https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

# Improving the simulation of El Niño Southern Oscillation within the ICON model by flux adjustments and parameter tuning

by Dietmar Dommenget[1], Helmuth Haak[2], Dakuan Yu[2] and Wolfgang A. Müller[2]

[1] Monash University, Clayton, VIC, Australia
[2] Max-Planck Institute for Meteorology, Hamburg, Germany

Corresponding authors: Dietmar Dommenget (dietmar.dommenget@monash.edu) and Wolfgang A. Müller (wolfgang.mueller@mpimet.mpg.de)

## Abstract

In this study we present the ICON model with strongly improved simulation of El Niño Southern Oscillation (ENSO). Starting from ENSO performance below the average of state-of-the-art (CMIP) climate models, the model now performs at a level comparable to, or exceeding, that of the best published CMIP models. We achieved this by a series of artificial corrections. First, we introduced flux adjustments of the net heat flux and wind stress to correct biases in the tropical mean sea surface temperature (SST) and mean momentum flux. We further optimized a set of atmospheric tuning parameters to minimize a cost function based on ENSO metrics. This has led to substantial improvements in large-scale SST amplitude, pattern, seasonality, non-linear behaviour, Bjerknes and cloud feedbacks. All three elements of our approach have contributed to improved ENSO simulations, but the largest improvements come from the heat flux adjustment that reduces mean SST biases. The tropical flux adjustment also improved tropical mean precipitation patterns but had only minor improvements in the mean climate outside the tropics. The tuning of atmospheric model parameters to improve ENSO shows negative side effects on the global mean climate, which were not part of our optimization cost function. The results suggest that flux adjustments of the tropical mean state can be beneficial for ENSO variability and should help to improve seasonal to decadal forecasting. Tuning of model parameters is more problematic, can lead to unwanted side effects and may require substantially larger computational resources than considered in this study.

## Short Summary

We present a climate model development approach with strongly improved simulation of the leading mode of natural climate variability, the El Niño Southern Oscillation (ENSO). Starting from performance below state-of-the-art climate models, the model is now comparable to the best published climate models. We achieved this by a series of correction methods. We introduced adjustments to the information exchange between the atmosphere and ocean. We also optimised a set of atmospheric processes.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

# 1. Introduction

State-of-the-art climate models, such as the Coupled General Circulation Models (CGCMs) of the Coupled Model Intercomparison Project (CMIP) phase 6, are highly complex models. They include many sub-models, with each being developed based on observations and theoretical considerations. The aim of these models is to simulate the current mean climate, its internal natural variability and its response to external forcings.

The current generation of CMIP6 models has improved simulations of the climate compared to previous generations in nearly all aspects of the climate system, including the mean state, its natural variability, and likely its response to external forcings. However, the current generation also still have significant biases in all these aspects, where many aspects of the models have common biases [Andrews et al. 2012; Flato et al. 2013; Luo et al. 2018; Byrne et al. 2025; Tian and Dong 2020; Tian et al. 2024]. These persistent common biases undermine the value and confidence we have in these models.

The complex interaction within a CGCM lead to biases in the climate simulations relative to the observation that are larger than expected from the development of the individual sub-models forced with observed boundary conditions. This is suggesting that a significant part of these model biases cannot be easily resolved by improving sub-models or by improving specific parameterisations in the models.

The past generations of CMIP models have seen several strategies to deal with CGCM model biases that do not directly focus on the improvement of the sub-models. Here we like to focus on two aspects: flux adjustments, also known as anomaly coupling, and model parameter tuning.

A flux adjustment, in general, is an artificial correction term added to model tendency equations to correct or adjust the mean tendencies to result into a mean state that is closer to observed or other desired mean state [Manabe and Stouffer 1988; Sausen et al. 1988; Schneider 1996; Collins et al. 2006; Irvine et al. 2013]. Previous generations of CMIP models, CMIP3 and earlier, have often used this approach to correct the mean Sea Surface Temperature (SST) to be closer to the observed state. Current generations of climate models also use this approach in the context of seasonal climate forecasts. Several studies have shown that introducing flux adjustments to improve the mean state has also improved the simulation of different aspects of climate variability [e.g., (Magnusson et al. 2013a; Vecchi et al. 2014; Li et al. 2026)].

The tuning of uncertain model parameters is often aiming at the same outcome as flux adjustments. It changes the values of different model parameters within any of the model's sub-models to achieve a better simulation of the climate as a whole or in specific aspects of the climate system [e.g., Mauritsen et al. 2012; Dommenget and Rezny 2018; Hourdin et al. 2017]. In essence, both methods represent artificial corrections of the CGCMs that are not desirable but are often necessary to improve the model performance.

The aim of this study is to test the skills these two concepts in improving the simulation of the El Niño Southern Oscillation (ENSO) in a CGCM. ENSO, while only being one specific aspects of the climate system, integrates many elements of the climate system. It therefore is a good proxy on how these methods may work for the large climate system.

The current generation of CGCMs can simulate ENSO with good skill in many different aspects of the phenomenon and have been improving in doing so over several generations [Bellenger et al. 2014; Vijayeta and Dommenget 2018; Planton et al. 2021]. However, they are still far from being an exact replica of the real world. Evaluations of the CMIP models in simulating

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

ENSO have pointed out many important biases in simulating ENSO, including biases that are common between models [Bellenger et al. 2014; Vijayeta and Dommenget 2018; Planton et al. 2021].

A number of studies have illustrated that some biases in simulating ENSO in CMIP models are directly related to biases in the mean state [Magnusson et al. 2013b; Bayr et al. 2018]. This would suggest that flux adjustment of the mean SST in models that have large mean state biases could improve the ENSO simulations. However, previous studies that implemented flux adjustments to improve the simulation of ENSO had mixed outcomes with some improvements and some degradations [Spencer et al. 2007; Kröger and Kucharski 2011; Magnusson et al. 2013b; Neelin and Dijkstra 1995].

Further, several studies have also illustrated how the simulation of ENSO variability is directly linked to specific processes in the climate system, such as equatorial (eq. hereafter) Pacific clouds, precipitation or atmospheric convection [Guilyardi et al. 2009; Watanabe et al. 2011; Bayr et al. 2019; Ma and Jiang 2021; Xu et al. 2025]. This in turn suggests that model parameter tuning related to such processes could potentially improve the simulation of ENSO.

This study presented here builds up directly on several previous studies. Dommenget and Rezny [2018] have tested the general idea of flux adjustments and tuning to improve a climate model with a much simpler Globally Resolved Energy Balance (GREB) model. The study found that both methods can improve model simulations, but both methods can also lead to some serious degradations of the simulations. It further also found that the tuning approach is far more complex and computationally expensive but does not achieve better outcomes than the much simpler and faster flux adjustment approach.

Yu et al. [2026; hereafter Y26] used the atmospheric model of the Icosahedral Nonhydrostatic eXtended Predictions and Projections (ICON-XPP) Earth System Model, which is a joint project between climate research institutes and the Deutscher Wetterdienst (DWD) [Müller et al. 2025] to systematically tune the atmospheric parameters of the model to improve ENSO simulations. They found that by using a systematic optimization approach they could substantially improve atmospheric ENSO metrics in atmospheric model simulations forced by observed SST (referred to as AMIP simulations). The concepts of optimizing the tuning parameters as described in Y26 will be the basis for the tuning approach in this study here. We further also use the same ICON-XPP model with the same 6 atmospheric tuning parameters which were found in Y26 to be potentially affecting ENSO variability.

This study is organised as follows: The following methods Section 2 will introduce all datasets, models and methods used in this study. Since this study is about introducing a model development approach, parts of the important results are within this method section. In particular, the Sections 2e-h outlining the tuning (optimization) approach present some important results. The performance of the model and the individual methods in simulating ENSO will be discussed in the results Section 3. The final part of this section will also evaluate how the model performs globally in aspects not related to ENSO. We conclude this study with a summary and discussion Section 4.

# 2. Methods

In the following sections we outline the model development approach of this study. We start by introducing the observational data sets and the ICON-XPP model configuration, which is the basis of this study. We then outline the development approach for improving the ENSO performance in Section 2c. This includes introducing, the flux adjustment method, ENSO

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

metric, estimating parameter sensitivities, our cost function and the optimization scheme in Sections 2d-h. In the final Section 2i we introduce further statistical methods used in the data analysis.

## a. Observational datasets

For the evaluation of the Planton et al. [2021; hereafter P21] ENSO metrics we use the same dataset as in the original publication. SST, zonal wind stress ($\tau_x$) and surface net heat flux are from the TropFlux data [Kumar et al. 2012]. Monthly tropical precipitation data is taken from the Global Precipitation Climatology Project (GPCPv2.3) dataset [Adler et al. 2003]. The Sea Surface Height (SSH) data are from the ocean reanalysis dataset GODAS [Saha et al. 2006]. All these observational reference datasets cover the time period from 1980 to 2018, which we use for our data analysis.

For additional SST analysis we also use the Hadley Centre global sea Ice coverage and SST (HadISST) for the time period 1950 to 2018 [Rayner et al. 2003]. The thermocline depth is computed on the bases of upper ocean temperature from ocean reanalysis datasets from Simple Ocean Data Assimilation version 3 (SODA3) for the period 1980 to 2017 [Carton and Giese 2008]. Additional data for zonal and meridional wind stress for the period 1955 to 2014 and air temperature in 200 hPa and Specific humidity in 400 hPa for the period 1979 to 2014 are taken from the ECMWF ReAnalysis version 5 (ERA5) [Hersbach et al. 2020].

## b. Model simulations

The ICON (Icosahedral Nonhydrostatic) model is the baseline of this study. It is a modelling framework for weather, climate, and environmental prediction (www.icon-model.org). Here, we use a newly designed Earth System configuration which provides the baseline for the next generation of climate predictions and projections (hereafter ICON-XPP – where XPP stands for eXtended Predictions and Projections) [Müller et al. 2025]. It is prepared for CMIP7 and operational seasonal to decadal climate forecasting. Here we use a lower resolution configuration, which is an atmospheric model with an unstructured triangular grid, R2B4 (~160 km, about 1.4°) and the ocean model at R2B6 (~40 km, about 0.36°). All simulations discussed here are integrations of historical forcings from 1950 to 2000 starting at the same initial condition.

## c. Development approach

We use three different empirical methods to improve the ENSO performance: First, we use net heat flux adjustments and second wind stress corrections to improve the mean SST and the mean momentum flux in the model simulations, see Section 2d. Third, we tune a set of atmospheric model parameters to optimise the ENSO performance. This approach is much more complex than the flux adjustment approach and includes several aspects that need to be considered.

Here we follow the approach of Y26. The optimisation approach includes a selection of ENSO metric, estimating parameter sensitivities, defining a cost function and an optimization scheme. The details are outlined in Sections 2e-h. Table 1 shows all sets of model simulations discussed in this study.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 1: List of ICON-XPP model simulations analysed in this study

| Model Simulations | | |
|---|---|---|
| **run** | **length [yrs]** | **Comments** |
| **ICON-ctrl** | 1950-1999 | control run; historical scenario of original model |
| **ICON-qflx** | 1950-1999 | as **ICON-ctrl** but with additional net heat flux adjustments for mean SST |
| **ICON-acpl** | 1950-1999 | as **ICON-qflx** but with additional wind stress corrections |
| **ICON-tune** | 1950-1999 | as **ICON-acpl** but with additional optimization of 6 atmospheric model parameters |
| **ICON-test-opt** | 3x[1950-1999] | as **ICON-tune** but with variations in the cost function weights |
| **Sensitivity runs** | 2x6x[1950-1999] | for each of the 6 parameters we have 2 perturbation runs with net heat flux and wind stress corrections starting from the **ICON-acpl** setup |

The ICON-ctrl applies the same experimental setup as in Müller et al. [2025] and serves as our reference control simulation that does not include any specific ENSO model development. The ICON-qflx simulation is identical to the ICON-ctrl with the addition of including the net heat flux adjustment terms as describe in the next Section. The ICON-acpl simulation further includes also the zonal and meridional wind stress corrections. With the main coupling fluxes now being corrected, we can also consider these simulations as an anomaly coupled (acpl) simulation, although salinity fluxes are not corrected. The ICON-tune also includes additional atmospheric parameter optimisations as outlined in Sections 2e-h.
The three simulations with different approaches of improving the ENSO performance allow us to evaluate how each of the different elements contribute to the improved ENSO performance.

## d. Flux adjustments

The coupling of the ocean to the atmosphere in the ICON-XPP model includes a net heat flux, zonal and meridional wind stress forcing. Imbalances in the exchange fluxes and model physics lead to biases in the mean SST and upper ocean circulation. Figure 1a shows the annual mean bias in SST in the ICON-ctrl run in the tropical region. The simulated tropical SST is generally too cold and has a bias pattern that is typical for CMIP6 models. The central eq. Pacific region is too cold and the eastern coastal upwelling regions in the Pacific and Atlantic are too warm. The annual mean SST bias has a root-mean-square (RMS) value of 0.9$^{o}$C, which is comparable other CMIP6 models.
The monthly climatology SST, with the annual mean bias removed, has RMS values relative to observations between 0.5$^{o}$C to 2$^{o}$C, with large values in the eastern eq. regions, coastal regions and higher latitudes, see Fig. 1b. This indicates that in addition to the annual mean biases, there are significant biases in the seasonal cycle at most locations.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

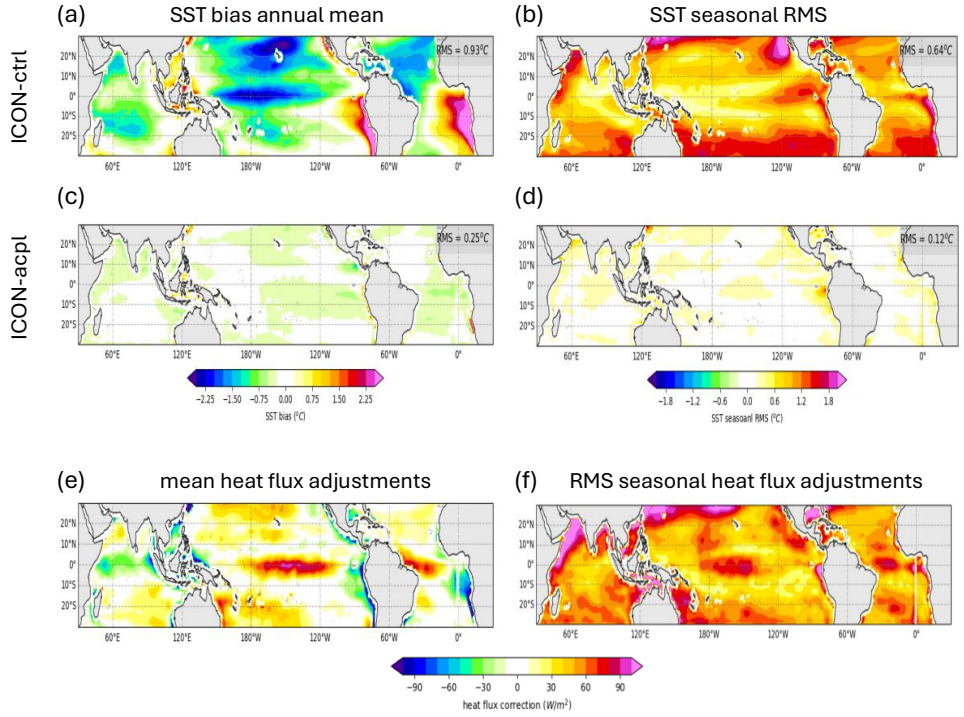



*Figure 1: SST biases and heat flux adjustments in ICON-XPP simulations: annual mean SST bias in ICON-ctrl (a) and ICON-acpl (b), RMS error of the 12 months SST climatology with the annual mean bias removed in ICON-ctrl (c) and ICON-acpl (d) relative to the observed values. Annual mean heat flux adjustments (e) and RMS value of the 12 months heat flux adjustments with the annual mean adjustments removed (f) in ICON-acpl.*

By adding an artificial correction term to the heat flux and wind stress received by the ocean model we can correct the mean SST and momentum flux in the coupled model. This can also be thought of as an anomaly coupling between atmosphere and ocean, since the models mean fluxes are not used, but are replaced by the adjusted mean fluxes.

Here we only focus on the tropical regions (30°S to 30°N), with regions outside this domain not being corrected in any way. We compute the heat flux and wind stress adjustments by an iterative procedure in which we run the coupled model for 5yrs to 30yrs to estimate the adjustments in heat flux and wind stress needed to get the correct mean 12-month climatology in SST and wind stress.

After each iteration the mean net heat flux is adjusted proportional to the local annual mean SST biases to the observed with a coupling constant of 10 W/m$^2$/°C. The seasonal cycle is corrected for each calendar month proportional to the mean bias in the calendar monthly mean surface layer heat tendency, assuming a 30m surface layer. After about 4-6 iterations, the SST bias and seasonal RMS values reduces to less than 0.3°C, see Fig. 1c-d.

The wind stress biases in the ICON-ctrl run (not shown) are strongly affected by the mean SST biases and are therefore different from the biases in ICON-acpl run. We therefore compute the wind stress correction terms in combination with the heat flux adjustments. The resulting

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

wind stress corrections shown in Fig. 2c-d are an indication of an atmospheric model bias in producing the correct wind stresses in the ICON-acpl run although the models see the correct mean SST.

The zonal wind stress correction (bias to observed; Fig. 2c) is mostly enhancing easterly wind stress for most of the tropical Pacific and in particular in the western part. The strength of the correction term is small if compared against the mean zonal wind stress (Fig. 2a). The meridional wind stress correction is also weak and primarily enhances wind stress convergences onto the eq. region. The sum of the wind stress computed by the ICON-XPP model and the wind stress correction terms give a mean wind stress that is very close to the observed, see Fig. 2e-f.

For each of the ICON-XPP runs discussed in this study we recompute the heat flux and wind stress correction terms (ICON-qflx has no wind stress correction). The resulting biases in SST and wind stress are similar in magnitude to the ones shown in Figs. 1c,d and 2e,f. The heat flux adjustment terms are also similar in the annual mean pattern, and strength of the monthly variations.

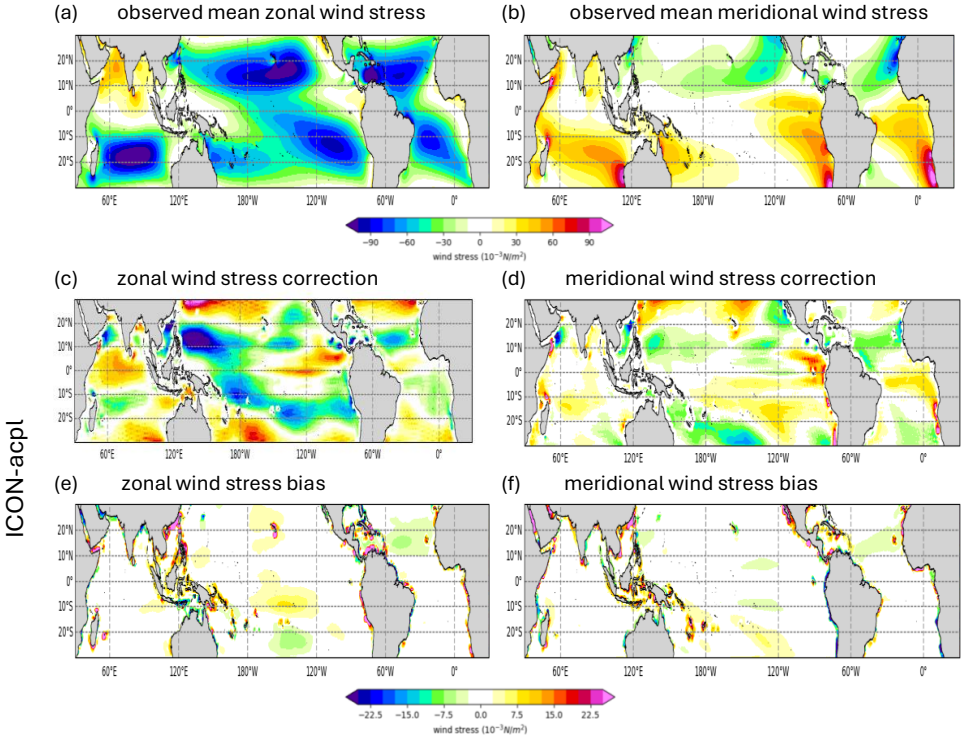



*Figure 2: mean observed zonal (a) and meridional wind stress (b). Annual mean zonal (c) and meridional (d) wind stress correction in the ICON-acpl simulation. Mean biases in the zonal (e) and meridional (f) wind stress in the ICON-acpl simulation.*

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

## e. ENSO metric

We use the P21 ENSO metrics to evaluate the ENSO performance in the ICON-XPP model simulations. It contains 24 difference metrics in four categories: ENSO variability, process, teleconnection and the tropical Pacific mean climate, see Table 2.

For each of these metrics a zero-dimensional bias value can be computed to quantify how well the model fits to observations. Each of these metrics has additional diagnostic levels to further evaluate the metrics performance as function of space or time dimensions. Most metrics have a level 1 or 2 diagnostic in which the ENSO performance is evaluated along one dimension (e.g., as function of latitude, longitude or lead time), which we will use for the development and evaluation of the model runs.

We select 13 of the 24 metrics to reduce the complexity in the analysis. Since the primary aim of this study is to improve ENSO dynamics, we primarily focus on the metrics of ENSO variability and processes. These are 12 metrics, of which two have the same level 1 diagnostic (EnsoSstTsRmse and EnsoDuration). We further do not consider the EnsoSstDiversity metric as it does not have a clear level 1 or 2 diagnostic as function of one dimension. We also do not consider EnsodSstOce metric since the level 2 diagnostic does not have a clear variation along the time dimension. In addition to the 9 remaining ENSO variability and processes we use 4 of the 8 tropical Pacific mean climate metrics, neglecting the 4 seasonal cycle metrics.

Figure 3 shows the 13 metrics we use to evaluate ENSO performance for the ICON-XPP runs. The metrics highlight some of the biases the ICON-ctrl experiment: ENSO amplitude is much weaker than observed (Fig. 3a), the ENSO pattern is reaching too far to the west (Fig. 3b), the seasonality in the ENSO variability is not clear and not peaking to the end of the year as observed (Fig. 3d), and the non-linear behaviour (skewness) of ENSO is mostly missing (Fig. 3e). Nearly all ENSO processes, including the three elements of the Bjerknes feedbacks [Bjerknes 1969; McPhaden et al. 2006] are underestimated (Fig. 3f-h). Most of these biases are common among the current generation of CMIP models. Nearly of all these ENSO metrics are substantially improved in the ICON-tune simulation, which will be discussed in more detail in the results sections.

We further combine 12 of the 13 ENSO metrics to one vector for a more compact evaluation of the ENSO performance, see Fig. 4. Here we do not include the SST mean metric in our model evaluations, since our model development includes flux adjustment to control the mean SST. We normalize each metric by subtracting the observed value and divide by the ICON-ctrl run bias value, and the x-dimension are normalized to unity, resulting into a non-dimensional x-axis of length 12. We further reduce the amplitude of the EnsoSstTsRmse and EnsoFbSshSst metrics by a factor of 0.5, as our analysis suggests that the statistical variations in these two metrics are less clear.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 2: List of all P21 ENSO metric. The metrics used for model development are marked by ✓. Alterations of some metrics are also noted.

| ENSO metric | | | | |
|---|---|---|---|---|
| | **metric** | **description** | **used** | **alterations** |
| variability | EnsoAmpl | SST amplitude | ✓ | |
| | EnsoSstLonRmse | eq. SST pattern | ✓ | |
| | EnsoSeasonality | Seasonal SST stdv | ✓ | SST normalised |
| | EnsoSstSkew | SST skewness | ✓ | |
| | EnsoSstTsRmse | SST auto-correlation | ✓ | Symmetric / smoothed |
| | EnsoDuration | ENSO event duration | x | |
| | EnsoSstDiversity | SST pattern diversity | x | |
| processes | EnsoFbSstTaux | Wind-SST feedback | ✓ | |
| | EnsoFbTauxSsh | SSH-Wind feedback | ✓ | |
| | EnsoFbSshSst | SST-SSH feedback | ✓ | |
| | EnsoFbSstThf | Heat-SST feedback | ✓ | |
| | EnsodSstOce | Oceanic part of SST tendencies | x | |
| climate | BiasSstLonRmse | eq. mean SST | ✓ | |
| | BiasPrLatRmse | zonal mean precip. | ✓ | |
| | BiasPrLonRmse | eq. mean precip. | ✓ | |
| | BiasTauxLonRmse | eq. mean zonal wind stress | ✓ | |
| | SeasonalPrLatRmse | Seasonal cycle zonal mean precip. | x | |
| | SeasonalPrLonRmse | Seasonal cycle eq. mean precip. | x | |
| | SeasonalSstLonRmse | Seasonal cycle eq. mean SST | x | |
| | SeasonalTauxLonRmse | Seasonal cycle eq. zonal wind stress | x | |
| teleconnections | EnsoPrMapDjfRmse | precipitation teleconnection map in DJF season | x | |
| | EnsoPrMapJjaRmse | precip. in JJA | x | |
| | EnsoSstMapDjfRmse | SST in DJF | x | |
| | EnsoSstMapJjaRmse | SST in JJA | x | |

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

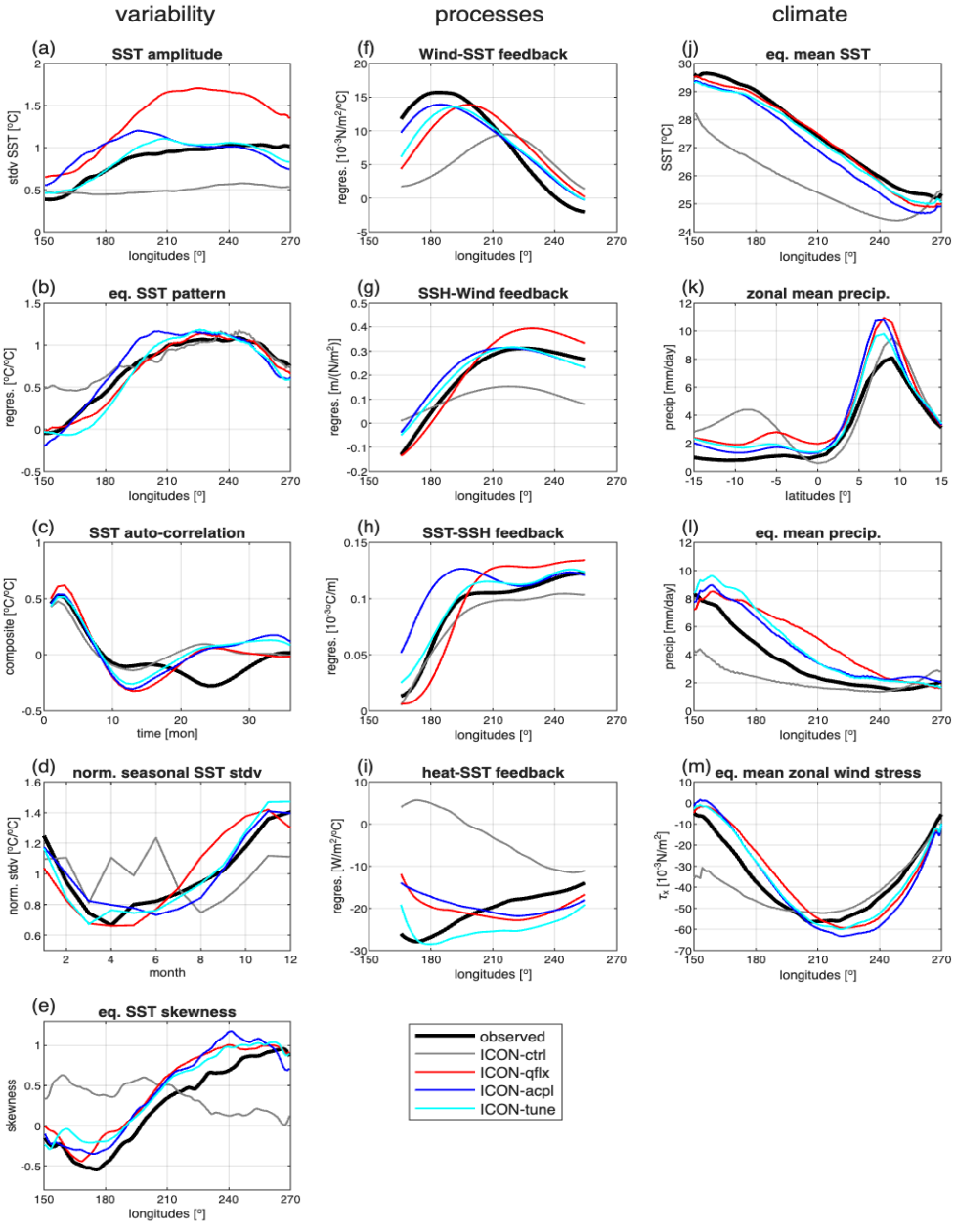



*Figure 3: Selected P21 ENSO metrics for variability (left column), processes (middle column) and mean climate (right column) for observations and ICON-XPP simulations. The mean zonal wind stress (m) does not include the correction term. See methods Section and Table 2 for more details on the ENSO metrics.*

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

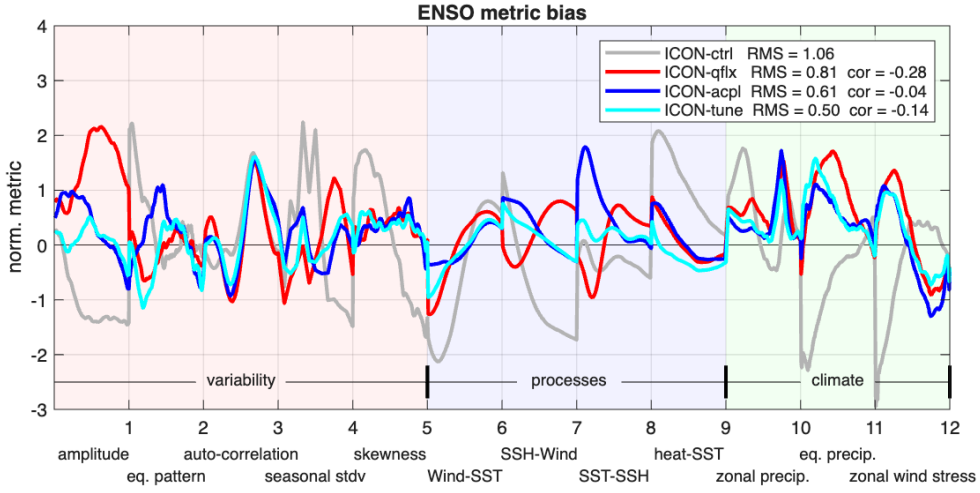



*Figure 4: Normalised bias vector for the ICON-XPP simulations. Each bias vector is based on the ENSO metric biases to observations in Fig. 3 combined to one normalised vector. Deviations from zero indicate biases relative to the observed value with the RMS value shown for each model in the legend. The correlation of the bias vector of ICON-ctrl with each of the other simulations are shown in the legend. Each unit one section on the x-axis presents a different P21 ENSO metric, with variability (0-5), processes (5-9) and mean climate (9-12), see x-axis labels. Scales of x- and y-axis are non-dimensional. See methods section and Table 2 for details.*

## f. Estimating parameter sensitivities

We use 6 parameters of the atmospheric model to tune the ENSO performance, see Table 3. The 6 parameters chosen are motivated by Y26, who found that these atmospheric process parameters are likely to influence ENSO dynamics. These parameters are part of the physical model of clouds, precipitation and small scale turbulence, see Table 3 for a short description. Parameters of the ocean model are not considered due to limitations in time and resources. The sensitivity of the ENSO-metric to each tuning parameter in ICON-acpl simulation is estimated by a set of two sensitivity runs. For each sensitivity run we vary a single parameter and recompute the flux correction (heat and wind stress), to assure that the mean SST in the sensitivity runs is the same as observed. Thus, we avoid that changes in process parameters affect the mean SST or mean wind stresses. We than conduct a 50yrs long sensitivity run to evaluate the ENSO metric. We use the ICON-acpl as our control run for which we change a single parameter by +/-1 one standard deviation (stdv) of the assumed parameters uncertainty, $\sigma_{para}$. The values of $\sigma_{para}$ are taken from Y26.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 3: List of atmospheric parameters tuned in this study.

| Tuning parameters | | |
|---|---|---|
| **short name** | **name in Y26** | **short description** |
| entrorg | tune_entrorg | Cloud physics: controls how much environmental air entrains into clouds. Also affects precipitation processes |
| pr0 | pr0 | Parameterisation of small-scale turbulence: controls how momentum and heat mix in neutral (non-stratified) turbulence |
| rprcon | tune_rprcon | Cloud physics: control on how cloud water turns into rainfall. |
| bliqasy | tune_box_liq_asy | Cloud physics: adjusts the asymmetrical response of liquid clouds to relative humidity |
| boxliq | tune_box_liq | Cloud physics: controls aspects of diagnosing cloud cover from liquid water content. affects how cloudy the model thinks it is. |
| tdecay | f_theta_decay | Parameterisation of small-scale turbulence: controls how quickly turbulent mixing decays at high atmospheric stability. |

Figs. 5 shows some results for the *pr0* (small-scale turbulence) and *rprcon* (cloud-precipitation physics) parameters for different ENSO metrics. For instance, we can see in Fig. 5a that an increase (decrease) in the pr0 parameter decreases (increase) the mean zonal wind stress all along the eq. Pacific.

Based on the perturbed parameter runs and the control run values we can estimate sensitivities of the ICON-acpl simulation ($\lambda$) to each parameter (p) and for each ENSO metric (m):

$$\lambda_p^m(\varphi) = \frac{d\xi}{dp}(\varphi) \approx \frac{\Delta\xi}{\Delta \mathrm{p}}(\varphi) \qquad [1]$$

Here $\xi$ the physical variable measured by the ENSO metric m (e.g., mean zonal wind stress as in Fig. 5a) and $\varphi$ is the physical dimension the metric is depending on (e.g., longitudes as in Fig. 5a). $\Delta\xi/\Delta$p is the change in $\xi$(p) relative to the control run. Note, $\lambda_p^m$ does not consider the observed values and it is therefore not measuring a sensitivity in the biases. It only estimates if $\xi$ would increase or decrease relative to the control run.

Fig. 5 right column, for instance, shows the estimated values of $\lambda_p^m(\varphi)$ and the ensemble mean for the perturbation runs with *pr0* or *rprcon* parameters for three different ENSO metric. $\lambda_p^m$ for the mean eq. zonal wind stress (Fig. 5b), for instance, is negative all along the eq. Pacific in both sensitivity runs and in the mean. The overall strength of $\lambda_p^m$ can be quantified by the root-mean-square (RMS) value along $\varphi$ ($\Delta_{par}$; see values in right columns of Fig. 5).

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

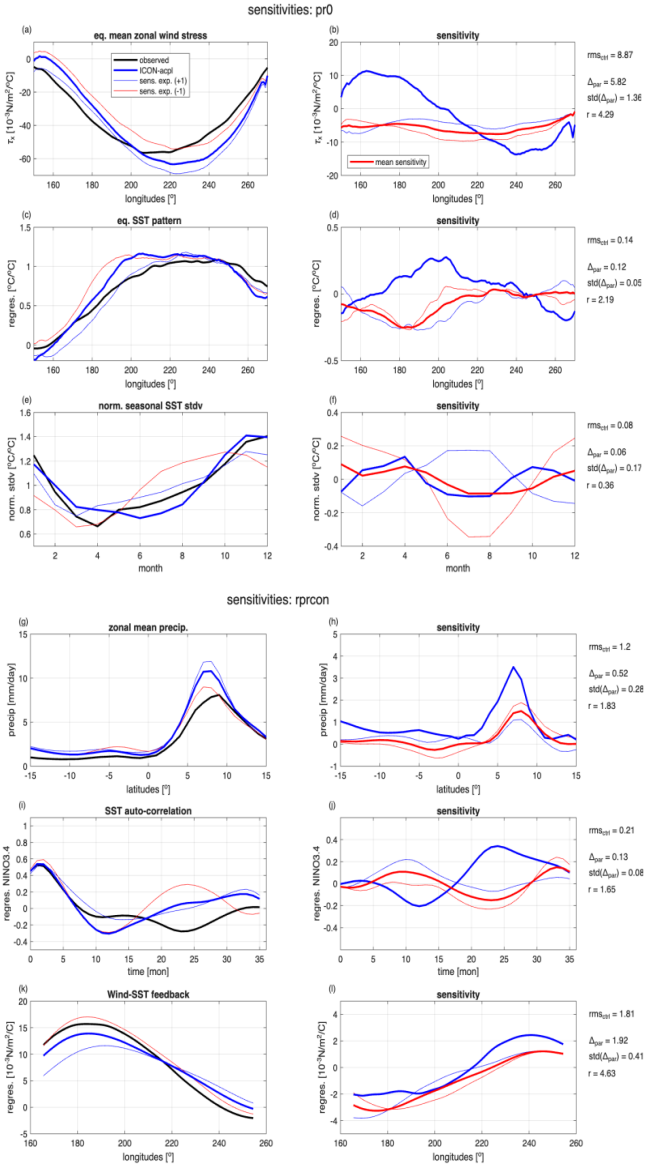



*Figure 5: Selected examples of ENSO metrics for different sensitivity experiments (left column). The right column shows the changes in ENSO metric (sensitivity) for the two sensitivity experiments, the mean of the two and the ICON-ctrl bias to observed for references. The upper three rows are for the pr0 parameter and the lower ones are for the rprcon parameters. The mean zonal wind stress (a,b) does not include the flux correction term.*

323

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

A comparison with the RMS-error of the control simulation ($\mathrm{rms}_{ctrl}$; see values in right columns of Fig. 5) gives a good indication if the model's sensitivity to the tunning parameter is strong enough to significantly affect the ENSO metric. The *pr0* parameter for the mean zonal wind stress (Fig. 5b), for instance, has a $\Delta_{par}$= 5.8 that is comparable in magnitude to the $\mathrm{rms}_{ctrl}$ = 8.9. Thus, this parameter could potentially significantly alter the model's bias in this metric.

We can further see that the two ensemble members for each parameter do not result into the same $\lambda_p^m$, leading to some spread. This spread gives us some indication of how uncertain these sensitivity estimates are. It is quantified by the stdv between the two estimates (*std*($\Delta_{par}$)). For the pr0 parameter for the mean zonal wind stress (Fig. 5b), for instance, *std*($\Delta_{par}$)*=1.36,* which is much smaller $\Delta_{par}$= 5.8, indicating this parameter has a significant impact on the mean zonal wind stress with only small uncertainties. The ratio, $\mathrm{r} = \Delta_{par}/std(\Delta_{par})$ gives a good first guess on how significant the sensitivity of an ENSO metric to a parameter is. These values range from relatively low or uncertain sensitivities with r = 0.36 (Fig. 5f) to strong or clear sensitivities with r = 4.63 (Fig. 5l).

The scale of $\lambda_p^m$ depends on the physical variable measured $\Delta\xi$ and the model parameter values Δp. Since each parameter will have a different range over which we can vary it, we need to normalize $\lambda_p^m$ to be able to compare sensitivities to different parameters. Similarly, comparing the sensitivity a parameter has on different ENSO metrics requires to normalise the ENSO bias by the scale of the biases. We therefore define a normalised sensitivity as:

$$\widetilde{\lambda_p^m}(\varphi) = \sigma_{para}/\mathrm{rms}_{ctrl} \cdot \lambda_p^m(\varphi) \qquad [2]$$

For each parameter we can again combine the 12 different sensitivities in the ENSO metrics to one vector for a more compact evaluation by normalizing the physical x-dimension $\varphi$ of each ENSO metric to unity, resulting into a non-dimensional x-axis of length 12, see Fig. 6.

Figure 6 shows all sensitivities vectors of all 6 parameters and the estimated uncertainty ranges. Note, unlike Fig. 4 the values do not relate to biased to the observed. Here zero values mean no sensitivity, positive (negative) values mean the parameter is increasing (decreasing) the ENSO metric relative to its value in the ICON-acpl run irrespective of what the observed values are. We can notice a few important findings here.

First, we can note that for nearly all parameters some ENSO metrics show sensitivities larger than the uncertainty ranges, indicating these parameters can significantly alter ENSO characteristics. For instance, the *entrorg* parameter (environmental air entrainment into clouds) clearly influences ENSO variability in terms of amplitude, pattern and skewness (Fig. 6a). It further also influences ENSO process and mean state in terms of heat-SST feedback and mean precipitation. In turn, some of the other ENSO metrics are not clearly affected by this parameter. The *tdecay* parameter (affecting turbulent mixing decay), however, has no clear impact on any of the ENSO metrics. We can also compare the sensitivities in the ICON-acpl simulation here with those found in Y26 for the ENSO metrics that can be evaluated in AMIP simulations (e.g., mean climate and atmospheric feedbacks). There is in general good agreements with most sensitivities having the same signs and strength.

Second, the RMS value of the normalised sensitivity vector gives us an overall indication of how strongly the corresponding parameter can affect ENSO. For most parameters the sensitivity RMS values are about 0.5 to 0.8, which is comparable in magnitude to the bias RMS of the ICON-acpl run (0.6; see Fig. 4). This suggests that changes in these parameters will have a significant impact on ENSO in the ICON-XPP simulations.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

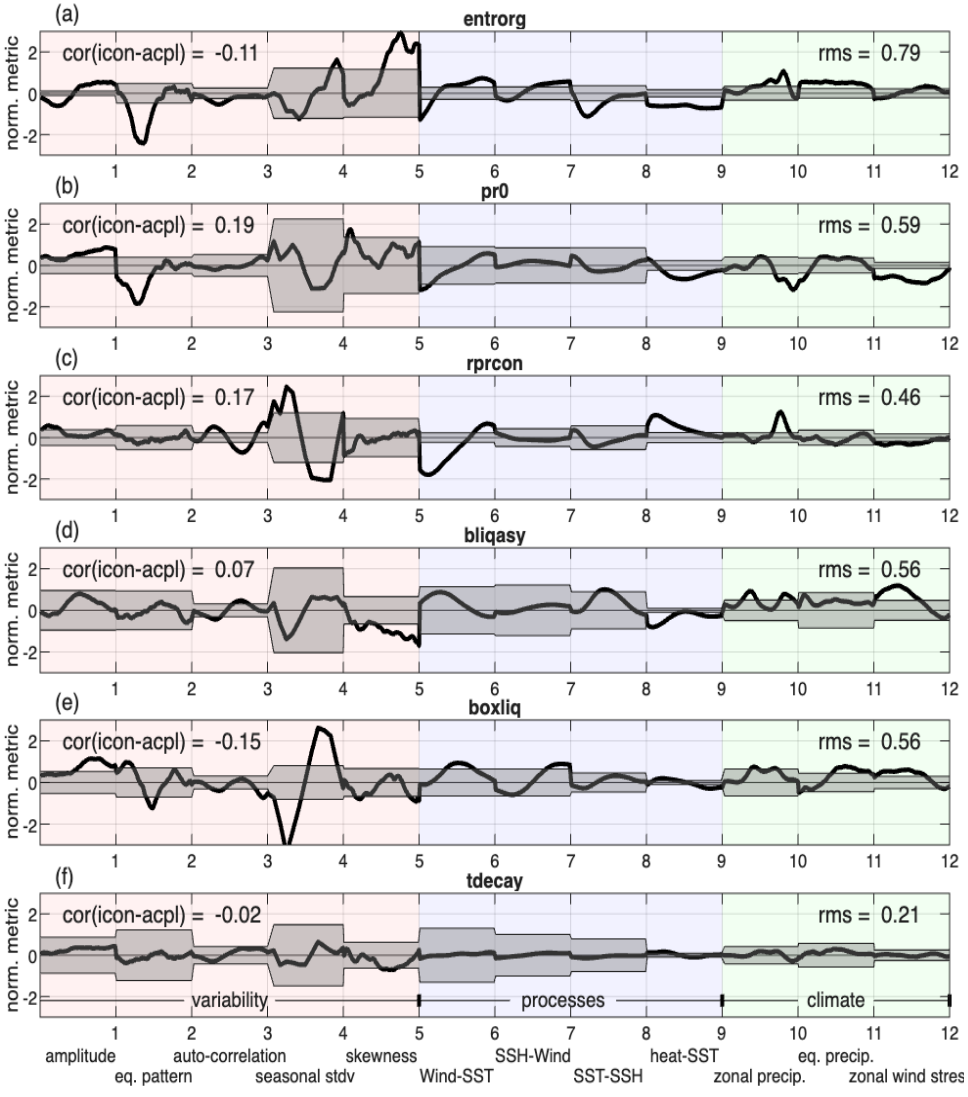



*Figure 6: Normalised combined ENSO metric parameter sensitivity ($\widetilde{\lambda_p^m}$) for each of the 6 tuning parameters. The grey shaded area mark the sensitivity uncertainty of +/- std($\Delta_{par}$). The correlation with the ICON-acpl bias vector (Fig. 4) and the RMS value for each sensitivity vector are shown in each plot. Each unit one section on the x-axis presents a different P21 ENSO metric, with variability (0-5), processes (5-9) and mean climate (9-12), see x-axis labels. Scales of x- and y-axis are non-dimensional. See methods section and Table 2 for details.*

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Here it is important to note the ICON-acpl bias vector (Fig. 4) is different from all the parameter sensitivity patterns (vectors), see correlation values in Fig. 6. This suggests that none of the tuning parameter can effectively correct the ICON-acpl bias pattern. However, the sensitivity patterns of the 6 different parameters are also different to each other, with absolute correlation values mostly <0.4. Thus, each parameter has different ways on affecting ENSO. This indicates that a combination of changes in several parameter can improve ENSO biases.

## g. Cost function

The aim of the cost function is to quantify how well the model simulation fits to the observations. The function is positive definite, with larger values when it does not fit well and smaller when it fits well. The cost function is defined by three elements:

$$f_{cost} = \Delta_{limit}(\Delta_{metric} + \Delta_{para}) \quad [3]$$

$\Delta_{metric}$ quantifies how well the ENSO metrics are simulated:

$$\Delta_{metric} = \sum_{i=1}^{M} \omega_i \frac{\delta_i}{\delta_{ctrl_i}} \; / \; \sum_{i=1}^{M} \omega_i \quad [4]$$

with $\delta_i$ the RMS-error of the metric and $\delta_{ctrl_i}$ the RMS-error of the metric in the ICON-ctrl ($rms_{ctrl}$ in Fig. 5 right). We further weight each metric with $\omega_i$. This way we can force that some metrics are more important than others (e.g., higher values for ENSO pattern than for ENSO skewness). $\Delta_{para}$ quantifies how much the model parameters deviate for the control run values:

$$\Delta_{para} = \alpha_{para} \frac{1}{N} \sum_{i=1}^{N} (\Delta\tilde{p}_i)^4 \quad [5]$$

with N=6; the total number of perturbed parameters. $\Delta\tilde{p}_i$ is the normalized parameter change relative to the control run:

$$\Delta\tilde{p}_i = \frac{\Delta p_i}{\sigma_{para_i}} \quad [6]$$

In eq. [5] we have the power of 4 to allow for small changes in the parameter without increasing the cost function significantly. Changes larger than $\sigma_{para}$ will increase the cost function quite strongly. $\alpha_{para} = 0.01$, is a scaling parameter to set the relative importance of model parameter deviations from the control with respect to the importance of the metrics. $\Delta_{limit}$ is a constraint which forces the parameter changes to not go outside meaningful values (e.g. negative values for positive definite parameters).

$$\Delta_{limit} = n_{limit} + 1 \quad [7]$$

with $n_{limit}$ the number of parameters outside the allowed range as in Y26.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

## h. Parameter optimisation

We use a Nelder-Mead scheme [Nelder and Mead 1965] to minimize the cost function eq. [3]. The Nelder-Mead scheme varies all 6 model parameters until a minimum in the minimization function is found. It will typically test about 1000 different parameter combinations until a minimum is found.

To estimate $\Delta_{metric}$ we need, in principle, to run the ICON-XPP model to compute $\delta_i$. This makes computing the cost function very expensive. We can however use the sensitivity estimates $\lambda_p^m$ to approximate $\Delta_{metric}$ by assuming that any metric bias $\delta_m$ is the result of the linear combination of the sensitivities $\lambda_p^i$:

$$\delta_m \approx \delta_m^{approx} = \sum_{i=1}^{N} \frac{r_i^m}{r_i^m + 0.3} \Delta\tilde{p}_i \cdot \lambda_i^m \qquad [8]$$

Here each sensitivity is scaled by $r_i^m$ to consider the uncertainty in the sensitivity ($r$ in Fig. 5 right). The larger the uncertainty in the sensitivity the less this parameter can influence the cost function.

This approximation is a strong simplification of optimizing the model parameters. Y26 showed that this linear approximation does have some skill in improving the atmosphere only model in ENSO metrics that only evaluates atmospheric processes. However, here we have a more complex problem with the fully coupled climate system and ENSO metrics that include fully coupled processes (e.g., SST variability).

We conducted 4 different optimization runs of the ICON-acpl model to evaluate the predictive skill of eq. [8]. For each of the optimization runs, we varied the values of $\omega_i$ in eq [4] to focus on different aspects in the parameter optimisation, resulting into different optimal parameters for ICON-acpl. We then recomputed the flux corrections for these parameters and conducted a 50yrs run with the ICON-acpl model. Figure 7 shows the bias patterns for the different test runs together with those predicted by eq. [8] and the ICON-acpl run as the control reference.

The correlation between the predicted bias pattern and the ICON-acpl results ranges between 0.3 to 0.6, suggesting the linear approximation of eq. [8] has moderate to good skill in predicting the ICON-XPP model performance. However, this is mainly due to the optimization runs being all similar to the ICON-acpl control run. If we only consider the bias changes relative to the ICON-acpl run, then the correlations are much weaker (see values in Fig. 7). This suggests the linear approximation of eq. [8] has only limited skill.

While the overall performance of the approximation is not that great in these 4 optimization runs, there are some indications that this approach is helpful. The first optimization run focussed on improving only the ENSO amplitude, and the resulting ICON-acpl run indeed improved the ENSO amplitude as predicted (Fig. 7a). Similarly, the second optimization focussed on the reducing the mean precipitation biases, and the resulting optimization run has indeed clearly reduced precipitation biases (Fig. 7b). The last two optimizations focussed on improving on most or all ENSO metrics together and in both cases the ENSO metric biases decreased in the optimization runs relative to the ICON-acpl control reference (Fig. 7c-d).

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

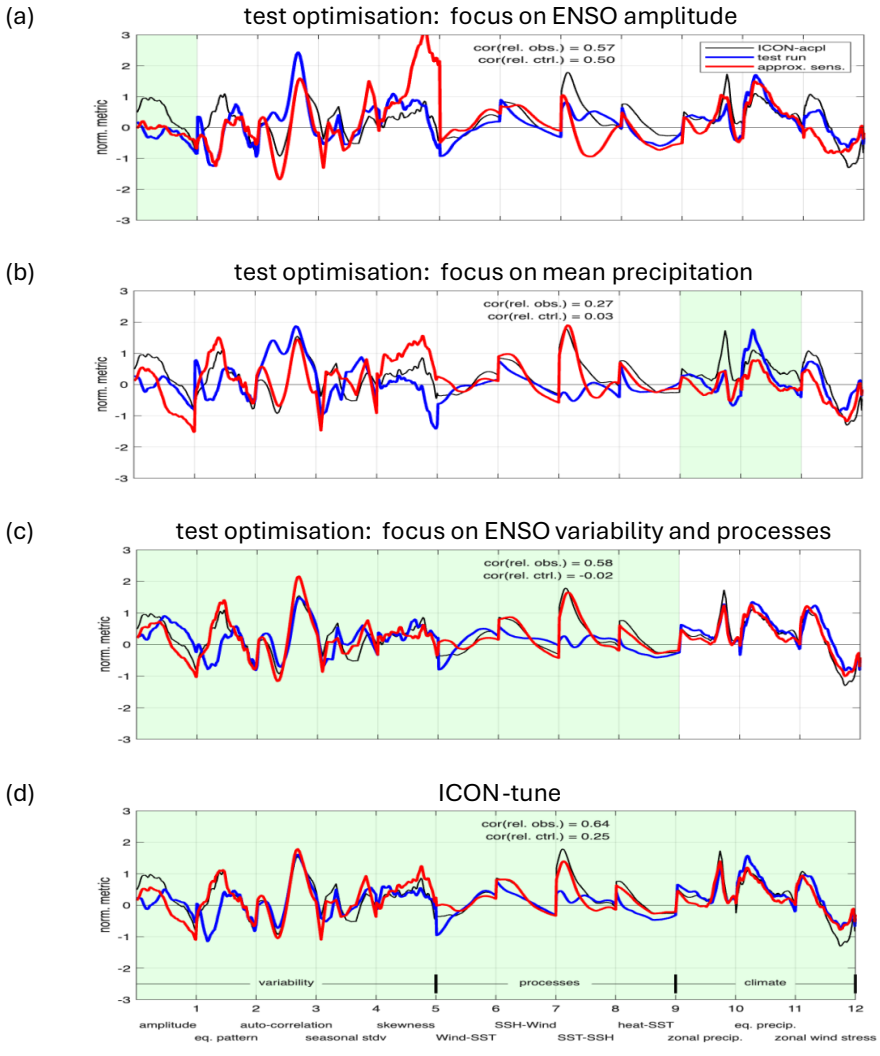



*Figure 7: Normalised bias vectors for different ICON-XPP simulations with optimised parameters (blue), the approximated sensitivities (red) and the ICON-acpl (black) for reference. Deviations from zero indicate biases relative to the observed value. The correlation between blue and red lines are shown in each panel (cor(rel. obs.)). In addition, the correlation between red and blue lines relative to the ICON-acpl (black) are shown too (cor(rel. ctrl.)). Each unit one section on the x-axis presents a different P21 ENSO metric, with variability (0-5), processes (5-9) and mean climate (9-12), see x-axis labels. Scales of x- and y-axis are non-dimensional. See methods section and Table 2 for details.*

458

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

### i. Statistical methods

All data analysis is performed on monthly mean data, with all datasets interpolated onto a common 1°x1° grid. Anomalies are always defined relative to the 12 month climatology. The magnitude of linear trends in the ICON-XPP simulation for global mean SST are all < 0.5°C per 100yrs.

The ENSO phase space statistics are calculated following Dommenget et al. [2023]. The x-axis is normalised SST of NINO3 (150°W–90°W, 5°S–5°N) and the y-axis is normalised eq. mean thermocline depth defined as the averaged over the eq. Pacific (130°E–80°W, 5°S–5°N), which is calculated for each ocean grid point as the maximum gradient in the upper ocean temperature profile. In the polar coordinate system, we can define the radius as the amplitude of ENSO and the angular velocity as the phase speed.

Motivated by the P21 ENSO metric EnsoSstDiversity we define ENSO diversity as the zonal position of the ENSO events by the longitude of the centre of mass of eq. SST anomalies. ENSO events are defined by 6 month running means passing the +/-1 stdv of NINO3 SST. We fit a normal distribution to the set of El Niño and La Nina event longitudes. The non-linear ENSO diversity is then defined as the differences of the El Niño minus the La Nina fitted normal distributions.

## 3. Results

We evaluate the performance of the model development in different aspects: first we evaluate how each of the different steps of the developments have contributed to the improved ENSO performance. We then highlight how different aspects of the ENSO variability and dynamics have improved in the ICON-tune simulation relative to the ICON-ctrl. Finally, we evaluate other aspects of the model performance that were not the focus of this study but still require some discussion to assure that improving the ENSO performance does not seriously degrade the climate simulations in non-ENSO related aspects.

### a. Performance of the different elements of the development

Figure 8 shows the overall normalised ENSO bias metric [P21] performance of all ICON-XPP simulations in comparison with the CMIP model distributions and observed references for all ENSO variability and process metrics (Table 2). We can first note that the CMIP distribution is centred around zero (by construction) and has a spread around +/-1 (each individual ENSO metric is normalized to have unity stdv in the CMIP5/6 ensemble). The observed reference value in this metric is negative and can, by construction, not be reached by a model simulation even if the model is a perfect replica of the real world. Some indication of the uncertainty in the observations is indicated by how much the ERA values deviate from the observed.

The CMIP distribution is relatively far away from the observed reference values, indicating that CMIP models have substantial biases. The spread within the CMIP distribution is relatively large with the best models being much closer to the observed values but still showing significant biases in their ENSO behaviour relative to the observed. The average CMIP model performance is far from being consistent with the observed ENSO behaviour [Bellenger et al. 2014; Vijayeta and Dommenget 2018; Planton et al. 2021].

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

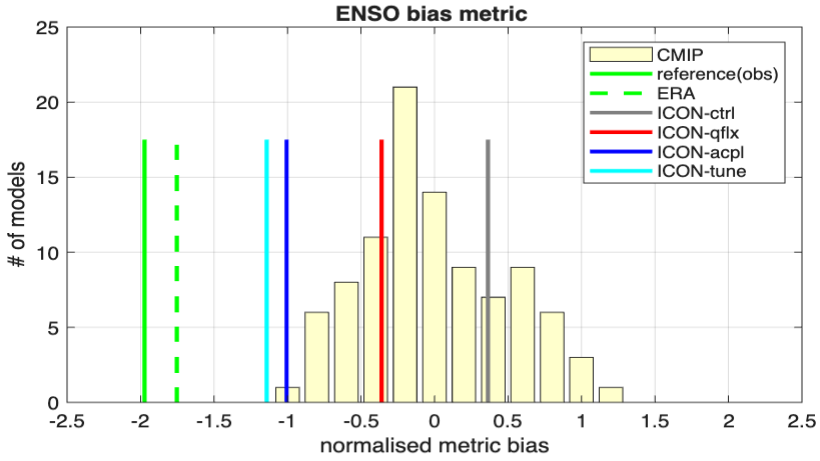



*Figure 8: Normalised mean ENSO biased metric in all variability and process metrics of Table 2. Zero values correspond to the CMIP ensemble mean value, with lower values closer to the observed reference.*

The ICON-ctrl model is below the average value of the CMIP distribution, while the ICON-tune simulation bias metric is smaller than any of the CMIP models in the combined variability and process metrics. This suggests that this model simulation has relatively good ENSO performance, if compared to the CMIP ensemble and the ICON-ctrl.

Figs. 3 and 4 show the 13 ENSO metrics we focussed on in the development. In essentially all metrics the ICON-tune simulation is closer to the observed value than the ICON-ctrl, with some of the metrics showing very realistic performances in the ICON-tune simulations. Most striking improvements can be seen in the ENSO pattern, seasonality, skewness, wind-SST and netheat-SST feedbacks. All these metrics are generally not well simulated in current CMIP models and are now substantially closer to the observed in the ICON-tune simulations.

We can further notice that with the introduction of each development step the ICON-XPP simulations reduce the biases in the ENSO metric (see RMS values in Fig. 4 and results in Fig. 8). This suggests that each of the three elements in our model development improves ENSO simulations in the ICON-XPP model. The bias vectors of all three development steps also have correlation below zero with the ICON-ctrl bias, suggesting the biases are now substantially different in nature to that of the ICON-ctrl (see correlation values in Fig. 4).

The introduction of heat flux adjustments to control the mean SST (ICON-qflx) substantially improves ENSO biases in the variability, processes and mean climate (Fig. 3). First, it reduces the eq. mean SST bias, which is what it is designed for (Figs. 1c,d and 3j). This improvement in the mean SST biases affects ENSO variability, process and the mean precipitation and wind stress substantially. It further changes the nature of the bias vector to be substantially different from the ICON-ctrl bias vector (see negative correlation values in Fig. 4).

The changes in the ENSO metrics from ICON-ctrl to ICON-qflx are similar to those found in Y26 when comparing the fully coupled model to the atmospheric AMIP simulations in terms of precipitation, mean zonal wind stress and atmospheric feedbacks. Thus both, the AMIP simulation of Y26 and the flux corrected ICON-qflx behave in a similar way, suggesting that the realistic mean SST make the atmospheric processes more realistic.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

The ENSO amplitude increased by a factor of two to three, changing from much weaker than observed to much stronger than observed (Fig. 3a). The eq. SST pattern becomes more realistic by reducing unrealistic amplitude in the western eq. Pacific (Fig. 3b). The seasonal cycle in SST stdv becomes more pronounced with a minimum variability around April and maximum around November, which is much closer to the observed than the ICON-ctrl simulation (Fig. 3d). The non-linear behaviour of ENSO as measured by the SST skewness has also substantially improved (Fig. 3e). It now shows the clear observed signature of positive skewness in the east and weakly negative skewness in the west.

All four ENSO feedback have been strengthened (Fig. 3f-i). In particular, the positive Bjerknes feedbacks of wind-SST and SSH-wind have become much stronger, and the negative heat-SST feedback has also become much stronger and more realistic. It is beyond this study to explain how each of the improvements ENSO metrics relate to physical processes, however we can see some indications in the results here.

The strengthening of the Bjerknes feedbacks strongly suggests that ENSO amplitude would increase, while the also strengthened negative heat-SST feedback will counteract this tendency. The shift in the eq. SST pattern (Fig. 3b) seem to be consistent with the shift in the SSH-wind, SST-SSH and heat-SST feedbacks, which all three are more strongly suppressed in the western Pacific than in the ICON-ctrl simulation (Fig. 3f-i). The additional strong changes in mean precipitation seem to clearly follow the changes in mean SST, which we will discuss in more detail in Section 3c.

The introduction of additional wind stress corrections in ICON-acpl does lead to some improvements relative to the ICON-qflx, although the improvements are not as clear as those from ICON-ctrl to ICON-qflx (Figs. 3,4 and 8). The correction of the mean wind stress only affects the ocean dynamics by improving the mean momentum flux into the ocean. It has no direct control on the mean SST, and therefore no direct control on atmospheric processes, as the mean SST is controlled by the heat flux adjustments. We can see that the SSH-wind and SST-SSH feedbacks both are reduced and more realistic in the eastern Pacific, while they are enhanced in the western Pacific (Fig.3g and h). This appears to reduce the ENSO amplitude and slightly changes the eq. SST pattern in the central to western Pacific (Fig.3a and b).

The final tuning of atmospheric processes further improves a number of aspects in ENSO variability, processes and mean climate (Figs. 3,4 and 8). However, the improvements from ICON-acpl to ICON-tune are not as strong as the improvements from ICON-ctrl to ICON-qflx. ENSO variability in amplitude, eq. SST pattern and seasonal cycle have improved, and all of the feedbacks process have improved in some ways (Fig.3a,b,d,f-i). Interestingly, this includes the SST-SSH feedback, which is not directly related to any atmospheric processes (Fig.3h). This is likely an indirect improvement by improving the other Bjerknes feedbacks (wind-SST and SSH-wind), and by improving the mean zonal wind stress (Fig.3m).

## b. ENSO performance

While the Planton et al. ENSO metric discussed above are an effective way to present ENSO performance, we now like to look at a few other ENSO statistics to better highlight the ENSO performance. Here we will primarily focus on the comparison of the ICON-tune and ICON-ctrl simulation with observations.

We start with the power spectrum of the NINO3 SST variability (Fig. 9), which is closely related to the ENSO amplitude and time scale metrics (Fig. 3a and c). The ICON-ctrl has too much high-frequency (intra-seasonal) variance, too little interannual to decadal variability

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

compared to the observed and no clear peak at interannual time scales. The ICON-tune simulation is close to the observed but the peak frequency in the observations is around 4yrs period and slightly shorter at 3yrs period for the ICON-tune simulation.

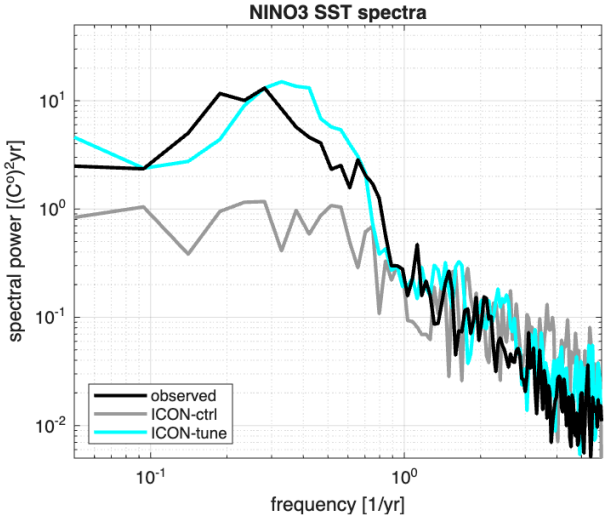



*Figure 9: Power spectra of NINO3 SST for observations and different ICON-XPP simulations.*

The ENSO non-linearity measured by the skewness (Fig. 3e) is an important aspect of ENSO that current generation of CMIP models fail to present well [e.g., P21]. Figure 10 shows maps of SST skewness in the tropical Pacific. The observed SST skewness is marked by strong positive skewness in the eastern eq. Pacific, suggesting stronger El Niño events than La Nina events. In the western eq. Pacific the observed SST skewness is weakly negative, suggesting stronger La Nina events than El Niño events in this region.
The ICON-ctrl simulation can essentially not capture the observed SST skewness with having mostly neutral or slightly positive skewness in most regions. The ICON-tune simulation has a much more pronounced SST skewness pattern with positive skewness in the eastern eq. Pacific and negative skewness in the western eq. Pacific similar to the observed pattern.
ENSO non-linear behaviour is not only affecting ENSO amplitude, as measured by the SST skewness, but also affects the ENSO time evolution in the phase space [e.g., Ohba et al. 2010; Okumura and Deser 2010; Dommenget et al. 2013; Kessler 2002; Dommenget et al. 2023], see Fig. 11. In the observed ENSO phase space we have a clear asymmetric evolution, with the El Niño to discharge as the stronger phases in average and the La Nina to recharge as the weaker phases in the mean. The observed phase transition speed is relatively slow around the La Nina phases and fastest between the El Niño to discharge phase (Fig. 11b).

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

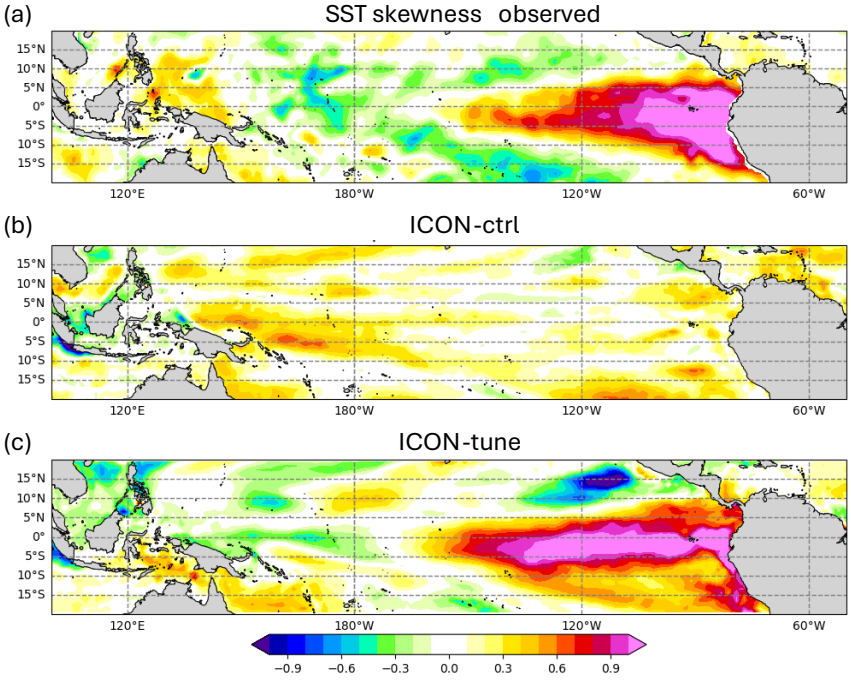



*Figure 10: Skewness of monthly mean SST anomalies in observations and ICON-XPP simulations.*

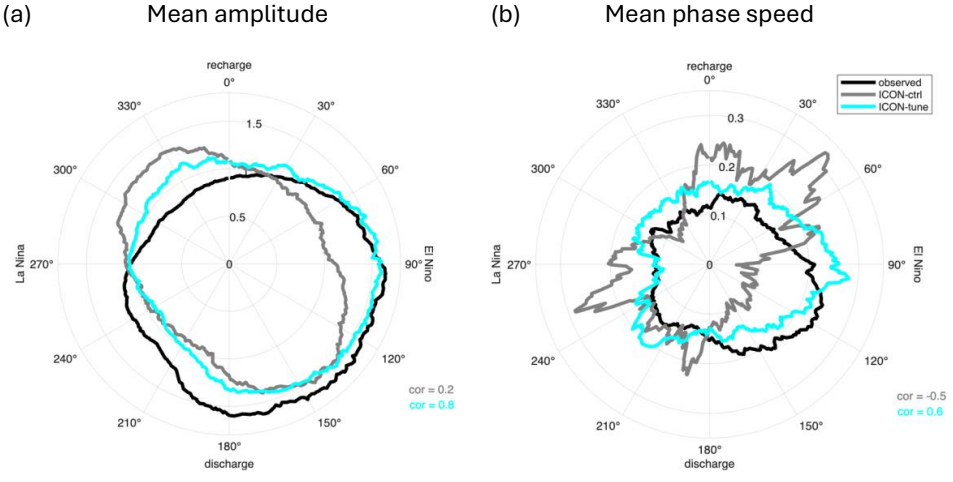



*Figure 11: ENSO phase space diagrams of the mean amplitude (a) and mean phase speed (b). The x-axis represents normalised T values and y-axis normalised h values. Phase of 0° corresponds to the recharge state, 90° to El Niño, 180° to discharge and 270° to La Nina. The correlation values of the ICON-XPP simulations with the observed phase-dependent statistics are shown in both panels with corresponding colors. See methods for details.*

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

The ICON-ctrl simulation does not capture the observed phase space characteristics. The ENSO magnitude is mostly symmetric for opposite phase and show a strong enhancement along the diagonal from 150 ° to 330°, representing a negative in-phase anti-correlation between T and h, which is not observed. This leads to a mean phase speed that is slow near these phases and fast if far away from these phases. The correlation in mean magnitude and phase speed with the observed pattern are weak or even negative (Fig. 11).

The ICON-tune simulation is much closer to the observed with correlation in mean magnitude and phase speed of 0.8 and 0.6, respectively. It can capture the stronger El Niño to discharge phase and weaker La Nina to recharge phases, although the differences are not as strong as observed. It can also capture the observed faster phase transitions during El Niño phases and slower phase transitions during La Nina phases. However, for most phases the ICON-tune phase speed is a bit faster than observed consistent with the power spectral peak being at shorter periods (Fig. 9).

Figure 12 shows the two leading empirical orthogonal function (EOF) modes of the tropical Pacific SST to illustrate the spatial organisation of the large-scale variability. The observed EOF-1 is a dominant east eq. pattern with a sign reversal on the equator at about 160°E and in the off-equatorial western Pacific explaining about 44% of the total variance in the domain. The EOF-2 mode is much weaker but has a clear dipole signature on the equator that, in superposition with the EOF-1 pattern, can represent shifts in eq. position of ENSO SST peaks, an important aspect of ENSO diversity [Takahashi et al. 2011; Dommenget et al. 2013; Capotondi et al. 2015].

The ICON-ctrl simulation has similar leading EOF-modes but shows significant deviations from the observed. The EOF-1 mode is much weaker, with only 14% explained variance and the pattern stretches out too far to the west. This is consistent with the ENSO metric (Fig. 3b), which also shows an eq. SST pattern to be too far to the west. The EOF-2 is also too weak and lacks the observed strong dipole on the equator.

The leading EOF modes of the ICON-tune are much closer to the observed than those of the ICON-ctrl. The EOF-1 mode is more dominant, but not as dominant as observed, and the eq. SST pattern has a sign reversal in the western Pacific at about the same longitude as observed, consistent with the ENSO metric (Fig. 3b). The EOF-2 mode also has a clear eq. dipole structure as observed but has some stronger loading in the southern western Pacific that are much weaker in observations.

ENSO diversity is strongly related to the leading EOF-modes [Takahashi et al. 2011; Dommenget et al. 2013; Capotondi et al. 2015] and is often described by variations of the position of the ENSO events along the eq. Pacific [e.g., Capotondi et al. 2015, P21]. An important aspect in ENSO diversity is the non-linear nature of it, with clear differences in the SST pattern of El Niño and La Nina events [Takahashi et al. 2011; Dommenget et al. 2013]. In Figure 13 we illustrate the non-linear aspect of ENSO diversity by showing the difference in the probability distributions of the position of the ENSO events along the eq. Pacific for El Niño minus La Nina events (see methods).

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

637

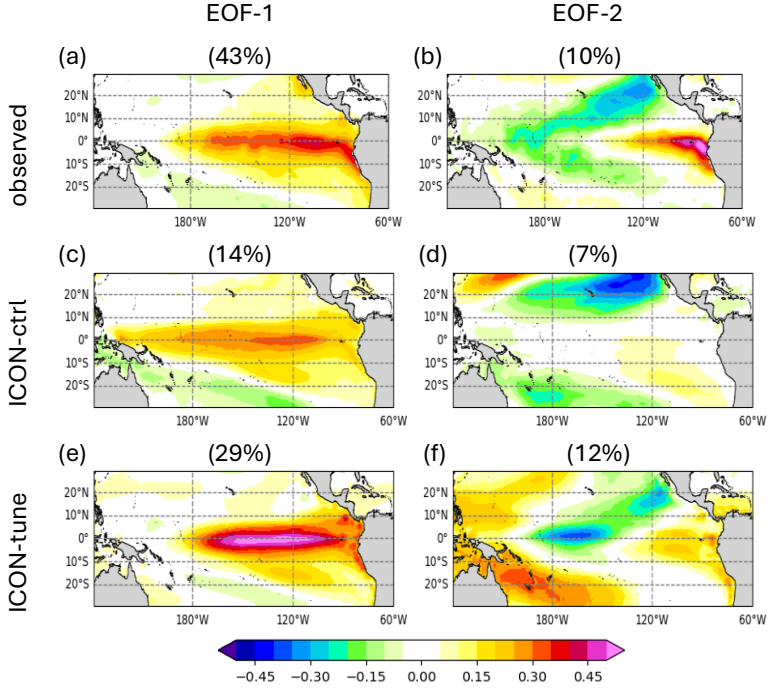



*Figure 12: Leading EOF-modes of the tropical Pacific monthly mean SST anomalies. Explained variance of each mode are shown in the heading.*

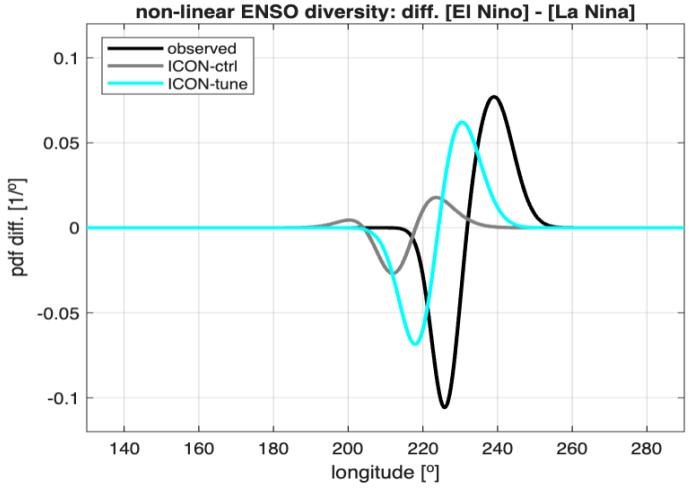



*Figure 13: Non-linear ENSO diversity in the zonal position of the ENSO event peak. Positive values indicate that El Niño events are more likely to peak at this longitude than La Nina events, vice versa for negative values. See methods for details.*

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

The observed ENSO El Niño events tend to peak further to the east than La Nina events by about 15° longitudes. Both ICON-XPP simulations do have the same signature, but it is much weaker in the ICON-ctrl simulation and shifted to the west by about 20° longitudes. The ICON-tune simulation is much closer to the observed but underestimates the non-linear aspects slightly.

We now focus on some key ENSO processes, by analysing the correlation map of some feedback processes of ENSO. Figure 14 show three different elements of ENSO feedbacks: the correlations between NINO3 SST and zonal wind stress, and second the thermocline depth vs. SST. Both are elements of the Bjerknes feedbacks, which is a positive feedback cycle that is key for the strength of ENSO. These two relationships are also measured by the ENSO metrics (Fig. 3f and h). The third process is the relationship between cloud cover and NINO3 SST, which is a negative feedback damping ENSO variability. It is the main process dominating the negative net heat flux feedback (Fig. 3i).

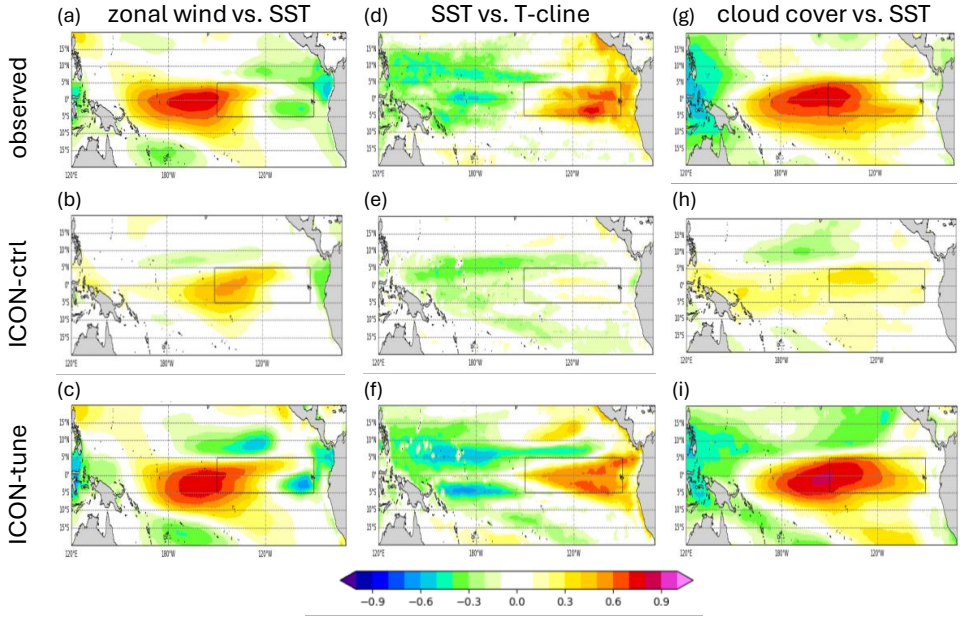



*Figure 14: correlation maps between NINO3 SST vs. zonal wind stress (left column), thermocline depth (middle column) and total cloud cover fraction (right column) for observed (upper row), ICON-ctrl (middle row) and ICON-tune (lower row). The black boxes mark the NINO3 region.*

The observed correlation between zonal wind stress and NINO3 SST peaks in the central eq. Pacific with a strong positive correlation indicating westerly zonal wind stress anomalies correlate with positive NINO3 SST (Fig. 14a). This relationship is much weaker in the ICON-ctrl simulations and is significantly shifted further to the east (Fig. 14b). The ICON-tune simulation is very close to the observed pattern and strength, but the positive correlation pattern of the eq. Pacific does not extend as far to the west as observed (Fig. 14c).

The observed correlation between the thermocline depth and NINO3 SST is marked by positive correlations in the far east eq. Pacific, illustrating that deeper thermocline depth here are associated with warmer SST anomalies, which is a key element of the positive Bjerknes

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

feedback loop (Fig. 14d). In the western half of the Pacific the sign is reversed, completing a dipole structure known for ENSO variability [Meinen and McPhaden 2000].

The ICON-ctrl simulation lacks the positive correlation at the eastern side of the Pacific, indicating it may miss an important element of the positive feedback loop, in which the ocean dynamics drive SST anomalies. This is not apparent in the related ENSO metric (Fig. 3h), which seem to suggest that the strength the thermocline depth versus SST interaction is at the right strength. Here the difference may be due to the different statistical measure, where regression (Fig. 3h) has physical units and correlation (Fig. 14) is normalised, focussing on coherence rather than physical strength.

The ICON-tune simulation is very similar to the observed both in strength and pattern in both the ENSO metrics: regression and in the correlation map. This indicates that this simulation has likely a more ocean dynamics driven ENSO dynamics than the ICON-ctrl simulation.

The observed cloud cover correlation with NINO3 SST is strongly positive in the central eq. Pacific. The ICON-ctrl simulation strongly underestimates this negative feedback, lacking a clear and strong correlation pattern consistent with the weak negative net heat feedback (Fig. 3i). The ICON-tune simulation in turn is quite similar to the observed pattern and even stronger in amplitude than observed.

The last element of the ENSO dynamics that we like to discuss is the relationship between precipitation and NINO3 SST, see Fig. 15. The tropical Pacific precipitation co-variability in ENSO is generally not considered as a process within the ENSO feedback cycle, but is a good indicator of the strength of the atmospheric circulation response, which is important for ENSO feedbacks, and it is also a good indicator for ENSO teleconnections.

The observed precipitation response to NINO3 SST shows a clear shift from the warm pool regions to the central eq. Pacific, with strongly enhance precipitation near the dateline. The ICON-ctrl simulation only moderately captures this signature. The precipitation response here is much weaker and the pattern show strong deviations from the observed, with very weak precipitation response on the eq. Pacific and the enhance precipitation response stretching too far into the warm pool region. This is likely related to the cold tongue bias in ICON-ctrl (Fig. 1a) and the related lack of precipitation in this region and all along the equator, as will be discussed in more detail in Section 3c.

The ICON-tune simulation has a much more realistic precipitation response than the ICON-ctrl, with a clear shift from the warm pool region to the central eq. Pacific and strongly enhance precipitation east of the dateline. The response amplitude is stronger than observed, which is likely related to the too strong precipitation in this simulation, as will be discussed in more detail in Section 3c.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

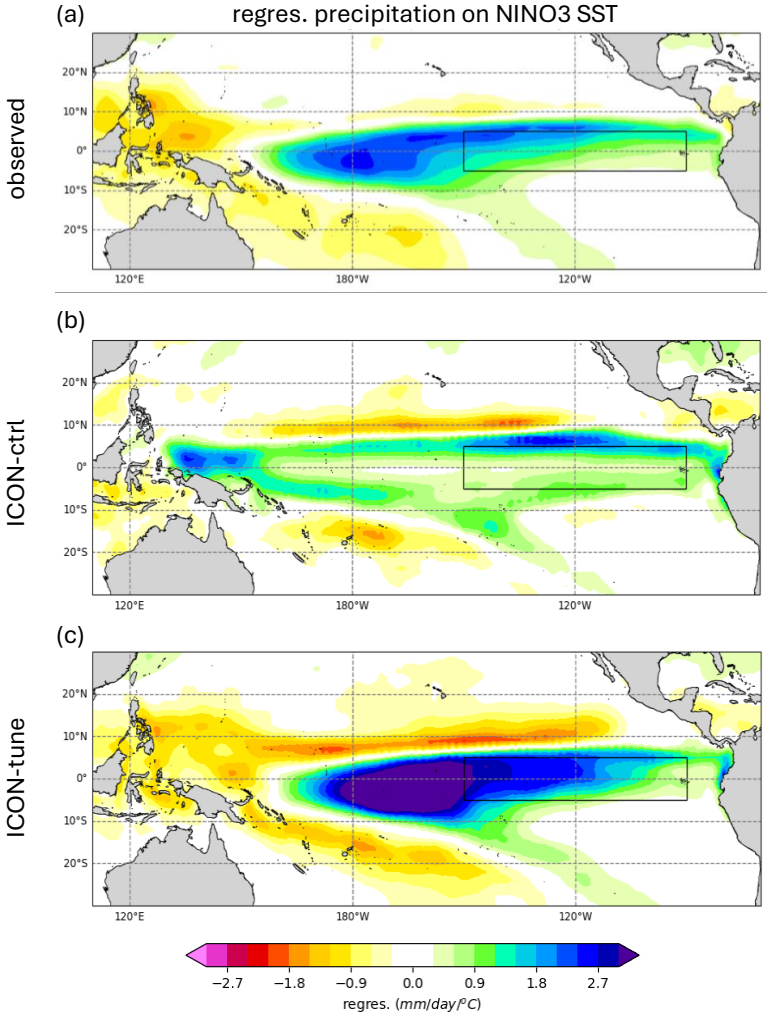



*Figure 15: regression of precipitation on NINO3 SST. The black boxes mark the NINO3 region.*

## c. Global climate performance

While the above discussion clearly illustrates that the ENSO performance has improved in the ICON-XPP model, we also like to have a short look on how the ENSO-focussed model development has affected the global climate performance. Figure 16 shows the mean precipitation in the tropical regions. While the tropical Pacific mean precipitation (150°E to 90°W; 15°S to 15°N) is part of the ENSO metrics used to develop the model, most of the tropical domain is not part of the ENSO metrics.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

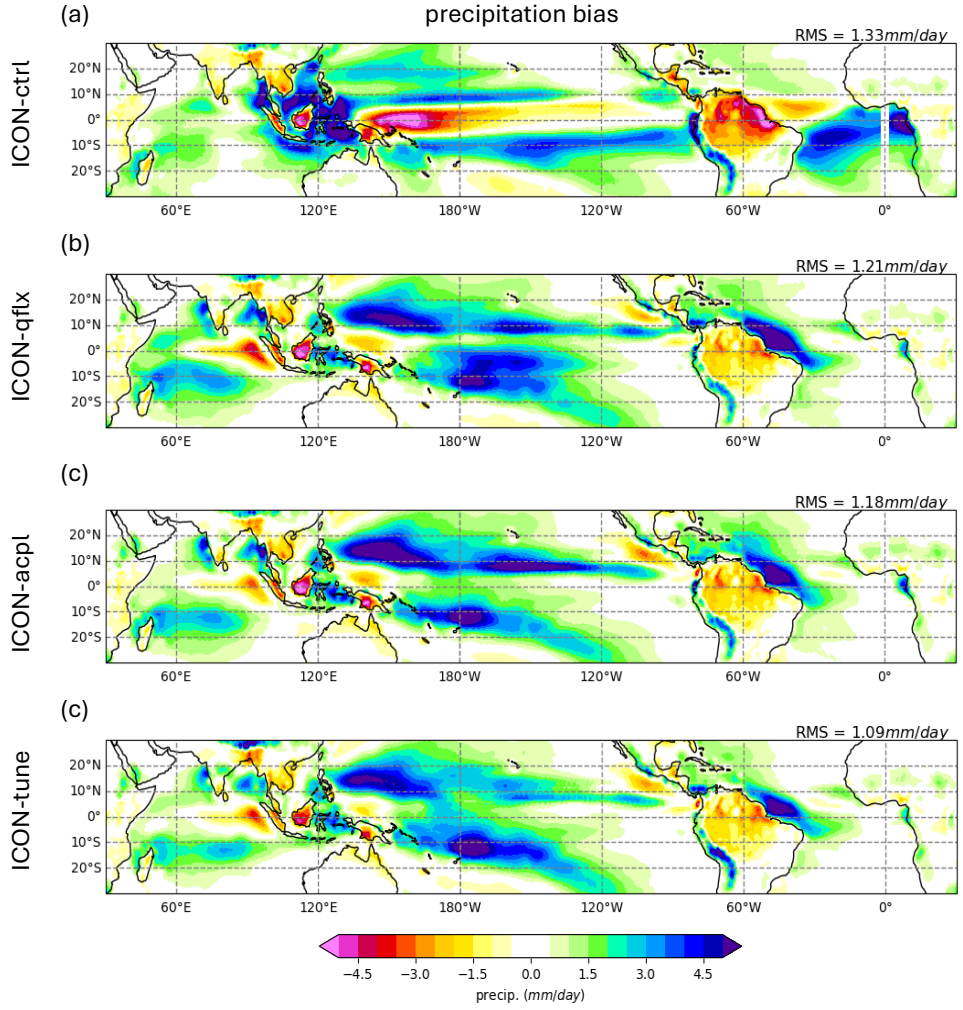



*Figure 16: mean precipitation bias in ICON-XPP simulations. The tropical RMS bias is shown in each panel.*

The ICON-ctrl bias map shows strong precipitation bias with mostly too much precipitation over oceans and too little over land in the Amazona basin and the Islands of the Maritime continent. Some of the biases over oceanic regions appear to be related to SST biases in these regions (Fig. 1a), such as the eq. Pacific cold SST bias reducing the precipitation.
The precipitation bias map in the three ICON-XPP setups with heat flux adjustments are all very similar to each other, but different from the ICON-ctrl. They all have slightly lower precipitation biases in the tropical domain, and also outside the tropical Pacific domain. This suggests that the changes in the mean SST climatology have a substantial impact on the precipitation. The too weak eq. Pacific precipitation (Fig. 3l) has now turned into too much precipitation, and the double ITCZ precipitation bias has essentially disappeared (Fig. 3k). We also see a reduction in precipitation biases over land and the tropical Atlantic. The ICON-tune

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

simulation has the lowest RMS error of all simulations, indicating that the atmospheric parameter tuning has further improved the tropical precipitation.

The left column of Figure 17 shows the global SST bias maps for all four ICON-XPP simulations. The flux corrections were only applied within the tropics (30°S to 30°N), and we can clearly see how the SST biases have been reduced in the three simulations with flux corrections. The SST bias poleward of the tropics are mostly unaffected but do show some small reductions. Removing the cold SST bias in most of the tropics does lead to some warming in the North Pacific and North Atlantic but has no significant impact on the Southern Ocean. It is remarkable how the fine details of the Southern Ocean SST biases are similar in all four simulations, suggesting these fine scale biases are statistically significant.

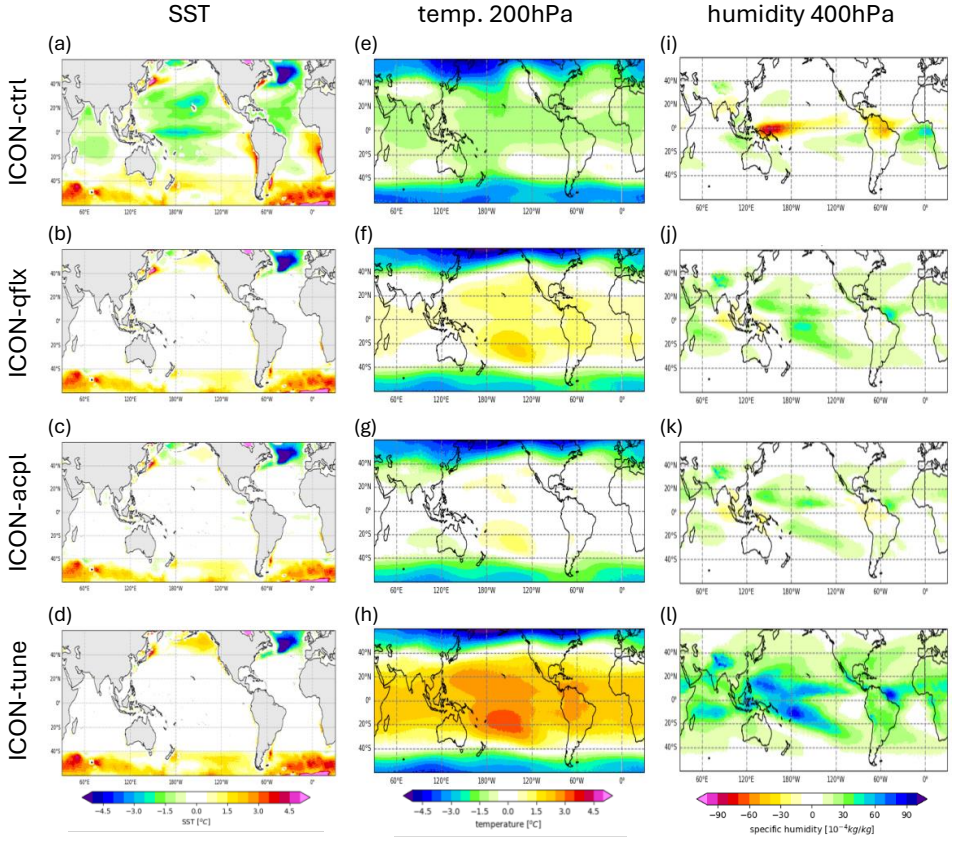



*Figure 17: Biases in SST (left column), atmospheric temperatures at 200hPa (middle column) and humidity at 400hPa (right column).*

We can see some more distinct differences between the model simulations at higher level tropospheric temperatures and mid-troposphere humidity, see Fig. 17 middle and right columns. The ICON-ctrl has a moderate cool bias in most regions for the 200hPa temperature and a clear dry bias in the 400hPa specific humidity in some tropical regions that also have dry biases in precipitation.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

The flux corrected runs ICON-qflx and ICON-acpl are similar to each other, but different compared to the ICON-ctrl. The correction of the tropical SST cold biases seems to have also reduced or even reversed the 200hPa temperature cold bias of the ICON-ctrl. The 400hPa specific humidity is also very different in pattern and sign compared to the ICON-ctrl. Overall, the biases at higher levels are generally slightly smaller in the ICON-qflx and ICON-acpl compared to the ICON-ctrl, suggesting that the flux correction of SST and wind stress do have a small but positive effect.

The ICON-tune is substantially different in the higher-level biases from the other simulations. Here, we see a clear increase in biases at 200hPa temperature and 400hPa specific humidity. Most of the tropics and midlatitudes see a clear warm bias and too high humidity that is significantly larger than in the ICON-ctrl. This suggests that the tuning of atmospheric parameters to improve ENSO does have degrading effects at higher level atmospheric climate.

# 4. Summary and discussion

In this study we applied three different kinds of artificial corrections to the ICON-XPP model to improve ENSO simulations. First, we introduce a net heat flux adjustment to the tropical oceans to correct the simulated mean tropical SST. Second, we corrected the mean wind stress forcing to the tropical oceans to match the observed mean wind stress forcing. Finally, we optimised 6 atmospheric model parameters to minimise biases in ENSO metrics in the ICON-XPP simulations in combination with the heat and wind stress flux corrections.

The combined effect of all three methods has led to substantial improvements of the ENSO simulation in nearly all aspects of ENSO variability and dynamics, transforming the uncorrected ICON-ctrl simulations from slightly weaker than the mean CMIP5/6 model performance to the ICON-tune simulation that is as good or better than the best CMIP6 model. The ICON-tune simulation is marked by substantially improved large-scale SST pattern, seasonality, non-linear behaviour, Bjerknes and cloud feedbacks relative to the ICON-ctrl simulation.

The largest improvements come from the net heat flux adjustment to correct the mean tropical SST biases in the ICON-qflx simulation. The strongly reduced mean SST biases resulted into more realistic large-scale SST pattern, seasonality, non-linear behaviour, Bjerknes and cloud feedbacks relative to the ICON-ctrl simulation. The addition of the wind stress corrections in the ICON-acpl simulation further improved the Bjerknes feedbacks and therefore improve the overall amplitude of ENSO variability relative to the ICON-qflx simulation. The final addition of tuning atmospheric model parameters in the ICON-tune lead to some smaller improvements in the cloud feedbacks, SST amplitude, pattern and seasonality.

The ICON-qflx improvements relative to the ICON-ctrl are very much in line with what Y26 has found in the AMIP simulation of the same ICON-XPP model. That is, the ICON-qflx ENSO metrics are much more similar to the AMIP simulation of Y26 than the ICON-ctrl, indicating that correcting the mean SST leads to a coupled model that behaves more like the AMIP simulation. Given that the AMIP ICON-XPP simulation has much smaller ENSO metric biases than ICON-ctrl, the success of the net heat flux adjustment results from relatively good performance of the atmospheric model, suggesting that large parts of the ENSO metric biases in the ICON-XPP model result from the SST biases (the coupling to the ocean).

While all three elements of the model development improved the ENSO performance it should be noted here that the flux adjustment approach is more effective and require less

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

computing resources than the tuning approach. The computing resources of climate models are an important consideration for further model development [Acosta et al. 2024].

Computing the flux corrections of both net heat and wind stress is done within an iterative approach that requires to run the model for about 50-70yrs in total. The tuning methods, in turn, requires computing sensitivities and their uncertainties for all tuned model parameters in combination with the iterative approach to compute the flux corrections. This is required to run the model for more than 1000yrs in total. The finding that the flux correction is more effective and cheaper than the tuning of parameter largely agrees with the outcomes of Dommenget and Rezny [2018].

The artificial corrections have led to some improvement in simulating the tropical mean precipitation, which is also improving the tropical precipitation response to ENSO variability. Outside the tropical domain, the flux corrections had only minor impacts but lead to some small reduction in extra-tropical SST biases, and slightly improved tropospheric mean climate. However, the parameter tuning did degrade tropospheric mean climate by increasing biases in upper tropospheric temperature and mid-tropospheric humidity. This negative side effect of tuning a cost function that does not consider large parts of the climate system is consistent with the finding of Y26 and Dommenget and Rezny [2018].

The results of this study open some discussion points: Should we use artificially corrected climate models? If so, are there some limits to it? What do we learn from this for further model developments?

We should note here that all published climate models have some artificially corrections included, mostly related to some parameter tuning, while flux corrections are usually not applied. The clear improvement of the model simulations using flux corrections here strongly suggest we should use these flux corrected models for climate studies. While it is unclear from this study where the limits are, we would suggest applying these models for studying natural climate variability on time scales of seasons to decades.

We further would suggest testing these models in climate change scenarios with external forcing. While we have no evidence that this model would perform better than the alternative, we also have no evidence against it. However, some studies suggest that some mean state biases do affect climate trends [e.g., Dommenget 2012; Samanta et al. 2019; Wills and Dinezio 2025; Zhuo et al. 2025]. This would indicate that the flux corrected models with much improved mean state could perform better in simulating the climate change response.

Alternatively, even if we do not want to use artificially corrected climate models to study the climate system they can clearly help to understand model biases and further improve the models. The flux corrected simulations, for instance, clearly point out how important aspects of the SST mean state are and they provide seasonally changing net heat flux maps that essentially indicate where problems in the climate system exist.

Further, we found clear changes in the tropical precipitation bias that suggest a clear pattern of too much precipitation over oceans and too little over land. This clear bias map can help to find the processes and parameterisations that need to be improved in the model to improve the precipitation biases [e.g., Kroll et al. 2025].

The significant sensitivities of the ENSO metrics to the tuning parameters give an indication on what processes are important to control ENSO and the tropical mean state. This can guide further model development. Tuning of model parameters is a widely used and accepted approach to improve model simulations. While the study here found only limited success and some serious problems, this does not mean that such methods cannot be successful. Considering the limited resource used in this study, we do suggest that given much larger

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

resources model tuning of low-resolution climate models can substantially improve the performance of the models. This is also likely to benefit higher-resolution climate models that often have similar model biases [e.g., Moreno-Chamarro et al. 2022; Liu et al. 2022; Respati et al. 2024; Müller et al. 2025; Kroll et al. 2025].

## Code and data availability

All scripts, processed data required to reproduce the results of this study are archived at Zenodo [Dommenget et al. 2026]: https://doi.org/10.5281/zenodo.22657661

## Author contribution

DD led the study, conducted all experiments and analysis, and wrote the first draft. HH introduced the flux adjustment method in the ICON-XPP model. All authors contributed to the final manuscript.

## Competing interests

There are no competing interests for any of the authors.

## Acknowledgements

This study was funded by the European Union's Horizon Europe programme through ASPECT (grant no.~101081460), the German Federal Ministry of Education and Research (BMBF) through the Coming Decade project (grant no.~01LP2327E) and the Australian Research Council (ARC) Centre of Excellence for Climate Extremes (grant no. CE170100023). We acknowledge the German Climate Computing Center (DKRZ) for providing computational resources. We like to thank Jürgen Kröger, Clarissa Kroll, Yann Planton, Pothapakula Praveen, Kira Rehfeld and Pham van Trang for helpful comments, discussion and feedback.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

# References

Acosta, M. C., and Coauthors, 2024: The computational and energy cost of simulation and storage for climate science: lessons from CMIP6. *Geosci. Model Dev.*, **17**, 3081–3098, https://doi.org/10.5194/gmd-17-3081-2024.

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *J. Hydrometeorol.*, **4**, 1147–1167, https://doi.org/https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor, 2012: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophys. Res. Lett.*, **39**, 1–7, https://doi.org/10.1029/2012GL051607.

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, https://doi.org/10.1007/s00382-017-3799-2.

——, C. Wengel, M. Latif, D. Dommenget, J. Lübbecke, and W. Park, 2019: Error compensation of ENSO atmospheric feedbacks in climate models and its influence on simulated ENSO dynamics. *Clim. Dyn.*, **53**, 155–172, https://doi.org/10.1007/s00382-018-4575-7.

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, https://doi.org/10.1007/s00382-013-1783-z.

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

Byrne, H., R. Seager, and J. E. Smerdon, 2025: CMIP6 models cannot capture long-term forced changes in the tropical Pacific sea surface temperature gradient. *Nat. Commun.*, **17**, 142, https://doi.org/10.1038/s41467-025-66839-w.

Capotondi, A., and Coauthors, 2015: Understanding enso diversity. *Bull. Am. Meteorol. Soc.*, **96**, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1.

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, https://doi.org/10.1175/2007MWR1978.1.

Collins, M., B. B. B. Booth, G. R. Harris, J. M. Murphy, D. M. H. Sexton, and M. J. Webb, 2006: Towards quantifying uncertainty in transient climate change. *Clim. Dyn.*, **27**, 127–147, https://doi.org/10.1007/s00382-006-0121-0.

Dommenget, D., 2012: Analysis of the Model Climate Sensitivity Spread Forced by Mean Sea Surface Temperature Biases. *J. Clim.*, **25**, 7147–7162, https://doi.org/10.1175/jcli-d-11-00600.1.

——, and M. Rezny, 2018: A Caveat Note on Tuning in the Development of Coupled Climate Models. *J. Adv. Model. Earth Syst.*, **10**, 78–97, https://doi.org/10.1002/2017MS000947.

——, T. Bayr, and C. Frauen, 2013: Analysis of the non-linearity in the pattern and time evolution of El Niño southern oscillation. *Clim. Dyn.*, **40**, 2825–2847, https://doi.org/10.1007/s00382-012-1475-0.

——, P. Priya, and A. Vijayeta, 2023: ENSO phase space dynamics with an improved estimate of the thermocline depth. *Clim. Dyn.*, **61**, 5767–5783, https://doi.org/10.1007/s00382-023-06883-8.

——, H. Haak, D. Yu, and W. A. Müller, 2026: Open Source data and scripts for publication

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



"Improving the simulation of El Niño Southern Oscillation within the ICON model by flux adjustments and parameter tuning" [Dataset].

Flato, G., and Coauthors, 2013: Evaluation of climate models [Evaluación de los modelos climáticos]. *Clim. Chang. 2013 Phys. Sci. Basis Work. Gr. I Contrib. to Fifth Assess. Rep. Intergov. Panel Clim. Chang.*, **9781107057**, 741–866.

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

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.*, **146**, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803.

Hourdin, F., and Coauthors, 2017: The Art and Science of Climate Model Tuning. *Bull. Am. Meteorol. Soc.*, **98**, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1.

Irvine, P. J., L. J. Gregoire, D. J. Lunt, and P. J. Valdes, 2013: An efficient method to generate a perturbed parameter ensemble of a fully coupled AOGCM without flux-adjustment. *Geosci. Model Dev.*, **6**, 1447–1462, https://doi.org/10.5194/gmd-6-1447-2013.

Kessler, W. S., 2002: Is ENSO a cycle or a series of events? *Geophys. Res. Lett.*, **29**, 40-1-40–44, https://doi.org/10.1029/2002GL015924.

Kröger, J., and F. Kucharski, 2011: Sensitivity of ENSO characteristics to a new interactive flux correction scheme in a coupled GCM. *Clim. Dyn.*, **36**, 119–137, https://doi.org/10.1007/s00382-010-0759-5.

Kroll, C. A., R. C. Jnglin Wills, L. Kornblueh, U. Niemeier, and A. Schneidereit, 2025: Parameterization adaption needed to unlock the benefits of increased resolution for the ITCZ in ICON. *Atmos. Chem. Phys.*, **25**, 16915–16943, https://doi.org/10.5194/acp-25-16915-2025.

Kumar, B. P., J. Vialard, M. Lengaigne, V. S. N. Murty, and M. J. McPhaden, 2012: TropFlux: Air-sea fluxes for the global tropical oceans-description and evaluation. *Clim. Dyn.*, **38**, 1521–1543, https://doi.org/10.1007/s00382-011-1115-0.

Li, G., G. Zhou, J. Jin, R. Guo, and Y. Zhang, 2026: Improving ENSO hindcast skill by adjusting the mean background state in a coupled general circulation model (CAS-ESM2.0). *Ocean Model.*, **204**, 102812, https://doi.org/https://doi.org/10.1016/j.ocemod.2026.102812.

Liu, B., and Coauthors, 2022: Will Increasing Climate Model Resolution Be Beneficial for ENSO Simulation? *Geophys. Res. Lett.*, **49**, 1–11, https://doi.org/10.1029/2021GL096932.

Luo, J. J., G. Wang, and D. Dommenget, 2018: May common model biases reduce CMIP5's ability to simulate the recent Pacific La Niña-like cooling? *Clim. Dyn.*, **50**, 1335–1351, https://doi.org/10.1007/s00382-017-3688-8.

Ma, L., and Z. Jiang, 2021: Sensitivity of ENSO Simulation to the Convection Schemes in the NESM3 Climate System Model: Atmospheric Processes. *Front. Earth Sci.*, **Volume 9**-, https://doi.org/10.3389/feart.2021.596442.

Magnusson, L., M. Alonso-Balmaseda, S. Corti, F. Molteni, and T. Stockdale, 2013a: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. *Clim. Dyn.*, **41**, 2393–2409, https://doi.org/10.1007/s00382-012-1599-2.

Magnusson, L., M. Alonso-Balmaseda, and F. Molteni, 2013b: On the dependence of ENSO simulation on the coupled model mean state. *Clim. Dyn.*, **41**, 1509–1525,

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

https://doi.org/10.1007/s00382-012-1574-y.

Manabe, S., and R. J. Stouffer, 1988: Two Stable Equilibria of a Coupled Ocean-Atmosphere Model. *J. Clim.*, **1**, 841–866, https://doi.org/10.1175/1520-0442(1993)006<0175:COSEOA>2.0.CO;2.

Mauritsen, T., and Coauthors, 2012: Tuning the climate of a global model. *J. Adv. Model. Earth Syst.*, **4**, 1–18, https://doi.org/10.1029/2012MS000154.

McPhaden, M. J., S. E. Zebiak, and M. H. Glantz, 2006: ENSO as an Integrating Concept in Earth Science. *Science (80-. ).*, **314**, 1740–1745, https://doi.org/10.1126/science.1132588.

Meinen, C. S., and M. J. McPhaden, 2000: Observations of warm water volume changes in the equatorial Pacific and their relationship to El Nino and La Nina. *J. Clim.*, **13**, 3551–3559, https://doi.org/10.1175/1520-0442(2000)013<3551:OOWWVC>2.0.CO;2.

Moreno-Chamarro, E., and Coauthors, 2022: Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models. *Geosci. Model Dev.*, **15**, 269–289, https://doi.org/10.5194/gmd-15-269-2022.

Müller, W. A., and Coauthors, 2025: The ICON-based Earth System Model for climate predictions and projections (ICON XPP v1.0). *Geosci. Model Dev.*, **18**, 9385–9415, https://doi.org/10.5194/gmd-18-9385-2025.

Neelin, J. D., and H. A. Dijkstra, 1995: Ocean-atmosphere interaction and the tropical climatology. Part I: the dangers of flux correction. *J. Clim.*, **8**, 1325–1342, https://doi.org/10.1175/1520-0442(1995)008<1325:OAIATT>2.0.CO;2.

Nelder, J. A., and R. Mead, 1965: A Simplex Method for Function Minimization. *Comput. J.*, **7**, 308–313, https://doi.org/10.1093/comjnl/7.4.308.

Ohba, M., D. Nohara, and H. Ueda, 2010: Simulation of asymmetric ENSO transition in WCRP CMIP3 Multimodel Experiments. *J. Clim.*, **23**, 6051–6067, https://doi.org/10.1175/2010JCLI3608.1.

Okumura, Y. M., and C. Deser, 2010: Asymmetry in the duration of El Niño and la Niña. *J. Clim.*, **23**, 5826–5843, https://doi.org/10.1175/2010JCLI3592.1.

Planton, Y. Y., and Coauthors, 2021: Evaluating climate models with the CLIVAR 2020 ENSO metrics package. *Bull. Am. Meteorol. Soc.*, **102**, E193–E217, https://doi.org/10.1175/BAMS-D-19-0337.1.

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. Atmos.*, **108**, https://doi.org/10.1029/2002jd002670.

Respati, M. R., D. Dommenget, H. Segura, and C. Stassen, 2024: Diagnosing drivers of tropical precipitation biases in coupled climate model simulations. *Clim. Dyn.*, **62**, 8691–8709, https://doi.org/10.1007/s00382-024-07355-3.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Clim.*, **19**, 3483–3517, https://doi.org/https://doi.org/10.1175/JCLI3812.1.

Samanta, D., K. B. Karnauskas, and N. F. Goodkin, 2019: Tropical Pacific SST and ITCZ Biases in Climate Models: Double Trouble for Future Rainfall Projections? *Geophys. Res. Lett.*, **46**, 2242–2252, https://doi.org/10.1029/2018GL081363.

Sausen, R., K. Barthel, and K. Hasselmann, 1988: Coupled ocean-atmosphere models with flux correction. *Clim. Dyn.*, **2**, 145–163, https://doi.org/10.1007/BF01053472.

Schneider, E. K., 1996: Flux correction and the simulation of changing climate. *Ann. Geophys.*, **14**, 336–341, https://doi.org/10.1007/s00585-996-0336-8.

https://doi.org/10.5194/egusphere-2026-5299
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



Spencer, H., R. Sutton, and J. M. Slingo, 2007: El Niño in a coupled climate model: Sensitivity to changes in mean state induced by heat flux and wind stress corrections. *J. Clim.*, **20**, 2273–2298, https://doi.org/10.1175/JCLI4111.1.

Takahashi, K., A. Montecinos, K. Goubanova, and B. Dewitte, 2011: ENSO regimes: Reinterpreting the canonical and Modoki El Nino. *Geophys. Res. Lett.*, **38**, 1–5, https://doi.org/10.1029/2011GL047364.

Tian, B., and X. Dong, 2020: The Double-ITCZ Bias in CMIP3, CMIP5, and CMIP6 Models Based on Annual Mean Precipitation. *Geophys. Res. Lett.*, **47**, e2020GL087232, https://doi.org/https://doi.org/10.1029/2020GL087232.

——, E. J. Fetzer, and J. Teixeira, 2024: Assessing the Tropospheric Temperature and Humidity Simulations in CMIP3/5/6 Models Using the AIRS Obs4MIPs V2.1 Data. *J. Geophys. Res. Atmos.*, **129**, e2023JD040536, https://doi.org/https://doi.org/10.1029/2023JD040536.

Vecchi, G. A., and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. *J. Clim.*, **27**, 7994–8016, https://doi.org/10.1175/JCLI-D-14-00158.1.

Vijayeta, A., and D. Dommenget, 2018: An evaluation of ENSO dynamics in CMIP simulations in the framework of the recharge oscillator model. *Clim. Dyn.*, **51**, 1753–1771, https://doi.org/10.1007/s00382-017-3981-6.

Watanabe, M., M. Chikira, Y. Imada, and M. Kimoto, 2011: Convective Control of ENSO Simulated in MIROC. *J. Clim.*, **24**, 543–562, https://doi.org/https://doi.org/10.1175/2010JCLI3878.1.

Wills, R. J., and P. Dinezio, 2025: Emergent climate projection discrepancies stem from spurious rainband. *ReserachSquare*, **reprint**, https://doi.org/10.21203/rs.3.rs-7860860/v1.

Xu, W., X.-T. Zheng, and K. Hu, 2025: Revisiting the modelled cloud radiative feedback on ENSO — the source of model uncertainty. *Clim. Dyn.*, **63**, 237, https://doi.org/10.1007/s00382-025-07728-2.

Yu, D., D. Dommenget, H. Pohlmann, and W. A. Müller, 2026: A systematic atmospheric parameter optimization method to improve ENSO simulation in the ICON XPP Earth system model. *Geosci. Model Dev.*, **19**, 5531–5552, https://doi.org/10.5194/gmd-19-5531-2026.

Zhuo, J. Y., C. Y. Lee, A. Sobel, R. Seager, S. J. Camargo, Y. H. Lin, B. Fosu, and K. A. Reed, 2025: A More La Niña–Like Response to Radiative Forcing after Flux Adjustment in CESM2. *J. Clim.*, **38**, 1037–1050, https://doi.org/10.1175/JCLI-D-24-0331.1.