Interactive comment on “The Fractional Energy Balance Equation for Climate projections through 2100” by Roman Procyk et. al.

Roman Procyk et. al.

Figures 9 to 11 show the current and future evolutions of the global temperature as compared with the observations and the CMIP5 simulations. These projections are provided as ensembles in order to cover the range of uncertainties related to the model settings. An observational uncertainty is also provided. It is obvious that the spread of the ensemble is much smaller for the FEBE model than for the CMIP5 ensemble in Figure 9, and very often the observation band is not falling into the range of predicted projections of FEBE, while it is the case for the CMIP5. This suggests that the projections are unreliable in the sense that it does not cover all possible situations that can be observed. In weather and climate sciences, this aspect is key when making forecasts, predictions and projections using ensembles. This unreliability should be first acknowledged in the manuscript. Second this unreliability is maybe related to the use of a weak stochastic forcing. Some experiments with larger stochastic forcing would be desirable in order to clarify the under-dispersion of the ensemble. Stronger stochastic forcing could also maybe lead to larger climate sensitivity. This should be checked too.

Author: The basic purpose of Figures 9 to 11 was to compare the FEBE and GCM projections. In both cases, the projections are deterministic but with uncertainty limits due to their respective model uncertainties. Both yielded an estimate of the forced response but with qualitatively different uncertainty bounds. In the case of GCMs, the uncertainty is termed “structural” while for the FEBE it is parametric uncertainty. Unlike a probabilistic forecast, the results cannot be interpreted with the help of stochastic forecast notions such as reliability.

We can see why the referee may have misunderstood, since we compared the forced and internal components with the observed temperature series. This was intended as a quick visual validation of the forced component but since it contains the (stochastic) internal variability, strictly speaking, it should not be directly compared to the forced component.

In order to make a proper comparison with data, we first removed the observations from the figures so as to only compare the forced component (new figs. 10, 11 to replace the original figs. 10, 11). All the forced components are deterministic (the internal variability has been averaged out), for the GCMs the uncertainty is structural uncertainty, while for the FEBE it is parametric uncertainty. We then replaced the original figure 9 with one that represents the ensemble average over all the parametric and internal variability of the FEBE and the mean observational temperature (shown in new figs. 9a, 9b). This was achieved since the statistical dependence of the internal forcing and the parametric uncertainty are independent: the errors therefore add in quadrature. For this, it is sufficient to take the globally averaged yearly temperature anomaly ($\approx \pm 0.11C$) and combined it with the annual resolution parametric forced component from figs. 10, 11 over the historical period (1880-2020). The new figs. 9a,9b shows this result. This can also be done at monthly resolution following the same procedure but using the globally averaged monthly temperature anomaly ($\approx \pm 0.14C$), shown in figs. 9c, 9d.

The temperature observations do indeed fall within the 90% confidence limits of the FEBE historical reconstruction (i.e. the ensemble average of the response to both internal and external forcing). In both figures at annual resolution shown below (figs. 9a,9b), the historical mean temperature (red) is within the 90% CI of the FEBE forced response (with internal variability added) 92% of the years using the RCP scenario, and 94% using the SSP scenario. At the monthly resolution shown in figs. 9c, 9d, the historical mean temperature (red) is within the 90% CI of the FEBE forced response (with internal variability added) 90% of the months using the RCP scenario or the SSP scenario. The uncertainty is therefore compatible with the data. This is not the same as the reliability but it is an analogous validation of probabilistic aspects of the projection.

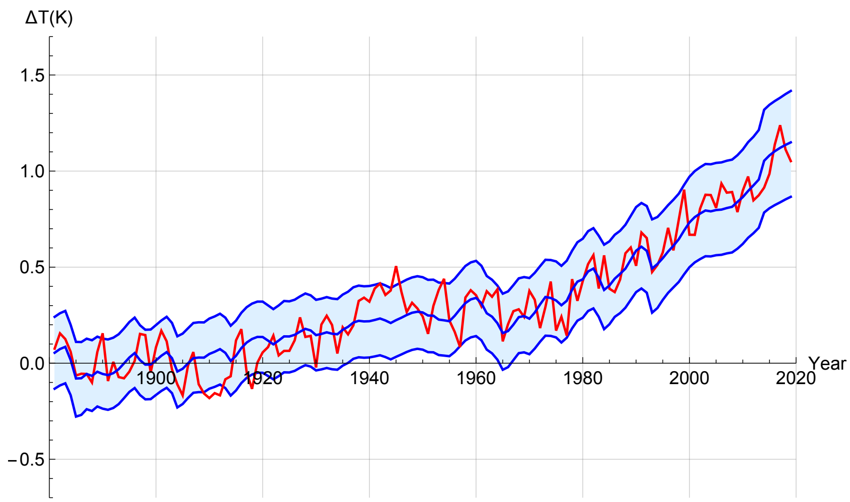

Fig. 9a: The historical reconstruction (forced temperature response and internal variability) of the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue) alongside mean of 5 observational temperature series (red) at yearly resolution; 90% CI (due to parametric uncertainty and internal variability) are indicated (shaded).

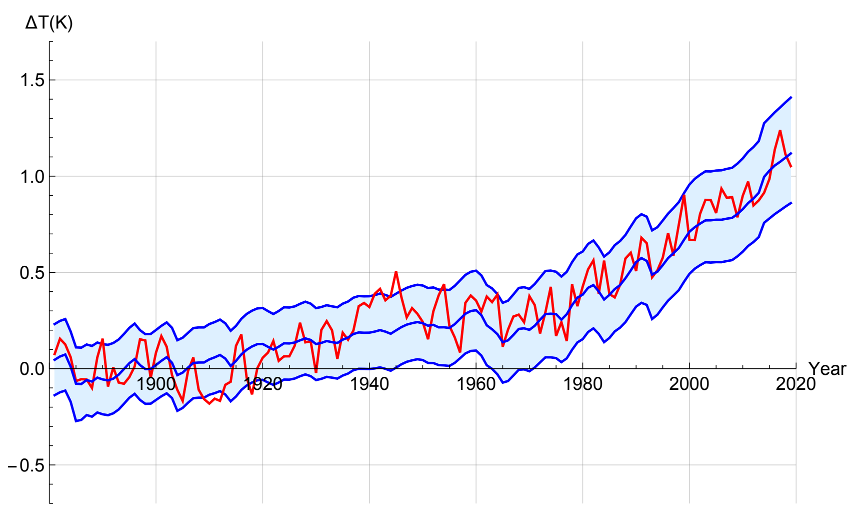

Fig. 9b: Same as fig. 9a except using $F_{Aer_{SSP}}$ parameters and forcing.

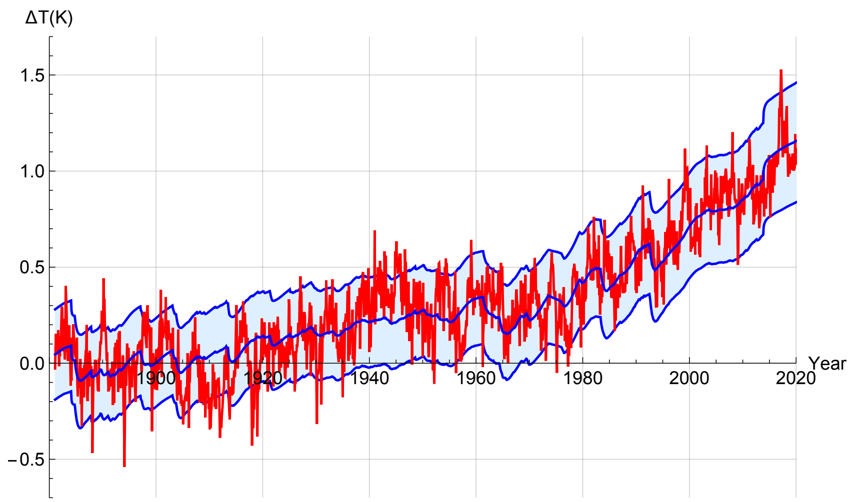

Fig. 9c: The historical reconstruction (forced temperature response and internal variability) of the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue) alongside mean of 5 observational temperature series (red) at monthly resolution; 90% CI (due to parametric uncertainty and internal variability) are indicated (shaded).

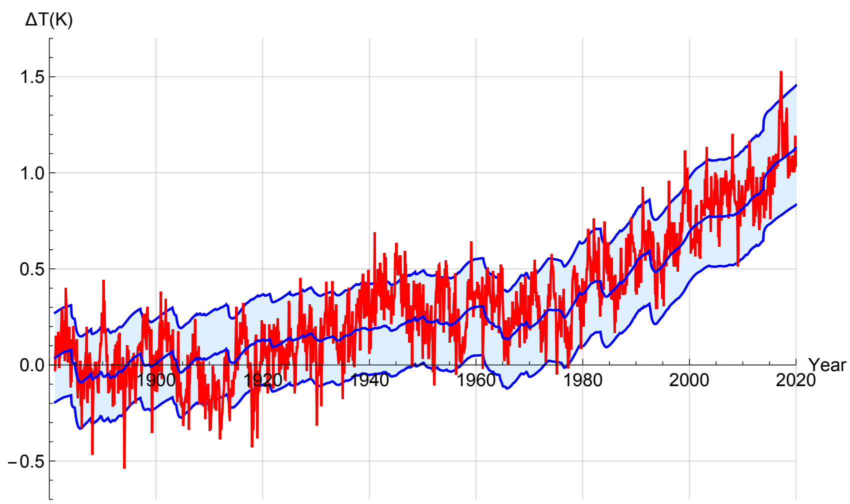

Fig. 9d: Same as fig. 9c except using $F_{Aer_{SSP}}$ parameters and forcing.

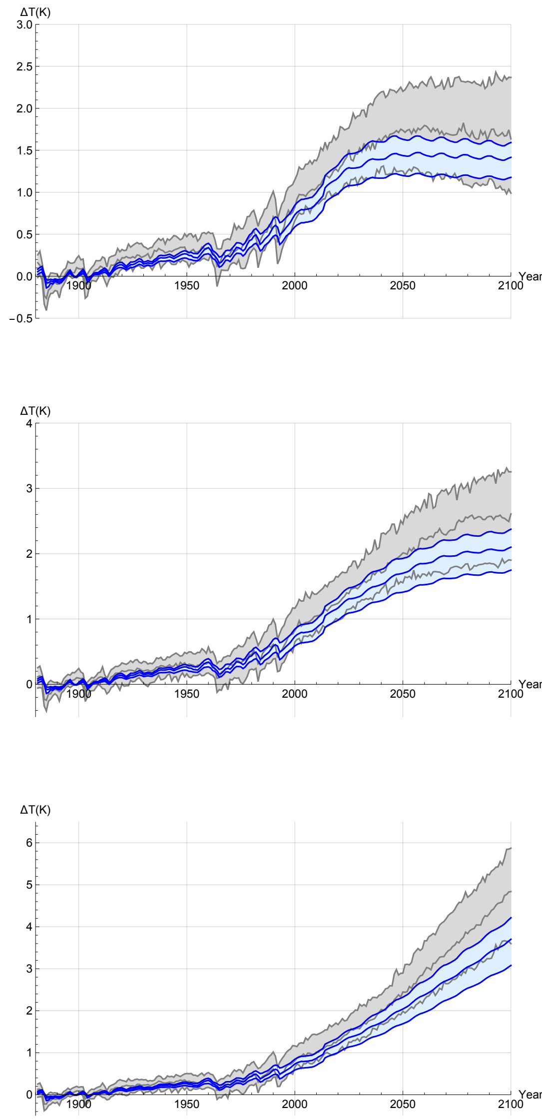

Fig. 10: The deterministic forced temperature response projected using the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue) compared with the CMIP5 MME projection (black); 90% CI from the parametric uncertainty are indicated (shaded). The projections until 2100, for RCP 2.6 (top), RCP 4.5 (middle) and RCP 8.5 (bottom), are shown.

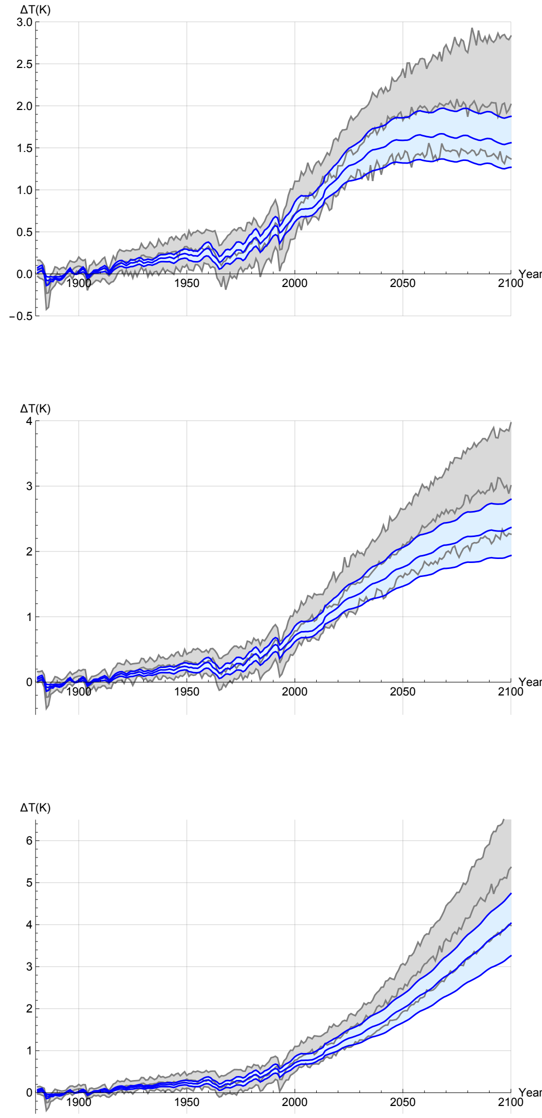

Fig. 11: The deterministic forced temperature response projected using the FEBE, with parameters calibrated using $F_{Aer_{SSP}}$ (blue) compared with the CMIP6 MME projection (black); 90% CI from the parametric uncertainty are indicated (shaded). The projections until 2100, for SSP 126 (top), SSP 245 (middle) and SSP 585 (bottom), are shown.

At page 20, line 421, the authors claim that the hiatus is better represented in the FEBE than in the CMIP5 ensemble. Well to me this is not true as the observations are most of the time out of the range of FEBE. The CMIP5 looks better at capturing the observations. So I suggest to modify these comments and try to be more objective in the comparison, maybe by using measures of reliability.

Author: We discussed the reliability in the previous responses, by showing the observed temperature series along with the projection of the forced response plus the internal variability response. With the internal variability we expect the data to lie within the 90% CI, 90% of the time. We also discussed how historical reconstructions were made (figs. 9a, 9b, 9c, 9d). We now discuss the latter over the hiatus period. In the insets of figs. 1a, 1b, we show the comparison of the forced median FEBE projection (blue) (using $F_{Aer_{RCP}}$ in fig. 1a, and $F_{Aer_{SSP}}$ in fig. 1b), the CMIP5/6 MME median (black) and the mean of 5 observational temperature series (red) with the 90% CI over the historical period with the inset showing the hiatus period (a blow-up of 1998-2015). We see that indeed, the FEBE median forced component in both cases captures the hiatus rather accurately (see Lovejoy (2015) for a stochastic forecast with a similar high frequency limit).

A quantitative comparison between the amount of time the FEBE median response is within the bounds of the observational temperature series 90% CI and the same for the CMIP5/6 MME was performed at the annual resolution data. The amount of time the median FEBE forced component using $F_{Aer_{RCP}}$ is within the 90% CI of the observational temperature series over the whole historic period is 47% and over the hiatus is 70% in comparison to the CMIP5 MME median which is within the whole historic period 39% and over the hiatus is 17%. When using the median FEBE forced component using $F_{Aer_{SSP}}$ similar results are found, over the whole period: 45% and over the hiatus: 35%, in comparison to the CMIP6 MME median which is within the whole historic period 39% and over the hiatus is 30%. In can be seen in both cases that the CMIP MME is generally warmer than the FEBE forced component notably over the period of the hiatus.

Lovejoy, S. (2015), Using scaling for macroweather forecasting including the pause, Geophys. Res. Lett., 42, 7148– 7155, doi:10.1002/2015GL065665.

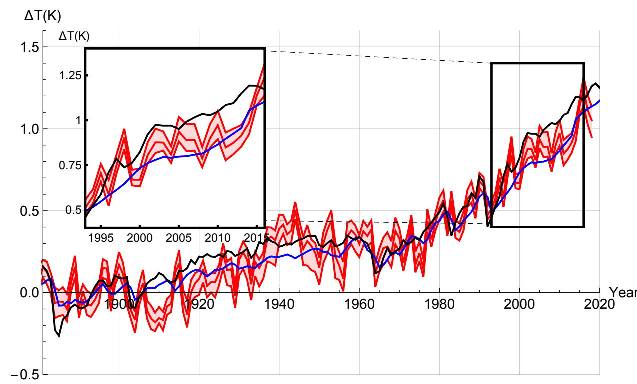

Fig. 1a: The median historical forced component of the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue), and the median of the CMIP5 MME (black) alongside mean of 5 observational temperature series (red) with their 90% CI indicated (shaded).

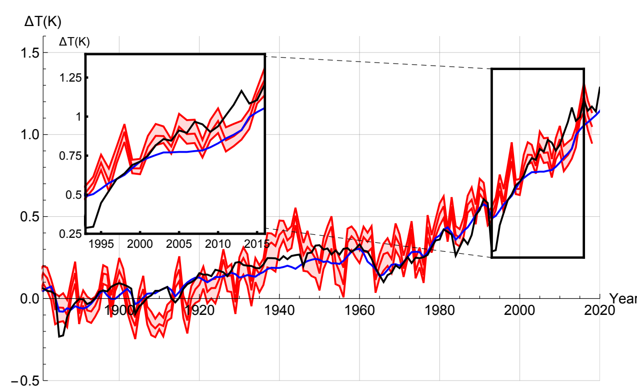

Fig. 1b: The median historical forced component of the FEBE, with parameters calibrated using $F_{Aer_{SSP}}$ (blue), and the median of the CMIP6 MME (black) alongside mean of 5 observational temperature series (red) with the 90% CI indicated (shaded).

One key conclusion of the manuscript is the lower sensitivity of the FEBE model on long time scales as shown for instance in Figure 10. What is happening when the FEBE model is fitted with a stronger noise that could maybe help in increasing the spread of the ensemble and its reliability? The results and the conclusions should then be revisited once these experiments are done and the impact of the stochastic noise clarified.

Author: The the amplitude of the noise is not an adjustable parameter, it is determined from the data. Although the empirical amplitude of the internal variability does not alter the projections of the forced temperature response, it does affect the uncertainties of the parameters (once these have been estimated, the resulting projection is purely deterministic).

Equation 14: Please clarify what is the variance of gamma(t)

Author: The variance of $\gamma(t)$ is the amplitude of the internal forcing assumed to be a Gaussian white noise. The internal variability of the observational temperature is equal to the observed series with the forced temperature response removed. If we take the global annually averaged monthly temperature anomaly to be $\sigma_{T,\tau_r} \approx \pm 0.14C$, we can determine the variance of $\gamma(t)$ from Lovejoy et. al (2021):

$$K_h = \sqrt{\frac{\pi}{2cos(\pi\left(h-\frac{1}{2}\right))\Gamma(-1-2h)}},$$
$$\sigma_{f,\tau_r} = \frac{\sigma_{T,\tau_r}K_h}{s}\left(\frac{\tau}{\tau_r}\right)^h.$$

Where K_h is a standard normalization constant chosen for convenience, τ is the relaxation time, τ_r is the resolution (taken to be monthly in this case), s is the climate sensitivity parameter, σ_{T,τ_r} is standard deviation of the globally averaged monthly temperature anomaly at resolution τ_r, h the scaling exponent of the temperature fluctuations, and $\sigma_{f,\tau_r} = \gamma(t)$ is the standard deviation of the internal forcing. Using our F_{RCP} (and F_{SSP}) parameter estimates, we find a mean estimate of the variance of $\gamma(t)$ to be 1.15 Wm^{-2} (1.29 Wm^{-2}) and 90% CI of $[0.89, 1.42]$ Wm^{-2} ($[0.99, 1.65]$ Wm^{-2}). If we introduce a white noise forcing, $\gamma(t)$, with the variance calculated above to the FEBE we will be able to recreate the the amplitude of the internal temperature variability response. This will be included in the revision.

Harries et. al. (2010) sets out to examine the net energy flux balance at the top of atmosphere (TOA) measured using observations from polar-orbiting spacecraft. The early observations, using the Nimbus experiments, show an internal variability of the $4.1 \pm 4.0 Wm^{-2}$, while more modern measurements (CERES) in the 2000s show variability of between ± 2 and ± 4 Wm^{-2} generally laying a few Wm^{-2} of zero. Thus our estimate of the internal forcing variability is within estimates of the TOA net energy flux balance.

Lovejoy S, Procyk R, Hébert R, Rio Amador L. The fractional energy balance equation. QJR Meteorol Soc. 2021;1–25. https://doi.org/10.1002/qj.4005

Figure 6: The response of FEBE is compared with an IPCC two-box model. What is this model? Maybe I missed the place where it is described. Please describe this model in more details in the text

Author: The two-box model we are referring to is the classical linear two-layer energy-balance model described in Held et al. (2010) and found in IPCC AR5 (2013, section 8.SM.11.2):

$$C\frac{dT}{dt} = F - \lambda T - \gamma(T - T_0),$$
$$C_0\frac{dT}{dt} = \gamma(T - T_0).$$

We graphically showed the comparison in fig. 6. The parameters for the two-box model are the best estimates from Geoffroy et al. (2013).

Held, I. M., M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, (2010): Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. J. Climate, 23, 2418–2427. https://doi.org/10.1175/2009JCLI3466.1

Geoffroy O, Saint-Martin D, Olivié DJ, Voldoire A, Bellon G, Tytéca S (2013) Transient climate response in a two-layer energy-balance model. part I: analytical solution and parameter calibration using CMIP5 AOGCM experiments. J Clim 26:1841–1857. https://doi.org/10.1002/env.2140

Figure 12: Please correct the colors of the curves. I cannot figure out what is plotted

Author: Thank you for pointing this out, the colours will be changed as to be easily readable.