The paper provides an interesting analysis on the connection between GWP timescale and discounting rates. However, I think that before publication, some major issues need to be addressed.

We would like to thank the referee for their comments. For responses comment by comment, see below. All author replies are in red. There is also a summary of new sensitivity analyses that is included at the end of the reply to William Collins.

First, while the paper acknowledges a lot of recent articles that discuss GWP, it does not adequately discuss recent articles that look at climate metrics in an economic frame- work (such as Tol et al. 2012 and Mallapragada and Mignone 2017). There needs to be more incorporation of these types of studies to show how this study builds on the existing literature.

We propose to move material that was originally in the SI into the main text, as well as incorporating more references such as the ones you suggest. See the following proposed text addition:

> While we argue that quantitative justifications for timescales within the GWP are rare, as reflected by the judgment of the IPCC authors that no scientific arguments exist for selecting given timescales, there is a rich literature addressing many aspects of climate metrics. Deuber et al. (2013) presents a conceptual framework for evaluating climate metrics, laying out the different choices involved in choosing the measure of impact of radiative forcing, temperature, or damages, and temporal weighting functions that can be integrative (whether discounted or time-horizon based) or based on single future time points. Deuber et al. conclude that the Global Damage Potential (GDP) could be considered a "first-best benchmark metric", but recognize that the time-horizon based GWP has advantages based on limited value-based judgments to a choice of time horizon, reducing scientific uncertainty by limiting the calculations of atmospheric effects to radiative forcing, and eliminating scenario uncertainty by assuming constant background concentrations. Mallapragada and Mignone (2017) present a similar framework which also notes that metrics can consider a single pulse of a stream of pulses over multiple years. Several authors have noted that under certain simplifying assumptions, the GWP is equivalent to the integrated GTP, and therefore any timescale arguments that apply to analyses of one metric would also apply to the other (Sarofim GTP paper and related references).

Second, the authors seem to misunderstand the messages of several recent studies, such as Shindell et al. 2017 and Ocko et al. 2017. These studies are not advocating for a shorter time horizon for GWP, as this paper implies in both the main text and the supplemental information. Rather, they are advocating for using BOTH short AND long-term time horizons to capture the full scope of climate impacts over all timescales – a key distinction that is not depicted in the text. The paper in its current forms criticizes these studies for something that they are not doing. Further, the authors frame their motivation around the fact that studies are advocating GWP20 to then show that GWP100 fits better with discount rates, but because these studies are not simply advocating GWP20, it makes the authors appear naïve to the existing literature. Further, there is a strong reason behind why other timescales are not promoted which needs to be acknowledged (it is not simply a lack of quantification in research efforts) – that just as it is difficult to move the policy community away from the comfortable GWP, it is reasonable to

believe that it will be equally as difficult to move the community away from 20 and 100 year timescales of which they are also most familiar with.

We will add clarification that in some cases authors are often suggesting presentation of short (about 20 years) timescales alongside, not in place of, 100 year timescales. However, we do feel that "promoting more emphasis on shorter time horizons" is an accurate description, as providing both GWP20 and GWP100 (for example) is "more emphasis" on shorter time horizons than just providing GWP100. Presented with values for 2 time horizons, one might also expect a decision-maker to have some probability of choosing to use only the shorter one, or to weigh them equally, which would look like a GWP60 (for comparison, a GWP60 would have an median equivalent discount rate of 5% using our methodology). Meanwhile, some authors (e.g., Howarth) do explicitly state that use of the 20 year GWP would better account for relevant climate impacts than 100 year GWPs, and a number of NGOs have followed suit. We propose adding a sentence to recognize some of these nuances:

> Recently, some researchers and NGOs have recently been promoting more emphasis on shorter time horizons, such as 20 years, which would highlight the role of short-lived climate forcers such as CH_4 (Howarth et al. 2011, Edwards and Trancik, 2014, Ocko et al. 2017, Shindell et al. 2017). These studies each have different nuances regarding their recommendations – for example, Ocko et al. (2017) suggest pairing the GWP100 with the GWP20 to reflect both long-term and near-term climate impacts – and therefore there is no simple summary of the policy implications of this body of literature, but it is plausible that more consideration of short-term metrics would result in policy that weights near-term impacts more heavily than would result from consideration of the GWP100 in isolation.

Third, it would be great if the damages function description went into more details about what is included in "damages." For example, I believe the authors make it clear later on that health or agriculture impacts from methane were not included. So what is included? Those damages are part of what makes near-term impacts so important to reduce, which justifies the use of a shorter time horizon.

Damages in this case are only climate-related damages. Inclusion of damages due to, for example, the health impacts of O_3 are relevant for policymaking purposes as one of the authors has argued elsewhere (see, e.g., Sarofim, Waldhoff, and Anenberg), but we would argue are not appropriate for the timescale discussion. One support for keeping timescale & non-climate effects separate is that if the timescale is adjusted to increase the value of methane to account for the methane-ozone effects, then it will equally increase the value of HFC-134a which has a similar lifetime: but HFC-134a does not have an equivalent ozone-effect. Therefore, instead, the relative climate effects of gases should be calculated using this timescale approach, and then the value of reducing methane can be increased to account for its ozone effects (and the value of reducing CO_2 can be adjusted to account for fertilization and acidification effects).

We propose to add a sentence in the conclusions noting that the timescale question is related to, but separate, from the relative value of abating methane and CO_2. Issues like ozone-damages resulting from

methane emissions should be addressed by a post-hoc adjustment, like how the ozone and stratospheric water vapor forcing effects are added to the methane GWP, rather than by using a method which truncates valuation of any climate impacts more than 20 years into the future.

Finally, I wonder about the argument that we should select a time horizon based off of appropriate discount rates. What if the GWP timescale tells us which discount rates are more appropriate? Why is it necessarily the other way around? The literature on appropriate discount rates is vast and its value is debated as much as GWP timescale selection. The paper makes it seem like there is solid agreement on appropriate discount rates but not GWP timescales, but both are subject to similar challenges and debates.

Our analysis can be used in either direction, as is discussed in the paper:

> Here we focus on what discount rates are consistent with a GWP time horizon in order to show the discount rates implied by common choices of GWP timescales. The converse calculation is relevant for an audience that has a preferred discount rate and is interested in the implied GWP timescale.

While we agree that there is no single consensus on an appropriate discount rate, we do think that that the framing of discount rates is good way to formally demonstrate the implications of GWP timescales for valuation over time. Part of the impetus of this paper was due to the possibility that some proponents of shorter timescales might not recognize the implicit discount rate embodied in the timescale choice, and might not consider a high discount rate to be desirable. We recognize that the discount rate implications calculated in this paper are a result of many decisions regarding parameters and analytical approaches, and that other approaches might yield different results, but we think that this is an important discussion to have.

We have also added a sensitivity analysis on the use of a Ramsey discounting approach as recommended by the National Academies for use in Social Cost of Carbon calculations (see response to William Collins) as that is an important alternate approach to discounting.

Minor comments:

1.26: Key criticisms also include the reliance of GWP value on a specified time horizon (that is a value judgement) (e.g. Ocko et al. 2017) and that emissions are not continuous (Alvarez et al. 2012). Would also include citations for each point of criticism that you mention.
http://www.pnas.org/content/109/17/6435

We propose to modify this paragraph as follows:

> Key criticisms of the metric are wide ranging. Some examples include: that radiative forcing as a measure of impact is not as relevant as temperature or damages (Shine et al. 2005); that the assumption of constant future GHG concentrations is unrealistic (Wuebbles et al. 1995, Reisinger et al. 2011); that discounting is preferred to a constant time period of integration (TBD); disagreements regarding the choice of time horizon in the absence of discounting (Ocko et al. 2017); that dynamic approaches would lead to a more optimal resource allocation over time (Manne and Richels, 2001); that the GWP

> does not account for non-climatic effects such as carbon fertilization (TBD); and that pulses of emissions are less relevant than streams of emissions (Alvarez et al. 2012). .

2.1: Definitely one of the reasons, stronger than "likely."

We will delete "likely".

2.3: Please explain upfront *why* you assess the choice of time horizon – as it wasn't even listed in your list of criticisms other than in reference to discounting (and it is problematic aside from discounting as well).

We propose starting the 3rd paragraph with the following sentence:

> In this paper, we focus on the choice of time horizon in the GWP as a key choice that can reflect decision-maker values, but where additional clarity regarding the implications of the time horizon could be useful. We also investigate the extent to which the choice of time horizon can incorporate many of the complexities of assessing impacts described in the previous paragraph.

2.3: 100 year was also selected as middle ground from IPCC FAR as values for 20, 100, and 500 years were given.

We suggest modifying the sentence as follows:

> The 100-year time horizon of the GWP (GWP_{100}) is the has been time horizon most commonly used in many venues, for example in trading regimes such as under the Kyoto Protocol, perhaps in part because it was the middle value of the three time horizons (20, 100, and 500 years) analyzed in the IPCC First Assessment Report.

2.8: Not sure why the word "therefore" is here. A description of why 100 year was selected does not in itself provide justification for why scientists are promoting 20 years. It is because 100 years does not adequately capture near-term impacts as it masks the importance of short-lived climate pollutants in the near-term. There needs to be a better transition from the 100 year discussion to the 20 year discussion.

We will replace "therefore" with "recently".

2.10: Papers such as Ocko et al. 2017 are not pushing for shortened time horizon, they are pushing for a two-valued GWP metric that includes BOTH 20 and 100 year time horizons. Very important distinction that needs to be clarified, as there are efforts (some livestock groups) that push for short time horizon only.

As noted above, we have attempted to present a more nuanced summary.

2.13: Part of the reason that other timescales are not suggested is because of the climate policy community's familiarity with 20 and 100 years. Just as they don't want to adopt a whole new metric, it is

very plausible that they will reject a new time horizon. Since 20 and 100 years are adequate for near- and long-term, pushing for say 30 and 200 year time horizons may be counter-productive.

We propose to clarify that we want more quantitative justifications of timescales in general (whether 20, 100, or 500, or anywhere in between (though also we are including more discussion of some papers such as Boucher 2012 where some analysis along these lines has already been done)

> We argue that more focus on quantitative justifications timescales within the GWP structure would be of value, as opposed to qualitative justifications such as a need for urgency to avoid tipping points as in Howarth et al. (2012).

2:20: There are more recent papers that need to be cited that look at the intersection of climate metrics and economics (Tol et al. 2012; Mallapragada and Mignone 2017).
http://iopscience.iop.org/article/10.1088/1748-9326/7/4/044006/meta
http://iopscience.iop.org/article/10.1088/1748-9326/aa7397

See above for our proposed language including some of these references and adding better context, and also our responses to other referees.

2.31: Why CO2 and CH4 only? Justification needed, such as represent the largest long-lived and short-lived climate pollutant contributor's to today's radiative forcing.

We propose adding the following sentence:

> The paper focuses on CO_2 and CH_4 as the two most important historical anthropogenic contributors to current warming, but the methodology is applicable to emissions of other gases and sensitivity analyses consider N_2O and some fluorinated gases.

3.5: Why is a pulse of 28.3 Mt of CH4 used, just bc of 10ppb? Why not today's annual emissions of methane from human activities (around 300-400 Mt)?

See our reply to William Collins for a description of the sensitivity analysis we performed in response to this comment, showing that sensitivity to the size of the pulse is small compared to other uncertainties.

3.12: What radiative efficiencies are used? Should specify this since you go into so much detail of other parameter values. I'm assuming radiative efficiencies are from IPCC AR5 but as you cite in your references, there are more recent calculations in Etminan et al. 2016.

The source of the radiative calculations is described here:

> The perturbation of radiative forcing from additional GHG concentrations are based on the equations in Table 8.SM.1 from IPCC AR5. CH_4 forcing is adjusted by a factor of 1.65 to account for effects on tropospheric ozone and stratospheric water vapor, as is standard in GWP calculations. N_2O forcing is adjusted by a factor of 0.928 to account for

N_2O's impacts on CH_4 concentrations, as is also standard in GWP calculations. Baseline radiative forcing is derived from the RCP scenario database.

As noted, Etminan has been cited as an example of updated information on radiative efficiency. Referee 3 also cited Modak et al. as showing the efficacy of methane forcing being lower than for CO_2 forcing.

See our reply to William Collins for a description of the sensitivity analysis we performed in response to this comment, showing that sensitivity to even a doubling of methane's radiative forcing would be very small compared to other uncertainties.

3:28: What damages are included by using this function?

This is meant as a simple approximation of all climate damages (sea level, health, ecoystems, etc.).

3:30: Please include citations for the first alternative.

We can cite the National Academies Social Cost of Carbon assessment here.

4.2: Suggest mentioning how these results fit in with scientific literature that has looked at these tradeoffs for decades.

We're not sure what tradeoffs the referee is referring to here: we discuss the Nordhaus GDP growth rates in the context of Gillingham et al.