Error Compensation of ENSO Atmospheric Feedbacks in Climate Models and its Influence on Simulated ENSO Dynamics

Tobias Bayr[1], Christian Wengel[1], Mojib Latif[1,2], Dietmar Dommenget[3], Joke Lübbecke[1] and Wonsun Park[1]

[1] GEOMAR Helmholtz Centre for Ocean Research Kiel,
Düsternbrooker Weg 20, 24105 Kiel, Germany.
Corresponding author (tbayr@geomar.de)

[2] Cluster of Excellence "The Future Ocean", University of Kiel, 24105 Kiel, Germany

[3] School of Mathematical Sciences, Monash University, Clayton, Victoria, Australia.

Abstract

Common problems in state-of-the-art climate models are a cold sea surface temperature (SST) bias in the equatorial Pacific and the underestimation of the two most important atmospheric feedbacks operating in the El Niño/Southern Oscillation (ENSO): the positive, i.e. amplifying wind-SST feedback and the negative, i.e. damping heat flux-SST feedback. To a large extent, the underestimation of those feedbacks can be explained by the cold equatorial SST bias, which shifts the rising branch of the Pacific Walker Circulation (PWC) too far to the west by up to 30°, resulting in an erroneous convective response during ENSO events. Based on simulations from the Kiel Climate Model (KCM) and the 5th phase of Coupled Model Intercomparison Project (CMIP5), we investigate how well ENSO dynamics are simulated in case of underestimated ENSO atmospheric feedbacks (EAF), with a special focus on ocean-atmosphere coupling over the equatorial Pacific. While models featuring realistic atmospheric feedbacks simulate ENSO dynamics close to observations, models with underestimated EAF exhibit fundamental biases in ENSO dynamics. In models with too weak feedbacks, ENSO is not predominantly wind-driven as observed; instead ENSO is driven significantly by a positive shortwave radiation feedback. Thus, although these models simulate ENSO, which in terms of simple indices is consistent with observations, it originates from very different dynamics. A too weak oceanic forcing on the SST via the positive thermocline, the Ekman and the zonal advection feedback is compensated by weaker atmospheric heat flux damping. The latter is mainly caused by a biased shortwave-SST feedback that erroneously is positive in most climate models. In the most biased models, the shortwave-SST feedback contributes to the SST anomaly growth to a similar degree as the ocean circulation. Our results suggest that a broad continuum of ENSO dynamics can exist in climate models and explain why climate models with less than a half of the observed EAF strength can still depict realistic ENSO amplitude.

1. Introduction

El Niño/Southern Oscillation (ENSO) is the most prominent climate variability on interannual timescales. Its warm (cold) phase, El Niño (La Niña), is associated with warm (cold) sea surface temperature (SST) anomalies in the central and eastern equatorial Pacific (Philander 1990). ENSO has huge socio-economic impacts, as it causes extreme weather events in the Pacific region and beyond (McPhaden 1999). As ENSO is a coupled atmosphere-ocean phenomenon, atmospheric and oceanic amplifying and damping feedbacks are involved in the generation and termination of ENSO (e.g. Jin et al. 2006). The two most important atmospheric feedbacks are the positive wind-SST feedback and the negative heat flux-SST feedback (Lloyd et al. 2009, 2011, 2012), with the former prominent over the western equatorial Pacific (Niño4 region, 160°E-150°W, 5°S-5°N) and the latter over the western and eastern equatorial Pacific (Niño4 and Niño3 (90°W-150°W, 5°S-5°N) regions). The strength of both feedbacks strongly depends on the position of the rising branch of the PWC, as both are tied to the convective response to SST anomalies (SSTa) over the equatorial Pacific (Bayr et al. 2018, hereafter B18). Figure 1 schematically illustrates the relation between equatorial Pacific SST, PWC and atmospheric feedbacks. In observations, the rising branch of the PWC is roughly at 150°E and shifts to the east (west) during an El Niño (La Niña) event (Fig. 1a), which causes a positive (negative) surface wind anomaly in the Niño4 region and a negative (positive) downward net heat flux anomaly across the equatorial Pacific due to more (less) clouds and less (more) solar radiation reaching the surface (Fig. 1b).

Despite improvements during the last decades, state-of-the-art coupled general circulation models (CGCMs) still show a large diversity in simulated ENSO under present-day conditions (e.g. Bellenger et al. 2014; Vijayeta and Dommenget 2018). Also in the projections for the 21st century the ENSO response to global warming remains highly uncertain (e.g. Stocker et al. 2013; Zheng et al. 2016; Li et al. 2017; Chen et al. 2017). A cold SST bias in the western equatorial Pacific, which is a common problem of many CGCMs, has been identified to account for a large part of model diversity in both oceanic and atmosphere feedbacks (Kim et al. 2014, B18). The cold SST bias causes a La Niña-like mean state with a too westward position of the rising branch of the PWC (Fig. 1c) that is associated with too strong descent and too little precipitation over the Niño4 region (B18). As a result, the convective response to SSTa is too weak and hampers both atmospheric feedbacks (Fig. 1d). The earlier studies of Guilyardi et al. (2009a) and Kim et al. (2014b) suggested error compensation between the too weak wind response and the too weak atmospheric heat flux damping in many CGCMs. Further, it has been shown that the biased shortwave feedback contributes strongest to the underestimated net heat flux feedback on SSTa, as the shortwave feedback is too weak in the Niño4 region and can even become positive in the Niño3 region in the presence of a strong cold bias. This can be explained by an overestimation of low-level stratus clouds which dissolve, for example, when SST is rising during an El Niño event (Lloyd et al. 2009, 2011, 2012; Dommenget et al. 2014; B18).

The study of B18 focused on the underestimated ENSO atmospheric feedbacks (EAF) and their relation to the mean-state biases in SST and PWC position. Too weak EAF may also bias the simulated ENSO dynamics, as suggested in Dommenget et al. (2014). This study further investigates the biases in ENSO dynamics in models with underestimated EAF. Three different methods are used: first, we use the framework of the Bjerknes feedback (Bjerknes 1969), which describes the basic positive feedback in ENSO. Second, we apply the Bjerknes Stability (BJ) index (Jin et al. 2006), which is a more sophisticated method and scales the positive and negative feedbacks in a way that they are directly comparable to each other. As a third method we exploit a Slab-Ocean SST calculated offline from the CGCM data to get an estimate of how much of the SST tendency over the tropical Pacific is caused by the heat fluxes.

The aim of this study is to enhance the understanding of how ocean-atmosphere coupling and ENSO dynamics are captured in climate models with underestimated EAF and to explain why climate models with a strongly underestimated wind-SST feedback still can depict realistic ENSO amplitude. We investigate the same multi-model ensemble of phase 5 of the Coupled Model Intercomparison Project (CMIP5) and perturbed physics ensemble of the KCM as in B18. This manuscript is organized as follows: In section 2 we give an overview of the data and methods used in this study, and in section 3 we analyze the atmospheric feedbacks. The Bjerknes feedback is investigated in section 4, while a more detailed analysis of the positive and negative feedbacks in ENSO is given by means of the BJ index in section 5. In section 6, we analyze the relative roles of wind and heat flux feedback in ENSO dynamics, and the effect of error compensation among the EAF on ENSO amplitude is addressed in section 7. A summary and discussion of the major results are given in section 8.

2. Data and Methods

A perturbed physics ensemble is performed with the Kiel Climate Model (KCM) (Park et al. 2009), the same as used in B18 (Tab. 1A in B18). KCM consists of the ECHAM5 atmosphere model (Roeckner et al. 2003) and the NEMO ocean-sea ice model (Madec et al. 1998; Madec 2008). ECHAM5 is run with a T42 horizontal resolution (~2.8°). NEMO is integrated on a 2°-Mercator mesh (ORCA2), with enhanced meridional resolution of 0.5° in the equatorial region and 31 vertical levels. In total we performed 40 “present day” integrations (employing an atmospheric CO_2-concentration of 348 ppm) that differ in vertical resolution (19, 31, and 62 levels, all have 10 hPa as the top level) and in the parameters of the convection scheme. The “convective mass-flux above level of non-buoyancy”, “entrainment rate for shallow convection” and “convective cloud conversion rate from cloud water to rain” are varied. All three parameters can be used to tune climate models, as discussed in detail in Mauritsen et al. (2012) and the chosen values lie within the suggested range. In each sensitivity experiment the ocean is initialized from Levitus climatology (Levitus et al. 1998). The experiments are 100 years long and the last 80 years were analyzed. The vertical resolution only has a small influence on the EAF strength (B18). Further, B18 could show by dedicated atmosphere-only experiments that the chosen parameters only have a minor direct effect on EAF strength, while the indirect effect by changing the equatorial SST bias explains most of the spread in EAF in the KCM perturbed physics ensemble.

For comparison two types of atmosphere-only experiments are performed: In “AMIP-type” experiments we force ECHAM5 by observed monthly SSTs (1980 - 2009) from HadISST using the three vertical resolutions (19, 31 and 62 levels) (Tab. 1B in B18). In the “KCM Biased-Slab-Ocean experiment” (Tab. 1E in B18) ECHAM5 is coupled to a slab ocean that is controlled to mimic a SST climatology with a large equatorial Pacific cold SST bias (see Dommenget, 2010 for details). In such an experiment Dommenget (2010) found a heat flux-driven El Niño-like SST variability.

From the CMIP5 database (Taylor et al. 2012) we use historical simulations (1900-1999) and, if available, the AMIP experiments with these models (see Tab. 1 for a list of the models). For consistency, we choose here the same CMIP5 models as in B18, but the oceanic data was only available for 21 of the 24 models used in B18 and AMIP experiments for 18 of the 21 models (Tab. 1). The data is interpolated on a regular 2.5°×2.5° grid.

We use observed SSTs for the period from 1958 to 2015 from HadISST (Rayner et al. 2003). From ERA40 reanalysis (Uppala et al. 2005) and ERA-Interim reanalysis (Simmons et al. 2007) we use for the period from 1958 to 2001 and from 1979 to 2015, respectively, the 10m zonal wind (U10), zonal wind stress (τ_x) and surface heat fluxes. Observed subsurface ocean temperatures for the period from 1979 to 2015 is taken from HadEN4 (Good et al. 2013), and subsurface ocean temperatures and velocities for the period from 1958 to 2001 from SODA 2.0.2 reanalysis (Carton and Giese 2008). The thermocline depth is defined as the depth of the 20°C isotherm. As ENSO characteristics vary on decadal timescales (Lübbecke and McPhaden 2014; Guan and McPhaden 2016) and observed records are short, observational estimates of ENSO strength are subject to some uncertainty. However, ENSO characteristics derived from the different observational datasets and reanalysis products do not strongly differ among each other, as shown below.

We consider monthly-mean values with the climatological seasonal cycle removed, and all data is linearly detrended for each month separately. The wind-SST feedback is calculated here by linear regression of zonal wind stress anomalies in the Niño4 region on SSTa in the Niño3.4 region (120°W-170°W, 5°S-5°N). As this study focuses on the coupled ocean-atmosphere ENSO dynamics we use zonal wind stress (τ_x) instead of 10m zonal wind (U10) as in B18, as it is the wind stress that drives ocean circulation. Data from ERA-Interim and ERA40 suggest that a τ_x-feedback of 100 Pa/K corresponds to a U10-feedback of roughly 1 m/s/K (as indicated by the dashed line in Fig. 2a,b). The perturbed physics ensemble of KCM has a stronger τ_x feedback than U10 feedback (all experiments are below the dashed line). In the CMIP5 multi-model ensemble, there is some spread in this relation (Fig. 2b): some models have roughly the same τ_x and U10 feedback and other models exhibit nearly a factor 2 difference (e.g. MRI-CGCM3 and IPSL-CM5A-MR, see also Tab. 1). The reason for this difference between the τ_x feedback and U10 feedback arises from the different τ_x calculation in the climate models, which origin in different drag coefficient parametrizations or non-linear effects (e.g. Zhai and Greatbatch 2007). Further, it is important to note that using fixed regions to calculate EAF may not capture the nature of the feedbacks in the models, as the feedbacks may operate in a different region in comparison to observations. An alternative would be the use of individual boxes for each model. As both methods have advantages and disadvantages, we use fixed boxes, because this is the more conservative.

From the output of the models we calculate the Bjerknes feedback, which has three elements: the zonal wind stress response over the western equatorial Pacific to SSTa in the eastern equatorial Pacific, the thermocline response to zonal wind stress changes, and the response of eastern equatorial SSTa to local thermocline perturbations. As a quantitative measure of the positive and negative feedbacks operating in ENSO the Bjerknes Stability (BJ) index is used. The BJ index derivation is described in Jin et al. (2006) and we apply modifications following Wengel et al. (2018b). The BJ index consists of three positive feedbacks, the zonal advection feedback, the Ekman feedback and the thermocline feedback, and two negative feedbacks, the dynamical damping and thermal damping. The three positive feedbacks describe the sensitivity of the zonal ocean currents, upwelling and thermocline tilt, respectively, to zonal wind stress changes and their impact of SST. The two negative feedbacks represent the damping of SSTa by the mean ocean currents and atmospheric heat fluxes (see Table 3 in Wengel et al. 2018b for a detailed formulation of the BJ index). The BJ index is defined as the sum of all positive and negative feedbacks. It is derived here from ERA40/SODA reanalyses and the individual KCM experiments separately for each calendar month. It is important to note that the BJ index is based on linear regression and thus cannot capture nonlinear ENSO dynamics (Graham et al. 2014). The nonlinearity of ENSO, however, is underestimated in the majority of current climate models (Bellenger et al. 2014; Karamperidou et al. 2017; Timmermann et al. 2018), especially in climate models with weak EAF (B18). Further, a major advantage of the BJ index is that the individual feedbacks are scaled such that they can be directly compared to each other. Therefore the BJ index is an appropriate tool for our purpose, as it is able to capture the main aspects of ENSO dynamics.

To derive a quantitative measure of how similar the simulated ENSO is to the heat flux-driven El Niño-like variability in the Biased-Slab-Ocean experiment from Dommenget (2010), we mimic a slab ocean SST for the CGCM runs with the KCM and the CMIP5 models (i.e. the SST in the absence of anomalous ocean circulation). We integrate the net surface heat flux, similar to Drews and Greatbatch (2016), but start six month before the maximum SSTa of the ENSO event:

$$dSST = \frac{1}{c_p \cdot \rho \cdot H} \int_{t=-6}^{t=0} Q_{net} dt$$

Here c_p = 4000 J kg^{-1} K^{-1} is the specific heat capacity at constant pressure of sea water, ρ = 1024 kg m^{-3} the average density of sea water, H = 50 m the depth of the slab ocean and t the time in months. We integrate the 6 months before the ENSO event, as this is the average growth period in HadISST, CMIP5 and KCM and we do the integration for each ENSO event individually, before averaging over all ENSO events. The difference in the 6 months before the peak of the ENSO event between the simulated SST change in the CGCMs and the change in the calculated slab ocean SST provides an estimate of how much of the SST change is caused by an anomalous ocean circulation. Similar calculations are performed separately for the net surface shortwave radiation, net surface longwave radiation and surface sensible and latent heat fluxes to obtain the main contributors to the heat flux-driven SST change.

ENSO events are defined using the criterion of Trenberth (1997): an El Niño (La Niña) event occurs if the 5-month running mean SSTa averaged over the Niño3.4 region is above 0.5 (below -0.5) times the standard deviation for at least six consecutive months. We define the maximum of an El Niño (La Niña) event for each event individually as the month of maximum (minimum) in 5-month running mean Niño3.4 SSTa.

3. Atmospheric feedbacks

In this section, we look at the strength of the atmospheric feedbacks in both the KCM perturbed physics ensemble as well as the CMIP5 experiments and their relation to the equatorial Pacific SST bias. The wind feedback, which is the atmospheric component of the Bjerknes feedback, describes the wind stress response over the western part of the basin to an SSTa in the central-to-eastern equatorial Pacific. There is a strong anticorrelation between the wind feedback and the heat flux feedback in the perturbed physics ensemble of the KCM (Fig. 3a), consistent with B18 where U10 was used. The strong anticorrelation of -0.93 indicates an error compensation between the two feedbacks. In the CMIP5 multi-model ensemble, the anticorrelation between the two atmospheric feedbacks is smaller (Fig. 3c) and amounts to -0.60 (compared to -0.76 in B18 using U10), but still significant. This weaker correlation can be explained by the differences among the models in τ_x calculation from U10 (Fig. 2), including the effect of non-linearities in the τ_x calculation. The red, blue and green color of the numbers in the scatter plots (e.g. Fig. 3) denote the three sub-ensembles consisting of models with STRONG, MEDIUM and WEAK atmospheric feedbacks, respectively. The sub-ensembles are defined according their total EAF strength (x-axis in Fig. 3b,d), i.e. the average of the wind and heat flux feedback after normalizing each by the observed value. In the STRONG sub-ensemble the individual members have EAF larger than 55% of the observed total EAF strength, in the WEAK sub-ensemble the members have feedbacks smaller than 35%, and in the MEDIUM sub-ensemble they are in between (Tab. 1). These three sub-ensembles also are used in the following.

The strength of total EAF is underestimated in all climate models, and it has a significant relationship to the SST bias in the Niño4 region (Fig. 3b,d). This link is weaker in the CMIP5 ensemble than in the KCM ensemble, which is somehow expected given that the CMIP5 models differ in many more aspects than the KCM experiments, such as different atmosphere and ocean models with different physical parametrizations and resolutions. We note that ECHAM5, the atmosphere model used in the KCM, forced by observed SSTs has a heat flux feedback strength comparable to observations and overestimates the wind feedback, and that both feedbacks increase in strength with higher vertical resolution (Fig. 3a: downward, sideward and upward pointing magenta triangles for L19, L31, L62, respectively). The KCM Biased-Slab-Ocean experiment (cyan circle) yields the smallest atmospheric feedbacks (Fig. 3a) and the largest cold SST bias (Fig. 3b). There is a large spread in total EAF strength among the CMIP5 models (Fig. 3d). Yet none of the coupled models depicts a total EAF strength as strong as in reanalysis, even if they exhibit no equatorial cold bias or even a warm bias. A similar spread in EAF strength as in CMIP5 is obtained with the KCM, which can be traced back to the cold SST bias (B18). Further, it is important to note that ECHAM5 by itself can generate a large spread in EAF strength. When driven with observed SSTs, as in the AMIP-type runs, ECHAM5 depicts similar feedback strengths as ERA-Interim. On the other hand, with a large superimposed cold SST bias ECHAM5 exhibits very weak EAF, that generate a purely heat flux-driven El Niño-like SST variability, as a slab ocean contains by definition no ocean dynamics (Dommenget 2010).

4. Bjerknes Feedback

To investigate how the coupled feedbacks operating in ENSO are simulated in the presence of underestimated EAF, we also analyze the other two components of the Bjerknes feedback. Figure 4 shows all three components of the Bjerknes feedback calculated from observations/reanalysis products and the three KCM sub-ensembles (as denoted by the red (STRONG), blue (MEDIUM) and green (WEAK) numbers in Fig. 3a). In the observations the strongest wind-SST feedback is located in the Niño4 region, with an average regression coefficient of 1.30 10^{-2} Pa/K, where it explains 40% of the wind stress variability linked to the SST variability in the Niño3 region (Fig. 4a). The regression of the thermocline depth (Z20) anomalies on the τ_x anomalies in the Niño4 region shows the largest positive regression values and explained variances in the Niño3 region (Fig. 4b). We note the regressions of opposite sign over the western equatorial Pacific with extremes off the equator. Overall, the anomaly structure in Z20 is reminiscent of wind stress-forced Rossby and Kelvin wave modes which drive changes in thermocline tilt. The local regression of SSTa on Z20 anomalies depicts large positive values and explained variances in the eastern equatorial Pacific (Fig. 4c), which is expected since SSTa is strongest coupled to Z20 in this region due to the shallow thermocline.

In the KCM sub-ensembles the wind-SST feedback decreases from STRONG to WEAK, as do the explained variances (Fig. 4d,g,j). Members of WEAK underestimate the wind-SST feedback strength and explained variances by more than 50%. The link between the wind stress anomalies in the Niño4 region and Z20 anomalies also weakens from STRONG to WEAK, as expressed by the explained variances (Fig. 4e,h,k). The local relationship between SSTa and Z20 anomalies in the east weakens from STRONG to WEAK (Fig. 4f,i,l). At the same time the relation between SSTa and the thermocline anomalies in the Niño4 region becomes more negative and significant. This untypical behavior (SST gets warmer when Z20 gets shallower) can be explained by the westward propagation of the SST signal during ENSO events in the WEAK sub-ensemble (Fig. 3fkp in B18), which is similar to the heat flux-driven El Niño-like variability in the KCM Biased-Slab-Ocean experiment and thus independent of Z20. An overall similar picture with respect to all three components of the Bjerknes feedback is obtained from the three sub-ensembles derived from the CMIP5 models (Fig. 5).

To underline the findings obtained from the KCM and CMIP5 sub-ensembles (Fig. 4,5), we analyze the Bjerknes feedback in the individual models. Regression coefficients and explained variances are shown for the three components of the Bjerknes feedback (Fig. 6), averaged over the Niño4 region for the wind-SST feedback and over the Niño3 region for the thermocline-wind and the SST-thermocline feedback, as indicated by the black boxes in Fig. 4,5. The feedback strengths and explained variances calculated for all individual KCM experiments and CMIP5 models confirm the results found in the sub-ensembles: In models with a stronger wind-SST feedback also the explained variance of the regression is larger (Fig. 6a,e), indicating that the variability in SST and τ_x is more determined by the wind-SST feedback than in models with weaker EAF. Further, in both ensembles models with stronger EAF tend to have a stronger thermocline-wind and SST-thermocline feedback, and the explained variance of the regression is larger (Fig. 6b,c,f,g). This becomes clearer when considering the total Bjerknes feedback strength and the averaged explained variance (Fig. 6d,h), which are defined as the product of the three regression coefficients and the arithmetic mean of the individual explained variances, respectively. Clearly, models with weaker (stronger) EAF tend to simulate a weaker (stronger) total Bjerknes feedback strength and smaller (larger) averaged explained variance, with correlations of 0.83 and 0.90 in the KCM and CMIP5 ensemble, respectively. As more of the variability in τ_x, Z20 and SST can be explained by the three components of the Bjerknes feedbacks in models with stronger EAF, this suggests that the three variables are more strongly linked to each other by the Bjerknes feedback in climate models with stronger EAF.

5. Bjerknes Stability Index

The Bjerknes Stability Index (BJ index), a measure for the SSTa growth rate, allows a more detailed analysis of the positive and negative feedbacks operating in ENSO, as the feedbacks are scaled in a way that they can be directly compared to each other. The BJ index is calculated for each calendar month separately and Figure 7 depicts the zonal advection feedback (ZAF), Ekman feedback (EF), thermocline feedback (TF), dynamical damping (DD), thermal damping (TD) and the sum of all five feedbacks, the BJ index, calculated from ERA40/SODA reanalysis as well as the KCM STRONG, MEDIUM and WEAK sub-ensembles. Relative to reanalysis, all three positive feedbacks are underestimated in the annual mean in the three sub-ensembles of the KCM. In the WEAK sub-ensemble, EF and TF and to a lesser extent ZAF are more strongly underestimated than in MEDIUM or STRONG. The DD is overestimated in the KCM, but there is virtually no difference among the three sub-ensembles. The TD, on the other hand, is most strongly underestimated in WEAK while it is close to the value derived from reanalysis in STRONG, as expected from Fig. 3. The small deviations between the TD and heat flux feedback shown in Fig. 3a are due to the different spatial domains. The

BJ index (Fig. 7f) is very similar for the three sub-ensembles of the KCM, illustrating that the individual positive and negative feedbacks, which exhibit noticeable differences among the sub-ensembles, are error compensating. Too weak forcing by ZAF, EF and TF is compensated by too weak TD, resulting in a quite similar BJ index (i.e. SSTa growth rate) in the three sub-ensembles.

The largest differences in the positive feedbacks between the three sub-ensembles of the KCM are observed during September to February (SONDJF) while those in the negative feedbacks occur during January to May (JFMAM). In SONDJF, TF contributes most to the underestimated positive feedbacks (difference between STRONG and WEAK is 1.1 yr^{-1}), EF only half as much as TF (difference between STRONG and WEAK is 0.5 yr^{-1}) and ZAF only little (difference between STRONG and WEAK is 0.1 yr^{-1}). In JFMAM, the difference in TD between STRONG and WEAK is of similar size as in TF (-1.2 yr^{-1}).

In the KCM, the strengths of ZAF, EF, TF and TD are strongly related to the equatorial Pacific SST bias (Fig. 8). The significant correlations of 0.86, 0.61, 0.75 and -0.84 between the equatorial Pacific SST bias and ZAF, EF, TF and TD, respectively, suggest that the strength of all four feedbacks strongly depends on the equatorial SST bias. It is important to note that also other factors can bias the individual feedbacks, as the EAF are already biased in AMIP experiments (Lloyd et al. 2011; Li et al. 2015; Ferrett et al. 2017a,b). Nevertheless, the KCM results further support the finding by B18 that a substantial part of the error compensation in climate models can be attributed to their equatorial cold bias.

Consistent with Wengel et al. (2018b) and B18, the BJ index results shown in Fig. 7 help to explain why climate models with too weak EAF have problems in simulating the phase locking of ENSO to the seasonal cycle: the ENSO phase locking can be explained by the positive feedbacks being strongest in boreal autumn and winter and the negative feedbacks being strongest in boreal spring, as derived from reanalysis products (ERA40/SODA). In the WEAK sub-ensemble of the KCM, TD is most strongly underestimated during the first half and ZAF, EF and TF during the second half of the year, resulting in too little seasonal variation of the BJ index.

6. Wind-driven vs. heat flux-driven ENSO dynamics

In the previous sections, we have shown that in models with a strong equatorial cold SST bias the wind-driven ENSO dynamics are considerably weaker compared to observations. The question arises if the too weak wind forcing is compensated by other processes, as climate models with a strongly underestimated wind-SST feedback can still exhibit realistic ENSO amplitude (Bellenger et al. 2014). The studies of Dommenget (2010) and Dommenget et al. (2014) suggest that the simulated ENSO in models with a large equatorial cold SST bias can, at least to some extent, be driven by a positive heat flux-SST feedback caused by a positive shortwave (SW) radiation-SST feedback over the eastern tropical Pacific. A positive SW-SST feedback is possible when low-level stratus clouds are overestimated and dissolve when the SST is rising and vice versa – a phenomenon that in observations is only found close to the South American coast (Lloyd et al. 2009). This process exists further off the coast in climate models with a large equatorial cold SST bias (Lloyd et al. 2009, 2011, 2012, B18). A large negative correlation between the wind-SST feedback strength in the Niño4 region and the SW-SST feedback in the Niño3 region is observed in the KCM ensemble as well as in the CMIP5 models (Figure 9). As shown in B18, the position of the rising branch of the PWC determines the strength of both feedbacks. A gradual change in the ENSO dynamics with increasing equatorial cold SST bias is indicated in climate models, as a decreasing ocean-atmosphere coupling by a weaker wind-SST feedback is compensated by a decreasing SW damping. The SW feedback can even shift from negative, i.e. damping to positive, i.e. amplifying in the presence of a large enough cold SST bias. In such a model, ENSO may largely become heat flux-driven.

To obtain a measure of the similarity of the ENSO simulated in the KCM and CMIP5 ensembles and the heat flux-driven ENSO in the Biased-Slab-Ocean experiment of Dommenget (2010), we calculate an offline Slab Ocean SST change by integrating the coupled models' heat fluxes over the six months prior to the maximum of an ENSO event (see section 2). We normalize the offline Slab Ocean SST change by the full SST change, which yields the heat flux-driven SST change per full SST change (Fig. 10). The difference between the offline Slab Ocean SST change and full SST change gives us an estimate how much of the warming is caused by ocean dynamics. We can test the approach with the Biased-Slab-Ocean experiment, and this yields roughly +1 K heat flux-driven warming per K warming during an ENSO event (Fig. 10a,c), as expected for this entirely heat flux-driven El Niño-like variability. The small deviations from +1 K/K can be explained by the usage of monthly-mean values instead of sub-daily data.

We repeat the integration separately for the SW radiation, the longwave radiation (LW), the sensible heat (SH) and latent heat (LH) fluxes to quantify the contribution of each heat flux component to the Slab Ocean SST change calculated from the net heat flux. The heat flux El Niño in the KCM Biased-Slab-Ocean experiment in the Niño3 region (Fig. 10a) is mainly driven by the SW (+4.1 K/K) and to a much lesser extent by the SH (+0.2 K/K) feedbacks, and is damped by the LW (-1.9 K/K) and LH (-1.4K/K) feedbacks. In observations, a +1 K SST warming in the Niño3 region is damped by the atmosphere by -1.3 K, i.e. ocean dynamics roughly contribute +2.3 K to the SST increase of +1 K and the remaining input is damped away by the atmospheric heat fluxes. The damping of the atmosphere of -1.3 K/K in the Niño3 region can be attributed to the SW (-0.5 K/K) and LH feedback (-0.8 K/K). The KCM AMIP-type simulation underestimates the net heat flux damping in the Niño3 region, as the SW feedback is slightly positive (+0.1 K/K) while damping by LH is a little stronger (-1.0 K/K) than in observations/reanalysis.

In the KCM sub-ensembles, the net heat flux damping decreases from STRONG to WEAK (Fig. 10a), as the SW feedback becomes more positive (+0.4 K/K, +0.7 K/K and +1.1 K/K in STRONG, MEDIUM, WEAK, respectively), and the LH damping becomes weaker (-1.2 K/K, -1.0 K/K and -0.8 K/K in STRONG, MEDIUM, WEAK, respectively). Consistent with the results of the BJ index analysis, the warming due to ocean dynamics decreases from STRONG to WEAK: in STRONG the ocean dynamics are responsible for +1.9 K/K in the Niño3 region and in WEAK for +1.0 K/K. The latter only is about 50 % of the observed dynamical heating. Further, in WEAK the SW feedback contributes +1.1 K/K to the SST warming, which is slightly more than the warming by ocean dynamics of +1.0 K/K. This indicates that ENSO in KCM is a hybrid of wind-driven and SW-driven ENSO dynamics, with a continuous transition to a more SW-driven ENSO from STRONG to WEAK. We observe a similar tendency towards a more SW-driven ENSO in the CMIP5 models (Fig. 10b). The corresponding AMIP experiments exhibit the strongest net heat flux damping, whereby we averaged over all atmosphere models irrespective of their coupled atmospheric feedback strength. In the CMIP5 models, the heat flux damping becomes weaker from STRONG to WEAK, and again the SW feedback is the major contributor to this shift. Thus the warming by ocean dynamics decreases from STRONG to WEAK and ENSO is also partly SW-driven albeit to a lesser extent than in KCM.

In the Niño4 region (Fig. 10c), observations/reanalysis show a net heat flux damping of -1.6 K/K that is mainly caused by SW damping (-2.6 K/K) and to a lesser extent by SH damping (-0.2 K/K), which is opposed by a positive LW (+0.4 K/K) and LH (+0.7 K/K) feedback. In KCM AMIP-type experiments the heat flux damping is stronger than in observations/reanalysis, mainly due to a too strong SW damping. The heat flux damping decreases from STRONG to WEAK, mainly caused by a decrease in the SW damping, but the SW and net heat flux feedback stay negative in all three sub-ensembles. Thus, the SST tendency in the Niño4 region is driven by anomalous ocean circulation. A similar behavior is found in the CMIP5-AMIP STRONG, MEDIUM and WEAK sub-ensembles (Fig. 10d), but with smaller differences between the sub-ensembles than in KCM. In summary we see a similar shift in ENSO dynamics in KCM and in CMIP5, but this shift seems to be more pronounced in KCM than in CMIP5.

In the Niño3 region, a significant negative correlation between the warming caused by ocean dynamics (deviation of the gray bars from the +1 K line in Fig. 10) and the SST change caused by the SW feedback is found in the two coupled model ensembles (Fig. 11). A +1 K warming can be caused by largely varying contributions of ocean dynamics and SW feedback: in some models a warming of 2.5 K by the ocean dynamics is required to realize +1 K warming while the SW feedback acts as a damping (CMIP5) or very weak forcing (KCM). In other models, the ocean dynamics only contributes a +1 K warming as the positive SW feedback is rather large. In summary, we find that a broad spectrum of ENSO dynamics exists in climate models, ranging from mostly wind-driven to mixed wind- and SW-driven ENSO dynamics with similar contributions.

7. Error compensation and ENSO amplitude

As shown above, the underestimated heat flux and wind feedbacks tend to compensate each other in many climate models, producing ENSO dynamics quite different from the observed. The question arises if this has an effect on the simulated ENSO amplitude. First, the level of underestimation in comparison to the observations/reanalysis is quantified. The wind feedback in the KCM (Fig. 12a) amounts to 62%, 54% and 43% of the observed in STRONG, MEDIUM and WEAK, respectively, the net heat flux feedback only to 58%, 39% and 10%. Thus the heat flux feedback is more strongly underestimated than the wind stress feedback, which also is the case for the CMIP5 models (Fig. 12b). The wind feedback and ENSO amplitude are positively correlated, while the net heat flux feedback and ENSO amplitude are negatively correlated (Fig. 13). This is consistent with the study of Vijayeta and Dommenget (2018) who analyze the important factors determining ENSO amplitude in CMIP3 and CMIP5 models in a Recharge Oscillator framework. They test in a CMIP3 and CMIP5 multi model ensemble the sensitivity of ENSO amplitude to the biases in the different parameters used in the Recharge Oscillator and show that the wind feedback is the dominant feedback for ENSO amplitude, while the net heat flux feedback has only half of the influence in comparison to the wind feedback (see Fig. 8 in Vijayeta and Dommenget 2018). Thus, the negative correlation between the net heat flux feedback and ENSO amplitude (Fig. 13b,d) may be caused indirectly by the strong correlation between the wind feedback and the net heat flux feedback. The rather moderate correlation between the wind feedback and ENSO amplitude in the KCM ensemble (Fig. 13a) and the even smaller and insignificant correlation in the CMIP5 ensemble (Fig. 13c) also suggests compensating effects in ENSO dynamics, as one would expect a larger ENSO amplitude with increasing wind feedback from ENSO theory. Thus, the error compensation between the too weak wind and net heat flux feedbacks, that we discuss here, may also explain the relatively small correlations shown in Fig. 13a,c).

8. Summary and discussion

In this study we have analyzed biases in ENSO dynamics in CGCMs in the presence of compensating errors between the two most important ENSO atmospheric feedbacks (EAF), the positive wind-SST and the negative net heat flux-SST feedback. Our results can explain why climate models with strongly underestimated EAF (in many models less than 50% of the observed feedback strength) can still have realistic ENSO amplitude due to error compensation between the two feedbacks. In addition and in agreement with Dommenget et al. (2014), our results provide further evidence that a broad range of

ENSO dynamics exists in climate models: dependent on the strength of the equatorial cold SST bias, there is a gradual change from a mostly wind-driven ocean-atmosphere coupling, as in observations, to a more SW-driven ocean atmosphere-coupling, as in the Biased-Slab Ocean experiment. An underestimated wind-SST feedback is linked to weaker positive feedbacks, i.e. weaker thermocline, Ekman and zonal advection feedbacks, resulting in a weaker subsurface heating of the SST by ocean dynamics. This is compensated by a weaker thermal damping by the atmosphere (i.e. net heat flux damping), resulting in a Bjerknes Stability (BJ) index that is not too different across different EAF strengths. This error compensation arises because both EAF strongly depend on the equatorial cold SST bias which determines the position of the rising branch of the Pacific Walker Circulation (PWC), as described in B18.

The equatorial Pacific cold SST bias is a common problem in CGCMs and its sources are still under debate. Possible contributors are too strong equatorial mean zonal surface winds, too large oceanic vertical mixing, and too little net surface SW radiation due to overestimated cloud cover and optical thickness (Davey et al. 2002; Guilyardi et al. 2009b; Vannière et al. 2013). However, it is important to note that another contributor to the underestimated EAF in CGCMs originates in their atmospheric components. In uncoupled AMIP simulations with the AGCMs, i.e. in the absence of SST biases, the EAF are already underestimated in most AGCMs, as there are uncertainties in the parametrizations of physical processes that are not resolved in coarse resolution models (Lloyd et al. 2011; Ferrett et al. 2017a,b).

With respect to ENSO amplitude, the error compensation may explain why climate models with a strongly underestimated wind feedback (less 50% of the observed value) still depict realistic ENSO amplitude, as often the net heat flux feedback also is strongly underestimated. The wind feedback strength appears to be more important than the net heat flux feedback, consistent with Vijayeta and

Dommenget (2018), even though the net heat flux feedback is more strongly underestimated in the climate models than the wind feedback. A recent study of Wengel et al. (2018a) suggests that the stochastic forcing of the SST and thermocline depth strongly influences ENSO amplitude and may also explain the low correlation between the wind stress feedback and ENSO amplitude in CMIP5 models (Fig. 13c).

The underestimated EAF also can explain why many climate models still show severe deficits in simulating important aspects of ENSO such as the phase locking of ENSO to the seasonal cycle or asymmetry between El Niño and La Niña events, as both aspects are better represented in climate models with strong EAF (B18). It still has to be investigated whether the large spread in simulated EAF could also be a major factor for the diversity in the ENSO response to global warming.

Our analyses suggest three possible pathways to improve the EAF in climate models: first, via an enhanced mean-state SST by either flux correction or tuning model parameters. Flux correction appears at first glance to be the simpler and more promising solution. However, Ferrett and Collins (2016) only report a moderate improvement in the EAF and ENSO dynamics when applying flux correction. On the other hand, the perturbed physics ensemble with the KCM suggests a large potential for tuning physical parameterizations, as the ensemble reproduced a similar spread in EAF strength and ENSO dynamics as that observed in the CMIP5 multi-model ensemble. However, Dommenget (2016) and Dommenget and Rezny (2018) note that it may be difficult to determine whether tuning makes a model more realistic for the right reasons. The second possibility for improving EAF is to improve the AGCMs so that they have a more realistic EAF in AMIP simulations with specified observed SSTs. Most of such uncoupled AGCM simulations already show too weak EAF, mainly due to biases in cloud cover, moisture and heat fluxes (Lloyd et al. 2011; Li et al. 2015; Ferrett et al. 2017a,b). Improvements in simulating these variables in AMIP-type simulations with AGCMs hold large potential for better representation of EAF in CGCMs. The third possibility to enhance EAF is to identify the processes and components of the climate models responsible for the cold SST bias and to improve them. As mentioned above, the surface winds, the oceanic vertical mixing and the cloud and convection schemes are promising candidates.

In summary our results suggest that many climate models have ENSO variability that is statistically not too different from observations, but for the wrong reasons. The equatorial mean state SST biases and EAF seem to be crucial to improve ENSO dynamics in current climate models.

Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive comments. We acknowledge the World Climate Research Program’s Working Group on Coupled Modeling, the individual modeling groups of the Climate Model Intercomparison Project (CMIP5), the UK Met Office and ECMWF for providing the data sets. The climate model integrations of the KCM and ECHAM5 were performed at the Computing Centre of Kiel University and the North-German Supercomputing Alliance (HLRN). This work was supported by the SFB 754 “Climate-Biochemistry Interactions in the tropical Ocean”, the European Union’s InterDec project, the ARC Centre of Excellence for Climate System Science (Grant CE110001028), the ARC project ‘‘Beyond the linear dynamics of the El Niño Southern Oscillation’’ (Grant DP120101442). This is a contribution to the Cluster of Excellence “The Future Ocean” at the University of Kiel.

References

Bayr, T., M. Latif, D. Dommenget, C. Wengel, J. Harlaß, and W. Park, 2018: Mean-State Dependence of ENSO Atmospheric Feedbacks in Climate Models. *Clim. Dyn.*, **50**, 3171–3194, doi:10.1007/s00382-017-3799-2.

Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, 2014: ENSO representation in climate models: From CMIP3 to CMIP5. *Clim. Dyn.*, **42**, 1999–2018, doi:10.1007/s00382-013-1783-z.

Bjerknes, J., 1969: Atmospheric Teleconnections from the Equatorial Pacific. *Mon. Weather Rev.*, **97**, 163–172.

Carton, J. A., and B. S. Giese, 2008: A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA). *Mon. Weather Rev.*, **136**, 2999–3017, doi:10.1175/2007MWR1978.1. http://journals.ametsoc.org/doi/abs/10.1175/2007MWR1978.1.

Chen, L., T. Li, Y. Yu, and S. K. Behera, 2017: A possible explanation for the divergent projection of ENSO amplitude change under global warming. *Clim. Dyn.*, **49**, 3799–3811, doi:10.1007/s00382-017-3544-x.

Davey, M., and Coauthors, 2002: STOIC: A study of coupled model climatology and variability in tropical ocean regions. *Clim. Dyn.*, **18**, 403–420, doi:10.1007/s00382-001-0188-6.

Dommenget, D., 2010: The slab ocean El Niño. *Geophys. Res. Lett.*, **37**, L20701, doi:10.1029/2010GL044888.

——, 2016: A simple model perturbed physics study of the simulated climate sensitivity uncertainty and its relation to control climate biases. *Clim. Dyn.*, **46**, 427–447, doi:10.1007/s00382-015-2591-4.

——, and M. Rezny, 2018: A Caveat Note on Tuning in the Development of Coupled Climate Model. *J. Adv. Model. Earth Syst.*, 335–356, doi:10.1002/2014MS000358.Received.

——, S. Haase, T. Bayr, and C. Frauen, 2014: Analysis of the Slab Ocean El Nino atmospheric feedbacks in observed and simulated ENSO dynamics. *Clim. Dyn.*, **42**, 3187–3205, doi:10.1007/s00382-014-2057-0.

Drews, A., and R. J. Greatbatch, 2016: Atlantic Multidecadal Variability in a model with an improved North Atlantic Current. *Geophys. Res. Lett.*, **43**, doi:10.1002/2016GL069815.

Ferrett, S., and M. Collins, 2016: ENSO feedbacks and their relationships with the mean state in a flux adjusted ensemble. *Climate Dynamics*, November 1 http://link.springer.com/10.1007/s00382-016-3270-9 (Accessed March 1, 2017).

——, ——, and H.-L. Ren, 2017a: Diagnosing relationships between mean state biases and El Niño shortwave feedback in CMIP5 models. *J. Clim.*, JCLI-D-17-0331.1, doi:10.1175/JCLI-D-17-0331.1. http://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0331.1.

——, ——, ——, S. Ferrett, M. Collins, and H.-L. Ren, 2017b: Understanding Bias in the Evaporative Damping of El Niño Southern Oscillation Events in CMIP5 Models. *J. Clim.*, JCLI-D-16-0748.1, doi:10.1175/JCLI-D-16-0748.1. http://journals.ametsoc.org/doi/10.1175/JCLI-D-16-0748.1.

Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Ocean.*, **118**, 6704–6716, doi:10.1002/2013JC009067.

Graham, F. S., J. N. Brown, C. Langlais, S. J. Marsland, A. T. Wittenberg, and N. J. Holbrook, 2014: Effectiveness of the Bjerknes stability index in representing ocean dynamics. *Clim. Dyn.*, 1–16, doi:10.1007/s00382-014-2062-3.

Guan, C., and M. J. McPhaden, 2016: Ocean processes affecting the twenty-first-century shift in ENSO

SST variability. *J. Clim.*, **29**, 6861–6879, doi:10.1175/JCLI-D-15-0870.1.

Guilyardi, E., and Coauthors, 2009a: Atmosphere Feedbacks during ENSO in a Coupled GCM with a Modified Atmospheric Convection Scheme. *J. Clim.*, **22**, 5698–5718, doi:10.1175/2009JCLI2815.1.

——, A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van Oldenborgh, and T. Stockdale, 2009b: Understanding El Niño in ocean-atmosphere general circulation models: Progress and challenges. *Bull. Am. Meteorol. Soc.*, **90**, 325–340, doi:10.1175/2008BAMS2387.1.

Jin, F. F., S. T. Kim, and L. Bejarano, 2006: A coupled-stability index for ENSO. *Geophys. Res. Lett.*, **33**, 2–5, doi:10.1029/2006GL027221.

Karamperidou, C., F. F. Jin, and J. L. Conroy, 2017: The importance of ENSO nonlinearities in tropical pacific response to external forcing. *Clim. Dyn.*, **49**, 2695–2704, doi:10.1007/s00382-016-3475-y.

Kim, S. T., W. Cai, F.-F. Jin, and J.-Y. Yu, 2014a: ENSO stability in coupled climate models and its association with mean state. *Clim. Dyn.*, **42**, 3313–3321, doi:10.1007/s00382-013-1833-6. http://link.springer.com/10.1007/s00382-013-1833-6 (Accessed October 26, 2016).

——, ——, F. F. Jin, and J. Y. Yu, 2014b: ENSO stability in coupled climate models and its association with mean state. *Clim. Dyn.*, **42**, 3313–3321, doi:10.1007/s00382-013-1833-6.

Levitus, S., and Coauthors, 1998: *World Ocean Data Base*. 346 pp.

Li, L., B. Wang, G. J. Zhang, L. Li, B. Wang, and G. J. Zhang, 2015: The Role of Moist Processes in Shortwave Radiative Feedback during ENSO in the CMIP5 Models. *J. Clim.*, **28**, 9892–9908, doi:10.1175/JCLI-D-15-0276.1. http://journals.ametsoc.org/doi/10.1175/JCLI-D-15-0276.1 (Accessed September 28, 2016).

Li, Y., J. Li, W. Zhang, Q. Chen, J. Feng, F. Zheng, W. Wang, and X. Zhou, 2017: Impacts of the

Tropical Pacific Cold Tongue Mode on ENSO Diversity Under Global Warming. *J. Geophys. Res. Ocean.*, **122**, 8524–8542, doi:10.1002/2017JC013052.

Lloyd, J., E. Guilyardi, H. Weller, and J. Slingo, 2009: The role of atmosphere feedbacks during ENSO in the CMIP3 models. *Atmos. Sci. Lett.*, **10**, 170–176, doi:10.1002/asl.227.

——, ——, and ——, 2011: The role of atmosphere feedbacks during ENSO in the CMIP3 models. Part II: using AMIP runs to understand the heat flux feedback mechanisms. *Clim. Dyn.*, **37**, 1271–1292, doi:10.1175/JCLI-D-11-00178.1.

——, ——, and ——, 2012: The role of atmosphere feedbacks during ENSO in the CMIP3 models. Part III: The shortwave flux feedback. *J. Clim.*, **25**, 4275–4293, doi:10.1175/JCLI-D-11-00178.1.

Lübbecke, J. F., and M. J. Mcphaden, 2014: Assessing the twenty-first-century shift in enso variability in terms of the bjerknes stability index. *J. Clim.*, **27**, 2577–2587, doi:10.1175/JCLI-D-13-00438.1.

Madec, G., 2008: NEMO ocean engine. *Note du Pole modélisation 27, Inst. Pierre-Simon Laplace*, 193 pp.

——, P. Delecluse, M. Imbard, and C. Lévy, 1998: OPA 8.1 Ocean General Circulation Model Manual. *Note du Pole modélisation 11, Inst. Pierre-Simon Laplace*, 91 pp.

Mauritsen, T., and Coauthors, 2012: Tuning the climate of a global model. *J. Adv. Model. Earth Syst.*, **4**, n/a-n/a, doi:10.1029/2012MS000154. http://doi.wiley.com/10.1029/2012MS000154.

McPhaden, M. J., 1999: The child prodigy of 1997-1998. *Nature*, **398**, 559–562.

Park, W., N. S. Keenlyside, M. Latif, A. Ströh, R. Redler, E. Roeckner, and G. Madec, 2009: Tropical Pacific Climate and Its Response to Global Warming in the Kiel Climate Model. *J. Clim.*, **22**, 71–92, doi:10.1175/2008JCLI2261.1.

Philander, S., 1990: *El Niño, La Niña, and the southern oscillation*. Academic Press, San Diego, USA, 293 pp.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407.

Roeckner, E., and Coauthors, 2003: *The atmospheric general circulation model ECHAM5. PART I: Model description, Report 349*. Max Planck Institute for Meteorology, Hamburg, Germany, 140 pp.

Simmons, A., S. Uppala, D. Dee, and S. Kobayashi, 2007: ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsl.*, **110**, 25–35.

Stocker, T., D. Qin, G. Plattner, M. Tignor, and S. Allen, 2013: *IPCC 2013: Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York, 1535 pp.

Taylor, K. E., R. J. Stouffer, and G. a. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bull. Am. Meteorol. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Timmermann, A., and Coauthors, 2018: El Niño-Southern Oscillation Complexity. *Nature*, 1–25, doi:10.1038/s41586-018-0252-6. http://dx.doi.org/10.1038/s41586-018-0252-6.

Trenberth, K. E., 1997: The definition of El Niño. *Bull. Am. Meteorol. Soc.*, **78**, 2771–2778.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.*, **131**, 2961–3012, doi:10.1256/qj.04.176.

Vannière, B., E. Guilyardi, G. Madec, F. J. Doblas-Reyes, and S. Woolnough, 2013: Using seasonal hindcasts to understand the origin of the equatorial cold tongue bias in CGCMs and its impact on ENSO. *Clim. Dyn.*, **40**, 963–981, doi:10.1007/s00382-012-1429-6.

Vijayeta, A., and D. Dommenget, 2018: An evaluation of ENSO dynamics in CMIP simulations in the framework of the recharge oscillator model. *Clim. Dyn.*, **0**, 1–19, doi:10.1007/s00382-017-3981-6. http://dx.doi.org/10.1007/s00382-017-3981-6.

Wengel, C., D. Dommenget, M. Latif, T. Bayr, and A. Vijayeta, 2018a: What controls ENSO-amplitude diversity in climate models? *Geophys. Res. Lett.*, 1–8, doi:10.1002/2017GL076849.

Wengel, C., M. Latif, W. Park, J. Harlaß, and T. Bayr, 2018b: Seasonal ENSO phase locking in the Kiel Climate Model: The importance of the equatorial cold sea surface temperature bias. *Clim. Dyn.*, doi:10.1007/s00382-017-3648-3.

Zhai, X., and R. J. Greatbatch, 2007: Wind work in a model of the northwest Atlantic Ocean. *Geophys. Res. Lett.*, **34**, 1–4, doi:10.1029/2006GL028907.

Zheng, X. T., S. P. Xie, L. H. Lv, and Z. Q. Zhou, 2016: Intermodel uncertainty in ENSO amplitude change tied to Pacific Ocean warming pattern. *J. Clim.*, **29**, 7265–7279, doi:10.1175/JCLI-D-16-0039.1.

Tables

Tab. 1: List of 10m zonal wind feedback, zonal wind stress feedback, heat flux feedback and total atmospheric feedback strength (average of column 6 and 8) in ERA-Interim and ERA40 reanalysis and CMIP5 models. The normalized feedbacks are divided by the average feedback of ERA-Interim and ERA40 (first and second row).

Nr.	Model	AMIP	10m wind feedback [m/s/K]	wind stress feedback [10^{-2} Pa/K]	normalized wind stress feedback [%]	Heat flux feedback [W/m²/K]	normalized heat flux feedback [%]	total atmosheric feedback [%]	sub-ensemble
	ERA-Interim		1.46	1.53	104.6	-16.6	98.7	101.7	
	ERA40		1.34	1.40	95.4	-17.0	101.3	98.3	
1	ACCESS1-0	yes	0.81	0.72	48.8	-12.2	72.3	60.5	STRONG
2	ACCESS1-3	yes	0.73	0.73	49.7	-3.8	22.9	36.3	MEDIUM
3	BCC-CSM1-1	yes	0.65	0.75	51.1	-8.3	49.3	50.2	MEDIUM
4	BNU-ESM	yes	0.91	0.76	51.7	-10.0	59.6	55.6	STRONG
5	CanESM2	yes	0.58	0.82	56.0	-7.3	43.6	49.8	MEDIUM
6	CMCC-CM	yes	0.96	0.86	58.5	-10.7	63.7	61.1	STRONG
7	CNRM-CM5	yes	1.13	0.98	66.8	-14.6	86.5	76.7	STRONG
8	CSIRO-Mk3-6-0	no	0.46	0.66	45.3	1.3	-8.0	18.6	WEAK
9	GFDL-CM3	yes	0.61	0.65	44.2	-6.8	40.4	42.3	MEDIUM
10	GFDL-ESM2G	no	0.52	0.75	50.9	-3.5	20.5	35.7	MEDIUM
11	GFDL-ESM2M	no	1.00	0.98	66.9	-10.2	60.3	63.6	STRONG
12	GISS-E2-R	yes	1.29	1.02	69.8	-10.4	61.9	65.8	STRONG
13	HadGEM2-CC	yes	0.66	0.67	45.6	-5.1	30.4	38.0	MEDIUM
14	HadGEM2-ES	yes	0.64	0.68	46.3	-3.4	20.2	33.2	WEAK
15	IPSL-CM5A-LR	yes	0.45	0.58	39.9	-1.2	7.2	23.6	WEAK
16	IPSL-CM5A-MR	yes	0.61	0.83	56.7	-2.7	15.8	36.2	MEDIUM
17	MIROC5	yes	0.98	1.01	69.2	-9.7	57.4	63.3	STRONG

18	MPI-ESM-LR	yes	0.45	0.39	26.5	-5.4	32.2	29.3	WEAK
19	MPI-ESM-MR	yes	0.60	0.66	44.9	-4.1	24.1	34.5	WEAK
20	MRI-CGCM3	yes	0.57	0.47	32.1	-4.1	24.4	28.3	WEAK
21	NorESM1-ME	yes	1.24	1.03	70.6	-9.1	54.1	62.3	STRONG

Figure Captions

Figure 1: Schematic of the mean state and atmospheric feedbacks in the tropical Pacific: a) mean state in observations: at the surface easterly winds blow along the equator from the cold tongue in the east to the warm pool in the west, where the rising branch of the PWC is situated close to Niño4. It shifts eastward (westward) during El Niño (La Niña), as indicated by the red arrow. b) This eastward (westward) shift of PWC during El Niño (La Niña) causes a weakening (strengthening) of zonal wind in Niño4, thus a positive zonal wind feedback that further amplifies the SST anomaly via the positive Bjerknes feedback. On the other hand causes the eastward (westward) shift of the PWC during El Niño (La Niña) more (less) convection over the equatorial Pacific and therefore a negative heat flux feedback that damps the SST anomaly. c) In the presence of a large equatorial cold SST bias, the PWC is due to the La Niña-like mean state too far in the west. d) As both feedback strengths strongly depend on position of the PWC, this causes a too weak positive wind-SST feedback and a too weak negative heat flux-SST feedback, thus an error compensation between the too weak positive and too weak negative feedback.

Figure 2: a) Zonal wind stress feedback in Niño4 region (local zonal wind stress regressed on SST of Niño3.4 region) on the x-axis vs. 10m zonal wind feedback in Niño4 region (local zonal surface wind regressed on SST of Niño3.4 region) on the y-axis, for ERA-Interim, ERA40 and individual experiments of the perturbed physics ensemble of KCM (numbers), KCM AMIP-type experiment with 19, 31 and 62 vertical levels; b) same as a) but here for the individual CMIP5 models (numbers); the correlation is given in the upper left corner and two stars indicate significant correlation on a 99% confidence level; the black dashed line marks the ratio of 100 Pa/K / 1 m/s/K.

Figure 3: a) Zonal wind stress feedback in Niño4 region (local zonal wind stress regressed on SST of

Niño3.4 region) on the x-axis vs. heat flux feedback in Niño3 and Niño4 (local heat flux regressed on SST of Niño3.4 region) on the y-axis, for ERA-Interim, ERA40 and individual experiments of the perturbed physics ensemble of KCM (numbers), KCM AMIP-type experiment with 19, 31 and 62 vertical levels and the biased KCM Slab Ocean experiment; b) atmospheric feedback strength (average of wind stress and heat flux feedback, after normalizing each by the average reanalysis value) on x-axis vs. relative SST bias in the Niño4 region (modeled SST minus observed SST, after subtracting the tropical Indo-Pacific area mean SST from each); c-d) same as a-b) but here for the individual CMIP5 models (numbers); the color of the numbers indicates the sub-ensembles of STRONG (red), MEDIUM (blue) and WEAK (green) EAF, as used in the following; the correlation is given in the upper right/left corner and two stars indicate significant correlation on a 99% confidence level; the black line is the regression.

Figure 4: Bjerknes feedback in observations/reanalysis data, in a) local zonal wind stress regressed on SST in the Niño3 region, b) local thermocline depth regressed on zonal wind stress in Niño4, c) local SST regressed on local thermocline depth; d-f) same as a-c) but here for KCM STRONG sub-ensemble; g-i) same as a-c) but here for the KCM MEDIUM sub-ensemble; j-l) same as a-c) but here for the KCM WEAK sub-ensemble; values of $r^2 > 0.2$ are indicated by shading and the values in the header are the average of regression and explained variance in the Niño4 region (first column) and Niño3 region (second and third column), as indicated by the black box.

Figure 5: Same as Fig. 4, but here for CMIP5 STRONG, MEDIUM and WEAK sub-ensembles.

Figure 6: Bjerknes feedback in the individual experiments of KCM, a) average regression coefficient in Niño4 region of local wind stress regressed on SST in Niño3 region on the x-axis vs. average explained variance in Niño4 on y-axis; b) same as a) but here the average in Niño3 region of local

thermocline depth regressed on wind stress in Niño4 region on the x-axis vs. average explained variance in Niño3 on the y-axis; c) same as b) but here for local SST regressed on local thermocline depth; d) product of regression coefficients of a-c) on the x-axis vs. the average explained variance of a-c) on the y-axis; e-h) same as a-d) but here for the CMIP5 models; The correlation is given in the upper right/left corner and two stars indicate a significant correlation on a 99% confidence level.

Figure 7: The individual feedbacks of the Bjerknes Stability Index for each calendar month in ERA40/SODA reanalysis and the KCM STRONG, MEDIUM and WEAK sub-ensembles, a) the zonal advection feedback, b) Ekman feedback, c) thermocline feedback, d) dynamical damping, e) thermal damping and in f) the total Bjerknes Stability Index, which is the sum of a-e); the error bars indicate the 90% confidence interval for reanalysis, estimated from the standard error of the regression slop.

Figure 8: For ERA40/SODA reanalysis (orange dot) and the individual KCM experiments (numbers), in a) the zonal advection feedback on the x-axis vs. the relative SST bias in Niño4 region on the y-axis, both for the calendar months SONDJF; b) same as a) but here on the x-axis the Ekman feedback; c) same as a), but here on the x-axis the thermocline feedback; d) same as a) but here on the x-axis the thermal damping and both for the calendar months JFMAMJ; The seasonal mean is calculated from the monthly resolved BJ index, as shown in Fig. 7. The correlation is given in the upper right/left corner and two stars indicate a significant correlation on a 99% confidence level.

Figure 9: a) zonal wind stress feedback in the Niño4 region vs. SW feedback in Niño3 region in ERA-Interim, ERA40 and KCM experiments; b) same as a) but here for the CMIP5 multi model ensemble; The correlation between the individual experiments is shown in upper left/right corner and two stars indicate a significant correlation on a 99% confidence level.

Figure 10: Offline slab ocean SST calculated by integrating the net heat flux (Q_{net}), short wave radiation (SW), long wave radiation (LW), sensible heat flux (SH) and latent heat flux (LH), respectively, to a 50m water column six months before the peak of the ENSO events and then normalized by the ENSO amplitude, in a) for reanalysis, KCM AMIP-type experiment, KCM STRONG, MEDIUM and WEAK sub-ensembles and Biased-Slab-Ocean experiment in the Niño3 region; b) same as a) but here for CMIP5 AMIP, STRONG, MEDIUM and WEAK sub-ensembles; c) same as a) but here for the Niño4 region; d) same as b) but here for the Niño4 region.

Figure 11: a) For reanalysis and the individual KCM experiments, on the y-axis the SST change in Niño3 region due to ocean circulation (i.e. 1 - $dSST_{Qnet}$ as shown in Fig. 10a) vs. the SST change by SW feedback (i.e. $dSST_{SW}$ as shown in Fig. 10a) on the x-axis; b) same as a) but here for the individual CMIP5 models; The color of the numbers indicate the sub-ensembles with STRONG (red), MEDIUM (blue) and WEAK (green) atmospheric feedbacks; The correlation between the individual experiments is shown in upper right corner and two stars indicate a significant correlation on a 99\% confidence level.

Figure 12: a) Wind stress feedback in Niño4 (top) and heat flux feedback in Niño3 and Niño4 (bottom) in ERA-Interim/ERA40 reanalysis and KCM experiments; b) same as a) but here for CMIP5 models; The strength of the feedbacks relative to the observed feedbacks is shown as numbers in % at the bars.

Figure 13: a) Zonal wind stress feedback in the Niño4 region on the x-axis vs. standard deviation of SST anomalies in the Niño3.4 region on the y-axis for ERA-Interim/ERA40 reanalysis and the individual KCM perturbed physics experiments; b) same as a), but here the heat flux feedback in Niño3 and Niño4 on the x-axis; c-d) same as a-b), but here for the CMIP5 models. The correlation between the individual experiments is shown in upper right/left corner and one (two) stars indicate a significant

correlation on a 95% (99%) confidence level.

Table captions:

Tab. 1: List of 10m zonal wind feedback, zonal wind stress feedback, heat flux feedback and total atmospheric feedback strength (average of column 6 and 8) in ERA-Interim and ERA40 reanalysis and CMIP5 models. The normalized feedbacks are divided by the average feedback of ERA-Interim and ERA40 (first and second row).

Error Compensation of ENSO Atmospheric Feedbacks in Climate Models and its influence on Simulated ENSO Dynamics

October 8, 2018

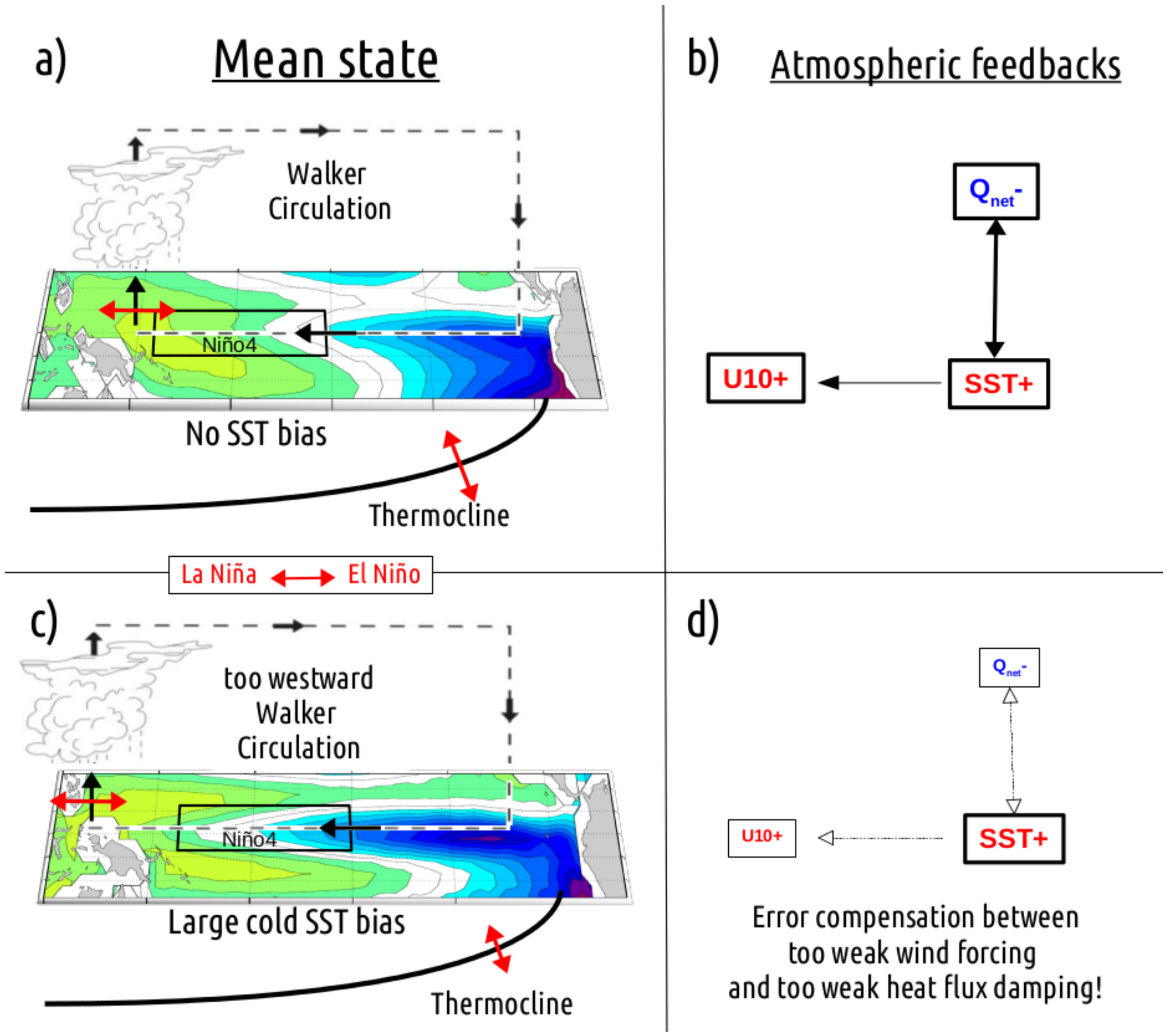

Figure 1: *Schematic of the mean state and atmospheric feedbacks in the tropical Pacific: a) mean state in observations: at the surface easterly winds blow along the equator from the cold tongue in the east to the warm pool in the west, where the rising branch of the Walker Circulation is situated close the Nino4 region. It shifts eastward (westward) during El Nino (La Nina), as indicated by the red arrow. b) This eastward (westward) shift of the rising branch of the Walker Circulation during El Nino (La Nina) causes a weakening (strengthening) of zonal wind in the Nino4 region, thus a positive zonal wind feedback, that further amplifies the SST anomaly via the positve Bjerknes feedback. On the other hand causes the eastward (westward) shift during El Nino (La Nina) more (less) convection over the equatorial Pacific and therefore a negative heat flux feedback, that damps the SST anomaly. c) In the presence of a large equatorial cold SST bias, the rising branch of the Walker Circulation is due to the La Nina-like mean state too far in the west. d) As both feedback strengths strongly depend on position of the Walker Circulation, this causes a too weak positive wind-SST feedback and a too weak negative heat flux-SST feedback, thus an error compensation between the too weak positive and too weak negative feedback.*

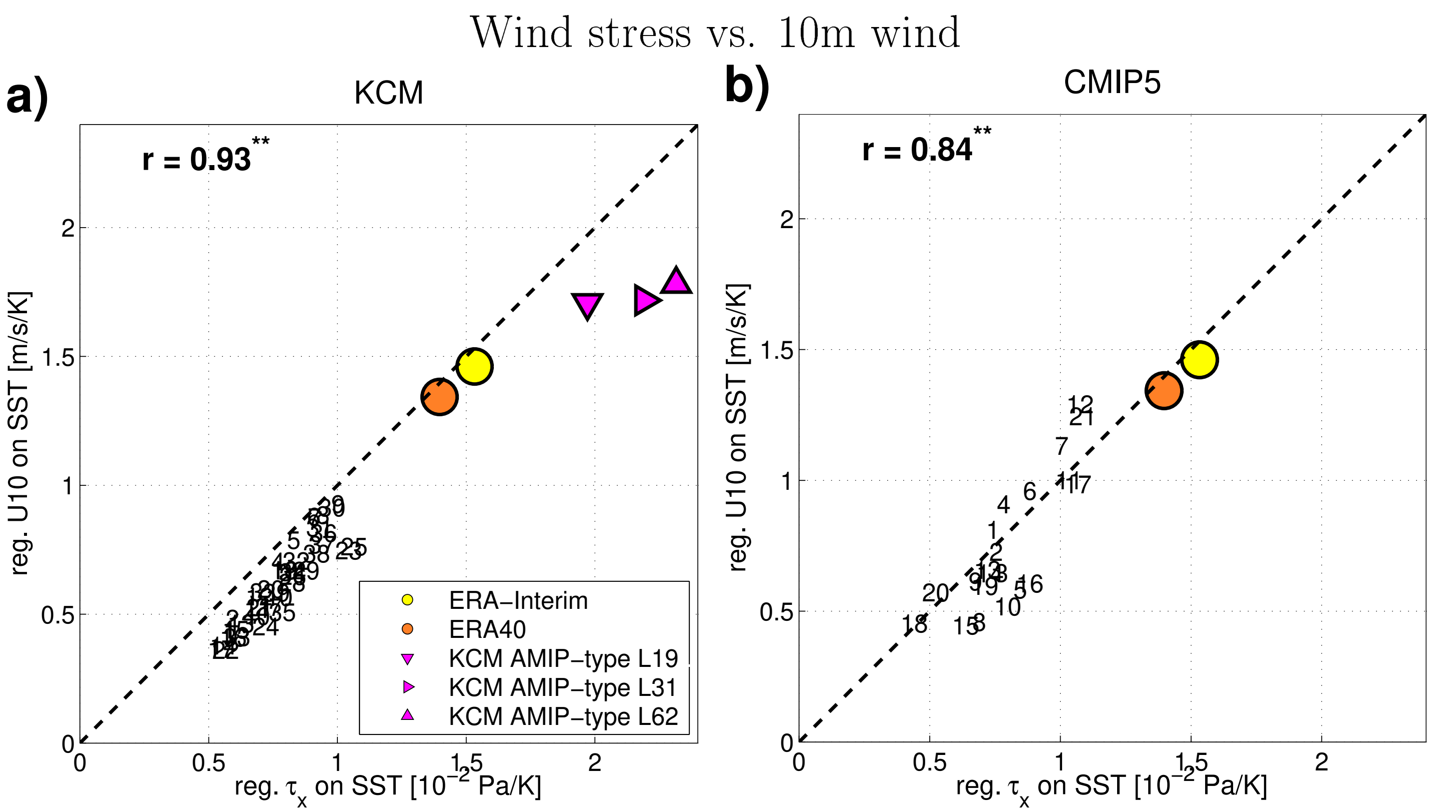

Figure 2: *a) Zonal wind stress feedback in Nino4 region (local zonal wind stress regressed on SST of Nino3.4 region) on the x-axis vs. 10m zonal wind feedback in Nino4 region (local zonal surface wind regressed on SST of Nino3.4 region) on the y-axis, for ERA-Interim, ERA40 and individual experiments of the perturbed physics ensemble of KCM (numbers), KCM AMIP-type experiment with 19, 31 and 62 vertical levels; b) same as a) but here for the individual CMIP5 models (numbers); the correlation is given in the upper left corner and two stars indicate significant correlation on a 99% confidence level; the black dashed line marks the ratio of 100 Pa/K / 1 m/s/K.*

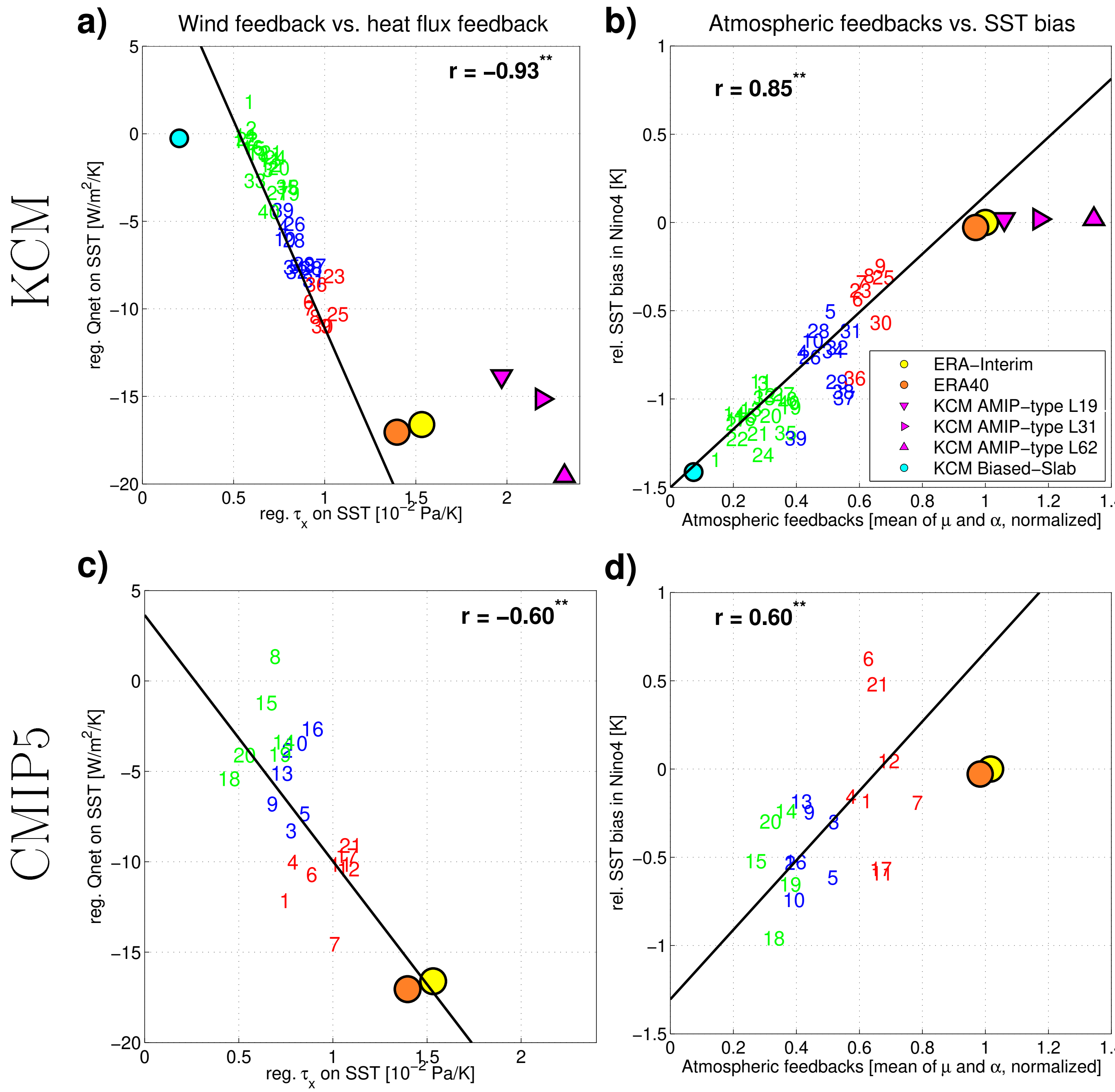

Figure 3: *a) Zonal wind stress feedback in Nino4 region (local zonal wind stress regressed on SST of Nino3.4 region) on the x-axis vs. heat flux feedback in the Nino3 and Nino4 region (local heat flux regressed on SST of Nino3.4 region) on the y-axis, for ERA-Interim, ERA40 and individual experiments of the perturbed physics ensemble of KCM (numbers), KCM AMIP-type experiment with 19, 31 and 62 vertical levels and the biased KCM Slab Ocean experiment; b) atmospheric feedback strength (average of wind stress and heat flux feedback, after normalsizing each by the average reanalysis value) on x-axis vs. relative SST bias in the Nino4 region (modeled SST minus observed SST, after subtracting the tropical Indo-Pacific area mean SST from each); c-d) same as a-b) but here for the individual CMIP5 models (numbers); the color of the numbers indicates the sub-ensembles of STRONG (red), MEDIUM (blue) and WEAK (green) atmospheric feedbacks, as used in the following; the correlation is given in the upper right/left corner and two stars indicate significant correlation on a 99% confidence level; the black line is the regression.*

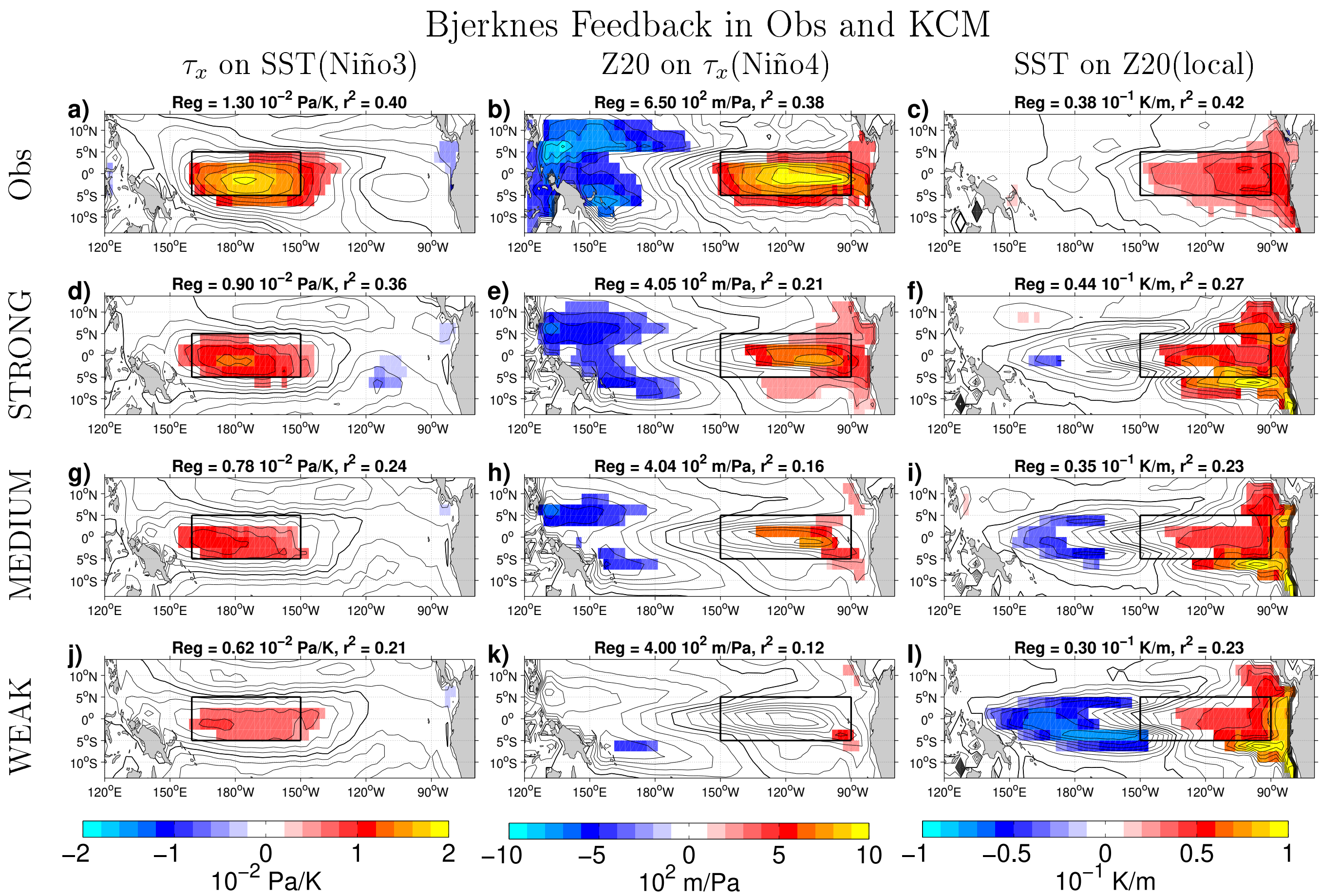

Figure 4: *Bjerknes Feedback in observations/reanalysis data, in a) local zonal wind stress regressed on SST in the Nino3 region, b) local thermocline depth regressed on zonal wind stress in the Nino4 region, c) local SST regressed on local thermocline depth; d-f) same as a-c) but here for KCM STRONG sub-ensemble; g-i) same as a-c) but here for the KCM MEDIUM sub-ensemble; j-l) same as a-c) but here for the KCM WEAK sub-ensemble; values of $r^2 > 0.2$ are indicated by shading and the values in the header are the average of regression and explained variance in the Nino4 region (first column) and Nino3 region (second and third column), as indicated by the black box.*

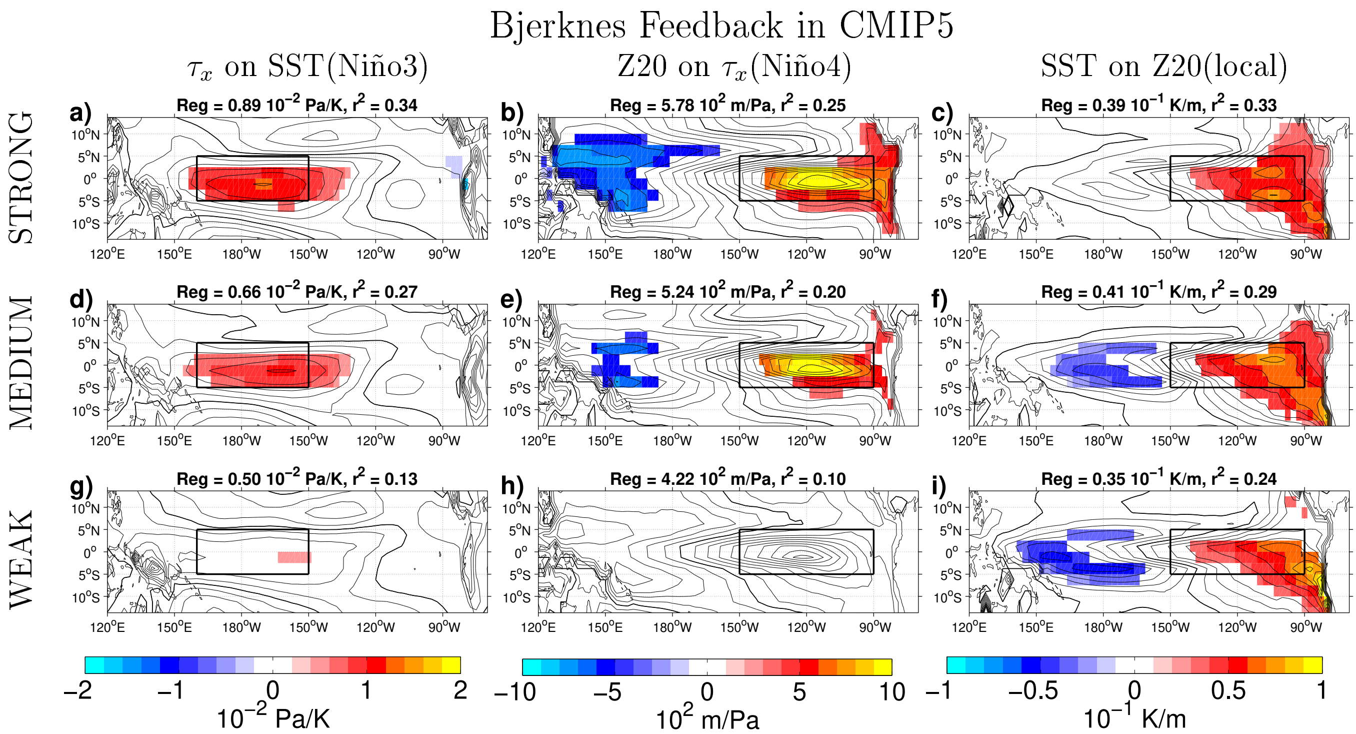

Figure 5: *Same as Fig. 4, but here for CMIP5 STRONG, MEDIUM and WEAK sub-ensembles.*

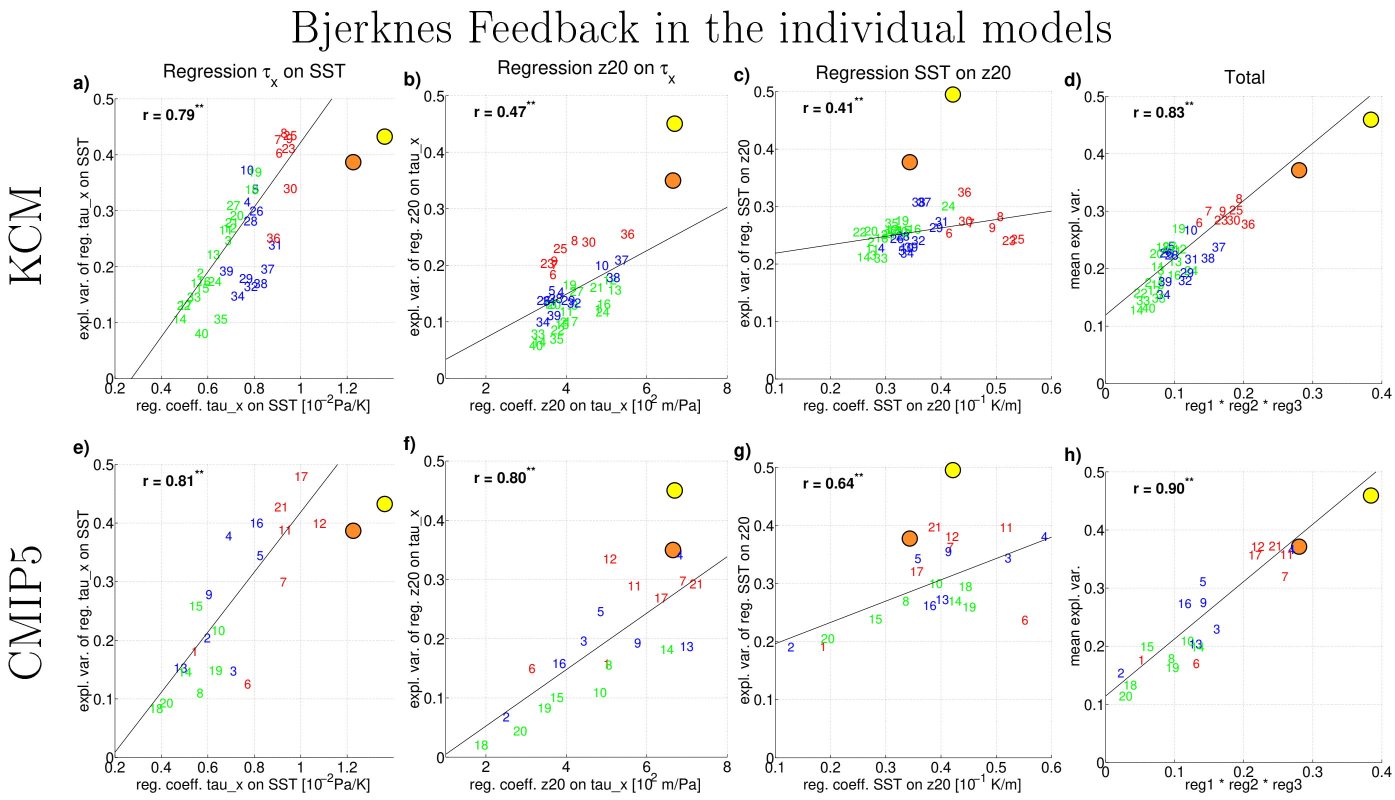

Figure 6: *Bjerknes feedback in the individual experiments of KCM, a) average regression coefficient in Nino4 region of local wind stress regressed on SST in Nino3 region on the x-axis vs. average explained variance in Nino4 region on y-axis; b) same as a) but here the average in Nino3 region of local thermocline depth regressed on wind stress in Nino4 region on the x-axis vs. average explained variance in Nino3 region on the y-axis; c) same as b) but here for local SST regressed on local thermocline depth; d) product of regression coefficients of a-c) on the x-axis vs. the average explained variance of a-c) on the y-axis; e-h) same as a-d) but here for the CMIP5 models; The color of the numbers indicates the sub-ensembles of STRONG (red), MEDIUM (blue), WEAK (green) atmospheric feedbacks; The correlation is given in the upper right/left corner and two stars indicate a significant correlation on a 99% confidence level.*

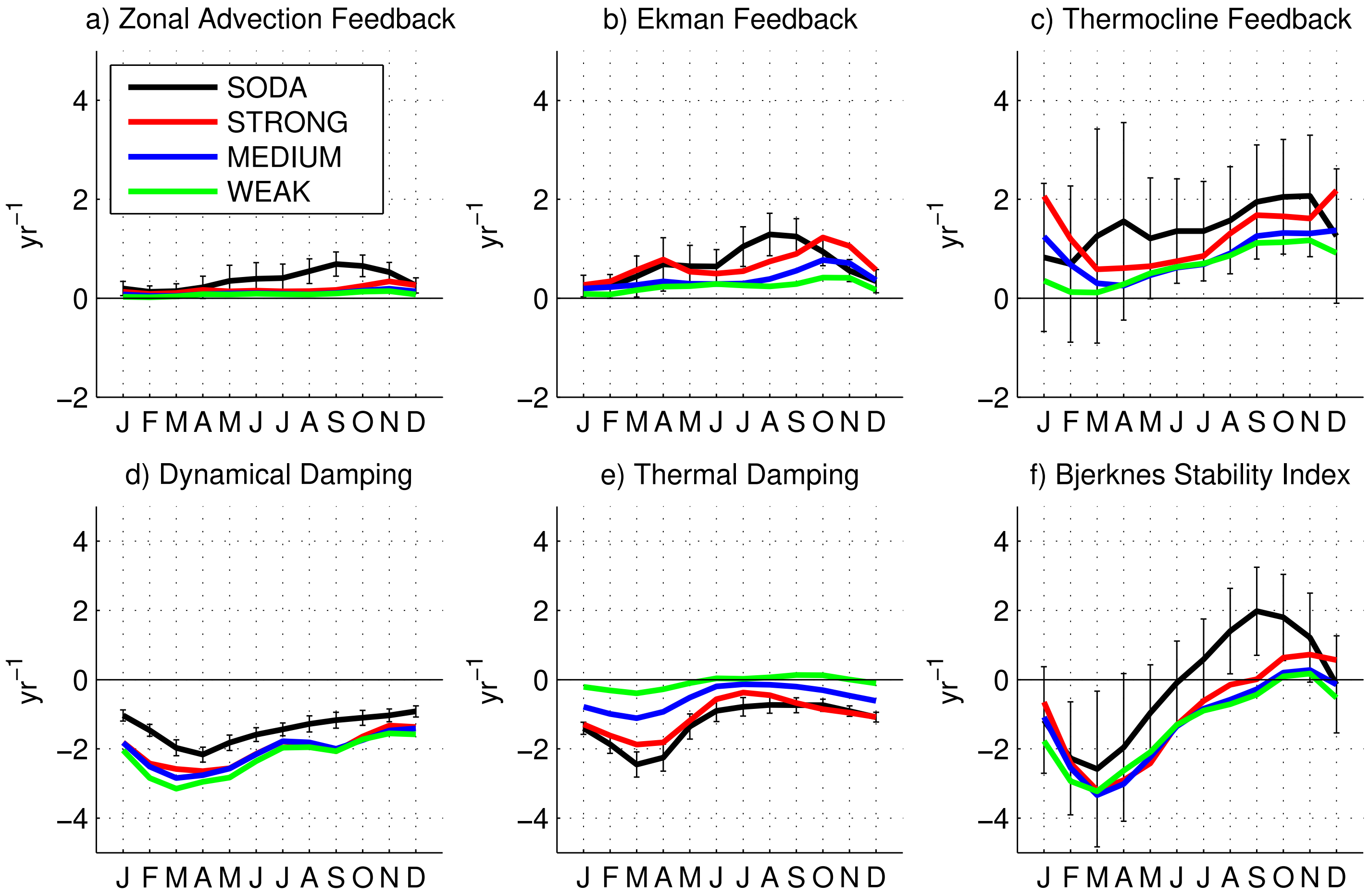

Figure 7: *The individual feedbacks of the Bjerknes Stability Index for each calendar month in ERA40/SODA reanalysis and the KCM STRONG, MEDIUM and WEAK sub-ensembles, a) the zonal advection feedback, b) ekman feedback, c) thermocline feedback, d) dynamical damping, e) thermal damping and in f) the total Bjerknes Stability Index, which is the sum of a-e); the error bars indicate the 90% confidence interval for reanalysis, estimated from the standard error of the regression slop.*

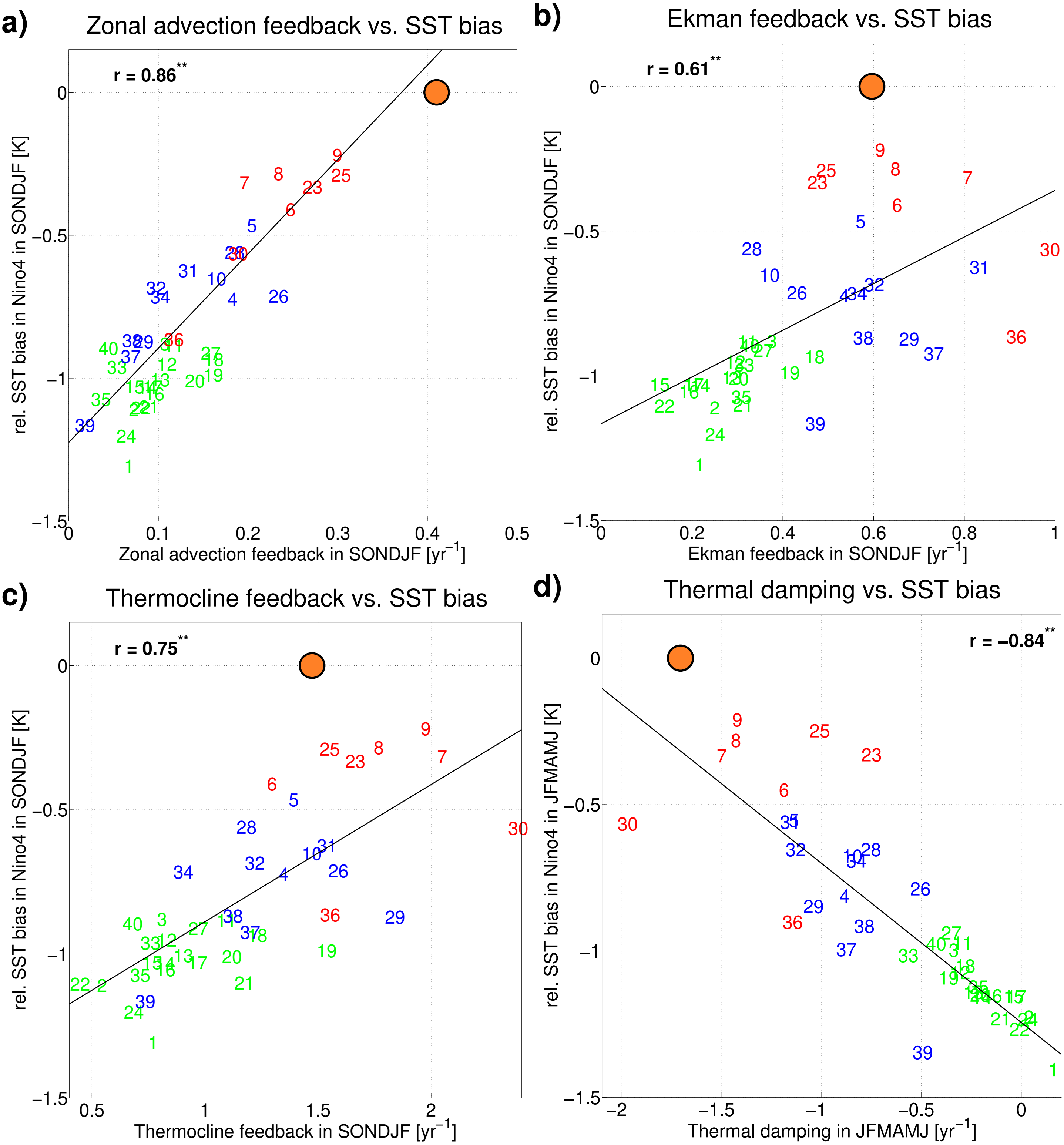

Figure 8: *For ERA40/SODA reanalysis and the individual KCM experments, in a) the zonal advection feedback on the x-axis vs. the relative SST bias in Nino4 region on the y-axis; b) same as a) but here on the x-axis the Ekman feedback; c) same as a), but here on the x-axis the thermocline feedback; d) same as a) but here on the x-axis the thermal damping; The annual mean is calculated from the monthly resolved BJ index, as shown in Fig. 7. The colors indicate the sub-ensembles with STRONG (red), MEDIUM (blue) and WEAK (green) atmospheric feedbacks. The correlation is given in the upper right/left corner and two stars indicate a significant correlation on a 99% confidence level.*

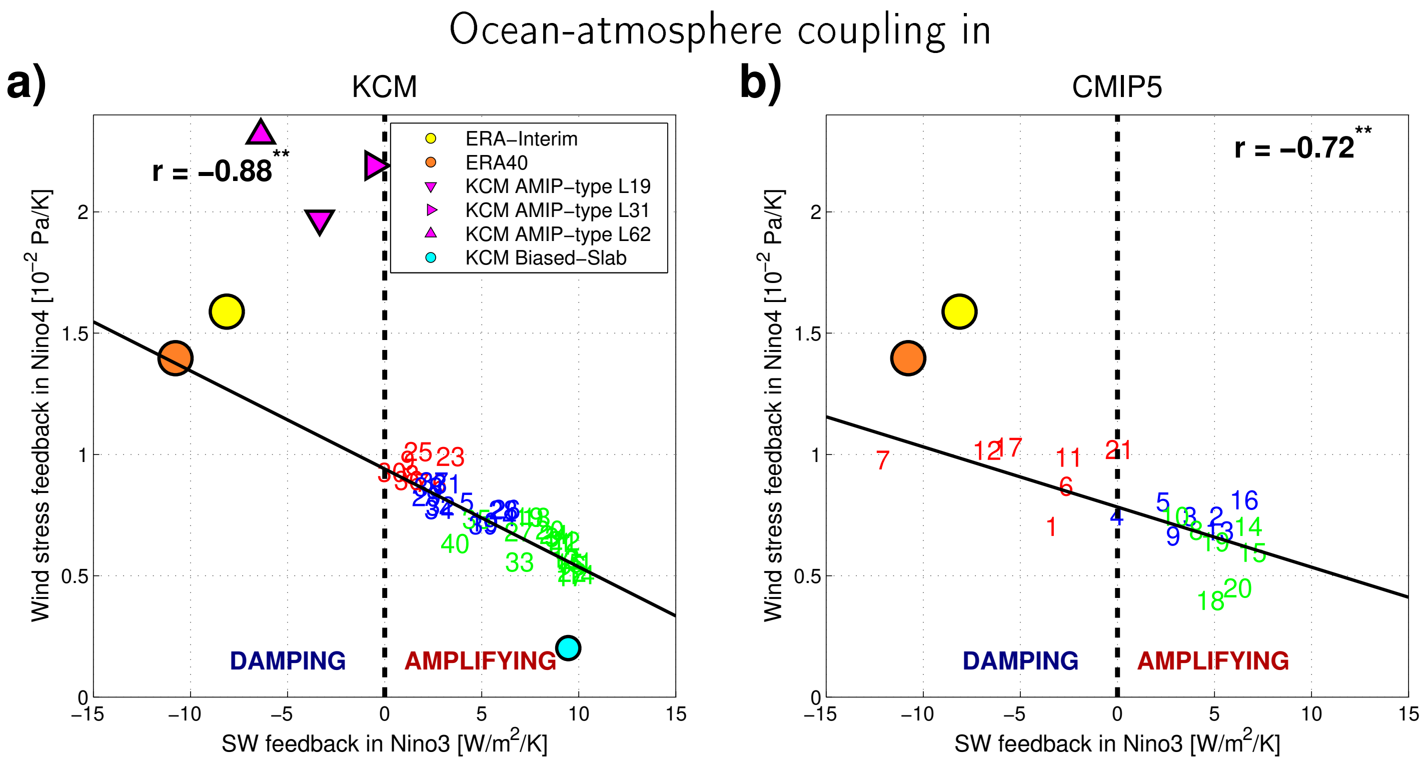

Figure 9: *a) zonal wind stress feedback in the Niño4 region vs. SW feedback in Niño3 region in ERA-Interim, ERA40 and KCM exeriments; b) same as a) but here for the CMIP5 multi model ensemble; The color of the numbers indicate the sub-ensembles with STRONG (red), MEDIUM (blue) and WEAK (green) atmospheric feedbacks. The correlation between the individual experiments is shown in upper left/right corner and two stars indicate a significant correlation on a 99% confidence level.*

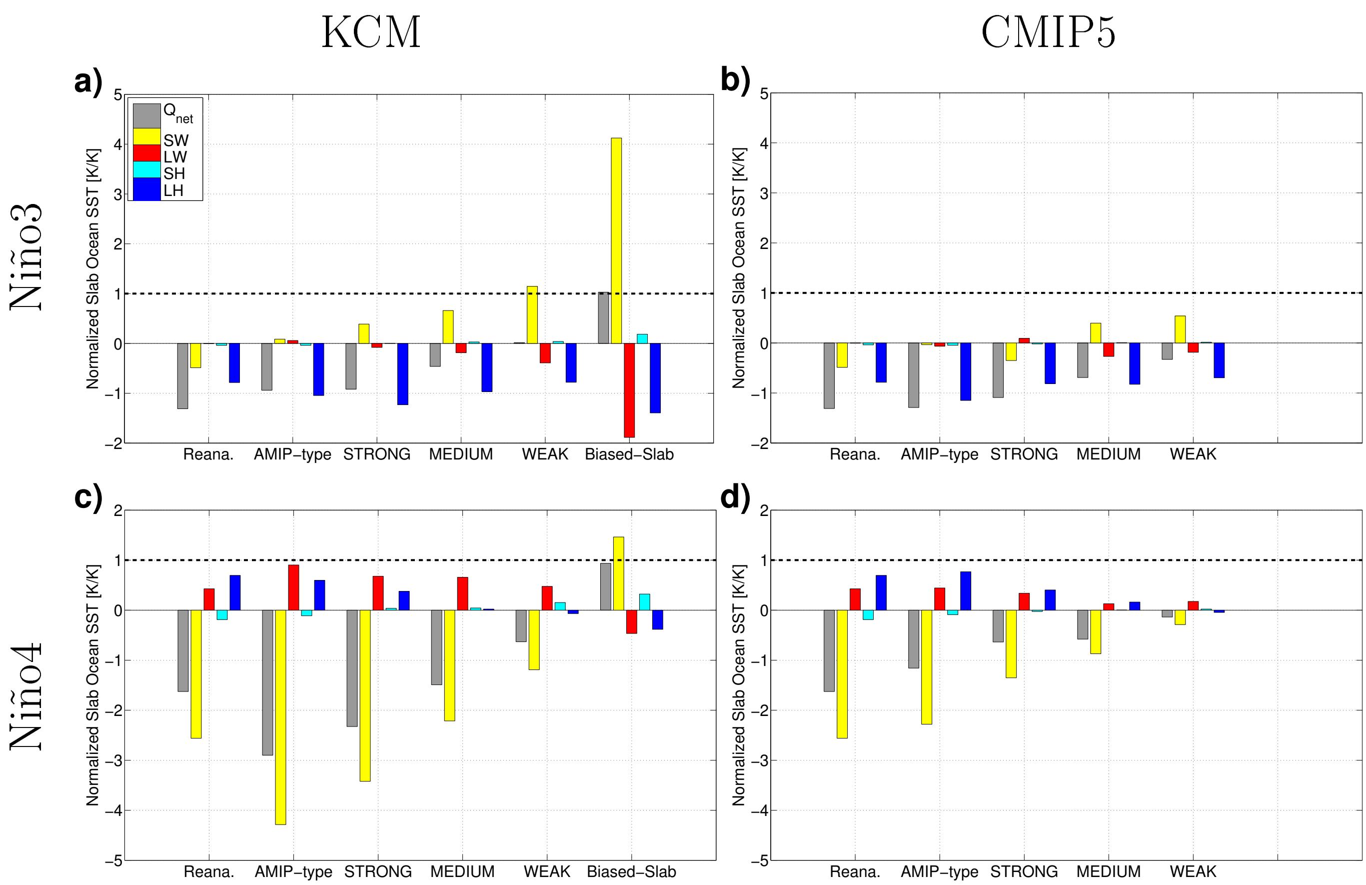

Figure 10: *Offline slab ocean SST calculated by integrating the net heat flux (Q_{net}), short wave radiation (SW), long wave radiation (LW), sensible heat flux (SH) and latent heat flux (LH), respectively, to a 50m water column six months before the peak of the ENSO events and then normalized by the ENSO amplitude, in a) for reanalysis, KCM AMIP-type experiment, KCM STRONG, MEDIUM and WEAK sub-ensembles and Biased-Slab-Ocean experiment in the Niño3 region; b) same as a) but here for CMIP5 AMIP, STRONG, MEDIUM and WEAK sub-ensembles; c) same as a) but here for the Nino4 region; d) same as b) but here for the Nino4 region.*

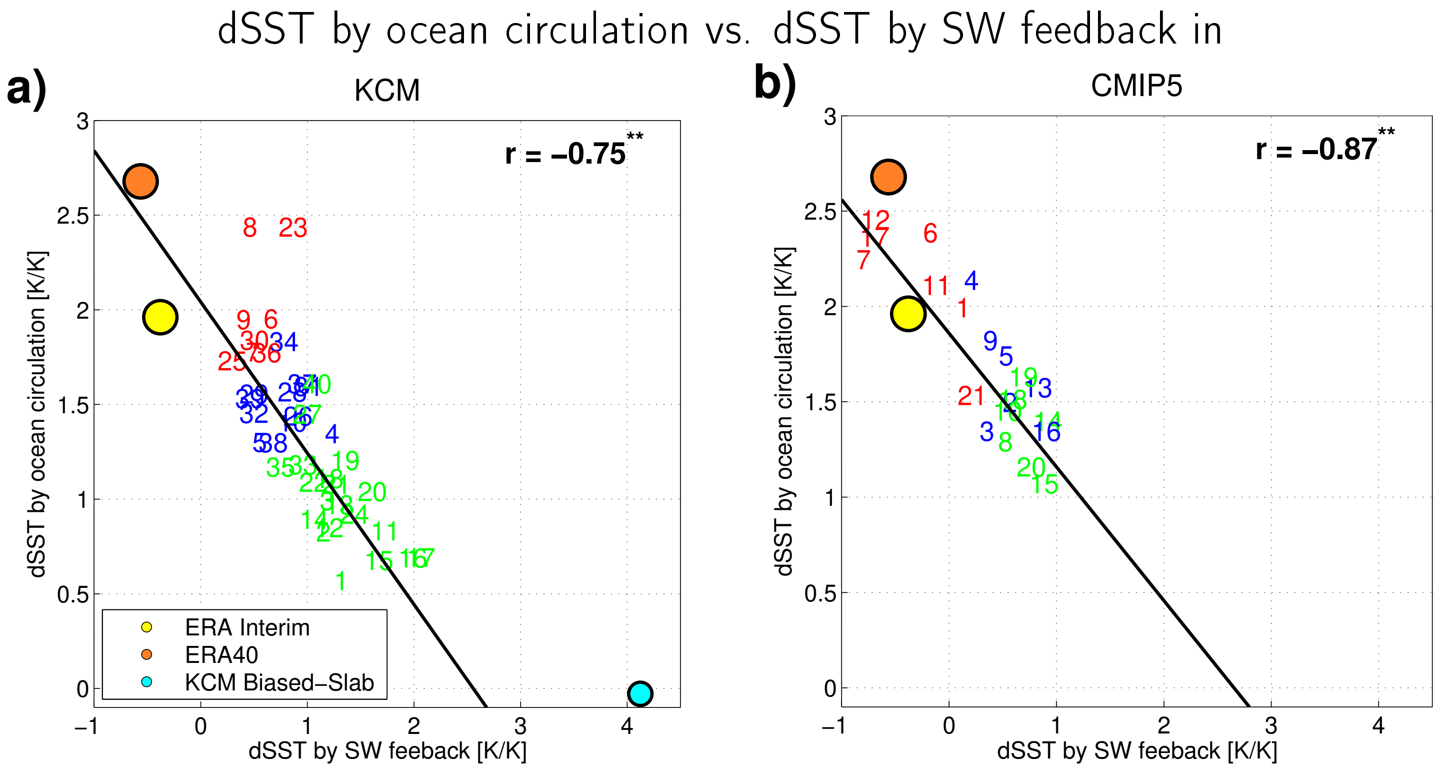

Figure 11: *a) For reanalysis and the individual KCM experiments, on the y-axis the SST change in Niño3 region due to ocean circulation (i.e.* $1 - dSST_{Qnet}$ *as shown in Fig. 10a) vs. the SST change by SW feedback (i.e.* $dSST_{SW}$ *as shown in Fig. 10a) on the x-axis; b) same as a) but here for the individual CMIP5 models; The color of the numbers indicate the sub-ensembles with STRONG (red), MEDIUM (blue) and WEAK (green) atmospheric feedbacks; The correlation between the individual experiments is shown in upper right corner and two stars indicate a significant correlation on a 99% confidence level.*

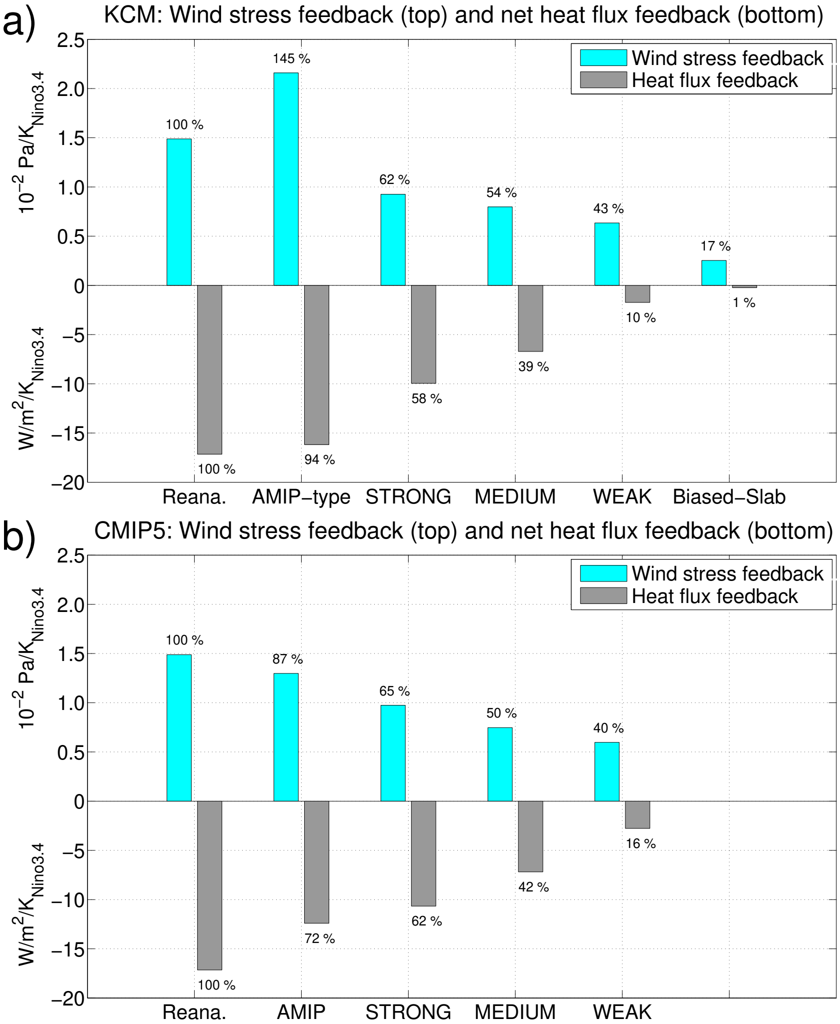

Figure 12: *a) Wind stress feedback in the Nino4 region (top) and heat flux feedback in the Nino3 and Nino4 region (bottom) in ERA-Interim/ERA40 reanalysis, KCM AMIP-type, STRONG, MEDIUM, WEAK sub-ensembles and Biased-Slab Ocean experiment; b) same as a) but here for CMIP5 AMIP, STRONG, MEDIUM and WEAK sub-ensembles; The the strength of the feedbacks relative to the observed feedbacks is shown at the bars as number in %.*

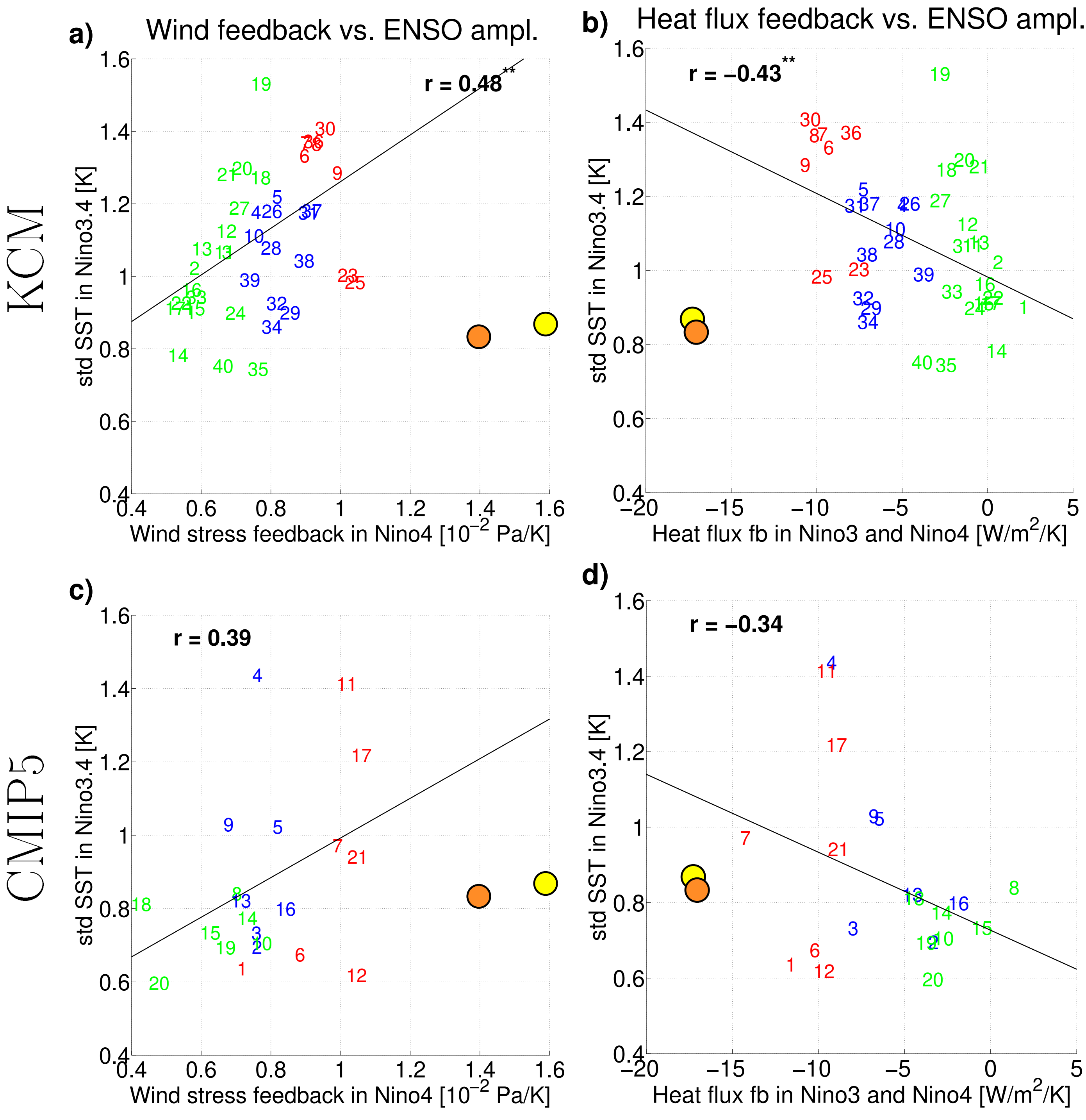

Figure 13: *a) Zonal wind stress feedback in the Nino4 region on the x-axis vs. standard deviation of SST anomalies in the Nino3.4 region on the y-axis for ERA-Interim/ERA40 reanalysis and the individual KCM perturbed physics experiments; b) same as a), but here the heat flux feedback in the Nino3 and Nino4 region on the x-axis; c-d) same as a-b), but here for the CMIP5 multi model ensemble. The color of the numbers indicate the sub-ensembles with STRONG (red), MEDIUM (blue) and WEAK (green) atmospheric feedbacks; The correlation between the individual experiments is shown in upper right/left corner and one (two) stars indicate a significant correlation on a 95% (99%) confidence level.*