

We thank both reviewers for their kind and constructive comments and suggestions. We revised the manuscript to address the reviewer comments. Below, we provide responses to the detailed comments of each reviewer. Reviewer comments are shown in black, while our responses are blue.

Reviewer 1

The manuscript presents a comprehensive analysis of the global mean surface anomalies observed in 2023 and 2024. It builds on previous studies that focussed on individual drivers of these anomalies by jointly estimating the contributions of anthropogenic activity, volcanic eruptions, 11-year cycle variability, El Nino Southern oscillation, Atlantic Multidecadal variability, Pacific Decadal Oscillation and Indian Ocean Dipole, using a multi linear regression energy balance model.

My only concern relates to section 2.1 and the Appendix, which would benefit from improvements in the presentation order and reproducibility of the results, as suggested in the specific comments. After these corrections, the manuscript is suitable for publication.

Thank you for the kind words! We revised Sect. 2.1 and the Appendix as suggested. The specific changes made are given below as responses to the individual comments.

Specific Comments:

87-88: the paragraph starts explaining the first term, then the time grid and then the other terms of equation 1. This sentence applies to all terms of the equation and breaks the flow of the description of each term of the equation. I would first mention the time grid and then each term.

We moved the sentence describing the time grid to the end of the paragraph that precedes Eq. 1.

89: equation 2 is briefly mentioned here before the explanation of equation 1 is complete.

We moved the sentence describing $\sigma_{OBS,i}$ to the end of the paragraph, just after the explanation of the individual terms of Eq. 1.

113-115: It is unclear whether AAWR represents the slope of dT_{ANTH} that is obtained from a linear fit to temperature anomalies shown in panel 1a or the slope is obtained from a linear fit to dT_{ANTH} . In the latter case, how is dT_{ANTH} retrieved? Please rephrase the sentence.

We made the following changes to increase the clarity of our explanation of ΔT_{ANTH} and AAWR. The sentence describing Eq. (1), which is the first mention of ΔT_{ANTH} , that had read:

This value is computed by the EBM component of our model from the magnitude of time-invariant climate feedback, the ERF of the climate due to GHGs, tropospheric aerosols and land-use change, as well as the export of heat to oceans.

now reads:

In Eq. (1), ΔT_{ANTH} corresponds to the change in GMST due to anthropogenic activity. This value is computed by the EBM component of our model, as described in Sect. 2.1.1 and Appendix A, from the magnitude of time-invariant climate feedback, the ERF of the climate due to GHGs, tropospheric aerosols and land-use change, as well as the export of heat to oceans.

We also revised the sentence on the original L113-115 that had read:

Following McBride et al., (2021), the rate of rise in GMST due to anthropogenic activity, termed the Attributable Anthropogenic Warming Rate (AAWR), is computed as the slope of a linear fit to ΔT_{ANTH} between 1975 and 2014 and is given in Fig. 1b.

to now read:

Following McBride et al., (2021), we use the slope of the linear fit to ΔT_{ANTH} (orange line in Fig. 1b) between 1975 and 2014 as the quantity describing the rate of rise in GMST due to anthropogenic activity, and term this metric Attributable Anthropogenic Warming Rate (AAWR).

119-120: The text says to “see caption” of figure 1, while the caption of Figure 1 similarly and repeatedly refers the reader back to the text (“see text”). In neither case is it clear what specific information the reader is expected or how it provides additional clarification. Please make these references explicit in the text and summarize the relevant information clearly in the caption of figure 1. Figure 1 contains a large amount of detail, and repeatedly moving back and forth between the text and the figure makes it difficult to follow the main point.

The original caption of Fig. 1 contained “see text” four times. We removed three of these four instances, and replaced one occurrence with a reference to a specific section of the paper. The sentence that had read:

The uncertainty in climate feedback or ERF_{AER} is accounted for using an ensemble method (see text).

now reads:

The uncertainty in climate feedback or ERF_{AER} is accounted for using an ensemble method (Sect. 2.1.2).

In these same lines: 1) it is unclear whether the term “single fit” refers to all panels in Figure 1 or only to panel a; 2) include the values of λ here as well, as is done in panel a, and refer to the section where you explained how you got this value; 3) does “the single time series of ERF_{aer} ” refer specifically to panel (b)?

1) Indeed, the term “single fit” refers to the entire of Fig. 1, where panels b-f represent the anthropogenic and natural components of this fit, panel (g) shows the evolution of ocean heat content for this fit, and the red line on panel (a) is the sum of the natural and anthropogenic components, that is, ΔT_{MDL} from Eq. (1).

To better highlight the above points, and to provide additional clarity for the readers regarding the above, we made the following changes:

i) We added the following sentence to the caption of Fig. 1, before the description of the individual panels:

Single fit to the observed GMST and its decomposition to anthropogenic and natural contributions.

ii) We rephrased the paragraph describing the figure. The sentence that had read:

Figure 1a shows a single modelled fit (red line) to the observed GMST record obtained from version 5 of the Hadley Centre Climatic Research Unit (HadCRUT5 (Morice et al., 2021), black), over 1850 to 2024.

now reads:

To illustrate the separation of ΔT_{MDL} to its anthropogenic and natural components that is described by Eq. (1), we show a single decomposition in Fig. 1. Panel a shows the modelled fit (ΔT_{MDL} , red line) to the observed GMST record obtained from version 5 of the Hadley Centre Climatic Research Unit (HadCRUT5 (Morice et al., 2021), black), over 1850 to 2024.

2-3) Fig. 1 shows a single fit to the GMST record, and its decomposition to natural and anthropogenic components (panels b-f). The value of λ_{Σ} is associated with this single fit, and the value of λ_{Σ} given in Fig. 1a is applicable to the entire figure. Therefore, we decided not to add the value of λ_{Σ} to avoid showing redundant information.

Similarly, the term “single time series of ERF_{AER} ” refers to the ERF_{AER} input that is used to simulate the anthropogenic component of GMST (ΔT_{ANTH} in Eq. (1), orange line in Fig. 1b) using the EBM component of the model as described in Sect. 2.1.1 and Appendix A.

We updated the contents of Sect. 2.1 to clarify the above:

i) The sentence that had read:

The single fit shown in Fig. 1 assumes a certain value of time-invariant climate feedback (λ_{Σ} , panel (a)) and a single time series of ERF_{AER} (see caption).

was changed to:

The single fit shown in Fig. 1 corresponds to a given level of time-invariant climate feedback (λ_{Σ} , panel (a)) and a single time series of ERF_{AER} . The specific time series that is used as the input for the creation of Fig. 1 is the best estimate for the temporal evolution of ERF_{AER} provided by Annex III of the IPCC AR6 report (IPCC, 2021a; Smith et al., 2021b), also shown by the solid black line in Fig. S2 of Farago et al., (2025b).

ii) The sentence that had read:

The uncertainty in the magnitude of climate feedback and ERF_{AER} is considered using an ensemble method (McBride et al., 2021; Farago et al., 2025b). For the ensemble, regressions are performed for 160,000 members as described in Sect. 2.1.2.

now reads:

To account for the uncertainty in the magnitude of climate feedbacks and aerosol cooling, the model performs the decomposition illustrated by Fig. 1 for a total of 160,000 ensemble members, as described in Sect. 2.1.2, where ensemble members differ in the strength of time-invariant climate feedback (λ_{Σ}) and the time series of ERF_{AER} .

134: While EBM acronym is introduced at line 60, the meaning of EBM-1 is not defined. It is only explained later in the Appendix (but not even at the beginning of the Appendix) and it is not straightforward what the number 1 means. A brief explanation should be provided here.

135: It only explains the temperature of the upper layer. What about the lower layer? You should also mention here how this approach is an improvement? I think that lines 591-600 belong here.

Thank you for these recommendations. We revised the corresponding section (Sect. 2.1.1) in response to the two comments above in the following manner: The contents of the original L591-600 were moved to Sect. 2.1.1, and merged with the existing text. We also added an additional sentence to introduce the EBM-1 and EBM- ϵ terms. The original Sect. 2.1.1. that had read:

In this section, we briefly summarize the updates to the energy balance component of the EM-GC model adapted for this paper, with additional details provided in Appendix A. To capture the short-term response to a sudden increase of ERF by a constant magnitude (hereafter termed step forcing), we employ the two-layer EBM formulation from Held et al. (2010). The two-layer EBM is sufficiently simple for use in reduced complexity climate models, and provides a temperature response under step forcing scenarios that is consistent with the response of Earth System Models (ESMs) (Geoffroy et al., 2013a; Geoffroy et al., 2013b; Gregory et al., 2015; Tsutsui and Smith, 2025). In this paper, we use the EBM-1 formulation described by Geoffroy et al., (2013b), and associate the temperature of the upper layer in the two-layer EBM with ΔT_{ANTH} in Eq. (1). The two-layer EBM approximation was also used extensively by the authors of the IPCC AR6 report (Sect. 7.SM.2 of Smith et al., (2021a)) in emulators calibrated using the output of CMIP6 models. Importantly, while two-layer EBMs are usually calibrated using CMIP model output, we use the observed rise in GMST and OHC for calibration, also described in Appendix A.

now reads:

In this section, we briefly summarize the updates to the energy balance component of the EM-GC model adapted for this paper, with additional details provided in Appendix A. One weakness of the representation of Ocean Heat Export (OHE) used in earlier versions of EM-GC (Canty et al., 2013; Hope et al., 2020; McBride et al., 2021; Farago et al., 2025b) is that the temperature response to a sudden increase of ERF by a constant magnitude (hereafter termed step forcing) leads to an immediate response of GMST, similar to the “deep-layer model” formulation described in Gregory et al., (2015). To better capture the short-term temperature response to a step forcing, such as that caused by the IMO2020 regulations, we employ the two-layer EBM formulation from Held et al. (2010) in this paper. The two-layer EBM is sufficiently simple for use in reduced complexity climate models, and provides a temperature response under step forcing scenarios that is consistent with the response of Earth System Models (ESMs) (Geoffroy et al., 2013a; Geoffroy et al., 2013b; Gregory et al., 2015; Tsutsui and Smith, 2025). The updated ocean module presented in Appendix A provides a more realistic short-term temperature response to sudden changes in ERF, relative to previous versions of the EM-GC model, which were primarily used to quantify the long-term response of GMST to changes in ERF under various Representative Concentration Pathway (RCP) (Canty et al., 2013; Mascioli et al., 2012; Hope et al., 2020) and Shared Socioeconomic Pathway (SSP) scenarios (McBride et al., 2021; Farago et al., 2025b).

The two-layer EBM can be implemented via two slightly different mathematical formulations, termed EBM-1 and EBM- ϵ (Geoffroy et al., 2013a), which differ in the

treatment of the efficacy of deep-ocean heat uptake (parameter ε), as described in Sect. A1 of Appendix A, and Sect. 2 of Geoffroy et al., (2013a). In this paper, we use the EBM-1 formulation described by Geoffroy et al., (2013b), because EBM-1 provides a highly similar temperature response to EBM- ε on the timescale of our simulations (Fig. 4 of Geoffroy et al., (2013a)), but requires the calibration of one less parameter. The temperature of the upper layer of this EBM formulation is commonly considered to be equal to the global mean surface temperature (e.g. Cummins et al., (2020)). Therefore, the temperature of the upper layer that is simulated by the EBM component of the model is used as ΔT_{ANTH} in Eq. (1).

The two-layer EBM approximation was also used extensively by the authors of the IPCC AR6 report (Sect. 7.SM.2 of Smith et al., (2021a)) in emulators calibrated using the output of CMIP6 models. Importantly, while two-layer EBMs are usually calibrated using CMIP model output, we use the observed rise in GMST and OHC for calibration, as described in Sect. A2 of Appendix A. Appendix A also describes the technical implementation of EBM-1 (Held et al., 2010; Geoffroy et al., 2013b) into the EM-GC model and the results of the benchmark simulations performed using the updated EBM component of our model.

279-280: same as comment for lines 113-115.

We revised Sect. 2.1 to provide better clarity regarding how AAWR is computed. Since the text in question refers to Sect. 2.1, this particular sentence was not changed.

284-285: This sentence is difficult to follow and would benefit from being rephrased.

We revised the sentence as follows. The sentence that had read:

Panels (b) and (c) differ in that, for the simulation shown in Fig. 2b, the high frequency component of the AMV input time series, that is frequencies greater than $1/9 \text{ yr}^{-1}$, was removed using a Fourier-filter (Sect. 2.2.5)

now reads:

The simulation shown in Fig. 2b uses an input time series for AMV in which frequencies greater than $1/9 \text{ yr}^{-1}$ were removed using a Fourier-filter, as described in Sect. 2.2.5.

286-288: Panel a should either be discussed first or moved to the last position in the figure. The discussion of Figure 2 begins at line 281 with panels b and c, which made me wonder whether I had missed something, as panel a is only discussed later.

We moved the sentence describing Fig. 2a, such that it is now in front of the sentences describing Fig. 2b-c.

303: EffCS is briefly mentioned here and in the introduction at lines 61-, where it is stated that “Our model [...] provides an estimate of Effective Climate Sensitivity (EffCS)”. The purpose of this sentence (L303) is unclear in its current form. The sentence should be expanded to provide substantive information about the EffCS estimate, as suggested in the Introduction, or it should be removed from this section.

We removed this sentence on the original L303. On L61, we added the specific values of EffCS.

326-328: It is unclear how the colors should be interpreted or how the probabilities were calculated. Additionally, the choice of colormap is confusing: the reds and blues appear to represent higher and lower probabilities, respectively, while white seems to indicate the most probable value, but it is actually the opposite.

We understand the confusion and have made changes, as described below, to clarify the interpretation of the color scale. The color scale represents the likelihood that the simulated GMST *exceeds* a certain threshold at a time. The white line (50th percentile) corresponds to the “central estimate” of the model, given that the likelihood of the GMST being lower than this value is the same as the likelihood of GMST being higher. To reflect these points, we expanded the section describing the figure.

We added the word “cumulative” to the original sentence describing the figure, and added additional content to the description of the figure.

The original description of the Fig. 3a-c (L319 in the original manuscript) that had read:

Colors denote the EM–GC simulated probability of the GMST being greater or equal than a given value at a time, as indicated by the color bar to the right.

now reads:

Colors denote the EM–GC simulated cumulative probability of the GMST being greater or equal than a given value at a time, as indicated by the color bar to the right. The probabilities are obtained from weighting the EM–GC ensemble using an asymmetrical Gaussian function (Fig. 2a) as summarized in Sect. 2.1.2, and detailed in McBride et al., (2021) and Farago et al., (2025b). The shading represents the uncertainty range of simulated GMST, with the white line corresponding to the 50th percentile of simulated GMST, while the red and blue regions are associated with ensemble members of high and low climate sensitivities, respectively.

394-400: can you provide a value for the respective change in TSI that corresponds to the change in GMST?

We added the following sentence to the text:

In 2023 and 2024, the annual mean TSI was 0.35 and 0.40 W m⁻² higher than in 2022, respectively.

519-525: This paragraph discusses the anthropogenic warming rate, which was already covered in Section 3.2. Its placement in the section on the Indian Ocean Dipole feels abrupt; it would be more appropriately included in the discussion in Section 3.2.

Conclusion: The conclusion presented here is weaker than the one in Section 3.4 (lines 523–525). Consider expanding this section by incorporating the points made in those lines.

We address these two comments together. We agree that the inclusion of warming rates in Sect. 3.4 is better suited elsewhere, and that the sentences on the original L523-525 would be more suitable for the Conclusions section. The following change was made:

We moved the contents of the final paragraph of Sect. 3.4 (L519-525 in the original manuscript) to the Conclusions section (L535 and L538 in the original manuscript).

Appendix: Please provide the values or a range of values that have been used throughout the appendix to allow reproducibility of the results (gamma L 655; value used in equation A8, only few values are defined at L675-677; initial value of gamma used in L705, three values of C_u and C_d at L726-729).

1) The value of γ on the original L655 and in Eq. (A8) does not refer to a specific number, but rather, the notion that once model quantifies the value of γ as described in Sect. A2, the time series of T_u and T_d can be computed for a given ensemble member. We updated the sentence in question that had read:

Consequently, for each ensemble member, a value of the parameter γ defines a pair of T_u - T_d time series.

to now read:

For each ensemble member, the value of the parameter γ is quantified based on the observed rise in OHC (Sect. A2). Using these values for γ , the model then computes time series for T_u and T_d .

We made the following, additional changes to the Appendix:

2) We added the specific values of C_u and C_d to the original L675-677, via the following, new sentences:

For the simulations presented in the main paper, we use values of 7.3 and 106 $W\ yr\ m^{-2}\ K^{-1}$ for C_u and C_d , respectively, which correspond to equivalent ocean depths of 77 and 1105 meters. These depths are the multi-model mean values, obtained by Geoffroy et al., (2013b) by fitting the two-layer EBM to the temperature output of a set of CMIP5 models. We describe the testing of the sensitivity of our model setup to the specific values of C_u and C_d in Sect. A3.

3) We added the initial value of γ (original L705)

4) We added the values of C_u and C_d (original L726-729)

Supplementary, L51: Define the latitudinal and longitudinal boundaries of the four regions to ensure data reproducibility

We added the boundaries of the four regions to the text of the Supplement (original L53).

Technical corrections:

118: remove the brackets: (gamma, Geoffroy et al., 2013)

Thank you, we made the change.

Reviewer 2

Farago et al. use multiple linear regression to attribute the exceptionally high global mean surface temperatures observed in 2023 and 2024, combining an empirical energy-balance framework with time series of multiple external forcings and modes of internal variability. They constrain the regression by historical observations from 1850 onward based on suitable datasets. They find that the extreme warmth of 2023/24 is largely explained by the superposition of long-term anthropogenic warming with strong interannual variability, including ENSO, but the study also highlights a large contribution associated with the positive phase of the Indian Ocean Dipole (IOD) in 2023/24. The work provides a timely and clearly structured quantitative synthesis of drivers behind the recent record temperatures. I think that some aspects could benefit from clarification, but overall I think that this work is a valuable contribution to the discussion around the record-high recent temperatures and their attribution to various drivers, with one key strength of the study being the comprehensive inclusion of numerous external and internal drivers in a common quantitative framework. I recommend publication subject to minor revisions.

Thank you for these very kind words. We sincerely appreciate the constructive comments of this review.

###

Before providing specific comments, I'd like to highlight three aspects that remain somewhat unclear to me:

(i) The large contribution of the IOD is surprising and also leaves some questions open about its (in)dependence from ENSO. If I am not mistaken, the IOD has thus far not been considered a significant contributor to past (annual) GMST anomalies - an amplitude around values typical for ENSO is astonishing. Can you provide an estimate for how much of the unforced GMST variability the index might explain before 2023, e.g., in terms of R-squared, and, more generally, provide more explanation why this large contribution in 2023/24 might be plausible, possibly including some hints at mechanism involved?

We added two new paragraphs to Sect 3.4, alongside an additional supplementary figure (new Fig. S10) that highlight strength of the change in IOD from 2022 to 2023. We also added an additional paragraph to the Supplement, that describes the figure.

The new paragraphs that we added to Sect. 3.4 reads:

Finally, we compare the IOD-related rise in annual mean GMST from 2022 to 2023 to IOD events of the past. Over the 1851–2019 time period, IOD explains only a small fraction of the variability in GMST ($R^2 = 0.027$, Fig. S1a). However, the rise in IOD between 2022 and 2023 is nearly unprecedented since 1850. Figure S10 shows a scatterplot of the rise in the annual mean value of the IOD index relative to the previous year, computed from two separate IOD indices that differ in the underlying SST datasets: the COBE2 IOD index used for the simulations of this paper (horizontal axis), and the NOAA DMI IOD index (vertical axis), which is based on HadISST SSTs (Sect. 2.2.3). The shaded areas in Fig. S10 represent the highest 5th percentile values of these IOD anomalies. Only three years fall into the 5th percentile of both datasets: 1961, 2017, and 2023. The rise in IOD from 2022 to 2023 is the highest on record for the COBE2 index, and the second highest in the NOAA DMI index. To test the sensitivity of our conclusions to the choice of IOD dataset, we performed a model

simulation that uses the NOAA DMI index as the input for IOD. The contributions of natural and anthropogenic factors to the GMST obtained from this simulation are given in Table S1, in a manner similar to Table 2. Using the NOAA DMI index as the input for IOD, the rise in GMST from 2022 to 2023 that is associated with IOD was found to be 0.072 °C [0.004 to 0.105 °C, 5–95% range], which is highly similar to the values found using the COBE2 index of 0.075 °C [0.036 to 0.096 °C]. Consequently, our conclusion regarding the contribution of IOD to the rise in GMST between 2022 and 2023 is largely independent of the choice of IOD dataset. Additional analysis of the contributions of other factors for the model run that is based on the NOAA DMI Index is provided in the Supplement.

Figure S10 highlights the nearly unprecedented rise in IOD from 2022 to 2023. Understanding the mechanisms of such a strong shift in the IOD is beyond the scope of this paper. However, the amplifying effect of a co-occurring strong El Niño event may have been a strong contributor to the conditions in the Indian Ocean that led to the development of this pIOD event, as suggested by Xie et al., (2025). In addition, as described above, the spatial pattern of surface temperature anomalies observed in 2023 resembles patterns similar to those associated with strong IOD events over the past half century.

(ii) AMV is considered in two flavours, once with and once without high-frequency variability. However, with the "M" in AMV standing for multi-decadal, is it at all reasonable to consider a monthly AMV index with a strong high-frequency component?

Indeed, AMV is used with and without the high-frequency component. We have decided to retain the "M" in both definitions of AMV, because this convention is consistent with our earlier paper (Sect. 3.2.3 of Canty et al., 2013). Furthermore, unfiltered AMV index shown in Fig. 1 still has a significant multidecadal component.

(iii) The IMO-related aerosol effect on GMST in 2023/24 is determined to a high degree by the choice of the corresponding ERF, for which two values are assumed, based approximately on existing estimates. The EBM parameters somewhat influence the temperature response, but it's still strongly determined by the ERF. Please check related formulations to avoid the impression that the new estimates provided here might be more independent than they actually are from the "educated guesses" about the ERF.

We agree, and we made following change to the manuscript:

The Conclusions section was adjusted; the sentences that had read:

We find that the IMO2020 regulations are responsible for an increase in GMST (ΔT_{IMO}) of 0.028 °C [0.025 to 0.031 °C, 5–95% range] and 0.043 °C [0.038 to 0.046 °C] from the start of 2020 to the end of 2024, for increases in ERF due to IMO2020 of +0.1 and +0.15 W m⁻², respectively.

now read:

We simulate the response of GMST to increases in ERF due to IMO2020 ($\Delta \text{ERF}_{\text{IMO}}$) of +0.1 and +0.15 W m⁻², which were chosen to be close to recent estimates of $\Delta \text{ERF}_{\text{IMO}}$ available in literature. We find that the IMO2020 regulations are responsible for an increase in GMST (ΔT_{IMO}) of 0.028 °C [0.025 to 0.031 °C, 5–95% range] and 0.043 °C

[0.038 to 0.046 °C] from the start of 2020 to the end of 2024, for these two cases, respectively.

###

Specific comments (including minor technical ones):

Title: "Factors" -> "factors"

We made the change.

First paragraph of introduction: When it comes to the role of internal variability in 2023/24 temperatures, you mention the Raghuraman et al. 2024 paper, but I think you should also mention the Terhaar et al. 2025 paper.

We added citations to both Terhaar et al., 2025, as well as Blanchard-Wrigglesworth et al., 2025.

L47/48: Actually global SO₂ emissions have been declining already since around 1980.

We reworded the sentence. The sentence that had read:

Efforts to improve air quality have resulted in a gradual reduction of sulfur emissions in the recent two decades.

now reads:

Efforts to improve air quality have resulted in a gradual reduction of sulfur emissions since the 1980s.

L57/58: It's a bit strange to point to the table with recent estimates of ERF_{AER} here without mentioning any of them explicitly, for example their range.

We added an additional sentence that specifies the range of ERF_{AER} estimates:

Most of these recent studies estimated the increase in ERF_{AER} due to IMO2020 to be in the range of 0.06 to 0.2 W m⁻².

L62/63: Also here I'd consider it helpful to mention the EffCS explicitly.

We added the values of EffCS to this sentence. The sentence that had read:

Our model is trained on 170 years of historical climate data from various measurements, and provides an estimate of Effective Climate Sensitivity (EffCS) that is consistent with recent literature values (Farago et al., 2025b).

now reads:

The model is trained on 170 years of historical climate data from various measurements and provides an estimate of Effective Climate Sensitivity (EffCS) of 2.63 °C [1.77 to 3.55 °C, 5–95% range], which is consistent with recent literature values of EffCS described in Farago et al., (2025b).

L89: Do I understand correctly that sigma_OBS is large in the earlier and small in the later part of the historical record, and that the formulation thereby simply makes sure that the more uncertain past temperature observations are more weakly constraining the model parameters than the more certain recent observations? If so, maybe that can be mentioned explicitly.

Indeed, this is the case. We added the following sentence to the paper to highlight this point:

The use of $\sigma_{\text{OBS},j}$ in Eq. (2) ensures that past temperature observations that correspond to higher observational uncertainty act as weaker constraints than more recent observations.

Fig. 1: Why is a single ensemble member shown here instead of the ensemble mean? If I understand it correctly, the ensemble members considered here sample parameter uncertainties, not different realizations of internal variability (like ensemble members of a climate model in CMIP), so also the ensemble mean here should show similar variability, right? Why would that not be more suitable to show?

Indeed, the model samples the uncertainty in time-invariant climate feedbacks (parameter λ_{Σ}) and aerosol cooling (the time series of ERF_{AER}). The purpose of showing a single ensemble member is to provide an illustration to the reader of how the model separates natural and anthropogenic influences based on Eq. (1). We felt that providing such an illustration is important for the understanding of how the model works, prior to the introduction of the ensemble approach, similarly to previous papers that used EM-GC, such as McBride et al., (2021).

Additionally, the single fit shown by our Fig. 1 is the ensemble member where the input time series of ERF_{AER} is the same as the best estimate for the time series of ERF_{AER} given by Chapter 7 and Annex III of IPCC AR6, and λ_{Σ} is the value of climate feedbacks that yields the lowest χ^2_{ATM} for this time series of ERF_{AER} .

To better highlight these points, we made the following changes to Sect. 2.1:

i) We rephrased the paragraph describing the figure. The sentence that had read:

Figure 1a shows a single modelled fit (red line) to the observed GMST record obtained from version 5 of the Hadley Centre Climatic Research Unit (HadCRUT5 (Morice et al., 2021), black), over 1850 to 2024.

now reads:

To illustrate the separation of ΔT_{MDL} to its anthropogenic and natural components that is described by Eq. (1), we show a single decomposition in Fig. 1. Panel a shows the modelled fit (ΔT_{MDL} , red line) to the observed GMST record obtained from version 5 of the Hadley Centre Climatic Research Unit (HadCRUT5 (Morice et al., 2021), black), over 1850 to 2024.

ii) The sentence that had read:

The single fit shown in Fig. 1 assumes a certain value of time-invariant climate feedback (λ_{Σ} , panel (a)) and a single time series of ERF_{AER} (see caption).

was changed to:

The single fit shown in Fig. 1 corresponds to a given level of time-invariant climate feedback (λ_{Σ} , panel (a)) and a single time series of ERF_{AER} . The specific time series that is used as the input for the creation of Fig. 1 is the best estimate for the temporal evolution of ERF_{AER} provided by Annex III of the IPCC AR6 report (IPCC, 2021a; Smith et al., 2021b)

Eq. 3: Why is there a sum (capital sigma) sign in front of $ERF_GHG(t)$? Is that not redundant with the plus signs, given that there's only one $ERF_GHG(t)$ time series, if I'm not mistaken?

We agree, and we removed the sum sign.

L161: If you mention the -1.1W/m^2 central estimate, I recommend to mention the left and right sigma values of the asymm. Gaussian, too, to define the weighing function completely.

We added the following sentence to provide this information:

Our 1σ (-0.75 and -1.4 W m^{-2}) and 2σ (-0.4 and -1.7 W m^{-2}) bounds of this Gaussian function are based on the “very likely” range for ERF_{AER} (-0.4 to -1.7 W m^{-2}) that is provided by Chapter 7 of AR6 (Forster et al., 2021).

L207: Consider quantifying what is meant by "insensitive".

We added additional detail to the paper to provide this quantification. We added a new paragraph to Sect. 3.4 that describes the contribution of IOD to the GMST anomalies, when the NOAA DMI IOD index is used as the input for IOD, instead of the COBE2 index. The following sentences were added to Sect. 3.4 to describe the sensitivity of our conclusions to the choice of the IOD dataset.

To test the sensitivity of our conclusions to the choice of IOD dataset, we performed a model simulation that uses the NOAA DMI index as the input for IOD. The contributions of natural and anthropogenic factors to the GMST obtained from this simulation are given in Table S1, in a manner similar to Table 2. Using the NOAA DMI index as the input for IOD, the rise in GMST from 2022 to 2023 that is associated with IOD was found to be $0.072\text{ }^\circ\text{C}$ [0.004 to $0.105\text{ }^\circ\text{C}$, 5–95% range], which is highly similar to the values found using the COBE2 index of $0.075\text{ }^\circ\text{C}$ [0.036 to $0.096\text{ }^\circ\text{C}$]. Consequently, our conclusion regarding the contribution of IOD to the rise in GMST between 2022 and 2023 is largely independent of the choice of IOD dataset. Additional analysis of the contributions of other factors for the model run that is based on the NOAA DMI Index is provided in the Supplement.

In addition, we added a new Supplementary table (Table S1), which shows the 50th percentile contributions to GMST in a manner similar to Table 2, for a simulation that used the NOAA DMI IOD index as the input for IOD. We also added a brief analysis of the values given in this table to our supplement.

Finally, on the original L207, we added a reference to Sect. 3.4. The sentence that had read:

Our findings regarding the contribution of the IOD to the anomalously high GMST in 2023 is insensitive to the use of the NOAA Dipole Mode Index (DMI) (Saji and Yamagata, 2003) for IOD, which is based on the HadISST1.1 SST dataset (Rayner et al., 2003).

now reads:

As described in Sect. 3.4, our findings regarding the contribution of the IOD to the anomalously high GMST in 2023 are insensitive to the use of the NOAA Dipole Mode Index (DMI) (Saji and Yamagata, 2003) for IOD, which is based on the HadISST1.1 SST dataset (Rayner et al., 2003).

L220-221: Misleading formulation: The CMIP6 TSI time series certainly covers more than until 1978, you just use it only up until then.

We agree, and made the following change. The sentence that had read:

We use this time series between 1979 and the end of 2024 to create a time series of monthly average TSI, which we append to the CMIP6 TSI input time series from Matthes et al., (2017), that covers 1850 to 1978.

now reads:

We use this time series between 1979 and the end of 2024 to create a time series of monthly average TSI, which we append to the 1850 to 1978 subset of the CMIP6 TSI input time series from Matthes et al., (2017).

L241: "Given that the net effect of the eruption of Hunga on GMST is small, we chose to use SAOD as a proxy for the impact of the Hunga volcano on GMST, while neglecting the additional radiative forcing from the injection of stratospheric water vapor."; The first part is per se not a good justification of the second part, given that a small net effect could in principle result from two large counteracting contributions. I think the assumption you make is OK, but the formulation here is misleading, so that should be clarified. Related, does the omission of the warming effect of the stratospheric water vapour not affect the "explanation gap"? Specifically, could the "Residual" in Tab. 2 be partly explained by the omission of that warming component of the HTHH eruption?

We agree, and clarified the statement written on the original L241 by removing the first part of the sentence. The sentence that had read:

Given that the net effect of the eruption of Hunga on GMST is small, we chose to use SAOD as a proxy for the impact of the Hunga volcano on GMST, while neglecting the additional radiative forcing from the injection of stratospheric water vapor.

now reads:

We chose to use SAOD as a proxy for the impact of the Hunga volcano on GMST, while neglecting the additional radiative forcing from the injection of stratospheric water vapor.

In addition, we added additional content to this section:

Further, the recently published, comprehensive report entitled The Hunga Volcanic Eruption Atmospheric Impacts Report, hereafter Hunga-report (APARC, 2025), suggested a Hunga-induced cooling of about -0.05 °C during 2022–2023. Our simulations, which are based on a proxy that omits the temperature effects of stratospheric water vapor, estimate a weaker cooling effect than given by the Hunga-report. Consequently, the use of our proxy does not result in an overestimation of Hunga-induced cooling, and the small residual between the modelled and observed GMST shown in the Results section are unlikely to be related to the use of our proxy for the effects of Hunga. Finally, Sects. 7.3 and 7.4 of the Hunga-report suggested that the surface temperature response prior to 2024 was dominated by the effects of stratospheric aerosols over water vapor, reinforcing our assumption for the use of SAOD as a proxy for the Hunga eruption.

Fig. 3+4: I am wondering how coherent the positioning of different ensemble members would be within the "uncertainty ranges" denoted in the upper panels: E.g., would a member with relatively large anomalies in 2019 also exhibit relatively large anomalies in 2024, that is, does the parameter sampling largely reflect the long-term behaviour or also (or more) the amplitude of short-term fluctuations, so that ensemble members would fluctuate more around the ensemble mean? Maybe an additional figure in the supplement could clarify that?

Indeed, the shading does correspond to long-term behavior, as ensemble members of high climate sensitivity are clustered in the higher-warming part of the uncertainty range (red region), while low climate sensitivity members are on the lower part (blue region) of the uncertainty range. We added additional details to the description of the figure to better highlight the details above.

The original description of the Fig. 3a-c (L319 in the original manuscript) that had read:

Colors denote the EM–GC simulated probability of the GMST being greater or equal than a given value at a time, as indicated by the color bar to the right.

now reads:

Colors denote the EM–GC simulated cumulative probability of the GMST being greater or equal than a given value at a time, as indicated by the color bar to the right. The probabilities are obtained from weighting the EM–GC ensemble using an asymmetrical Gaussian function (Fig. 2a) as summarized in Sect. 2.1.2, and detailed in McBride et al., (2021) and Farago et al., (2025b). The shading represents the uncertainty range of simulated GMST, with the white line corresponding to the 50th percentile of simulated GMST, while the red and blue regions are associated with ensemble members of high and low climate sensitivities, respectively.

L377-380: Also here, it would be useful to consider the Terhaar et al. (2025) paper.

We agree. We added the following sentence to this paragraph:

Our conclusions regarding the influence of ENSO on the 2023 GMST anomalies align much better with those of Terhaar et al., (2025), who suggested that a strong El Niño event is a necessary, but not sufficient condition for the development of temperature spikes similar to that observed in 2023.

L475/476: "whether this variability was influenced substantially by the IMO2020 regulations remains unclear"; consider reformulation, given that the permanent IMO regulations would not really "influence variability" (their effect could just be mistaken for variability).

We reworded the sentence. The sentence that had read:

Our results suggest that short-term variability in the North Atlantic SSTs was responsible for a portion of the observed rise in GMST between 2022 and 2023, but whether this variability was influenced substantially by the IMO2020 regulations remains unclear.

now reads:

Our results suggest that short-term variability in the North Atlantic SSTs was responsible for a portion of the observed rise in GMST between 2022 and 2023, but

whether the effects of the IMO2020 regulations could be mistaken for variability remains unclear.

L479: "During the positive phase of IOD" -> I suggest "During positive phases of IOD"

We made the change.

L503: It is unclear in what way / why Swapna et al. is cited here. (I'm not necessarily suggesting to remove the reference, it should only be clear what it's about / what it is stating).

We decided to remove this citation.

Tab. 2: Is there a particular reason why "Non-IMO anthropogenic" is lumped together, instead of showing its components (GHG, AER, TSI, ...)?

We group together the Non-IMO anthropogenic (GHG, AER, LUC) components, as this sum provides a quantification of long-term anthropogenic warming trends that we consider useful for policymaking purposes. We decided to make no changes here.

L532: I would avoid the term "projections" here, making one think of future scenarios...

We agree. We replaced the term "projections" with "estimates".

###

References:

Terhaar, J., Burger, F.A., Vogt, L. et al. Record sea surface temperature jump in 2023–2024 unlikely but not unexpected. *Nature* 639, 942–946 (2025). <https://doi.org/10.1038/s41586-025-08674-z>

References:

Blanchard-Wrigglesworth, E., Bilbao, R., Donohoe, A., and Materia, S.: Record Warmth of 2023 and 2024 was Highly Predictable and Resulted From ENSO Transition and Northern Hemisphere Absorbed Shortwave Anomalies, *Geophysical Research Letters*, 52, e2025GL115614, <https://doi.org/10.1029/2025GL115614>, 2025.

Canty, T., Mascioli, N. R., Smarte, M. D., and Salawitch, R. J.: An empirical model of global climate-Part 1: A critical evaluation of volcanic cooling, *Atmospheric Chemistry and Physics*, 13, 10.5194/acp-13-3997-2013, 2013.

McBride, L. A., Hope, A. P., Canty, T. P., Bennett, B. F., Tribett, W. R., and Salawitch, R. J.: Comparison of CMIP6 historical climate simulations and future projected warming to an empirical model of global climate, *Earth System Dynamics*, 12, 10.5194/esd-12-545-2021, 2021.