Zhu Zhiwei (Orcid ID: 0000-0002-4885-1299)

Internal Atmospheric Variability of Net Surface Heat Flux in Reanalyses and CMIP5 AMIP Simulations

Hua Chen[1], Edwin K Schneider[2] and Zhiwei Zhu[1,*]

1 Key Laboratory of Meteorological Disaster, Ministry of Education (KLME)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD)/Joint International Research Laboratory of Climate and Environment Change (ILCEC), Nanjing University of Information Science and Technology, Nanjing, China

2 Department of Atmospheric Oceanic and Earth Sciences, George Mason University, Fairfax, Virginia, U.S.

*Corresponding author: Zhiwei Zhu, Nanjing University of Information Science and Technology, 219 Ningliu Road, Meteorology Bldg. Jiangsu 210044, China.

E-mail: zwz@nuist.edu.cn

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/joc.7232

Abstract

The internal atmospheric variability (IAV) of the net surface heat flux (NHF) in the observed $20^{th}/21^{st}$ century atmosphere is estimated as the residual after removing the SST and externally forced atmospheric response derived from the four atmospheric models of the Atmospheric Model Intercomparison Project (AMIP) simulations under phase 5 of the Coupled Model Intercomparison Project (CMIP5). The mean NHF of four atmospheric reanalysis datasets is an estimate of the observed NHF.

Although the AMIP models are forced with the same SST and external forcing, the forced responses differ significantly among AMIP models, suggestive of uncertainty in the models. Besides, the uncertainty of IAV in the reanalyses could also arise from the uncertainty in reanalyses as observations contain errors and reanalysis includes interpolation by models. It is concluded that: 1) The SST/NHF and SST/forced NHF correlations are significantly negative over most of world ocean in the AMIP models, indicating damping of the SST anomalies by the NHF. 2) The IAV of the AMIP models is not correlated with SST, while the positive IAV/SST correlations in the reanalyses suggests the role of IAV in forcing the SST variability in the extra-tropics. 3) The standard deviation (STD) of the IAV of AMIP models is indistinguishable from that of the mean reanalysis over a majority of world ocean, and the STD of the NHF of the AMIP models is larger than that of the mean reanalysis in the subtropics and midlatitudes. 4) The IAV in the mean reanalysis plays a role in forcing the SST variability in the extra-tropics (e.g., Atlantic Multidecadal Variability), while it may not be an important forcing in the tropical oceans (e.g., ENSO).

KEYWORDS:

Internal Atmospheric Variability, AMIP Simulations, Uncertainty

1 INTRODUCTION

In climate models, the atmospheric circulation can be separated into two parts, one part being the response to the forcing external to the atmosphere (e.g., sea surface temperature, solar input, volcanic aerosols), and the other being the internal atmospheric variability (Rowell *et al.*, 1995). In conceptual climate models, the internal atmospheric variability (referred to hereafter as IAV) is commonly parameterized as a specified stochastic forcing applied to the ocean. Hasselmann (1976) demonstrated in a one-variable slab ocean model that the parameterized IAV could play a major role in forcing the low-frequency sea surface temperature (SST) variability. Using a simple linear coupled atmosphere-ocean model, Barsugli and Battisit (1998) showed that the properties of the full coupled model and an uncoupled version (atmosphere with specified SST) had important differences in extratropical regions, which were attributed to the role of IAV in forcing the SST variability in the coupled model but not in the uncoupled model.

Chen *et al.* (2013) decomposed the atmospheric circulation of a coupled GCM (CGCM) simulation with constant external forcing into the SST-forced response and the IAV (named as atmospheric noise in their paper). They applied the methodology of extracting IAV from a model simulation by estimating the SST-forced component using auxiliary simulations and subtracting it from the model results. The SST forced response was found as the average of an ensemble of Atmospheric Model Intercomparison Project (AMIP)-type atmospheric GCM (AGCM) simulations, using the atmospheric component of the CGCM. Each of the ensemble members was forced by the CGCM simulation's SST but with different initial atmosphere and land surface conditions (Harzallah and Sadourny, 1995; Saravanan, 1998; Frankignoul, 1999; Compo and Sardeshmukh, 2009; Hurrell *et al.*, 2009; Fan and Schneider, 2012). The IAV was found to be statistically indistinguishable between the CGCM and the AGCM simulations in this perfect model framework. Colfescu and Schneider (2017) reached a similar conclusion for the IAV in a CGCM simulation with historical external forcing.

Some studies found that the IAV was the dominant driving mechanism for the SST variability. Yeh and Kirtman (2004) used the interactive ensemble technique to control the amplitude

of IAV at the air-sea interface in the coupled model and found that most of the SST variability in the North Pacific was the response to IAV. Applying the IAV of the surface fluxes to the ocean surface of an interactive ensemble CGCM, Chen *et al.* (2016) found that the IAV forcing was responsible for the SST variability in the Atlantic Ocean (in particular, the Atlantic multidecadal variability, AMV). Colfescu and Schneider (2020) added historical external forcing to the same model and found the IAV was important in driving the internal AMV, except for the high latitude North Atlantic. Besides driving the SST variability, the IAV also played a key role in the Atlantic Meridional Overturning Circulation (Delworth and Greatbarch, 2000; Dong and Sutton, 2005; Jungclaus *et al.*, 2005; Kwon and Frankignoul, 2012) and ENSO (Moore and Kleeman, 1999; Thompson and Battisti, 2001; Zavala-Garay *et al.*, 2003; Yeh and Kirtman, 2006).

In this paper, we extend the methodology of Chen *et al.* (2013) to decompose the net surface heat flux in analyses of the observed atmosphere into the SST/externally forced response and the IAV. The SST/externally forced response is estimated from an AMIP ensemble. Then the IAV in the observed atmosphere is estimated as the residual of the observed variable after removing the average of AMIP simulations forced by the observed SST and external forcing. To our knowledge, this is the first attempt to quantify and document the IAV in the observed atmosphere. However, the forced response in the AGCM may not be same as the real forced response in observations (Colfescu *et al.*, 2013; Chen and Schneider, 2014). The estimate of the SST/externally forced response of the observed atmosphere then includes model bias. Consequently, the IAV of the observed atmosphere, computed as the residual after removing the forced response, contains model bias as well as uncertainty.

On one hand, uncertainty comes from the "observation." Reanalysis datasets are widely used as "observation" to represent the actual in situ and satellite observations, due to superior spatial and temporal coverage. Reanalyses are observation-based, but studies have shown that reanalyses from different research centers differ, particularly in regions with little or no data, due to the choice of in-situ observations, changes in the observing systems, differing models, differing data assimilation configurations and so on (Newman *et al.*, 2000; Smith *et al.*, 2001; Bosilovich *et al.*, 2008; Hodges *et al.*, 2011; Chaudhuri *et al.*, 2014; Grist *et al.*, 2014).

On the other hand, model uncertainty comes from the models used to estimate the forced response. Even forced with the same boundary conditions, the forced response from state-of-the-art AGCMs from various modeling centers differ from each other in many aspects (Lau *et al.*, 1996; Boyle, 1998). Various differences in the model formulations, including in physical parameterizations and numerics, are responsible for differences in the simulations (Taylor *et al.*, 2012), although direct attribution of how the differences in formulation affect the results is difficult (Schneider, 2002; Cash *et al.*, 2005, 2007).

The motivation of the current study is to apply the procedures of Chen *et al.* (2013) and Colfescu and Schneider (2017) to investigate IAV in the observed atmosphere. There are several reanalyses available for estimating the observed atmosphere for the recent past. Also, the data necessary to calculate estimates of the SST/externally forced response of the atmospheric circulation for the same period have been contributed by several modeling groups to phase 5 of the Coupled Model Intercomparison Project (CMIP5). These datasets allow us to compute a number of estimates of IAV. In the following, Sec. 2 describes the models, datasets and methodology used in the study. Sec. 3 examines the IAV estimates and associated uncertainties that arise from trying to combine results from models and reanalyses in a consistent manner. Sec. 4 contains discussion and conclusions.

2 DATA AND METHODS

The reanalysis, models and SST datasets that are used for our estimates of the IAV are described in this section. The SST data is monthly means of the Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST1; Rayner *et al.*, 2003) merged with National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation Sea Surface Temperature (NOAA OI SST; Reynolds *et al.*, 2007). The data processed consists of 360 monthly mean fields from 1979 to 2008. The field analyzed for IAV is the net surface heat flux (NHF). The NHF and SST are detrended and the annual cycles are removed.

2.1 Reanalysis datasets

The reanalyses are internally consistent but distinct estimates of the true atmospheric state for the recent past. Four widely-used global atmospheric reanalysis datasets are employed: 1) European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (ERA; Dee *et al.*, 2011), 2) Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker *et al.*, 2011) from the National Aeronautics and Space Administration (NASA), 3) Climate Forecast System Reanalysis (CFSR; Saha *et al.*, 2010) from National Centers for Environmental Prediction (NCEP), and 4) NCEP/National Center for Atmospheric Research (NCAR) reanalysis (NCEP; Kistler *et al.*, 2001). The net heat fluxes in the reanalyses are not directly observed, rather they are model output constrained to be dynamically consistent with the observations.

2.2 CMIP5 AMIP ensembles

The CMIP5 AMIP (Atmospheric Model Intercomparison Project, Gates *et al.*, 1999) ensembles are free running AGCM simulations with SST and sea ice boundary conditions specified from the HadISST1 merged with NOAA OI SST. Other time dependent conditions obtained from observations are also specified in the AGCM simulations, such as concentrations of gases including CO_2, incoming solar radiation at the top of the atmosphere, emissions or concentrations of short-lived species and natural and anthropogenic aerosols or their precursors, and land use (Taylor *et al.*, 2009).

The four AMIP ensembles that have at least five ensemble members out of the seventeen available from the CMIP5 data archive are used in the analysis. The ensembles are CSIRO, GISS, CCSM4, and GFDL: Commonwealth Scientific and Industrial Research Organization (CSIRO)-Mark 3 (MK3)-6-0 (CSIRO for short; Gordon *et al.*, 2002) from CSIRO-Queensland Climate Change Center of Excellence, Goddard Institute for Space Studies (GISS)-E2-R (GISS for short; Schmidt *et al.*, 2006) from the National Aeronautics and Space Administration, Community Climate System Model version 4 (CCSM4 for short; Meehl *et al.*, 2012) from National Center for Atmospheric Research, and Geophysical Fluid Dynamics Laboratory (GFDL)-CM3 (GFDL for short, Donner *et al.*, 2011) from

NOAA-GFDL. There are 10 members in the CSIRO ensemble, 6 members in the GISS ensemble, 5 members in the CCSM4 ensemble and 5 members in the GFDL ensemble.

Atmospheric models forced by observed SST and external forcings produce estimates of the model atmospheric response to those forcings. The forced responses from these AMIP simulations can be considered to be estimates of the hypothetical actual forced response in the observed system that is contained in each of the reanalyses. However, the actual forced response is not known and cannot be directly compared with the model estimates. Errors due to unrealistic physical or numerical representations of various processes cannot be excluded in the forced response of the AGCM that was forced by observed SST (Chen *et al.*, 2013; Colfescu and Schneider, 2017) and these errors will then transfer into the IAV estimates for the observed atmosphere, as described in the next section.

The four reanalyses are used to construct a 4-member ensemble of reanalyses denoted 4Rean. A 20-member ensemble of AMIP simulations, denoted 4AMIP is constructed by choosing 5 members from each of the model AMIP ensembles.

2.3 IAV estimation and statistical bias corrections

The information contained in reanalyses and models can be better understood by comparing the forced responses estimated for separate models and by comparing the IAV implied by the forced responses in reanalyses and models simulations. We introduce briefly the estimation of IAV for the reanalyses and AMIP simulations. Following Chen *et al.* (2013) an atmospheric variable is assumed to be the sum of two parts, the forced response (the atmospheric response to SST/external forcing) and the IAV. The procedure to estimate IAV of an atmospheric variable is then to subtract the time-varying forced response computed from an AMIP ensemble from the atmospheric variable in a model simulation or a reanalysis. The estimates of the forced response described below are ensemble means of the NHF produced by the four CMIP5 AMIP ensembles as well as the forced response from the 4AMIP ensemble.

While the forced responses in the CMIP5 ensembles can be considered to provide estimates of the forced response in the reanalyses, the IAV in the model ensemble members is uncorrelated with

the actual IAV in the reanalyses, since the evolution of the weather systems in the free running models is not constrained to be that seen in the observations. Our comparison of the IAV between models and reanalyses is then restricted to overall statistics, including correlations and comparisons of variance.

The IAV for a reanalysis is found by subtracting from that reanalysis the SST/externally forced response estimated from an AMIP ensemble. Model data of ensembles of *M* AMIP simulations are considered, with each member forced by the same SST and external forcing over the historical period. The SST/externally forced solution from that model ensemble, denoted *F*, is estimated as the ensemble mean of the *M* members. Due to the finite number of members in an ensemble, statistical bias exists between the unbiased and the estimated variances of the forced responses, as well as the variances of the unbiased and the estimated IAV of the observed atmosphere (Rowell *et al.*, 1995; Chen *et al.*, 2013 supplementary material). Details of the decomposition and the statistical bias corrections used in the analysis of the variance-related statistics are given in the Appendix.

3 RESULTS

This section examines both the forced response and the IAV and their uncertainties in the observed atmosphere for the NHF, to compare among the reanalysis datasets and AMIP simulations.

3.1 Local correlation of NHF with SST

Before investigating the IAV, we compare the reanalyses with AMIP simulations in capturing the SST-NHF relationship for the undecomposed NHF.

Figure 1 shows the local correlations between the observed SST anomaly and the NHF from four reanalyses, with dotted regions statistically significant at 5% level. It can be seen that the SST-NHF correlations agree among the four reanalyses to a large degree, with negative correlations in the tropical Pacific, eastern tropical Atlantic and northwest Indian Ocean, and positive correlations in the subtropical Pacific, North Atlantic and southern Indian Ocean. Discrepancies exist in the Arctic Ocean between 120°E and 160°W, where correlations are significantly positive for ERA and CFSR but negative for MERRA, and more complicated for NCEP. Furthermore, CFSR shows a larger

magnitude of positive correlation in the subtropical Pacific and North Atlantic Ocean and a smaller magnitude of negative correlation in the tropical oceans than other three reanalyses. Note that different reanalyses do not behave the same everywhere, due to differences in the choice of in-situ observations, the interpolating models, and the data assimilation schemes.

Figure 2 shows the local correlations between the observed SST anomaly and NHF from AMIP runs from the four models. They are significantly negative for all of the AGCMs over a majority of the globe. Additionally, although the four AMIP models are forced by the same observed SST and external forcing, their atmospheric NHF fields have regions where they disagree in terms of their relationship with SST. For instance, the SST-NHF correlation in the equatorial eastern Pacific near the coast is negative for GISS (Fig. 2c) but is positive for other three models. The correlation in the equatorial central Pacific is positive for CSIRO (Fig. 2d) but is negative for other three models.

Instead of using the undecomposed NHF from a single simulation of an AMIP model, Fig. 3 displays the local correlations between the observed SST anomaly and the forced response of NHF (referred to as NHF_F, i.e., the average of a full AMIP model ensemble) for AMIP runs from the four models. The spatial patterns are consistent with Fig. 2 for each AMIP model, but the magnitude of the negative values and the significant regions are much larger for the SST-NHF_F correlations (Fig. 3) than the SST-NHF correlations (Fig. 2).

The negative SST-NHF and SST-NHF_F correlations indicate damping of SST anomalies by the NHF, which is consistent in the tropical oceans between the reanalyses and AMIP runs, except for CSIRO in the equatorial central Pacific. However, the most striking difference between Fig. 1 and Figs. 2-3 lies in the extra-tropics, with positive correlations for reanalyses (Fig. 1) but negative correlations for AMIP runs (Figs. 2-3). The positive IAV/SST correlation is expected if the IAV in the observed atmosphere plays an important role in forcing the SST variability (Chen and Schneider 2014).

Using the IAV estimated as in Section 2.3, correlations between SST anomaly and the IAV of NHF (referred to as NHF_I) are examined for the reanalyses and AMIP runs. For AMIP simulations, the NHF_I is not significantly correlated with SST (figure not shown). The lack of correlation occurs

since the atmosphere cannot force SST anomalies in the AMIP simulations, while the forcing of the atmosphere by the SST is captured by the forced response. Thus, the increased damping correlation in Fig. 3 compared to Fig. 2 can be explained as follows: as NHF_I is uncorrelated with SST for AMIP runs, the covariance of the NHF and SST is the same for each ensemble member in an AMIP model. However, the variance of NHF_F is smaller than that of the undecomposed NHF, leading to larger magnitude of SST-NHF_F correlation than the SST-NHF correlation.

Then the SST-NHF_I correlations for the reanalyses, where the IAV estimates are found by removing the forced responses from the AMIP simulations, are investigated (Fig. 4). The correlations are significantly positive over a majority of the extratropical oceans in four reanalyses, indicative of the role of IAV in forcing the SST variability in the extra-tropics. Since the reanalyses represent the NHF/SST relationships in the coupled atmosphere-ocean system, this agrees with Chen *et al.* (2013), which attributed the differences between the CGCM and the AGCM forced with SST from the CGCM to local IAV forcing on the SST in the CGCM but not in the AGCM. It is also noted that the correlations are weak negative over the tropical Pacific and tropical Atlantic in the four reanalyses, suggesting that IAV may not be an important forcing for the SST variability in these regions. Similar to Fig. 1, there are some differences among reanalyses, such as the magnitude and spatial extension of the negative regions.

3.2 Standard deviations and standard deviation ratios

The standard deviations (STD) and standard deviation ratios (STDR) between the AMIP runs, between the reanalyses and between the reanalyses and the AMIP runs are investigated in this section. Firstly, the STD of the NHF_F are compared among the AMIP models (Fig. 5). The structures of the NHF_F STD in each of the four model ensembles generally resembles that of the 4AMIP ensemble (Fig. 5a), with largest variability in the tropical Pacific, Northwest Pacific and North Atlantic Current regions. For comparison, the STDR of the 4AMIP NHF_F to the bias-adjusted NHF_F of the individual models is shown in Figs. 5b-e. The bias corrections to the 4AMIP NHF_F are reduced compared to those of the individual model ensembles as a consequence of using the large ensemble.

The NHF_F variances and bias adjustments for the individual models are calculated using the full ensembles of each model. In comparison with 4AMIP, the NHF_F variance in GFDL (Fig. 5c) is larger than that of the other models along the western coast of North/South America and in the west Pacific and northern Indian Ocean. The NHF_F variance in GISS (Fig. 5d) is smaller than that of the other models in these regions, as well as in the equatorial Atlantic. CSIRO has larger variability in NHF_F over much of the Atlantic as well as the eastern South Pacific (Fig. 5e). There are belts of low variability with the STD less than 10 W m^{-2} in the 4AMIP Arctic and Southern Oceans (Fig. 5a), and very large or small ratios in these regions may be misleading as they can be due to small absolute differences in NHF_F variability between the models. Then, although the AMIP models are forced with the same SST and external forcing, the significant differences in the forced responses among the AMIP models are indicative of uncertainty in the models due to intrinsic model differences.

The STD of the unbiased NHF in the 4Rean mean and the STDR of NHF between 4Rean and the individual reanalysis are shown in Fig. 6. The 4Rean mean NHF is an estimate of the NHF of the observed atmosphere including both the forced and IAV components, and can be considered as a noise-filtered representation of the observed information in the 4 reanalyses. The large-scale structure of the STD of the 4Rean NHF (Fig. 6a) resembles that of the 4AMIP NHF-F (Fig. 5a); however, the reanalysis mean NHF STD is generally larger than the model mean NHF_F STD in the extra-tropics and smaller in the tropics excepting the equatorial eastern Pacific. The STDR of the individual reanalysis relative to the 4Rean mean shown in Figs. 6c-f shows that the individual reanalysis has generally larger NHF variability than the mean. A decrease in the variability in the ensemble mean is expected if, for example, the differences between the reanalyses are random noise. The largest increases in STD relative to the mean occur in the tropics and the Arctic Ocean, both regions of small mean NHF STD. While the large ratios indicate regions dominated by inter-reanalysis noise, the result is due to small NHF STD differences in regions of small NHF STD in the mean reanalysis. The MERRA reanalysis has the smallest NHF STD and is also closest to 4Rean in the extra-tropics. The comparison in Fig. 6 shows that there is significant uncertainty in NHF between the reanalyses.

The STD of the NHF_I in the reanalyses is compared with that in the models in Fig. 7. A representative NHF_I for the reanalyses is found by removing the 4AMIP NHF_F from the 4Rean NHF (Fig. 7a). The unbiased variance of the 4Rean NHF_I is computed from Eq. (A12). The NHF_I is also found for the individual members of the 4AMIP and the full ensembles of the individual models, with unbiased NHF_I variances computed using Eq. (A11). The average of the NHF_I variances of the members of the 4AMIP and AMIP ensembles is used to compute the STDR in Figs. 7b-f. The 4Rean NHF_I STD pattern (Fig. 7a) resembles that of the 4Rean NHF (Fig. 6a). The 4AMIP NHF_I variance, is indistinguishable from that of 4Rean, except for the deep tropics, the western Indian Ocean, the eastern South Atlantic, coastal western North America, and the Arctic Ocean (Fig. 7b). The NHF_I of the individual model AMIP ensembles show similar characteristics. As mentioned above, the deep tropics and Arctic ocean are regions of low NHF variance where the STDR can be misleading. Thus, it appears that the NHF_I in the models is generally consistent with that in the reanalyses.

Figure 8a shows the STDR of NHF in 4Rean to 4AMIP, with the 4AMIP variance computed as the mean of the NHF variances of the 20 ensemble members (no bias corrections). The reanalysis NHF has substantially smaller variance than the AMIP NHF in the subtropics and midlatitudes, except in a few isolated regions. Figure 8b shows that the variance of the NHF_I in 4Rean found from removing 4AMIP NHF_F is larger than the 4Rean NHF in belts from 10 to 60 latitude in both hemispheres. In contrast, Fig. 8c shows that the NHF_I in 4AMIP found using the 4AMIP NHF_F has smaller variance than the 4AMIP NHF everywhere.

The results in Figs. 1-8 are consistent with the view that the models taken together are able to produce realistic simulations of statistical aspects the NHF_F and NHF_I of the observed atmosphere, and indicate NHF_I forces much of the SST variability over the world ocean.

3.3 Regressions onto SST indices

Chen and Schneider (2014) proposed an indirect way, without explicit calculation of the IAV, to test whether the SST-forced response is the same between two AGCM simulations forced by the

same SST. This is to compare the time-lagged regressions, with SST leading the atmosphere by more than the decorrelation time of IAV. They used this test to demonstrate that the SST-forced responses in a perfect model framework were the same in a CGCM and an AGCM forced with SST from the CGCM. This method can be extended to compare reanalysis with reanalysis, reanalysis with model, or model with model. In this section, we will employ it to compare the SST/externally forced response in the reanalyses and AMIP simulations taking CCSM4 AMIP as an example. Two SST indices are used, the AMV index and Niño3.4 index, which are defined as the area-averaged SST anomalies over (0°-60°N, 80°W-0°) and (5°S-5°N, 170°W-120°W), respectively.

Figure 9 displays the lagged regressions of NHF from both ERA reanalysis and CCSM4 AGCM onto the AMV index. When SST leads the atmosphere, the NHF is negatively correlated with AMV in the North Atlantic to the east of 30°W (Figs. 9a, b), and the differences between reanalysis and model are not significantly different from zero over most of North Atlantic Ocean, except for a region in the eastern North Atlantic at 30°-40°N (Fig. 9c), indicating that the SST-forced responses of the NHF are indistinguishable between ERA and CCSM4 AGCM in most of North Atlantic Ocean. For simultaneous regressions or when SST lags the atmosphere, the NHF-AMV relation becomes positive in most of North Atlantic in reanalysis (Figs. 9d, g), while it does not change sign in the AGCM (Figs. 9e, h). Differences between reanalysis and AGCM are significant in the North Atlantic (Figs. 9f, i), indicating that IAV forcing of the observed SST is probably important in the North Atlantic. This is consistent with the IAV forcing SST in the coupled system but not in the uncoupled model (Chen and Schneider, 2014).

Also note that there are regions in the Pacific with a negative NHF-AMV relationship in the reanalysis (Figs. 9a, d, g), suggestive of possible teleconnections from the AMV. However, the correlation is positive in these regions in the AGCMs (Figs. 9b, e, h), which might be explained as due to model bias.

Unlike the AMV index, the lagged linear regressions of NHF anomalies onto the Niño3.4 index in the reanalysis (Figs. 10a, d, g) show the same features for SST leading, simultaneous with, or lagging the atmosphere, with the sign negative in the central and eastern tropical Pacific and positive

in the western tropical Pacific and tropical Atlantic Ocean. The patterns are also the same in the CCSM4 AGCM (Figs. 10b, e, h). This indicates that IAV is unlikely to play an important role in forcing the ENSO SST variability. The differences between the reanalysis and AGCM (Figs. 10c, f, i) can be attributed to model bias as well.

The lagged regressions then suggest that the IAV forcing is an important forcing for the AMV, but not for ENSO.

4 DISCUSSION AND CONCLUSIONS

The 1979-2008 monthly mean net surface heat fluxes (NHF) in four atmospheric reanalyses and a set of CMIP5 AMIP simulations forced by observed SST and external forcing for the same period were decomposed into the externally/SST forced and internal atmospheric variability (IAV). The AMIP simulations were used to estimate the externally/SST forced NHF variability, which was removed from the total NHF to estimate the IAV in the reanalyses and the AMIP simulation. The 4Rean mean was decomposed into a forced solution, taken as the NHF_F from the 4AMIP ensemble and a residual NHF_I, representing the IAV in the mean reanalysis. In both reanalysis and model, the decomposition is written as

$$NHF = NHF\_F + NHF\_I. \quad (1).$$

Then taking variances of NHF,

$$Var(NHF)=Var(NHF\_F)+Var(NHF\_I)+2Cov(NHF\_F,NHF\_I) \quad (2).$$

In the AMIP simulations, the SST is specified and is not affected by the IAV. There is no relationship between the NHF_F_{4AMIP} and NHF_I_{4AMIP} in the 4AMIP ensemble members used to compute NHF_F_{4AMIP}, so that $Cov(NHF\_F_{4AMIP},NHF\_I_{4AMIP})=0$ as would be expected if NHF_I_{4AMIP} were random noise. As a consequence, $Var(NHF_{4AMIP})>Var(NHF\_I_{4AMIP})$ (Fig. 8c). $Var(NHF\_I_{4REAN})$ in the 4Rean mean, computed using the NHF_F_{4AMIP}, is close to $Var(NHF\_I_{4AMIP})$ (Fig. 7b). However, $Var(NHF_{4REAN})<Var(NHF_{4AMIP})$ over most of the world ocean (Fig. 8a), so that $Cov(NHF\_F_{4AMIP},NHF\_I_{4REAN})<0$ in those regions, a relationship that can also be inferred in the regions where $Var(NHF\_I_{4REAN})>Var(NHF_{4REAN})$ in Fig. 8b. If NHF_I_{REAN} can be thought of as

random noise, the relationship between the NHF_F and the NHF_I_{REAN} can be explained as due to NHF_I_{REAN} forcing SST anomalies, while NHF_F is the response to these SST anomalies. The correlations shown in Figs. 2-4 are consistent with NHF_I_{REAN} forcing on the SST, with the major exception of the equatorial oceans.

On the other hand, a systematic bias in the magnitude of the NHF_F response in the AMIP ensembles could lead to an erroneous conclusion that $Cov(NHF\_F_{4AMIP}, NHF\_I_{4REAN})<0$ in the decomposition of the reanalyses. This cannot be ruled out in the present study.

In a perfect model framework, where an AGCM AMIP ensemble was forced by the SSTs from a CGCM simulation, Chen *et al*. (2013) showed that the CGCM/AGCM STDR of NHF was significantly less than one for most oceans, and the variances of IAV are indistinguishable between the CGCM and AGCM. As model bias is eliminated using the perfect model strategy, the less-than-one ratio can be confidently interpreted as due to the destructive interference between the IAV and the forced response in the coupled model, whereas the IAV and forced component are temporally uncorrelated in the uncoupled simulations. The same interpretation is consistent with the results presented here. Additional evidence of forcing of SST anomalies by NHF_I_{REAN} was shown in lag correlation.

The uncertainty of the IAV in the reanalyses can arise from two causes, the uncertainty in reanalyses (as observations contain errors and reanalysis includes interpolation by models) and the uncertainty in models (as the forced responses are different when estimated from different AGCMs). Both of these sources are shown to be non-negligible by comparison of various reanalyses and AMIP ensemble simulations made with different atmospheric models.

The local correlations between the observed SST and the NHF from AMIP runs are significantly negative for most oceans, indicating damping of SST by the NHF in the AMIP models, which is as expected if the IAV is forcing the SST variability. However, the SST-NHF correlations from reanalyses are largely positive in the extra-tropics and negative in the tropics, suggestive of the role of IAV in forcing the SST variability in the extra-tropics. The reanalyses are generally consistent in their SST-NHF correlations, yet discrepancies exist, which must be due to differences in the choice

and correction of in-situ observations, interpolating models, and data assimilation schemes. We have addressed these differences by constructing a 4-reanalyses mean reanalysis and bias correcting the variance of this mean using the differences of the individual reanalyses from the mean. The bias corrected variance of this mean reanalysis represents the NHF agreed on by all of the reanalyses

Time-lagged regressions of NHF from reanalyses and AMIP models onto the AMV index showed that the SST/externally forced responses of the NHF were indistinguishable between ERA reanalysis and CCSM4 AGCM in most of North Atlantic Ocean and that the IAV was an important forcing for the SST variability in the North Atlantic Ocean (i.e., AMV). In contrast, the IAV forcing was probably not important in forcing the SST variability for ENSO. The same pattern of the difference field between reanalysis and AGCM, for the SST leading/lagging or simultaneous with atmosphere, probably resulted from model bias in the forced response.

Since AMIP models forced with the observed SSTs and sea ice are necessary to estimate the SST/externally forced response of the observed atmosphere, model bias cannot be excluded in the forced response and the derived IAV in the observed atmosphere. Furthermore, differences in the reanalyses from various sources contribute to uncertainty in the IAV estimates. Despite these uncertainties, the IAV variances in the combined model ensembles and the mean reanalysis were very similar. The current work suggests that the decomposition into the forced response and IAV could be a useful diagnostic both to understand climate variability and to isolate and diagnose differences between reanalysis datasets and between models.

ACKNOWLEDGEMENT

Chen was supported by the National Key R&D Program (2016YFA0600402). Schneider was supported by NSF grants AGS-1338427 and 1558821, NOAA NA14OAR4310160, and NASA NNX14AM19G. Zhu was supported by the National Natural Science Foundation of China (No. 42088101).

APPENDIX

The corrections derived below are applied to remove statistical biases in the externally forced and internal variances that occur due to the use of small ensembles. An atmospheric variable in the reanalysis is decomposed as

$$O_k = H + r_k = \tilde{H} + \tilde{r}_k \qquad (A1)$$

where O_k is the undecomposed atmospheric variable in the k^{th} reanalysis, H is the actual field, which is the same in each reanalysis, and r_k is the residual. The associated biased estimated variables are $\tilde{H}$, and $\tilde{r}_k$. Taking $\tilde{H}$ to be estimated as the mean of the N reanalyses,

$$\tilde{H} = \frac{1}{N}\sum_{j=1}^{N} O_k = \frac{1}{N}\sum_{j=1}^{N}(H + r_j) = H + \frac{1}{N}\sum_{j=1}^{N} r_j \qquad (A2).$$

The actual residuals r_k are assumed to be realizations of random noise, each with equal variance, uncorrelated with each other or with H. Denoting the variance by Var and defining $Var(r) = Var(r_k)$, it follows that

$$Var(\tilde{H}) = Var\left(H + \frac{1}{N}\sum_{j=1}^{N} r_j\right) = Var(H) + \frac{1}{N}Var(r) \qquad (A3).$$

From Eqs. (A1) and (A2),

$$\tilde{r}_k = O_k - \tilde{H} = O_k - \left(H + \frac{1}{N}\sum_{j=1}^{N} r_j\right) = r_k - \frac{1}{N}\sum_{j=1}^{N} r_j \qquad (A4).$$

Taking the variance of Eq. (A4),

$$Var(\tilde{r}_k) = Var\left(r_k - \frac{1}{N}\sum_{j=1}^{N} r_j\right) = Var(r) + \frac{1}{N}Var(r) - 2Cov\left(r_k, \frac{1}{N}\sum_{j=1}^{N} r_j\right) = \frac{N-1}{N}Var(r) \qquad (A5).$$

From Eqs. (A3) and (A5), the unbiased variances of the mean reanalysis and the residuals of the individual reanalysis are given by

$$Var(H) = Var(\tilde{H}) - \frac{1}{N-1}Var(\tilde{r}) \qquad (A6)$$

$$Var(r) = \frac{N}{N-1}Var(\tilde{r}_k).$$

Since in practice the $Var(\tilde{r}_k)$ are not equal, the average is used instead of $Var(\tilde{r}_k)$ to estimate the unbiased $Var(r)$:

$$Var(r) = \frac{N}{N-1} Var(\tilde{r}) \quad (A7),$$

with $Var(\tilde{r}) = \frac{1}{N}\sum_1^N Var(\tilde{r}_k)$.

A similar set of manipulations leads to the unbiased variances of the forced response and the IAV in an ensemble of AMIP ensemble. An atmospheric variable in an AMIP ensemble is decomposed as

$$A_k = F + \eta_k = \tilde{F} + \tilde{\eta}_k \quad (A8),$$

where A_k is the undecomposed atmospheric variable in the k^{th} ensemble member, F is the unbiased forced response, which is the same in each ensemble member, and η_k is the residual, in this case the ensemble member's IAV. The associated biased estimated variables are $\tilde{F}$ and $\tilde{\eta}_k$. The IAV η_k are assumed to be realizations of random noise, each with equal variance $Var(\eta)$, uncorrelated with each other or with F. Then

$$F = \tilde{F} - \frac{1}{M}\sum_{j=1}^{M} \eta_j \quad (A9),$$

and the unbiased variances of the forced response and the IAV in the ensemble members are

$$Var(F) = Var\left(\tilde{F}\right) - \frac{1}{M-1} Var(\tilde{\eta}) \quad (A10)$$

$$Var(\eta) = \frac{M}{M-1} Var(\tilde{\eta}_k).$$

Since in practice the $Var(\tilde{\eta}_k)$ are not equal, the average is used instead of $Var(\tilde{\eta}_k)$ to estimate the unbiased $Var(\eta)$:

$$Var(\eta) = \frac{M}{M-1} Var(\tilde{\eta}) \quad (A11),$$

is used with $Var(\tilde{\eta}) = \frac{1}{M}\sum_1^M Var(\tilde{\eta}_k)$.

The IAV in the mean reanalysis is estimated by removing the estimated forced response from the estimated reanalysis mean. Denoting the actual IAV in the mean reanalysis as I and the estimated IAV as $\tilde{I}$, and for an N-member reanalysis ensemble and an M-member AMIP ensemble,

$$\tilde{I} = \tilde{H} - \tilde{F}$$

Then from Eqs. (A2) and (A9)

$$\tilde{I} = I + \frac{1}{N}\sum_{j=1}^{N} r_j + \frac{1}{M}\sum_{j=1}^{M} \eta_j.$$

The actual IAV in the mean reanalysis is assumed to be uncorrelated with the IAV in the models or the inter-analyses residuals, so that, using Eqs. (A7) and (A11), the unbiased variance of the IAV in the mean reanalysis is

$$Var(I) = Var(\tilde{I}) - \frac{1}{N-1}Var(\tilde{r}) - \frac{1}{M-1}Var(\tilde{\eta}) \qquad (A12).$$

If the IAV is computed by removing forced response from a single reanalysis, Eq. (A12) applies with $\tilde{I} = O_k - \tilde{F}$ and $Var(\tilde{r}) = 0$.

The above results are also used in estimates of correlations of unbiased quantities. Assuming that the covariance between the SST anomaly and the IAV in the AMIP simulations are zero, the bias-corrected correlation between the SST anomaly and the SST/externally forced response in an AMIP ensemble is

$$Corr(F, SST) = \frac{Cov(\tilde{F}, SST)}{\sqrt{Var(\tilde{F}) - \frac{1}{M-1}Var(\tilde{\eta})}\sqrt{Var(SST)}} \qquad (A13).$$

The bias corrected correlation between the IAV in a reanalysis and the SST anomaly is:

$$Corr(I, SST) = \frac{Cov(\tilde{I}, SST)}{\sqrt{Var(\tilde{I}) - \frac{1}{N-1}Var(\tilde{r}) - \frac{1}{M-1}Var(\tilde{\eta})}\sqrt{Var(SST)}} \qquad (A14).$$

REFERENCES

Barsugli, J.J. and Battisti, D.S. (1998) The basic effects of atmosphere–ocean thermal coupling on midlatitude variability. *Journal of the Atmospheric Sciences*, 55, 477–493.

Bosilovich, M.G., Chen, J., Robertson, F.R. and Adler, R.F. (2008) Evaluation of global precipitation in reanalyses. *Journal of Applied Meteorology and Climatology*, 47, 2279–2299.

Boyle, J.S. (1998) Intercomparison of interannual variability of the global 200-hPa circulation for AMIP simulations. *Journal of Climate*, 11, 2505–2529.

Cash, B.A., Schneider, E.K. and Bengtsson, L. (2005) Origin of regional climate differences: Role of boundary conditions and model formulation in two GCMs. *Climate Dynamics*, 25, 709-723.

Cash, B.A., Schneider, E.K., and Bengtsson, L. (2007) Origin of climate sensitivity differences: Role of selected radiative parameterizations in two GCMs. *Tellus*, 59, 155-169.

Chaudhuri, A.H., Ponte, R.M. and Nguyen, A.T. (2014) A comparison of atmospheric reanalysis products for the Arctic Ocean and implications for uncertainties in air–sea fluxes. *Journal of Climate*, 27, 5411–5421.

Chen, H., Schneider, E.K., Kirtman, B.P. and Colfescu, I. (2013) Evaluation of weather noise and its role in climate model simulations. *Journal of Climate*, 26, 3766-3784.

Chen, H. and Schneider, E.K. (2014) Comparison of the SST-forced responses between coupled and uncoupled climate simulations. *Journal of Climate*, 27, 740-756.

Chen, H., Schneider, E.K. and Wu Z. (2016) Mechanisms of internally generated decadal-to- multidecadal variability of SST in the Atlantic Ocean in a coupled GCM. *Climate Dynamics*, 46, 1517-1546.

Colfescu, I., Schneider, E.K. and Chen, H. (2013) Consistency of 20th century sea level pressure trends as simulated by a coupled and uncoupled GCM. *Geophysical Research Letters*, 40, 3276-3280.

Colfescu, I. and Schneider, E.K. (2017) Internal atmospheric noise characteristics in twentieth century coupled atmosphere–ocean model simulations. *Climate Dynamics*, 49, 2205-2217.

Colfescu, I. and Schneider, E.K. (2020) Decomposition of the Atlantic Multidecadal Variability in a historical climate simulation. *Journal of Climate*, 33, 4229-4254.

Compo, G.P. and Sardeshmukh, P.D. (2009) Oceanic influences on recent continental warming. *Climate Dynamics*, 32, 333-342.

Copsey, D., Sutton, R. and Knight, J.R. (2006) Recent trends in sea level pressure in the Indian Ocean region. *Geophysical Research Letters*, 33, L19712. http://doi.org/10.1029/2006GL027175.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L.,

Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut J.N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

Delworth, T., and Greatbatch, R. (2000) Multidecadal thermohaline circulation variability driven by atmospheric surface flux forcing. *Journal of Climate*, 13, 1481–1495.

Dong, B., and Sutton, R.T. (2002) Mechanism of interdecadal thermohaline circulation variability in a coupled ocean-atmosphere GCM. *Journal of Climate*, 18, 1117-1135.

Donner, L.J., Wyman, B.L., Hemler, R.S., Horowitz, L.W., Ming, Y., Zhao, M., Golaz, J.C., Ginoux, P., Lin, S.J., Schwarzkopf, M.D., Austin, J., Alaka, G., Cooke, W.F., Delworth, T.L., Freidenreich, S.M., Gordon, C.T., Griffies, S.M., Held, I.M., Hurlin, W.J., Klein, S.A., Knutson, T.R., Langenhorst, A.R., Lee, H.C., Lin, Y., Magi, B.I., Malyshev, S.L., Milly, P.C.D., Naik, V., Nath, M.J., Pincus, R., Ploshay, J.J., Ramaswamy, V., Seman, C.J., Shevliakova, E., Sirutis, J.J., Stern, W.F., Stouffer, R.J., Wilson, R.J., Winton, M., Wittenberg, A.T. and Zeng, F. (2011) The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. *Journal of Climate*, 24, 3484–3519.

Fan, M. and Schneider, E.K. (2012) Observed decadal North Atlantic tripole SST variability. Part I: weather noise forcing and coupled response. *Journal of the Atmospheric Sciences*, 69, 35-50.

Frankignoul, C. (1999) A cautionary note on the use of statistical atmospheric models in the middle latitudes: Comments on "Decadal variability in the North Pacific as simulated by a hybrid coupled model." *Journal of Climate*, 12, 1871-1872.

Gates, W.L., Boyle, J.S., Covey, C., Dease, C.G., Doutriaux, C.M., Drach, R.S., Fiorino, M, Gleckler, P.J., Hnilo, J.J, Marlais, S.M, Phillips, T.J., Potter, G.L., Santer, B.D., Sperber, K.R., Taylor, K.E. and Williams, D.N. (1999) An overview of the Atmospheric Model Intercomparison Project (AMIP I). *Bulletin of the American Meteorological Society*, 80, 29-55.

Grist, J.P., Josey, S.A., Marsh, R., Kwon, Y.O., Bingham, R.J. and Blaker, A.T. (2014) The Surface-Forced Overturning of the North Atlantic: Estimates from Modern Era Atmospheric Reanalysis Datasets. *Journal of Climate*, 27, 3596–3618.

Gordon, H.B., Rotstany, L.D., McGregor, J.L., Dix, M.R., Kowalczyk, E.A., O'Farrell, S.P., Waterman, L.J., Hirst, A.C., Wilson, S.G., Collier, M.A., Watterson, I.G. and Elliott, T.I. (2002) The CSIRO Mk3 climate system model. Tech. Paper 60, CSIRO Division of Atmospheric Research, 130 pp. http://www.cmar.csiro.au/e-print/open/gordon_2002a.pdf.

Harzallah, A. and Sadourny, R. (1995) Internal versus SST-forced atmospheric variability as simulated by an atmospheric general circulation model. *Journal of Climate*, 8, 474-495.

Hasselmann, K. (1976) Stochastic climate models, Part I. Theory. *Tellus*, 28, 473–485.

Hodges, K.I., Lee, R.W. and Bengtsson, L. (2011) A Comparison of extratropical cyclones in recent reanalyses ERA-Interim, NASA MERRA, NCEP CFSR, and JRA-25. *Journal of Climate*, 24, 4888–4906.

Hurrell, J., Meehl, G.A., Bader, D., Delworth, T.L., Kirtman, B. and Wielicki, B. (2009) A unified modeling approach to climate system prediction. *Bulletin of the American Meteorological Society*, 90, 1819-1832.

Jungclaus, J.H., Haak, H., Latif, M., and Mikolajewicz, U. (2005) Arctic–North Atlantic interactions and multidecadal variability of the meridional overturning circulation. *Journal of Climate*, 18, 4013–4031.

Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R. and Fiorino, M. (2001) The NCEP–NCAR 50–Year Reanalysis: Monthly Means CD–ROM and Documentation. *Bulletin of the American Meteorological Society*, 82, 247–267.

Kumar, K., Hoerling, M. and Rajagopalan, B. (2005) Advancing dynamical prediction of Indian monsoon rainfall. *Geophysical Research Letters*, 32, L08704. http://doi.org/10.1029/2004GL021979.

Kwon, Y-O., and Frankignoul, C. (2012) Stochastically-driven multidecadal variability of the Atlantic meridional overturning circulation in CCSM3. *Climate Dynamics*, 38, 859-976.

Lau, K.M., Sud, Y. and Kim, J.H. (1996) Intercomparison of hydrologic processes in AMIP GCMs. *Bulletin of the American Meteorological Society*, 77, 2209–2227.

Meehl, G.A., Washington, W.M., Arblaster, J.M., Hu, A., Teng, H., Tebaldi, C., Sanderson, B.N., Lamarque, J.F., Conley, A., Strand, W.G. and White III, J.B. (2012) Climate system response to external forcings and climate change projections in CCSM4. *Journal of Climate*, 25, 3661-3683.

Meng Q., Latif, M., Park, W., Keenlyside, N.S., Semenov, V.A. and Martin, T. (2011) Twentieth century Walker Circulation change: data analysis and model experiments. *Climate Dynamics*, 38, 1757-1773.

Moore, A.M., and Kleeman, R. (1999) Stochastic forcing of ENSO by the intraseasonal oscillation. *Journal of Climate*, 12, 1199-1220.

Newman, M., Sardeshmukh, P.D. and Bergman, J.W. (2000) An assessment of the NCEP, NASA and ECMWF reanalyses over the tropical west Pacific warm pool. *Bulletin of the American Meteorological Society*, 81, 41–48.

Rayner, N.A., Parker, D.E., Horton, E.B., Folland, C.K., Alexander, L.V., Rowell, D.P., Kent, E.C. and Kaplan, A. (2003) Global analyses of sea surface temperature, sea ice, and night marine

air temperature since the late nineteenth century. *Journal of Geophysical Research*, 108, No. D14, 4407. http://doi.org/10.1029/2002JD002670.

Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S. and Schlax, M.G. (2007) Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*, 20, 5473-5496.

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M. and Woollen, J. (2011) MERRA - NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, 24, 3624–3648.

Rowell, D.P., Folland, C.K., Maskell, K. and Ward, M.N. (1995) Variability of summer rainfall over tropical north Africa (1906-92): Observations and modelling. *Quarterly Journal of the Royal Meteorological Society*, 121, 669-704.

Saha S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.T., Chuang, H.Y., Juang, H.M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P.V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Dool, H.V.D., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G. and Goldberg, M. (2010) The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, 91, 1015–1057.

Saravanan, R. (1998) Atmospheric low-frequency variability and its relationship to midlatitude SST variability: studies using the NCAR Climate System Model. *Journal of Climate*, 11, 1386-1404.

Schmidt, G.A., Ruedy, R., Hansen, J.E., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., Genio, A.D., Faluvegi, G., Friend, A.D., Hall, T.M., Hu, Y., Kelley, M., Kiang, N.Y., Koch, D., Lacis, A.A., Lerner, J., Lo, K.K., Miller, R.L., Nazarenko, L., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Russell, G.L., Sato, M., Shindell, D.T., Stone, P.H., Sun, S., Tausnev, N., Thresher, D. and Yao, M.S. (2006) Present-Day Atmospheric Simulations Using GISS ModelE: Comparison to In Situ, Satellite, and Reanalysis Data. *Journal of Climate*, 19, 153–192.

Schneider, E.K. (2002) The causes of differences between equatorial Pacific SST simulations of two coupled ocean-atmosphere general circulation models. *Journal of Climate*, 15, 449-469.

Smith, S.R., Legler, D.M. and Verzone, K.V. (2001) Quantifying uncertainties in NCEP Reanalyses using high-quality research vessel observations. *Journal of Climate*, 14, 4062-4072.

Taylor, K.E., Stouffer R.J. and Meehl, G.A. (2009) A summary of the CMIP5 experiment design. *PCDMI Report*, 33pp. http://cmip-pcmdi.llnl.gov/cmip5/experiment_design.html

Taylor, K.E., Stouffer R.J. and Meehl, G.A. (2012) An overview of CMIP5 and the experimental design. *Bulletin of the American Meteorological Society*, 93, 485-498.

Thompson, C.J., and Battisti, D.S. (2001) A linear stochastic dynamical model of ENSO. Part II: Analysis. Journal of Climate, 14, 445-466.

Yeh, S.-W., and Kirtman, B.P. (2004) The impact of internal atmospheric variability on the North Pacific SST variability. *Climate Dynamics*, 22, 721–732.

Yeh, S.-W., and Kirtman, B.P. (2006) Origin of decadal El Nino–Southern Oscillation–like variability in a coupled general circulation model. *Journal of Geophysical Reseach*, 111, C01009, doi:10.1029/2005JC002985.

Zavala-Garay, J., Moore, A.M., Perez, C.L., and Kleeman, R. (2003) The response of a coupled model of ENSO to observed estimates of stochastic forcing. *Journal of Climate*, 16, 2827-2842.

FIGURE CAPTIONS

FIGURE 1 Pointwise correlations between the SST anomaly and the undecomposed net surface heat flux (NHF) from four reanalyses: (a) ERA, (b) MERRA, (c) CFSR, and (d) NCEP. The values between -0.1 and 0.1 are not plotted. Regions are dotted that are statistically significant at 5% confidence level using the t test.

FIGURE 2 Pointwise correlations between the SST anomaly and the undecomposed NHF from the AMIP AGCMs of (a) CCSM4, (b) GFDL, (c) GISS, and (d) CSIRO. Each AGCM is a single simulation chosen arbitrarily from the AGCM ensemble. The values between -0.1 and 0.1 are not plotted. Regions are dotted that are statistically significant at 5% confidence level using the t test.

FIGURE 3 Pointwise correlations between the SST anomaly and the forced response of NHF (NHF_F) from the AMIP AGCMs of (a) CCSM4, (b) GFDL, (c) GISS, and (d) CSIRO. The forced response in each AMIP model is estimated as the ensemble mean of AGCM ensemble from that model. The correlation is bias corrected using Eq. (A13). The values between -0.1 and 0.1 are not plotted. Regions are dotted that are statistically significant at 5% confidence level using the t test.

FIGURE 4 Pointwise correlations between the SST anomaly and the IAV of NHF (NHF_I) from four reanalyses: (a) ERA, (b) MERRA, (c) CFSR, and (d) NCEP. The NHF_I in each reanalysis is estimated as the NHF from that reanalysis minus the forced response from the average of the forced responses of the four AMIP ensembles. The correlation is bias corrected using Eq. (A14). The values between -0.1 and 0.1 are not plotted. Regions are dotted that are statistically significant at 5% confidence level using the t test.

FIGURE 5 Top row: SST/externally forced NHF (NHF_F) in the AMIP ensembles. (a) Standard deviation (STD) of the 4-model 20 member mean (4AMIP) NHF_F (W m^{-2}). Lower rows: Ratios of the standard deviation (STDR) of the 4AMIP NHF_F to the NHF_F in the model ensemble of (b) CCSM4, (c) GFDL, (d) GISS, (e) CSIRO. The NHF_F in each AMIP model is estimated from the mean of a 5-member AGCM ensemble from that model, while the 4AMIP NHF_F treats the 20 members of the 4 model ensembles as a single ensemble. The variances are bias corrected using Eq. (A10). The STDR in shaded areas are significantly different from 1 at the 5% level. Land is masked.

FIGURE 6 Top row: (a) The standard deviation (STD) of the unbiased 4-reanalysis mean (4Rean) NHF, and (b) the bias removed from the mean (W m^{-2}). Bottom rows: Ratio of the standard deviations (STDR) of the 4Rean NHF to the standard deviation of the individual reanalysis NHF in (c) ERA, (d)

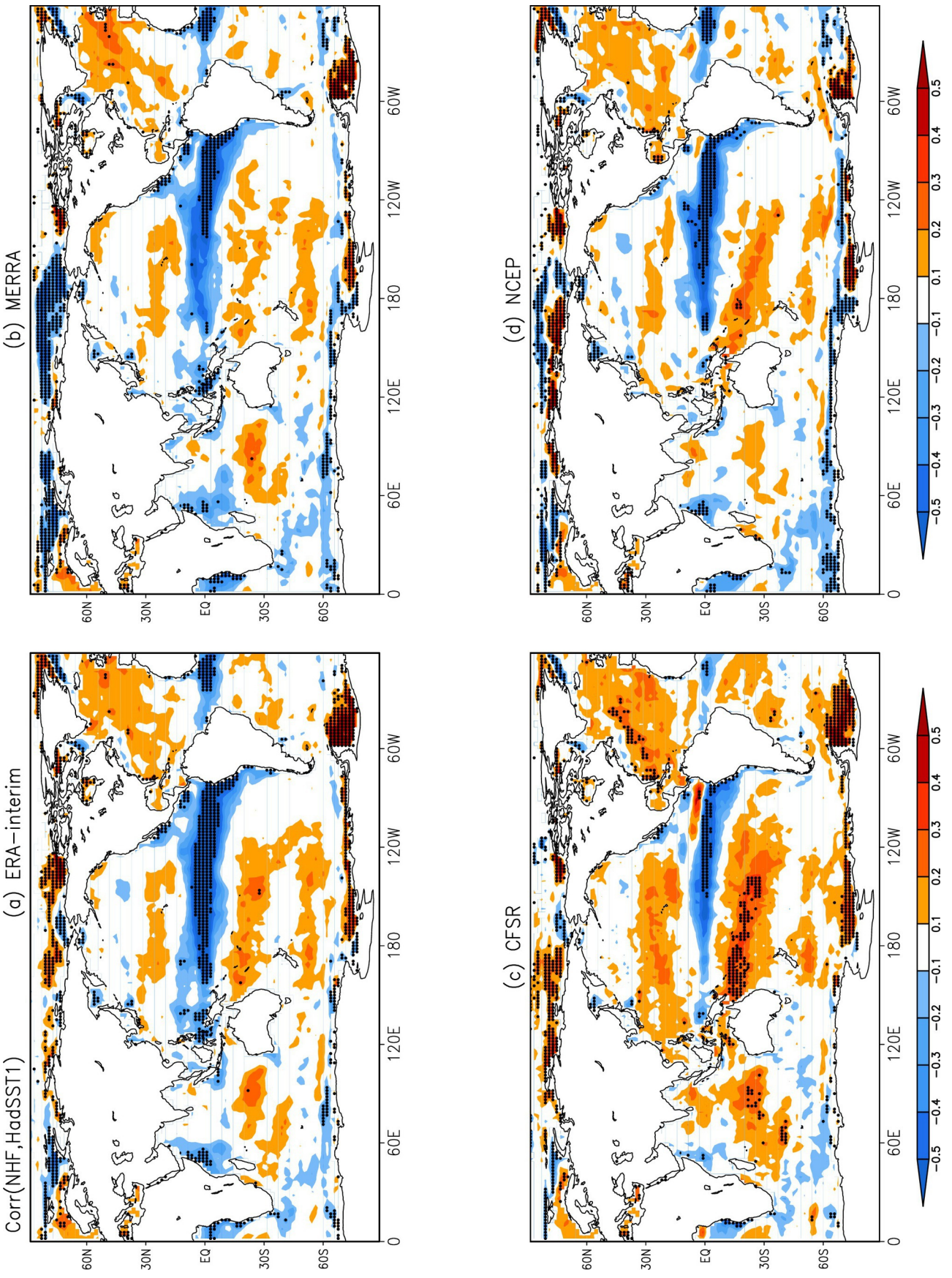
MERRA, (e) CFSR, and (f) NCEP. The variance of the 4Rean NHF is bias corrected using Eqs. (A6)-(A7). The STDR in shaded areas are significantly different from 1 at the 5% level. Land is masked.

FIGURE 7 (a) The standard deviation (STD) of NHF_I (W m^{-2}) in the 4Rean mean, estimated by removing NHF_F of 4AMIP (so 4Rean4AMIP). The IAV estimates are bias corrected using Eq. (A12). (b) The ratio of standard deviations (STDR) of the 4Rean4AMIP NHF_I to the 4AMIP NHF_I. The 4AMIP IAV is the average of the NHF_I of the 20 4AMIP ensemble members and is bias corrected using Eq. (A11). Bottom two rows: The STDR of the 4Rean4AMIP NHF_I to the NHF_I of individual model ensembles: (c) CCSM4, (d) GFDL, (e) GISS, (f) CSIRO. The STDR ratios in shaded areas are significantly different from 1 at the 5% level. Land is masked.

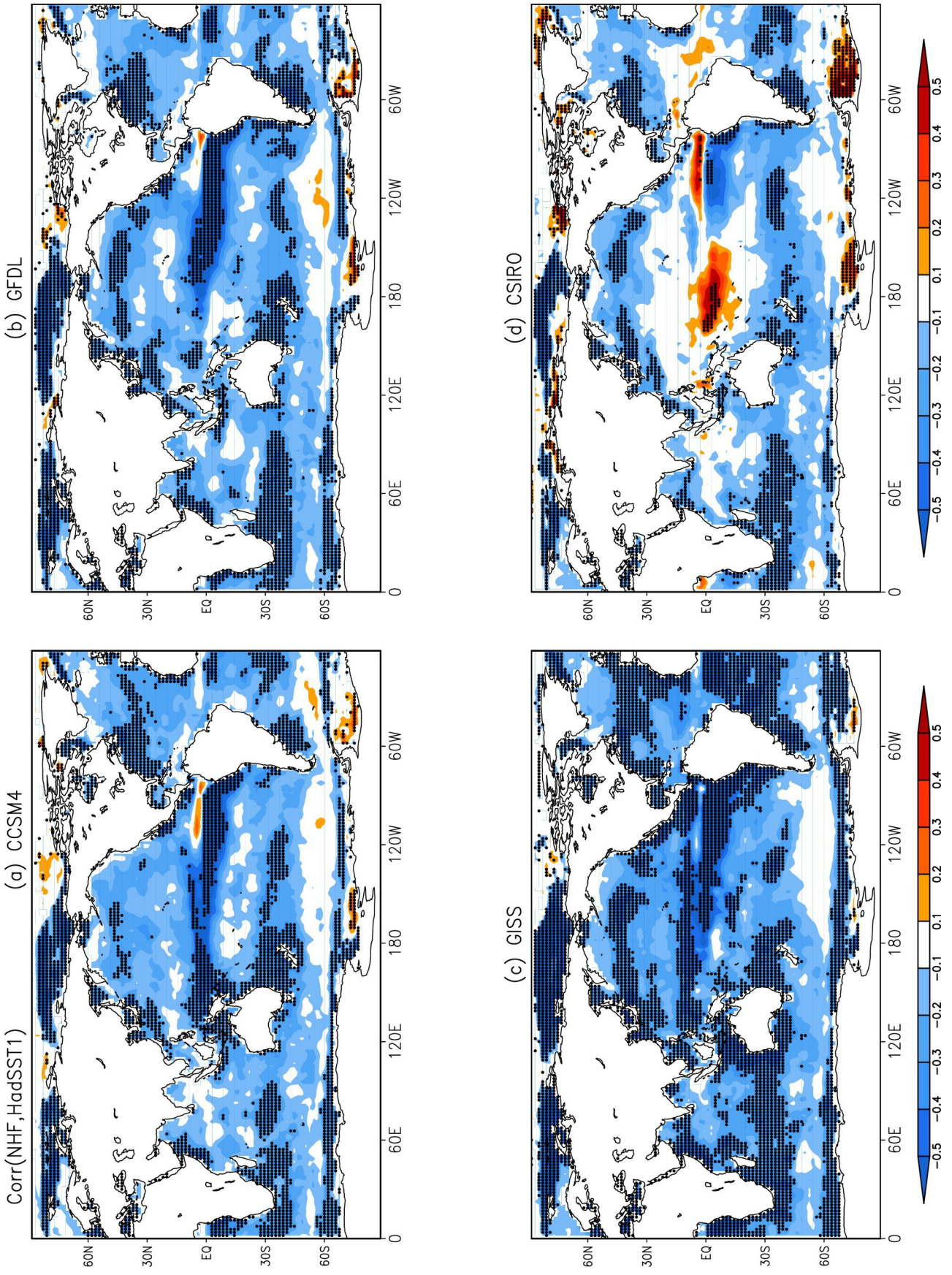
FIGURE 8 Standard deviation ratio (STDR): (a) The 4Rean NHF (Fig. 6a) to the 4AMIP NHF, computed as the square root of the average of the NHF variances of the 20 4AMIP members. (b) The 4Rean4AMIP NHF_I (Fig. 7a) to the 4Rean NHF. (c) The 4AMIP NHF_I to the 4AMIP NHF. The STDR ratios in shaded areas are significantly different from 1 at the 5% level. Land is masked.

FIGURE 9 Lagged linear regressions of the undecomposed NHF anomalies onto the normalized Atlantic multidecadal variability (AMV) index [W m^{-2} (std dev)$^{-1}$] in the (top) ERA reanalysis and (middle) CCSM4 AGCM, and (bottom) the difference between top and middle, with the AMV leading the atmosphere by one month (a-c), simultaneous with (d-f), and lagging the atmosphere by one month (g-i). The shaded regions are significant at 5% level. Land is masked.

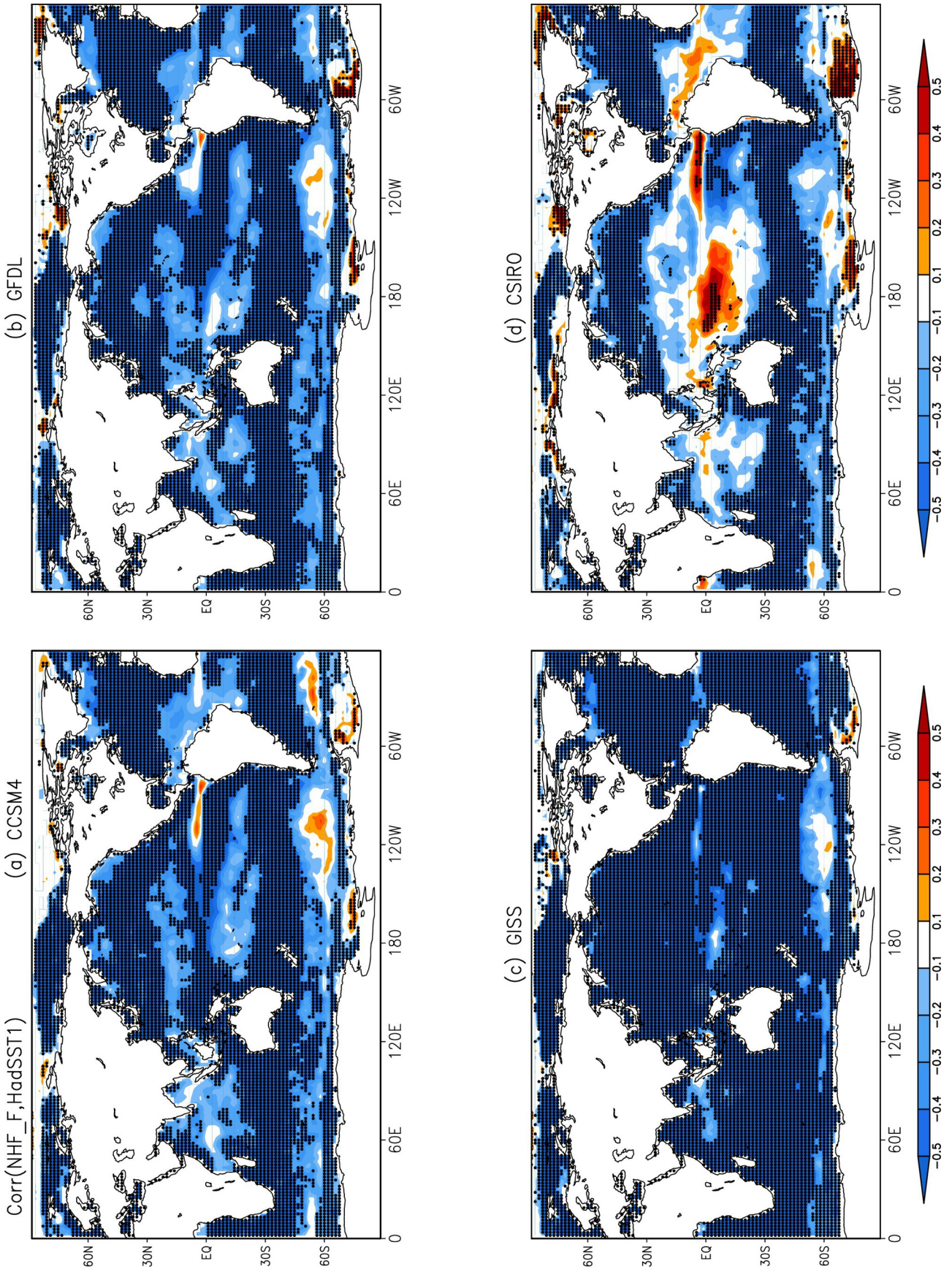
FIGURE 10 As in Figure 9, but for the regressions onto the normalized Niño3.4 index.



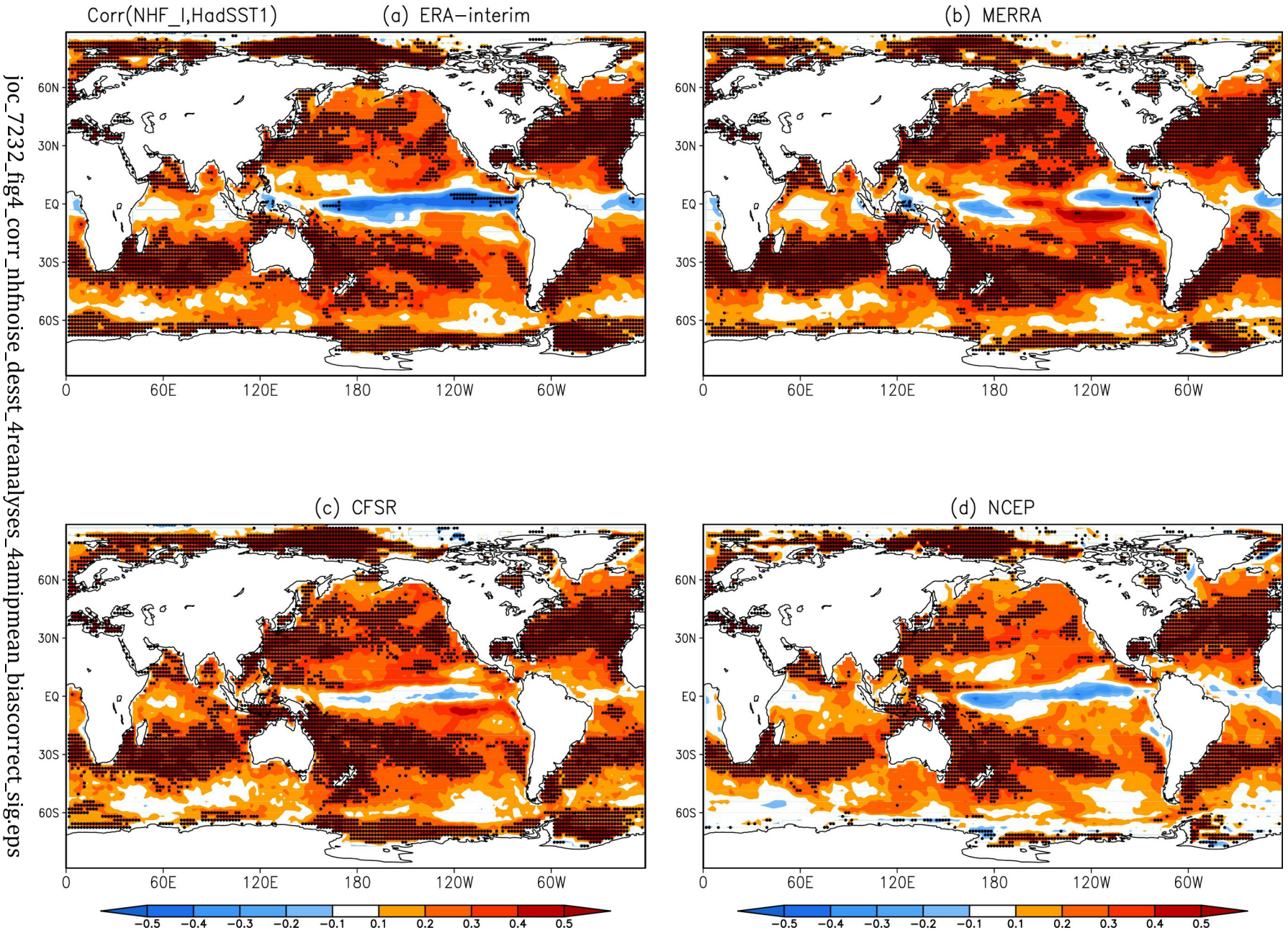
joc_7232_fig1_corr_nhf_desst_4reanalyses_sig.eps



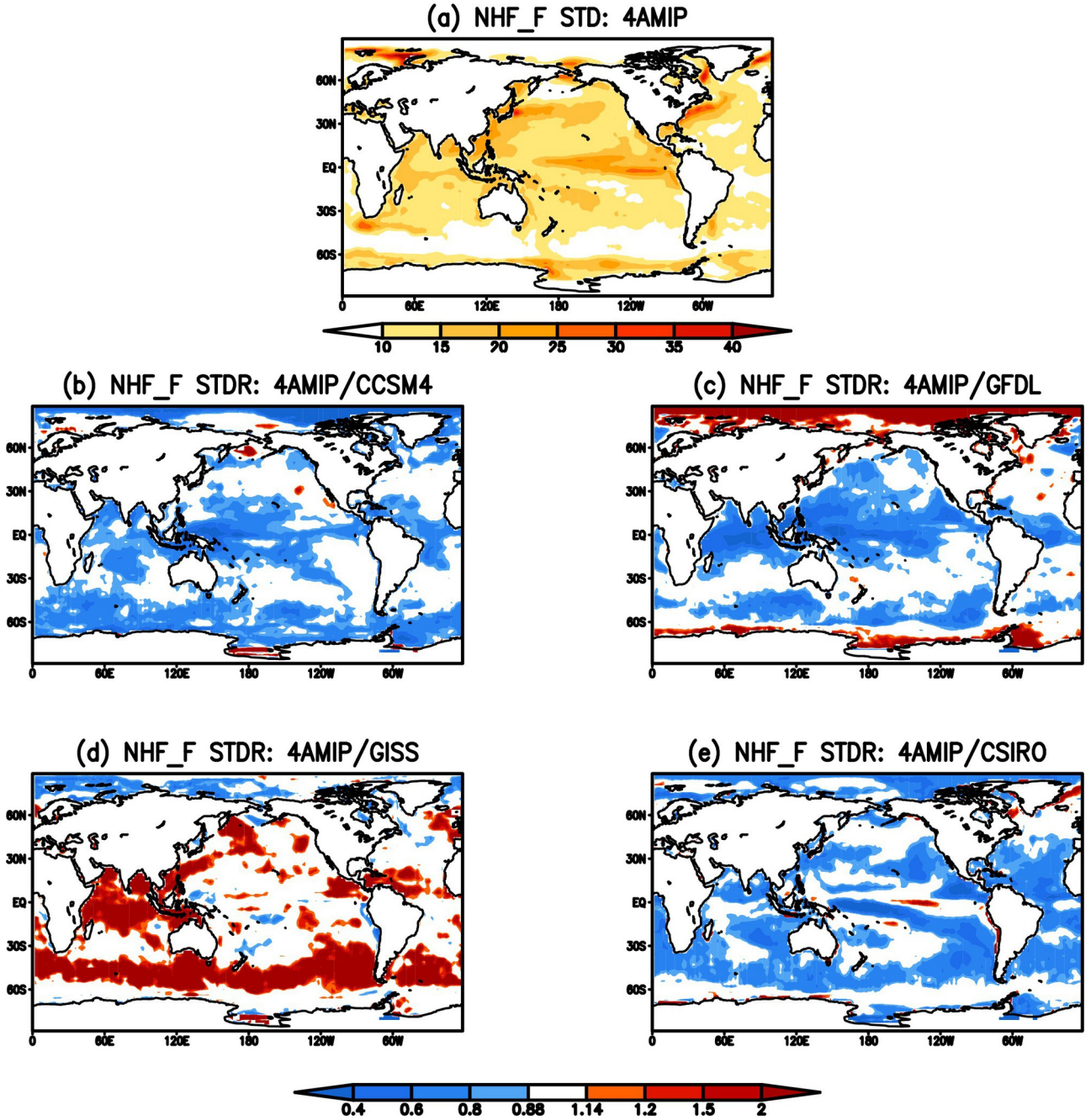
joc_7232_fig2_corr_nhf_desst_4amipmodels_sig.eps



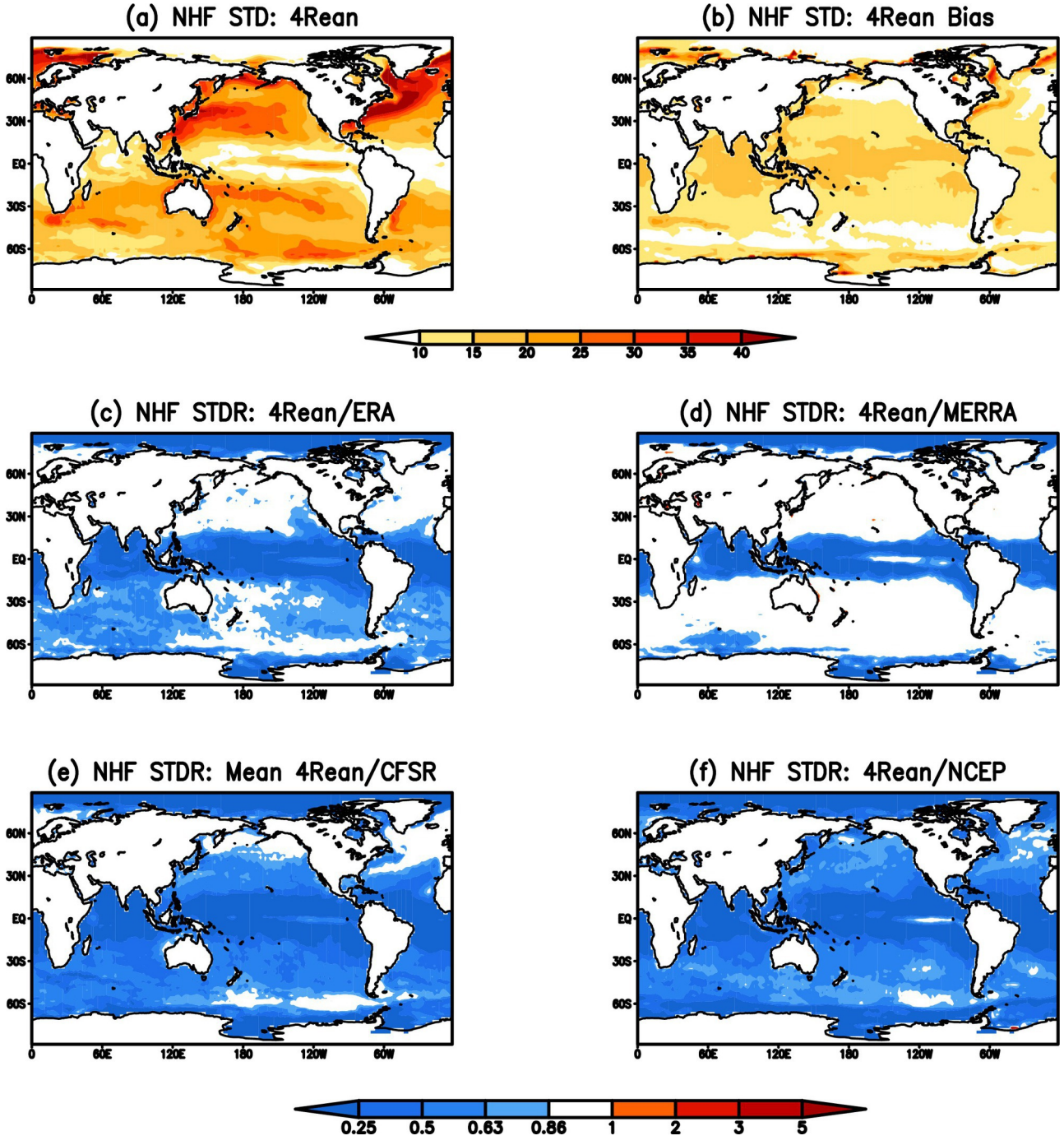
joc_7232_fig3_corr_nhfforced_desst_4amip_biascorrect_sig.eps



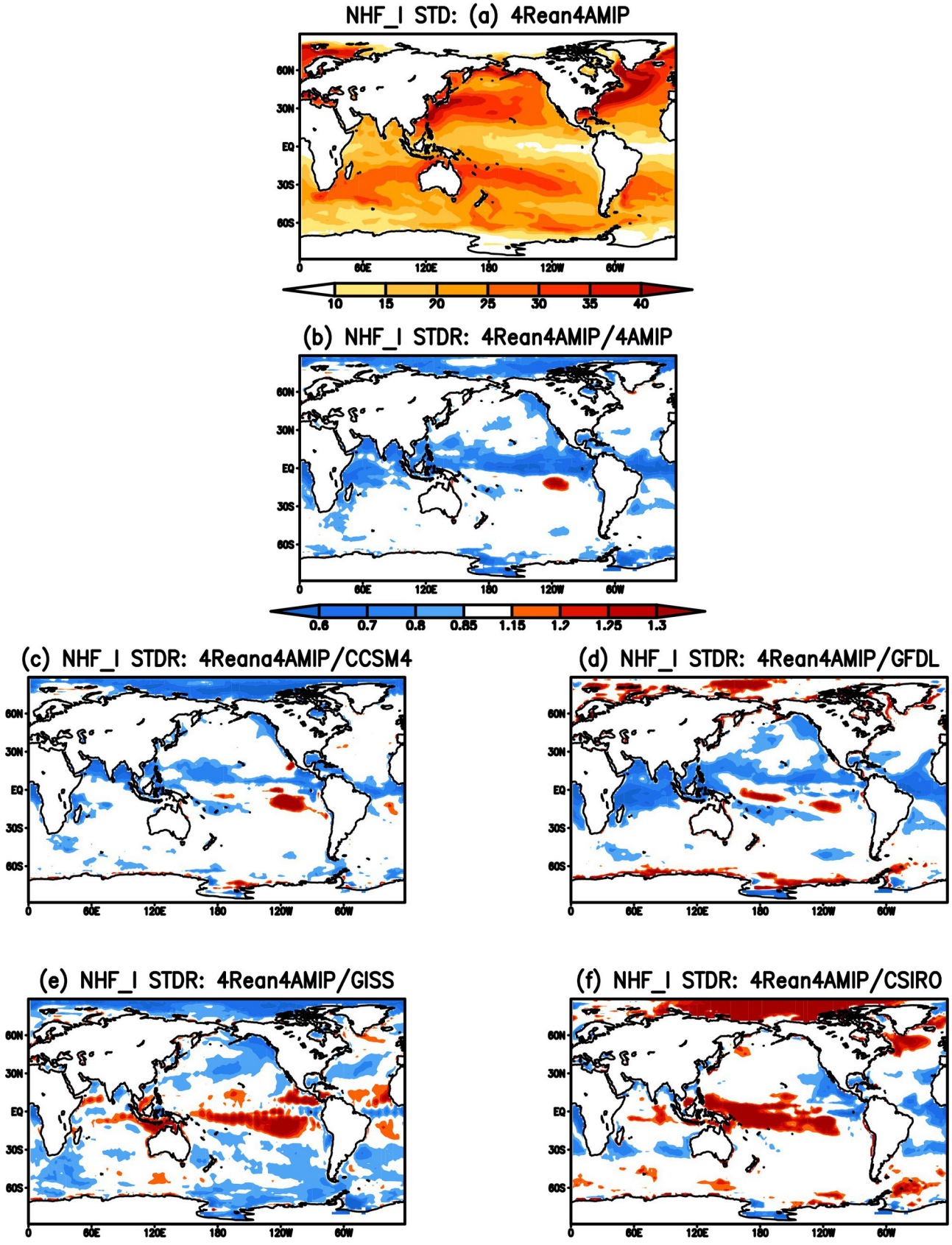
joc_7232_fig4_corr_nhfnoise_desst_4reanalyses_4amipmean_biascorrect_sig.eps



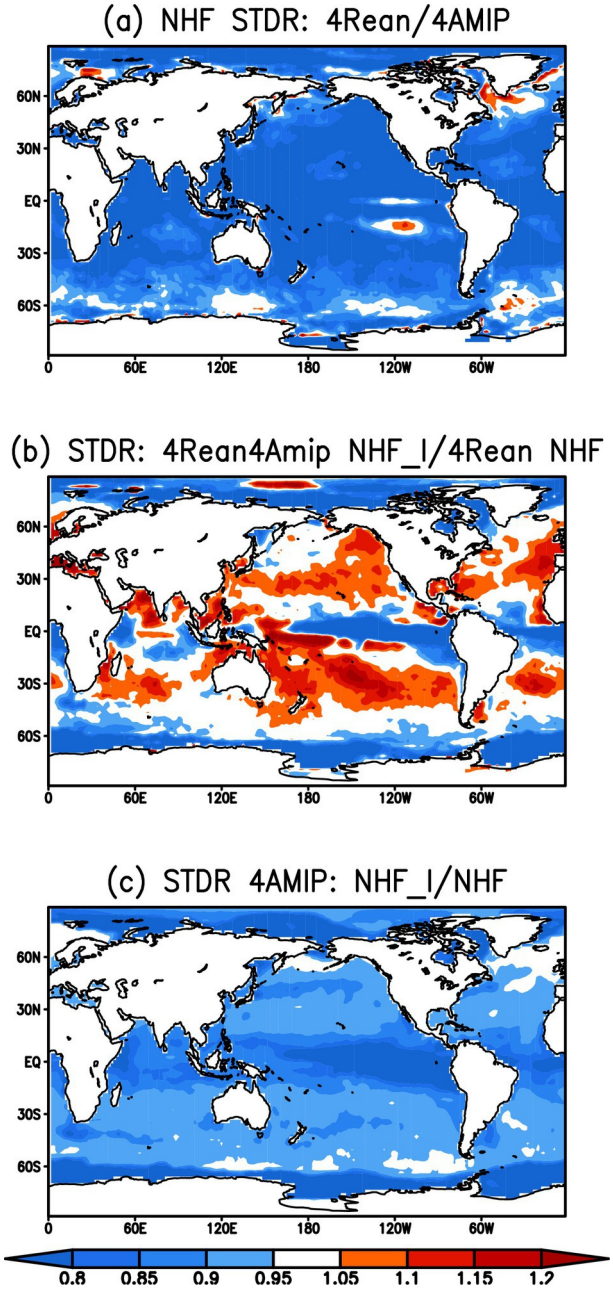
joc_7232_fig5_frcd_amip.eps



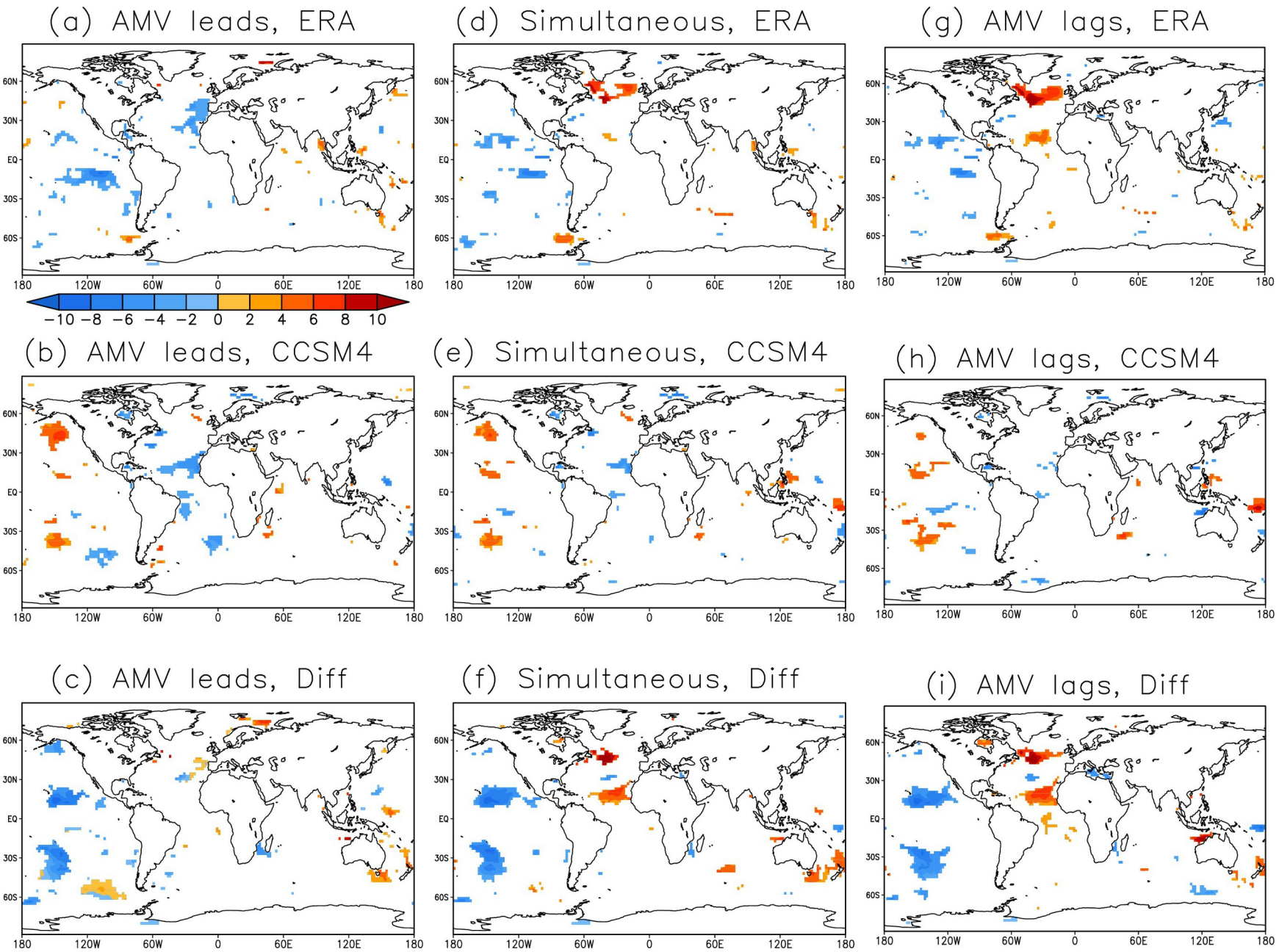
joc_7232_fig6_var_rean_nse.eps



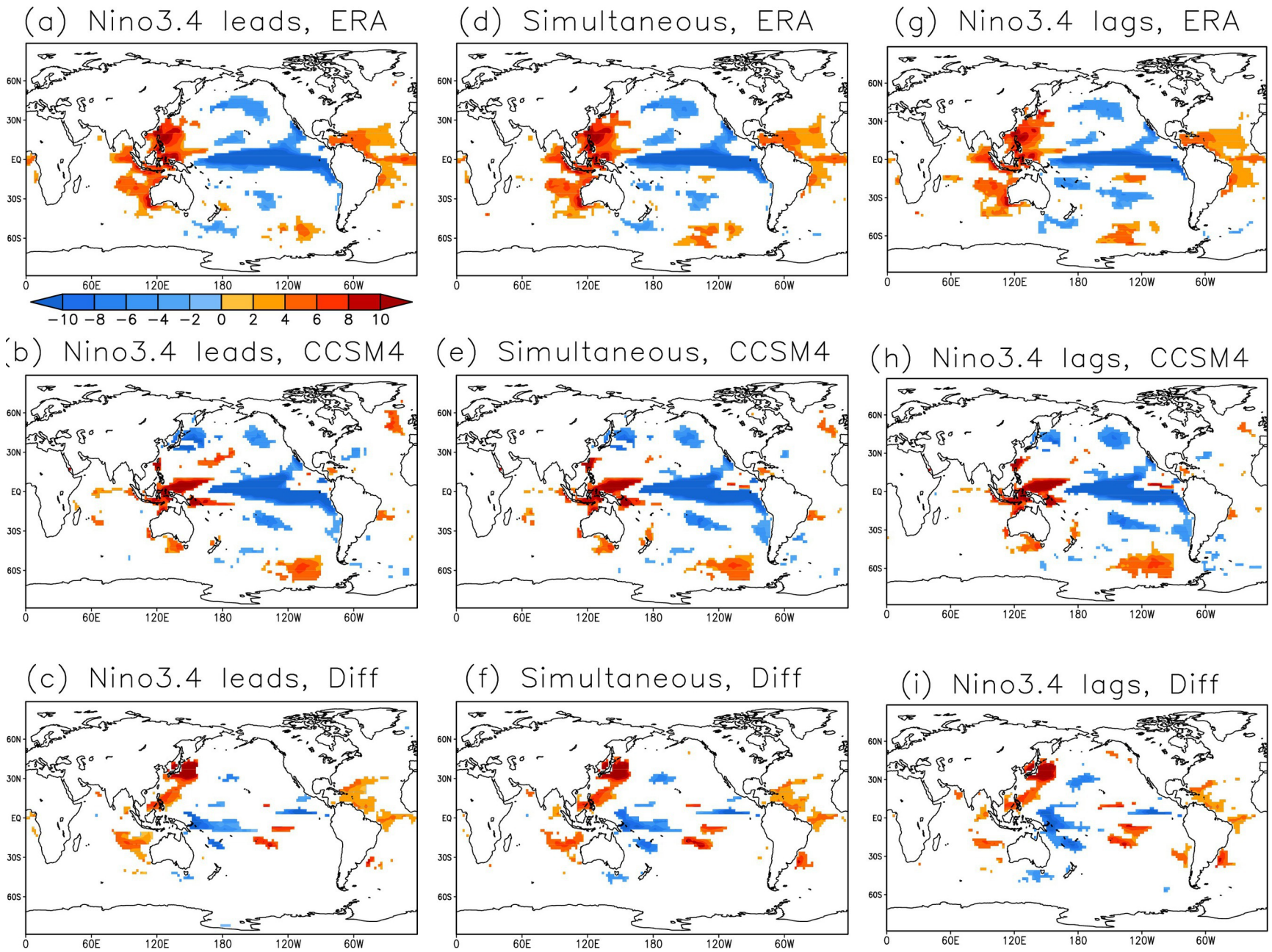
joc_7232_fig7_iav_rean_amip.eps



joc_7232_fig8_4rean_4amip.eps



joc_7232_fig9_reg_deamv_leadlag_nhf_eraint_ncar.ccsm4_sig2.eps

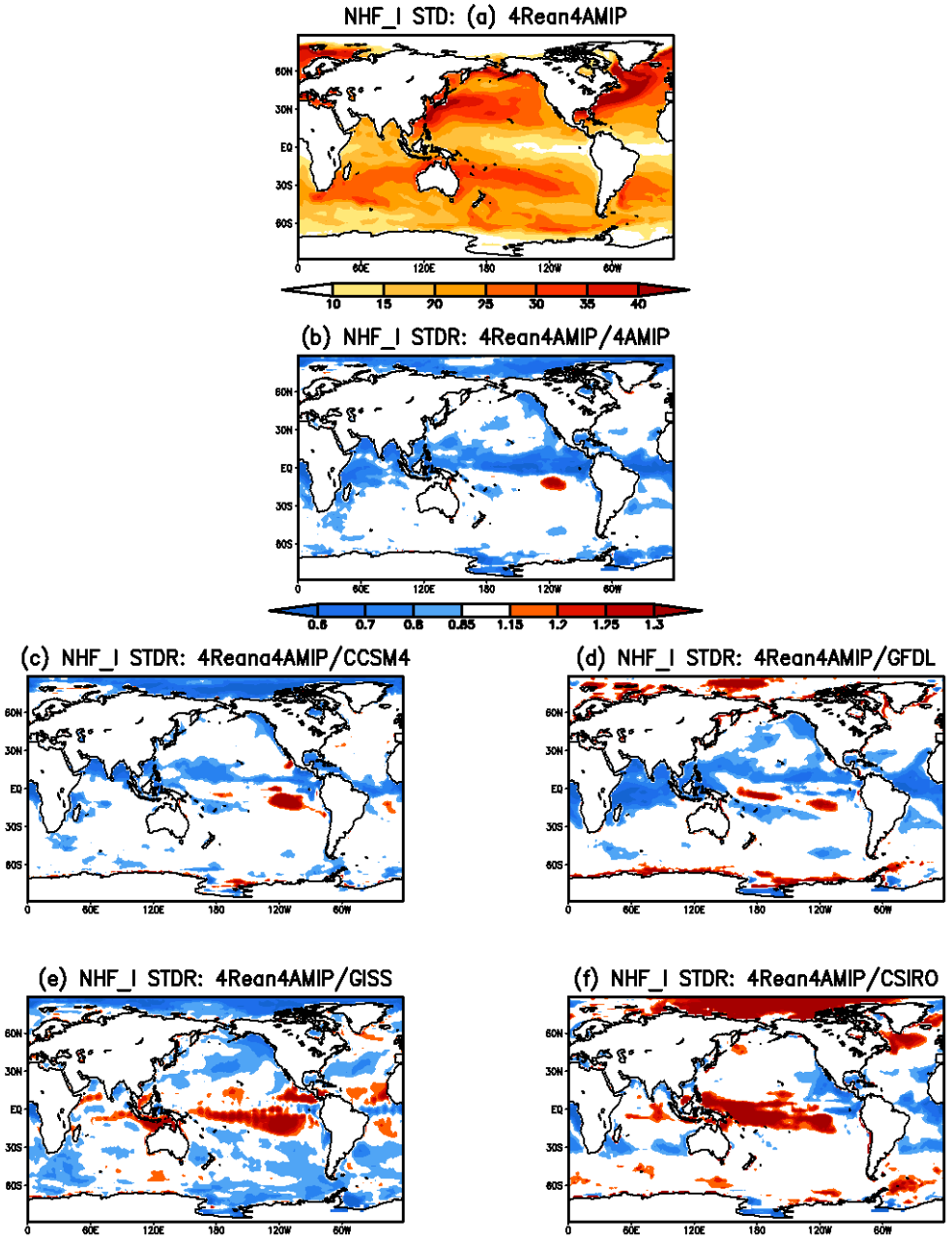


joc_7232_fig10_reg_denino34_leadlag_nhf_eraint_ncar.ccsm4_sig2.eps

Internal Atmospheric Variability of Net Surface Heat Flux in Reanalyses and CMIP5 AMIP Simulations

Hua Chen, Edwin K Schneider and Zhiwei Zhu*

The internal atmospheric variability (IAV) of the net surface heat flux is estimated as the residual after removing the SST and externally forced atmospheric response derived from AMIP simulations. IAV in the mean reanalysis plays a role in forcing the SST variability in the extra-tropics, while it may not be an important forcing in the tropical oceans. The standard deviation of IAV of AMIP models is indistinguishable from that of the mean reanalysis in the subtropics and midlatitudes.



Top row: (a) The standard deviation (STD) of NHF_I (W m^{-2}) in the 4Rean mean, estimated by removing NHF_F of 4AMIP (so 4Rean4AMIP). The IAV estimates are bias corrected using Eq. (A12). Second row: (b) The ratio of standard deviations (STDR) of the 4Rean4AMIP NHF_I to the 4AMIP NHF_I. The 4AMIP IAV is the average of the NHF_I of the 20 4AMIP ensemble members and is bias corrected using Eq. (A11). Bottom two rows: The STDR of the 4Rean4AMIP NHF_I to the NHF_I of individual model ensembles: (c) CCSM4, (d) GFDL, (e) GISS, (f) CSIRO. The STDR ratios in shaded areas are significantly different from 1 at the 5% level. Land is masked.