The reply to the anonymous referee #2 (RC4)

We are thankful to the referee for the careful reading of our manuscript and for the valuable comments. We appreciate all the comments; we took them into account while preparing the revised version of the manuscript.

Below, the actual comments of the referee are given in **bold courier font and blue colour**. The text added to the revised version of the manuscript is marked by red colour.

> **Section 2.1: This section should contain information about the calibration procedure of the HATPRO. How often where calibrations done? Is there any correlation between time since last calibration and quality of agreement between HATPRO and SEVIRI?**

We added the information about calibrations in the end of Section 2.1:

> "There were 13 calibrations of the instrument during this period of time including 7 absolute calibrations with liquid nitrogen and 6 sky-tipping calibrations. The interval between absolute calibrations varied from 2 to 4 months."

We did not estimate the correlation between time since last calibration and the quality of agreement between HATPRO and SEVIRI because the HATPRO measurements of LWP for the whole period of time under consideration have already been validated. And this fact has been mentioned in the original text: "…(3) the measurement data have already been validated and analysed for this time period (Kostsov et al., 2018)."

> **Line 159: Is sampling equal to averaging? If not, what is meant by 'sampling interval'? Does it mean taking just a single 1-s value every X seconds? If yes, why is this done? To safe disk space? To get into better agreement to the temporal resolution of SEVIRI?**

Sampling interval is the time interval between instantaneous measurements by HATPRO. Sampling period is the time of averaging of the signal coming from the atmosphere; it may be also called the integration time. In order to clarify this point we changed the text in the middle of Section 2.3. Now the text reads:

> "The sampling interval (the interval between instantaneous measurements) of routinely performed ground-based MW observations is about 1-2 s since the sampling period (the integration time of the incoming atmospheric signal) is equal to 1 s."

We have the feeling that the confusion was caused by the sentence:

"It has been noted by Rose et al. (2005) that the integration time (**or** the sampling interval) should not be greater than 20 s…"

Now this sentence reads:

> "It has been noted by Rose et al. (2005) that the integration time (**and also** the sampling interval) should not be greater than 20 s…"

In order to explain why the problem of variable sampling interval is important, we added the following text:

> "However, there are situations when keeping the sampling interval as small as possible is problematic. If an instrument is functioning in the mode of azimuth scanning or zenith scanning, some time is needed to change

pointing. Also, an instrument can be set to make the mixed mode observations. In this case the interval between measurements made in a certain mode can be rather large."

Lines 195-200: Suddenly the 'Petergof' site is introduced, but where does it come from? Is this the location of the MW? If yes, it would fit to introduce Petergof in line 195: "...measurement site of Petergof."

Yes, the radiometer is located near the railway station Petergof. In order to avoid confusion, we replaced "Petergof" by "St.Petersburg" in the revised text.

Lines 262-263: Is there any possible explanation for the found extremely large discrepancy? Such a statement leaves a very curious reader...

We pointed that the explanation will be given below:

"Since there was only one day with extremely large discrepancy between the results (day No 52), we excluded this day from the calculations (this is 14 May 2014 and the reasons for large discrepancies are discussed below)."

Line268ff: Which Seviri pixel was used for the comparison shown in Figure 8? Also at other positions in the text it was not always clear whether the shown results of SEVIRI are an average of the whole domain or only of a single pixel. If there was a general procedure applied, the authors should present it in Section 2.3.

We added the following remark in Section 2.3:

"It should be noted that all comparisons have been made for SEVIRI ground pixel which is the nearest to the radiometer site. In case other pixels are considered, it will be mentioned explicitly."

Line 312: Is 'not less' the same as 'at least'? If so, at least would lead to less confusion while reading.

In order to avoid confusion the sentence now reads:

"Based on these values and accounting for the higher latitude of the St.Petersburg measurement site, we can expect the parallax effect for the St.Petersburg measurement site to be about 3 km or more in terms of the displacement to the North direction."

Line 483: What determines 'systematic' and 'unsystematic' discrepancies?

To our opinion, the answer to this question asked by the referee can be placed in Section 5.2 "Analysis of discrepancies" before the discussion of unrealistically high LWP values produced by SEVIRI:

"The sources of systematic and unsystematic discrepancies are multiple. They can be related to the retrieval algorithms, parameters of time-averaging of HATPRO data, viewing conditions of SEVIRI, and also to weather conditions, type, height, spatial and temporal evolution of clouds, and the magnitude of parallax effect. The analysis of the details of retrieval algorithms is beyond the scope of the present study. The variation of the averaging period was shown to have minor influence on the results of comparison. Therefore, while making the analysis we focused on weather and cloudiness conditions provided by the SEVIRI observations simultaneously with LWP data."

Lines 487-489, or Line 510: Can the authors present ideas, which parameters should be investigated to identify the reasons for the occasionally observed discrepancies? Would the analysis of rain radar data help? Or can it be attributed to strong small-scale fluctuations of LWP?

We presented one of the ideas in the end of item 5) in Section 7:

> "To our opinion, in order to further analyse the reasons of these systematic differences, it would be useful to combine the HATPRO and SEVIRI data with collocated LWP data produced by the AVHRR instrument. Though the LWP measurement over St.Petersburg site is made by AVHRR only once per day, the size of the ground pixel of AVHRR is smaller than of SEVIRI and this fact would be very helpful."

Line 165: situation .... situations

Corrected. Now the text reads:

> "…for constant atmospheric conditions…"

Line 166: form ... from

Corrected.

Line 215: superscript m^-2

This sentence has been rewritten.

Line 231: "...not one, but two SEVIRI pixels..."

Corrected.

Line 239: "...20 min and 60 min, respectively."

Corrected.

Line 346: mono-model ... mono-modal

Corrected.

Lines 406-407: "...for the summer months..."

Corrected.

...there are much more typos present in the manuscript...

We checked the text and tried to correct all typos.

Figures 2-5: The same scale should be used for all figures. At the moment, each figure shows a different range of values, making a visual comparison impossible. Also, it would be of interest, how many data points are contained in each map. Could the authors provide a map of the number of data points per pixel?

We do not agree with the referee that the same scale should be used for all maps. The primary goal was to demonstrate the land-sea LWP gradient. We specially chose the scale and colours in order to reach this goal and to emphasize that the gradient area corresponds to the coastline. Therefore we kept the Figures unchanged in the revised version of the manuscript.

We indicated the number of data points per pixel not in the plot but in the text in Section 3:

> "Fig. 2 presents the maps of the mean LWP values calculated for the large and small terrains and for the whole considered 2-year period of observations (about 20000 data points per pixel)."

and

> "In order to assess whether this gradient can influence the results of the comparison of the SEVIRI and the HATPRO data, we plotted similar maps only for the data selected for comparison and considered the WH and the CD periods separately (about 4000 and 2000 data points per pixel respectively), see Fig. 3."

> **Figure 14: Please show a legend.**

We do not understand this comment. Figure 14 and Figure 15 are similar and both have the description of lines in the figure caption. Since the plots are simple with only two types of lines, the legend is not necessary.