Originally published as:

Harmsen, M. J. H. M., Berg, M. van den, Krey, V., Luderer, G., Marcucci, A., Strefler, J., Vuuren, D. P. van (2016): How climate metrics affect global mitigation strategies and costs: a multi-model study. - Climatic Change, 136, 2, 203-216

DOI: 10.1007/s10584-016-1603-7

How climate metrics affect global mitigation strategies and costs: a multi-model study

Mathijs JHM Harmsen[a,b], Maarten van den Berg[b], Volker Krey[c], Gunnar Luderer[d], Adriana Marcucci[e], Jessica Strefler[d], Detlef P Van Vuuren[a,b]

[a] Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, The Netherlands
[b] PBL Netherlands Environmental Assessment Agency, The Netherlands.
[c] International Institute for Applied Systems Analysis (IIASA), Vienna, Austria
[d] Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany
[e] Center of Economic Research, ETH Zurich, Switzerland

Abstract

In climate policy, substitutions metrics are used to determine exchange ratios for different greenhouse gases as part of a multi-gas strategy. The suitability of the metric depends on the policy goals and considerations regarding its practical use. Here, we present a multi-model comparison study to look at the impact of different metrics on the mitigation strategies and global climate policy costs. The study looks into different Global Warming Potentials (GWP) and the Global Temperature change Potential (GTP). The study shows that for all the models, varying between GWPs - from different IPCC reports, with different integration periods: 20 or 100 years - has a relatively small influence on policy costs (< 2.2% spread across scenarios with a 2.8 W/m^2 target) and climate outcomes. Metrics with a constant low substitution value for methane (effectively reducing its abatement), in contrast, lead to higher-cost mitigation pathways (with an average cost increase of 32.8% in a 2.8 W/m^2 scenario). If implemented efficiently, a time-varying GTP leads to a limited cost reduction compared to GWP. However, under imperfect foresight in combination with inertia of CH_4 abatement options, or if implemented sub-optimally, time-varying GTP can result in higher costs than a 100-year GWP. At the same time, given a long-term radiative forcing target, a time-varying GTP results in slightly higher maximum global temperature change rates.

1. Introduction

In addition to CO_2, several other greenhouse gases (GHGs) contribute to climate change. To express the contribution of each of these gases in a common indicator, several emission and climate equivalent metrics have been developed. Such metrics are necessary in tracking total emission trends and to compare the contribution of different countries or sectors to climate change. Yet, arguably the main application of metrics is to allow for substitution of gases in multi-gas emission trading schemes (O'Neill, 2003). It has been shown repeatedly that multi-gas mitigation strategies, which allow substitution across different gases based on marginal costs, are able to achieve targets at lower costs than a CO_2 only approach (van Vuuren et al., 2006b; Weyant et al., 2006).

However, developing a common metric for different climate forcers is far from straightforward, because of the large differences in physical properties of GHGs, such as atmospheric lifetime and radiative potency (Myhre et al., 2013). This is further complicated by the fact that a metric implicitly or explicitly includes value judgments concerning the overall goal of specific climate policy (Deuber et al., 2013; IPCC, 2009). Different climate policy goals may be pursued, such as limiting long-term or short-term temperature change.

In 2011, the UNFCCC requested further research on the impact of the choice of global warming potentials (GWPs) on climate policy strategies (UNFCCC, 2011). In response, several single-model studies were carried out to analyse different metrics (See Section 2). These studies generated some robust results, but also led to varying conclusions. Notably, they do not always agree on the relation between metrics and policy costs in mitigation scenarios. As a result, different most cost-effective metrics have been proposed. Related to this, some studies show large cost differences between metrics, while others indicate a small spread. The problem with these study results is that it is not directly possible to trace differences in outcomes to underlying model assumptions, to input data or to the setup of the experiment. These are all factors that are relevant in a policy context.

With this study, we assessed the use of common metrics in a multi-model comparison study with Integrated Assessment Models (IAMs) to provide insights into key uncertainties and the difference in outcomes of earlier studies and to identify robust results across all IAMs. This analysis was part of the EU FP7 AMPERE project (Kriegler et al., 2014), using the following four IAMs: IMAGE, MERGE_ETL, MESSAGE and REMIND. IAMs are particularly suitable to use for such an inter-disciplinary analysis, because they simulate the interplay of atmospheric-chemical and socio-economic mechanisms.

Our study specifically focussed on assessing the influence of different metrics on mitigation strategies and costs, and how differences in results can be explained by different modelling approaches. For specific metrics, models were compared on GHG emission pathways, policy costs and global mean temperature profiles in achieving the same climate target in 2100, while using the same scenario setup. This study shows the impact of using other metrics than the 100 year GWPs from the IPCC assessment reports for climate policy and related fields (see section 2).

2. Earlier assessments of the influence of different metrics

2.1 Different types of climate metrics

A large number of metrics to convert GHGs to a common unit have been proposed, based on very different principles. One class of metrics are based exclusively on physical parameters. The best known examples are the global warming potential (GWP) and the global temperature change potential (GTP). The GWP is based on the GHG induced integrated total radiative forcing (RF) over a certain timespan. The GTP, one of the most suggested possible alternatives to the GWP, is aimed at optimally reducing temperature change in a specific target year (Shine et al., 2007; Shine et al., 2005). Other metrics focus more on climate change related damages, such as the global damage potential (Kandlikar, 1995; Tol, 1999) or the global cost potential that accounts for the contribution of each gas to overall mitigation costs (Johansson, 2011; Manne and Richels, 2001). Already in 1990, the IPCC stressed that there is no unambiguous methodology for combining all factors in one metric (IPCC, 1990). The clear advantage of physical metrics is that they can be derived in a relatively transparent way. This has the additional advantage that socio-economic uncertainties can be treated separately (Deuber et al., 2013; O'Neill, 2000). Therefore, for actual 'real world' purposes so-far only physical metrics have been discussed. Here, we also focus on a selected set of physical metrics. Economic considerations, such as social discounting and GHG emission mitigation costs, are included in the IAMs that are involved in this study and the combined effect is analysed.

The 100 year GWP from the Second Assessment Report (SAR) is without question the most widely used metric in climate policy. It is used in most climate policies to-date, including for the first commitment period under the Kyoto Protocol as well as in different assessment reports (e.g. the Fourth Assessment Report (AR4)). More recently, it was replaced by the values of 100 year GWPs of

AR4. At the same time, GWPs have also been extensively criticised by natural scientists and economists (Fuglestvedt et al., 2003; Manne and Richels, 2001; O'Neill, 2000; Shine, 2009). Main points of critique have been the arbitrary timespan used as a basis for the metric (20, 100 or 500 years), the lack of rooting in economic theory, and the metrics' inability to reflect damages caused by (the temperature increase due to) climate change.

The time varying global temperature potential (GTP(t)) has been proposed as a suitable candidate for cost-optimal climate policy (Shine et al., 2007). It differs with the GWP in two ways: 1) It compares gases on the basis of the induced temperature change instead of radiative forcing, 2) It focuses on a certain target year (a so-called "snapshot" approach). The GTP(t) does not have a single numerical value for a specific GHG but its value varies over time. A short-lived GHG such as methane thus has a low value (normalized by CO_2) early on, but a very high one when the target year is approached (here: 2100). Depending on the overall goal of climate policy, the GTP(t) can be more cost-efficient as it provides the largest incentives to reduce emissions when it really matters most.

2.2 Effects on policy costs and strategies

Climate metrics are used both to facilitate comparison of past and future emission trends of individual gases as well as to facilitate substitution as part of actual climate policy. In this paper we concentrate on the latter use, as only this influences actual emission reduction strategies. The choice of the metric can influence the overall costs, the emission reduction strategies of individual gases as well as the overall timing of emissions, and thus temperature.

The current literature gives some common conclusions, but also some clearly different messages. In most studies the cost differences as a result of the choice of different metrics are found to be relatively small, when considering the same prescribed climate target (Ekholm et al., 2013; Reisinger et al., 2013; Strefler et al., 2014; van den Berg et al., 2015). However, if relatively long time horizons are assumed, such as with the 500 year GWP or the 100 year GTP, policy costs are likely to increase considerably (in the order of 5% to 20%) (Ekholm et al., 2013; Reisinger et al., 2013; van den Berg et al., 2015). The reason is that with long timescales, methane reduction becomes unattractive because of a low metric value, and other gases have to be abated more at a higher cost. The use of time-dependent, GTP(t) metric, leads to different results: while some studies have reported a cost reduction (Johansson et al., 2006; Reisinger et al., 2013), others have reported equal costs (Strefler et al., 2014) or even higher costs (van den Berg et al., 2015). This study aims to understand the underlying reasons for this diversity in cost estimates resulting from GTP (t) and to conclude the ongoing debate about this topic.

Although global differences in policy costs can be small, the choice of a metric has large implications for the timing and amount of methane emissions (Reisinger et al., 2013; Smith et al., 2013; Strefler et al., 2014; van den Berg et al., 2015; Van Vuuren et al., 2006a). Next to emission reduction profiles, another strategic consideration for the choice of a metric is the induced temperature profile. Temperatures could potentially overshoot unacceptably high before a target year or change too rapidly (Ekholm et al., 2013). By comparing several metrics, used to reach the same two degree climate target, Strefler et al. (2014) found very small differences in maximum transient temperatures ($<= 0.05^oC$), with the slightly lower temperatures generally corresponding with slightly higher policy costs.

The insights discussed above mostly result from single model studies that often use somewhat different assumptions. By comparing multiple models following exactly the same approach it is possible to thoroughly check these results and to consolidate the understanding of the impact of metric choice on transformation pathways. A model intercomparison approach, as the one adopted in this paper, can be a powerful tool in deriving robust conclusions required for informing decision makers.

3. Research methods

For this study, results have been used from several IAMs that were involved in the EU FP7 AMPERE project: IMAGE, MERGE_ETL, MESSAGE and REMIND (see Table 1)(Kriegler et al., 2014). These represent a range of different models. One distinction between models is how they describe relevant economic processes. These can be based on the concept of economic equilibrium, when aiming for a minimum overall cost from a centralized perspective, taking into account price-elasticity in supply and demand. Some of these models focus on specific sectors (partial equilibrium), while others focus on overall macro-economic impacts (general equilibrium). Another important distinction is the focus on optimisation versus simulation. The optimization models MERGE_ETL, MESSAGE and REMIND include foresight of future supply and demand to reach an optimal least cost solution. The simulation model included in this study (IMAGE) has no foresight, but is still able to derive least-cost climate policies in a recursive dynamical way by means of the sub-model FAIR-SiMCaP (Den Elzen et al., 2007). The Supplementary Material provides additional information about the climate modules used by the models and shows these perform well in emulating climate mechanisms.

Table 1: Integrated Assessment Models used in this study

Model	*Model category*	*Solution dynamics*	*Policy costs*
IMAGE	Partial equilibrium	Recursive dynamic	Area under MAC curve
MERGE_ETL	General equilibrium	Intertemporal optimization	GDP loss
MESSAGE	General equilibrium	Intertemporal optimization	Consumption loss
REMIND	General equilibrium	Intertemporal optimization	Consumption loss

All models have been used to generate cost-optimal trajectories towards the same global radiative forcing (RF) targets for the year 2100. The models aimed to meet two RF targets: 2.8 and 3.7 W/m^2 increase compared to pre-industrial levels. The former is a stringent climate goal associated with a 2 degree temperature rise, while the second leads to an approximate 2.5 degree increase, which allows for more flexibility in emission reduction strategies[1]. In combination with each target, 12 scenarios were prescribed, which differed in the use of a climate metric and the overall climate target (see Table 2). The employed metrics are: the 100 year GWP based on the SAR, the 100 and 20 year GWPs based on AR4, the 100 year GTP and the time-varying GTP. In addition, MERGE_ETL made use of an additional time-varying metric, MERGE_RF (t). This metric aims for a least-cost solution to reach the RF target, based on the Global Cost Potential (GCP) approach by Manne and Richels (2001) (see Table 2). The behaviour of this metric is very comparable to GTP(t). Models were allowed to simulate an overshoot in RF (and temperature) in the years before 2100. Furthermore, scenarios were based on the assumption of full globally integrated carbon markets, implying equal marginal CO_2 and non-CO_2 GHG abatement costs across all regions. Full technology availability was assumed. The model projections were compared in terms of climate policy costs; CO_2, CH_4 and N_2O emission

[1] For the radiative forcing target the so-called "AN3A" metric was used to generate results comparable to the widely used Representative Concentration Pathways (RCPs) (Van Vuuren et al., 2011). This metric includes all anthropogenic forcing agents except direct forcing from land albedo changes, mineral dust and nitrate aerosols. This means that the total radiative forcing target is +/- 0.2 W/m^2 higher.

profiles; carbon price profiles; global temperature change in 2100 (compared to the pre-industrial level); and maximum global temperature change before 2100.

Table 2: Climate metric scenarios

Climate metric	*Target*	*CH4 (CO2 = 1)*	*N2O (CO2=1)*	*Information*
GWP 100 (SAR)	2.8 W/m^2 or 3.7 W/m^2	21	310	GWP metric used in 1st commitment period of the Kyoto Protocol
GWP 100 (AR4)		25	298	GWP metric used in 2nd commitment period of the Kyoto Protocol
GWP 20 (AR4)		72	289	
GTP 100		0.4	265	Numerical values based on (Shine et al., 2005) *
GTP (t)		0.8 (2010) - 102 (2100)	280 (2010) - 216 (2100)	Numerical values based on (Shine et al., 2005) *
MERGE_RF (t)		RF target dependent	RF target dependent	Approach based on (Manne and Richels, 2001)**

* In this study, the simple approach from Shine et al. 2005 to calculate GTP100 and GTP(t) was used. Note that the resulting values for methane are markedly lower (by up to a factor of 10 for GTP100) than the values for GTP presented in the recent IPCC assessment (Myhre et al 2013). The implications of this approach to calculating GTP are discussed where relevant in the main text; the specific metric values generally do not alter the main conclusions of this study, but they have been shown to influence calculated mitigation costs for some models in specific circumstances.

** The Global Cost Potential (GCP) proposed by Manne and Richels, used as a basis for the "MERGE_RF" metric, also included economic considerations (as represented in the MERGE model) and was originally used in combination with a temperature change target. Substitution in MERGE_ETL is derived from radiative forcing expressions from the IPCC Third Assessment Report (IPCC, 2001)(Table 6.2).

4. Results

4.1 Emission profiles

We first look at the impacts of the different metrics on emissions (Figure 1). In the discussion of results, we focus on the 3.7 W/m^2 scenarios. The results for the 2.8 W/m^2 are very similar, and are therefore only briefly summarised while detailed results are found in the Supplementary Material.

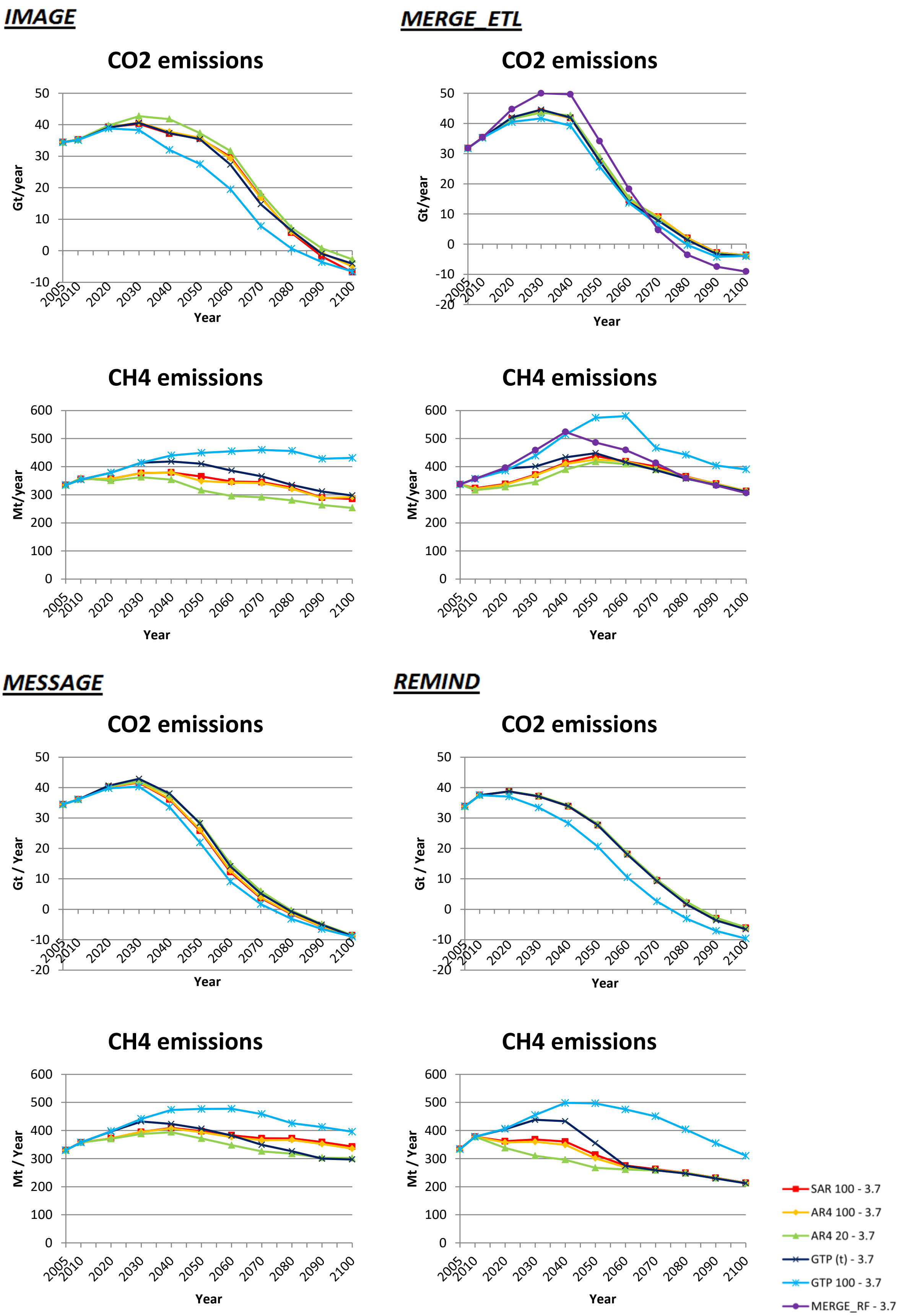
For all models, the use of different metrics results in clear differences in reduction strategies for methane. Typically, higher methane emission reductions correspond with higher metrics values for methane, as this increases the relative value of the gas in reduction strategies. In the case of MERGE_ETL and REMIND differences between scenarios at the end of the century are only small. The reason is that unlike in the other models, maximum reductions are effectively reached in all scenarios (except GTP-100). The marginal abatement cost (MAC) curves, used in the model to calculate emission reductions at various price levels, limit the maximum reduction potential (see Supplementary Material for analysis of the methane MAC curves). At a certain high amount of emission reductions, higher prices hardly influence reduction rates. For REMIND, this can be seen in Figure 1 where all scenarios (except GTP-100) show almost equal methane emissions in 2060 (close to 243 Mt) until 2100 (close to 210 Mt). This increase in reduction potential over time can be attributed to technological learning and is based on Lucas et al., 2007 (Lucas et al., 2007).

In the GTP-100 scenario, models consistently show relatively high methane emissions compared to other metrics. This is a result of the low metric value. The difference of GTP-100 to GWP-100 methane emissions is somewhat greater in REMIND and IMAGE than in the other models. This can be attributed to the higher methane abatement potential considered in these two models. With GTP(t), methane emission reduction also starts late because of flexible valuation, yet it generally leads to the lowest emissions in the long run compared to other scenarios. Here, IMAGE forms an exception. In the model, GTP(t) does not lead to minimum methane emissions before the climate target is reached, despite high metric values later in the century. This is caused by the assumed limitation of year-to-year changes in methane emission reduction levels, as described by Van den Berg et al. (2015). Annual changes in methane reduction rates are assumed to be limited to 2.5% to 5% of the baseline emissions, depending on the source. This inertia effect implies that in extreme delay scenarios, it is not possible to achieve similar reduction rates by the end of the century. Therefore, reduction measures have to be taken early enough to fully exploit the potential in later years. This assumption is also relevant because IMAGE is a simulation model, and does not have perfect foresight in its annual investment decisions (that are thus purely guided by the current value of the metric and gas-specific reduction curves).

In general, CO_2 emission profiles are relatively similar for the different metrics in the different models. Scenario differences in CO_2 emissions are mainly the result of compensating for higher or lower methane emissions. Only the MERGE_RF metric leads to high CO_2 *and* methane emissions early in the century. With this metric, MERGE strongly favours late century mitigation of GHG emissions, including a much larger deployment of bioenergy in combination with carbon capture and storage (BECCS) to ensure negative carbon emissions.

The choice of a metric has a very small effect on the N_2O emission profile (shown in the Supplementary Material). Differences between scenarios are never larger than 5 % (except GTP-100 at 2.8 W/m^2 in IMAGE due to a very high carbon price) and are smaller than 0.5% for most years and all models.

Figure 1: Scenario results: Emissions



4.2 Policy costs

Figure 2 shows the carbon price and policy cost profiles in all models for the 3.7 W/m^2 scenarios. The different methods used in the models to calculate policy costs imply that absolute values cannot be compared among them. Therefore, we concentrate on analysis of the relative differences in costs across scenarios. In Table 3, this is shown for all scenarios (including those with a 2.8 W/m^2 target), with the total integrated discounted policy costs in 2100 expressed in relative difference to SAR GWP-100 (discount rate = 5%).

Overall, cost differences as a result of the use of different GWP metrics are small at the global scale. Particularly, substitution across different values for GWP-100 has hardly any impacts: The AR4 GWP seems to lead to a slightly more cost-efficient climate policy than the SAR GWP, considering the similar result for both forcing targets. Models disagree on the effect of changing the GWP time horizon from 100 to 20 years, but again, the overall effect on policy costs is relatively modest. Only IMAGE in the 3.7 W/m^2 projects significantly lower costs for the GWP 20 (-5.7%), due to an early start in methane abatement leading to a higher reduction potential (in contrast to GTP (t), further explanation below).

The only metric that consistently leads to higher policy costs is the 100 year GTP. All models agree that using this metric would lead to much more expensive climate policy. Because of very low valuation in the metric, mitigation of methane reduces considerably and as compensation, CO_2 emissions are reduced much more at a far higher carbon price (see Figure 2). This clearly shows the advantage of a multi-gas strategy. For IMAGE in the 2.8 W/m^2 scenario, this effect almost leads to a doubling of policy costs. The reason is that in this extreme case, the target can only be met with a very early start of CO_2 emission reduction at a much less discounted (and thus higher) carbon price. Note that the numerical value for methane in this metric is uncertain and very low in this study. If a higher value is used, e.g. 4 in Myhre et al. (2013), the policy costs are expected to be lower.

The GTP(t) is closely related to the Global Cost Potential (Tol et al., 2012), and therefore can be expected to be close-to-optimal in terms of the costs to reach a prescribed climate target (Shine et al., 2007). This was confirmed by the earlier modelling study by Reisinger et al. (2013), which made use of MESSAGE, but not found by the other models participating in this study Figure 3 shows the methane abatement and policy cost profiles in all the models for the GTP(t) 2.8 W/m^2 scenario. Using the GTP(t) metric leads to a shift of methane abatement closer to the target year, which could reduce the mitigation costs by avoiding too early reductions of methane. MESSAGE, and to a lesser degree MERGE_ETL do indeed show lower costs due to optimal timing of emission reductions (up to 4.7% in the 2.8 W/m^2 scenario). The main reason that MESSAGE shows the lowest costs for GTP(t) is the assumed increase in methane reduction potential at higher carbon prices towards the time horizon. This is shown in a detailed analysis of the methane marginal abatement cost (MAC) curve analysis in the Supplementary Material. This is in line with the earlier findings from Reisinger et al. (2013).

However, there are several factors that might counteract the cost advantage of GTP(t). One reason is that the metric is aimed at a temperature target instead of a radiative forcing target, although that effect is small given the large correlation between these parameters. The different scenarios that focused on the same forcing target using different metrics lead to very similar temperature levels (see section *temperature*).

Another factor is the lack of methane reduction potential (as explained in the MAC analysis in the Supplementary Material). REMIND shows costs that are almost equal to GWP 100 for GTP(t) (in the 3.7 W/m^2 scenario) up to slightly higher policy costs (in the 2.8 W/m^2 scenario). The main reason is that mitigation is limited as maximum reductions are effectively reached, leading to similar CH_4 and CO_2 trajectories (see section *emission profiles*).The slightly higher cost for GTP(t) is caused by stronger CO_2 mitigation to compensate for higher CH_4 emissions at the beginning of the century. It is important to note that the values used in the GTP metric to ensure optimal substitution depend on climate model assumptions such as the CO_2 background concentration. Here, we used the numerical values of GTP(t) as provided in Van den Berg et al. (2015) based on the original study (Shine et al., 2005). The results would have looked different if the climate modules native to the integrated assessment models had been used to derive the GTP(t) values. In fact, in a single model study in which climate metrics were compared with REMIND, GTP(t) did lead to slightly lower policy costs, due to optimal tuning to the model (Strefler et al., 2014).

For IMAGE, GTP(t) clearly leads to higher costs, particularly in the 2.8 W/m^2 scenario. This is a result of an assumed inertia effect in the upscaling of methane abatement measures, which prevents a fast increase in emission reductions (as explained in the *emission profiles* section). Without this effect GTP(t) is shown to lead to least cost trajectories in IMAGE (van den Berg et al., 2015). The rapid change in methane emission related policies might therefore be another reason why GTP(t) is less cost-efficient in practice. For the same reason, the 20 year GWP is the most cost-efficient metric in IMAGE. In that scenario, methane emission reductions are maximized because of an early start in implementing abatement measures.

Figure 2: Scenario results: Carbon price and Policy costs (note: different scales on y-axis)

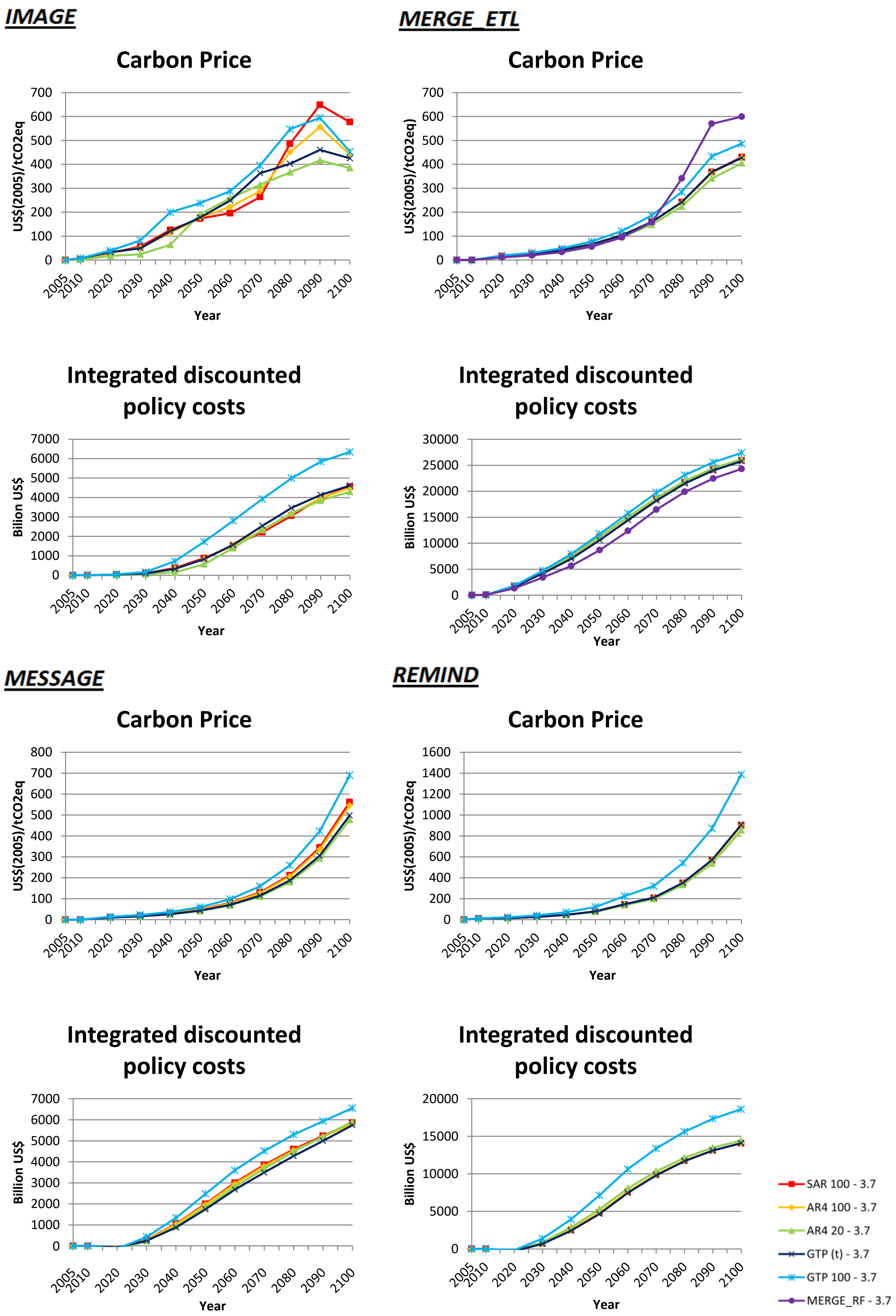
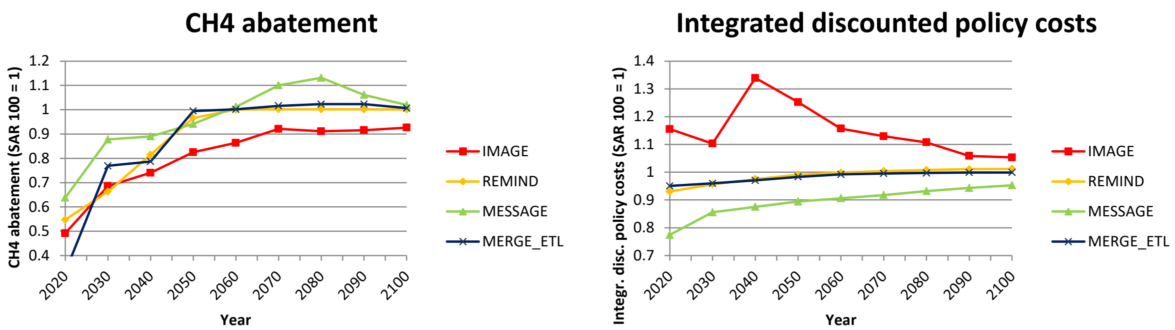


Table 3: Integrated discounted policy cost in 2100, for all models and all scenarios, relative to SAR-100 (= 100%)

RF target	Metric /Scenario	IMAGE	MERGE_ETL	MESSAGE	REMIND	Average
3.7 W/m2	GWP 100 (AR4)	-1.4%	0.2%	-0.8%	0.0%	-0.5%
	GWP 20 (AR4)	-5.7%	1.0%	0.8%	2.6%	-0.3%
	GTP 100	39.2%	5.6%	11.8%	32.2%	22.2%
	GTP (t)	1.4%	-0.6%	-2.0%	0.2%	-0.3%
	MERGE_RF (t)		-6.3%			-6.3%
2.8 W/m2	GWP 100 (AR4)	-1.2%	0.3%	-1.9%	0.1%	-0.7%
	GWP 20 (AR4)	-2.2%	2.0%	0.9%	1.4%	0.5%
	GTP 100	89.6%	7.3%	13.7%	20.6%	32.8%
	GTP (t)	5.3%	-0.1%	-4.7%	1.2%	0.4%
	MERGE_RF (t)		-1.3%			-1.3%

Figure 3: Methane abatement and integrated discounted policy cost profiles for the GTP (t) scenario, relative to SAR-100 (= 1), for all models (RF target = 2.8 w/m^2)



The MERGE_RF metric, only used in the MERGE model, leads to considerably lower policy costs, especially in the 3.7 W/m^2 scenario. This metric is, however, optimally tuned to the MERGE model and will not be optimal in another model. In addition, the potential limitations associated with GTP(t) will also apply to this metric.

4.3 Temperature

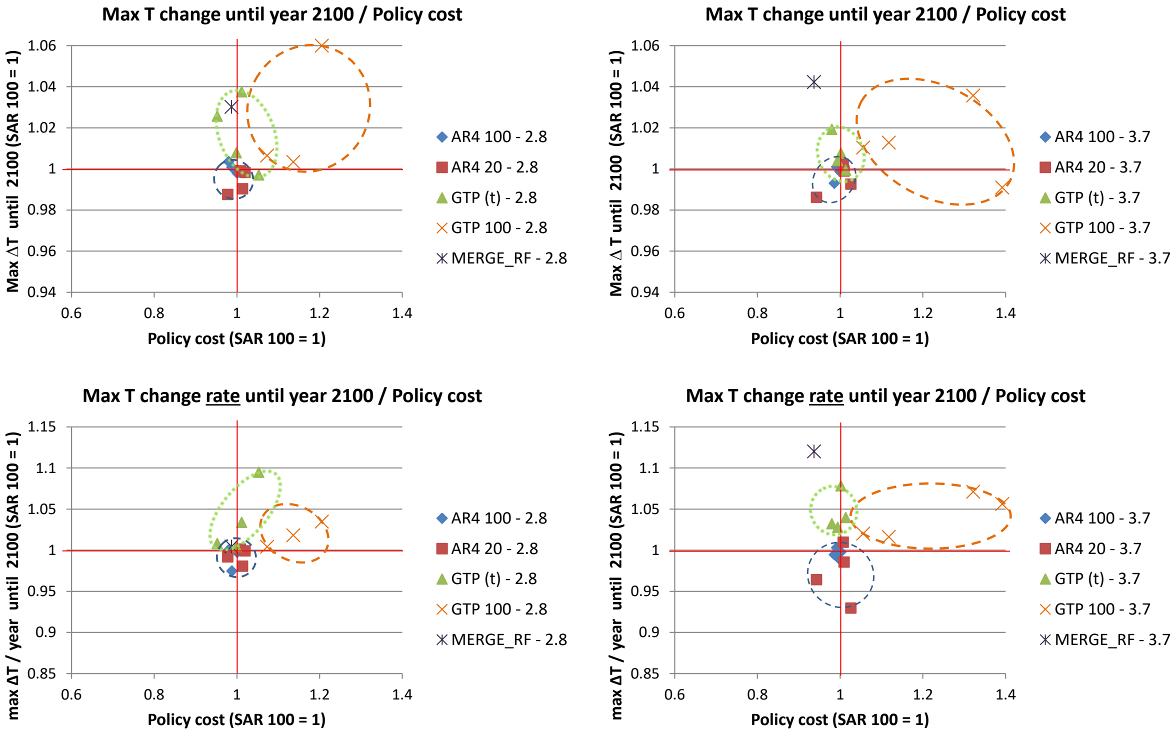
The effect of the choice of a certain metric on temperature change, compared to pre-industrial temperatures, is very small. However, GTP(t), GTP-100 and the MERGE_RF metric do lead to slightly higher than average maximum transient temperatures and temperature change rates. In Figure 4, this is shown with a policy cost / temperature change plot, based on the maximum temperature until the year 2100 (upper panel) and the maximum temperature change *rate* until 2100. The scenario results have been normalized by the values for SAR-100 and are given for all models and both forcing targets.

When only considering the maximum temperature change (upper panel), the highest temperature change levels result from GTP-100 in the 2.8 W/m^2 scenario in REMIND (6%) and from MERGE_RF in the 3.7 scenario (4.2%). This roughly corresponds to a 0.1°C higher maximum temperature than SAR-100. In all models, except IMAGE, GTP(t) also leads to higher maximum temperatures. For GTP-100, GTP(t) and MERGE_RF, the slightly higher temperature is the result of late methane mitigation. Only in the case of MERGE_RF this is further aggravated by late mitigation of CO_2 emissions. One exception to this effect is the GTP-100 result from IMAGE, particularly in the 2.8 W/m^2 scenario. There, due to very early CO_2 mitigation, the maximum temperature change before 2100 is actually 5% lower, but at a policy cost of 1.9 times that of SAR-100. For that reason, the metric falls outside the scale of the figure and using it for reducing temperature change can be considered highly unrealistic. All GWP metrics (blue circles in Figure 4), as well as the GTP(t) metric are within 1.5% or 0.04^{o}C of the SAR-100 result. Within that range, the slightly lower temperatures in the AR4-20 scenarios are seen across the models. This can be explained because of an early methane mitigation resulting in early radiative forcing reduction and a decrease in temperature.

A roughly similar result emerges when considering the maximum temperature *rate* until 2100 (lower panel). GTP(t), GTP-100 and MERGE_RF generally lead to higher temperature change rates due to a lack of methane mitigation in the first decades (up to 12% higher for MERGE_RF in the 3.7 W/m^2). The maximum ΔT / year is approximately 0.03 ^{o}C/year, so the expected difference in temperature change rate with these metrics is small in absolute terms, not exceeding 0.0036 ^{o}C/year and generally leading to less than half that value. GWP-20 can potentially lead to a slightly lower maximum temperature change rate, due to higher short-term methane abatement.

The small differences between metrics imply that the choice of a metric hardly needs to be motivated by the effect it has on the maximum temperature change. This also holds for temperature differences in the target year (see Supplementary Material). Related to this, it can be argued that a metric based on radiative forcing is quite suitable for policy involving temperature change targets, given the proportionality between radiative forcing and temperature levels in the target year.

Figure 4: Policy cost / Temperature change plot. Shown for both forcing targets and all models. Normalized by SAR-100 (= 1), indicated by the red lines. GWP metrics are encircled by blue lines, GTP (t) by green lines and GTP-100 by orange lines. **Upper panel:** Integrated discounted policy cost / Maximum temperature change until the year 2100. **Lower panel:** Integrated discounted policy cost / Maximum temperature change *rate* until the year 2100 (note the different y-axis scale).
The GTP-100 2.8 W/m^2 scenario is not shown for IMAGE on this scale (with a policy cost of 1.9 times that of SAR_100, a 5% lower max T and a 5% higher max T rate)



Discussion and conclusion

By using a multi-model analysis, this study aims to understand how different climate metrics influence climate change mitigation strategies and costs. The multi-model approach avoids differences in results due to different experimental set-ups. The models in this study base their projections on a large body of work and aim to include most factors that are relevant to climate change mitigation. The conclusions of this study relate to policy relevant parameters such as emission pathways, policy costs and temperature changes. As such, some conclusions have been found in earlier, single-model studies, and are found to be robust despite the differences across the models. In this paper, we show how model assumptions on emission reduction potentials, inertia and foresight can differently affect the resulting effects of various metrics

In the past, models have reported different findings with respect to the impact of metric choice on mitigation strategies. Despite model differences, they all find that the different consequences of metrics are primarily caused by diversity in methane emission abatement strategies. Metrics have a clear impact on methane emission reduction levels. Only when models reach their maximum abatement potential, different metrics lead to similar emission levels. This also implies that different model assumptions on methane abatement potential lead to different projected outcomes for the same metric (e.g. as described below for GTP (t)). N_2O emissions reductions are hardly influenced by the metric choice (usually less than 0.5 % difference between scenarios). Differences in CO_2 emissions across scenarios are relatively small and tend to compensate for higher or lower methane emissions.

The time varying GTP(t) can lead to cost optimality under perfect world conditions, but could lose this advantage when implemented. Differences in the projected cost-effectiveness of the metric trace back to model assumptions on methane abatement. Two models showed the time-dependent GTP metric as defined in this study to lead to slightly lower costs than other metrics for achieving a long-term climate target, in line with previous work. An analysis of the models' marginal abatement cost curves for methane showed that the possibility of additional methane reductions at higher mitigation costs contribute to the cost-effectiveness of the metric. However, other model outcomes indicated that the advantage of avoiding too early methane reductions might be counteracted by technical limitations in combination with imperfect foresight. This inertia effect implies that methane reductions have to start long before the target year, making the metric ineffective and costly. Another reason for increased policy costs might be a deviation from the optimal time-variant trajectory of CH_4-to-CO_2 exchange ratios. In all models, we found the cost difference between the time-varying GTP(t) and 100-year GWP to be relatively small (<5%).

Models consistently show that most GWP metrics that are considered for policy making lead to very similar global mitigation costs. The reason is that all GWP metrics allow for sufficient non-CO_2 emission reduction. Especially substitution between different values for GWP-100 from different Assessment Reports (AR) does not lead to important changes in overall global cost levels (with an average difference in policy costs of -0.7% to -0.5% between the second and fourth AR). The same is true for changes in the metrics with different time-horizons (with an average difference of -0.3% to 0.5% between a 100 and 20 year horizon).

The 100 year GTP with a low valuation of methane emissions leads to high policy costs in all model projections. Compared to GWP 100 it led to an increase of 6% to 40%, and in a single case 90%. This high cost increase can partly be explained by the very low methane valuation used in this study and implies that constant time horizon metrics need to valuate methane mitigation sufficiently in order to be cost-efficient.

Models agree that GTP(t) and GTP-100 would lead to a small increase in the maximum temperature *rate of change.* However, the effect of the use of different metrics on maximum temperature change is very limited. Although the induced temperature profile could be a relevant strategic consideration in climate policy, the choice of a metric does not have to play a large role.. Given the proportionality between radiative forcing and temperature levels in the target year, it can be argued that a metric based on radiative forcing would not lead to ineffective policy aimed at reaching temperature targets.

From a global perspective, and in the long term, the 100-year GWP metric seems to lead to relatively attractive outcomes in terms of mitigation costs and climate outcomes, and no reason is found to replace it as the most common metric used in climate policy. However, there are possible considerations that could lead to alternatives. For policy making, the choice of timing of methane reductions impacts short-term co-benefits with respect to air pollution, costs and temperature. As such, the choice of metric can be used to influence policy decisions in the short term. Alternatively, it is also possible to consider separate abatement strategies for long-lived and short-lived greenhouse gases so that independent choices can be made with respect to the different advantages and disadvantages of reducing short-lived GHG emissions. This would lose, however, the advantage of common framework for short and long-lived Kyoto gases.

Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme FP7/2010 under grant agreement n°265139 (AMPERE) and under grant agreement n° 282846 (LIMITS)

Supplementary Material

1) Climate modules in IAMs in this study

IMAGE, MESSAGE and REMIND make use of the MAGICC model (Model for the Assessment of Greenhouse-gas Induced Climate Change) for their climate calculations. This is a reduced-complexity coupled global climate-carbon cycle model that emulates the behaviour of complex atmospheric chemistry and climate models. MERGE_ETL uses its own simple climate model (SCM), as an integral part of its modelling framework, that makes use of equations based on the IPCC's Third Assessment Report (TAR) (IPCC, 2001). See below for a detailed description of the MAGICC versions and the climate submodel in MERGE[2].

All four models are expected to follow the patterns of complex climate models well. The behaviour of the relevant MAGICC versions and the MERGE SCM has been compared to that of complex models in terms of CO_2 (Van Vuuren et al., 2011) and non-CO_2 GHGs (Harmsen et al., 2015). It was shown that the SCMs are within the expert model range, which is also quite large due to scientific uncertainty. The models were found to be similar in the representation of climate mechanisms related to CO_2 and CH_4 (the most important gases in this study).

For these reasons, the comparison in this study can also be considered relevant and lead to robust conclusions. Only small differences between models can be explained by differences in the climate related assumptions.

IMAGE

IMAGE uses the simple climate model MAGICC 6.0 (Meinshausen et al., 2011a; Meinshausen et al., 2011b) to simulate the effects of changing greenhouse gas emissions on atmospheric composition, radiative forcing and global mean temperature. MAGICC was used extensively in the Third, Fourth, and Fifth assessment reports of IPCC (Intergovernmental Panel on Climate Change) in assessing a range of greenhouse gas concentration scenarios. Since publication of these reports, MAGICC has been updated in line with results from Atmosphere-Ocean General Circulation Models (AOGCMs).

There is still considerable uncertainty in climate change simulations, as illustrated by differences in results from various AOGCMs, in terms of mean global temperature, and even more so in geographical patterns of surface temperature and precipitation. By adjusting the values of a few of the model parameters, MAGICC 6.0 can reproduce time dependent responses of AOGCMs (Meinshausen et al., 2011a; Meinshausen et al., 2011b). This allows IMAGE to reflect the uncertainty in AOGCM results, and to provide plausible projections of future climate-change feedbacks and impacts.

The analysis of climate impacts and feedbacks requires location-specific temperature and precipitation changes. Thus, a pattern scaling technique is applied in IMAGE by combining MAGICC results with maps on climate change from the same AOGCMs assessed in AR4 (IPCC, 2007) and used for calibrating MAGICC. The consistent combination of AOGCM-specific parameter settings for MAGICC and matching geographical patterns of climate change make the dynamic results from IMAGE physically more consistent, and extend the range of uncertainties that can be covered to include future climate change.

MESSAGE

The MESSAGE modeling framework makes use of the MAGICC 5.3 reduced-complexity global climate model. MAGICC stands for the "Model for the Assessment of Greenhouse-gas Induced Climate Change." It is a reduced-complexity coupled global climate-carbon cycle model in the form of a user-friendly software package that runs on a personal computer (Wigley 2008). In its standard form, MAGICC calculates internally consistent

[2] The text for IMAGE, MESSAGE and REMIND is taken from the following model documentation website (under review), created as part of the FP7 project ADVANCE: https://wiki.ucl.ac.uk/display/ADVIAM/Models. The description of the climate submodel in MERGE_ETL is from: Marcucci, A. and Turton, H. "The MERGE-ETL model: Model Documentation", PSI report, 2012. http://www.psi.ch/eem/ModelsEN/2012MergeDescription.pdf

projections for atmospheric concentrations, radiative forcing, global annual-mean surface air temperature and other metrics, given emissions trajectories of a range of gases (CO_2, CH_4, N_2O, CO, NO_x, VOCs, SO_2, and various halocarbons, including HCFCs, HFCs, PFCs, and SF_6), all of which are outputs from MESSAGE. The time horizon of MAGICC extends as far back as 1750 and can make projections as far forward as 2400. The climate model in MAGICC is an upwelling-diffusion, energy-balance model, which produces output for global- and hemispheric-mean temperature and for oceanic thermal expansion. Climate feedbacks on the global carbon cycle are accounted for through the interactive coupling of the climate model and a range of gas-cycle models. As of mid-2013, version 5.3 of MAGICC is being used in conjunction with MESSAGE. This version is consistent with the IPCC Fourth Assessment Report (AR4), Working Group 1, except that the model has been slightly modified to permit the explicit treatment of black and organic carbon (BC and OC) and their impacts on the global climate. (The MESSAGE team gratefully acknowledges Dr. Steve Smith of the Pacific Northwest National Laboratory (USA) for sharing a modified version of MAGICC (v5.3), which explicitly takes user-specified trajectories of BC and OC as inputs.)

In contrast to how a typical user would normally operate MAGICC, it has become common for the MESSAGE team to run the model stochastically in order to generate probabilistic estimates of climate system responses (e.g., temperature increase or atmospheric GHG concentrations), a methodology first described in Keppo et al. (2007). In short, a single set of emissions trajectories is run in MAGICC under 100 different sets of parameter assumptions. This allows one to explore the uncertainty in climate system responses for a single set of emissions trajectories (from the MESSAGE scenario output) by using a probability density function (PDF) to describe the following parameters: climate sensitivity, ocean diffusivity, and aerosol forcing. Therefore, instead of simply saying that, for a given mitigation scenario and emissions trajectory, "the projected maximum global temperature increase over the course of the twenty-first century is estimated at X ºC", one can say something like "the probability of staying below X ºC maximum global temperature increase is Y%."

The reason the projections of climate system responses are estimated probabilistically is because of the large amount of uncertainty in the three climate system parameters mentioned above. Perhaps the most important among these, and one of the most uncertain, is climate sensitivity, which refers to the equilibrium global average warming expected if CO_2 concentrations were to be sustained at double their pre-industrial values. This value is estimated by the IPCC AR4 "as likely to be in the range 2 to 4.5 ºC with a best estimate of about 3 ºC" (IPCC 2007). Contributing to the IPCC AR4 were a number of studies that estimate PDFs for climate sensitivity (see Meinshausen et al. (2009) and O'Neill et al. (2010) for good reviews). And as Figure 1 illustrates, the shape of these PDFs can be quite different. A typical probabilistic application of the joint MESSAGE-MAGICC set-up involves dividing each of these PDFs into 100 steps between 0.1 and 10 ºC. PDFs for ocean diffusivity and aerosol forcing – two other important, though uncertain, climate parameters – are also generated by correlating them with climate sensitivity at each step (Meinshausen 2006). The Forest et al. (2002) distribution with uniform priors (bold line in figure) is often chosen because it is near the middle of the range found in the literature: a climate sensitivity value of 3 ºC has a (cumulative) likelihood of 53.9% using the PDF from Forest et al. (2002).

REMIND

By default, REMIND is coupled to the MAGICC 6 climate model to translate emissions into changes in atmospheric composition, radiative forcing and temperature increase. Due to numerical complexity, after running REMIND we perform the evaluation of climate change using MAGICC. Iterative adjustment of emission constraints or carbon taxes allows meeting specific temperature or radiative forcing limits in case of mitigation scenarios (see Section 1.4).

In addition, REMIND includes a reduced-form climate model similar to the one used in DICE (Nordhaus and Boyer 2000) which can be used within the REMIND optimization to enable direct formulation of temperature or radiative forcing targets in climate mitigation scenarios. It comprises (1) an impulse-response function with three time scales for the carbon cycle, (2) an energy balance temperature model with a fast mixed layer, and (3) a slow deep ocean temperature box. Equations in the carbon-cycle temperature model describe concentration and radiative forcing that result from CH_4, N_2O, sulphate aerosols, black carbon, and organic carbon (Tanaka and Kriegler 2007). The climate module determines the atmospheric concentrations of CO_2,

CH_4, and N_2O and computes the resulting radiative forcing and mean temperature at the global level. Its key parameters are calibrated to reproduce MAGICC, with a climate sensitivity of around 3.0°C.

MERGE

(From: Marcucci, A. and Turton, H, 2012)

The climate submodel represents carbon and non-CO_2 gases cycles to estimate atmospheric concentration of GHGs, and then calculates the radiative forcing and global temperature change.

Emission to concentration parameter

	CO2	**CH4**	**N2O**
Expression from IPCC the science basis, 2001.	1ppmv co2 = 2.12 Gt C	1745 ppb = 4850 Tg (Mt) CH_4	314 ppb = 1510 Tg N 314 ppb = 1510*44/28 Tg N_2O
Factor	0.472 ppm/Gton C	359.794 ppb/Mt CH_4	132.33ppb/Mt N_2O

CO_2

The carbon cycle in MERGE is based on the atmospheric CO_2 impulse-response estimated by Maier Reimer and Hasselman (1987), representing five independent atmospheric reservoirs, with A capacity and T time constant of absorption. The atmospheric CO_2 response $y\delta(t)$ to an unitary impulse $\delta(t)$ is given by,

$$y_\delta(t) = A_0 + \sum_{j=1}^{4} A_j e^{-t/T_j}$$

The CO2 concentration in each time period corresponds to the sum of the impulse responses during the period plus the remaining carbon in each reservoir. The parameters Aj and Tj used in MERGE-ETL correspond to the 2x fit (Maier-Reimer and Hasselman, 1987):

	Reservoir				
	0	1	2	3	4
A_j	0.142	0.241	0.323	0.206	0.088
T_j [years]	∞	313.8	79.8	18.8	1.7

CH_4 and N_2O

The behaviour in the atmosphere of the other greenhouse gases is modelled using a single reservoir representation, based on Intergovernmental Panel in Climate Change (IPCC) (1997, App. 1). The atmospheric concentration is calculated using the impulse response of the reservoir, thus,

$$c_g(t) = Me^{-t/\tau_g}$$

where M is the magnitude of the impulse. Based on Intergovernmental Panel in Climate Change (IPCC, 2001), the atmospheric lifetime of the non-CO_2 GHG are: τCH_4=12 years; τN_2O = 114 y; τSLF = 13.8 y; and τLLF = 3200 y.

Concentration to forcing

For the greenhouse gases, the change in radiative forcing is calculated using the simplified expressions based on the IPCC Third Assessment Report (Intergovernmental Panel in Climate Change (IPCC), 2001, Table 6.2). These expressions depend on the current concentration of each greenhouse gas: CO_2 [ppm], CH_4 [ppb], N_2O [ppb], SLF [ppb] and LLF [ppb]; and the pre-industrial concentration indicated with the subindex $_o$:

Gas	Change in net flux [W/m^2]
CO_2	$5.35\ln\left(\frac{CO_2}{CO_{2o}}\right)$
CH_4	$0.036\left(CH_4^{0.5}-CH_{4o}^{0.5}\right)-f(CH_4,N_2O)^*-f(CH_{4o},N_2O)$
N_2O	$0.12\left(N_2O^{0.5}-N_2O_o^{0.5}\right)-f(CH_{4o},N_2O)-f(CH_{4o},N_2O_o)$
SLF†	$0.15\ (SLF-SLF_o)$
LLF‡	$0.52\ (LLF-LLF_o)$
CFC-11	$0.25\ (CFC\text{-}11-CFC\text{-}11_o)$
CFC-12	$0.32\ (CFC\text{-}12-CFC\text{-}12_o)$

$^* f(CH_4,N_2O)=0.47\ln\left[1+2.01\times10^{-5}(CH_4\cdot N_2O)^{0.75}+5.31\times10^{-15}CH_4(CH_4\cdot N_2O)^{1.52}\right]$

†Corresponds to the HFC-134a value

‡Corresponds to the SF_6 value

References

Intergovernmental Panel in Climate Change (IPCC). "An introduction to simple climate models used in the IPCC second assessment report.", 1997.

Intergovernmental Panel in Climate Change (IPCC). Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report. Cambridge University Press, 2001b

Maier-Reimer, Ernst, and Klaus Hasselman. "Transport and storage of CO2 in the ocean." Climate Dynamics 2: (1987) 63–90.

Marcucci, A. and Turton, H. "The MERGE-ETL model: Model Documentation", PSI report, 2012. http://www.psi.ch/eem/ModelsEN/2012MergeDescription.pdf

Van Vuuren DP, Lowe J, Stehfest E, Gohar L, Hof AF, Hope C, Warren R, Meinshausen M, Plattner GK (2011b) How well do integrated assessment models simulate climate change? Climatic Change 104:255-285.

Mathijs JHM Harmsen, Detlef P van Vuuren, Maarten van den Berg, Andries F Hof, Chris Hope, Volker Krey, Jean-Francois Lamarque, Adriana Marcucci, Drew T Shindell, Michiel Schaeffer.
(2015) How well do integrated assessment models represent non-CO_2 radiative forcing? Climatic Change (in press)

2) Scenario results in the 2.8 W/m^2 scenarios

Figure S1: Emissions in the 2.8 W/m^2 scenarios

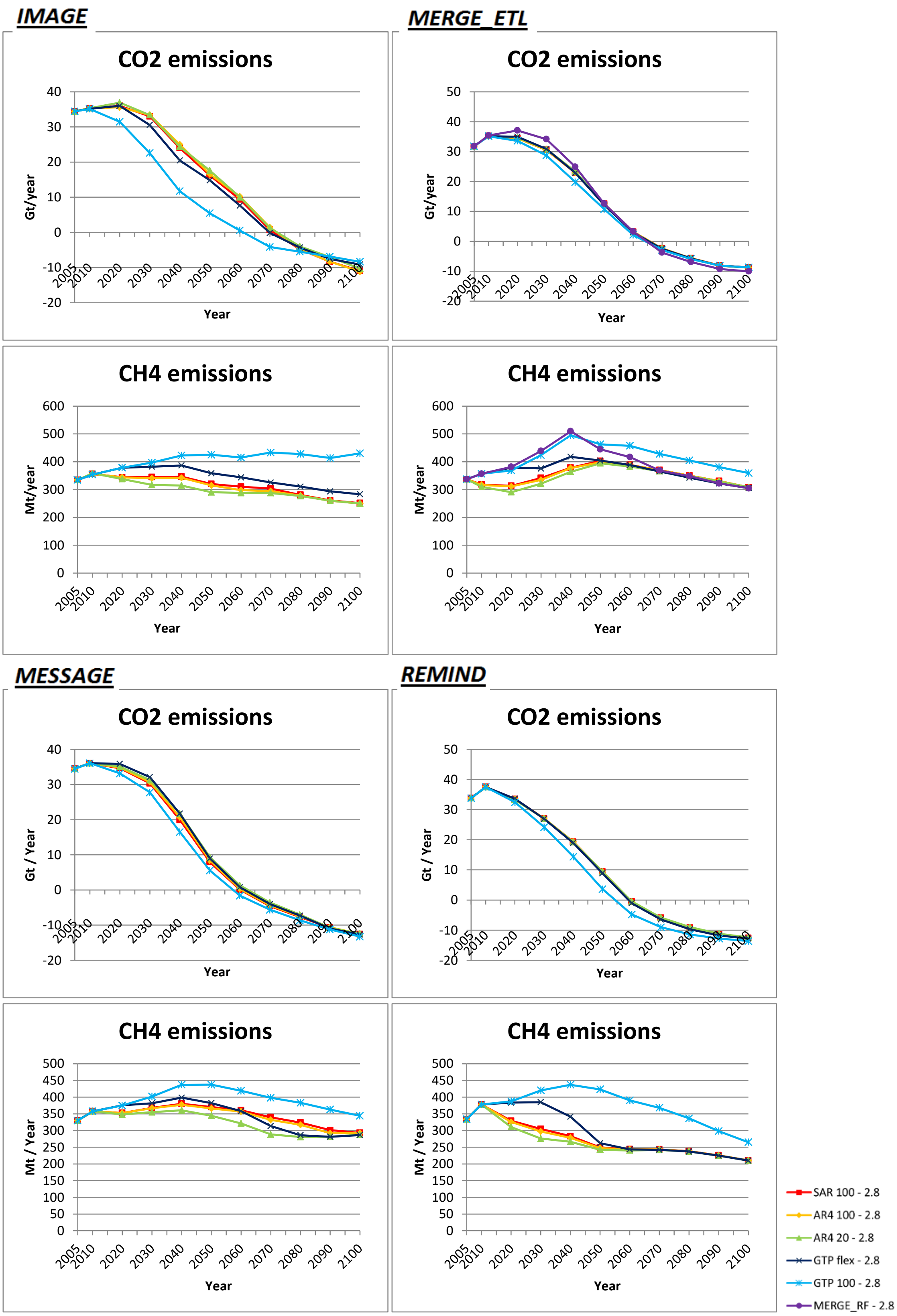
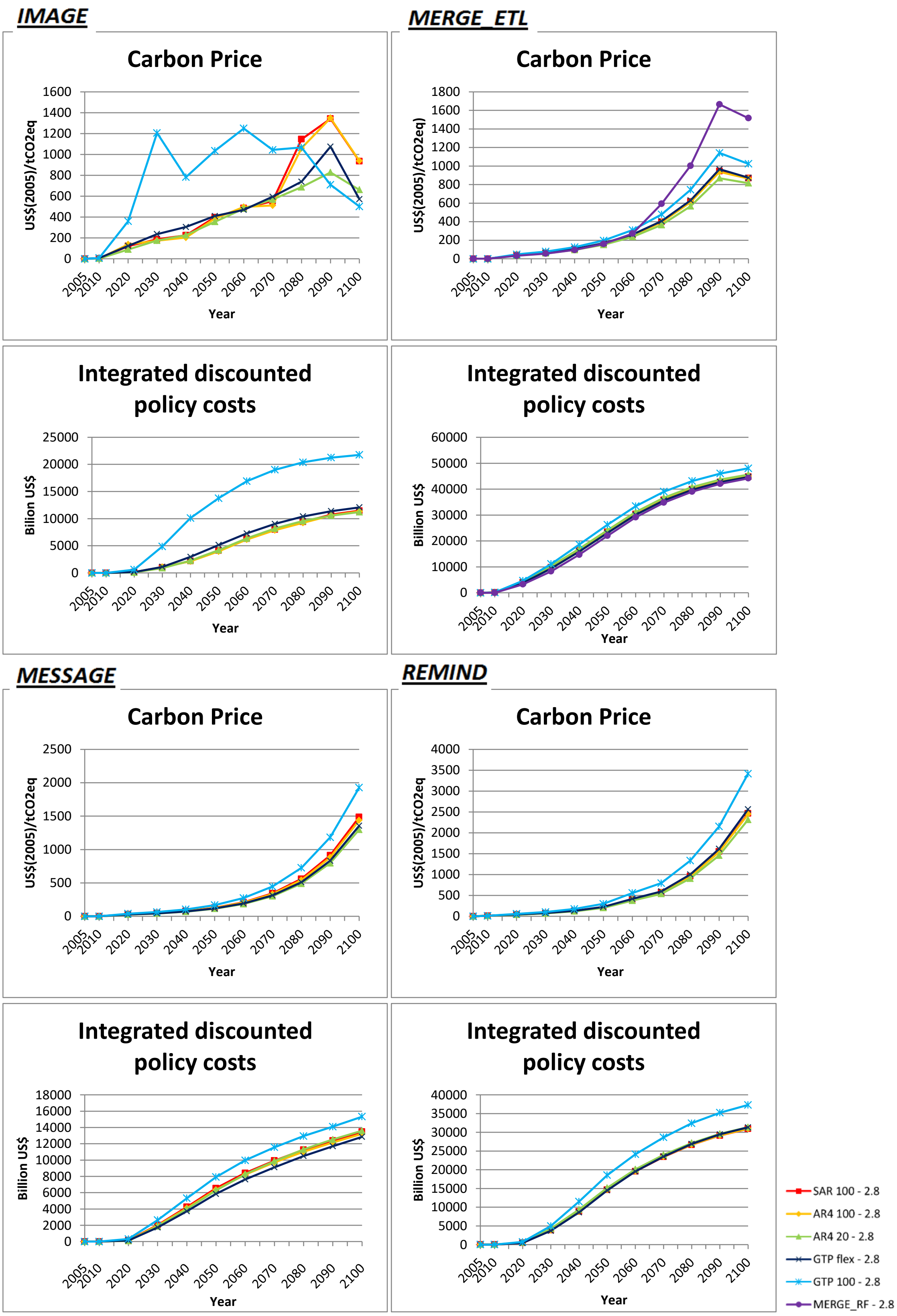


Figure S2: Carbon price and Policy costs in the 2.8 W/m² scenarios (note: different scales on the y-axis)



3) Analysis of the methane marginal abatement cost (MAC) curves in the models

In order to construct the marginal abatement cost (MAC) curves for methane in the integrated assessment models (IAMs) used in this study, we made use of the diagnostic runs that were performed in the AMPERE project (Kriegler et al., 2015). These runs were constructed to determine the response of IAMs to different carbon price trajectories. For this analysis, we included two scenarios with a low and an a high carbon price increase profile (see table S1), and looked at the relative methane emission reductions at different carbon prices in the two scenarios.

Table S1: Two AMPERE diagnostic run scenarios: low and high carbon price increase

	Low carbon price increase	High carbon price increase
Year	*Carbon price (2005$ / tCO2)*	*Carbon price (2005$ / tCO2)*
2005	0	0
2010	0	0
2020	19	74
2030	27	110
2040	41	162
2050	60	240
2060	89	355
2070	131	526
2080	195	779
2090	288	1152
2100	426	1706

The effect of the carbon price on methane abatement is shown in Figure S3 (for IMAGE, MESSAGE and REMIND) and in Figure S4 (for MERGE_ETL). The upper graphs show the relative methane emission reductions compared to the baseline at various carbon prices, while the lower graphs show the relative reductions at the corresponding methane price (this is 24 times the carbon price, since the models made use of AR4 in the diagnostic runs). The lower graph with the methane price is relevant for this study, because the different metrics lead to different CH_4/CO_2 price ratios, which changes the scale of the upper graph. In table S2 the methane prices are given for all the scenarios and all the models (derived from the carbon price and the methane value in the metric). In Figure S3 we have combined the two diagnostic runs, because the reduction at a certain methane price is relatively constant in IMAGE, MESSAGE and REMIND. For MERGE_ETL this is not the case, as the reduction potential is much more time dependent than price induced. This means that high prices in early years do not considerably increase the maximum emission reduction. In Figure S4 it can be seen that this results in very different MAC curves in different scenarios.

The figures explain why REMIND and MERGE_ETL show very similar emission profiles in the last decades across all scenarios except GTP-100. In both models, the maximum methane reduction potential is reached at the end of the century. This is shown by the green vertical line (for REMIND) in Figure S3 and the black vertical line (for MERGE_ETL) in Figure S4 (indicating that the same abatement is reached at different prices), as well as in Table S2 where blue shaded cells show scenarios in which the maximum reduction potential is reached in 2100. In REMIND, this maximum reduction potential is reached about 10 years earlier than in MERGE_ETL, which is illustrated by the earlier convergence of methane emissions in the different scenarios.

MESSAGE, on the other hand, does not reach its full methane reduction potential in 2100 in most scenarios. For that reason, it has a larger spread in scenario outcomes. This is caused by a steadily increasing reduction potential at higher prices, shown by the diagonal line between the 10000 $/t and 41000 $/t methane price. In table S2 this is also indicated by the larger number of scenarios that do not reach their full reduction potential in 2100 (orange shaded cells). Compared to REMIND, MESSAGE generally also has a lower carbon price, which translates into a lower methane price. This therefore slightly adds to the larger variety in scenario outcomes.

The cost-effectiveness of GTP(t) in MESSAGE can also be explained by the possibility to further reduce methane emissions at higher costs at the end of the time period. Because of this, the "potential" of GTP(t) can be fully used until the end of the century.

Table S2 shows that in IMAGE, all scenarios except GTP 100 reach the maximum methane reduction price in 2100 (which is 3750 $/tCH4 or 500 $/tC (Van den Berg et al, 2015)). However, as explained in the main text, the scenarios differ greatly in the amount of abated methane, because of the inertia effect. Without that, the scenarios would also show the same emission trajectories towards the end of the century (similar as in MERGE_ETL and REMIND).

Figure S3: Methane MAC curve profile in IMAGE, MESSAGE and REMIND

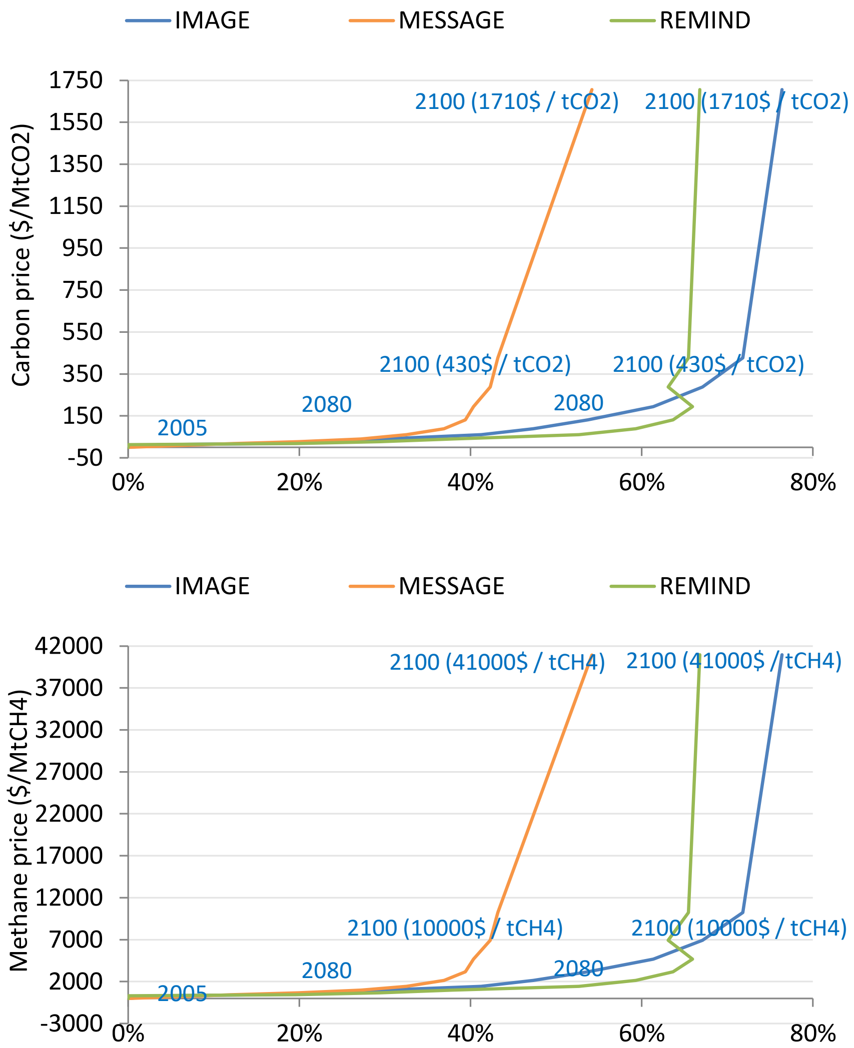
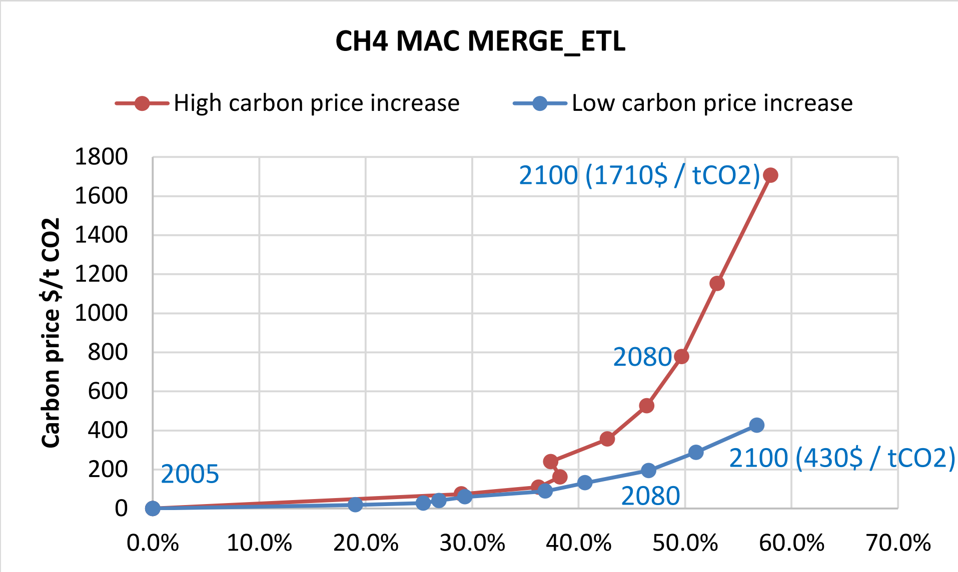


Figure S3: Methane MAC curve profile in MERGE_ETL



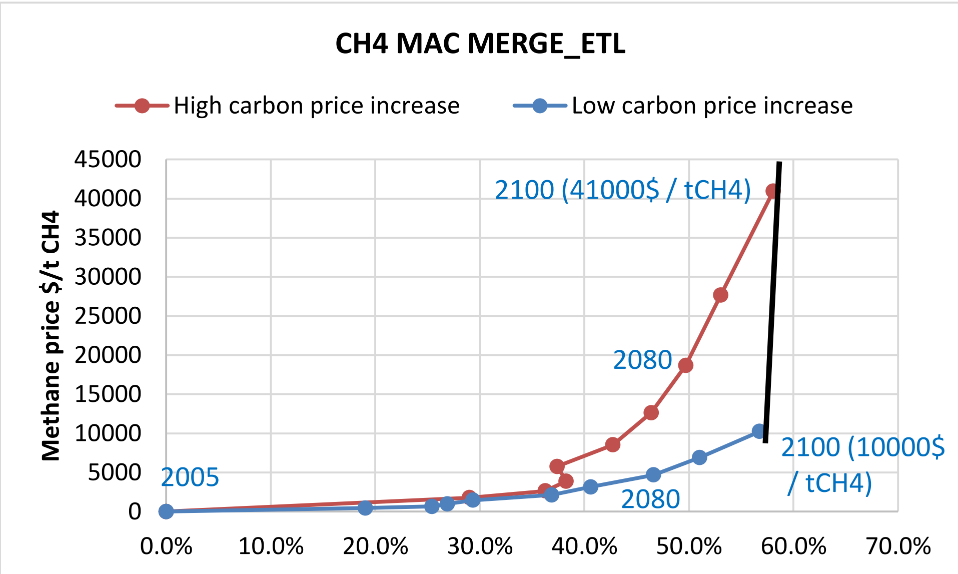


Table S2: Methane prices per scenario in the four IAMs (in 2005$/tCH4). Blue shaded cells in 2100 indicate that the maximum methane price is reached in that year and that no further reductions can be achieved at higher prices. Orange shaded cells indicate that more emission reductions are possible at higher methane prices.

	2005	2010	2020	2030	2040	2050	2060	2070	2080	2090	2100
IMAGE											
SAR 100 - 3.7	0	57	605	1194	2640	3645	4110	5533	10206	13624	12107
AR4 100 - 3.7	0	64	708	1220	2714	4263	5328	6911	10832	13343	10595
AR4 20 - 3.7	0	160	1247	1726	4577	13772	18614	22654	26403	29975	27692
GTP (t) - 3.7	0	6	53	161	781	2209	5547	13796	24261	40658	43264
GTP 100 - 3.7	0	3	16	33	80	95	115	158	219	238	181
SAR 100 - 2.8	0	57	2568	3912	4729	8336	10203	11586	24072	28187	19645
AR4 100 - 2.8	0	64	3330	4172	4865	9250	11922	12274	25431	32375	22528
AR4 20 - 2.8	0	160	6407	12604	16339	25539	34631	40913	49446	59610	47570
GTP (t) - 2.8	0	6	198	780	1981	5075	10468	22601	44561	94916	58189
GTP 100 - 2.8	0	3	144	482	313	414	500	417	427	284	201
MERGE_ETL											
SAR 100 - 3.7	0	0	335	547	882	1393	2178	3347	5079	7718	9034
AR4 100 - 3.7	0	0	380	621	1001	1581	2472	3799	5766	8762	10263
AR4 20 - 3.7	0	0	1067	1741	2808	4438	6930	10641	16148	24535	29123
GTP (t) - 3.7	0	0	26	86	273	823	2312	6050	14586	32488	43497
GTP 100 - 3.7	0	0	7	12	20	31	49	75	114	174	195
MERGE_RF - 3.7	*Low*			*(Not precisely known due to dependence on forcing in 2100)*							*High*
SAR 100 - 2.8	0	0	838	1361	2187	3453	5406	8338	12943	19849	18277
AR4 100 - 2.8	0	0	949	1541	2478	3912	6124	9445	14660	22482	20743
AR4 20 - 2.8	0	0	2644	4295	6905	10897	17049	26287	40798	62557	58689
GTP (t)- 2.8	0	0	65	218	691	2080	5855	15406	37971	85272	88914
GTP 100 - 2.8	0	0	19	31	50	79	124	191	297	456	409
MERGE_RF - 2.8	*Low*			*(Not precisely known due to dependence on forcing in 2100)*							*High*
MESSAGE											
SAR 100 - 3.7	0	0	238	388	632	1029	1676	2730	4447	7244	11800
AR4 100 - 3.7	0	0	264	430	701	1142	1860	3030	4935	8038	13093
AR4 20 - 3.7	0	0	693	1129	1838	2995	4878	7946	12943	21083	34341
GTP (t) - 3.7	0	0	16	54	173	538	1577	4378	11317	26993	50691
GTP 100 - 3.7	0	0	6	9	15	24	39	64	104	169	276
SAR 100 - 2.8	0	0	630	1026	1671	2721	4432	7220	11761	19157	31204
AR4 100 - 2.8	0	0	696	1134	1847	3008	4900	7981	13001	21177	34495
AR4 20 - 2.8	0	0	1885	3070	5002	8147	13270	21616	35211	57354	93424
GTP (t) - 2.8	0	0	44	147	472	1466	4293	11917	30803	73472	137976
GTP 100 - 2.8	0	0	16	25	41	67	109	178	290	473	770
REMIND											
SAR 100 - 3.7	0	231	315	559	971	1654	3062	4366	7354	11911	18927
AR4 100 - 3.7	0	263	359	635	1104	1882	3483	4966	8368	13537	21525
AR4 20 - 3.7	0	790	1024	1814	3151	5373	9951	14193	23918	38696	61539
GTP (t) - 3.7	0	9	24	89	303	986	3281	7969	21296	50479	92469
GTP 100 - 3.7	0	4	9	16	29	49	90	129	216	349	555
SAR 100 - 2.8	0	231	883	1552	2678	4533	8439	12030	20201	32589	51721
AR4 100 - 2.8	0	263	1002	1762	3040	5146	9582	13660	22940	37009	58738
AR4 20 - 2.8	0	790	2842	4992	8617	14595	27171	38756	65088	105018	166679
GTP (t) - 2.8	0	9	70	255	864	2789	9337	22665	60359	142574	260917
GTP 100 - 2.8	0	4	23	41	71	120	223	318	533	861	1366

References

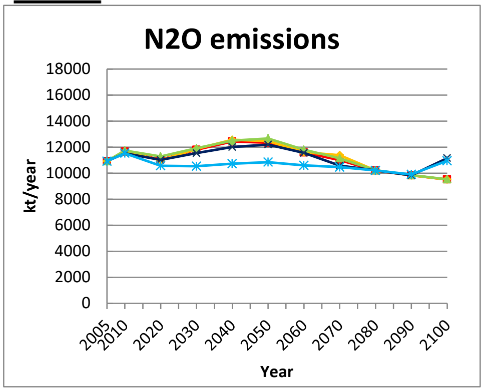
Van den Berg M, Hof AF, van Vliet J, van Vuuren DP (2015) Impact of the choice of emission metric on greenhouse gas abatement and costs. Environmental Research Letters 10:024001.

Kriegler, Elmar, Petermann, Nils, Krey, Volker, Schwanitz, Valeria Jana, Luderer, Gunnar, Ashina, Shuichi, Bosetti, Valentina, Eom, Jiyong, Kitous, Alban, Méjean, Aurélie, Paroussos, Leonidas, Sano, Fuminori, Turton, Hal, Wilson, Charlie, Van Vuuren, Detlef P., (2015), Diagnostic indicators for integrated assessment models of climate policy, Technological Forecasting and Social Change, 90, 45-61

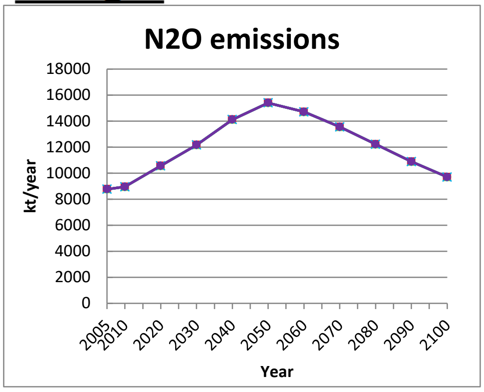
4) N_2O emission trajectories

Figure S5: N_2O emission trajectories (upper panels: 2.8 W/m^2, lower panels: 3.7 W/m^2)

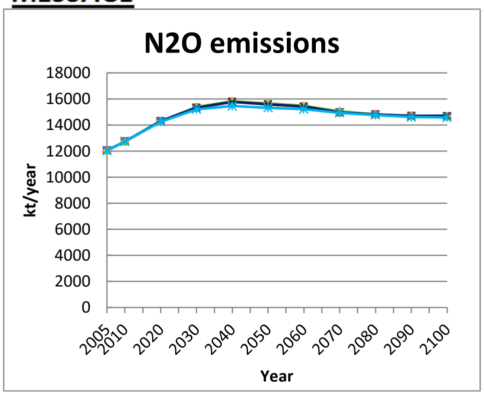
IMAGE



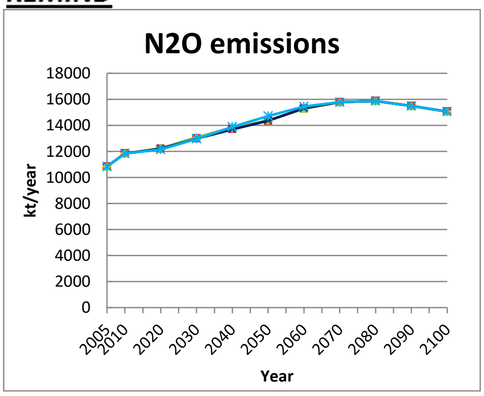
MERGE_ETL

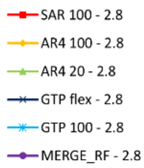


MESSAGE

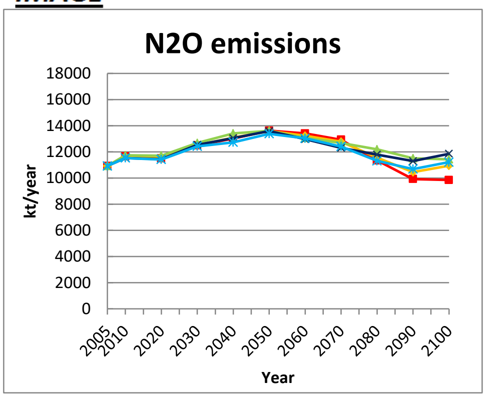


REMIND

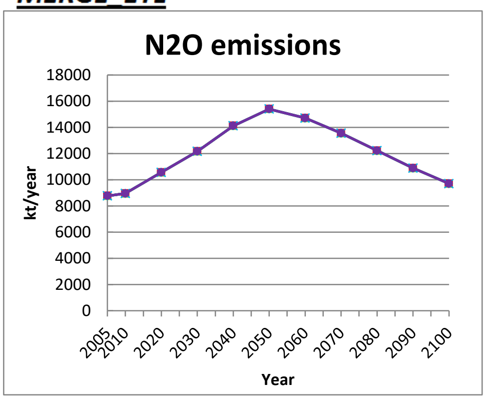




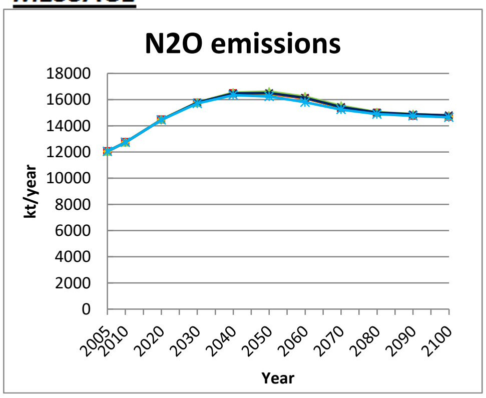
IMAGE



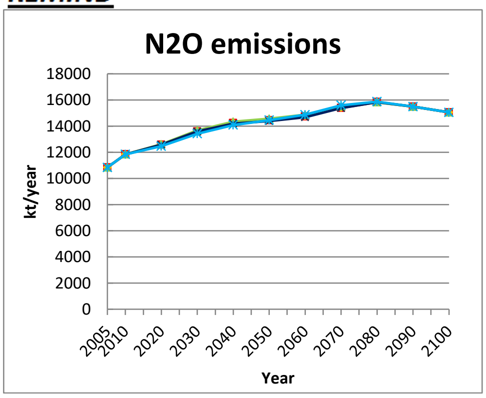
MERGE_ETL

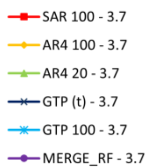


MESSAGE



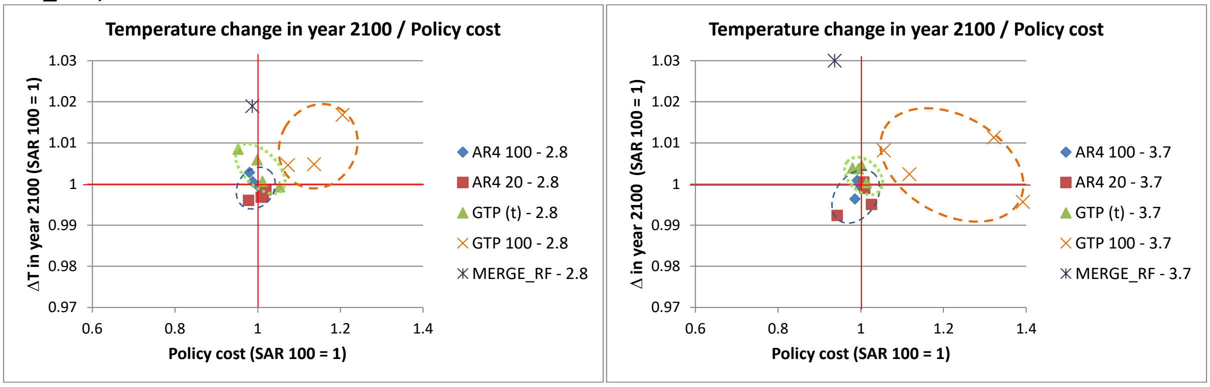
REMIND





5) Integrated discounted policy cost / Temperature change in the year 2100.

Figure S6: Integrated discounted policy cost / Temperature change in the year 2100. Shown for both forcing targets and all models. Normalized by SAR-100 (= 1), indicated by the red lines. GWP metrics encircled by blue lines, GTP (t) by green lines and GTP-100 by purple lines. The GTP-100 2.8 W/m² scenario is not shown for IMAGE on this scale (with a temperature change difference of 2% (in 2100) at a policy cost of 1.9 times that of SAR_100)



References

Den Elzen MGJ, Meinshausen M, Van Vuuren DP (2007) Multi-gas emission envelopes to meet greenhouse gas concentration targets: costs versus certainty of limiting temperature increase. Global Environmental Change 17:260-280.

Deuber O, Luderer G, Edenhofer O (2013) Physico-economic evaluation of climate metrics: A conceptual framework. Environmental Science & Policy 29:37-45.

Ekholm T, Lindroos TJ, Savolainen I (2013) Robustness of climate metrics under climate policy ambiguity. Environmental Science & Policy 31:44-52.

Fuglestvedt JS, Berntsen TK, Godal O, Sausen R, Shine KP, Skodovin T (2003) Metrics of climate change: assessing radiative forcing and emission indices. Climate Change 58:267–331.

IPCC (1990) Climate Change: the Intergovernmental Panel on Climate Change Scientific Assessment. Cambridge University Press, Cambridge.

IPCC (2001) TAR WG1, Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

IPCC (2009) Meeting Report of the Expert Meeting on the Science of Alternative Metrics. Plattner G-K, Stocker TF, Midgley P et al (eds). IPCC WGI Technical Support Unit, Bern, Switzerland, pp75.

Johansson DJA (2011) Economics- and physical-based metrics for comparing greenhouse gases. Climatic Change 110:123-141.

Johansson DJA, Persson UM, Azar C (2006) The Cost of Using Global Warming Potentials: Analysing the Trade off Between CO2, CH4 and N2O. Climatic Change 77:291-309.

Kandlikar M (1995) The relative role of trace gas emissions in greenhouse abatement policies. Energy Policy 23:879-883.

Kriegler E, Riahi K, Bauer N, Valeria Jana Schwanitz, Nils Petermann, Valentina Bosetti, Adriana Marcucci, Otto S, Leonidas Paroussos, Shilpa Rao, Tabaré Arroyo Currás, Shuichi Ashina, Johannes Bollen, Jiyong Eom, Meriem Hamdi-Cherif, Thomas Longden, Alban Kitous, Aurélie Méjean, Fuminori Sano, Schaeffer M (2014) Making or breaking climate targets: The AMPERE study on staged accession scenarios for climate policy. Technological Forecasting and Social Change.

Lucas PL, Van Vuuren DP, Olivier JGJ, Den Elzen MGJ (2007) Long-term reduction potential of non-CO_2 greenhouse gases. Environmental Science & Policy 10:85 - 103.

Manne AS, Richels RG (2001) An alternative approach to establishing trade-offs among greenhouse gases. Nature 410:675-677.

Myhre G, Shindell D, Bréon F-M (2013) Chapter 8: Anthropogenic and Natural Radiative Forcing. In: IPCC (2013) WGI Climate Change: The scientific basis. Cambridge University Press, Cambridge, UK, and New York, USA.

O'Neill BC (2000) The Jury is still out on Global Warming Potentials. Climate Change 44:427–443.

O'Neill BC (2003) Economics, natural science and the costs of Global Warming Potentials. An Editorial Comment. Climate Change 58:251–260.

Reisinger A, Havlik P, Riahi K, Vliet O, Obersteiner M, Herrero M (2013) Implications of alternative metrics for global mitigation costs and greenhouse gas emissions from agriculture. Climatic Change 117:677-690.

Shine K, Berntsen T, Fuglestvedt J, Skeie R, Stuber N (2007) Comparing the climate effect of emissions of short- and long-lived climate agents. Philosophical Transactions of The Royal Society of London Series A: Mathematical, Physical & engineering Sciences 365:1903–1914.

Shine KP (2009) The global warming potential – the need for an interdisciplinary re-trial. An editorial comment. Climate Change 96:467–472.

Shine KS, Fuglestvedt JS, Hailemariam K, Stuber N (2005) Alternatives to the Global Warming Potential for Comparing Climate Impacts of Emission of Greenhouse Gases Climatic Change 68:281–302.

Smith SJ, Karas J, Edmonds J, Eom J, Mizrahi A (2013) Sensitivity of multi-gas climate policy to emission metrics. Climatic Change 117:663-675.
Strefler J, Luderer G, Aboumahboub T, Kriegler E (2014) Economic impacts of alternative greenhouse gas emission metrics: a model-based assessment. Climatic Change 125:319-331.
Tol RSJ (1999) The marginal costs of greenhouse gas emissions. Energy Journal 20:61-81.
Tol RSJ, Berntsen T, O'Neill BC, Fuglestvedt JS, Shine KP (2012) A unifying framework for metrics for aggregating the climate effect of different emissions. Environmental Research Letters 7:044006.
UNFCCC (2011) Greenhouse gases, sectors and source categories, common metrics to calculate the carbon dioxide equivalence of anthropogenic emissions by sources and removals by sinks, and other methodological issues (Durban: United Nations Framework Convention on Climate Change).
van den Berg M, Hof AF, van Vliet J, van Vuuren DP (2015) Impact of the choice of emission metric on greenhouse gas abatement and costs. Environmental Research Letters 10:024001.
Van Vuuren DP, Eickhout B, Lucas PL, den Elzen MGJ (2006a) Long-term multi-gas scenarios to stabilise radiative forcing - Exploring costs and benefits within an integrated assessment framework. Energy Journal 27:201-233.
van Vuuren DP, Weyant J, de la Chesnaye F (2006b) Multi-gas scenarios to stabilize radiative forcing. Energy Economics 28:102-120.
Weyant J, Delachesnaye P, Blanford G (2006) An overview of EMF-21: multigas mitigation and climate change. Energy Journal.

Van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque JF, Masui T, Meinshausen M, Nakicenovic N, Smith SJ, Rose SK (2011) The representative concentration pathways: an overview. Climatic Change 109:5–31.