Variability and trends in dynamical forcing of tropical lower stratospheric temperatures

Response to reviewers by S. Fueglistaler on behalf of all co-authors.

We thank the 3 reviewers for their detailed, constructive and critical reviews. Based on the reviews, and our own assessment, we concluded that a major revision is required. Compared to the previous version, we have changed the focus of the discussion (and added/removed some of the analyses), but the overall ‘message’ remains similar. However, the text is new (except for a few paragraphs), and consequently highlighting text changes makes little sense. To facilitate the reviewer’s work, in the point-by-point response provided below, we - wherever possible - point to page/line number of the revised manuscript.

We realize that we’ve made more changes than usual for ‘major revisions’, and before we provide the point-by-point response it may be helpful to have a brief overview of the main changes, and the motivation behind.

The original manuscript was motivated in no small part by the question to what extent the eddy heat flux proxy can be used to estimate temperature variations. That is, the proxy was shown to provide reasonable results before (Fueglistaler, 2012), but in that paper no systematic comparison with other reanalyses, or with the more physical momentum balance calculation, was provided. The original manuscrip then also mentioned the use of 100hPa eddy heat fluxes, without providing much explanation except for vague hints pointing to CCMVal2 data. We indeed have analysed the CCMVal2 model data, but these results are beyond the scope of the present paper. We agree with the reviewers that this left the reader wondering.

The revised version is more focused on the main results of this work, namely the similarities and differences between the eddy heat flux calculation and the momentum balance calculation, and between the reanalyses. To this end, we have largely removed results based on NCEP and focus on the two more modern reanalyses, ERA-Interim and MERRA. We have added the momentum balance calculation for MERRA, and most of the discussion now concerns these 4 calculations (i.e. eddy heat flux and momentum balance, using ERA-Interim and MERRA). We have also decided to discuss results more in context with previously published results. It is not possible to discuss all previously published papers, and quantitative comparisons can be difficult because each study evaluates slightly different quantities. Bearing these considerations in mind, we have decided to focus on two studies, namely Fu et al. (2010) and Polvani and Solomon (2012). These two studies arrive at somewhat contradicting conclusions (Fu et al. emphasise the dynamical forcing for temperature trends, while Polvani and Solomon emphasise the contribution from ozone trends). We show in our paper that the eddy heat flux proxy gives results similar to the Fu et al. study, which is interesting because Fu et al. also work with eddy heat fluxes (but different reanalyses, and different specification where eddy heat fluxes are taken). Conversely, our momentum balance calculations seem to fit more with the results of Polvani and Solomon inasmuch as with these calculations a large contribution to temperature trends from radiatively active trace constituents is required. The revised manuscript now shows more clearly where the differences in the trends originate. Although trends differ, the agreement in the evolution of dynamical forcing of temperature (and how much of the observed variability is explained by the dynamical forcing) is remarkable, and is certainly one of the main new results of the paper.

Finally, we have changed two minor technical details of the calculations, and hence numbers may differ slightly between the two versions of the manuscript. First, we now eliminate 3 full calendar years for the volcanic periods (i.e. starting with January of the year where the volcano erupts). This ensures that the trends calculated with annual means, and for months individually, are consistent (i.e. they draw from the same data in the 'no-volc' case). Second, the regression slopes between actual temperature and estimates of dynamical temperature are calculated with an ordinary least squares fit; we have also experimented with a total least squares fit (one can find arguments in favour of each method). Generally, the slopes are about 20% larger when calculated with a total least squares fit for all estimates, which translates directly to a trend magnitude difference of 20%. The revised manuscript now discusses this source of uncertainty (and also states that the 20% uncertainty is smaller than other uncertainties).

In the following **point-by-point responses** the reviewer comments are in *italics*, and are numbered for further reference.

Reviewer 1:

R1-1: *In this work, the relation between lower stratospheric heat fluxes and tropical lower stratospheric temperatures is examined, in three reanalysis. Not clear if the aim is the inter-comparison itself, of if new knowledge is expected from the inter-comparison exercise. Or if the main aim is the comparison of the two different methods used. What is also unclear is the role of "variability" and "trends" in the investigation (see below). So, although the material is technically sound, the manuscript can be improved by clarifying the aim of the work and its scientific goals. My recommendation is possibly publication after a major revision of the aims and interpretations, following the general and specific comments below.*

We agree (see also general response).

R1-2: *Once the whole paper is read, it seems that the aim is on defying trends in tropical lower stratospheric temperatures that are due to dynamical processes - hence a (high frequency wrt the trend time scale) "variability" proxy is used to characterize dynam- ical processes. Fine. If this is indeed the case, I would remove "variability" from the title - because it is not an objective of the work to determine causes of "variability in dynamical forcing", something distinct to "trends in dynamical forcing"*

With variability we refer to the evolution of the time series (i.e. from monthly to interannual), rather than variability in the processes that contribute to the dynamical forcing. The revised manuscrip (abstract, introduction) hopefully clarify this point.

R1-3: *Methods are generally well described and sound (scientific quality excellent), however not new: previously used by the Author(s), and variants by others as well, and as appropriately cited. Hence this work is a useful application, while not proposing/exploring new approaches. The applicative nature of this work is acknowledged by the Authors at page 7, discussion of the proxy. One is therefore lead to ask, what specific aspects of the results and conclusions are new? Unfortunately, this is not clear from the current conclusions.*

While similar methods have been used before, our study is, to the best of our knowledge, the first to discuss the full time series in detail, and provide an overview of similarities and differences between different methods, and different reanalyses. The revised manuscript hopefully makes these points clearer.

R1-4: *Given that, it is disturbing that the aim of the work is not well defined. "Variability and trends" may cover a variety of scales - I could* *deduce those of interest after reading the methods (2.2) section - but the timescale of interest is a scientific aim that should be stated explicitly, best in the Introduction. Given the applicative nature of this work, and its focus on the dynamical proxy, the calculation of the correlation could actually be better explained. In the Northern hemisphere, the correlation I would expect from eq.(3) page 8 would actually be negative (larger positive heat flux anomaly, stronger circulation, negative tropical temperature anomaly). But this is not the case (figure 1)*

(See previous responses.) With respect to the correlation between tropical temperatures and northward eddy heat flux: yes, the correlation is negative, but the proxy is positively correlated (the sign change is from the slope of the regression).

R1-5: *The main result seems to be that ERA-Interim and MERRA mostly agree (for the carried out comparison), while NCEP deviates. Is this a new result? The inter-comparison is useful. Not consistent trends are found in the three reanalysis for the dynamically forced trend. But - unfortunately - no reason/speculation is given for the noted discrepancy!*

The revised manuscript does not discuss the NCEP trends anymore; no claim is made that any method or reanalysis is better than the other(s). Generally, we try to not speculate why different methods/reanalyses give different results, but to identify where the differences originate from - this allows future work to focus on these aspects, and perhaps establish causes for the differences. While not explicitly stated in the manuscript, we note that (a) targeted primitive equation model calculations may be used to address the question whether eddy heat flux calculations give biased trend results, and (b) that problems in reanalysis data are almost impossible to identify with the very limited information available in the public domain.

R1-6: *P 3, LL 15-20: I would suggest to introduce the purpose of the work not with ",.. and . . ." but a proper sentence. Also not clear what is the aim of the reanalysis inter-comparison: Is it the interest of the author only to do the inter-comparison - or do they aim at a gain in learning? If so, what learning?*

(See responses above.)

R1-7: *P 3, L 25: "The dynamically forced upwelling in the tropical stratosphere links temperatures and dynamics via radiation" unclear.*

This text has been removed. The mechanism is now discussed in Section 2.1 (on page 5).

R1-8: *P 5, LL 1-5: Unclear the time scales of the trends considered*

- and/or "trends" is used in a more general sense, that is including "low-frequency variations" - that may looks like a trend if the time-series is not long enough. The paragraph is vague. One is left with asking: Is the "identification of possible reasons for differences in trends" between re-analysis the main aim? So, is the focus on disentangling if some re-analysis might mis-represent some processes? Or is the focus on the origin (radiative versus dynamical) of observed trends?

(See responses above.)

R1-9: *P5, L 12: Please specify "tropical mean"*

We explain the choice of 35S-35N now in Section 2.5.

R1-10: *Page 6 LL 1-4: Not clear the context of these not shown approximations. "Variations" on which time scales?*

We used radiative transfer calculations with the fixed dynamical heating assumption; we deliberately keep this paragraph short since aerosol and ozone are dominant, which is all what is needed to follow the discussion in the paper.

R1-11: *Page 9 LL 16-24: The results of this paragraph sounds a bit circular, given that tau70 days is selected to obtain maximum correlations between Wstar and temperature variations. Please explain better what is expected by construction and what is not.*

We agree with the reviewer's concern. The text in the revised manuscript (page 8, lines 14-20) should be clearer.

R1-12: *Page 10 L 21: "is a coincidence" is too strong and misleading. The ozone induced effect highlighted by Fueglistaler et al 2011 is of dynamical origin - hence the strength of the residual circulation is heavily implicated. The residual circulation is indeed behind the relationship assessed by the authors (as they know well).*

Yes, the temperature variations are forced by the dynamics, but the fact that the variations in the combined extratropics and tropics nearly cancel is an artefact of the MSU-4 weighting channel - if the weighting function were different, no such cancellation would be observed. (I.e. the anticorrelation is not an artefact, but the similarity of the magnitudes in MSU-4 data is.)

R1-13: *Page 17/18 Conclusion, seasonality of the trends: Any new result with respect to Fu et al ACP 2010?*

The revised manuscript puts results now explicitly in context with Fu et al. (2010).

R1-14: *Page 18, LL7-11: This extension of the Fueglistaler 2012 result is interesting. However, it is not clear what is meant with "dynamical response" - are the Authors attributing this dynamical disturbance to Pinatubo? What* *would be the evidence, based on this work, for this claim? Could the Authors also please discuss how this tropical cooling (and the implied circulation increase) relates to the stratospheric-tropospheric response in the Northern hemisphere claimed (e.g., Robock/Graf/Stenchikov works) to occur after volcanic eruption? And what could be the role of other phenomena unrelated to the effects of the Pinatubo aerosols?*

The dynamical response following Pinatubo is one of my (sf) main interests, and all points raised here by the reviewer are questions that I am working on. In the Fueglistaler [2012] paper I only show one dataset (ERA-Interim) and one method, namely the v'T'-proxy. I have always been concerned that the stationarity assumption between the proxy and temperature might break during an unusual situation such as the Pinatubo period. Hence it is very important that we show in this paper that the momentum balance calculation gives a very similar response, and that results do not depend on choice of reanalysis. Further discussion of the Pinatubo period, i.e. how/why/if the aerosol heating induces the anomalous dynamical situation, is beyond the scope of this paper but is subject to ongoing work in my group.

R1-15: *In general, the figure captions should be more precise, for example: Figure 1: correlations: are those reported for "no-volc" for the 1995-2011 period as well? Why the correlation goes down, for "no-volc"?*

We have revised the figure captions, and have been careful to explicitly state throughout the manuscrip that 'no-volc' refers to the full period. Also, the revised manuscript discusses these correlations in more detail. (Section 3.1.1 on page 12ff.)

R1-16: *Figures 8: caption has letters for panels, but these letters are missing in the figure. Description of what seen in panels a, b is insufficient. What are the signatures?*

The reviewer is referring to the manuscript submitted, not the version that was published in ACPD. The missing labels were corrected during the 'technical review' stage of the manuscript; but the reviewer is correct that the caption was incomplete; which is corrected now. (In the revised manuscript, this figure is now Figure 1.)

Reviewer 2:

R2-1: *This study examines variations in tropical lower stratospheric temperatures in a number of reanalysis datasets (ERA-Interim, MERRA and NCEP). The study addresses several issues: (1) dynamical contributions to shorter timescale variations; (2) long-term trends; (3) consistency between multiple reanalyses; (4) consistency between different methods for estimating dynamical contributions in a single reanalysis. I have no major issues with the scientific analysis in the paper, although it would be nice to also have a comparison of the momentum budget estimate for MERRA and a discussion of whether this is consistent with ERA-Interim. My issues are rather with the overall clarity of what the study aims to do and why, and what the reader should expect to take away from it. Because of the multiple foci, I was left feeling that whilst there are some interesting results, the manuscript lacked a coherent narrative that clearly highlights the novel contributions and key messages. Many related topics have been covered in previous studies and since the attempts to explain why differences arise are limited, the results read at times like a bland documentation of characteristics of reanalyses. Although such a documentation could be of use to other reanalysis users, my view is that it is of limited value to the ACP community in its present form. I think that the manuscript could be significantly improved by providing more background motivation and well de- fined goals in the Introduction and clearer outcomes and links to the wider literature in the Conclusions. I have made some suggestions for how to improve the manuscript below and these points would need to be addressed before I would recommend the paper for publication in ACP.*

(We agree with the reviewer - see our general response above; the revised manuscript now also discusses the results for the MERRA momentum balance calculation.)

R2-2: *P3L20-21 the purpose of this paper is to document common features and differences therein in several modern reanalyses. is this the main goal? If so, what is the new information coming out and how does it fit into the existing literature? See also P4 L8-12.*

(See general response.)

R2-3: *P14L14-20 It seems a bit contradictory as to what to take away from the paper. In some respects, the proxy does a remarkably good job of capturing the interannual variability, and there is some consistency between the reanalyses as to the long-term trends (e.g. the vT estimate from ERA-I and MERRA). But this result for ERA-I casts some doubt on how these trends should be interpreted. So what is the overarching message?*

The revised manuscript cannot resolve this contradiction, but shows that it is the trend in high latitude eddy heat fluxes in December/January that account for much of the differences between the two methods. To the best of our knowledge this problem has not been pointed out before, but affects also all previous work based on eddy heat flux considerations. The overarching message is that as long as this discrepancy is not resolved, we don't know whether we can trust trends in observations (reanalyses).

R2-4: *Section 3.4 This section is again giving rather mixed messages. After making the case in section 3.3 that the momentum balance and eddy heat flux methods disagree in their trends for ERA-Interim. This section gives a detailed description of the seasonality of the hemispheric trends in the eddy heat flux proxy and states that there are strong consistencies between MERRA and ERA-I. However, it then goes on to say that the momentum balance estimate for ERA-I has different seasonality to the others. There is no estimate of the momentum balance calculation for MERRA, so we cannot know whether this proxy might be consistent across the two reanalyses. The reader is therefore left wondering what to take away from these calculations, other than a strong sense of skepticism and some confusion. These contrasting findings need to be brought together more clearly in a summary paragraph or the conclusions.*

The revised manuscript is hopefully clearer in the presentation of the results, and now also shows the MERRA momentum balance calculation. But that calculation does not resolve the problem - i.e. the message remains "mixed" - but it helps to clarify that the main difference is between the methods, not reanalyses. An important result from our work is indeed that the message is mixed - uncertainties remain large, and we need to be very open about this.

R2-5: *Conclusions: This needs a paragraph which brings all the findings together and links them back to the literature. What is the key message a reader should take away? This is not clear to me in the current manuscript. At several points in the manuscript the authors make reference to the limited dynamical output available from model intercomparison projects. However, the potential value of their simple analytical approach for modeling studies is not strongly emphasised. Do the findings here suggest that a significant amount of information can be extracted from the limited available data? If so, that should be made clear in the conclusions.*

The revised manuscript is hopefully clearer; we emphasise that both methods actually give surprisingly good results, and much can be said about the evolution of tropical lower stratospheric temperature, specifically also for specific periods of interest (such as the Pinatubo and year 2000 periods discussed in the paper).

R2-6: *P3 L4 Thompson et al. (Nature, 2012) show that models forced with all known forcings also do not simulate the observed cooling in the tropical lower stratosphere.*

Yes; we will discuss this in the context of work in progress where we look at GCM results (as archived in CCMVal2, and new runs).

R2-7: *P3 L4 Consequently, trends in stratospheric dynamics . . . may play a major role for temperature trends the GCMs referred to in the previous sentence could simulate dynamical trends due to increasing CO2 alone, and indeed most of them do, so this statement is not clear.*

We would argue that the situation is even worse - GCM's produce a trend in dynamical forcing with CO2, but do we know whether there really is a trend in dynamical forcing? We wish to leave the statement (which is quite general) as is. Since we do not analyse GCM results here, we will not say much about trends in GCMs except for the discussion of the Polvani and Solomon results (see page 18, lines 12-15), where we note that none of the four calculations give an upwelling trend similar to that reported by Polvani and Solomon (2012) in their GCM calculation.

R2-8: *P3 L2-8 It might be worth clarifying the potential for dynamical and radiative contributions to regional and global stratospheric temperature trends. i.e. dynamical contributions small on global scale.*

If we understand correctly, the reviewer refers here to the fact that dynamically forced temperature trends are of opposite sign in the tropics and extratropics. In Fueglistaler et al. (2011) we argue that the tropical temperature change is - roughly! - twice the extratropical temperature change (due to ozone amplification and latitudinal structure of stratification), such that an amplification of the residual circulation should lead to a global average cooling of the lower stratosphere. However, in order to keep the discussion focused on the analyses presented in this paper (we do not show analyses for the extratropical temperatures), we prefer not to discuss this aspect here.

R2-9:*P3 L11 delete what*

Text changed.

R2-10: *P3 L25 four uses of different/differences in one sentence!*

Text changed.

R2-11: *P3 L26-28 The dynamically forced upwelling in the tropical stratosphere links temperatures and dynamics via radiation. Would it be clearer to* *make this statement by saying that temperatures reflect the balance between dynamical and radiative processes?*

Yes, we agree with the reviewer; but the new text does not have that statement anymore.

R2-12: *P4L1 to processes not directly coupled to dynamics e.g. ?*

Text changed; we refrained from listing possibilities because the list would be long - ranging from chemical ozone depletion to changes in stratospheric composition due to tropospheric chemistry, aerosol processes etc.

R2-13: *P4L9-10 The TLS time series is considered to be of high quality, but may suffer also from smaller artifacts. References for these statements?*

We have removed this statement; it was based on a poorly-quantitifed general feeling within the community. Since the statement is not critical for our work, we eliminate it and spare us a longer excursion into reliability of lower stratospheric temperature trends. For the same reason, we focus our paper on the estimates of the dynamical temperature, and do not discuss the differences in actual temperature between ERA-Interim and MERRA.

R2-14: *P4L29 P5L5 These are all relevant, however, we never learn which of these, if any, are most important for the discrepancies described later in the results section.*

We decided to keep this overview so readers know what may contribute to errors; as also said before, with the limited information available from reanalyses, combined with the fact that reanalyses are free to violate energetic constraints, it is almost impossible to *exactly* pinpoint what the dominant errors are. Compared to the original manuscript, we now provide more information about the source of differences from the data perspective.

R2-15: *P5L23 tropical means in this work later on (P11L1) is it stated that the tropics are defined up to 35. Is this the tropical mean referred to here? If so, it would be helpful to make this clear.*

This part is now on page 5, lines 6ff; at this point, tropical mean is sufficient, we later specify what exactly we mean with 'tropical mean'.

R2-16: *P5L24-25 where variations in T_E are a consequence of variations in radiatively active tracers. Could changes in surface temperature also affect T_E? These might not be coupled to radiatively active tracers.*

Yes, the reviewer is correct that the discussion here is not complete. Once one allows variations of T_{E} in the Newtonian cooling framework, these variations may not be only due to local tracer changes, but also due to anything that changes the incoming up and downwelling radiative fluxes at that layer. Surface temperatures have an impact on lower stratospheric temperature via radiation in the ozone bands, where the troposphere is relatively transparent. However, we feel that a complete discussion of the radiative transfer aspects at this point would confuse the reader, not least also because in the remainder of the paper we do not discuss what contributes to the changes in T_{E}; that is, we seek to quantify changes in T_{E}, but do not seek to attribute these changes to any specific process.

R2-17: *P6L19 Its not clear to me what negative time means in this context - can you elaborate on this.*

We've followed standard conventions for the convolution.

R2-18: *P6L9-12 You should state which altitude range this refers to, since it might not be true at all levels.*

Done.

R2-19: *P7L7-8 Out of curiosity, does the choice of resolution make much of a difference for the non-linear terms? e.g. if you regrid ERA-Interim to 2.5 do the values change much?*

We have not done this calculation (for the admittedly prosaic reason of lack of time); since the revised manuscript focuses on ERA-Interim and MERRA, this is less of an issue. However, we agree with the reviewer that this is an interesting aspect, and we hope to have a look into this in the near future.

R2-20: *P7L24 should be very similar this is cryptic. Have you checked that they are similar? If so, you should say they are similar, if not then this is speculative.*

Yes, we agree with the reviewer that this is too vague. The revised manuscript is now more precise in that it explicitly compares with Fu et al. (2010).

R2-21: *P8 L2 Li and Thompson (JGR, 2013) also use a similar proxy to relate midlatitude dynamical forcing to tropical cirrus clouds and tropical temperatures in Era Interim.*

Yes, there are more studies that use eddy heat flux proxies in one way or another - we've trimmed the selection to studies with a similar focus.

R2-22: *P8 L6-7 Is it important that this metric will integrate across the turnaround latitudes during some seasons and will therefore include a part of the tropical upwelling region?*

The revised manuscript discusses the contributions from different latitudes in detail; the main problem is probably not that we average over a band that sometimes includes part of the upwelling region, but that the reanalyses do not agree about the trends in eddy heat fluxes at lower latitudes (see Figure 8 of the revised manuscript).

R2-23: *P8L10-11 I consider 25 to be in the subtropics. Does this therefore imply that the subtropics dont matter for tropical upwelling at 70hPa, since you state that you results are insensitive to whether or not the subtropical component from 25-40 is included? If so, is this surprising given the suggested important role of the subtropical wave driving for simulated long-term trends in tropical upwelling (e.g. Shepherd and McLandress, 2011)?*

The revised manuscript shows that the role of the subtropics is unclear because the trends there differ between ERA-Interim and MERRA. Hence, it is difficult to make a statement about the role of the subtropics. We plan to analyse this in more detail in future work, also in light of the results shown by Shepherd and McLandress (2011).

R2-24: *P8L22 Does the shorter optimal timescale for NCEP suggest a fundamental difference from the other reanalyses, e.g. a greater occurrence of shorter vertical wavelength perturbations, or some fundamentally different radiative behavior? Or is this within the uncertainty range?*

This is an interesting question that we intend to study in more detail in the near future. The quick answer is that it probably reflects systematic differences in the eddy heat fluxes - for example due to different resolution - but that because the uncertainty is large (see Figure 1c of the revised manuscript) it may be difficult to make a strong case.

R2-25: *P9 L26-27 Does the use of fixed bounding latitudes rather than the turnaround latitudes make any difference to the results? Past studies have shown that diagnosis of trends in tropical upwelling in reanalyses can depend upon the use of fixed latitudes vs. turnaround latitudes (e.g. Seviour et al., 2012).*

The momentum balance calculation provides the mean upwelling between the bounding latitudes, and as such the dynamical forcing that acts on the temperature between the bounding latitudes. Hence, changes in the turnaround latitude do not affect the momentum balance calculations. Note that we do not interpret the results in terms of net overturning mass. The fact that for temperature we average between 35S-35N is of no importance, results for temperature are virtually identical between 35S-35N and 30S-30N.

The situation is different for the eddy heat flux proxy. The intensification of high latitude eddy heat fluxes and decreases over mid-latitudes (see Figure 8 of revised paper) corresponds to a poleward shift of eddy activity, which could occur in the situation of a broadening of the upwelling region. As such, one could speculate that the eddy heat flux and momentum balance calculation may not contradict each other: overall, there may be a larger mass turnover (as suggested by the net increase of eddy heat fluxes), but that this is due to a widening of the upwelling region, but not intensification of upwelling between 30S-30N (hence the momentum balance calculations show no trend). If this were the case, the momentum balance calculation would give the correct forcing for the tropical temperature. However, with the available data we cannot prove/disprove this speculation, and we therefore do not wish to elaborate on this possibility in the paper.

R2-26: *P10L12-13 It would be worth emphasizing that any changes in T_E due to circulation driven changes in ozone would be included in this estimate.*

This text is newly formulated.

R2-27: *P11L12-15 You should clarify here that the step-wise changes follow the two major volcanic eruptions, I think.*

Yes, done.

R2-28: *P12 L11 What is the effect? State that the trends are slightly smaller without the volcanoes included.*

The new text briefly points out the main differences between the full time series, and the 'no-volc' time series. The subsequent discussion then is on the 'no-volc' time series, since results including volcanic periods are dominated by the distribution of volcanic events in the period considered.

R2-29: *P13L1-5 This is interesting, but it begs the question as to whether there is any understanding of where the difference in the annual mean trend in NCEP might be coming from?*

In the revised manuscript, we have eliminated most of the NCEP results to allow for a clearer structure of the paper. The eddy heat flux trend differences shown in Figure 8 of the revised manuscript motivate us to look into the full latitude/height structure of eddy heat fluxes of more reanalyses in the near future; but this is beyond the scope of this paper.

R2-30: *P14L22-23 by generally good agreement are you referring to the shorter timescale variability? Please make clear in the text.*

The new text is now more precise.

R2-31: *P14L24-27 This agreement seems like cherry picking. There is no explanation as to what determines the periods for which the two methods do agree and the periods for which they dont agree, and more importantly why.*

The revised manuscript is more precise; specifically, we note that the large amplitude variations are very similar, the differences are for smaller amplitude variations, and for trends.

R2-32: *P17L5-7 This is written as though this fact invalidates the estimates of the trends in the proxy timeseries, but does this have to be the case? Surely the trends in the residuals could be driven by different processes (e.g. volcanoes) which dont have to be represented consistently in MERRA and ERA-I for the long-term dynamically driven trends to be consistent? You need to better explain what you mean by this sentence.*

This text has been removed.

R2-33: *P18L14-16 But since both calculations are subject to errors, should we believe any of the estimates discussed here? Particularly given that the reanalyses assimilate much common data.*

The revised text is more precise; indeed, what we find (and is now explicitly stated) is that the dynamical information of the two reanalyses is fairly similar (as may be expected due to common input data), and that the differences between the methods are systematic.

R2-34: *Figure 1: You should change the figure labels to show 80:11 actually means 1980:2011.*

The figure (now Figure 3) is now drawn without this information, instead the information is given in tables.

R2-35: *Figure 5: Its interesting that MERRA indicates a significant warming trend from the NH in JJASO, which partly balances the SH driven cooling. This isnt apparent in ERA- Interim, so is worth pointing out.*

Yes, we agree. Unfortunately, in order to focus the paper we decided to drop the most of the discussion of hemispheric contributions. We hope that we can return to this aspect in the near future in the context of the eddy heat flux trend comparisons mentioned above.

R2-36: *Figure 8: There is a lot of information in this figure could it be simplified. I dont for example know what the difference is between the triangle and square lines in panels (a) and (b).*

Our apologies that the caption for this figure was incomplete; this is now fixed in the revised version. (Where it is now Figure 1.)

Reviewer 3:

R3-1: *The manuscript by Fueglistaler et al. discusses the contribution of dynamical forcings to variability and trends of tropical lower stratospheric temperatures at 70hPa. The analysis is based on meteorological reanalyses from ERA-Interim, MERRA and NCEP over the period 1980-2011, together with MSU satellite temperature data. Dynamical forcings are calculated (a) from a simple proxy based on eddy heat flux at 70hPa integrated over the extra-tropics of both hemispheres and (b) the momentum balance evaluated at 30N/S, following Randel et al. (2002). It is shown that both methods to calculate the dynamical forcing describe the inter-annual variability of tropical lower strato- spheric temperatures well, with high consistency between ERA-Interim and MERRA, but give very different results for the long-term trends. The paper addresses the im- portant topic of attribution of changes in temperatures and upwelling in the tropical lower stratosphere and makes important contributions to this topic by evaluating and discussing the dynamical forcing. I recommend publication in Atmos. Chem. Phys. after consideration of the following remarks.*

–

R3-2: *1. The temperature proxies based on extra-tropical heat-flux and based on the momentum balance have very different long-term trends. Is there any chance to explore in more detail why the trends are so different, including the very different seasonality of the trends? As discussed in the manuscript, the momentum balance is the theoretically more complete description, but may also be more sensitive to errors in the reanalyses. In any case it would be good to provide a bit more background here on the details of the calculations and a discussion of how the two proxies are theoretically related. E.g., is the term u'w' in the momentum balance neglected?*

The revised manuscript shows in more detail when/where the two estimates differ, but we have not been successful (with the information available) to clarify unambiguously which terms of the calculations may be affected most by errors in the reanalyses (we continue working on this problem, but from what we know at present it is clear that much more time is needed, and it is not clear whether there will ever be a clear result). In our calculation of the EP-fluxes we include the $u'w'$-term (we have clarified this in the revised manuscript). Hence, for the momentum balance calculation the problems are probably dominated by errors in the reanalys data. Conversely, the eddy heat flux proxy calculation may be subject to errors in the method - i.e. since the calculation ignores where the waves dissipate, it is possible that the implicit stationarity assumption (i.e. that the typical dissipation pattern does not change over time) is violated (for example, the eddy heat flux trends are latitude dependent; shown in Figure 8 of the revised manuscript). We intend to analyse this problem with primitive equation model calculations, but much more work is needed and we feel that this is beyond the scope of this paper (i.e. this paper's contribution is to show that both eddy heat flux proxy and momentum balance calculation can be used to study the contribution from dynamics to the time series of lower stratospheric temperatures, but that due care is required with respect to trends). In addition to the potential problem in the method (which we don't discuss much), the revised manuscript now shows clearer that there exists an uncertainty in the data - namely the mid/low latitude eddy heat fluxes differ between ERA-Interim and MERRA, and this difference is responsible for the difference in the dynamically forced trend between the two reanalyses according to the eddy heat flux proxy.

R3-3: *2. Would it be possible to show seasonal trends also for MSU data for comparison with Figs. 2, 5 and 6? It is noted at several places in the text that 70hPa temperatures are strongly constrained by MSU data. On the other hand MSU data have a relatively broad weighting function and may be influenced by upper tropospheric changes.*

We prefer to refrain from showing MSU-4 temperatures in more detail exactly because of the deep vertical scale of the measurement.

R3-4: *3. Tropical mean temperatures, integrated heat fluxes and the momentum balance calculations all use different bounding latitudes: temperatures are integrated over 35N/S, heat fluxes starting at 25N/S and the momentum balance at 30N/S. Although it is noted in the text that results are very similar for different latitudes, it would be reassuring if the authors could show that the long-term trends are not affected by different choices of the latitudes.*

We have compared the temperature trends for 30S-30N versus 35S-35N, and the difference is marginal. In the manuscript we only provide the numbers for the 35S-35N latitude belt to remain consistent with latitude belt chosen to establish the correlation; adding the numbers for 30S-30N might unnecessarily distract the reader. We have considered to take the 30S-30N latitude range instead (which would be consistent with the latitude belt for the momentum balance calculation), but then one would have to go into some longer discussion of why the correlations are lower than they really are. Since temperature trends are not sensitive to 30 versus 35degree bounds, we chose the 35S-35N band.

The momentum balance calculation refers to 30S-30N, but as said since the actual temperature trends are almost identical, this is not a problem.

Finally, for the eddy heat flux proxy it does not really matter because that relation is basically not much more than an empirical relation. The revised manuscript now shows in more detail the sensitivity of trends in the eddy heat flux proxy to choice of bounding latitudes (this is the new Figure 9).

R3-5: *p.13390, l.7-16: It is reassuring that the slope is close to 1, but I am not fully convinced that the implication is that changes in the equilibrium temperature are small. Is the 70 day radiative relaxation time derived from an independent analysis, or empirically derived from the fit? In other words: can changes in the equilibrium temperature be hidden by an effective relaxation time? Anyway it would be good to spell out explicitly which "two time series" (l.12) are fitted here.*

Yes, the reviewer is correct that this statement was not carefully worded, and we have changed the text (now on page 8, line 13ff). More can be said about this aspect, but for the purpose of this paper we decided it is better to simply present results for the best fit for τ.

R3-6: *Section 2.5: It was not fully clear to me what you do to "filter" or remove the QBO influence. My understanding after 2nd reading is, that you average tropical temperatures over 35S to 35N, as this minimizes the influence of the QBO. Is this right, or do you do any other fitting or filtering to remove the QBO signal? I suggest to spell this out more clearly. Anyway there is much overlap and redundancy between Section 2.5 and Appendix A1. I suggest to combine the information and remove Appendix A1.*

We meant to say simply that we average 35S-35N (which is also a filter, albeit not a very sophisticated one). We have followed the reviewer's advice and moved the information from the appendix to the main text.

R3-7: *p.13393, l.28: Again, does the difference in latitudes (30N/S for the momentum balance, 35N/S for the temperatures) play a role here?*

No, this does not make a difference. We have re-written the program that plots the time series and calculates the correlations (Figure 3 and Tables 1-3 of the new manuscript), and to my (sf) embarrassment there was a small glitch in the original code; the correlations are actually very similar for the eddy heat flux proxy and the momentum balance calculation (see table 2 of the new manuscript).

R3-8: *p.13394, l.15: Can you expand the discussing of how/why the two calculations give imperfect estimators of the true dynamical forcing? Which* *terms (processes) are neglected in the momentum balance? What is the theoretical relationship between extra-tropical heat flux and tropical temperatures?*

The short - and arguably unsatisfactory answer - to this suggestion is that so far we have not found a single obvious source for the problems; see also our response to question R3-2. For example, we suspect that errors in the momentum balance calculations are dominated by drifts in the reanalyses due to changes in the input data, but in the absence of a 'gold standard' it is almost impossible to show that there are drifts (in Section 3.3 we present evidence for a drift in eddy heat fluxes of ERA-Interim relative to MERRA, but the argument that it is ERA-Interim rather than MERRA that drifts is based essentially on an educated guess - one would have to re-run ERA-Interim (not possible for us!) without assimilating GPS temperature data to prove/disprove our hypothesis). The situation is really unsatisfactory, and there is good reason not to trust trends in reanalyses. However, as also argued in the paper, reanalyses still represent the best available information, and we believe that our paper provides a fair description of what is robust, and what not (and in doing so also contributes to put previous work into context).

On the upside, the revised manuscript is now more specific about the sources of differences, and follow-up studies with a focus more on stratospheric dynamics (rather than the evolution of the time series) may use this information to address the problems presented in this paper.

R3-9: *p.13394, l.28: I suggest to split this sentence in two, as these are two different ideas.*

Done (i.e. second part dropped in this context; now on page 14/line 9).

R3-10: *p.13395, l.25: Better say larger contribution to the trends than contribution to the forcing.*

This text has been removed.

R3-11: *p.13395, l.26: "Positive trend": this is slightly confusing. Do you mean a positive trend in the dynamical forcing, that leads to a negative trend in tropical temperatures as seen in Fig. 5?*

Yes, that was the meaning of the sentence; this text no longer exists in the revised manuscript. (We have paid attention in the revised manuscript to make it clear whether we refer to trends in dynamical forcing, or dynamically forced temperature, but there probably remains potential for confusion due to the sign difference of the two.)

R3-12: *p.13399, l.24: what is area weighted here? I suppose it is an area weighted integration of the slopes, not an area weighted slope, that is* *integrated.*

Yes; we have removed "area weighted" as this should be self-evident. (This text is now on page 11, line 5.)

R3-13: *p.13400, l.19: I am confused here about the sign of the correlation. What is correlated here, the dynamical forcing and the tropical temperatures (which are generally anti correlated with colder temperatures for larger extra tropical dynamical forcing), or the temperature with the temperature proxy (which should ideally be positively correlated)? This applies also to other places in the manuscript. Clearly spelling out which time series are considered would help.*

The revised manuscript is hopefully clearer (see also response to question above).

R3-14: *Fig. 6: You may include in the caption to panel (a) that the trends are for 1980-2011.*

This figure has been removed.

R3-15: *p.13388, l.20: "a coefficient" p.13389, l.16: "Te" should be "TE" p.13392, l.7 "are removed"*

Corrected, or text changed in revised version.