Please note that there are some small mistakes in the manuscript, coming from the use of different latex packages. E.g. the authors names are not shown correctly and in composite units spaces are missing between the composites. These issues will vanish, when the right latex packages will be used and it's better to leave them as they are now, so that the manuscript does not need to be set back to the state as it is now later in the typesetting process.

Answer to remarks of Referee #1:

This paper describes new retrievals of sodium number densities from SCIAMACHY limb scans in the mesosphere and lower thermosphere. The paper discusses basic ideas and challenges of the retrieval, and presents results in terms of seasonal and latitudinal climatologies. For details of the retrieval algorithms, the authors refer to an earlier paper on SCIAMACHY Mg/Mg+ retrievals (Langowski et al., AMT, 2014).
I find the paper suitable for publication in AMT after some clarifications.

I find the discussion of the retrieval sound, including the spectroscopy and the radiative transfer of the sodium lines. Much consideration is given to the treatment of multiple scattering (or albedo contributions). Unfortunately, not all theoretical concepts work out when it comes to the practical retrieval. Nevertheless, the authors manage to define a useful way of handling multiple scattering. Here are the topics that I would like to see discussed more:

Langowski et al. (AMT, 2014) is a good reference explaining the details of the retrieval algorithms used here. Nevertheless, some more information should be given in the present paper in order to make it more readable, without the need of going back to the original paper too often. Hence, I recommend the authors to add some more information about the basic retrieval ideas in the current paper, without overloading it. One important example is the discussion of smoothing and the constraint parameter (section 2.3.1, figures 3-5). It would greatly enhance the readability if more basic explanations of the smoothing and its effect on the retrieval were given in the present paper.

More details of the retrieval algorithm were added to the paper using parts of Langowski et al. (AMT, 2014). However as the algorithm was already presented there, the discussion is kept rather short. Please note, that this comes along with reuse of same or very similar formulation as used in the original paper.

The handling of multiple scattering takes much place in the paper, including six figures. Sections 2.3.1 and 2.3.2 discuss two different methods. The detailed discussions in both sections are at times hard to follow. It might help the reader to give a brief summary of the methods at the beginning of each section before going into all the details.

A paragraph as a short summary is added to both sections.

The operational retrieval handles multiple scattering by invoking regular (lower atmospheric) SCIAMACHY scans from days before and after the MLT measurements. As the variation of multiple scattering (albedo) is strongly controled by the variability of tropospheric cloudiness, this approach would likely lead to rather large uncertainties when attempting to retrieve individual Na profiles.
For the zonally averaged approach
chosen here, the effect of tropospheric cloud variability may not be so critical. Nevertheless, I very much would like to see an uncertainty discussion addressing this.

This is a good point, which we thought was already treated in the discussion, but maybe not formulated well enough. The lower atmospheric variation of the actual measurement scene is considered due to the fit of the factor B. A more thorough discussion is added to this section.

How much uncertainty is contributed to the retrieval by basing the multiple scattering analysis on albedo data that are not coincident with the sodium measurements?

Before the retrieval is performed, the limb radiance data are daily and zonally averaged.
This raises some questions that should be discussed:
- To what extend is the averaging necessary, given that the signal-to-noise ratio presumably is substantially better for Na than for Mg/Mg+ as retrieved in the earlier paper.
Can individual Na profiles be retrieved?
It needs to be checked but it is highly likely that individual Na profiles can be retrieved. This is also mentioned in the paper now. Using nearly the same condition for the Na and the Mg/Mg+ retrieval has the advantage, that all three data sets were retrieved for the same conditions and are better inter-comparable. The first approach of calculating the multiple scattering fraction, that is not used in the end, needed a better signal to noise ratio than the now used approach. Therefore, investigation where concentrated on averaged data.
- How does the averaging affect the retrieval results, as compared to retrieving individual number density profiles first and performing the averaging of the individual profiles then?
Due to the non linear forward model, averaging before and after the retrieval step will not produce the same results. There are two effects that need to be considered. The first one is, that a low signal-to-noise ratio produces profiles that are shifted to higher densities. This was discussed for Mg in Langowski, AMT 2014. This shouldn't be an issue for Na for individual measurements. The second issue is that a high natural variability also causes a bias to higher densities, if considered. So the averaging before the retrieval step can lead to too small densities. However, we expect this effect to be small enough compared to our estimated error interval to result in significant changes.
We added a note to the paper.
- Do individual profiles show signs of deviations from well-behaved Na profiles, such as sporadic sodium layers, that may affect the averaging?
We checked individual spectra for sporadic layers, but only found very few cases, where these were clearly visible. In half of this cases a higher Na column density above peak altitude could only be seen for one of the Na lines, which rather points to a error in the measurement process, than a sporadic layer.
These events can also be seen in Figure 10, where in February, Mai and June sporadic signatures could be identified in the north at around 130 to 140 km. However, the number of spectra is rather small, and it is hard to tell if these features are true and significant. Therefore we omitted the discussion of these features.
No computational check for strong sporadic events in individual profiles was yet programmed, but the unsystematic checks by eye rather suggest a very low occurrence rate of these kind of events.

I also have some other minor comments:
Abstract, line 9: Please state the years for which data are presented (2008-2012).
corrected
Abstract, line 11: It may be clearer to write "peak altitude" instead of "altitude".
corrected
page 7914, line 11: "seperation" should read "separation".
corrected
pag 7925, line 3: Make clearer what you mean by "four to eight individual day measurements".
corrected... as most measurements of 2008 are only starting in autumn and there are only measurements for 2012 in Winter and Spring, there are more individual days available e.g. in december, than in june.
page 7926, line 23: "errors bars" should read "error bars".
corrected
page 7927, line 12: "lidar's" should read "lidar".
corrected
page 7927, line 21: You might consider calling the agreement for "good" rather than

"quite good".
corrected .... also for line 27 on page 7926
page 7928, line 18: "at summer" should read "in summer".
corrected
figures 3 and 4: In order to mark the lines in a consistent way, you may want to change the black dashed line into a red dashed line.
This was intended, as in Fig. 4 both lines are not separable anymore if they have the same color.

figure 6: Please specify in the figure caption what wavelengths have been used here. Is it the integrated wavelength range 650-660 nm or an individual wavelength pixel?
The wavelength interval is the one shown in Fig 7. We added a note of this in Fig 6 (649 to 661) nm

Answer to remarks of Referee #2:

General comments This paper deals with retrieval of sodium atom number densities between 50 and 150 km altitudes from SCIAMACHY limb emission measurements. The topic is interesting, analysis is sound and the paper is reasonably well written. However, in my mind it relies a little bit too much on the earlier paper by the first author. I hope that the authors could somewhat lessen this binding (see comments below). Anyhow, I would like to recommend this paper to be published in AMT. I have the following comments:
Specific comments

page 7916, line 7: In this section authors rely strongly on the earlier paper Langowski et al., 2014. I would recommend that you wrote few sentences about your 2-D-method. For example, is the method truly tomographic where latitude-altitude cells are probed from different direction by several LOSs?
Additional information on the calculation of the considered Paths is now added, which includes a short discussion how strong the overlap of the used measurements is. Different latitude and altitude are probed by different LOS. For a 4° separation of two limb measurements at the same tangent altitude, both LOS overlap lower than 5 km above the tangent point. It is probably a question of definition, what is meant with "truly tomographic".
Enough information is provided now, that the reader can draw his own conclusions based on his own assumptions what "truly tomographic" is.
page 7918, line 18: Same as previous. Open up a little bit more the differences to your earlier paper. Why is the Ring effect not important here? And open what do you mean by “etc”.
The same Ring correction as for Mg+ was tested and found to be negligibly small. As for Mg+, the error duo to the kind of Ring correction used, was larger, when the correction is applied than vice versa. The reason for the small Ring effect correction lies in the fact, that the resonance fluorescence to rayleigh scattering ratio is much higher for Na than for Mg.
page 7918, line 1: Please, explain why do you need an artificial model for the solar spectrum. Could you use some fine resolution measured solar spectrum?
The spectrum used is a measured photometer spectrum by McNutt and Mack, which has a higher resolution, than the width of the line. The formula is a good approximation for the measured center region of the line, which is the small region of interest (Fig 2) but of course not for the whole line as in Fig 1. There is a note in Fig 1 that says that.
The Doppler-shifts need to be calculated for the actual date and local time.
page 7918, line 16: Why do you mention only red shift here?

Some of the shifts are only red shifts, other can go in both directions.
There is the gravitational red shift. The Lindholm shift due to pressure is also a red shift but is already included in the used value. For the total red shift we used the value in McNutt and Mack's paper, which also listed other small shift contributions to the red shift. The Doppler shift can be red or blue shifts, which we noticed with the +- sign.

Sec. 2.2: Retrieval algorithm: Clarify by a few sentences what is your algorithm. You are mainly talking about forward model details.
The retrieval is the nversion of the forward model by a iterative Gauss-Newton algorithm. We only mentioned the details of the forward model, as the retrieval equation is the principially same as for the Mg/Mg+ retrieval, just that there is a different forward model used in the Jacobian J.
Sec. 2: The order of things is not optimal in this section. Usually the whole forward model is presented first and only then the retrieval algorithm (as I said above, algorithmic aspects are not really mentioned).
This is right, we actually totally omited the actual retrieval step. Beside that other matrices for the forward model are formed, it is exactly the same as for Mg/Mg+.
The original paper was a very long paper and even in this some parts are hard to understand, as they would have needed more space. Therefore the intention was to just mention the sodium specific new things and actually present results of Na here.
However, more information on the actual retrieval steps are provided now.
Fig. 2.: Mention your blue line identification in the caption
This has been done
Fig. 5.: Why do you extend the y-axis up to 150 km? Above 110 km profiles look like noise.
Measurements as well as retrieved densities are available between 50 and 150 km. Above 110 km the measurement error is larger than the mean. Nevertheless there is information on the upper limit of the density in the data. We rather like to show the full data, than only the peak region.
It also shows that no spurious artifacts occur above the peak region.

Fig. 8.: Say more clearly that the line colors used are the same as in Fig. 6.
The line colors are actually different, but we think, as the plots are very similar, it can be seen, which color was used for which plot. We added a note.