Response to Reviewer 1 are structured as follow: (1) 1.X: comments from Reviewer 1, (2) Response to 1.X: author's response and author's changes in manuscript when any. For sake of clarity, line and page numbering from the revised version are used.

Reviewer#1

Dear Reviewer#1 many thanks for reviewing the manuscript and for highlighting its relevance and interest. Your comments and suggestions led to an improved version of the manuscript. Below is a point by point answer to your specific comments, all your editorial and technical comments were accounted for in the revised version of the manuscript.

1.1 [It would have been interesting to see a comparison of analysis vs. open-loop root-zone soil moisture skill (compared to the International Soil Moisture Network), as this could have a longer memory than the surface zone soil moisture, however, this is not crucial for the conclusions of this study.]

Response to 1.1

Thank your for your highly relevant comment. Following it and similar comments from the other Reviewers, it has been decided to revisit the soil moisture evaluation part of the study:
(1) we have added an evaluation of soil moisture from LDAS-Monde fourth layer of soil (10 to 20 cm) against in situ measurements of soil moisture at 20 cm depth when available (10 networks and 685 stations),
(2) for surface soil moisture (SSM), correlation values (R) were calculated for both absolute and anomaly time-series in order to remove the strong impact from the SSM seasonal cycle on this specific metric,
(3) a 95% Confidence Interval (CI) has been added to R values.
(4) we have added the number of stations for which correlations differences are significant (significant improvement or degradation from the analysis) as well as a map over North America for illustration.

It involves several changes in the revised version of the manuscript, they are listed below.

Methodology section, 2.5 Evaluation datasets and metrics

P.11, Lines 358-365: "In situ measurements of surface soil moisture from 19 networks across 14 countries available from the ISMN are also used to evaluate the performance of the soil moisture analysis. They represent 782 stations with at least 2 years of daily data over 2010-2018. Sensors at 5 cm depth (SSM) are compared with soil moisture from LDAS_ERA5 third layer of soil (4-10 cm), sensors at 20 cm depth with the fourth layer of soil (10-20 cm, 685 stations from 10 networks). Beside 11 stations located in 4 countries of Western Africa (Benin, Mali, Sénégal and Niger) and 21 stations in Australia, most stations are located in North America and Europe, see Table S3."

P.12, Lines 374-377: "For global estimates, Normalized RMSD (NRMSD, Eq.(2)) was used, also. Finally, for surface soil moisture, R was calculated for both absolute and anomaly time-series in order to remove the strong impact from the SSM seasonal cycle on this specific metric (see e.g. Albergel et al., 2018a, 2018b)."

Result section, 3..1.2 Ground-based datasets

P.17-18, Lines 548-580: "The statistical scores for soil moisture from LDAS_ERA5 open-loop and analysis (third and fourth layers of soil, 4-10 cm depth, 10-20 cm depth, respectively) over 2010-

2018 when compared with ground measurements from the ISMN (5 cm depth and 20 cm depth) are presented in Table S2 for each individual network. Averaged statistical metrics (ubRMSD, R,$R_{anomaly}$ and bias) are similar for both LDAS_ERA5 analysis and open-loop even if local differences exist. For the analysis, averaged R ($R_{anomaly}$) values along with its 95% Confidence Interval (CI) using in situ measurements at 5 cm (782 stations from 19 networks) are 0.68±0.03 (0.53±0.04) (0.67±0.03(0.53±0.04) for the open-loop) with averaged-network values going up to 0.88±0.01 (0.58±0.04) for the analysis (SOILSCAPE network, 49 stations in the USA) and always higher than 0.55 except for one network, ARM (10 stations in the USA) presenting an averaged R value of 0.29±0.05. Averaged ubRMSD and bias (LDAS_ERA5 minus in situ) are 0.060 m^3m^{-3} and 0.077 m^3m^{-3} for the analysis, 0.060 m^3m^{-3} and 0.076 m^3m^{-3} for the open-loop, respectively. NIC (Eq.1) has also been applied to R values, 65% of the pool of stations present a neutral impact from the analysis (511 stations at NIC ranging between -3 and +3), 12% present a negative impact (91 stations at NIC < -3) and 23% present a positive impact at (180 stations at NIC > +3).
The number of stations where R differences between the analysis and the openloop are significant (i.e. their 95% CI are not overlapping) is 186 out of 782 (about 26%). There is an improvement from the analysis w.r.t. the openloop for 128 stations (out of 186, i.e. about 69%) and a degradation for 58 stations (about 31%). Figure 7 illustrates R differences between the analysis and the openloop runs. When differences (analysis minus openloop) are not significant stations are represented by a small dot. When they are significant, large circles have been used, blue for positive differences (an improvement from the analsysis) and red for negative differences (a degradation from the analysis). For most of the stations where a significant difference is obtained, it represent an improvement from the analysis.
Averaged analysis R (95%CI), bias and ubRMSD for the fourth layer of soil (685 stations from 10 networks) are 0.65±0.03, 0.049 m^3m^{-3} and 0.055 m^3m^{-3}, respectively. For the open-loop, they are 0.064±0.03, 0.048 m^3m^{-3} and 0.056 m^3m^{-3}, respectively. For soil moisture at that depth, about 60% of the stations present a neutral impact from the analysis (410 stations at NIC ranging between -3 and +3), 28% a positive impact (189 stations at NIC > +3) and 12% a negative impact (86 stations at NIC < -3). Although differences between the openloop run and the analysis are rather small, these results underline the added value of the analysis with respect to the model run. Figure S6 represents the distribution of the scores values for LDAS_ERA5 open-loop and analysis using boxplots centred on the median value. They look very similar and from this figure, it is difficult to see either improvement or degradation from the analysis."

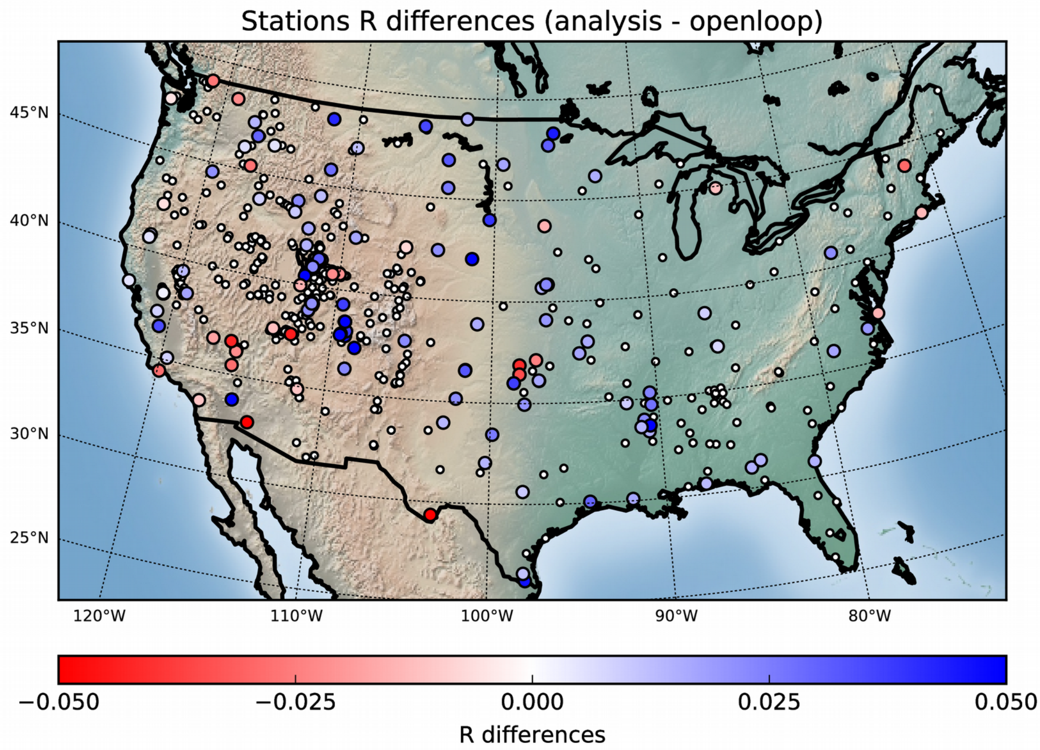

Figure 7: Map of correlations (R) differences (analysis minus openloop) for stations available over North America. Small dots represent stations where R differences are not significant (i.e. 95% confidence intervals are overlapping), large circles where differences are significant.

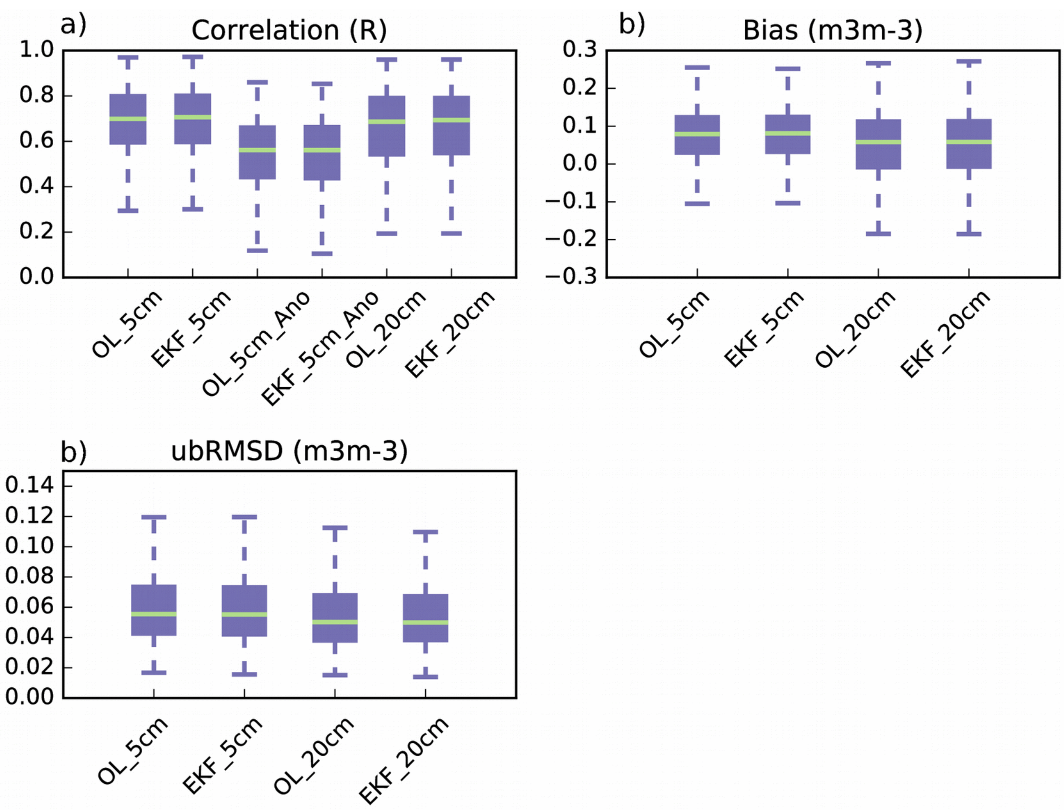

Figure S6: a) Boxplots representing the distribution of the correlation values on absolute time-series and anomaly time-series ("Ano") between the stations with in situ measurements of soil *moisture either 5cm depth or 20 cm depth and soil moisture from LDAS_ERA5 openloop and analysis over 2010-2018 (third and forth layer of soil, respectively). Correlation values are presented for surface soil moisture (5 cm depth measurements against third layer of soil), only. Distribution are centred on the median values. b) Distribution of the Bias values between the stations with in situ measurements of soil moisture either 5cm depth or 20 cm depth and soil moisture from LDAS_ERA5 openloop and analysis over 2010-2018 (third and forth layer of soil, respectively).c) Same as b) for ubRMSD.*

1.2 [L107-117: Is it necessary to include such details about the datasets in the introduction?]

Response to 1.2

We agree that a lot of information is provided in this bullet. However we believe acronyms should be detailed and appropriate references should be used the first time they appear in the text.

1.3 [L180: Please specify what you mean by flow dependency between the prognostic variables and the observations.]

Response to 1.3

This sentence has been rephrased: "Flow dependency between the model control variables and the observations are generated using finite differences from perturbed simulations" is now (P.6, Lines 192-195): "The flow-dependency (dynamic link) between prognostic variables and the observations is ensured in the SEKF through the observation operator Jacobians, which propagate information from the observations to the analysis via finite-difference computations (de Rosnay et al., 2013)"

1.4 [L198-200: Difficult to interpret the difference in LAI error when you use a mix of percent and m2/m2. Please could you clarify this?]

Response to 1.4

We agree that this sentence could be improved. Setting up the observed and modelled LAI standard deviation to 20 % of the LAI value is an empirical option coming from previous studies by Jarlan et al. (2008) and Rudiger et al. (2010), which have underlined the need for a variable LAI error definition. Barbu et al. (2011) further explored the impact of LAI model and background errors on the assimilation results by using diagnostics on model and observation errors (e.g. Desroziers and Ivanov, 2001) on different setups (see figure 2 of Barbu et al., 2011). They found that for small LAI values, it is necessary to use a fixed error standard deviation. This value was set to 0.04 m^2m^{-2} for LAI values lower than 2 m^2m^{-2} and is also used in this study.

The following sentence: "The standard deviation of errors for the observed LAI is assumed to be 20% and a similar assumption is made for the standard deviation of errors of the modelled LAI values higher than 2 m2m−2. For modelled LAI values lower than 2 m^2m^{-2}, a constant error of 0.4 m2m−2 is assumed (Barbu et al., 2011). More details can be found in Albergel et al, 2017 or Tall et al., 2019." as been reformulated and is now (P.7, Lines 220-224): "Based on previous results from Jarlan et al., 2008, Rüdiger et al., 2010, Barbu et al., 2011, observed and modelled LAI standard deviation errors are set to 20 % of the LAI value itself for values higher than 2m^2m^{-2}. For LAI values lower than 2 m^2m^{-2}, a fixed value of 0.04 m^2m^{-2} has been used. More detailed can be found in Barbu et al., 2011 (section 2.3 on data assimilation scheme and figure 2)."

Reference (not added to the manuscript):

Desroziers, D. and Ivanov, S.: Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation, Q. J. Roy. Meteorol. Soc., 127, 1433–1452, 2001.

Reference (added to the manuscript):
Jarlan, L., Balsamo, G., Lafont, S., Beljaars, A., Calvet, J.-C., and Mougin, E.: Analysis of leaf area index in the ECMWF land surface model and impact on latent heat on carbon fluxes: Application to West Africa, J. Geophys. Res., 113, D24117, doi:10.1029/2007JD009370, 2008.

Reference (already in the manuscript):

Rüdiger, C.; Albergel, C.; Mahfouf, J.-F.; Calvet, J.-C.; Walker, J.P. Evaluation of Jacobians for leaf area index data assimilation with an extended Kalman filter. J. Geophys. Res. 2010.

1.5 [L251: Could you please include why you don't consider assimilating surface soil moisture observations from the Soil Moisture and Ocean Salinity (SMOS) and/or from the Soil Moisture Active Passive (SMAP) satellite missions? As these satellites are expected to be more sensitive to surface soil moisture than the C-band observations from ASCAT.

Response to 1.5

We find it difficult at this stage to include why a specific dataset has not been used. The development of LDAS-Monde at CNRM has been made possible through different externally funded project including the Copernicus Gobal Land Service providing, amongst other datasets, the ASCAT Soil Wetness Index used in this study. ASCAT is from 2007 onward an operational product obtained from sensors onboard the METOP satellites and has been used at CNRM for many years. However, it is true than any satellite surface soil moisture products can be assimilate into LDAS-Monde. At CNRM, Albergel et al., 2017 have assimilated the ESA CCI (European Space Agency, Climate Change Initiative) combined surface soil moisture product (e.g. Dorigo et al., 2015), Parrens et al., 2014 have assimilated SMOS surface soil moisture. Future work will assimilate the most recent ESA CCI surface soil moisture dataset (v4.5) up to 2018. It includes the SMOS data.

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbu, A. L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Le Moigne, P., Decharme, B., Mahfouf, J.-F., and Calvet, J.-C.: Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area, Geosci. Model Dev., 10, 3889–3912, https://doi.org/10.5194/gmd-10-3889-2017, 2017.

Dorigo, W.A., A. Gruber, R.A.M. De Jeu, W. Wagner, T. Stacke, A. Loew, C. Albergel, L. Brocca, D. Chung, R.M. Parinussa and R. Kidd: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sensing of Environment, http://dx.doi.org/10.1016/j.rse.2014.07.023, 2015.

Parrens, M., Mahfouf, J.-F., Barbu, A. L., and Calvet, J.-C.: Assimilation of surface soil moisture into a multilayer soil model: design and evaluation at local scale, Hydrol. Earth Syst. Sci., 18, 673–689, https://doi.org/10.5194/hess-18-673-2014, 2014.

1.6 [Furthermore, as I understand ASCAT data are already assimilated in the production of the ERA5 dataset. Will the LDAS-Monde assimilation not lead to a "double" counting or usage of the ASCAT data and what are the potential consequences for your analyses results?]

Response to 1.6

Thank you for your comment. ASCAT soil moisture is indeed assimilated in the ERA5 LDAS. However, previous studies showed that its impact is confined to the soil and that it is neutral on the

IFS atmospheric analysis and forecasts (de Rosnay et al 2014, Munoz-Sabater et al 2019). In our study we use the ERA5 atmospheric analysis as forcing but we do not use any of the ERA5 soil analysis variables as input of our system. So, we consider the ASCAT SM contribution to the ERA5 atmospheric forcing to be negligible.

Reference (already in the manuscript):
de Rosnay, P.; Balsamo, G.; Albergel, C.; Muñoz-Sabater, J.; Isaksen, L. Initialisation of land surface variables for numerical weather prediction. Surv. Geophys., 35, 607–621, doi: 10.1007/s10712-012-9207-x, 2014.
Reference (not added to the revised version of the manuscript):
Muñoz-Sabater, J. , Lawrence, H. , Albergel, C. , de Rosnay, P. , Isaksen, L. , Mecklenburg, S. , Kerr, Y. and Drusch, M. (2019), Assimilation of SMOS brightness temperatures in the ECMWF Integrated Forecasting System. Q J R Meteorol Soc. Accepted Author Manuscript. doi:10.1002/qj.3577

1.7 [L268: Please could you specify the difference between linear rescaling and CDF-matching (if any)? To my understanding linear rescaling is correction of the mean and standard deviation, while CDF-matching corrects the whole CDF (i.e., all moments of the probability distribution function), hence linear rescaling is not the same as CDF-matching.]

Response to 1.7
We use in this paper a seasonal linear rescaling. Linear rescaling was introduced by Scipal et al. (2008) and has been shown giving results that are very similar to an exact CDF matching. Nevertheless, to avoid any confusion, we have rewritten the sentence as follows (P.9-10, Lines 294-301): “This is done through a linear rescaling as proposed by Scipal et al. (2007), where the observations mean and variance are matched to the modelled soil moisture mean and variance from the second layer of soil (1-4 cm depth). This rescaling gives in practice very similar results to CDF (cumulative distribution function) matching. The linear rescaling is performed on a seasonal basis (with a 3-month moving window) as suggested by Draper et al., (2011), Barbu et al., (2014).” Further mentions of CDF matching in the manuscript have been replaced by “seasonal linear rescaling”.

1.8 [L292-294: Could you please discuss how this short spinup period could affect your results?]

Response to 1.8

Nine months can be perceived as a short period to spin up the system. Unfortunately, HRES atmospheric forcing is only available from April 2016 and the LDAS-HRES experiment ends in December 2018. We have considered this 9 months period for the spin up in order to have the longest possible time series for land surface variables, thus giving more strength to statistics. We could have considered a longer period for spin up (April 2016 to December 2017) and studied only 2018. This gives very similar results on surface soil moisture and LAI (not shown). While not being fully spun-up, results obtained with LDAS-HRES can be considered as representative of the system response to data assimilation. Note that most initial values of the LDAS-HRES run are taken from the ECOCLIMAP-II database. For instance, initial LAI is set from a 1999-2005 climatology derived from MODIS.

Another possibility to initialise LDAS-HRES could have been to downscale the state of LDAS-ERA5 run in April 2016 to 0.10°x0.10° spatial resolution. LDAS-ERA5 runs have been set to an equilibrium spinning up 20 times the first year (2010).

The following sentence: "The period 2017-2018 is presented, HRES is available at this spatial resolution from April 2016, only, and the time period from April to December 2016 is used as a short spinup." has been modified and is now (P.10, Lines 327-332): "HRES is available at a 0.1° x 0.1° resolution only from April 2016. April to December 2016 is used as a short period for spinup and results are presented for the period 2017-2018. Although a 9-month spinup period can be seen as rather short, evaluating LDAS-HRES on either 2017-2018 or 2018 (using instead a 21-month spinup) leads to similar results on surface soil moisture and LAI (not shown). While the system is not fully spun-up, it can be considered as representative of the system response to data assimilation."

1.9 [L383: Could you please provide more details on how this can explain the differences seen between ISBA and GLEAM?]

Response to 1.9

GLEAM is an hydrological model and vegetation is mainly driven by observations. On the contrary, ISBA also represents plant growth and leaf-scale physiological processes, models key vegetation variables like LAI and above ground biomass (see section 2.1.1 on ISBA Land Surface Model). Within GLEAM, each grid cell comprises four different land-cover types: (1) bare soil, (2) low vegetation (e.g. grass), (3) tall vegetation (e.g. trees), and (4) openwater (e.g. lakes). Except for the fraction of open water, these fractions are sourced from the Global Vegetation Continuous Fields product (MOD44B), based on observations from the Moderate Resolution Image Spectroradiometer (MODIS). While in ISBA, each grid cell can be composed of up to 12 generic land surface types, bare soil, rocks, and permanent snow and ice surfaces as well as nine plant functional types (needle leaf trees, evergreen broadleaf trees, deciduous broadleef trees, C3 crops, C4 crops, C4 irrigated crops, herbaceous, tropical herbaceous and wetlands). Those types depart from prevalent land cover products such as CLC2000 (Corine Land Cover) and GLC2000 (Global Land Cover) by splitting existing classes into new classes that possess a better regional character by virtue of the climatic environment (latitude, proximity to the sea, topography).

Work is undergoing at CNRM to better understand the differences between terrestrial evaporation from ISBA and GLEAM. In particular, the different components of terrestrial evaporation, i.e. transpiration, bare soil evaporation and, interception loss are investigated.

Paragraph: "However GLEAM only estimates (root-zone) soil moisture and terrestrial evaporation, while ISBA in LDAS_ERA5 is a physically-based land surface model, accounting for more processes linked to vegetation."
is now (P.14-15, Lines 458-471):
"However GLEAM is an evaporation model designed to be driven by remote sensing observations only. GLEAM only estimates (root-zone) soil moisture and terrestrial evaporation while the CO_2-responsive version of ISBA in LDAS_ERA5 is a physically-based land surface model, accounting for more processes linked to vegetation (see section 2.1.1). It has to be noted that the auxiliary dataset used to e.g. represent the different land cover types are different also. Within GLEAM, the land cover types fractions are sourced from the Global Vegetation Continuous Fields product (MOD44B), based on observations from the Moderate Resolution Image Spectroradiometer (MODIS). Four land cover types are considered, bare soil, low vegetation (e.g. grass), tall vegetation (e.g. trees), and openwater (e.g. lakes). In ISBA the 12 land cover types fraction depart from prevalent land cover products such as CLC2000 (Corine Land Cover) and GLC2000 (Global Land Cover). It can potentially impact the distribution of the terrestrial evaporation between GLEAM and ISBA."

Further work at CNRM will focus on understanding the differences between ISBA and GLEAM, in particular investigating the sub-components of terrestrial evaporation.”