Response to reviewer #1

General comments:

Overall: The authors have made an effort to address my comments, but I still think the paper is not suitable for ACP. In particular, the paper is poorly organized; the presentation could be improved; and results and/or methodology are possibly questionable. Examples follow:

Organization: For example, the information presented in the Introduction. Among other things, it is not clear to me from the Introduction why this work is important and what it will bring to the scientific discussion. In general, it is not straightforward to follow the paper.

A: Introduction was shortened in line with request of reviewer #2 and partly re-done.

Presentation: The tables are overly complicated and difficult to follow.

A: We tried to simplify Tables but we did not succeed to do it without loss of information, therefore we keep Tables as they were. We are sorry for that.

Questionability of results and/or methodology: (i) The title suggests a discussion of mid latitude phenomena, but this is based on the range 49N – 56N. (ii) I understand most of the trends results in Table 1 are not significant at the 99% level. Why show them? (iii) The response from the authors indicates that the cal-val of the reanalyses is done at 3 stations; the text just mentions one station. I do not find this a satisfactory way of evaluating the quality of reanalyses. But maybe I have misunderstood what the authors are doing regarding evaluation of the reanalyses.

A: ad1) Title is modified to "higher midlatitudes" and some results from 42.5 and 62.5^{o}N are added (new Figure 2).

ad2) The standard significance level for analyses in meteorology (wind, temperature, etc.) used by majority of authors is 95%. That is why we prefer using this threshold. When the results are predominantly significant at the 99% level, like in Table 3, the 99% significance is highlighted. This was added in section 2 Data and methods.

ad3) According to Kozubek et al. (2014), stratospheric winds from the NCEP/NCAR reanalysis reveal better results compared to ERA-40 and ERA-Interim. Nevertheless, this part of Introduction was deleted (as recommended by reviewer #2) during shortening of Introduction. Evaluation of reanalyses was more complex, for instance it was shown that compared to ERA-Interim and NCEP/NCAR, EAR-40 winds at 10 hPa over the last four winters of EAR-40 period are wrong – however, this discussion does not concern the paper in question, it concerns already published paper Kozubek et al. (2014).

Before the paper is suitable for ACP, the authors should address the above issues, as well as the specific issues below. The English should also be looked at.

Specific comments (not exhaustive, but illustrative):

P. 1

L. 12: Use reanalysis data to do what?

A: Deleted from abstract.

L. 13: Why do the reanalyses data provide a more consistent dataset than observations?

A: Deleted from abstract.

L. 17: Why is this result the most important one?

A: This is discussed in the paper, no space for reasoning in brief abstract.

L. 20-21: Use of “appears” here (and elsewhere) is vague.

A: Can you please suggest us which word to use?

P. 2

L. 38: What is the importance of the UTLS, e.g., regarding the atmospheric circulation?

A: Deleted in shortening of Introduction.

L. 51: Make sure you introduce acronyms when first used – both in the abstract and in the main text, e.g., GHG.

A: We apologize, GHG is now spelled out.

P. 3

L. 61-62: This should be mentioned toward the beginning of the Introduction. L. 69: Avoid “chatty” language in the text.

A: Introduction was shortened and partly rearranged.

P. 5

L. 115-121: This is confusing to me. What is important in this paper? How are you addressing key questions? The Introduction needs rewriting.

A: Introduction is about introducing the problem and what is the paper about. How we address questions is written in Section 2 and 3, maybe 4.

L. 128: You use the reanalysis to do what?

A: We use reanalysis for our study. It is added to the text.

P. 6

L. 138: Evaluation of the reanalyses is against one site mentioned here (and a total of 3 according to the response of the authors). I do not find this satisfactory.

A: That part is much reduced. According to Kozubek et al. (2014), stratospheric winds from the NCEP/NCAR reanalysis reveal better results compared to ERA-40 and ERA-Interim. Evaluation of reanalyses was more complex, for instance it was shown that compared to ERA-Interim and NCEP/NCAR, EAR-40 winds at 10 hPa over the last four winters of EAR-40 period are wrong – however, this discussion does not concern the paper in question, it concerns already published paper Kozubek et al. (2014).

L. 153: The latitude band is 49N – 56N, a range of 7 degrees. Is this representative of mid latitudes?

A: *Title is modified to "higher midlatitudes" and some results from 42.5 and 62.5°N are added (new Figure 2). The text is modified; you are right that such a relative narrow band need not be sufficiently representative for middle latitudes.*

P. 7

L. 170: Significance calculated at 99% only for some cases? Why? According to the author response, most of the trends in Table 1 are not significant at 99% (but are at 95%). Why show the Table 1 data?

A: The standard significance level for analyses in meteorology (wind, temperature, etc.) is 95%. That is why we prefer using this threshold. It was added in Section 2. See also response (ad2).

P. 9

L. 216: What do you mean by "some"? A small percentage?

A: We have changed some. Now it is "the". We would like to express that statistically significant trends occur mainly in the Atlantic sector.

P. 10

L. 243: Only 4 trends (out of 192, I understand) are significant at the 99% level. And as mentioned by the authors, this is likely due to the limited length of the datasets. Which begs the question, why calculate trends with a dataset limited in length?

A: Well, any dataset is limited in length. The longer data series is not available and reliable (before 1970). We started in 1970 and we split this period in the mid-1990s to see possible impact of overturning of stratospheric ozone trend, which does exist. The significance level of 95 % is usually used in meteorology.

P. 13

L. 325: Your results only concern a small range in mid latitudes (unless I have misunderstood something).

A: The text is modified. Our results on trends in total horizontal wind concern higher middle latitudes (49°N-57°N) with focus on European sector, where trends are best pronounced. The results for 42.5 and 62.5°N reveal essentially no detectable trend. On the other hand, longitudinal distribution of meridional wind is constructed for all longitudes and latitudes 20-60°N.

L. 330: What is the significance level of the trends discussed?

A: The significance level is 95 %. Now in the text.

P. 22

Table 1: Why show this table if significance at 99% only occurs for 4 cases?

A: The standard significance level for analyses in meteorology (wind, temperature, etc.) is 95%. That is why we use this threshold. It was added in section 2 Data and methods. Data series are probably too short for getting significance at 99% in noisy atmosphere but due to change of trends in the mid-1990s they cannot be longer.

Tables 1, 2: By total wind speed you mean the absolute magnitude of the horizontal wind component?

A: It is exactly as you comment. We mean total horizontal wind. It was corrected in the text, Figure 1 and 2 captions, and description of Tables 1 and 2.

P. 23-24
The Tables shown look complicated and difficult to follow. Is it possible to simplify their design?

A: *We tried to simplify Tables but we did not succeed to do it without loss of information, therefore we keep Tables as they were. We are sorry for that.*

P. 25

Fig. 2: Where is the dashed line?

A: It was a mistake; old version instead of present version was submitted. A new figure has been added and dashed line is now on it.

P. 26+

Figs. 3-6: I suggest that the authors indicate what the end points of the colour scale mean, e.g., red positive values, blue negative values.

A: Description and clarification for original Figs. 3-5 (now 4-7) have been done.

Response to reviewer #2

The authors have made a few improvements to their manuscript but several major comments made in the previous round of reviews have not been addressed. I do not believe the improvement have been substantial enough to reach the high quality requirements for publications in ACP at this time. While the paper presents some interesting and relevant results, the manuscript needs further revisions in order to be considered for publication in ACP. These revisions would essentially require improving/expanding the analysis, reorganizing the manuscript and improving the discussion section. These are major revisions, but the manuscript would greatly improve in quality. I have listed several major revisions (by order of the paper not importance) that the authors need to make before I can recommend their manuscript for publication:

-The introduction was clearly improved. It remains however quite long (around 1200 words or a little more than 25% of the paper) so it should be more concise. It also generally lacks proper organization. I would suggest aiming preferably for closer to 15% (or around 800 words). For example, I do not see how the paragraph on reanalysis products (line 90 to 100) is relevant to the introduction: analyzing reanalysis data for stratospheric dynamics studies is certainly not novel and most people are familiar with reanalysis products and their general shortcomings (at least, this might be more relevant in the methodology section).

A: Introduction has been shortened (from 1160 to 750 words) and it is now more relevant to the paper. The paragraph on reanalysis products has been deleted as recommended.

-I am not convinced by the explanation for using only one particular re-analysis dataset, in this case the NCEP/NCAR.

• First, the Kozubek et al. (2014) study only compares to three observation stations and only after 1989 (only Prague-Libus is mentioned on line 138). I don't see how it is possible to generalize the quality of (or even rank) the three reanalysis products based on so few samples.

A: The reanalyses were compared mutually and as additional information with Prague station. Nevertheless this discussion concerns already published paper. The text is reduced and modified.

• Second, the authors claim that the data is more reliable from 1957 onwards without citations to back it up. I am not sure I would consider reanalysis data before 1979 as "reliable", especially in the stratosphere.

A: Wind data used by Kozubek et al. (2014) do not display a jump in 1979, data before and after are mutually consistent. We are not interested in fine structures, we analyse gross features (trends). We would not expect such consistency at the Southern Hemisphere with much less ground-based stations.

• I think it would be preferable to simply write that you chose the NCEP/NCAR reanalysis, even though you could have chosen others (including the JRA55 data, the NCEP/DOE reanalysis 2…). That this reanalysis product shows reasonable agreement with observations (for example, based on Kozubek et al., 2014), cite some studies that have analyzed the representation of the QBO, SSW… in the NCEP/NCAR reanalysis, an state as a limitation the fact that this study uses only one reanalysis and thus could lead to different results if using another reanalysis product.

A: In the revised paper the meridional wind longitudinal distribution analysis is done for three reanalyses (NCEP/NCAR, ERA Interim and MERRA) and all three reanalyses show very similar results in main features. That is why we keep NCEP/NCAR reanalysis for other analyses (the longest available data series). In conclusions we mention that most analyses are made for NCEP/NCAR only and that the results based on other reanalyses may partly differ (even though we do not expect qualitative differences, only quantitative differences).

-I am not convinced by the explanation for the choice of the three latitude bands.

- The authors claim that the lack of ocean makes the data "better than in majority of other latitudinal band" without any citation to back up that claim. I certainly don't think it would be an issue after 1979, since satellite data is assimilated. But if it were an issue before 1979, it would make sense to simply start the analysis in 1979 then and analyze a larger latitudinal band. In addition, since the authors analyze longitudinal sectors, some with more ocean than the others, if the ocean does indeed lead to a degradation of the quality of the data and thus of the analysis, that could be a very worrisome issue for this study (although I don't believe it is the case).

A: *Title is modified to "higher midlatitudes" and some results from 42.5 and 62.5°N are added (new Figure 2). Ocean could play a role only in pre-satellite era; our results do not indicate substantial differences related to ocean but some influence cannot be a priori excluded.*

- I think it is necessary that the analysis be extended to more than the three chosen latitudes (see further down). This comment was already made during the first round of review but not addressed by the authors.

A: We have analysed the total horizontal wind in 42.5°N and 62.5°N as well. We do not want to do latitude averages because it could affect the results. The results for 42.5°N and 62.5°N do not show such a pronounced trend as it is the case for higher mid-latitudes. We decided to do other analyses only for higher midlatitudes (except for longitudinal distribution).

-The biggest flaw of the manuscript is its lack of consistency in three different aspects of the analysis: 1) <u>The region of study</u>: three latitude bands for the study of trends, QBO, SSW and solar cycle (only one for the NAO impact) but the mean of the 20-60ºN for the analysis of the longitudinal distribution of the meridional wind, the analysis of tide and geopotential height. It is unclear to me how the authors can make general concluding statement of the impact of NAO, QBO… on stratospheric winds based on just three latitudes (that are next to each other, thus only covering a 7.5º band). This issue was pointed out in the previous round of review but not addressed by the authors.

A: *Title is modified to "higher midlatitudes" and some results from 42.5 and 62.5°N are added (new Figure 2). The results for 42.5°N and 62.5°N do not show such a pronounced trend as it is the case for higher mid-latitudes. Thus the pronounced trend is confined to latitudes of about 49-56°N (maybe a bit more) and to Atlantic sector, which suggests a role of NAO in trends. We do not want to do latitude averages because it could affect the results. Therefore we decided to do other analyses only for higher midlatitudes. For longitudinal distribution analysis we need larger range of latitudes (we use 20-60°N) due to latitudinal size of longitudinal structures (two* *cores at 10 hPa).*

2) The total wind speed versus meridional wind: the analysis of the longitudinal distribution of wind is focusing only on the meridional wind, even though it only contributes a small part to the total wind speed. There is some inconsistency in several places: the authors analyze the trends in total wind speed in Table 3 to identify trends in the " two-core" pattern of meridional wind even though the trend in the total wind might not be related at all to the meridional wind (it is most likely driven by changes in the zonal wind). The authors also state that Monier and Weare (2011b) "found some strengthening of northern jet in October-December and weakening in January-March" but that the authors' analysis of "meridional wind at northern higher middle latitudes reveals similar trends for October-December and January-March". The jet is mainly driven by zonal wind, not meridional wind, so the analysis of Monier and Weare relates to changes in zonal wind, not meridional wind. The authors should analyze the total wind separately between Oct-Dec and Jan-Mar, not the meridional wind… It is also unclear how to link trends in total wind speed to the meridional wind component.

A: It was a mistake in description of Table 3. This table shows trend of meridional wind, not total horizontal wind.
We would like to focus on meridional wind because this part of total horizontal wind has not been studied in previous studies. Moreover, generally meridional wind contributes only a small portion to the total horizontal wind but this general statement is not valid everywhere. Due to two core structure of meridional wind at 10 hPa, which is probably evoked by distribution of geopotential height, for example at 60^oN, 135^oE the zonal and meridional wind are equal each other, about 20-25 m/s each (see Figs. 4 and 9). So it makes sense to study the longitudinal distribution of meridional wind. This study also documents limitations of zonal mean wind analyses.

3) The authors claim that the "well-pronounced longitudinal structure of wintertime meridional wind at 10 hPa to be the most important result" of their study. Yet, because the trend analysis is done before the analysis of the longitudinal distribution of meridional wind, the authors do not really discuss how this structure has evolved in the last few decades (see point above about inconsistency between analysis of meridional wind and total wind speed).
This issue was pointed out in the previous round of review but not addressed by the authors.

A: The longitudinal structure of meridional wind is now presented for NCEP/NCAR, ERA-Interim and MERRA, which provide very similar result. We explain the 2-core structure by a well-pronounced blocking Aleutian anticyclone at 10 hPa. Zonal wind behaviour in new Fig. 9 fully confirms this explanation. Trends in peak regions of two cores are presented in Table 3 – they are opposite before and after the mid-1990s.

I suggest the authors make the following revisions to their manuscript in order to bring it to the high quality requirements for publication in ACP: Present the analysis of the longitudinal distribution of the meridional wind (section 3.3) first, and add to it an analysis of the longitudinal distribution of zonal wind (essentially redo Figure 3 and Figure 4 but for the zonal wind). Then present the trend analysis of the total wind speed (section 3.1), but on the average over latitudinal bands similar to section.

A: We prefer to keep the current structure as the trend and solar cycle effect analysis is done for the total horizontal wind at higher midlatitudes, whereas longitudinal distribution is analysed for meridional wind over substantially broader latitudinal range. However, if you still will request to change the order of subsections 3.1-3.3, we shall do it in the next revision.

3.3. Averaging over 20-60ºN might be too large, so 30-40, 40-50 and 50-60 might be more appropriate, following the same “three latitudinal band” structure of Figure 1, Table 1 and Table 2. You should then discuss how the trends affect the dipole pattern of meridional wind in more details, using Table 3. Finally present the role of the solar cycle on the trends (section 3.2), once again on the same latitudinal bands (30-40, 40-50 and 50-60). This would provide an improved flow and add much needed consistency in the analysis.

A: Instead of making analysis for suggested latitudinal bands, we added two other latitudes, 42.5 and 62.5^{o}N. Their results support focus on 49-56^{o}N.

-The authors mention that the reversal of trends in winds in the mid-1990s is “in accord with change in trend of ozone” on line 364. However, the authors state in the response to reviewers that “the total ozone trend change in the mid-1990s was caused mainly by dynamics (e.g. Harris et al., 2008), therefore role of factors like NAO, QBO etc. might be really important.” This last statement warrants a more complex discussion (and expanding analysis) on the role of ozone on the trends in total wind speed that the single sentence on line 364. What is the role of ozone versus NAO/QBO/SSW… in controlling the trends of total wind? No particular explanation is given on how ozone is affecting trends in the total wind speed, and in particular, why it only affects the Atlantic sector (as shown in Figure 1 and mentioned in line 187). The authors could, for example, plot the ozone concentration along Figure 2, showing that the different trends in wind speed are not only explained by NAO, QBO… but also by ozone concentrations. In any case, the discussion and analysis of the impact of ozone is seriously missing, and that’s a clear shortcoming of this manuscript.

A: *As concerns the role of ozone, this is rather indicator than driver of changes in wind. A paragraph about that is added in Discussion. Both sections Discussion and Conclusions are partly modified.*

-Generally, the manuscript is too descriptive (Fig X shows that such and such is occurring) but lacks explanations (such and such is occurring because of this potential mechanism…). There are a lot of questions that are neither answers nor even raised: By what mechanism are the NAO, QBO, solar cycle impacting trends in the wind speed? What is causing the trends in wind speed to start with? Is it changes in ozone? Is it decadal variability in NAO? Why would the phase of the QBO, solar cycle, impact these trends? Are NAO trends driving ozone changes? Are ozone changes driving NAO changes? How? Why are there different trends pre-1995 and post-1995 in NAO if the trends in ozone are driven by dynamics? Why is there also a reversal of trends in 1970 (based on Figure 2)? The manuscript would benefit significantly from a better discussion/conclusion section, instead of just simply presenting the results without any real discussion. This issue was raised in the first round of review but not addressed by the reviewers (except to add a few general statements on the importance of ozone for stratospheric dynamics in the introduction).

A: We are afraid that full answer to your above questions would need several years of investigations by a broad team and a book, not a paper. We add some information in Discussion, particularly as concerns the role of ozone, which is rather indicator than driver of changes in wind.

Minor comments:

Figure 1: the top panel (100 hPa) starts in 1968 while the bottom panel (10 hPa) starts in 1958. Both should start at the same time. In addition, the difference between sectors is difficult to see. The different markers should be replaced by different colours instead.

A: We apologize, it was a technical mistake. It is corrected, now both panels cover the same period. The sectors are displayed in different colours.

Figure 2: The artificial discontinuity between the three time periods should be removed. Please plot a single continuous line. This was a comment in the previous round but was not addressed by the authors.

A: We apologize, by mistake we included old version of Fig. 2. The artificial discontinuity between the three time periods is now removed.