



1 **A Systematic Atmospheric Parameter Optimization method to Improve**
2 **ENSO Simulation in the ICON XPP Earth System Model**

3

4 **Dakuan Yu¹, Dietmar Dommenges², Holger Pohlmann¹, and Wolfgang A. Müller¹**

5

6 ¹Max Planck Institute for Meteorology, Hamburg, Germany

7 ² ARC Centre of Excellence for Climate Extremes, School of Earth, Atmosphere and Environment,
8 Monash University, Clayton, Victoria, Australia

9

10 Corresponding author: Dakuan Yu (dakuan.yu@mpimet.mpg.de)

11



12 Abstract

13 The El Niño–Southern Oscillation (ENSO) is a dominant mode of interannual climate
14 variability, yet accurately simulating ENSO in climate models remains a major challenge due to
15 its complex coupled dynamics. In this study, we present a novel linear optimization methodology
16 and systematically adjust atmospheric parameters to improve ENSO fidelity in the Icosahedral
17 Nonhydrostatic eXtended Predictions and Projections (ICON XPP) Earth System Model of the
18 Max-Planck-Institute for Meteorology. The optimization approach is based on the superposition
19 of parameter sensitivities and a Nelder–Mead algorithm that reduces the ENSO cost function.
20 The cost function account for ENSO-related tropical climatology, variability, and feedbacks, which
21 are estimated with the ENSO metric package. We firstly assess the sensitivity of ENSO metrics to
22 21 atmospheric parameters in atmosphere-only simulations. The optimization approach reduces
23 the ENSO cost function by 30% in the optimized atmosphere-only runs. Key improvements
24 include reduced precipitation bias and strengthened atmospheric feedbacks such as the Bjerknes
25 and thermal damping feedbacks. These results demonstrate the effectiveness of our method in
26 improving ENSO metrics within the atmosphere-only configuration. Six parameters identified as
27 most impactful from atmosphere-only tuning experiments are subsequently tuned in fully
28 coupled simulations. The optimized fully coupled run yields moderate improvements in ENSO
29 amplitude, cold tongue SST bias, seasonal phase-locking, ocean-atmosphere coupling and
30 teleconnection patterns. However, isolated ENSO tuning introduces unrealistic global warming,
31 which is further corrected by adjusting turbulence-related parameters without degrading ENSO
32 skill. These results demonstrate that systematic ENSO tuning can yield performance gains but
33 must be balanced with broader climate stability constraints. Our method offers a scalable,
34 physically grounded optimization strategy, with strong potential for tuning ENSO in climate
35 model configurations.

36

37

38

39

40

41

42

43



44 1 Introduction

45 The El Niño-Southern Oscillation (ENSO) is the most significant mode of interannual climate
46 variability, exerting profound influences on the global atmospheric circulation, temperature
47 patterns, and extreme weather events (McPhaden et al., 2006; Timmermann et al., 2018). ENSO-
48 driven anomalies impact monsoon systems (Kumar et al., 2006), alter precipitation patterns
49 (Trenberth et al., 1998), and modulate the frequency of droughts, floods, and hurricanes
50 worldwide (Cai et al., 2020). Given its far-reaching consequences, accurately simulating ENSO
51 within climate models is essential for reliable climate predictions (Collins et al., 2010; Ham et al.,
52 2019). However, despite decades of model development, challenges remain in fully capturing
53 ENSO phenomenon in climate models, particularly in reproducing its feedback processes and
54 teleconnections (Bellenger et al., 2014; Planton et al., 2021, 2024). Here, we provide a novel
55 methodology that optimizes ENSO in the newly developed Icosahedral Nonhydrostatic eXtended
56 Predictions and Projections (ICON XPP) Earth System Model (Müller et al., 2025a, b), by direct
57 assessment of an ideal atmospheric parameter setting. The presented optimization approach
58 serves as a blueprint for model developers to explicitly account for ENSO simulation in climate
59 models.

60 Climate models have historically suffered with systematic biases in ENSO simulations,
61 including the well-documented cold tongue bias, which produces excessively cold sea surface
62 temperatures (SSTs) in the equatorial Pacific and affects ENSO amplitude (Li & Xie, 2014). Many
63 models also exhibit an excessive westward extension of ENSO anomalies, leading to unrealistic
64 spatial distributions of SST variability and misrepresentations of atmospheric convection patterns
65 (Jiang et al., 2021). Additionally, atmospheric feedbacks, such as the Bjerknes feedback (the
66 coupling between zonal wind stress and SST in tropical Pacific), tend to be too weak, limiting the
67 amplification of ENSO events (Lloyd et al., 2009). ENSO frequency and amplitude biases are also
68 persistent issues, with some models producing ENSO events that are too strong, too weak, or
69 overly periodic, failing to capture the observed irregularity of ENSO cycles (Guilyardi, et al.,
70 2009b; Bellenger et al., 2014). Although different generations of climate models, particularly
71 those participating in the Coupled Model Intercomparison Projects (CMIP), have made progress
72 in reducing these biases, significant uncertainties still remain, especially from CMIP5 to CMIP6
73 (Planton et al., 2021).

74 For example, The Community Earth System Model (CESM) has undergone significant
75 revisions across different versions, yet challenges persist. CESM1 exhibits excessive ENSO
76 variability and an unrealistic persistence of El Niño events (Zhang et al., 2017). CESM2 introduces
77 improvements in cloud physics and oceanic processes, leading to a more realistic ENSO
78 amplitude, yet biases remain in tropical Pacific convection and wind stress feedbacks
79 (Danabasoglu et al., 2020). Similarly, the Max-Planck-Institute Earth System Model (MPI-ESM)
80 sees improvements over successive versions. While MPI-ESM1 in CMIP5 displayed a strong cold
81 tongue bias and weaker ENSO variability (Jungclaus et al., 2013), its successor, MPI-ESM1.2 in
82 CMIP6, improved ENSO frequency and amplitude by refining ocean-atmosphere coupling and
83 tropical convection schemes (Mauritsen et al., 2019). However, the model still struggles to
84 accurately represent ENSO phase-locking, asymmetry, and ENSO-related teleconnections and



85 feedbacks (Müller et al., 2018; Bayr et al., 2019). Other major climate models also exhibit notable
86 ENSO simulation biases (Bellenger et al., 2014; Planton et al., 2021), such as the duration of El
87 Niño events bias in the HadGEM3 model from the UK Met Office (Kuhlbrodt et al., 2018; Williams
88 et al., 2018) and the ENSO-related precipitation bias in GFDL-CM4 model from NOAA's
89 Geophysical Fluid Dynamics Laboratory (Held et al., 2019). These examples highlight the ongoing
90 challenges in ENSO simulation across leading climate models, despite continued advancements
91 in model physics and resolution.

92 Moreover, although some climate models successfully simulate ENSO amplitude, this may
93 result from error compensation—where the weak Bjerknes positive feedback and thermal
94 damping negative feedback counteract each other, leading to seemingly accurate ENSO
95 variability (Lloyd et al., 2009; Guilyardi, et al., 2009a; Bayr et al., 2019; Planton et al., 2021). This
96 highlights a critical issue in climate modeling: a well-simulated ENSO amplitude does not
97 necessarily indicate a well-represented ENSO feedback process. As a result, direct tuning of
98 ENSO-related processes remains necessary in climate models to ensure a physically consistent
99 representation of ENSO variability.

100 Given these persistent biases and the limited success of ENSO tuning strategies, there is a
101 clear need for a systematic and targeted method to improve the ENSO simulation in climate
102 models. Here, we use the newly developed configuration ICON XPP which serves as a baseline
103 for the next generation climate predictions (Müller et al., 2025a, b). We focus on 21 atmospheric
104 parameters related to cloud physics, microphysics, and turbulence schemes, which are known to
105 influence ENSO dynamics. The tuning is guided by the ENSO Metrics Package (Planton et al.,
106 2021), which provides a comprehensive evaluation of ENSO simulation in climate model across
107 four key dimensions: tropical climatology, variability, feedback processes, and teleconnection
108 patterns. To efficiently explore the parameter space, we employ a Nelder–Mead optimization
109 algorithm that leverages the linear superposition of parameter sensitivities (Luersen & Riche,
110 2004). The optimization is first conducted in an atmosphere-only configuration to isolate the
111 atmospheric contribution to ENSO biases, and then extended to fully coupled simulations to
112 assess robustness and ocean–atmosphere interaction effects.

113 This study aims to: (1) Quantify the sensitivity of ENSO metrics to individual atmospheric
114 parameters; (2) Optimize the parameter set using a linear superposition approach; (3) Evaluate
115 the effectiveness of the method in both atmosphere-only and fully coupled configurations,
116 including high-resolution experiments and adjustments for global climate stability.

117 The structure of this paper is as follows: Section 2 describes the observations and the ICON
118 XPP model, experimental setup and simulations. The optimization method and the elements
119 needed for it are described in section 3. Section 4 presents result from atmosphere-only ICON
120 XPP model, including parameter sensitivity and optimization outcomes. Section 5 discusses the
121 application of the method to the fully coupled ICON XPP model, and evaluates improvements in
122 ENSO performance and teleconnections, and explores high-resolution and global temperature
123 results. Section 6 provides a summary and outlook for future work.

124



125 **2 Data and Model**

126 **2.1 Observation data**

127 The observational data for the ENSO metrics are chosen to be the same as in Planton et al.
128 (2021). The monthly precipitation data are from GPCPv2.3 dataset, which combines observations
129 and satellite data (Adler et al., 2003). The SST, zonal wind stress, and surface net heat flux data
130 are from the TropFlux dataset, which is mainly derived from a combination of ERA-Interim and
131 ISCCP corrected fields using Global Tropical Moored Buoy Array data (Kumar et al., 2012). The
132 Sea Surface Height (SSH) data are from the GODAS dataset, which is an ocean reanalysis dataset
133 forced by the momentum flux, heat flux and fresh water flux from NCEP2 (Saha et al., 2006). All
134 the reference data range from 1980 to 2018. The model data and observational data are regarded
135 to the same 1°x1° global grids. The Niño3, Niño3.4, and Niño4 regions are defined in 210°E-270°E,
136 5°S-5°N; 190°E-240°E, 5°S-5°N; and 160°E-210°E, 5°S-5°N, respectively.

137

138 **2.2 ICON XPP Earth System Model**

139 The ICON XPP Earth System Model is mainly developed by the Max Planck Institute for
140 Meteorology (MPI-M) and the German Weather Service (DWD). ICON XPP integrates components
141 from numerical weather prediction and earth system modeling into a unified framework capable
142 of addressing both scientific and operational forecasting challenges (Müller et al., 2025 (a)). It
143 builds upon core elements of ICON, including the atmospheric component (ICON-NWP) (Zängl et
144 al., 2015), the ICON ocean model (ICON-O) (Korn et al., 2022), and the land component (ICON-L)
145 (Reick et al., 2021). The ocean and atmosphere are coupled through the YAC coupler (Yet Another
146 Coupler, Hanke et al., 2016). ICON XPP will serve as the next-generation earth system model from
147 MPI-M and deploy for CMIP7-class simulations. More information about the ICON XPP Earth
148 System Model can be found in Müller et al., 2025a, b. In this study, ICON XPP is employed in both
149 atmosphere-only and fully coupled configurations. The resolutions for atmosphere-only and fully
150 coupled experiments are 160 km atmosphere and 160 km atmosphere / 40 km ocean,
151 respectively, which provides a balance between computational tractability and the ability to
152 resolve ENSO-related dynamics. This work presents one of the first systematic assessments and
153 optimization efforts targeting ENSO performance in ICON XPP.

154

155 **2.3 Parameters perturbation experiments**

156 The parameters perturbation experiments are initially conducted using an atmosphere-only
157 simulations with twenty-one different parameters, wherein each parameter is tested across ten
158 different values. The atmosphere-only simulations are performed at a spatial resolution of 160
159 km and cover the historical period from 1979 to 1997, totaling 18 years per simulation. Following
160 the atmosphere-only tuning experiments, six parameters that exhibited a significant impact on
161 ENSO simulation are selected for further tuning within the fully coupled experiments. In this



162 phase, each parameter is tested across six different values. The ranges of value of each
 163 parameters in atmosphere-only and fully coupled experiments are determined based on
 164 documented guidelines from the ICON model parameter documentation ([https://www.cosmo-
 166 model.org/content/support/icon/tuning/default.htm](https://www.cosmo-

 165 model.org/content/support/icon/tuning/default.htm)). A comprehensive overview of the tuned
 167 parameters is provided in Table 1. The fully coupled reference run is taken from the low-
 168 resolution configuration of ICON XPP version 1.0 (Müller et al., 2025b). These simulations follow
 169 a pre-industrial control (PiControl) setup, spanning 100 years, with the last 50 years used for
 170 ENSO diagnostics. To contextualize the impact of parameter optimization against the effect of
 171 model resolution, we analyze a fully coupled ICON XPP high-resolution (80 km atmosphere / 20
 172 km ocean) simulation (Müller et al., 2025b), this high-resolution result will be discussed in Section

173

174 **Table 1.** The tuning experiments for ENSO simulation in ICON XPP. If Reference Values both
 175 contain atmosphere-only (AO) and fully coupled (FC) simulations, it means this parameter is both
 176 tuned in AO and FC simulations. The AO and FC reference runs utilize different parameters values
 177 to reach the stable state of atmosphere thermodynamic (Müller et al., 2025b).

Name	Definition Simple Description	Reference Values	Optimized Values	Range
tune_entrorg	Controls how much environmental air mixes into clouds—affects cloud growth and rainfall	AO: 0.00195 FC:0.0021	AO:0.0017 FC: 0.00227	[0.00165, 0.00225]
pr0	Controls how momentum and heat mix in neutral (non-stratified) turbulence—affects air mixing efficiency	AO: 0.87 FC: 0.68	AO: 0.8573 FC:0.793	[0.6, 1.1]
tune_rprcon	Determines how easily cloud water turns into rain—controls rain formation efficiency	AO: 0.00165 FC:0.0014	AO: 0.001 FC:0.00113	[0.0008, 0.002]
tune_eiscrit	Critical estimated inversion strength above which to switch shallow convection	AO: 7	AO: 13.2992	[1, 8]
tune_sc_eis	Critical estimated inversion strength above, which to use modify cloud diagnostic	AO: 7	AO: 9.6581	[1, 8]
tune_zvz0i	Variability range for terminal fall velocity of cloud ice	AO: 1	AO: 0.6974	[0.4, 1]
tune_box_liq_asy	Adjusts the asymmetry in diagnosing liquid cloud cover—can shift how clouds respond to environmental changes.	AO: 3 FC: 3.35	AO: 1.2381 FC: 3.393	[1, 5]



tune_box_liq	Sets the width for diagnosing cloud cover from liquid water—affects how cloudy the model thinks it is	AO: 0.047 FC: 0.05	AO: 0.1008 FC: 0.0265	[0.01, 0.1]
tune_zceff_min	Minimum value controlling how cloud particles stick together—affects formation of larger droplets	AO: 0.025	AO: 0.0481	[0.01, 0.075]
tune_rhebc_land	Relative Humidity (RH) threshold for onset of evaporation below cloud base over land	AO: 0.825	AO: 0.9863	[0.4125, 1.2375]
tune_rcucov	Convective area fraction used for computing evaporation below cloud base	AO: 0.075	AO: 0.1081	[0.01, 0.1]
tune_rdepths	Maximum allowed depth of shallow convection	AO: 20000	AO: 11200	[10000, 30000]
tune_rhebc_land_trop	RH threshold for onset of evaporation below cloud base over land in the tropics	AO: 0.75	AO: -0.538	[0.6, 1]
tune_rhebc_ocean_trop	RH threshold for onset of evaporation below cloud base over sea in the tropics	AO: 0.8	AO: 1.2452	[0.6, 1]
f_theta_limit_fraction	Limits how much the decay of heat mixing is suppressed in very stable conditions	AO: 0.2	AO: 0.323	[0, 0.3]
f_tau_limit_fraction	Similar to above but for momentum mixing (wind-related turbulence)	AO: 0.25	AO: 0.4774	[0, 0.6]
f_theta_decay	Influences how quickly turbulent mixing decays at high atmospheric stability—affects heat transport in stable layers	AO: 1 FC: 1	AO: 0.3493 FC: 1.663	[2, 7]
f_tau_decay	Controls how fast turbulent momentum mixing fades away—affects wind stress decay	AO: 4	AO: 4.5782	[0, 5]
fsl	Surface stress tuning factor	AO: 0.8	AO: 0.3689	[0.1, 0.9]
z0m_min	Minimum roughness length for momentum	AO: 0.000015	AO: 0.0000346	[0, 0.00005]
f_tau0	neutral non-dimensional stress factor	AO: 0.17	AO: 0.2711	[0.1, 0.5]

178

179 3 Optimization method

180 In this section, we lay out a systematic approach in optimizing model parameters in ICON
 181 XPP with the aim to improve the ENSO simulations. We first define the ENSO evaluation metrics
 182 (Section 3.1), then describe how we estimate their sensitivities to parameter perturbations



183 (Section 3.2). We formulate the cost function (Section 3.3), explain its approximation using
184 parameter sensitivities (Section 3.4), and conclude with a description of the optimization scheme
185 (Section 3.5).

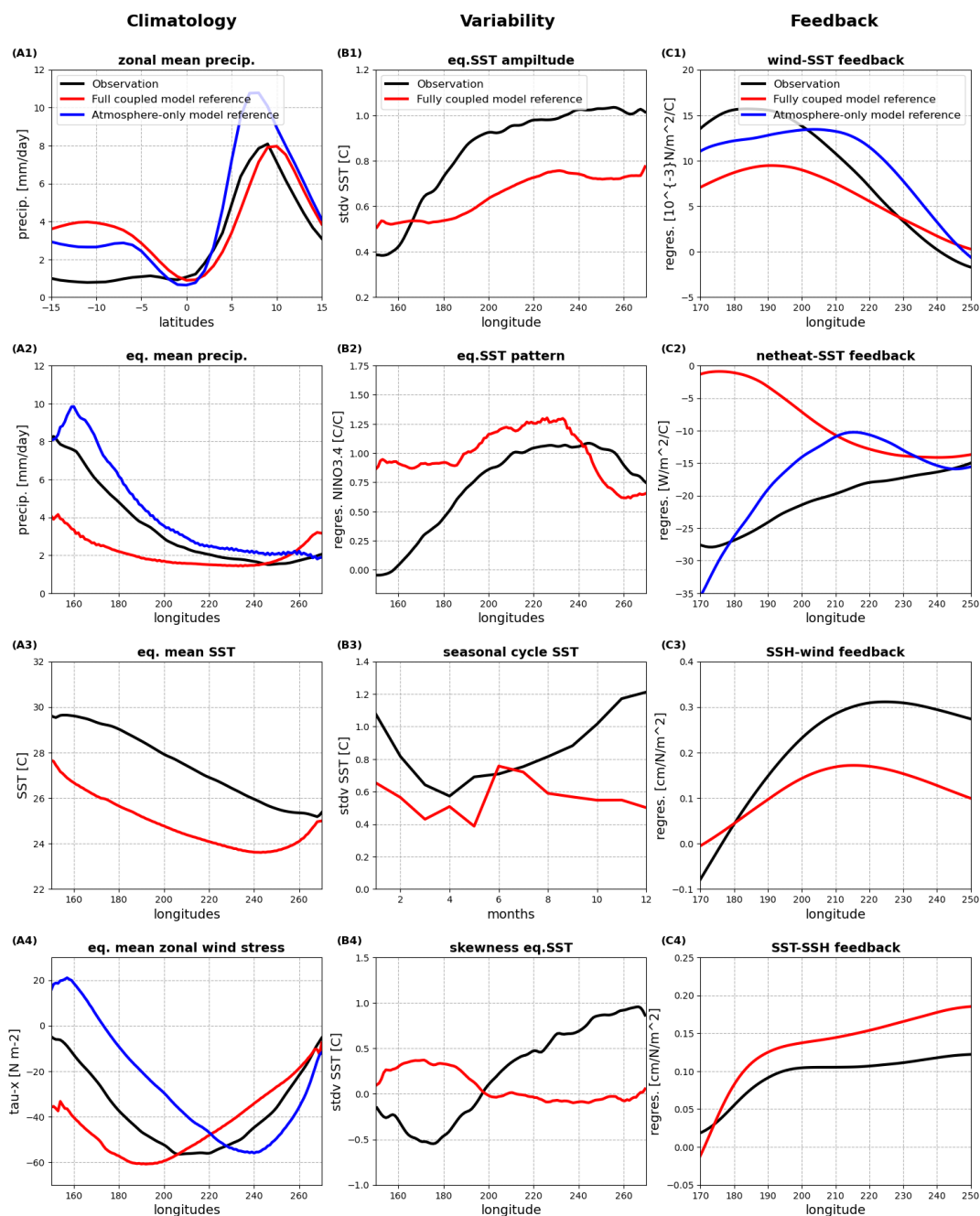
186

187 **3.1 ENSO metrics**

188 The ENSO metric by Planton et al. (2021) is the basis for the model evaluation in this study.
189 It includes 21 different metrics in four different categories (tropical climatology, variability,
190 feedback processes and teleconnections). 17 of these metrics can be expressed as a function of
191 a single spatial (e.g., longitude or latitude) or time dimension, while the remaining four are
192 teleconnection metrics and which will therefore not used as a metric for our model evaluation.
193 The detail definitions of these metrics can be found in Planton et al. (2021) and also the website
194 (https://github.com/CLIVAR-PRP/ENSO_metrics/wiki).

195 Figure 1 shows four examples of the ENSO metrics in rows for each of the three categories
196 in columns. The climatology metrics show common model errors, such as the double ITCZ bias
197 (Fig. 1A1), which appears as excessive precipitation in the southern hemisphere tropics, both in
198 atmosphere-only and fully coupled reference run. An important caveat is that a pronounced
199 double-ITCZ persists in the atmosphere-only configuration, even with observed SSTs prescribed.
200 This implies that a significant fraction of the bias is generated within the atmospheric and land-
201 surface components themselves. Equatorial precipitation errors (Fig. 1A2) illustrate further
202 discrepancies in tropical rainfall patterns. Fully coupled simulations show less precipitation
203 compared with observation, and the atmosphere-only experiments overestimate the
204 precipitation. The cold tongue bias, a persistent issue where fully coupled models underestimate
205 equatorial SST, is apparent in Fig. 1A3. The minimum of equatorial SST is also shifted west in the
206 fully coupled model results. Zonal wind stress biases (Fig. 1A4) reveal substantial discrepancies
207 in both the magnitude and position of the trade-wind maximum, which appears displaced too far
208 west in the fully coupled simulation and too far east in the atmosphere-only experiment.

209



210

211 **Figure 1.** Comparison of some ENSO Metrics for Observations (black), ICON atmosphere-only model (blue), and ICON
 212 fully coupled (red) reference simulations for climatology (left), variability (middle) and feedback (right). The metrics
 213 definition can be found in Planton et al. (2021).



214

215 In terms of ENSO variability, the ICON XPP fully coupled reference run underestimates key
216 features of ENSO behavior. The ENSO amplitude is too weak (Fig. 1B1), the spatial pattern of SSTA
217 is too zonally confined and shifted westward (Fig. 1B2), and the seasonality of ENSO variance is
218 misrepresented, particularly in showing unrealistic peak during May to September and failing to
219 capture the sharp September–February peak (Fig. 1B3). The ENSO skewness (Fig. 1B4), reflecting
220 nonlinear asymmetry between El Niño and La Niña, is longitudinally out-of-phase and strongly
221 underestimated. Most ENSO feedbacks (Fig. 1C1–C3) are persistently underestimated.

222 Critically, ENSO feedbacks (Fig. 1C1–C4) are persistently underestimated. The Bjerknes
223 feedback (zonal wind stress response to SSTA) and thermal damping feedback (net surface heat
224 flux response to SSTA) are weaker than observed (Fig. 1C1–C2). Additionally, ocean coupling
225 metrics involving SSH-SST and SSH-wind stress correlations (Fig. 1C3–C4) show that coupled ICON
226 XPP model fails to reproduce realistic thermocline-related feedbacks. These biases likely limit
227 ENSO amplitude (Fig. 1B1) and skewness (Fig. 1B4) and are tied to overly strong mean upwelling
228 and cold SSTs (cold tongue bias, Fig. 1A3).

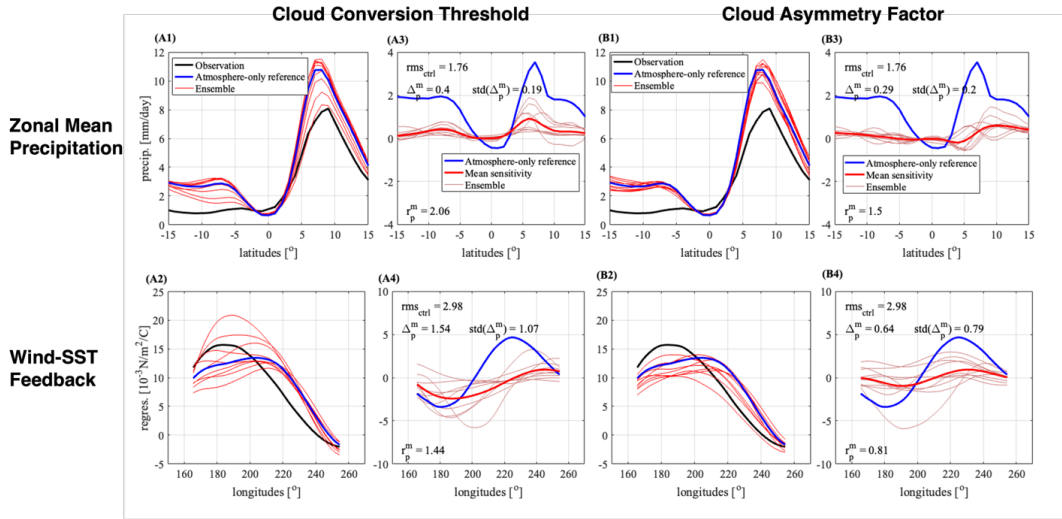
229 In the atmosphere-only simulations, only 8 ENSO metrics that measure atmospheric
230 variables can be used to evaluate the model performance. Five of these are shown in Fig. 1, and
231 the other three are the seasonal cycles of the zonal and eq. mean precipitation, and the eq. mean
232 zonal wind stress. We can notice that the model biases in the atmosphere-only simulations are
233 substantially different from those in the fully coupled simulations (Fig. 1), suggesting that the
234 coupling between ocean and atmosphere lead to substantial changes in model biases.

235

236 **3.2 Parameter sensitivities**

237 The sensitivity of the ENSO-metrics to each tuning parameter of the ICON model is estimated
238 by a set of perturbed parameter runs (detail see Section 2.3). For example, Figure 2 illustrates
239 the sensitivity in atmosphere-only experiments associated with the cloud conversion threshold
240 for cloud water to rain (tune_rprcon) and the cloud asymmetry factor parameters
241 (tune_box_liq_asy). Here the atmosphere-only control run is our reference for which we change
242 a single parameter relative the value of the control run.

243



244

245 **Figure 2:** ENSO metric of Pacific zonal mean precipitation (upper row) and zonal wind-SST regression (lower row) for
 246 observations (black), reference atmosphere-only run (blue) and perturbed cloud conversion threshold (tune_rprcon)
 247 parameters for atmosphere-only runs (red). Column 2 (A3, A4): correspond sensitivity estimate for each perturbation
 248 runs (thin red lines); the ensemble means of them (thick red line) and the reference bias (blue line). B1, B2, B3 and B4
 249 are for cloud asymmetry factor parameter (tune_box_liq_asy).

250

251 We can estimate the sensitivities of the ICON model (λ) to each parameter (p) and for each
 252 ENSO metric (m):

253
$$\lambda_p^m(\varphi) = \frac{d\xi}{dp}(\varphi) \approx \frac{\Delta\xi}{\Delta p}(\varphi) \quad [1]$$

254 Here ξ is the physical variable measured by the ENSO metric m (e.g., precipitation as in Figs.
 255 2A1 and B1, and wind-SST relationship in Figs. 2A2 and B2) and φ is the physical dimension the
 256 metric is depending on (e.g., latitudes). $\Delta\xi$ is the change in ξ relative to the control run, which
 257 means $\frac{\Delta\xi}{\Delta p}(\varphi) = \frac{\xi_{p_j}^m(\varphi) - \xi_{ctrl}^m(\varphi)}{p_j - p_{ctrl}}$, p_j is the value of parameter p value in one of the ensemble
 258 experiments, p_{ctrl} is the value of parameter p in the control experiment.

259 The Root Mean Square (RMS) error of the control run to the observation is given by rms_{ctrl} :

260
$$rms_{ctrl} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\xi_{ctrl}(\varphi_i) - \xi_{obs}(\varphi_i))^2} \quad [2]$$

261 where N is the number of grid points or time steps along the dimension φ_i (e.g., latitude,
 262 longitude, or month), ξ_{ctrl} is the metric value in control simulation, and ξ_{obs} is the corresponding
 263 observed value. Hence, a smaller RMS error value means the model agrees better with
 264 observations.



265 The RMS amplitude of the mean sensitivity $\lambda_p^m(\varphi)$ of ENSO metric m to parameter p is given
 266 by Δ_p^m :

$$267 \quad \Delta_p^m = \sqrt{\frac{1}{N} \sum_{i=1}^N (\overline{\lambda_p^m(\varphi_i)})^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{1}{P} \sum_{j=1}^P \lambda_{p_j}^m(\varphi_i) \right)^2} \quad [3]$$

268 where $\lambda_p^m(\varphi_i) = \frac{\Delta \xi_p^m(\varphi_i)}{\Delta p}$ is the sensitivity of metric m to parameter p at grid point φ_i , the
 269 overbar ($\overline{\cdot}$) denotes the ensemble mean over multiple perturbation runs, and P means the
 270 number of ensemble runs for each parameter p . Hence, in simpler terms, Δ_p^m quantifies how
 271 strongly the model metric changes on average, when parameter p is perturbed. Larger Δ_p^m means
 272 stronger parameter influence on that ENSO metric.

273 The uncertainty of the sensitivities among individual perturbed ensemble members is given
 274 by $std(\Delta_p^m)$:

$$275 \quad std(\Delta_p^m) = \sqrt{\frac{1}{P-1} \sum_{j=1}^P (\Delta_{p_j}^m - \overline{\Delta_p^m})^2} \quad [4]$$

276 , where $\Delta_{p_j}^m = \sqrt{\frac{1}{N} \sum_{i=1}^N (\lambda_{p_j}^m(\varphi_i))^2}$ is the sensitivity for one ensemble member j and $\overline{\Delta_p^m} =$
 277 $\frac{1}{P} \sum_{j=1}^P \Delta_{p_j}^m$ is the ensemble mean of sensitivity over multiple perturbation runs. Hence, $std(\Delta_p^m)$
 278 represents how consistent the sensitivity is across different perturbations. If $std(\Delta_p^m)$ is small,
 279 the response of metric m to parameter p is consistent and reliable across different
 280 perturbations. If it's large, the model's response changes unpredictably between perturbations,
 281 which may result from the model nonlinearities or interactions with other parameters. Finally,
 282 the signal to noise ratio is given by $r_p^m = \frac{\Delta_p^m}{std(\Delta_p^m)}$. It compares the mean signal strength (the
 283 systematic sensitivity) to the ensemble spread (randomness or uncertainty). If $r_p^m > 1$, it means
 284 parameter p has strong, robust, and consistent sensitivity on ENSO metric m . If $r_p^m < 1$, it means
 285 the sensitivity is weak or inconsistent, likely obscured by noise.

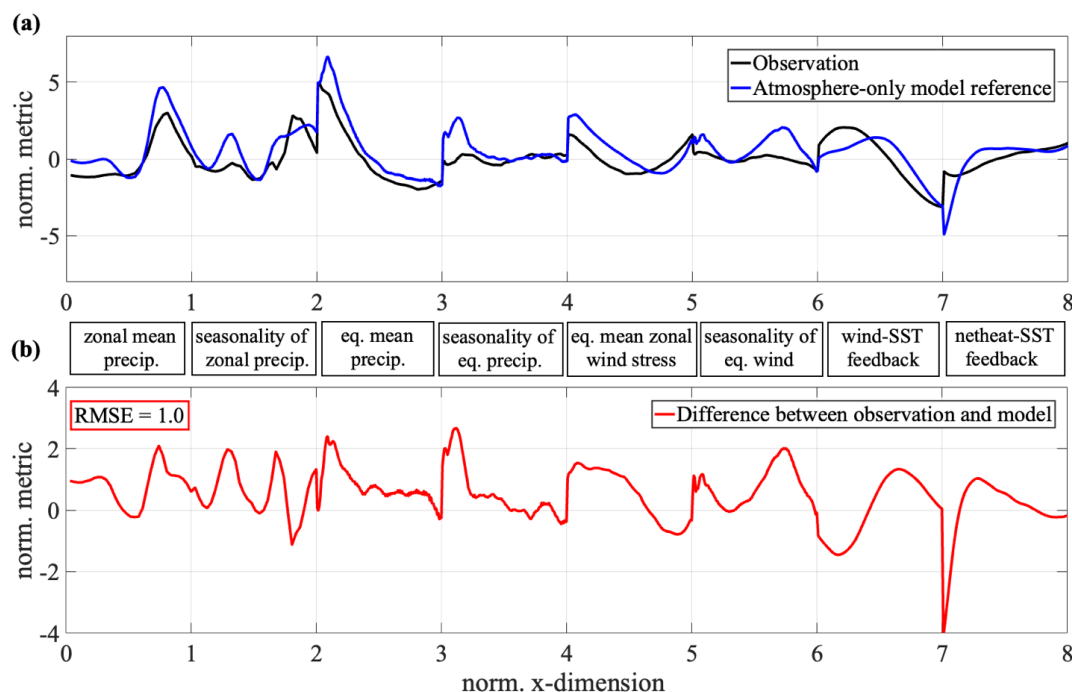
286 Fig. 2A3 shows the estimated values of $\lambda_p^m(\varphi)$ on the meridional precipitation structure for
 287 the cloud conversion threshold parameter perturbation runs and the ensemble mean. Note,
 288 $\lambda_p^m = \frac{\xi_{p_j}^m(\varphi) - \xi_{ctrl}^m(\varphi)}{p_j - p_{ctrl}}$ is a sensitivity in ξ/p relative to the control atmosphere-only run. It does not
 289 consider the observed values and is therefore not measuring a bias. It only estimates if ξ/p would
 290 increase or decrease relative to the control run. We can further see that not all perturbed
 291 parameter runs result into the same λ_p^m , leading to some spread. This spread gives us some
 292 indication of how uncertain these sensitivity estimates are which we can use later for the
 293 optimization scheme. In the example shown in Fig. 2A1 we can note that the tuning parameter
 294 has a significant influence on the tropical precipitation simulation, in particular at latitudes
 295 around 7°N, and a $r_p^m > 1$, which suggests a significant signal to noise ratio. The sensitivity of
 296 different ENSO metrics varies across parameters. For instance, the cloud conversion threshold



297 for cloud water to rain shows a stronger impact on meridional precipitation structure (Fig. 2A3,
 298 $r_p^m = 2.06$) than on the wind–SST correlation (Fig. 2A4, $r_p^m = 1.44$). In contrast, the cloud
 299 asymmetry factor exhibits weaker sensitivities r_p^m for both metrics (Figs. 2B3, B4) compared to
 300 the cloud conversion threshold (Figs. 2A3, A4).

301 Since the ENSO metrics differ in both magnitude and physical dimension, all metrics are
 302 combined into a single composite normalized curve to allow the misfit to be minimized through
 303 a single scalar quantity. For example, for each metric m in the atmosphere-only experiments, we
 304 first compute the model–observation difference $\Delta\xi_m(\varphi)$ along its natural one-dimensional axis
 305 φ (e.g., latitude for zonal-mean quantities in Fig. A1, longitude for equatorial averages in Figs. A2,
 306 A4, C1 and C2), then (i) normalize the amplitude by the RMSE of the atmosphere-only control,
 307 $\widehat{\Delta\xi}_m(\varphi) = \Delta\xi_m(\varphi)/rms_{ctrl}$; and (ii) normalize the axis φ to unit length, $\widehat{\varphi} \in [0,1]$. The eight
 308 normalized metric segments are concatenated sequentially to construct a single composite curve
 309 $\chi(\psi)$ defined over $\psi \in [0,8]$ (Fig. 3). By design, the composite RMSE of the control simulation
 310 equals 1.0 (Fig. 2b). The optimization scheme therefore aims to minimize the amplitude (RMSE)
 311 of $\chi(\psi)$, representing the aggregated ENSO-model misfit across all metrics.

312



313

314 **Figure 3.** Normalized and combined ENSO metric for the atmosphere-only simulations. (a) shows the observed (black)
 315 and atmosphere-only simulation (blue) values of the combined ENSO metric. The x-axis is a combined dimension of
 316 the 8 ENSO metrics used for atmosphere-only simulations with the labels marking which part of the x-axis
 317 corresponds to which ENSO metric. (b) shows the difference between the atmosphere-only model run and the



318 observations (red) with the RMSE value shown in the upper left corner. For 0-1, zonal mean precipitation; 1-2,
319 seasonality of zonal precipitation; 2-3, equatorial mean precipitation; 3-4, seasonality of equatorial precipitation; 4-
320 5, equatorial mean zonal wind stress; 5-6, seasonality of equatorial wind; 6-7, wind-SST feedback; 7-8, net heat-SST
321 feedback.

322

323 The sensitivity λ_p^m will also be normalized by $\widehat{\lambda}_p^m = \frac{\Delta \xi}{\Delta p}(\widehat{\varphi})$. Because $\widehat{\lambda}_p^m$ depends on the
324 specific range of each model parameter p , direct comparison across parameters is not
325 meaningful without further normalization. To ensure comparability, the sensitivities are
326 standardized according to

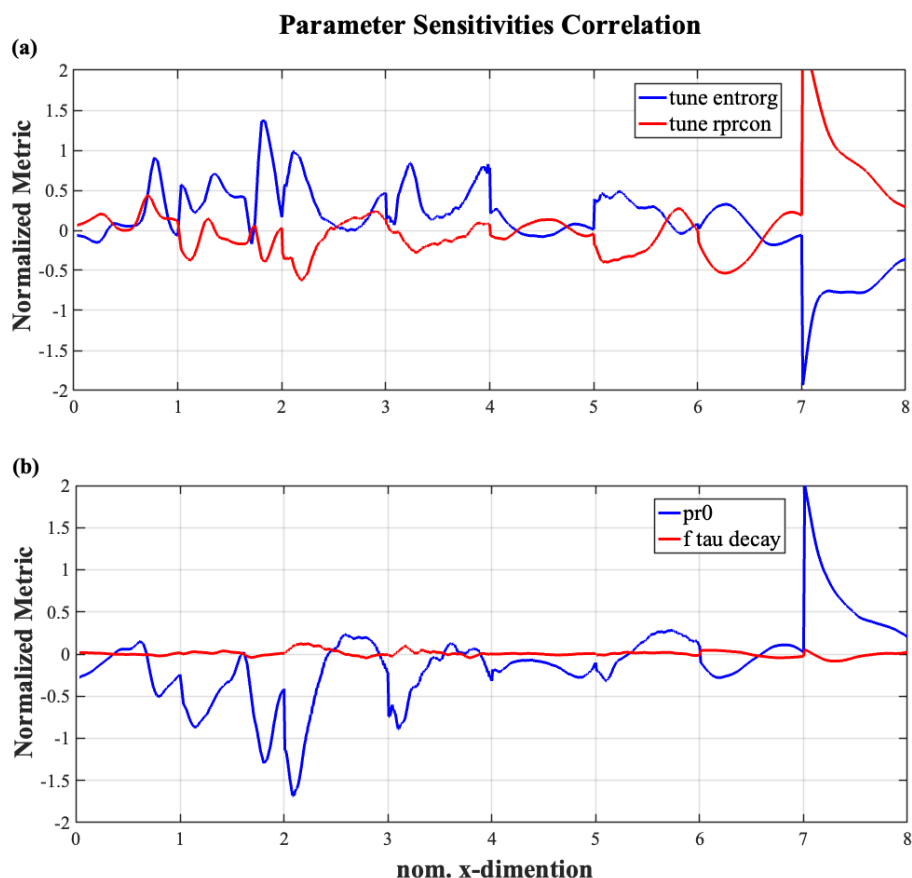
327
$$n\widehat{\lambda}_p^m = \frac{\widehat{\lambda}_p^m}{\sigma_p} \quad [5]$$

328 here σ_p is the standard deviation of the perturbed parameters p , which is approximated as
329 $\frac{1}{4}$ of the parameter range in Table 1.

330 Figure 4a–b shows examples of the normalized sensitivities $n\widehat{\lambda}_p^m$ for four representative
331 tuning parameters across the eight combined ENSO metrics in the atmosphere-only simulations.
332 Each sensitivity is scaled by the RMS error of the corresponding metric in the control run (e.g.,
333 rms_{ctrl} in Figs. 3A2 and A4), enabling a direct comparison across different ENSO metrics. The
334 sensitivities reveal that different parameters influence ENSO-related processes with distinct
335 amplitudes and spatial structures, yet many exhibit broadly similar patterns of response. This
336 similarity is quantified by the strong cross-correlations between parameter sensitivities (-0.83 in
337 Fig. 4a and -0.55 in Fig. 4b), indicating that several parameters act in comparable directions
338 within the ENSO-related error space. In practical terms, such strong correlations imply that the
339 effective degrees of freedom available for independent tuning are reduced: modifying one
340 parameter may partially replicate the effect of another, limiting the uniqueness of the
341 optimization solution.

342 Conversely, parameters that display weak or spatially incoherent sensitivities (e.g.,
343 f_tau_decay in Fig. 4b) contribute less to ENSO variability and exert limited leverage on the
344 overall ENSO metrics. This suggests that not all tunable parameters are equally influential, and
345 that focusing on the few parameters with large, structured sensitivities provides a more efficient
346 path for improving ENSO performance. Together, these results demonstrate that the ENSO-
347 related model behavior in the atmosphere-only configuration is controlled by a small subset of
348 interdependent parameters rather than by the full ensemble of 21 tested parameters.

349



350

351 **Figure 4:** Normalized sensitivity of the combined ENSO metrics in atmosphere-only runs for four different tuning
 352 parameters, (a): `tune_entrorg` and `tune_rprcon`; (b): `pr0` and `f_tau_decay`. The correlation between the pair of
 353 sensitivities is shown in the heading of each panel. The norm. X-dimension has same definition with in Fig. 3.

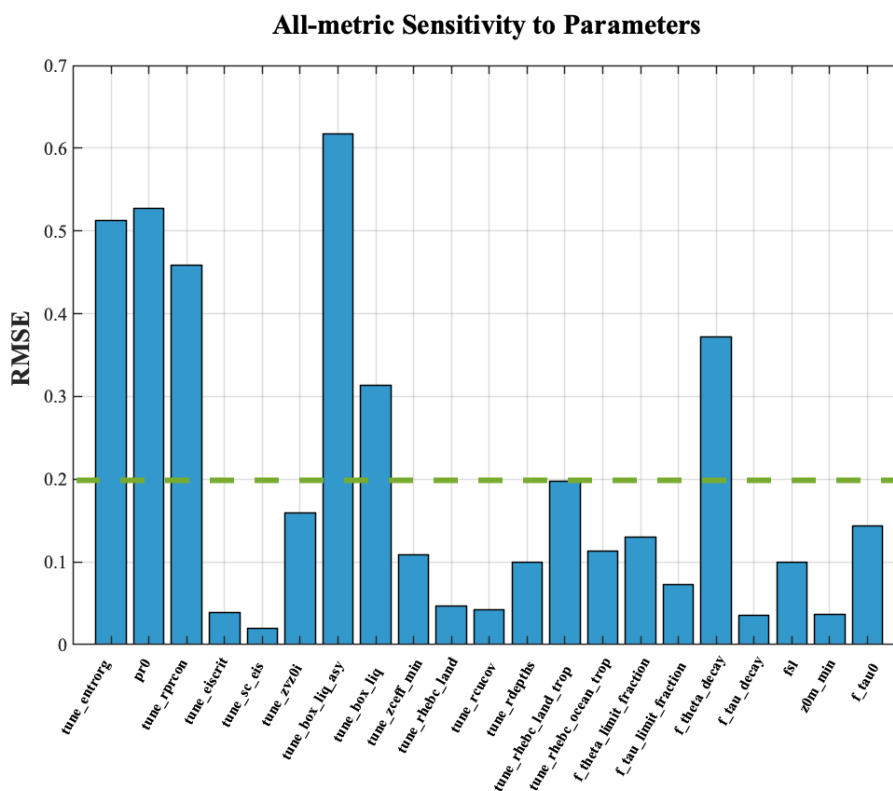
354

355 Figure 5 shows the RMS value of the scaled $n\hat{\lambda}_p^m$, averaged over all eight ENSO metrics for
 356 the atmosphere-only simulations. A mean RMS value of 1.0 corresponds to a sensitivity
 357 magnitude comparable to the ENSO-related bias of the control run, implying that perturbing a
 358 parameter by one standard deviation (σ_p) would alter the ENSO metric by an amount similar in
 359 strength to the model bias, though not necessarily with the same spatial pattern (cf. Figs. 3A2,
 360 A4). Among the 21 tested parameters, only about six exhibit pronounced impacts on the ENSO
 361 metrics, exceeding an RMS threshold of 0.2. This threshold was empirically determined as a
 362 practical cutoff corresponding to roughly 20% of the control-run bias amplitude—large enough
 363 to isolate parameters with meaningful physical influence, yet small enough to avoid spurious
 364 noise from weak sensitivities.



365 Given that these six parameters also display high inter-correlation in their sensitivity patterns
 366 (e.g., Fig. 4a–b), the effective number of independent parameters controlling ENSO behavior in
 367 the atmosphere-only configuration is likely much smaller than 21, and probably closer to two.
 368 This finding underscores that ENSO biases in the atmospheric component are governed by a
 369 limited subset of strongly interacting parameters, rather than by many independent degrees of
 370 freedom.

371



372

373 **Figure 5:** mean RMS values of the sensitivity of the combined ENSO metrics in atmosphere-only runs for all 21
 374 parameters. A mean RMS sensitivity of 1.0 refers to a sensitivity as strong in amplitude as the atmosphere-only
 375 control run biases. Parameters exhibiting RMSE values above the threshold of 0.2 (horizontal green dashed line) are
 376 considered to have significant impacts on ENSO simulation during the atmosphere-only model tuning phase.

377

378 3.3 Cost Function

379 The cost function quantifies the overall model–observation misfit while penalizing large
 380 deviations in the parameter space. It is defined as a positive-definite function, with smaller values



381 indicating a better fit to observations. For both atmosphere-only and fully coupled experiments,
382 the cost function consists of three components:

$$383 \quad f_{cost} = \Delta_{limit}(\Delta_{metric} + \Delta_{para}) \quad [6]$$

384 Δ_{metric} quantifies how well the ENSO metrics are simulated:

$$385 \quad \Delta_{metric} = \frac{\sum_{m=1}^M \omega_m \frac{\delta_m}{\delta_{ctrl,m}}}{\sum_{m=1}^M \omega_m} \quad [7]$$

386 Here, $\delta_m = \sqrt{\frac{1}{N} \sum_{i=1}^N (\xi_{ens,m}(\varphi_i) - \xi_{obs,m}(\varphi_i))^2}$ is the RMS error of metric m (as defined in
387 Planton et al., 2021) in the ensemble experiments, and
388 $\delta_{ctrl,m} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\xi_{ctrl,m}(\varphi_i) - \xi_{obs,m}(\varphi_i))^2}$ is the corresponding RMS error in the control run (cf.
389 Fig. 3b), M is the number of ENSO metrics, which is 8 in the atmosphere-only and 17 in the fully
390 coupled experiments. This normalization is necessary because the ENSO metrics differ in units
391 and amplitude. Although using the control-run uncertainty as a normalization factor is a
392 pragmatic choice rather than an optimal one, it provides a first-order approximation. Each metric
393 is weighted by ω_m , which determines its relative importance (set to 1 in this study for equal
394 weighting).

395 The second term, Δ_{para} , penalizes large deviations of the tuning parameters from their
396 control values:

$$397 \quad \Delta_{para} = \alpha_{para} \frac{1}{K} \sum_{k=1}^K (\Delta np_k)^4 \quad [8]$$

398 with K the total number of perturbed parameters, which is 21 in the atmosphere-only and
399 6 in the fully coupled experiments. p_k are the parameters that are tuned in this study (e.g.
400 tune_entrog, pr0, tune_rprcon in Table 1). Δnp_k is the normalized parameter change relative to
401 the control run:

$$402 \quad \Delta np_i = \frac{\Delta p_k}{\sigma_{para_k}} = \frac{p_k - p_{ctrl}}{\sigma_{para_k}} \quad [9]$$

403 In eq. [8] we have the power of 4 to allow for small changes in the parameter without
404 increasing the cost function. $\alpha_{para} = 0.01$, is a scaling parameter to determine the relative
405 importance of model parameter deviations from the control with respect to the importance of
406 Δ_{metric} . In short, Equation [8] acts as a regularization term that constrains the optimized
407 parameter values from deviating excessively from their control-run settings, thereby ensuring
408 physical plausibility and preventing overfitting of the model to the ENSO metrics.

409 Finally, the constraint term Δ_{limit} ensures that parameters remain within physically
410 meaningful bounds (e.g., positive-definite quantities):

$$411 \quad \Delta_{limit} = n_{limit} + 1 \quad [10]$$



412 where n_{limit} counts the number of parameters that exceed their prescribed physical bounds
 413 within a given parameter combination p_k . For example, if in one candidate parameter set p_k , two
 414 parameters take negative values that violate their positive-definite constraints, then $\Delta_{limit} = 3$,
 415 which substantially increases f_{cost} . Consequently, such an unphysical parameter combination is
 416 automatically discarded during the optimization process.

417

418 3.4 Approximating the ENSO metrics

419 Directly estimating Δ_{metric} requires rerunning the ICON model to compute each δ_m , which is
 420 computationally expensive. As an alternative, we approximate Δ_{metric} using the pre-calculated
 421 parameter sensitivities $n\widehat{\lambda}_p^m$, under the assumption that the bias in each metric, δ_m , can be
 422 represented as a linear combination of these sensitivities:

$$423 \quad \delta_m \approx \delta_m^{approx} = \sum_{k=1}^K \frac{r_{p_k}^m}{r_{p_k}^m + 0.3} \Delta n p_k \cdot n\widehat{\lambda}_p^m \quad [11]$$

424 Each sensitivity is weighted by its confidence ratio $r_{p_k}^m$, as shown as r_p^m in Fig. 3b, to account
 425 for uncertainty in the sensitivity estimates. Parameters with low signal-to-noise ratios ($r_{p_k}^m$)
 426 contribute less to the cost function, ensuring that uncertain or noisy sensitivities exert limited
 427 influence on the optimization.

428

429 3.5 Optimization scheme

430 The optimization process seeks to minimize f_{cost} using the Nelder–Mead simplex method
 431 (Nelder & Mead, 1965; Luersen & Riche, 2004), a derivative-free algorithm that efficiently
 432 searches parameter space through iterative geometric transformations—reflection, expansion,
 433 contraction, and shrinkage. In this implementation, f_{cost} is evaluated using the linear
 434 approximation δ_m^{approx} , which avoids the need to rerun the ICON model for each parameter
 435 combination and thereby reduces computational expense by several orders of magnitude. The
 436 algorithm progressively adjusts the parameter ensemble until a local minimum of f_{cost} is reached,
 437 typically requiring around 1,000 candidate parameter evaluations.

438

439 4 Results for atmosphere-only experiments

440 4.1 Performance of the Optimized Atmosphere-only Configuration

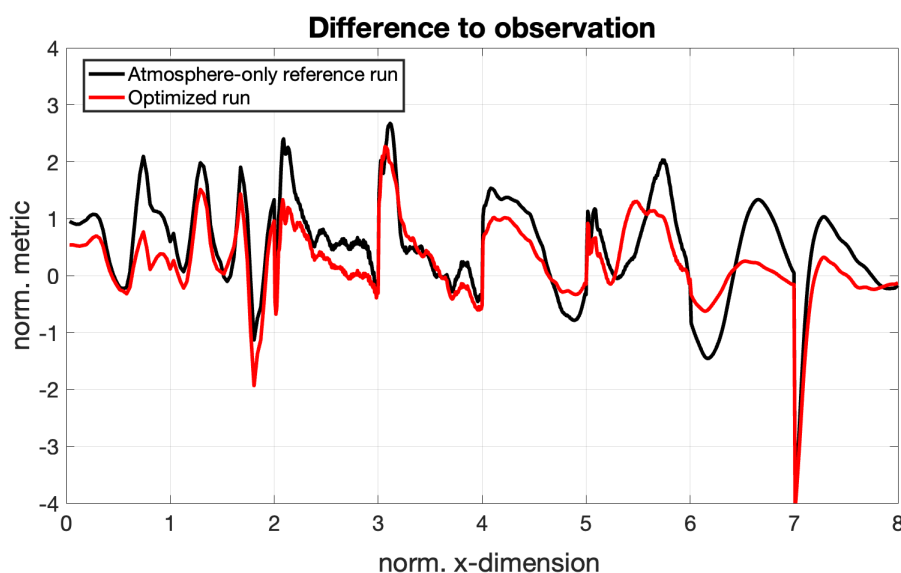
441 The initial tuning phase in this study involves perturbing 21 atmospheric parameters
 442 individually to systematically evaluate their sensitivities using atmosphere-only simulations. Each
 443 parameter's sensitivity is estimated by analyzing its impact on various ENSO metrics, following



444 the method outlined in Section 3.2. The optimization process (Section 3) finds a set of optimal
 445 values for all 21 parameters (Table 1) that minimize the cost function eq. [3]. Figure 6 shows the
 446 normalized and combine ENSO metrics performance for the atmosphere-only run with the
 447 optimized parameters and for the atmosphere-only reference run.

448 The ENSO metrics of the optimizes atmosphere-only run are nearly always closer to zero
 449 than the control run, suggesting improvements in nearly all metrics for nearly all regions. The
 450 overall RMSE error reduced substantially from 1.0 to 0.73. We can further notice the two ENSO
 451 metric curves are similar (corr. = 0.85; Fig. 6), suggesting that the structures of the biases are the
 452 same in the control and the optimized run. This means that persistent biases such as the double
 453 ITCZ or the too weak wind-SST feedback are still present in the optimized run, but have
 454 significantly reduced magnitudes.

455



456

457 **Figure 6:** Normalized and combined ENSO metrics in atmosphere-only simulations for a simulation with the
 458 atmosphere-only reference parameters (black) and for a run with 21 optimized parameters (red). A zero value suggest
 459 a match to the observed reference. The norm. X-dimension has same definition with in Fig. 3.

460

461 4.2 Improvement in global climate beyond the cost function

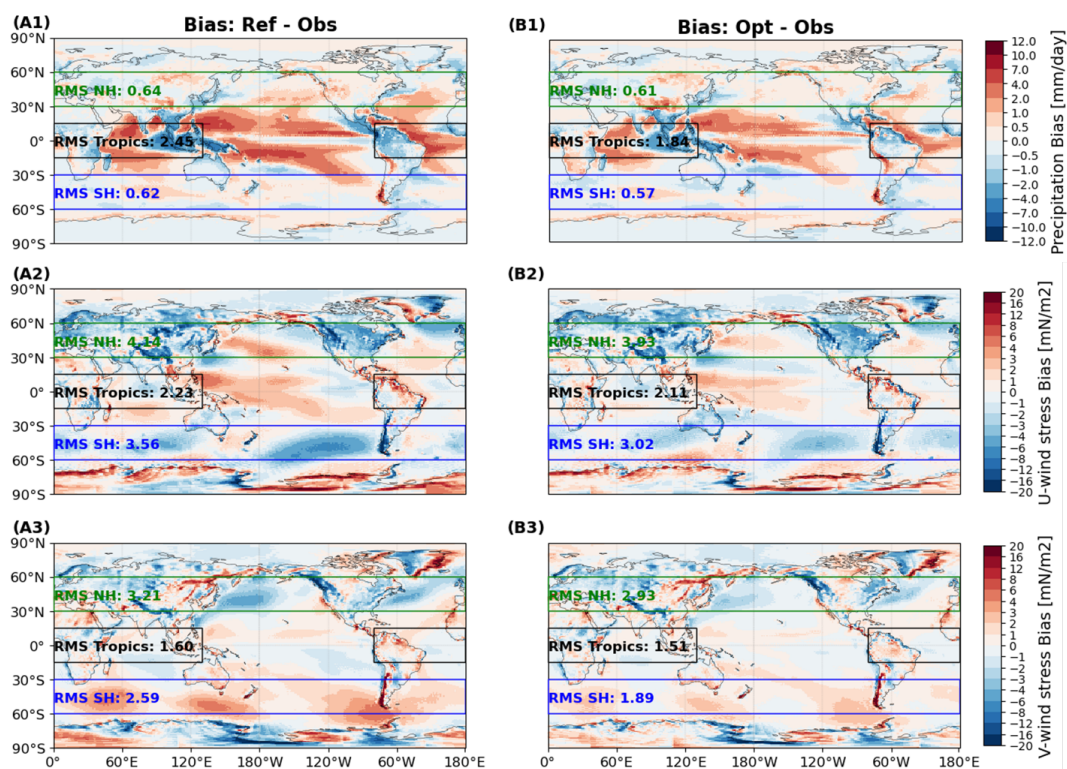
462 The optimization scheme by construction reduces the value of the f_{cost} (eq. 3), which can
 463 only be achieved by reducing the value of Δ_{metric} (eq. 4). Thus, it is by construction that the
 464 optimizes atmosphere-only run will have a smaller RMSE than the control atmosphere-only run
 465 in ENSO metric values shown in Fig. 6. While the improvements in the ENSO metric are
 466 substantial, it is still important to verify that the model simulations improve beyond the



467 characteristics of the climate directly quantified by Δ_{metric} . This can be done by evaluating the
 468 climate simulations beyond the Pacific domain, as none of the metrics included in Δ_{metric}
 469 consider climate variables beyond the Pacific domain.

470 Figure 7 compares the annual-mean biases in precipitation and surface wind stress between
 471 the reference and optimized atmosphere-only runs. The analysis also summarizes area-weighted
 472 RMS errors across three regions: the tropics excluding the Pacific (15°S – 15°N , outside 150° – 270°
 473 E), the Northern Hemisphere extratropic (30° – 60°N), and the Southern Hemisphere extratropic
 474 (30° – 60°S). Although the optimization was explicitly designed to reduce ENSO-related biases
 475 within the tropical Pacific, the results show clear global improvements beyond the target tropical
 476 region. Across all three regions and for all variables—precipitation, zonal wind stress, and
 477 meridional wind stress—the RMSE decreases by approximately 5–27%. The largest fractional
 478 improvements occur in the tropical Atlantic–Indian sector for precipitation, along Northern
 479 Hemisphere storm-track latitudes for zonal wind stress, and over the Southern Ocean for
 480 meridional wind stress. These results demonstrate that the ENSO-focused tuning yields emergent
 481 global benefits, improving large-scale precipitation and momentum flux patterns well beyond the
 482 Pacific basin metrics directly included in the cost function.

483



484



485 **Figure 7:** Global maps of annual-mean biases in precipitation (A1 B1), zonal (A2 B2) and meridional (A3 B3) wind
486 stress for the ICON XPP atmosphere-only reference run ("Bias: Ref - Obs", left column) and optimized atmosphere-
487 only run ("Bias: Opt - Obs", right column), relative to observational datasets. Panel annotations summarize area-
488 weighted RMSE in three extratropic regions: (black) tropics excluding the Pacific (15°S–15°N, outside 150°–270°E),
489 (green) Northern Hemisphere extratropic (30°–60°N), and (blue) Southern Hemisphere extratropic (30°–60°S). RMS
490 NH, RMS Tropics, and RMS SH represent the root mean square values of bias in Northern Hemisphere extratropic,
491 tropics excluding the Pacific, and Southern Hemisphere extratropic, respectively. Decreases in the regional RMS
492 metrics denote improvement.

493

494 5.1 Parameters sensitivity for ENSO metrics in fully coupled experiments

495 Despite substantial improvements achieved in atmospheric processes in atmosphere-only
496 tuning experiments, the ENSO variability and key ENSO feedback mechanisms are naturally not
497 represented, indicating the necessity of coupled ocean-atmosphere tuning for comprehensive
498 ENSO simulation improvement. To address this, a second-stage tuning is conducted using fully
499 coupled fully coupled experiments. Based on the atmosphere-only sensitivity results (Fig. 5), six
500 atmospheric parameters with the strongest impacts on ENSO metrics are selected for fully
501 coupled tuning. Each of these parameters is perturbed individually across a range of values
502 informed by prior sensitivity analysis listed in Table 1.

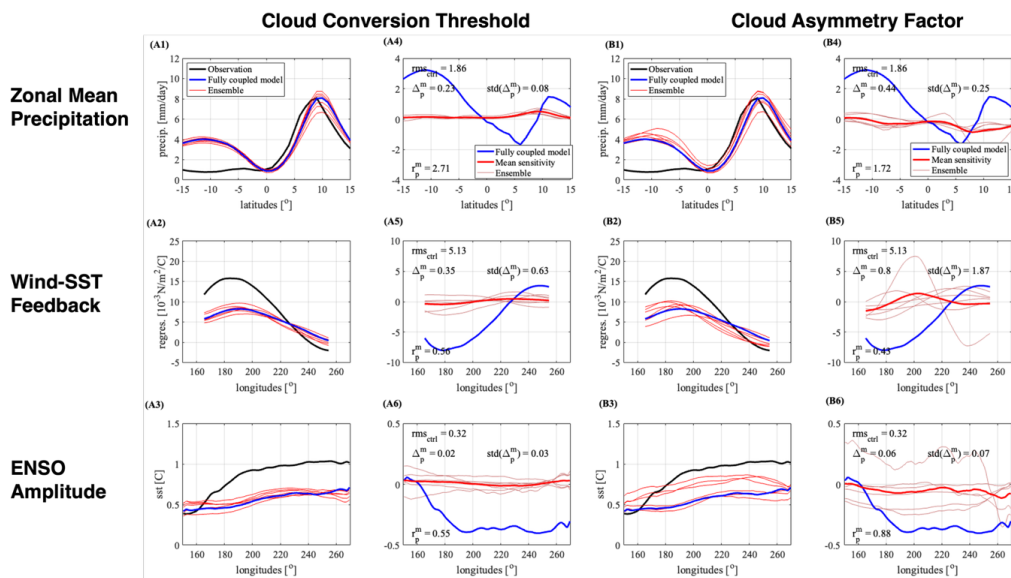
503 To further examine the necessity of direct coupled-model tuning, Figure 8 shows the
504 parameter sensitivity analysis to fully coupled simulations, focusing on the cloud conversion
505 threshold (tune_rprcon) cloud asymmetry factor (tune_box_liq_asy) parameters. This analysis
506 mirrors the atmosphere-only-based results shown in Figure 2, enabling direct comparison of
507 parameter impacts across configurations. In contrast to the atmosphere-only case, where
508 perturbations to the cloud asymmetry factor showed relatively consistent impacts on tropical
509 precipitation and feedback strength, the fully coupled simulations exhibit more muted and
510 spatially variable responses. For example, for the cloud conversion threshold (row A in Fig. 8),
511 the meridional precipitation structure in the eastern Pacific shows muted and latitude-shifted
512 responses: ensemble spread is modest and the signal concentrates in the southern tropics (Fig.
513 8A1), whereas the atmosphere-only case showed a stronger response with a northern-
514 hemisphere maximum (Fig. 2A1). The zonal wind-stress–SST coupling exhibits only moderate
515 ensemble divergence in the fully coupled model (Fig. 8A2), with a weaker and less coherent mean
516 sensitivity than in the atmosphere-only counterpart (Fig. 2A2). Once the ocean is interactive,
517 thermocline and background-wind adjustments partly compensate the atmospheric
518 perturbation, reducing—and sometimes relocating—the effective sensitivity. For the cloud
519 asymmetry factor parameter (row B in Fig. 8), fully coupled responses are smaller and more
520 spatially variable than in the atmosphere-only case. The precipitation metric (Fig. 8B1) shows
521 weaker, patchy sensitivities, while the wind-stress–SST coupling (Fig. 8B2) has a low signal-to-
522 noise ratio, again contrasting with the clearer signal in Fig. 2B2.

523 The results suggests that the cloud asymmetry factor and cloud conversion threshold
524 interact differently with coupled ocean-atmosphere dynamics compared with atmosphere-only
525 experiments, potentially due to compensating oceanic adjustments or altered mean state



526 climatology. This divergence underlines the importance of conducting sensitivity analysis directly
 527 in the coupled configuration rather than relying solely on atmosphere-only-informed
 528 expectations. Actually, when the atmosphere-only optimized parameters are directly applied in
 529 coupled fully coupled simulations, they resulted in unrealistic warming of global mean surface
 530 temperature, approximately 7.3 °C higher than observations (results not shown). Therefore,
 531 direct sensitivity analyses and parameter tuning within the coupled model context are not only
 532 beneficial, but essential for realistic and stable ENSO simulation.

533



534

535 **Figure 8.** Sensitivity of ENSO-relevant metrics to perturbations in the cloud conversion threshold (*tune_rprcon*) (A1-
 536 A6) and cloud asymmetry factor parameter (*tune_box_liq_asy*) (B1-B6) from fully coupled simulations. Panels (A1-
 537 A3 and B1-B3) show metric values from observations (black), ICON XPP fully coupled reference run (blue), and
 538 perturbed parameters ensemble (ensemble) simulations (red) same with atmosphere-only sensitivity results in Figure
 539 3. Corresponding sensitivity estimates are shown in panels (A4-A6, and B4-B6), including ensemble mean sensitivity
 540 (thick red), spread (thin red), and reference bias (blue). The third row (A3, A6, B3 and B6) are for ENSO amplitude,
 541 standard deviation of Sea Surface Temperature Anomalies (SSTA) in the central equatorial Pacific, which is not
 542 included in the atmosphere-only results.

543

544 Figure 8 also introduces ENSO amplitude sensitivity, which is not available in atmosphere-
 545 only configuration. For the cloud asymmetry factor, the fully coupled model shows a consistent
 546 increase in ENSO variability across ensemble members (Fig. 8B3), with a positive and coherent
 547 ensemble-mean sensitivity (Fig. 8B6). This indicates that strengthening cloud asymmetry tends
 548 to amplify central-Pacific SSTA variance, i.e., it acts to increase ENSO amplitude. In contrast, the
 549 cloud conversion threshold exhibits a smaller and noisier response (Fig. 8A3, A6): the ensemble



550 spread is larger and the mean sensitivity is weakly positive at best, implying a limited leverage of
551 this parameter on ENSO amplitude once the ocean is interactive.

552

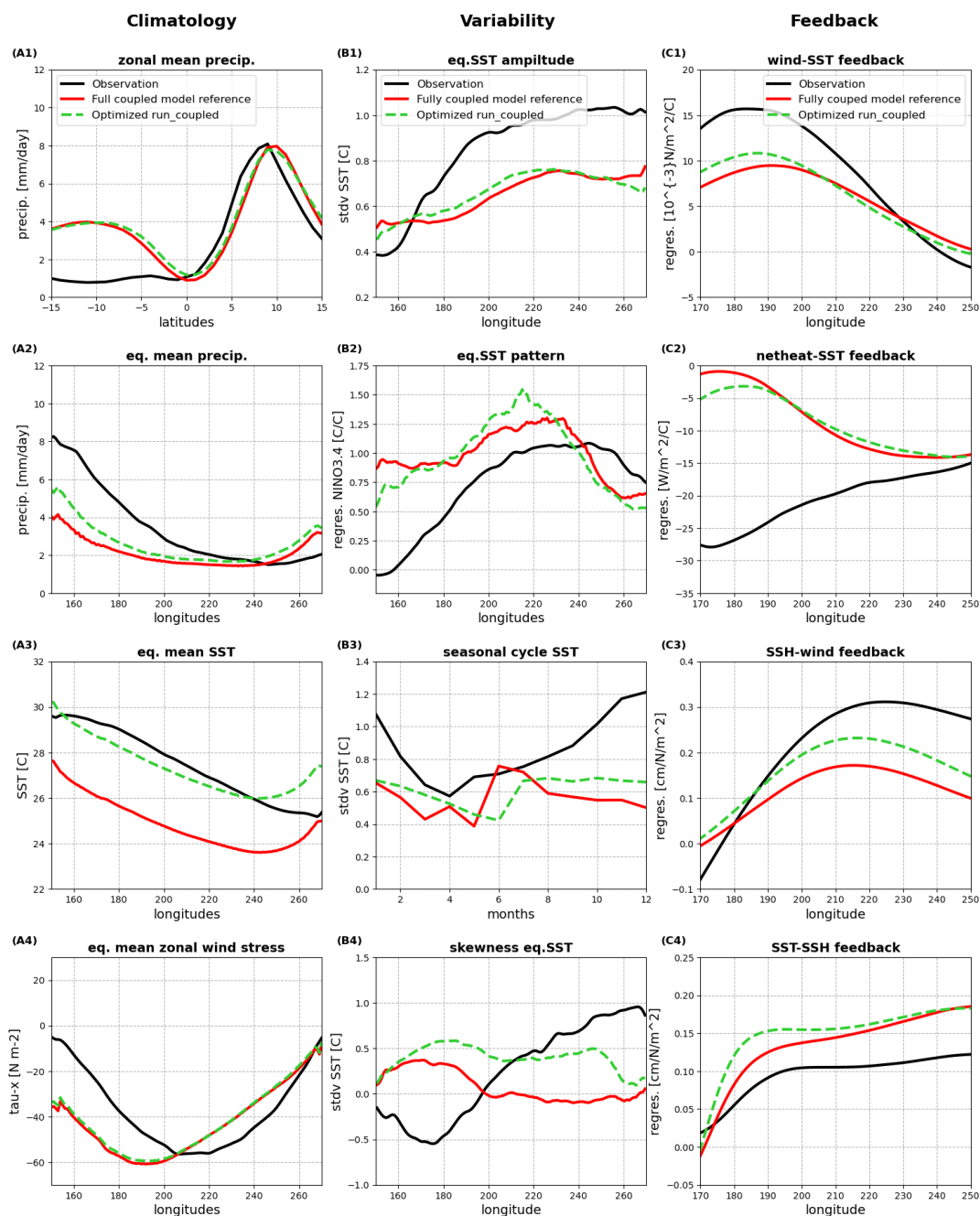
553 **5.2 Performance of the Optimized Coupled Configuration**

554 In the fully coupled optimization, we adopt the same linear sensitivity-based tuning
555 framework used for the atmosphere-only experiments but recalculate the parameter sensitivities
556 within the fully coupled ICON XPP model. The cost function remains structurally consistent with
557 the atmosphere-only case, comprising a weighted sum of normalized RMSEs across selected
558 ENSO metrics. A key difference in this stage is the broader availability of ENSO-relevant
559 diagnostics. In contrast to the atmosphere-only case—where oceanic processes cannot be
560 resolved and certain feedbacks cannot be computed—the fully coupled model allows us to
561 include a complete set of ENSO metrics, as defined in the CLIVAR ENSO Metrics Package (Planton
562 et al., 2021). Each metric is normalized by the standard deviation across the perturbed ensemble
563 and equally weighted in the total cost. This approach enables direct comparison with
564 atmosphere-only results while appropriately accounting for coupled dynamics.

565 Figure 9 illustrates the impacts of coupled-model parameter tuning on ICON XPP's ENSO-
566 related metrics. Observations, fully coupled reference run, and optimized coupled-model
567 experiment are compared, providing insight into how tuning selected atmospheric parameters
568 affects ENSO simulation fidelity in the coupled framework. The optimized parameters' values
569 V_p^{opt} can be found in Table 1. A notable and concerning feature of the tuning results is that nearly
570 all six parameters selected for optimization exhibit opposite directional adjustments between
571 atmosphere-only and fully coupled simulations. Specifically, parameter values that required
572 increases in atmosphere-only tuning (e.g., entrainment rate or cloud asymmetry factor) are
573 found to require reductions in the fully coupled tuning, and vice versa. This reversal suggests that
574 the nature of ENSO-related biases in atmosphere-only simulations differs substantially from
575 those in coupled configurations.

576

577



578

579 **Figure 9.** Comparison of ENSO metrics from the ICON XPP fully coupled reference simulation
 580 (“Fully coupled model reference”, red line), observations (“Observation”, black line), and the



581 optimized coupled-model experiment ("Optimized run_coupled", green dash line). The metrics
582 are identical to those shown in Fig. 1.

583

584 Climatological precipitation metrics show mixed outcomes. The meridional precipitation
585 structure (Fig. 9A1) remains largely unchanged from the reference run, indicating limited
586 progress in correcting the double ITCZ bias. Zonal precipitation bias (Fig. 9A2) is slightly reduced
587 over the western Pacific (150°–200° E), though substantial errors persist further east. In contrast,
588 climatological SST (Fig. 9A3) exhibits a moderate reduction in the cold tongue bias across the
589 central-to-eastern equatorial Pacific, representing a partial improvement achieved through the
590 optimization. However, zonal wind stress (Fig. 9A4) shows minimal change, with the optimized
591 simulation closely resembling the reference case.

592 For ENSO variability, some aspects show improvement while others remain deficient. The
593 amplitude of SST anomalies (Fig. 9B1) increases toward observed levels, although the overall
594 variance remains underestimated. Seasonal phase-locking (Fig. 9B3) shows improvement,
595 particularly during boreal summer, with a reduction of the unrealistic early peak seen in the
596 reference run. The spatial structure of ENSO-related SST anomalies (Fig. 9B2) displays modest
597 improvement between 150° and 200° E, while skewness (Fig. 9B4) becomes more realistic in the
598 eastern Pacific (200°–260° E) but degrades in the central region, suggesting regional trade-offs
599 associated with the atmospheric-only tuning.

600 Feedback processes display a similar pattern of selective improvement. The Bjerknes
601 feedback (Fig. 9C1) and the thermal damping feedback (Fig. 9C2) show moderate enhancement
602 in the central Pacific (170°–200° E). The wind–thermocline feedback (Fig. 9C3) exhibits the
603 clearest improvement, indicating stronger coupling between wind stress and subsurface
604 variability. However, the thermocline–SST feedback (Fig. 9C4) remains largely unchanged,
605 highlighting that key oceanic processes are still inadequately represented despite the
606 atmospheric parameter tuning.

607 In summary, the coupled-model tuning yields valuable improvements in SST climatology,
608 ENSO amplitude, seasonal phase-locking, and air-sea coupling feedbacks. However, persistent
609 deficiencies in precipitation structure, wind stress patterns, and oceanic feedbacks highlight the
610 need for additional tuning to fully capture the dynamics of ENSO within the coupled system.

611

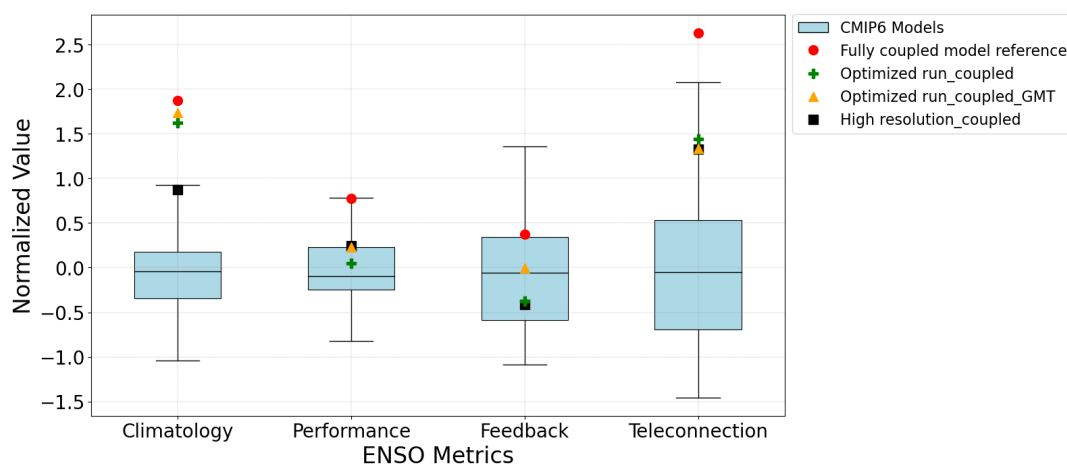
612 **5.3 Comparison of ENSO metrics between Optimized Run and CMIP6 results**

613 Figure 10 summarizes the ICON XPP model's performance across four categories of ENSO
614 metrics—tropical climatology, ENSO characteristics, feedbacks, and teleconnections—comparing
615 the fully coupled reference run, the optimized coupled-model experiment, and a high-resolution
616 ICON configuration (Müller et al., 2025b). The CMIP6 multi-model ensemble distribution (light
617 blue box, Planton et al., 2021) is also shown for context. In these normalized metrics, a lower
618 value indicates smaller deviation from observations and therefore better performance, with



619 values below zero representing improvement relative to the CMIP6 ensemble mean. This
 620 comparison provides a quantitative benchmark for assessing how the ICON XPP configurations
 621 perform relative to state-of-the-art CMIP6 models.

622



623

624 **Figure 10.** Summary of ICON XPP model performance across different categories of ENSO metrics—tropical
 625 climatology, ENSO characteristics, feedbacks, and teleconnections—comparing the fully coupled reference run (“Fully
 626 coupled model reference”, red point), optimized coupled-model experiment (“Optimized run_coupled”, green cross),
 627 fully coupled high resolution run for ICON XPP model (“high resolution_coupled”, black rectangle), optimized coupled-
 628 model experiment with global mean temperature (GMT) correction (“Optimized run_coupled_GMT”, yellow triangle).
 629 The CMIP6 distribution is shown as a boxplot with the box representing the interquartile range (25th–75th percentile),
 630 the central line indicating the median, and whiskers showing the full model spread (minimum to maximum). See
 631 Planton et al. (2021) for detailed definitions of all ENSO metrics and different categories. The ENSO metrics values
 632 for observation are -2.43, -1.77, -2.27, and -6.82 for Climatology, Performance, Feedback, and Teleconnection,
 633 respectively.

634

635 Despite showing limited improvement in overall tropical climatology biases, as
 636 demonstrated by persistent precipitation errors, the optimized fully coupled run achieves
 637 measurable advancements in key aspects of ENSO performance and feedback processes relative
 638 to the reference simulation. Among the four ENSO metric categories, the most notable progress
 639 occurs in the feedback metrics, where the optimized ICON XPP outperforms the CMIP6 multi-
 640 model mean. However, for the other categories—tropical climatology, ENSO characteristics, and
 641 teleconnections—the ICON XPP simulations still exhibit substantial room for improvement to
 642 reach or exceed the CMIP6 ensemble average. Importantly, teleconnection metrics were not
 643 explicitly included in the optimization process; their modest improvement therefore represents
 644 an emergent response to ENSO-targeted tuning, suggesting that systematic parameter
 645 optimization can enhance broader model performance beyond the specific metrics included in
 646 the cost function.



647 In addition to evaluating the reference and optimized runs, Figure 10 further presents the
648 results from an ICON XPP higher-resolving configuration (80 km atmosphere / 20 km ocean), as
649 described in Müller et al. (2025b). Previous studies have consistently demonstrated that
650 increasing model resolution generally leads to improved ENSO simulation performance, including
651 more realistic representations of ocean-atmosphere feedback processes, SST patterns, and ENSO
652 variability (Shaffrey et al., 2009; Roberts et al., 2018). Comparing the high-resolution
653 configuration to the low-resolution reference run confirms that increasing resolution will also
654 substantially improve ENSO Metrics in ICON XPP Earth System Model, especially in tropical
655 climatology and ENSO teleconnection simulations. Most notably, the impact of our optimization
656 result at standard resolution is comparable in magnitude to the improvement gained from
657 doubling model resolution. This equivalence—observed across several ENSO categories—
658 demonstrates the power of systematic linear sensitivity-based tuning as a computationally
659 efficient method for improving climate model performance. Given that our method matches the
660 effectiveness of a much more resource-intensive high-resolution setup, applying this tuning
661 strategy in high-resolution simulations has the potential to yield even greater performance gains.
662 Thus, the combination of systematic parameter tuning and increased resolution represents a
663 potential path forward for advancing ENSO realism in coupled climate models.

664

665 **5.4 Global Mean Temperature Bias and Model Stability adjustment**

666 While the primary objective of this study is to optimize ENSO-related processes, an
667 important unintended consequence emerged during coupled model simulations. Specifically, the
668 optimized fully coupled run ("Optimized run_coupled")—derived from ENSO-focused parameter
669 tuning—exhibits an unrealistic global mean surface warming, with a global mean 2-meter
670 temperature bias reaching +1.58 °C and local maxima up to +15.65 °C. This excessive warming is
671 not present in the reference fully coupled simulation, which remains closer to observed
672 climatology.

673 This warming bias is attributed to a destabilization of turbulent mixing processes, particularly
674 those governed by the heat and momentum mixing parameters ($f_{\text{theta_limit_fraction}}$ and
675 $f_{\text{tau_limit_fraction}}$), which regulate limiting thresholds in the ICON turbulence scheme under
676 strongly stratified conditions. These parameters influence the reduction of turbulent exchange
677 coefficients in stable layers, helping to prevent excessive vertical mixing that could otherwise
678 lead to unrealistic near-surface warming (Dipankar et al., 2015; Heinze et al., 2017). In a revised
679 fully coupled experiment, we addressed this issue by reducing both of these parameters to 0.1.
680 The resulting simulation shows that global mean warming is effectively suppressed (global mean
681 bias reduced to +0.09 °C), while ENSO-related metrics ("Optimized run_coupled_GMT", yellow
682 triangle in Fig. 10) remained largely unchanged and stable compared to the original optimized
683 run.

684 These findings underscore a critical lesson: although targeted tuning of ENSO metrics can
685 lead to meaningful improvements in tropical Pacific variability, it may simultaneously destabilize
686 other components of the global climate system if broader metrics (e.g., global mean



687 temperature, AMOC index) are excluded from the cost function. Therefore, to achieve both
688 regional and global fidelity, future tuning efforts—especially within coupled models—should
689 include constraints based on large-scale climate indicators. This example demonstrates that our
690 linear sensitivity optimization is a powerful tool, but one that benefits from holistic design.
691 Including global-scale constraints will be a crucial step forward in enhancing the physical realism
692 and overall climate consistency of optimized Earth System Models like ICON XPP.

693

694 **6 Summary and Discussion**

695 Although the ENSO phenomenon has distinct impact on global climate patterns and extreme
696 weather events, accurate model representation of ENSO phenomenon is paramount for climate
697 research and prediction. By leveraging the comprehensive ENSO Metrics Package (Planton et al.,
698 2021), this study introduces a systematic, linear sensitivity-based optimization method to
699 enhance ENSO representation in the ICON XPP Earth System Model. The foundation of the
700 optimization method is a targeted ensemble of experiments of span the hyperspace of model
701 parameters. Here we consider 21 atmosphere parameters that are related to cloud cover,
702 microphysics and turbulence parameterization. Based on these experiments, the optimization
703 procedure estimates the sensitivity of the ENSO metrics to the parameters and generates the
704 best possible parameter combination for the optimization experiments.

705 The atmosphere-only optimization experiments demonstrate the effectiveness of the linear
706 sensitivity-based tuning approach in improving ENSO-related performance in ICON XPP. The
707 resulting optimized simulation achieved a ~30% reduction in the ENSO metrics cost function
708 (from 1.0 to 0.73), highlighting significant improvement relative to the reference run. Physically,
709 the optimized parameters lead to meaningful corrections in several key biases, including a
710 reduced double ITCZ pattern in meridional precipitation, improved zonal wind stress, and
711 enhanced representation of the Bjerknes and thermal damping feedbacks. Interestingly,
712 although the optimization targeted only ENSO-specific metrics, broader improvements are
713 observed in global precipitation and wind stress patterns. Hence, the atmosphere-only results
714 confirm that systematic, sensitivity-guided parameter tuning can effectively reduce model biases.

715 Following the atmosphere-only optimization, we conduct a second stage of tuning with fully
716 coupled ICON XPP simulations, repeating the perturbation experiments for the selected
717 parameters. Then, we recalculate parameter sensitivities within the coupled system and apply
718 the same linear sensitivity-based optimization framework. The improvement of ENSO
719 simulations by our optimized method is comparable with the effect of increase resolution in
720 model, and the ENSO feedback enhancements now position the optimized ICON XPP
721 performance better than the average CMIP6 model feedback metrics. Specifically, the
722 optimization yields targeted improvements in several critical ENSO metrics, such as the
723 climatological SST distribution, ENSO amplitude, seasonality, and the wind–thermocline
724 feedback. However, certain ENSO-related characteristics—such as precipitation seasonality, SST
725 variance, and thermocline–SST feedback—remain largely unaffected by atmospheric parameter



726 changes. These limitations underscore the importance of extending systematic tuning strategies
727 beyond the atmosphere, particularly toward ocean model parameters and background state
728 biases.

729 An important insight emerges during fully coupled tuning: while optimizing for ENSO metrics
730 improved regional performance, it inadvertently induces unrealistic global mean temperature
731 (GMT) warming. This warming bias, with global mean 2-meter temperature exceeding
732 observations by over 1.5 °C and maxima above 15 °C, is traced to destabilized vertical mixing due
733 to insufficient constraints on turbulent damping in stably stratified layers. The issue is resolved
734 by reducing the heat and momentum turbulence mixing parameters, restoring realistic GMT
735 values while preserving ENSO improvements. This result underscores that ENSO-focused tuning
736 alone may destabilize other components of the climate system if global metrics are excluded
737 from the tuning process. In future applications, constraints such as global energy balance, AMOC
738 strength, or GMT could be incorporated into the optimization framework to ensure broader
739 physical consistency.

740 In conclusion, this study demonstrates that our newly proposed linear sensitivity-based
741 optimization method effectively identifies and tunes key atmospheric parameters, to improve
742 representation of critical ENSO characteristics and feedback processes fidelity within the ICON
743 XPP Earth System Model. With greater computational resources, the approach can be extended
744 to multi-objective optimization that simultaneously targets ENSO metrics and large-scale climate
745 constraints (e.g., tropical mean SST, global energy balance, AMOC, GMT), and to broader
746 parameter spaces including ocean and sea-ice schemes. Doing this well will require a team effort
747 that combines climate process expertise, numerical optimization and uncertainty quantification
748 (e.g., ensemble design, regularization, and cross-validation against out-of-sample periods).

749

750 **Acknowledgments**

751 This study was funded by the German Ministry of Education and Research (BMBF) through
752 the Coming Decade project (WM, DY; grant 01LP2327E). DD is supported by the Australian
753 Research Council (ARC) Centre of Excellence for Climate Extremes (Grant Number:
754 CE170100023). DY also acknowledges partial support from the Shanghai Jiao Tong University
755 Outstanding Doctoral Graduate Development Scholarship. HP has received funding from the
756 European Union's Horizon Europe research and innovation program under grant agreement No
757 101081460. Views and opinions expressed are, however, those of the author(s) only and do not
758 necessarily reflect those of the European Union. Neither the European Union nor the granting
759 authority can be held responsible for them. We further acknowledge the German Climate
760 Computing Center (DKRZ) for providing the computational resources necessary for this work. We
761 thank Dian Putrasahan at the Max Planck Institute for Meteorology for reviewing the manuscript
762 prior to submission.

763



764 Code and Data Availability Statement

765 The source data and scripts used to run the optimization process for this study are available in
766 the Open Research Data Repository of the Max Planck Society
767 (<https://doi.org/10.17617/3.KT6XVF>) (Yu et al., 2025). ICON model is available to the
768 community under a permissive open-source license (BSD-3C). Please follow the instructions on
769 the ICON web-page (<https://www.icon-model.org/>). The ENSO metrics package is available in
770 https://github.com/CLIVAR-PRP/ENSO_metrics/wiki. The GPCPv2.3, TropFlux, and GODAS
771 dataset can be obtained from <https://psl.noaa.gov/data/gridded/data.gpcp.html>,
772 <https://services.incois.gov.in/tropflux/index.jsp>, and
773 <https://psl.noaa.gov/data/gridded/data.godas.html>, respectively.

774

775 Authors Contribution

776 Conceptualization: DY, DD, PH, WM. Writing - first draft: DY. Writing - review: DY, DD, PH, WM
777 Visualization: DY, DD. Supervision: WM.

778

779 Competing Interests

780 The contact author has declared that none of the authors has any competing interests.

781

782 References

- 783 Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider,
784 U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., & Nelkin, E. (2003). The Version-2
785 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–
786 Present). *Journal of Hydrometeorology*, 4(6), 1147–1167. [https://doi.org/10.1175/1525-7541\(2003\)004<1147:TVGPCP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2)
- 788 Bayr, T., Wengel, C., Latif, M., Dommenges, D., Lübbecke, J., & Park, W. (2019). Error
789 compensation of ENSO atmospheric feedbacks in climate models and its influence on
790 simulated ENSO dynamics. *Climate Dynamics*, 53(1–2), 155–172.
791 <https://doi.org/10.1007/s00382-018-4575-7>
- 792 Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M., & Vialard, J. (2014). ENSO representation in
793 climate models: From CMIP3 to CMIP5. *Climate Dynamics*, 42(7–8), 1999–2018.
794 <https://doi.org/10.1007/s00382-013-1783-z>
- 795 Cai, W., McPhaden, M. J., Grimm, A. M., Rodrigues, R. R., Taschetto, A. S., Garreaud, R. D.,
796 Dewitte, B., Poveda, G., Ham, Y.-G., Santoso, A., Ng, B., Anderson, W., Wang, G., Geng, T., Jo,
797 H.-S., Marengo, J. A., Alves, L. M., Osman, M., Li, S., ... Vera, C. (2020). Climate impacts of the



- 798 El Niño–Southern Oscillation on South America. *Nature Reviews Earth & Environment*, 1(4),
799 215–231. <https://doi.org/10.1038/s43017-020-0040-3>
- 800 Collins, M., An, S. Il, Cai, W., Ganachaud, A., Guilyardi, E., Jin, F. F., Jochum, M., Lengaigne, M.,
801 Power, S., Timmermann, A., Vecchi, G., & Wittenberg, A. (2010). The impact of global
802 warming on the tropical Pacific Ocean and El Niño. *Nature Geoscience*, 3(6), 391–397.
803 <https://doi.org/10.1038/ngeo868>
- 804 Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J.,
805 Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G.,
806 Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J.,
807 ... Strand, W. G. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of*
808 *Advances in Modeling Earth Systems*, 12(2). <https://doi.org/10.1029/2019MS001916>
- 809 Dipankar, A., Stevens, B., Heinze, R., Moseley, C., Zängl, G., Giorgetta, M., & Brdar, S. (2015).
810 Large eddy simulation using the general circulation model ICON. *Journal of Advances in*
811 *Modeling Earth Systems*, 7(3), 963–986. <https://doi.org/10.1002/2015MS000431>
- 812 Guilyardi, E., Braconnot, P., Jin, F. F., Kim, S. T., Kolasinski, M., Li, T., & Musat, I. (2009a).
813 Atmosphere feedbacks during ENSO in a coupled GCM with a modified atmospheric
814 convection scheme. *Journal of Climate*, 22(21), 5698–5718.
815 <https://doi.org/10.1175/2009JCLI2815.1>
- 816 Guilyardi, E., Wittenberg, A., Fedorov, A., Collins, M., Wang, C., Capotondi, A., van Oldenborgh,
817 G. J., & Stockdale, T. (2009b). Understanding El Niño in ocean-atmosphere general circulation
818 models: Progress and challenges. *Bulletin of the American Meteorological Society*, 90(3),
819 325–340. <https://doi.org/10.1175/2008BAMS2387.1>
- 820 Ham, Y. G., Kim, J. H., & Luo, J. J. (2019). Deep learning for multi-year ENSO forecasts. *Nature*,
821 573(7775), 568–572. <https://doi.org/10.1038/s41586-019-1559-7>
- 822 Hanke, M., Redler, R., Holfeld, T., & Yastremsky, M. (2016). YAC 1.2.0: New aspects for coupling
823 software in Earth system modelling. *Geoscientific Model Development*, 9(8), 2755–2769.
824 <https://doi.org/10.5194/gmd-9-2755-2016>
- 825 Heinze, R., Dipankar, A., Henken, C. C., Moseley, C., Sourdeval, O., Trömel, S., Xie, X., Adamidis,
826 P., Ament, F., Baars, H., Barthlott, C., Behrendt, A., Blahak, U., Bley, S., Brdar, S., Brueck, M.,
827 Crewell, S., Deneke, H., Di Girolamo, P., ... Quaas, J. (2017). Large-eddy simulations over
828 Germany using ICON: a comprehensive evaluation. *Quarterly Journal of the Royal*
829 *Meteorological Society*, 143(702), 69–100. <https://doi.org/10.1002/qj.2947>
- 830 Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., Shevliakova, E., Winton,
831 M., Zhao, M., Bushuk, M., Wittenberg, A. T., Wyman, B., Xiang, B., Zhang, R., Anderson, W.,
832 Balaji, V., Donner, L., Dunne, K., Durachta, J., ... Zadeh, N. (2019). Structure and Performance
833 of GFDL's CM4.0 Climate Model. *Journal of Advances in Modeling Earth Systems*, 11(11),
834 3691–3727. <https://doi.org/10.1029/2019MS001829>
- 835 Hwang, Y. T., & Frierson, D. M. W. (2013). Link between the double-intertropical convergence
836 zone problem and cloud biases over the Southern Ocean. *Proceedings of the National*



- 837 *Academy of Sciences of the United States of America*, 110(13), 4935–4940.
838 <https://doi.org/10.1073/pnas.1213302110>
- 839 Jiang, W., Huang, P., Huang, G., & Ying, J. (2021). Origins of the Excessive Westward Extension of
840 ENSO SST Simulated in CMIP5 and CMIP6 Models. *Journal of Climate*, 34(8), 2839–2851.
841 <https://doi.org/10.1175/JCLI-D-20-0551.1>
- 842 Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz,
843 D., & Von Storch, J. S. (2013). Characteristics of the ocean simulations in the Max Planck
844 Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model.
845 *Journal of Advances in Modeling Earth Systems*, 5(2), 422–446.
846 <https://doi.org/10.1002/jame.20023>
- 847 Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., et
848 al. (2021). JSBACH 3 - The land component of the MPI Earth System Model: documentation
849 of version 3.2. *Berichte zur Erdsystemforschung*, 240. doi:10.17617/2.3279802.
- 850 Karamperidou, C., Jin, F. F., & Conroy, J. L. (2017). The importance of ENSO nonlinearities in
851 tropical pacific response to external forcing. *Climate Dynamics*, 49(7–8), 2695–2704.
852 <https://doi.org/10.1007/s00382-016-3475-y>
- 853 Korn, P., Brüggemann, N., Jungclaus, J. H., Lorenz, S. J., Gutjahr, O., Haak, H., Linardakis, L.,
854 Mehlmann, C., Mikolajewicz, U., Notz, D., Putrasahan, D. A., Singh, V., von Storch, J. S., Zhu,
855 X., & Marotzke, J. (2022). ICON-O: The Ocean Component of the ICON Earth System Model—
856 Global Simulation Characteristics and Local Telescoping Capability. *Journal of Advances in
857 Modeling Earth Systems*, 14(10). <https://doi.org/10.1029/2021MS002952>
- 858 Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., Hill, R., Graham, T.,
859 Ridley, J., Blaker, A., Calvert, D., Copsey, D., Ellis, R., Hewitt, H., Hyder, P., Ineson, S., Mulcahy,
860 J., Siahann, A., & Walton, J. (2018). The Low-Resolution Version of HadGEM3 GC3.1:
861 Development and Evaluation for Global Climate. *Journal of Advances in Modeling Earth
862 Systems*, 10(11), 2865–2888. <https://doi.org/10.1029/2018MS001370>
- 863 Kumar, B. P., Vialard, J., Lengaigne, M., Murty, V. S. N., & McPhaden, M. J. (2012). TropFlux: Air-
864 sea fluxes for the global tropical oceans-description and evaluation. *Climate Dynamics*, 38(7–
865 8), 1521–1543. <https://doi.org/10.1007/s00382-011-1115-0>
- 866 Kumar, K. K., Rajagopalan, B., Hoerling, M., Bates, G., & Cane, M. (2006). Unraveling the Mystery
867 of Indian Monsoon Failure During El Niño. *Science*, 314(5796), 115–119.
868 <https://doi.org/10.1126/science.1131152>
- 869 Li, G., & Xie, S. P. (2014). Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial
870 pacific cold tongue and double ITCZ problems. *Journal of Climate*, 27(4), 1765–1780.
871 <https://doi.org/10.1175/JCLI-D-13-00337.1>
- 872 Liu, C., Wang, F., Köhl, A., Wang, X., Wang, C., & Richards, K. J. (2025). Subsurface ocean turbulent
873 mixing enhances central Pacific ENSO. *Nature Communications*, 16(1).
874 <https://doi.org/10.1038/s41467-025-57058-4>



- 875 Lloyd, J., Guilyardi, E., Weller, H., & Slingo, J. (2009). The role of atmosphere feedbacks during
876 ENSO in the CMIP3 models. *Atmospheric Science Letters*, 10(3), 170–176.
877 <https://doi.org/10.1002/asl.227>
- 878 Luersen, M. A., & Le Riche, R. (2004). Globalized Nelder–Mead method for engineering
879 optimization. *Computers & Structures*, 82(23–26), 2251–2260.
880 <https://doi.org/10.1016/j.compstruc.2004.03.072>
- 881 Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M.,
882 Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S.,
883 Haak, H., Hagemann, S., Hedemann, C., ... Roeckner, E. (2019). Developments in the MPI-M
884 Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO₂. *Journal
885 of Advances in Modeling Earth Systems*, 11(4), 998–1038.
886 <https://doi.org/10.1029/2018MS001400>
- 887 McPhaden, M. J., Zebiak, S. E., & Glantz, M. H. (2006). ENSO as an Integrating Concept in Earth
888 Science. *Science*, 314(5806), 1740–1745. <https://doi.org/10.1126/science.1132588>
- 889 Müller, W. A., Früh, B., Korn, P., Potthast, R., Baehr, J., Bettems, J.-M., Bölöni, G., Brienen, S.,
890 Fröhlich, K., Helmert, J., Jungclaus, J., Köhler, M., Lorenz, S., Schneidereit, A., Schnur, R.,
891 Schulz, J.-P., Schlemmer, L., Sgoff, C., Pham, T. V., ... Marotzke, J. (2025a). ICON: Towards
892 vertically integrated model configurations for numerical weather prediction, climate
893 predictions and projections. *Bulletin of the American Meteorological Society*.
894 <https://doi.org/10.1175/BAMS-D-24-0042.1>
- 895 Müller, W. A., Lorenz, Stephan., ... Marotzke, J. (2025b). The ICON-based Earth System Model for
896 Climate Predictions and Projections (ICON XPP v1.0), EGU sphere [preprint],
897 <https://doi.org/10.5194/egusphere-2025-2473>
- 898 Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch,
899 M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D.,
900 Pohlmann, H., Roeckner, E., Stemmler, I., ... Marotzke, J. (2018). A Higher-resolution Version
901 of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *Journal of Advances in
902 Modeling Earth Systems*, 10(7), 1383–1413. <https://doi.org/10.1029/2017MS001217>
- 903 Nelder, J. A., & Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer
904 Journal*, 7(4), 308–313. <https://doi.org/10.1093/comjnl/7.4.308>
- 905 Oueslati, B., & Bellon, G. (2015). The double ITCZ bias in CMIP5 models: interaction between SST,
906 large-scale circulation and precipitation. *Climate Dynamics*, 44(3–4), 585–607.
907 <https://doi.org/10.1007/s00382-015-2468-6>
- 908 Planton, Y. Y., Guilyardi, E., Wittenberg, A. T., Lee, J., Gleckler, P. J., Bayr, T., McGregor, S.,
909 McPhaden, M. J., Power, S., Roehrig, R., Vialard, J., & Voldoire, A. (2021). Evaluating Climate
910 Models with the CLIVAR 2020 ENSO Metrics Package. *Bulletin of the American
911 Meteorological Society*, 102(2), E193–E217. <https://doi.org/10.1175/BAMS-D-19-0337.1>
- 912 Planton, Y. Y., Lee, J., Wittenberg, A. T., Gleckler, P. J., Guilyardi, É., McGregor, S., & McPhaden,
913 M. J. (2024). Estimating Uncertainty in Simulated ENSO Statistics. *Journal of Advances in
914 Modeling Earth Systems*, 16(9). <https://doi.org/10.1029/2023MS004147>



- 915 Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., Van Den Dool, H. M., Pan, H.-L.,
916 Moorthi, S., Behringer, D., Stokes, D., Peña, M., Lord, S., White, G., Ebisuzaki, W., Peng, P., &
917 Xie, P. (2006). *The NCEP Climate Forecast System*.
- 918 Timmermann, A., An, S. Il, Kug, J. S., Jin, F. F., Cai, W., Capotondi, A., Cobb, K., Lengaigne, M.,
919 McPhaden, M. J., Stuecker, M. F., Stein, K., Wittenberg, A. T., Yun, K. S., Bayr, T., Chen, H. C.,
920 Chikamoto, Y., Dewitte, B., Dommenges, D., Grothe, P., ... Zhang, X. (2018). El Niño–Southern
921 Oscillation complexity. *Nature*, 559(7715), 535–545. [https://doi.org/10.1038/s41586-018-](https://doi.org/10.1038/s41586-018-0252-6)
922 0252-6
- 923 Trenberth, K. E., & Hoar, T. J. (1997). El Niño and climate change. *Geophysical Research Letters*,
924 24(23), 3057–3060. <https://doi.org/10.1029/97GL03092>
- 925 Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P.,
926 Graham, T., Hewitt, H. T., Hill, R., Hyder, P., Ineson, S., Johns, T. C., Keen, A. B., Lee, R. W.,
927 Megann, A., Milton, S. F., Rae, J. G. L., Roberts, M. J., ... Xavier, P. K. (2018). The Met Office
928 Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations. *Journal of Advances in*
929 *Modeling Earth Systems*, 10(2), 357–380. <https://doi.org/10.1002/2017MS001115>
- 930 Yu, Dakuan; Dommenges, Dietmar; Pohlmann, Holger; Wolfgang. A. Müller, 2025, "Source data
931 and scripts for publication "A Systematic Atmospheric Parameter Optimization method to
932 Improve ENSO Simulation in the ICON XPP Earth System
933 Model"", <https://doi.org/10.17617/3.KT6XVF>, Edmond, V2
- 934 Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2015). The ICON (ICOsahedral Non-hydrostatic)
935 modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical
936 core. *Quarterly Journal of the Royal Meteorological Society*, 141(687), 563–579.
937 <https://doi.org/10.1002/qj.2378>
- 938 Zhang, T., Shao, X., & Li, S. (2017). Impacts of atmospheric processes on ENSO asymmetry: A
939 comparison between CESM1 and CCSM4. *Journal of Climate*, 30(23), 9743–9762.
940 <https://doi.org/10.1175/JCLI-D-17-0360.1>
- 941