Response to the Reviewer #2

We thank the Reviewer for the constructive review and the valuable comments. We carefully considered all suggestions and made improvements related to almost all of them. In the following we address the comments point by point. *The reviewer's comments are in italics.* Changes in the text are highlighted.

General Comments

One thought was whether this paper was more suitable for Atmospheric Measurement Techniques (AMT) rather than ACP. Well, then again, I suspect that the authors want to target the broader climate science community and not just measurement scientists. Anyway, they should just give some thought to this.

As the Reviewer correctly assumes we would like to target the broader climate community, which we explicitly state in the abstract and in the introduction section. We intentionally designed the whole study to focus directly on RO climatological products, in order to quantify their long-term stability and uncertainty regarding atmospheric trends. This is especially of interest for climate-interested scientists who might download and use such RO climatological fields for their research. That's why we chose ACP rather than AMT.

The figures in this paper need work. There are only 4 figures in the paper and they each consist of 20 panels or more. The panels are way too small to see much detail. For example, I would suggest splitting Figure 1 into 5 figures, one for bending angle, one for refractivity, one for pressure, one for geopotential height and one for temperature.

We thoroughly considered the concerns of the Reviewer and agree that the panels are kind of small for the ACP discussion paper format.

The figures were optimally designed for A4 format which is the final ACP paper format. In Fig. 1 we intentionally strived to provide the reader with an overview on RO atmospheric climatology products as provided by different centers. We therefore show the difference of each center to the all-center mean for all RO products, from the more basic bending angle via refractivity to the more derived parameters pressure, geopotential height, and temperature. We think that the reader can visualize the main differences easier in an overall plot on one page and gets a feeling on the basic behaviour of the products related to the processing steps of the centers. It is not the details but the big picture which should be transported here. Therefore we find the figure layout adequate and would like to keep it as is. Again, it will come better in the final ACP paper since we optimally designed the figures for that A4 format.

Specific Comments

Page 2, line 17: I am not sure how to interpret this "<0.06 K for temperature". Do you mean a temperature trend of 0.06K over 7 years i.e. around 0.1K/decade?

Page 2, line 19: But surely this depends on the expected temperature trend. Chemistry-climate models suggest expected trends in stratospheric temperature on the order of 0 to 1K/decade.

These comments relate to the following sentence in the abstract (page 2, line 16 to 19):

"In this region, the structural uncertainty in trends over 7 yr is < 0.03% for bending angle, refractivity, and pressure, < 3 m for geopotential height of pressure levels, and < 0.06 K for temperature; low enough for detecting a climate change signal within about a decade."

Here we state, that the **structural uncertainty in trends** is 0.06 K over 7 yr. We **do not** discuss climatological trends but the structural uncertainty in trends due to the different centers' processing schemes. (Please see the definition of structural uncertainty at the end of section 2.)

Thus, the structural uncertainty of the trend **does not** depend on the climate trend magnitude itself.
Regarding climatological trends, we briefly discuss former results based on RO data at the end of section 4.3 and relate them to the findings of this study.

Page 3, line 2: I think that you need to be more specific here. For example I suspect that the newly revised SSU time series (Wang, L.; Zou, C.-H. and Qian, H., Construction of Stratospheric Temperature Data Records from Stratospheric Sounding Units, J. Climate, 25, 2931-2946, 2012) could be considered as coming close to providing an upper atmosphere climate data record for temperature. Maybe you should just say that the radiosonde record is not sufficient? But at the very least you should state what the requirements are for a climate data record before saying that none of the existing data sets (other than RO) meet that standard.
Besides the requirements for a fundamental climate data record, we also included the requirements for essential climate variables. We now included further explanations and elaborated the definition of structural uncertainty in the introduction section (paragraph 1 and paragraph 2):
"Fundamental climate data records are defined by the Global Climate Observing System programme (GCOS, 2010) as homogeneous records derived from a series of instruments with sufficient calibration and quality control. Observation requirements are defined for essential climate variables (ECVs) such as upper air temperature, including for this variable a resolution of 500 km horizontally, 0.5 km vertically in the upper troposphere, 1.5 km vertically in the lower stratosphere, and a root mean square (RMS) accuracy of < 0.5 K.

Upper-air climate records from conventional observations such as weather balloons and microwave soundings currently fail to fulfill GCOS requirements (Immler et al., 2010). These observations were primarily intended for short-term weather forecasting and not for climate monitoring, the latter demanding accurate and long-term stable measurements. A substantial degree of uncertainty arises from changes in instrumentation and observation practice over time. Demanding homogenization and intercalibration procedures are required for the construction of a climate record and uncertainties in the trend rates and their vertical structure remain large in the upper troposphere and stratosphere (e.g. Randel et al., 2009; Thorne et al., 2011).

For a specific observational record, structural uncertainty arises due to different choices in processing and methodological approaches for constructing a data set from the same raw data (Thorne et al., 2005). Structural uncertainty for upper-air records decreases by increasing the number of independent observational data sets. Thus, multiple independent efforts should be undertaken to create climate records."

Regarding to the SSU time series there are still some basic issues to be clarified.
Thompson, D. W. J., et al.: The mystery of recent stratospheric temperature trends, *Nature* 491, 692–697, doi:10.1038/nature11579, 2012.

Page 3, line 28: I don't think that you have expanded the CHAMP acronym anywhere?
We included an explanation of the acronym "CHAllenging Minisatellite Payload for geoscientific research (CHAMP) satellite" in section 1.

Page 6, line 11: Even after reading this line a couple of times I still could not understand what the 'impact altitude' is. Maybe it's because I don't know what the 'impact parameter' is. I think that this needs to be explained for people such as myself who are not experts in RO.
We added an explanation for the impact parameter:

"The impact parameter is defined as the perpendicular distance between the center of local curvature and the ray path asymptote (e.g. Fjeldbo et al., 1971; Kursinski et al., 1997)."

"The impact parameter is defined as the perpendicular distance between the center of local curvature (which is close to the Earth's center) and the ray path from the GPS satellite (e.g. Kursinski et al., 1997)."

Page 6, line 14: I am very confused by this. The native vertical coordinate for RO is geometric height about the ground. Now you say you calculate geopotential height, assuming a scale height of 7km which is only appropriate when the atmospheric temperature is 239K, from pressure, but where does the pressure come from?

This is a misunderstanding. Geopotential height itself is a non-meteorological quantity. It is computed using the acceleration of gravity as a function of latitude and height.

We inserted the following explanation at the processing description in section 2 (last but one paragraph).

"Geopotential height is computed by integrating the latitude- and height-dependent acceleration of gravity over the RO derived altitude, divided by the standard acceleration of gravity."

In order to set geopotential height into meteorological context we relate it to pressure. We compute the so-called "pressure altitude" via the barometric height formula using the actual RO pressure field, a standard surface pressure, and a standard scale height of 7 km.

We clarified this in the text:

"In order to set geopotential height into meteorological context we relate it to pressure levels. We express it as a function of pressure in form of a "pressure altitude". The pressure altitude is defined as $zp[\mathrm{m}] = (7000\ \mathrm{m}) \cdot \ln(1013.25\ \mathrm{hPa}/p[\mathrm{hPa}])$."

Page 6, line 31: I don't know what you mean by 'suitably available'.

For the computation of bending angle sampling error it is necessary to apply a forward model to compute bending angle from refractivity. This forward model is currently not available for atmospheric reference fields (used to estimate the sampling error). We improved to the respective sentence to:

"(Bending angle requires a more complex forward model that depends on a different altitude coordinate, impact altitude; such a forward model is not yet suitably available for this study)."

Page 8, line 1: While I could unscramble all of the equations listed in the paper, I thought that in many cases it would be much easier for the reader if they were just replaced with narrative e.g. equation 5 could be replaced with something like "Time series were deseasonalized by subtracting the mean annual cycle from the data".

We prefer to keep the equations since they give an unambiguous description of the computation methodology. We improved the respective sentence to:

"In the following, to have the methodology clearly set, explicit equations are provided for the computations performed…"

Page 8, line 13: I think that you need to say a bit more about the regression model that you used to calculate trends. For example, what basis functions, other a trend basis function, were included in the model? How were uncertainties on the trends derived? This, for example, would be important if you are using trend differences to say something about the structural uncertainty in the different data sets. If they do not differ within their uncertainties then that's quite a different matter to if they do. That uncertainty will be affected by how you dealt with auto-correlation in the residuals when running the regression model. This is why I would like to see more details about the regression model.

We do not analyse the trends climatologically. We performed a simple linear fit to the anomaly time series, which we now explicitly state in the text. The trend analyses of anomaly differences is intended to show the uncertainty in trends derived from GPS RO data with regard to different processing implementations. Correlations are of no relevance in this context. They cancel out anyway since we work with anomaly differences.
We clarified this now and reformulated the last paragraph of section 2 accordingly:
"Trends in the anomaly difference time series are computed using a linear fit. The spread of anomaly difference trends and finally the standard deviation of the all-center mean trend are used as an estimate of the structural uncertainty of RO records (Wigley, 2006). We note that we do not analyze climatological trends but the uncertainty in trends from RO data with regard to different processing implementations."
We already have stated at the end of section 4.3:
"We note that the purpose of our trend analysis here is to provide an estimate of the structural uncertainty of climate trends from RO but not to interpret the trends climatologically. A thorough analysis of climate trends requires a full multivariate regression analysis also accounting for natural variability, e.g. El Niño Southern Oscillation and Quasi-Biennial-Oscillation (see e.g. Steiner et al., 2009; 2011). For this study we just performed a standard linear regression for the 7-yr CHAMP data set in order to derive the uncertainty of trends but the trends themselves are not meaningful in a climatological sense. "

Page 9, line 4: I really don't like the use of parentheses to create 'parallel sentences'. When I read 'Mean pressure (geopotential height)' I read this as geopotential height being another form of mean pressure which is certainly not the case.
We changed this sentence. It now reads:
"Mean pressure differences are within ±0.1 % and mean geopotential height differences are 5 m to 10 m below about 20 km altitude, increasing up to ±0.4 % for pressure and to ~25 m for geopotential height at 30 km."

Figure 2 caption: The figure caption is incorrect. It refers to panels (a) through (e) but there are only columns (a) and (b) on the figure. The caption needs to say rather that (a) is for the sampling error not removed and (b) is for the sampling error removed.
This figure caption has already been corrected for the ACPD online version. The caption now reads:
"De-seasonalized anomaly difference time series and trends of **(a)** refractivity and of **(b)** refractivity where sampling error was subtracted, shown for the (left) UT and (right) LS for (top to bottom) five zonal regions."

Page 9, line 16: I would suggest that rather than showing all of the panels without the sampling error removed, just show e.g. two pairs of panels with and without sampling error removed, as a completely separate and new figure, to demonstrate the value of removing the sampling error and then say that similar gains in noise reduction are achieved across all other regions. In the end, it is the two rightmost columns of Figures 2 and 3 that convey the key messages to the reader. If you drop the two leftmost columns the figures will be less cramped and also you can then expand the vertical axes to show more of the detail. I think that doing this would significantly improve
Figures 2 and 3. On line 19 you even say 'In the following we discuss the results after sampling error subtraction' so why spend half of Figure 2 and 3 showing results without sampling error subtraction?
We thoroughly considered the suggestion and re-checked the figures. As pointed out in the paper text, we find it important to show the plots with and without the sampling error in order to underpin the benefit of sampling error removal. Due to smaller atmospheric variability the

sampling error is found smallest at low latitudes equatorwards of 40°. It is larger at high latitudes, especially in winter due to a higher atmospheric variability. Subtracting the sampling error from raw RO climatologies still leaves a residual sampling error, which is estimated to be approximately 30 % or less of the original sampling error for all parameters (Scherllin-Pirscher et al., 2011a). Thus the noise reduction is better at low- to mid-latitudes than at high latitudes, which can be seen in Fig. 2 and 3. We also stated this in section 4.3: "Higher structural uncertainty at high latitudes is also caused by a larger residual sampling error (Scherllin-Pirscher et al., 2011a); at SHL the standard deviation does not decrease much after sampling error subtraction in contrast to all other regions (Fig. 4)."
Concerning the axes ranges, they were chosen to consistently display the range of the anomaly trend differences for refractivity and temperature, for the case where sampling error is included in the data (panels a) as well as for the case where the sampling error was subtracted from the data (panels b). In order to make the effect of the sampling error subtraction directly visible, the y-axis ranges have to be kept the same for panels a and panels b.
In regions where the trend differences are very small, e.g., in the tropics, trend lines may become hard to distinguish (since data of six centers are over plotted), but these cases are not the interesting ones anyway as the key of visually conveyed message here is the high consistency. The important cases are those where the trend differences are larger, e.g., at high latitudes, and there the trend lines are separating and are well visible. Keeping as-fixed-as-possible axes ranges supports to see this whole context. Thus we decided to keep the layout of the figures in order to transport what we find as main important points.

Page 9, line 30: Wouldn't it be more precise to say that these results indicate the advantages of the use of bending angle climatologies in the retrievals for climate trend studies?
We included the following reformulation:
"These results indicate the advantage of the use of bending angle climatologies for climate trend studies (Ringer and Healy, 2008)."

Figure 4: I would suggest splitting this into 5 separate figures, one for each variable.
We also designed Fig. 4 to fit the A4 format which is the final ACP paper format (Please see our comment above; general comment to Fig. 1). For comparison purposes of the trend uncertainty we think that it is more instructive to have one overview plot, showing how the trend uncertainties change for different latitude regions from bending angle to more derived atmospheric products. Details are given in Table 2 and 3 anyway.

Page 11, line 8: It doesn't appear that the MSIS acronym has been expanded anywhere?
We included an explanation of the acronym here:
"(Mass Spectrometer and Incoherent Scatter radar (MSIS) climatology (Hedin, 1991) …"

Page 11, line 31: But surely temperature changes are also a factor here since the scale height depends on temperature? I really think that you need to explain more carefully in this paper your use of geopotential height as a stability metric.
This comment relates to the following sentence: "Furthermore, geopotential height change (m decade^{-1}) relates to relative pressure change (% decade^{-1}) via the atmospheric scale height by a factor of about 70 m%$^{-1}$ (e.g. Leroy et al., 2006b; Scherllin-Pirscher et al., 2011b) …"
Please see our detailed explanation above on the computation of geopotential height and pressure altitude. Geopotential height is related to pressure levels via pressure altitude, which is computed using a constant scale height of 7 km.

Page 12, line 21: I don't understand what you mean by 'allows isolating the above

trends'. Can you please explain this more clearly.
We changed the formulation to:
"allows for a detection of the above trends as found by Lackner et al. (2011)."

Page 13, line 13: I don't understand what you mean by 'including different high altitude background information'. Can you please explain this more clearly.
We included the following explanation in section 2 at the retrieval description (paragraph#4):
"The bending angle decreases exponentially with altitude but the measurement noise stays relatively constant in the stratosphere, which leads to an exponential decrease in signal-to-noise ratio with altitude. In order to calculate refractivity, an initialization of bending angles with a priori information is performed at high altitudes to reduce the effect of error propagation downward."
We included further explanations in paragraph#3 of section 4.3. It now reads:
"Overall, structural uncertainty is lowest where the measurement information content is highest. Structural uncertainty increases with increasing altitude and at high latitudes. The larger differences between centers are regarded to mainly stem from increased sensitivity to the different bending angle initialization at high altitudes in the centers' processing schemes. Different centers use different a priori information, i.e. climatology models (DMI, GFZ, UCAR), exponential extrapolation (JPL), or numerical weather prediction forecasts (WEGC). The initialization approach affects uncertainty from about 25 km upwards and the atmospheric products become increasingly sensitive to a priori information as the altitude increases."

Grammar and typographical errors

Page 2, line 4: Replace 'provides' with 'has provided'. Likewise on page 3, line 6.
Replaced.
Page 3, line 22: Replace 'has been' with 'was'.
Replaced.
Page 5, line 30: Replace 'An overview on' with 'An overview of'.
Replaced.
Page 7, line 22: Replace 'center as function of' with 'center as a function of'
Replaced.
Page 8, line 13: Replace 'Trends of the' with 'Trends in the' and replace 'using standard' with 'using a standard'.
Replaced.
Page 8, line 22: Replace 'Figure 1 pictures' with 'Figure 1 shows'.
Replaced.
Page 9, line 8: Replace 'Fig. 1 pictures' with 'Fig. 1 demonstrates'.
Replaced.
Page 9, line 21: Delete 'basically'.
Deleted.
Page 11, line 9: Replace 'strato-' with 'stratosphere'.
Replaced.
Page 11, line 23: Replace 'program' with 'programme'.
Replaced.
Page 13, line 16: Replace 'were found consistent' with 'were found to be consistent'.
Replaced.
Thank you very much also for this editorial care.