Referee 1

Comments on the ACPD manuscript entitled “Variability of water vapour in the Arctic stratosphere by Laura Thölix, Leif Backman, Rigel Kivi and Alexey Karpechko.

The present paper discusses the variability of water vapour in the Arctic stratosphere. For that mainly simulations from FinROSE and ECMWF are used. In addition also observations from Aura/MLS, frost point hygrometers and CALIPSO are considered. The manuscript is quite comprehensive in the meaning that it touches upon a wide range topics that concern Arctic water vapour. This comprises water vapour variability in general, but focuses also on the winter season and polar stratospheric clouds. Particular in that sense it is an interesting manuscript. My main criticism is that it just touches upon all these interesting topics, scratching at the surface, many details and discussions are missing. There is much potential and I can easily see the manuscript being split in two, as there is so much interesting stuff. For time being I go along with major revisions. Please find my detailed comments below:

We thank the reviewer for his/her encouraging and detailed review. Following this review we have substantially revised the manuscript providing more insights into the processes behind water vapour variability in the Arctic stratosphere.

Comments:

• Abstract in general: There are a few abbreviations here, but none of them is defined. As this is typically the first encounter I prefer to have them here already. Later in the manuscript there some abbreviations that remain undefined. This should not be the case. Abbreviations in the abstract have been defined

• page 1, line 5 to 7: I presume you are referring here to the top panel of Fig. 4. That is the only time observations are involved at where the timing matches. This should be made clear.
  Yes, this part refers to Fig. 4. We have rewritten the sentence to make the message clearer:
  ‘A FinROSE chemistry climate model simulation covering years 1990–2014 is compared to observations (satellite and frostpoint hygrometer soundings) and the sources of stratospheric water vapour are studied. In the simulations, the Arctic water vapour shows decadal variability with a magnitude of ~0.6 ppm. Both observations and the simulations show an increase in the water vapour concentration in the Arctic stratosphere after year 2006, but around 2012 the concentration started to decrease.’

• page 1, line 9 to 12: There are two sentences in a row that convey almost the same information. In the first sentence the message comes across more uncertain or speculative though.
  The first sentence have been taken off.

• page 2, lines 35 and 36: Somehow the sentence concerning NAT clouds does not fit the natural flow here.
  The sentence have been reformulated with the previous sentences

• page 2, line 41 to 44: You could add a reference to the model evaluation study by
Gettleman et al. (2009) that shows the simulated change of the tropical tropopause. Also increasing methane could be added.

We added a sentence: 'However, stratospheric water vapour is expected to increase due to climate change caused warming of the tropical tropopause (Gettleman et al., 2009) and to an increase in the atmospheric methane concentration.'

- page 2, line 51: "... is therefore controlled by ..." - Somehow I would squeeze in here "largely" or "to a first order" as there are other processes that can contribute on different scales.

Done, the sentence is now: "The variability in the entry of water vapour into the stratosphere is largely controlled by the variability in the tropical cold point temperature."

- page 3, line 61 to 63: Satellite measurement across the tropopause are challenging, not only for water vapour. But this is not the reason for why there are no long-term observations.

The sentence have been rewritten and is now:

The exceptional dryness of the stratosphere makes observation of stratospheric water vapour challenging. Long-term time series of stratospheric water vapour are rare, which complicates the study of concentration trends.

- page 3, line 69: "... continuous ..." - This is really a question how do you define that. If you use the daily coverage of Aura/MLS as reference, then UARS/MLS was not as continuous. Typically there was coverage between 34° on one hemisphere and 80° on the other hemisphere, switching roughly every five weeks by a 180° yaw manoeuvre of the spacecraft. Hence in the tropics and subtropics there was "continuous" (daily) coverage while at mid- and high latitudes this was not the case. The balloon measurements at Boulder are typically performed only once a month, just to give another example.

- page 3, lines 69 and 70: "... since 2004 ..." - This only applies for Aura/MLS but not for UARS/MLS.

The sentence is modified, word 'continuous' has been taken off.

'Additionally, global data is available from satellite instruments, but only for a limited time span, for example from Microwave Limb Sounder (MLS) on board the Upper Atmosphere Research Satellite (UARS) (1991—1993) and the Earth Observing System on Aura (EOS-Aura) (2004—today) allows a monitoring of the northern hemispheric stratospheric water vapour distribution.'

- page 3, line 70 to 72: In the list also POAM III (1998 - 2005), SAGE III (2001 - 2005) and SOFIE (since 2007) are missing. POAM and SOFIE focused actually only on the polar regions. MIPAS and SMR (at least in boreal winter) had/have coverage all the way to the poles, while MLS only goes to 82° latitude.

POAM III, SAGE III and SOFIE have added to the list.

- page 3, line 87 to 89: There is more to these negative trends as they are related to the sudden water vapour in 2000 (e.g. Hegglin et al., 2014). After a few years with very low water vapour mixing ratios a recovery started in 2005. In 2011 again a substantial drop was observed, but this one was more short-lived (Urban et al., 2014).

The sentence in the manuscript have been modified and references have added:
Based on satellite and sounding measurements e.g. Solomon et al. (2010) reported negative trends in Boulder and generally in the mid-latitudes at 18 km altitude between 2000 and 2009, while e.g. Hegglin et al., (2014) showed that these negative trends are mainly related to the sudden drop in the water vapour in 2000 and that after a few years with very low water vapour mixing ratios a recovery started in 2005. Recently Urban et al. (2014) reported another drop in the tropical water vapour during 2011-2012.

page 3, line 89 to 91: The study by Hegglin et al. (2014) should me mentioned here. Even though there is no focus on polar latitudes, they show at least water vapour trends up to 80° latitude for the time period between the late 1980s and 2010. The reported trends in the lower stratosphere are actually negative. On the other hand one should acknowledge that the coverage before 1998 using HALOE and SAGE II was far from optimal and that some caution is warranted. Following sentence have inserted:

Hegglin et al. (2014) show water vapour trends up to 80° latitude for the time period between the late 1980s and 2010. On the other hand one should keep in mind that the coverage before 1998, which was based on HALOE and SAGE II, was not good, which warrants some caution to the results. The reported trends in the lower stratosphere are negative.

page 4, line 108: “ctm” - Does this refer to CTM = Chemistry Transport Model? Corrected

page 4, line 117: For me prescription of the number density profile for the individual PSC types seems like a profound restriction. Has that been quantified? The PSC scheme is not a microphysical scheme, but it is based on thermodynamic equilibrium equations. In this case there are basically two choices, either prescribe the particle size or the number density. We have chosen the latter: i.e. when the temperature decreases there becomes more particles, in the other case the particles would become larger. In both cases the sedimentation increases and the surface area increases. The number density profile for LBA and STS is estimated from McLinden et al 1999. The number density of NAT particles is initially assumed to be 1 cm-3 (Krämer at al 2003), the number density is reduced for large NAT particles. The ICE number density is assumed to be 0.04 cm-3, estimated from synoptic scale PSCs (Dye et al. 1992). More detail on the PSC scheme were added to the text:

The composition of LBA and STS are calculated using the method by Carslaw et al. (1995). The STS are not considered below the ICE PSC formation temperature. The number density profile for LBA and STS is estimated from McLinden et al (1999) and the sulphuric acid distribution [μm² cm⁻³] is based on 2-D model data Bekki and Pyle (1992). NAT formation is based on the thermodynamic equilibrium equations by Hanson and Mauersberger (1988). The model includes an option to include a supersaturation requirement for NAT and ICE formation, but this option was not used in the simulations reported in this paper. The choice was made due to the relative modest resolution of the model. Co-existence of NAT and STS is allowed. A scheme for growth of NAT particles is included based on Fahey et al (2001). The number density of NAT particles is initially assumed to be 1 cm-3 (Krämer at al 2003), the number density is reduced for large NAT particles. The temperature threshold for ICE particle formation is based on expressions by Marti and Mauersberger (1993). The equilibrium pressure of nitric acid above ICE is calculated according to Hanson and
Mauersberger (1988). The ICE number density is assumed to be 0.04 cm⁻³, estimated from synoptic scale PSCs (Dye et al. 1992).

- **page 4, line 123 to 125:** Where is the boundary for the prescribed tropospheric water vapour? If it is too close to the tropopause, in particular in the tropics, you may get a dry bias due to the cold bias of ECMWF there. A more general question at this point regards the complexity of the methane oxidation scheme? Depending on that you may get different trend estimate, in particular in the upper part of the stratosphere were methane oxidation is more effective. Does FinROSE include water vapour production from the molecular hydrogen reservoir (Wrotny et al., 2010)? Methane is oxidized by OH, O₁D, Cl and photodissociation is also included. The intermediate formaldehyde is also included as well as molecular hydrogen. The text was revised...

  ‘The tropopause height is calculated at every time step using potential vorticity as defining parameter. Model levels below +2 PVU are considered to be in the troposphere. The 380 K potential temperature level further is used to define the tropopause height near the equator. The tropopause is thus changing with time depending on meteorological conditions. The tropospheric concentrations of the chemical species are not calculated in the model but prescribed via model boundary conditions. Tropospheric water vapour and ozone were obtained from the ECMWF ERA-Interim reanalysis (Simmons et al. 2007, Dee et al. 2011).’

- **page 5, line 138 to 140:** This sentence seems to imply that satellite measurements are not accurate, which I would definitely argue against. Please rephrase. Word accurate have been changed to ‘High resolution sounding’

- **page 5, line 143 to 159:** This whole part needs a better structure. There is some jumping between campaigns and instruments; back and forth. Maybe a summary table with the two campaigns and the relevant instruments could be helpful. CFH observations are mentioned for both campaigns but then only used for the second campaign. Why? In the paragraph following these lines Aura/MLS and CALIPSO are described, but they only contribute to the second campaign. That should be made clear.

  We agree that the part needs to be rewritten and we will provide a better version. In the paper we have used measurements from both campaigns and in addition CFH observations in between the campaigns. In this paper we have only used the CFH observations, although also the FLASH measurements have been made (Khaykin et al., 2013).

- **page 6, line 176 to 178:** Given this statement here, the section description in the final part of the introduction and earlier experiences of mine with ECMWF water vapour I wondered a couple of times along the manuscript why ECMWF water vapour is shown at all. I have to admit that compared to earlier incarnations of this data product the current water vapour data set looks relatively fair but you should have in mind that is based on a simple methane oxidation parametrisation and relaxes to 6 ppmv at the stratopause. Overall, for this analysis here, I do not see the value and would focus more on the observations.

  We fully agree with the reviewer that the quality of the ECMWF water vapour distribution suffers from a too simplified treatment of chemistry. However we see no reason to conceal this from the readers and we’d rather keep it in the figures. Note
that the reasons for the discrepancy between ECMWF and observations are discussed in the text. Moreover we believe the transport in ECMWF is well represented. Therefore any trends arising from changes in the transport would be captured by ECMWF even if absolute values are biased.

- page 7, line 205 to 207: I guess the standard deviation is simply derived from all the profiles that fitted your coincidence criteria. Maybe the approach could be made clearer. The standard deviation can be quite tricky to interpret. For the observations measurement and retrieval characteristics contribute to the standard deviation (it is not all natural variability); for the simulations the model setup plays a role. That should be kept in mind!

  Thank you for this comment. Yes, the standard deviation is calculated across the individual profiles. To make it clearer a corresponding statement is added to the revised manuscript:

  ‘…From MLS all the profiles measured in the Sodankylä gridpoint and flagged as good quality are used. Standard deviation is calculated across the individual profiles for each data set and thus represents uncertainty due to natural variability and random errors. Figure 2 compares January–February-mean water vapour…’

- page 8, lines 252 and 253: Why does the time axis in Fig. 4 starts first in 1994?

  Figure 4 (and new figure 5) have redrawn for years 1990-2014 and text about it have been modified. The water vapour tracer has recalculated using better upper boundary conditions, and the upper levels in this figure are now better and can be shown. (The upper boundary condition of the tracer is now always 65% of water vapour. The value is the average of the upper level in the long simulation.)

- page 8, line 257 to 260: This discrepancy between the simulations and observations certainly warrants more attention. Effects of the coarser altitude resolution of Aura/MLS compared to the simulations certainly, I think, can be neglected here. The only exception in general would be if there are dehydration features or if you would look closer to the hygropause. But 56 hPa seems fine given a typical MLS altitude resolution of about 3 km for water vapour in this altitude range. Also in the figure it looks like as there may be a time delay between the simulations and the observations that may hint on different transport time scales.

  We agree with reviewer that the time delay suggest different transport time scales. We also note that the water vapour increase in FinROSE and ERA-Interim, as well as the discrepancy with MLS, is most pronounced in the lower stratosphere, while the difference is smaller in the upper stratosphere. This suggest that the difference maybe associated with transport processes within the stratosphere as represented by ERA-Interim during this period. In the revised manuscript we discuss the discrepancy between the simulations and observations in more detail.

- page 8 and 9, line 270 to 288: I definitely think there should be more analysis and discussion of the variability here. After all it is the main key topic of the manuscript. The analysis may involve a thorough regression analysis. In terms of discussion there is more than these linear short-term changes (the term trend feels somewhat exaggerated), like those drops in water vapour (e.g. Randel et al., 2006; Urban et al., 2014), influence of sudden stratospheric warmings (e.g. Straub et al., 2012; Tschanz and Kämpfer, 2015) or the QBO that shows up quite prominently. Also considering more the Aura/MLS measurements would be great.
Following this comment, MLS anomalies have been added to the figure. We have also performed regression analysis where we considered influences on the Arctic water vapour of BD circulation, QBO as well as tropical tropopause temperatures. Please see more discussion on this below. We have also tried to identify the impact of sudden stratospheric warming on the water vapour based on our timeseries. However one should note that we consider monthly mean values whereas analysis of SSW require daily values. Therefore we cannot detect any clear signal of SSWs in our long term timeseries based on monthly values.

- **page 9, lines 277 and 278:** After the drop in 2000/2001 there was a recovery afterwards.
  Thank you for this specification. We reformulated the sentence as follows:
  ‘Solomon et al. (2010) found a positive trend in the water vapour data until about year 2000 and a negative one after that over Boulder at the mid-latitudes around 80 hPa. The negative trend was mainly caused by a drop of the concentrations in 2000/2001 (Randel et al. 2006) which was somewhat compensated by a slower recovery afterwards.’

- **page 9, line 278 to 280:** “As for ...” - Somehow this does not sound the right way. The implication by Hegglin et al. (2014) that observations at Boulder should not be generalised in to describe the global stratosphere has in my opinion two sides. There could be local effects that would make this location different from the global behaviour. However, for the time being, it seems that there is more a discrepancy between two sets of data, i.e. the data from the frost point hygrometers and the satellite data merged using a nudged model. The last word has certainly not been spoken regarding this inconsistency.
  Thank you for this note of caution. We reformulate the interpretation of the Hegglin et al. result as follows:
  ‘However, Hegglin et al. (2014) showed that the water vapour trends over Boulder might be associated with local patterns of changes and not necessarily representative of the global stratosphere.’

- **page 9, line 283 and 288:** The QBO signal is more general due to its influence on the Brewer-Dobson circulation. That affects the tropical tropopause temperatures but also along the way there are variations of the transport. Mixing is certainly less important, except when the polar vortex breaks up.

Following these comments we have elaborated the discussion of the water vapour variability. We performed regression analysis following Dessler et al. (2014). We used three proxies: qbo index (QBO, equatorial winds at 50hPa), Brewer Dobson circulation index (BD, residual vertical winds at 70hPa averaged from 30S to 30N), and cold point temperature (CPT). Unlike Dessler et al. (2014) we found that the use of tropical temperatures at 500 hPa was not enough to explain the variability of the cold point temperature, and therefore used it as one of proxies. Although, there is some correlation between CPT and QBO (0.36) QBO also affect the transport of the water vapour not directly influenced by CPT; therefore the use of both proxies is justified. We apply multiple regression analysis with all three proxies to water vapour time series averaged north of 70N and at 82 hPa and 56 hPa. Cross-correlation analysis shows bread peaks at lags 6-12 months for the proxies. The maximum of the correlations of QBO and CPT with water vapour at 56 hPa is at about 10 months lag,
and with 82 hPa is at 8-9 month lags, suggesting that propagation of the tropical anomalies in the lower stratosphere is faster than that in the middle stratosphere, likely due to more efficient mixing. We use 10 month lag for all proxies for regression at 56 hPa and 9 month lag for the regression at 82 hPa.

The individual correlation coefficients with our proxies are shown in the Table below. The main contribution to the polar water vapour variability is CTP, followed by QBO. We found very weak contribution of BD proxy to the variability of the water vapour. One reason is that the effect of BD contribution is accumulated over time and this is not well represented by the monthly proxy. The multiple regression coefficients are 0.57 and 0.51 at 82 hPa and at 56 hPa correspondingly, showing that our models only explain 25-30% of the variability. This is considerably less than that of Dessler et al. suggesting that different processes contribute to the polar water vapour variability in comparison to those in the tropics. Note that the regression somewhat explains the increase of the water vapour from 2005 to 2010, which is more clear at 82 hPa. However the peak of the water vapor during 2011-2013 is not explained by these proxies.

Table: Correlation coefficients

<table>
<thead>
<tr>
<th></th>
<th>CPT</th>
<th>QBO</th>
<th>BD</th>
<th>Multiple</th>
</tr>
</thead>
<tbody>
<tr>
<td>56 hPa</td>
<td>0.454073</td>
<td>0.315499</td>
<td>-0.209597</td>
<td>0.51</td>
</tr>
<tr>
<td>82 hPa</td>
<td>0.518588</td>
<td>0.399678</td>
<td>-0.180337</td>
<td>0.57</td>
</tr>
</tbody>
</table>

*page 10, line 328 to 339: I seek more discussion on Fig 6. I am definitely not a PSC expert. Hence my first order expectation would be a linear relationship between the cold area and the PSC area. Hence the large cold areas with small PSC areas stick out for me. Why is that? I do not think that water vapour is the decisive factor here and there are reddish, greenish and orange colours visible there. The relationship between large cold areas and PSC areas accompanied by high water vapour seems more clear. A large cold area means you have a stable vortex where moist air from the stratopause is brought down to the lower stratosphere. This has been recently addressed by Khosrawi et al. (2015). In that regard it makes sense to me to look at the vortex average water vapour volume mixing ratio. Beyond that I really wondered why this was used given the more localised dehydration layers in connection with PSC. What temperatures do you use to get the area for T < 190 K for CALIPSO? At least I am not aware that CALIPSO provides temperatures.

Following this comment, and a related comment by R2, Figure 6 has been updated and more discussion has been added. Instead of using area with T<190K we use area with T<188K which is tighter related to ice PSC. We get correlations between temperature and ice PSC of 0.93 for the model and 0.72 for CALIPSO, consistent with expectations. To better illustrate the dependence of ice PSC on water vapour we add panel (b) and also support the dependence by providing mean values of ice PSC area conditioned on water vapour and temperature. In statistical sense the dependence clearly emerges, so that increase of water vapour leads to larger PSC. The correlations between water vapour and ice PSC is 0.21 in FinROSE and 0.41 in CALIPSO. Although we agree that the link between local water vapour and PSC should be clearer, the use of mean vortex water vapour is justified by the obtained correlation. It is important to demonstrate this relation in the context of the paper. In
the revised version we use ERA-Interim temperatures in connections with ice PSC from both FinROSE and CALIPSO. Water vapour is from FinROSE.

- page 11, line 346: "... the vortex moved to the south ..." - Really to the south? At least my simple expectation would be different. Or was this something filament-like?
- page 11, lines 346 and 347: "... mixing with moister mid-latitude air ..." - Under typical conditions, leaving dehydration aside, the water vapour mixing ratios in the lower stratosphere are lower outside the polar vortex than inside (e.g. Nassar et al., 2005; Lossow et al., 2009). If there is dehydration than things may be the opposite, but it is unclear if there was still dehydration by 11 February 2011. Thanks for noticing the inaccuracy. Meteorological conditions of winter 2009/2010 have discussed lot (e.g. Khaykin et al. 2013, Pitts et al. 2011, Dörnbrack et al. 2012). We added references to those studies, and changed these sentences as: 'A major warming around 24 January started the vortex break up.'

- page 12, line 390 to 394: I find that that the changes in water vapour are very difficult to see. Frost point temperatures are certainly not optimal here. Frost point temperatures have changed to water vapour mixing ratios in the maps. Profile pictures still show frost point temperature.

- figure 2: Could you use the x-axis range more efficient? Also a legend would be very helpful!
  X-axis ranges have modified and also a legend is inserted.

- figure 4: I wondered if it is an idea to split Fig. 4 into two figures, one showing the absolute volume mixing ratios at Sodankylä (top panel) and the other showing the residuals for 70°N - 90°N. This change in latitude, data sets shown and lacking legends made it difficult to digest. I do not know how often I read the caption. Having Aura/MLS data in the residual plots would be great too.
  Thanks for suggestion. Figure 4 have now split into two. Also MLS data is inserted. Legends are also inserted.

- figure 7: The colour bars have no annotation of what is shown. Also there is some overlap of the x-axis labels, which does not look particular nice.
  The units have been added to the colour bars and also the x-axis labels have now more space.

Technical corrections:

- page 4, line 94: replace "... polar regions and the mesosphere ..." by "polar regions, the mesosphere...".
  Corrected
- page 4, lines 127 and 128: “Carbon dioxide CO2 ...” should likely be “Carbon dioxide (CO2) ...”.
  Corrected
- page 5, line 143: Something seems to be missing here at the beginning of the sentence.
- page 9, line 205: “... ration ...” should read “... ratio ... ” or “... ratios ...”.
Corrected
• page 10, line 336: “... allways ...” should read “... always ...”.
Corrected
• page 11, line 367: “... at level 56 hPa ...” should read “... at the level of 56 hPa ...”.
Corrected
• page 11, line 368: “... timeperiod ...” should read “... time period ...”.
Corrected
• page 13, line 435: “... AURA ...” should read “... Aura ...”.
Corrected
• page 25, line 435: “All the maps are from 56 hPa altitude” could read “All maps consider the 56 hPa pressure level.”.
Corrected

References:


