Referee 3

GENERAL COMMENTS
This is a well written and interesting paper describing a neural network approach for retrieving soil moisture from SMOS in near-real-time. The results are highly relevant for operational applications in hydrology, meteorology and other earth sciences. The results are realistic and I recommend publication after minor revisions.

We thank you the referee, Prof. Wagner, for his constructive comments that have allowed us to improve the manuscript, in particular putting the new SMOS NRT SM in the broader context of other NRT SM processing chains such as the ASCAT one. We answer below the specific comments.

SPECIFIC COMMENTS
Page 3, lines 8-9: Explain why the NRT requirements cannot be met by the operational SMOS Level 2 processor. Is it just a matter of timeliness?

The operational SMOS level 2 processor performs a very detailed modelling of the Earth emission at 1.4 GHz at two polarizations and a large number of incidence angles. The surface is modelled at 4 km resolution and the processor includes the aggregation of the contributions of those 4 x 4 km^2 cells within a 123 x 123 km^2 square. The processor included simulations of the instrumental response to model the signals that are detected by SMOS. The vegetation optical depth and the soil moisture content are considered as free parameters that vary to minimize the difference of the simulated Tbs and those that were actually been measured by SMOS. This detailed modelling and minimization is done for every position of the ISEA grid with a 15 km spacing, i.e., for a half-orbit those operations have to be repeated ~1e5 times. Currently, the inversion of a half orbit takes around 6 hours using a cluster of computers, which is too long to be useful for NRT applications. In contrast, in the case of the NN algorithm, the training is done once, offline, performing a minimization that uses at once ~5e5 samples. But once that the NN is trained, the inversion of a whole orbit takes a few minutes using a single CPU.
A sentence has been rephrased in the introduction and the reader is referred to Sect. 2, where more detailed information on the level 2 processor is given in the new manuscript.

Page 3: Please also discuss possible disadvantages of the neural network approach already here. One topic is certainly the difficulties caused by changes in the sensor characteristics and Level 1 algorithms. Refer to experiences from other operational NRT services.

In the new version of the introduction we have cited the ASCAT NRT processing chain and that it is actually used for data assimilation at ECMWF. At the end of the same paragraph, we also comment on the fact that ASCAT and SMOS NRT processing chains are based on models whose parameters are fixed offline using a large amount of data and that those parameters should be updated if there are significant changes in the input data.

Page 4, line 5: What exactly do you mean by “arboreous component”? Do you mean the forest canopy?

We meant “trees”. We replaced “arboreous” by “trees”.

Page 4, line 16: “. . . was obtained by training . . .”

Done

Chapters 3.1 and 3.2: Please explain why the neural network relies on normalised data instead of the absolute brightness temperature values.

The local normalized index (or linear expectation) were showed to improve the retrieval results by Rodriguez-Fernandez et al. (2015). As explained in that paper, the use of those indexes as predictors was inspired by the “change detection” approach used by scatterometers. In the new version of the manuscript those informations are reminded explicitly in Section 3.1.

Page 8, line 23: 50 % is a very large value. Please explain.
Page 8, line 32: Here you allow no open water (0 %), which is in stark contrast to the 50 % threshold from above.

Thank you for pointing this out. Actually no additional filters regarding the water fraction were used to train the NN. The sentence has been removed. The maximum of 50 % of open water was used because it is the same value for ECMWF land products. Even with less than 50 % of open water the SM retrievals can be overestimated. However, L2SM retrievals are needed to define the local linear expectation index I, and the total number of L2SM retrievals with a free water fraction higher than 20 % is very low. Therefore, the same applies to the NRT NN retrieval. Still, SM values provided for footprints with open water are less reliable but it was decided to keep the values to allow the users to evaluate and decide by themselves. In the new version of the manuscript the readers are directed to the sea land surface mask aggregated into the ISEA grid in order to filter out mixed footprints if needed.

Page 8, bottom: Reformat list of put into table.

We guess that the referee remark was due to the double spacing of the “review” format for the manuscript but we have reformatted the list into a Table.

Chapter 4.1: Please explain why you decided not to use more advanced metrics.

We are not sure what are the metrics that the referee was thinking of. However, this paper concerns the presentation of the processor and the first evaluation results. In this context, we reckon that a global evaluation comparing the means of the two data sets, the bias, the Pearson correlation, RMSD and STDD, gives a good first idea of the properties of the new dataset. In addition, the evaluation of those two datasets has been compared to in situ measurements using the former metrics plus, in addition, the correlation of the anomalies time series to get further insight into the abilities of each dataset to capture the short term dynamics. We do reckon that the manuscript contains a good overview of quality metrics.

Discussion of Table 1 and Figure 5: The fact that the correlation R is overall slightly better for NRT than the L2 processor is noteworthy. Please discuss this in more detail and provide possible hypothesis why this is the case.

The reviewer is right. The NRT product shows a lower STDD and a higher R for the central two quartiles of the distribution (green boxes in Fig. 5). This behaviour was already found in previous studies such as Rodriguez-Fernandez et al. 2015. In the new version of the manuscript this result is discussed and interpreted in Sect. 5.3. Provided that the training is done with a large number of statistically representative samples, the NN will not be significantly affected by outliers or inconsistent values during the training phase and the NN output will the most likely (in the sense of the Bayes theorem) SM value taking into account a given set of input data. Thus, a good NN model can show slightly better quality metrics when compared to in situ measurements than the dataset used to as reference to train the NN.

Page 16, line 15: Explain what you mean by “similar”. What are the differences in implementation to the approach introduced by Rodriguez-Fernandez (2015)?

The main differences are :

- the training data is SMOS L2 and not ECMWF models
- a limited number of incidence angles is used in order to increase the swath width.
- Input Tbs are NRT Tbs and not Level 3 Tbs

Taking into account a comment by reviewer 1, which apparently understood that the NN was trained on ECMWF as in Rodriguez-Fernandez et al. 2015, we have decided to better explain the differences in Sect. 3 and to remove the sentence in the summary that could lead to misunderstandings for people not reading the whole paper.