#Reviewer 1

This study estimated a large set of marginal abatement cost (MAC) curves based on the output of IAMs in the ENGAGE Scenario Explorer and the GET model. The MAC curves were then applied to the emulator for Integrated Assessment Models (emIAM) and coupled to a simple climate model, ACC2. The test results showed that emIAM was able to reproduce the original IAM emission outcomes under similar conditions. The topic provided rich information about MAC curves under various IAMs, as well as different regions evaluated in the manuscript. While I agree with the authors that the analysis provided by the authors is certainly of general interest to climate-economic model developers and climate-focused researchers, I unfortunately cannot recommend publication of the manuscript in its present form. Here are my concerns:

[Response] We thank the reviewer for taking the time to read our manuscript and for providing useful comments. We have carefully revised the manuscript based on the reviewer's comments.

First, the authors reviewed a range of existing literature about the categories of MAC curves and different MAC curves estimated under various backgrounds. However, the results and analysis generally focused on the outcome of this study. I recommend adding a comparison between the estimated MAC curves in this study and those presented in existing studies, including differences in function forms, appropriate interpretation of parameters, and other major differences compared to existing estimates.

[Response] We selected several previous studies using MAC curves and compared them with ours (see Figure 4 and Figure 7). We added the following text to the manuscript:

The functional form of the MAC function used by Su et al. (2017) is consistent with our study, and Tanaka et al. (2021) used equation (2) in Table S2. Harmsen et al. (2019) considered time-dependent MAC curves and no explicit function is provided. Despite some differences in the form of the functions, the MAC curves for energy-related CO_2 used in Su et al. (2017) and Tanaka et al. (2021) are within the range of the MAC curves from ENGAGE IAMs, but the MAC curves for CH_4 and N_2O used in Tanaka et al. (2021) show a higher level of marginal carbon price. Harmsen et al. (2019) show that CH_4 MAC curve in 2050 is also to the left of our results, but that in 2100 are close to our study. Meanwhile, their results for N_2O are in the middle of our results and not much different between 2050 and 2100.

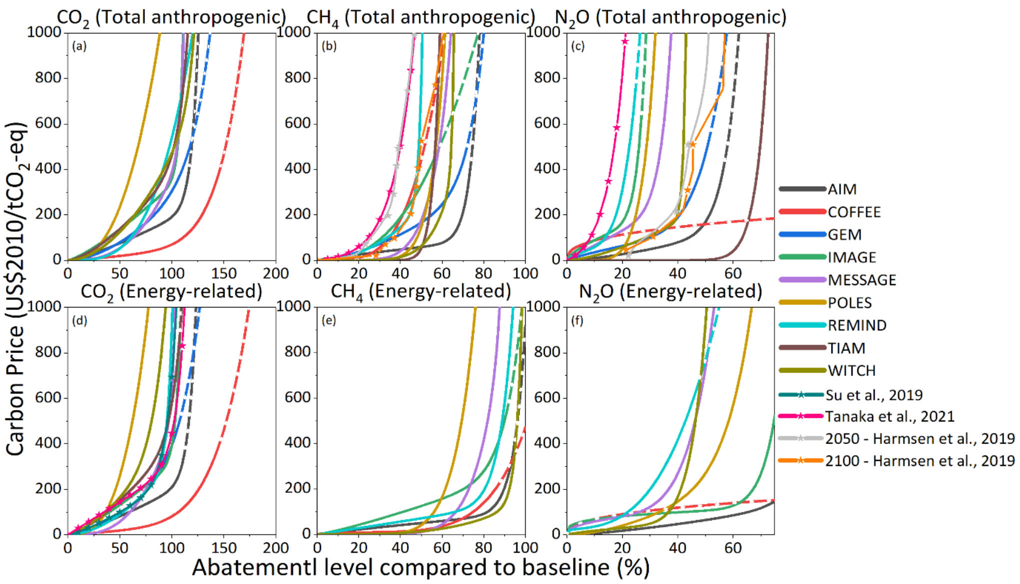

Figure 1. Global MAC curves for total anthropogenic and energy-related CO_2, CH_4, and N_2O emissions derived from nine ENGAGE IAMs. In panels (a) to (f), the solid line indicates that the MAC curve is within the applicable range; the dashed line means that it is outside the applicable range (i.e., above the maximum abatement level indicated from underlying IAM simulation data or above the range of carbon prices considered for fitting the MAC curve; see Tables 1 and 2). Different colors indicate different IAMs. The MAC curves from selected previous studies (Su et al., 2017; Harmsen et al., 2019; Tanaka et al., 2021) are shown for comparison. The MAC curves from Harmsen et al., (2019) are time-dependent and the figure shows those for the years 2050 and 2100.

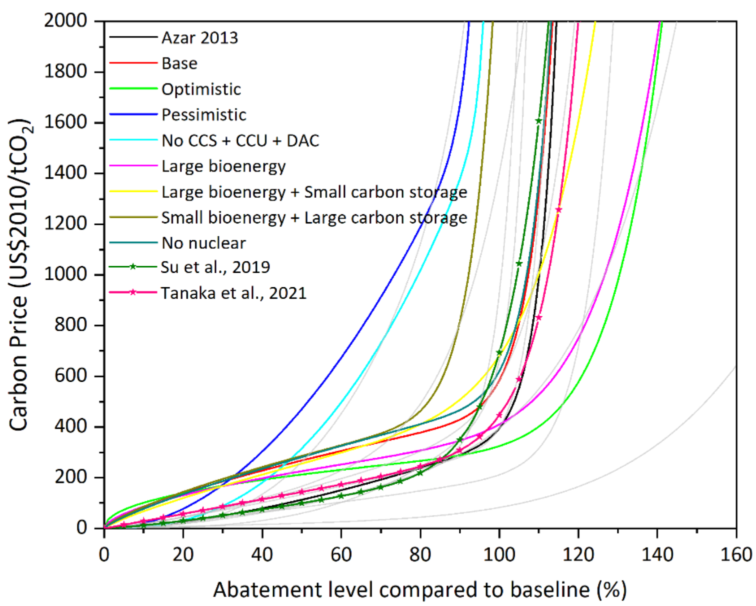

Figure 7. Global MAC curves for energy-related CO_2 emissions derived from the GET model with different portfolios of available mitigation technologies. Different colors indicate

different technology portfolios (see Section 2.2 for details). Global MAC curves for energy-related CO_2 emissions from ENGAGE IAMs are shown as a comparison in gray lines, and the MAC curves from selected previous studies (Su et al., 2017; Tanaka et al., 2021) are shown in star-shaped lines.

Second, the ENGAGE Scenario dataset includes a wide range of outputs from various IAMs and regions. I am not quite sure about the reasons why the output of a separate GET model was also used for estimating the MAC curves. An explanation of the necessity of adding the output of the GET model is needed for readers to understand the framework of this study more clearly.

[Response] It is correct that the ENGAGE scenario database covers a broad range of output from various IAMs. However, we still simulated the GET model and used the output to additionally explore the effect of technological assumptions (e.g. CCS capacity) on the MAC curves. This was possible only with GET because we have a capability to run the GET model as needed, but this was not possible with ENGAGE IAMs because no such output is included in the ENGAGE Scenario Explorer (i.e. technological assumptions are kept the same in each model when it is simulated under different carbon prices). The motivation for using GET was already stated in the initial manuscript:

We further apply the emIAM approach to the GET model (Lehtveer et al., 2019), an IAM that did not take part in the ENGAGE project. We can directly simulate GET to derive MAC curves under different model configurations, which complements the existing data from IAMs simulated under single configurations for the ENGAGE project.

Third, the manuscript mentions that the emIAM-ACC2 model minimized total abatement costs to obtain possible emission pathways for reproducing the outcomes from other IAMs. More information about how this process works is needed, including the necessary equations and the objective function for minimizing.

[Response] We thank the reviewer for the suggestion. We have added a more detailed description of this process in Section 4.1, as in the following statements:

More specifically, ACC2 uses equation (2) to calculate the abatement costs (ABC) of regions (or global total), gases, and years.

$$ABC_{t,r,g} = Eb_{t,r,g} \cdot \int_0^x f_{t,r,g}(x)dx \qquad (2)$$

where t, r, g represent year, region, and gas, respectively. x is the abatement level compared to the baseline scenario. $f_{t,r,g}(x)$ is the MAC function. Eb is the baseline emission level for the IAM. The objective of the model is to minimize the net present value of the total abatement cost ($TABC$), that is:

$$\min TABC = \sum_{t,r,g} \frac{ABC_{t,r,g}}{(1+DSC)^{t-t0}} \tag{3}$$

where DSC is the discount rate and $t0$ represents the base year used for abatement cost calculations (2010 in this study).

In this study, we replace the existing set of MAC curves in ACC2 with the global and regional MAC curves obtained in this study. We also replace the limits on abatement (i.e., upper limits on abatement levels and their first and second derivatives) with those obtained from this study. We assume a 5% discount rate in the validation tests, a rate commonly assumed in IAMs (Emmerling et al., 2019), which is also consistent with some of the IAMs analyzed here such as MESSAGE and GET (Figures SI 1.2-1 and 1.2-2 of Riahi et al. (2021)). But we were unable to find the discount rates used in the other IAMs. Note that a 4% discount rate was used as default in recent studies using ACC2 (Tanaka and O'Neill, 2018; Tanaka et al., 2021) We consider the mitigation costs through 2100 in scenario optimizations.

Fourth, this study provided many figures (some of which are similar) to present the estimations of the MAC curves and the emulating results, especially in the Supplement. While these figures provide visual information to present relevant results, there are too many figures stacked together, making it difficult for readers to find the information they need. An appropriate way to manage these figures, such as indexing them using tables or other means of relevance, should be added.

[Response] We thank the reviewer for the suggestion. We have added a list of tables and figures to Supplement so that it is easier for readers to find the relevant content.

#Reviewer 2

Xiong et al have developed an emulator for integrated assessment models (IAMs) using a "marginal abatement cost (MAC)" approach. The emulator uses a large set of MACs derived from IAM-based scenarios in an existing database to reproduce most original IAM emission outcomes at a much lower computational cost than the original model. Additionally, the emulator can be coupled to a simple climate model to generate emission pathways for a specific temperature target. In general, this is a positive modeling development, as emulators are common in various fields, including climate models, but are currently lacking in the IAM field. As IAMs continue to advance in complexity, emulators could be valuable in scenario discovery.

[Response] We thank the reviewer for recognizing the future potential of our work and for providing comments that were useful for improving the quality of our manuscript.

While this study represents one of the first attempts to develop an IAM emulator, there are three main areas for improvement, as summarized below and discussed in detail.

Firstly, the overall flow should be better. Sometimes, the details are provided before a general overview, creating challenges for readers.

[Response] We thank the reviewer for pointing out this problem. We have substantially revised the paper structure to improve the flow of the paper.

Secondly, the visualization could be improved. Many figures are too busy to deliver the critical message.

[Response] We thank the reviewer for the suggestion. We have made systematic refinement of the visualization of the figures so that they present the desired information more clearly. For example, we improved the layout of Figure 1 (see below), with panel c presenting results for each decade (rather than every five years) in larger subpanels. Panel d, which shows MAC curves, is now shown in a larger format. In addition, we have also improved the legend of the figure. We think the legend is now easier to understand.

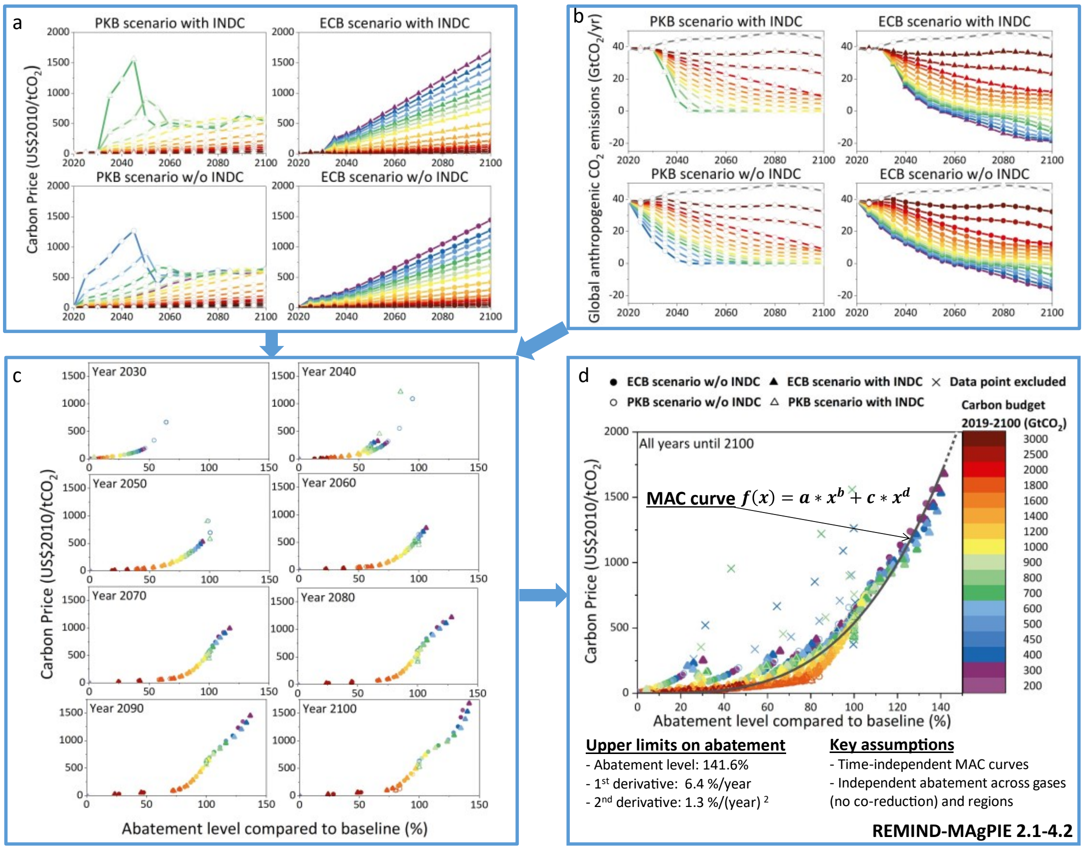

Figure 2. Overview of the methods to derive MAC curves and limits on abatement (upper limits on abatement levels and their first and second derivatives). The figure uses the data for global total anthropogenic CO_2 emissions from REMIND for illustration. The chromatic colors indicate the respective carbon budgets for the period 2019 – 2100 in $GtCO_2$. The grey color indicates the baseline scenario ("NPi2100" in the original scenario name). Scenarios without INDC consider currently implemented national policies (circle; indicated as "NPi2020" in the original scenario name); scenarios with INDC further consider national emission pledges until 2030 (triangle; indicated as "INDCi2030" in the original scenario name). ECB scenarios consider carbon budgets till the end of this century, with a possibility of temporal budget overspending (filled circles; with "f" in the original scenario name); PKB scenarios consider carbon budgets without allowing temporal budget overspending (open circles; without "f" in the original scenario name). Crosses indicate data points from scenarios that were not considered in the derivation of the MAC curve (i.e., EN_INDCi2030_700, EN_INDCi2030_800, EN_NPi2020_400, and EN_NPi2020_500 for REMIND (see Table 1)). In the equation of the MAC curve, a, b, c, and d are the parameters to be optimized; x is the variable representing the abatement level in percentage relative to the assumed baseline level). Note that panel c shows data only for every ten years for the sake of presentation.

Lastly, while I appreciate the massive details and results provided by authors, most of the result text was just purely describing the results, without a high-level generalization or explanation of the reason behind the findings. There is little discussion about the model structures, which might help explain the results.

[Response] We thank the reviewer for the suggestion. We have substantially expanded Section 5 to provide a high-level generalization. Regarding the model structures, we have the following discussion in Section 3:

The results vary in terms of the range of carbon prices, the range of abatement levels, and the dispersion of data points. For example, the carbon prices of AIM and COFFEE remain below \$500/tCO_2, while the carbon prices of POLES and MESSAGE can exceed \$5,000/tCO_2. The maximum abatement levels of COFFEE and REMIND are over 150%, while others are in the range of 100%-120%. AIM provides a limited amount of data at low abatement levels. IMAGE and POLES produce more dispersed data distributions than other models, which may be related to the fact that these models are recursive dynamic models (Table 1); however, the other recursive dynamic models, AIM and GEM, produce less dispersed data distributions that can be well captured by MAC curves. POLES can be seen as an example where our time-independent MAC curve approach does not work well. The MAC curve, if taken every five years, shifts to the right over time (Figure S4).

Detailed comments:

1. Line 108: please explain what's the NPi2100 scenario. Previous sentences mentioned other scenarios like NPi2020 and INDCi2030, but not NPi2100.

[Response] We have added an explanation about what NPi2100 is in that sentence. Here, NPi2100 is our reference scenario that assumes a continuation of the current stated policies until 2100.

2. Section 2 breaks the entire flow. First, it's unclear what precisely the MAC is in this context. Some experienced readers might generally know a MAC as a function between the carbon price and % emission reductions. Still, different kinds of literature might have different definitions (i.e., carbon price or emission price) or sectoral and gas specifications. This critical "background" information did not show up until Section 3.1. So before diving into the IAM and overwhelming scenarios definitions, this paper could benefit from a high-level schematic showing the entire working flow. (BTW, the current Fig.1 is overwhelming, with many texts and details but somewhat unclear logic).

[Response] We thank the reviewer for the comment. Given the general structure of our paper, we think that the discussion on scenarios (Section 2) should come before the discussion of MAC curves (Section 3). In Section 1, we have a paragraph that introduces the general concept of MAC curves and why the MAC curve approach was used to conduct these studies in the Introduction. We have further added the following text in Section 1: "*In the context of climate change mitigation, a MAC generally represents the incremental cost of reducing an additional unit of emissions; a MAC curve illustrates these costs as the level of emission reductions increases relative to the baseline.*" We have modified Figure 1 to more clearly present the methodological flow. We put the description of the paper structure at the end of Section 1. In revising the manuscript, we kept in mind that the paper structure should be clearer (we further made use of footnotes where necessary). We hope that our revision adequately addresses the reviewer's concern.

3. From section 2, It's unclear why this paper needs the GET model in addition to the ENGAGE scenario database.

[Response] This point was also raised by reviewer #1. The reason why we use the GET model is that we can directly simulate the GET model to explore the effect of technological assumptions on the MAC curves. Though the ENGAGE Scenario Explorer provides a large number of scenarios that show the carbon price pathways under different carbon budgets, it is not suited for the type of analyses that can be possible with GET. While we have a capability of simulating GET, we cannot directly simulate the IAMs in the ENGAGE Scenario Explorer and can only use existing output from these IAMs. The GET model provides a set of CO_2 emission pathways due to the change of carbon prices under different technical portfolios, which can complement the output of the ENGAGE IAMs. Therefore, we use both ENGAGE IAMs and the GET model for this study.

4. Line 157: "if there are non-zero carbon prices in baseline, we subtracted them from the carbon price in mitigation scenarios", is this implicitly assuming a linear relationship between CO2 price and emission reductions? i.e., a linear MAC?

[Response] We thank the reviewer for the question. All our MAC curves are nonlinear as described in Section 3.1. The carbon price for each case is also the relative level to the baseline scenario. The small corrections that the reviewer pointed out should not influence the functional form of the MAC curve.

5. Line 163: I know the term "portfolio" is clearly defined in the GET modeling part in Section 2.2, but what does the "portfolio" mean in the ENGAGE scenario database?

[Response] A portfolio is a set of technological assumptions in the GET model. In the ENGAGE Scenario Explorer, there is only one portfolio for each IAM, so the portfolio is irrelevant to ENGAGE IAMs. Therefore, we revised the statement as follows:

for all cases (i.e., models, gases, regions, and sources in ENGAGE, and portfolios in GET).

6. Line 165 and below: what exactly does this functional form mean? Again, this is breaking the flow, as I saw additional explanations 30 lines below in line 192.

[Response] We have revised this section to improve the flow of the argument. Meanwhile, we have further explained this function, and the definition of each parameter has been clarified as well.

7. Line 165, where is the carbon price in this equation (1)? I guess the carbon price is f(x), but the text below, albeit with many details embedded, did not indicate which term represents the carbon price.

[Response] Yes, the carbon price is f(x), which means that the carbon price is a function of the abatement level. We added a further explanation for this equation to the following statements:

a, b, c, and d are the parameters to be optimized in each case. x is the variable representing the emission abatement level in percentage relative to the assumed baseline level. The carbon price (i.e., $f(x)$ in equation (1)) is expressed in per ton of CO_2-equivalent emissions, using GWP100 (28 and 265 for CH_4 and N_2O, respectively (IPCC, 2013)) to convert CH_4 and N_2O emissions, as assumed in the IAMs emulated here (Harmsen et al., 2016). GWP100 is effectively the default emission metric used to convert non-CO_2 GHG emissions to the common scale of CO_2 and has been used for decades in multi-gas climate policies and assessments, including the Paris Agreement (Lashof and Ahuja, 1990; Fuglestvedt et al., 2003; Tanaka et al., 2010; Tol et al., 2012; Levasseur et al., 2016; UNFCCC, 2018, 2023).

8. Line 199, "performing consistently the best for all IAMs (see the Zenodo repository)". This crucial result needs at least a supplementary table or figure or even a main figure/table.

[Response] We thank the reviewer for the suggestion. We have added the table below in Supplement to show this result:

Table S3. Statistics for function choices

Function	Count	Percentage (%)
T1	127	51.42
T2	15	6.07
T3	45	18.22
T4	60	24.29
Total	**247**	**100**

9. Line 208-209: why do the maximum first and second derivatives of temporal change in abatement levels correspond roughly to the limit of the technological change rate and the socio-economic inertia?

[Response] We thank the reviewer for the question. The technological change rate of mitigation measures shows limitations in the speed of implementation, while socio-economic inertia interferes with the rate of technology change by revealing that some systems need more time to change and adapt (Schwoon and Tol, 2006; Harmsen et al., 2019; Hof et al., 2021). In our study, we interpret technological change rate as the first derivative of the abatement levels, and socio-economic inertia as the second derivative. Therefore, the upper limits of the first and second derivatives, derived from individual model behavior, respectively represent the peak rate of technological change and socio-economic inertia for the entire sample of the IAM.

10. Can you show the x- and y-axis in the same scale? (so that we can see how MACs shift in time)

[Response] We thank the reviewer for the comment. We have expanded the sizes of the subpanels in panel c by showing the results for each decade, and we have also used the same scale of the x- and y-axis for all subpanels in panel c. It can now be easier to see how the data shifts over time. The revised figure was copied as part of our response to the second general comments.

11. Line 248: "crosses in the right panel of Figure 1" --- I cannot find crosses in the right panel because they are too small.

[Response] We thank the reviewer for the comment. In order to improve the clarity of the information displayed in Figure 1, we have adjusted Figure 1 so that the subpanels are larger and the information is easier to see. Please refer to the revised Figure 1 above.

12. In figure 2, the authors pointed out different models show very different carbon prices for the same level of reduction. For example, when reaching a 100% reduction, the corresponding price is about \$150, while POLES is about \$1000. However, this could be the masked effect of the single fitted line on a wide range of scenarios. Even the fitted value for POLES indicated an ~\$1000 to achieve a 100% reduction, there are individual data points (scenarios) reaching 100% reduction with much lower prices. For this type of data and distribution, perhaps the MAC approach is not suitable because of the nature of some particular models.

[Response] We thank the reviewer for the comment. We examined the data from POLES in more detail. We realized that the POLES model offers a sort of failed example of our MAC curve approach, in which the MAC curve, if taken every five years, shifts to the right over time (see Figure S4). In such cases, a time-independent MAC curve is not a proper approach to capturing the emission behavior of the model. We nevertheless present the results because a motivation of this study is to understand to what extent our general MAC curve approach can emulate the behavior of various IAMs.

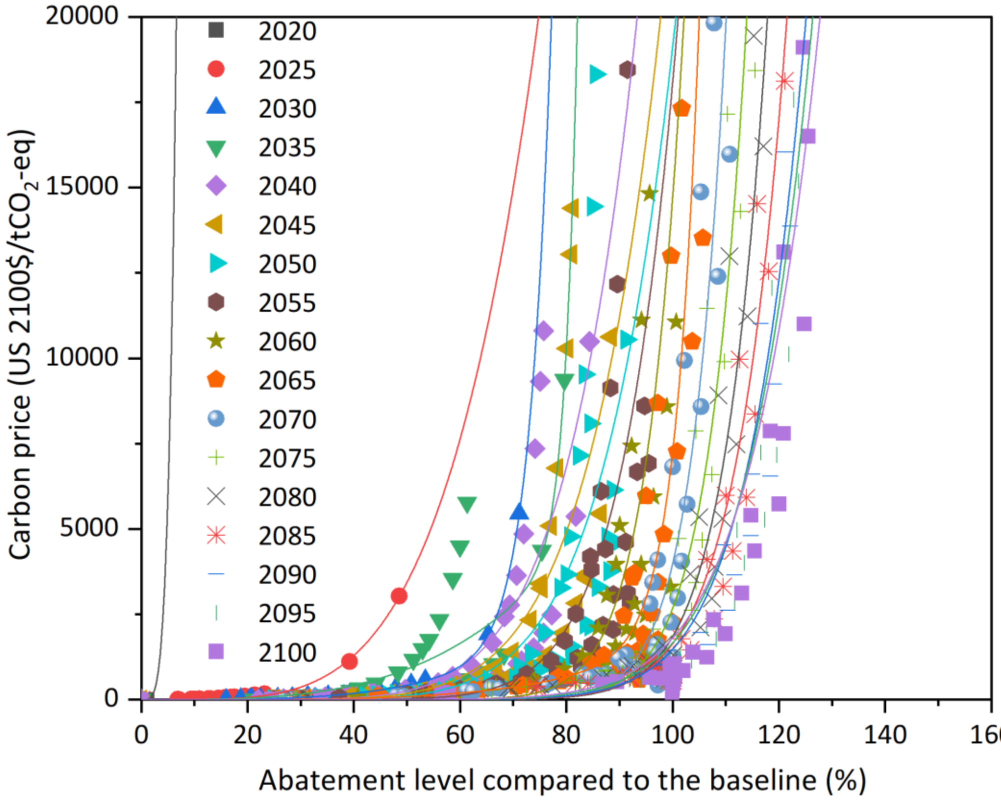

Figure S4. MAC curves of total anthropogenic CO_2 emissions per five years for POLES. The dots are original data from the POLES model, and the lines are MAC curves derived from these data for different years.

13. Section 3.2.2 discussed the role of the first and second derivatives of the abatement changes, which is interesting, but I still don't fully understand its value. I.e., do they have physical meanings? (see my comment # 8). Also, what if those upper limits for the first and second derivatives were removed? How could that change the fitted models?

[Response] If the first and second derivatives are taken into account, the rate of increase in the abatement level will rise slowly in the near term. It can reach its upper limit when society and technology have adapted to the policy requirements of climate mitigation (see red and blue lines of Figure R1). However, if this constraint is removed, it means that the upper limit of emission reductions (e.g. net zero for CO_2 or even negative CO_2 emissions) can be reached immediately (see black line of Figure R1), which is clearly not the case in the original model output.

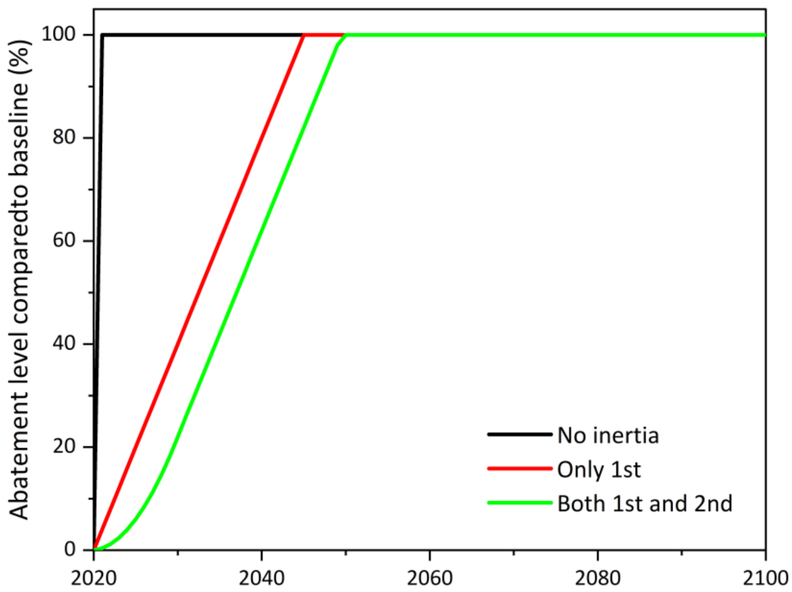

Figure R1. Abatement level considering the most growth rate of mitigation. Here we assume the maximum potential of mitigation is 100%. If no social-economic inertia is considered for the MAC curve (black line), then the abatement level can reach 100% quickly. If only technological change inertia is considered for the MAC curve (red line), then the abatement level will grow with a fixed slope before it reaches 100%. If both technological change and social-economic inertia are considered for the MAC curve (green line), the abatement will increase slowly because the technology also needs time to change and adapt.

14. Fig 4 seems to capture the model differences. However, the true question is to what extent are these differences because of the model's structural differences or differences in the scenarios simulated by different models? Each model may contribute varying numbers of scenarios to the database with unevenly distributed scenario narratives. Thus, the differences here might be driven by the artificial selection of the training sample. I hope the authors can

share some thoughts on this. Also, is there any notable structural differences that might be helpful to explain the observations in line 320-329?

[Response] We thank the reviewer for the question. We aim to extract the relationship between abatement levels and carbon prices for models from a large number of scenarios, so the number of available scenarios can influence how well the MAC curves can be fitted. The larger the number of carbon budget scenario is, the more accurate the fitted curves will generally be. Therefore, the distribution of carbon budgets tested by individual models (Table S7; see below) can be a potential source of bias. Meanwhile, the model structure could also be a reason as well. For example, the five IAMs (COFFEE 1.1, MESSAGEix-GLOBIOM 1.1, POLES-JRC ENGAGE, REMIND-MAgPIE 2.1-4.2, and WITCH 5.0) have very similar carbon budget ranges and number of scenarios while different solution concepts and solution methods (Table 1 and Table S7). Thus, we chose these IAMs and filtered the scenarios that they all provided (19 scenarios in total, see Table S7). The MAC curves for anthropogenic CO_2 emissions are given in Figure S37. The MACs between the different IAMs still vary considerably, but the results of the three general equilibrium models are close to each other, while those of the two partial equilibrium models are far apart, although we do not have a further insight into why this occurs. Note that the results of the MAC curves are not very sensitive to the number of scenarios, as the results for the subsample of scenarios we used here are very similar to the results for the full sample.

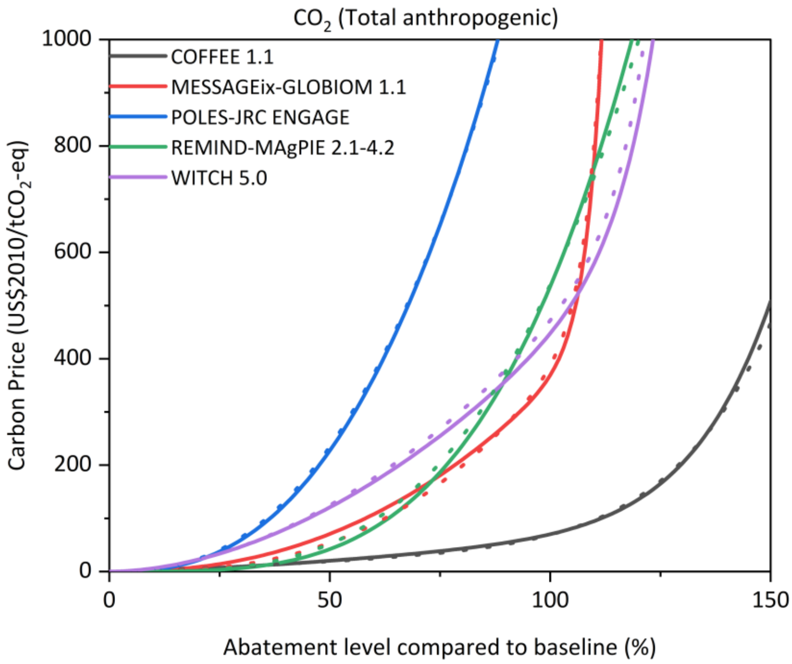

Figure S37. Global MAC curves for total anthropogenic CO_2 emissions derived from the same scenarios for five ENGAGE IAMs. The solid lines are MAC curves derived from the subsample, and the dotted lines are MAC curves derived from the full sample. No upper limit of abatement level is shown for MAC curves.

Table S7. Available scenarios for each model in the ENGAGE Scenario Explorer. 0 means that the model does not provide this scenario, while 1 means that the model provides this scenario.

Scenarios	AIM/CGE V2.2	COFFEE 1.1	GEM-E3 V2021	IMAGE 3.0	MESSAGEix-GLOBIOM 1.1	POLES-JRC ENGAGE	REMIND-MAgPIE 2.1-4.2	TIAM-ECN 1.1	WITCH 5.0	Count
EN_INDCi2030_300f	0	0	0	0	0	1	1	0	0	2
EN_INDCi2030_400	0	0	0	0	0	0	0	0	0	0
EN_INDCi2030_400f	0	1	0	0	0	1	1	0	0	3
EN_INDCi2030_500	0	1	0	0	0	0	0	0	0	1
EN_INDCi2030_500f	0	1	0	0	0	1	1	0	1	4
EN_INDCi2030_600	0	1	0	0	0	0	0	0	0	1
EN_INDCi2030_600f	0	1	1	0	0	1	1	0	1	5
EN_INDCi2030_700	0	1	0	0	0	0	1	0	0	2
EN_INDCi2030_700f	0	1	0	0	1	1	1	0	1	5
EN_INDCi2030_800	0	1	1	0	0	0	1	0	1	4
EN_INDCi2030_800f	1	1	1	1	1	1	1	0	1	8
EN_INDCi2030_900	0	1	0	0	0	1	1	1	1	5
EN_INDCi2030_900f	1	1	0	0	1	1	1	1	1	7
EN_INDCi2030_1000	0	1	1	1	1	1	1	1	1	8
EN_INDCi2030_1000f	1	1	1	1	1	1	1	1	1	9
EN_INDCi2030_1200	1	1	0	1	1	1	1	1	1	8
EN_INDCi2030_1200f	1	1	0	1	1	1	1	1	1	8
EN_INDCi2030_1400	1	1	1	1	1	1	1	1	1	9
EN_INDCi2030_1400f	1	1	1	1	1	1	1	1	1	9
EN_INDCi2030_1600	1	1	0	0	1	1	1	1	1	7
EN_INDCi2030_1600f	1	1	0	0	1	1	1	1	1	7
EN_INDCi2030_1800	1	1	1	0	1	1	1	0	1	7
EN_INDCi2030_1800f	1	1	1	0	1	1	1	0	1	7
EN_INDCi2030_2000	0	1	0	0	1	1	1	1	1	6
EN_INDCi2030_2000f	0	1	0	0	1	1	1	1	1	6
EN_INDCi2030_2500	0	1	0	0	1	1	1	1	1	6
EN_INDCi2030_2500f	0	1	0	0	1	1	1	1	1	6
EN_INDCi2030_3000	0	0	0	1	1	1	1	1	1	6
EN_INDCi2030_3000f	0	0	0	1	1	1	1	1	1	6
EN_NPi2020_200f	0	0	0	0	1	0	1	0	0	2
EN_NPi2020_300	1	0	0	0	0	0	0	0	0	1
EN_NPi2020_300f	1	0	0	0	1	1	1	0	0	4
EN_NPi2020_400	0	1	0	0	1	0	1	0	0	3
EN_NPi2020_400f	1	1	1	0	1	1	1	0	1	7
EN_NPi2020_450	0	0	0	0	1	0	0	0	0	1
EN_NPi2020_450f	0	0	0	0	1	0	1	0	0	2
EN_NPi2020_500	0	1	1	0	1	1	1	0	1	6
EN_NPi2020_500f	1	1	1	0	1	1	1	0	1	7
EN_NPi2020_600	1	1	1	0	1	1	1	0	1	7
EN_NPi2020_600f	1	1	1	1	1	1	1	0	1	8
EN_NPi2020_700	1	1	0	0	1	1	1	0	1	6
EN_NPi2020_700f	1	1	0	0	1	1	1	0	1	6
EN_NPi2020_800	1	1	1	1	1	1	1	1	1	9
EN_NPi2020_800f	1	1	1	1	1	1	1	1	1	9
EN_NPi2020_900	1	1	0	0	1	1	1	1	1	7
EN_NPi2020_900f	1	1	0	0	1	1	1	1	1	7
EN_NPi2020_1000	1	1	1	1	1	1	1	1	1	9
EN_NPi2020_1000f	1	1	1	1	1	1	1	1	1	9
EN_NPi2020_1200	1	1	0	1	1	1	1	1	1	8
EN_NPi2020_1200f	1	1	0	1	1	1	1	1	1	8
EN_NPi2020_1400	1	1	1	1	1	1	1	1	1	9
EN_NPi2020_1400f	1	1	1	1	1	1	1	1	1	9
EN_NPi2020_1600	1	1	0	0	1	1	1	1	1	7
EN_NPi2020_1600f	1	1	0	0	1	1	1	1	1	7
EN_NPi2020_1800	1	1	1	0	1	1	1	0	1	7
EN_NPi2020_1800f	1	1	1	0	1	1	1	0	1	7
EN_NPi2020_2000	0	1	0	0	1	1	1	1	1	6
EN_NPi2020_2000f	0	1	0	0	1	1	1	1	1	6
EN_NPi2020_2500	0	1	0	0	1	1	1	1	1	6
EN_NPi2020_2500f	0	1	0	0	1	1	1	1	1	6
EN_NPi2020_3000	0	0	0	1	1	1	1	1	1	6
EN_NPi2020_3000f	0	0	0	1	1	1	1	1	1	6
EN_NPi2100	1	1	1	1	1	1	1	1	1	9
Count	**34**	**52**	**23**	**21**	**52**	**53**	**58**	**35**	**51**	

15. Table 2, why do some models have huge coefficients for a and c? For example, the "a" parameter for REMIND CH4 or the "c" parameter for WITCH CH4 and N2O? Is this because of the model itself, or were the scenarios chosen for fitting? Also, the main text did not make any comment on Table 2.

[Response] We thank the reviewer for the question. In a single power function $y = a * x^b$, a determines the position of the function curve in the vertical direction, and b determines its shape. That is, when b>1, the curve is flatter near the origin and then rises sharply. When 0<b<1, the curve is steeper near the origin and then flattens out. A large a implies a large y value. The function $y = a * x^b + c * x^d$, which has two power functions, allows us to capture more complex trends of MAC curves.

Therefore, we think the phenomenon that the reviewer raised is due to a combination of the chosen function and the data distribution of models. The reason for the very high value of *a* for REMIND CH_4 and *c* for WITCH CH_4 and N_2O is that the mitigation price is very low at a low abatement level, but rises sharply when the abatement level is close to the upper limit (nearly vertically for the REMIND and WITCH models (see Figure S7(g), S7(i), and S10(i)).

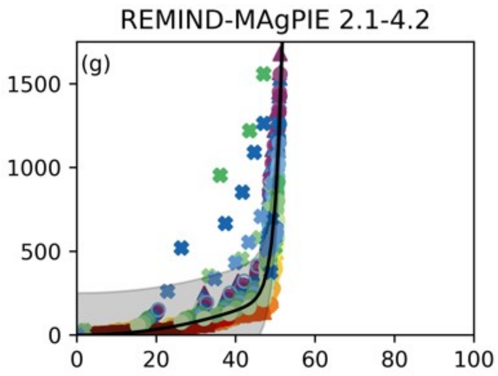

Figure S7(g) Global total CH_4 MAC

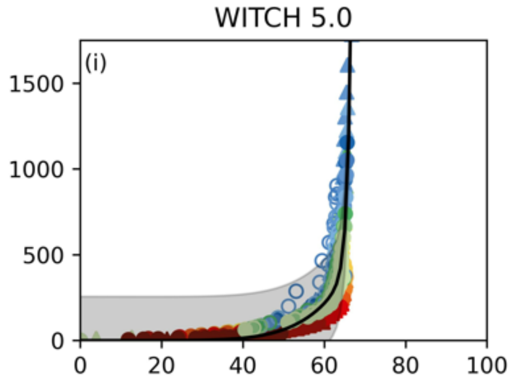

Figure S7(i) Global total CH_4 MAC

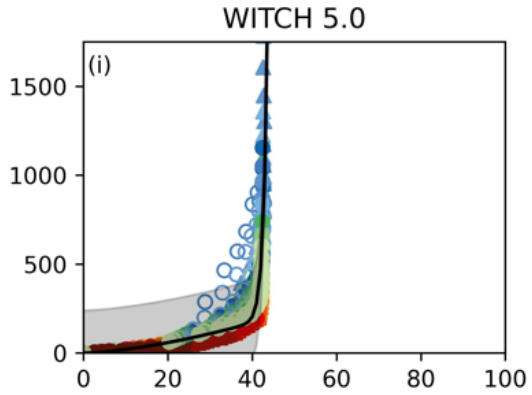

Figure S10(i) Global total N_2O MAC

16. Line 366: "They are further compared with the Global MAC curves for energy-related CO2 emissions from ENGAGE IAMs."

[Response] This review comment looks incomplete, but we guess that the reviewer asks why we compare the GET model's results with ENGAGE IAMs. We aim to see if there is any significant difference between the two datasets for MAC curves. The results show that the range is nearly as wide as that from ENGAGE IAMs (i.e., inter-technology portfolio range ≈ inter-model range) if we disregard the MAC curve from COFFEE.

17. Figures 7 and 8: Given the current presentation, there's no way to check the model performance for emIAM-ACC2 visually. Please avoid showing so many lines/dots in one figure; this busy chart provides minimal information.

[Response] We thank the reviewer for the suggestion. We assume that the reviewer refers to Figures 8 and 9. We think that Figure 8 in the original manuscript was clear enough for the comparison because it shows only a subset of simulations (i.e. results of EBC scenarios without INDC). However, Figure 9 in the original manuscript was more difficult to read because all scenarios are shown. Therefore, we have modified Figure 9 to present only EBC scenarios without INDC (consistent with Figure 8).

18. Technically, the entire validation test (section 4) is performed in the "training set". Ideally, this should be done in a validation set outside the training set. Authors could 1) try to select scenarios from another scenario database, such as IPCC AR6, with the same set of models and selected scenarios as validation, or 2) just randomly choose a part of the ENGAGE scenarios as the training set to fit MACs (if there's enough sample size), then use the remaining ENGAGE scenarios as the validation set.

[Response] We thank the reviewer for the suggestion. It is an interesting idea, but since our framework considers both the MAC function and the baseline scenario, as well as the constraints of the first and second derivative of abatement rates, selecting scenarios from other projects in AR6 can lead to inconsistency between the model used to train the emulator and the model that give test scenarios. The second point is also a useful suggestion, but the number of scenarios used to generate MAC curves in the dataset in this study is limited, so we decided to stick to our current approach. Nevertheless, we thank the reviewer for sharing the thoughts, which could be applied to our future study.

19. Comparing Figures 10 and 11, I wonder why COFFEE performed well in the global test but poorly for most regions for CO2. Are they consistent?

[Response] We thank the reviewer for the question. We double checked and updated the results. Because of our mistakes, we provided the wrong emission pathways for scenarios with INDC from ACC2-emIAM. Now we have updated these figures with the correct results. The revised results show that the COFFEE model also performed well in reproducing the regional CO_2 emissions.

20. Line 623 "The results showed that the original emission pathways were reproduced reasonably well in a majority of cases." This is oversimplified. The performance depends on

the gas, model, and maybe other features (if Figures 7 and 8 could have been clearer). Here needs a better summary of the findings.

[Response] We thank the reviewer for the suggestion. We have substantially expanded the discussion in Section 5. The new Section 5 has a high-level summary of the findings as follows:

- *The validation results for the two long-lived gases CO_2 and N_2O did not strongly differ across all four Tests, even though for Tests 2 to 4, there is a difference in the model setup between the original IAMs (GHG aggregation using GWP100) and ACC2-emIAM (individual gas cycle modeling without using GWP100). On the other hand, the validation results for the short-lived gas CH_4 in Tests 2 to 4 were not as good as those in Test 1. Test 4, with the additional mid-century temperature target, yielded higher reproducibility for CH_4 than Tests 2 and 3.*
- *Overall, the global emissions were better reproduced than the regional emissions. CO_2 emission pathways were generally better reproduced than CH_4 and N_2O pathways. Specific pathway features such as CO_2 emissions in 2030, 2050, and 2100, cumulative negative CO_2 emissions from 2020 to 2100, the year to net zero for CO_2, and that for GHG were reproduced to varying degrees across models and carbon budgets (Figure 10). While certain biases for some models were found for certain pathway features, as reported earlier, no general conclusions can be drawn.*
- *Some IAMs were more easily emulated than other IAMs, reflecting specific model features such as solution methods, technology assumptions, and abatement inertia. The emulator can usually reproduce the emission pathways of an IAM better if the model response to carbon price are well fitted with a MAC function.*
- *Certain data points were difficult to capture by MAC curves. In particular, PKB scenarios with low carbon budgets can give very large carbon prices in the near-term. Such data points tend to deviate from the trend of other data points and were manually removed from the MAC curve fitting where appropriate (Figure 1 and Table 1). Except for these "outliers," no discernible difference in the data trend was found between ECB scenarios and PKB scenarios, supporting the use of common MAC curves for ECB and PKB scenarios. Note also that certain data points from GET at high abatement levels do not follow the trend of other data points and were also removed from the MAC curve fitting where appropriate. We speculate that these data points are affected by the limit on CCS capacity assumed in GET.*
- *The overall good reproducibility of emIAM relies on our novel approach: time-independent MAC curves for percentage emission reductions. The behaviors of IAMs that contain various time-dependent processes were generally well captured by the time-*

independent MAC curves. A plausible explanation is that the use of percentage abatement levels relative to rising baseline can offset the effect of lowering mitigation costs through learning.

21. Line 626, "Materials that are required for making such decisions are systematically presented in Supplement and our Zenodo repository." This is essential information; the authors should provide a couple of high-level bullet points.

[Response] As we responded above, the revised manuscript provides several high-level bullet points.

22. Line 627, "Some IAMs were more easily emulated than other IAMs. The goodness of fit of the MAC curves depends on gases and regions." Again, this is another place that should have provided richer information beyond the current simple comment (which readers would even know before reading this paper).

[Response] This also relates to the two previous comments. We hope that the newly added bullet points address the reviewer's concern.

#Reviewer 3

Summary:

The paper describes an emulator for Integrated Assessment Models (IAMs) based on an aggregation of MAC curves of different models, regions, time points and greenhouse gases. The idea is interesting and useful, because it allows for quick assessments of abatement given different carbon prices, for which running IAMs may be computationally costly. The paper focuses on the calculation of these MAC curves, on which the authors are thorough, and on the validation of the resulting emulator in comparison to the output of the IAMs that the authors started with.

[Response] We thank the reviewer for taking the time to read our manuscript and for providing useful comments. We also thank the reviewers for recognizing the usefulness and thoroughness of our work.

General comments

While the idea of this emulator is interesting and useful, I unfortunately do not recommend publication in the paper's current form and am providing a few suggestions below that may be used for major revision.

1. The scenarios used as input are merely listed, but little motivation is given why the ENGAGE database is chosen, while I think this is key to the resulting MAC curves in the emIAM. My suggestion would be to at least motivate why the ENGAGE database is suitable for this exercise, and why the authors are not using the full AR6 scenario database that came out last year.

[Response] We thank the reviewer for the comment. The AR6 Scenario database includes a large ensemble of scenarios from different projects, including the ENGAGE project. However, the AR6 Scenario database was not available at the time of our analysis. Meanwhile, we chose only the ENGAGE project because this project adopts the same socioeconomic assumptions (i.e. second marker baseline scenario from the Shared Socioeconomic Pathways (SSP2), which reflect middle-of-the-road socioeconomic conditions (Riahi et al., 2017)) and provides plenty of cases to derive the MAC curves (see Figure 3.2 of (Riahi et al., 2022)). Thus, we argue that the ENGAGE Scenario Explorer is the best dataset for our application as it gives a range of scenarios under different carbon budgets for many models with consistent configurations. We have added some text to explain why we only used the scenarios from the ENGAGE project instead of the full dataset of AR6:

The ENGAGE Scenario Explorer is now part of the larger IPCC Sixth Assessment Report (AR6) Scenario Explorer (Byers et al., 2022), which was not available at the time of our analysis. Although the use of the entire AR6 scenario dataset could be advantageous in terms of the number of IAMs and scenarios available for analyses (189 IAMs (including different model versions) and 1389 scenarios in the AR6 Scenario Explorer; 20 IAMs (including different model versions) and 231 scenarios in the ENGAGE Scenario Explorer), an advantage of using the ENGAGE Scenario Explorer is that the data from IAMs were obtained under a common experimental protocol, allowing consistent analyses.

2. A major concern is the lack of discussion in this paper. The paper contains a lot of detailed description of results, along with many detailed figures, but lacks broader discussion. For example, where to the gas differences in Fig. 4 or the regional differences in Fig. 5 come from? Could we have expected them beforehand? And what do the significant model differences in Fig. 2 imply for the ultimate results?

[Response] We thank the reviewer for pointing out this problem. However, we are afraid that it is nearly impossible to directly answer these questions because the IAMs we are dealing with are very different from each other and we do not have deep insight into each of these IAMs (we are not taking part in the ENGAGE project. We are merely using the publicly available database of ENGAGE). In this sense, we argue that our study is not designed to provide explanations for the differences found. Rather, our study aims to explore to what extent our generic MAC curve approach works for different models, gases, regions, etc, although we discuss possible reasons of good/poor MAC curve fitting and reproducibility where possible.

3. The paper can be written more concise and requires a bit more flow to guide the reader throughout the steps. Also, the paper contains too many figures/panels which are not well readable, especially when it comes to symbols (circles/triangles, etc.) and scenario labels. The authors may consider moving some to the SI.

[Response] We thank the reviewer for the comment. We have made efforts to streamline the content and improve the flow throughout the manuscript. We have also polished the manuscript figures for better readability (e.g., font size and color scheme). For example, we have modified Figure 1 to more clearly present the overall methodological flow (e.g. reduced the number of small panels over time; changed the figure legend to something more intuitive). In Figure 9, we reduced the number of scenarios presented so that each scenario can be read more clearly. In

Section 3, we merged the content of the MAC functions from different paragraphs and moved it after the introduction of data processing. We also made use of footnotes where necessary to shorten the text and avoid breaking the flow of the manuscript.

4. More details on the uncertainty of this approach is needed. Clearly, the results are gas, model, region and time dependent, while some of these things are actually aggregated into a single MAC in emIAM. What does this imply for the end results? Perhaps work with uncertainty bars in a summarizing plots in the end to give the reader a feeling for the uncertainty of emIAM. Similar for the parametric uncertainties in the values of a, b, c and d when fitting, which may require a sensitivity analysis.

[Response] We thank the reviewer for the suggestion. Though the relationship between the carbon price and the CO_2 abatement level can be well captured by MAC curves for most IAMs we considered, the results vary in terms of the range of carbon prices, the range of abatement levels, and the dispersion of data points. We have added 95% confidence intervals of the fitted MAC curves in Figure 2 and Figure 6 (and more figures in Supplement).

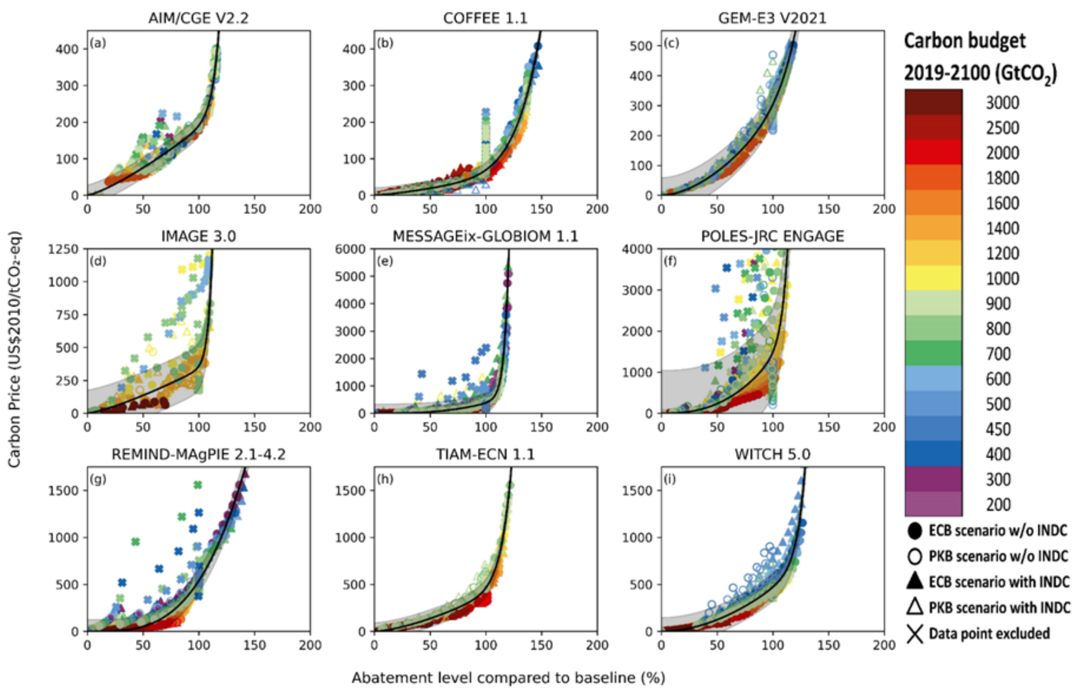

Figure 3. Relationships between the carbon price and the global total anthropogenic CO_2 abatement level obtained from nine ENGAGE IAMs. Each panel shows the results from each ENGAGE IAM. Data were obtained from the ENGAGE Scenario Explorer and are shown in colors and markers as designated in the legend. Black lines are the MAC curves. Crosses are the

data points that were not included in the derivation of MAC curves (Table 1). The shaded bands are the 95% confidence intervals of the fitted curves calculated by $\hat{y} \mp t_{\frac{\alpha}{2}} * S_{\varepsilon} * \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - \frac{(\sum x)^2}{n}}}$ (Thomson and Emery, 2014), where $S_{\varepsilon} = \sqrt{\frac{\sum(y-\hat{y})^2}{n-2}}$, n is the sample size, $t_{\frac{\alpha}{2}}$ is the critical value of t-distribution, $\bar{x}$ is the mean of samples, $\hat{y} = f(x)$, and x, y are the original abatement level and carbon price result from the IAM, respectively.

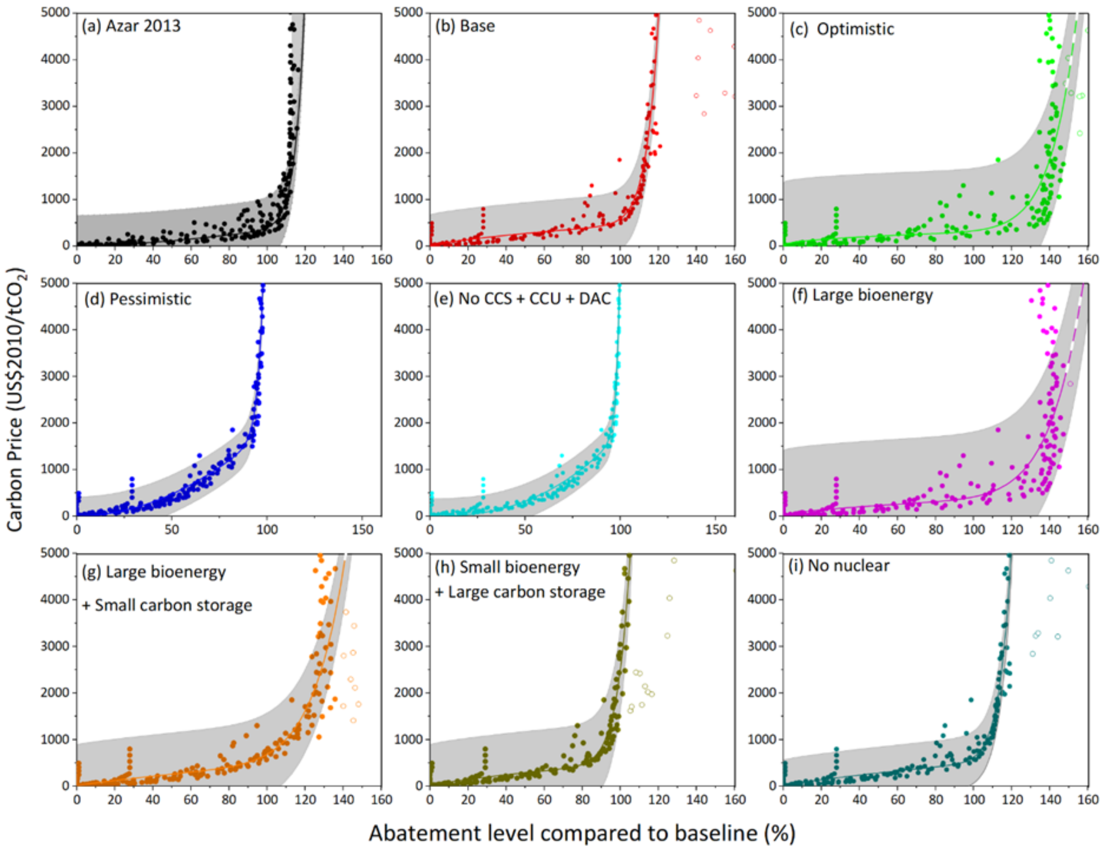

Figure 4. Relationships between the carbon price and the global energy-related CO_2 abatement level obtained from GET with different portfolios of available mitigation technologies. Panel (a) shows the results obtained from an older version of GET (Azar et al., 2013) for the sake of comparison. Panels (b) to (i) show the results from GET (Lehtveer et al., 2019) with different technology portfolios. See Section 2.2 for the definitions of technology portfolios. Points are the data obtained from GET; lines are the MAC curves calculated based on our approach. Open circles are the data that were not considered in the derivation of MAC curves (Table 1) and are typically found after 2100, in some cases above the abatement level of 160% (not shown). Note that we have converted the unit in Panel (a) from US$2010/tC, which is used in the older version of GET, to US$2010/t$CO_2$, the commonly used unit here. The shaded bands are the 95% confidence intervals of the fitted curves calculated (see the caption of Figure 2)

5. The authors have chosen to work with percentage abatement w.r.t. baselines rather than absolute abatement. I understand the reasoning, but it is not trivial that this choice fully counteracts the lack of temporal dependency in the analysis (e.g., in the form of learning by doing), even though this is (perhaps even coincidentally) visible when comparing the percentage versions versus the absolute versions. Moreover, baselines significantly differ among models, which introduces another source of uncertainty. A discussion on this would be helpful in the paper.

[Response] We thank the reviewer for the comment. We compared the data distribution of relative and absolute abatements for three models (AIM, REMIND, and MESSAGE) (see Figure S3). The figure shows the results for relative abatement are more concentrated, supporting the use of relative abatements for our MAC functions.

Although there are large differences in the baselines of the models, many models assume a rising baseline scenario (especially for CO_2). Rising baseline scenarios counteract, at least to a certain extent, the increasing abatement level over time at the same carbon price.

We have the following relevant discussions in the manuscript (Sections 3.1 and 5):

"Learning by doing" and "learning with time," which reduce the mitigation cost with abatement (endogenously) and time (exogenously), respectively (Hof et al., 2021), are not explicitly considered in our MAC curve approach, but are partially captured in our approach, which describes percentage reduction rates relative to rising baseline scenarios. For example, constant emission reductions in absolute terms can appear smaller over time in relative terms and thus become less costly in our approach.

- *The overall good reproducibility of emIAM relies on the use of time-independent MAC curves for percentage emission reductions. The behaviors of IAMs that contain various time-dependent processes were generally well captured by the time-independent MAC curves. A plausible explanation is that the use of percentage abatement levels relative to rising baseline can offset the effect of lowering mitigation costs through learning.*

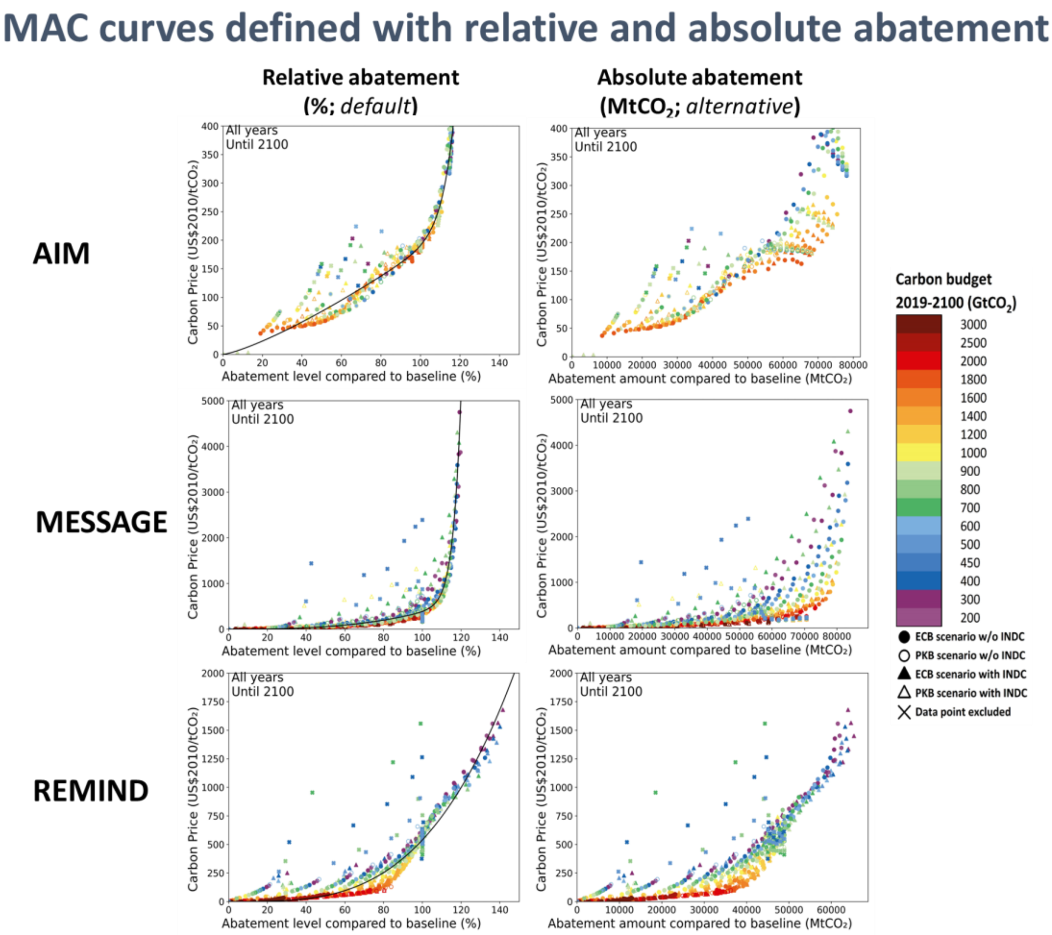

Figure S3. MAC curves defined with relative and absolute abatement for three models. The left three panels show the relationship between carbon price and relative abatement level, while the right three panels show the relationship between carbon price and absolute abatement. Black lines in the left three panels are MAC curves in percentage used in our study.

6. The numbers in Fig. 10 and 11 are difficult to judge purely on their numerics. It would be useful to provide an example and focus on a number of key ingredients of emission pathways rather than pure correlations: how do the 2030 emissions differ, the netzero years, and the required negative emissions in overshoot scenarios? I guess that it is almost trivial to have a high correlation in general, because in all scenarios, emissions go down over time. Hence, to convince the reader, focusing on comparisons beyond mere correlation metrics would be useful.

[Response] We thank the reviewers for their comments. This suggestion is very useful, and we have incorporated this idea into our manuscript. We considered several new indicators that would be useful for comparing emission pathways of ACC2-emIAM and EMGAGE IAMs. Specifically, we considered the difference in carbon emissions between our reproduced results and the original results for 2030, 2050, and 2100, as well as the difference in cumulative negative CO_2 emissions during the period 2020-2100, the difference in the year in which net zero of CO_2 and net zero of GHGs (CO_2 + CH_4 + N_2O) are achieved. We added Figure 10, which presents the results for the ECB scenarios without INDC from Test 4, as well as associated discussions as follows:

Furthermore, we examine several selected features of the original and reproduced emission pathways from Test 4 (ECB scenarios without INDC only), such as CO_2 emissions in 2030, 2050, and 2100, cumulative negative CO_2 emissions from 2020 to 2100, the year to net zero for CO_2, and that for GHG. Figure 10a-c indicates that the reproducibility of CO_2 emissions for three different points in time varies across models and carbon budgets, but it is worth noting that ACC2-emIAM nearly consistently overestimates and underestimates 2030 CO_2 emissions from AIM and REMIND, respectively. Cumulative negative CO_2 emissions are negatively underestimated for COFFEE (Figure 10d), which is related to the general overestimation of 2100 CO_2 emissions for COFFEE (Figure 10c). The year to net zero for CO_2 tends to be overestimated (later than the original year) for REMIND with the carbon budget at or below 800 $GtCO_2$.

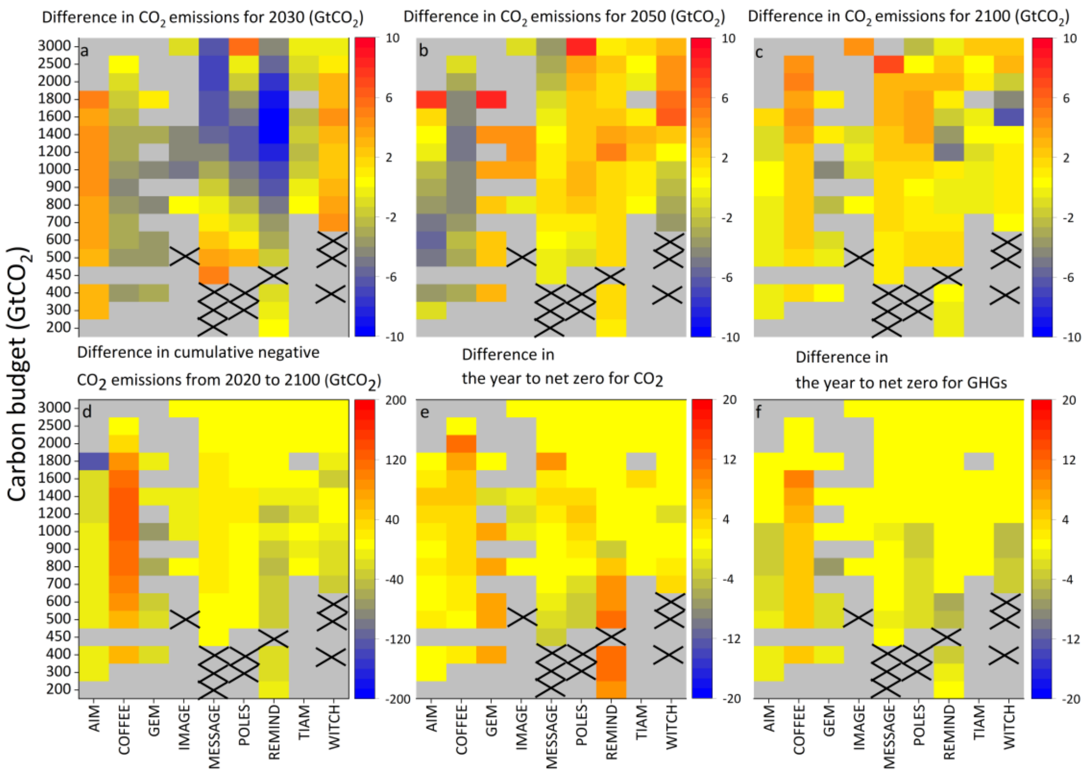

Figure 5. Differences in the pathway features between ENGAGE IAMs and ACC2-emIAM.
Panels a to c show the difference in CO_2 emissions for 2030, 2050, and 2100, respectively. Panel

d shows the difference in cumulative negative CO_2 emissions. Panel e shows the difference in the year to net zero for CO_2. Panel f shows the difference in the year to net zero for GHGs (for CO_2, CH_4, and N_2O). Positive values indicate that the features in the original pathways (from ENGAGE IAMs) are larger than those in the reproduced pathways (from ACC2-emIAM), while negative values indicate the opposite. Gray boxes without black crosses indicate that the corresponding scenarios were not available in the ENGAGE Scenario Explorer, while those with black crosses indicate that the corresponding scenarios were available in the ENGAGE Scenario Explorer but not successfully reproduced by ACC2-emIAM (i.e., infeasible solutions).

7. Perhaps more generally and related to aforementioned points: the MAC curve deductions themselves are interesting and a lot of insights can be obtained from them. However the analysis also reveals that "We do not provide specific recommendations on the appropriateness of the use of each MAC curve and leave the users to decide which MAC curves to apply" (p. 31), suggesting that the many differences between the MAC curves limit the universal applicability of emIAM. Potential users need to be guided better: which results are generalizable, what are the main uncertainties? A discussion section, looking at this question from a helicopter point-of-view may help in this respect, which is currently missing. This paper may be a first step in the direction of IAM-emulators, but then the authors are invited to write a bit more about what the next steps should be.

[Response] We thank the reviewer for the suggestion. Another reviewer expressed the same concern. In the revised manuscript, we have added several bullet points that give a high-level summary of the findings. We hope this generalization can better guide potential users.

- *The validation results for the two long-lived gases CO_2 and N_2O did not strongly differ across all four Tests, even though for Tests 2 to 4, there is a difference in the model setup between the original IAMs (GHG aggregation using GWP100) and ACC2-emIAM (individual gas cycle modeling without using GWP100). On the other hand, the validation results for the short-lived gas CH_4 in Tests 2 to 4 were not as good as those in Test 1. Test 4, with the additional mid-century temperature target, yielded higher reproducibility for CH_4 than Tests 2 and 3.*
- *Overall, the global emissions were better reproduced than the regional emissions. CO_2 emission pathways were generally better reproduced than CH_4 and N_2O pathways. Specific pathway features such as CO_2 emissions in 2030, 2050, and 2100, cumulative negative CO_2 emissions from 2020 to 2100, the year to net zero for CO_2, and that for GHG were reproduced to varying degrees across models and carbon budgets (Figure 10). While certain biases for some models were found for certain pathway features, as reported earlier, no general conclusions can be drawn.*

- *Some IAMs were more easily emulated than other IAMs, reflecting specific model features such as solution methods, technology assumptions, and abatement inertia. The emulator can usually reproduce the emission pathways of an IAM better if the model response to carbon price are well fitted with a MAC function.*
- *Certain data points were difficult to capture by MAC curves. In particular, PKB scenarios with low carbon budgets can give very large carbon prices in the near-term. Such data points tend to deviate from the trend of other data points and were manually removed from the MAC curve fitting where appropriate (Figure 1 and Table 1). Except for these "outliers," no discernible difference in the data trend was found between ECB scenarios and PKB scenarios, supporting the use of common MAC curves for ECB and PKB scenarios. Note also that certain data points from GET at high abatement levels do not follow the trend of other data points and were also removed from the MAC curve fitting where appropriate. We speculate that these data points are affected by the limit on CCS capacity assumed in GET.*
- *The overall good reproducibility of emIAM relies on the use of time-independent MAC curves for percentage emission reductions. The behaviors of IAMs that contain various time-dependent processes were generally well captured by the time-independent MAC curves. A plausible explanation is that the use of percentage abatement levels relative to rising baseline can offset the effect of lowering mitigation costs through learning.*

Minor comments

1. The portfolios for GET are described only qualitatively. The choices (p. 5) even seem arbitrary – e.g., why did the authors use the numbers of 100% larger and 50% smaller bioenergy constraints in the respective portfolios?

[Response] This is just an arbitrary assumption used to illustrate our purpose.

2. Unclear: in Section 4, also the regional MAC curves from emIAM are used, while on p. 6 a regional independence is assumed. What is it you are actually using in section 4?

[Response] Thank you for your comment. When we derived MAC curves for regions using the ENGAGE project, we assumed regions are independent. That is, we do not consider the correlation (or inter-dependency) between the abatement level of a region and that of another region. In Section 4, we used the regional MAC curves derived from the ENGAGE project. Here, the trade-off of abatement levels between regions with the least-cost emission pathways can be seen. The model will decide which region should remove a certain level of gases considering its carbon price. Therefore, the work in section 4 does not conflict with that assumption.

3. Could you elaborate a bit on Fig. 6 and where these points come from?

[Response] We thank the reviewer for the comment. As stated in the figure caption, different panels in Figure 6 present the relationship between the carbon price and the global energy-related CO_2 mitigation level under different technology portfolios. The original results from the GET model are shown in points. The process to calculate the abatement level can be seen in Section 3.1.

4. Second derivative unit should be % / year^2 I guess, or are the numbers of the fractional order 1e-4?

[Response] We thank the reviewer for pointing out this issue. The second derivative unit is %/(year)^2, and this has been corrected in the text.

References

Azar, C., Johansson, D. J. A., and Mattsson, N.: Meeting global temperature targets—the role of bioenergy with carbon capture and storage, Environ. Res. Lett., 8, 034004, https://doi.org/10.1088/1748-9326/8/3/034004, 2013.

Byers, E., Krey, V., Kriegler, E., Riahi, K., Schaeffer, R., Kikstra, J., Lamboll, R., Nicholls, Z., Sandstad, M., Smith, C., van der Wijst, K., Al -Khourdajie, A., Lecocq, F., Portugal-Pereira, J., Saheb, Y., Stromman, A., Winkler, H., Auer, C., Brutschin, E., Gidden, M., Hackstock, P., Harmsen, M., Huppmann, D., Kolp, P., Lepault, C., Lewis, J., Marangoni, G., Müller-Casseres, E., Skeie, R., Werning, M., Calvin, K., Forster, P., Guivarch, C., Hasegawa, T., Meinshausen, M., Peters, G., Rogelj, J., Samset, B., Steinberger, J., Tavoni, M., and van Vuuren, D.: AR6 Scenarios Database, https://doi.org/10.5281/zenodo.7197970, 2022.

Emmerling, J., Drouet, L., Wijst, K.-I. van der, Vuuren, D. van, Bosetti, V., and Tavoni, M.: The role of the discount rate for emission pathways and negative emissions, Environ. Res. Lett., 14, 104008, https://doi.org/10.1088/1748-9326/ab3cc9, 2019.

Fuglestvedt, J. S., Berntsen, T. K., Godal, O., Sausen, R., Shine, K. P., and Skodvin, T.: Metrics of climate change: assessing radiative forcing and emission indices, Clim. Change, 58, 267–331, https://doi.org/10.1023/A:1023905326842, 2003.

Harmsen, J. H. M., van Vuuren, D. P., Nayak, D. R., Hof, A. F., Höglund-Isaksson, L., Lucas, P. L., Nielsen, J. B., Smith, P., and Stehfest, E.: Long-term marginal abatement cost curves of non-CO2 greenhouse gases, Environ. Sci. Policy, 99, 136–149, https://doi.org/10.1016/j.envsci.2019.05.013, 2019.

Harmsen, M. J. H. M., Berg, M. van den, Krey, V., Luderer, G., Marcucci, A., Strefler, J., and Vuuren, D. P. V.: How climate metrics affect global mitigation strategies and costs: a multi-model study, Clim. Change, 136, 203–216, https://doi.org/10.1007/s10584-016-1603-7, 2016.

Hof, A. F., Wijst, K.-I. van der, and Vuuren, D. P. van: The Impact of Socio-Economic Inertia and Restrictions on Net-Negative Emissions on Cost-Effective Carbon Price Pathways, Front. Clim., 3, 785577, https://doi.org/10.3389/fclim.2021.785577, 2021.

IPCC: Summary for Policymakers, 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, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Lashof, D. A. and Ahuja, D. R.: Relative contributions of greenhouse gas emissions to global warming, Nature, 344, 529–531, https://doi.org/10.1038/344529a0, 1990.

Lehtveer, M., Brynolf, S., and Grahn, M.: What Future for Electrofuels in Transport? Analysis of Cost Competitiveness in Global Climate Mitigation, Environ. Sci. Technol., 53, 1690–1697, https://doi.org/10.1021/acs.est.8b05243, 2019.

Levasseur, A., Cavalett, O., Fuglestvedt, J. S., Gasser, T., Johansson, D. J. A., Jørgensen, S. V., Raugei, M., Reisinger, A., Schivley, G., Strømman, A., Tanaka, K., and Cherubini, F.: Enhancing life cycle impact assessment from climate science: Review of recent findings and recommendations for application to LCA, Ecol. Indic., 71, 163–174, https://doi.org/10.1016/j.ecolind.2016.06.049, 2016.

Riahi, K., Vuuren, D. P. van, Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Silva, L. A. D., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M.: The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview, Glob. Env. Chang, 42, 153–168, https://doi.org/10.1016/j.gloenvcha.2016.05.009, 2017.

Riahi, K., Bertram, C., Huppmann, D., Rogelj, J., Bosetti, V., Cabardos, A.-M., Deppermann, A., Drouet, L., Frank, S., Fricko, O., Fujimori, S., Harmsen, M., Hasegawa, T., Krey, V., Luderer, G., Paroussos, L., Schaeffer, R., Weitzel, M., Zwaan, B. van der, Vrontisi, Z., Longa, F. D., Després, J., Fosse, F., Fragkiadakis, K., Gusti, M., Humpenöder, F., Keramidas, K., Kishimoto, P., Kriegler, E., Meinshausen, M., Nogueira, L. P., Oshiro, K., Popp, A., Rochedo, P. R. R., Ünlü, G., Ruijven, B. van, Takakura, J., Tavoni, M., Vuuren, D. van, and Zakeri, B.: Cost and attainability of meeting stringent climate targets without overshoot, Nat. Clim. Change, 11, 1063–1069, https://doi.org/10.1038/s41558-021-01215-2, 2021.

Riahi, K., Schaeffer, R., Arango, J., Calvin, K., Hasegawa, T., Jiang, K., Kriegler, E., Matthews, R., Robertson, S., Sebbit, A. M., Steinberger, J., Khourdajie, A. A., Brutschin, E., Byers, E., Carleton, T., Clarke, L., Chaturvedi, V., Chen, W., and Martínez, J. T.: Mitigation pathways compatible with long-term goals, in: IPCC, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2022.

Schwoon, M. and Tol, R. S. J.: Optimal CO_2-abatement with Socio-economic Inertia and Induced Technological Change, Energy J., 27, https://doi.org/10.5547/ISSN0195-6574-EJ-Vol27-No4-2, 2006.

Su, X., Takahashi, K., Fujimori, S., Hasegawa, T., Tanaka, K., Kato, E., Shiogama, H., Masui, T., and Emori, S.: Emission pathways to achieve 2.0°C and 1.5°C climate targets, Earths Future, 5, 592–604, https://doi.org/10.1002/2016EF000492, 2017.

Tanaka, K. and O'Neill, B. C.: The Paris Agreement zero-emissions goal is not always consistent with the 1.5 °C and 2 °C temperature targets, Nat. Clim. Change, 8, 319–324, https://doi.org/10.1038/s41558-018-0097-x, 2018.

Tanaka, K., Peters, G. P., and Fuglestvedt, J. S.: Policy Update: Multicomponent climate policy: why do emission metrics matter?, Carbon Manag., 1, 191–197, https://doi.org/10.4155/cmt.10.28, 2010.

Tanaka, K., Boucher, O., Ciais, P., Johansson, D. J. A., and Morfeldt, J.: Cost-effective implementation of the Paris Agreement using flexible greenhouse gas metrics, Sci. Adv., 7, eabf9020, https://doi.org/10.1126/sciadv.abf9020, 2021.

Thomson, R. E. and Emery, W. J.: Statistical Methods and Error Handling, in: Data Analysis Methods in Physical Oceanography, Elsevier, 219–311, https://doi.org/10.1016/B978-0-12-387782-6.00003-X, 2014.

Tol, R. S. J., Berntsen, T. K., O'Neill, B. C., Fuglestvedt, J. S., and Shine, K. P.: A unifying framework for metrics for aggregating the climate effect of different emissions, Environ. Res. Lett., 7, 044006, https://doi.org/10.1088/1748-9326/7/4/044006, 2012.

UNFCCC: "Report of the Conference of the Parties serving as the meeting of the Parties to the Paris Agreement on the third part of its first session, held in Katowice from 2 to 15 December 2018. Addendum 2. Part two: Action taken by the Conference of the Parties serving as the meeting of the Parties to the Paris Agreement" (FCCC/PA/CMA/2018/3/ Add.2 2019), 2018.

UNFCCC: "In-session technical workshop on findings on emission metrics contained in the Sixth Assessment Report of the Intergovernmental Panel on Climate Change" (FCCC/SBSTA/2023/INF.2), Bonn, 2023.