Response to Reviewer #1

This manuscript presents a process-based evaluation of ENSO simulation sensitivity to horizontal resolution in the CAS FGOALS-f3 climate system model. By comparing low-resolution (~100 km) and high-resolution (~25 km) configurations, the authors diagnose differences in ENSO amplitude, oscillation regularity, and underlying air–sea feedback processes using a reproducible framework including BJ index decomposition and high-frequency wind diagnostics. The study is well structured, methodologically transparent, and aligns with the scope of *Geoscientific Model Development*, particularly under the "Model Evaluation Papers" category. The process-oriented approach and the explicit tracing of resolution-sensitive feedback pathways are meaningful for both model developers and modeler users. However, several issues still need clarification or strengthening before publication. Overall, I find the manuscript suitable for publication after minor revision. Below I outline specific comments and suggestions.

We thank the reviewer for his/her valuable and insightful comments and suggestions that help improve the manuscript. Following are the point-to-point replies to the comments (blue indicates original comment, and black indicates our reply).

1. One of the central arguments follows the logical chain "TC->HF westerlies->stochastic forcing-> ENSO irregularity", which is physically plausible and well motivated. However, the manuscript does not quantify the relative magnitude of HF wind variance versus ENSO growth rate. It would be helpful to evaluate whether the stochastic forcing amplitude differs significantly relative to the linear growth rate (e.g., using a simple signal-to-noise ratio metric). Even a simple variance ratio metric or growth rate comparison would further strengthen this section.

Reply: We thank the reviewer for this helpful suggestion. Following this comment, we have introduced a noise-to-signal ratio (NSR) metric to quantify the relative magnitude of stochastic atmospheric forcing compared to the ENSO signal. The NSR is defined as:

$$NSR = \frac{\sigma(u'_{HF})}{\sigma(SSTA_{Niño3.4})} \quad (1)$$

where $\sigma(u'_{HF})$ is the standard deviation of 90-day high-pass-filtered zonal wind anomalies averaged over the western equatorial Pacific (5°S–5°N, 120°E–180°), and $\sigma(SSTA_{Niño3.4})$ denotes the standard deviation of SSTA averaged over the Niño3.4 region (5°S–5°N, 170°–120°W). A larger *NSR* indicates stronger stochastic forcing relative to the ENSO signal, implying a greater potential for HF atmospheric activity to disrupt the regularity of the ENSO cycle.

The *NSR* values are 2.67, 1.98, and 4.73 (m s^{-1} K^{-1}) for the observation, f3-L, and f3-H, respectively. The substantially smaller NSR in f3-L reflects the combination of its weaker high-frequency wind activity and stronger ENSO amplitude, confirming that the stochastic forcing in f3-L is insufficient to disrupt its overly intense ENSO oscillation. In contrast, the larger NSR in f3-H indicates that stronger stochastic forcing acts on a weaker ENSO signal, facilitating the irregular oscillation that more closely resembles the observation. These quantitative results further support our interpretation in the manuscript.

To respond to the reviewer's concern, the corresponding context have added in the revised manuscript.

2. The BJ framework in Section 2.3.2 should be presented more clearly to meet GMD's reproducibility standards. Specifically, every symbol should be defined explicitly, units of each term should be provided, and the areas used for the eastern and western box regions in the BJ index calculation need to be specified. The full formulation can be provided either in the main text (with complete equations) or in an Appendix with a clean, self-contained mathematical definition.

Reply: Thanks for the comment. We agree that a self-contained and explicit mathematical definition is crucial for GMD standards.

In the revised manuscript, we have thoroughly updated Section 2.3.2 to explicitly define all symbols, specify units for every term, and clearly state the eastern and western box regions. These changes ensure the methodology is clear and fully reproducible as requested.

3. For a GMD audience, it would be helpful to briefly discuss the computational cost increase from f3-L to f3-H and provide the implications for CMIP7 model development strategy. This would enhance model-development relevance of the manuscript.

Reply: We thank the reviewer for this constructive comment, which enhances the practical relevance of this study for the model development community. We have added a brief discussion of computational costs and implications for future model development in the last section of the revised manuscript. The key information is summarized below.

The computational cost increases substantially from f3-L to f3-H. The low-resolution version (f3-L) runs on 384 processor cores and achieves a throughput of approximately 15–20 model years per wall-clock day, whereas the high-resolution version (f3-H) requires 6,144 processor cores and achieves only ~0.25 model years per wall-clock day. This represents a vast increase in computational cost (accounting for both the reduced throughput and the 16-fold increase in processor usage), making century-scale ensemble simulations with f3-H considerably more demanding.

This cost–benefit trade-off carries important implications for CMIP7 model development strategy. First, our results demonstrate that the atmospheric resolution of ~25 km is a critical threshold for realistically simulating TC activity and its associated HF wind forcing on ENSO. This finding lends support to the emerging development of variable-resolution modeling frameworks, such as the Model for Prediction Across Scales (MPAS) and the ICOsahedral Nonhydrostatic (ICON) model, which can selectively refine the grid over the tropical Pacific to capture these resolution-sensitive processes while maintaining coarser resolution elsewhere, thereby achieving a favorable balance between physical fidelity and computational affordability. Second, the process-based diagnostic framework employed in this study provides a model-agnostic and reproducible toolkit that can be readily applied to systematically evaluate ENSO simulation across different models and resolutions participating in CMIP7 and HighResMIP2. We encourage the community to adopt such process-oriented diagnostics as a complement to conventional statistical metrics, so that resolution-induced improvements (or degradations) in ENSO simulation can be traced back to specific physical mechanisms rather than assessed solely by outcome-based indices.

4. Consider adding a short graphical summary (schematic) figure illustrating the two key pathways:

(1) “resolution->wind stress structure->feedback->amplitude”,

(2) “resolution->TC->HF noise-> irregularity”).

Such a conceptual figure would help readers quickly grasp the paper’s main messages.

Reply: Thanks for the comment. As suggested, we have added a schematic figure to synthesize the two pathways discussed in the manuscript. For the reviewer’s convenience, this figure (Figure 15 in the revised manuscript) is also copied below.

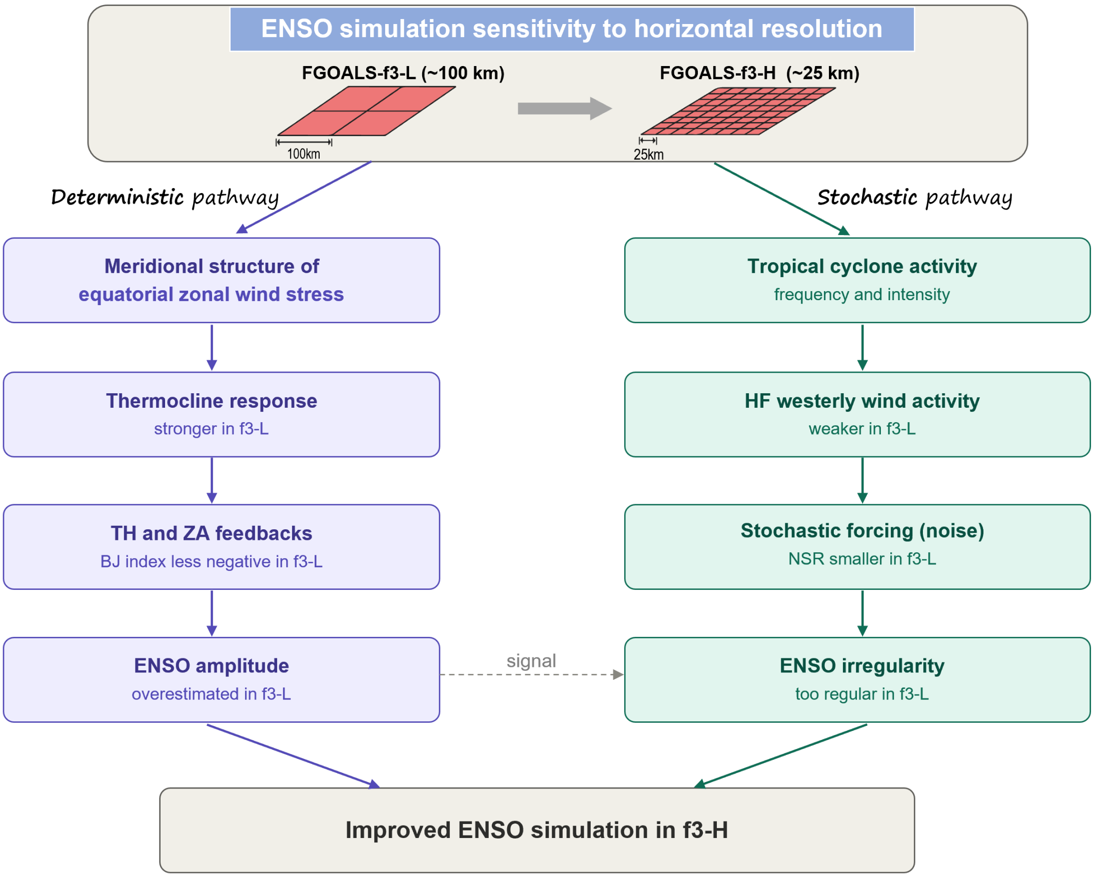

Figure A1 Schematic diagram illustrating how increased horizontal resolution (~100 km to ~25 km) improves ENSO simulation in FGOALS-f3 via both the deterministic feedback processes and the stochastic atmospheric forcing pathways.

5. Line 271-274: ENSO regularity is currently discussed mainly based on qualitative inspection of the Niño3.4 time series. It would be helpful to complement this with a simple quantitative metric of regularity (e.g., spectral peak sharpness/width, autocorrelation-based periodicity, coefficient of variation of event intervals, or an "irregularity index"). This would make the comparison more objective.

Reply: Thanks for this valuable suggestion. As suggested, we have proposed an ENSO irregularity index to quantitatively evaluate the irregularity of ENSO oscillation. This metric is based on the coefficient of variation of inter-event time intervals (CVT), and is computed through the following steps.

First, ENSO events are identified using the 3-month running mean of the Niño3.4 index (monthly SSTA averaged in 5°S–5°N, 170°–120°W). A warm (cold) event is defined when the 3-month running mean Niño3.4 index exceeds 0.5 (falls below –0.5) standard deviation of the Niño3.4 index, and the event is considered to terminate when the

Niño3.4 index returns to within the ± 0.5 standard deviation range for at two consecutive months.

Second, the time interval between two successive events of the same sign is defined as the time separation between adjacent event peaks (i.e., the month of maximum warming for warm events or maximum cooling for cold events).

Third, the CVT is computed as the ratio of the standard deviation to the mean of these inter-event intervals:

$$CVT = \frac{\sigma_T}{\mu_T} \tag{1}$$

where T denotes the set of all inter-event intervals, and μ_T and σ_T denote the mean and the standard deviation of these intervals, respectively.

Finally, the CVT is calculate separately for warm events (CVT_{warm}) and cold events (CVT_{cold}), and their average is taken as the final ENSO irregularity index used in this study. A larger CVT indicates more irregular ENSO oscillation with highly variable inter-event spacing, whereas a smaller CVT (approaching zero) indicates a more periodic and regular oscillation.

Based on the proposed ENSO irregularity index, we calculated the CVT for the observation, f3-L and f3-H. As shown in Table A1, the CVT values are 0.61, 0.17, and 0.53 for the observation, f3-L and f3-H, respectively. The results clearly indicate that ENSO oscillation in f3-L is excessively regular compared to the observation (CVT of 0.17 vs. 0.61), whereas f3-H produces a degree of irregularity much closes to the observation.

Table A1. The ENSO irregularity index (CVT) for the observation, f3-L and f3-H.

	Observation	**f3-L**	**f3-H**
CVT	0.61	0.17	0.53

In summary, these quantitative results corroborate our previous qualitative assessment that ENSO variability in f3-L is overly regular and self-sustained compared to that in f3-H and the observation. To respond to the reviewer's concern, the corresponding analysis and description have been added into the revised manuscript.

6.Several typos and grammatical refinements are still needed.

6.1 Line 184-185: Consider removing the full name after the abbreviation "HF" if it has already been defined earlier.

6.2 Replace “key influencing ENSO simulation” with “a key factor influencing ENSO simulation”.

6.3 Line 134: “use” should be “uses”.

6.4 Line 160: In Table 1, the land component abbreviation should be “CLM4.0”, not “CLIM4.0”.

6.5 Line 174: “are” should be “is”.

6.6 Data availability section contains a duplicated DOI string: https://doi.org/https://doi.org/... Please correct it.

Reply: We are very grateful to the reviewer for your careful and detailed reading of this manuscript. We have corrected all of the listed errors in the revised manuscript. We appreciate your great help in improving the manuscript’s overall quality and readability.