

Response to William Collins, 02/08/2020

Our response is given in standard typeface, with the original review in italic.

This paper proposes a set of equations to be used as a simple tool to compute temperature changes from given emission scenarios. This is certainly a worthwhile concept and this paper will make an important contribution to this goal. However, if the aim is to encourage wide uptake, considerable work is needed to make the paper more readable. The paper needs to explain the concepts more fully (one example out of many is the central role of iIRF100) but in a simpler way without assuming so much familiarity with the model.

We thank the reviewer for their thorough and instructive review of the manuscript. In the following response, we address all the concerns raised on a point-by-point basis. In each case, we will either state what amendments are made to the revised manuscript following each comment, or explain any reason for disagreement.

My first point is rather minor, but it seems the justifications put forward for GIS are very similar to those originally put forward for FAIR. Did the authors find in developing FAIR that it became more complex than expected? Line 54 states “. . . representation of other greenhouse gases, significantly increasing the structural complexity of the FAIR model. . .”. So is the main difference between GIS and FAIR the treatment of non-CO₂ gases? Is this sufficiently different to be a new model?

We agree with this comment. This is largely resolved in our revision, since we now present an update to FAIR, rather than an entirely new model. In terms of the motivation for this update; we feel that while FAIR is an extremely useful and still relatively simple tool for exploring climate change, we believe that it is possible to simplify it even further without losing robustness and make it even more transparent with this update.

My bigger criticism is on the “transparency” of the r_0 , r_u , r_T and r_a coefficients. It is not at all clear that these translate readily into physically meaningful parameters that can be compared with and between models. This contrasts with the “Gas pools” and “Thermal boxes” that can be understood in terms of reservoirs and easily be compared eg with Joos et al. (2013) and Geoffroy et al. (2013). The formulation of alpha as a sinh is very non-intuitive, and so there is a concern that this will discourage the extent of the uptake of this model that the authors hope for. The defining principle in Millar et al. 2017 was the concept of the iIRF100. However, it was not obvious (or explained in Millar et al.) why this should have been a fundamental quantity for CO₂, and it isn't explained here why this should be a fundamental quantity for methane and N₂O. Table 1 in Millar et al. 2017 gives guiding analogues for the r-terms, but it is not clear that they mean anything physical for methane or N₂O. For instance the text refers to natural methane emissions being accounted for in the fit to get the r_0 term, but it is not obvious there should be any connection between emissions and r_0 .

Some sentences better discussing the physical reasoning behind r parameters in CO₂, CH₄ and N₂O formula have been added to the text.

The physical justification behind the alpha value in Millar et al's FAIRv1.0 is that it acts as a means to provide a state dependence to the timescales in the carbon cycle. The alpha value multiplied by the nominal time constants $\tau_1 \leftrightarrow \tau_4$ produce real time carbon cycle timescales which are dependent on the carbon accumulated in the land, ocean and biosphere, and on the global average temperature anomaly.

The iIRF100 is the integrated airborne fraction over a 100 year period, and represents the extent to which a pulse emission of CO₂ remains in the atmosphere over a 100 year period. We integrate the airborne fraction to get a parameter which distills the average airborne fraction over a 100 year period, as opposed to an instantaneous airborne fraction

at year 100. The iIRF100 can be exactly calculated with the FaIRv1.0 framework, and estimated using a parameterisation in terms of global average temperature and accumulated CO₂ in the pools of the carbon cycle in order to estimate the value of the alpha parameter, i.e. how much has the current climate state impacted the carbon sink behaviours.

In FaIRv2.0 we have followed this framework but argued an impulse response structure is more generally applicable to describe the evolution of concentrations for a wide range of pollutants. The necessary step is to define adequate parameterisations such that the correct physics is captured for each gas. In the case of FaIRv1.0, the requirements were for CO₂ a global atmospheric residence time for ~40% of input CO₂ emissions of several centuries, and a carbon cycle feedback which depends on the accumulated carbon stock in the non-atmospheric pools of the carbon cycle (r_C) and the global average temperature anomaly (r_T).

For methane the requirements are different; we want a globally averaged atmospheric residence time of around a decade, and feedbacks which are dependent on the atmospheric concentration (r_a) and the global average temperature (r_T). Further there is only one major decay pathway for methane out of the atmosphere, and so one pool is sufficient ($\tau_2 \leftrightarrow \tau_4 = 0$). This parameterisation is adequate, as is demonstrated in the paper text over historical period, to capture the globally averaged emissions to concentrations relationship for an SLCP such as methane.

For N₂O the story is similar to CH₄, except we require only a feedback which is dependent on the atmospheric concentration of N₂O (r_a), again requiring a single pool ($\tau_2 \leftrightarrow \tau_4 = 0$).

For minor contributing pollutants, such as HCFCs, we work with a simple parameterisation with a fixed lifetime ($\alpha \tau_1$) which is representative of values quoted in the literature. This is analogous to the most simple exponential decay model.

As for the fit to get the r_0 requiring accounting for natural emissions of CH₄/N₂O: this is required because the r_0 parameter sets the baseline atmospheric decay lifetime for anthropogenic CH₄/N₂O. This lifetime is dependent on the background of natural emissions because the CH₄/N₂O lifetime is dependent on the atmospheric concentration of these gases.

Related to the above, the fitting procedures are not clear, particularly the value in fitting to the historical observations. With 14 parameters in table 1, it is not surprising that the model can fit the historical record well, but is it for the right reasons? It would be more useful to fit to idealised experiments (as is done in Joos et al., and Geoffroy et al., and in the C4MIP experiments for beta and gamma). Then it should be clear which terms in the models are being represented by parameters in GIR.

We will clarify the fitting procedures within the text and in accompanying publicly available code. Specifically addressing the 14 parameter fitting procedure: we do not fit to all 14 parameters simultaneously at any point (though this would presumably lead to an excellent fit). We fit groups of parameters in what we believe are physically motivated ways, basing our fits on either available data or literature. However, we entirely agree that the original submission was not clear enough with exactly how we fit each parameter, and we will expand on this section fully in the revision. In the revision, we have tried to avoid fitting parameters ourselves except where necessary; instead taking values from the literature or using published methodologies.

The sections describing calculations of e.g. species lifetime “Emission-driven historical simulations”, and climate responses metrics (TCR, TCRE) “Idealised experiments” need to be clearer as to what extent these quantities are inputs to the models and to what extent new information is provided through the fits to historical timeseries. Similarly for the emission metrics such as GWP.

We do not wish to suggest that the GIR metric values represent any advances in their estimation with this model study; since they are really just a function of the tuned parameters (and as stated other methodologies of parameter fitting would be equally valid to those chosen). We provided these values as they are useful for any consumers of the model who wish to use the default configuration without having to calculate these metrics themselves. As such, we have moved all the metric calculation sections of the paper to the supplementary material and made it clear that they are for model benchmarking purposes only.

It would be useful to compare SSP3-70 results from FAIR and GIR since this scenario has very different levels of ozone precursors. This would confirm (or not) that the treatment of ozone forcing in GIR was sufficient.

This is an extremely helpful suggestion. In the revision we include a comparison of FaIR v1.5 and GIR over the SSP3-70 scenario. We have made some adjustments to the ozone parameterisation presented, and a comparison with FaIR v1.5 and MAGICC7.0.1-alpha is now included.

Line 49: What was “not quite adequate” about AR5-IR? This needs to be more explicit.

This has been clarified in the text. While AR5-IR is a transparent, simple tool for calculating metrics at the present day- something it does extremely accurately, it cannot be used robustly to simulate either historical or future emission scenarios. This is due to the lack of a state dependence in the carbon cycle: the airborne fraction within AR5-IR remains constant throughout a simulation, so AR5-IR cannot emulate any changes in the efficacy of carbon sinks with rising atmospheric concentrations or temperatures (as we observe in Earth-system-models).

Line 80: This section needs to start with some introduction and explanation of the concepts rather than immediately diving into the equations.

Thank you for the advice, we have attempted to clarify and restructure the text to improve its readability.

Line 85: Equation (3) is very non-intuitive, it is the solution to Eq 7 in Millar et al. 2017, but it seems to overcomplicate very uncertain relationships. While iIRF100 might have been a useful concept in Millar et al. 2017, it is not at all obvious that it is the most useful formulation for GIR. This is particularly true for methane and N2O since later on the equation needs to be linearised. Why not just leave it in a linear form? What is “h”? 100 years?

The formulation of this equation is essentially directly inherited from the Millar et al paper. They proposed that the iIRF100 was a linear function of accumulation carbon stock in the land and ocean and temperature. Within the original FaIR framework, this linear dependence resulted in the state-dependence parameter, alpha, being found with a root finding routine (since alpha is a nonlinear function of the iIRF100, as equation 7 in Millar et al). However, this solution for alpha (for a very wide range of iIRF100 values) is very well approximated by the sum of a linear and exponential function. This is what results in the admittedly rather menacing looking equation 3.

We linearise the equation for N2O and CH4 only to enable a more transparent procedure for fitting the r coefficients for these gases, rather than just using an arbitrary optimisation routine. As these terms are small for realistic CH4/N2O concentrations, we believe this linearisation is defensible.

“h” is indeed 100 years- we have made this clear in the text.

Line 93: In what way is the analytical equation an approximation, what terms have been neglected?

The solution for alpha offered in Millar et al. (2017) requires solving for the root of a non-linear equation. The solution offered here is that of the form $\alpha = g_0 \cdot \sinh(iIRF100/g_1)$. The two are not identical, and the latter is an approximation of the true non-linear solution for alpha over a wide range of iIRF100 values (though this non-linear solution is itself a subjective parameterisation and therefore which is more justifiable is debatable). The difference between the approximation (equation 3 in text) and the exact solution found in Millar et al. (2017) is shown in figure S1 of the supplementary material. The difference, although larger at very low and very high iIRF100 values, is small over the range of iIRF100 values encountered in real world scenarios, even in very high emissions scenarios such as RCP85.

Line 116: Since GIR is representing carbon-cycle models, it would seem much more sensible to fit to emission-driven models, rather than bottom-up emissions to observations. If there is any discrepancy between emission-driven complex models and observed trends, then that represents a process we don't understand. Whereas in fitting to observations, any discrepancy will get folded into the fitted parameters in an unknown way and hidden. Can the rT and ru terms be related to the more physically relatable beta and gamma (either capital or lower case) of C4MIP?

For clarity, we have altered the CO2 cycle tuning to match that in Jenkins et al (2018). We believe that fitting carbon-cycle parameters to observed emissions is a defensible way to select default parameters such that the model will reproduce historical observations; this has been done in previous model literature (Millar et al, 2017 and Smith et al, 2018 both fit to match historical emissions and concentrations). While we entirely agree that this could run the risk of folding up discrepancies into the parameters, such discrepancies likely exist in future scenarios that are simulated by GIR, and so we argue that it is better to have a model that by default can reproduce historical trends; rather than one that may represent our current understanding of processes involved but cannot reproduce historical trends. However, we emphasise that this is only one method for tuning the model, and others are equally valid. For example, one could tune GIR to reproduce the carbon cycles of individual CMIP6 models, something we aim to do in the future. We shall explore relating the our feedback terms to those used in C4MIP: this is a very helpful suggestion.

Line 130: Presumably some fixed relationship between temperature change and water vapour change is used? This should be stated.

We don't use a fixed relationship explicitly. Instead, we convert the best literature estimates of the CH4 lifetime sensitivity to both water vapour and tropospheric air (Holmes 2013) temperature to the GIR parameterised temperature feedback (r_T). In Holmes, these two variables are computed using fixed relationships with surface air temperatures; here we take those same relationships to compute the CH4 lifetime dependence on just those two, and fit the r_T coefficient to both simultaneously. We will clarify and state this in the text.

Line 131: As for CO2, this seems very dangerous. The suggestion from Smith et al. 2017 is that bottom-up emission estimates are not consistent with the observed concentrations. This fitting hides that by folding the inconsistencies into the fitted parameters. How many parameters are fitted? The sentence suggests the pre-industrial concentration is "specified" rather than fitted.

In line with our core aim that the fitting we do here is as minimal as possible, we chose (as in FaIR v1.0 and v1.3) to specify a pre-industrial value. In the revision, we have decided to fit only two parameters: r_0 (analogous to the pre-industrial iIRF100) and r_C , keeping the r_T/r_C ratio the same as in Millar et al (2017). This tuning procedure is identical to that in Jenkins et al (2018), which used a previous iteration of FaIR model. We will fully clarify the tuning procedures used in the revised text.

Line 146: What do these tuned parameters mean? In particular what does a r_0 of 9.079 (years?) mean physically? The integral of a pulse of methane is equal to $\tau_1=9.15$ years, so it seems as if τ_1 and r_0 are degenerate. Similarly for N2O: r_0 is just $\tau_1(1-$

$\exp(-100/\tau_1)$). Does r_0 have any meaningful property that is different to τ_1 ?
Table 1. These parameters all need units and guiding analogues.

We have somewhat simplified the tuning procedure in the revision to aid clarity. The reviewer is correct that for single pool gases, r_0 is essentially degenerate with τ_1 ; and as such we specify r_0 values such that $\alpha=1$ for these gases at the start of the integration. We will provide guiding analogues and units in the revised text. To physical meaning behind r_0 would be that it is the average airborne fraction over 100 years from a pulse emission.

Line 167 to 175: This explanation of the natural emissions needs expanding. Indeed in GIR, the natural emissions must be fixed at C_0/r_0 for methane and N_2O . Smith et al. 2017 showed that the bottom-up emissions are inconsistent with the observed concentrations, so it is not clear how GIR can reproduce the historical concentrations from these emissions (figure 2). How is Supplementary figure 2 generated if GIR can reproduce the observed concentration with constant natural emissions?

Figure 2 shows the GIR response when the “best-estimate” bottom-up emission timeseries (from PRIMAP-HISTtp) are input. As other models have tended to benchmark their response to the RCP timeseries (eg. FaIRv1.5), the supplementary figure 2 shows the residual when the RCP emissions are subtracted from the GIR inversion of RCP concentrations. This figure therefore demonstrates the incompatibility (at least, without additional “natural” emissions as are included in FaIRv1.5) of the RCP database emission and concentration scenarios (which has been shown previously). Within the natural emissions section we aimed to demonstrate that GIR is not vastly different to other SCMs (since the emission residual is similar to the specified natural emissions in FaIR); and we suggest that the problems associated with these internally specified natural emission pathways (or similarly concentration driven runmode to the present day) require more attention- since, for example, they result in historical pathways being treated differently to future scenarios, possibly with a discontinuity at the present day. As present day trends are key in robust estimation of a wide range of policy-relevant quantities (one example being the carbon budget), we argue that it is worth at least exploring the idea of a a single consistent model over the full (historical and future) scenario, rather than changing model internals between the two. We will expand on this section and clarify the points we attempt to make.

195: It is not clear whether the lifetimes presented here contain any new data, given the τ_1 and r_0 are specified from the Prather and Holmes studies. Are the present and pre-industrial lifetimes just extrapolations based on those coefficients – if so, it is not surprising they agree. Again there is a suggested dependence on r_0 - this one parameter seems to do a lot of work so there really does need to be a physical justification for it.

Yes they are- the idea behind their inclusion is to demonstrate that our parameterisation of the Holmes/Prather studies has not significantly affected any of the relevant physical quantities. We will include more explanation of the r_0 parameter.

Line 220 -240: Again, this parameter fitting hides the science. The Etmian formulae are transparent, whereas the formulae in table 2 have different coefficients and additional terms. Are the non-primary coefficients significantly different from zero? The calculations for the f_2 factors for the non-direct effects of methane on ozone and others need to be shown. I suggest sticking with the Etmian formulae and explicitly adding in extra terms only when necessary to represent physical processes (CH_4 - N_2O overlap, ozone production). The calculations for these extra terms need to be provided.

We have re-parameterised the model in the revision, and will be explicit about the formulae used.

Line 245: While ozone is historically correlated with CH_4 . The assumption of the same correlation continuing in future might not be valid. This could be quickly tested using

the FAIR parameterisation for ozone.

We have slightly adjusted the ozone forcing calculation in line with a parameterisation from Ehhalt (2001) in the revision, which is much more similar to the parameterisation in FaIR. However, we explicitly compare the two in the supplement now.

Line 257: Where does this value of 60 MtSO₂/yr come from?

It was from Stevens (2015). We have adjusted the aerosol cloud interaction scheme in the revision.

Line 337: What is the full forcing for N₂O? There is no additional forcing attributed to N₂O in table 8.SM.6 of Myhre et al. 2013. The calculation used needs to be shown in Full.

This arose from a misunderstanding of mine- we provided the "full" and "direct-only" based on the with/without cc feedback GWP values for N₂O. Clearly this is not the same as direct and indirect forcing so the "full" value has been removed.

Line 339 "We find values comparable to the current literature": These calculations of IRF100, iIRF100, AGWP100 and GWP100 and their methodology do not add value to the paper if all that can be said is that they are comparable to current literature. I suggest this section is removed. The values for CO₂ come from the a_i and τ_i parameters which come from Joos et al, so it is not surprising that these agree with Joos. Similarly for the methane and N₂O metrics, these are determined by the τ_1 and the f_i parameters which are derived from the same Prather, Holmes and Etminan papers as used in the literature. It is possible that the added temperature dependence of some of these parameters could affect the metrics - if the authors think this is worthy of discussion then the difference between the metrics for variable alpha could be compared with alpha=1 values (which is implicitly what is assumed in the literature).

We agree that these do not add academic value to the paper. However, as they are useful for consumers of the model; and for model intercomparison, they are still included, but in the supplementary material.

Lines 354-373: Again it is not clear that TCRE is a new result from GIR, rather than a consequence of the parameters adopted in GIR. The text seems to suggest that the TCRE agreement with literature is a validation of the model, whereas it seems mostly driven by the same inputs as the literature (Joos-like carbon response, and Geoffroy-like climate response).

We are aware that some consumers (non-experts who want a reasonable physical model for their study) may use this paper when deciding which model to employ. We wanted to display a number of key results, which although unspectacular in terms of differences to other studies, demonstrate consistency with the other standard SCM choices i.e. that our simplifications have not come at the expense of performance in these key metrics

Line 368-369 ". . .lowers the upper end of the TCRE distribution and raises the median." If the authors are implying that using GIR can provide new information on these metrics, then this needs much more explanation on where this new information is coming from.

We will revisit this part of the text and revise / move to SI. These sections aren't aiming to introduce new information, but instead provide estimates of the key parameters to compare against other parameters.

Line 380: The IIASA database has harmonised historical and future emissions, so it would make sense to use those rather than redoing this independently.

We had problems with the IIASA database having a complete set of resources to sample from. We have shifted to the RCMIP database, which contains complete scenarios that

match the IIASA database and fill in the missing gaps, and can now provide this, and what is requested below.

Line 388: CO₂, N₂O and CH₄ concentrations are available on the IIASA SSP database (from MAGICC). Presumably the authors could quickly generate these from FAIR too.

We had problems with the IIASA database having a complete set of resources to sample from. We have shifted to the RCMIP database of simple model scenarios instead, and can now provide this, and what is requested above.

Line 389: Why is the CO₂ concentration slightly lower than FAIR given that GIR and FAIR use the same formulation for CO₂?

While GIR is extremely similar to FAIR, the analytic approximation of alpha as function of the iIRF100 implemented in GIR results in very slightly different concentrations when compared to the original FAIR model, which solves for alpha at each timestep. On a more detailed level, it is because the pre-industrial iIRF100 value has to be slightly lower to give the same value of alpha in GIR when compared to FAIR; this lowered iIRF100 (and correspondingly, r_0) results in the marginally lower concentrations observed. The differences are still small when compared to the differences between different SCMs.

We use different parameters so expect different output. We have retuned the parameter set using updated datasets for emissions inputs and concentration outputs. FAIRv1.0 is an old tuning for CO₂ (from 2017 datasets), so we update as part of the FAIRv2.0 release.

Line 478: Given that the more complex chemistry and carbon-cycle models can't reproduce atmospheric concentrations from bottom-up emissions, it is extraordinary that a simple parameterisation of these complex models can do so.

They certainly can't do it perfectly, but they can do it reasonably. We have some global understanding of the processes governing emissions inputs and concentrations outputs. We have to make an assumption as to the baseline pre-industrial concentration, but every simple model has to do this (or specify natural emissions), including FAIRv1.3

Line 483: This GWP100 increase is not discussed earlier in the text. It shouldn't appear first in the conclusions. The value (35.3) is also different from any in table 4.

Our apologies, this was erroneously left in from an earlier stage of the analysis. We have removed this value and the metric discussion from the conclusions, instead including metrics for the model in the SI.

Line 484: The timing of peak warming is not discussed earlier in the text (apart from briefly for CO₂). This shouldn't form part of the conclusions unless it is discussed more fully earlier. – is it not just that N₂O has a longer lifetime (~120 years) than the ~10 years for methane and ~4 and 36 years for CO₂, rather than any new finding from GIR?

We will remove this from the conclusions and add to the supplementary information. In essence the differences are due to the different lifetimes of the gases combined with the thermal response timescales, but we thought that the result was interesting and potentially policy relevant. It is essentially the same computation as was done in Ricke (2014) for CO₂ (peak warming occurs about a decade after emission), but applied to CH₄ and N₂O also, which we have not seen published before.

References

Millar, R. J., Nicholls, Z. R., Friedlingstein, P., & Allen, M. R. (2017). A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions. *Atmospheric Chemistry and Physics*, 17(11), 7213–7228. <https://doi.org/10.5194/acp-17-7213-2017>

Holmes, C. D., Prather, M. J., Søvde, O. A., & Myhre, G. (2013). Future methane, hydroxyl, and their uncertainties: Key climate and emission parameters for future predictions. *Atmospheric Chemistry and Physics*, 13, 285–302.

<https://doi.org/10.5194/acp-13-285-2013>

Jenkins, S., Millar, R. J., Leach, N., & Allen, M. R. (2018). Framing Climate Goals in Terms of Cumulative CO₂-Forcing-Equivalent Emissions. *Geophysical Research Letters*.

<https://doi.org/10.1002/2017GL076173>