Dear Reviewer,

We would like to thank you for the thorough review of the paper you have provided. As you will see below, we think that many of your comments are pertinent and we have used them to improve the manuscript. For others, we think that there has been a misunderstanding and we have tried to explain our view on the subject.

Section 1:

1. Line 24-26, Page 11, It really depends on how you define the hydraulic properties. Nevertheless, in soil physics, no one say hydraulic properties -> saturated and residual soil water content. You can say, saturated and residual soil water content are part of soil hydraulic properties (only loosely).
Thank you very much for this remark, we have replaced property by conductivity which is more precise and corresponds better to what we wish to express.

Section 2.2:

2. The incidence angles for TB_HT and TB_OR are at 40deg and 42.5deg, respectively. Could the authors explain the possible inconsistency caused by this different configuration? This is suggested to do so to convince readers that such difference in incidence angels will not affect the results and the conclusions.
We selected the range of incidence angles 42.5±5° to build the TB maps from SMOS. Angular averaging of the observations is needed to ensure continuous maps covering the whole swath of the MIRAS instrument (see Fig. 1 in Piles et al., 2009). The choice of an angle of 40 or 42.5 degrees in the forward models is not affecting the presented results, since the two are contained in the selected SMOS angular range, and variations of the modelled TB between 40 and 42.5 are minimal (see Fig. 1 in Piles et al. 2010).

M. Piles, A. Camps, M. Vall-llossera and M. Talone, "Spatial-Resolution Enhancement of SMOS Data: A Deconvolution-Based Approach," in IEEE Transactions on Geoscience and Remote Sensing, vol. 47, no. 7, pp. 2182-2192, July 2009.

Piles, M.; Vall-llossera, M.; Camps, A.; Talone, M.; Monerris, A. Analysis of a Least-Squares Soil Moisture Retrieval Algorithm from L-band Passive Observations. Remote Sens. 2010, 2, 352-374.

3. For “simulation 2”, the “Soil Discretization” is merely the average of ORCHIDEE’s soil temperature simulation into 3 soil layers. This is actually not physical “soil discretization”, which used to configure ORCHIDEE runs. This means the “soil discretization” as the current manuscript proposed didn’t change any simulated soil states by ORCHIDEE, while just using the same ORCHIDEE simulation with averaging the finer layering into coarser layering. As such, this “simulation 2” doesn’t seem provide add-value information, unless the authors use different soil discretization configuration to run the ORCHIDEE one more time.
The state variables from the same simulation of ORCHIDEE were used to estimate TB_{OR} and $TB_{OR(SD)}$.The difference between them, is that $TB_{OR(SD)}$ was estimated using a coarser layering. For this, CMEM's configuration was modified and provided with a lower number of soil layers (3 instead of 11) and thus, different soil moisture and soil temperature profiles. The aim was to study if a coarser layering had an effect in TB simulation using CMEM. In order to avoid any confusion we have replaced “discretization” by “representation”.

Section 2.3

4. I appreciated authors’ efforts in explaining this with more details. It is quite understandable to use E-OBS or other precipitation data, LandSAF’s LST or

other LST products, as independent data to verify ERA-Interim data. Nevertheless, it is still needed to describe a bit more details on those reference datasets. For example, for E-OBS data over IP, you can further demonstrate how many rain gauge stations were included, while for LandSAF's LST, you can provide further information on how it was validated by others, with what in-situ data/other means, over IP.
The phrase was indeed ambiguous. We have reformulated it as follows : "For example, E-OBS data can be over-smoothed depending on the station network density (Hofstra et al., 2009) or for LST sensor noise, emissivity uncertainties, etc. are errors which can propagate in the LandSAf algorithm (Freitas et el., 2007)". We believe that the interested reader can refer to the cited papers in order to evaluate if he agrees with our choice of reference data sets for rainfall and land surface temperature. For your indication we have verified and about 200 stations cover the Iberian Peninsula in the E-OBS product. That is clearly insufficient to attain a reliable description of rainfall events at small scales (< 100km) but good enough to cover larger scales as indicated at the end of the paragraph.

5. It seems the conclusion in the manuscript will not change too much if you used different reference data? If it is the case, please indicate it in the manuscript. Or if different reference data will change the conclusion here, could the author help to illustrate/discuss such difference in terms of how and what?
We used E-OBS precipitation and LandSAF's LST because these are the most reliable data sets for the region which are available currently. This issue was already addressed in the previous review and included in the text (Page 11, lines 2 to 4). We do not believe it would be scientifically sound to pretend that our conclusions would not change with a different reference data set. We have no way of proving such a statement.

Section 3.1.2.

6. For Table 3, I see the TB_HT is filtered based on orography, while there is no similar filtering rules for TB_OR, why is that?
TB_{HT} was filtered by ECMWF using the criteria explained in section 3.1.2 and listed in the second column of Table 3. We filtered TB_{OR} following the same objectives as those from ECMWF's criteria to filter TB_{HT}. Although the filtering rules are different, they aim the same objectives. This is already explained in the text (Page 12, lines 3 to 5). However, we have included the following phrase to clarify this point "In the end, only data which is not masked in either case is retained".

Section 3.2.1

7. It is not very clear how the spatial correlation analysis was performed. Please give more details. For example, how was the "5-day window" applied? Could you give a concreate example to demonstrate?
Section 4.1
The text has been rewritten as "An averaging window of 5 days was applied to the data before performing the spatial correlation analysis to ensure the highest coverage possible".

8. For Table 4, why not temporal correlation calculated for the three simulations (VC, SD, FW)?
A good agreement was found between models and observations when analysing the temporal correlation. In the text, this is explained by the fact that this diagnostic is mainly driven by the "quick responses of temperature and emissivity to precipitation events, which drive TB's fast variations and correspond to the synoptic variability of the signal". A similar result was obtained by Polcher et al. (2015) regarding surface soil moisture. Bearing this in mind we decided not to compute the temporal correlation of $TB_{OR(VC)}$, $TB_{OR(SD)}$, and $TB_{OR(VFW)}$ with TB_{SM}, because as occurred for TB_{OR} and TB_{HT}, high values were expected and thus, no added value would be gained with this result. We did compute the spatial correlation between the 3 simulations and TB_{SM}, because poor spatial

correlations were obtained between models and observations (Figure 2), implying that these are not spatially consistent. Figure 2 also shows a marked seasonality, which suggests that different processes are behind the results for the temporal and spatial correlation. This was confirmed by the EOF study, which showed that the largest spatially coherent error between models and observations is dominated by the slow varying component of the TB signals (which is driven by the annual cycle).

Section 4.2.1

9. I believe that the Table 5 here is not correct and messed up somehow. Please clarify and correct.
Table 5 is as we intended it. Our aim is to show i) the high resemblance between the EOF's spatial patterns of the first variation mode and ii) the lower resemblance between the EOF's spatial patterns of the second variation mode compared to the first one, of:

- The spatial correlations of the TB error of ORCHIDEE in horizontal and vertical polarization (row 2 of Table 5).
- The spatial correlations of the TB error of H-TESSEL in horizontal and vertical polarization (row 3 of Table 5).
- The spatial correlation of TB errors of ORCHIDEE and H-TESSEL in horizontal polarization (row 4 of Table 5).
- The spatial correlation of the TB errors of ORCHIDEE and H-TESSEL in vertical polarization (row 5 of Table 5).

To do so, the spatial correlation between the EOF's spatial patterns of the TB errors (Figures 3 and 4) was computed.

10. It is not clear what is the difference between EC and the spation correlation in 4.1.2. Please elaborate.
Section 4.1.2 is about the spatial correlation between models and observations and there is no reference to EC (Expansion Coefficient).
In section 3.2.2 it is explained that the temporal evolution of a mode of variation is referred to as the Expansion Coefficient (EC) (Page 13, lines 6 to 9).

11. Line 18-24, page 25, it seems to me from Figure 2 I can understand easier for this paragraph.?
This paragraph is about one of the conclusions that can be drawn from the EOF analysis of the TB error, not the spatial correlation. Therefore, it refers to Figures 3 and 4. More precisely, to the temporal evolution of the EOF's spatial patterns. On page 13, lines 5 to 11, the Expansion Coefficients are defined (the reference Björnsson and Venegas (1997) is given, in case the reader wants to read more about this methodology). As said in the text, the ECs are the temporal evolution of a mode of variation. Figure 5 shows this evolution for the first EOF mode of variation. In this case, the behaviour describes an annual variation. Therefore, the main dominant error shown in Figures 3 (a and c) and 4 (a and c) is driven by the TB's annual cycle.

12. Line 21, "… evidences that …" , from what results it is evident?
The spatio-temporal correlation analysis between measured and modelled TBs showed high values in the temporal correlation and poor values in the spatial one. High temporal correlation values are mainly explained by the quick responses of temperature and emissivity to precipitation events, while the poor spatial correlation values, by the TB's annual cycle.

13. Line 22-24, "… confirm …", why need to confirm here and how it was confirmed? Is it not already demonstrated in Figure 1? And, the hypothesis you mentioned is not clearly explicitly explained in earlier sections, or please use a reference to point out where you put forward this hypothesis.
We agree with you and have moved this part of the text to Section 4.1.1. In addition, we have referred to section 3.2.1 to refer to the section where we introduce this hypothesis.

14. In general, I didn't see clear link between the statements in this paragraph (Line 18-24) and results you were referring to (Figure 3 & 4)
In Figures 3 and 4 the spatial patterns of the first (P1) and second (P2) variation modes of the TB errors are shown. Those from the first EOF variation mode show the largest spatially coherent error between measured and modelled TBs. It is the largest, because when performing the EOF analysis, this variation mode corresponds to the biggest eigenvalue, and thus, it explains the largest fraction of the TB error. However, there is no temporal dimension in the spatial patterns. To study their temporal evolution, we need to compute the Expansion Coefficients (ECs) of each variation mode, as explained in Section 3.2.2. These are shown in Figure 5 and allow us to conclude that the largest spatially coherent error identified in Figures 3 and 4 (a and c → P1) is dominated by the slow varying component of the TB's signal, its annual cycle.

15. Line 27, page 25, What does it mean by the ECMWF'S mean first guess departure? Could the authors present an equation to demonstrate how it was calculated? For the time being, this is confusing for readers. For example, the first guess departure is expressed as "obs-model", why not using "model-obs" to keep it consistent with Figure 3, 4 etc.
In numerical weather prediction jargon "first guess departure" is different from a model error as the difference is taken only a few hours after the assimilation process. Thus the initial conditions, obtained from data assimilation, are still the dominant factor in the error. In order to make the text easier to understand we have replaced "first guess departure" by "error".

16. Line 5-7page 26, what is the add-value of this statement, in the context of this study?
Discrepancies in the spatial structures between models and observations occur at a global scale, not only over the Iberian Peninsula. This is reminded in the conclusions (Section 6), when we draw the reader's attention to the fact this issue occurs at a global scale.

17. Line 8, page 26, "two TB errors" refer to TB_HT and TB_OR?
The text has been rewritten as "the ORCHIDEE and H-TESSEL TB errors".

Section 4.2.2

18. Line 17, page 26, "two TB errors" refer to TB_HT and TB_OR?
The text has been rewritten as "the ORCHIDEE and H-TESSEL TB errors".

19. Line 29-30, page 26. I assumed this is mainly due to the difference in simulated LST by ORCHIDEE and by HTESSEL? However, here you claimed this is most likely explained by forcing induced bias? This statement as in line 29-30 onwards to me is wrong.
This paragraph analyses the EOF's first variation mode spatial patterns of the LST and precipitation errors:

- P from the ERA-Interim forcing – P from the E-OBS dataset
- LST modelled by ORCHIDEE – LST provided by the LandSAF product
- LST modelled by H-TESSEL – LST provided by the LandSAF product

The patterns shown by the LST errors differ between them. On the one hand, the spatial pattern of the LST error computed using ORCHIDEE's modelled LST (Figure 7 a) shows a North-South gradient, which suggests that there is an issue with the available energy provided during the simulation. This suggests that it may be related to forcing induced biases. On the other hand, the spatial pattern of the LST error computed using H-TESSEL's modelled LST (Figure 7 c) shows a structure which, in principle, is not related to an issue of the available energy provided, as occurred in the other LST error.
None of these spatial patterns resemble those of the TB errors. Therefore, we discarded forcing induced biases as the main cause of the dominant TB error.

20. Line 4-10, page 27, this paragraph to me is too abrupt without any supports or proofs. Or you can refer to Figure 10. Anyhow, all statements should be backed up by results (data, table or figures).

This paragraph analyses the Expansion Coefficients (ECs) from the LST and precipitation errors. This is, the temporal evolution of the EOF's first variation mode spatial patterns of these errors (explained in the previous paragraph).
The manuscript follows the same strategy as the one followed for the TB error EOF : first the spatial patterns are analysed (Page 15, line 10) and then their ECs (Page 16, line 5). As done for the TB error, there is a figure showing the spatial patterns (Figure 7) and another one representing their Ecs (Figure 8). Therefore, we are not sure to have understood this comment, because the text from this paragraph refers to Figure 8, as said in Page 18, line 4.

21. Line 16 page 27 “higher frequency” , is this corresponding to first or second model? How to see it from Figure 7 not clear to me.
The text is about the ECs of the precipitation and LST errors. Therefore, we refer to Figure 8, as said in Page 18, line 4, not Figure 7.
If your doubt is about which is the EOF's variation mode whose spatial patterns are shown in Figure 7, it is the first one, as explained in Figure 7's caption.
We understand that by “model” yo mean “mode”. We would like to point out that a high frequency is not identified to a first or second mode. Depending on the ECs of a field's variation mode, we will observe a higher or lower frequency variation of the spatial patterns.

22. Line 20-23 page 27, “… The EOF analysis …” I don’t agree with this statement. You actually present in earlier section that precipitation can explain the error source for spatial patterns? Other than this, the claim on the “.., errors in their surface temperature …” is not correct as what I mentioned in comment nr. 18.
When dealing with the EOF analysis, we have to consider both, spatial patterns and ECs. The spatial patterns must not be isolated. The study shows a dominant structure of the TB error which is driven by the annual cycle of the TB signal. Therefore, the objective is not to identify the error source of the spatial patterns. In any case, precipitation is not presented as the error source of the main dominant structure of the TB error. We refer to precipitation in the following sections:
Section 2.3: It is explained that biases in the imposed atmospheric conditions can affect modelled TBs. Therefore, we decided to validate two important variables, one of them being the precipitation (Page 10, lines 11 to 15).
Section 4.1.1: High values in the temporal correlation between modelled and measured TBs are likely explained by the quick responses of temperature and emissivity to precipitation events (Page 14, lines 6 and 7).
Section 4.2.2: After analysing the spatial pattern of the EOF's first variation mode of the precipitation error (Figure 7) and its ECs (Figure 8) we discarded precipitation as the responsible for the dominant mode of the TB error, because neither the precipitation's spatial pattern, nor its ECs resemble to the spatial patterns and ECs shown by the EOF's first variation mode of the TB errors.
Regarding the claim “.., errors in their surface temperature …”, the reason for this is the same as the one explained to exclude precipitation: neither the LST errors' spatial patterns (Figure 7), nor their ECs (Figure 8) resemble to the spatial patterns and ECs shown by the EOF's first variation mode of the TB errors. We are sorry to say that we do not understand the link of this comment to comment number 18.

23. Line 7-10, page 28, not surprised, no big differences as stated, because it is not physical “soil discretization’ as I mentioned in comment nr. 2.
To reply to this comment, we refer to the reply given to comment number 3.

Section 4.3

24. Line 19, page 28, “The slow varying component” means P1 mode?
When we mention “the slow varying component of the TB signal” we are not referring to a variation mode, but to the TB's annual cycle. There is no temporal dimension in the EOF's spatial patterns of a variation mode (P1 and P2). To analyse their temporal evolution, their associated Expansion Coefficients (ECs) have to be computed, as explained in Section 3.2.2.

The slow varying component of the TB signal is driven by the TB's annual cycle, as said in Page 16, line 19 and line 21. In page 19, lines 21 to 23 it is explained that to analyse the TB's slow varying components their annual cycles are computed and shown in Figure 10.

25. Line 21, page 28, " … (Fig.5)" , Fig.2 can be put in this bracket as well?
Figure 2 cannot be included in this bracket, because the text refers to the "largest spatially coherent error structure between measured and modelled TBs" and that is shown by the EOF study of the TB error, not the spatial correlation between measured and modelled TBs.

26. Line 1, page 29. Please add a sentence here to say something like "the annual cycle variation of TBs do correspond well as shown in Figure 1?"
Figure 1 shows the temporal correlation between modelled and measured TBs. In the text it is said that the high values obtained are due to "the quick responses of temperature and emissivity to precipitation events, which drive TB's fast variations and correspond to the synoptic variability of the signal" (Page 14, lines 6 to 8). The temporal correlation is thus, mainly driven by the fast varying component of the TB signal and not its annual cycle.

27. Line 12, page 29, How it was different not clear. It looks similar in Fig.10.
The plot corresponding to Southwest IP shows a difference between modelled and measured TBs during Winter and Summer. This is not the case for the Northwest IP plot, where measured TBs are within the range of modelled ones. We have replaced the sentence with a more explicit one : "The plot shows that the TB errors behave differently over the annual cycle in both regions and are not related to the LST errors obtained when comparing to the LandSAF product."

28. Line 22-25. "This might explain why …" onwards. In my view, emissivity is not the sole factor affecting the TB calculation, effective temperature is another one.
To reply to this comment, we refer to the reply given to comment number 33.

29. Line 1-3, page 30, this sentence is too long and not clear what do you mean.
The regions where the EOF's first variation mode of the TB error show the largest discrepancies between measured and modelled TBs are over southwestern and northwestern IP. Their ECs show that these are maximum during the Winter season. Figure 10 shows that the annual cycles of measured and modelled TBs differ most during the Winter season. We have rewritten the sentence as : "The strong difference in behaviour between the two selected regions in winter is already visible in the dominant EOF mode of the TB error. The spatial patterns (Figures 3 and 4) have different signs in the North-west and South-west regions."

30. Line 4-6. Page 30, Again, the TB is affected by both emissivity and effective temperature. I would suggest authors to compare effective temperatures as calculated from both HTESSEL and ORCHIDEE. Perhaps the difference in TBs are induced by different Teff calculation from the two models.
To reply to this comment, we refer to the reply given to comment number 33.

31. Line 6-8, page 30, "on the other hand, … tose of TBs." these two sentences seem conflicting with each other.
We have clarified the sentence as follows : "On the other hand, the differences in annual amplitudes of LST could contribute to the differences in the amplitude of the simulated TB annual cycle."

Section 5

32. Line 13-15 Page 31, English is awkward here.
Sorry, we cannot see where the English could be improved. The authors agree that their ideas are clearly expressed in this paragraph.

33. For the explanations to the discrepancies, I suggested two more points, worthy of discussion:
We would like to thank you for the suggestions.
a) Is it possible due to different Plant Functional Types (PFT) as defined in ORCHIDEE and HTESSEL?
The possibility of discrepancies due to different PFTs was already taken into account in the study.

- TB_{OR} and TB_{HT} were simulated using the same vegetation cover: Ecoclimap
- $TB_{HT(VC)}$ was simulated using H-TESSEL's prescribed vegetation, which differs from Ecoclimap as explained in Page 7, lines 17 to 22.

The mean spatial correlation of the three sets of modelled TB with TB_{SM} provided similar values (Table 4: between 0.17 and 0.24 [horizontal pol.] and between 0.30 and 0.36 [vertical pol.]) and an EOF analysis of the TB error computed using $TB_{HT(VC)}$ showed similar EOF spatial patterns and ECs as those from H-TESSEL's TB error. This is explained in Page 19, lines 3 to 6.

b) The role of Teff in calculating TB can be another important factor, please see following references:
Lv, S., Wen, J., Zeng, Y., Tian, H., & Su, Z. (2014). An improved two-layer algorithm for estimating effective soil temperature in microwave radiometry using in situ temperature and soil moisture measurements. Remote Sensing of Environment, 152, 356-363

Lv, S., Zeng, Y., Wen, J., & Su, Z. (2016). A reappraisal of global soil effective temperature schemes. Remote Sensing of Environment, 183, 144-153

The role of T_{eff} has been already taken into account in the study, by comparing TB_{OR} with $TB_{OR(FW)}$. In TB_{OR}, The soil temperature profile of the 11 layer soil discretization is used as T_{eff}, while in $TB_{OR(FW)}$ the parametrization proposed by Wigneron (2001) is used to compute T_{eff}. Therefore, T_{eff} is computed in two different ways. However, the mean spatial correlation of $TB_{OR(FW)}$ with TB_{SM} provided the same value as that from the mean spatial correlation of TB_{OR} with TB_{SM} (Table 4: 0.2 [horizontal pol.] and 0.3 [vertical pol.]). In addition, an EOF analysis of the TB error computed using $TB_{OR(FW)}$ showed similar EOF spatial patterns and ECs as those from ORCHIDEE's TB error. This is explained in Page 19, lines 11 to 14.

Taking into account these results, we think that these two points are not responsible for the largest spatially coherent error identified in the EOF analysis of the TB errors.