-6-0	**10**	**0.997**	**1.089**	**42.54**	**0.13**	**-0.933**	**0.078**	**0.936**
FGOALS-g2	**1**	**0.976**	**0.683**	**55.89**	**1.04**	**0.707**	**0.169**	**0.727**
FIO-ESM	3	0.965	1.624	31.21	1.37	-0.134	0.121	0.181
GFDL-CM3	**1**	**0.984**	**0.935**	**46.97**	**0.82**	**0.607**	**0.259**	**0.660**
GFDL-ESM2G	1	0.951	0.948	46.62	2.45	-0.224	0.007	0.225
GFDL-ESM2M	1	0.945	0.975	45.76	2.76	-0.232	-0.025	0.233
GISS-E2-H	1	0.971	1.125	41.53	1.42	-0.314	0.305	0.438
GISS-E2-R	**2**	**0.973**	**1.236**	**38.82**	**1.30**	**-0.556**	**0.318**	**0.640**
HadGEM2-CC	3	0.978	1.074	42.90	1.11	0.252	0.300	0.391
HadGEM2-ES	4	0.989	1.118	41.76	0.56	-0.084	0.366	0.375
inmcm4	1	0.967	1.036	43.96	1.65	-0.001	0.085	0.085
IPSL-CM5A-LR	4	0.995	1.085	42.64	0.24	-0.089	0.293	0.306
IPSL-CM5A-MR	1	0.982	1.116	41.80	0.87	-0.051	0.225	0.231
IPSL-CM5B-LR	1	0.955	0.840	50.21	2.18	0.199	-0.127	0.236
MIROC-ESM	**1**	**0.973**	**0.939**	**46.86**	**1.36**	**0.636**	**0.123**	**0.648**
MIROC5	**3**	**0.963**	**0.989**	**45.33**	**1.87**	**0.450**	**0.622**	**0.768**
MPI-ESM-LR	3	0.988	1.028	44.20	0.59	-0.019	0.000	0.019
MPI-ESM-MR	1	0.965	1.064	43.16	1.76	0.201	0.347	0.401
MRI-CGCM3	1	0.962	0.927	47.25	1.90	0.241	0.158	0.288
NorESM1-M	**1**	**0.953**	**1.038**	**43.88**	**2.34**	**0.431**	**0.344**	**0.552**
Multi-model mean	***29***	***0.999***	***1.043***	***43.79***	***0.04***	***0.783***	***0.171***	***0.801***