

Reviewer 1:

General comments:

The paper assesses consistency among global land surface emissivity products. Four products are considered from respective sources - SSM/I, AMSR-E, TMI, and WindSat. Comparison is done at a monthly-average time scale. The paper illustrates that there are large inconsistencies, and shows where these are in a global, monthly-average sense. Unfortunately not much more than this can be concluded from the paper, which is a limitation. The authors cannot trace the inconsistencies to specific reasons. A long list of possible reasons is given, but this provides little real insight. The paper thus illustrates a conundrum for numerical weather prediction (NWP), which most likely needs instantaneous and not monthly-average emissivity anyway, in that microwave surface emissivity is not a well-defined or well-measurable property, and its use in NWP is still unreliable and, as indicated by this paper, may likely remain so for the foreseeable future.

This paper aims to investigate the consistencies and inconsistencies among microwave land surface emissivity products. As the reviewer mentioned many reasons could contribute to such inconsistencies. However, it is often seen that emissivity retrievals are used for the validation of physical model emissivity estimates. This paper, however, shows that even retrievals consist discrepancies and attention should be paid when they are used for such validation. This paper for the first time compares emissivity retrievals in global scale where all other current studies are limited to point comparisons. Moreover, it reveals places that have more consistencies as an indicator of confidence in using them in applications such as NWP, and the places that still need to be provided better estimation and certainty. The reviewer is correct that ideally instantaneous estimates are needed for NWP applications. First, not all retrievals have their estimates available in instantaneous scale, although monthly products are based on instantaneous retrievals which made such analysis impossible for all four sensors. Secondly, the instantaneous emissivity retrievals even have more discrepancies within each sensor. For instance, it has been shown that day and night estimates in arid regions from one sensor may have up to 10 percent difference that it is not expected in stable surface conditions (e.g. Norouzi et al, 2012 and Moncet et al, 2011). Therefore, the conclusion in this paper that these inconsistencies should be resolved at least in monthly scale would make sense. In final paper, we try to discuss these valid points/issues are raised by the reviewer.

Specific comments:

It would be helpful to include a brief discussion on how the NWP models use microwave emissivity, and whether this emissivity is required at specific microwave frequencies, viewing angles, and diurnal sampling. What frequencies, angles, times are these?

NWP models rely on Radiative Transfer (RT) models (e.g. Community radiative Transfer Model (CRTM)) that are used to determine the state of the atmosphere and account for the radiative transfer at different spectral ranges, among others the microwave frequencies (window and sounding channels). Estimates of the emissivities are particularly important as NWP models attempt to assimilate passive microwave observations over land (Prigent et al 2006). Specifically, when it comes to window channels, we do need to determine the surface radiance i.e. that is controlled by the land

emissivity. The simulations using the RT models should be carried across a wide range of angles and frequencies. Sensors can only provide estimates at specific angles and frequencies values. An interpolation along the obtained values can be considered. This description was added to the manuscript.

It seems that NWP models are needed in the emissivity retrieval to make atmospheric humidity and skin temperature corrections, and the emissivity is then used as a boundary condition for NWP models. Isn't this a circular process? Some explanation is needed.

Global microwave land surface emissivities are as a first step for data assimilation over the continents. Reliable emissivity retrieval could directly affect the atmospheric temperature and humidity retrievals (English, 1999). Therefore, the retrieved emissivities are used as a good first estimate to assimilate microwave radiances using microwave sounder data. Accurate estimates of microwave land-surface emissivities are essential to extract such information in any inversion scheme such as statistical or 1D-Var retrievals or within complex 4D data assimilation system used in numerical weather prediction (NWP) models. However, the required atmospheric information for emissivity retrievals doesn't come necessarily from microwave sounders. They might be from infrared-based sounders or reanalysis.

The paper suggests that a blended land emissivity product among the existing ones might be an ultimate step. However, based on the paper, this would not be a recommended path to pursue, and this should be stated in the conclusion. There would be no way to determine how the blending would be done given the lack of knowledge of the error structures in the product.

We agree with the reviewer that the lack of a specific structure in the relative "errors"/differences among the tested products may not foster the use of a single blended product in the retrieval. However, we do believe that a blended product may balance the errors amongst the distinct products as such product could be obtained using a weighted regression among all emissivity retrievals (where weights should vary spatially to account for the spatial variability of the consistency among the products) (Sahoo et al., 2011). The use of such blended product in NWP model for instance may lead to weather forecasts that are closer to the outcome when any individual land emissivity product is used. We agree with the reviewer that this remain an interpretation/expectation that should be corroborated with further analysis and assimilations. A discussion was added to the paper to elaborate on this topic.

The emissivity in low vegetation areas is affected significantly by soil moisture. The authors do not indicate what the sensitivity of emissivity to soil moisture is. Soil moisture varying from dry to wet can have a large effect on emissivity (there are many papers on this in the literature) and ignoring this fact can introduce large error into a NWP boundary condition. What is the sensitivity of NWP error to neglecting soil moisture variability effects on emissivity?

For sure, emissivity is affected by soil moisture, that's why we analyzed the responses of emissivity products against soil moisture in Figure 3. Evaluation of the products showed that as soil moisture increases the MPDI decreases. In previous publication we have amply addressed the relationship between emissivity and soil moisture. Our findings were in line with those stated in previous study, like the one published by Basist et al 1998 in which vertical polarization emissivity were found to decline across frequencies with soil wetness. By including reliable emissivity estimates in NWP models

the effect of soil moisture is embedded. Therefore, the effect of soil moisture in emissivity is NOT neglected. All emissivity products are based on microwave observations and they are NOT modeled emissivities and automatically the soil moisture has affected the signals. We have expanded our discussion of the point and we tried to be concise to avoid any potential redundancy with the previous publications.

The data processing method mentions resampling the emissivity products to the same spatial resolution. This is not a simple process to do correctly. It's not clear that the spatial resolutions have indeed been normalized. Further, there seems to be an assumption in the paper that the brightness temperatures of the various sensors (SSM/I, AMSR-E, TMI, WindSat) over land have been cross-calibrated, which if not would lead to differences in derived emissivity no matter how accurately the temperature estimation was done. Do the authors know if a cross-calibration was done?

The only adjustment to the products in this study was to re-project them to the same spatial resolution. This was a necessary step to perform such analysis. We have acknowledged such error in our intercomparison by *"This step is required to make the intercomparison possible despite the systematic differences that it may introduce."*

The cross calibration among the sensors may affect the comparisons of emissivity products. However, this effect is not significant based on many studies such as Holmes et al, 2013 that suggested the variability is less than 1K. This level of variability will have a low impact on emissivity retrieval. This explanation is at the end of discussion part of the manuscript *"A cross-calibration that involves data from all the used sensors in this study is necessary to detect the magnitude of the discrepancies in the raw data and the determined brightness temperatures and apply appropriate corrections to mitigate its impact on land emissivity retrievals."*

The MPDI depends on both soil moisture and vegetation but in the analysis of the MPDI the soil moisture effect is hardly discussed. The soil moisture dynamics tends to be somewhat smoothed out at the monthly timescale, but it is still there. What is the effect of diurnal sampling time differences amongst sensors? Emissivity varies with time of day as the surface moisture changes diurnally. Is this impact significant?

Similar to the previous comment, we added some explanation about the effect of soil moisture on emissivity estimates. This is a valid point that the difference in sampling time of sensors may change the emissivity estimates comparison where soil moisture has diurnal variation. We added this to our discussion as well.

Reviewer 2:

The paper provides a global intercomparison of different emissivity retrievals. There are various retrieval algorithms available for emissivity estimation, and still our understanding of surface emissivity products and their uncertainties are limited. One important question is how emissivity products can be used in precipitation retrieval algorithms to improve precipitation estimation. The paper performs a comprehensive evaluation of emissivity estimates from various sensors. The paper provides insights into

the quality of the available data sets, especially given that the analysis is performed at the global scale. Overall, I am in favor of publication after addressing the below comments:

Main Comments:

1- Although the purpose of the paper is to study the inconsistencies among emissivity retrievals, it would be helpful to mention the level of accuracy that is needed for different applications. For example, what level of accuracy of emissivity is needed to make the data useful for use in satellite precipitation retrieval algorithms?

For satellite precipitation algorithms also emissivity estimates with one percent accuracy is required based on many studies. Moreover, these algorithms require having emissivities that are dynamically accurate. The text is revised to address this comment.

2- It is not clear how emissivity can be used in other retrieval algorithms. Please add a discussion.

The application of emissivity estimates in NWP models to predict atmospheric humidity and temperature and total Precipitable water was added to the manuscript.

3- The Radiative transfer models used in the selected retrievals are different. It would be good to mention the differences of these models. In fact, this could help explain some of the observed differences.

All products use relatively similar general algorithm to estimate emissivity on top of canopy or the surface. However, the differences in ancillary data and the ways they may be interpolated for using in the retrieval may affect the retrieved emissivity values. Different radiative transfer models to account atmospheric contribution are among the sources of differences. A discussion was added to the paper to highlight such effects.

3- As mentioned in the paper, the WindSat removes the effect of vegetation while others retrieve the emissivity at the top of canopy. Doesn't this explain some of the observed discrepancies between and WindSat and other algorithms in northern regions? This is worth a discussion.

This statement in our manuscript was not accurate. The WindSat emissivity does use the same algorithm for emissivity estimation. We revise and eliminate this statement in the manuscript. Therefore, the higher differences between WindSat and other sensors still could be just based on filtering snowy pixels.

4- It would be helpful to quantify and report the maximum discrepancies observed for all the sensors as a general (metric) of uncertainty/error in Section Conclusions. Based on the presented results, the maximum error appears to be around 4 percent, except for AMSR-E that exhibits higher error at higher frequency (i.e., 89 GHz).

This is a good suggestion and we added this to the manuscript.

5- What is the level of uncertainty expected in "physical modeling" emissivity retrievals compared to "retrievals"? This is an important issue, and it is worth including a discussion on this.

Similar to previous comment this also was added to the manuscript.

Reviewer 3:

General Comments:

This paper presents an inter-comparison of four global emissivity retrieval products. The products each use different passive microwave sensors (AMSR-E, TMI, SSM/I, and Windsat), algorithms, and ancillary data. The differences are explored in terms of global monthly mean and standard deviation, as well as sensitivity to vegetation and soil moisture. Significant differences are found, particularly over desert and snow covered surfaces.

In section 1, the authors make a great point about the difficulty of validation in emissivity studies, and make a case for the role that inter-comparisons can play in assessment. I am not convinced that this is the best experimental setup for such a comparison, however. The point is made throughout the paper that the instruments and retrieval techniques are very different. In comparing emissivities for different sensors using different retrieval/radiative transfer schemes and input data for each sensor, there is little to be learned, as there is no way to separate out the individual differences/sensitivities.

The paper presented here builds upon an initial effort that LSWG team has taken to evaluate the effect of the surface on precipitation retrievals. The working group members acknowledged the need of analyzing the consistency of the available products to enhance the retrieval of precipitation. A first attempts was led by Ferraro et al. 2013 (cited in the paper) made the use of three-point measurements across the globe. Our study expands the analysis to cover the entire globe over five-year time period. We expected to find discrepancies as stated by the reviewer but the goal was to quantify the discrepancies and their spatial and temporal change across the products.

Moreover, this study evaluates the dynamic responses of the products to the surface condition changes in terms of soil moisture and vegetation density variations. Therefore, as mentioned by the reviewer, there are some limitations due to lack of a ground truth measurement in global scale. It was learned that although there are some relatively significant differences among the products, they capture the dynamic surface changes well. This will help to focus on some bias correction methods to eliminate the observed systematic differences. Therefore, we believe that the paper addresses questions raised by the community through the members of the Land Surface Working Group (LSWG) and builds upon previous attempts and brings the right answers in quantifying the discrepancies of land emissivity products and determining their seasonality and sensitivity to parameters like soil moisture and vegetation density. It is a work in progress and we agree with the reviewer that further analysis can be done in this topic in the future.

The use of different cloud clearing algorithms, for example, could lead to large differences and it is possible that much of this comparison is a comparison of the input data, with no way to conclude which is correct in an absolute sense. It would be much more useful to apply different retrieval techniques to a single sensor, or conversely the same retrievals to multiple sensors, to get at what is behind the

differences. Use of consistent surface temperature and cloud data sets could also lead to important information about differences in the retrieval schemes themselves. It is known that these products differ, and the next step is to determine why. The comparisons are a good starting point, but the work would benefit from applying this in some way, such as by offering a clear experimental plan for understanding the differences based upon what was learned through the comparison.

This is a relevant suggestion. The idea of using similar inputs for different algorithms or apply different inputs for a single approach makes sense and could help the uncertainty analysis of emissivity products. This study took advantage of dynamic analysis using standard deviation of emissivity values to evaluate them against known surface condition changes which helped to understand their variability. Moreover, it was found that the differences are somehow systematic across different surface types for various sensors (Figure 1). We added a statement to the manuscript regarding possible future studies that was suggested by the reviewer. Later suggestion (applying different inputs for the same algorithm), though, has been investigated in many previous studies (such as Norouzi et al 2011). However, conducting such study is out of the scope of this work. We were more interested to see where the consistencies and inconsistencies exist among the available land emissivity products, rather than conducting an error analysis which can be planned to perform in the future. Thank you for the relevant suggestion.

Specific Comments:

In the introduction, the early discussion could benefit by expanding discussion of how land surface emissivities are used in NWP and retrievals.

This was added to the manuscript as it was suggested.

Physical emissivity modeling is mentioned very briefly in passing. Is there a reason that a model product has not been included in the comparison? It may be helpful to mention this.

Physical models are required many inputs and less available in global scale. Moreover, the idea in this paper was to evaluate retrievals as they are often used as “ground truth” validation information in physical models. Therefore, it was intended to test their quality first. The level of uncertainty as it was added to the paper is higher in emissivity models. A note was included in the manuscript regarding this issue.

Section 2 – Data sets: a sentence about the importance of clouds would be useful here, and clear highlighting of the different masking techniques used.

This has been included in the paper.

Section 3 – more specifics are needed to describe the resampling and re-projection process.

A description regarding the resampling process and its effect on our comparison was added to the revised paper.

The use of monthly means and standard deviation of monthly estimates limits the scale of this comparison. There is no ability to compare how each technique responds to large changes in soil moisture following a rain event, for example, which would be useful in diagnosing differences. This method of comparison should be justified in the text.

As it was mentioned above, the paper intended to perform a consistency study on different emissivity products. Therefore, it was found that monthly means will reveal the systematic differences, and standard deviations will to evaluate the dynamic changes of emissivity products. The existence of differences among monthly averages highlights the existence of higher differences among instantaneous products. Not all products also were available in daily temporal scales as a product.

References:

English, S. J.: Estimation of Temperature and Humidity Profile Information from Microwave Radiances over Different Surface Types, *Journal of Applied Meteorology*, 38, 1526-1541, 1999.

Sahoo, A., Pan, M., Troy, T., Vinukollu, R., Sheffield, J., and Wood, E.: Reconciling the global terrestrial water budget using satellite remote sensing, *Remote Sensing of Environment*, 115, 1850–1865, 2011.