

# Near-Surface Salinity as Nature's Rain Gauge to Detect Human Influence on the Tropical Water Cycle

LAURENT TERRAY AND LOLA CORRE

*Sciences de l'Univers au CERFACS, CERFACS/CNRS, URA1875, Toulouse, France*

SOPHIE CRAVATTE AND THIERRY DELCROIX

*IRD, LEGOS, and Université de Toulouse, UPS (OMP-PCA), Toulouse, France*

GILLES REVERDIN

*LOCEAN/IPSL, CNRS/UPMC, Paris, France*

AURÉLIEN RIBES

*CNRM-GAME, Météo-France/CNRS, Toulouse, France*

(Manuscript received 17 December 2010, in final form 12 August 2011)

## ABSTRACT

Changes in the global water cycle are expected as a result of anthropogenic climate change, but large uncertainties exist in how these changes will be manifest regionally. This is especially the case over the tropical oceans, where observed estimates of precipitation and evaporation disagree considerably. An alternative approach is to examine changes in near-surface salinity. Datasets of observed tropical Pacific and Atlantic near-surface salinity combined with climate model simulations are used to assess the possible causes and significance of salinity changes over the late twentieth century. Two different detection methodologies are then applied to evaluate the extent to which observed large-scale changes in near-surface salinity can be attributed to anthropogenic climate change.

Basin-averaged observed changes are shown to enhance salinity geographical contrasts between the two basins: the Pacific is getting fresher and the Atlantic saltier. While the observed Pacific and interbasin-averaged salinity changes exceed the range of internal variability provided from control climate simulations, Atlantic changes are within the model estimates. Spatial patterns of salinity change, including a fresher western Pacific warm pool and a saltier subtropical North Atlantic, are not consistent with internal climate variability. They are similar to anthropogenic response patterns obtained from transient twentieth- and twenty-first-century integrations, therefore suggesting a discernible human influence on the late twentieth-century evolution of the tropical marine water cycle. Changes in the tropical and midlatitudes Atlantic salinity levels are not found to be significant compared to internal variability. Implications of the results for understanding of the recent and future marine tropical water cycle changes are discussed.

## 1. Introduction

Evidence is building that human-induced climate change is changing the global water cycle (Zhang et al. 2007; Santer et al. 2007; Willett et al. 2007) and precipitation frequency, intensity, and spatial distribution with consequences on hydrological extreme events such

as floods and droughts (Trenberth 2011; Dai 2010; Min et al. 2011). Increased heating will enhance land evaporation and surface drying, leading to more severe droughts, while increased air moisture due to an increase in saturation vapor pressure with temperature will produce more intense precipitation, enhancing the risks of floods. While much uncertainty still remains regarding the prediction of how much warming will occur through greenhouse gas buildup, observed and projected changes in the hydrological cycle have recently been undergoing increased levels of scrutiny. While there appear to be robust

---

*Corresponding author address:* Laurent Terray, CERFACS, 42 Ave. Gaspard Coriolis, 31057 Toulouse CEDEX 01, France.  
E-mail: terray@cerfacs.fr

physical controls on the global hydrological cycle changes due to atmosphere energy balance constraints (Allen and Ingram 2002; Held and Soden 2006; Richter and Xie 2008; Stephens and Ellis 2008), it is much more problematic to project how regional precipitation and evaporation patterns might be altered within the context of climate change. This is particularly relevant for the tropical oceans where projected hydrological cycle trends appear to be influenced by spatial variations of sea surface temperature (SST) change in addition to the wet get wetter and dry get drier pattern associated with quasi-uniform tropical SST warming (Xie et al. 2010). Despite common features among projected spatial patterns of SST change, deviations from uniform tropical SST warming differ between models leading to large uncertainty with regard to the future distribution of tropical precipitation and evaporation changes. Furthermore, early detection and attribution of these changes is also hampered by the difficulty and lack of long-term freshwater flux observations over the oceans and their high space and time variability. Present evidence for a changing tropical marine hydrological cycle must then be searched for elsewhere. It is now well established that surface ocean salinity provides nature's largest possible rain gauge and can be efficiently used as an indicator of the changing marine water cycle (Schmitt 2008). Large-scale salinity variations are mainly shaped by the evaporation minus precipitation patterns and oceanic circulation. While the former mechanism acts to create salinity contrasts, ocean circulation and small-scale mixing act to attenuate them. Furthermore and in contrast with SST, there are to first order no strong direct and local feedbacks between sea surface salinity (SSS) and surface freshwater fluxes. This is in particular relevant for the tropical oceans where SSS seasonal to decadal variability mainly reflects to first order the freshwater flux changes linked to dominant variability modes such as El Niño–Southern Oscillation (ENSO) and the Pacific decadal oscillation (PDO) for the Pacific Ocean (Delcroix 1998; Delcroix et al. 2007). SSS thus appears to be an efficient integrated indicator of the marine hydrological changes. Here, we use observed SSS datasets and large ensembles of climate model simulations to identify possible human-induced changes in the tropical marine hydrological cycle evolution over the past few decades. The observed and model data used in this study are briefly described in the following section. Section 3 details the two detection methodologies applied to infer the nature of the surface salinity changes. Section 4 then reviews and compares the observed and simulated (past and future) changes in tropical surface salinity. Section 5 examines the question of the anthropogenic influences on the recent SSS changes. This section begins by comparing

the signal-to-noise ratio between SSS and freshwater fluxes. It then compares the basin-averaged and interbasin observed SSS with model estimates of internal variability. It further examines the relevance of the pattern scaling or scalability assumption for the changes in response to anthropogenic forcing. Finally, two detection methods are applied to assess whether a human influence can be detected in the recently observed SSS changes. The paper concludes with a brief summary and a discussion of the results within the context of the future marine water cycle and ocean changes.

## 2. Observed and model data

### *a. Observed salinity data*

For this investigation, we use a compilation of Pacific (30°S–30°N, 1950–2008) and Atlantic (30°S–50°N, 1970–2002) SSS datasets, which consist of in situ salinity observations gridded on a monthly basis on a  $1^\circ \times 1^\circ$  grid [see Reverdin et al. (2007) and Delcroix et al. (2011) for full details about the various data sources and postprocessing]. Due to the use and mix of multiple instruments, observed sea surface salinity in this study is representative of near-surface salinity, meaning the 0–10-m depth-averaged salinity. Given that 90% of the time the observed vertical gradient of salinity in the upper 10 m is less than 0.05 (Hénocq et al. 2010) and taking into consideration our data processing and large-scale averaging, the mix of instruments and 0–10-m depth sampling is highly unlikely to bias the present trend analysis. An optimal interpolation analysis is applied to the quality controlled SSS data to produce a monthly SSS gridded field at  $1^\circ$  resolution. The algorithm solves a local problem at each grid point by using closely located data and accounting for the spatial and temporal scales of the physical variable. In addition to the SSS values, the algorithm also yields an error field at each grid point, which is given as a percentage of the interannual variance. This allows us to estimate at least qualitatively the confidence we can have when computing multidecadal trends. Complete description of the datasets and optimal interpolation methodologies, as well as additional references, can be found online (<http://www.legos.obs-mip.fr/observations/sss>). We only consider data with sufficient coverage and reasonable error estimates. We use observed SSS data from the full observational period, meaning January 1950–December 2008 and January 1970–December 2002 for the Pacific and Atlantic, respectively. We use SSS monthly means only at locations where data are present with an SSS error that is less than 90% of the SSS standard deviation. For the mean SSS climatology, we only consider grid cells that have at least 20% time coverage. For the overlapping 1970–2002 period trend

TABLE 1. Summary of the various models and experiments used together with the relevant lengths of the control integrations: X, model experiment used; N, no model data available. Here picntrl is a control simulation with constant preindustrial external forcings, 20c3m is a twentieth-century simulation with observed historical forcings, and sresa1b is a twenty-first-century simulation with the Special Report Emission Scenario A1B scenario for anthropogenic forcings.

CMIP3 model	picntrl	20c3m	sresa1b	CMIP3 model	picntrl	20c3m	sresa1b
Bergen Climate Model, version 2 (BCM2.0)	250	X	X	Flexible Global Ocean–Atmosphere–Land System Model (FGOALS)	350	X	X
Canadian Centre for Climate Modelling and Analysis (CCCma) Coupled General Circulation Model, version 3.1 (CGCM3.1)	1001	X	X	Istituto Nazionale di Geofisica e Vulcanologia (INGV) ECHAM4	100	X	X
CGCM3.1-T63	400	X	X	L’Institut Pierre-Simon Laplace Coupled Model, version 4 (IPSL CM4)	500	X	X
Centre National de Recherches Météorologiques Coupled Global Climate Model, version 3 (CNRM-CM3)	500	X	X	Model for Interdisciplinary Research on Climate 3.2, high-resolution version [MIROC3.2(hires)]	100	X	X
CSIRO Mark version 3.0 (CSIRO-MK3.0)	380	X	X	MIROC 3.2, medium-resolution version [MIROC3.2(medres)]	500	X	X
CSIRO Mark version 3.5 (CSIRO-MK3.5)	950	X	X	Meteorological Institute of the University of Bonn, ECHO-G Model (MIUBECHOG)	341	X	X
Geophysical Fluid Dynamics Laboratory Climate Model, version 2.0 (GFDL-CM2.0)	500	X	X	Max Planck Institute (MPI) ECHAM5	N	X	X
GFDL CM, version 2.1 (GFDL-CM2.1)	300	X	X	Meteorological Research Institute Coupled General Circulation Model, version 2.3.2a (MRI CGCM2.3.2a)	350	X	X
Goddard Institute for Space Studies Atmosphere–Ocean Model (GISS-AOM)	251	X	X	CCSM3.0	455	X	X
GISS Model E-H (GISS-EH)	400	X	X	National Center for Atmospheric Research (NCAR) Parallel Climate Model (PCM)	N	X	X
GISS Model E-R (GISS-ER)	500	X	X	HadCM3	422	X	X
Hadley Centre Global Environmental Model (HadGEM)	70	X	N	CCSM3.0	N	N	40 mem.

estimation, we only use grid cells containing at least five SSS observations during each of the seven pentads. This leads to the definition of an observational spatial mask, which is then used in the trend detection analysis. For the temporal detection analysis, we use regionally averaged SSS annual means constructed from the selected monthly means. Parts of the southeastern tropical Pacific, as well as a large area of the southern Atlantic, are poorly sampled and are therefore excluded from the detection analysis. The practical salinity scale (pss-78) is used throughout the paper, meaning that the salinity is a ratio and does not have physical units (Millero 1993).

### b. Model data

The modeled data used in this study are those of the coupled general circulation models archived by the World Climate Research Programme’s (WCRP) Coupled Model Intercomparison Project, Phase 3 (CMIP3), dataset

(available online at <http://www-pcmdi.llnl.gov/>). The three datasets used here are referred to by the Program for Climate Model Diagnosis and Intercomparison (PCMDI) as the picntrl, 20c3m, and sresa1b multimodel experiments (see Table 1 to see the model list and integrations used). The sresa1b simulations followed the Special Report Emission Scenario (SRES) A1B scenario for the anthropogenic forcing evolution. The 20c3m simulations have a wide variety of anthropogenic influences (such as well-mixed greenhouse gases, ozone, sulfate, and black carbon aerosols). Roughly half of them also include changes in natural forcings (including total solar irradiance and stratospheric aerosols following volcanic eruptions). The picntrl or control integrations have constant and preindustrial values for the anthropogenic and natural forcings. We consider only one realization from each of the 23 models under these experiments. It is worth noting that not all 23 models contribute to all calculated

quantities studied in this paper when relevant data are missing or a specific subset of models is being examined.

We also use a 40-member ensemble for the period 2000–61 performed with the Community Climate System Model version 3 (CCSM3) forced by the SRES A1B greenhouse gas emission and stratospheric ozone recovery scenarios (Branstator and Teng 2010). Each member undergoes the same external forcing and starts from the same ocean, land, and sea ice initial conditions given by those on 1 January 2000 from a single twentieth-century CCSM3 integration. All members have different atmospheric initial conditions taken from days during December and January 2000 from the same CCSM3 integration, and simulation differences are thus only due to the model internal variability.

We only analyze model data between 30°S and 50°N to fit the observational geographical domain. To ease the analysis, all model SSS results have been first interpolated onto a common regular grid (similar to the observational  $1^\circ \times 1^\circ$  grid) accounting for the model differences in the land–sea mask. Before the trend analysis and detection part of the process, observed and simulated SSS data were further gridded onto a common  $5^\circ \times 5^\circ$  grid in order to remove the small-scale noise and variability present in both the observations and simulations. Model data were then processed in the same way as were the observations, so that the model values are only used at the same places as those indicated by the observational mask.

### c. Freshwater flux observed estimates

We only use one precipitation and one evaporation dataset. The monthly mean precipitation totals from the Global Precipitation Climatology Project dataset version 2.1 are based on both satellite and gauge datasets and cover the 1979–2008 period (Huffman et al. 2009). Monthly mean evaporation results from the Objectively Analyzed air–sea Fluxes (OAFlux) objective analysis dataset are based on a blending of satellite retrievals and three atmospheric reanalyses, as well as on the version 3.0 of the Coupled Ocean–Atmosphere Response Experiment (COARE) bulk algorithm, for the estimation of the turbulent fluxes and cover the 1958–2008 period (Yu and Weller 2007). Here, we use data over the period 1979–2008 to get the maximum overlap between the two datasets.

## 3. Detection methods

To assess the extent to which significant SSS changes may be detected as being significantly different from those expected from internal variability, we apply two different optimal detection methodologies. The first one, the optimal fingerprint approach (OF hereafter),

has already been extensively used to detect changes in atmospheric temperature and precipitation among other parameters (Allen et al. 2006; Hegerl et al. 2007). Here, we use the OF approach to compare quantitatively current and future trends in observed and simulated SSS, respectively. The second approach, named here the temporal optimal detection approach (TOD thereafter), is a recently developed alternative to the standard optimal fingerprint analysis (Ribes et al. 2010). It consists of searching for a smooth temporal pattern in the observations. While any pattern can theoretically be used, we focus here on temporal patterns best suited to representing the mean response of the upper tropical and subtropical ocean to anthropogenic forcing. Consideration of the two approaches is interesting as they make very different assumptions and use different ways of enhancing the signal-to-noise ratio.

### a. Optimal fingerprint methodology

The OF method can be simply described as an optimal regression of the response guess patterns (the fingerprints) to some given external forcings against the corresponding observed patterns. Here, we are focusing on one (anthropogenic) signal. The optimal fingerprint method then assumes the following statistical model:

$$\Psi_{s,t} = \Psi_s^0 + \beta \mathbf{X}_{s,t} + \Phi_{s,t}, \quad (1)$$

where  $\Psi$  is an observed climate variable (here annual mean SSS) and  $\Psi^0$  its climatological mean,  $\mathbf{X}$  is the response to anthropogenic forcing,  $\Phi$  is a centered noise term representing internal variability, and  $s = 1, \dots, S$  and  $t = t_0, \dots, T$  are the spatial and temporal indices, respectively. The observations are thus represented as the sum of a scaled simulated response to anthropogenic forcing and internal variability. The fingerprints and observations are a priori depending on space and time. Due to the relatively short length of the tropical observed SSS record, we follow the common procedure and mask the time dimension by using linear trend and epoch difference spatial patterns for the observations and fingerprints, respectively. The observed patterns are taken to be 33-yr SSS linear trends over the 1970–2002 period, thereby masking the time dimension from (1). Using a scalability hypothesis (see below), the fingerprint patterns are taken as multimodel ensemble mean epoch differences between the last 30 yr of the twenty-first century (2070–99) and that of the twentieth century (1970–99). To account for sampling noise not only in the observations but also in the signal estimates, the total least squares (TLS) approach is used to derive unbiased estimates of the scaling regression coefficients and confidence intervals (Allen et Stott 2003).

The optimal detection also requires two independent estimates of internal variability needed for optimization ( $\mathbf{C}_{n1}$ ) and significance testing ( $\mathbf{C}_{n2}$ ). Simulated and observed spatial patterns are then projected onto a reduced space spanned by the leading  $p$  empirical orthogonal functions of  $\mathbf{C}_{n1}$  (see the appendix). The optimal detection then calculates the scaling factor (regression coefficient) by which the simulated response guess pattern can be scaled up or down while remaining consistent with the observed pattern. The 5th–95th percentile uncertainty range in the scaling factor is then assessed from the independent internal variability estimate  $\mathbf{C}_{n2}$  and the anthropogenic forcing response pattern is detected if the scaling factor is found to be positive and inconsistent with zero at the 10% significance level.

#### b. Temporal detection methodology

The statistical model behind the temporal optimal detection approach can be written as

$$\Psi_{s,t} = \Psi_s^0 + \mathbf{g}_s \mu_t + \Phi_{s,t}, \quad (2)$$

with the same definitions as above. Within this approach, the climate change signal is decomposed as a product  $\mathbf{g}_s \mu_t$ , where  $\mu$  and  $\mathbf{g}$  are the climate change temporal and spatial patterns, respectively, and therefore use the scalability assumption. The idea behind the TOD approach is to assume that  $\mu$  is known from model simulations, while  $\mathbf{g}$  is estimated from the observations via generalized regression. The TOD approach can then be seen as a two-step process: first, one needs to estimate  $\mu_t$ . Here, rather than using the global mean surface air temperature (Ribes et al. 2010), we prefer to use an oceanic integrated variable that can better reflect upper-ocean thermal changes linked to changes in radiative forcing. The second step is the statistical test of the null hypothesis  $H_0$ :  $\mathbf{g}_s = 0$  against  $H_1$ :  $\mathbf{g}_s \neq 0$  assuming knowledge of  $\mu_t$  (see Ribes et al. 2010). TOD also uses one assumption on the nature of  $\Phi$ , namely that  $\Phi$  is a stationary autoregressive process of order one in time [AR(1) or red noise,  $\Phi_{s,t} = \alpha \Phi_{s,t-1} + \Theta_{s,t}$ , with  $\Theta_{s,t}$  independent identically distributed random variables]. The TOD hypothesis testing procedure requires  $T \geq S + 2$ , implying that it cannot be applied over the global scale taking each individual grid point into account. Instead, we consider a small number of regions in each basin (see section 5 and Fig. 9a) based on the mean SSS distribution and data availability and use the spatially averaged annual mean SSS over these regions as pseudogridpoints in the detection test. Results from the TOD approach are presented in terms of the  $p$  value, which is the probability of obtaining a value of the test statistic more extreme than that actually observed under the null hypothesis  $H_0$ . The

time evolution of the  $p$  value is presented in order to illustrate the evolution of the anthropogenic signal and its emergence. The  $p$  value for a given date, say 1990, is then the  $p$  value of the detection test applied to the data over the periods 1956–90 and 1970–90 for the Pacific and Atlantic, respectively.

### 4. Observed and future salinity changes

#### a. Observed and recent salinity changes

SSS changes over the past decades exhibit a strong Pacific freshening and Atlantic salinity increase leading to a strengthening of the mean SSS interbasin contrast, which reflects to a large extent the mean pattern of freshwater fluxes (Fig. 1). The salty Atlantic due to net evaporation contrasts with a low-salinity Pacific associated with a net input of freshwater. Regional SSS changes of large amplitude seem to follow the rich gets richer paradigm: the warm and fresh pool of the western Pacific has freshened (Cravatte et al. 2009) whereas the salty subtropical North Atlantic has become saltier (Gordon and Giulivi 2008). The western equatorial Pacific and South Pacific convergence zone (SPCZ) salinity fronts have both migrated eastward during the last decades, leading to a spatial extension of the Pacific fresh pool (Cravatte et al. 2009). Salinity increases occurred in the western Coral Sea and the southeastern and central North Pacific. Atlantic changes show more contrast with freshening of the deep tropical central and north-western part and saltier waters over the northern subtropics and southern regions. Whether the equatorial band freshening is due to decadal variability or represent a lower-frequency signal is still a matter of debate (Grotsky et al. 2006). Although the density of SSS observations is scarce in the South Atlantic, the available data show increased salinity on the eastern and western sides. The main spatial features of the tropical SSS trend seem to be robust to the inclusion or not of the recent Argo profiling float data (Hosoda et al. 2009; Durack and Wijffels 2010). Furthermore, the recent (2003–08) Argo-derived South Atlantic positive SSS anomaly suggests that the corresponding low-frequency signal has a basin spatial scale. Further examination of Argo data also suggests that the recent salinity changes are not restricted to the surface layer and extend at depth, in particular in the subtropical Atlantic (Von Schuckmann et al. 2009; Durack and Wijffels 2010).

The observed SSS evolution thus shows intensified spatial contrasts, suggesting an increase in the marine tropical hydrological cycle strength. These SSS changes have been documented in recent studies using related or independent datasets (Curry et al. 2003; Boyer et al.



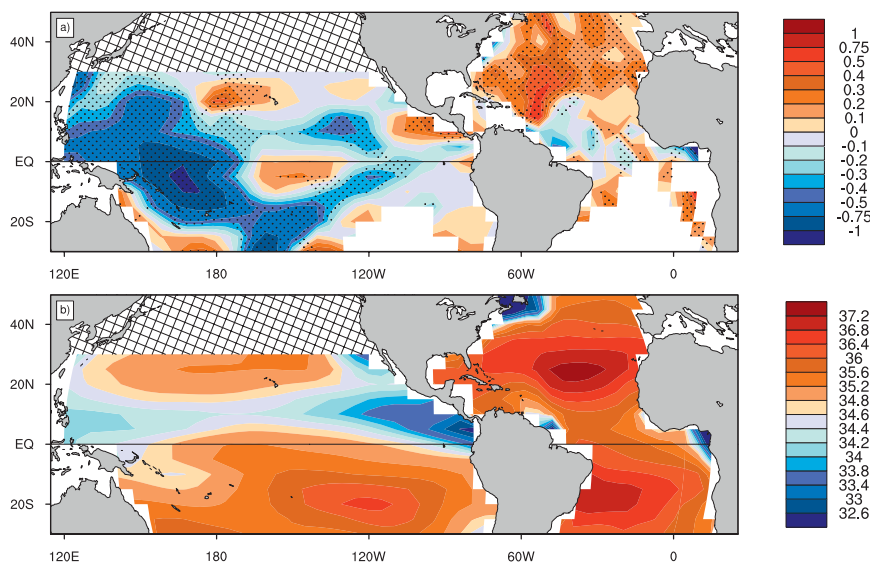


FIG. 1. Observed SSS trends and means. (a) The 33-yr (1970–2002) linear surface salinity trend ( $\text{century}^{-1}$ ) computed from monthly anomalies. White grid boxes indicate regions with insufficient or no data and stippling denotes areas where trends are statistically significant from 0 at the 5% level using a two-sided Student's  $t$  test. (b) Climatological mean observed surface salinities estimated over 1950–2008 and 1970–2002 for the Pacific and Atlantic, respectively. Geographical masks were applied in order to prevent mixing of data across the land boundaries.

2005; Bindoff et al. 2007; Reverdin et al. 2007; Cravatte et al. 2009; Durack and Wijffels 2010). However, none of these studies has tried to formally detect an anthropogenic fingerprint in the observed tropical SSS changes. A first attempt using a one-model detection and attribution study has recently been done and focused only on the Atlantic, using as detection vector SSS zonal means from a coarser observed dataset (Stott et al. 2008).

To support the idea that recent tropical SSS evolution may reflect changes in the tropical marine hydrological cycle, we compare recent observed trends in freshwater flux estimates with that of SSS. As we recognize that there are still large uncertainties regarding the recent evolution of the two main components of the ocean freshwater flux, we are only interested in the presence or lack thereof of qualitative agreement with SSS changes. Consequently, we view differences in observed SSS and freshwater trend estimates as indicative of observed datasets biases and/or the importance of ocean circulation changes in addition to that of the freshwater forcing. Coherent changes with recent SSS evolution include increased freshwater fluxes in the western Pacific and along the south and intertropical convergence zones, as well as decreased ones in the central equatorial Pacific and central-eastern subtropics (Fig. 2). The observed salinity increase along the east Australian coast seems to be partly related to an increase in evaporation combined with reduced precipitation. Changes in precipitation usually dominate evaporation ones in particular in the

convergence zones. The much larger zonal and meridional extents of the negative SSS anomalies in the western Pacific may be explained through the spread of SSS anomalies due to the influence of mean advection, including the Ekman divergent drift near the equator and weakening of the south equatorial current (Huang et al. 2005). The comparison is even more difficult for the Atlantic as the overlap period between the SSS and freshwater component datasets is smaller and the decadal variability is strong in particular in the tropics, which raises questions concerning the use of a linear trend to represent the observed changes. Nevertheless, both freshwater components seem to contribute to more saline surface waters in the northern subtropics and midlatitudes and to a freshening in the western tropical Atlantic.

#### b. Projected and future salinity changes

Before formally addressing the detection question, we first compare observed SSS trends with twenty-first-century climate model simulations. A first guess of the SSS response to anthropogenic forcing is simply the changes over the twenty-first century as simulated by the CMIP3 multimodel average. The multimodel mean SSS change displays large-scale features very similar to the observed trend with an increased interbasin contrast, as well as a fresher western Pacific warm pool and a saltier subtropical North Atlantic (Figs. 3a,b). Main differences are located in the equatorial and midlatitude North Atlantic as well as in the equatorial and subtropical

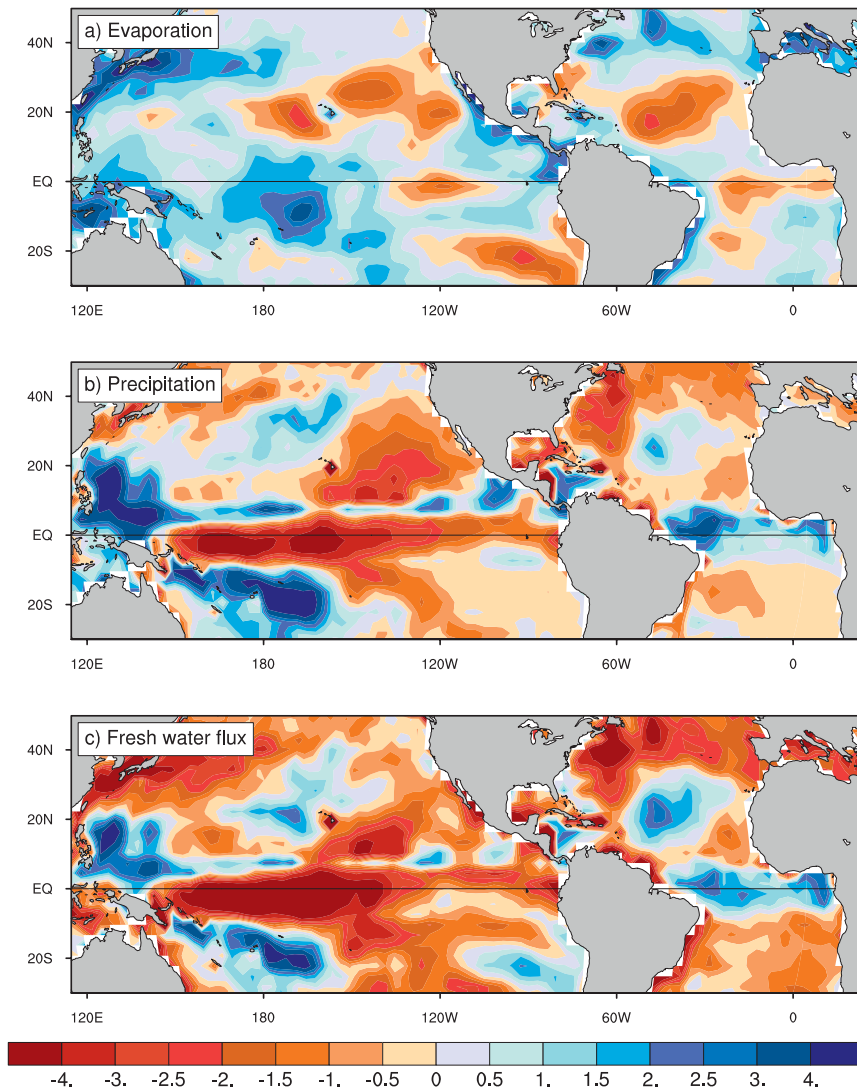


FIG. 2. Observed freshwater flux trends: 30-yr (1979–2008) linear trends using annual means for (a) evaporation from OAFlex and (b) precipitation from the Global Precipitation Climatology Project (GPCP). (c) Total freshwater flux defined as precipitation minus evaporation. Units are  $\text{mm day}^{-1} \text{ century}^{-1}$ .

central Pacific. The projected twenty-first-century equatorial Atlantic salinity increase and midlatitude freshening suggests that the recent and corresponding observed trend may not necessarily reflect the long-term response to anthropogenic forcing. We further note that most individual models consistently exhibit quite similar large-scale spatial patterns, suggesting that the multimodel mean can be used as a robust estimate of the response to anthropogenic forcing, at least in the tropical and subtropical oceans (Figs. 3 and A1). Although most models indicate freshening north of  $50^\circ\text{N}$ , it has been suggested that the twenty-first-century evolution of the North Atlantic SSS may not be necessarily linear due to the competing influences between the advection of

saltier water from the tropics and local precipitation and evaporation changes (Stott et al. 2008). Finally, there is no systematic difference in the projected SSS changes between models with and without freshwater flux corrections, suggesting that they can both be used to characterize the response to anthropogenic forcing. This also indicates that there is no clear relationship between model skill levels in simulating the mean state and in capturing the response to anthropogenic forcing (Santer et al. 2009).

### c. Evaluation of model performance in simulating variability

Most detection studies use climate model control simulations in which there is no change in forcing as

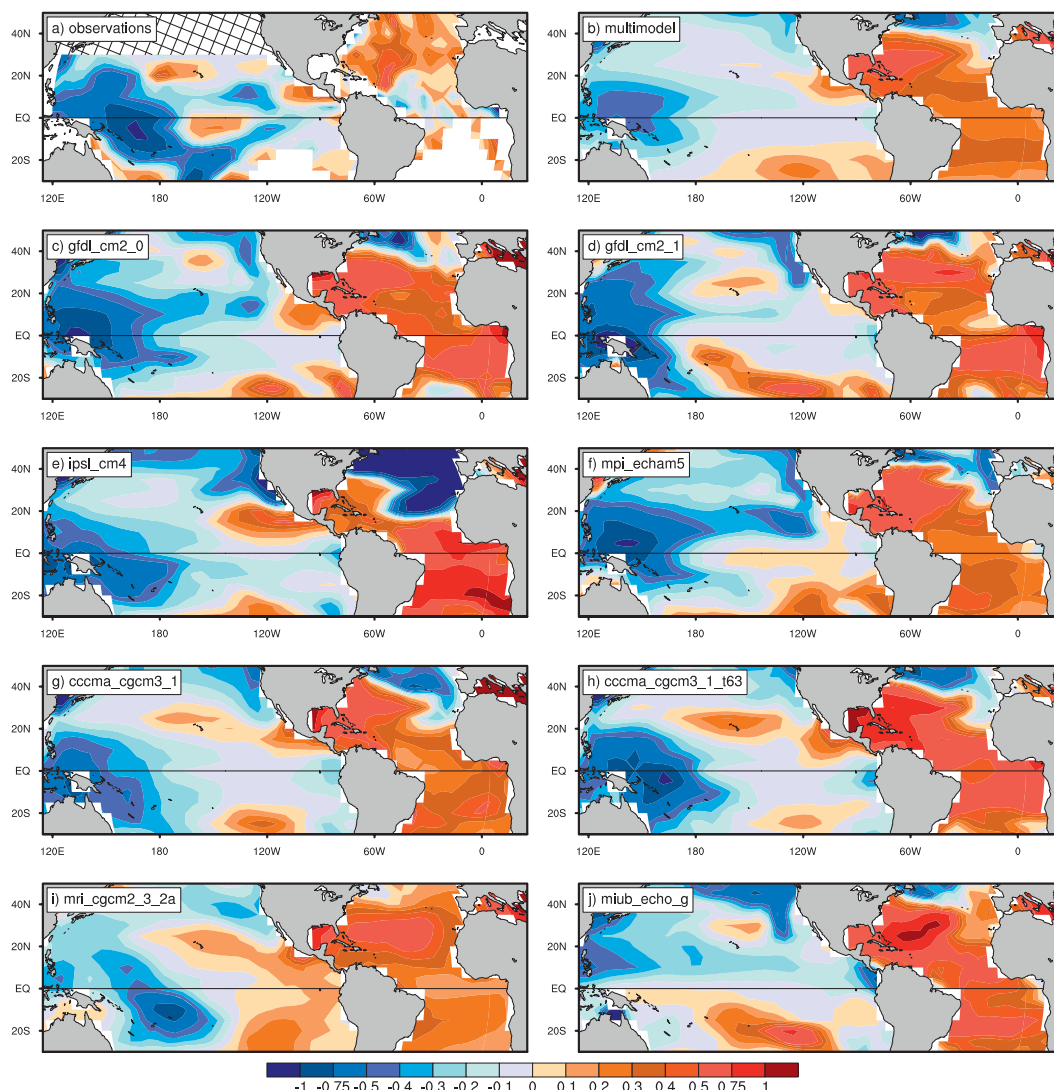


FIG. 3. Twenty-first-century SSS changes. (a) Recently (1970–2002) observed SSS trends and (b) multimodel average SSS changes ( $\text{century}^{-1}$ ) between the end of the twenty-first and twentieth centuries (2070–99 and 1970–99). Simulated changes for individual models: (c), (d), (e), (f) without freshwater flux adjustments and (g), (h), (i), (j) with freshwater flux adjustments. The individual models without flux adjustments are the ones selected in Delcroix et al. (2011).

pseudo-observations to provide an estimate of the distribution of the detection statistic under scrutiny under the null hypothesis of no anthropogenic change. It then would be relevant to assess the confidence in the realism of the model representation of internal climate variability. However, the SSS observed period is marked by increasing external forcing precluding any direct comparison with model internal variability. Here, we prefer to evaluate the model's ability to realistically simulate tropical SSS variability (see also Delcroix et al. 2011). We compare observed and twentieth-century-simulated SSS changes averaged over three different regions characterized by large observed monthly and interannual variabilities. The first two are in the western Pacific and

characterize the variations of the western equatorial and SPCZ salinity fronts, which exhibit variability at all time scales due to natural modes such as ENSO and PDO. The third one is located in the subtropical North Atlantic and is under the influence of the North Atlantic Oscillation (NAO) and Atlantic decadal variability. The model-average temporal standard deviation is actually close to (similar in the SPCZ front and underestimated by a majority of models in the western equatorial Pacific) or even larger (Atlantic) than the observed variability (Fig. 4). There is no sign of a systematic model underestimation of the monthly to interannual variability nor is there any sign of a uniform bias across models in term of the mean state. However, the model spread is large, suggesting that



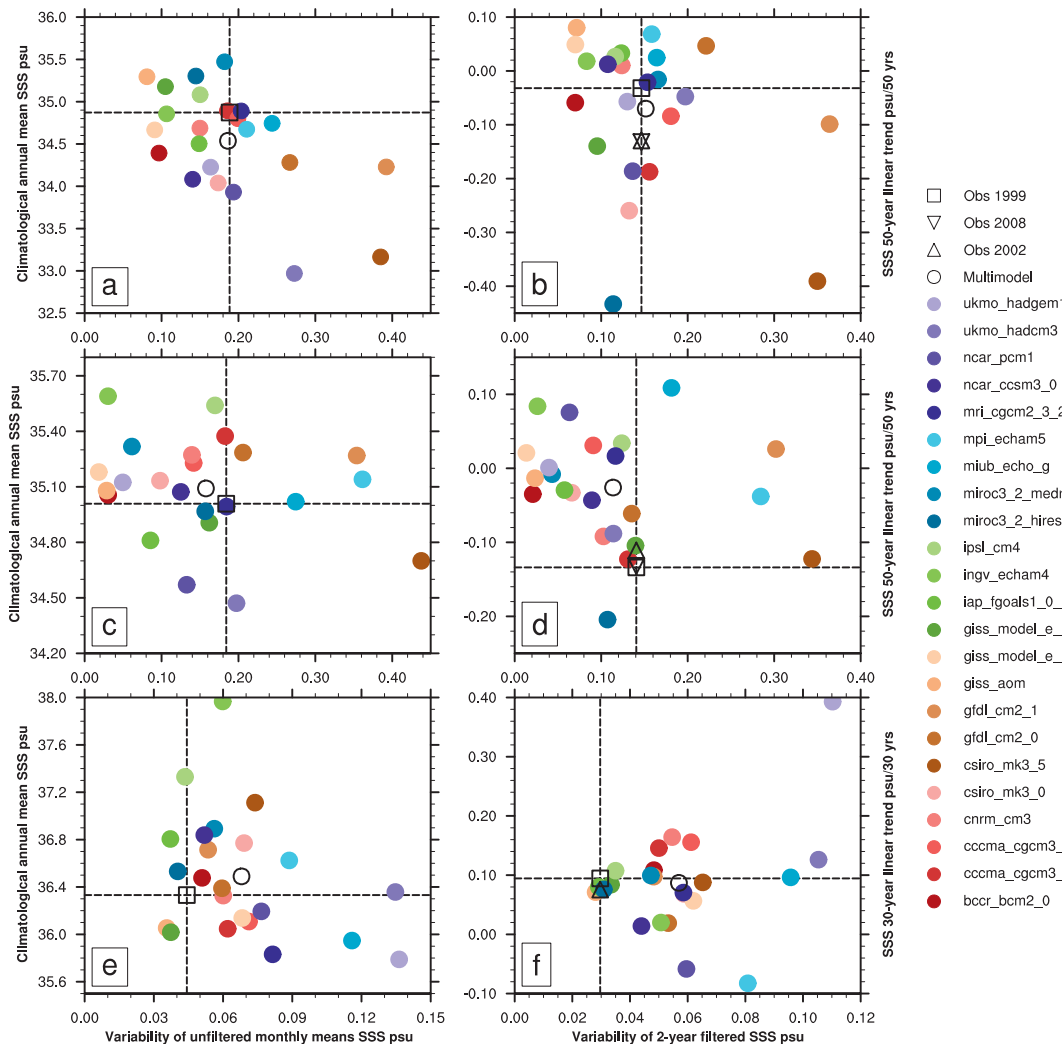


FIG. 4. Observed and simulated twentieth-century SSS mean and variability. Comparison of simulated and observed means, variabilities, and trends for SSS averaged over three tropical regions: (a),(b) western South Pacific ( $20^{\circ}$ – $10^{\circ}$ S,  $160^{\circ}$ E– $180^{\circ}$ ) over the 1950–99 period, (c),(d) western equatorial Pacific ( $5^{\circ}$ S– $5^{\circ}$ N,  $170^{\circ}$ E– $170^{\circ}$ W) over the 1950–99 period, and (e),(f) subtropical North Atlantic ( $15^{\circ}$ – $30^{\circ}$ N,  $70^{\circ}$ – $50^{\circ}$ W) over the 1970–99 period. Scatterplots show the relationships between the temporal standard deviation of unfiltered SSS monthly means anomalies and the climatological annual mean SSSs in (a),(c), and (e), and the temporal standard deviation of 2-yr-filtered SSS monthly mean anomalies and 50-yr (30 yr) linear trends in SSS over 1950 (1970)–1999 for the Pacific (Atlantic) regions in (b), (d), and (f). All standard deviations were estimated from linearly detrended data. The dashed lines indicate the location of the observed values to ease the visual comparison of observed and simulated values. The multimodel average values are indicated by an open circle and the observed values by an open square. In (b),(d), and (f) additional trend observed values (open triangles) over periods with the same initial date but ending in 2002 (both basins) and 2008 (Pacific only) have been added.

a number of models suffer from mean state biases possibly due to inadequate freshwater fluxes. A majority of models underestimate the variability in the South Pacific region likely due to a misrepresentation of the SPCZ mean and shifts related to ENSO or PDO. Interestingly, we find there is sign agreement between the multimodel average and the observed trends. While there is no indication of a systematic bias in low-frequency trend

amplitude for the equatorial Pacific and Atlantic regions, most models have smaller trend amplitudes than are observed (Figs. 4b,d,f). As expected from independent model initialization and internal variability, the modeled Pacific trends show a large spread while most models agree on a salinity increase trend (with different amplitudes) for the Atlantic. The similarity of the observed trend values with different ending dates illustrates that

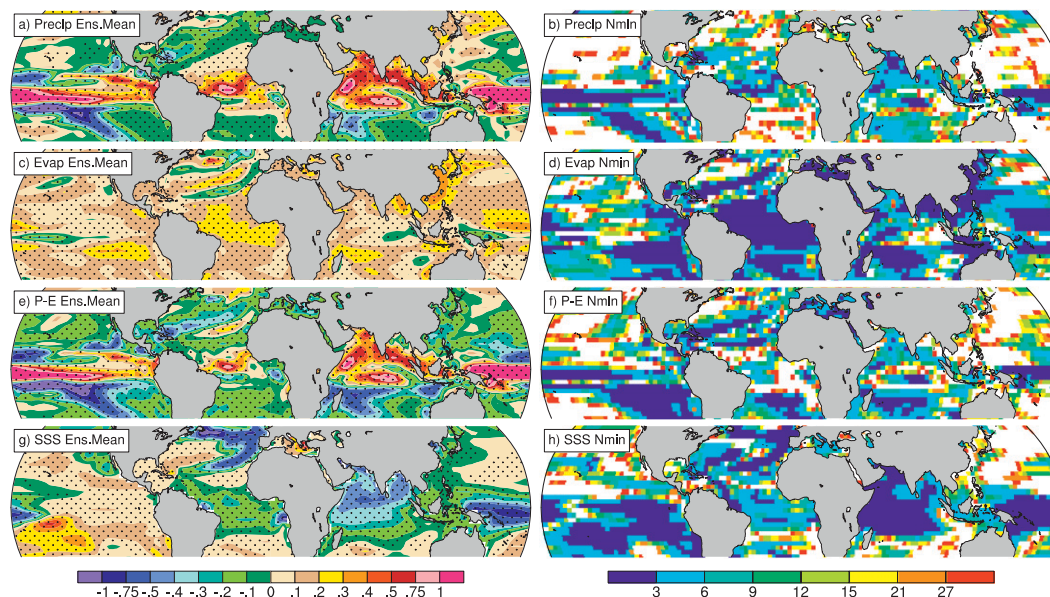


FIG. 5. Signal-to-noise ratio: SSS vs freshwater fluxes. (left) CCSM3 40-member ensemble mean 50-yr linear trend (2011–60) and (right) minimum number of ensemble members needed to detect a nonzero ensemble mean trend (taken as the forced response) at the 5% level at each grid point. Shown are (a),(b) precipitation, (c),(d) evaporation, (e),(f) precipitation minus evaporation, and (g),(h) SSS. In (a),(c),(e), and (g), stippling indicates 50-yr linear trends significantly different from 0 at the 5% confidence level relative to the spread of the 40 50-yr linear trends using a two-sided Student's  $t$  test. In (b),(d),(f), and (h), white areas indicate locations where the 40-member ensemble mean response is not significantly different from 0 at the 5% confidence level. Units in (a),(c), and (e) are  $\text{mm day}^{-1} (50 \text{ yr})^{-1}$  and  $50 \text{ yr}^{-1}$  in (d).

the 1998 ENSO is not strongly impacting the low-frequency trends. Furthermore, the Pacific regions have very similar observed SSS trends while their linearly detrended and 2-yr filtered time series are actually negatively correlated at an interannual time scale [correlation of  $-0.57$  significant at the 1% level using a phase-scrambling bootstrap test accounting for time series autocorrelation; Davison and Hinkley (1997)]. This suggests that the two regional SSS trends may be caused by a common low-frequency mechanism different from those related to ENSO or PDO. Alternatively, this inverse relationship between the different time scales could also indicate a change in the ENSO-related SSS pattern due to instance to an increased frequency of central Pacific El Niño events, which could occur in response to global warming (Yeh et al. 2009).

## 5. Detection of human influences on salinity changes

### a. Signal-to-noise ratio: Salinity versus freshwater fluxes

In studying the human influences on the tropical water cycle, it is relevant to first consider the question of the signal-to-noise ratio between different climate variables.

Here, we compare a simple measure of the signal-to-noise ratio for SSS, freshwater flux, and its different components. We use the 40-member CCSM3 ensemble to assess the signal-to-noise ratios associated with a linear trend diagnostic applied to the different variables of the marine hydrological cycle. Here, the signal refers to the SSS forced response to the anthropogenic forcing and is estimated by the 40-member ensemble mean linear trend. The noise refers to pure internal variability and is estimated by the ensemble dispersion of individual trends. As all members share the same ocean and land initial conditions, it is possible that the ensemble provides a biased low estimate of internal variability. However, previous analysis of the 40-member ensemble has suggested that the memory associated with the ocean initial conditions is lost within less than a decade (Branstator and Teng 2010). To account for this effect in the signal-to-noise ratio analysis, we use annual means and calculate 50-yr linear trends covering the 2011–60 period (similar results are obtained if epoch differences are used). The left panel of Fig. 5 shows the forced response for precipitation; evaporation; total freshwater flux, defined here as precipitation minus evaporation; and sea surface salinity. There is a good spatial correspondence between regions of decreasing salinity with regions with an increased total freshwater flux (and vice versa),

in particular over the western Pacific, the tropical Atlantic, and the Indian Ocean. The SSS change also exhibits a much larger spatial scale pattern than the freshwater fluxes, suggesting the influences of ocean circulation in spreading the freshwater anomaly. Statistical significance of the forced response is reached over a large fraction of the tropical oceanic regions for the four variables. Yet the precipitation and total water flux exhibit a significantly larger area with no significant forced response other than evaporation or surface salinity. This is more clearly seen in the right panel of Fig. 5, which is shown as a simple measure of the signal-to-noise ratio,  $N_{\min}$ , the minimum number of members needed to detect a significant forced response (meaning a nonzero ensemble mean trend). The value of  $N_{\min}$  is computed by inverting the formula for the standard error of the mean:  $N_{\min} \sim 8/(X/\sigma)^2$ , where  $X$  is the ensemble mean trend and  $\sigma$  is the standard deviation of the 40 individual trends (Deser et al. 2012). Only a few members ( $N_{\min} < 3$ ) are needed to detect a significant SSS response over the quasi-majority of the tropical oceans while precipitation and total freshwater flux generally present lower signal-to-noise ratios or no detectable response, in particular over the western Pacific, the tropical Atlantic, and the equatorial Indian Ocean.

#### *b. Observed and internal variability basin-averaged salinity trends*

A first simple approach to the detection question is to assess whether the recent SSS trends are outside of the range due to internal variability alone. Here, we use the CMIP3 control simulations to provide a trend distribution estimate purely based on the internal variability. We deliberately focus on a few integrated SSS signals: the Atlantic and Pacific basin averages and the basins contrast. We consider averages over the entire observation domain for each basin to estimate the mean basin and interbasin contrast SSS time series. We then seek evidence for the rejection of the null hypothesis that the observed basin average and interbasin contrast SSS trends could be due to internal variability alone. As several models exhibit significant nonlinear drifts, we first detrend the control simulations by systematically removing a second-order polynomial trend to each grid-point time series to prevent artificially inflating the width of the distribution. We then calculate the three annual mean time series from each CMIP3 model.

From each model, the distribution of 33-yr trends is obtained by fitting linear trends to all 33-yr overlapping segments (years 1–33, 2–34, ...). We then pool the results from the various models to form a multimodel distribution of unforced trends. The control simulations used here compose a total of 8620 years of unforced

data. This yields 7914 nonindependent samples of overlapping 33-yr linear trends only driven by internal variability. Assuming that the model-generated internal variability is reliable, it is unlikely to be the sole factor of the observed mean Pacific SSS decrease and increase in SSS interbasin contrast, as very few of the unforced trends exceed the observed 1970–2002 trend (Fig. 6). However, the null hypothesis cannot be rejected for the Atlantic where the observed trend is well within internal variability estimates. This is coherent with the sign differences noted previously between recent observed and future simulated Atlantic SSS trends in response to anthropogenic forcing. This lack of evidence for the Atlantic does not rule out the possible detection of a spatial pattern instead of an integrated measure such as a basin average.

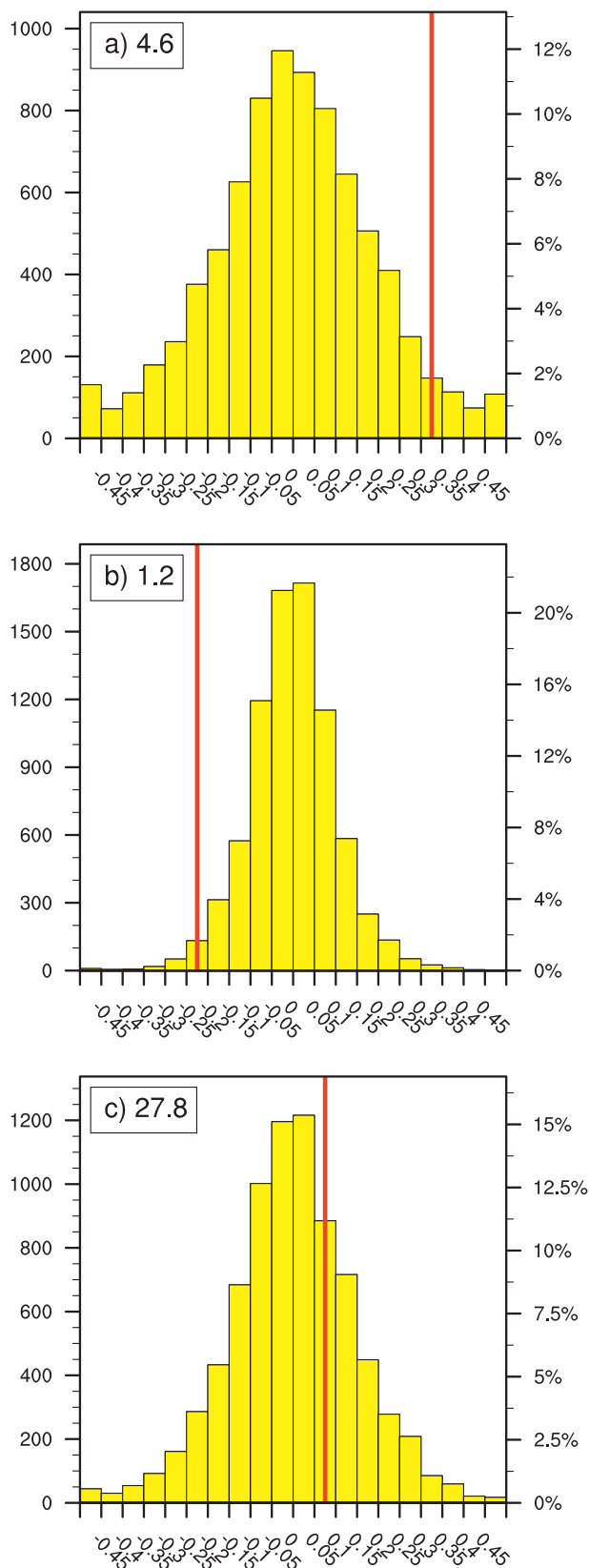
#### *c. The scalability hypothesis*

Our implementation of both detection approaches relies upon the scalability hypothesis, namely saying that the spatial structure of the transient forced response is quasi-invariant with time given a homothetic transformation (note that this hypothesis is likely too severe for very long time scales, as in stabilization experiments, where the slow adjustment of long memory components such as the deep ocean and ice sheets will usually lead to different response patterns). While it may seem to be a rather strong hypothesis, it is possible to test its validity for the twentieth- and twenty-first-century transient integrations using the CMIP3 data. Figure 7 shows that the hypothesis is fairly reasonable for tropical SSS. The centered spatial patterns of the different epoch changes are similar, which is to be expected if the scalability hypothesis holds. We further note that it is likely to be also the case for surface air temperature, as suggested by the simple scaling relating patterns of simulated temperature change between different SRES scenarios (Meehl et al. 2007). Consideration of global SSS maps shows that this assumption is more doubtful for subpolar latitudes due to nonlinearities in particular linked to strong sea ice melting. It is also worth pointing out that the simulated mean tropical Atlantic (averaged over 30°S–50°N) across the twentieth and twenty-first centuries is continuously getting saltier while the Pacific (with the same latitudinal average) is getting fresher.

#### *d. Internal variability and temporal pattern estimation*

We now examine some assumptions of the nature of internal variability in the two detection methodologies. We also describe the rationale behind the choice of the temporal patterns within the TOD framework.

## 1) OPTIMAL FINGERPRINT



The two noise covariance matrices ( $\mathbf{C}_{n1}$  and  $\mathbf{C}_{n2}$ ) are provided by sampling 33-yr linear trends from CMIP3 control simulations with constant preindustrial forcings. As was done previously, we first detrend and then divide each of the CMIP3 control integrations into nonoverlapping 33-yr samples, which yields a set of 254 unforced linear trends of 33-yr duration. We then split the control integrations in such a way that every model is evenly represented in the covariance matrix estimates  $\mathbf{C}_{n1}$  and  $\mathbf{C}_{n2}$ . We want to assess if model uncertainty is biasing our estimates of internal variability as represented by the  $\mathbf{C}_{n1}$  and  $\mathbf{C}_{n2}$  matrices. We compare the main modes of the SSS 33-yr linear trend variability due to pure internal variability deduced from the 40-member CCSM3 ensemble with those estimated from the  $\mathbf{C}_{n1}$  and  $\mathbf{C}_{n2}$  matrices. We use the period 2021–53 to calculate the set of 40 trends from each individual CCSM3 member. An empirical orthogonal function (EOF) analysis based on an area-weighted covariance matrix is then computed on the set of 40 trend maps separately for each ocean basin. We note that due to the removal of the ensemble mean trend in the EOF analysis, the first EOF modes depict the dominant patterns of the internal variability component of the trends. A similar EOF analysis is also applied to the set of 254 SSS 33-yr linear trends from the  $\mathbf{C}_{n1}$  and  $\mathbf{C}_{n2}$  matrices. The leading modes have similar structures in the two sets for each basin, suggesting that the model uncertainty is not strongly biasing our noise estimates (Fig. 8). The spatial structure of the Pacific dominant noise pattern is reminiscent of meridional (zonal) displacements of the ITCZ (SPCZ) while that of the Atlantic shows a dipolar structure between the tropical and subtropical regions and the midlatitudes.

FIG. 6. Distribution of unforced basin-average SSS trends: comparison of observed SSS trends ( $\text{century}^{-1}$ ) with model simulations of unforced trends for (a) the interbasin SSS contrast, (b) the basin-averaged (30°S–30°N) Pacific SSS, and (c) the basin-averaged (30°S–50°N) Atlantic SSS. The red line indicates the observed values. The sampling distribution of the unforced 33-yr SSS trends was calculated as described in the text. The left and right vertical axes indicate the numbers and percentages of trend occurrences for each bin, respectively. The number given in the top-left box is the percentage of simulated trends smaller (Pacific) or greater (interbasin or Atlantic) than the observed value. Using a two-sided test leads to  $p$  values that are twice as large (0.024 and 0.092 for the Pacific and interbasin trends).



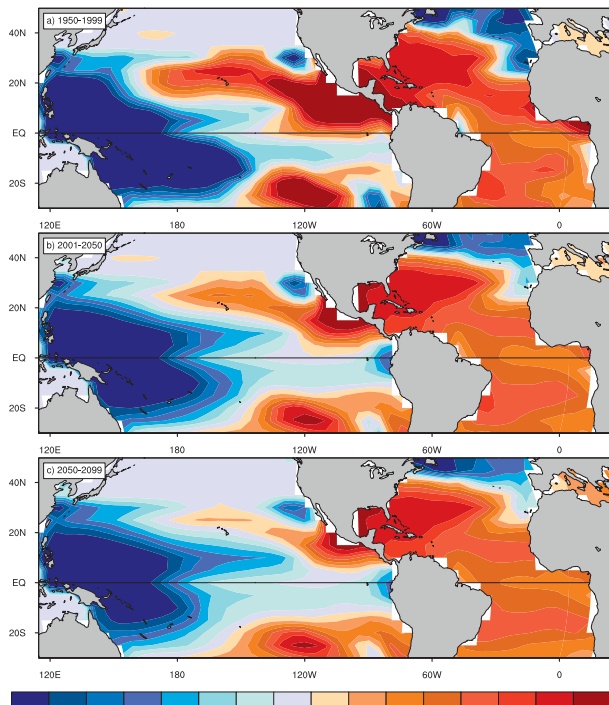


FIG. 7. Validity of the scalability hypothesis: time evolution of the multimodel climate change spatial distribution defined as the centered pattern difference between three 50-yr periods: (a) 1950–99, (b) 2001–49, and (c) 2050–99 and 1900–50. The Atlantic and Pacific SSS changes are further normalized by the mean SSS change averaged over the Atlantic and Pacific (30°S–50°N) domains, respectively.

## 2) TEMPORAL DETECTION METHOD

Internal variability in the TOD method is assumed to have an AR(1) structure. It is first required to estimate the persistence parameter, that is, the  $\alpha$  parameter. The choice of the  $\alpha$  value is addressed by using the CMIP3 control simulations and fitting AR(1) processes to grid-point annual mean SSS for each model. Figure 9a shows the spatial distribution of the ensemble mean value for  $\alpha$ . Low values are present in the deep tropics while the subtropics and midlatitudes are characterized by higher values (thus greater persistence). However, and as mentioned in section 3b, TOD is applied to a small set of spatial domains (Fig. 9a) rather than at the gridpoint scale. Table 2 provides a multimodel mean estimate of  $\alpha$  over the 12 regional domains involved into the analysis. Based on these results, we choose to use the  $\alpha = 0.7$  value for the analysis over the full domain. The next question is whether such a parametric model for internal climate variability allows us to properly account for the multidecadal variability or for a potential long-range memory effect in SSS variability. To address this question, we verify that over 33-yr periods (the period used

in our full tropical analysis), the internal variability from the CMIP3 control simulations is consistent with the one assumed in the TOD method. This is shown in Fig. 9b, by estimating the distribution of the  $p$  values obtained by applying the TOD method to the full set of overlapping 33-yr periods taken from CMIP3 control simulations. If the TOD's assumptions are appropriate, this distribution is expected to be uniform (within sampling errors) over the  $[0, 1]$  interval. If the internal variability is underestimated (overestimated) in the TOD method, then the probability to find small  $p$  values will be increased (decreased). Figure 9b shows that our assumption is fairly reasonable, with a multimodel distribution that is close to uniform. However, it must be noted that this multimodel mean result is somewhat model dependent, with some individual models simulating a lower (higher) persistence and then a conservative (permissive) testing procedure with the  $\alpha = 0.7$  value (not shown). Similar diagnoses have been performed to estimate appropriate values of  $\alpha$  when the analysis is applied over different domains (such as those used in Fig. 10; see legend).

We now examine the choice of the temporal pattern  $\mu_t$ . As suggested by recent studies (Palmer et al. 2007, 2009), we choose the mean temperature between the surface and the 14°C isotherm as our proxy for upper-ocean warming (Tiso14). Using Tiso14 filters out part of the internal variability from climate simulations and better captures the externally forced component of tropical and subtropical upper-ocean temperature evolution over the twentieth century. We then calculate  $\mu_t$  using simulated annual mean values of Tiso14 from the CMIP3 twentieth-century integrations with anthropogenic forcing only and impose a smoothing constraint using four degrees of freedom in the minimization procedure, as in Ribes et al. 2010 (Fig. 10). The smoothness constraint allows us to filter out a large part of the remaining internal variability, thereby further increasing the signal-to-noise ratio. The patterns have similar shapes, namely a very slow increase for the first 70 yr followed by a sharp rise over the last three decades. Note that the detection test does not depend on the mean or amplitude of  $\mu_t$  but only on its shape.

### e. Detection results

We estimate the effects of anthropogenic forcing on observed SSS evolution first over both basins together and then separately. When considering the full Atlantic-Pacific domain, the SSS response to anthropogenic forcing is detected using the OF method (Fig. 11a;  $p$  value  $\approx 10\%$ ), but not detected with the TOD method (Fig. 11b;  $p$  value  $\geq 20\%$ ). Such a difference may be explained as follows: in addition to the temporal information used



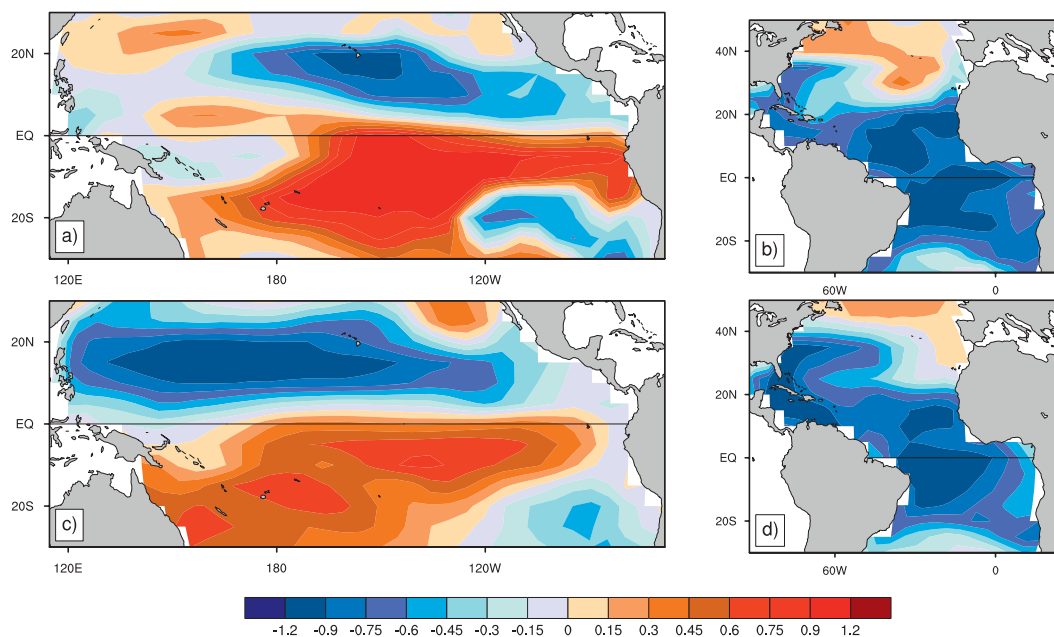


FIG. 8. Model uncertainty and internal variability leading modes: the leading EOF of SSS 33-yr trends (50 yr<sup>-1</sup>) from (a),(b) the 40-member CCSM3 ensemble and (c),(d) the 254-member trend set from the control CMIP3 simulations ensemble. The analysis is done separately for the two basins.

in both methods, OF also uses the expected spatial pattern of change, which is very similar to the observed trends (see Fig. 3). In such a case, OF does a better job of discriminating between anthropogenic change and internal variability, while TOD is expected to be efficient when the expected spatial pattern of change is inaccurately estimated. This result from the OF method suggests that anthropogenic forcing acts to increase the interbasin SSS contrast, which is confirmed by the TOD results when TOD is applied to only two large boxes covering the Pacific and Atlantic basins (Figs. 11c,d). These results also confirm the simple linear trend distribution analysis (Fig. 6). An anthropogenic influence upon Pacific SSS alone can also be detected with both approaches (Figs. 11a,e). Furthermore, residual and internal variability are found to be consistent within the OF framework, and the two-basin and Pacific scaling factors are found to be consistent with one another, indicating consistency between the simulations and observations (Allen and Tett 1999; Allen et al. 2006). The temporal approach suggests that the anthropogenic signal is robustly detected in the Pacific starting in the late 1990s. Applying the temporal approach to individual tropical Pacific regions shows robust detection in the whole western Pacific fresh pool (regions 1–3; see Fig. 9a) and no detection in other regions. No anthropogenic forcing response can yet be robustly detected in the observed SSS for the full Atlantic domain (Figs. 11a,f). The scaling factor 5%–95% uncertainty range includes 0, meaning

that one cannot reject the null hypothesis. Further insights can be gained using the temporal approach. While  $p$  values are below the 10% confidence threshold during the 1980s, that is no longer the case for the most recent period where they are very close to or above it. This holds with most temporal patterns due to their similarity over the 1970–2002 period. It is worth noting that the third climate configuration of the Met Office Unified Model (HadCM3) temporal pattern (which is the lowest value in the model range in Fig. 11f) is detected at the end of the period although not very strongly, in agreement with a previous detection study (Stott et al. 2008). However, we find that an anthropogenic forcing response can be detected for the subtropical North Atlantic alone (Figs. 11a,g) or in combination with the Pacific while detection fails for the other regions when taken separately (not shown). These results are coherent with our previous findings regarding the influence of internal decadal variability upon the recent observed Atlantic trends and the spatial differences between the observed trends and projected twenty-first-century changes.

## 6. Summary and discussion

Together with previous related studies on atmospheric water vapor (Santer et al. 2007) and surface humidity (Willett et al. 2007), our findings suggest that there is an emerging anthropogenic signal in the recent evolution of

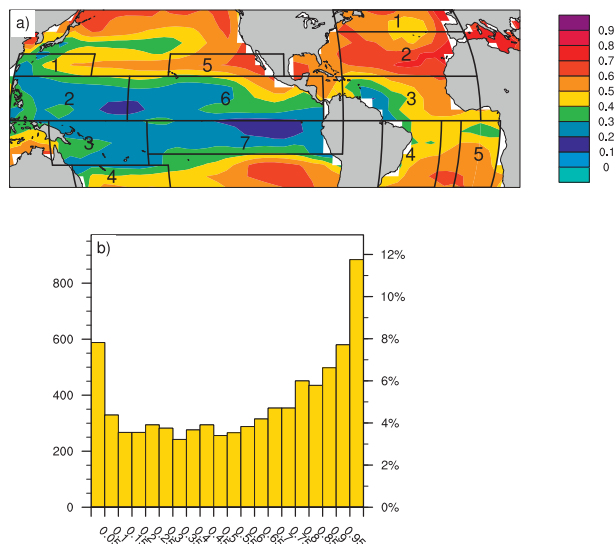


FIG. 9. Temporal detection method. (a) Mean (across the 22 CMIP3 models) value of the coefficient  $\alpha$  of the AR(1) process fitted to annual mean SSS grid point time series derived from the CMIP3 control integrations. The black boxes indicate regions where annual mean SSS values were spatially averaged before being used in the detection test of the temporal detection approach. (b) Distribution of the  $p$  values provided by the TOD procedure when applied to all overlapping 33-yr periods from CMIP3 control simulations, with the same parameters ( $\alpha$  and  $\mu_t$ ) as for the observations.

the tropical marine hydrological cycle. The main results of our study can be summarized as follow:

- (i) Near-surface salinity observations during recent decades show a tropical Pacific freshening and Atlantic salinity increase, which suggests an enhanced interbasin salinity contrast. The mean-basin SSS 33-yr trends are significantly different from what is expected from internal variability alone for the Pacific and interbasin contrast while they are indistinguishable for the Atlantic.
- (ii) Twenty-first-century projected SSS changes in response to anthropogenic forcing display a freshening of the western tropical Pacific and a widespread tropical and subtropical Atlantic salinity increase. Two regions have changes that contrast with the mean basin trend: the southeast Pacific is getting saltier while the midlatitude Atlantic is getting fresher.
- (iii) Anthropogenic forcing contributed significantly to the observed tropical Pacific freshening and to the subtropical Atlantic salinity increase. The robustness of the detection is suggested by the consistency of the results from two different methodologies. This enhanced SSS contrast is consistent with the projected enhanced moisture transport from the tropical

TABLE 2. Mean (across the 21 CMIP3 models) value of the coefficient  $\alpha$  of the AR(1) process fitted to the time series of the annual mean SSS averaged over the 12 regions defined in Fig. 9a, derived from the CMIP3 control integrations.

No.	Atlantic	Pacific
1	0.73	0.50
2	0.80	0.55
3	0.64	0.54
4	0.61	0.47
5	0.68	0.67
6		0.55
7		0.44

North Atlantic to the Pacific under increased greenhouse gas forcing (Richter and Xie 2010).

- (iv) The spatial patterns of observed and projected SSS changes reflect the rich get richer paradigm (Held and Soden 2006). The Pacific fresh pool and the subtropical North Atlantic experience the largest SSS changes, in agreement with projections of freshwater atmospheric forcing (Seager et al. 2010). The observed and projected extension of the western Pacific fresh pool likely involves additional influences from ocean dynamics, including changes in the South Equatorial Current, divergent Ekman flows, and poleward western boundary currents (Huang et al. 2005).

These results may have important implications on the future evolution of climate variability modes. A fresher and warmer western Pacific pool may potentially lead to changes in mixed layer depth and barrier layer thickness although quantitative estimates require consideration of additional factors such as mixing induced by wind changes (Carton et al. 2008). These potential changes in

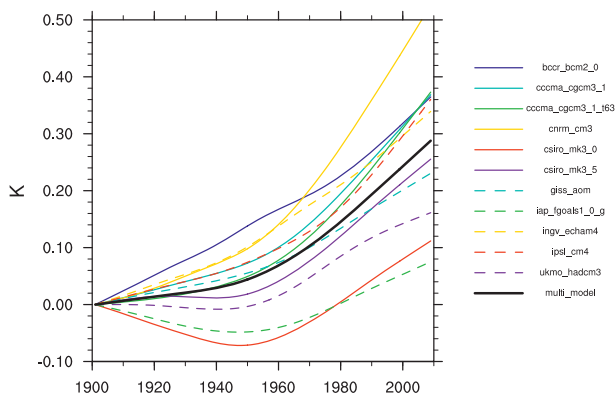


FIG. 10. Estimates of the temporal patterns  $\mu_t$  (in K) derived from annual mean Tiso14 from individual models using the CMIP3 twentieth- and twenty-first integrations with 1900 as a time reference for all curves. We only consider models with only anthropogenic forcings prescribed during the twentieth century (no natural forcings). The black line shows the multimodel temporal pattern.

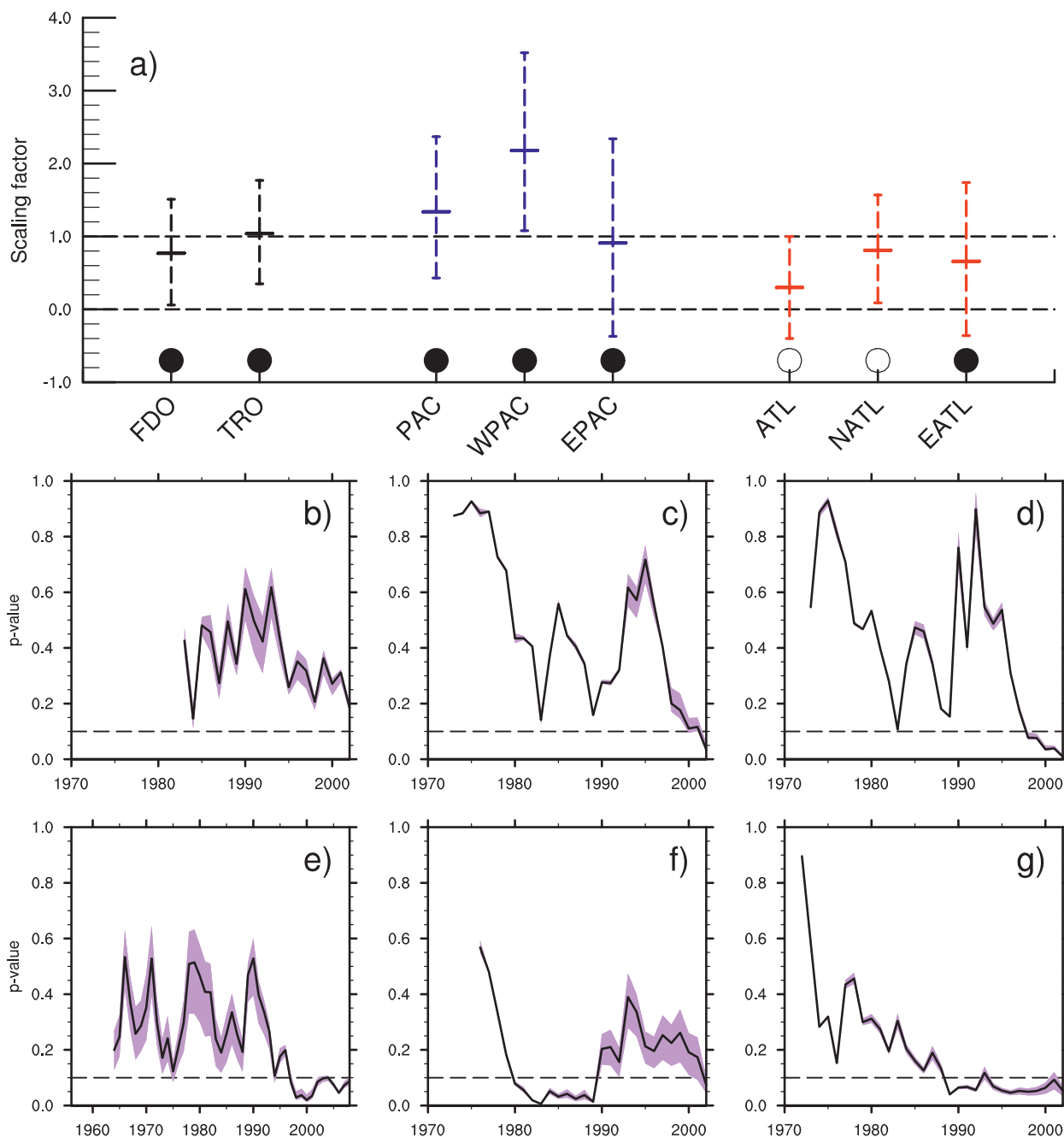


FIG. 11. (a) Results from detection analysis of SSS 33-yr trends. Scaling factor best estimates and their 5%–95% uncertainty ranges are given from one-signal (anthropogenic) fingerprint analysis using the TLS approach for different regions: the full domain (FDO), the tropics (TRO, 30°S–30°N), Pacific (PAC, 30°S–30°N), western Pacific (WPAC, 120°E–160°W), eastern Pacific (EPAC, 160°–80°W), Atlantic (ATL, 30°S–50°N), subtropical North Atlantic (NATL, 20°–40°N), and equatorial Atlantic (EATL, 20°S–20°N). Filled black dots indicate when the residual consistency test passes with a truncation of 16, whereas open circles indicate a needed higher truncation to pass the test. (b) Results from the temporal detection analysis applied to the observed annual mean SSS for the entire Pacific and Atlantic regions (with  $\alpha = 0.7$ ). (c) Results from two basin-scale regions, over the entire Atlantic and Pacific domains (with  $\alpha = 0.75$ ). (d) As in (c), but with centered data. (e) For the entire Pacific region ( $\alpha = 0.65$ ). (f) For the entire Atlantic region ( $\alpha = 0.75$ ). (g) For the subtropical Atlantic (region 2) only ( $\alpha = 0.8$ ). In (b)–(g), the statistical test p value is shown for the multimodel temporal pattern by a solid black line. Shading shows the minimum and maximum of the p value range of all model temporal patterns. The dashed black line indicates the 10% significance threshold and rejection of  $H_0$  for a given model is achieved when the associated curve is below this line. Note that the time axis differs between the Pacific only and other regions.

upper-ocean stratification may influence Pacific variability modes such as ENSO. A saltier subtropical Atlantic Ocean suggests that increasing northward salinity transport may influence the meridional overturning circulation (MOC) multidecadal variability and compete with an externally forced or internal variability high-latitude enhanced freshwater input (Pardaens et al. 2008). As the time scales of these processes may be different and vary among climate models, there is still a large uncertainty as to the future MOC evolution. Our study also suggests that sea surface salinity is a good candidate for further detection and attribution of an anthropogenic fingerprint in the oceans and marine hydrological cycle changes, including the Indian Ocean. The Argo salinity program as well as ongoing (Soil Moisture and Ocean Salinity, SMOS) and future (the National Aeronautics and Space Administration's planned AQUARIUS project) remote sensing measurements will be key players in documenting current and future salinity variations on a global scale (Lagerloef et al. 2010). However, one of the forthcoming challenges will be to design a multidecadal salinity dataset that minimizes the effects of spatiotemporal nonhomogeneity associated with the use of the different data sources. This highlights the strong need for the continuation and extension of existing long time series based on weather ships and repeat hydrographic lines, which currently are the only data sources with the capability of providing observed records of multidecadal salinity changes.

**Acknowledgments.** The authors thank Jim Hurrell for his comments and suggestions and Adam Phillips for helping with the model datasets, as well as graphic tricks. This work was completed when L.T. was on sabbatical leave at NCAR. He thanks CGD/NCAR for providing partial support and Jim Hurrell and Bill Large for hosting him in 2010. Thanks also go to C. Deser and J. Fasullo for stimulating discussions. The authors would also like to thank the two anonymous reviewers for their useful feedback and suggestions, which have greatly helped to improve the manuscript. This work was supported by the European Community via the seventh framework program COMBINE project under Contract EVK2-CT-2001-00156. Some of the research was also conducted with the support of the National Center for Atmospheric Research, which is sponsored by the National Science Foundation. We acknowledge the following modeling groups for their roles in making the WCRP CMIP3 dataset available: the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modeling. Support of this dataset is provided by the Office of Science, U.S. Department of Energy.

## APPENDIX

### Optimal Fingerprint Methodology

#### *a. Signal coherency in multimodel estimates*

Figure A1 shows the signal estimates for the CMIP3 models not included in Fig. 3. Although there are important intermodel differences as far as regional features of the simulated changes, the main large-scale changes (freshening of the western Pacific and salinity increase in the tropical and subtropical Atlantic) are present in all models apart from two of the Goddard Institute for Space Studies (GISS) models.

#### *b. Choice of the truncation space*

Guidance on the number  $p$  of EOFs to retain in the optimal fingerprint analysis can be obtained by performing a residual consistency test (Allen and Tett 1999). As mentioned in Allen et al. (2006), there is always ambiguity when interpreting the results from these consistency tests. We thus do not use them as an essential constraint on our analysis. We prefer focusing on the signal detection and then we use the consistency tests to specify the range over which uncertainty estimates are likely to be biased. Carrying out the detection analysis on the full domain gives robust and positive detection results when  $p$  is greater than 12. Figure A2 shows that the observed and model variability are consistent at the 10% significance level within the  $p = 13$ –16 EOF range. More than 80% (90%) of the observation (signal and noise) variance is retained at the truncation  $p = 16$ , which is then used in detection tests for the other domains. At this truncation, the analysis over almost all domains shows consistency between the simulated and observed internal variabilities. Two domains, the full and subtropical North Atlantic, needed a larger truncation ( $p = 42$ ) in order to pass the test, suggesting that more modes are needed in the North Atlantic to represent a large amount of SSS internal variability as the major low-frequency modes (ENSO, NAO) only explain a small part of the total variance (Reverdin et al. 2007).

#### *c. Model uncertainty bias in signal estimates*

Use of a single-model ensemble mean to derive the forced signal has often been adopted in classical detection and attribution studies. The simple rationale behind it is that it allows us to reduce the bias in the signal estimate due to internal variability. However, even averaging across a large set of model members will not lead to a fully unbiased estimate of the forced signal due to the always finite ensemble size. A

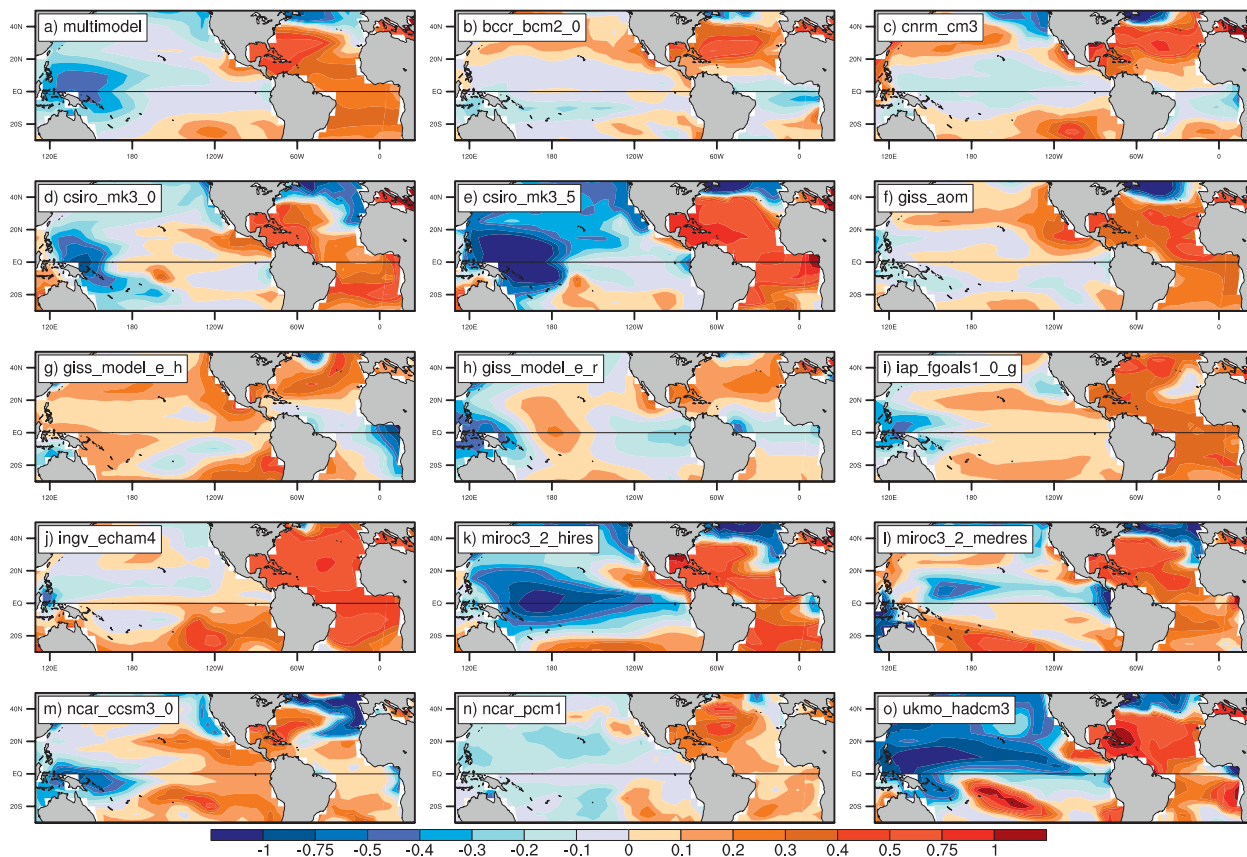


FIG. A1. Twenty-first-century SSS changes. (a) Full CMIP3 multimodel average SSS change ( $\text{century}^{-1}$ ) between the end of the twenty-first and twentieth centuries (2070–99 and 1970–99). (b)–(o) Simulated changes as above for CMIP3 individual models not included in Fig. 3.

conservative and standard approach is thus to use the total least squares (TLS) algorithm, which accounts for uncertainty in the signal estimates due to internal variability. Even when the ensemble size is large (say greater than four to six), use of TLS is still advised as it gives a more realistic estimate of the upper uncertainty range for the scaling coefficients. Recent studies have also suggested that the use of a multimodel ensemble is also beneficial to the specification of various climate responses to anthropogenic forcing due to the reduced multimodel bias compared to individual model ones (Gillett et al. 2002). However, this also introduces further complications as the signal estimate is then also influenced by the structural uncertainty associated with the model error in the response patterns. A slight modification of the total least squares algorithm (the error-in-variables variant) has recently been suggested to account for the two sources of bias in the signal estimates (Huntingford et al. 2006). Nevertheless, the standard total least squares approach can still be used based on the assumption that the covariance structure of the structural (model error) uncertainty is reasonably similar to that of the internal variability. We thus implicitly

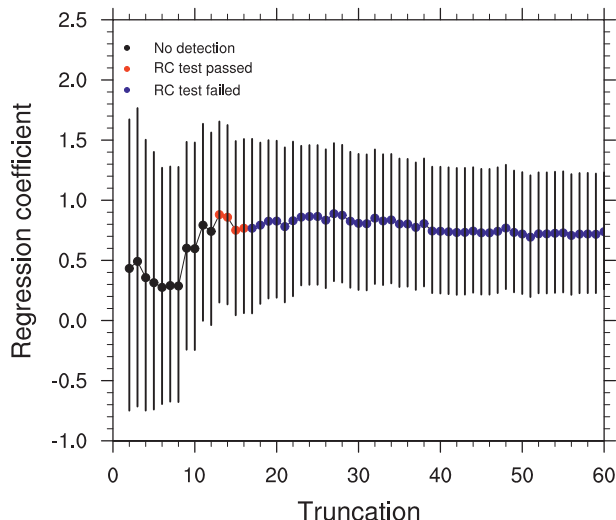


FIG. A2. Scaling factor best estimates (dots) and their 5%–95% uncertainty ranges (bars) are given from one-signal (anthropogenic) fingerprint analysis using the TLS approach for the full domain as a function of EOF truncation. Black dots indicate no detection while red and blue dots indicate a positive detection result. Simulated and observed residual variabilities are consistent for  $p = 13$ –16 (red dots) while simulated internal variability is overestimated for  $p$  greater than 16 (blue dots).



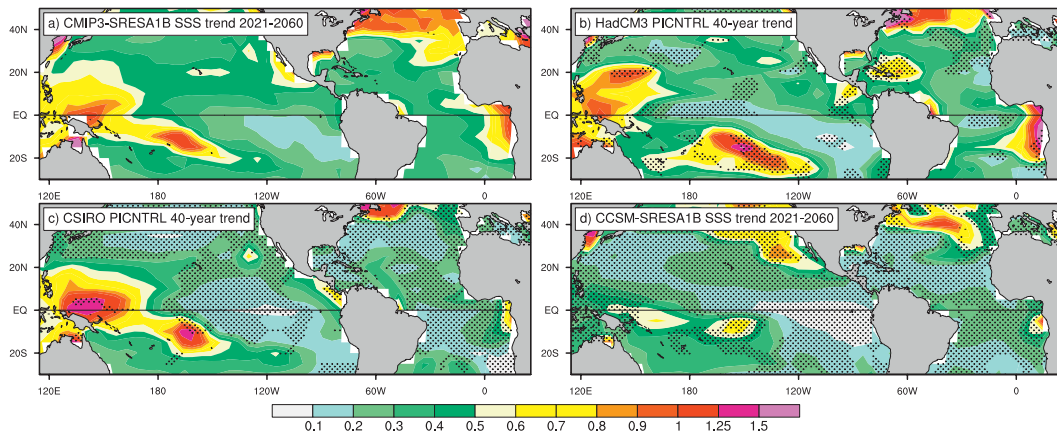


FIG. A3. Model uncertainties in signal and internal variabilities. Standard deviation of 40-yr SSS trends from (a) CMIP3 SRES A1B simulations for the 2021–60 period. (b) The 40-member CCSM3 ensemble for the 2021–60 period. (c) The full set of nonoverlapping 40-yr trends from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) control simulation. (d) The full set of nonoverlapping 40-yr trends from the HadCM3 control simulation. Variability in (a) comes a priori from both internal variability and model uncertainty while it is only due to the former for (b)–(d). Stippling in (b)–(d) indicates regions where variances differ from that in (a) using a two-sided F test at the 10% significance level.

assume that intermodel fluctuations about a common mean signal are mostly due to purely internal variability and not to different and plausible representations of the response to anthropogenic forcing. To ensure that this is a reasonable hypothesis, we compare the intermodel variability in the CMIP3 model-derived signals with various pure internal variability estimates and find no obvious and systematic differences (Fig. A3). The regions of intermodel maximum variance always have a counterpart in one or several of the internal variability estimates. Furthermore, the signal intermodel variability modes are closely related to the main modes of internal variability (Fig. 8).

#### REFERENCES

- Allen, M. R., and S. F. B. Tett, 1999: Checking for model consistency in optimal fingerprinting. *Climate Dyn.*, **15**, 419–434.
- , and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 224–232, doi:10.1038/nature01092.
- , and P. A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting. Part I: Theory. *Climate Dyn.*, **21**, 477–491, doi:10.1007/s00382-003-0313-9.
- , and Coauthors, 2006: Quantifying anthropogenic influence on recent near-surface temperature change. *Surv. Geophys.*, **27**, 491–544.
- Bindoff, N. L., and Coauthors, 2007: Observations: Oceanic climate change and sea level. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 385–432.
- Boyer, T. P., S. Levitus, J. L. Antonov, R. A. Locarnini, and H. E. Garcia, 2005: Linear trends in salinity for the World Ocean, 1955–1998. *Geophys. Res. Lett.*, **32**, L01604, doi:10.1029/2004GL021791.
- Branstator, G., and H. Teng, 2010: Two limits of initial-value decadal predictability in a CGCM. *J. Climate*, **23**, 6292–6311.
- Carton, J. A., S. A. Grodsky, and H. Liu, 2008: Variability of the oceanic mixed layer, 1960–2004. *J. Climate*, **21**, 1029–1047.
- Cravatte, S., T. Delcroix, D. Zhang, M. J. McPhaden, and J. Leloup, 2009: Observed freshening and warming of the western Pacific warm pool. *Climate Dyn.*, **33**, 565–589, doi:10.1007/s00382-009-0526-7.
- Curry, R., B. Dickson, and I. Yashayaev, 2003: A change in the freshwater balance of the Atlantic Ocean over the past four decades. *Nature*, **426**, 826–829, doi:10.1038/nature02206.
- Dai, A., 2010: Drought under global warming: A review. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 45–65, doi:10.1002/wcc.81.
- Davison, A. C., and D. V. Hinkley, 1997: *Bootstrap Methods and Their Application*. Cambridge University Press, 582 pp.
- Delcroix, T., 1998: Observed surface oceanic and atmospheric variability in the tropical Pacific at seasonal and ENSO time scales: A tentative overview. *J. Geophys. Res.*, **103**, 18 611–18 633.
- , S. Cravatte, and M. J. McPhaden, 2007: Decadal variations and trends in tropical Pacific sea surface salinity since 1970. *J. Geophys. Res.*, **112**, C03012, doi:10.1029/2006JC003801.
- , G. Alory, S. Cravatte, T. Corrège, and M. McPhaden, 2011: A gridded sea surface salinity dataset for the tropical Pacific with sample applications (1950–2008). *Deep-Sea Res. I*, **58**, 38–48.
- Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, doi:10.1007/s00382-010-0977-x, in press.
- Durack, P., and S. Wijffels, 2010: 50-year trends in global ocean salinities and their relationship to broad-scale warming. *J. Climate*, **23**, 4342–4362.
- Gillett, N. P., and Coauthors, 2002: Detecting anthropogenic influence with a multi-model ensemble. *Geophys. Res. Lett.*, **29**, 1970, doi:10.1029/2002GL015836.
- Gordon, A. L., and C. F. Giulivi, 2008: Sea surface salinity trends over fifty years within the subtropical north Atlantic. *Oceanography*, **21**, 20–29.

- Grodsky, S. A., J. A. Carton, and F. M. Bingham, 2006: Low frequency variation of sea surface salinity in the tropical Atlantic. *Geophys. Res. Lett.*, **33**, L14604, doi:10.1029/2006GL026426.
- Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 663–745.
- Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699.
- Hénoq, C., J. Boutin, G. Reverdin, F. Petitcolin, S. Arnault, and P. Lattes, 2010: Vertical variability of near-surface salinity in the tropics: Consequences for L-band radiometer calibration and validation. *J. Atmos. Oceanic Technol.*, **27**, 192–209.
- Hosoda, S., T. Suga, N. Shikama, and K. Mizuno, 2009: Global surface layer salinity change detected by Argo and its implication for hydrological cycle intensification. *J. Oceanogr.*, **65**, 579–586.
- Huang, B. Y., V. M. Mehta, and N. Schneider, 2005: Oceanic response to idealized net atmospheric freshwater in the Pacific at the decadal time scale. *J. Phys. Oceanogr.*, **35**, 2467–2486.
- Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu, 2009: Improving the global precipitation record: GPCP version 2.1. *Geophys. Res. Lett.*, **36**, L17808, doi:10.1029/2009GL040000.
- Huntingford, C., P. A. Stott, M. R. Allen, and F. H. Lambert, 2006: Incorporating model uncertainty into attribution of observed temperature change. *Geophys. Res. Lett.*, **33**, L05710, doi:10.1029/2005GL024831.
- Lagerloef, G., and Coauthors, 2010: Resolving the global surface salinity field and variations by blending satellite and in situ observations. *Proc. OceanObs'09: Sustained Ocean Observations and Information for Society*, Vol. 2, Venice, Italy, ESA Publ. WPP-306, doi:10.5270/OceanObs09.cwp.51.
- Meehl, G. A., and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 747–845.
- Millero, F. J., 1993: What is PSU? *Oceanography*, **6** (3), 67.
- Min, S. K., X. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more intense precipitation extreme. *Nature*, **470**, 378–381.
- Palmer, M. D., K. Haines, S. F. B. Tett, and T. J. Ansell, 2007: Isolating the signal of ocean global warming. *Geophys. Res. Lett.*, **34**, L23610, doi:10.1029/2007GL031712.
- , S. A. Good, K. Haines, N. A. Rayner, and P. A. Stott, 2009: A new perspective on warming of the global oceans. *Geophys. Res. Lett.*, **36**, L20709, doi:10.1029/2009GL039491.
- Pardaens, A., M. Vellinga, P. Wu, and B. Ingleby, 2008: Large-scale Atlantic salinity changes over the last half-century: A model-observation comparison. *J. Climate*, **21**, 1698–1720.
- Reverdin, G., E. Kestenare, C. Frankignoul, and T. Delcroix, 2007: Surface salinity in the Atlantic Ocean (30°S–50°N). *Prog. Oceanogr.*, **73**, 311–340, doi:10.1016/j.pocean.2006.11.004.
- Ribes, A., J. M. Azais, and S. Planton, 2010: A method for regional climate change detection using smooth temporal patterns. *Climate Dyn.*, **35**, 391–406, doi:10.1007/s00382-009-0670-0.
- Richter, I., and S.-P. Xie, 2008: Muted precipitation increase in global warming simulations: A surface evaporation perspective. *J. Geophys. Res.*, **113**, D24118, doi:10.1029/2008JD010561.
- , and —, 2010: Moisture transport from the Atlantic to the Pacific basin and its response to North Atlantic cooling and global warming. *Climate Dyn.*, **35**, 551–566.
- Santer, B. D., and Coauthors, 2007: Identification of human-induced changes in atmospheric moisture content. *Proc. Natl. Acad. Sci. USA*, **104**, 15 248–15 253.
- , and Coauthors, 2009: Incorporating model quality information in climate change detection and attribution studies. *Proc. Natl. Acad. Sci. USA*, **106**, 14 778–14 783.
- Schmitt, R. W., 2008: Salinity and the global water cycle. *Oceanography*, **21**, 12–19.
- Seager, R., N. Naik, and G. A. Vecchi, 2010: Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming. *J. Climate*, **23**, 4651–4668.
- Stephens, G. L., and T. D. Ellis, 2008: Controls of global-mean precipitation increases in global warming GCM experiments. *J. Climate*, **21**, 6141–6155.
- Stott, P., R. T. Sutton, and D. S. Smith, 2008: Detection and attribution of Atlantic salinity changes. *Geophys. Res. Lett.*, **35**, L21702, doi:10.1029/2008GL035874.
- Trenberth, K. E., 2011: Changes in precipitation with climate change. *Climate Res.*, **47**, 123–138, doi:10.3354/cr00953.
- Von Schuckmann, K., F. Gaillard, and P. Y. Le Traon, 2009: Global hydrographic variability patterns during 2003–2008. *J. Geophys. Res.*, **114**, C09007, doi:10.1029/2008JC005237.
- Willett, K. W., N. P. Gillett, P. D. Jones, and P. W. Thorne, 2007: Attribution of observed humidity changes to human influence. *Nature*, **449**, 710–712.
- Xie, S. P., C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg, 2010: Global warming pattern formation: sea surface temperature and rainfall. *J. Climate*, **23**, 966–986.
- Yeh, S. W., J. S. Kug, B. Dewitte, M. H. Kwon, B. Kirtman, and F. F. Jin, 2009: El Niño in a changing climate. *Nature*, **461**, 511–514, doi:10.1038/nature08316.
- Yu, L., and R. A. Weller, 2007: Objectively analyzed air–sea heat fluxes (OAFlux) for the global ocean. *Bull. Amer. Meteor. Soc.*, **88**, 527–539.
- Zhang, X. B., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa, 2007: Detection of human influence on twentieth-century precipitation trends. *Nature*, **448**, 461–464.