Reply to Anonymous Referee #2:
We are thankful for the detailed review. We like to point out two important comments, which can be summarized as a better description of limitations (similar to the reviewer #1) and a better explanation of the factor 1/2 in the final tagging equation. We have included a section on limitations to clarify the reviewer's questions and we included some more equations in the mathematical re-formulation concerning the factor 1/2. For a couple of other comments, however, we have the feeling that the reviewer assumes and even demands that the tagging diagnostics package performs identical to the perturbation approach. This is neither our intention nor do we pretend it. In contrast, we clearly described the difference and stated that a combination of both approaches gives a good insight in atmospheric processes.

Reviewer Comment:
This new model development by Grewe et alii continues the development of tagging tracers in various ways to attribute environmental degradation to specific emissions (or possibly other actions). The method is reasonable but certainly not a unique or singularly correct approach to derive the environmental damage for a given set of actions. Grewe has made some very specific assumptions about how to partition a key species like tropospheric O3 into a unique sum of causes. The choices made are plausible, but there are readily available other methods (e.g., the 'perturbation method', line 50, which goes back decades before the Grewe references listed here indicate). I tend to concur with the other RC1 review in that a diversity of approaches can always teach us something. Thus, with some chemical model revisions and with a recognition that tagging does not just give us the "contribution", the model should be published.
Authors' Comment:
We are happy that the reviewer acknowledge, in principle, the publication of this method. We think that it is important to stress that different methods often address different questions. And hence, we agree that no unique, best, or better technique exists in quantifying e.g., anthropogenic impacts on atmospheric chemistry and climate change. We already state in the introduction
"The combination of both approaches leads to much better insights in the reasons how emission changes lead to concentration changes",
and hence clearly agree and support both reviewers view that the more diagnostics we have, the more we learn. We agree that the "perturbation method" has been used previously, frequently, and by many more authors, which we thought is obvious. The reference to our own work was actually meant to strengthen the point that we use both approaches and that we are not trying to rate one over the other, in general. For different purposes, though, one method might be, and actually is, better suited than the other. On the other hand, our experience is that the different aspects of different diagnostic packages are sometimes not well differentiated. Hence, we tried to define a wording in the beginning of the paper, which clarifies our understanding of the wording "contribution" and "change". There is no unique definition. But to our understanding, our choice is at least meaningful. And it is important to acknowledge that there is a difference, which the reviewers are obviously aware of, but others may not.

Reviewer Comment:
There are several serious problems with this GMDD paper as written that might be cleared up with major revisions affecting both the (i) chemical modeling and (ii) the authors' choice to describe the tagging method as the true 'correct' method and thus discussing the errors, for example, in the perturbation method. There is an additional

(iii) potential problem here in that the extensively expanded TAGGING here may still lack the full chemical coupling across species and regions that has been demonstrated in perturbation experiments with fully coupled models. There are some very interesting results here, but the paper needs to be a bit better balanced and informative.

Authors' Comment:

We are happy to provide more information and equally happy to adapt the wording to better balance our enthusiasm on our own work.

Reviewer Comment:

(i) The chemical model has some clear problems in language or concept. For example, CH4 does NOT make NMHCs. The production of NMHCs (e.g., C2H6 etc) comes from the sources. CH4 is a major source for H2CO in the remote troposphere but that species is not listed here and is not an NMHC. At best it might be a VOC. The NMHC reactions in the table make no sense.

Authors' Comment:

We have the impression that there might be a misunderstanding, which we probably have caused by using the word 'tagging' for both, the conceptual approach and the way we implemented it in the current model version. The mathematical concept 'tagging' is a decomposition of the equations in order to derive, what we called, contributions. No further assumptions, linearizations, or other limiting processes are applied. Hence, we do not see any problems with the mathematical concept. On the other hand the implementation requires simplifications. These are now better addressed in a Section by its own. We think that this might have caused some irritations. We think that the way we have implemented the mathematical concept of tagging provides useful information. However, work is still required to deal with shortcomings.

Reviewer Comment:

Another mistake appears to be the lack of photolysis of O2 as an important source of O3 in the tropical upper troposphere. This is well established and I can only take it that the old photolysis lookup tables used here have cut off this process in the troposphere? Otherwise the explanation for tropospheric O3 sources does not makes sense.

Authors' Comment:

Ozone is often divided up into stratospheric and tropospheric ozone, which either means ozone produced in the stratosphere and troposphere, respectively, or ozone present in the stratosphere or troposphere, respectively. Here, we are referring to ozone production terms in a way that we call ozone produced by oxygen photolysis, regardless where it happens, as stratospheric ozone production, since it is a process typical for the stratosphere. In the same way, we name ozone production via other chemical reactions, e.g. $NO+HO_2 \longrightarrow NO_2+OH$, as tropospheric ozone. We focus on the type of production terms, regardless of where it happens. However, we know that the one is a typical stratospheric ozone production and the other a tropospheric ozone production term. (Note that we have one chemical mechanism, from the surface to the middle atmosphere.) We think that this is justified, since the primary goal is to discriminate ozone production from surface emissions from other sources. The tropopause region is actually characterized by both processes. A split into also ozone produced in the tropopshere and stratosphere is feasible, but beyond the scope of this work. We have adapted the text in the introduction to Section 3.

Reviewer Comment:
Putting
in N2O emissions is interesting, but I do not see where it is then listed as a source of tropospheric NOx (0.1 Tg-N/yr)? Moreover, the N2O-CH4-O3 system in the stratosphere is quite complex and controls the net trop influx of O3 and NOy. That system is also established as a coupled perturbation, but is not accurately represented in this tagging model.
Authors' Comment:
The mechanism requires a complete set of NO_x emissions and loss terms (see introduction to Section 3). Therefore stratospheric NOx production is included. We do not understand the comment with respect to the coupled system.

Reviewer Comment:
Almost all modern scavenging algorithms for species such as H2O2 and HNO3 and others follow the rainout and washout AND re-evaporation of these species in layers below the original uptake. The method of tagging here would seem to be inadequate to deal with this.
Authors' Comment:
No, the 3D-tendencies are obtained from the scavenging submodel. Negative tendencies are interpreted as losses due to washout, positive tendencies are re-evaporation.

Reviewer Comment:
(ii) The tagging method as best I understand is built to achieve a zero sum in that a certain level of, say O3 concentration, is partitioned into the different sources so as to sum correctly. This is clearly a personal choice, since I would prefer the linearized tangent approach in which the differential evaluated at the current atmospheric state is used to calculate the change in O3.
Authors' Comment:
This actually seems to be the source of a misunderstanding. We are not considering changes in ozone. We are not investigating, how ozone would change if we were changing the strength of any emission source. This question, as discussed in our manuscript in the introduction section, would be best answered by the perturbation approach. Diagnostic packages as that favored by the reviewer or the tagging, we are proposing are adding additional information so that the physical and chemical reasons for these ozone changes can be understood. Here, we are investigating one simulation with one specific chemical regime and attributing emissions to ozone concentrations. In our view that is best described by the wording 'contribution of an emission sector to the atmospheric concentration of ozone' and it is different from 'contribution of changes in emission sectors to change atmospheric concentrations of ozone', to which, at least as far as we understand the comments, the reviewer is referring to.

Reviewer Comment:
This of course will not lead to a sum of all the
components being zero because the chemistry is non-linear over the range from zero to full industrial emissions. This has a similar issue with CO2 and radiative forcing

attribution, since the RF of CO2 goes as the log of its concentration. In this case we have the "which came first?" or "straw that broke the camel's back" problem. Recent increase in CO2 has a lower impact than earlier ones. For many of us the only fair way is to do the perturbation experiment (flat slope at high CO2) and then realize that the sum of all tangent additions if scaled to zero would not be equal to the current state. This is the same problem with the attribution of say biomass burning, and there is no single correct answer. This paper needs to realize and carefully explain the arbitrary choices made.

Authors' Comment:

We agree that the fundamental problem is the "straw that broke the camel's back". Please note that there is a difference between ozone and carbon dioxide, which, in our opinion, questions the approach suggested by the reviewer. It is well established that for a certain NOy concentration any increase in NOy reduces the net-ozone production (Ehalt and Rohrer, 1994 and many others). Hence an increase in any NOx emission source potentially leads to a lower ozone concentration. And this is then the case for all emission sectors. Defining the contributions on the basis of this approach hence leads even to an overall negative contribution of emissions to ozone. So what process actually produces ozone to close the budget? Note that in this case, we are considering only the "last straw" as relevant ozone prodcution terms. In our approach, we argue that air chemistry is not making any difference from which source a NO molecule has been emitted. A NO molecule from industry or from traffic emissions has the same likelihood to react with HO2. The reaction kinetics are the same.

Reviewer Comment:

There is a worrisome statement in the introduction that somehow demonstrates the absurdity of the different approaches. The idea that tagging is the 'correct' answer is just incorrect. It is indeed one of the answers, but not necessarily the best: "For example, the change in ozone due to a 100% reduction in road traffic emissions is smaller by a factor of 5 than the contribution of the road traffic emissions to ozone." The tagging method is obviously defined here to be "the contribution of ", but as a policy maker who wants to know what happens if I reduce road traffic, I would prefer a 100% reduction as the correct answer, or I might choose the 5% reduction times 20. Clearly if I reduce road traffic by 5% or by 100%, the perturbation run, not the tagging run, is the correct value. Does the tagging method do any better than a series of 5% reductions scaled across different emission sources? Moreover, one would not be interested in attributing the background basic atmospheric state (e.g., lightning emissions) in a proportional basis with those of industry, since the background state is not billable for damages as an anthropogenic perturbation is.

Authors' Comment:

We clearly stated that the perturbation method is the adequate method to answer the question how changes in road traffic emissions affect ozone or any other atmospheric species (see Introduction). So we still are puzzled how the reviewer got another impression. We are interested in understanding atmospheric chemistry. We are interested in how much lightning is contributing to the ozone concentration. Not everything is about "billing for damages". Agreed, that is important. But we also think that understanding the simulated changes is equally important and our diagnostics package is contributing to this understanding.

Reviewer Comment:
The authors continue to use the English word contribution as a code word for their own specific method for dividing up species concentrations. Lines 530-535 argue that the perturbation method "underestimates this contribution" by a factor of 2. I would assert that the "tagging-contribution" is larger than the perturbation by a factor of two and that one of the causes may be the lack of full tagging. The other reason is the apparent need to tag everything including background processes in a similar way to pollution sources. Personally I am not sure that the treatment of the background atmosphere by tagging is done well (e.g., the stratosphere) and thus the partitioning here may be specific to the assumptions made.
Also the authors really need to explain their 1/2 factor throughout the equations (starting with eqn 3). It is far from obvious since a simple Taylor expansion would not give 1/2. Please help us out. If it does not apply to small perturbations but only when trying to ensure that the sums are balance, then explain.
Authors' Comment:
Again and similar to our previous comment, the reviewer implicitely assumes that we are investigating changes or perturbations, because she or he is mentioning the Taylor-approximation, which is based on estimating effects of a small perturbation. Since we have foreseen such a discussion, we already have included a whole Section on tagging basics, where we clearly state that the tagging principle does not include any approximations or linearizations. An additional explanation on the factor 1/2 is given above. At reaction level, looking at reaction kinetics, rather than concentration level, where we consider solutions of a ODE, the reaction rate for the above considerer reaction is P=k [NO] [HO_2]. The chemistry mechanism is not discriminating between NO from different sources. The reaction rate P is not adapted whether NO from road traffic or from lightning is reacting with HO_2. The molecules NO and HO2 are both equally important for this reaction. See also answer to reviewer 1. We have included a discussion on this point in Section 2.

Reviewer Comment:
(iii) There are clearly identified global chemical coupling patterns that reach across species and regions (strat vs. trop). They are readily identified in models through perturbation simulations. For example, these early chemical feedbacks of tropospheric OH-CH4 and the N2O-NOy-O3 in the stratosphere have been demonstrated to work across many models in various IPCC model comparisons. These are important because they affect the lifetime of a perturbation and hence the attributable damage of emissions. They are most surely in the full MESSY model. From the couplings of this tagging method, I do not believe that these fundamental couplings are present in the tagging model. If you could demonstrate that both of these feedbacks can be derived from the tagging then it would be convincing. Otherwise it shows that tagging really cannot include the dominant chemical feedbacks of the lower atmosphere. This lack of full coupling in the TAGGING model means that one cannot be sure what chemical feedbacks are not included.
Authors' Comment:
The tagging method, as a diagnostic package, is controlled by the MECCA chemistry, which includes

feedback processes. And can be analyzed, as suggested, by perturbations of the system. In the paper we have described the feedback of NOx emission changes to the ozone production efficiency: A decrease in, e.g., road traffic emission leads to lower NOx concentrations. This may lead to larger net-ozone production rates. Hence this is a negative feedback. The lower NOx emissions (with unchanged) net-ozone production rates leads to lower ozone concentrations. However, since the net-ozone production rates increase, the ozone reduction is reduced. This is a negative feedback process, which can be deduced only in the comparison of two simulations. Another interpretation is that the unchanged, e.g., lightning emissions, also experience the enhanced net-ozone production rates decrease, and hence the ozone produced by lightning NO is larger than in the previous simulation. During the simulation and more importantly also in reality, a NO molecule has no remembrance of its source of emission. Hence, no difference is made between the reaction of a NO molecule with HO2, whether it originates from lightning or road traffic. Feedbacks are always related to changes with respect to a base situation, whereas tagging is analysing a base situation by itself. So we think that asking for feedbacks to be included in the tagging is simply trying to achieve the same results from the tagging method as a perturbation method is providing. However, this is not the intention and not the case. Both methods are valid and usable, however, answering different questions.

Reviewer Comment:
Other issues: (iv) The method is described as low cost and non-intrusive, but the only global example given is for T42 resolution (2.8 deg). This is very low resolution for current global models, yet this is only a GMD paper to establish the development of the model. OK, but can you run at T159 (1.1 deg) for example with all the memory requirements for the tagged tracers to be transported? I had thought that tracer transport was one of the dominant costs of high-res CTMs. Indeed, line 564 seems to indicate that you already have memory limitations at T42.
Authors' Comment:
Actually, EMAC is not a CTM. I am not quite sure, with which coupled troposphere-to-mesosphere chemistry-climate model, multi-decadal simulations were performed at T159. The most recent models, which will participate in the upcoming IPCC via CCMI have actually a similar resolution as we have used in our study (Morgenstern et al., 2017; their table 3). Moreover, the sensitivity study presented in Section 5.1 has a horizontal resolution of 10 km or 0.1°.

Reviewer Comment:
(v) What was the STE flux of O3 and NOy as a function of latitude and season. This would seem to be very important since the background atmosphere shares the attribution in this scheme. Please denote.
Authors' Comment:
We have estimated the net ozone flux from the stratosphere in the EMAC model for the years 2000-2004. Using the residuum method, i.e. STE=Burden change-Prod+Loss+Deposition, the ozone STE is calculated to be 393 +- 25 TgO3 per year (Jöckel et al., 2006), which is in agreement with other modeling studies (Stevenson et al., 2006). Note that this method includes net fluxes, i.e. upward and downward fluxes of ozone through the tropopause. The calculation of the stratosphere-to-troposphere ozone flux was performed by using a diagnostic tracer Strat-O3, which is nudged to ozone in the stratosphere and experiences loss terms (chemistry and deposition) in the tropopshere, only. The accumulated loss terms provide a measure for the stratosphere-to-troposphere ozone flux and is 1198 +- 28 TgO_3 per year (see

also Jöckel et al., 2006, Table 2). The analysis of the impact of STE on the composition of the troposphere is indeed an important question, but beyond the scope of this paper. Generally, it was shown with a similar approach used here that stratospheric ozone changes and STE variations lead to variations in tropospheric ozone (e.g. Grewe, 2007). The concentration of the tagged tracer for stratospheric ozone is basically a result of the ozone produced by O_2 photolysis, transport to the tropopshere and tropspheric loss processes.

Reviewer Comment:
(vi) The idea that the Mediterranean Sea contains "pristine areas" anywhere is at least humorous - thanks.
Authors' Comment:
You're welcomed. Text adapted.

Reference:
Morgenstern, O., Hegglin, M. I., Rozanov, E., O'Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Marcal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K., and Zeng, G.: Review of the global models used within phase 1 of the ChemistryClimate Model Initiative (CCMI), Geosci. Model Dev., 10, 639-671, doi:10.5194/gmd-10-639-2017, 2017.