Response to comments of Referee #1

General comment:

This paper examines performance of a newly performed CMIP-style long-term free-run simulation of a newly developed atmosphere-ocean coupled climate model, named KIM, using a community established evaluation toolkit. In particular, the authors focused on the evaluation of ENSO. As the ENSO is one of most influential and significant modes of variability in the climate system, it is logically natural to examine the reproducibility of ENSO for the newly developed model and its simulation. The multiple aspects of ENSO were evaluated collectively from the KIM's CMIP-Historical-like simulation, and the results were compared to the observation and fellow CMIP6 models that are sharing the same type of ocean modeling component, i.e., NEMO. It was demonstrated for the majority of the ENSO aspects, including climatology, seasonality, and teleconnection, KIM's performance is in the range of the selected CMIP6 multi-model performance spread, indicating KIM shows reasonable ENSO reproducibility with identifying areas to improve. I acknowledge the importance of documenting the performance of developing model in a collective way, which could serve as an important resource for future development planning and directing. I believe the manuscript is well prepared, and it is within the scope of the journal, GMD. I only have a few minor suggestions, and recommend accepting the paper after those comments get addressed.

We sincerely thank the reviewer for the valuable time and effort devoted to evaluating our manuscript. We have revised the manuscript in accordance with the reviewer's constructive comments, all of which have helped improve the clarity and quality of the paper. In addition, following the suggestion of Reviewer #2, we changed the analysis period of the coupled KIM from 1952-1990 to 1976-2014 to be consistent with other models. We also added the zonal advection feedback to the analyses of ENSO intensity and seasonality to strengthen the dynamical interpretation of the results. We would appreciate it if the reviewer could also take these changes into consideration. We hope that our revisions adequately address the reviewer's points in the updated version of the manuscript.

Specific comments:

1. Line 110 "While this introduces an initial discontinuity between the 1850 and forcing and 1979 states, preliminary sensitivity tests confirmed that this reanalysis-based initialization strategy yielded robust performance, with negligible differences compared to simulations that include a standard spin-up period": I'd be interested to learn more about this result. Is there a citable literature for this sensitivity tests, or do the authors happen to have those results that could be shared maybe in a supplementary document of this paper?

We thank the reviewer for this comment, which led us to re-examine the statement in question. On reflection, the sentence at L110 did not accurately describe what was done, and we have removed it from the revised manuscript rather than retain a claim we cannot substantiate. We did not carry out a controlled comparison between the present integration and a companion integration branched from an equilibrated pre-industrial state, and we should not have implied that such a comparison had been performed.

The reason the standard CMIP6 protocol was not followed is computational rather than diagnostic. Equilibrating a pre-industrial control integration requires a multi-century simulation, which is not feasible at the horizontal resolution at which the coupled KIM is run (Table 1); KIM was developed as a high-resolution NWP system and the coupled configuration retains that resolution. A piControl integration was attempted at an earlier stage of development using a coarser (~100 km) configuration of an earlier model version. That integration does not correspond to the configuration evaluated here and was not retained for analysis, so we are unable to present diagnostics from it.

We have therefore revised the manuscript so that the design choice rests on its actual justification: reanalysis-based initialization is established practice in coupled prediction systems (Johnson et al., 2019); biases diagnosed in reanalysis-initialized coupled integrations have been shown to correspond closely to those of the same model in free-running climate mode (Vannière et al., 2013); the first 100 years are discarded as an adjustment period; and ENSO, as an interannual mode, is governed primarily by the simulated mean state rather than by the initial condition.

The limitations that follow are stated explicitly in the Discussion: no parallel piControl is available, the simulation cannot be de-drifted in the conventional manner, and long-term trends should not be interpreted as a forced climate response. Producing an equilibrated piControl integration, and thereby a standard piControl–historical pair, is a development priority for the next version of the coupled KIM.

[L106-113: The standard CMIP6 approach, in which the historical experiment is branched from an equilibrated pre-industrial control (piControl) integration, was not adopted here. Reaching equilibrium in a piControl integration requires a multi-century simulation, which is computationally prohibitive at the horizontal resolution used for the coupled KIM (Table 1). A reanalysis-based initialization was therefore used instead. This strategy is widely employed in coupled prediction systems (e.g., Johnson et al., 2019), and biases diagnosed in such initialized integrations have been shown to correspond closely to those of the same model in free-running climate mode (e.g., Vannière et al., 2013). Furthermore, since ENSO is an interannual mode whose characteristics are governed primarily by the simulated mean state rather than by the initial conditions, the influence of the initialization on the diagnostics presented here is expected to be small once the adjustment period has passed.]

[L473-484: The experimental design also differs, as the long-term simulation of the coupled KIM does not strictly follow the standard CMIP6 historical protocol. This departure from the standard protocol leads to a further caveat. No parallel piControl integration is available, so the conventional drift correction applied to CMIP6 historical simulations could not be performed. The long-term trends in the present simulation should therefore not be interpreted as a forced climate response. For this reason, the analysis focuses on ENSO characteristics, which are diagnosed from interannual anomalies. Model drift may, however, affect the mean state metrics, so we tested the robustness of the results to the choice of analysis period. We repeated the analysis using the model years 1952-1990, the period in which the simulated SST best matches the observed SST (see Fig. 1b). The overall results remained largely unchanged, which

indicates that the conclusions are robust to the choice of analysis period. Establishing an equilibrated piControl integration, and thereby a standard CMIP historical pair, is a priority for the next version of the coupled KIM, and will require either additional computational resources or a dedicated reduced-resolution configuration.]

2. Line 120 "Further details are given in the next sub-Sect..": This sentence is looking like it is incomplete because of the two periods (although one for acronym and another for ending the sentence). Maybe it could be revised to something like "Further details are given in the next sub-section" for further clarity.
Thank you for the careful comment. It has been reflected.

3. Figure 1a: Any chance to show observed TOA (or from reanalysis) as well in the upper panel?
We thank the reviewer for this constructive suggestion, which has substantially improved the figure. We have added the global-mean net radiation flux at the top of the atmosphere from the ERA5 reanalysis to the upper panel of Figure 1. One point requires clarification regarding the time axis. Following the suggestion of Reviewer #2, the model evaluation period has been changed to 1976-2014, the period used for the CMIP6 historical simulations. The observational references cover 1982–2020, a 39-year window of equal length beginning with the first full year of OISST; the ERA5 series uses the same window for consistency. Because the two periods are offset by six years, the observational series are plotted against the model time axis so that they align with the model evaluation period, and the caption now states this explicitly. The comparison shows that the simulated TOA net radiation is broadly consistent with the observations, whereas the global mean SST exhibits a systematic warm bias. In the revised manuscript, Figure 1 and its caption have been revised and the corresponding description in the manuscript has been updated as follows.

[L148-163: The analysis period of the coupled KIM is 1976-2014, the same period used for the CMIP6 historical simulation. Figure 1a shows the global-mean, annual-mean net radiative flux at the top of the atmosphere (TOA), defined as incoming solar radiation minus the sum of outgoing longwave and reflected shortwave radiation, from the coupled KIM simulation and the ERA5 reanalysis. The net radiative flux in the coupled KIM converges toward zero after about 50 years of integration, indicating a substantially reduced energy imbalance in the coupled system compared with the early stage of the integration. From the late 1970s onward, however, the net radiative flux becomes slightly positive. This is comparable to the ERA5 estimate, indicating that the simulated TOA energy imbalance during 1976-2014 is broadly consistent with the observed one. Figure 1b presents the annual-mean SST averaged over 60°S–60°N from the coupled KIM simulation and OISST. Consistent with the positive TOA net radiation flux, the simulated SST exhibits a pronounced warming trend after the 1970s. Over the evaluation period, the simulated warming rate is comparable to that of OISST. The simulated SST, however, is systematically warmer than the observations by about 0.4K. This mean state bias may arise from differences in the model's transient response to external forcing relative to the real climate system, including differences in adjustment timescale and effective

climate sensitivity. It may also partly reflect the fact that the long-term coupled simulation has not yet been fully optimized for historical transient forcing—a common challenge when adapting an NWP-focused model for extended climate integrations. Because key aspects of coupled atmosphere–ocean variability, including ENSO, are highly sensitive to the background mean state, this warm bias should be kept in mind when interpreting the results.]

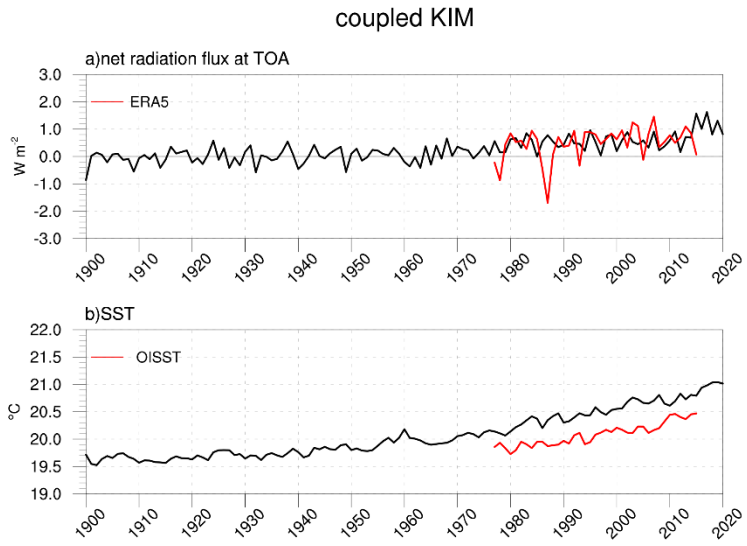


Figure 1: a) Annual-mean, global-mean net radiative flux at the top of the atmosphere (TOA) and b) annual-mean sea surface temperature (SST) averaged over 60°S–60°N in the coupled KIM simulation over the period 1900–2020. The TOA net radiative flux is defined as incoming solar radiation minus outgoing longwave and reflected shortwave radiation. The ERA5 and the OISST time series for 1982–2020 are overlaid in red in (a) and (b), respectively. The observational time axis is shifted by 6 years to align with the model evaluation period 1976–2014 (both 39 years).

4. Code and data availability: While I do understand the institutional restrictions regarding the model source codes, and the authors offer the KIM simulation output and analysis scripts upon request, as this is not comply with the GMD's Open Source rule, is it fair enough to expect that the simulation output (that might have been post-processed to include a subset of data used for the analysis) and the analysis scripts will become publicly available by the time of publication?
We thank the reviewer for raising this point. The same concern was raised by the Chief Editor during the review of our code and data availability statement, and we acknowledge that our original statement did not meet the requirements of the GMD policy. We have addressed this before the revision rather than deferring it to publication. All model output used in the analysis, together with the complete set of analysis scripts, is now publicly archived in a Zenodo repository and is no longer offered on request. With respect to the KIM source code, the institutional restriction remains in force and we are unable to release it publicly. We have revised the availability statement to describe the restriction explicitly rather than to offer the

code on request, and we are providing the supporting institutional documentation to the Editor separately so that this may be assessed under the GMD exception procedure.

[L509-525: The full source code of the Korean Integrated Model (KIM) is not publicly available due to institutional restrictions related to the development and maintenance of the operational numerical weather prediction system. All materials required to reproduce the results presented in this study are archived at https://doi.org/10.5281/zenodo.22003893 (Yeo et al., 2026). The archive comprises the processed KIM simulation output, the atmospheric, oceanic, and sea-ice initial conditions used to launch the coupled integration, the observational and reanalysis datasets used for model evaluation (subset to the variables, regions, and periods required to reproduce the published figures), and the NCL and Python analysis scripts that generate each figure. The redistributed subsets are provided under the license terms of the respective data providers, and the original, unmodified datasets remain available from the sources cited below. ERA5 monthly averaged data on pressure levels and on single levels were obtained from the Copernicus Climate Data Store (Hersbach et al., 2023a, https://doi.org/10.24381/cds.6860a573; Hersbach et al., 2023b, https://doi.org/10.24381/cds.f17050d7); the ERA5 system is documented in Hersbach et al. (2020). ORAS5 global ocean reanalysis data, from which the oceanic and sea-ice initial conditions were derived, were obtained from the Copernicus Climate Data Store (Copernicus Climate Change Service, 2021, https://doi.org/10.24381/cds.67e8eeb7); the system is documented in Zuo et al. (2019). NOAA Optimum Interpolation Sea Surface Temperature version 2.1 is available at https://doi.org/10.25921/RE9P-PT57 (Huang et al., 2021). The Global Precipitation Climatology Project monthly precipitation dataset version 2.3 is available at https://doi.org/10.7289/V56971M6 (Adler et al., 2018). NCEP/DOE Reanalysis 2 is available from the NCAR Research Data Archive at https://doi.org/10.5065/KVQZ-YJ93 (NCEP/NWS/NOAA/U.S. Department of Commerce, 2000); the system is documented in Kanamitsu et al. (2002). The CMIP6 model output used in this study is archived in the Earth System Grid Federation (ESGF). The persistent DataCite DOI of each model, experiment, and variant used is listed individually in Table 1, and each DOI resolves directly to the corresponding ESGF dataset landing page.]