Anonymous Referee #2
Received and published: 12 October 2012

We thank Anonymous Referee #2 for constructive comments. For simplicity, we present his comments in black and our responses in red.

General comments
I agree with the comments made by Daniel that the paper presents a valuable overview and summary of ways to calculate a broad range of physically-based metrics via analytical expressions. As such, it presents little novel science – the main novelty lies in the few, but valuable, applications that the paper demonstrates. This means its merit should not be judged merely by its the novelty of its scientific findings. I strongly support the publication of this article for the summary and consistent analytical expressions it offers, and the demonstrated applications, even though it offers little new scientific insights.

The focus on analytical expressions is in my view not a shortcoming but simply reflects the chosen focus of the paper. I do agree though with Daniel's comments that the title of the paper should better reflect this focus, and in several places a comment is warranted that the choice does not lie between either simple analytical expressions or GCMs to evaluate the climate impact of emissions of different gases. Simple or reduced complexity climate models have been used to calculate metrics and can do this fairly efficiently, and the paper should (fairly prominently, I suggest) recognise this, before then proceeding to lay out a coherent framework for analytical expressions. I think the focus on analytical expressions is a choice that does perhaps not need to be justified as Daniel argues, but it needs to be made clear that this is a choice, and other options exist to derive results of the same breadth and complexity as the authors present.
Both reviewers find the title somewhat misleading. To improve the title we have, therefore, changed the wording to focus more on simple emission metrics. Hence, the new title is: "Simple emission metrics for climate impacts."

We agree in that the chosen focus of this paper is simple analytical emission metrics based on numerical complex models. You correctly point out that the article was lacking an overview of the different metric approaches and an explanation of why our focus was on analytical metrics. We will, therefore, include a paragraph fairly early in the article (in Section 2) that comments the different metric approaches and that we focus in the paper on the simple analytical metrics:

"While we here focus on simple analytical metrics using simple parameterizations of the climate system, there is a variety of alternative approaches to develop emission metrics (Tanaka et al., 2010). One can conceptualize "climate models" as spanning from simple analytical models to complex general circulation models or earth-system models (IPCC, 2001; Held, 2005). Reisinger et al. (2010); Tanaka et al. (2009); Azar and Johanson (2012) are examples of studies that use reduced complexity carbon cycle and energy balance models. In general, the more complex the models are, the better they are to handle the processes in the climate system, but at the cost of increasing computational time making them unsuitable for most common metric applications. Simple climate models with faster solution times are often used as the basis of emission metrics (Tanaka et al., 2009; Wigley, 1998; Manne and Richels, 2001), but these are difficult to represent in reduced analytical form. We focus in this article on analytical expressions to be able to provide a single consistent and transparent analytical framework that can handle a broad range of metric calculations. Despite the simplicity of these metrics, the key parameters are based on more complex climate models ensuring the metric values are realistic."

Given the focus on analytical expressions, I would strongly urge the authors to make use of this computational efficiency and strive to be more comprehensive in the gases they cover. The value of

this article does not lie in its novel scientific insights but in being able to act as a reference document for the scientific (and applied science) community. To service this potential use and value of the article, I would like to also have results for SF6 and some selected HFCs and PFCs included. Not in all figures and applications, but mainly in the presentation of results that are currently shown in Figure 9. That would mean that the article could act as the go-to place for anybody wishing to obtain consistent metric values for a wide range of gases. Right now, given that the results are shown only in a Figure but not in a table, and that some important GHGs are not included, it fails to do that. Including other gases, including some longlived ones, will also help correct some incorrect generalisations that the authors draw from their results (see specific comments below).
We agree that the analytical expressions presented should cover as many gases as possible. Although not explicitly written in the article, our analysis of climate impacts presented in Section 4 includes 37 different species. We have only flagged those that are most important. The group "synthetic" consists of a range of species. In Figure 9, we will add the species SF6, HFC-134a, HFC-152a, HCFC-22, CFC-11, CFC-12, and PFC-14. At the end of Section 1 we have added a sentence to explain that the metrics are based on parameters from AR4. We would like to be consistent and not give values that are somewhere between AR4 and AR5. If we do so, the values will be outdated as soon as AR5 is published. Your proposal to include a table over the metric values is good; however, such tables already exist in AR4 and in Fuglestvedt et al. (2010). We would prefer to refer to literature (IPCC, 2007; Fuglestvedt et al., 2010), as all values shown here have been presented before in review material. Sentence added at the very end of Section 1:
"The parameters used in this paper are based on IPCC (2007) for consistency, but as new input becomes available, it is straight forward to update the metric values using the material presented in this paper."
In the first paragraph in Section 4, we are also clarifying that the metric values are consistent with AR4:
"All the parameters used in the metric calculations are consistent with IPCC (2007). Although new input exists for some species, we prefer to use values that are consistent with those given by IPCC (2007) and the ATTICA assessment (Fuglestvedt et al., 2010). The parameters used in the metrics presented here can be found in Forster et al. (2007); Fuglestvedt et al. (2010)."
We also clarify in the text where Figure 9 is presented, by including this sentence into the first part of Section 4.1:
"These metrics values are consistent with the parameters from IPCC (2007) and the metric values for some alternative time horizons are available in Forster et al. (2007); Fuglestvedt et al. (2010)."

Finally as general comment, and in keeping with this article being a review and hopefully one-stop shop for information on analytically-derived metrics, I would have liked a place somewhere in the manuscript (at the beginning and/or end of section 2) that summarises the issues canvassed in section 2, and more importantly, an assessment of the relative uncertainty introduced by the different choices and uncertainty in the parameters used. E.g. the authors devote substantial space to discussing the difference between small and marginal perturbations of CO2 (2.2.1), but it then becomes clear that the differences for pulses up to at least 10ppm pulse are almost irrelevant compared to other uncertainties and choices. So a summary of the issues explored that gives a clear ranking of which choices and uncertainties actually matter, and which ones are intellectually interesting but of less relevance in practice, would be important to avoid skewed messages in this article.
We agree that a summary of the issues presented in Section 2 is helpful for the article. Therefore, we will add such a summary at the end of Section 2. We will add one section called "2.7 Consistency issues", which discusses such issues for different metrics. How to treat indirect effects and feedbacks will be discussed. Further, we will add a section called "2.8 Relative uncertainty from parameters and choices". We will first present a summary on uncertainties in the different metric parameters. This summary will be followed by an equation on how to estimate metric uncertainty and also using CH_4 as an example.

The sum of those general comments in my view suggests that a major revision is required, though one could equally regard such revisions as 'extensive but minor' changes since I'm perfectly satisfied with the structure and scientific approach of the article.

Specific comments
(#1)
Most of my specific comments, although there are plenty of them, should be fairly easy to deal with. There is one major specific comment though, which is whether in the applications listed in section 4, the changing background concentrations for sustained emissions has been taken into account, and if not, to provide an estimate of the systematic error introduced into the results.
The analytical expressions for sustained emissions are based on the assumption of constant background concentrations. We agree with the reviewer that this is a critical issue. In the article, we present simple emission metrics as found and applied in the scientific literature. The concept of sustained emissions is not uncommon in the literature, and this concept is applied as in the literature. There are many assumptions in these emission metrics. The sustained emissions case is based on several unrealistic assumptions, such as constant emissions forever and a constant background forever. Despite its weaknesses, this concept is useful to show different perspectives. The magnitude of the step change will determine how large the error is (for example, a sustained emissions of 1kg/yr for 100 years will have negligible effect on the background). Hence, this method will introduce uncertainties in the climate impact due to sustained global emissions, while smaller perturbations will give only marginal errors. For national or sectorial emissions, or emissions for a project, the uncertainties will be smaller. We can't give explicit numbers on how large this error source is, as we haven't compared these simple metrics with complex climate models.

A paragraph in Section 2.2.1 has been improved, which includes a general discussion on changing RE: "Since the background concentration is constantly changing, the RE is technically a function of time (Figure 5). As c_0 increases, the RE decreases and, hence, CO_2 becomes relatively less important. Compared to pre-industrial times, the RE in 2005 is 40% lower and may be 50-100% lower in 2100 depending on the future scenario (Figure 5). Even if emission metrics are based on a constant background concentration, the background is usually different when metric values are updated (Reisinger et al., 2011) leading to a different RE. For a scenario background, the RE will change as a function of time within the metric calculation. In both constant and scenario backgrounds, the changes in concentration and hence RE are partially offset by changes in the IRF as a function of concentration (Caldeira and Kasting, 1993;Reisinger et al., 2011)."

We have also improved a paragraph in Section 3.1:
"Sustained emissions are a specific type of scenario that neglects changes due to economic growth, technology improvements, mitigation policies or the lifecycle of infrastructure. The scientific literature also assumes for simplicity that sustained emissions do not change the background concentrations (as for pulse emissions, they are marginal); hence, all factors influencing the metric calculations stay constant. From a policy perspective, sustained emission may seem more relevant, since in reality, emissions are unlikely to stop instantaneously as in a pulse emission."

We have also included a sentence in Section 4.3:
"The sustained emission scenario assumes constant 2008 emissions into the future, and we further assume consistent with the application of metrics (to compare GHGs) that this does not affect the background concentration."

P872/L19: spell out SLCF in the abstract since readers may not be familiar with the acronym
Added

P873/L5-6: the sentence starting "A limitation of using RF..." is not clear to me and seems incorrect since RF has a simple relationship with atmospheric concentration. Perhaps you meant "transient response in the climate system" rather than in the atmospheric concentration?
We agree that this sentence is confusing. We have, therefore, deleted that sentence and replaced with:
"The RF is often used in literature to compare the RF at two points in time, such as the change in RF between current and pre-industrial times (Forster et al., 2007, Figure 2.20)."

P873/L11-12: you could extend the quote and highlight that it was meant as illustrative example to demonstrate the difficulties of using a single metric!
We have extended the quote to:
"simple approach … to illustrative the difficulties inherent in the concept, to illustrate the importance of some of the current gaps in understanding and to demonstrate the current range of uncertainties"

P873/L15: not clear what is meant by "related to its interpretation". The interpretation of the GWP seems about the only thing that is straightforward and not dependent on value judgements!?
We agree that the GWP is straightforward, but that it can be misunderstood outside of the expert science community and whether GWP is suitable with the objectives of current climate policy. Hence, we have changed the wording to:
"Most critiques focus on the physical interpretation of the GWP in terms of the climate impact and if this is broadly consistent with the objectives of climate policy."

P873/L15: might be useful to state here that the 100-year GWP is use in UNFCCC reporting and accounting under the Kyoto Protocol, and is also used in most LCA analyses as well as industry reporting and carbon accounting tools.
We have added a sentence on P873/L10:
"The GWP with a 100 year time horizon is used for reporting of emissions under the UNFCCC and its Kyoto Protocol, and consequently is applied almost universally in life cycle assessment and other forms of GHG reporting at the national, regional, city, industry, and individual levels (Peters, 2010)."

P873/L16: I'd challenge the words "in use". It's being discussed in the literature and receiving attention, but there is very, very little 'use' of the GTP in any applied GHG reporting and accounting sense. If the authors disagree with this, it would be useful to have some of those actual applications (in a practical, applied policy or industry context) referenced.
We agree that GTP is mostly discussed in the science community and receiving increasingly more attention from outside, but there are still few examples of GTP actually being used. We change "in use" to "applied in the literature." When the first introduce GTP in the paper, we include this sentence:
"GTP is increasingly applied in the literature, but has yet to be broadly applied for emission accounting or policy applications."

P874/L10-12: add "to report emission trends of countries under the UNFCC"
Added the text:
"and to be able to report emission trends of countries under the UNFCCC"

P874/L15: insert "scientific" before "need", since clearly from a policy perspective, it would be rather helpful if there were just one single metric as it would avoid a lot of potential confusion.
Agreed, "scientific" added

P875/L17-19: Not convinced by this statement. One could use SF6 with its simple exponential decay rate, or even N2O, and still end up with very different metric values for CH4 depending on the value judgement of TH.
This sentence was not well written. You are right that there are many issues on time horizon regardless of if CO_2 is the reference gas or not. Our point, which we didn't spell out clearly, was to come back to the origin of GWP. The GWP concept was originally a climate analogue to the ozone depletion potential (ODP), which compares the steady-state ozone depletion for a sustained emission relative to a reference gas. The ODP integrated to infinity (steady state with no discounting). This is not possible with CO_2, as long as you parameterize the IRF for CO_2 to not decay to zero. Hence, a time horizon (less than infinity) was needed because CO_2 became the reference gas. We have improved the text by shortly discussing ODP and also state that choosing CO_2 as a reference gas is a value choice.

Paragraph added in Section 2 to describe ODP and the time horizon issue:
"Together with the climate impact of an emission metric, the time-horizon (or more generally the discount function) is one of the key value based choices in metric design. The Ozone Depletion Potential, which serves a similar purpose to the GWP for ozone depletion (IPCC, 1990), did not use any discounting in effecting choosing an infinite time horizon (Cox and Wuebbles, 1989). In contrast, the GWP requires a discount function for CO_2 to ensure the metric values are finite, and this is because it is generally assumed that a pulse emission of CO_2 does not decay to zero (Lashof and Ahuja, 1990;Archer et al., 2009)."

Sentence added to Section 2.6 describing that the choice of reference gas is always a value choice:
"The choice of reference gas is a value based choice, but an obvious choice is to use the trace gas of primary concern, namely carbon dioxide (IPCC, 1990). There is no obvious need to have only one reference gas or let CO_2 always be the reference gas."

P875/L20: a significant omission in this article is the recognition that not only one could use different metrics in different applications, but also that there is no reason to have only one reference gas. Two-basket approaches have been discussed in the literature and should be briefly flagged here as one way of dealing with widely differing lifetimes.
E.g. Smith, S.M., J.A. Lowe, N.H.A. Bowerman, L.K. Gohar, C. Huntingford, M.R. Allen, 2012: Equivalence of greenhouse-gas emissions for peak temperature limits. Nature Clim. Change. 2(7), 535-538.
Agreed. We add the following text in Section 2.6, as we move all the discussion of normalized metrics to Section 2.6:
"There is no obvious need to have only one reference gas or let CO_2 always be the reference gas. A two-basket or multi-basket approach to climate policy could be used to treat species with different lifetimes differently, and each basket may have a different reference gas (e.g. Smith et al., 2012; Daniel et al, 2012)."

P876/L13-22: Here would be an appropriate place to state very clearly that a variety of approaches exist and have been used in the literature to calculate metrics, ranging from analytical expressions based on parameterisations to simple and reduced-complexity models, to models of intermediate complexity. Give a very brief overview of the advantages and disadvantages of those approaches, and then state that in this article you focus on analytical expressions to be able to provide a single consistent numerical framework that allows the calculation of any metric as long as some key parameters are known.
As written previously, we agree that the paper should include a short presentation of different methods in metric calculations and an explanation that the analytical metrics is the focus here. We have included this paragraph before the last paragraph of the introduction in Section 2:

“While we here focus on simple analytical metrics using simple parameterizations of the climate system, there is a variety of alternative approaches to develop emission metrics (Tanaka et al., 2010). One can conceptualize “climate models” as spanning from simple analytical models to complex general circulation models or earth-system models (IPCC, 2001; Held, 2005). Reisinger et al. (2010); Tanaka et al. (2009); and Azar and Johanson (2012) are examples of studies that use reduced complexity carbon cycle and energy balance models. In general, the more complex the models are, the better they are to handle the processes in the climate system, but at the cost of increasing computational time making them unsuitable for common most metric applications. Simple climate models with shorter computational times are often used as the basis of emission metrics (Tanaka et al., 2009; Wigley, 1998; Manne and Richels, 2001), but these are difficult to represent in reduced analytical form. We focus in this article on analytical expressions to be able to provide a single consistent and transparent analytical framework that can handle a broad range of metric calculations. Despite the simplicity of these metrics, the key parameters are based on more complex climate models ensuring the metric values are realistic.”

P877/L20: This reads as if there was somehow a definition or otherwise requirement to use the Bern CC model to calculate IRFs for emission metrics. Make clear that in practice this has been a common choice, but note that others have used other models or even a range of CC models (e.g. Reisinger et al 2010, cited elsewhere in the article).
We agree that the formulation here was misleading. We change the text so that we write:
“The IRF for CO_2 that is used in emission metrics in the literature is usually based on the Bern carbon cycle model (Joos et al., 2001) with the IRF experimental setup described by Enting et al. (1994), also see Figure 1 in Joos et al. (2012). Other studies have used other models or a range of carbon cycle models (e.g. Reisinger et al., 2010; Gillett and Matthews, 2010).”

P878/L11: not clear from the discussion that follows that, or rather which aspects of, experimental setup have a large influence on the IRF. Perhaps replace by “some possible choices in the experimental setup”
We have added “some choices in the experimental setup”

P878/L27: The difference obviously depends on the scenario chosen. Can you indicate over which scenario range this holds true? E.g. if one assumes RCP8.5 scenario and extended to 2300, is it still only 0.1-0.2 over 500 years?
How the IRF evolves depending on the scenario chosen is an important issue. Some of the authors of this article are actually as we speak writing a paper on this very issue. We cite Figure 5.5 in IPCC (1994), which clearly states the scenarios in which the statement is valid for. We have changed the text to:
“…can lead to a difference in the IRF of 20% after 500 years…”

P878/L28: A very large pulse will also change the radiative efficiency of that pulse, which should be mentioned here as well
Pulse sizes are large to to get a good signal-to-noise ratio. However, metrics are used on the assumption that the emissions are marginal. We reference Joos et al. (2012), which has a longer discussion on this. The sentence has been improved to:
“Different pulse sizes also lead to different IRFs (Archer et al., 2009), but for use in metrics these are normalized to 1kg to represent a marginal perturbation (Joos et al., 2012).”

P880/L7-8: Here is another point where the text should make clear that using an IRF to derive temperature responses is a choice, not something in the definition of emission metrics. One could equally use climate models to derive the temperature response (even in combination with a parameterised IRF for concentrations).
We agree. We have altered the text to:

“For emission metrics that link from RF to temperature, an IRF can be used for the temperature response to an instantaneous unit pulse of RF, IRF_T (Shine et al., 2005; Boucher and Reddy, 2008). Alternatively, climate models can also derive the temperature response (e.g. Reisinger et al., 2010).”

P884/L5-15: Given the very small difference between a 1ppm and a marginal perturbation, I question the relevance of this paragraph. It might be more useful to derive a plausible threshold for when a pulse size is no longer appropriate and consistent with the definition of the metrics under discussion. It seems that a pulse up to 10ppm is defensible and introduces an effort in the order of only 1% - but what about pulses of 100ppm or even more? That would be useful guidance for future experiments, such as that in Joos et al (2012).
We agree with the reviewer that using the average or marginal method introduces negligible uncertainty; however, we find the intellectual discussion important. We argue that this difference in most cases has little practical relevance; however, we would like to keep some sentences on this issue. As a response to this comment, we have included error numbers for 100 ppm in Table 2. We have deleted the paragraph and replaced with a much shorter one:
“In the Fourth Assessment Report, *ΔC* is taken as 1ppm (Forster et al., 2007), while in the Third Assessment Report it is taken as the magnitude of the CO_2 pulse (IPCC, 2001; WMO, 1999). For small perturbations, the difference between approaches is negligible in comparison to the uncertainty in RF (see Table 2).”

P884/L16-27: With Daniel, I struggle with the relevance of this. The main argument against using pre-industrial IRFs and RE surely must be that the values are counterfactual to their actual climate impact. This discussion needs to be more nuanced and extended, or perhaps shortened and focus purely on the policy trade-off between a constant metric value, versus a metric value that changes over time but in a predictable manner, versus a metric value that changes over time but not in a predictable manner if it depends on future emissions and concentrations. It would also help to have an estimate of the amount of change over time, and compare this with other uncertainties and importance of other choices. This goes back to my general comment that Section 2 needs a summary that helps rank alternative choices with regard to the actual difference they make for the metric values.

You are correct in stating that pre-industrial values are not really of need today since the atmospheric composition is quite different today. We are, therefore, shortened the discussion on pre-industrial conditions. The paragraph mainly discusses general issues in changing RE and how to select a RE. The policy trade-off between different metrics is given indirectly by this improved paragraph. We keep Figure 5, which shows how much changing the CO2 concentration will affect the RE in different RCPs. The paragraph is changed to:

“Compared to pre-industrial times, the RE in 2005 is 40% lower and may be 50-100% lower in 2100 depending on the future scenario (Figure 5). Even if emission metrics are based on a constant background concentration, the background is usually different when metric values are updated (Reisinger et al., 2011) leading to a different RE. For a scenario background, the RE will change as a function of time within the metric calculation. In both constant and scenario backgrounds, the changes in concentration and hence RE are partially offset by changes in the IRF as a function of concentration (Caldeira and Kasting, 1993;Reisinger et al., 2011). For impact assessment, it can be argued to base the RE on a pre-determined fixed concentration such as pre-industrial concentrations (e.g., Huijbregts et al., 2011). This would ensure that the metric values only change due to updated scientific information, but would mean that the relative weights of GHGs are based on pre-industrial conditions.”

P885/L20: Would be useful to state here the range of uncertainties for REs, to help compare and contrast those with the impact of other uncertainties and experimental set-up choices.
We have added one sentence on RE uncertainty for LLGHGs:

"The uncertainty in the RE estimates for the LLGHGs is about 10% (Forster et al., 2007)."

P886/L8-17: It would be useful to emphasise here the order of magnitude greater uncertainties in calculating these simple parameters, and the reliance on much more complex models to derive those parameters. Also, it needs to be made clear up-front that deriving a single global exponential decay term for a species that is not globally well-mixed is a much more gross simplification than using similar exponential decay terms of LLGHGs, which puts a significant question mark behind the physical interpretation and uncertainty of parameters such as tau. This is another example why a summary comparison of the scale of uncertainties from the different issues considered in section 2 is important to avoid skewed messages.
We agree that the uncertainties for SLCFs are much larger than for LLGHGs. One global value with a single exponential decay is a large simplification. In the beginning of Section 2.2.3, we state that RE for SLCFFs are based on chemical transport models; hence, complex models have been used in calculating the simple parameters. We have improved Section 2.2.3 by including two sentences concerning extra uncertainties for SLCFs. Regional variability for the SLCFs is also discussed in Section 2.3. We also add references to Fry et al. (2012) and Collins et al. (2012, ACPD), which both presents new results on regional metrics. The added text is:
"When we assume global average values for the SLCFs, even though there are large regional variations, this simplification will lead to different REs and lifetimes (see Section 2.3)."

P888/L4: It would be helpful to quantify the important of those interactions with other gases on overall radiative forcing. Also you might wish to point out that given the complexity of interactions, there is a largely arbitrary cut-off point about which ones are included in the calculation of feedbacks and their contribution to the overall RF and temperature response. This even applies to the RF from CH4 emissions, since the interaction with aerosols can significantly add to the net RF caused by CH4 emissions
(Shindell, D.T., G. Faluvegi, D.M. Koch, G.A. Schmidt, N. Unger, S.E. Bauer, 2009: Improved Attribution of Climate Forcing to Emissions. Science. 326(5953), 716-718).
We think this comment is meant for P887/L4. Please correct us, if this assumption is wrong.
We agree that there are a number of interactions and that the cut-off point of which of those interactions is included can be arbitrary. We have improved the paragraph by including some more discussion on the indirect effects and made a comment on the cut-off point. We will include a comment that VOC, NOx, and CO have only indirect effects, while we will give quantify the ratio between direct and indirect effects for another gas. The paragraph is edited to:
"*Chemical Reactions*. Emissions of chemically active species can cause changes in concentrations of other species which can have radiative effects. For CH_4, the indirect effect is accounted for by an adjustment factor of 1.4 (Forster et al., 2007). For the ozone-precursors, the radiative forcing is exclusively due to chemical reactions. Shindell et al. (2009) modeled that including interactions with aerosols increases the best estimate of the GWP 100-year value for CH4 and CO and decreases the value for NOx. Given the complexity of interactions, there is a largely arbitrary cut-off point about which interactions are included in the calculation of feedbacks and contribution to RF. Most relevant in this context are emissions linked to tropospheric ozone formation or destruction, enhancement of stratospheric water vapor, changes in concentrations of the OH radical, and secondary aerosol formation. Most indirect effects change the atmospheric residence time of the species (see Section 2.1): e.g., CH_4, the ozone-precursors (NO_x, CO, NMVOC), and halocarbons. NO_x also has indirect effects on clouds, see below. "

P888/L8: state whether the impact depends on the location of emissions.
We think this comment is meant for P887/L8. Please correct us, if this assumption is wrong. Added: "depending on the location of emissions"

P888/L17: insert "is" after "contrails"

We think this comment is meant for P887/L17. Please correct us, if this assumption is wrong. Added "is"

P895/L5: include reference to Azar, C., D.J.A. Johansson, 2012: On the relationship between metrics to compare greenhouse gases – the case of IGTP, GWP and SGTP. Earth Syst. Dynam. Discuss. 3(1), 113-141.
Added

P895/L9: here also you should cite Azar and Johansson 2012 (cited above)
Added

P895/L17: one can argue about the "arguably better": it is a useful alternative interpretation, but there is no reason given why this should be universally better.
We have altered the text to:
"Given these quantitative relationships, it can be argued to interpret the AGWP as iAGTP."

P897/L11-16: As elsewhere above, I would like to challenge the word "use" here: most other metrics are not really used; they are presented and discussed in the scientific literature, but they are hardly ever used in an applied sense (e.g. emissions tracking, reporting, accounting by policy or industry; any use there seems to occur only in a context that is still primarily a scientific one).
OK, we have change the text so we don't say "used," but "discussed in the scientific literature"," used in literature", or "applied in the literature."

P898/L8: not sure I understand the expression "in a normalized gas" – what is that?
Here we mean the normalization of the metric, as in going from iAGTP to iGTP. This section has been removed.

P898/L20: not just for SLCFs but also for short-lived LLGHGs, most notably CH4, but also a range of HFCs. That should be made clear.
You are right; hence, we added "and shorter-lived LLGHGs" in the sentence.

P898/L22-23: I don't understand what the authors mean by "characteristic of moving towards a target, not necessarily a characteristic of the economic model". Can you reformulate? It clearly is a property of economic models that aim to achieve cost-effective abatement.
We have improved the sentence:
"Since this property occurs in the purely physical based metric, it may suggest that this property is a results of moving towards a target and not a consequence of including an economic model in the metric."

P899/L17-25: the authors seem to have an inconsistent use of SLCFs: the discussion here is highly relevant and appropriate to CH4 and short-lived HFCs, not just SLCFs. Or are those meant to be included here? Please spell this out.
Yes, LLGHGs such as CH4 are meant to be included. Since this has been expressed earlier in the paper, we have removed this section.

P899/L21-22: sentence repeated from P898/L22-23. Delete here I suggest.
Deleted

P899/L24-25: based on Peters et al (2011a), the iGTP isn't really a new metric but in fact was there before the GWP, even though not given this name. Also cite Azar and Johansson here.
This section is now removed.

P99/L1-11: Here would be another place where approaches using multiple baskets and using multiple reference gases would be highly appropriate.
We think this comment is meant for P900/L1-11. Please correct us, if this assumption is wrong. We agree that a short comment on multiple baskets and multiple reference gases is suitable here. The first part of the paragraph has been improved to:
"The normalized metric is dependent on the absolute metric of CO_2, since the absolute metric of CO_2 is the denominator. In a multiple baskets approach, several different reference gases could be used. We show the importance of the denominator, here CO_2, in the case of CH_4 for GWP in…"

P901/L10-11: I don't think it's a value judgement but rather it depends on context. If you only have a single emissions pulse to contemplate in a certain context then it's not a value judgement whether you use a pulse or sustained emissions metric.
We clarify that this choice becomes a value choices when the emissions consist of a variety of gases. Sentence is altered to:
"The choice between a pulse and sustained emission scenario for a mix of species is an important value judgment as they place very different weights on SLCFs and LLGHGs."

P902/L13-20: An important point that seems to be missing in this discussion, and in the analytical expressions derived, is the fact that a sustained emission can change the background concentration and hence the RE and possibly IRF of subsequent pulse emissions. The equations derived seem to be correct strictly only for sustained emissions that are small enough that they don't change the background concentration – which they clearly do, however!
Please see Specific comments (#1) for our response on how to parameterize metrics for sustained emissions. These equations hold as long as the system is linear and that the emissions are so small that they don't change the background concentration. You are right that an error is introduced if the background concentration is changed. However, we present metrics for sustained emissions as used in scientific literature. The emissions can also be seen as marginal; hence, they will have only marginal impact on the background concentration. In that context, the formulations presented in Section 3.2 are correct.

P903/L11-14: This seems to be somewhat irrelevant numerical detail
We have shortened the text and removed the most trivial details:
"The convolution can be estimated by numerical integration, though, most numerical integrations have problems with species with a short lifetime (e.g. BC), typically when the time step is larger than the residence time ($\Delta t > \tau$). This problem can be solved by reducing the time step."

P904/L5: Here again, recognition needs to be made of changing background concentrations. MAJOR: This is an important point: in the calculations that the authors show in the next section, did they assume a constant IRF and RE for CO2 and other gases (which would make the results strictly incorrect, and would require them to at least give an estimate of the error introduced by this approach; note that the error is not random but systematic for each gas, and hence it is rather significant to make this issue clear and if necessary quantify it).
Please see Specific comments (#1) for our response on how to parameterize metrics for sustained emissions. The calculations shown in the next section are based on the equations presented in this paper. Hence, we assume for simplicity a constant world. Both IRF and RE for the different species stay constant. Since we haven't compared with complex climate models, it is not possible to give exact error estimates of assuming a constant world.

P904/L8: revise wording: this reads as if climate metrics are "by definition" based on simple parameterisations. But that is not the case, as one could equally (and perhaps, more justifiably) use simple or even complex climate models to calculate metrics. Reword: "Since in the framework we present in this article, all climate metrics are..."

We have improved the text, so that we are not saying that the presented method is the only method: "The emission metrics presented in this article are based on simple parameterizations of more complex models …"

P904/L27: perhaps cross-ref back here to section 2.2.3
Cross-reference included

P905/L1: for readability and since the use of this acronym occurred a while ago, spell out AIE here again
Aerosol indirect effect is written out

P905/L18-P906/9 and Figure 9: Following from my general comment, I would strongly encourage the authors to make significant changes to the results presented. This overview is potentially extremely useful, but as presented at present it is impossible to obtain concrete numbers, and the list of species covered is highly incomplete in that is misses key LLGHGs. I would encourage the authors to substitute, or add to, Figure 9 a table that lists metric values for the standard time horizons (say, 20, 50, 100 and 500 years), and give values also for SF6 and some of the most important HFCs and PFCs. Providing this information in a numerical and comprehensive form would make this article much more useful. I imagine that this should not be an impossible amount of work given that the key parameters for those gases are readily available in the sources already used for this study. I'm not sure if there are page limitations for ESD, but I would argue that there would be a lot of value in providing rather extensive tabulated results.
Please see earlier comment (your general comment) and response. In that response, we argue that such a table is not needed, as we present numbers that are already been published by review article (Fuglestvedt et al., 2010) or IPCC (Forster et al., 2007). The results we present in Section 4 are based on 37 species. For presentation purposes, we have only flagged the most relevant species. We will added a few more species to Figure 9.

P905/L19-22: This strikes me as a tautology, since by definition an integration of something remembers what happened before the end of the integration period. As worded, it gives the sense that this is somehow wrong, whereas it is purely a matter of choice: do you want an integrating metric, or not?
You are right that this sentence is self evident. The sentence was included to be sure that those not being experts in metrics understand that GWP/iGTP are integrated metrics, which results in a behavior quite different from GTP. Hence, we would like to an improved version of this sentence: "Since both GWP and iGTP integrate the effects over time, both these metrics have full "memory" of the climate effects that occurred at previous times, while the GTP does not have full "memory" as…"

P905/L21: this statement seems to causally mix two things: replace "as" with "and", since the decaying RF and the energy being radiated back into space are two different processes. The latter is only of relevance to the GTP, whereas the former matters for both GWP and GTP.
We have clarified the sentence:
"Since both GWP and iGTP integrate the effects over time, both these metrics have full "memory" of the climate effects that occurred at previous times, while the GTP does not have full "memory" as the RF eventually decays to zero because the and energy is eventually radiated back to space."

P905/L27-28: This sentence needs correcting, since it is only true for the limited gases covered at present, and for the limited 100-year time horizon. Certainly not only for N2O but also for other long-lived GHGs such as SF6 and some PFCs, their value increases over several centuries.
We agree that N_2O is not the only species showing that characteristics. Hence, we have improved the sentences to:

"Almost all species become less important with time relative to CO_2, with the exception of N_2O and other LLGHGs with similar or longer lifetimes."

P906/L1-3: With Daniel, I can't make sense of the claim that the climate impact is governed by species with strong/short-lived and weak/long-lived impacts, since species with strong/long-lived impacts clearly matter even more! It becomes a somewhat empty statement though, since it's obvious that species with weak/short-lived impacts don't matter! Delete?
You are right that this sentence is misplaced. We have moved the sentence to Section 4.3, as this sentence refers to Figure 11. What we meant was that the emission weighted metrics typically have this behavior.

P906/L8-9: Again, you seem to use LLGHGs inconsistently: very clearly, for short-lived LLGHGs such as CH4, and some HFCs, the metric values certainly aren't constant throughout the period. This needs to be revised.
As for previous comments, we split between short-lived (like CH_4) and other LLGHGs. Sentences altered to:
"As the target year is approached, the metric values of species with shorter lifetimes then the mean lifetime of CO_2 increase. For the other LLGHGs, the metric values are rather constant throughout the period."

P906/L10: perhaps insert "emissions from" to make clear what basis the ranking has.
We have improved the title to:
"Ranking of countries by total emissions using different metrics"

P906/L11 and Figure 11: I would find Figure 11 more helpful if it used stacked bars, as it would more easily show (in my view) how the total effects add up. But that is a matter of judgement.
How to make the plots is a matter of personal taste. You are right that stacked bar could be as fine, but we (or a majority of us) decided to go for the existing figure. A problem with a stacked bar is that all the small contributions are less visible.

P906/P17-26: Might be worth stating that which countries are the top 10 emitters is also almost independent of metric – only France gets pushed into the top if GWP20 is used (presumably because of their relatively large fraction of CH4 emissions). In summary, it seems that metrics affect the ranking of countries less than one might have assumed – this is actually quite an important conclusion that might be worth even bringing into the abstract?
We have added a sentence:
"The top ten emitters are almost independent of the metric, with the only change being the inclusion of France with a GWP20."
We have also added this sentence to the abstract:
"Further, the ranking of countries by emissions changes very little with different metrics despite large differences in metric values."

P907/L1: You need to specify which metric you use to determine the top 10 emitters since under GWP20, France would be in that list.
The top ten emitters are given according to GWP100. The sentence has been altered to:
"Table 4 shows the relative share of CH_4 in the total emissions using different emission metrics for the top ten emitters (as for GWP100)."

P907/L10: Explain why/how China can have more than 100% of its emissions from CH4. I assume this is because of the negative emissions from SLCFs outweighing the emission of CO2, leaving CH4 to be effectively the only net GHG emission. Make explicit?

You are right, since warming and cooling effects cancel each other out. For GWP20, the SLCFs are comparable more important relative to CO2 and other LLGHGs. In the case of China and GWP20, all other effects than CH4 cancel each other out. Thus, CH_4 seems to stand for approximately 100%. There is a comment of this in the caption of Table 4. We also add this sentence in the article text:

"The contribution of CH_4 is largest for GWP20 due to the shorter perturbation lifetime of CH4, and for GWP20 the contribution of CH4 can be greater than 100% due to the presence of SLCFs with cooling effects."

P907/L11: The heading would make more sense to me if it said "ranking of gases and emissions from sectors by metric and delta T"
We agree that the title could be improved. We have changed the title to:
"Application of metrics to sectorial and regional emissions"

P907/L13: Spell out that you are assuming constant 2008 emissions into the future for the sustained emissions scenario. It's almost clear, but might help to say this explicitly.
We have added this sentence:
"The sustained emission scenario assumes constant 2008 emissions into the future."

P907/L12-28: MAJOR Following up on earlier comment: please state clearly whether you included changing background concentrations resulting from sustained emissions in your calculations, and if you haven't, provide an estimate of the error introduced by this.
Please see Specific comments (#1) for our response on how to parameterize metrics for sustained emissions. The calculations are based on the analytical expressions presented, which assume for simplicity constant background concentrations. We are not able to give good quantative numbers on the error introduced by this simplification.

P908/L17: this summary provides a false advertisement since at present, you have included only a very limited subset of GHGs with simple exponential decay; as suggested in other comments, I would strongly encourage you to make this a more comprehensive analysis and provide data for a more complete set of gases (and time horizon out to 500 years)
Please see earlier comment and response. Although not explicitly written in the article, our analysis of climate impacts includes 37 different species. We have only flagged those that are most important. The group "synthetic" consists of a range of species. The actual metric values can be found in Forster et al. (2007) and Fuglestvedt et al. (2010).

P908/L23-25: Rather than saying you showed their utility, you might want to say that you showed examples of where metrics matter in policy-relevant applications such as the ranking of emissions from countries, sectors, and different gases, but also where they don't actually matter all that much. In my experience there is a lot of unsubstantiated assumption in some policy circles that different metrics would radically change our view of the mitigation challenge (which could be a good or a bad thing, but that's another matter), but your analysis, along with others, seems to demonstrate that while metrics matter in some important details and aspects, they don't matter in some others. Spelling this out in some detail here would make the conclusions you present far more relevant to a wide audience. You may also wish to consider whether some more of the key results you obtain are worthwhile bringing into the abstract.
We have changed the wording from "utility" to "examples" and have altered and added some more text in line with your comment in the conclusion:
"Finally, we applied the metrics in a variety of different applications to show the importance of metrics and the related choices in policy-relevant applications, such as ranking of emissions from countries, sectors, and different species."

"The ranking of the top ten countries by emissions varies little with different metrics."

P908/L23-25: the statement creates a false dichotomy in that it implies that there is a choice between either using metrics calculated based on analytical expressions, or using GCMs without any metrics. But there are many other ways of addressing the question of how different emissions compare, namely using simple or reduced complexity climate models, either directly or to compute metrics and then use those metrics – or one could use GCMs to calculate metrics and then use those metrics for further applications.

We have altered the sentence to:

"We have focused on simple reduced form emission metrics based on simple analytical expressions with parameters derived from more complex models. There are a range of alternative approaches to develop emission metrics that include more detailed representations of the climate or economic system."