Reply to Review 1 of EC-Earth3-HR paper submitted to Earth system Dynamics:

Main comments:

In this manuscript from Karami and colleagues a newly tuned version of EC-Eath3 at high resolution (EC-Earth3-HR) is presented. The authors compare this new model version with observations, with the ensemble of EC-EARTH3 at standard resolution, and with the previous less tuned high-resolution model performed for HighResMIP. In addition, there is a special focus on understanding AMOC changes by analysing the deep mixed layer and deep-water formation. This is a very nice set of simulations, and its description and evaluation should certainly be published. The focus on the North Atlantic, the Arctic and the AMOC allows a deeper analysis for these regions and processes which is a good idea. However, before I endorse the publication of the manuscript there are three major points that should be improved.

We sincerely thank the reviewer for their thorough and constructive comments. We provide a detailed, point-by-point response below (our responses are shown in blue). We have revised the manuscript accordingly.

First, since it is argued by the authors that the tuning is the difference with the previous high-resolution model version, I suggest that the authors explain the tuning better and reflect on it in the discussion section. For the tuning of the atmosphere, which sea surface temperature is used as boundary condition? Are the greenhouse gases set to 1850 values? What does “minimising the climate drift” mean? Which variables should not drift? How are the 5 parameters in Table 1 chosen? Are those known to be particularly important based on previous experiments with this model? How are those parameters influencing the radiative balance? Was the value of other parameters also tested?

We have expanded the description of the tuning procedure in Section 2. The revised text is included below for the reviewer’s reference.

In brief, the atmospheric tuning was performed in AMIP-type simulations forced with prescribed SST and sea-ice concentrations from the PCMDI AMIP boundary condition dataset (Durack et al., 2022), using historical greenhouse gas concentrations. The term “minimising climate drift” has been removed, and the text now explicitly refers to targeting a TOA radiative imbalance close to observational estimates (Hansen et al., 2011), while avoiding physically unrealistic trends in surface air temperature and precipitation during the tuning integrations. The five atmospheric parameters listed in Table 1 were selected based on previous tuning and sensitivity experiments with EC-Earth3-SR (Döscher et al., 2022), where they were identified as particularly influential for cloud and microphysical convection processes and, consequently, the radiative balance. Due to the high computational cost of the T511 configuration, tuning was restricted to this subset of parameters.

The revised text reads as follows:

> “These AMIP tuning simulations were forced with prescribed SST and sea-ice concentrations from the PCMDI AMIP boundary condition dataset (Durack et al., 2022), and greenhouse gas concentrations followed historical values. Each tuning integration covered 20 years starting in 1990, with the final 15 years used for evaluation. The objective of the tuning was to target a

TOA radiative imbalance close to observational estimates (Hansen et al., 2011). Consistent with Döscher et al. (2022), the tuning further targeted reduced global-mean biases in the net radiative flux at the surface and in the TOA longwave flux. In addition, surface air temperature (TAS) and precipitation were monitored to avoid physically unrealistic trends during the relatively short tuning integrations. Further details of the tuning strategy and target metrics are described in Döscher et al. (2022). A major challenge was achieving a stable model state with relatively short tuning runs. To meet these radiative constraints and to reproduce the observed climate reasonably well, the model was fine-tuned by adjusting selected parameters of sub-grid-scale parameterizations. Specifically, the five atmospheric parameters listed in Table 1 were selected based on previous EC-Earth3-SR tuning and sensitivity studies (Döscher et al., 2022), where these parameters were identified as particularly influential for cloud and microphysical convection processes, which effectively control the radiation balance. Due to the very high computational cost of the T511 configuration, only this restricted subset of atmospheric parameters was retuned, while the remaining atmospheric parameters were kept at their EC-Earth3-SR default values."

Some more information is also needed for the coupled model tuning. For the choice of advection scheme, what does "overall better performance" (l.139) mean? Is it the results that are better? Were specific metrics used or did the authors look at maps of key variables to make a choice? Or was it the computational time or stability of the model that was important for this choice?
We have revised the description of the coupled-model tuning procedure by including more details about the tuning exercise and clarifying the criteria used to evaluate different tuning options, including the choice of ocean advection scheme.

The revised text reads as follows:

"Further tuning focused on the ocean component using coupled simulations. We carried out a set of 15 coupled EC-Earth3-HR simulations (50–180 years each) with different ocean and sea-ice parameter values, forced with fixed radiative conditions representative of 1981. These fixed-forcing simulations were designed to facilitate direct comparison with observations. In parallel, we conducted a coupled simulation under pre-industrial forcing conditions, in which selected tuning parameters were iteratively updated based on insights from the fixed-forcing experiments. Each new configuration was continued from the end of the previous run, under the assumption that parameter changes would induce only incremental adjustments to the model state. Except for the initial experiment initialized from Levitus climatology, all subsequent tuning experiments were concatenated. This approach limited model drift during tuning and reduced computational cost. Therefore, the concatenation of the retained tuning experiments can be regarded as a long pre-spin-up (> 200 years) for the coupled EC-Earth3-HR configuration. Following this concatenated pre-spin-up, we performed a 105-year spin-up using the final parameter set, from which a 350-year pre-industrial control simulation was initialized. Figure A1 shows that, over the final 100 years of this simulation, the global-mean TAS, SST, and Arctic sea-ice area exhibit negligible trends, indicating no ongoing drift in the simulated surface climate. The vertically averaged upper-ocean temperature (0–2000 m) and the AMOC also show negligible drift, providing no indication of ongoing drift in the deeper ocean.

The parameters tested during the tuning are listed in Table 2, with the aim of reducing model biases. The selection of adjustable parameters follows earlier EC-Earth configurations, including EC-Earth3-SR (Döscher et al., 2022) and EC-Earth3P-HR in its second configuration (Haarsma et al., 2020). The evaluation of the different tuning options was based on quantitative diagnostics and climatological bias assessments. Global-mean and map-level comparisons of key variables were used to assess the overall differences relative to observational datasets and reanalysis products. We selected the tuning choices that, overall, reduced the model biases, as discussed below.

With regards to the performance of the different advection schemes, we tested different advection schemes, including the Upstream-Biased Scheme (UBS) and, Total Variations Diminishing (TVD) approach. The UBS configuration showed a persistent warming drift in the global ocean temperature (100–1000 m) and surface heat fluxes that did not stabilise even after more than 130 years, while the TVD simulations showed a marked slowdown of drift and approached a quasi-steady state after about 60 years. The TVD scheme was therefore selected based on its long-term stability rather than on computational cost."

L.139-141: "The turbulent kinetic energy (TKE) mixing below the mixed layer was set to zero (nn_etau=0), as in EC-Earth3-SR (Doscher et al., 2022), which would otherwise lead to a significant reduction in AMOC"
Given that the AMOC changes between the time it is tunned (1850) and the time it can be compared to the RAPID array this is challenging. Was there a specific target for AMOC strength? The authors mention in the results section that the AMOC is larger in the model than in the observations. In retrospect should that value have been non zero?
The choice of nn_etau = 0 in EC-Earth3-HR follows the EC-Earth3-SR configuration, from which the high-resolution model development was initiated. In EC-Earth3-SR, setting nn_etau = 0 led to a stronger and more realistic AMOC (Döscher et al., 2022), and this motivated us to test the same choice in EC-Earth3-HR.

For the coupled tuning, our target was the observed AMOC strength of around 17–18 Sv, consistent with RAPID estimates (~17–18 Sv at 26.5°N; Smeed et al. 2018), based on the assumption that PI-to-present AMOC long-term trend is relatively small. This assumption is supported by reconstructions indicating changes of roughly 0–15% over this period (corresponding to ~0–2 Sv), as summarized in IPCC AR6 WG1. With nn_etau = 0, EC-Earth3-HR attains an AMOC slightly stronger than present observations but close to both the observational range and our pre-industrial target. We emphasize that this choice was not driven by the AMOC alone. Switching to nn_etau = 1 did not improve the overall coupled climate in EC-Earth3-HR, and sensitivity experiments with nn_etau = 0 yielded overall less bias in large-scale SST and sea-ice patterns in the North Atlantic–Arctic sector.

We have added clarifications for this choice in Section 2:

"The turbulent kinetic energy (TKE) mixing below the mixed layer was set to zero (nn_etau=0), as in EC-Earth3-SR (Döscher et al., 2022), which would otherwise lead to a significant reduction in AMOC. The tuning target was an observed AMOC strength in the range 17–18 Sv, consistent with RAPID estimates at 26.5°N (Smeed et al., 2018). It should be emphasized that this choice was not driven by the AMOC alone. Switching to nn_etau = 1

> did not improve the overall coupled climate in EC-Earth3-HR, and sensitivity experiments with nn_etau = 0 produced preferable large-scale SST and sea-ice patterns in the North Atlantic–Arctic sector."

The eddy diffusivity for tracers was increased a lot compared to EC-Earth3P-HR. What was the rationale behind that choice? It is rather counter-intuitive that this parameter is now the same in the standard resolution and in the high resolution because with a higher resolution one would expect that less eddy diffusion needs to be parameterised.
We agree with the reviewer that, from a purely theoretical perspective, one would generally expect a lower lateral eddy diffusivity to be required at higher resolution. The increase of the tracer diffusivity from 300 $m^2 s^{-1}$ (EC-Earth3P-HR) to 1000 $m^2 s^{-1}$ (EC-Earth3-HR) was therefore not motivated a priori by resolution arguments, but was the outcome of targeted sensitivity experiments during the coupled tuning phase. Starting from the EC-Earth3-SR parameter set, we tested different values of rn_aht in EC-Earth3-HR in order to assess the sensitivity of the large-scale circulation and North Atlantic climate to lateral tracer mixing. Compared to a configuration with rn_aht = 300 $m^2 s^{-1}$, the configuration with rn_aht = 1000 $m^2 s^{-1}$ exhibits a stronger AMOC and deeper convection in the Labrador Sea. It also reduced the model bias for key surface climate features in the North Atlantic–Arctic sector. We therefore retain rn_aht = 1000 because of its overall better performance, while acknowledging that this improvement may partly reflect error compensation.

We have added the following sentence in Section 2 for clarity:

> "Sensitivity experiments showed that rn_aht_0 = 1000 $m^2 s^{-1}$ yields a stronger AMOC and deeper Labrador Sea convection than rn_aht_0 = 300 $m^2 s^{-1}$, with an overall improved North Atlantic–Arctic surface climate."

What about the viscosity? Which type of viscosity and which coefficient is used? How does it differ between the different model versions and resolution?
We have added a description of the viscosity formulation and parameter values in Section 2, including the type of lateral momentum mixing, the reference coefficients used, and how these differ between the high-resolution and standard-resolution model configurations.

The revised manuscript text reads:

> "For the lateral momentum mixing, EC-Earth-HR employs 2D-varying bilaplacian eddy viscosity with the reference value at the Equator of -6.4×10^{11} $m^4 s^{-1}$, whereas EC-Earth-SR uses a 3D-varying Laplacian eddy viscosity with the reference value of 2×10^4 m^2s^{-1}. We set the background vertical mixing parameters to 10^{-4} m^2s^{-1} and 10^{-5} m^2s^{-1} for the vertical eddy viscosity and diffusivity, respectively, compared with 1.2×10^{-4} m^2s^{-1} and 1.2×10^{-5} m^2s^{-1} in EC-Earth-SR."

The second point to improve is the connection between the main conclusions and the analysis. The abstract is written in a precise way but in the discussion and conclusion section some claims are not supported by the analysis:
l.551-552: "the updated configuration shows reduced radiative imbalances and minimal deep-ocean drift." This is not shown. I suggest adding a figure to show it. There is a clear drift of the surface

temperature in the model (Figure A1) so I expect that the deep ocean will also still be drifting.
We agree that this statement required supporting evidence. To address this, we have added two new panels to Appendix Fig. A1 from the same pre-industrial control simulation: panel (c) showing the global-mean SST and panel (d) showing the vertically averaged upper-ocean temperature (0–2000 m). While an initial adjustment is evident in the early stages of the integration, the drift in these quantities is negligible over the final 100 years of the simulation. In addition, the AMOC (panel b) exhibits no systematic drift over the full simulation, indicating no systematic drift in the deeper ocean layers. We also note that panel (a) of Figure A1 has been corrected to include the final 12 years of the simulation, which were missing in the original submission due to a plotting error.
The updated Figure A1 is shown below and has been included in the revised manuscript.

The following text is also added in Section 2, for concision and alignment, at the first mention of Figure A1:

> "Figure A1 shows that, over the final 100 years of this simulation, the global-mean TAS, SST, and Arctic sea-ice area exhibit negligible trends, indicating no ongoing drift in the simulated surface climate. The vertically averaged upper-ocean temperature (0–2000 m) and the AMOC also show negligible drift, providing no indication of ongoing drift in the deeper ocean."

Revising the sentence referred to by the reviewer to remove overstatement:

> "Compared to EC-Earth3P-HR simulations under HighResMIP, which suffered persistent biases and oceanic drift (Haarsma et al., 2020; Roberts et al., 2019), the updated configuration shows no evidence of long-term drift in the surface climate and has negligible drift in the deep ocean during the final 100 years of the pre-industrial control simulation (Figure A1)."

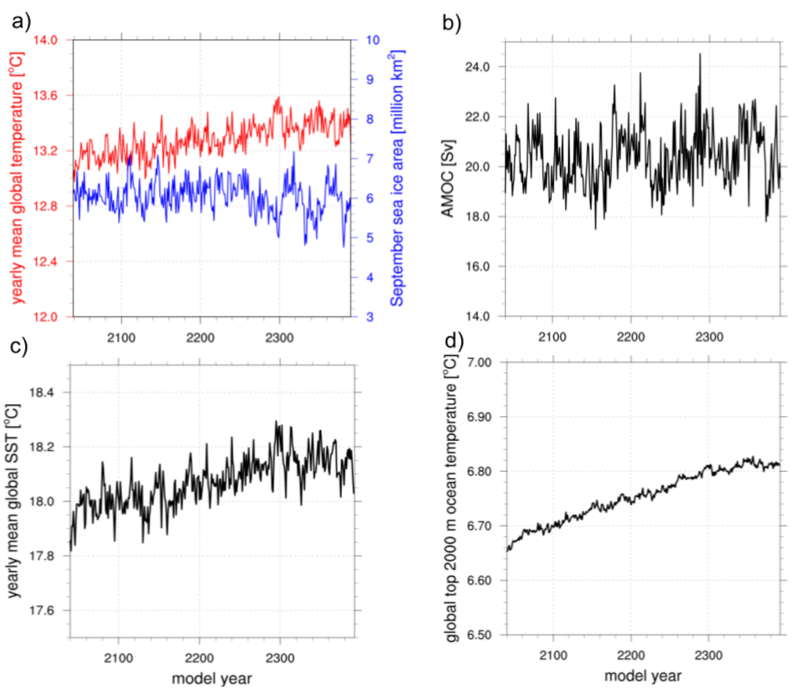

Figure A1 *(UPDATED). Time series from the 350-year EC-Earth3-HR pre-industrial (PI) control run showing: (a) annual-mean global surface air temperature (°C, red) and September sea ice area (million km², blue); (b) annual-mean AMOC strength, defined as the maximum overturning streamfunction between 20°–40°N at 800–1100 m depth. (c) annual-mean global sea-surface temperature (SST); and (d) annual-mean global vertically averaged upper-ocean temperature (0–2000 m). An initial adjustment is evident early in the integration, while the residual drift becomes negligible over the final 100 years shown.*

The enhanced abilities or improved performance of the HR model compared to SR is not argued in the result section. Yet the following is written in the discussion and conclusion section:
l.559-561: “Time series of global and Arctic temperatures, AMOC strength, SST and sea ice area show better agreement with observations and reanalysis, indicating that higher resolution enhances the model’s ability to capture transient responses and low-frequency variability”
We agree with the reviewer that clearer qualification of the HR–SR comparison was required. As stated in the manuscript, our analysis is based on a single-member simulation, which limits the robustness of comparisons between one EC-Earth3-HR realization and the ensemble mean of EC-Earth3-SR. We have therefore revised the wording of this sentence and the preceding one to better qualify the comparison and to avoid overinterpretation. The comparison is further illustrated by Figure A6, which was added to the revised manuscript in response to a subsequent reviewer comment.
The revised text now reads:

> “EC-Earth3-HR shows modest improvements in the climatological means of key variables (e.g., SST, TAS, precipitation, and sea ice) compared to EC-Earth3-SR. Time series of global and Arctic TAS, global SST, and Arctic sea ice area also show better agreement with observations and reanalysis (Figure A6).”

l.568-569: “The improved performance of the EC-Earth3-HR model in simulating Arctic–North Atlantic climate extends to key components such as Arctic sea ice, deep convection, and AMOC variability” …. l.586-587: “This is while the horizontal and vertical distribution of water masses improves in EC-Earth3-HR.”
We thank the reviewer for pointing this out. Demonstrating improved performance in deep convection, AMOC variability, and the distribution of water masses would require additional targeted analyses that go beyond the scope of the present study. We have therefore removed these statements from the revised manuscript.

Based on Figure 6b it seems difficult to argue that HR performs better than LR for the Artic temperature. The improvements in AMOC trend and variability are also difficult to see Figure 9a. I suggest that the authors write more precisely what improves and what doesn’t in the HR model and define clear metrics to backup claims of improvements: rate of change, standard deviation, root-mean-square-error…
We agree with the reviewer that qualitative inspection of Figures 6b and 9a alone does not allow a robust assessment of improved performance in EC-Earth3-HR. To address this, we have introduced explicit quantitative metrics, including the correlation coefficient, root-mean-square error, and mean bias relative to observations and reanalysis.
The results are summarized in the new Figure A6, which is shown below and has been added to the revised manuscript. These metrics indicate improvements in EC-Earth3-HR for global mean surface air temperature, global SST, and Arctic sea-ice area. Improvements in Arctic surface air temperature are modest, while for AMOC strength the applied metrics do not indicate improved performance in EC-Earth3-HR relative to EC-Earth3-SR (AMOC metrics are not shown). Accordingly, we have revised the manuscript text as follows:

> “Time series of global and Arctic TAS, global SST, and Arctic sea ice area also show better agreement with observations and reanalysis (Figure A6).”

The clause “indicating that higher resolution enhances the model’s ability to capture transient responses and low-frequency variability” has been removed, as it is not directly supported by the applied quantitative metrics.

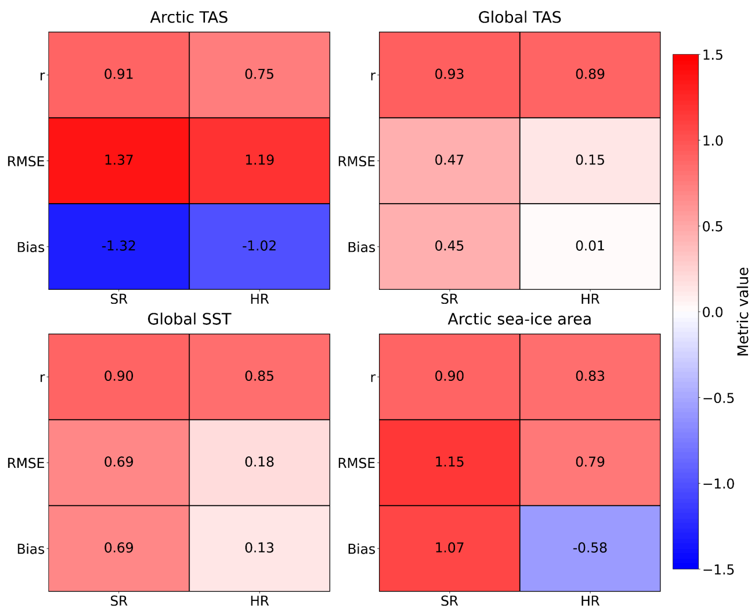

Figure A6 (Newly added). Quantitative evaluation of EC-Earth3-SR and EC-Earth3-HR performance relative to observations and reanalysis for Arctic and global surface air temperature (TAS), global sea surface temperature (SST), and Arctic sea-ice area. Shown are correlation (r), root-mean-square error (RMSE), and mean bias for each variable. Metrics are computed over the historical period: 1950–2014 for TAS, 1870–2014 for SST, and 1979–2014 for sea ice.

The third aspect to be improved is the data and code availability. This is important for other researchers to build on this work. The code used for the analysis and plot of this manuscript is not available. It is mentioned that the CDFTOOLS are used for the DWF analysis but this is not precise enough. Which script from the CDFTOOLS were used? With which options? Enough information should be provided for someone to be able to reproduce the DWF analysis. Especially since this analysis is a novelty from this manuscript.
We have made all CDFTOOLS codes used for the DWF diagnostics publicly available in this repository: https://github.com/enerle/CDF-analysis/tree/main
Specifically, the DWF analysis relies on the cdftransport operator, which was executed for each section and time slice using explicit command-line calls documented in the repository. The repository includes the exact scripts used, all options passed to cdftransport, and the post-processing steps required to compute the diagnostics. This information is sufficient to fully reproduce the DWF analysis presented in the manuscript.

The following paragraph is now added to “Calculation of Deep Water Formation (DWF) analysis” in the revised manuscript:

> “In particular, volume transports across each section are calculated with the cdftransport operator applied to the EC-Earth HR output for each time slice. The exact command-line calls, including all non-default options, section definitions, and post-processing scripts used to derive the mean diagnostics, are provided in a publicly available repository (https://github.com/enerle/CDF-analysis/tree/main). This information is sufficient to reproduce the DWF analysis.”

Concerning the data availability I think the following claim is misleading:
l.615-616: “Data from the EC-Earth3-HR historical and SSP2-4.5 simulations are available through any ESGF data node as part of CMIP6”
Isn’t this data from the HighResMIP version of the model (called EC-Earth3P-HR in this manuscript)? How and where is the data from the simulations analysed in this paper available?
Most ESFG nodes include data of the EC-Earth3P-HR simulations, but only a few of them include the EC-Earth3-HR used in this study.

We have updated this sentence for clarity:

> “Data from the EC-Earth3-HR historical and SSP2-4.5 simulations are available through any ESGF-CoG data node (e.g. https://esg-dn1.nsc.liu.se/search/cmip6-liu/) as part of the CMIP6 project. Search for source_id=“EC-Earth3-HR” and experiment_id=“historical” or “ssp245”. ”

Minor comments:
l.148: “the freshwater correction value was slightly modified to oas_mb_fluxcorr=1.07945”
Modified compared to what? The value is the same in the two models compared in Table 2. Also, a short explanation of what this parameter is would be useful. The name “oas_mb_fluxcorr” is not used in Doscher et al. (2022). Is this needed because IFS does not conserve water? Heat is also probably not conserved in IFS, is it also corrected for?
We did not change this parameter; the sentence and table have been corrected accordingly. oas_mb_fluxcorr is indeed needed to counteract the P-E imbalance of IFS.

l.160-162: How long is this “long concatenated pre-spin-up run”?
It spans more than 200 years, which has been added to the revised text.

Fig. 1a: Could you compare with longer reconstructions of TAS? For example, ERA5 is now available from 1940 while here only data from around 1980 is used.
Updated figure is added to the revised version.

l.243: “CAA” is not defined.
Corrected to Canadian Arctic Archipelago

l.321: “PrI” is not defined
Corrected to pre-industrial

Fig. 9b: Is that the density change at the surface? It would be good to mention it in the caption.
Added

l.412: It is not clear what “potential” means here. I would understand if “potential” was used for a claim about the real world but here it is about models.
We removed “potential” from the sentence.

l.464 and l.533: These correlation coefficients need more context. Are they computed on the yearly averaged data? Are the time series detrended? How? The trend is not linear. Are the correlations reflecting the co-variability at inter-annual time scale or a trend common to the time series?

To assess how strongly the long-term AMOC weakening co-evolves with regional DMV and DWF changes, we computed the correlations from the full, non-detrended time series. Our focus was therefore on trend-related co-variation rather than detrended variability. While these correlations largely reflect shared trends, they do not, by themselves, demonstrate physical coupling. To avoid overinterpretation, we removed the DMV correlations and only retained the DWF analysis in the revision. Our aim was not to imply a tight dynamical link, but rather to examine whether DWF diagnostics evolve consistently with the AMOC index under anthropogenic forcing.

Citation: https://doi.org/10.5194/egusphere-2025-2653-RC1

References

Döscher, R., Acosta, M., Alessandri, A., et al. (2022). The EC-Earth3 earth system model for the climate model intercomparison project 6. Geoscientific Model Development, 15, 2973–3020. https://doi.org/10.5194/gmd-15-2973-2022

Durack, Paul J.; Taylor, Karl E.; Ames, Sasha; Po-Chedley, Stephen; Mauzey, Christopher (2022). PCMDI AMIP SST and sea-ice boundary conditions version 1.1.8.Earth System Grid Federation. https://doi.org/10.22033/ESGF/input4MIPs.16921

Hansen, J., Sato, M., Kharecha, P., and von Schuckmann, K.: Earth's energy imbalance and implications, Atmos. Chem. Phys., 11, 13421–13449, https://doi.org/10.5194/acp-11-13421-2011, 2011.