Manuscript Ref: acp-2018-26

Rapid and reliable assessment of methane impacts on climate

Ilissa B. Ocko, Vaishali Naik, and David Paynter

We sincerely appreciate the time the referees have taken to carefully review the manuscript revisions in response to the initial comments. The additional suggestions have certainly strengthened the paper. Below, we respond point-by-point to comments (reviewer comments in blue, responses in black).

Responses to Anonymous Referee #1:

Comment 1: I have been rather sensitised by a comment from another reviewer to the fact that this paper should not be seen to specifically endorse MAGICC (as it can be read to be doing so in several places) but more generally endorse the use of reduced complexity models, of which MAGICC is just one. So for example, at 1:21 it would be better written as "reduced complexity models such as MAGICC", and similarly at 4:8-9, 13:18-19 and 14:20.

> **Response:** We appreciate the feedback and agree with the reviewers. We have tweaked the language in all of the locations identified.

Comment 2: In their response to my original comment 11:21, the authors said they would note that MAGICC is not open source. They say they have acted on this, but I couldnt find where they had done so (unless it is the cryptic (to some) statement that the executable is available (14:26)). For some people, this is a major limitation of using MAGICC because, frankly, one doesnt fully know what is really in it. I suggest that in the "limitations" around 5:27 and again at 14:25 this issue is explicitly noted. It may help to encourage the custodians of MAGICC to enter the 21st century!

> **Response:** We wholeheartedly agree that the closed source code is a big limitation and frustration to using MAGICC, and have more clearly stated that it is not open source beyond the available executable statement. We have added statements in both locations identified by the reviewer (both in the methods and conclusions sections). The reason why we use MAGICC is because of its already widespread use in the community for mitigation analyses (such as Shoemaker et al., 2013 and World Energy Outlook 2017).
>
> Shoemaker, J. K., Schrag, D. P., Molina, M. J. and Ramanathan, V.: What role for short-lived climate pollutants in mitigation policy?, Science, 342(6164), 1323–1324, 2013.

Comment 3: At 6:10 the authors say that the ECS is higher than MAGICC, but as I understand, it is in the range of the calibration models - so strictly it is higher than the MAGICC median and mean? This could be usefully clarified.

> **Response:** This is an excellent point and we have clarified. The text now reads (lines 6:12-13 in tracked changes version): *"The equilibrium climate sensitivity of CM3 is 4.8 K (Paynter et al., 2018), which is in the range of the MAGICC calibration models but higher than the median and mean."*

Comment 4: Regarding my original 6-15 (I dont think this really needs acting on, as it is arguably "geeks corner") I still do not understand why MAGICC needs methane radiative efficiencies. It uses the methane radiative forcing expression, and these efficiencies are implicit in that expression. It could be that it only uses them to compute the impact of ozone change on methane.

Response: The MAGICC model allows the user to specify the coefficient in the radiative forcing expression, which, although is in units of Wm-2/ppb CH4, we had incorrectly referred to it as the radiative efficiency. The coefficient remains unchanged from IPCC AR4 to AR5 and therefore we did not update this parameter. We have clarified this in the text.

Response: We removed the word independent in the original manuscript 8-28 ("The two independent observational datasets are perfectly correlated (r = 1.00)." reads "The two observational datasets are perfectly correlated (r = 1.00)." in the submitted revised version.) However, the word "independent still exists in the previous paragraph, which we believe is what the referee is indicating. We have now removed this "independent" as well.

Response: We thank the referee for the careful reviews and have revised the manuscript accordingly.

Comment 1: My recommendation is to return the manuscript to the authors once again, at least to do some more 'polishing' on the presentation. I also strongly support to make a title change along the line of anonymous referee1's suggestion: "The difficulty of using small ensembles of simulations of an ESM with large interannual variability to validate simple climate models in cases of small forcings"

> **Response:** We appreciate the careful review of our manuscript and the helpful suggestions. We have considered several titles and title changes, but ultimately prefer a simpler and more general title to a longer and more specific one. Although the referee's suggestion is ultimately one of our results, it is not the main motivation of the paper and does not capture the full scope of our study, which is to identify a quick and reliable method for methane mitigation analysis.

Comment 2: I regard the main conclusions of the paper, 1) 'Well trained simplified models like MAGICC are able to provide a representation of the global mean climate well enough to provide assessment studies', and 2) 'Complex climate models have limited ability to identify the response to small forcings in cases where the expected response and the simulated internal variability have similar orders of magnitude', as basically correct. However, the authors have allowed repeatedly to let themselves get carried away by their enthusiasm, and have inflicted into their paper a number of exaggerations of alleged complex model disadvantages that ought to be toned down for the final manuscript. (Beyond this, some confusing formulations need to be rectified.)

> **Response:** We have attempted to tone down the language involving disadvantages of complex models. See responses to Comment 4 for specific changes in the abstract, results, and conclusion text relevant to this comment.

Comment 3: I also feel that the ESM simulations have not been optimally setup for the fairest possible comparison with MAGICC. Besides the possibilities, which specified dynamics simulations (e.g., Lamarque et al., 2012; Kooperman et al., 2012) would have been offered for reducing ESM internal variability, the possibility to calculate radiative forcing (RF) rather than effective radiative forcing (ERF) by radiation double calling (e.g., Chung and Soden, 2015; Dietmüller et al., 2016) has also not been exploited. This implies that when discussing radiative forcing calculations, the authors' comparison is often not so much between a noisy ESM and a noise-free simplified model, but rather between a noisy ERF and a noise-free RF (see extensive discussion in Forster et al., 2016). However, as the main referees have not been so strict, I won't go nitpicking either, here. Yet, I request that the distinction between RF and ERF ought to be clear throughout the paper, and that the consequences of using ERFs from the ESM, but RFs from MAGICC are openly discussed.

> **Response:** We calculated ERF for a few reasons. First, it is the preferred measure of forcing since AR5 and the standard way in which forcing will be estimated for CMIP6/AR6. Therefore, we do not have radiation double calling implemented in

AM3. Second, a limitation of the double call is that it does not account for any fast adjustments to the climate system that occur due to the forcing agent (i.e. stratospheric cooling, fast cloud responses). These can impact the value of forcing considerably. So while using the double call would have resulted in much less noise in the forcing signal for sure, it would have also been an estimate of the forcing which is not in keeping with how forcing is generally estimated from GCM experiments or the CMIP6 protocol.

We have gone through the manuscript to clarify the distinction between AM3 ERF and MAGICC RF throughout, and have added text about the implications of using ERF in AM3 and RFs in MAGICC (lines 8:24-27 in tracked changes version): *"While RF does not capture the full alterations in the energy balance, ERF is more uncertain than RF because it involves multiple climate interactions (Forster et al., 2016). However, several studies have found that ERF and RF are nearly equal for many situations, and especially for increased concentrations in CO2 and methane (Myhre et al., 2013)."*

Comment 4: Specific Remarks

p. 1, l. 13 (Abstract): I recommend to phrase more carefully as follows: 'Using basic knowledge from observations and complex Earth system models, reduced-complexity climate models offer an ideal compromise in that they provide quick reliable insights into climate responses, with only a limited computational infrastructure needed. They are particularly useful for simulating the response to forcings of small changes in different climate pollutants, due to the absence of mentionable internal variability.'

Response: We have modified this sentence in the abstract based on the suggestion. The text now reads (lines 1:13-16 in tracked changes version): *"On the other hand, reduced-complexity climate models that use basic knowledge from observations and complex Earth system models offer an ideal compromise in that they provide quick, reliable insights into climate responses, with only limited computational infrastructure needed. They are particularly useful for simulating the response to forcings of small changes in different climate pollutants, due to the absence of internal variability."*

p. 2, l. 23: Please, cite Fuglestvedt et al. (2010) in addition.

Response: We have added this reference.

p. 4 l. 6, 7: '… if … the response is comparable..' ; '… and (iii) whether the lack of internal variability …'

Response: We have made this revision.

p. 6, l. 17: 'climate sensitivity of MAGICC'; confusing as it is variable and arbitrarily set. Either recall the mean value from p. 5, l.22, or state as the reason is that CM3

has a higher sensitivity than the majority of AOGCMs used to define the MAGICC climate sensitivity).

> **Response**: The text now reads (lines 6:12-13 in tracked changes version): *"The equilibrium climate sensitivity of CM3 is 4.8 K (Paynter et al., 2018), which is in the range of the MAGICC calibration models but higher than the median and mean."*

p. 8 , l. 16 (major issue): ERF is introduced here as an extra abbreviation but is not used later when the AM3 forcings (p. 10, l. 7) are presented which are confusingly designated with RF. Please, rectify throughout the paper.

> **Response**: We thank the reviewer for this observation, and have gone through the text to make it clear that AM3 forcings when presented are ERF.

l. 18: delete one bracket behind Myhre et al. and, please, cite Shine et al. (2003) in addition.

> **Response:** We have added this reference and made this revision.

p. 8, l. 30 Please, notice a technical error with the equation setting.

> **Response:** We have made this revision.

p. 10, l. 7: (major issue): Following Forster et al. (2016), ERF can hardly be expected to take robust values when calculated on a one-year or 5-year basis. Thus, the MAGICC vs. AM3 comparison must remain on a plausibility level within this paper, which should be emphasized.

> **Response**: We have revised the text to read (lines 10:19-21 in tracked changes version): *"While these are the standard forcing calculation methods for both types of models, we emphasize that comparing values of AM3 ERF to MAGICC RF can only allow for comparisons in broad patterns and relative magnitudes, especially because of the large variability in ERF values when averaged over one to five year timescales (Forster et al., 2016), as discussed above."*

p. 10 l. 25: Change CO2 to CO2 (several further examples afterwards).

> **Response:** We have made these revisions.

p.10, l. 20: I guess it's AM3 rather than CM3?

> **Response:** It is supposed to be AM3. We have made this revision.

p. 11, l. 1-3: Confusing, as you have given another number (0.97 W/m2) for the methane forcing on p. 10, l. 25. Is this exclusively due to the different reference year? If

yes, please state so clearly; if no, please give other potential origin(s) for the issue.

Response: The value cited on 10-25 is the IPCC value for RF in response to methane emissions, which accounts for both direct and indirect effects. This is the same value that is provided on 11-3. We have clarified that the value on 10-25 refers to IPCC value and not MAGICC's, as we believe this may have been the source of the confusion.

p. 11, l. 13 : In my understanding, here we are not dealing with 'responses', but with 'forcings' and 'adjustments' from concentration changes; the 'response' is coming only in the next subsection, so please adjust the phrasing.

Response: We have modified the text to read (lines 11:26-27 in tracked changes version): *"The results from AM3 further highlight the role of unforced variability in complicating perceived forcings from small concentration changes."*

p. 11, l. 20: Throughout this subsection, the issue of differing climate sensitivity between CM3 and the MAGICC mean is not raised, hence creating the impression that you would regard identical surface temperature responses between the models as an optimal evaluation result. This is obviously not the case, please reconsider.

Response: We have modified the text to read (lines 12:26-28 in tracked changes version): *"Further, recall that the equilibrium climate sensitivity in CM3 is larger than the mean/median in MAGICC, and therefore we expect differences in the ensemble member-averaged responses from this characteristic alone."*

p.11, l. 29 (major issue): "MAGICC has much higher correlation coefficients [with observed data], likely through the absence of internal variability". This I an odd sentence which urgently needs to be set in an adequate context. Of course, reality has no "internal variability" because there is but one realization of it! Hence, it is absolutely necessary that such an argument must not even hint at the fact that MAGICC provides better agreement with reality than CM3 (or any other complex model) does.

Response: In this sentence, we meant that the lack of internal variability in MAGICC, and thus a smoother response, is likely why the correlation between MAGICC and the observational datasets is higher. We have rephrased the text as to just state objectively that MAGICC r is higher than CM3 r (lines 12:11-13 in tracked changes version): *"MAGICC and CM3 both have high correlations with NOAA and NASA data, although MAGICC's are higher (MAGICC r = 0.92 (NOAA) and 0.93 (NASA); CM3 r = 0.76 (NOAA) and 0.75 (NASA))."*

p. 12, l. 6: 'correlation of the ensembles means', recall that the MAGICC ensemble is for 19 different model representations, while CM3 has 3 independent realizations of the same model.

Response: We have qualified the text via the following (lines 12:22-23 in tracked changes version)*: "The correlations of the ensemble-means (19 physics-driven ensemble members for MAGICC and three initial condition-driven ensemble members for CM3) are extremely high (CO2 r = 0.98; methane r = 0.92)."*

p. 12, l. 19 (major issue): Given that the temperature response lags the radiative forcing, I deem it not surprising that (spurious) negative radiative forcing, while leading to temperature decrease after 1895, may not correlate with negative temperature response during the same time period, but occurs only somewhat delayed.

Response: This is a good point by the referee, and we have modified the text accordingly (lines 13:7-10 in tracked changes version): *"This cooling is likely a lagged response to negative methane ERF (at most -0.15 W m^{-2}) from 1895 to 1900, seen both in the direct and indirect methane forcings (Figs. 3 and 4)."*

p. 12, l. 27 until end of this section: As the authors (correctly) claim a far-reaching influence of internal variability on the exact evolution of the temperature response time series simulated by CM3, I see no much sense in looking for mechanistic reasons to explain details of the actual evolution. If you like to keep this, please motivate it more convincingly. Another example of negative temperature responses simulated in case of CO2 increase has been given by Huszar et al. (2013, their Fig. 10). There, too, mechanistic explanation attempts would have been obsolete.

Response: We have removed this explanation from the text.

p. 13, l 21 (major issue): "sophisticated coupled chemistry-climate models … are generally unsuitable for analysis of methane mitigation strategies"; expressed with such dogmatic universality, this statement has to be rejected. It only holds, if assessments of many mitigation options – especially with relatively little expected impact -, or a large extent of parameter sensitivities are to be investigated. Then application of complex models becomes clearly unfeasible and resorting to simplified models is doubtlessly required; best may be a combination of both as, e.g., described by Dahlmann et al. (2016) for an example of aviation impact mitigation.

Response: We understand where the referee is coming from, and have rephrased the text (lines 14:12-16 in tracked changes version)*: "Given that sophisticated coupled climate-chemistry models are generally inaccessible, time-intensive, and often employ high internal variability, they can be unsuitable for analysis of methane mitigation strategies when emissions changes are small, when many mitigation scenarios are considered, and when a large extent of parameter sensitivities are to be investigated."*

Rapid and reliable assessment of methane impacts on climate

Ilissa B. Ocko[1], Vaishali Naik[2] and David Paynter[2]

[1]Environmental Defense Fund, Washington DC, 20009, USA
[2]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, 08540, USA

5   *Correspondence to*: Ilissa B. Ocko (iocko@edf.org)

Abstract. It is clear that the most effective way to limit global temperature rise and associated impacts is to reduce human emissions of greenhouse gases, including methane. However, quantification of the climate benefits of mitigation options are complicated by the contrast in the timescales at which short-lived climate pollutants, such as methane, persist in the atmosphere

10 as compared to carbon dioxide. Whereas simple metrics fail to capture the differential impacts across all timescales, sophisticated climate models that can address these temporal dynamics are often inaccessible, time-intensive, require special infrastructure, and include high unforced interannual variability that makes it difficult to analyse small changes in forcings. On the other hand, Rreduced-complexity climate models that use basic knowledge from observations and complex Earth system models offer an ideal compromise in that they provide quick, reliable insights into climate responses, with only to

15 relatively small changes in forcings of different climate pollutants, due to the absence of strong internal variability, using basic knowledge and limited computational infrastructure needed. They are particularly useful for simulating the response to forcings of small changes in different climate pollutants, due to the absence of internal variability. In this paper, we build on previous evaluations of the freely-available and easy-to-run reduced-complexity climate model MAGICC by comparing temperature responses to historical methane emissions to those from a more complex coupled global chemistry-climate model, GFDL CM3.

20 While we find that the overall forcings and temperature responses are comparable between the two models, the prominent role of unforced variability in CM3 demonstrates how sophisticated models are potentially inappropriate tools for small forcing scenarios. On the other hand, we find that MAGICC can easily and rapidly provide robust data on climate responses to changes in methane emissions with clear signals unfettered by variability. We are therefore able to build confidence in using reduced complexity climate models such as MAGICC for purposes of understanding the climate implications of methane mitigation.

25 ## 1 Introduction

Reduced-complexity climate models offer an ideal framework for evaluating greenhouse gas mitigation options if they can accessibly and rapidly reproduce the results of the more complex global chemistry-climate models (CCMs) that include more advanced and comprehensive treatments of chemistry and physics (Meinshausen et al., 2011a). However, there is a critical need to build confidence in the ability of reduced-complexity models to simulate temperature responses to individual

greenhouse gases rather than just the suite of climate pollutants, because greenhouse gases have vastly different radiative properties and atmospheric lifetimes (Myhre et al., 2013; Fiore et al., 2015); it is important to confirm that individual species are represented appropriately if reduced-complexity models are to serve as an effective tool for assessing climate benefits of mitigation actions. This is especially central for the analysis of methane (CH_4) mitigation actions, of which the climate policy

5   community has been increasingly focused on (e.g., Shindell et al., 2012; Collins et al., 2018). Therefore, this paper builds on previous evaluations by comparing forcing and temperature responses to historical methane and carbon dioxide (CO_2) concentrations from a widely-used reduced-complexity climate model, with that from a state-of-the-art coupled global chemistry-climate model. While it is difficult to compare climate responses of simple models with that of complex ones because of the presence of unforced variability in the latter, we ultimately seek to determine if general temporal patterns and

10   magnitudes are consistent enough to justify the use of a reduced-complexity climate model as a reliable tool for rapid assessment of methane mitigation measures.

Climate change impacts have now been observed on every continent and in every ocean (Stocker et al., 2013). If we want to reduce near- and long-term anthropogenically caused warming, then we need to reduce emissions of several climate pollutants. While limiting long-term climate warming requires drastically reducing CO_2 emissions, reducing emissions of short-lived

15   climate pollutants (SLCPs)—specifically, methane and black carbon (BC)—has been identified as one of the most ~~effective~~ impactful ways to reduce near-term warming (e.g., Ramanathan and Xu, 2010; Shindell et al., 2012; Rogelj et al., 2013; Shoemaker et al., 2013). Methane emissions in particular account for a quarter of the excess energy trapped by human emissions, and today's global anthropogenic methane emissions will have a larger impact on near-term warming than today's global fossil fuel CO_2 emissions (based on forcing data provided in Myhre et al., 2013 and references therein; methane

20   emissions provided in EPA, 2012; CO_2 emissions provided in IEA, 2015; and radiative efficiency estimates of methane provided in Etminan et al., 2016). Sustained methane emissions will also impact long-term warming (Allen et al., 2016). Furthermore, reducing methane emissions has air quality, health, and food security co-benefits (Shindell et al., 2012; West et al., 2013; Zhang et al., 2016; Melvin et al., 2016).

Most methane mitigation measures are assessed as a comparison to carbon dioxide warming impacts; almost all policy analyses

25   rely on the simple metric Global Warming Potential (GWP) because of its simplicity and ease of use (Fuglestvedt et al., 2010; Ocko et al., 2017). However, GWP is limited in its ability to quantify climate effects because it relies on the integrated impact of a pulse of emissions over a specified time horizon. Because methane and CO_2 have vastly different atmospheric lifetimes, their respective climate impacts occur over different timescales. Due to the inherent selection of a single time horizon, GWP is incapable of capturing these important temporal distinctions (e.g., Solomon et al. 2010; Alvarez et al., 2012) unless two time

30   horizons that represent near- and long-term impacts are reported simultaneously (Ocko et al., 2017).

Assessment of SLCP climate impacts over different timescales can be performed using comprehensive global chemistry-climate models (CCMs), however, a full assessment of various SLCP scenarios using sophisticated CCMs is computationally

intensive and time-consuming, and forcing perturbations from slight changes in individual species are often too small for the response signal to be detected among the high unforced internal climate variability present in CCMs (e.g., Ocko et al., 2014). Determining robust climate responses to small-forcing scenarios using CCMs therefore requires a large number of ensemble simulations (Deser et al., 2012). Given that many institutions do not have access to CCMs nor the technical capacity or expertise to run these models, they must rely on partnerships with modelling centres that are often focused on model development. These characteristics of CCMs reinforces the use of the simple GWP metric for assessments of climate pollutant mitigation measures.

While detailed assessment of regional climate responses can only be provided by complex CCMs, reduced-complexity climate models offer a useful alternative for global changes in major climate characteristics that is far more advanced than GWP but avoids the need for the tremendous amount of computational resources required to perform CCM simulations (and especially with enough ensemble members to average out unforced variability). These simpler models can rapidly analyse global average climate responses because they are easily accessible and quick to run, thereby providing immediate scientific guidance for mitigation assessments. Further, because they do not include unforced internal variability, they provide clear responses to small forcing scenarios without any noise. .

There are several models that have been developed that fall within this intermediate complexity class—more advanced than simple metrics but far less sophisticated than CCMs. They range from simplified expressions (e.g., Shine et al. 2005) to more complex chemistry and physics but computations of only a few climate indicators averaged over large spatial domains (e.g., Meinshausen et al. 2011a; Hartin et al. 2014). One of the latter is the freely available Model for the Assessment of Greenhouse-gas Induced Climate Change (MAGICC), initially developed in the late 1980s (Wigley and Rapper, 1987, 1992) and routinely updated since (e.g., Meinshausen et al., 2011a). While not meant to replace atmosphere-ocean global climate models (AOGCMs) and carbon cycle models, MAGICC is a complementary, computationally-inexpensive tool that is capable of efficiently analysing basic climate responses (such as radiative forcing, surface air temperature, and ocean heat uptake) to a suite of emission scenarios. Confidence in MAGICC results comes from a comprehensive effort to match several AOGCMs and carbon cycle models (Meinshausen et al., 2008, 2011a). Evaluations show that MAGICC closely reproduces temperature responses to aggregated forcings from the sophisticated Coupled Model Intercomparison Project CMIP3 atmosphere-ocean and C4MIP carbon cycle models (Meinshausen et al., 2011c).

While not the only model of its class, the reduced complexity model MAGICC is an especially great resource for mitigation analysis because of its widespread use in international climate reports, and the ability of the user to modify future emissions of every radiatively active species. Therefore, when numerous scenarios exist and need to be evaluated for decision-making, a rapid tool like MAGICC can provide rapid insight into the climate impacts of various options. However, to build confidence in MAGICC's evaluation of greenhouse gas mitigation strategies, we need to adequately assess its ability to reproduce climate responses to individual greenhouse gases beyond the aggregated forcings. Here, we analyse the capability of MAGICC in simulating climate responses to historical increases (1860-2014) in methane and CO_2 by comparing the results with that from

a state-of-the-art coupled chemistry-climate model, the National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Laboratory (GFDL) CM3 model, which has been shown to adequately reproduce historical temperature trends (Golaz et al. 2013, Griffies et al., 2011; Donner et al., 2011; Winton et al., 2012; John et al., 2012; Levy et al., 2013). While it is difficult to compare simpler models with sophisticated ones for scenarios with small forcings – due to high interannual variability built into the latter – it is nevertheless important to do so because of the more advanced and comprehensive chemistry and physics in the more complex models.

We compare the response of the two models to assess similarities and differences, seeking to determine (i) if the forcings/temperature response is comparable; (ii) if the complexity of the CCM provides any benefits over the simple model; and (iii) ~~does~~ whether the lack of variability in the simple model provide any advantages over the CCM when looking at small forcing amounts. Our goal is to build confidence in the simulation of the climate response to methane in order to justify future use of reduced complexity climate models, such as MAGICC, to assess the climate impact of methane emissions mitigation scenarios. In this analysis, we add to previous evaluations by showing a high correlation between CM3's and MAGICC's radiative forcing and surface air temperature responses to changes in either CO_2 or methane in isolation, despite large unforced variability in CM3, thereby strengthening confidence in MAGICC's simulation of climate responses to individual greenhouse gases with vastly different radiative properties and lifetimes.

2 Models and Simulations

2.1 MAGICC model description

We use MAGICC v.6 version developed in 2011 (http://www.magicc.org/download). MAGICC represents the complex coupled carbon-cycle climate system as a hemispherically averaged upwelling-diffusion ocean coupled to an atmosphere layer and a globally averaged carbon cycle model. The atmosphere has four boxes (one over land and one over ocean for each hemisphere) and is coupled to the mixed layer of the ocean hemispheres. The default number of ocean layers in each hemisphere is 50 including the mixed layer (though users can select the number of levels), and heat exchange is driven by vertical diffusion and advection. The terrestrial carbon cycle model is a globally integrated box model with one living plant box and two dead biomass boxes (one for detritus and one for organic matter in soils). The terrestrial carbon cycle does not feedback into carbon dioxide concentrations in the atmosphere. The sea-to-air carbon flux is determined by the partial pressure differential for carbon dioxide between the atmosphere and surface layer of the ocean.

From 1765–2005, the MAGICC v.6 radiative forcing is driven by global-mean concentrations of greenhouse gases (carbon dioxide, methane, nitrous oxide, ozone-depleting substances and their replacements); prescribed regional direct aerosol radiative forcings (sulphate, black and organic carbon, sea salt, mineral dust); land-use, volcanic, and solar radiative forcings; prescribed black carbon on snow radiative forcings; emissions of tropospheric ozone precursors (carbon monoxide, nitrogen

4

oxides, non-methane volatile organic carbon); and indirect (first and second) aerosol forcings calculated from prescribed regional aerosol optical depths (parameterizations described in detail in Meinshausen et al. (2011a)). For 2006 to 2014, the model is driven by emissions of gases and aerosols taken from the Representative Concentration Pathway (RCP8.5) scenario to capture a business-as-usual trajectory. Climate responses (such as surface air temperature) are provided as global annual
5 averages and also across four spatial boxes (over land and ocean and by hemisphere).

Historical greenhouse gas concentrations are from Meinshausen et al. (2011b); prescribed aerosol forcings and land-use historical forcings are from the National Aeronautics and Space Administration (NASA) GISS model (http://data.giss.nasa.gov/); solar irradiance is provided by Lean et al. (2010); and historical emissions of ozone precursors are from Lamarque et al. (2010). Present-day and future (2005–2100) forcings are driven by emissions of gases and aerosols, and
10 are taken from the Representative Concentration Pathway (RCP8.5) scenario to capture a business-as-usual trajectory, though we restrict our analysis here to climate responses from 1860–2014. Carbon dioxide radiative forcings are calculated using a standard simplified expression (Shine et al. 1990 with updated scaling parameter from Myhre et al. 1998). Methane radiative forcings are calculated using a radiative efficiency parameter in conjunction with standard simplified expressions from Myhre et al. (1998), and accounts for overlap between methane and nitrous oxide absorption bands.

15 For the most recent version of MAGICC, seven key climate parameters were calibrated to match 19 AOGCMs used in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report AR4 (see Meinshausen et al., 2011a). The parameters include: equilibrium climate sensitivity, land-ocean warming ratio at equilibrium, vertical diffusivity in the ocean, sensitivity of feedback factors to radiative forcing change, sensitivity of vertical diffusivity at mixed layer boundary to global-mean surface temperatures (i.e., thermal stratification), land-ocean heat exchange coefficient, and an amplification factor for
20 the ocean to land heat exchange. The MAGICC parameter set that best reproduces surface air temperatures and heat uptake of each AOGCM is determined via an optimization routine with 1000 iterations to find the combination that minimizes the squared differences between low-pass filtered time series. The effective climate sensitivities in MAGICC v.6 vary over time due to spatially non-homogeneous varying feedbacks, until they reach the equilibrium climate sensitivity. The equilibrium climate sensitivity input into MAGICC depends on which AOGCM calibration is used; they range from 1.9 to 5.73 °C across
25 all 19 models, with a mean (median) of 2.88 °C (2.59 °C). Multi-model-ensembles are generated by running each simulation for all 19 AOGCM calibrations, which we refer to as "physics-driven ensemble members." The user of the downloaded MAGICC model can select which parameters to use for each simulation.

While the MAGICC model is particularly well-calibrated to more sophisticated models, the realism of MAGICC results relies on the realism of GCMs, which have their own sets of limits and uncertainties. Further limitations of MAGICC include
30 incomplete knowledge of forcing patterns, unknown responses outside of the calibrated range, limited set of climate responses evaluated (such as temperature and heat uptake but not precipitation), reliance on a high level of parametrization (such as cloud feedbacks tuned to match those of more sophisticated GCMs), and possible errors in the data used for calibration. In addition, although the model is freely available, it is not open source. However, despite these limitations, MAGICC has been shown to

5

reasonably reproduce climate responses to all-forcing scenarios (Meinshausen et al., 2011a) and is one of the most prominent reduced complexity climate models in use.

2.2 CM3 model description

We employ the GFDL global coupled atmosphere-ocean-chemistry model (GFDL-CM3; Donner et al., 2011; Griffies et al.,

5   2011) to assess the climate response to historical changes in methane and CO_2. CM3 uses a finite-volume dynamical core on a cubed-sphere horizontal grid composed of six faces; each face includes 48×48 grid cells. The size of the grid cells range from 163 km at the corners to 231 km near the face centres. In the vertical, the model domain extends from the surface up to 0.01 hPa (86 km) with 48 vertical hybrid sigma pressure levels. The model simulates tropospheric and stratospheric chemistry interactively over the full vertical domain, with simulated ozone and aerosols influencing radiation calculations (Naik et al.,

10   2013; Austin et al., 2013). Ensemble members for CM3 are generated by employing different sets of stochastically-selected initial conditions (discussed in more detail in Sect. 2.3.), which we refer herein as "initial condition-driven ensemble members." The equilibrium climate sensitivity of CM3 is 4.8 K (Paynter et al., 2018), which is in the range of the MAGICC calibration models but higher than that ofthe median and mean MAGICC.

Global mean concentrations of well-mixed greenhouse gases (WMGHGs), including carbon dioxide, nitrous oxide, methane,

15   and ozone-depleting substances (ODSs) are specified for radiation calculations for the historical period (1860-2005) from Meinshausen et al. (2011b) and for the period 2006 to 2014 following the Representative Concentration Pathway (RCP8.5) scenario. Within the chemistry module, global mean concentrations of methane are prescribed at the surface as the lower boundary condition and are allowed to undergo chemistry everywhere else in the model domain. Radiation calculations do not see the full three-dimensional methane field (simulated in the chemistry module) and only employ the global-mean

20   concentrations, however, changes in ozone and water vapour are seen by the radiation. Further, CM3 CO_2 concentrations do not get altered by reactions that occur in the model.

CM3 is forced with emissions of short-lived species including ozone precursors, and aerosols and their precursors, volcanic aerosols, solar radiation, and land-use change as described in detail by Donner et al. (2011) and Naik et al. (2013). Anthropogenic emissions, including from biomass burning and ships, for the time period 1860–2005 are from the dataset of

25   Lamarque et al. (2010) developed in support of the Couple Model Intercomparison Project Phase 5 (CMIP5). For years 2006– 2014, anthropogenic emissions follow the RCP8.5 scenario. Natural emissions of all precursor species, except isoprene, are included as described by Naik et al. (2013). Biogenic isoprene emissions are calculated interactively, as described by Lin et al. (2012), based on the Model of Emissions of Gases and Aerosols in Nature (MEGAN) (Guenther et al., 2012). 'Explosive' volcanic eruptions are imposed via a time series of volcanic optical properties rather than from direct injection of sulphur into

30   the stratosphere (Stenchikov et al., 2006; Donner et al., 2011).

Shortwave and longwave radiation algorithms in CM3 are described in Freidenreich and Ramaswamy (1999) and Schwarzkopf and Ramaswamy (1999), respectively, with some modification to enhance computational efficiency (GAMDT 2004). The shortwave algorithm includes 18 bands in the solar spectrum, and the longwave algorithm includes eight bands. Shortwave radiation parameterizations account for absorption by water vapour, carbon dioxide, ozone, molecular oxygen; molecular

5 scattering; and absorption and scattering by aerosols and clouds. The longwave radiation parameterizations account for absorption and emission by water vapour, carbon dioxide, ozone, nitrous oxide, methane, halocarbons (CFC-11, CFC-12, CFC-13 and HCFC-22), aerosols, and clouds. Aerosols included are sulphate, carbonaceous (black and organic carbon), dust, and sea salt.

CM3 includes explicit representation of both the direct and indirect aerosol effects on radiation. For the calculation of the
10 direct effect of aerosols on radiation, physical and optical properties of sulphate, black carbon, organic carbon, sea salt, and dust are considered in the model (Donner et al., 2011). Sulphate and black carbon are assumed to be internally mixed while all other aerosols are assumed to be externally mixed for radiation calculations. To account for the indirect effect of aerosols via aerosol-water cloud interactions, the model treats water soluble aerosols, including sea salt, and organic aerosols as cloud-condensation nuclei (CCN) allowing a physically based parameterization of CCN activation (Ming et al., 2006). The model
15 does not consider the reduction in surface albedo caused by the deposition of black carbon on snow-covered surfaces.

2.3 Simulations

Three historical simulations are run for both MAGICC and GFDL CM3 to derive climate responses to isolated CO_2 and methane concentrations, respectively. MAGICC was run from 1750 to 2100 by default, and CM3 was run from 1860 to 2014. As shown in Table 1, the direct runs for both models include an all-forcing simulation with all forcings varying with time
20 except land-use; a simulation with CO_2 concentrations held at 1860 levels; and a simulation with methane concentrations held at 1860 levels. Subtracting temperature responses of the two latter runs from the former yield CO_2-only and methane-only climate responses, respectively (see Eqs. (1) and (2)). The same equations hold for the forcings as well.

$$\Delta T_{CO_2} = T_{AllForc} - T_{CO_2 1860} \; , \tag{1}$$

$$\Delta T_{CH_4} = T_{AllForc} - T_{CH_4 1860} \; , \tag{2}$$

25 For MAGICC, each simulation is run for all 19 AOGCM-calibrated configurations; each 350-year integration took approximately one second to run on a modern PC with a three GHz CPU processing speed. We use default MAGICC gas and aerosol properties, but update ~~methane and~~ tropospheric ozone radiative efficienc~~y~~ies and methane atmospheric lifetime to IPCC Fifth Assessment Report (AR5) values (Myhre et al., 2013; Stevenson et al., 2013) to reflect the latest science. (Note that the updated atmospheric lifetime only impacts the model from 2006-2014 as it is driven by emissions and not
30 concentrations during this period.) However, we do not include newer estimates of methane radiative ~~efficiency~~ effects that

7

account for shortwave absorption in addition to longwave absorption (Etminan et al., 2016) to be consistent with the CM3 model that only includes longwave effects. Including the shortwave component increases methane's radiative efficiency by over 20%. Further, we specifically do not tune MAGICC model climate and forcing properties to match that of CM3 because we are assessing how a "standard version" of the reduced-complexity climate model compares with CM3; the goal is not to match MAGICC to CM3 but to assess whether a downloaded version of MAGICC broadly behaves similarly to CM3. However, two of MAGICC's physics-driven ensemble members are derived from two predecessors of CM3: CM2.0 and CM2.1 (Delworth et al., 2006).

The set-up of GFDL-CM3 simulations conducted here was similar to that adopted for simulations performed in support of the CMIP5, except we obtained initial conditions from a longer preindustrial control (3000 years). Three-member initial condition-driven ensembles of transient CM3 simulations were performed with each ensemble member initialized stochastically at different points in the preindustrial control simulation. Each 155-year integration of CM3 took about 15 days to complete on the NOAA's Remotely Deployed High Performance Computing System (RDHPCS) machine known as "Gaea" running on 464 processors. While three ensemble members is relatively small, we are limited by computational resources and studies have shown that forced changes in air temperature, as opposed to changes in atmospheric circulation and precipitation, can be detected with fewer ensemble members (Deser et al., 2012).

To compute CM3 radiative forcings for CO_2 and methane (direct and indirect) that are closest to the definition used by MAGICC (the forcing at the tropopause after stratospheric temperature adjustment), we performed simulations with the atmosphere-only version of CM3—AM3. The model configuration of AM3 was exactly the same as CM3 except AM3 model integrations over the period 1870 to 2014 were performed with observed sea-surface temperature and sea-ice cover (Rayner et al., 2003), and therefore do not include an ensemble driven by different initial conditions. Through the additional AM3 simulations, we were able to diagnose transient effective radiative forcing (ERF) (the change in net radiation balance at the top-of-atmosphere (TOA) following a perturbation to the climate system taking into account any rapid adjustments (Shine et al., 2003; Myhre et al., 2013)) due to CO_2 and methane. Transient ERF calculated in this way follow the proposed protocol for the AerChemMIP (Collins et al., 2016). While RF does not capture the full alterations in the energy balance, ERF is more uncertain than RF because it involves multiple climate interactions (Forster et al., 2016). However, several studies have found that ERF and RF are nearly equal for many situations, and especially for increased concentrations in CO_2 and methane (Myhre et al., 2013).

To separate the effect of methane due to its influence on ozone and water vapour (indirect effects) from its effect on radiation (direct effect), we ran two more simulations for MAGICC with methane chemistry turned off (an all-forcing run and methane held at 1860 levels run with methane chemistry turned off for both). Equation (3) shows how the direct methane forcings were calculated for MAGICC; subtraction between methane-only forcings and the direct forcings yielded the indirect responses to methane. We also ran two more simulations for AM3/CM3 with methane radiation calculations or chemistry held constant beyond 1860, respectively. Equations (4) and (5), respectively, show how the direct and indirect methane forcings were

calculated in AM3.While we only show forcing calculations here via AM3 simulations with fixed sea surface temperatures, we also ran the simulations for the fully coupled CM3 model.

$$\Delta F_{CH_4,direct\,(MAGICC)} = F_{CH_4 nochem} - F_{CH_4 1860 nochem} \,, \qquad (3)$$

$$\Delta F_{CH_4,direct\,(AM3)} = F_{AllForc} - F_{CH_4 rad1860} \,, \qquad (4)$$

5 $\quad\Delta F_{CH_4,indirect\,(AM3)} = F_{AllForc} - F_{CH_4 chem1860} \,, \qquad (5)$

The global mean historical concentrations of CO_2 and methane used by the models to calculate radiative forcings and therefore temperature changes are shown in Fig. 1. (Note that concentrations are prescribed for MAGICC only through 2005, and then emissions inputs drive the model thereafter; however, the resulting concentrations from these emissions are consistent with that input into CM3.) Results for both models are presented as an average of the individual ensemble members (initial 10 condition-driven ensemble members for CM3 and physics-driven ensemble members for MAGICC). Surface air temperatures are taken to be 2 meters above the surface. For both models, we calculate temperature changes as the difference between temperatures in year t compared to that in 1860.

A key difference between AM3/CM3 and MAGICC is that the full GCM has internally generated unforced variability. This occurs both when the model is coupled (CM3) and run with prescribed sea surface temperatures and sea ice (AM3). The 15 variability can be dampened by applying a smoothing to the annual time-series. However, too long of a smoothing period removes much of the decadal level forcing that we hope to uncover in this study. Therefore, we employ a five-year smoothing average to AM3/CM3 results to filter out some of the internal variability. Additionally, to better quantify and isolate the role of unforced variability in theAM3/CM3, we run control experiments of each with fixed forcing. For CM3 we ran a 500 year control simulation with all radiative forcing held constant at 1860 level. For AM3 we ran a shorter 200 year control run, with 20 all radiative forcing held fixed at 1860 with annually repeating monthly averaged sea surface temperatures and sea ice characteristics taken from 30 years of the CM3 control run.

3 Results and Discussion

Here we analyse AM3/CM3's and MAGICC's radiative forcing and surface air temperature responses to changes in either CO_2 or methane in isolation.

25 Given that an important difference between AM3/CM3 and MAGICC results is the role of unforced variability in AM3/CM3, we first analyse the magnitudes of unforced variability in both AM3 and CM3. Although initial condition-driven ensemble member means and/or running averages are employed to dampen out some of the variability in AM3/CM3, it still plays a large

role in forcing and temperature responses. CM3, in particular, has been shown to produce magnitudes of variability on the upper end of CMIP5 models (Brown et al., 2015).

The results of the control simulations with constant preindustrial (1860) external ~~radiative~~ forcings are shown in Fig. 2. In the case of AM3 unforced changes in net radiation ~~radiative forcings~~ at the top-of-atmosphere for all-sky conditions range from -

5   0.18 to 0.21 W m^{-2} with a standard deviation of 0.07 W m^{-2} for a five-year running mean. We find the maximum swing between two consecutive five-year means to be 0.35 W m^{-2}. Sources of unforced variability in AM3 include a mixture of land snow/ice cover variability, clouds, and just year-to-year variability in the ~~weather~~meteorology; soil moisture may also play a role. For CM3, unforced internal dynamics yield temperature responses ranging from -0.27 to 0.24 °C for five-year running means with a standard deviation of 0.1 °C. We find the maximum swing between two consecutive five-year means to be 0.2 °C. The

10 variability is driven by interactions among the ocean-atmosphere-land systems. While unforced variability is a key component to modelling the climate system, it can mask or amplify responses to external forcings over short timescales (e.g., Brown et al., 2017). This makes it difficult to clearly assess responses to small external forcings, and provides further motivation for using simpler models like MAGICC for analysis of small forcing scenarios.

3.1 Radiative Forcing

15 Figure 3 shows the global-mean radiative forcings (RF) in response to the all-forcing scenario as well as forcings attributed to isolated CO_2 and methane concentrations, respectively. Note that AM3 forcings are taken as top-of-atmosphere and include rapid adjustments in the troposphere in addition to the stratosphere, and therefore are considered an effective ~~RF~~radiative forcing, while MAGICC derived RF is calculated at the tropopause and only considers stratospheric temperature adjustment. While these are the standard forcing calculation methods for both types of models, we emphasize that ~~c~~omparing values of

20 AM3 ERF to MAGICC RF can only allow for comparisons in broad patterns and relative magnitudes, especially because of large variability in ERF values when averaged over one to five year timescales (Forster et al., 2016) as discussed above.

MAGICC methane and CO_2 isolated forcings are much smoother than that of AM3 because of the lack of unforced variability in MAGICC. Some unexpected features in AM3 forcings (such as negative forcings in the earlier years despite increasing atmospheric concentrations) are likely due to unforced variability. Using the MAGICC forcings as a benchmark for a signal

25 due to forced changes only, we find that nearly all of the deviations of AM3 fall within the range of internal variability as derived from the control simulation: 0.35 W m^{-2}. However, despite the slightly different forcing definitions and the unforced variability in AM3, all results are strongly correlated between AM3 and MAGICC (All-forcing r = 0.81, CO_2-only r = 0.96, CH_4-only r = 0.93).

In the present-day (model year 2014), AM3 and MAGICC yield an all-forcing ERF and RF of 2.0 and 2.5 W m^{-2}, respectively;

30 note that land use is held constant in this analysis. This is consistent with the IPCC (2013) values that show an all-forcing ~~effective~~ ERF of 2.3 W m^{-2} in 2011 (Myhre et al., 2013). The magnitudes for the ~~AC~~M3 and MAGICC all-forcing radiative

forcings are offset after 1960 (-1 W m^{-2} in 1960). This is due to AM3's strong aerosol indirect forcing (Golaz et al., 2010) beginning around this time when aerosol emissions in the mid-latitudes increased rapidly (Lamarque et al., 2010).

Isolating CO_2 and methane's contribution to overall forcings (Fig. 3), MAGICC ~~RF~~ is reasonably consistent with the AM3 ~~forcing~~ ERF evolutions throughout the 20th Century. Preindustrial to present-day forcings for CO_2 and methane simulated by

5 AM3 and MAGICC are similar to those given by IPCC (Myhre et al., 2013), (IPCC values: 1.68 from CO_2 emissions and 0.97

W m-2 from methane emissions in 2011 relative to 1750 levels), albeit there are important differences, including baseline years (1750 for IPCC and 1870 for AM3 and MAGICC in this study to match that from AM3) and time series of atmospheric concentrations (Myhre et al., 2013). While the same radiation expressions are used for IPCC and MAGICC for CO_2 and

methane atmospheric concentrations, the representation of tropospheric ozone chemistry and its radiation effects in MAGICC

10 is extremely simplified due to hemispheric averages in a four-box atmosphere. For a short-lived climate forcer that is highly spatially variable, this is a vastly different treatment than that by the IPCC, which employs multi-model assessments for tropospheric ozone forcings. We find that our direct methane forcing in MAGICC in model year 2011 is 0.45 W m^{-2}, extremely close to the IPCC's forcing of 0.48 W m^{-2} from changes in methane concentrations alone (recall however different baselines) (Myhre et al., 2013). However, when methane interactions with other chemical species are accounted for, MAGICC estimates

15 a forcing of 0.7 W m^{-2} attributed to changes in methane compared to the IPCC's value of 0.97 W m^{-2} in 2011 (Myhre et al., 2013).

In MAGICC, methane's RF is consistently around half the value of that by CO_2. In AM3, methane's ERF is much closer to that of CO_2 until the year 2000 and beyond where they diverge. While this divergence is consistent with global atmospheric methane concentrations levelling off for about a decade in the mid-1990s to mid-2000s, before rapidly increasing from 2007

20 onwards (Fig. 1), further simulations are required (such as more ensemble members or adjustments to input conditions) in order to determine if the close methane and CO_2 ERFs before 2000 are an artefact of unforced variability or a substantiated feature. Based on our analysis of unforced variability in AM3, it is quite possible that they are features of internal variability.

Methane's role in radiative forcing can be divided into direct contributions via warming by methane as a greenhouse gas, and indirect contributions via production of other greenhouse gases (mainly tropospheric ozone) as it oxidizes to CO_2 in the

25 atmosphere. Figure 4 compares the direct and indirect methane forcings from MAGICC and AM3, calculated via Eqs. (3) and (4), respectively. The results from AM3 further highlight the role of unforced variability in complicating perceived ~~responses~~ forcings ~~to~~ from small ~~forcing~~ concentration changes; the seemingly large swings in AM3 forcings deviate from that of MAGICC by around 0.25 W m^{-2} at most, which is within the realm of unforced variability (see Fig. 2). While correlation coefficients show consistency between MAGICC and AM3 (direct r = 0.87; indirect r = 0.78), the strong variability in AM3

30 makes comparisons of magnitude difficult. MAGICC attributes around 35% of methane's present-day total radiative forcing to indirect effects, similar to the IPCC's attribution of 34% (Myhre et al., 2013). AM3 shows magnitudes of indirect forcings

11

in the present-day that are around 30-50% of the total methane forcing, depending on the year; this variation is due to unforced variability.

3.2 Global Surface Air Temperature Change

To build confidence in the simulation of surface air temperature by both MAGICC and CM3, we compare the model results with 20th Century reconstructions of surface air temperature, of which several ~~independent~~ datasets are available. Figure 5 shows the historical global-mean surface air temperature responses to changes in all-forcings in MAGICC and CM3 compared with NOAA and National Aeronautics and Space Administration (NASA) time series of global surface temperature anomalies, freely available online (https://www.ncdc.noaa.gov/cag/time-series/global and https://data.giss.nasa.gov/gistemp/). Following NOAA's methodology (NOAA, 2017), we compute the 20th Century average temperatures in MAGICC, CM3, and NASA, and calculate the annual temperature departures from this baseline.

The two observational datasets are perfectly correlated (r = 1.00). ~~While~~ MAGICC and CM3 both have high correlations with NOAA and NASA data, although MAGICC's ~~are has~~ much higher ~~correlation coefficients~~ (MAGICC r = 0.92 (NOAA) and 0.93 (NASA); CM3 r = 0.76 (NOAA) and 0.75 (NASA)), ~~likely due to the absence of internal variability~~. Consistent with Fig. 3, CM3 shows lower temperature responses post-1960 due to the strong effect of aerosols (Golaz et al. 2013). We note, however, that the 'lingering' temperature response in CM3 to major volcanic eruptions is an artefact of the ~~5~~five-year running mean smoothing process; this is why CM3 temperature responses to volcanic eruptions persist longer than what is seen in the observational records and by MAGICC. This is not found, however, to considerably impact the correlation coefficients between the CM3 data and the NOAA/NASA data. Overall, the general temperature anomaly temporal patterns reveal that both models adequately reproduce surface air temperature, providing confidence in both climate models of differing complexity levels.

The global mean surface air temperature responses attributed to CO_2 and methane forcings are shown in Fig. 6, calculated via Eqs. (1) and (2), respectively. The correlations of the ensemble-means (19 physics-driven ensemble members for MAGICC and three initial condition-driven ensemble members for CM3) are extremely high (CO_2 r = 0.98; methane r = 0.92). Figure 6 also shows individual CM3 initial condition-driven ensemble members and the range of MAGICC responses from all 19 AOGCM calibrations; however, we do not include MAGICC's highest climate sensitivity physics-driven ensemble member as the responses were a clear outlier to the rest of the members. Further, recall that the equilibrium climate sensitivity in CM3 is larger than the mean/median in MAGICC, and therefore we expect differences in the ensemble member-averaged responses from this characteristic alone.

We find that both CM3 and MAGICC attribute a nearly 1 °C rise in temperature from 1860 to 2014 from rising CO_2 concentrations (CM3: 0.9 °C; MAGICC: 0.9 °C). For methane, CM3 suggests a rise of 0.5 °C and MAGICC suggests a rise of 0.4 °C, consistent with the larger methane forcing in CM3 (Fig. 3). It is important to note that cooling from aerosols mask

some of the warming that we otherwise would be experiencing from CO_2 and methane, which is why the combined warming from CO_2 and methane is larger than today's observed warming.

Two major features of the temperature response to methane in CM3, that are not present in MAGICC, further highlight the difficulty of extracting a small signal (and with a small ensemble) given the size of the unforced variability (Fig. 6); methane's forcing is considerably smaller than that of CO_2, making it difficult to extract a temperature response from the variability. The first feature is a global mean cooling response to methane forcings around 1900 to 1915, which is strongly apparent in two of the three initial condition-driven ensemble members. This cooling ~~response~~ is ~~not clearly reflected in the forcings~~likely a lagged response to ~~of both the direct and indirect methane responses, and while total methane forcings in AM3 are slightly~~ negative methane ERF (at most -0.15 W m^{-2}) from 1895 to 1900, seen both in the direct and indirect methane forcings ~~they are positive (around 0.2 W m^{-2} on average) from 1900 to 1915~~ (Figs. 3 and 4). The second feature is a strong warming signal in response to methane from 1980 to 1995, followed by cooling through 2000; while this is consistent with AM3 RFs (Fig. 3), the feature is more pronounced in the temperature response. Both of these features fall within the range of annual temperature swings due to unforced variability in CM3 (at most around 0.2 °C for a five-year running mean). Therefore, we cannot conclude that they are robust responses to methane, but rather serve as a further example of why CCMs are difficult to employ for small individual forcings and the need for large ensembles.

To dig into these features further, we analyse regional surface air temperature responses to CO_2 and methane isolated forcings (Fig. 7). Methane-induced cooling between 1900 and 1915 is ~~mostly attributed to~~strongest in the Southern Hemisphere and especially over Southern Hemisphere oceans. ~~This is likely due to the southern ocean polynya (Gordon and Comiso, 1988), which can be very strong in CM3, leading to very large unforced multidecadal time-scale variability over the southern ocean that propagates throughout the Southern Hemisphere (e.g., de Lavergne et al., 2014).~~ On the other hand, the large methane warming in CM3 around 1990 is most prominent in the Northern Hemisphere, over both land and ocean.

When the global mean responses are parsed out by region (Fig. 7), the highest surface air temperature responses to methane and CO_2 are found over land in the Northern Hemisphere, with temperatures from CO_2 rising by well over 1 °C from 1860 to 2014 in both models. There is high correlation between MAGICC and CM3 for all regions. We expect and find methane correlations between the two models to be slightly lower than CO_2 because methane has more complex chemical interactions in the atmosphere than CO_2 that introduce more degrees of freedom than CO_2, and are also potentially more simplified in MAGICC. We also find that correlations in the Southern Hemisphere are lower than in the Northern Hemisphere, especially for methane.

As seen and discussed earlier in Fig. 3 forcings, there are several time periods when the methane temperature responses are comparable in magnitude to that by CO_2 in CM3 global mean and regional responses (Fig. 6 and 7). We see this for all initial condition-driven ensemble members, and it is consistent with AM3 RFs (Fig. 3). In the ensemble mean, the comparable

13

warming magnitudes between 1940 and 1950 are consistent with the rate of growth of CO_2 concentrations slowing down while methane concentrations consistently increase (Fig. 1).

Also discussed earlier and in contrast to the CO_2 and methane concentration trends from 1940-1950, the methane concentration growth rate slows down in the 1990s while the CO_2 concentrations consistently increase (Fig. 1). This is reflected in the CM3
5 temperature trends in addition to forcings (Fig. 3) as a divergence in the magnitude of temperature responses between methane and CO_2 to where they stand in the present-day, with CO_2 yielding twice as much warming in 2014 as methane (Fig. 6).

4 Conclusions

The purpose of this study is to enhance confidence in reduced complexity climate models, a freely available and computationally efficient reduced complexity climate model, MAGICC, in the context of simulating temperature responses to
10 methane and CO_2 atmospheric concentrations. We use the freely available (but closed source code) and computationally efficient model, MAGICC, for our analysis, Our analysis is motivated by the need to determine a quick and accessible, yet reliable, method for analysing impacts of future changes in methane emissions on climate warming. Given that sophisticated coupled climate-chemistry models are generally inaccessible, time-intensive, and often employ high internal variability that obscures the response signal, they are generally can be unsuitable for analysis of methane mitigation strategies when emissions
15 changes are small, when many mitigation scenarios are considered, and when a large extent of parameter sensitivities are to be investigated. Employing a model like MAGICC, rather than resorting to simple GWP metrics, would significantly enhance the accuracy of mitigation assessments while still using basic infrastructure and providing immediate guidance for decision making.

To determine MAGICC's reliability for methane analysis, we performed several sets of experiments using MAGICC and
20 CM3—all forcing with both time-varying natural and anthropogenic forcings but land-use held constant; simulations where CO_2 and methane concentrations are held constant at 1860 levels, respectively; and a simulation to isolate methane indirect effects resulting from its influence on ozone and water vapour (for MAGICC, we turned off methane chemistry; for CM3, we held methane radiative effects at 1860 levels). We also ran simulations using the atmosphere-only version of CM3, AM3, to calculate radiative forcings in response to the four sets of experiments. Finally, we ran control simulations for AM3 and CM3
25 to determine the role of unforced variability in influencing climate responses.

Both CM3 and MAGICC models adequately reconstruct surface air temperature records from NOAA and NASA from 1860 through 2014, especially for 1950 onwards. For isolated forcings, overall temporal patterns were consistent between MAGICC and CM3 temperature responses to methane and CO_2, including for indirect effects via methane chemical reactions. Correlation coefficients were very high at 0.98 and 0.92 in the global mean for CO_2 and methane, respectively, with overall magnitudes
30 consistent. We therefore conclude that MAGICC is able to reproduce the general isolated greenhouse gas forcing results

14

(temporal patterns and magnitudes) of a more sophisticated coupled global climate model, providing confidence in the use of MAGICC for understanding the climate implications of methane mitigation analyses.

Further, we find that methane accounts for a considerable fraction of 20th Century and early 21st Century warming—roughly half that of CO_2's warming response. However, there are some features present in CM3 results without parallels in MAGICC.

5   The features are, however, consistent in magnitude with forcing and temperature fluctuations due to unforced variability, and therefore are unable to be classified as robust responses. ~~A good example of this is that CM3 exhibits a cooling response to methane from 1900 to 1915 likely due to the formation of the southern ocean polynya leading to very large unforced multidecadal time-scale climate variability.~~ This highlights how unforced variability present in sophisticated models can make it difficult to ascertain robust responses to small changes in multiple forcings individually, further justifying the use of a model

10   such as MAGICC beyond pure accessibility. To overcome this challenge, a larger number of ensembles could be employed or simulations can be run with a quasi-chemistry-transport model (Deckert et al. 2011).

Overall, we find that reduced complexity climate models, with the MAGICC model as an example, ~~a reduced complexity climate model, is~~are able to satisfactorily match the global mean temperature response to increases in isolated greenhouses

15   gases as simulated by the GFDL-CM3, a complex chemistry-climate model. Furthermore, we find that the prominent role of unforced variability in AM3 and CM3 makes it difficult to clearly assess climate responses to small forcing changes, ultimately supporting further use of models like MAGICC, that have little to no unforced variability, for analysing climate responses to future changes in methane emissions.

Code availability

20   The MAGICC v6 model ~~(the executable file)~~ is available for download at: http://www.magicc.org/download upon registration. Only the executable file is available, and not the source code. The user manual can be accessed at: http://wiki.magicc.org/index.php?title=Manual_MAGICC6_Executable. Full model details along with nineteen sets of AOGCM-calibrated parameters used here for the physics-driven ensemble members are found in Meinshausen et al. (2011a). We update the default values of methane and tropospheric ozone radiative efficiency and methane atmospheric lifetime to

25   values in Myhre et al. (2013).

The atmospheric model component (AM3) source code for GFDL CM3 is available here: https://www.gfdl.noaa.gov/am3-model/. The ocean model component (MOPM5) source code for GFDL CM3 is available here: https://www.gfdl.noaa.gov/mom-ocean-model/.

Data availability

15

Results from CM3/AM3 simulations and from the MAGICC model are available from Vaishali Naik (vaishali.naik@noaa.gov) and Ilissa Ocko (icko@edf.org), respectively, upon request.

References

Allen, M. R., Fuglestvedt, J. S., Shine, K. P., Reisinger, A., Pierrehumbert, R. T. and Forster, P. M.: New use of global warming potentials to compare cumulative and short-lived climate pollutants, Nature Climate Change, 6, 773–776, doi:10.1038/nclimate2998, 2016.

Alvarez, R. A., Pacala, S. W., Winebrake, J. J., Chameides, W. L. and Hamburg, S. P.: Greater focus needed on methane leakage from natural gas infrastructure, Proc. Nat. Acad. Sci., 109(17), pp.6435–6440, 2012.

Austin, J., Horowitz, L. W., Schwarzkopf, M. D., Wilson, R. J. and Levy, H.: Stratospheric ozone and temperature simulated from the preindustrial era to the present day, J. Clim., 26(11), pp.3528–3543, 2013.

Brown, P. T., Li, W. and Xie, S. P.: Regions of significant influence on unforced global mean surface air temperature variability in climate models, J. Geophys. Res.: Atmos, 120(2), pp.480-494, 2015.

Brown, P. T., Ming, Y., Li, W. and Hill, S. A.: Change in the magnitude and mechanisms of global temperature variability with warming, Nature Climate Change, 7(10), p.743, 2017.

Collins, W. J., Lamarque, J. F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D. and Smith, S. J.: AerChemMIP: Quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev. Discuss., doi: 10.5194, 2016.

Collins, W. J., Webber, C. P., Cox, P. M., Huntingford, C., Lowe, J., Sitch, S., Chadburn, S. E., Comyn-Platt, E., Harper, A. B., Hayman, G. and Powell, T.: Increased importance of methane reduction for a 1.5 degree target, Environ. Res. Lett., 13(5), p.054003, 2018.

~~De Lavergne, C., Palter, J. B., Galbraith, E. D., Bernadello, R., and Marinov, I.: Cessation of deep convection in the open Southern Ocean under anthropogenic climate change, Nature Clim. Change, 4, 278–282, doi:10.1038/nclimate2132, 2014.~~

Deckert, R., Jöckel, P., Grewe, V., Gottschaldt, K.D. and Hoor, P.: A quasi chemistry-transport model mode for EMAC. Geoscientific Model Development, 4(1), pp.195-206, doi: 10.5194/gmd-4-195-2011, 2011.

Delworth, T. L., Broccoli, A. J., Rosati, A., Stouffer, R. J., Balaji, V., Beesley, J. A., Cooke, W. F., Dixon, K. W., Dunne, J., Dunne, K. A. and Durachta, J. W.: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics, J. Clim., 19(5), pp.643-674, 2006.

Deser, C., Phillips, A., Bourdette, V. and Teng, H.: Uncertainty in climate change projections: the role of internal variability, Clim. Dyn., 38(3-4), 527–546, doi:10.1007/s00382-010-0977-x, 2012.

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., Golaz, J. C., Ginoux, P., Lin, S. J., Schwarzkopf, M. D. and Austin, J.: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3, J. Clim., 24(13), 3484–3519, 2011.

Environmental Protection Agency (EPA): Global Anthropogenic Non-CO2 Greenhouse Gas Emissions: 1990 – 2030, 2012.

Etminan, M., Myhre, G., Highwood, E. J. and Shine, K. P.: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, Geophys. Res. Lett., 43(24), 2016.

Fiore, A. M., Naik, V., and Leibensperger, E. M.: ir quality and climate connections, J. of Air & Waste Management, 65:6, 645–685, doi: 10.1080/10962247.2015.1040526, 2015.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., Andrews, T., Pincus, R. and Schulz, M.: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6. Journal of Geophysical Research: Atmospheres, 121(20), 2016.

Freidenreich, S. M. and Ramaswamy, V.: A new multiple-band solar radiative parameterization for general circulation models, J. Geophys. Res., 104, 31, 389-31, 409, 1999.

Fuglestvedt, J. S., Shine, K. P., Berntsen, T., Cook, J., Lee, D. S., Stenke, A., Skeie, R. B., Velders, G. J. M. and Waitz, I.A.: Transport impacts on atmosphere and climate: Metrics. Atmospheric Environment, 44(37), pp.4648-4677, 2010.

Geophysical Fluid Dynamics Laboratory Global Atmospheric Model Development Team (GAMDT): The new GFDL global atmosphere and land model AM2-LM2: Evaluation with prescribed SST simulations, J. Clim., 17(24), 4641-4673, 2004.

Griffies, S. M., Winton, M., Donner, L. J., Horowitz, L. W., Downes, S. M., Farneti, R., Gnanadesikan, A., Hurlin, W. J., Lee, H. C., Liang, Z. and Palter, J. B.: The GFDL CM3 coupled climate model: characteristics of the ocean and sea ice simulations, J. Clim., 24(13), 3520–3544, 2011.

Golaz, J.C., Horowitz, L.W. and Levy, H., 2013. Cloud tuning in a coupled climate model: Impact on 20th century warming. Geophysical Research Letters, 40(10), pp.2246-2251, 2013.

Golaz, J.-C., Salzmann, M., Donner, L. J., Horowitz, L. W., Ming, Y., and Zhao, M.: Sensitivity of the aerosol indirect effect to subgrid variability in the cloud parameterization of the GFDL Atmosphere General Circulation Model AM3, J. Clim., 24(13), doi:10.1175/2010JCLI3945.1, 2011.

Gordon, A. L., and Comiso, J. C.: Polynyas in the Southern Ocean, Sci. Amer., 256, 90–97, 1988.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K. and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions, Geoscientific Model Development, 5(6), 1471–1492, 2012.

17

Hartin, C.A., Patel, P., Schwarber, A., Link, R.P. and Bond-Lamberty, B.P.: A simple object-oriented and open source model for scientific and policy analyses of the global carbon cycle-Hector v0. 1. Geoscientific Model Development Discussions, 7(5), 2014.

John, J. G., Fiore, A. M., Naik, V., Horowitz, L. W. and Dunne, J. P.: Climate versus emission drivers of methane lifetime against loss by tropospheric OH from 1860–2100, Atmos. Chem. Phys., 12(24), 12021–12036, 2012.

International Energy Agency (IEA): World Energy Outlook, 586, 2015.

Lamarque, J. F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B. and Schultz, M. G.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmos. Chem. Phys, 10(15), 7017–7039, 2010.

Lean, J. L.: Cycles and trends in solar irradiance and climate, Wiley interdisciplinary reviews: climate change, 1(1), 111–122, 2010.

Levy, H. II, Horowitz, L. W., Schwarzkopf, M. D., Ming, Y., Golaz, J.-C., Naik, V. and Ramaswamy, V.: The roles of aerosol direct and indirect effects in past and future climate change, J. Geophys. Res. Atmos., 118, 4521–4532, doi:10.1002/jgrd.50192, 2013.

Lin, M., Fiore, A. M., Horowitz, L. W., Cooper, O. R., Naik, V., Holloway, J., Johnson, B. J., Middlebrook, A. M., Oltmans, S. J., Pollack, I. B. and Ryerson, T. B.: Transport of Asian ozone pollution into surface air over the western United States in spring, J. Geophys. Res. Atmos., 117(D21), 2012.

Meinshausen, M., Meinshausen, N., Hare, W., Raper, S. C. B., Frieler, K., Knutti, R., Frame, D. J. and Allen, M.R.: Greenhouse-gas emission targets for limiting global warming to 2°C, Nature, 458(7242), 1158, 2009.

Meinshausen, M., Raper, S. C. B. and Wigley, T. M. L.: Emulating IPCC AR4 atmosphere-ocean and carbon cycle models for projecting global-mean, hemispheric and land/ocean temperatures: MAGICC 6.0, Atmos. Chem. Phys., 8(2), 6153–6272, 2008.

Meinshausen, M., Raper, S. C. and Wigley, T. M.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6–Part 1: Model description and calibration, Atmos. Chem. Phys., 11(4), 1417–1456, 2011a.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J. F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic change, 109(1-2), 213, 2011b.

Meinshausen, M., Wigley, T. M. L. and Raper, S. C. B.: Emulating atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6–Part 2: Applications, Atmos. Chem. Phys., 11(4), 1457–1471, 2011c.

Melvin, A. M., Sarofim, M. C., and Crimmins, A. R.: Climate Benefits of U.S. EPA Programs and Policies That Reduced Methane Emissions 1993–2013, Environ. Sci. Technol., 50(13), 6873–6881, DOI: 10.1021/acs.est.6b00367, 2016.

Ming, Y., Ramaswamy, V., Donner, L. J. and Phillips, V. T.: A new parameterization of cloud droplet activation applicable to general circulation models, J. Atmos. Sci., 63(4), 1348–1356, 2006.

Myhre, G., Highwood, E. J., Shine, K. P., and Stordal, F.: New estimates of radiative forcing due to well mixed greenhouse gases, Geophys. Res. Lett., 25, 2715–2718, 1998.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocke,.:/,,,r, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Naik, V., Voulgarakis, A., Fiore, A. M., Horowitz, L. W., Lamarque, J. F., Lin, M., Prather, M. J., Young, P .J., Bergmann, D., Cameron-Smith, P. J. and Cionni, I.: Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13(10), 5277–5298, 2013.

NOAA National Centers for Environmental Information, Climate at a Glance: Global Time Series, published May 2017, retrieved on June 16, 2017 from http://www.ncdc.noaa.gov/cag/.

Ocko, I. B., Hamburg, S. P., Jacob, D. J., Keith, D. W., Keohane, N. O., Oppenheimer, M., Roy-Mayhew, J. D., Schrag, D. P. and Pacala, S. W.: Unmask temporal trade-offs in climate policy debates, Science, 356(6337), 492–493, 2017.

Ocko, I. B., Ramaswamy, V. and Ming, Y.: Contrasting climate responses to the scattering and absorbing features of anthropogenic aerosol forcings, J. Clim., 27(14), pp.5329-5345, 2014.

Paynter, D., Frölicher, T.L., Horowitz, L.W. and Silvers, L.G.: Equilibrium climate sensitivity obtained from multimillennial runs of two GFDL climate models. Journal of Geophysical Research: Atmospheres, 123(4), pp.1921-1941, 2018.

Ramanathan, V. and Xu, Y.: The Copenhagen Accord for limiting global warming: Criteria, constraints, and available avenues, Proc. Nat. Acad. Sci., 107(18), 8055–8062, 2010.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C. and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res. Atmos., 108(D14), 2003.

Rogelj, J., McCollum, D. L., Reisinger, A., Meinshausen, M. and Riahi, K.: Probabilistic cost estimates for climate change mitigation, Nature, 493(7430), 79–83, 2013.

Schwarzkopf, M. D. and Ramaswamy, V.: Radiative effects of CH4, N2O, halocarbons and the foreign-broadened H2O continuum: A GCM experiment, J. Geophys. Res., 104, 9467-9488, 1999.

Shindell, D. T., Faluvegi, G., Koch, D. M., Schmidt, G. A., Unger, N. and Bauer, S. E.: Improved attribution of climate forcing to emissions, Science, 326(5953), 716–718, 2009.

Shindell, D., Kuylenstierna, J. C., Vignati, E., van Dingenen, R., Amann, M., Klimont, Z., Anenberg, S. C., Muller, N., Janssens-Maenhout, G., Raes, F. and Schwartz, J.: Simultaneously mitigating near-term climate change and improving human health and food security, Science, 335(6065), 183–189, 2012.

Shine, K. P., Cook, J., Highwood, E. J. and Joshi, M. M.: An alternative to radiative forcing for estimating the relative importance of climate change mechanisms. Geophysical Research Letters, 30(20), 2003.

Shine, K., Derwent, R., Wuebbles, D., and Morcrette, J.-J.: Radiative forcing of climate, in: Climate Change: The IPCC Scientific Assessment, edited by Houghton, J., Jenkins, G., and Ephraums, J., Cambridge University Press, New York, USA, 41–68, 1990.

Shine, K.P., Fuglestvedt, J.S., Hailemariam, K. and Stuber, N.: Alternatives to the global warming potential for comparing climate impacts of emissions of greenhouse gases. Climatic Change, 68(3), pp.281-302, 2005.

Shoemaker, J. K., Schrag, D. P., Molina, M. J. and Ramanathan, V.: What role for short-lived climate pollutants in mitigation policy?, Science, 342(6164), 1323–1324, 2013.

Solomon, S., Daniel, J. S., Sanford, T. J., Murphy, D. M., Plattner, G.-K., Knutti, R., Friedlingstein, P.: Persistence of climate changes due to a range of greenhouse gases, Proc. Natl. Acad. Sci.; 107(43), 18354–18359, doi: 10.1073/pnas.1006282107, 2010.

Stenchikov, G., K. Hamilton, R. J. Stouffer, A. Robock, V. Ramaswamy, B. Santer, and H.-F. Graf, 2006: Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models. J. Geophys. Res., 111, D07107, doi:10.1029/2005JD006286

Stevenson, D. S., Young, P. J., Naik, V., Lamarque, J.F., Shindell, D. T., Voulgarakis, A., Skeie, R. B., Dalsoren, S. B., Myhre, G., Berntsen, T. K. and Folberth, G. A.: Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13(6), 3063–3085, 2013.

Stocker, T. F., Qin, D., Plattner G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M.: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge: Cambridge University Press, 2013.

West, J. J., Smith, S. J., Silva, R. A., Naik, V., Zhang, Y., Adelman, Z., Fry, M. M., Anenberg, S., Horowitz, L. W., Lamarque, J.-F.: Co-benefits of mitigating global greenhouse gas emissions for future air quality and human health, Nature Climate Change, 3(10), 885–889, http://doi.org/10.1038/NCLIMATE2009, 2013.

Wigley, T. M. L. and Raper, S. C. B.: Thermal-Expansion of Sea-Water Associated with Global Warming, Nature, 330, 127–131, 1987.

Wigley, T. M. L. and Raper, S. C. B.: Implications for climate and sea level of revised IPCC emissions scenarios, Nature, 357, 293–300, 1992.

Zhang, Y., Bowden, J. H., Adelman, Z., Naik, V., Horowitz, L. W., Smith, S. J. and West, J. J.: Co-benefits of global and regional greenhouse gas mitigation for US air quality in 2050. Atmospheric Chemistry and Physics, 16(15), 9533–9548, 2016.

	Experiments	Abbreviation	MAGICC v6	GFDL CM3
	All-Forcing	AllForc	X	X
	CO_2 concentrations held constant at 1860 levels	$CO_2$1860	X	X
	Methane concentrations held constant at 1860 levels	$CH_4$1860	X	X
Direct Simulations	All-Forcing with methane chemistry turned off	CH_4nochem	X	
	Methane concentrations held at 1860 levels with methane chemistry turned off	$CH_4$1860nochem	X	
	Methane concentrations held at 1860 for radiation	$CH_4$1860chem		X
	Methane concentrations held at 1860 for chemistry	$CH_4$1860chem		X
	CO_2-only		AllForc – $CO_2$1860	
Derived Simulations	CH_4-only		AllForc – $CH_4$1860	
	CH_4-direct		CH_4nochem – $CH_4$1860nochem	AllForc – $CH_4$1860rad
	CH_4-indirect		CH_4-only – CH_4-direct	AllForc – $CH_4$1860chem

Formatted: Subscript

Table 1. Direct experiments performed by MAGICC and GFDL-CM3 models, as well as derived simulations. All-Forcing simulations include time-varying natural and anthropogenic forcings but land-use held constant. Each experiment is run for 19 physics-driven ensemble members for MAGICC and three initial condition-driven ensemble members for GFDL-CM3 over the period 1860 – 2014.
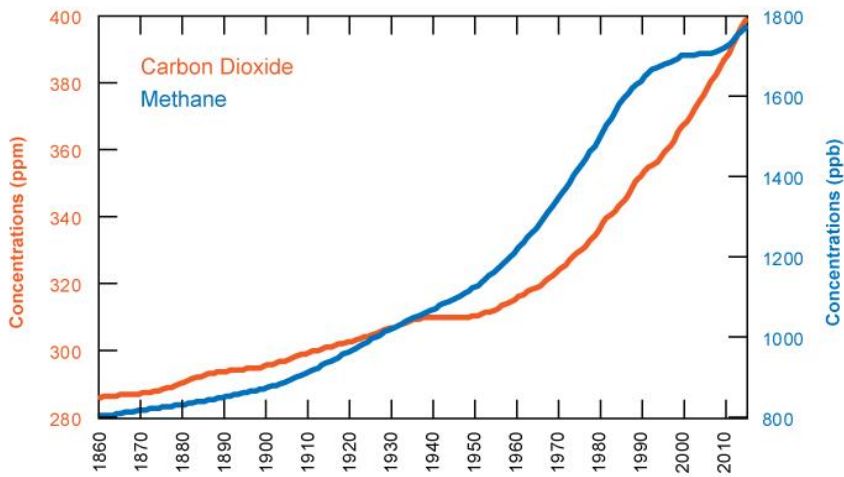
Carbon Dioxide
Methane

Concentrations (ppm)

Concentrations (ppb)

Figure 1. Atmospheric concentrations of carbon dioxide in parts per million (orange) and methane in parts per billion (blue) used in this study (Meinshausen et al. 2011b). Note that concentrations are prescribed for CM3 throughout this time period, but only prescribed for MAGICC through 2005, of which methane emissions inputs drive the model from 2006-2014. The resulting concentrations are plotted here.
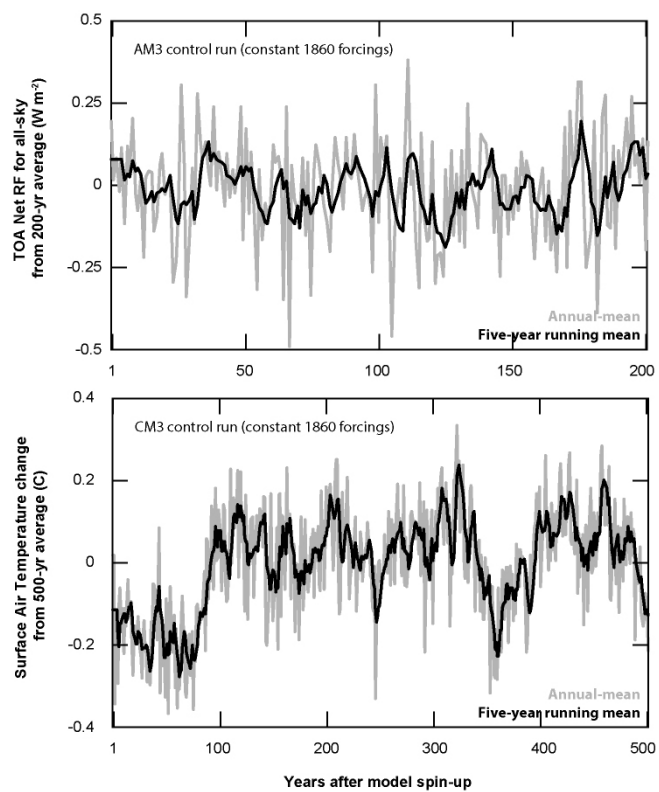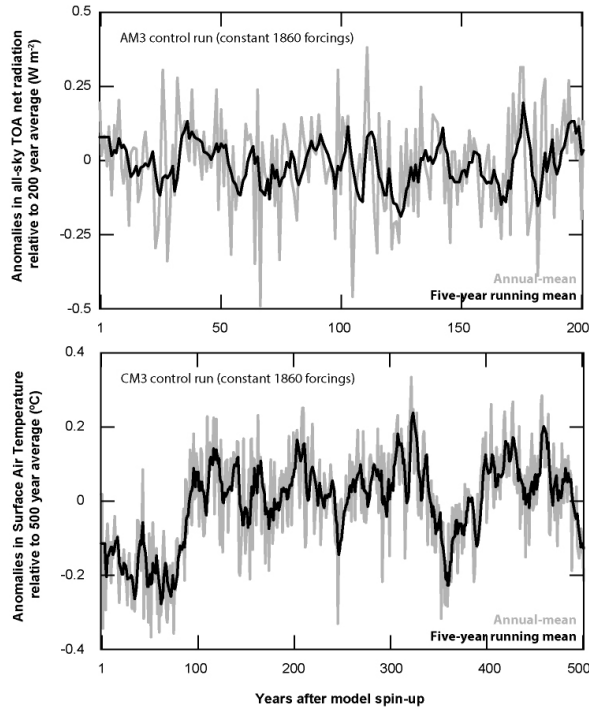
Figure 2. Anomalies in (a) all-sky top-of-the-atmosphere (TOA) net radiation (Wm^{-2}) with respect to 200 year mean and (b) surface air temperature (ºC) with respect to 500 year mean of the control simulation of AM3 and CM3 with preindustrial (1860) forcings held constant post spin-up, respectively. Fluctuations indicate unforced variability. ~~Control run simulations for AM3 and CM3 with preindustrial (1860) forcings held constant for 200 and 500 years post- spin-up, respectively. Top-of-atmosphere~~ effective ~~radiative forcing fluctuations for all-sky conditions in W m^{-2} introduced by unforced variability in AM3. Surface air temperature fluctuations in ºC introduced by unforced variability in CM3.~~ **Results** are **shown for annual averages** (grey line) **and five-year running means** (black line).

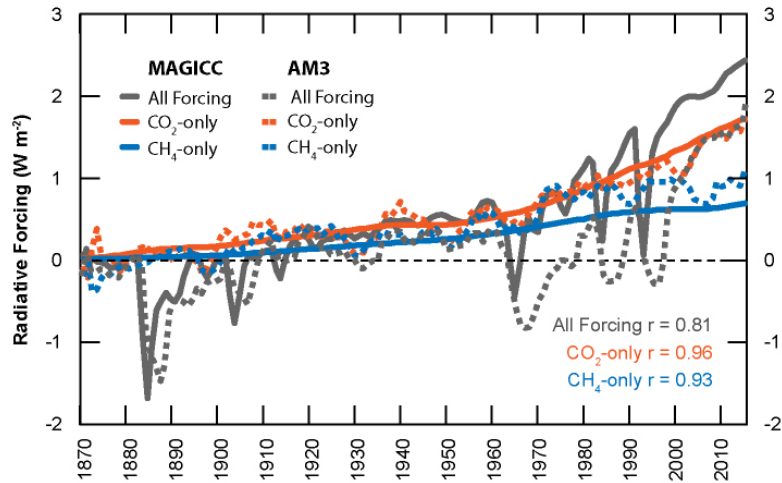**Figure 3. Radiative forcings (W m^{-2}) after stratospheric adjustment due to all forcing (grey), CO$_2$-only (orange), methane-only (blue), for both AM3 (dashed) and MAGICC (solid) model simulations. Methane forcing includes its direct as well as indirect effect from influences on chemistry. AM3 radiative forcings are 'effective' radiative forcings (ERF), and include tropospheric adjustments as well, and are calculated at the top-of-atmosphere (TOA). MAGICC radiative forcings are calculated at the tropopause. AM3 data are 5-year running means. Correlation coefficients between MAGICC and AM3 ~~radiative~~ forcings are shown inset.**
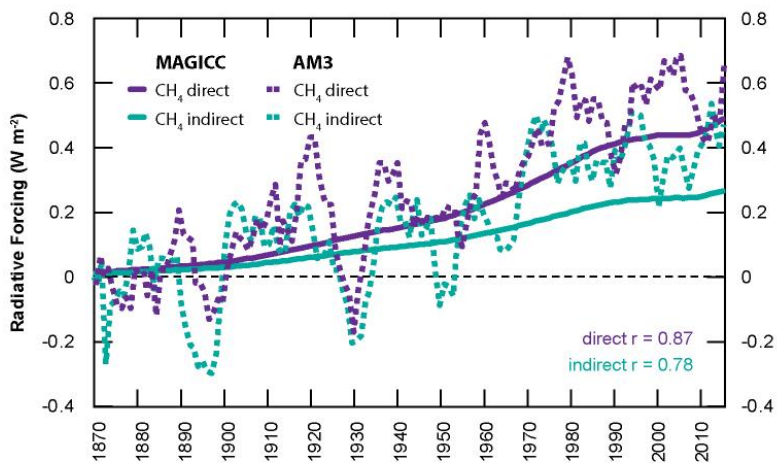
Figure 4. Same as Figure 3, but showing Ddirect (purple) and indirect (from methane's influence on ozone and water vapour, green) radiative forcings (W m^{-2}) after stratospheric adjustment, for both AM3 (dashed) and MAGICC (solid) model simulations. AM3 radiative forcings are technically 'effective' radiative forcings, and include tropospheric adjustments as well, and are calculated at the top-of-atmosphere. MAGICC radiative forcings are calculated at the tropopause. AM3 data are 5-year running means. Correlation coefficient between MAGICC and AM3 forcings are also shown.
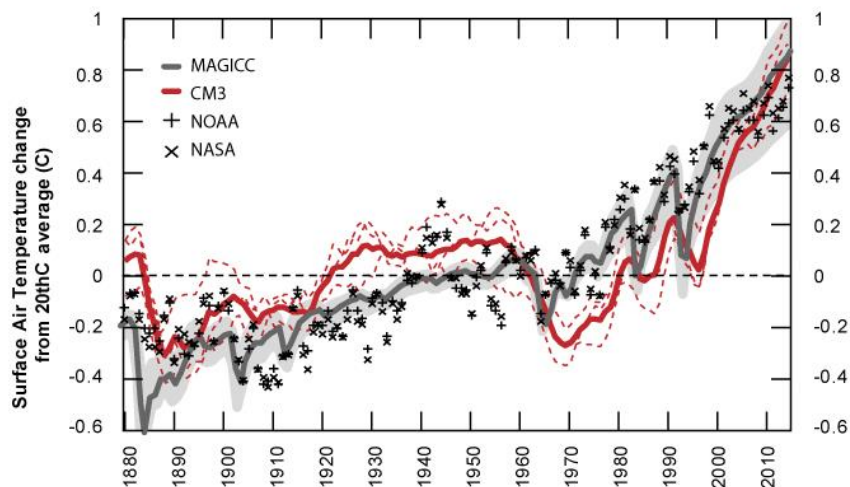
Figure 5. All forcing global-mean surface air temperature responses in ºC for CM3 (solid red line) and MAGICC (solid grey line) model simulations as compared to observations by NOAA (+) (https://data.giss.nasa.gov/gistemp/) and NASA (x) (https://www.ncdc.noaa.gov/cag/time-series/global). All annual temperature anomalies shown as change from 20th Century average for each dataset. Individual initial condition-driven ensemble members for CM3 runs shown in thin dashed red lines. Physics-driven ensemble-member range for MAGICC shown as shaded grey. CM3 data are 5-year running means.
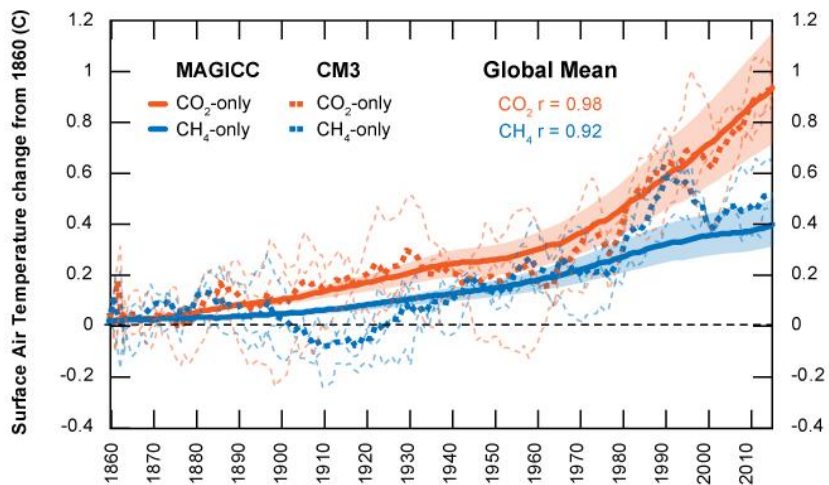
Figure 6. Global mean surface air temperature responses in ºC for CM3 (dashed line) and MAGICC (solid line) model derived simulations – CO₂-only (orange) and methane-only (blue). Individual initial condition-driven ensemble members for CM3 runs shown in thin dashed lines. Range for MAGICC physics-driven ensemble members shown in shaded colours. CM3 data are 5-year running means. Correlation coefficient between MAGICC and CM3 temperature responses are also shown.
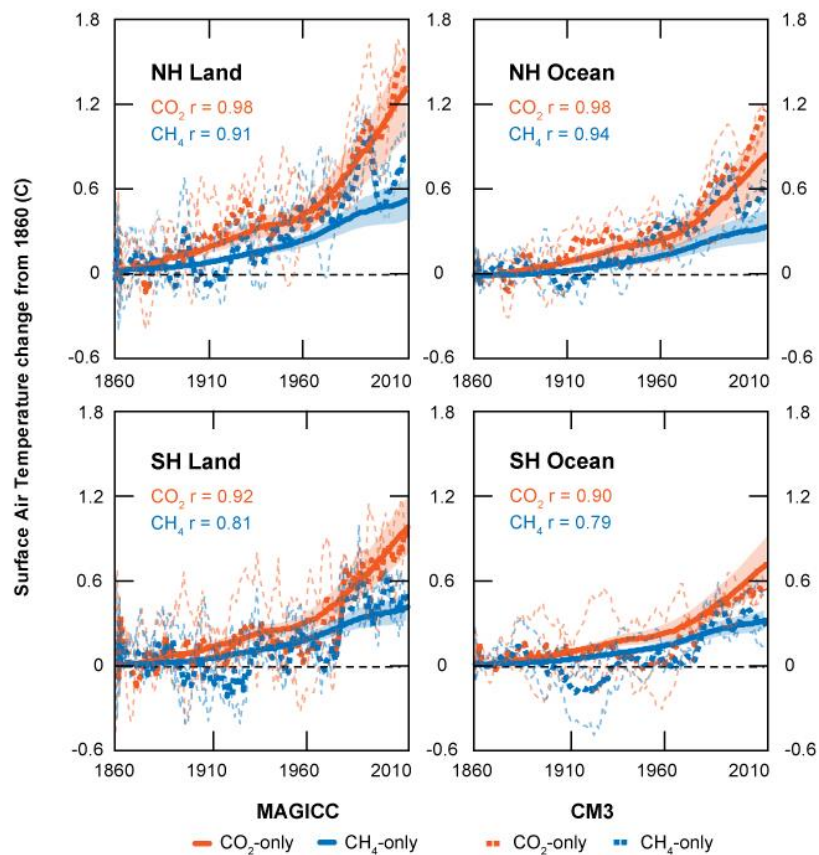
Figure 7. Regional surface air temperature responses in ºC for CM3 (dashed line) and MAGICC (solid line) model indirect simulations – CO₂-only (orange) and methane-only (blue). Individual initial condition-driven ensemble members for CM3 runs shown in thin dashed lines. Range for MAGICC physics-driven ensemble members shown in shaded colours. CM3 data are 5-year running means. Correlation coefficient between MAGICC and CM3 temperature responses are also shown.

30